Study of rare b-baryon decays and test of lepton universality at LHCb

Vitalii Lisovskyi

université PARIS-SACLAY

NNT : 2019SACLS208

Thèse de doctorat

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir

Study of rare *b*-baryon decays and test of lepton universality at LHCb

Études de désintégrations rares de baryons beaux et test de l'universalité du couplage aux leptons avec LHCb

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud au sein du Laboratoire de l'Accélérateur Linéaire

Ecole doctorale n°576 Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: Physique des particules

Thèse présentée et soutenue à Orsay, le 09 Septembre 2019, par

Vitalii Lisovskyi

Composition du Jury :

Yasmine Amhis Chargée de recherche (LAL)	Co-directrice de thèse
Marcella Bona Senior Lecturer (Queen Mary University of London)	Rapportrice
Fabien Cavalier Directeur de recherche (LAL)	Président
Ulrik Egede Professor (Monash University)	Rapporteur
Gino Isidori Professor (University of Zürich)	Examinateur
Marie-Hélène Schune Directrice de recherche (LAL)	Directrice de thèse
Achille Stocchi Professeur (Université Paris-Sud, LAL)	Examinateur

Contents

Synthèse

Les transitions rares $b \to s\ell^+\ell^-$ sont un excellent laboratoire pour les tests de précision du Modèle Standard (MS). Dans le MS, elles se produisent uniquement dans des diagrammes dit en boucle ou en boîte (Fig. 1), ce qui les rend sensibles à de potentiels effets Nouvelle Physique.

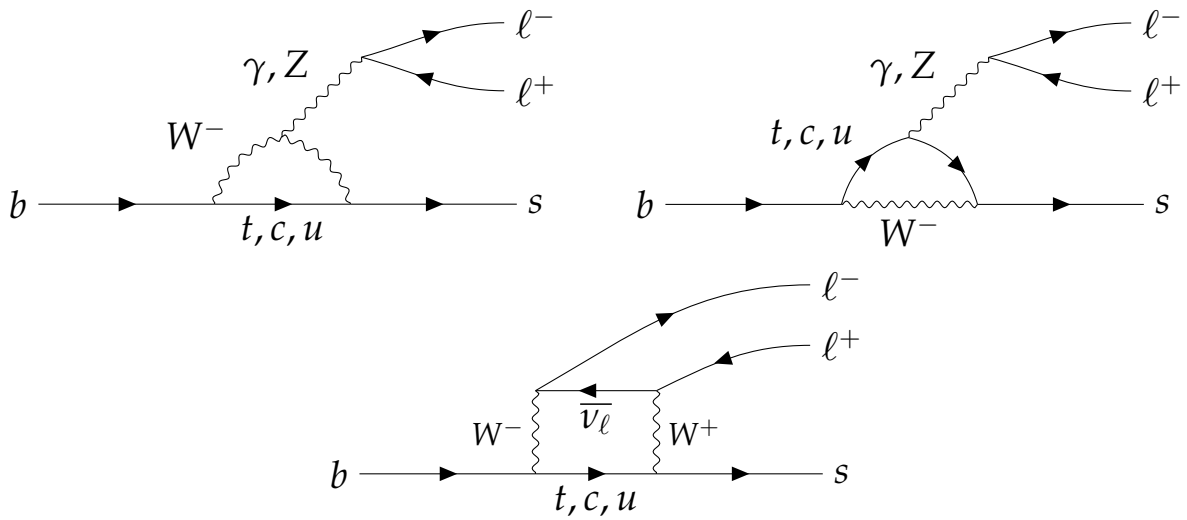


Figure 1: Diagrammes de Feynman contribuant aux transitions $b \to s\ell^+\ell^-$.

Des déviations par rapport aux prédictions théoriques ont été observées dans un certain nombre d'observables décrivant les transitions $b \to s\ell^+\ell^-$, comme par exemple les rapports d'embranchements différentiels, des paramètres angulaires ou les tests d'universalité du couplage leptoniques. Ces dernières mesures testent si les processus impliquant différents types de leptons dans l'état final – électrons ou muons – ont le même taux de désintégration, une fois corrigés les effets purement cinématiques dus aux differences de masse des leptons mis en jeu.

L'universalité du couplage aux leptons a été testé avec une précision inférieure à un pour cent dans des modes de désintégration au niveau de l'arbre tels que $Z \to \ell^+\ell^-$ ou $J/\psi \to \ell^+\ell^-$. Cependant, ce n'est que récemment que des tests précis ont été effectués dans les transitions $b \to s\ell^+\ell^-$. La collaboration LHCb a récemment annoncé deux résultats, testant l'universalité du couplage aux leptons dans les désintégrations $B^+ \to K^+\ell^+\ell^-$ et $B^0 \to K^{*0}\ell^+\ell^-$. Ces rapports, appelés R_K et $R_{K^{*0}}$, se sont révélés être à environ 2.5 écarts types des prédictions du MS [1, 2]. Pour déterminer avec certitude si l'universalité du couplage aux leptons est vérifiée dans les désintégrations $b \to s\ell^+\ell^-$, deux approches peuvent être adoptées : analyser une plus grande quantité de données, ou explorer de nouveaux modes de désintégration.

Cette thèse présente donc le premier test de l'universalité du couplage aux leptons dans les désintégrations des b-baryons. Les modes de désintégration $\Lambda_b^0 \to pK\ell^+\ell^-$ ont été choisies car ils sont accessibles expérimentalement à LHCb. La désintégration $\Lambda_b^0 \to pK\mu^+\mu^-$ a été observée dans l'expérience LHCb [3], mais son rapport d'embranchement n'a pas été mesuré. Le mode de désintégration $\Lambda_b^0 \to pKe^+e^-$ n'a pas encore été observé. Par conséquent, avant d'effectuer le test d'universalité du couplage aux leptons, il est nécessaire d'effectuer d'abord la première observation du mode de désintégration mettant un jeu un di-électron.

Cette thèse présente en détails toutes les étapes de l'analyse. L'ensemble de données LHCb, collecté en 2011 (s'élevant à 1 fb^{-1}), 2012 (2 fb^{-1}) et 2016 (1.7 fb^{-1}), à 7, 8 et 13 TeV, respectivement, est utilisé. L'analyse est effectuée dans l'intervalle de masse carrée invariante du dilepton $0.1 < q^2 < 6 \ \ \text{GeV}^2/c^4$, et de masse invariante du dihadron $m(pK) < 2600 \ \ \text{MeV}/c^2$.

Un ensemble spécifique de critères est développé afin de maximiser l'éfficacité de sélection et de supprimer les bruits de fonds dominants. Cela inclut notamment les critères d'identification des particules et des veto de masse dédiés à la suppression des contributions dues à des particules mal identifiées. Des classificateurs multivariés dédiés sont développés pour supprimer le bruit fond constitué de combinaisons aléatoires de traces. L'efficacité de ces critères de sélection est estimée à l'aide d'échantillons de simulation dédiés, qui sont étalonnés à l'aide de données de contrôle. Les incertitudes systématiques sont estimées en prenant en compte les imperfections de la simulation, le manque de connaissances sur la composition du bruit de fond ainsi qu'en faisant varier les fonctions décrivant les distributions en masses des différentes composantes de l'ajustement.

La principale difficulté de l'analyse est due aux spécificités de la reconstruction et du système de déclenchement des électrons dans LHCb. Comme ce sont des particules légères, les électrons émettent des photons de bremsstrahlung lorsqu'ils interagissent avec le matériel du détecteur. Malgré l'utlisation d'algorithmes dédiés à la récupération de ces photons, ce rayonnement entraine une dégradation de la résolution en masse. Par ailleurs, le système de déclenchement de l'acquisition de données a une efficacité limitée sur des électrons en raison de la forte occupation du calorimètre électromagnétique. Cela motive l'utilisation de deux catégories de déclenchement exclusives dans l'analyse en cours : les événements déclenchés indépendamment du signal, et les événements déclenchés sur les

électrons du signal (non comptabilisés dans la première catégorie).

Le bon étalonnage de la simulation est vérifié sur le mode de contrôle $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$, en profitant de l'universalité bien testée du couplage aux leptons dans les désintégrations $J/\psi \to \ell^+\ell^-$. La valeur du rapport $r_{J/\psi}^{-1}$, défini comme le rapport des rapports d'enbranchement des désintégrations $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ et $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$, étant connue, sa compatibilité avec l'unité est exploitée pour valider la methode. Ce test est effectué en l'aveugle : la valeur réelle de $r_{J/\psi}^{-1}$ n'est pas examinée jusqu'à ce que l'ensemble de la procédure soit etablie, et que toutes les vérifications soient effectuées. En particulier, l'absence de la dépendance du rapport $r_{J/\psi}^{-1}$ avec les variables cinématiques ou géométriques est vérifiée. Enfin, la valeur du $r_{J/\psi}^{-1}$ sans aveuglement se trouve dans un bon accord avec l'unité :

$$r_{J/\psi}^{-1} = 0.962 \pm 0.048. \tag{1}$$

Une autre étape de ce travail est la première observation du mode de désintégration $\Lambda_b^0 \to pKe^+e^-$. Tous les critères de sélection nécessaires à cette observation ont été préparés en aveugle, c'est à dire sans regarder les données dans la région du signal. L'ajustement de masse invariante a été préparé, en tenant compte des bruit de fonds les plus importants, et l'ajustement a été testé avec des pseudo-expériences. De nombreuses études systématiques ont été réalisées afin de quantifier les biais potentiels sur l'extraction du signal, en raison de la connaissance limitée de la composition du bruit de fond. L'effet systématique le plus important provient de la contribution de la désintégration partiellement reconstruite $\Lambda_b^0 \to pKe^+e^-\pi^0$, dont le modèle de désintégration est inconnu. Le mode de désintégration $\Lambda_b^0 \to pKe^+e^-$ a été finalement été observé avec un taux de $122 \pm 17(\text{stat}) \pm 6(\text{syst})$ candidats. L'ajustement sur la masse invariante du $\Lambda_b^0 \to pKe^+e^-$ est présenté en Fig. 2. Le signal observé a une signification supérieure à sept écarts-types.

Le nombre d'événements de signal observés en mode électronique est suffisamment important pour permettre le test de l'universalité du lepton R_{pK}^{-1}. Afin de lever certaines des incertitudes systématiques liées à un état final spécifique, le paramètre R_{pK}^{-1} est mesuré comme un double rapport :

$$R_{pK}^{-1} = \frac{\mathcal{B}(\Lambda_b^0 \to pKe^+e^-)}{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)} \times \frac{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))}. \tag{2}$$

La mesure est effectuée au moyen d'un ajustement simultané des données $\Lambda_b^0 \to pK\mu^+\mu^-$ et $\Lambda_b^0 \to pKe^+e^-$ dans toutes les catégories de déclenchement. Le nombre d'événements des modes $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$, ainsi que les rapports d'efficacité, entrent dans l'ajustement comme paramètres externes. Les paramètres R_{pK}^{-1} et $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))$ sont donc extraits directement de l'ajustement. La stablité de l'ajustement a été testé avec des pseudo-expériences. Les incertitudes systématiques sont réparties en deux catégories : celles qui ne sont pas corrélées entre les jeux de données et celles qui le sont entièrement. Le premier ensemble d'incertitudes

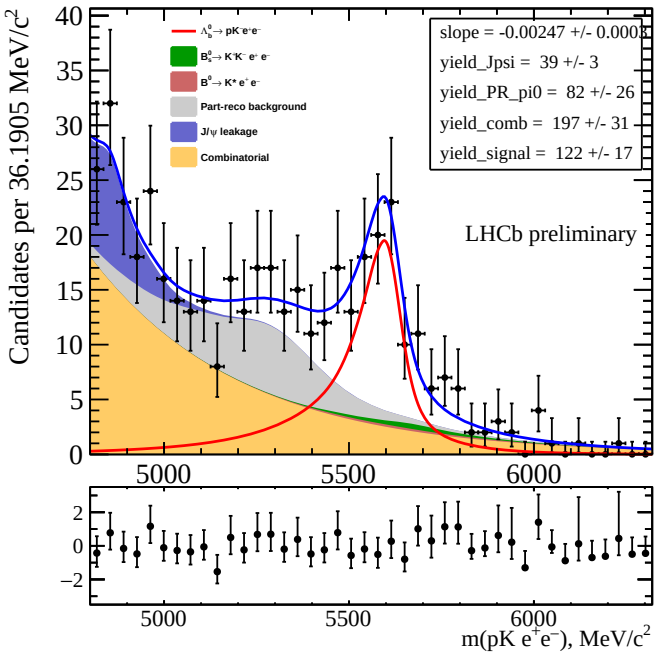


Figure 2: L'ajustement à la masse invariante $m(pKe^+e^-)$ dans les données LHCb.

entre en tant que contraintes gaussiennes sur les rapports d'efficacité. Le second ensemble d'incertitudes est convolué avec le logarithme de la fonction de vraisemblance finale.

Au moment de la présentation de ce travail, les jeux de données entrant dans l'ajustement étaient accessibles ainsi que les efficacités de sélection *sauf* celles liées à la désintégration $\Lambda_b^0 \to pKe^+e^-$ qui étaient modifiées de manière à contraindre R_{pK}^{-1} à l'unité. Les incertitudes sur R_{pK}^{-1} sont ainsi représentatives des incertitudes relatives attendues sur le paramètre.

Les deux paramètres libres de l'ajustement ont les valeurs suivantes :

$$R_{pK}^{-1}(\text{aveuglée}) = 1.00^{+0.15}_{-0.14}(\text{stat}) \pm 0.06(\text{syst}); \tag{3}$$

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} = (8.45^{+0.45}_{-0.43}(\text{stat}) \pm 0.26(\text{syst})) \times 10^{-4}. \tag{4}$$

Les deux mesures sont effectuées dans l'intervalle $0.1 < q^2 < 6$ GeV$^2/c^4, m(pK) < 2600$ MeV$/c^2$ uniquement. Les courbes des logarithmes de la fonction de vraisemblance finale en fonction des deux paramètres d'intêret sont présentés en Fig. 3.

En conclusion, le test d'universalité du couplage aux leptons R_{pK}^{-1} sera mesuré avec une précision d'environ 16%, et la première mesure du rapport des rapports d'enbranchement $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))$ est faite avec une précision d'environ 6%. Par ailleurs cette analyse permet aussi d'examiner le spectre de masse invariante proton-kaon dans le mode de désintégration $\Lambda_b^0 \to pK\mu^+\mu^-$, comme indiqué dans la Fig. 4.

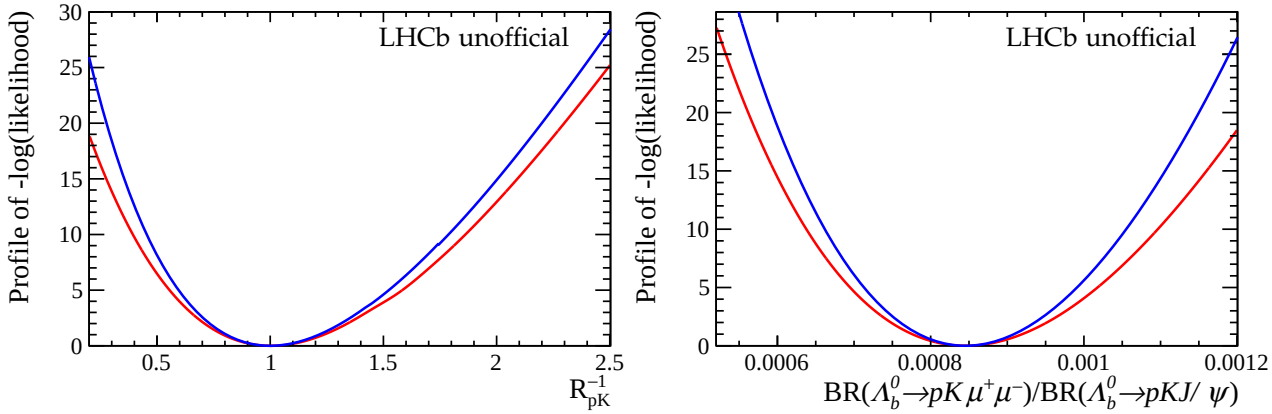


Figure 3: Courbe rouge : profil du logarithme de la fonction de vraisemblance pour (à gauche) R_{pK}^{-1} et (à droite) le rapport des rapports de branchement de $\Lambda_b^0 \to pK\mu^+\mu^-$ et $\Lambda_b^0 \to pKJ/\psi$, en tenant compte de toutes les incertitudes systématiques. Courbe bleue : mêmes profils, mais sans incertitudes systématiques.

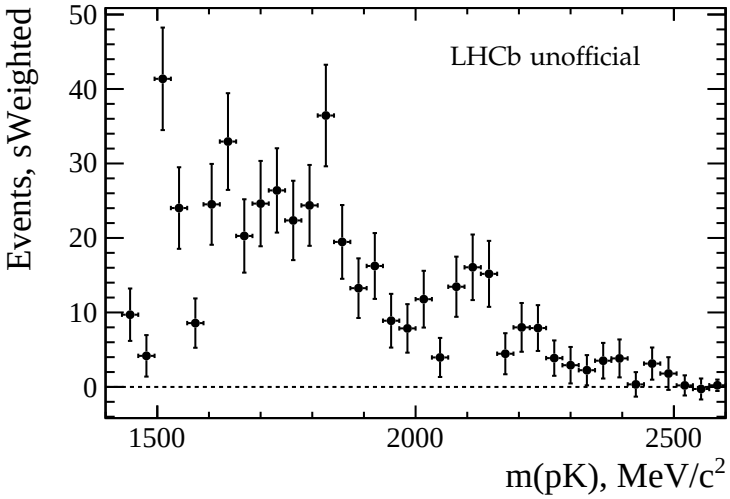


Figure 4: Spectre de masse invariante proton-kaon dans le mode de désintégration $\Lambda_b^0 \to pK\mu^+\mu^-$.

Au moment de soutenir cette thèse, l'analyse était en revue interne dans la Collaboration LHCb, les dernières étapes de l'analyse et les résultats sont attendus rapidement. Cette mesure fournira une première mesure de l'universalité du couplage aux leptons dans le secteur des baryons beaux et constituera un test supplémentaire du MS.

Le dernier chapitre de cette thèse est consacré à la première observation du mode de désintégration $\Xi_b^0 \to pKJ/\psi$. Ce mode de désintégration a le même état final que le $\Lambda_b^0 \to pKJ/\psi$, où des candidats "pentaquark" ont été découverts par la Collaboration LHCb [4,5]. Cependant, l'observation de ce canal est potentiellement compliquée par un faible nombre d'événements qui s'explique à la fois par une suppression due aux éléments de la matrice CKM entrant en jeu et par un taux de production du baryon Ξ_b^0 plus faible

que celui du baryon Λ_b^0. Par conséquent, une sélection dédiée a dû être développée afin d'obtenir la meilleure sensibilité pour ce faible signal. Elle a permis l'observation avec une significance statistique supérieure à dix écarts types du mode de désintégration $\Xi_b^0 \to pKJ/\psi$ puisqu'un clair signal de 118 ± 13 candidats a été obtenu comme l'illustre la Fig. 5.

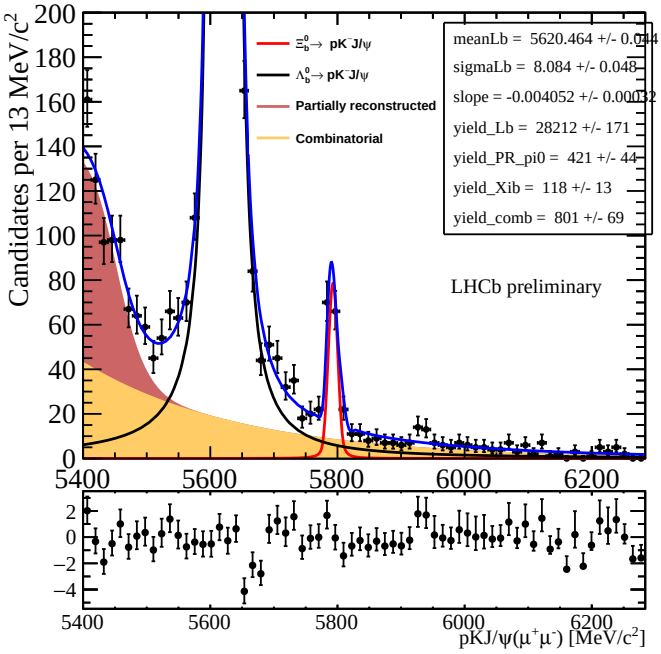


Figure 5: Ajustement de la distribution de masse invariante $m(pKJ/\psi(\mu^+\mu^-))$ dans les données LHCb.

Le rapport d'embranchement de ce mode de désintégration est mesuré par rapport au mode de normalisation $\Lambda_b^0 \to pKJ/\psi$. Le rapport des fractions d'hadronisation en baryons Λ_b^0 et Ξ_b^0 ($f_{\Lambda_b^0}/f_{\Xi_b^0}$) n'est pas mesuré à ce jour et peut à priori être différent pour différentes énergies. Les résultats sont :

$$\frac{\mathcal{B}(\Xi_b^0 \to pK^-J/\psi(\to \mu^+\mu^-))}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} \times \frac{f_{\Xi_b^0}}{f_{\Lambda_b^0}} = \begin{cases} (3.77 \pm 0.55) \times 10^{-3}, \text{ à 7-8 TeV}, \\ (3.05 \pm 0.51) \times 10^{-3}, \text{ à 13 TeV}. \end{cases} \quad (5)$$

CHAPTER 1

Theoretical and experimental overview

This chapter covers the theoretical and experimental background and motivation, on which the work presented in this thesis is based.

1.1 The Standard Model of particle physics

The idea that matter surrounding us consists of some elementary, discrete constituents, appears in philosophy since very ancient times. In the end of 18^{th} century, John Dalton was studying the patterns of mass proportions of various elements in chemical compounds. He proposed that each chemical element is represented by *atoms* of a certain type, which cannot be changed by any chemical reactions. Atoms were then considered as the smallest indivisible pieces of matter for almost a century. In the second half of the 19^{th} century, various physicists, notably J.J. Thomson, were studying the so-called cathode rays, luminescence rays emitted by an electrode located inside a high-vacuum tube. These rays were finally identified as charged particles and named ”electrons”. The first elementary (in the current understanding) particle was discovered. Studies of atomic structures and then-newly-discovered phenomenon of radioactivity not only led to the discovery of nucleons, but served also as a kickoff to the development of quantum mechanics. Studies of cosmic rays delivered several more discoveries – muons, pions and ”strange particles”. This led to some attempts of classification – the terms ”leptons”, ”mesons” and ”baryons” were born. Thanks to the development of accelerator and detector technologies throughout the following decades, an extensive list of seemingly elementary particles was discovered, with most of them being very unstable. It became evident that a deeper classification was required to sort out this ”particle zoo”. This was achieved with the development of the quark model and the theory of the strong interaction (quantum chromodynamics, QCD). In the meantime, studies of the radioactive β-decay and discovery of parity violation led to the development of the electroweak theory. The combination of the electoweak

theory, QCD, and the Higgs mechanism (described below) was nicknamed as "Standard Model" (SM). This unified theory allowed to predict the existence and properties of the $W^{\pm}$ and Z^0 bosons, and their discovery in 1983 was its greatest triumph. It was followed by discoveries of other missing building blocks – notably, the top quark in 1995, and finally the Higgs boson in 2012, the last piece of the SM puzzle. During the last decades, the SM has been tested by numerous precision measurements, the overwhelming majority of which has been able to verify its accuracy.

The SM is currently the most accurate theory describing the fundamental interactions and particles. It relies on the mathematical framework of a renormalisable quantum field theory. The SM incorporates the electromagnetic, strong and weak sectors. All known natural phenomena can be attributed at the microscopic level to one of these interactions. For example, the forces that bind protons and neutrons in the atomic nuclei are mediated by the strong interaction. The binding of electrons to nuclei in atoms, or of atoms in molecules (and therefore, entire variety of chemical phenomena) is caused by electromagnetism. Finally, the radioactive beta decay and the energy production in the Sun involve processes induced by weak interactions. Gravity is not incorporated in the SM, however, compared to the other three forces, the gravitational interaction is so weak at the scale of elementary particles, that it can be neglected.

The elementary constituents of the SM are classified as *fermions*, particles with half-integer values of spin (the intrinsic angular momentum), and *bosons*, particles having integer spin. In the SM, elementary fermions make up all visible matter of the Universe, while bosons are the force-carriers responsible for interactions between particles.

1.1.1 Elementary particles

Within the SM description, all visible matter in the Universe is made up of elementary fermions. They are divided into two categories, *quarks* and *leptons*. The fundamental difference between the two is that quarks are sensitive to all known interactions, whereas leptons are not sensitive to the strong interaction.

The SM includes six types, or *flavours* of quarks: up (u), down (d), charm (c), strange (s), top (t), and bottom, sometimes referred to as *beauty* (b). Quarks can be classified into three *generations*: the two lightest quarks, u and d, compose the first generation, c and s quarks – the second one, and finally two heaviest quarks, t and b, enter into the third generation. In each generation, there is an *up-type* quark with the electric charge of $+2/3$ units of elementary electric charge (u, c, t); and a *down-type* quark with a charge of $-1/3$ units (d, s, b). Quark masses range from a few MeV/c^2 for the light quarks, up to 173 GeV/c^2 for the top quark, as it is shown in Table 1.1. In the SM, quark masses are free parameters. However, the definition of the quark masses is not unique, as it depends on the quantum corrections and so on the energy scale. The most common renormalisation scheme for the quark masses in QCD is the so-called $\overline{\text{MS}}$ scheme, which is used to quote quark masses in this work. Quarks have non-zero quantum number of colour, therefore they cannot be observed directly, and almost instantly (at a timescale of about 10^{-23} s) *hadronise*, *i.e.* form states called hadrons, bound by the strong interaction. This is however not the case for the top quark, as due to its huge mass it decays even faster than hadronisation can

occur. To date, no hadrons formed by a top quark were observed. Possible colourless quark combinations define the classification scheme of hadrons: *mesons* are the states of quark content $q\bar{q}$, *baryons* are qqq states; also more exotic combinations such as tetraquarks ($qq\bar{q}\bar{q}$) and pentaquarks ($qqqq\bar{q}$) are allowed. Hadrons formed by heavy quarks (c and especially b) provide an excellent laboratory for tests of the SM, as their large mass allows for a rich variety of possible decay modes, each coming with its set of observables such as branching fractions, angular distributions or asymmetries.

There are also six types of leptons, and they can also be classified in three generations. Each generation (Table 1.1) includes a charged lepton (e^-, μ^-, τ^-) carrying one unit of elementary electric charge, and its neutral partner called neutrino (ν_e, ν_μ, ν_τ), which cannot participate in electromagnetic interactions. The number of leptons belonging to each generation is conserved in any interaction: this rule is known as the *lepton flavour number conservation*. It is believed to be exact for the charged leptons (to date, no lepton-flavour-violating decays were observed), however it is violated in the case of neutrino oscillations. There is no such constraint for quarks, where certain transitions between generations are allowed by the weak interaction.

No sign of existence of a fourth and further generations of quarks or leptons was found [6].

Table 1.1: Fermions in the SM. Mass values are based on Ref. [6] and are reported in the $\overline{\text{MS}}$ scheme. No uncertainty is shown when it is much smaller than the last shown digit.

	Quarks			Leptons		
Gen.	Flavour	Mass	Charge	Flavour	Mass	Charge
1st	u	$2.2^{+0.5}_{-0.4}$ MeV/c^2	$+2/3$	ν_e	< 2 eV/c^2	0
	d	$4.7^{+0.5}_{-0.3}$ MeV/c^2	$-1/3$	e^-	0.511 MeV/c^2	-1
2nd	c	$1.275^{+0.025}_{-0.035}$ GeV/c^2	$+2/3$	ν_μ	< 190 keV/c^2	0
	s	95^{+9}_{-3} MeV/c^2	$-1/3$	μ^-	105.66 MeV/c^2	-1
3rd	t	173.0 ± 0.4 GeV/c^2	$+2/3$	ν_τ	< 18.2 MeV/c^2	0
	b	$4.18^{+0.04}_{-0.03}$ GeV/c^2	$-1/3$	τ^-	1776.9 ± 0.1 MeV/c^2	-1

All of the stable matter surrounding us is made of the particles of the first generation: u and d quarks form protons, neutrons, and therefore atomic nuclei; which together with electrons group into atoms. This raises the question, why does Nature have not one, but three generations, which have very similar properties, except for the masses of the particles? Trying to find an answer to this question is one of the goals of flavour physics, in particular the LHCb physics program.

In addition to quarks and leptons, which constitute matter, there are anti-quarks and anti-leptons, constituting *antimatter*. Charges (of strong, weak and electromagnetic interactions) for anti-particles have an opposite sign with respect to the ones of corresponding particles, however they have the same mass, spin and other properties. Search for differences between matter and antimatter is a core part of the modern flavour physics research. In particular, it is not understood, why all the Universe seems to be made of regular matter, with almost no signs of antimatter.

1.1.2 Fundamental interactions

The theoretical pillar of the SM is the gauge invariance under local transformations of the symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{1.1}$$

where C denotes the colour, L – the left-handed chirality, and Y – the weak hypercharge. The $SU(3)_C$ symmetry is associated with QCD and the strong interaction. The $SU(2)_L \otimes U(1)_Y$ term describes the electroweak (EW) interaction, *i.e.* a combination of the weak and electromagnetic interactions.

The symmetries under the transformations of this gauge group determine the interactions and the number of gauge bosons which correspond to the generators of the group. All these gauge bosons have spin 1.

The gauge bosons mediating each of the interactions are:

- **Strong interaction:** eight massless neutral gluons g, each carrying a different combination of colour and anti-colour;
- **Electromagnetic interaction:** a massless neutral colourless photon (γ);
- **Weak interaction:** two charged massive colourless bosons $W^{\pm}$, and one neutral massive colourless boson Z^0.

Each interaction is characterised by its *coupling constant*, which determines the strength of the force mediated by this interaction. Despite its name, the coupling is not exactly constant, as it can change with the energy scale ("running coupling"). At low energies, the strong interaction has a coupling constant about two orders of magnitude larger than the electromagnetic one. The weak coupling constant is about four times larger than the electromagnetic one, however, the huge mass of weak gauge bosons makes this interaction effectively much weaker.

The different behaviour of the strong, electromagnetic and weak interactions as a function of the distance derives from the properties of the gauge bosons mediating them. Gluons carry nonzero colour, making interactions between themselves possible. On contrary, photons carry no electric charge, thus they do not interact with each other. The existence of the self-interactions of gluons leads to the different behaviour of the strong and electromagnetic interactions: while the strong coupling increases with distance, the electromagnetic interaction becomes weaker. At the same time, the short-range property of the weak interaction is due to the huge mass of the $W^{\pm}$ and Z^0 bosons (as compared to massless gluons or photons).

Finally, the SM includes a spin-zero particle, the Higgs boson, the leftover of the Higgs mechanism, that allows for the generation of the mass of particles from the spontaneous symmetry breaking of the gauge group of the electroweak interaction.

In order to describe the SM interactions, the total Lagrangian density is used:

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} \tag{1.2}$$

Quantum chromodynamics

The charge of QCD is the so-called *colour* charge. It can take the values commonly referred to as *red, green and blue*, as well as the corresponding anti-colours. There are eight gluons which correspond to the eight generators of $SU(3)_C$. Gluons couple to quarks and to other gluons via their colour charges.

The strong coupling α_s is a function of the transferred square of four-momentum q^2 in the reaction. At small transferred momenta values (equivalent to large distances), the value of the coupling becomes large, leading to non-perturbative phenomena. It becomes impossible to separate individual coloured particles (quarks and gluons), and this property is known as *confinement*. Only bound states of quarks, hadrons, which are colourless, can be observed. At the same time, the coupling decreases with increasing the q^2, this property is commonly called *asymptotic freedom*. The two regimes of large and small coupling values are separated by a reference scale, Λ_{QCD} which is of the order of 1 GeV [7]. For processes with a characteristic scale larger than Λ_{QCD}, perturbative methods can be used. This simplifies calculations for beauty hadrons, compared to the hadrons made of light quarks.

Electroweak theory

The generators of the electroweak group, $SU(2)_L \otimes U(1)_Y$, are the massless W_i ($i = 1, 2, 3$) bosons from $SU(2)_L$, and a massless B from $U(1)_Y$. These gauge fields couple to the characteristic charges of the electroweak interaction: the weak isospin T, with its third component T_3, and the weak hypercharge Y, related to the electric charge Q via the relation $Y = Q - T_3$.

An important property of the weak interaction is that it is only sensitive to left chirality component of fields, known as left-handedness. The W_i fields couple only to left-handed particles and right-handed antiparticles, while the B field couples to particles carrying weak hypercharge, regardless of the weak isospin. The left-handed fermionic fields are divided into doublets of one up-type quark and the corresponding down-type quark, or one neutrino and one charged lepton. Left-handed fermions carry weak isospin. Right-handed fermions, which have weak isospin T = 0, exist as singlets, for all particles except neutrinos. The SM does not include the right-handed neutrino, as it does not participate in any of three interactions, *i.e.* is "sterile", it has not been observed experimentally to date.

Table 1.2 reports the SM fermions and their important quantum numbers. For antiparticles, the values of T_3, Y and Q change sign. It is common to refer to particles in terms of their multiplicities in $SU(3)_C$ and $SU(2)_L$, and the value of the weak hypercharge. For example, the Q_L can be referred to as $(3, 2, +1/6)$, while $\overline{Q_L}$ as $(\overline{3}, 2, -1/6)$. Alternatively, it can be said that Q_L is a triplet under $SU(3)_C$ and a doublet under $SU(2)_L$.

Higgs mechanism

In the SM, at energies below the electroweak scale, the electroweak $SU(2)_L \otimes U(1)_Y$ symmetry is spontaneously broken to $U(1)_Q$. The fields corresponding to the physical

Table 1.2: Fermions in the electroweak theory, and their properties under $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$ groups.

Field	Generations			T	T_3	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$	$Q = T_3 + Y$
Q_L	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$1/2$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	3	2	$+1/6$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$
U_R	u_R	c_R	t_R	0	0	3	1	$+2/3$	$+2/3$
D_R	d_R	s_R	b_R	0	0	3	1	$-1/3$	$-1/3$
L_L	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	$1/2$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	1	2	$-1/2$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
E_R	e_R	μ_R	τ_R	0	0	1	1	-1	-1

mediators of the electromagnetic and weak forces are not the W_i and the B boson fields, but rather linear combinations of them:

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

where θ_W is the weak mixing angle, also known as the Weinberg angle.

Electromagnetic interactions are mediated by the massless photon, having equal couplings to the left- and right-handed fermions, with a coupling strength equal to the electric charge. Weak interactions are classified in *neutral currents* mediated by the Z^0 boson, and *charged currents* mediated by the $W^\pm$ bosons, which are linear combinations of the $W_{1,2}$ electroweak mediators.

The mediators of the weak interactions, contrary to the photon, are massive: $m(W^\pm) = 80.379 \pm 0.012\,\text{GeV}/c^2$, and $m(Z) = 91.1876 \pm 0.0021\,\text{GeV}/c^2$ [6]. This is a consequence of the so-called Higgs mechanism. It introduces the Higgs field ϕ, the only scalar field in the SM, with a non-zero *vacuum expectation value* v, *i.e.* the average value of the Higgs field in vacuum.

Additionally, this mechanism predicts a massive spin-0 particle, the so-called Higgs boson, H, discovered in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [8, 9]. The Higgs boson couples to the SM fermions with a strength proportional to their masses. In the SM, the Higgs field is defined as a complex $SU(2)$ doublet. Its Lagrangian contains the gauge interaction term (coupling to the electroweak gauge bosons), the self-interaction term (the Higgs potential) and the Yukawa interaction with fermions which is responsible for the generation of the fermion masses.

$$\mathcal{L}_H = |D_\mu \phi|^2 - V(\phi) + \mathcal{L}_Y, \tag{1.4}$$

with the scalar potential defined as

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{1.5}$$

with the mass term having $\mu^2 < 0$, and where λ is a positive dimensionless constant (representing the coupling of four-boson vertex of Higgs self-interactions). This potential

has minima at non-zero value of the field

$$\frac{-\mu^2}{\lambda} = \frac{v^2}{2} \tag{1.6}$$

providing a non-zero vacuum expectation value v. This induces the spontaneous symmetry breaking (SSB). Through this SSB, the Higgs gauge interaction provides a mass to the electroweak gauge bosons. The Higgs boson H arises from fluctuations about the non-zero minimum of the Higgs potential.

Finally, the Yukawa term of the Higgs Lagrangian, responsible for the fermion mass generation, is

$$\mathcal{L}_Y = -Y^D_{ij}\overline{Q}_{Li}\phi D_{Rj} - Y^U_{ij}\overline{Q}_{Li}\phi^* U_{Rj} - Y^L_{ij}\overline{L}_{Li}\phi E_{Rj} + \text{h.c.}, \tag{1.7}$$

where the quark and lepton left- and right-handed fields discussed in the Table 1.2 enter in pairs, multiplied by Yukawa couplings Y_{ij}, and indices i, j run over three generations. Higgs mechanism cannot generate masses for neutrinos as they have no right-handed partners. It should also be noted that the values of fermion masses and their hierarchy cannot be predicted from the SM.

1.1.3 A note on the units and dimensions

In particle physics, a common convention is to use units $\hbar = c = 1$. Under this convention, all physical quantities can be represented in units of energy (or mass), to some power. Counting these powers leads to so-called "dimensional analysis", which is widely used in theoretical considerations (*e.g.* Sec. 1.3.1). It is straightforward to see that in powers of mass, $[energy] = [mass] = [momentum]$, so these quantities are regarded as having *dimension* 1. Furthermore, using the well-known relationships between the energy and frequency ($E = \omega$ in the natural units), as well as between the momentum and the wavelength ($p = 2\pi/\lambda$), it is clear that $[time] = [distance] = [energy]^{-1}$, or, equivalently, time and distance have dimension -1. The derivative with respect to time and distance consequently have a dimension 1.

The action, which is the integral of Lagrangian density over the trajectory

$$S = \int d^4x \mathcal{L}, \tag{1.8}$$

is dimensionless, therefore the SM Lagrangian density has a dimensionality 4.

This allows to derive the dimensions of the fermionic and bosonic fields. Bosonic fields always have their kinetic terms entering as quadratic in $\partial_\mu(\text{field})$, and quadratic in the mass term. For example, the scalar field Φ has Lagrangian density $\mathcal{L} = \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{1}{2}m^2\Phi^2$, where m is the mass. The massive vector field A_μ has $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}A_\nu A^\nu$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This implies that any bosonic field has a dimension 1.

On contrary, fermionic fields Ψ enter with kinetic terms having two fields but only one derivative: $\mathcal{L} = \overline{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi$. Consequently, the dimension of a fermionic field is $\frac{3}{2}$.

1.1.4 Discrete symmetries

The concept of symmetry is one of the fundamental pillars of modern physics. A "symmetry operation" is a transformation which leaves a system unchanged. Three important discrete symmetries in particle physics are parity (P), charge conjugation (C) and time reversal (T).

The parity operation performs a spatial inversion. For a wave function $\psi(\vec{r})$,

$$P\psi(\vec{r}) = \psi(-\vec{r}). \tag{1.9}$$

The parity operation applied twice restores the original state. The parity of the wave function can be either even ($+1$) or odd (-1), or it can have no defined parity. The quantum number of parity allows to distinguish spin-0 particles into scalars (even) and pseudo-scalars (odd), while spin-1 particles can be vectors (odd) or axial vectors (even). Fermions have the even intrinsic parity, while anti-fermions – odd.

Charge conjugation operation reverses the signs of the characteristic charges of each interaction: electric charge, colour charge and the weak isospin (or weak hypercharge). This means, charge conjugation operation transforms a particle in its antiparticle.

The time inversion operation is defined as

$$T\psi(t) = \psi(-t). \tag{1.10}$$

This operation inverts the momentum and orbital momenta of a particle, but keeps unchanged its energy. Under the T symmetry, process $a \to b$ has the same probability as $b \to a$ (assuming both of them are allowed by other laws of physics).

Each of these three discrete symmetries is preserved in the strong and electromagnetic interactions, however they are (maximally, for C and P) violated in weak interactions. Even more, the combined CP symmetry is also violated in the weak interaction. The CPT theorem states that in relativistic quantum field theories formulated on flat space-time, all interactions are invariant under the combined transformations of C,P and T symmetries. So far, no experimental evidence of the CPT violation has been observed.

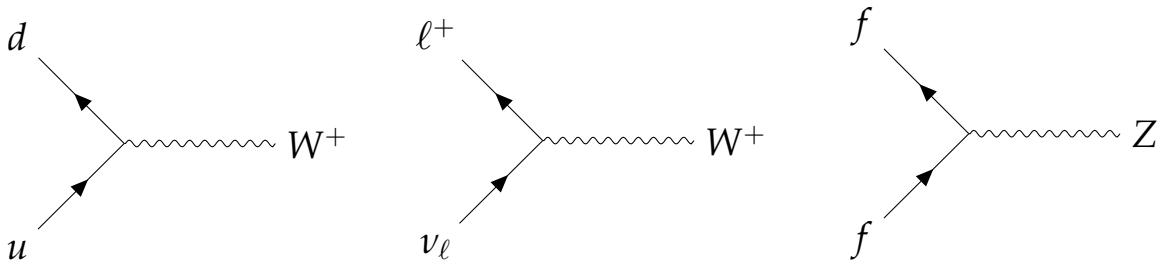
1.1.5 Flavour and weak interactions

The SM fermions couple to the $W^{\pm}$, Z gauge bosons via vertices shown in Fig. 1.1. It can be seen that charged weak current changes the up-type fermions to down-type ones, and vice versa. The neutral weak current does not change the flavour of the fermion.

For the leptons, charged currents can only change a lepton to another lepton from the same generation. This is the already discussed lepton flavour conservation rule. To first order approximation, the same holds for quarks. However, as it will be discussed below, certain transitions between the generations are also possible in the quark sector.

The vertex coupling (Feynman rule) for the charged weak vertices relating quarks or leptons within one generation, presented in Fig. 1.1(a,b) is

$$-i\frac{g_W}{\sqrt{2}}\gamma_\mu P_L, \tag{1.11}$$



(a) (b) (c)

Figure 1.1: (a-b): Charged weak vertices for quarks u, d and leptons ℓ, ν_ℓ; (c): Neutral weak vertex for any fermions f.

where $g_W = \frac{e}{\sin\theta_W}$, e is the elementary charge, and $P_L = \frac{1-\gamma_5}{2}$ is the left-handed chiral projector.

The neutral current vertex can be represented by the coupling

$$-i\frac{g_Z}{2}\gamma_\mu(v_f - a_f\gamma_5), \tag{1.12}$$

where $g_Z = \frac{e}{\sin\theta_W\cos\theta_W}$; $v_f = \frac{1}{2}T_3^f - Q^f\sin^2\theta_W$; and $a_f = \frac{1}{2}T_3^f$.

It can be seen that the couplings do not depend neither on the number of the generation, nor on the masses of fermions. The coupling is the same for $\mu\mu Z$ and eeZ vertices; it is also the same for $\mu\nu_\mu W$ and $e\nu_e W$ vertices. It should be also noted that the situation is similar in the electromagnetic interaction, where the $\gamma\ell\ell$ vertex has a factor $-iQ^f e\gamma^\mu$ independent of the generation. An important conclusion is that the weak and electromagnetic interactions are *lepton-flavour-universal*, *i.e.* their couplings are independent of the flavour of the lepton. This property is called lepton universality (LU). These couplings have been probed directly at Large Electron-Positron Collider (LEP) experiments, via precision measurements of the partial widths at the Z-pole, and the ratios of these partial widths are in a good agreement with unity [6, 10]:

$$\frac{\Gamma_{Z\to\mu^+\mu^-}}{\Gamma_{Z\to e^+e^-}} = 1.0009 \pm 0.0028; \tag{1.13}$$

$$\frac{\Gamma_{Z\to\tau^+\tau^-}}{\Gamma_{Z\to e^+e^-}} = 1.0029 \pm 0.0032. \tag{1.14}$$

Measurements also exist (at LEP and LHC) comparing the $W^\pm$ decays:

$$\frac{\mathcal{B}(W^-\to e^-\overline{\nu}_e)}{\mathcal{B}(W^-\to \mu^-\overline{\nu}_\mu)} = 1.004 \pm 0.008. \tag{1.15}$$

It should be noted that the LU holds in the limit of massless leptons, as mass enters into the phase-space factor. In reality, corrections due to non-zero lepton masses have to

be applied. Normally, these corrections are small for electrons and muons which are light, but can be large for the heavy tau lepton.

LU can be tested not only directly studying couplings of leptons to the gauge bosons, but also in the decays of hadrons. For example, a charged pion decay, mediated by a weak charged current, can be used to measure the following ratio [11]:

$$\frac{\Gamma_{\pi^- \to e^- \overline{\nu}_e}}{\Gamma_{\pi^- \to \mu^- \overline{\nu}_\mu}} = (1.230 \pm 0.004) \times 10^{-4}, \tag{1.16}$$

which is in a good agreement with the SM prediction of $(1.2352 \pm 0.0001) \times 10^{-4}$ [12]. It should be noted that in this particular case, LU does not imply the ratio to be equal to unity. The π^- is a spin-zero particle, and any spin-zero decay in two fermions implies that helicities of the two are the same (*e.g.*, right-handed anti-neutrino and a right-handed charged lepton). The vectorial nature of the weak interaction implies a helicity suppression of the electron mode, as in the massless lepton limit any of the two decay modes would be forbidden. As the weak interaction is sensitive only to the left-chirality fields, this decay is allowed only if there exists a reference frame, in which the charged lepton is left-handed. There is no possibility to define such a frame in the limit of a massless charged lepton, as massless particles move at the speed of light.

As it has already been mentioned, the situation is somewhat more complicated for quarks. The weak charged-current interaction is the only force in the SM which allows the changing of the quark flavour, and the generation. Weak-interaction quark eigenstates (also called flavour eigenstates) are not equal to the mass eigenstates. By convention, they are chosen to be equal for the up-type quarks, whereas the down-type quarks are chosen to be rotated between the flavour and mass basis:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.17}$$

V_{CKM} is the unitary Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [13, 14]. Phenomenologically, the CKM matrix is hierarchical, with diagonal elements close to one, and smaller (but non-zero) contributions outside the diagonal. The off-diagonal elements show a strong hierarchical order: $|V_{us}|$ and $|V_{cd}|$ are about 0.22, $|V_{cb}|$ and $|V_{ts}|$ of order 0.04, and $|V_{ub}|$ and $|V_{td}|$ of order 0.005. The fact that its non-diagonal elements are not zero allows for transitions between different generations of quarks through charged current interactions, although the small values of these elements imply the relative suppression of such transitions, known as *Cabibbo-suppression*.

The vertex factor for the charged quark currents of a type $i_u j_d W$ gets a form

$$-i\frac{g_W}{\sqrt{2}}\gamma_\mu P_L V_{ij}, \tag{1.18}$$

where V_{ij} is the relevant CKM matrix element. The transition probability between an up-type quark i_u and a down-type quark j_d is then proportional to the squared matrix element $|V_{ij}|^2$.

The hierarchy of the CKM matrix can be visualised by the Wolfenstein parametrisation [15,16]. The matrix can be expressed as

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.19}$$

The four real parameters are $\lambda = |V_{us}| \approx 0.23$, $A \approx 0.80$, $\rho \approx 0.14$ and $\eta \approx 0.34$. (As the matrix is unitary, and global phases cannot be observed, there are four free parameters). It should be noted that the Wolfenstein parametrisation is an approximation and is in fact not unitary in this exact formulation, however, this can be eliminated by redefining ρ and η.) The complex off-diagonal matrix elements give rise to CP violation in the SM, causing the different behaviour of quarks and anti-quarks under the weak interaction.

Expansion up to $\mathcal{O}(\lambda^3)$ terms is sufficient to describe the CP-violation in the B^0 meson sector. However, when discussing the CP-violation in B_s^0 or charm meson sectors, it is useful to expand V_{CKM} up to the $\mathcal{O}(\lambda^5)$ or even $\mathcal{O}(\lambda^6)$ terms, in order to get access to the smaller imaginary terms in the matrix elements V_{ts}, V_{cd}, V_{cs}. To date, the CP-violation has been observed in decays of kaons (1964) [17,18], B^0 (2001) [19–22], B^+ [23,24], B_s^0 (2013) [25], and, very recently, D^0 mesons (2019) [26]. No CP-violation has been observed in baryon decays to date [27].

The hierarchical structure of the CKM matrix explains, for example, a relatively long lifetime of the b hadrons. As the b quark decay to a top quark (involving the matrix element $V_{tb} \approx 1$) is forbidden by the energy conservation, the b-quark can only decay to lighter flavours, with transitions that are suppressed to a different order by the CKM elements. Such a protection does not exist for the charm hadrons, as the c-quark is allowed to decay to the s-quark via the CKM-favored ($V_{cs} \approx 1$) transition.

1.1.6 Shortcomings of the SM and searches beyond

Despite its enormous success in passing numerous precision tests, the SM is thought to be a low-energy effective approximation of a more global theory. Notably, the SM does not explain the matter-antimatter asymmetry of the Universe, and does not incorporate Dark Matter, which is believed to dominate over the usual matter in the Universe. It has quite a large number of free parameters (18 or 19), notably the masses of quarks and charged leptons, however not providing an explanation or the mass hierarchy between the different generations [28]. It does not explain the relation between the electrical charges of quarks and leptons, *i.e.* does not provide any reason why the charges of a proton and an electron have the same absolute value. As it has already been mentioned, it does not incorporate gravity.

These unanswered questions remain among the core topics of the current research activities in particle physics. Searches for signs of New Physics (NP), *i.e.* effects beyond the SM, can be performed in two ways. The so-called *direct searches* profit from an increase in the energy of the collisions in accelerators, allowing to produce heavier particles and therefore probe a higher mass range. This allows to set direct limits on the

masses (and existence) of the NP particles. Alternatively, one could exploit the *indirect* method, performing precision measurements of low-energy processes which involve virtual particles. In this case, the masses of virtual mediators which are probed, can be orders of magnitude larger than the scale of a studied process. Indirect measurements provide constraints relating the mass of NP particles and their coupling to the SM ones. Under a given assumption of the value of the coupling, an effect of the NP particles on the SM observables decreases with the increase of their mass.

Flavour physics probes the decays and interactions of b, c and also s hadrons, investigating the rates and angular distributions of the processes which could be affected by NP mediators. This technique is most successful when the SM process is very suppressed or even forbidden, so that even tiny NP effects become prominent. However, this requires a collection of huge data samples, and the relatively low statistical power is the main limitation of such studies.

Flavour physics has a notable record of particles predicted indirectly through their effect on low-energy processes: starting from the radioactive β-decay predicting the existence of the W boson; multiple observations in the kaon physics predicting the c quark and the existence of the third generation; and all the way to the prediction of the very large top quark mass from the frequency of B^0 meson oscillations [29].

1.2 Quark model and beauty baryons

Quarks are fermions, and, according to the spin-statistics theorem, the total wave function of a baryon consisting of three quarks should be antisymmetric under an exchange of any two quarks. The spectrum of allowed baryon states is constrained by the Pauli exclusion principle.

The total wave function of a baryon can be represented as a product of spin, colour, flavour and spatial wave functions:

$$\psi = \phi_{spin}\chi_{colour}\xi_{flavour}\eta_{space} \tag{1.20}$$

For ground state baryons, the angular momentum L is equal to zero. The spatial part of the wave function, proportional to $(-1)^L$, is therefore symmetric.

All particles which can be observed are colour-neutral, or *colour-singlet*. For baryons, it is only possible to create a colour-singlet combination of three quarks if all three have different colour:

$$\chi^{singlet}_{colour} = \frac{1}{\sqrt{6}}\left(RGB - RBG + BRG - BGR + GBR - GRB\right) \tag{1.21}$$

This combination is antisymmetric under exchange of an arbitrary pair of quarks.

This implies that $\chi_{colour}\eta_{space}$ is always an antisymmetric combination for the ground-state baryons. Consequently, the remaining part of wave function, $\phi_{spin}\xi_{flavour}$, is required to be always symmetric in order to create the antisymmetric total wave function.

The state of three quarks (each having a spin 1/2) can have the following combinations of the total spin and its projection $|S, m_s\rangle$, where an $\uparrow$ denotes the spin projection $+1/2$ and a $\downarrow$ denotes the spin projection $-1/2$ [30,31]:

$$\left|\frac{3}{2}, +\frac{3}{2}\right\rangle = \uparrow\uparrow\uparrow \tag{1.22}$$

$$\left|\frac{3}{2}, +\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left(\uparrow\uparrow\downarrow + \uparrow\downarrow\uparrow + \downarrow\uparrow\uparrow\right) \tag{1.23}$$

$$\left|\frac{3}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left(\downarrow\downarrow\uparrow + \downarrow\uparrow\downarrow + \uparrow\downarrow\downarrow\right) \tag{1.24}$$

$$\left|\frac{3}{2}, -\frac{3}{2}\right\rangle = \downarrow\downarrow\downarrow \tag{1.25}$$

which is a quadruplet of states symmetric under the exchange of any two quarks;

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = -\frac{1}{\sqrt{6}}\left(2\downarrow\downarrow\uparrow - \uparrow\downarrow\downarrow - \downarrow\uparrow\downarrow\right) \tag{1.26}$$

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}\left(2\uparrow\uparrow\downarrow - \downarrow\uparrow\uparrow - \uparrow\downarrow\uparrow\right) \tag{1.27}$$

which are states with a mixed symmetry, symmetric only for exchange of $1 \leftrightarrow 2$ but with no definite symmetry under interchange $1 \leftrightarrow 3$ and $2 \leftrightarrow 3$; and finally

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left(\uparrow\downarrow\downarrow - \downarrow\uparrow\downarrow\right) \tag{1.28}$$

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left(\uparrow\downarrow\uparrow - \downarrow\uparrow\uparrow\right) \tag{1.29}$$

are the states of a mixed symmetry, antisymmetric under interchange of $1 \leftrightarrow 2$ but with no definite symmetry under interchange $1 \leftrightarrow 3$ and $2 \leftrightarrow 3$.

It should be noted that the three-quark combination cannot provide a spin wave function antisymmetric under interchange of any two quarks.

For what concerns the flavour part of the wave function, the available combinations depend on the number of quarks accounted for. In the simplest case when only up and down quarks are considered, the flavour symmetry is usually referred to as *isospin* symmetry, and has the algebra identical to the spin symmetry described above, where the u replaces the $\uparrow$, and d replaces the $\downarrow$. However, the formulae become much more complicated when accounting for more quark flavours. To describe the baryons having one b-quark and two light (u, d, s) quarks, one has to consider the possible combinations of four quarks.

The requirement that the total baryon wave function must be antisymmetric poses a constraint on the spectrum of ground baryon states.

For example, let's consider the simplest three-quark combination uuu. Its flavour part is symmetric under exchange of any quarks, as all three quarks have an identical flavour. To satisfy the requirement of an antisymmetric total wave function, the spin part must be symmetric. This implies that the ground state uuu cannot have the spin $1/2$, as the corresponding spin wave functions have a mixed symmetry. Consequently, the ground uuu state has a spin $3/2$ which is symmetric, and this particle is known as Δ^{++}. In general, any ground state baryon composed of three identical quarks should have a spin $3/2$.

On the other hand, a state formed of up and down quarks, uud, can have a spin $1/2$. This is achieved when the combination of the both spin and flavour wave functions is symmetric, while spin and flavour parts alone have a mixed symmetry [32,33]:

$$\frac{1}{\sqrt{18}}\left[|uud\rangle\left(\uparrow\downarrow\uparrow+\downarrow\uparrow\uparrow-2\uparrow\uparrow\downarrow\right)+|udu\rangle\left(\uparrow\uparrow\downarrow+\downarrow\uparrow\uparrow-2\uparrow\downarrow\uparrow\right)\right.$$
$$\left.+|duu\rangle\left(\uparrow\downarrow\uparrow+\uparrow\uparrow\downarrow-2\downarrow\uparrow\uparrow\right)\right] \quad (1.30)$$

Such a wave function is symmetric under exchange of any quark pairs, while the spin of such a state is $1/2$. This state is the proton. It is also possible to construct the state of the spin $3/2$ with the same quark content, which will correspond to the Δ^{+} baryon. So, there are two ground state (*i.e.* neither orbitally nor radially excited) baryons of the quark content uud, one with a spin $1/2$ and one with a spin $3/2$. This conclusion can be generalised to any baryons having two identical quarks, and a different third one.

Following similar considerations, it can be shown that for the baryons consisting of three different quarks, there are three ground states, *two* with a spin $1/2$ and one with a spin $3/2$ [32]. For example, in the b-baryon sector, the quark content usb is shared by three ground state baryons: Ξ_b^0 and $\Xi_b^{0\prime}$ have a spin $1/2$, while the Ξ_b^{0*} baryon has a spin $3/2$.

Having the knowledge about the allowed baryon ground states, one can now categorize baryons into multiplets. In the original quark model proposed by Gell-Mann in 1964 [34], only the u, d and s were involved, leading to an SU(3) flavour symmetry. The SU(3) baryon multiplets are presented in Fig. 1.2, where all the members of the *baryon octet* have spin-parity $1/2^{+}$, while members of the *baryon decuplet* have $3/2^{+}$. It can be seen that grouping baryon states into the multiplets effectively classifies them by the electric charge, strangeness and projection of the isospin. It should also be noted that the SU(3) flavour symmetry is approximate, as the mass of the s-quark is about 90 MeV/c^2 larger than the masses of u and d quarks, as it was shown in Table 1.1. The mass of the baryon states is almost equal on the horizontal axis ($u \leftrightarrow d$), but changes with strangeness.

A natural extension of the SU(3) multiplets to the case of four quarks, needed to describe the known b-baryon states (composed of one b and two light quarks), is provided by the SU(4) symmetry, as presented in Fig. 1.3. The symmetry is badly broken due to the large mass of the b-quark. The picture is very similar for singly-charmed baryons, if the b-quark is replaced by a c-quark [6].

Finally, when discussing baryons, one should note that there is a different class of

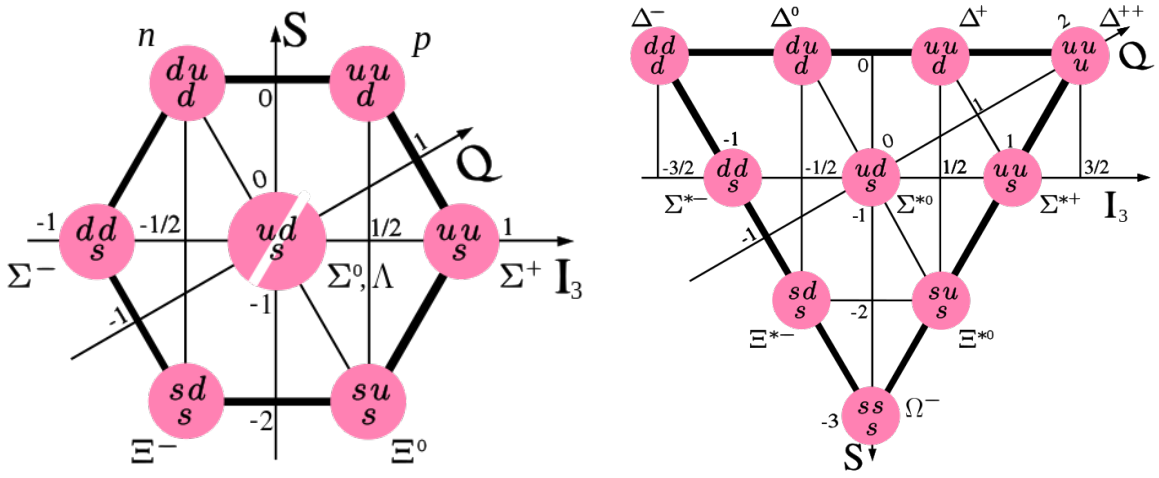


Figure 1.2: SU(3) multiplets of baryons made of u, d and s quarks. Left: SU(3) octet[1]. Right: SU(3) decuplet[2].

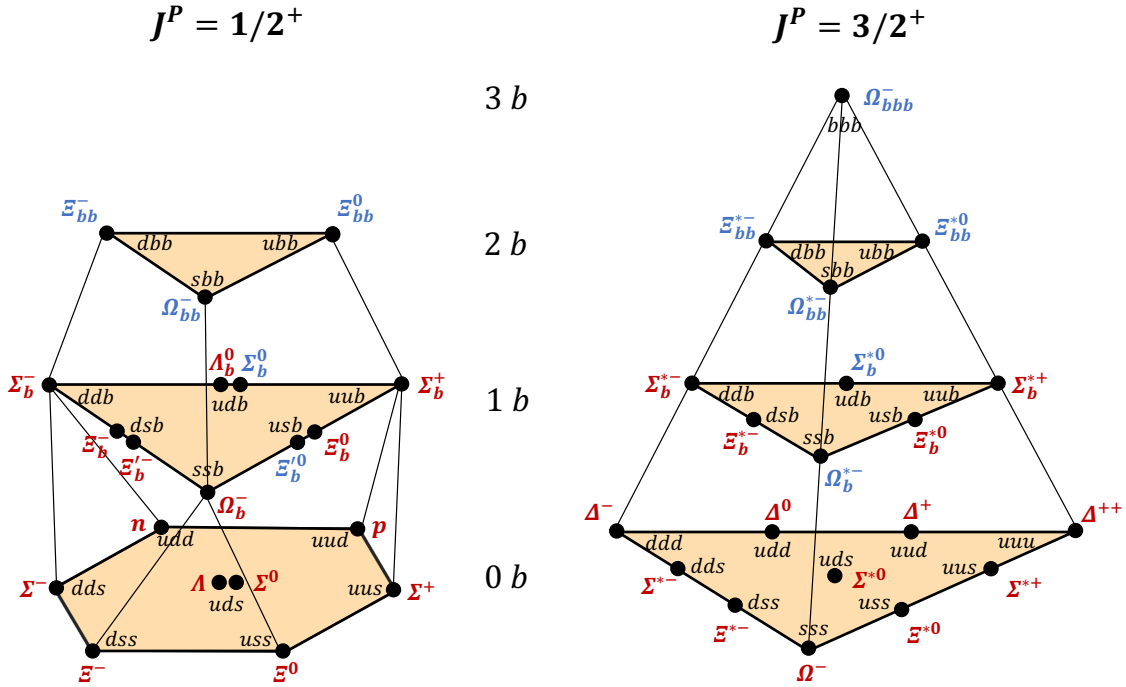


Figure 1.3: SU(4) multiplets of baryons made of u, d, s and b quarks. Left: the multiplet with an SU(3) octet on the lowest level. Right: the multiplet with an SU(3) decuplet on the lowest level. Observed states are shown in red, while predicted unobserved states – in blue. Prepared based on Ref. [6].

particles having non-zero baryon number: the pentaquarks.

[1]Source: https://commons.wikimedia.org/wiki/File:Baryon-octet.svg
[2]Source: https://commons.wikimedia.org/wiki/File:Baryon-decuplet-small.svg

The states of quark content $qqqq\bar{q}$ were predicted in the quark model [34]. Throughout decades, several unconfirmed claims for light pentaquark states were proposed. The landscape has changed recently with the observation of pentaquark candidates in the $\Lambda_b^0 \to pKJ/\psi$ channel by the LHCb Collaboration [4,5]. These states decay to the pJ/ψ final state, which implies their minimal quark content to be $uudc\bar{c}$. The exact internal structure of these candidates is not yet known. Among the models proposed, there are ideas of meson-baryon molecules [35,36], tightly-bound states [37,38], or kinematical "rescattering" effects [39].

1.2.1 b-baryon spectroscopy: experimental status

It is useful to make a correspondence between the states predicted by the quark model, and the ones actually observed in experiment. Table 1.3 presents the state of the art in the exploration of the lowest b-baryon states, which are ground states in the sense that they are neither orbitally nor radially excited. The first feature which can be figured out from this table is that there are only *four* b-baryons which decay weakly, the Λ_b^0, Ξ_b^0, Ξ_b^- and Ω_b^-, and among those there are no positively charged states. These four baryons have a lifetime of an order of picoseconds, which is typical for weakly-decaying b hadrons. Other, heavier states can decay either via strong or electromagnetic interaction to one of these four states: for example, the dominant decay mode of $\Sigma_b^{(*)\pm}$ baryons is the strong decay to the Λ_b^0 and a charged pion [40]. Four of such ground states, namely Σ_b^0, $\Xi_b'^0$, Σ_b^{*0} and Ω_b^{*-}, have not been observed to date. This is explained by the fact that, according to the predicted values of their masses [41], their only allowed decays are those to another b-baryon and a soft photon, or a soft π^0. Such decay modes are particularly difficult to observe at hadron colliders.

The lightest b-baryon, the Λ_b^0, was discovered in 1991 by the UA1 experiment [42] at the SPS at CERN, in the decay mode $\Lambda_b^0 \to \Lambda J/\psi$. LEP experiments contributed to the first studies of the properties of the Λ_b^0 baryon, such as its lifetime. The current experimental knowledge on the b-baryon decays is dominated by the studies performed at hadron colliders: Tevatron and LHC. b baryons cannot be produced at B-factories operating at $\Upsilon(4S)$ resonance, as the mass of $\Upsilon(4S)$ is too small to decay in a pair of b baryons. Consequently, a thorough exploration of their decay modes has started only recently. Profiting from the large production at LHCb (discussed in more detail in the Sec. 2.2), the Λ_b^0 baryon has been studied relatively well in the past few years. Other b-baryons have a much smaller production than the Λ_b^0, so a larger dataset is needed to overcome the statistical limitations.

Another b-baryon relevant for this thesis, the Ξ_b^0, was directly observed for the first time only in 2011 by the CDF collaboration [43], in the channel $\Xi_b^0 \to \Xi_c^+ \pi^-$. Since then, LHCb measured its mass and lifetime exploiting the same decay channel [44]. In addition [45], a new decay mode $\Xi_b^0 \to D^0 p K^-$ was observed by the LHCb experiment, and an evidence for the decay $\Xi_b^0 \to \Lambda_c^+ K^-$ was seen. Recently [46], the charmless four-body decay modes $\Xi_b^0 \to pK^-\pi^+\pi^-$ and $\Xi_b^0 \to pK^-\pi^+K^-$ were observed as well, and the latter was exploited to search for CP violation [47]. So, with only five decay channels well established, Ξ_b^0

Table 1.3: Properties of the lightest b-baryon states. Based on Ref. [6].

Name	Quark content	Mass, MeV/c^2	Lifetime or width	J^P	(I,I_3)	S
Λ_b^0	udb	5619.60 ± 0.17	1.470 ± 0.010 ps	$1/2^+$	$(0,0)$	0
Ξ_b^0	usb	5791.9 ± 0.5	1.479 ± 0.031 ps	$1/2^+$	$(1/2,1/2)$	-1
Ξ_b^-	dsb	5797.0 ± 0.9	1.571 ± 0.040 ps	$1/2^+$	$(1/2,-1/2)$	-1
Σ_b^+	uub	5810.55 ± 0.25	4.83 ± 0.48 MeV/c^2	$1/2^+$	$(1,1)$	0
Σ_b^0	udb	not yet observed		$1/2^+$	$(1,0)$	0
Σ_b^-	ddb	5815.64 ± 0.28	5.33 ± 0.56 MeV/c^2	$1/2^+$	$(1,-1)$	0
$\Xi_b'^0$	usb	not yet observed		$1/2^+$	$(1/2,1/2)$	-1
$\Xi_b'^-$	dsb	5935.02 ± 0.05	< 0.08 MeV/c^2	$1/2^+$	$(1/2,-1/2)$	-1
Ω_b^-	ssb	6046.1 ± 1.7	1.64 ± 0.18 ps	$1/2^+$	$(0,0)$	-2
Σ_b^{*+}	uub	5830.28 ± 0.28	9.34 ± 0.54 MeV/c^2	$3/2^+$	$(1,1)$	0
Σ_b^{*0}	udb	not yet observed		$3/2^+$	$(1,0)$	0
Σ_b^{*-}	ddb	5834.73 ± 0.30	10.68 ± 0.68 MeV/c^2	$3/2^+$	$(1,-1)$	0
Ξ_b^{*0}	usb	5949.8 ± 1.4	0.90 ± 0.18 MeV/c^2	$3/2^+$	$(1/2,1/2)$	-1
Ξ_b^{*-}	dsb	5955.33 ± 0.13	1.65 ± 0.33 MeV/c^2	$3/2^+$	$(1/2,-1/2)$	-1
Ω_b^{*-}	ssb	not yet observed		$3/2^+$	$(0,0)$	-2

is much less well studied than the Λ_b^0 baryon. The production of Ξ_b^0 has never been measured. To date, Ξ_b^0 is the only of the four weakly-decaying b-baryons, which has no established decay modes involving the J/ψ meson in the final state. This can be explained by the fact that potentially the dominant of such decay modes, $\Xi_b^0 \to J/\psi \Xi^0$, has a Ξ^0 in the final state which is a long-lived neutral baryon decaying in about 99.5% cases to a long-lived neutral Λ and a soft π^0 [6], making this final state difficult to reconstruct at the hadron colliders. A part of this thesis is dedicated to the first observation and study of the decay mode $\Xi_b^0 \to pKJ/\psi$.

1.3 Rare $b \to s\ell^+\ell^-$ decays as probes of New Physics

One of the promising channels for NP searches is the $b \to s\ell^+\ell^-$ transition. This process has several reasons to be an important probe of the SM. It is a flavour-changing neutral current (FCNC), arising only at the loop level in the SM (the dominant contributing diagrams are presented in Fig. 1.4), making the SM contribution very small. This results in very small branching ratios of FCNC decays: for example, $\mathcal{B}(K_L \to \mu^+\mu^-) = (6.84 \pm 0.11) \times 10^{-9}$, or $\mathcal{B}(B^+ \to K^+\mu^+\mu^-) = (4.41 \pm 0.23) \times 10^{-7}$ [6]. This allows to spot even tiny NP effects, which might enter via new tree-level diagrams, or via new particles involved in the loops. In addition, the b-quark mass is much larger than the QCD scale Λ_{QCD}, thus allowing for a precise SM predictions for a selection of NP-sensitive observables. Yet, the b-quark mass is still much smaller than the masses of the electroweak

bosons and of the top quark, which allows the construction of an effective low-energy theory describing this transition, by separating the two mass scales.

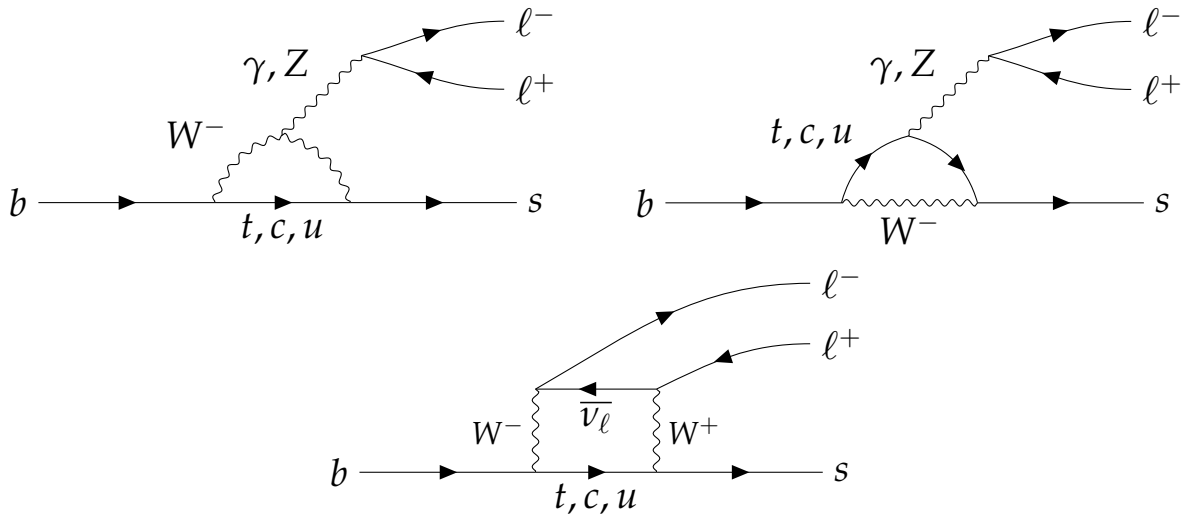


Figure 1.4: Feynman diagrams[3]contributing to the $b \to s\ell^+\ell^-$ transition.

1.3.1 Theoretical description

The basic framework for the theoretical description of weak decays of b hadrons is the effective field theory (EFT) [49–51], relevant for scales $\mu \approx \mathcal{O}(1-5)$ GeV, which is much smaller than m_W, m_Z, m_t. It represents a generalisation of the Fermi theory of weak interactions, and allows to handle simultaneously the (high-energy, or "short-distance") electroweak and (low-energy, or "long-distance") QCD effects.

An amplitude of a decay of a given particle M to a final state F is given by

$$\mathcal{A}(M \to F) = \langle F|\mathcal{H}_{eff}|M\rangle\,, \tag{1.31}$$

where $\mathcal{H}_{eff}$ is the relevant Hamiltonian.

The SM usually provides several ways (diagrams) for a given $M \to F$ transition to occur, which can be governed by different interactions. Therefore, the most general Hamiltonian for the case of b-hadron decays involves several operators $\mathcal{O}_i$ with various structures, each multiplied by a perturbatively calculable coefficient $\mathcal{C}_i$:

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{CKM} \sum_i \mathcal{C}_i(\mu)\mathcal{O}_i, \tag{1.32}$$

where G_F is the Fermi constant, and V_{CKM} is the set of relevant CKM factors describing the transition of interest. The whole procedure is commonly called "Operator Product Expansion" (OPE). The operators $\mathcal{O}_i$ describe the low-energy physics below the energy

[3]All the diagrams in this work are drawn using the `tikz-feynman` tool described in Ref. [48].

scale μ. As these operators describe four-fermion vertices, they have dimension 6. The coefficients $\mathcal{C}_i$ are called Wilson coefficients and cover the high-energy part, above the scale μ. The values of these coefficients depend on the scale μ. They can be evaluated at the characteristic scale of the weak interaction (m_W) by matching the effective theory to the full SM theory. At this scale, QCD corrections are small and can be calculated perturbatively. Then, the coefficients are propagated down to the scale μ using the renormalisation group equations.

Hadronic matrix elements, which allow to pass from the description of the transition at the quark level, to the description at the meson (baryon) level, are parametrised in terms of *form factors*, which are in general q^2-dependent. Their expressions can be predicted from non-perturbative QCD methods: lattice QCD at high values of q^2, or light cone sum rules at low values of q^2 [52]. These predictions have theoretical uncertainties, which have to be taken into account when predicting the differential decay rate and related observables.

For the $b \to s\ell^+\ell^-$ transitions, the relevant CKM factors can be deduced from the Feynman diagrams presented in Fig. 1.4: $V_{tb}V_{ts}^*$, $V_{cb}V_{cs}^*$ and $V_{ub}V_{us}^*$, the latter being Cabibbo-suppressed and so considered as a higher-order correction. Out of the remaining two, the contributions proportional to $V_{tb}V_{ts}^*$ are related to the processes above the scale μ (as the top quark is heavy), while the ones proportional to $V_{cb}V_{cs}^*$ are related the scale below μ. This results in a different treatment of the two. The CKM unitarity relations are used to express the contributions proportional to $V_{cb}V_{cs}^*$ via the other products of the CKM elements, leading to four-quark operators proportional to $(\bar{c}b)(\bar{s}c)$ (usually referred to as $\mathcal{O}_2$). These contributions receive non-factorisable QCD corrections that cannot be absorbed into form factors [53], and are another important source of theoretical uncertainties.

Consequently, for the $b \to s\ell^+\ell^-$ transitions, the effective Hamiltonian takes the form

$$\mathcal{H}_{eff}(b \to s\ell^+\ell^-) = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_i \mathcal{C}_i \mathcal{O}_i. \tag{1.33}$$

Operators $\mathcal{O}_1 ... \mathcal{O}_6$ are four-quark operators, classified into current-current tree-level W exchange operators (1-2) and QCD penguins mediated by gluons (3-6). Operators $\mathcal{O}_7 ... \mathcal{O}_{10}$ are electroweak penguin operators, classified in electro- and chromo-magnetic penguins (7-8), and semileptonic operators (9-10). Three of them are the most relevant for the description of the $b \to s\ell^+\ell^-$ transition: $\mathcal{O}_7$ describing the $b \to s\gamma$ transition with an on-shell photon (Fig. 1.5(a)), and $\mathcal{O}_9$, $\mathcal{O}_{10}$ are the vectorial and axial operators describing the $b \to s\ell^+\ell^-$ transition itself (Fig. 1.5(b)).

Their expressions are [10]:

$$\mathcal{O}_7 = \frac{e}{16\pi^2} m_b \left(\bar{s}\sigma_{\mu\nu}P_R b\right) F^{\mu\nu}; \tag{1.34}$$

$$\mathcal{O}_9 = \frac{e^2}{16\pi^2} \left(\bar{s}\gamma_\mu P_L b\right) \left(\bar{\ell}\gamma^\mu \ell\right); \tag{1.35}$$

$$\mathcal{O}_{10} = \frac{e^2}{16\pi^2} \left(\bar{s}\gamma_\mu P_L b\right) \left(\bar{\ell}\gamma^\mu\gamma_5 \ell\right); \tag{1.36}$$

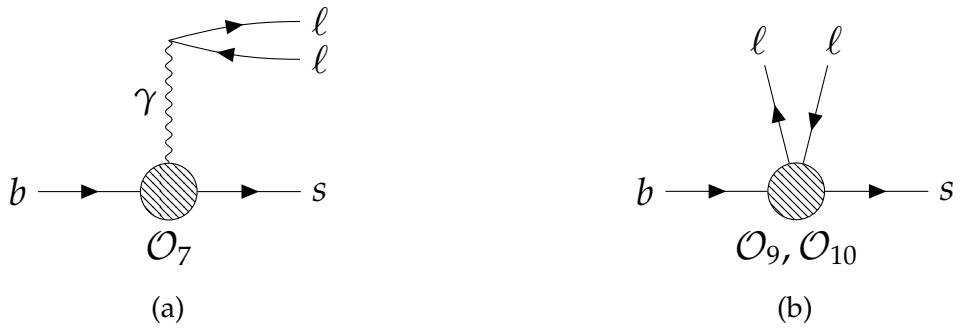


Figure 1.5: Diagrams for the radiative $b \to s\gamma$ and semileptonic $b \to s\ell^+\ell^-$ operators. Short-distance contributions are hidden in the vertex.

where $P_{R,L} = \frac{1\pm\gamma_5}{2}$ are the right- and left-handed chiral projectors.

In addition, there are operators which are very suppressed in the SM, but can be important in various NP models. Those include scalar, pseudoscalar and tensor operators $\mathcal{O}_S$, $\mathcal{O}_P$ and $\mathcal{O}_{T(5)}$. Finally, there are right-handed operators $\mathcal{O}'_{7,9,10}$ which differ from the $\mathcal{O}_{7,9,10}$ by replacing $P_L \leftrightarrow P_R$. They are suppressed in the SM as the weak interaction is left-handed, but can become important in the NP models having a different helicity structure.

1.3.2 Phenomenology of the $b \to s\ell^+\ell^-$ transitions

As free quarks cannot be studied, discussion of the $b \to s\ell^+\ell^-$ transition implies hadronisation to a meson or a baryon transition, for example $B^+ \to K^+\ell^+\ell^-$ or $\Lambda_b^0 \to pK\ell^+\ell^-$. These decays proceed through FCNC and are therefore sensitive to potential NP contributions.

The transferred momentum in the $b \to s$ transition, usually denoted as q^2, is an important parameter for both experimental and theoretical descriptions. Technically, it is equal to the dilepton invariant mass squared. A sketch of a typical q^2 distribution in the $b \to s\ell^+\ell^-$ transitions is presented in Fig.1.6. Two prominent resonant peaks can be seen, and they correspond to the tree-level $b \to c\bar{c}s$ transitions of b-hadrons to charmonia (J/ψ, $\psi(2S)$), with subsequent charmonium decays to a dilepton pair. Charmonium peaks are centered around q^2 values of about 9.6 GeV^2/c^4 (J/ψ) and 13.6 GeV^2/c^4 ($\psi(2S)$). These contributions are not FCNC, but they cannot be separated experimentally from the $b \to s\ell^+\ell^-$ transitions as they share the same initial and final states, and charmonium decays promptly.

At very low q^2, the enhancement in rate is due to the photon pole. This region is dominated by the $b \to s\gamma$ transition with the nearly-on-shell photon, and so is described by the Wilson coefficient $\mathcal{C}_7$. There is no such enhancement for B decays to pseudoscalar mesons, *e.g.* $B^+ \to K^+\ell^+\ell^-$, because the $B^+ \to K^+\gamma$ transition is spin-violating. The region around q^2 =1.05 GeV^2/c^4 is polluted by the narrow $\phi(1020)$ resonance. However, due to the small value of $\mathcal{B}(\phi \to \ell^+\ell^-) \approx 2.9 \times 10^{-4}$ [6], the contribution of the $\phi(1020)$ resonance is almost negligible – this is the reason why it is not visible on the sketch. The

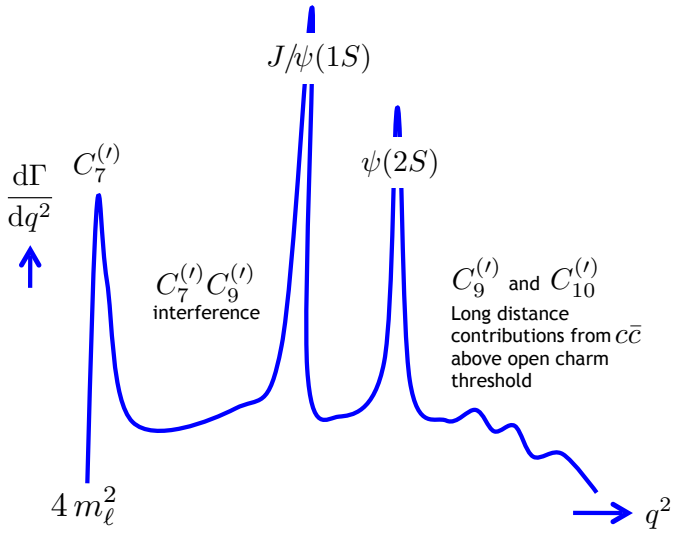


Figure 1.6: Differential rate as a function of q^2 in $b \to s l^+ l^-$ transitions, for the example of the $B^0 \to K^{*0}\ell^+\ell^-$ decay.

q^2 region below the ϕ, dominated by the $\mathcal{C}_7$ Wilson coefficient, is usually called the `low`-q^2 region. The region between the ϕ and J/ψ– the `central`-q^2 – is characterised by the interference of the coefficients $\mathcal{C}_7$, $\mathcal{C}_9$ and $\mathcal{C}_{10}$. Finally, the region above $\psi(2S)$ is called the `high`-q^2, or the region of low hadronic recoil. The latter is polluted by broad charmonium states (such as the $\psi(3770)$) which have fairly small branching fractions of their decays to dileptons. An interesting feature of this region from the theoretical point of view is that the energy of the hadron is smaller than the energy scale of QCD interactions Λ_{QCD}. This allows to handle the theoretical calculations in this region by working in the heavy quark limit ($m_b \to \infty$).

1.3.3 Experimental status

The major classes of analyses performed in $b \to s\ell^+\ell^-$ transitions are LU tests, differential branching fraction measurements, and angular analyses.

1.3.3.a Differential branching fraction measurements

The LHCb Collaboration has performed the measurements of differential branching fractions of $B \to K\mu^+\mu^-$, $B \to K^*\mu^+\mu^-$, $B_s^0 \to \phi\mu^+\mu^-$, $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays [54–57] mediated by the $b \to s\mu^+\mu^-$ transition. As it can be seen from a selection of these measurements presented in Fig. 1.7, they show some deficit of events in the central-q^2 region with respect to the SM predictions. However, the size of the theoretical uncertainties due to hadronic form factors is quite large. Another particular issue is the size of non-factorisable effects, such as quark loop contributions ($b \to s(q\bar{q})$), when $q = c$ ("charm loops").

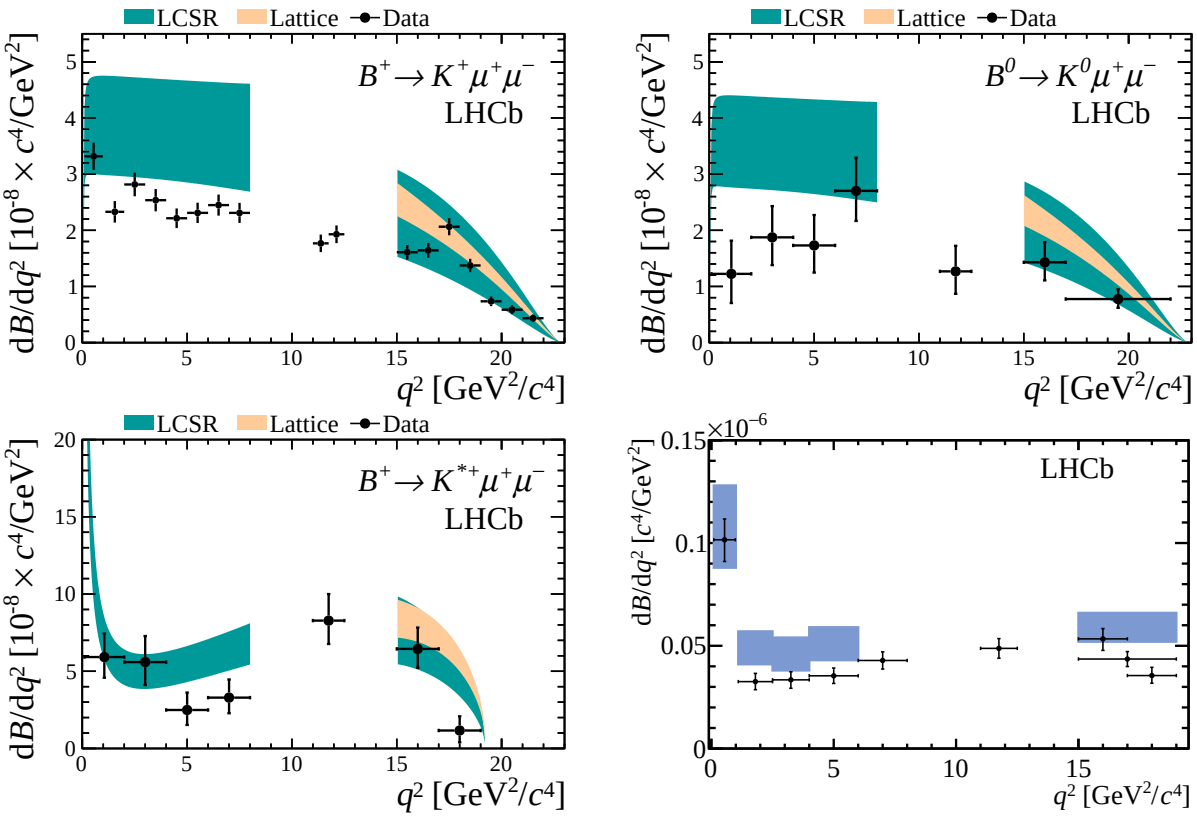


Figure 1.7: Differential branching fractions for $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^0_S\ell^+\ell^-$, $B^+ \to K^{*+}\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays, compared to the relevant theoretical predictions. Taken from Refs. [54, 56].

1.3.3.b Lepton universality tests

To overcome the issue of the large theoretical uncertainties in the branching fraction predictions, ratios of branching fractions can be constructed, sharing the same hadronic part but comparing different leptons in the final state. The measurement of such ratios provides a test of LU.

The LU ratio [10, 58], R_X, is in general defined as

$$R_{H_s} = \frac{\int \frac{d\Gamma(H_b \to H_s \mu^+\mu^-)}{dq^2}}{\int \frac{d\Gamma(H_b \to H_s e^+e^-)}{dq^2}}, \tag{1.37}$$

where H_b is the b hadron decaying into the s hadron H_s and a dilepton pair; and the integration is performed over a given q^2 region which is identical for muon and electron mode. There are several possible hadronisation options of the b-quark. To date, LU tests have only been performed in decays of B-mesons to a strange meson (K or K^*) and a dilepton pair. This approach allows to cancel the QCD contributions such as knowledge of the form-factors and resulting uncertainties in the theoretical predictions, bringing the SM prediction close to unity within an uncertainty at the percent level [58, 59]. At very low-q^2, well below $1\,\text{GeV}^2/c^4$, the lepton mass effects become important (due to

the difference in allowed phase-space leading to rapid lepton non-universal variation of $\frac{d\Gamma}{dq^2}$ close to the dimuon kinematic threshold, as well as due to light-hadron effects described in Ref. [59]), so the theoretical prediction can be by few percent smaller that exactly one [60–62]. However, this effect is only about two percent, if the region below $0.1\,\mathrm{GeV}^2/c^4$, where the dimuon mass threshold implies rapid and flavour-non-universal variation of the differential branching fraction, is not considered in the measurement [59]. The summary of some prominent theoretical predictions for the R_K and $R_{K^{*0}}$ ratios is presented in Tab. 1.4.

Table 1.4: Theoretical predictions for R_K and $R_{K^{*0}}$ observables.

Observable	Ref. [59]	Ref. [60]	Ref. [61]
R_K $(1 < q^2 < 6\ \mathrm{GeV}^2/c^4)$	1.00 ± 0.01	1.00 ± 0.01	$1.0004^{+0.0008}_{-0.0007}$
$R_{K^{*0}}$ $(0.045 < q^2 < 1.1\ \mathrm{GeV}^2/c^4)$	0.906 ± 0.028	0.92 ± 0.02	$0.920^{+0.007}_{-0.006}$
$R_{K^{*0}}$ $(1.15 < q^2 < 6\ \mathrm{GeV}^2/c^4)$	1.00 ± 0.01	1.00 ± 0.01	$0.996^{+0.002}_{-0.002}$

The formulation of the LU test as a ratio also allows to cancel the experimental effects related to the hadronic part of the decay, b-hadron production, luminosity measurement, *etc.*.

As it has already been mentioned, all the LU measurements to date were performed by the LHCb experiment in the regions of q^2 below the charmonia region: low q^2 below the $\phi(1020)$ region, and the central q^2 between $\phi(1020)$ and J/ψ resonances. The region with larger q^2 are experimentally challenging to probe, due to the complicated description of backgrounds.

The LHCb collaboration has reported two measurements: R_K and $R_{K^{*0}}$, in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays, respectively. These two complementary measurements allow to probe the LU in cases where the strange hadron spin is either zero (K) or one (K^{*0}). The latest R_K result [1,63] is

$$R_K = 0.846^{+0.060+0.016}_{-0.054-0.014}, \tag{1.38}$$

for $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$, using a dataset of integrated luminosity about 5 fb^{-1}. The ratio $R_{K^{*0}}$ was found to be equal to

$$R_{K^{*0}} = \begin{cases} 0.66^{+0.11}_{-0.07}\,(\mathrm{stat}) \pm 0.03\,(\mathrm{syst}) & \text{for } 0.045 < q^2 < 1.1\,\mathrm{GeV}^2/c^4 \\ 0.69^{+0.11}_{-0.07}\,(\mathrm{stat}) \pm 0.05\,(\mathrm{syst}) & \text{for } 1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4, \end{cases} \tag{1.39}$$

using about 3 fb^{-1} of LHCb data [2].

These measurements are in some tension with the SM predictions, at a level of 2.5 standard deviations for R_K; and 2.1-2.4 standard deviations for $R_{K^{*0}}$. Similar measurements have been performed by Belle and BaBar collaborations [64–67], but they have significantly larger statistical uncertainties.

1.3.3.c Angular analyses

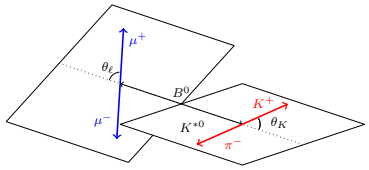
NP contributions could also be identified by their modification of the angular distributions of the $B^0 \to K^{*0}\ell^+\ell^-$ decay products, where K^{*0} denotes the $K^{*0}(892)$ meson. Profiting from the vectorial nature of the K^{*0} resonance ($J^P = 1^-$) decaying to $K^+\pi^-$, the decay $B^0 \to K^{*0}\ell^+\ell^-$ offers a rich set of observables which can be measured experimentally. The differential decay rate of the $B^0 \to K^{*0}\ell^+\ell^-$ decay can be expressed in terms of three helicity angles $\theta_K, \theta_\ell, \phi$ which are presented in Fig. 1.8. The angle θ_ℓ is defined as the angle between the direction of the ℓ^+ (ℓ^-) and the direction opposite to that of the B^0 ($\bar{B}^0$) in the dilepton rest frame. The angle θ_K is defined as the angle between the direction of the kaon and the direction opposite to that of the B^0 ($\bar{B}^0$) in in the K^{*0} ($\bar{K}^{*0}$) rest frame. The angle ϕ is the angle between the plane containing the two leptons, and the plane containing the kaon and pion, in the B^0 ($\bar{B}^0$) rest frame. The basis is designed such that the angular definition for the $\bar{B}^0$ decay is a CP transformation of that for the B^0 decay. The exact expression of the differential decay rate can be found in Ref. [68]:

$$\begin{aligned}\frac{d^4\Gamma}{dq^2 d\cos\theta_K d\cos\theta_\ell d\phi} = \frac{9}{32\pi}\Big[& I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + \\ & I_2^s \sin^2\theta_K \cos 2\theta_\ell + I_2^c \cos^2\theta_K \cos 2\theta_\ell + \\ & I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_\ell \cos\phi + \\ & I_5 \sin 2\theta_K \sin\theta_\ell \cos\phi + I_6 \sin^2\theta_K \cos\theta_\ell + \\ & I_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + I_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi + \\ & I_9 \sin^2\theta_K \sin^2\theta_\ell \sin 2\phi\Big],\end{aligned} \tag{1.40}$$

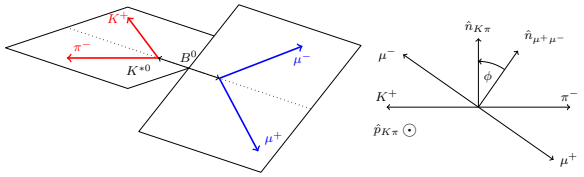
where combinations of helicity angles enter with eleven q^2-dependent coefficients $I_i(q^2)$, which are called *moments of the angular distribution*, and are bilinear combinations of K^{*0} decay amplitudes. Combining observables for B^0 and $\bar{B}^0$ decays, one can construct CP averages $S_i = (I_i + \bar{I}_i)/\left(\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}\right)$, or CP asymmetries $A_i = (I_i - \bar{I}_i)/\left(\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}\right)$ (where the bar over observable means it is the observable for the $\bar{B}^0$ decay). A normalisation of these observables by a differential decay rate allows to reduce dependencies on form factors. In the limit $q^2 > 4m_\mu^2$, CP averages $S_{1,2}^{s,c}$ are related to the fraction of longitudinal polarisation of the K^{*0} meson, F_L.

An optimised set of angular observables, which are 'clean' from the theoretical point of view due to being largely free from form factor uncertainties, was proposed in Refs. [69,70]. They are normalised in a way that form factors cancel exactly in the heavy quark limit. One of these observables, defined as $P_5' = \frac{S_5}{\sqrt{F_L(1-F_L)}}$, has attracted interest in recent years due to some tension between the experimental distribution and the theoretical predictions, notably also observed in central-q^2 region. Results from several experiments are presented in Fig. 1.9. [71–74].

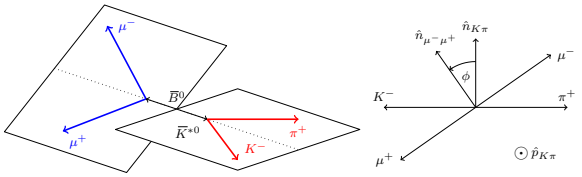
As it has been mentioned above, there are some non-factorisable QCD contributions, which could potentially mimic NP effects in the differential decay rate. These contributions



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 1.8: Definition of angles θ_L, θ_K and ϕ. Taken from Ref. [68].

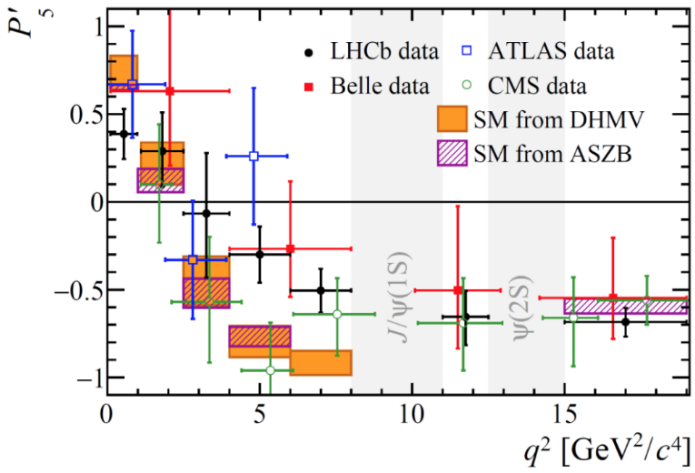


Figure 1.9: Measurements of P'_5 by different experiments, compared to the theoretical predictions [71–74].

do not cancel out in observables like P'_5. However, an important property of these contributions is that they are lepton flavor-universal, and cannot mimic NP effects which violate LU. In the light of possible hints of lepton non-universality described above, it can be interesting to perform a lepton-flavour-dependent angular analysis of $B^0 \to K^{*0}\ell^+\ell^-$

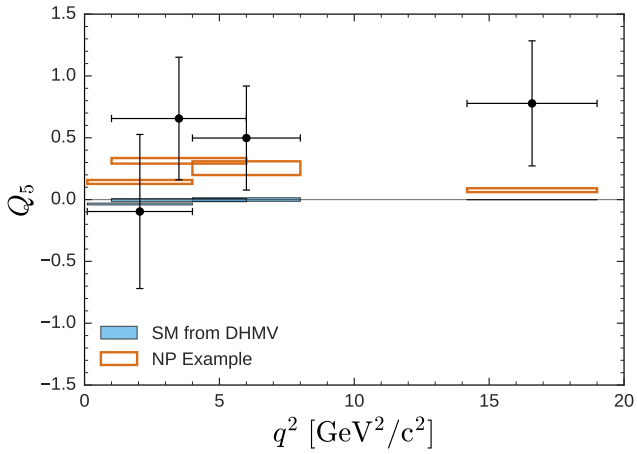


Figure 1.10: Measurement of Q_5 by the Belle collaboration [74]. Black points represent Belle data in different (overlapping) q^2 bins. They are to be compared with the SM predictions shown in blue, which is very close to zero for all q^2 regions.

decays, with $\ell^+\ell^- = e^+e^-$, $\mu^+\mu^-$. A set of new observables defined as a difference of angular observables for muon and electron modes, can be defined [75]. For example, a significant deviation from zero in the observable $Q_5 = P'_{5\mu} - P'_{5e}$ would be a direct hint of New Physics. Such as measurement has been to date performed only by the Belle Collaboration [74], and is statistically limited, as it can be seen in Fig. 1.10.

1.3.3.d Interpretations

Global fits [76–80] have been performed by different groups to the large set of observables in $b \to s\ell^+\ell^-$, $b \to s\gamma$ and $B_{(s)} \to \ell^+\ell^-$ transitions. The combined data shows a preference for NP contributions in the Wilson coefficients in the $b \to s\mu^+\mu^-$ transition, with a combined significance greatly exceeding 3σ. Various plausible scenarios can introduce NP contributions to $\mathcal{C}_9$ and $\mathcal{C}_{10}$ coefficients, but can also allow for right-handed $\mathcal{C}'_9$ and $\mathcal{C}'_{10}$ coefficients. These NP contributions can be different for the muon and electron channels, in case LU is violated.

There is a number of NP models trying to explain the observed deviations. However, direct searches for NP signatures, non-observation of proton decay, high-transverse-momentum observables in global-purpose detectors, searches for lepton-flavour-violating decays, and various indirect constraints from flavour measurements (such as the rate of B^0_s oscillations) impose quite stringent limits on these models. Out of the models which pass the current constraints, one notable proposed solution is introducing the vector leptoquark $U_1(3,1,2/3)$ of mass in the TeV range [81,82].

1.4 $\Lambda_b^0 \to pK\ell^+\ell^-$ decays as a probe of lepton universality

Decays $\Lambda_b^0 \to pK\ell^+\ell^-$, where $\ell^\pm$ are muons and electrons, are the baryonic analogs of the decays $B^0 \to K^{*0}\ell^+\ell^-$. Until now, LU tests at LHCb have only been conducted with b-mesons [1,2,63]. Given that large amounts of b-baryons are also produced at the LHC (as it will be discussed in Sec. 2.2), the decays $\Lambda_b^0 \to pK\ell^+\ell^-$ are suitable to perform the first test of LU in the baryonic sector. While to date there are no theoretical predictions for LU with $\Lambda_b^0 \to pK\ell^+\ell^-$ transitions, based on the acquired knowledge from the b-mesons there are no phenomenological reasons to believe that the ratio should be different than unity in the SM. Same as in the b-meson case, this statement is true in the q^2 region located sufficiently far from the invariant mass thresholds, so that the effects due to the difference in mass between muons and electrons are negligible. Testing LU in baryonic decays would allow to verify the presence of possible deviations from the SM prediction in the system of a different spin-structure. Therefore, this work can be considered as a null test for the SM.

In the absence of precise measurements from LEP or from the B-factories, there are significant uncertainties on the absolute branching fractions of all Λ_b^0 decays, coming from the uncertainty on the relative Λ_b^0 production rate. On the theory side, calculating form factors for baryons is more complicated than for mesons, making the theoretical uncertainties on the predicted branching fractions quite large as well. In practice, this implies that the absolute branching fractions of Λ_b^0 decays are less precise observables for comparisons to the SM predictions, compared to the LU tests.

The decay mode $\Lambda_b^0 \to pK\mu^+\mu^-$ has been observed by the LHCb experiment [3], however its branching fraction measurement was not reported. The decay channel $\Lambda_b^0 \to pKe^+e^-$ has never been observed. Therefore, in addition to the LU test, this work also covers the first observation of the $\Lambda_b^0 \to pKe^+e^-$ mode, and the measurement of the branching fractions of the $\Lambda_b^0 \to pKe^+e^-$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ decay modes.

The dominant Feynman diagrams contributing to the $\Lambda_b^0 \to pK\ell^+\ell^-$ decay are presented in Fig. 1.11, and include so-called "penguin" (top row) and "box" (bottom row) diagrams.

Multiple resonances, such as $\Lambda^*(1520), \Lambda^*(1860)$ *etc.*, populate the pK spectrum of the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays. This rich structure has been extensively studied and characterised in details by the LHCb Collaboration in the q^2 region corresponding to the J/ψ meson [4], as it is shown in Fig. 1.12. For the time being, there is no characterisation of this spectrum outside of the J/ψ region. Nevertheless, a background-subtracted distribution of $m(pK)$ in the non-resonant $\Lambda_b^0 \to pK\mu^+\mu^-$ decay can be found in the supplemental material of the paper studying the CP violation in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode [3], and is reproduced in Fig. 1.13.

To date, the theoretical exploration of the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays is very limited, which is partially explained by the complexity of the pK resonant structure. Several authors studied this decay in the region of the prominent $\Lambda^*(1520)$ resonance decaying to pK [83,84]. The latter work predicts the dilepton invariant mass distribution and several angular

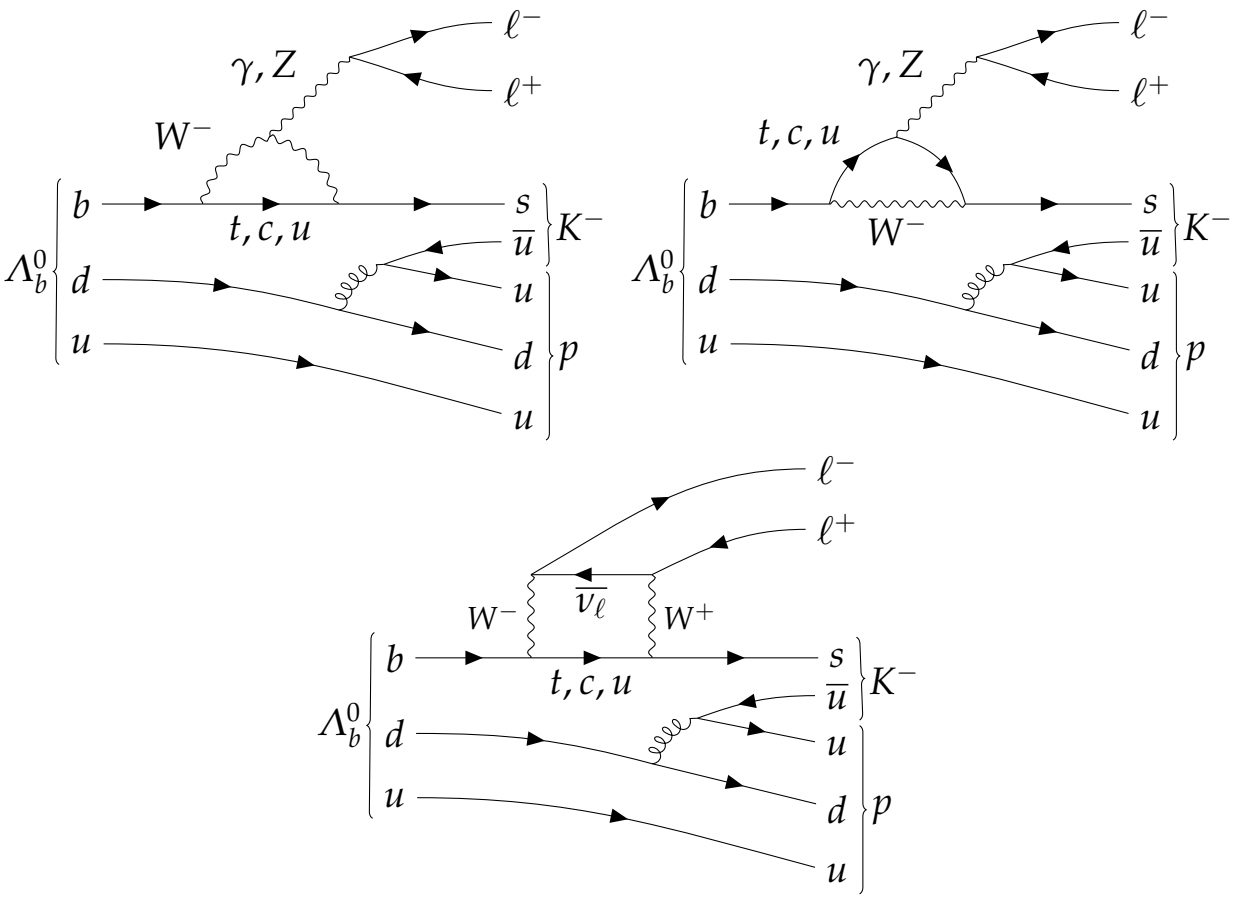


Figure 1.11: Feynman diagrams contributing to the decay $\Lambda_b^0 \to pK\ell^+\ell^-$.

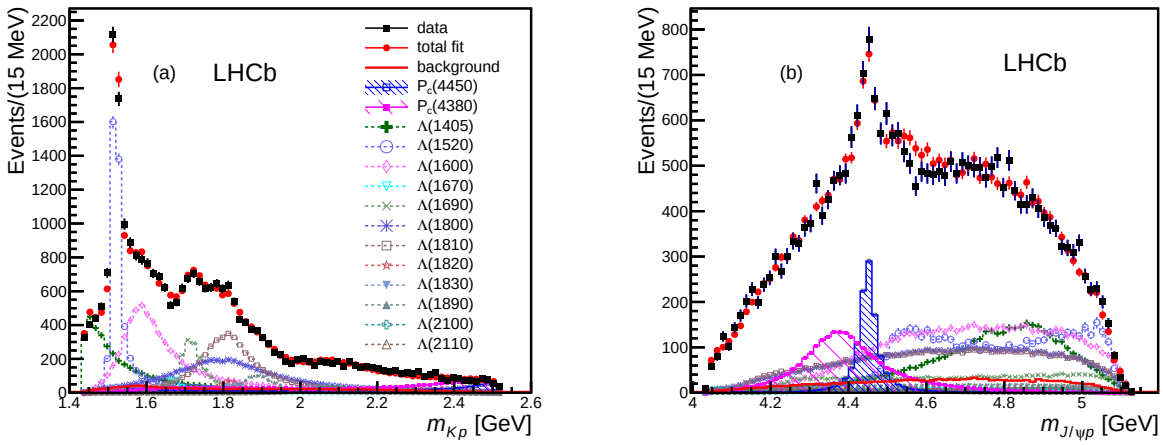


Figure 1.12: Invariant mass of $m(pK)$ and $m(pJ/\psi)$ in the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ data, with the dominant intermediate resonances. Taken from Ref. [4].

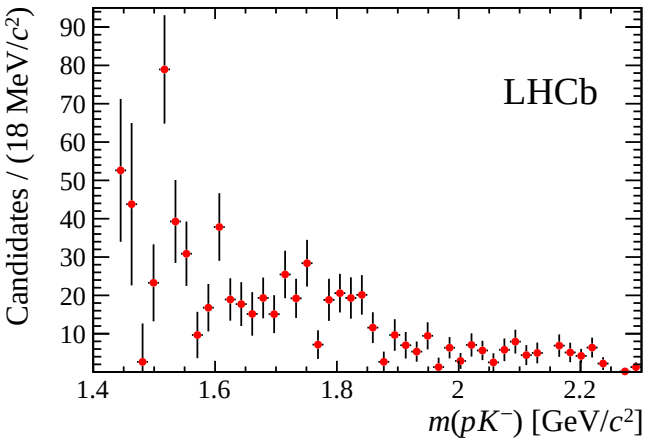


Figure 1.13: Invariant mass of $m(pK)$ in the $\Lambda_b^0 \to pK\mu^+\mu^-$ data. Taken from Ref. [3].

observables in the $\Lambda_b^0 \to \Lambda^*(1520)\ell^+\ell^-$ decay, showing the potential to discriminate the SM scenario from various NP models. It should be noted that a similar study is more complicated for the broader higher-mass resonances, as the narrow-width approximation (treating the Λ^* as a QCD-stable object) is not valid anymore. However, the authors of Ref. [85] perform a similar study for several prominent Λ^* resonances, and predict the $\frac{d\Gamma}{dq^2}$ distribution and the branching fraction of the $\Lambda_b^0 \to \Lambda^*\mu^+\mu^-$ decay for each resonance. In particular, it can be seen that the q^2 distribution changes depending on the mass and spin of the final state resonance (Fig. 1.14). No such study exists to date for the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays averaged over the complete $m(pK)$ spectrum.

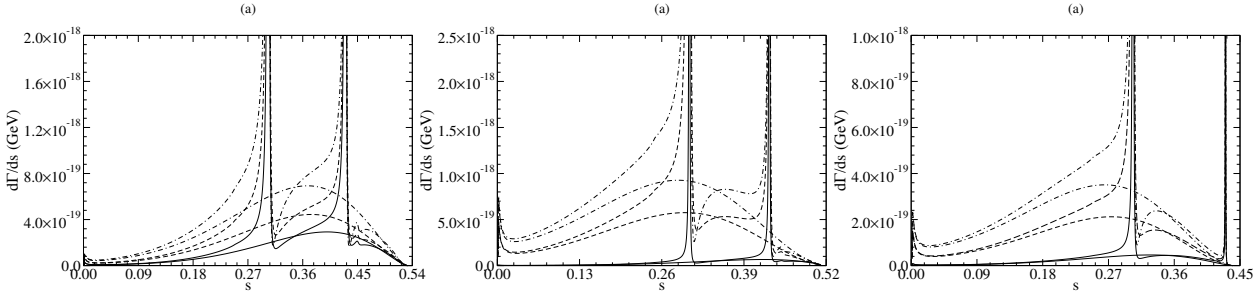


Figure 1.14: Predicted differential decay rate as a function of q^2 for different Λ^* resonances: (left) $\Lambda^*(1520)$ ($J^P = 3/2^-$), (centre) $\Lambda^*(1600)$ ($J^P = 1/2^+$) and (right) $\Lambda^*(1890)$ ($J^P = 3/2^+$). Taken from Ref. [85].

1.4.1 Experimental overview of the knowledge on relevant Λ_b^0 baryon decays

A thorough exploration of Λ_b^0 decay modes has only started recently, and many important decays are yet to be measured. The decay modes presenting a major interest for this

analysis are those having the $pK\ell^+\ell^-$ final state. This can be the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays themselves, but also decays $\Lambda_b^0 \to pKX$ with $X \to \ell^+\ell^-$. In this case, X can be a charmonium state (J/ψ, $\psi(2S)$). It can also be a photon undergoing a conversion to a dielectron pair. Out of these relevant decay modes, only some have been measured precisely. Table 1.5 provides a summary of the state-of-the-art in this field.

Table 1.5: Status of the decay modes of interest of the Λ_b^0 baryon.

Decay mode	Published observation	BR measured	Refs.
$\Lambda_b^0 \to pK\mu^+\mu^-$	Yes	No	[3]
$\Lambda_b^0 \to pKe^+e^-$	No	No	N/A
$\Lambda_b^0 \to pKJ/\psi$	Yes	Yes	[86]
$\Lambda_b^0 \to pK\psi(2S)$	Yes	Yes	[87]
$\Lambda_b^0 \to pK\gamma$	No	No	N/A

1.4.2 Comparison of b-baryon and b-meson decays

For the LU analysis presented in this thesis, in order to estimate the expected signal yields and sensitivities, it is important to understand not only the production ratio between the Λ_b^0 and B^0, but also to have a qualitative grasp of the relevant branching fractions. Based on the experimental results [6], it can be concluded that in the baryon world the branching fractions for the decays $\Lambda_b^0 \to pK\psi$, where ψ is a charmonium state, are in general lower than for their meson counterparts. Indeed, the $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ is measured to be $(3.17 \pm 0.04^{+0.57}_{-0.45}) \times 10^{-4}$ which is about four times smaller than the $\mathcal{B}(B^0 \to K^{*0}J/\psi) = (1.27 \pm 0.05) \times 10^{-3}$. The situation is even more unbalanced for the $\psi(2S)$: $\mathcal{B}(\Lambda_b^0 \to pK\psi(2S)) = (6.6^{+1.2}_{-1.0}) \times 10^{-5}$ is about nine times smaller compared to $\mathcal{B}(B^0 \to K^{*0}\psi(2S)) = (5.9 \pm 0.4) \times 10^{-4}$. These values are important to keep in mind when discussing the fit yields presented in the Section 3.7. Such a difference, multiplied by the difference in the production rate, has important consequences for the proportion of the misidentification backgrounds, as it is further discussed in detail in the Sec. 3.3. The branching fraction of the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays has never been measured, however, the signal yields in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, observed by the previous analysis [3], provide a rather optimistic prospect for observation of the analogous dielectron decay mode, which is a crucial part of the measurement presented in this work. In addition, the LHCb Collaboration has observed a Cabibbo-suppressed decay $\Lambda_b^0 \to p\pi\mu^+\mu^-$ [88] which has a very similar topology to our decays of interest, and measured its branching fraction to be $(6.9 \pm 2.5) \times 10^{-8}$. This value can be used to estimate roughly the expected branching fraction of the similar Cabibbo-favored decay $\Lambda_b^0 \to pK\mu^+\mu^-$.

Some supplementary knowledge might also be gained from the Λ_b^0 decays to the ground-state Λ baryon, where somewhat a larger number of measurements (and theoretical predictions) exist. As it has already been mentioned, the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ was studied by the LHCb Collaboration, its differential branching fraction was measured, and

an angular analysis was performed [57, 89]. It should be noted that the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays differ from the $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decays from several points of view, besides the available phase-space and spin-structure. First, the long-lived Λ has a lower reconstruction efficiency than the (pK) system, where all the tracks originate from the Λ_b^0 decay vertex. A direct analogy in the meson system are $B^0 \to K_S^0\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays, where the K^{*0} resonance has a higher reconstruction efficiency at LHCb [90]. The fact that all four tracks originate from the same vertex, allows to constrain the decay geometry and suppress the background coming from the combinations of random tracks. No less important, the dilepton invariant mass distribution is different. The $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ analysis shows the region of high dimuon invariant mass to be the dominant one [57]. However, in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay the available phase-space is reduced by the kinematics of the decay, because the (pK) system has a larger mass compared to the mass of the ground-state Λ baryon. As a consequence, the yields in the high-q^2 region are significantly smaller. This can be clearly seen comparing the q^2 distributions for ground-state Λ and various Λ^* resonances in Ref. [85] (Fig. 1.14).

CHAPTER 2

The LHCb experiment at the LHC

The LHCb (Large Hadron Collider beauty) experiment [91] is dedicated to study flavour physics at the Large Hadron Collider (LHC) at CERN. It is designed to perform precision measurements of CP violation and study rare decays of beauty and charm hadrons. These measurements allow for indirect searches for effects beyond the Standard Model.

The data used for the analysis presented in this work, has been collected at the LHCb experiment. This chapter introduces the experiment, focusing mainly on the aspects crucial to understand the experimental challenges of studies of rare decays at LHCb.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [92], located at CERN, is currently the world's most powerful particle collider. It is located in the circular tunnel originally built for the Large Electron-Positron (LEP) collider. The LHC was designed to accelerate protons and heavy nuclei such as lead ions. The design proton beam energy is 7 TeV, implying a 14 TeV center-of-mass energy in the collision. So far, this specification has not been met due to the limitations of the cryogenic system [93]: in the so-called Run I (2010-2012), the collision energy was progressively increased to 7 and then 8 TeV; and in the Run II (2015-2018) the LHC was run at 13 TeV. Proton beams start from a linear accelerator and are progressively accelerated through a chain of circular machines (Fig. 2.1). They achieve the energy of 450 GeV in the Super Proton Synchrotron (SPS) and are then injected to the LHC itself. The acceleration to the final energy is achieved by sixteen radio-frequency cavities. Magnetic field up to 8.3 T is needed to keep protons moving in a circular path along the ring. It is created by more than a thousand superconducting Nb-Ti dipole magnets. Opposite magnetic fields have to be applied to the two proton beams moving in opposite directions. Beams are focused using quadrupole magnets located along the ring. Protons are grouped in bunches, and the spacing between the bunches was 50 ns in Run I, and 25 ns in most of

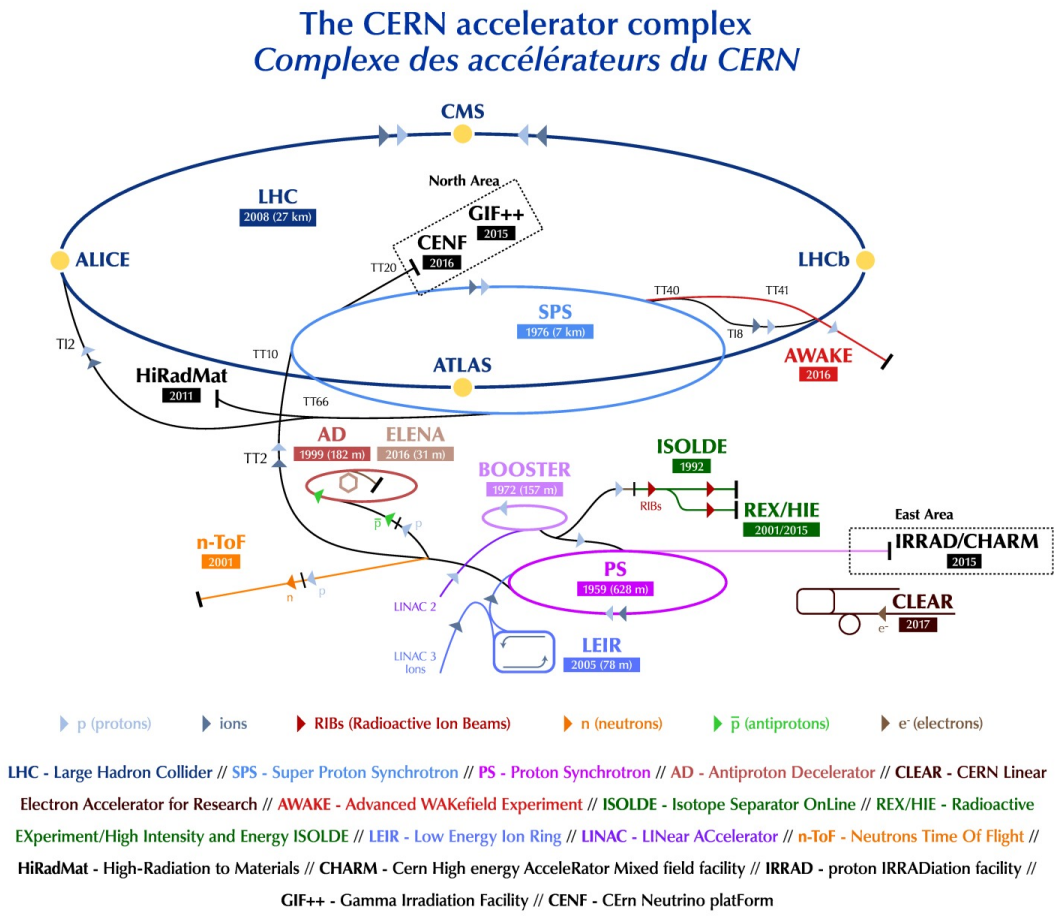


Figure 2.1: The CERN accelerator complex as of 2018. Taken from Ref. [94].

the Run II.

In four points along the ring, the particle detectors are installed. In these places the two proton beams are collided. The four largest experiments operating at LHC include two general-purpose detectors ATLAS and CMS, which surround their interaction points almost hermetically. They have a broad physics programme oriented towards high-transverse-momentum physics, precision tests of the Standard Model, top-quark physics, study of the Higgs boson properties which they have discovered in 2012 [8,9], as well as search for the phenomena beyond the SM. However, they also have certain capabilities to perform flavour physics studies, in particular the hadron spectroscopy programme. Their main limitations in this direction come from the absence of a charged hadron identification, tight trigger thresholds and a harsh environment with very high track multiplicity and number of interactions per collision (*pileup*). These factors are optimised in another experiment – LHCb, which is described in the following sections. The fourth large LHC experiment is ALICE, dedicated to the study of heavy ion collisions and quark-gluon plasma.

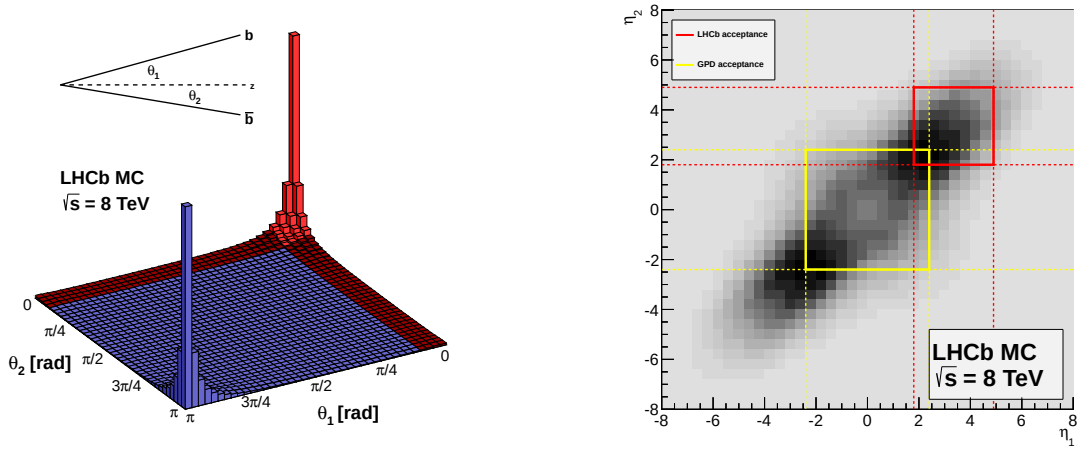


Figure 2.2: $b\bar{b}$ production rates as a function of their angles (left) and pseudorapidities (right). The regions fully covered by the LHCb, are presented in red. Taken from Ref. [95].

2.2 Production of b hadrons at the LHC

A proton is a not a fundamental particle, but a composite object. It can be considered as consisting of three valence quarks (uud), sea quarks, and gluons mediating the strong interaction and keeping the proton bound. This means the proton-proton collision is a very complex process, as compared to the e^+e^- collision. Up to the first order, the collision is mediated by the strong interaction which is flavour-conserving. This means, the leading order processes for creating b quarks are pair creation processes, creating a $b\bar{b}$ pair. The most important contributions arise from the quark-antiquark annihilation $q\bar{q} \to b\bar{b}$, and gluon-gluon fusion $gg \to b\bar{b}$, with the latter being the dominant at the LHC energy scale. Due to the large momentum asymmetries between the partons (gluons or quarks) interacting in the collision, and the fact the LHC energy is much larger than the mass of the b quark, the produced $b\bar{b}$ pair is usually boosted in the forward or backward direction along the beam axis. The closeness to the beam axis is usually characterized in terms of the *pseudorapidity* defined as

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{2.1}$$

where θ is the polar angle w.r.t. the beam axis. This variable is useful, as detectors are usually rather symmetric around the beam axis. It has an important property that pseudorapidity differences between two particles are invariant with respect to Lorentz boosts along the beam axis, which makes it crucial in high-energy physics.

This feature of the boosted b-quark production is useful as the larger boost provides a more displaced b-hadron decay vertex, allowing to identify b hadrons in the harsh environment of the hadronic collision.

When created, b quarks instantly hadronise, combining with one or two lighter (anti)quarks (u,d,s,c) to create a meson or a baryon. Probabilities to hadronise into a

meson, picking an (u,d,s,c) antiquark, usually denoted as f_u, f_d, f_s, f_c, are approximately scaled as (4:4:1:0.01). Probabilities to hadronise into a baryon are studied less precisely, but it is generally accepted that the Λ_b^0 is the most abundantly produced b baryon. The probability of b quark to hadronise into a Λ_b^0 baryon is usually denoted as $f_{\Lambda_b^0}$.

The Λ_b^0 baryon production at the LHC is about half of the B^0 meson production, and decreases with the Λ_b^0 transverse momentum p_T [96,97]. For example, at 13 TeV, LHCb has measured [97]:

$$\frac{f_{\Lambda_b^0}}{f_u + f_d} = 0.259 \pm 0.018 \text{ (averaged over kinematics)} \tag{2.2}$$

The Λ_b^0 production decrease with the transverse momentum can be parametrised as:

$$\frac{f_{\Lambda_b^0}}{f_u + f_d}(p_T) = A[p_1 + \exp(p_2) + p_3 \times p_T], \tag{2.3}$$

where $A = 1 \pm 0.061$, $p_1 = (7.93 \pm 1.41) \times 10^{-2}$, $p_2 = -1.022 \pm 0.047$ and $p_3 = -0.107 \pm 0.002\,\text{GeV}^{-1}$.

Compared to the e^+e^- machines such as B-factories, the LHC environment provides a higher $b\bar{b}$ production cross section and a larger boost. The cross section for producing b quarks in inelastic pp collisions $pp \to b\bar{b}X$ has been measured by LHCb to be $72.0 \pm 0.3 \pm 6.8$ μb at a centre-of-mass energies of 7 TeV, and $144 \pm 1 \pm 21$ μb at 13 TeV, in the LHCb acceptance (defined in the following section) [98]. This is to be compared to the total inelastic cross section measured by LHCb to be 55.0 ± 2.4 mb at 7 TeV and 62.2 ± 2.5 mb at 13 TeV in the same acceptance [99,100]. It can be concluded that the $b\bar{b}$ cross section is two or three orders of magnitude smaller than the total inelastic cross section, leading to a higher background level at the LHC compared to B-factories. This makes challenging the flavour tagging (identification of the flavor of neutral B mesons), and the reconstruction of the final states with neutral (photons, π^0) or missing (neutrinos) particles.

2.3 The LHCb detector

As it was already mentioned, LHCb is a dedicated heavy flavour physics experiment [91, 101,102], whose main goal is the study of the CP violation and indirect search for NP effects in rare decays of beauty and charm hadrons. It is located at Point 8 of the LHC. The large beauty and charm production cross sections at the LHC allowed to collect large samples of heavy flavour decays during the Runs I and II. The LHCb detector is a single-arm forward spectrometer covering only the forward pseudorapidity region $2 < \eta < 5$. This corresponds to geometrical coverage ranges from about 10 mrad to approximately 300 (250) mrad in the horizontal (vertical) plane. The horizontal coverage is designed to be larger than the vertical one, as the 4 Tm LHCb dipole magnet bends charged particles in the horizontal plane [103]. Although this coverage corresponds to only about 4% of the solid angle, the predominantly forward production of the $b\bar{b}$ pairs,

discussed in the previous section, allows to get about 25% of them in the LHCb acceptance. For comparison, general-purpose detectors are covering more than 90% of the solid angle, corresponding to $-2.4 < \eta < 2.4$ and get about 45% of the $b\bar{b}$ pairs (see the comparison in Figure 2.2), though they are less boosted at small pseudorapidities, so the decay vertices are less displaced.

LHCb uses a coordinate frame in which the z-axis is defined along the beam axis, with its positive direction pointing from the interaction point towards the muon system (see Fig. 2.5). The y-axis is vertical, oriented from the interaction point towards the surface and perpendicular to the LHC tunnel. Finally, the x-axis is defined in order to form a right-handed coordinate system.

Instrumenting a small solid angle for a dedicated flavour physics experiment is not only cost-efficient and allows to get a larger displacement of the decay vertices. This also allows to introduce crucial particle identification systems such as ring-imaging Cherenkov detectors (RICH) which are difficult to scale to the general-purpose detector size and shape. The detector design itself is also optimised, with the readout electronics placed out of the acceptance, to decrease the amount of material in the acceptance.

This chapter describes the LHCb detector as it was operated during the Run I and Run II of the LHC. In order to optimise the LHCb performance in the high-multiplicity hadronic environment, the LHC provides collisions for LHCb at a significantly lower rate than for ATLAS and CMS. The collision rate at the LHC is commonly characterised in terms of *instantaneous luminosity* $\mathcal{L}(t)$, which is proportional to the parameters of colliding beams. Luminosity relates the rate of a given process (number of events N) to its cross section σ:

$$N = \sigma \times \int \mathcal{L}(t) dt \tag{2.4}$$

As it is shown in Fig. 2.3, the instantaneous luminosity delivered to LHCb is kept almost constant ('leveled') during the LHC fill, by adjusting the beam overlap. This allows to maximize the *integrated* luminosity collected during the fill, to maintain the same trigger configuration and to reduce systematic uncertainties. The integrated luminosity collected by the LHCb to date is about 9 fb^{-1}, and the split per each year of the data taking is presented in Fig. 2.4. The average data taking efficiency is about 90%, and about 99% of the recorded data is good for analysis [102]. The dominant sources of inefficiency include the safety procedures, dead-time of some subdetector front-end systems, and short technical problems.

The LHCb detector is designed to perform the reconstruction of exclusive decays of beauty and charm hadrons in various final states involving charged leptons, charged and neutral hadrons, as well as photons. The following particles are usually called "stable" within the LHCb-related context as they live sufficiently long to cross the detector, and can be detected directly: charged pions ($\pi^{\pm}$), charged kaons ($K^{\pm}$), protons (p and $\overline{p}$), electrons ($e^{\pm}$), muons ($\mu^{\pm}$), photons (γ) and deuterons (d). Unstable particles of much shorter lifetime are formed by combining their stable decay products. A specific class of *long-lived* particles is also considered, which includes the weakly-decaying strange hadrons such as K^0_S, Λ or Ξ^-: they can cross several subdetectors before decaying.

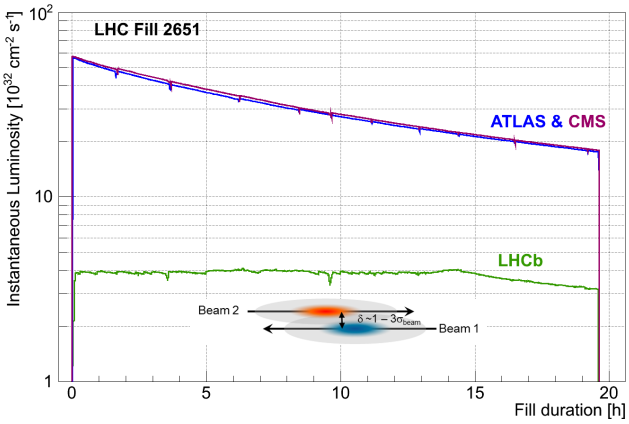


Figure 2.3: Development of the instantaneous luminosity for ATLAS, CMS and LHCb during one long LHC fill in Run I. Taken from Ref. [102].

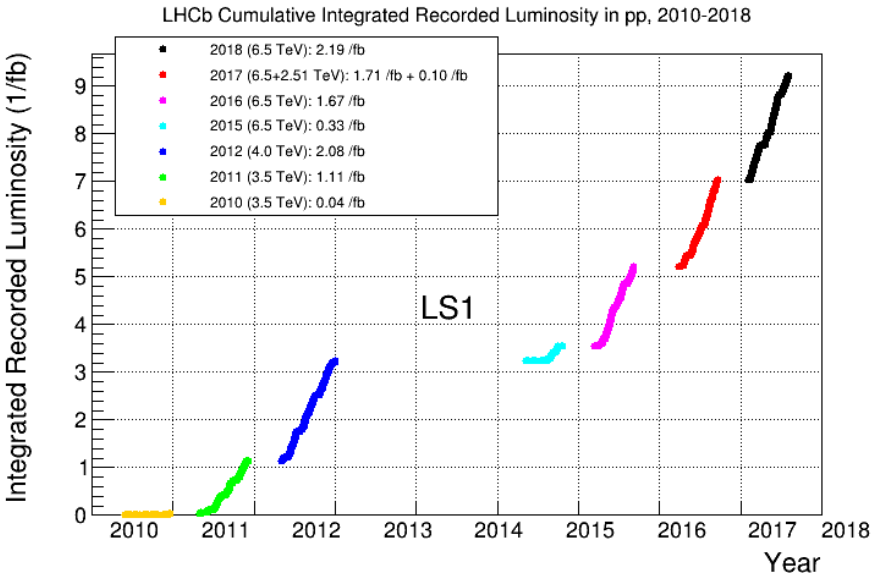


Figure 2.4: The cumulative integrated luminosity, recorded by the LHCb detector during each year of Run I and II. Taken from Ref. [104].

In order to reconstruct (*i.e.* measure the properties) and identify these particles, the LHCb detector incorporates several specialised sub-detectors, shown schematically in Fig. 2.5. The track reconstruction system includes a vertex detector, called the Vertex Locator (VELO) which is located very close to the collision point, and planar tracking stations: the Tracker Turicensis (TT) located upstream of the dipole magnet, and three tracking (T1-T3) stations located downstream of the magnet. VELO is a very special subdetector at the LHCb, which can locate precisely the collision point and decay vertices of b hadrons. This is crucial in order to separate the b-hadron decays from the promptly produced tracks, to measure the b- and c-hadron lifetime and to distinguish between

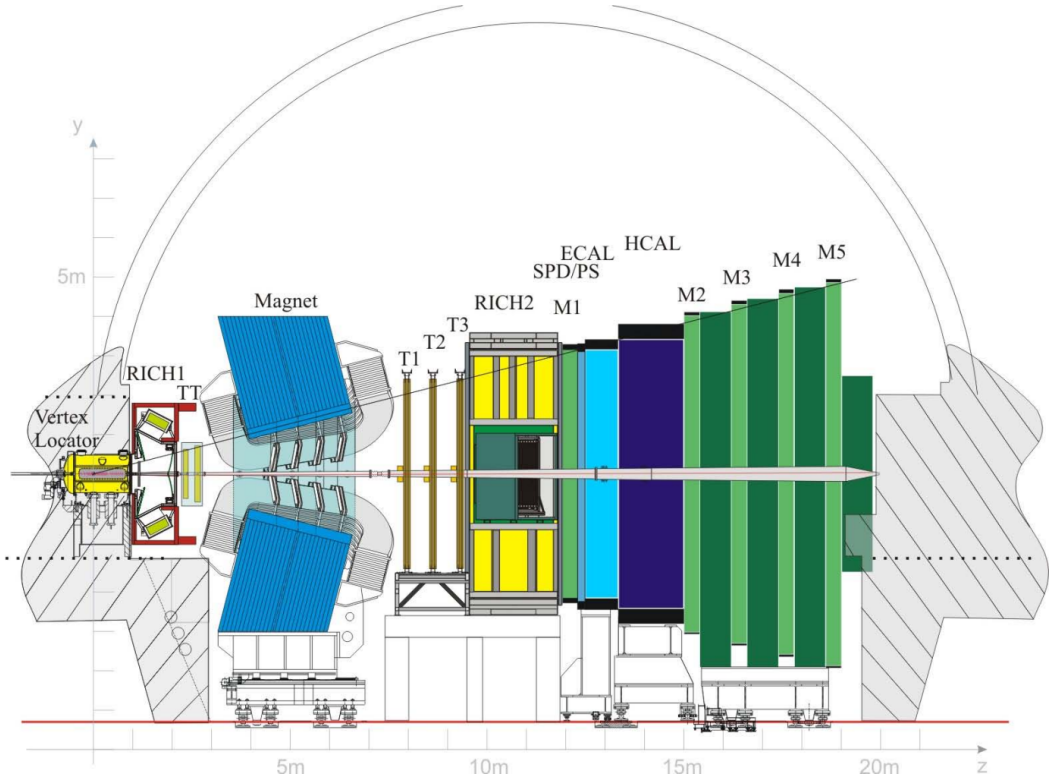


Figure 2.5: The schematic view of the LHCb detector. Taken from Ref. [101].

prompt and displaced event topologies. The magnetic field allows to determine the charge of the track, as well as estimate its momentum based on the curvature. The non-uniformity of the magnetic field was measured to be about 1% [91]. The magnet polarity can be reversed to reduce detection asymmetries. Further throughout the text the two possible polarity configurations are called "MagDown" and "MagUp". Polarities are regularly reversed, so that the samples of both polarities are approximately equal for each data-taking year. Charged particle identification (PID) is provided by two Ring Imaging Cherenkov detectors (RICH1 and RICH2), as well as by information from the calorimeter and the muon system. The calorimeter system is composed of an electromagnetic calorimeter (ECAL) with its Preshower (PS) and Scintillator Pad Detector (SPD); and a hadronic calorimeter (HCAL). The calorimeter system is used in the hardware trigger, and for the reconstruction of photons and electrons. The muon system is composed by five muon chambers (M1-M5), all but M1 located behind the calorimeter system.

The individual subdetectors are described in detail in the following sections.

2.4 Tracking system and vertex reconstruction

Precise vertex reconstruction and high momentum resolution are key ingredients for the detector performance. The tracking system is designed to reconstruct the trajectories of charged particles from the hits recorded by various subdetectors. It requires a high spatial

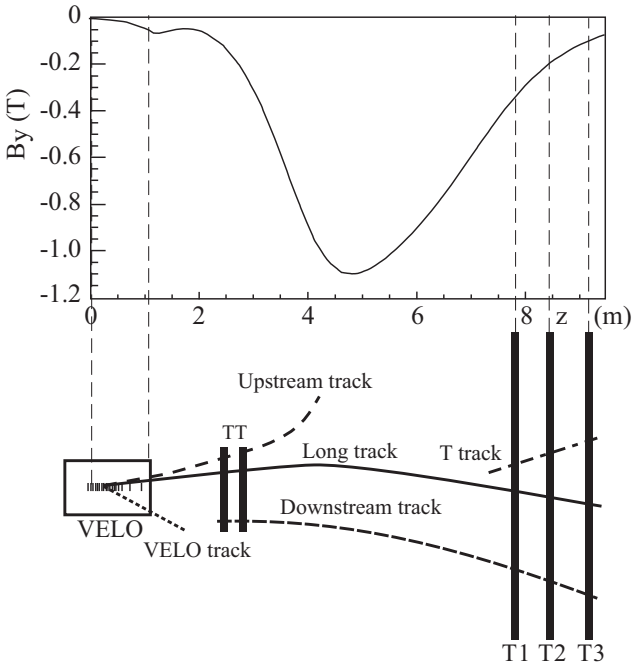


Figure 2.6: The types of tracks which can be reconstructed in the LHCb detector, and the main component of the magnetic field (B_y) as a function of the z coordinate. Taken from Ref. [102].

resolution and a low material budget.

The various track types in the LHCb detector are presented in Fig. 2.6. Most of the analyses, including this work, rely on the long tracks – those which have hits in the VELO and tracking stations T1-T3. Long tracks are reconstructed with about 96% efficiency. Analyses using long-lived neutral particles such as Λ can also rely on the downstream tracks – those without VELO hits. Other tracks are rarely used directly in physics analyses due to the poor momentum resolution or unreliable identification. However, they still play an important role in various reconstruction and track isolation algorithms. For instance, the VELO tracks, reconstructed via VELO hits, allow to locate the primary vertex (PV), in particular because it is the only subdetector having an acceptance in the backward direction. However, as it can be deduced from Fig. 2.6, the magnetic field in the VELO region is very small, therefore it is difficult to measure the momentum of VELO tracks. If a particle is reconstructed more than once, as different track types, only the track best suited for analysis is kept [105]. In that case, a track crossing a larger number of subdetectors are preferred. The number of unique tracks in an event, `nTracks`, is frequently used as a proxy for the event multiplicity.

Reconstruction of long tracks starts from forming the straight VELO tracks, exploiting the near absence of the magnetic field in the VELO region [105, 106]. To form long tracks, VELO tracks are combined with either hits in the T1-T3 tracking stations, or with already prepared T-tracks [107]. If there are hits in TT on the trajectory of the long track, they are added in order to improve the momentum resolution, and reject *fake* tracks, *i.e.* tracks which are created as a combination of random unrelated hits in tracking detectors.

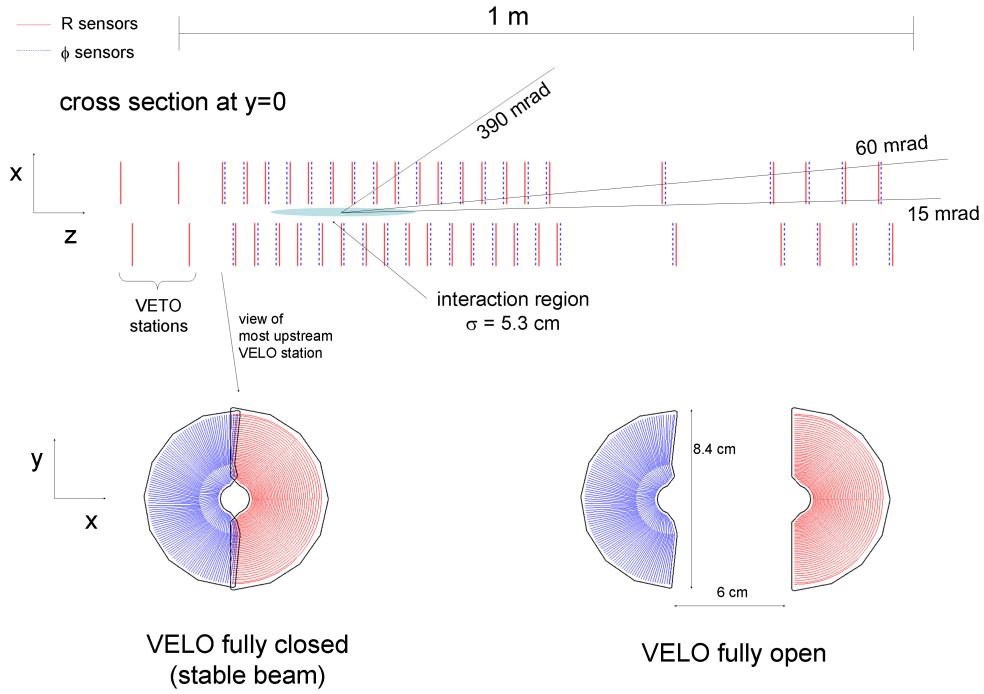


Figure 2.7: Overview of the VELO detector, showing spacing of modules along z-axis, and open and closed positions. Taken from Ref. [114].

LHCb tracking algorithms rely on a fit based on the Kalman filter [108], allowing to improve the resolution on the track parameters and reject fake ('ghost') tracks. This fit allows to reconstruct more precisely the trajectory of a particle, taking into account effects from multiple scattering and energy loss in the material interactions. The χ^2 per number of degrees of freedom, $\chi^2_{track}/nDOF$, of this fit, is a proxy for the quality of the track. Further fake track rejection is provided by a neural network considering the information from all subdetectors [109].

Finally, a removal of duplicate tracks, so-called clones, is performed by a dedicated `CloneKiller` algorithm. Clone tracks are tracks which share more than 70% of the hits among themselves [110, 111]. Only the track with the best quality is kept.

A more detailed overview of the design of each tracking subdetector is provided below.

The Vertex Locator

Surrounding the interaction point, the VELO detector [112, 113] aims at locating precisely the primary vertex position. This detector is composed of $R-\phi$ silicon sensors placed along the beam axis, as shown in Fig. 2.7.

R-type and ϕ-type sensors have their routing lines oriented in perpendicular and parallel directions to the silicon strips, respectively. This allows to measure the radial

distance to the beam axis, and the azimutal angle, respectively. During data taking, the sensors are located in 8 mm from the beam centre. This can be dangerous when the beam is not sufficiently stable, for example during injections. To prevent sensors from potential damage, they are split in two halves which are kept open during the injection, and are closed only when the beam is stable. There are 21 modules in each half of the subdetector. The entire VELO is placed in a vacuum container surrounded by a thin aluminium foil, which separates the beam vacuum from the vacuum inside the VELO box. It also shields the VELO sensors against radio-frequency pickup from the LHC beams, therefore is often called 'RF-foil'. This foil contributes significantly to the material budget, somewhat decreasing the vertexing performances [115, 116]. The VELO achieves a spatial vertex resolution down to 4 μm, which is the best vertex detector resolution at the LHC.

TT

The TT, which stands for either 'Tracker Turicensis' or 'Trigger Tracker', is one of two silicon tracker subdetectors at LHCb. It is composed of two stations, each of two layers, placed between the RICH1 and the magnet (Fig. 2.8). Two of the four layers are tilted by five degrees with respect to the vertical axis, allowing to get a three-dimensional track reconstruction. The vertical orientation of the silicon strips is needed to obtain the optimal spatial resolution in the horizontal plane (the bending plane of the magnet). The single hit resolution of the TT is about 50 μm. The TT is important to improve the momentum resolution of the long tracks, and even more important to form the downstream tracks. It is also used in Run II tracking algorithms in the software trigger, providing the first hit after VELO and allowing to narrow down the search window in the tracking stations, therefore speeding up the algorithms [117].

Tracking stations

Tracking stations are based on two different detector technologies. The inner tracker (IT) [118], is based on the same silicon-strip technology as the TT, and covers the region close to the beam axis. Same as TT, each tracking station has four layers, two of them tilted. The achieved resolution is similar to the TT one. The Outer Tracker (OT) [119–121], covering the outer regions of each tracking stations, is composed of the straw tubes filled with a mixture of argon, CO_2 and oxygen. The position of the hits in the OT is determined by measuring the drift-time of the ionisation clusters created in the gas to the anode wire. This is possible thanks to precise synchronisation with the LHCb clock and the collision time. The used gas mixture allows to have a drift time shorter than 50 ns and a spatial resolution of about 200 μm.

Tracking performance

The efficiency of the tracking algorithms can be estimated using the *tag-and-probe* techniques on abundant decay modes such as $J/\psi \to \mu^+\mu^-$ or $D^0 \to K^-\pi^+$. One track ("tag") is reconstructed using the nominal full reconstruction, while the other track ("probe") is reconstructed using only specific tracking subsystems, allowing to probe the

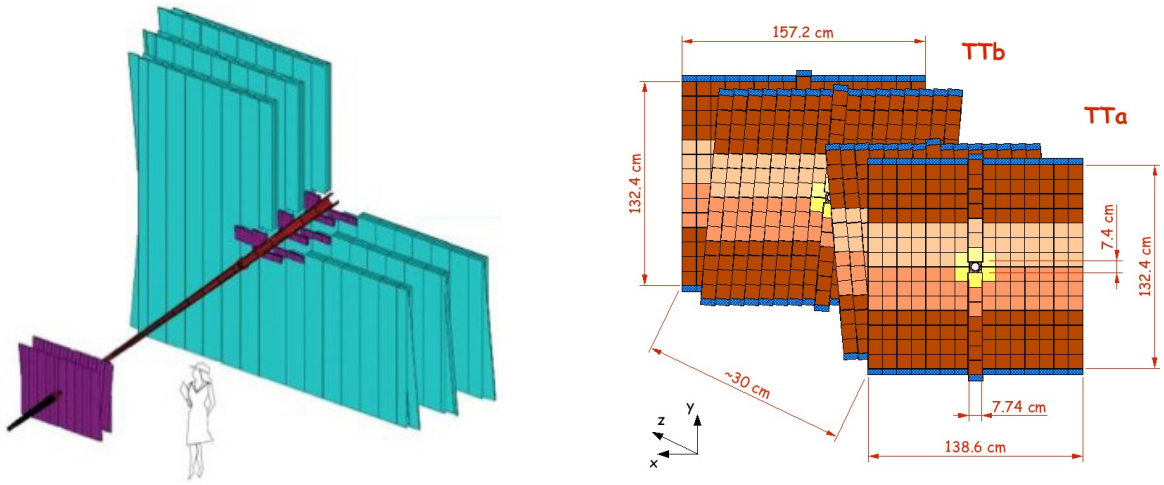


Figure 2.8: (Left): LHCb tracking detectors. The TT detector can be seen to the left, with three larger tracking stations (T1-T3) to the right. The silicon trackers are shown in violet, and the straw tubes – in green. Tracking stations T1-T3 combine both technologies: inner part (IT) is based on silicon sensors, while the outer part (OT) relies on straw tubes. (Right): the schematic view of the TT stations. Taken from Refs. [91, 101].

remaining subdetectors [105]. The decay $J/\psi \to \mu^+\mu^-$ is the best suited for this task, as muons leave hits not only in the tracking system, but also in the muon stations. The VELO-track-finding efficiency has been measured using tag-and-probe method to be about 98% in Run I data [113]. The total track reconstruction efficiency is above 95% [105, 117]. Fig. 2.9 (left) describes the tracking efficiency in Run I and Run II as a function of the track momentum. A small reduction in the tracking efficiency is observed in 2015 due to the fact that the bunch spacing of the LHC was reduced from 50 to 25 ns, resulting in larger spillover effects in the OT detector whose readout window is larger than 25 ns [117].

The momentum resolution in LHCb varies from $\delta p/p = 0.5\%$ at low momentum to 1% at 200 GeV/c, as it is shown in Fig. 2.10 for Run I data.

The performance of the ghost track rejection algorithm is described in Fig. 2.9 (right) on an example of Run II data. The default working point was chosen such that it rejects 60% of all fake tracks, while maintaining an efficiency of about 99%. However, a tighter selection is usually applied later on, at the analysis level.

2.5 Particle identification systems

Up to this point, it was shown how tracks are reconstructed in LHCb, however, the nature of these particles is not known. LHCb can profit from the information of multiple subdetectors in order to perform particle identification. This can either be based on the response of a specific subdetector, or on global likelihoods or neural networks which use information from all subdetectors to assign a particle type to a track. For this purpose, two Ring Imaging Cherenkov detectors (RICH1 and RICH2) are used, one located in front

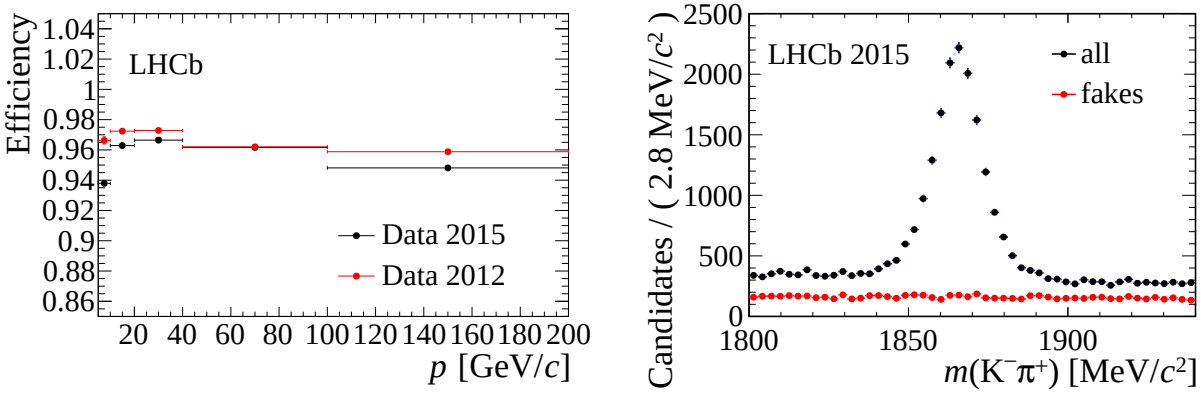


Figure 2.9: (Left): Comparison of tracking efficiency in 2012 and 2015 data, as a function of track momentum. (Right): Performance of the fake track rejection algorithm in 2015 data. Taken from Ref. [117].

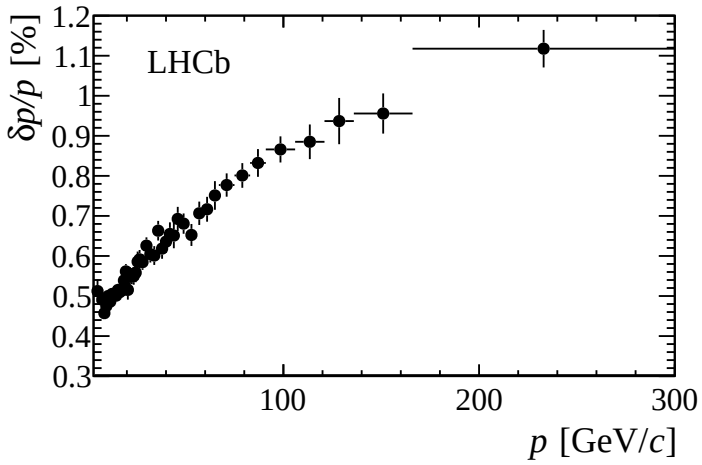


Figure 2.10: Relative momentum resolution as a function of momentum, for long tracks in Run I data, obtained using J/ψ decays. Taken from Ref. [102].

of the magnet, which identifies lower-momentum particles, and another located after the tracking stations and helps identifying higher-momentum particles. RICH detectors exploit the Cherenkov radiation emitted by a charged particle in order to discriminate between the different charged hadron species. They also play an important role in identification of muons and electrons, as it will be discussed below.

The calorimeter system aims at measuring the energy of the particles, by stopping and absorbing them. It provides identification of electrons and photons, and is sufficiently fast in order to be used in the hardware trigger.

The muon system is devoted to the detection of the muons. As it has already been mentioned, all the muon stations but one are located behind the calorimeter, and separated by additional iron shields. Muons of energies above a few GeV have sufficient penetration power to pass through this amount of material, which would stop most of other charged

particles.

RICH detectors

For precision heavy-flavour measurements, it is crucial to have a powerful tool able to separate protons, kaons and pions. At LHCb, this tool is based on the two RICH detectors [122, 123]. Let's consider charged particles crossing a certain dielectric medium with refraction index n, and assume the particle's velocity is larger than the phase velocity of light in this medium c/n. Such particles will emit photons in the direction defined by a cone with an opening angle $\cos\theta = \frac{1}{n\beta}$, where $\beta = v_{particle}/c$. This angle depends on the *velocity* of the particle, but not its momentum. Therefore, combining the measured velocity of the particle, with the track momentum measured by the tracking system, it is possible to deduce the mass of the particle:

$$\cos\theta = \frac{1}{n}\sqrt{1 + \left(\frac{mc}{p}\right)^2}. \tag{2.5}$$

Two RICH detectors in LHCb use different media (also called radiators), with different refraction indices. This allows to cover a wider momentum range. RICH1 exploits the C_4F_{10} gas with $n = 1.0014$, and RICH2 – the CF_4 gas with $n = 1.0005$. This allows RICH2 to be sensitive to particles of higher momentum. RICH1 covers the full LHCb acceptance, while the RICH2 covers only high-pseudorapidity range, where the majority of high-momentum particles are expected to end up. Cherenkov photons emitted inside radiators are reflected by the mirrors and are directed to the Hybrid PhotoDetectors (HPD), that detect photons in the optical range (200-600 nm). Dedicated algorithms assign single photons to circles and compute the corresponding Cherenkov angle. Reconstructed Cherenkov angle θ as a function of momentum of the tracks crossing the C_4F_{10} radiator, is presented in Fig. 2.11 for several charged particles. A dedicated algorithm associates tracks to rings reconstructed in the RICH system, and assigns a likelihood to each track of being a given particle type `DLLX`. Typically, differences of likelihoods of the kaon (proton) hypothesis and the pion one, called `PIDK` (`PIDp`), are used as discriminating variables in the offline selection. Alternatively, a neural network (multilayer perceptron with one hidden layer), is trained separately for each charged particle species. It is used to combine information from all subdetectors and return an output for each charged particle hypothesis (pion, kaon, proton, muon, electron or deuteron), `ProbNNX`, which is a more powerful discriminating variable than the difference of likelihoods [124].

However, as it can be seen from Fig. 2.11, there are kinematic regions where a reliable separation of different hadrons is not possible. This is visualised in Fig. 2.12 (left) on the example of proton-kaon separation, for the loose cut `(DLLp-DLLK)>0` and a tighter cut `(DLLp-DLLK)>5`. At low momentum, kaons and protons do not have a sufficient velocity to create a Cherenkov ring, so their identification is challenging. At the same time, at high momentum, the Cherenkov angle is behaving in a similar manner for all charged particles, making their separation difficult. Even in the intermediate momentum range, the particle identification efficiency is not equal to 100%. It can also be seen from Fig. 2.12 (right) that the particle identification performance for the cut `PIDK>0` degrades with event multiplicity.

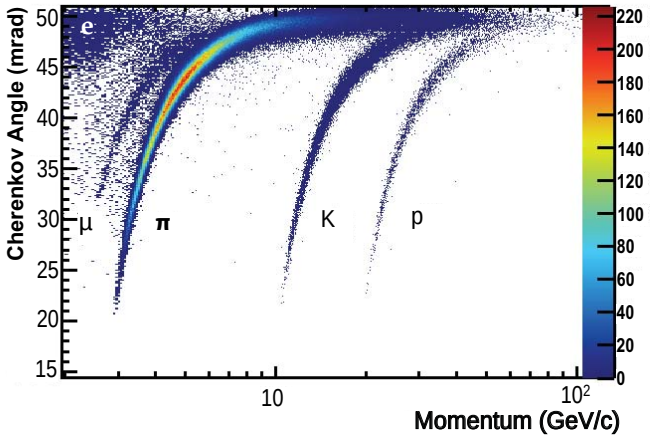


Figure 2.11: Reconstructed Cherenkov angle θ as a function of momentum of the tracks crossing the C_4F_{10} radiator, for several charged particles in 2011 LHCb data. Taken from Ref. [123].

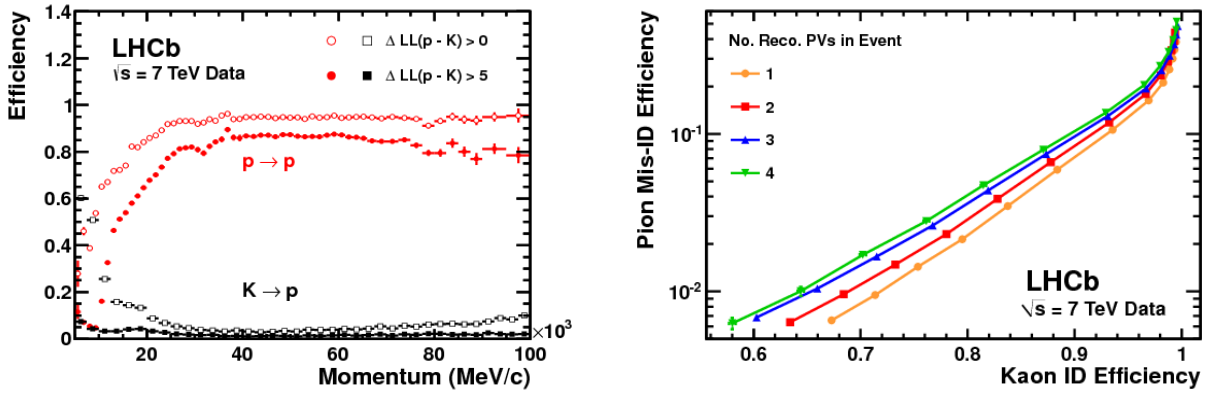


Figure 2.12: Typical proton and kaon PID performance as a function of (left) track momentum, and (right) event multiplicity. Taken from Ref. [123].

Calorimeter system

As it has already been mentioned, the LHCb calorimeter [125] provides separation of electrons, photons and hadrons, also allowing to measure their energies and positions. It plays a crucial role in the hardware trigger system, where high transverse energy (E_T) deposits of these particles are searched for, as it will be discussed in more detail in Sec. 2.6.1. It consists of four subdetectors - SPD, PS, ECAL and HCAL.

SPD and PS consist of scintillator pads, with a lead plate located between the two subdetectors. ECAL consists of stacks of alternating scintillator and lead absorber plates, this type of design is commonly called "shashlik". Finally, the HCAL is made of iron and scintillator tiles. All four subdetectors use wavelength-shifting fibers to propagate the scintillation light to the photomultipliers. The granularity of the four subdetectors

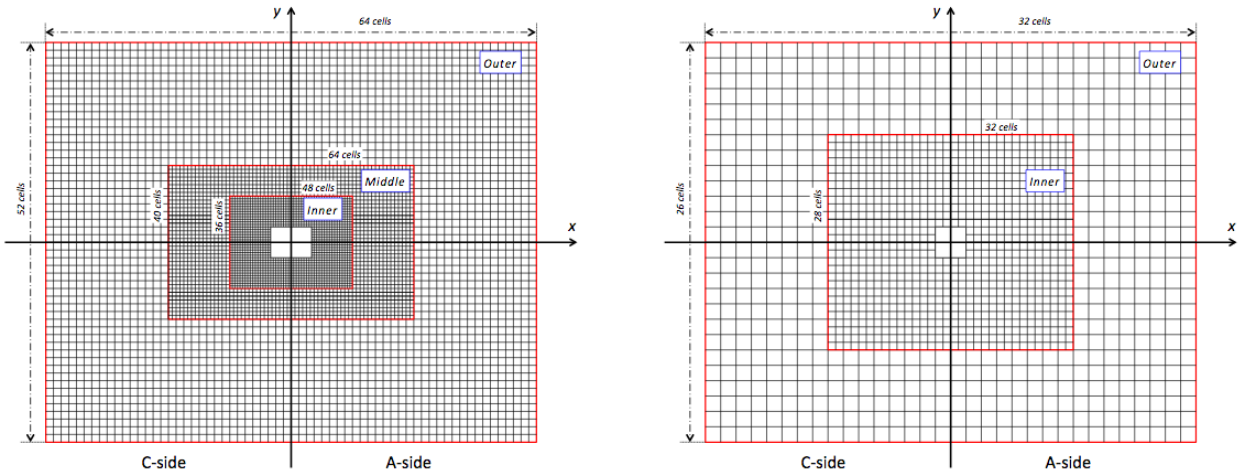


Figure 2.13: Regions of a different granularity in ECAL (left) and HCAL (right). Taken from Ref. [126].

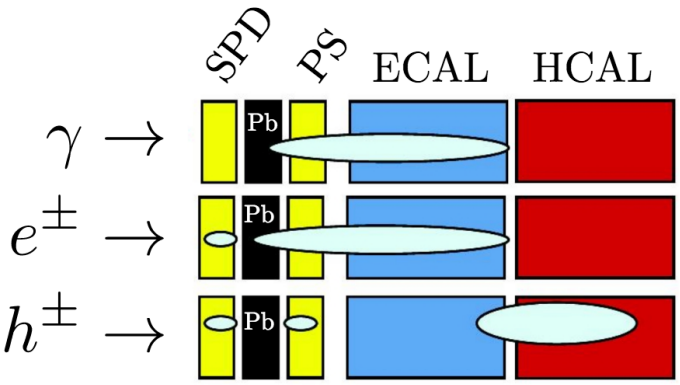


Figure 2.14: Schematic explanation of the principle of identification of photons, electrons and hadrons with the LHCb calorimeter system. Taken from Ref. [127].

increases closer to the beam pipe, in order to cope with the increase in the hits density. In particular, the ECAL is divided into three granularity regions, with cell widths of 40.4, 60.6, and 121.2 mm in the inner, middle, and outer regions respectively (Fig. 2.13). HCAL is only divided in two regions.

The thickness of the lead plate (about 2.5 radiation lengths X_0) between the SPD and PS is chosen in a way that photons start to initiate the electromagnetic shower in it, which is then detected by the PS and ECAL. Electrons in addition produce also hits in the SPD, and this signature allows to separate them from photons. The electron identification is described in detail in Sec. 2.7. Photons are identified as ECAL clusters which neither have associated tracks pointing to them nor have associated hits in the SPD system (Fig. 2.14). In addition, multivariate algorithms are used to separate ECAL clusters caused by single photons, from the clusters from the decays of highly boosted π^0

mesons. Photons reconstructed in the ECAL are also important in order to recover the bremsstrahlung radiation emitted by electrons due to their interactions with the LHCb material (Sec. 2.7.2). The ECAL energy resolution is about $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} + 1\%$, where the energy is given in GeV [128,129].

The ECAL thickness is about 25 radiation lengths, allowing to fully contain electromagnetic showers. However, this is still smaller than the hadronic interaction length. Therefore, hadrons leave relatively small deposit in the ECAL, and produce the main shower in the HCAL which has a thickness of only about 5.6 hadronic interaction lengths. By contrast to the ECAL, the HCAL is primarily used to identify high energy hadronic clusters for the hardware trigger stage. Its energy resolution is too poor to be useful for other purposes [128,129].

The muon system

The LHCb muon system [130–132] is composed of five muon stations (M1-M5), which are designed for identifying and triggering on muons. Muon identification relies on combining the information from the muon stations, RICH and tracking systems, and calorimeters [133]. In particular, some of the most abundant decay modes of b hadrons – semileptonic modes and modes with charmonia in the final state, can have muons in the final state, for which triggering is easy. The angular acceptance of the LHCb muon system ranges from 20 (16) mrad to 306 (258) mrad in the bending (non-bending) plane. The first station, M1 is located upstream of the calorimeters. This is designed in order to improve the transverse momentum resolution in the muon trigger, by minimizing effects of multiple scattering in the material of the calorimeter. The other four stations are located after the calorimeter. Absorbers made of 80-cm thick iron plates are placed between the muon stations. Together with the stopping power of calorimeters, this allows to suppress most of charged particles except for muons which are long-lived and minimum-ionising particles. In order to traverse all muon chambers and absorbers, muons need to have a momentum above 6 GeV/c.

Each of the muon stations is divided in four regions (R1-R4) of a different granularity (Fig. 2.15): having a finer segmentation in the regions close to the beam pipe, which experience a higher particle multiplicity. All the regions of each muon station use Multiwire Proportional Chambers (MWPC). This is although not the case for the innermost region of M1, which uses the Gas Electron Multiplier (GEM) technology, as it allows for a larger radiation tolerance [131].

The muon identification consists of several steps [133]. First, the information is collected on the muon penetration through the calorimeters and iron plates, and is used to construct a binary variable `isMuon`. It provides a high efficiency for muons with momentum above 3 GeV/c, reducing the hadron misidentifications to the percent level. At the next step, the tracking system is used to extrapolate the trajectories of charged particles to the muon stations with high precision, and study the pattern of the hits in the muon stations around the extrapolation. This allows to create a discriminating variable `muDLL` which is the logarithm of the ratio of muon and non-muon hypotheses. It is then plugged in a combined likelihood for the different charged particle hypotheses (or the

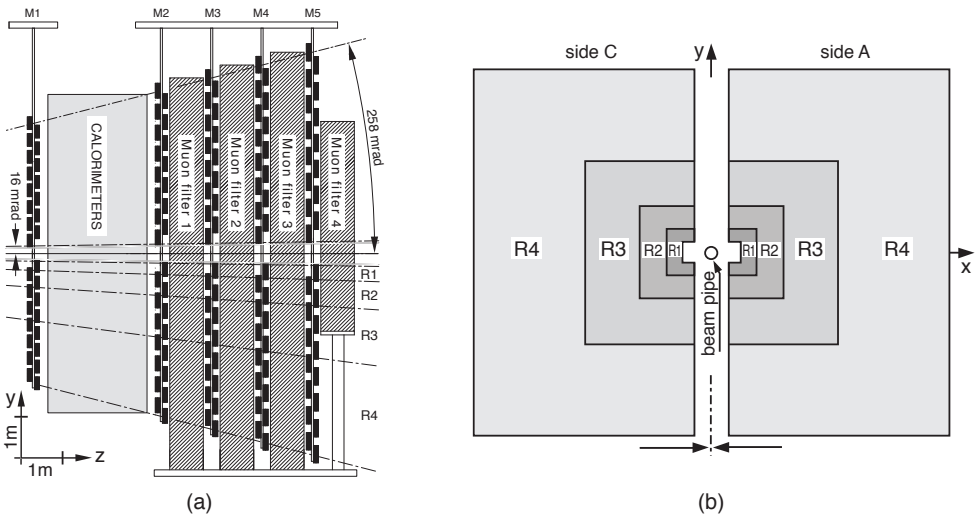


Figure 2.15: The layout of the LHCb muon system. Left: side view. Right: station layout featuring four regions of a different granularity. Taken from Ref. [132].

neural network described above), which also includes the information from the RICH system and the calorimeter. This provides the logarithm of the ratio between the muon and pion hypotheses `PIDmu` (or the neural network output `ProbNNmu`) which is used as a PID discriminating variable in the offline selection.

2.6 The LHCb data processing chain

The proton-proton bunch crossing rate in LHCb is 40 MHz, out of which the "visible" rate, *i.e.* the rate of collisions at the LHCb collision point, was about 15 MHz in Run I and almost twice larger in Run II [134]. It is not practical to store all the data, not only from the memory point of view, but also due to the fact that only a part of these events are interesting to study. During the design phase of the LHCb detector, there was no way to process all the data at such a high rate. Instead, recorded collision data goes through a dedicated data flow, designed to optimize the data-taking efficiency and the purity of recorded signals.

First, raw data from the detector is passing through the *trigger*. It consists of hardware (L0), and software (high-level trigger, HLT) steps. The hardware trigger applies very basic selections aiming to reduce the rate down to about 1 MHz, which are further discussed in Sec. 2.6.1. Software trigger runs on the Event Filter Farm, and is split in two stages. The first step, called HLT1, performs a partial event reconstruction and can apply selection requirements on the tracks. The second step, HLT2, performs a full event reconstruction (also called "online reconstruction") and is flexible enough to accommodate complicated selection requirements. The software trigger strategy has been changed between Runs I and II of LHCb, accounting for the higher rate of interesting events coming with the increase in the energy. This is reflected in the diagram presented in Fig. 2.16 and is

discussed below in Sec. 2.6.2. The overview of the performance of the hardware and software triggers in Run I and II is given in Refs. [135–137].

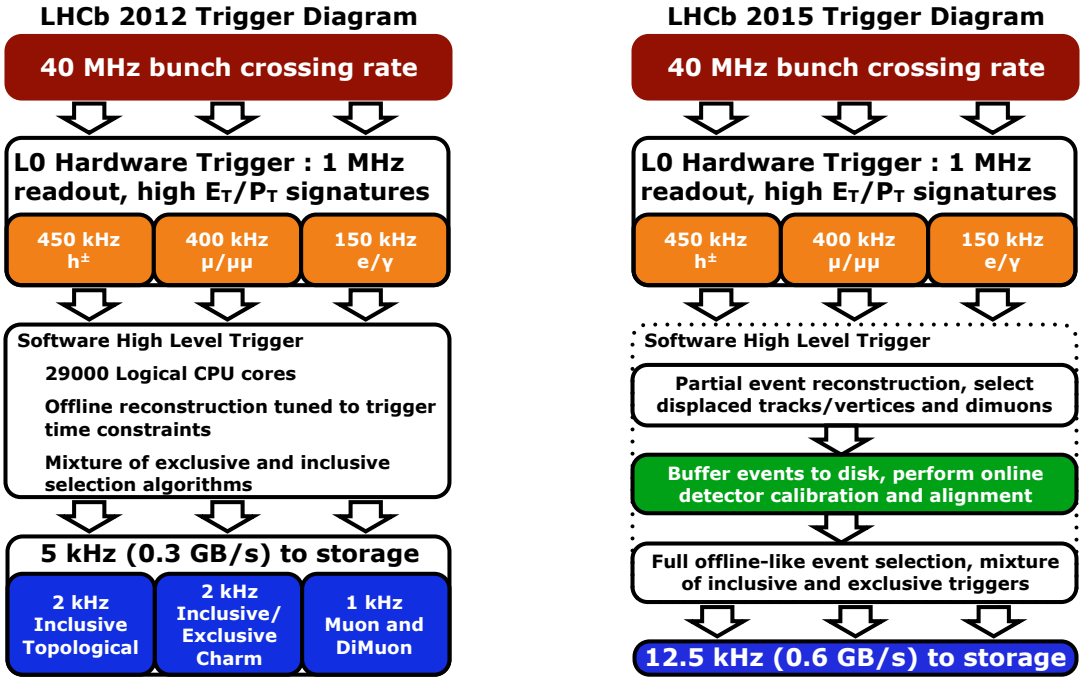


Figure 2.16: LHCb trigger processing diagram for Run I (left) and Run II (right). Taken from Ref. [138].

Below, the data processing steps are discussed in a more detailed manner.

2.6.1 L0 trigger

The hardware trigger works with basic detector information coming from the calorimeter and the muon system. However, the need to fit into the 1 MHz rate limit induces the low efficiency of the hardware trigger for many channels. Triggering on muons is relatively easy, due to the low occupancy of the muon stations. However, triggering on hadrons, electrons and photons is way more complicated, due to the high occupancy of the calorimeter and abundance of prompt hadrons coming from the collision point.

There are two classes of L0 triggers in LHCb: muon triggers and calorimeter triggers. Muon triggers include the single-muon `L0Muon` trigger, and few additional ones (dimuon, high-p_T muon) which are less important. The calorimeter trigger relies on the information from the SPD and PS systems to distinguish between hadron, electron and photon candidates. Consequently, it is split in `L0Hadron`, `L0Electron` and `L0Photon` triggers. Below, more details are provided on the triggers which are used in this work.

The `L0Muon` trigger searches for straight tracks in the five muon stations [137]. The track direction is used to estimate the transverse momentum p_T of a muon candidate, assuming that this muon candidate originated from the interaction point. If there are

several muon candidates in the event, the one with the largest p_T is used for the trigger decision. Event passes the trigger, if this largest p_T is above the pre-defined threshold p_T value (called `L0Muon` trigger threshold).

The L0-calorimeter system relies on the energies deposited in the SPD, PS, ECAL and HCAL detectors [137]. As it has already been discussed, all detector components are segmented transversely to the beam axis into cells of different size. The decision to trigger an event by the `L0Electron` trigger is based on the *transverse energy* deposited in clusters of $2x2$ cells in the ECAL. The transverse energy of a cluster is defined as

$$E_T = \sum_{i=1}^{4} E_i \sin\theta_i, \tag{2.6}$$

where E_i is the energy deposited in cell i, and θ_i is the angle between the z-axis and a line from the cell centre to the average proton-proton interaction point. The E_T of the candidate clusters is compared to a fixed threshold, and events containing at least one candidate above this threshold, are accepted [126]. A similar strategy is exploited by the `L0Hadron` trigger, which uses energy release in HCAL and ECAL cells.

In all hardware triggers, the busy events which could saturate the HLT processing farms, are rejected by applying a cut on the number of cells with hits in the SPD [139]. For most triggers, `nSPDHits` < 600 is applied in Run I, and `nSPDHits` < 450 in Run II. The effect of this cut can have some impact on the analysis, as it is further discussed in Sec. 3.8.2.

The summary of typical L0 thresholds for muons, electrons and hadrons is provided in Table 2.1. These thresholds are not constant during the data taking period, but are changed several times per year, in order to have the optimal setup following the changes in beam conditions and instantaneous luminosity. Therefore, this table summarizes those thresholds which were used in the largest part of each data sample.

Trigger	2011	2012	2016
`L0Muon`	$p_T > 1.48\,\text{GeV}/c$	$p_T > 1.76\,\text{GeV}/c$	$p_T > 1.8\,\text{GeV}/c$
`L0Electron`	$E_T > 2.5\,\text{GeV}$	$E_T > 3\,\text{GeV}$	$E_T > 2.4\,\text{GeV}$
`L0Hadron`	$E_T > 3.5\,\text{GeV}$	$E_T > 3.7\,\text{GeV}$	$E_T > 3.7\,\text{GeV}$

Table 2.1: Summary of typical L0 thresholds in 2011, 2012 and 2016 data-taking periods. Based on Refs. [135–137].

It can be seen that the `L0Hadron` trigger has much tighter thresholds compared to other triggers, which makes it quite inefficient. This has to be done in order to fit into the maximally allowed L0 rate in the harsh environment of the hadron collider.

Another important point is that the momentum or energy resolution in the hardware trigger is worse than in the further online and offline selections. In particular, the p_T resolution of the `L0Muon` trigger is about 20-25% averaged over the p_T range [136, 137]. This implies that in the dataset used for the analysis, the trigger decision does not appear as a sharp cutoff, but is rather smeared.

2.6.2 High-level trigger in Runs I and II

The main purpose of the HLT1 is to select beauty and charm decays. This requires a partial reconstruction of the event, searching for such signatures as displaced vertices and tracks, as well as dimuons. However, it does not have dedicated selections for photons which are calorimeter objects: event with photons can only be triggered at the HLT1 level by other particles in the event. For the same reason, there is no dedicated dielectron selection HLT1 algorithms, although electron tracks can still pass general HLT1 selections.

HLT2 performs the full event reconstruction. Ideally, this requires the fully calibrated and properly aligned detector. However, the calibration and alignment can change from one fill to another, so has to be re-optimised at every fill. In Run I, this was not performed in parallel to the data taking, but at a later stage. In parallel to the data taking, *i.e.*'online', only simplified reconstruction algorithms were used in order to fit into the timing constraints. This required yet another reconstruction step and reprocessing of the full dataset after the HLT2 step, 'offline'. These reasons caused an important mismatch between 'online' and 'offline' event reconstruction. Towards the end of Run I, an optimised HLT strategy was used, saving 20% of events which passed L0 to disk, and processing them through the HLT in the intervals between the LHC fills. This allowed to optimise the usage of computing resources, improve the track reconstruction in the HLT2 by lowering p_T thresholds of track reconstruction algorithms, and implement a dedicated track reconstruction for long-lived particles [136].

In Run II, the optimisations in the HLT software and the upgraded Event Filter Farm allowed to perform the full event reconstruction directly in the HLT2. The HLT1 was still run in parallel to the data taking, but its output was saved to the disk buffer. This allowed to perform the full alignment and calibration of the detector online during the fill, while data was kept in the buffer, and apply it before applying HLT2 selections. HLT2 could be run at a later stage, in the interfill periods, or in parallel with the HLT1 for the following fills. As a consequence of these important changes, the Run II online reconstruction is significantly improved compared to Run I, and is identical to the one performed offline.

Two types of HLT output (so-called "streams") were used:

- **Full stream**: entire raw event information is saved. This allows to reprocess the data at the later step, improving the reconstruction algorithms or selecting additional decay channels, if they passed some inclusive trigger. After the trigger, most of the data is passing through the full offline reconstruction. This means converting the raw detector hits in high-level objects such as tracks and clusters. Still, the resulting dataset is very large to be used directly for the analysis. This motivates an additional step – a set of preselection requirements which is called *Stripping*. The output data files are grouped in smaller *streams* depending on their final state and selection patterns. Finally, for a given analysis, only one or few particular decay channels are interesting, which are extracted by analysts from the output of the Stripping. The work presented in this thesis is based on the output of the Stripping framework.

- **Turbo stream**: was introduced in Run II in order to perform almost real-time analysis directly from the HLT2 candidates. There are several available configurations (saving

the entire event, saving only signal candidates, or saving signal candidates together with pre-defined tracks of interest) which have different event sizes. However, the event size in such framework is always smaller than in the full stream, as the raw detector information is not saved. Notably, this format was used to collect the dedicated samples for calibration of the particle identification calibration response in Run II, which will be discussed later.

2.6.3 Commonly used variables

Variables related to quantities measured by the LHCb detector are mentioned throughout this thesis, and are defined in Appendix A.

2.7 Electrons in LHCb

Reliable identification of electrons is crucial for the work presented in this thesis. Electrons behave in a very different manner with respect to muons, and understanding these differences is of paramount importance when performing lepton universality tests.

2.7.1 Calorimeter for electron identification

Identification of electrons is quite special compared to other charged particle species in LHCb. As it can be deduced from Fig. 2.11, due to the low mass of electrons, the RICH can be helpful to identify electrons only in the very low-momentum region. Additional information is required, based on the calorimeter information. Electrons leave quite a special signature in the calorimeter system (Fig. 2.14), leaving hits in the SPD (as any charged particle) and releasing most of their energy in the ECAL. The presence of the SPD hits allows to separate electrons from photons. Measuring the fraction of energy released in ECAL allows to distinguish electrons from charged hadrons, which release most of their energy in HCAL. To do so, tracking information is required. Looking for a track associated to the calorimeter deposit allows to determine the ratio of the deposited energy over the track momentum E/p, which is a powerful discriminating variable, as it is shown in Fig. 2.17(left). Indeed, for electrons this ratio is close to unity (modulo the resolution effects), while hadrons leave only a little fraction of their energy in the ECAL.

2.7.2 Bremsstrahlung

Electrons are very light particles, and therefore emit bremsstrahlung radiation in interactions with the material of the detector. Radiation is emitted colinearly to the electron direction. Bremsstrahlung emission complicates the reconstruction of electrons. It also affects the measurement of their momentum, which is performed by the track curvature in the magnetic field.

Fig. 2.19 illustrates the regions of the LHCb detector where electron interactions with the material happen. If the bremsstrahlung photon is emitted before the magnet (mainly

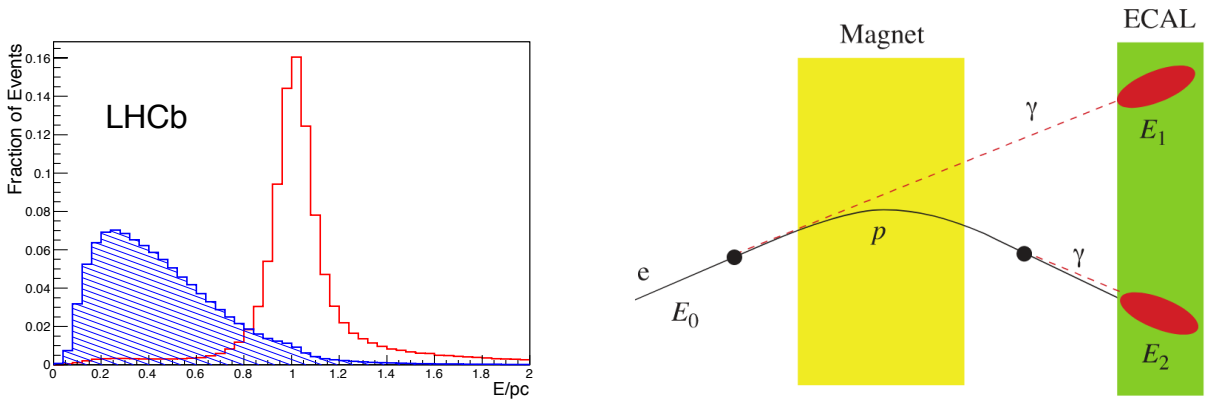


Figure 2.17: Left: Distribution of E/p for electrons (red) and hadrons (blue), as obtained from a part of the 2011 dataset. Taken from [102]. Right: Schematic representation of the bremsstrahlung photons emitted by an electron before and after the magnet. Taken from Ref. [101].

in interactions with the material of VELO, RICH1, TT, and the beam pipe), it propagates in the original direction of electron flight, while the electron trajectory itself is altered by the magnetic field. Therefore, the photon and electron end up in different cells of ECAL, as it is shown in Fig. 2.17 (right). The measured momentum of the electron is lower, as a part of it was transferred to the photon. At the same time, the energy released by the electron in ECAL is also lower for the same reason. This means that emitted bremsstrahlung does not affect the measured ratio E/p used for the electron identification.

If the photon is emitted after the magnet (but before the ECAL), the momentum measurement has already been performed before the bremsstrahlung emission and provides the original momentum of the electron. Photon and electron end up in the same ECAL cell, so that their combined energy release provides the original energy of the electron. Consequently, the ratio E/p is again not affected. Therefore, the bremsstrahlung emission after the magnet does not affect the electron reconstruction and identification.

Finally, a small amount of interactions can happen inside the magnet region (mainly in interactions with the air and the beam pipe supports). In this case, the bremsstrahlung photon is most probably lost, as the simple extrapolation from the region before the magnet will not give its correct position in ECAL. This case is usually neglected due to the small material budget of this region, compared to the one of the tracking system.

In order to improve the invariant mass resolution in decays involving electrons, a dedicated bremsstrahlung recovery algorithm is run within the LHCb software, for the electrons which radiate bremsstrahlung photons before the magnet. The `BremAdder` algorithm extrapolates the electron tracks from the region before the magnet to the ECAL, and performs a search for the ECAL clusters compatible with being bremsstrahlung photons. This procedure is illustrated in Fig. 2.18. ECAL measures only the energy deposited by the cluster. The origin of the bremsstrahlung photon is determined from the slope of the electron track, allowing to reconstruct the four-momentum of the photon. The four-momentum of the electron track is then corrected by the four-momentum of the photon. Another algorithm called `DiElectronMaker` is run in the Stripping for the decay

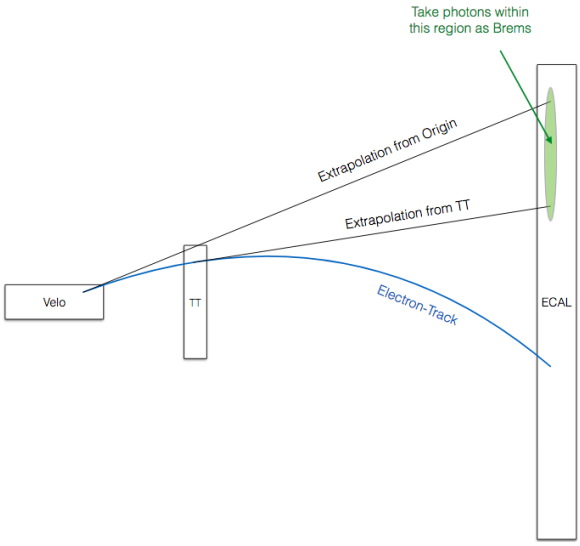


Figure 2.18: Schematic illustration of the `BremAdder` algorithm search window. Taken from Ref. [140].

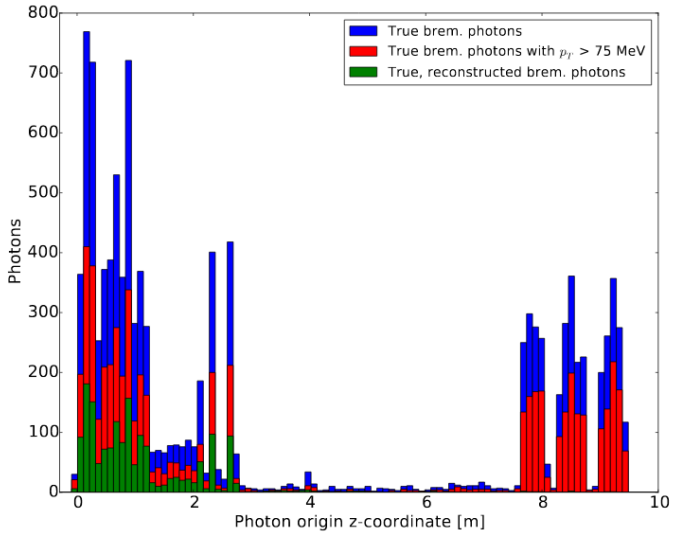


Figure 2.19: z-coordinates of bremsstrahlung photon origin in LHCb. Regions of VELO ($z < 1.5$ m), TT ($2 < z < 3$ m) and tracking stations T1-T3 ($7 < z < 10$ m) are clearly visible. Region of the magnet ($3 < z < 7$ m) has almost no material, except for the beam pipe, its support structures, and the air. All bremsstrahlung photons are shown in blue; potentially reconstructible photons (*i.e.* those having $p_T > 75$ MeV/c) are shown in red; and photons which are actually reconstructed properly are shown in green. Taken from Ref. [140].

channels containing the e^+e^- pair, in order to make sure that the same bremsstrahlung photon is not added to both electrons at the same time.

However, given the high occupancy in the electromagnetic calorimeter, it can be difficult to find the correct photon. Two extreme configurations can occur, when either

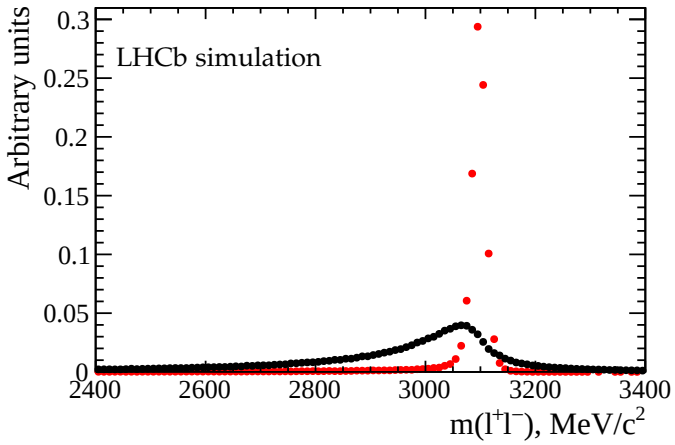


Figure 2.20: Invariant mass of $\ell^+\ell^-$ in the Run I simulation of $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$, where ℓ = (red) μ, or (black) e. Distributions are normalised to the unit area. The dielectron shape is wider and has long tails due to bremsstrahlung.

the bremsstrahlung photon is not found, or when a random photon is picked up as a bremsstrahlung photon. These two configurations lead to either underestimating or overestimating the electron momentum. This effect significantly degrades the resolution of the dielectron invariant mass, and consequently the b-hadron invariant mass resolution. Even in the case of perfect photon recovery, the resolution on the invariant mass will be degraded anyway due to the limited ECAL resolution, and the cut on the minimal photon transverse momentum ($p_T > 75$ MeV/c) below which the bremsstrahlung photons are not reconstructed in LHCb [140]. It can be seen from Fig. 2.19 that a significant fraction of the bremsstrahlung photons have a transverse momentum below this threshold, and therefore cannot be reconstructed.

Contrary to electrons, muons loose much smaller fraction of their energy when crossing the LHCb detector, and have therefore much better momentum resolution. As a consequence, the dielectron invariant mass (and so, the b-hadron invariant mass in the decays with electrons in the final state) has a shape very different for the one in the dimuon modes, as shown in Fig. 2.20.

Bremsstrahlung recovery provides a powerful discrimination variable for the electron identification. A track having bremsstrahlung photons associated is a specific signature of electrons. Fig. 2.21 compares the difference in likelihood between the electron and pion ID hypotheses, for the electrons in Run I data. It can be seen that the electrons with associated bremsstrahlung photons have a much more reliable identification.

2.8 Calibration of the particle identification response

In LHCb, there are several approaches for calibrating the particle identification (PID) response in the simulation. All of them rely on dedicated calibration samples recorded in parallel to the general data taking, and described in detail in Refs. [141,142]. The list of

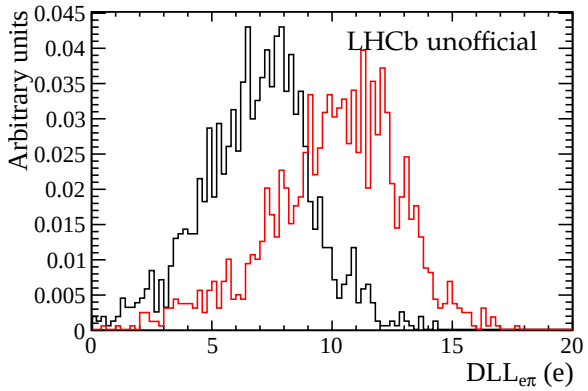


Figure 2.21: Distribution of $DLL_{e\pi}$ for electrons from $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ decay, in fully selected Run I data. Red: electrons which have at least one bremsstrahlung photon attached; black: electrons which have no photons attached.

principal calibration samples used for Run I and Run II datasets, is described in Table 2.2.

Table 2.2: PID calibration samples in LHCb.

Particle	Run I	Run II
p	$\Lambda \to p\pi^-$ $\Lambda_c^+ \to pK^-\pi^+$ inclusive	$\Lambda \to p\pi^-$ $\Lambda_c^+ \to pK^-\pi^+$ from semileptonic Λ_b^0 decays
K	$D^0 \to K^-\pi^+$ from $D^{*+} \to D^0\pi^+$	$D^0 \to K^-\pi^+$ from $D^{*+} \to D^0\pi^+$ $\phi \to K^+K^-$ from $D_s^+ \to \phi\pi$
μ	$J/\psi \to \mu^+\mu^-$ from B decays	$J/\psi \to \mu^+\mu^-$ inclusive $J/\psi \to \mu^+\mu^-$ from B decays
e	$J/\psi \to e^+e^-$ from B decays	$J/\psi \to e^+e^-$ from B decays

In this analysis, the `PIDCalib` package described in Ref. [143], was used. This tool allows to compute the efficiency of a certain set of PID requirements applied to a given track, using the corresponding calibration sample. This efficiency is subsequently applied as a per-event weight to the simulation samples. There are alternative techniques exploited in LHCb, allowing to resample the given PID variable in the simulation, described in Ref. [144], not covered in this work.

The usage of PID calibration samples and `PIDCalib` method rely on the tag-and-probe technique, similar to the one already discussed for tracking algorithms in Sec. 2.4. In a reasonably clean data sample of abundant two- or three-body decays, one particle is treated as a "tag" by applying certain tight requirements which ensure its correct identification. Another particle is treated as a "probe", and is used to determine the PID efficiency. The samples are statistically cleaned from the background contamination using the ${}_s\mathcal{P}lot$ technique described in Ref. [145]. To do so, invariant mass fits are centrally performed to the entire calibration sample. The PID efficiency is estimated, counting the number of events before and after applying the required PID cut, in the background-

subtracted sample. If the calibration sample is large enough, it is possible to study the PID efficiency in bins of certain kinematic variables. This results in a map of PID weights as a function of these variables.

2.8.1 Issues with calibration of the electron PID response

As it has been discussed above, the ${}_s\mathcal{P}lot$ tool is used in order to statistically subtract the background from calibration datasets. However, the ${}_s\mathcal{P}lot$ technique is strictly valid only when there is no correlation between the fitted parameter (invariant mass) and the studied observables [145]. However, the dielectron invariant mass correlates with the kinematic variables due to bremsstrahlung. The $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ signal shape is very different in various regions of phase-space, as it is shown in Fig. 2.22. It is also different depending on the number of bremsstrahlung photons associated with the probe electron track.

Another problem arises due to the different nature of the background in the cases of electrons which have no bremsstrahlung photons attached, and electrons which have at least one associated photon. As it has been already discussed, associating a photon with the electron track significantly improves the electron identification. An important background consisting of misidentified pions can be present almost exclusively for electrons which have no associated photons. However, at the time of writing this thesis, for the datasets used in this analysis, no separation of these two cases was performed in the `PIDCalib` tool. Instead, a single invariant mass fit was performed to the datasets with and without attached bremsstrahlung photons.

As a consequence of the problems discussed above, the `PIDCalib` tool is known to overestimate the particle identification efficiencies for electrons. This motivated the development of the alternative technique, discussed in the following section.

2.8.2 Fit-and-count technique

To estimate the PID efficiency for electrons, the alternative `fit-and-count` procedure was developed. While the approach described in the previous section relies on one single invariant mass fit to the entire calibration dataset, with background subtracted by means of the ${}_s\mathcal{P}lot$ technique, the fit-and-count approach implies repeating the invariant mass fits in each kinematic bin, and evaluating the PID efficiency directly from these fits. This technique is more reliable, as it allows to account properly for the difference in signal and background invariant mass shapes between kinematic regions. It also lifts the limitations of the ${}_s\mathcal{P}lot$ technique, which is not used. The procedure described below has been developed for this analysis. It is now in process of being released as a tool available for the entire LHCb Collaboration.

First, the default preselection was applied to the calibration samples, requiring a good identification of the kaon and the tag electron, as documented in Tab. 2.3.

The sample was split according to the number of bremsstrahlung photons added to the probe electron (zero, or at least one). These two subsamples were considered separately.

Table 2.3: Preselection applied to the electron calibration samples

Particle	Requirements
B^+	$\chi^2_{IP} < 9$
Kaon	`ProbNNk`> 0.2
Tag electron	$p > 6\,\text{GeV}/c$, $p_T > 1.5\,\text{GeV}/c$, PIDe > 5
Probe electron	`hasRich`, `hasCalo`, $p > 3\,\text{GeV}/c$, $p_T > 0.5\,\text{GeV}/c$
	`nSPDHits` < 600 (450) in Run I (II)

For each subsample, a simultaneous fit was performed to the two datasets: a) events passing the PID requirement, and b) events failing this requirement. The PID efficiency together with its uncertainty was obtained directly from this simultaneous fit.

The signal shape is described by a Bifurcated Crystal Ball probability density function (PDF), which consists of a Gaussian core with two power-law tails below a certain threshold. It can be defined using the following parameters: mean μ_B, width σ, and four tail parameters $\alpha_{L,R}, n_{L,R}$ – two for left (L) and right (R) tails. The free parameter of interest x is the invariant mass. The PDF is given by:

$$BCB(x, \alpha_{L,R}, n_{L,R}, \mu_B, \sigma) = N \times \begin{cases} A_L \times (B_L - \frac{x-\mu_B}{\sigma})^{-n_L}, \text{ for } \frac{x-\mu_B}{\sigma} \leq -|\alpha_L| \\ A_R \times (B_R - \frac{x-\mu_B}{\sigma})^{-n_R}, \text{ for } \frac{x-\mu_B}{\sigma} \leq -|\alpha_R| \\ \exp\left(-\frac{(x-\mu_B)^2}{2\sigma^2}\right), \text{ for } -|\alpha_L| \leq \frac{x-\mu_B}{\sigma} \leq -|\alpha_R| \end{cases} \tag{2.7}$$

where:

$$A_L = (\tfrac{n_L}{|\alpha_L|})^{n_L} \times \exp\left(-\tfrac{|\alpha_L|^2}{2}\right); \quad B_L = \tfrac{n_L}{|\alpha_L|} - |\alpha_L|;$$
$$A_R = (\tfrac{n_R}{|\alpha_R|})^{n_R} \times \exp\left(-\tfrac{|\alpha_R|^2}{2}\right); \quad B_R = \tfrac{n_R}{|\alpha_R|} - |\alpha_R|,$$

and N is a normalization factor.

The shape was tuned using the $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ simulation samples, with a selection applied which is identical to the one in the fitted dataset. Examples of resulting fits in different kinematic regions are shown in Fig. 2.22. The tail parameters were subsequently fixed in the fits to data.

The background shape consists of a combinatorial background, described by an exponential shape with a free slope and yield; and a partially reconstructed $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ component (with the pion from the K^{*0} decay not reconstructed) described by the kernel estimation method (`RooKeysPdf`) [146]. The latter invariant mass shape is obtained using the appropriate simulation samples, with the preselection requirements identical to the one in the fitted dataset. For both signal and background simulation, separate samples have been used for Run I and Run II conditions.

The fraction of yields between the partially reconstructed and signal components is required to be the same between the 'failed' and 'passed' samples. This is because within a given kinematic bin, there is no reason why the electron PID efficiency would be different for $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ decay modes.

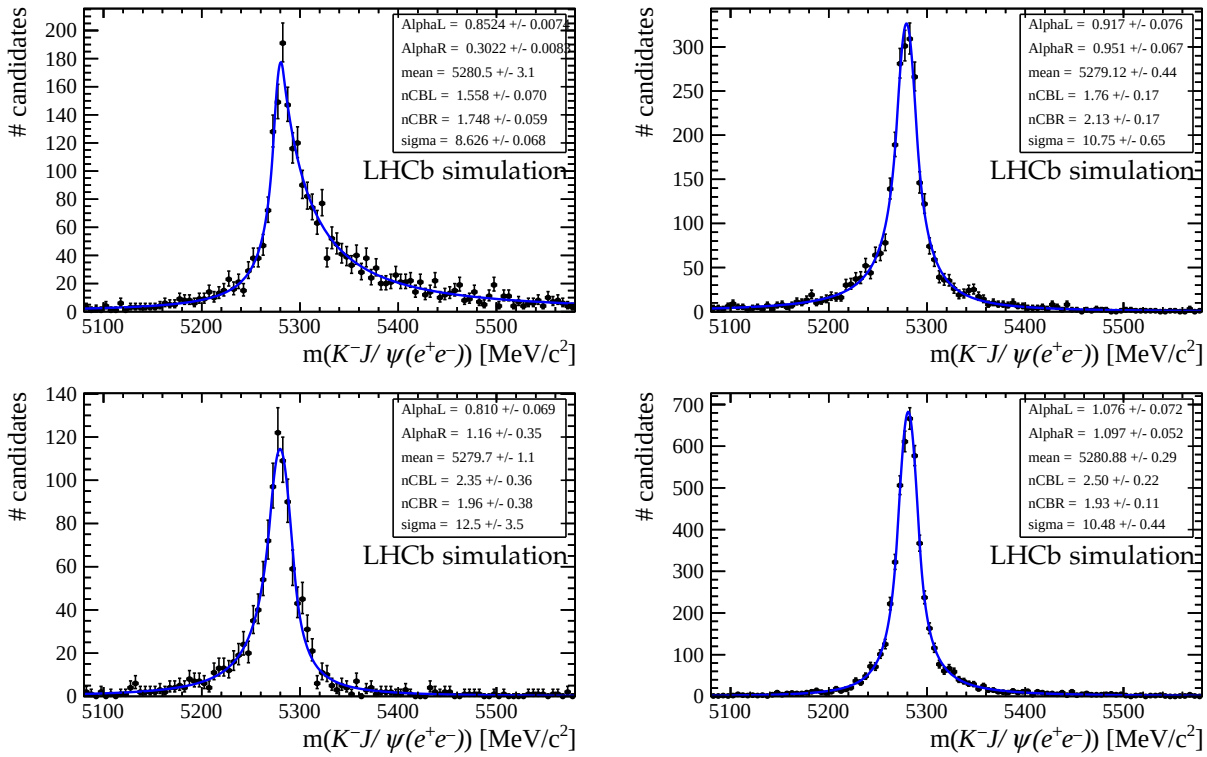


Figure 2.22: Invariant mass shape in the $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ simulation, at (left) low momentum and high pseudorapifity, and (right) high momentum and low pseudorapidity. Top row: probe electron has no associated photons; bottom row: probe electron has at least one photon associated.

An example of a simultaneous fit for 2016 calibration data, in one kinematic bin at central values of momentum and pseudorapidity, is shown in Fig. 2.23 for the category with zero bremsstrahlung photons added to the probe electron, and in Fig. 2.24 for the category with more than zero bremsstrahlung photons added to the probe electron. It can be seen that the category with 0 photons suffers from much higher background level. As it has already been discussed, this background is mainly populated by charged pions misidentified as electrons: requiring that the electron has bremsstrahlung photons attached removes most of the pion background. This also means that after applying the PID cut, this pion background will mostly disappear. So, the background conditions change drastically in this category before and after applying the PID cut. This is not accounted for when using the ${}_s\mathcal{P}lot$ of the calibration sample which has no PID cut applied.

The fit stability has been tested for different conditions. The corner-cases have been clarified and a special treatment has been implemented for them. For instance, if the fit hits the allowed boundary of a free parameter, the statistical uncertainty on the efficiency becomes unphysically small. This can happen, if the yield of the combinatorial background is small, and the slope of the exponential describing it hits the allowed boundary. A special check has been implemented for this case, and the fit is rerun again, after fixing the slope parameter.

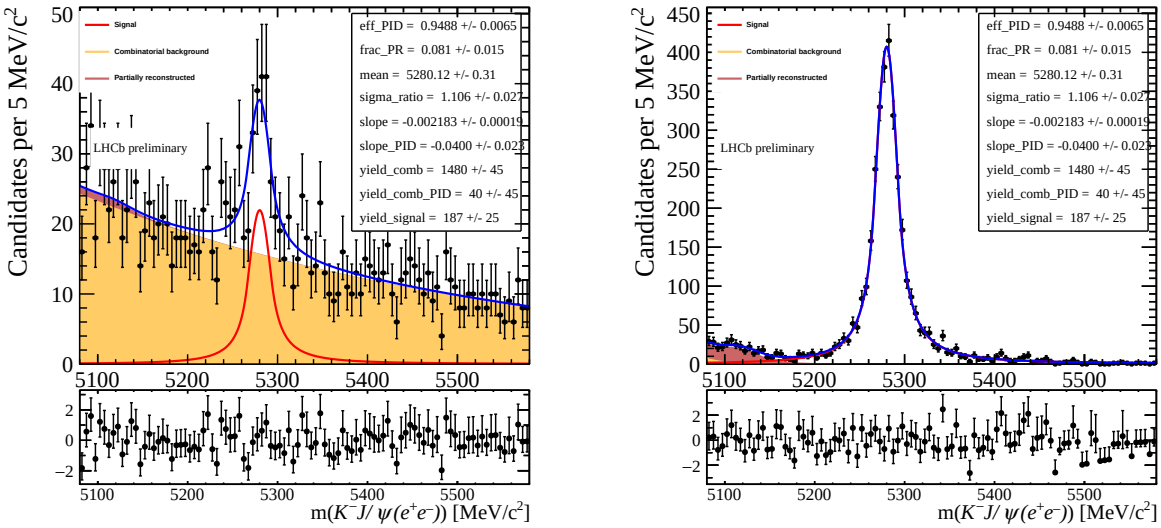


Figure 2.23: Example of simultaneous fits to 2016 calibration dataset, used to extract the fit-and-count electron PID efficiencies for electrons with 0 bremsstrahlung photons added (left – dataset failing to pass the PID requirement, right – dataset passing the PID requirement).

The difference between the fit-and-count efficiencies, and those obtained by the default method via ${}_s\mathcal{P}lot$ used in `PIDCalib`, is reported in Figs. 2.25-2.26, where the 2016 dataset is used as an example. It can be seen that the difference is larger for electrons having no bremsstrahlung photons associated, as in this case the background conditions are harsher, and are different for "passed" and "failed" categories. This means that in this case the ${}_s\mathcal{P}lot$ technique is not strictly valid. The same preselection as the one reported in Tab. 2.3 is used for the ${}_s\mathcal{P}lot$ method.

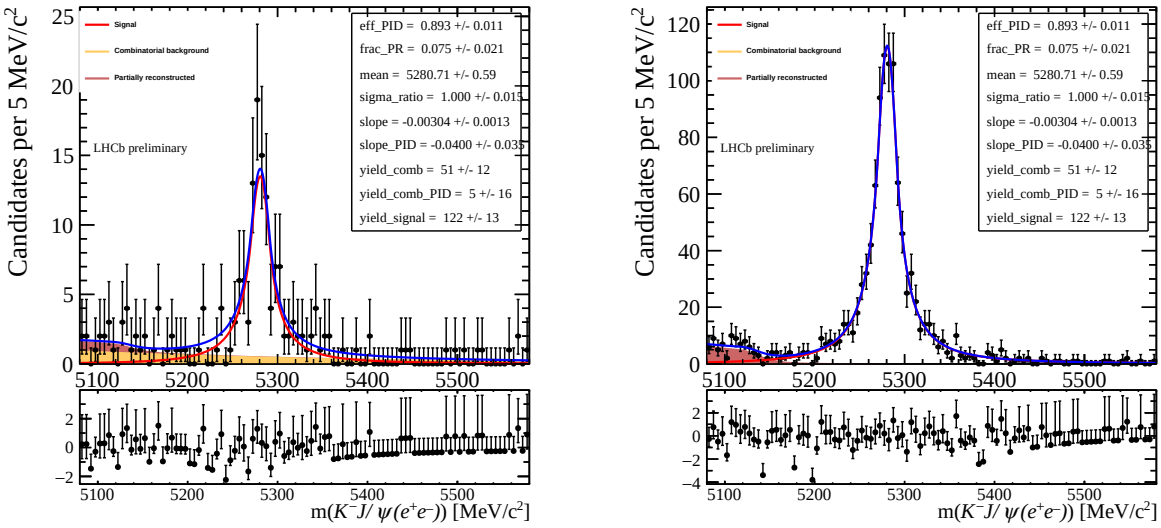
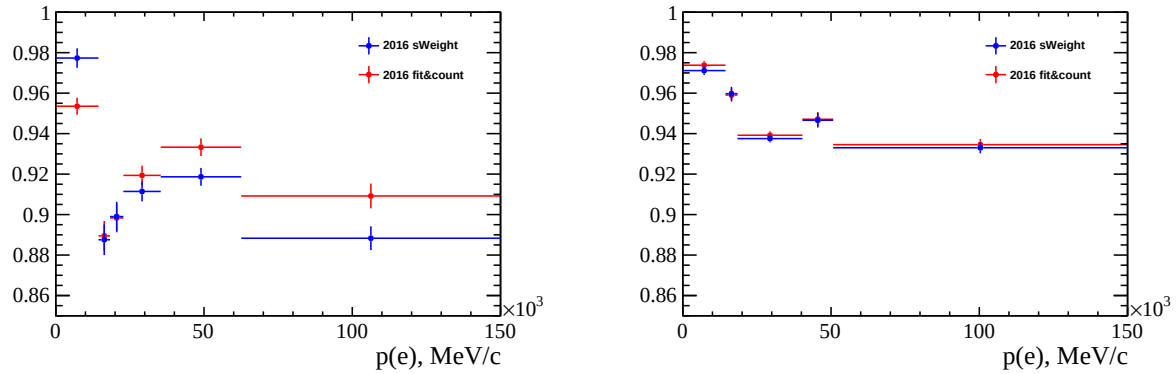
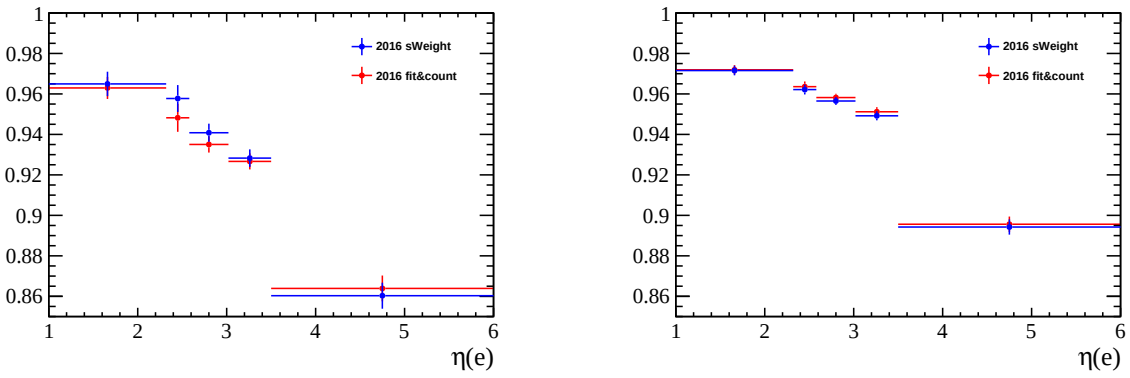


Figure 2.24: Example of simultaneous fits to 2016 calibration dataset, used to extract the fit-and-count electron PID efficiencies for electrons with >0 bremsstrahlung photons added (left – dataset failing to pass the PID requirement, right – dataset passing the PID requirement).



(a) 0 bremsstrahlung photons category (b) >0 bremsstrahlung photons category

Figure 2.25: Difference between the sWeight and fit-and-count PID efficiencies for electrons in 2016 dataset, as a function of electron momentum.



(a) 0 bremsstrahlung photons category (b) >0 bremsstrahlung photons category

Figure 2.26: Difference between the sWeight and fit-and-count PID efficiencies for electrons in 2016 dataset, as a function of electron pseudorapidity.

CHAPTER 3

Test of lepton universality using $\Lambda_b^0 \to pK\ell^+\ell^-$ decays

This chapter covers the measurement of the ratio R_{pK}^{-1}, which is the first test of lepton universality in rare b-baryon decays, using the $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ decays.

3.1 Strategy

In order to minimise the systematic uncertainties on the measurement of interest, R_{pK}^{-1} is measured using a double ratio defined as:

$$R_{pK}^{-1} = \frac{BR(\Lambda_b^0 \to pKe^+e^-)}{BR(\Lambda_b^0 \to pK\mu^+\mu^-)} \times \frac{BR(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}{BR(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))} \tag{3.1}$$

As it will be shown later, the yields of the rare electron mode $\Lambda_b^0 \to pKe^+e^-$ are significantly smaller than the ones in the rare muon mode $\Lambda_b^0 \to pK\mu^+\mu^-$. For this reason, it was decided to measure R_{pK}^{-1} and not R_{pK}, as having a smaller number in the numerator rather than the denominator ensures a better behaviour of the statistical uncertainties on the ratio of interest.

The approach of a double ratio relies on the well-tested lepton universality in the $J/\psi \to \ell^+\ell^-$ decays, with the $\mathcal{B}(J/\psi \to e^+e^-)/\mathcal{B}(J/\psi \to \mu^+\mu^-)$ being in agreement with one within sub-percent precision [6]:

$$\frac{\Gamma_{J/\psi \to e^+e^-}}{\Gamma_{J/\psi \to \mu^+\mu^-}} = 1.0016 \pm 0.0031 \tag{3.2}$$

Although having the same initial and final states as the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays, the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ transition is a tree-level decay, with a branching fraction larger by several orders of magnitude. The Feynman diagram for this transition is shown in Fig. 3.1.

Testing that in our dataset the lepton universality holds in the $\Lambda_b^0 \to pKJ/\psi$ decay, is a crucial cross-check of this analysis. The single ratio $r_{J/\psi}^{-1}$ is measured, as discussed in Sec. 3.8:

$$r_{J/\psi}^{-1} = \frac{BR(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))}{BR(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}. \tag{3.3}$$

3.1.1 Signal and normalisation modes

Both $\Lambda_b^0 \to pKe^+e^-$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ decays are referred to as `'signal'` or `'rare'` mode. Decays $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ and $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ are referred to as `'normalisation'` or `'resonant'` modes. Additional decay modes such as $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$, $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$ are the `'control'` modes used to validate various aspects of this analysis. Details of this validation can be found in Section 3.8. Throughout the manuscript of this thesis, depending on the context, both normalisation and control modes can be referred to as "signal" to discriminate them from various background modes discussed in the text. Fig. 3.2 features the invariant mass of $pK\ell^+\ell^-$ as a function of dilepton invariant mass, for muons and electrons. The two vertical bands corresponding to J/ψ and $\psi(2S)$ resonances, are well visible, and the horizontal band is the Λ_b^0 signal. Diagonal bands are due to the mass resolution and bremsstrahlung (for electrons).

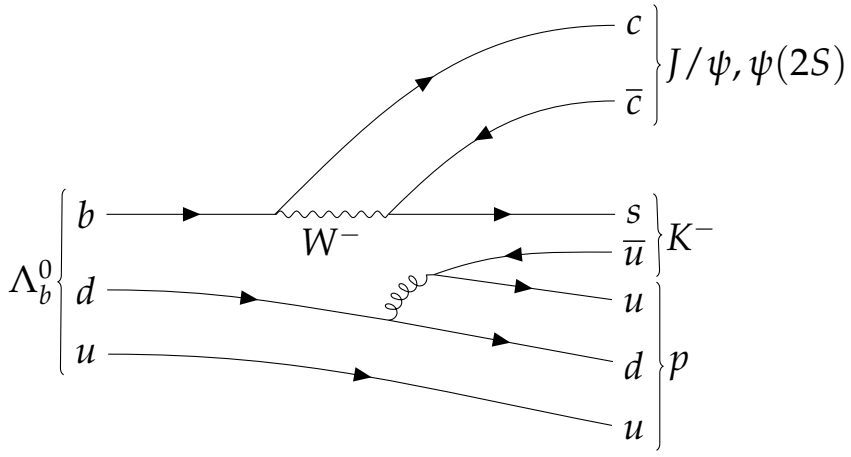
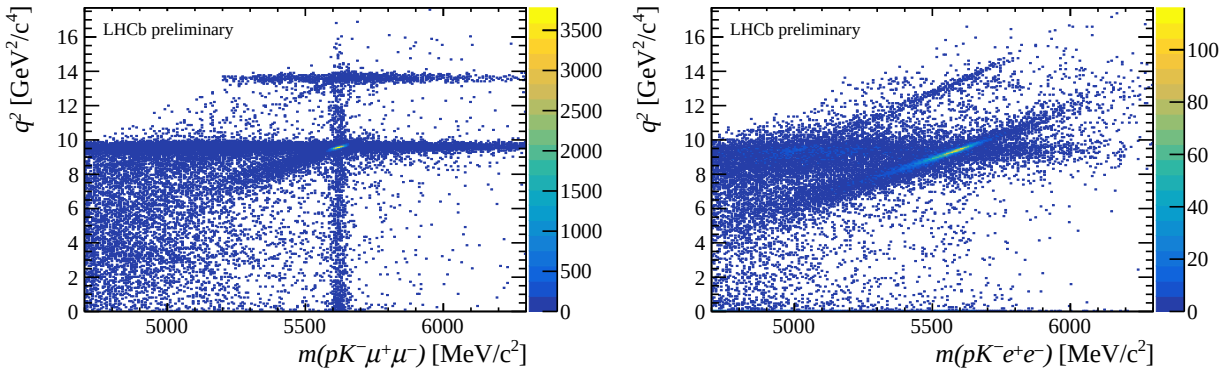


Figure 3.1: Feynman diagram of the decay $\Lambda_b^0 \to pK^-\psi$, where ψ denotes any charmonium state.

3.1.2 Definition of the dilepton invariant mass windows

The dilepton invariant mass windows in the scope of the lepton universality analyses, are usually defined in terms of q^2 variable, which corresponds to the dilepton invariant mass squared. The shape of the q^2 distribution in fully selected $\Lambda_b^0 \to pK\mu^+\mu^-$ data (in $\pm 45\,\mathrm{MeV}/c^2$ around the known Λ_b^0 mass) is shown in Fig. 3.3.



(a) Muons (b) Electrons

Figure 3.2: Invariant mass of $pK\ell^+\ell^-$ candidates as a function of dilepton invariant mass squared (q^2), in the data, for (a) $\ell = \mu$, (b) $\ell = e$. The complete dataset used for this analysis is exploited to prepare this plot. The complete selection is applied to both distributions, except for q^2 cuts and the HOP cut (Sec. 3.4.6). The right plot was unblinded after unblinding the fit to the $\Lambda_b^0 \to pKe^+e^-$ data.

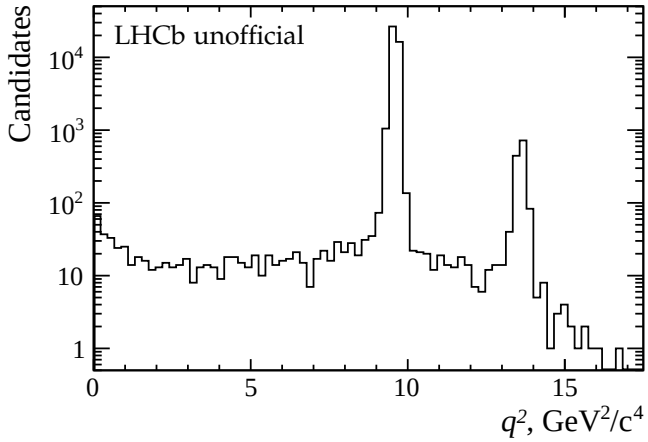


Figure 3.3: Shape of q^2 in the data (full Run I + Run II sample) under the $\Lambda_b^0 \to pK\mu^+\mu^-$ signal peak.

As it has already been discussed in Sec. 1.3.2, in previous LU analyses published by LHCb, R_K and $R_{K^{*0}}$, the following terminology was adopted: the q^2 bin between the $\phi(1020)$ and J/ψ resonances is referred to as `central`-q^2, the one between dimuon threshold (0.011025 GeV$^2/c^4$) and $\phi(1020)$ – `low`-q^2, and finally the region above the $\psi(2S)$ resonance is referred to as `high`-q^2. Theoretical calculations in Ref. [59] showed that closeness to the dimuon threshold can be a potential source of a bias, since the form factors describing the transition of interest do not cancel completely at very low q^2. Therefore, it was decided to increase the lower edge of the low-q^2 bin with respect to the previous $R_{K^{*0}}$ measurement. In this analysis, the lower cut is applied at 0.1 GeV$^2/c^4$.

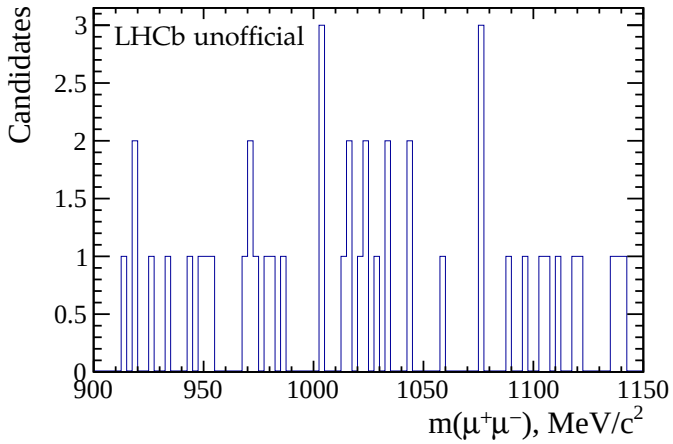


Figure 3.4: Invariant mass of $\mu^+\mu^-$ in the data (full Run I and Run II sample) under the $\Lambda_b^0 \to pK\mu^+\mu^-$ signal peak, zoomed in the region around the $\phi(1020)$ resonance

On the other hand, the $\phi(1020)$ resonance contribution is negligibly small, and no narrow peak is observed in the dimuon invariant mass plot around the $\phi(1020)$ mass region, shown in Fig. 3.4. The low and central q^2 bins are therefore merged to reduce the statistical uncertainty on the value of R_{pK}^{-1}: the current analysis is performed in the region $0.1 < q^2 < 6$ GeV$^2/c^4$.

At the same time, the kinematically limited phase-space of the decay leads to a small signal yield in the highest q^2 bin located above the $\psi(2S)$ resonance ($15 < q^2 < 17.5$ GeV$^2/c^4$). This can also be seen in Fig. 3.2(a) and Fig. 3.3, where only a few candidates are present above the $\psi(2S)$ resonance. Therefore, given the size of current datasets, this region is not expected to contribute significantly to the R_{pK}^{-1} measurement, and it was decided not analyse it.

Throughout this analysis, the following definitions of q^2 are used:

- Data and reconstructed-level simulation: $q^2 = |p_{\ell^+} + p_{\ell^-}|^2$, corresponding to a dilepton invariant mass squared;
- Generator-level simulation of the $\Lambda_b^0 \to pK\ell^+\ell^-$ modes: "true-q^2 " is defined as a difference of four-vectors $q^2_{\text{true}} = |p_{\Lambda_b^0} - p_{(pK)}|^2$.

The q^2_{true} is used to account for the final-state radiation of the leptons when computing signal efficiencies (Sec. 3.6.3), and make sure that the R_{pK}^{-1} is measured in the same exact q^2 region between the muon and electron modes, not being affected by the radiation and reconstruction effects.

3.1.3 Photon conversions

The decay $\Lambda_b^0 \to pK\gamma$ is expected to have a branching fraction about two orders of magnitude larger than the rare decays $\Lambda_b^0 \to pK\ell^+\ell^-$. It is known that high-energy

photons have a few-percent probability to be converted into a dielectron pair in the material of the detector. If this conversion happens inside the VELO, the event will have a topology similar to the $\Lambda_b^0 \to pKe^+e^-$ decay, with the dielectron mass being very close to zero (deviations from exactly zero are mainly driven by the resolution). This process creates a peak in the $m(pKe^+e^-)$ at very low q^2 (see Fig. 3.5), contaminating therefore the decay $\Lambda_b^0 \to pKe^+e^-$ in the q^2 region much below $0.1\,\text{GeV}^2/c^4$.

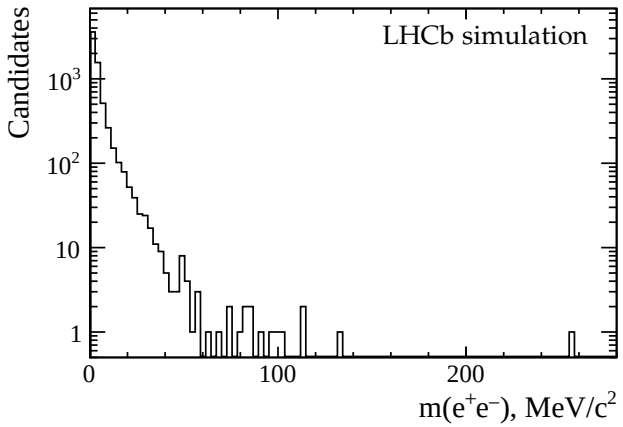


Figure 3.5: Distribution of $m(e^+e^-)$ in the simulation sample of $\Lambda_b^0 \to pK\gamma(\to e^+e^-)$ decay. It can be seen the photon conversions are located well below the lower threshold of $q^2 = 0.1\,\text{GeV}^2/c^4$ ($m(e^+e^-) \approx 316\,\text{MeV}/c^2$).

Given that this q^2 region is not used for the measurement of R_{pK}^{-1}, this contamination due to photon conversions is not causing a problem for the LU test.

3.1.4 On the $m(pK)$ invariant mass window

The kinematically allowed $m(pK)$ range extends from $m(p) + m(K) = 1432$ MeV/c^2 to the upper boundary which depends on the decay mode: for example, it is $m(\Lambda_b^0) - m(J/\psi) = 2523$ MeV/c^2 for the $\Lambda_b^0 \to pKJ/\psi$ decay, $m(\Lambda_b^0) - 2 \times m(\mu) = 5410$ MeV/c^2 and $m(\Lambda_b^0) - 2 \times m(e) = 5619$ MeV/c^2 for the $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ decays, respectively. However, the study performed in Ref. [4], shows that a simple phase-space decay model does not describe the $m(pK)$ distribution properly in the $\Lambda_b^0 \to pKJ/\psi$ decay. This is expected since this model does not include any resonant states. Instead, the physical model of the decay includes various Λ^* resonances in the $m(pK)$ spectrum, and several pentaquark candidates in the $m(pJ/\psi)$ spectrum (Fig. 1.12).

Moreover, it was observed that most of the signal candidates populate lower values of the $m(pK)$ distribution also for the decay $\Lambda_b^0 \to pK\mu^+\mu^-$, as it was shown in Fig. 1.13. This range is indeed dominated by various Λ^* resonances. The signal yield above 2.6 GeV/c^2 is very small, thus this region is background-dominated. In the current analysis, only the region $m(pK) < 2.6$ GeV/c^2 is considered.

The principal role of the LU test performed in this work is a null test of the SM. However, in order to be able to compare accurately our measurement of R_{pK}^{-1} with the

theoretical predictions for various New Physics models, one would need to be able to isolate a specific set of resonances in the pK spectrum, by performing an amplitude analysis similar to the one that allowed to discover the pentaquark candidates [4]. This would mean to be able to measure R_{pK}^{-1} separately for each Λ^* resonance, however given the limited size of the available datasets at the time of this analysis, such an exercise is not possible. This pilot analysis aims at measurement of an "inclusive" value of R_{pK}^{-1}, where no separation of the various Λ^* states will be attempted.

Finally, it should be noted that the measurement in Ref. [4] has recently been updated [5]. This update presents a more detailed investigation of the $m(pJ/\psi)$ spectrum with a larger dataset, and does not change any conclusion of the current analysis. These results are not directly used throughout the current analysis, as the update of the full amplitude model of the $\Lambda_b^0 \to pKJ/\psi$ decay has not been published at the moment.

3.1.5 Blinding and unblinding strategy

The complete analysis strategy has to be validated and frozen before looking at the most important results. This is done in order to avoid a potential bias, which could arise if the final result is matching (or not) our expectations. The procedure adopted to avoid such a bias is called "blinding".

The decay mode $\Lambda_b^0 \to pKe^+e^-$ has not been observed yet, so the signal window and the yield are kept blind throughout the analysis. The invariant mass window which is expected to contain about 90% of signal, is never plotted; and the signal yield is never shown. R_{pK}^{-1} is kept blind as well, and in addition to the blinded $\Lambda_b^0 \to pKe^+e^-$ signal yield, the $\Lambda_b^0 \to pKe^+e^-$ signal efficiencies are scaled by a random factor. This allows to perform the unblinding in steps – first, making sure that we indeed observe the $\Lambda_b^0 \to pKe^+e^-$ signal, and only as a second step unblinding the R_{pK}^{-1} itself.

In addition, a crucial cross-check – the single ratio $r_{J/\psi}^{-1}$ (defined in Eq. 3.3), is kept blind by scaling the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ efficiencies by a random number, until the entire analysis chain is settled, and all the consistency cross-checks are performed. It is also to be unblinded independently before unblinding the R_{pK}^{-1} itself.

To summarize, the unblinding is planned in three steps, with first two steps being independent of each other:

- Unblind the central value of $r_{J/\psi}^{-1}$ and make sure it is compatible with one and independent of the kinematical region, data taking period and other parameters;
- Unblind the fits to $\Lambda_b^0 \to pKe^+e^-$ data, make sure that a) we indeed observe the signal, b) the fits look reasonable and all the backgrounds are properly taken into account. Evaluate the systematics related to the fit model;
- Finally, unblind the efficiency of $\Lambda_b^0 \to pKe^+e^-$ and consequently the final value of the R_{pK}^{-1}. This can be done after the intermediate step: to unblind relative R_{pK}^{-1} values in Run I and Run II datasets separately, while keeping the absolute scale blind. This

extra step would allow to cross-check the compatibility between datasets before unblinding the final result.

By the time of writing this manuscript, the permissions to unblind the $r^{-1}_{J/\psi}$ and $\Lambda^0_b \to pKe^+e^-$ invariant mass fits, *i.e.* first two steps from the list, were granted. The R^{-1}_{pK} values and $\Lambda^0_b \to pKe^+e^-$ efficiencies are still blind.

3.2 Data and simulation samples

3.2.1 Data samples

The measurements in this work are performed using the LHCb data collected during Run I (2011-2012), corresponding to about 3 fb^{-1} taken at the center-of-mass energies of 7 and 8 TeV, and a part of the Run II data taking period (2016), corresponding to about 1.6 fb^{-1} taken at center-of-mass energy of 13 TeV where the b-production is larger. The data collected during 2017-2018 (3.8 fb^{-1} at 13 TeV) years are not used due to the lack of simulation samples. The expected statistics in the 2015 sample (corresponding to 0.3 fb^{-1} at 13 TeV) is about the half of the statistics in the 2011 sample, which makes the data-driven corrections to the simulation affected by large statistical uncertainties.

3.2.2 Simulation samples

Simulated data samples are used to describe and study the properties of the invariant $pK\ell^+\ell^-$ mass distributions of signal and background modes, used to parametrise signal shapes in the invariant mass fits (Sec. 3.7); to train multivariate classifiers designed to reject the combinatorial background (Sec. 3.4.3); and finally to compute the reconstruction and selection efficiencies (Sec. 3.6). In the context of the analysis, two different kinds of simulation samples are used. *Generation-level* simulation samples provide the distributions of kinematical variables as they were generated by the LHCb simulation framework, without accounting for any detector effects. *Reconstructed-level* simulation samples, on contrary, provide the distributions as they are expected to look in data, accounting for all detector effects. Both generation and reconstruction steps in the simulation can be slightly imprecise, calling the need for the dedicated data-simulation corrections described in the Sec. 3.5.

Table 3.1 summarises the simulation samples, used for the decay modes considered in this analysis as "signal" modes, in different dilepton invariant mass windows, including the samples used for the cross-checks. The numbers shown in this table correspond to the raw number of events used for the analysis. It should be noted that in order to save disk space, *filtering* scripts at the generation level were used for a part of the simulation production: only events passing the stripping requirements[1] (Sec. 3.2.6) have been saved to disk. This is accounted for when computing the signal efficiency, in Sec. 3.6.2.

[1] As it will be discussed below, particle identification cuts, present in the stripping lines, were removed

Table 3.1: List of the signal simulation samples. Two numbers shown for each cell represent the statistics available for the two magnet polarities (MagDown, MagUp).

Decay mode	2011	2012	2016
$\Lambda_b^0 \to pK\mu^+\mu^-$	303 972	689 570	525 188
	301 003	658 779	510 656
$\Lambda_b^0 \to pKe^+e^-$	601 886	1 011 851	1 013 677
	622 695	1 029 594	1 002 485
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	307 147	376 732	546 169
	313 002	372 986	500 363
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	609 939	655 149	1 002 649
	614 802	673 597	1 008 487
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	301 945	374 510	501 171
	308 929	355 496	594 895
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	672 769	659 248	501 827
	606 230	658 768	505 499

As it will be shown in Sec. 3.3, an important source of background originates from the misindentification of a certain type of particle with another type. Studies have shown that this source of background is mainly due to the misidentification of the proton with a kaon or a pion. Other important sources of background include partially reconstructed (including semileptonic) backgrounds, which are described in details in Sections 3.3.1.d and 3.3.2.d. To model the shape of all these backgrounds in the invariant mass fits, dedicated simulation samples have been used.

3.2.3 Trigger

As it has already been discussed, the LHCb trigger system is based on a hardware level (L0), and a software one made of two stages (HLT1 and HLT2). Because of the different interaction with the detector, the muon and electron channels are recorded and selected with different triggers paths.

3.2.4 Hardware trigger (L0)

In the L0 trigger selection, several trigger "categories" are considered, depending on the particle which triggered the event.

The L0Lepton (L0L) trigger category, named correspondingly `L0Muon` (denoted throughout the text as L0M) or `L0Electron` (L0E) for muons and electrons, requires the event be triggered on at least one of two leptons in the decay.

The fact that the event is triggered by at least one of the particles of the signal mode, is usually denoted as "TOS" (Triggered On Signal). For muons, this trigger

from the filtering scripts.

decision is defined as L1_L0MuonDecision_TOS || L2_L0MuonDecision_TOS, where L1 and L2 correspond to the two leptons. Similarly, for electrons, L1_L0ElectronDecision_TOS || L2_L0ElectronDecision_TOS decision are used.

Other trigger categories can be considered to increase the statistical power of the measurement, for example, L0TIS (L0I), when the event is Triggered Independently of the Signal decay (TIS). The L0TIS category is defined as Lambdab_L0Global_TIS, where "Global" means considering the decisions of all available triggers.

Throughout the text, the following terminology is adopted: inclusive trigger categories are denoted as L0M, L0E, L0I *etc.*, and include all the events triggered by the corresponding L0 trigger (even if this same event was in addition triggered by another trigger); and exclusive trigger categories, denoted as L0M!, L0E!, L0I! *etc.*, which require the events to be triggered by the given trigger, but not by the one already defined before. For example, if the first category defined is L0I, defining an exclusive L0E! category means requiring the event to be triggered by the L0Electron but not by the L0TIS trigger.

In this analysis, the following trigger categories are considered to measure R_{pK}^{-1}:

- For the muon modes: L0M;
- For the electron modes: L0I and L0E!.

The need to include more than one trigger category for the electron modes is explained by the relatively low L0Electron trigger efficiency (mainly due to the high hardware trigger thresholds, put in order to fit the output rate from the busy ECAL into the available bandwidth). This is not the case for the muon mode where the low occupancy of the muon stations allows for relatively loose thresholds, consequently providing larger statistics.

It can be seen that the two categories as defined for this analysis for the electron modes are mutually exclusive. With this definition of the two electron categories, they are expected to provide comparably equivalent statistics, thus allowing to obtain two statistically independent measurements of R_{pK}^{-1} (one for L0I and one for L0E!).

In addition, complementary categories (such as L0I for muon modes, or inclusive L0E for electron modes) may be used to compute data-simulation corrections for the trigger response (Sec. 3.5.6) or for the cross-checks where needed. They are not used to evaluate final results.

3.2.5 Software trigger (HLT)

High Level Trigger (HLT) selection consists of the HLT1 and HLT2 stages. HLT1 trigger runs in parallel to the data taking, and relies on rather simple and fast decisions. HLT2 stage can be run later and has a certain time margin, so allows for a more complicated event selections. The HLT selections are called "lines" throughout this text. If several lines are used, it implies the TOS requirement on each and the logical OR applied between them.

It was decided to keep the HLT selection as close as possible between the electron and muon modes, and Run I and Run II, to avoid potential bias. Some of the HLT selections

Electrons	Muons
HLT1 in Run I	
TrackAllL0	TrackAllL0
HLT1 in Run II	
TrackMVA	TrackMVA
HLT2 in Run I	
Topo(2,3,4)BodyBBDT	Topo(2,3,4)BodyBBDT
TopoE(2,3,4)BodyBBDT	TopoMu(2,3,4)BodyBBDT
HLT2 in Run II	
Topo(2,3,4)Body	Topo(2,3,4)Body
TopoE(2,3,4)Body	TopoMu(2,3,4)Body

Table 3.2: Summary of HLT lines used in the measurements. See explanations in the text.

have been re-optimised between Run I and Run II, leading to the larger signal efficiency in Run II. The list of HLT lines used is summarised in the Table 3.2, and is discussed in more detail below.

At the HLT1 step, the same trigger selection has been used for both muon and electron modes, based on generic track information. The HLT2 selection relies mostly on so-called topological lines (Topo), using the multitrack event topology. They are relying on the fast multivariate algorithms, such as bonsai BDT (BBDT) described in Ref. [147]. These algorithms use kinematic and topological variables as input, in order to select inclusively $B \to N$-body decays, where N can be 2, 3 or 4. There is a dedicated set of topological lines for decays involving muons or electrons: TopoE, TopoMu. The efficiency of these lines is increased by relaxing the requirement on the multivariate discriminant whenever at least one of the tracks associated with the topological vertex is identified as a muon or electron [137].

3.2.6 Stripping

Dedicated stripping lines named Bu2LLK[2] are defined for the lepton universality measurements. The stripping line used for dimuon channel is Bu2LLK_mmLine, while for dielectron channel the line Bu2LLK_eeLine2 is used. The latter relies on an algorithm called DiElectronMaker described in Sec. 2.7.2. The stripping cuts are identical between the different years of the data taking, and are reviewed in the Table 3.3.

[2]The name is slightly misleading, as this stripping line includes selections for several $b \to s\ell^+\ell^-$ processes, not only $B^+ \to K^+\ell^+\ell^-$. Some selections are common between channels, that explains why the mass window in Tab. 3.3 is defined around the B^+ mass.

Table 3.3: Summary of the requirements of the `Bu2LLKmmLine` and `Bu2LLKeeLine2` stripping lines.

	Requirement
Event	`nSPDHits` < 600
Λ_b^0	$\lvert m - m_{B^+}^{PDG} \rvert < 1500\,\text{MeV}/c^2$
	`DIRA` > 0.9995
	$\chi^2_{\text{IP}}(primary) < 25$
	end vertex $\chi^2/\text{ndf} < 9$
	$\chi^2_{DV \leftrightarrow PV} > 100$
(pK) system	$m < 2600\,\text{MeV}/c^2$
	$p_\text{T} > 400\,\text{MeV}/c$
	$\chi^2_{\text{IP}}(primary) > 9$
	origin vertex $\chi^2/\text{ndf} < 25$
K	$\texttt{PID}_K > -5$
	$\chi^2_{\text{IP}}(primary) > 9$
p	$\texttt{PID}_p > -5$
	$\chi^2_{\text{IP}}(primary) > 9$
$\ell^+\ell^-$ system	$m < 5500\,\text{MeV}/c^2$
	end vertex $\chi^2/\text{ndf} < 9$
	$\chi^2_{DV \leftrightarrow PV} > 16$
μ	`hasMuon && isMuon`
	$p_\text{T} > 300\,\text{MeV}/c$
	$\chi^2_{\text{IP}}(primary) > 9$
e	$\texttt{PID}_e > 0$
	$p_\text{T} > 300\,\text{MeV}/c$
	$\chi^2_{\text{IP}}(primary) > 9$

3.3 Background studies and signal selection

When studying rare decays, it is crucial to understand the background components that can pollute the signal. This implies understanding the origin of these backgrounds, quantifying their amount, describing their invariant mass shape, and finally identifying all the possible ways of minimising their impact on the signal extraction. In this section, the major background contributions to the signal and control modes are studied using both data and simulation samples.

A detailed description of all the identified backgrounds is given in this section. As a conclusion of this section, the actions taken to suppress them, such as cuts or vetoes, are described in Sec. 3.4.

The dominant type of background is the one originating from the combination of random tracks in the detector, this source is referred to as combinatorial background. Its properties vary between the different q^2 windows. Other important background sources are so-called misidentifications; partially reconstructed decays; and tree-level decays having exactly the same final state as signal modes. Because of their abundance, tree-level decays represent the dominant source of background for most of the decay modes of interest.

Throughout this section, lepton universality is assumed to be conserved for tree-level decays. Given the fact that the amount of expected signal candidates for the $\Lambda_b^0 \to pKe^+e^-$ mode is much lower with respect to the $\Lambda_b^0 \to pK\mu^+\mu^-$ mode; and that the mass resolution is significantly worse, the majority of the background studies are performed on the muon modes.

The strategy adopted is to carefully study all the dominant background components using the $\Lambda_b^0 \to pK\mu^+\mu^-$ data, keeping blind the $\Lambda_b^0 \to pKe^+e^-$ channel, and then simply transfer the acquired knowledge of the background composition from the muon mode to the electron mode, while taking into account the specificities of the electron mode, for example the worse mass resolution or the presence of $\gamma(ee)$ conversions.

However, one has to account for possible background components which could appear only in the electron mode, these could be due to specific features of the electrons such as bremsstrahlung.

Certain specific background sources are common to all the q^2 regions, while others are significant only to some of them. Therefore, for clarity purposes, this section is split in several subsections describing the background composition for each q^2 region separately.

3.3.1 Background composition in $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ normalisation modes

3.3.1.a Combinatorial background

In the J/ψ window, the combinatorial background originates predominantly from the combination of a true J/ψ with random hadron tracks. It is continuously distributed above the invariant mass threshold of $m(p) + m(K) + m(J/\psi) = 4530$ MeV/c^2. Multivariate techniques (MVA) are used in order to suppress the combinatorial background, as

discussed later in Sec. 3.4.3.

3.3.1.b Hadron misidentification backgrounds

One of the dominant sources of background emerges from misidentifying hadrons (protons, kaons and pions) between themselves. As it has been explained in Sec. 2.5, RICH detectors (together with the information with other subdetectors) allow to obtain a likelihood for a given track to correspond to a given particle species. This information is then used to assign a mass hypothesis to each final state particle, and compute the invariant mass of the parent particle. However, as it was mentioned, there are kinematic regions (at low and high momentum) where a reliable separation of different hadrons is not possible. Fig. 2.11 suggests that the proton has the highest probability to be misidentified: it has the highest momentum threshold for the creation of a Chereknov ring, so low-momentum protons (below the thereshold) have no reliable identification. However, low-momentum kaons can also be misidentified. As it was already discussed in the previous sections, the production of Λ_b^0 is smaller than the one of B mesons, while the branching fractions of the Λ_b^0 decays of interest are also smaller than the ones of similar B^0 and B_s^0 decays. This makes the contributions of misidentification backgrounds very important compared to the signal yield. As a consequence, one has to consider carefully the possible backgrounds emerging due to a wrong identification of final state hadrons. The dominant contributions are expected from the proton single misidentification due to a high RICH threshold on proton, while identification of kaon is more reliable. Single misidentifications should also happen more often than the double ones.

Having two hadrons (proton and kaon) in the final state, in general, eight possible misID combinations can be created for the pK^- final state, namely: K^+K^-, π^+K^-, $p\pi^-$, $p\bar{p}$, $K^+\bar{p}$, $K^+\pi^-$, $\pi^+\bar{p}$ and $\pi^+\pi^-$. In terms of specific decay modes, the summary of possible decays contributing is shown in the Table 3.4.

Table 3.4: Possible hadron misidentification decays. Backgrounds which have the largest contribution, are captioned as "important", and are further taken into account during the analysis, by including them in invariant mass fits (Sec. 3.7).

MisID	Decay	Comments
$K \to p$	$B_s^0 \to J/\psi K^+K^-$	Important
$\pi \to p$	$B^0 \to J/\psi K^+\pi^-$	Important
$\pi \to p$	$B_s^0 \to J/\psi K^+\pi^-$	CS[3], negligible
$p \to K$	$B^0 \to J/\psi p\bar{p}$	Rare, negligible
$p \to K$	$B_s^0 \to J/\psi p\bar{p}$	Rare, negligible
$\pi \to K$	$\Lambda_b^0 \to J/\psi p\pi$	CS, negligible
$p \leftrightarrow K$ swap ID	$\Lambda_b^0 \to J/\psi pK$	Important
$p \to K$ and $\pi \to p$	$\Lambda_b^0 \to J/\psi p\pi$	CS, negligible
$\pi \to K$ and $\pi \to p$	$B^0_{(s)} \to J/\psi \pi^+\pi^-$	Small, neglected
$\pi \to K$ and $K \to p$	$B^0 \to J/\psi K^+\pi^-$	Small, neglected

Another prominent example is the one of the $\Lambda_b^0 \to \Lambda J/\psi(\to \ell^+\ell^-)$ decay, with consequent weak $\Lambda \to p\pi^-$ transition, with a pion misidentified as a kaon. This one is not observed in our data for the reason that Λ is a long-lived particle which flies a significant distance in the detector before decaying weakly, compared to very short-lived Λ^* resonances in the pK final state. The two-body (pK) invariant mass, with an assumption that the kaon was really a pion, does not show any sign of a peaking structure around the Λ mass.

Illustration methods of misidentification backgrounds

There are several tests which can be performed in order to determine the presence of backgrounds due to the misidentification of a hadron.

One can start from two-body decays of a kind $A \to B + C$. In the scope of this analysis, this can be $\Lambda^* \to pK$ decay, which has to be separated from various misidentification modes such as $\phi(1020) \to KK$.

One approach which can be used to separate different two-body decays is the so-called Armenteros-Podolanski plot [148]. It is useful to visualise different two-body decay modes present in the dataset. Fig. 3.6 features the transverse momentum of the proton with respect to the Λ^*, as a function of the asymmetry of longitudinal (with respect to Λ^*) momenta of proton and kaon. The precise mathematical formulation of this relation is given in Appendix B. The plot discussed here (as well as most plots throughout this section) is prepared using the 2012 $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ dataset, with loose preselection requirements applied (minimal PID cuts `X_ProbNNX>0.1` on top of the stripping selection and a very loose BDT cut at -0.2 (Sec. 3.4.3)). In this plot, several arc-like structures can be seen. The most prominent one, symmetric around zero in x-axis, is due to the narrow $\phi(1020)$ resonance decaying to two kaons. other symmetric structures around zero are due to higher meson states decaying to K^+K^-, notably the $f_2'(1525)$ resonance. The broad arc displaced to the left originates from $K^*(892)$ resonance decaying to a pion and a kaon (so having negative mass asymmetry between decay products). Finally, the arcs which are displaced to the right, correspond to various Λ^* resonances decaying to a proton and a kaon, and having positive mass asymmetry between decay products. The only prominent of them corresponding to the narrowest $\Lambda^*(1520)$ resonance, while more broad resonances are not clearly visible in this plot.

One useful conclusion which can be drawn from this study is that the misidentified $\phi(1020) \to K^+K^-$ decay is exceptionally prominent in our data, and has to be suppressed. In the following discussion, this resonance is vetoed applying a requirement on the proton-kaon invariant mass, with a proton having a kaon mass hypothesis: $|m(pK)_{p\to K} - 1020| > 12\,\mathrm{MeV}/c^2$. The mass window of 12 MeV/c^2 is about three times the width of the $\phi(1020)$ invariant mass peak (which is driven by both natural width of $\phi(1020)$ meson and experimental resolution). This is a very powerful requirement suppressing not only fully-reconstructed, but also partially reconstructed B_s^0 decays with a ϕ meson in a final state. It is about 99% efficient for all signal modes, removing about 90% of the ϕ candidates. The

[3] Cabibbo-Suppressed decays.

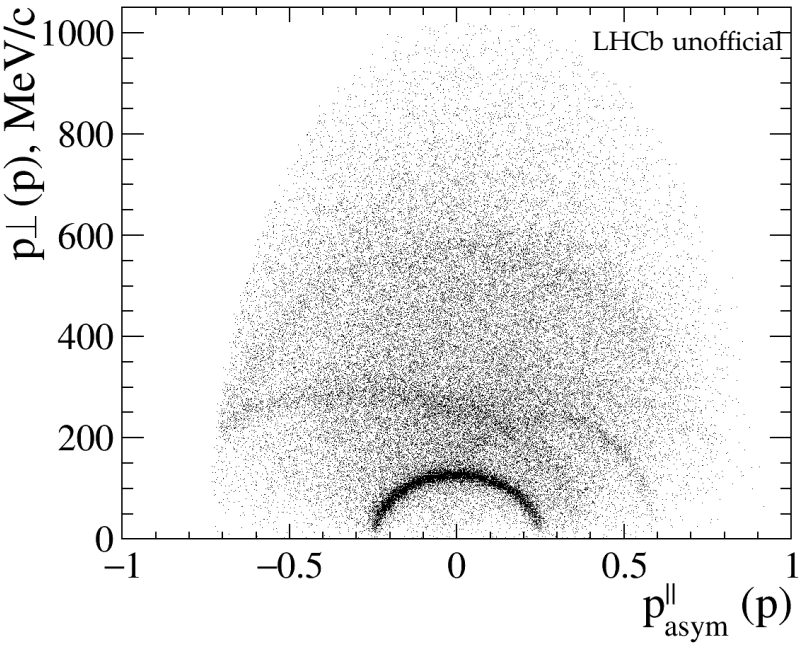


Figure 3.6: Armenteros-Podolanski plot for the two-body $\Lambda^* \to pK$ decay. Vertical axis features the transverse momentum of the proton with respect to the Λ^*, while the horizontal axis – the asymmetry of longitudinal (with respect to Λ^*) momenta of proton and kaon

remaining candidates are covered by the $KK\ell\ell$ invariant mass shapes in the fits.

A similar study can be performed for the three-body decays such as $\Lambda_b^0 \to pKJ/\psi$, as it is discussed in more detail in the Appendix B.

One has to bear in mind that the pK mass window used in this analysis is quite large, about 1300 MeV/c^2 wide. This means this analysis is missing a handle of a narrow hadronic resonance which could be useful to suppress misidentification contributions by limiting the allowed phase-space. The first step of the analysis will consist in suppressing these misidentification backgrounds as much as possible, and then accounting or any remaining ones with dedicated components in the final fits to invariant mass.

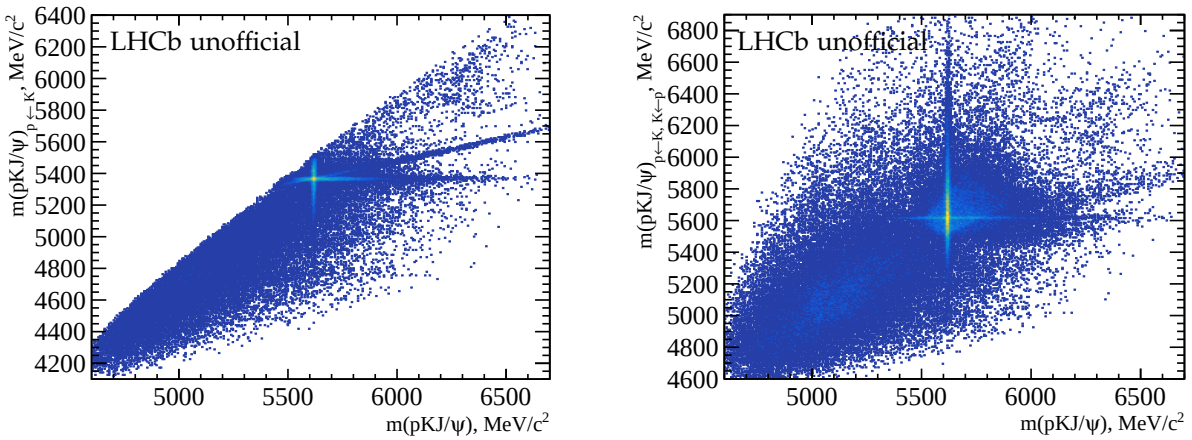
Substitution of the mass hypothesis

A thorough and complete investigation of the misidentification backgrounds can be performed using substitutions of the mass hypothesis. Applying a mass substitution means re-computing the invariant mass of the Λ_b^0, or of two- and three-body compositions, assuming a different mass value for a given particle.

First, a search is performed for all the possible single misidentifications in the hadron system. Then, the less pronounced double misidentification modes are looked for. To improve the mass resolution, a J/ψ mass constraint is applied throughout these studies. This means replacing the dilepton invariant mass by the known value of the J/ψ mass.

For simplification, only the background modes found to be dominant, are described here. In all the figures in this section, only a loose PID selection is applied (`X_ProbNNX>0.1`). This loose PID selection is used to figure out the dominant misidentification background contributions surviving it, and justify a further tightening of the PID cuts (or applying alternative selection requirements such as mass vetoes). Fig. 3.7(a) features the invariant mass of pKJ/ψ with proton having the mass hypothesis of a kaon, compared to the normal invariant mass of pKJ/ψ. Except for signal (visible as a vertical band), two misidentified contributions are clearly prominent: $B_s^0 \to K^+K^-J/\psi$ (horizontal band) and $B^0 \to K\pi J/\psi$ (diagonal band).

Fig. 3.7(b) features the invariant mass of pKJ/ψ with proton and kaon having mass hypotheses swapped, compared to the normal invariant mass of pKJ/ψ. Besides the signal visible as a vertical band, the horizontal band can be seen, corresponding to the signal with the swap of proton and kaon mass hypotheses through a double misidentification.



(a) (b)

Figure 3.7: Dominant hadron misidentification backgrounds in the 2012 $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset. Vertical band is a signal, horizontal and diagonal bands are misidentification backgrounds. See detailed description in the text.

Other backgrounds coming from misidentification contributions are negligible. The most prominent of them, the double misidentification background $B^0 \to K^{*0}J/\psi$ with a pion misidentified as a kaon, and a kaon as a proton, can be only visible after vetoing single misID modes.

For what concerns the hadron system in the $B_s^0 \to K^+K^-J/\psi$ and $B^0 \to K\pi J/\psi$ background channels, one should consider their rich resonant structure, a glimpse on which could already been seen in Fig. 3.6. Fig. 3.8 features the two-hadron (K^+K^- or $K^+\pi^-$) invariant mass distributions in 50 MeV/c^2 mass window around the (B_s^0 or B^0) invariant mass peak. For these studies, in order to suppress cross-feeds from misidentifications, a 25 MeV/c^2 veto is applied on the J/ψ-constrained Λ_b^0 invariant mass, as well as on the J/ψ-constrained B_s^0 invariant mass when studying the hadron system

in B^0 decays; or on the J/ψ-constrained B^0 invariant mass when studying the hadron system in B_s^0 decays. In the $B_s^0 \to K^+K^-J/\psi$ decay, one can indeed observe a rich resonant structure consistent with the one observed in [149]: on top of the very prominent and narrow $\phi(1020)$ resonance (vetoed in this analysis, but a part of its upper tail survives), some broader structures are seen, notably the broader $f_2'(1525)$ meson. The situation is quite different for the $B^0 \to K\pi J/\psi$ decay, where the K^{*0} resonance is dominating the landscape (as it is not vetoed), with only minor contributions of higher states such as $K_2^*(1430)$.

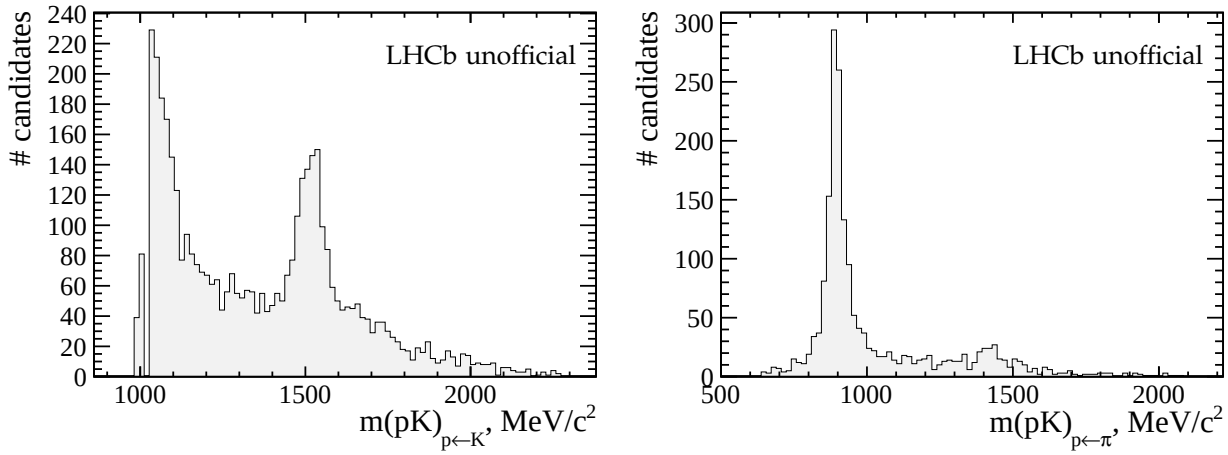


Figure 3.8: The two-body hadronic invariant mass under the peak of a) $B_s^0 \to K^+K^-J/\psi$ (with the $\phi(1020)$ resonance vetoed), b) $B^0 \to K\pi J/\psi$, in the 2012 $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset.

As a conclusion, particle identification requirements are necessary to suppress the hadron misidentification backgrounds. The remaining background contaminations have to be described in the invariant mass fit to data.

3.3.1.c Lepton misidentification backgrounds

Lepton misidentifications are expected to be negligible in the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ mode, given that the narrow J/ψ peak proves a clean dimuon signature. This is more complicated for the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, where the dielectron mass window is much larger, therefore some amount of non-J/ψ backgrounds, described in the Section 3.3.2.c, such as semileptonic decays, might penetrate inside the J/ψ window. The most important of them, arise from decays with Λ_c^+ particles in the final state (further discussed in Sec. 3.3.2.c), which can indeed be seen in the dielectron data sample within the J/ψ window. In all q^2 windows, it is easy to veto these backgrounds without a significant signal efficiency loss (see Sec. 3.6). For consistency, this veto is applied to both the electron and muon samples (see Sec.3.4).

3.3.1.d Partially reconstructed backgrounds

Another important source of background is due to multibody decays, where some of the final particles are lost because they were not reconstructed or they fell out of the detector acceptance. The topology of these decays can be almost identical to the decay of interest.

Often these lost particles are neutral (photon, neutral pions), but it can also be a pair of charged particles. It is nearly impossible to have only one charged pion lost in a decay, since there is no weakly-decaying positively-charged b baryons which could decay to $pK^-\pi^+\ell^+\ell^-$ final state. Furthermore, any potential decays to $pK^-\pi^-\ell^+\ell^-$ final state are suppressed by both the Cabibbo suppression on one hand and the small production rate of the Ξ_b^- and Ω_b^- baryons on the other hand. Lost particles heavier than pions are not considered here, since such background candidates will be located significantly further away from the signal peak in the invariant mass distribution.

There are two classes of partially reconstructed backgrounds: those emerging from the hadronic system (pK) and those emerging from the leptonic system. The latter can be observed only in the resonant modes, in particular $\Lambda_b^0 \to pKJ/\psi$, while the first one had to be considered for both resonant and rare decay modes.

Partially reconstructed backgrounds from the hadron system

For what concerns partially reconstructed backgrounds of hadronic nature, two cases are considered as important: loosing a neutral pion, or loosing a photon. Non-resonant multi-body decays $\Lambda_b^0 \to pKJ/\psi\,\pi^0$ or $\Lambda_b^0 \to pKJ/\psi\,\gamma$ are expected to be very suppressed, however there are quite important possibilities of such final states emerging from excited hadronic resonances. The most prominent contribution might arise from the decay $\Lambda_b^0 \to pK^{*-}J/\psi$ (not observed to date), with a consequent $K^{*-} \to K^-\pi^0$ transition. Similar partially reconstructed backgrounds have been encountered in both the R_K and $R_{K^{*0}}$ analyses. The decay $\Lambda_b^0 \to \Delta^+ KJ/\psi$, with $\Delta^+ \to p\pi^0$ is also not forbidden, but has never been observed. Potential partially reconstructed contributions coming from higher Λ^* states are suppressed: decays of excited Λ^* states in lower Λ^* states and π^0 are isospin-suppressed. At the same time, analogous radiative decays have a small probability due to large phase space available for various strong decays of excited Λ^* states. The same stands for the radiative decays of excited K^{*0} states. Therefore, the radiative contributions are considered to be negligible.

As a summary, only one mode of partially reconstructed background of a hadronic nature is expected to be dominant ($\Lambda_b^0 \to pKJ/\psi\,\pi^0$), where a missing π^0 can come either from excited kaon resonances, or from Δ baryons. Other options, including any decays with a missing photon, are suppressed.

Partially reconstructed backgrounds from the dilepton system

For the partially reconstructed backgrounds from the leptonic system, the J/ψ can originate from an excited charmonium state (e.g. χ_{c1}, χ_{c2} or $\psi(2S)$), with a consequent decay $\chi_{c1,2} \to J/\psi\,\gamma$ or $\psi(2S) \to J/\psi\,\pi^+\pi^-$, where the particles accompanying the J/ψ meson are not reconstructed. Out of other charmonium(-like) states, the so-called $X(3872)$ may

have a significant contribution via its decays to $J/\psi X$, however it would be located very low in invariant mass of pKJ/ψ (below 4900 MeV/c^2 according to the simulation), so is not expected to affect this analysis, as this low mass region is not considered in invariant mass fits.

Invariant mass distributions

Usually, the partially reconstructed decays are located in the lower mass sideband with respect to the signal peak, due to missing particle(s) in the final state. This does not create large problems for the muon modes which have a good mass resolution, though it is of a higher importance for electron modes, where the resolution is significantly worse.

In Fig. 3.9, the distribution of pKJ/ψ invariant mass is shown, with additional very tight PID requirements applied (`Proton_ProbNNp>0.7, Kaon_ProbNNk>0.6`). This ensures the absence of misidentification backgrounds, and allows to investigate a very pure pKJ/ψ final state. At the same time, a tight requirement on the output of the BDT classifier (Sec. 3.4.3) is applied to minimise the impact of the combinatorial background. On the left sideband, a bump (note the logarithmic scale) can be seen. This structure originates from a mixture of partially reconstructed and combinatorial backgrounds. Several smaller sub-contributions can be observed, they originate from the leptonic system (various excited charmonium resonances). At the same time, hadronic partially reconstructed backgrounds with lost π^0 contribute to the broad distribution seen left to the signal peak, however it is impossible to distinguish them clearly from combinatorial and other non-peaking backgrounds. The knowledge acquired from this study is later used in the fits to the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ data in a broad invariant mass range, used for systematic studies of the fit model (mentioned in Sec. 3.8.4).

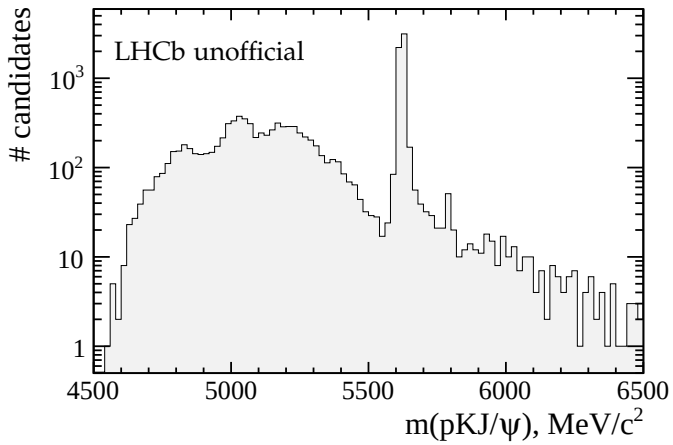


Figure 3.9: pKJ/ψ invariant mass in the wide range in the 2012 $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset, with very tight PID cuts and misID vetoes applied, shown in the logarithmic scale. The bump of partially reconstructed backgrounds is visible in the lower mass sideband (see Sec. 3.3.1.d), in particular those coming from excited charmonium resonances. The tiny peak of the $\Xi_b^0 \to pKJ/\psi$ decay is visible in the upper sideband (see Sec. 3.3.5).

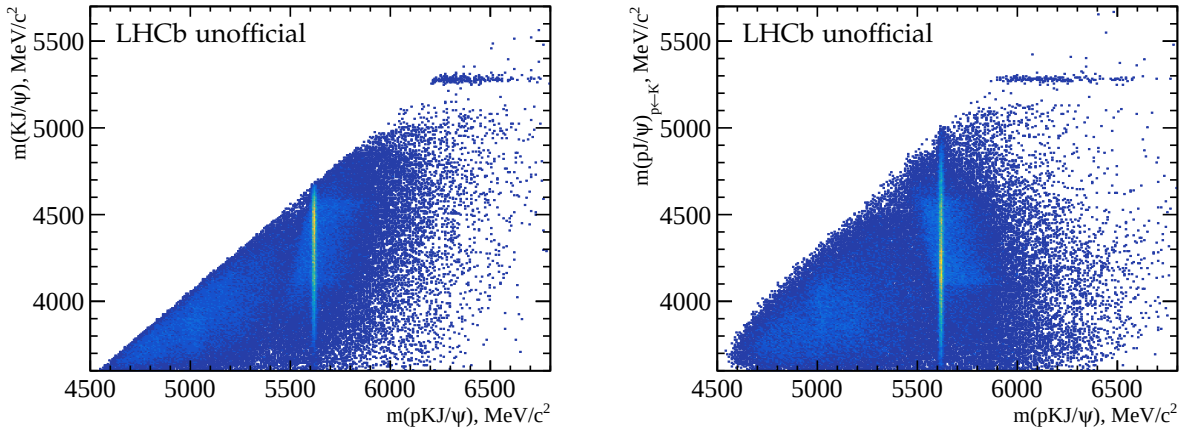
3.3.1.e Over-reconstructed backgrounds

An important type of background emerges from the combination of real $B^+ \to J/\psi K^+$ candidates with random protons, as shown in Fig. 3.10(a). Due to the large proton mass, these candidates will always have a (pKJ/ψ) invariant mass larger than 6200 MeV/c^2. This implies that they are harmless for the muon mode, while they could be a more important issue for the electron mode due to a poor mass resolution creating long tails. To ensure the fit stability in the upper sideband region, and to control properly the shape of the combinatorial background, these events are simply vetoed by applying a cut $m(Kl^+l^-) < 5200$ MeV/c^2. In addition, the events with a $p \leftrightarrow K$ swap were found among these background events (Fig. 3.10(b)), so to remove this contribution an additional cut is applied: $m(pl^+l^-)_{p\to K} < 5200$ MeV/c^2. This cut is of larger importance because such events can penetrate much lower in $m(pKJ/\psi)$.

This over-reconstructed background is visible in the J/ψ mode, however is not seen in other q^2 regions due to the lower yields (although technically they exist there as well). For consistency, the cuts are applied throughout all the decay modes.

It is important to emphasize that due to the large proton mass, both of these cuts are completely harmless to the signal modes (in both muon and electron channels), having an almost 100% efficiency, as shown in Sec. 3.6. This is an important difference between this analysis and the similar analyses with mesons such as $R_{K^{*0}}$, where the small pion mass makes these over-reconstructed backgrounds located much closer to the signal peak.

It should also be noted that no b-hadron can decay to $J/\psi p$ (there is no weakly decaying positively charged b-baryons), so analogous backgrounds of the kind "b-baryon + a random kaon" do not exist.



(a) (b)

Figure 3.10: Dominant over-reconstructed backgrounds. Vertical band is the signal in both plots; horizontal band corresponds to (a) combinations of $B^+ \to K^+ J/\psi$ with random proton, (b) combinations of $B^+ \to K^+ J/\psi$ with random proton, with proton and kaon mass hypotheses swap. See detailed description in the text.

It has been checked that the cuts suppressing over-reconstructed backgrounds do not sculpt the shape of the combinatorial background. To do so, a sample of data was selected from the negative BDT output region (Sec. 3.4.3) mostly dominated by the combinatorial background, and the effect of the cut studied. As it is shown in Fig. 3.11, only the slope of the background slightly changes in the region of upper mass sideband, but no fake peaking structures or threshold effects are created. This check has been performed in all q^2 regions of interest. It should be noted that similar checks have been performed for other selection requirements which cut directly on mass variables.

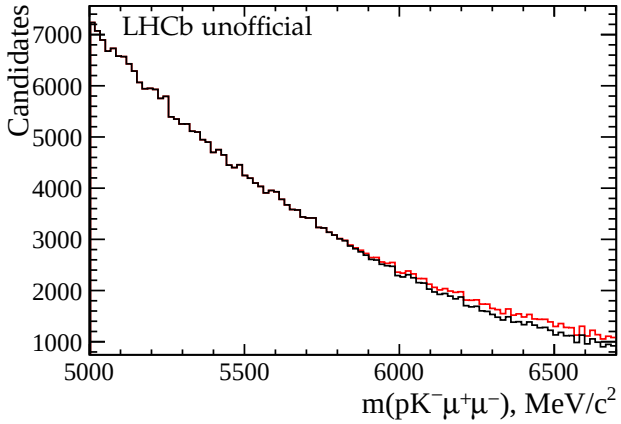


Figure 3.11: Effect of cuts suppressing over-reconstructed backgrounds, on the shape of the combinatorial background in the $0.1 < q^2 < 6$ GeV$^2/c^4$ region. Red: original distribution before applying the cut; black: distribution after the cut is applied. No potentially dangerous structures are created.

3.3.2 Background composition in the rare $\Lambda_b^0 \to pK\ell^+\ell^-$ modes

3.3.2.a Resonances in the dilepton invariant mass

The $\Lambda_b^0 \to pK\ell^+\ell^-$ decay can proceed through several resonant modes $\Lambda_b^0 \to pKX$ with $X \to l^+l^-$. In Fig. 3.3, the dimuon spectrum is shown for events selected in the window of 50 MeV/c^2 around the known Λ_b^0 mass, with the complete selection applied (Sec. 3.4). On top of the continuous component, which corresponds to the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay, two resonant components are clearly observed, they correspond to the $\Lambda_b^0 \to pKJ/\psi\,(\mu^+\mu^-)$ and $\Lambda_b^0 \to pK\psi(2S)(\mu^+\mu^-)$. The contribution of other resonances decaying to dilepton (in particular, $\rho(770)$, $\omega(782)$, $\phi(1020)$ and $\psi(3770)$) are expected to be tiny due to small branching fractions of these mesons into a dilepton final state [6].

The resonant decays $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$ are used as normalisation modes, but they can also be a source of background in the electron mode, where long bremsstrahlung tails create a leakage of candidates out of the resonant to a central-q^2 window, as it will be discussed in Sec. 3.3.4.a.

3.3.2.b Hadron misidentification backgrounds

Since the hadron system is similar between the q^2 regions, the backgrounds with hadron misidentifications affecting the data in the J/ψ window (Sec. 3.3.1.b) have similar analogs here, changing the $J/\psi \to l^+l^-$ to simply l^+l^-. Some difference in their relative contribution may occur, because of the difference in relative branching fractions, spin-dependent effects such as relative contributions of different resonances, or differences in the kinematics.

3.3.2.c Lepton misidentification backgrounds

Electrons and muon candidates can also be misidentified. As it was already discussed, in the case of resonant charmonia modes, a clean selection of J/ψ or $\psi(2S)$ peaks assumes the misidentification rate to be relatively low. However, lepton misidentifications could be an important source of the background for the rare decay modes $\Lambda_b^0 \to pK\ell^+\ell^-$.

There are two important cases of lepton misidentifications.

- The first one arises from the misidentification of a pion as a lepton. Pions are very abundant in proton-proton collisions at LHCb, and have a mass close to the muon mass. This makes pion misidentification as muon happening rather often. Less often, they may be misidentified as electrons. Having two leptons in the decay means there are two possibilities: a single or a double lepton misidentification. Such combinations can be suppressed by applying the tighter PID requirements or, as it is shown later, invariant mass cuts.

  In addition, some pions and kaons can decay in flight ($(\pi, K) \to \mu\nu_\mu$) before reaching the muon stations, and are then identified as muons. This creates a background irreducible by the PID requirements. This background is relevant mostly for the muon modes, as pion or kaon decays to electrons are helicity-suppressed.

- The second case is a swap of particle identifications of a hadron (most often a kaon, due to its mass being smaller than of a proton), and a lepton, in the resonant (thus abundant) $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$, $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ or $\Lambda_b^0 \to pK\gamma(\to e^+e^-)$ decays. The effect of these swaps can be easily seen as a corresponding peak in the distribution of the invariant mass of a hadron and a lepton, with the hadron is assumed to be identified as a lepton. These backgrounds can be suppressed by applying tighter PID requirements, or mass vetoes.

Let's consider these cases on the most prominent examples, starting with pion-lepton misidentification first.

Pion-lepton misidentification

One important decay mode of the Λ_b^0 baryon is with a Λ_c^+ and a π^- in the final state, with a consequent $\Lambda_c^+ \to pK^-\pi^+$ decay. Assuming a misidentification of both pions as leptons, one gets the same final state as the one of the signal decay $\Lambda_b^0 \to pK\ell^+\ell^-$. Given the

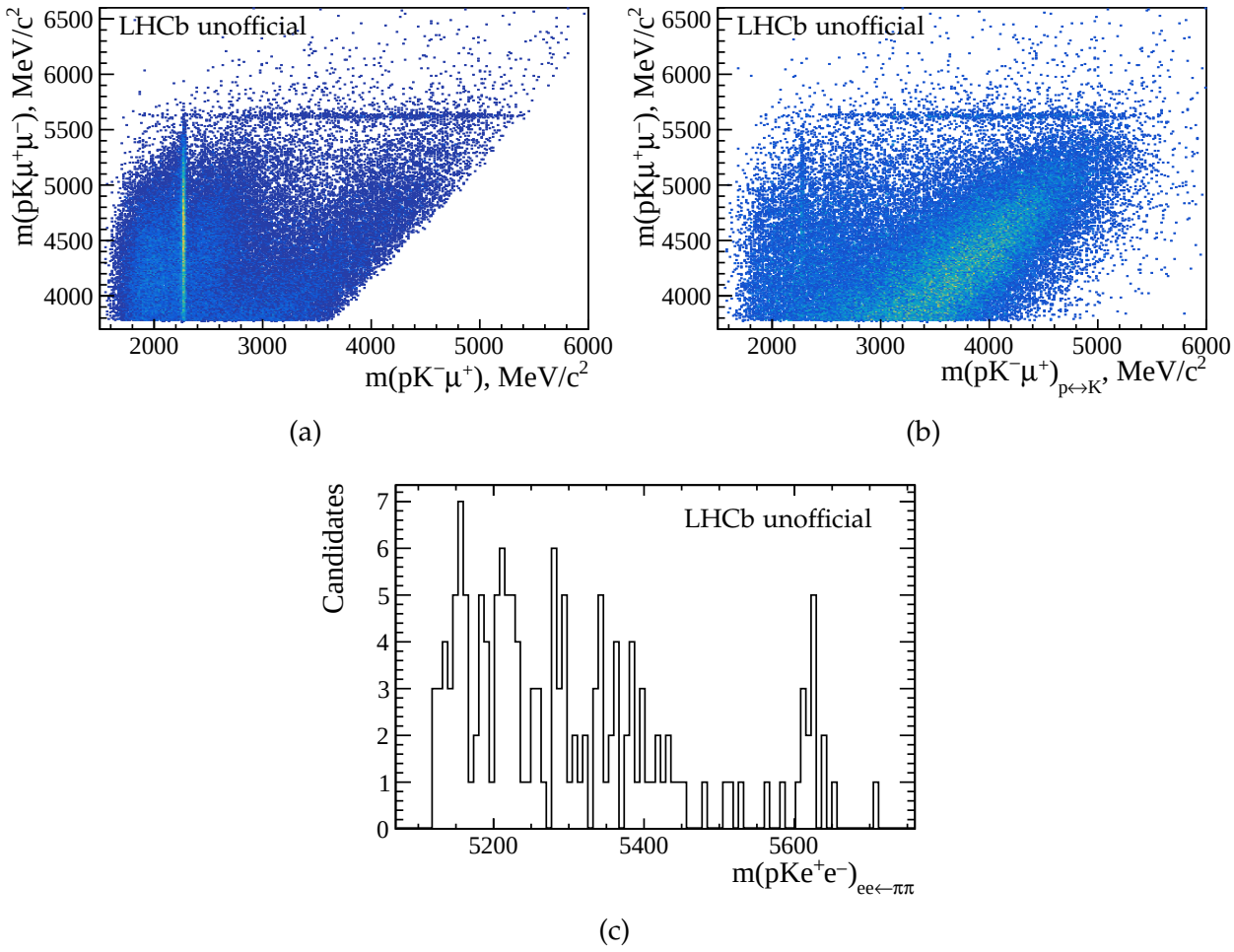


Figure 3.12: Dominant backgrounds involving Λ_c^+. In the top plots, the narrow horizontal band is a signal $\Lambda_b^0 \to pK\mu^+\mu^-$, and narrow vertical bands around 2300 MeV/c^2 correspond to (a) decay $\Lambda_b^0 \to \Lambda_c^+(\to pK\pi)\mu\nu_\mu$, with pion misidentified as a muon, and (b) same decay $\Lambda_b^0 \to \Lambda_c^+(\to pK\pi)\mu\nu_\mu$, with pion misidentified as a muon, but with proton and kaon mass hypotheses swap. Bottom (c): pKe^+e^- invariant mass, with both electons having the pion mass hypothesis, in the window of the $\Lambda_c^+ \to pK\pi$ peak. See discussion in the text.

closeness of the pion and muon masses, this background can be especially troublesome for the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, being located just underneath the signal peak.

Table 3.5 summarizes the possible background modes involving a Λ_c^+ particle, assuming the correct proton and kaon identifications. From top to bottom, these channels populate the regions located lower in the invariant mass distribution, and present less danger to the signal, with the top one being located underneath the signal peak and the bottom one creating a long tail in the lower sideband. The 4th contribution in the table is actually not a misidentification, but a double semileptonic background, discussed in the following section, which is however mentioned here for completeness.

Table 3.5: Background modes with Λ_c^+.

#	Λ_b^0 decay	Λ_c^+ decay	Comment	$m(pK^-\ell^+)$
1	$\Lambda_b^0 \to \Lambda_c^+ \pi$	$\Lambda_c^+ \to pK^-\pi^+$	Double pion misID	close to $m(\Lambda_c^+)_{\rm PDG}$
2	$\Lambda_b^0 \to \Lambda_c^+ \pi$	$\Lambda_c^+ \to pK^-\ell^+\nu_\ell$	Semileptonic with pion misID	below $m(\Lambda_c^+)_{\rm PDG}$
3	$\Lambda_b^0 \to \Lambda_c^+ \ell\nu_\ell$	$\Lambda_c^+ \to pK^-\pi^+$	Semileptonic with pion misID	close to $m(\Lambda_c^+)_{\rm PDG}$
4	$\Lambda_b^0 \to \Lambda_c^+ \ell\nu_\ell$	$\Lambda_c^+ \to pK^-\ell^+\nu_\ell$	Double semileptonic	below $m(\Lambda_c^+)_{\rm PDG}$

In Fig. 3.12(a), the prominent Λ_c^+ peak in the $m(pK^-\mu^+)$ invariant mass distribution can be seen as a narrow vertical band, in addition to the signal $\Lambda_b^0 \to pK\mu^+\mu^-$ (narrow horizontal band). The background contribution #1 from Table 3.5 is located at the intersection of Λ_c^+ and Λ_b^0 bands; contribution #3 constitutes the rest of Λ_c^+ band, while contributions #2 and #4 are located on the lower sideband with respect to the Λ_c^+ peak. In addition, Fig. 3.12(c) allows to see the contribution #1 as a mass peak consistent with the Λ_b^0 mass. In that plot, the Run I $\Lambda_b^0 \to pKe^+e^-$ data is taken, with both electrons assigned the pion mass hypothesis, and the $m(pK^-\pi^+)$ is taken within 50 MeV/c^2 of the known Λ_c^+ mass. A little but prominent peak of Λ_b^0 candidates can be seen.

In addition to this, partially reconstructed modes can be present in our dataset, such as coming from the rather abundant decay $\Lambda_c^+ \to pK^-\pi^+\pi^0$, with the π^0 not being detected (and Λ_c^+, as before, originating from Λ_b^0 decays). In terms of invariant mass, they behave similarly to the contributions #2 and #4.

From Fig. 3.12(a), it can be seen that applying the requirement $m(pK^-\ell^+) > 2320\,\text{MeV}/c^2$ is sufficient to suppress all these background modes coming from various decay chains with the Λ_c^+ involved (assuming the proper identification of both proton and kaon), preserving very high signal efficiency.

In this analysis, the cut $m(pK\ell^+) > 2320$ MeV/c^2 allows to remove the $\Lambda_c^+ \to pK\pi$ as well as $\Lambda_c^+ \to pK\ell\nu_\ell$ candidates from the decays $\Lambda_b^0 \to \Lambda_c^+\pi$ and $\Lambda_b^0 \to \Lambda_c^+\ell\nu_\ell$. This veto is applied to all channels. It has a signal efficiency about 99% on the rare $\Lambda_b^0 \to pK\ell^+\ell^-$ modes, and about 96% on the $\Lambda_b^0 \to pKJ/\psi$ modes which have slightly different kinematics. For the misidentified $\Lambda_c^+ \to pK\pi$ mode, it removes about 100% of this background. For the semileptonic $\Lambda_c^+ \to pK\mu\nu_\mu$ mode, it is also about 100% efficient, however, is a bit less efficient (98%) for the $\Lambda_c^+ \to pKe\nu_e$ channel due to the bremsstrahlung tail extending over the cut value.

In addition, the case of proton-kaon ID hypotheses swap should be considered: in Fig3.12(b), the tiny Λ_c^+ contribution can be also seen in $m(pK^-\mu^-)$ (note the sign of the muon charge) invariant mass distribution with proton-kaon ID hypothesis swap. By analogy, four cases can be considered, and a cut $m(pK^-\ell^-)_{p\leftrightarrow K} > 2320\,\text{MeV}/c^2$ can be applied in order to get rid of them.

For the electron mode, it is important that the misidentification has a much higher chance to happen if the electron has no bremsstrahlung photon attached. As it is shown in Fig, 3.13, the misidentification Λ_c^+ peak is visible only for such electrons, while there is no sign of it for the electrons with one or more bremsstrahlung photons attached.

Cases when the proton or the kaon alone has a wrong identification, are also possible,

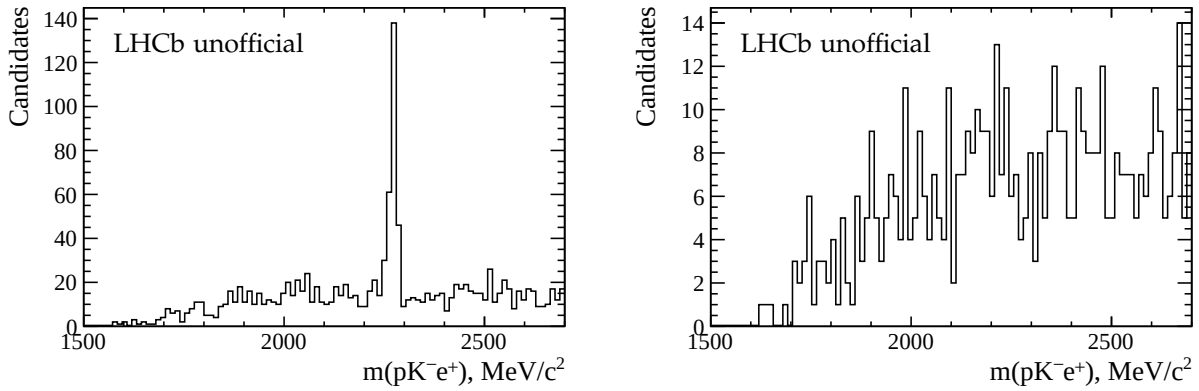


Figure 3.13: Invariant mass of pKe system. Left: electron has no brem photon attached; right: electron has brem photons attached. The prominent $\Lambda_c^+ \to pK\pi$ misID peak is seen only in the 0brem plot.

but represent a second-order effect, and the corresponding contributions in data are much smaller than those discussed above.

Given the relative closeness of the mass of a pion and a muon, or a pion and an electron, it is very important to make sure that the decay $\Lambda_b^0 \to pK^-\pi^+\pi^-$ does not create a peaking background under the signal of $\Lambda_b^0 \to pK\ell^+\ell^-$ via a double misidentification of both pions as a leptons. This can be achieved by applying the particle identification requirements on the leptons, however, one has to make sure they are sufficient to suppress this background.

This decay has the same final state as the above-discussed contribution #1 of the backgrounds originating from the Λ_c^+, *i.e.* $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)\pi^+$. To distinguish between the two, we will call the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decay as "charmless" in the following discussion.

The decay $\Lambda_b^0 \to pK\pi^+\pi^-$ is studied in detail in Ref. [47], and the decay $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)\pi^-$ is used as a control channel. Thanks to that, one can estimate the expected $\Lambda_b^0 \to pK^-\pi^+\pi^-$ charmless yield in our dataset. They report $\frac{N(non-res)}{N(\Lambda_c^+)} = \frac{19877}{113612} \approx 0.17$.

In Fig. 3.12(c) representing the full Run I $\Lambda_b^0 \to pKe^+e^-$ dataset with the full selection (Table 3.6) applied (except for the Λ_c^+ mass veto), there are 18 $\Lambda_b^0 \to \Lambda_c^+(pK^-\pi^+)\pi^-$ candidates falling in the 100 MeV/c^2 window around the known Λ_b^0 mass. These events are then suppressed by the Λ_c^+ mass veto. Assuming that the ratio of charmless channel over the Λ_c^+ channel is the same as in Ref. [47], one should expect about 3 $\Lambda_b^0 \to pK^-\pi^+\pi^-$ events in our $\Lambda_b^0 \to pKe^+e^-$ Run I dataset. However, this estimate does not take into account that a) in the current analysis, the cut $m(pK) < 2600$ MeV/c^2 is applied, and the Dalitz projections shown in the Ref. [47], suggest that this cut would suppress more than a half of the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ events; b) the cut $m(pK^-\ell^+) > 2320$ MeV/c^2 used in this analysis to suppress various backgrounds involving Λ_c^+ is also removing a significant fraction of the phase space of the charmless background mode. An additional check is performed based on the fact that the Dalitz plot of the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decay at low

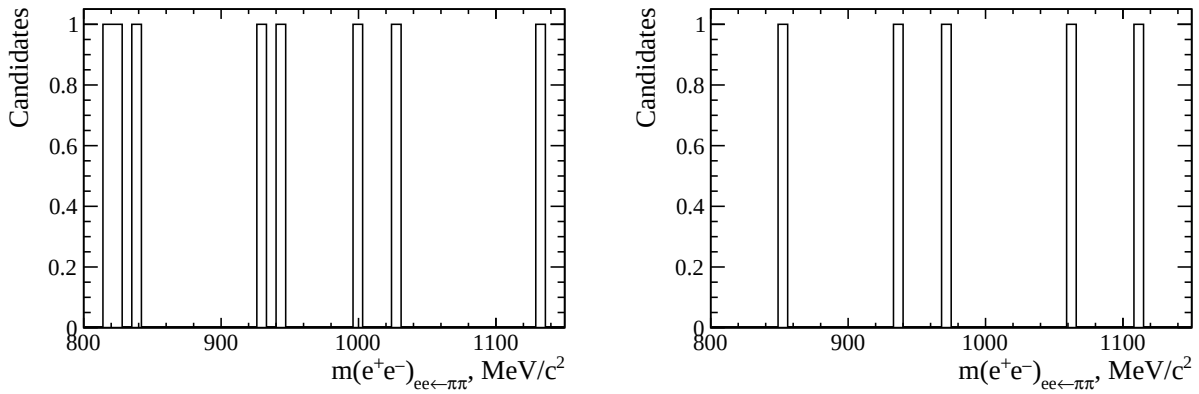


Figure 3.14: $m(e^+e^-)_{ee\leftarrow\pi\pi}$ in the $f_0(980)$ region. Left plot is for Run I, and right plot is for Run II.

$m(pK)$ is dominated by the $f_0(980) \to \pi^+\pi^-$ decay, while no sign of such contribution is seen in our data. In Fig. 3.14, we consider only the category where both electrons have no bremsstrahlung photons attached (as it was shown above that misidentified pions do not tend to pick bremsstrahlung photons), after the full selection applied. No accumulation of events is seen in the proximity of $f_0(980)$ mass. As a conclusion, it can be concluded that this double misidentification mode should not present any significant problem for the current analysis, when the default particle identification requirements are applied.

Another, less important case is the one of the $\Lambda_b^0 \to pD^0\pi^-$ decay, with a subsequent $D^0 \to K^-\pi^+$ transition. It can be seen as a prominent vertical band in Fig. 3.15(a), where pion is misidentified as a lepton. This contribution is vetoed by applying a narrow 20 MeV/c^2 mass veto around the known D^0 mass, which removes about 93% of the $D^0 \to K\pi$ candidates preserving about 98% signal efficiency.

The case of a single lepton misidentification assumes the decay contains one true lepton and one pion. This can happen if the above-mentioned Λ_c^+ or D^0 decay semileptonically, to $pK^-l^+\nu_l$ or $K^-l^+\nu_l$ respectively. The essential difference between the single and the double lepton misidentifications is that in the first case one neutrino is missing, so the invariant mass of this background is located in the lower sideband of the signal peak.

Hadron-lepton swaps

For what concerns the lepton-hadron ID swap, as it was already discussed, the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ decays are several orders of magnitude more abundant than the rare $\Lambda_b^0 \to pK\ell^+\ell^-$ decays. In this context, even if the probability of an ID swap between a hadron and a lepton is rather small, such swaps in the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ become an important source of background for the rare modes. Similar considerations apply for $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ and $\Lambda_b^0 \to pK\gamma(\to e^+e^-)$, but to a lesser extent.

For example, a contribution from such a misidentification can be seen in Fig. 3.15(b) as a narrow vertical band in the right part of the plot. It can be better visible on the one-dimensional projection shown in Fig. 3.16(left). However, the amount of data is not

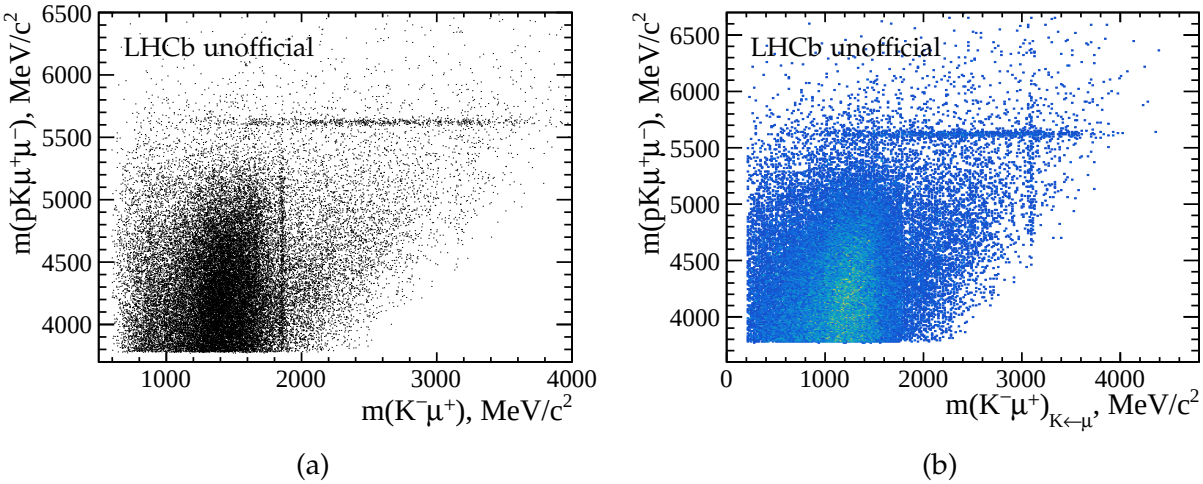


Figure 3.15: Backgrounds having signatures in $K^-\mu^+$ invariant mass. Narrow horizontal band is a signal $\Lambda_b^0 \to pK\mu^+\mu^-$, and narrow vertical bands (a) around 1900 MeV/c^2 represent D^0 from the decay $\Lambda_b^0 \to pD^0(\to K\pi)\mu\nu_\mu$, with pion misidentified as a muon, and (b) around 3100 MeV/c^2 represent J/ψ from the decay $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ with the ID swap of kaon and negative muon. In these plots, PID requirements are applied for all final state particles.

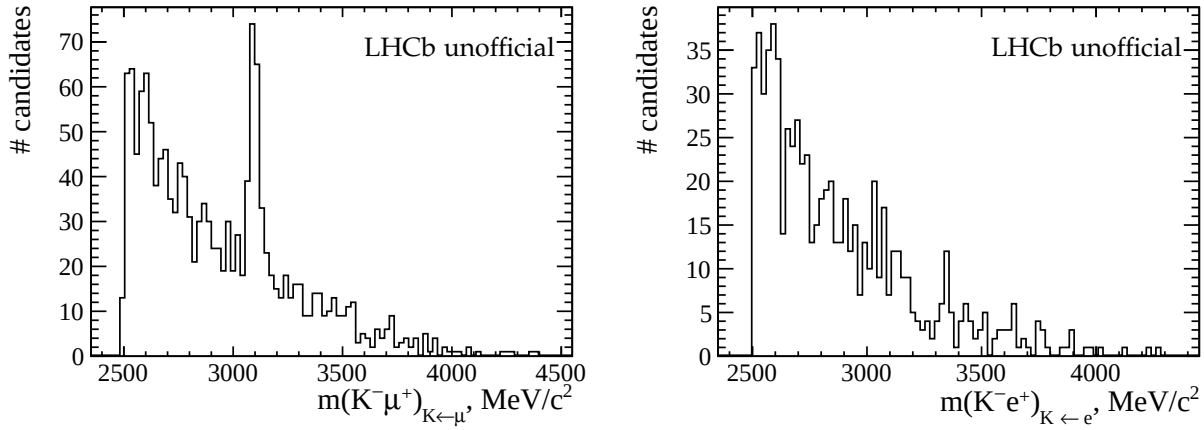


Figure 3.16: Kaon-lepton invariant mass, with the kaon having the lepton mass hypothesis. Left – $\Lambda_b^0 \to pK\mu^+\mu^-$, right – $\Lambda_b^0 \to pKe^+e^-$ data (Run I + Run II datasets merged). The peak of J/ψ can be seen in the left plot. See discussion in the text.

sufficient to see clearly the contribution from the $\psi(2S)$. In Fig. 3.16(right), the same distribution is shown for the $\Lambda_b^0 \to pKe^+e^-$ dataset, and no obvious peaking structure is seen. This is because the expected resolution of the J/ψ is larger, with the yields being smaller (in addition to the lower electron efficiencies in general, the probability for kaon to be misidentified as an electron is lower than as muon, due to the larger difference in mass.) Fig. 3.17 uses the complete Run I+II $\Lambda_b^0 \to pKe^+e^-$ dataset with a full selection applied, except for the kaon-electron misidentification veto. No visible J/ψ peak is seen

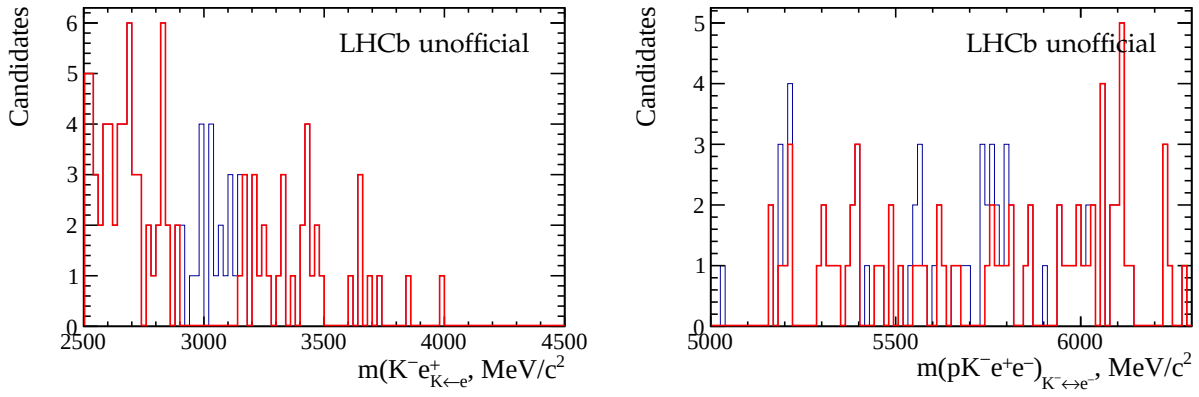


Figure 3.17: Left: $m(K^-e^+)$ when kaon has an electron mass hypothesis. Right: $m(pKee)$ with the same-charge kaon-electron mass hypotheses swapped, and the J/ψ mass constraint applied on the $m(K^-e^+)_{K\leftarrow e}$. Blue points – without the J/ψ veto, red points – after the veto.

in the left plot, and no visible Λ_b^0 peak is seen in the right plot. All these checks allow to conclude that the pollution from the kaon-electron swaps is negligible in the current analysis.

It is important to suppress such backgrounds arising from swaps of leptons and hadrons, because they end up in the region of invariant mass located underneath the signal peak.

To remove the $J/\psi \to \mu^+\mu^-$ candidates with identity swap between kaon and one of the muons, the veto $|m(K^-\mu^+)_{K\leftarrow\mu} - 3097| > 35$ MeV/c^2 is applied only to the rare mode window, which is about 98% efficient for the signal and removes about 92% of background events.

The veto $!(2900 < |m(K^-e^+)_{K\leftarrow e}| < 3150)$ MeV/c^2 is applied to the electron mode, which has about 93% signal efficiency and removes about 60% of the background events. Indeed, given the larger resolution in the electron mode, a broader cut is needed. However, no sign of any J/ψ peak was seen in data (Fig. 3.16(b) and Fig. 3.17), so this background is believed to be very small. This is because of the rather tight PID requirement on the kaon. According to the estimations from the simulation, the total yield of the misidentification backgrounds in the $\Lambda_b^0 \to pKe^+e^-$ data, due to the kaon-electron swap, will be at the level of about 0.5 events per trigger category, after this cut is applied.

3.3.2.d Doubly semileptonic backgrounds

The doubly semileptonic background modes were already mentioned in the previous section, here they are discussed in more details.

A specific case is of semileptonic backgrounds having the same initial and final state, but missing neutrinos. For example, the following decays can be observed: $\Lambda_b^0 \to \Lambda_c^+(\to pK\ell\nu_\ell)\ell\nu_\ell$ and $\Lambda_b^0 \to pD^0(\to K\ell\nu_\ell)\ell\nu_\ell$. Both of these decays have a proton, kaon, dilepton pair in the final state, together with neutrino and antineutrino which are lost. The distinctive feature of the semileptonic backgrounds consists in the absence of associated

mass peaks: their invariant mass distributions are very broad.

However, the $\Lambda_b^0 \to \Lambda_c^+(\to pK\ell\nu_\ell)\ell\nu_\ell$ decay always has a $(pK^+\ell^-)$ invariant mass lower than the known mass of the Λ_c^+ baryon, due to the missing neutrino. So, the same cut $m(pK^+\ell^-) > 2320$ MeV/c^2 is applied to fully remove this component together with the $\Lambda_b^0 \to \Lambda_c^+(\to pK\pi)\ell\nu_\ell$ component discussed in the previous section. This cut is about 98% efficient on the signal. The same idea is applied for the case of $p \leftrightarrow K$ swap in the $\Lambda_b^0 \to \Lambda_c^+(\to pK\ell\nu_\ell)\ell\nu_\ell$ decay. In this case, the cut applied is $m(pK^+\ell^+)_{p\leftrightarrow K} > 2320$ MeV/c^2.

For what concerns the $\Lambda_b^0 \to pD^0(\to K\ell\nu_\ell)\ell\nu_\ell$ background, vetoing it by applying a cut $m(K^+\ell^-) > m_{PDG}(D^0)$ would result in a significant loss of signal efficiency (about 50%), therefore this requirement is not applied. This background is not peaking in the $pK\ell^+\ell^-$ invariant mass, and is located at low invariant mass sideband with respect to the signal, therefore is not dangerous. It is significantly reduced by the momentum balance requirement discussed in Sec. 3.4.6.

3.3.3 Background composition in $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ modes

Everything discussed previously regarding the background composition in the J/ψ window is in general applicable here for both the muon and the electron modes. As it has been stated earlier, some differences may occur, for example, in the relative fractions of the misidentification backgrounds between the J/ψ and $\psi(2S)$ cases, since the relative branching fractions or kinematics are not expected to be identical. In particular, as it has been discussed in Sec. 1.4.2, the relative branching fractions of Λ_b^0 and B_s^0 decay modes are very different between J/ψ and $\psi(2S)$ cases, leading to a much larger relative yields of misidentification backgrounds (as compared to the J/ψ region), comparable or even higher than the signal yields. In addition, backgrounds coming from excited charmonium states are negligible in this case, because most of the states above $\psi(2S)$ have small branching fractions of decays into $\psi(2S)X$. The small phase space of the $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ decay leaves rather limited possibilities for the partially reconstructed decays of hadronic nature. On the other hand, given that the size of the $\psi(2S)$ data sample is much smaller with respect to the J/ψ mode, a reasonably simple description of background model can be adopted.

3.3.4 Backgrounds specific to electron modes only

There are some important aspects which are notably different between the electron and muon modes, and may lead do additional background contaminations:

- The resolution is worse on the electron modes, leading to "leakages" of candidates from one q^2 region to another;
- The ability of photons to convert into a dielectron pair when interacting with material.

For what concerns the conversions of photons coming from the $\Lambda_b^0 \to pK\gamma(\to e^+e^-)$ mode, it contaminates the signal $\Lambda_b^0 \to pKe^+e^-$ only at very low q^2 region ($m(\ell^+\ell^-) < 50\,\mathrm{MeV}/c^2$) which is excluded from the current analysis. The similar reasoning can be used for the decays of the kind $\Lambda_b^0 \to pK\pi^0$, with $\pi^0 \to \gamma\gamma$, where a) one of the photons undergoes a conversion to a dielectron pair, or b) the π^0 decays via a Dalitz decay. In this case, the invariant mass of the dielectron pair cannot exceed the mass of π^0, which is far below the lower boundary of the q^2 region considered in the current analysis. Some other possibilities are described below.

3.3.4.a Leakages from one q^2 region to another

Due to the poor resolution in the electron modes, tails from the invariant mass distributions of various modes can "leak" from one q^2 region to another. Such an effect can be seen in Fig. 3.18: the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ events are present also within the q^2 region used for the R_{pK}^{-1} measurement. However, given the nature of this effect which is intrinsically related to the dielectron system itself, a strong correlation is observed between the dielectron and Λ_b^0 invariant masses. This implies that this leakage to the q^2 region below J/ψ is located in the lower m(pKe^+e^-) sideband with respect to the signal, – effectively, below 5000 MeV/c^2 in the region used for the R_{pK}^{-1} measurement. This is taken into account in invariant mass fits (Sec. 3.7.4).

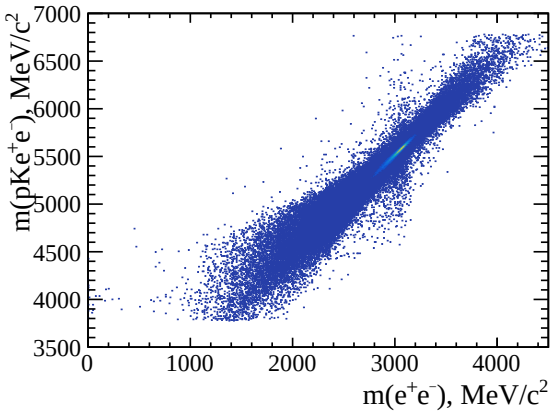


Figure 3.18: Invariant mass distribution of the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ events in the simulation.

3.3.4.b Backgrounds due to random photon conversions

It was observed in data, that when computing the invariant mass of (hadron+electron) of opposite charges, and assuming that hadron is actually a misidentified electron, a sharp peak around zero can be seen in any q^2 region (Fig. 3.19(a)). Properties of this peak are compatible with it originating from the photon conversions, as checked using the $\Lambda_b^0 \to pK\gamma(\to e^+e^-)$ simulation (*e.g.* pseudorapidities of two conversion tracks are very close), with one of conversion electrons later being misidentified as a hadron. A majority

of these events are also compatible with being clone tracks of opposite charges (analogous to the same-sign clones described in Sec. 3.3.6).

It was checked that the events under such a peak do not create any peaking structure in the four-body pKe^+e^- invariant mass (Fig. 3.19(b)), but rather behave like a combinatorial background (both in terms of the $m(pKe^+e^-)$ invariant mass shape, and in terms of being suppressed by tight cuts on the multivariate classifier trained to suppress combinatorial background, and described in Sec. 3.4.3.). In addition, these events can be easily suppressed applying a cut on the two-body invariant mass to be just above zero (10 MeV/c^2), and such a cut is harmless to the signal, preserving about 100% signal events.

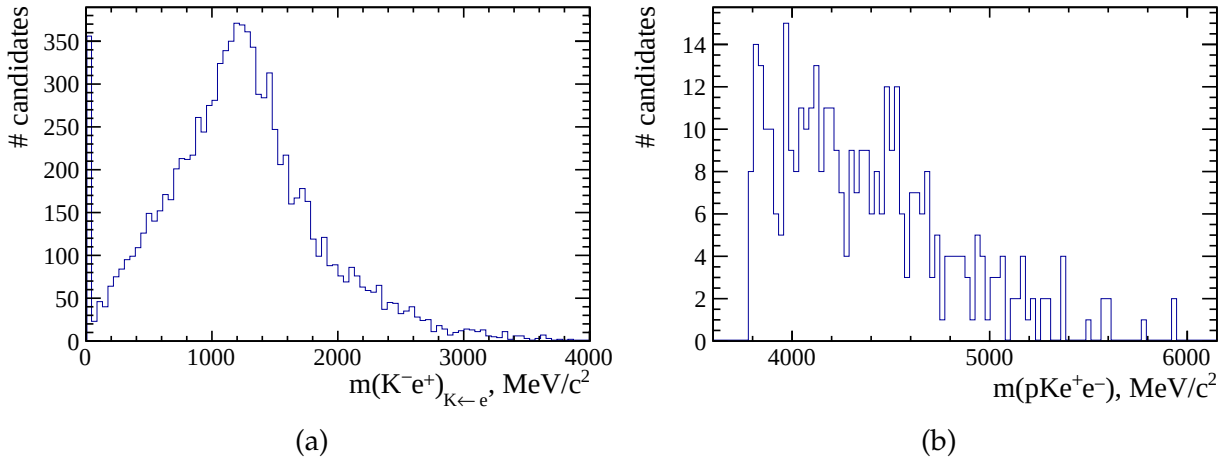


Figure 3.19: a) Invariant mass of K^-e^+, where K has an electron mass hypothesis, in $\Lambda_b^0 \to pKe^+e^-$ data, in the $0.1 < q^2 < 6$ GeV$^2/c^4$ window. The peak at zero can be seen, presumably coming from the photon conversions. b) Invariant mass $m(pKe^+e^-)$ shape for the candidates having $m(K^-e^+)_{K\leftarrow e} < 10$ MeV/c^2.

3.3.5 Other possible specific backgrounds

Potentially, many other decay modes could contribute to the background composition. Contributions from other b-baryons are possible and discussed below. The most important contribution arises from a Cabibbo-suppressed decay $\Xi_b^0 \to pKJ/\psi$, which can be observed in our data when applying very tight PID requirements, as seen in Fig. 3.9. This decay mode is observed for the first time in the current analysis, and it is studied in more detail in the Chapter 4. The yield of the $\Xi_b^0 \to pKJ/\psi$ decay mode is of sub-percent order with respect to that of the $\Lambda_b^0 \to pKJ/\psi$ mode, thus it is considered as negligible for the purpose of this measurement. Same stands for the not-yet-observed $\Xi_b^0 \to pK\ell^+\ell^-$ rare mode. Another neutral b-baryon, the Σ_b^0 (not yet observed), is expected to decay strongly to $\Lambda_b^0\pi^0$ in about 100% of cases, assuming its mass is close to the masses of observed isospin partners $\Sigma_b^\pm$. Therefore, it does not decay to the $pK\ell^+\ell^-$ final state and cannot create a background for this analysis. Charged b-baryons cannot decay to the pKl^+l^- final state,

however, can be a source of partially reconstructed backgrounds, when decaying to a more complicated final state. No such decay has been observed to date. However, the small production of other b-baryon states makes these possible contributions sufficiently small to be ignored in the current analysis.

Several excited states are known for the Λ_b^0 baryon: the $\Lambda_b^0(5912)$ and $\Lambda_b^0(5920)$, which both can decay to $\Lambda_b^0\pi^+\pi^-$ [150] and $\Lambda_b^0\gamma$ final states. The latter decay mode assumes a possibility of Dalitz decays (or photon conversions) to $\Lambda_b^0\ell^+\ell^-$, creating a potential possibility of a swap between a bachelor lepton and a lepton from the Λ_b^0 decay. However, given that these decay modes are expected to have tiny branching fractions, and are unlikely to create peaking structures mimicking the signal peak, this family of backgrounds is not considered in this analysis as well.

3.3.6 Backgrounds due to the clone tracks

As it has been discussed in Sec. 2.4, a dedicated `"Clone Killer"` algorithm is run as a part of the standard tracking procedure. However, in certain cases a non-negligible amount of clone tracks survive. This may be a more important effect for electrons: due to the bremsstrahlung radiation changing the direction of the track, it is sometimes possible to associate one set of VELO hits to multiple sets of hits in the tracking stations. In most of the cases, the clone tracks are of the same charge; however it is also possible to have clones of opposite charges if only the VELO hits are shared, and completely different sets of hits in the tracking stations are considered.

The common test of presence of clone tracks in the data sample, broadly exploited in various spectroscopy analyses in LHCb, is based on calculating the angle between the two tracks. For the clone tracks, this angle is usually very close to zero (< 0.5 mrad). The momenta and PID hypotheses can be different between the two clone tracks, and the clones are potentially a dangerous source of background.

The presence of clone tracks in our datasets is tested separately in q^2 regions.

Figures 3.20 and 3.21 present the angle between the hadron and lepton tracks of the same charge, in the $\Lambda_b^0 \to pKe^+e^-$ and $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ datasets (2012 data, full selection applied except for the clone veto). A clear peak at zero can be seen in the non-resonant dataset, and a much less prominent peak in the J/ψ mode. It can be seen that as much as 20% of the $\Lambda_b^0 \to pKe^+e^-$ rare mode dataset (which has passed the entire selection but the anti-clone cut) is composed of these clone tracks. For the J/ψ mode, this percentage is only at 1% level, as this window is dominated by real J/ψ meson candidates: the absolute number of clone tracks is of about the same order, but the total amount of data is much larger here. The amount of clone tracks is also very small in the muon mode. For example, the distribution of the angles between the muon and hadron tracks of the same charge, are presented in Fig. 3.22. As it can be seen from Fig. 3.20 (bottom), the clone candidates populate the region of angles very close to zero, while candidates of signal and other backgrounds have significantly larger angles. Therefore, the angle>0.5 mrad requirement is applied in this analysis in order to get rid of the clone track background, to both muon and electron datasets in all q^2 regions.

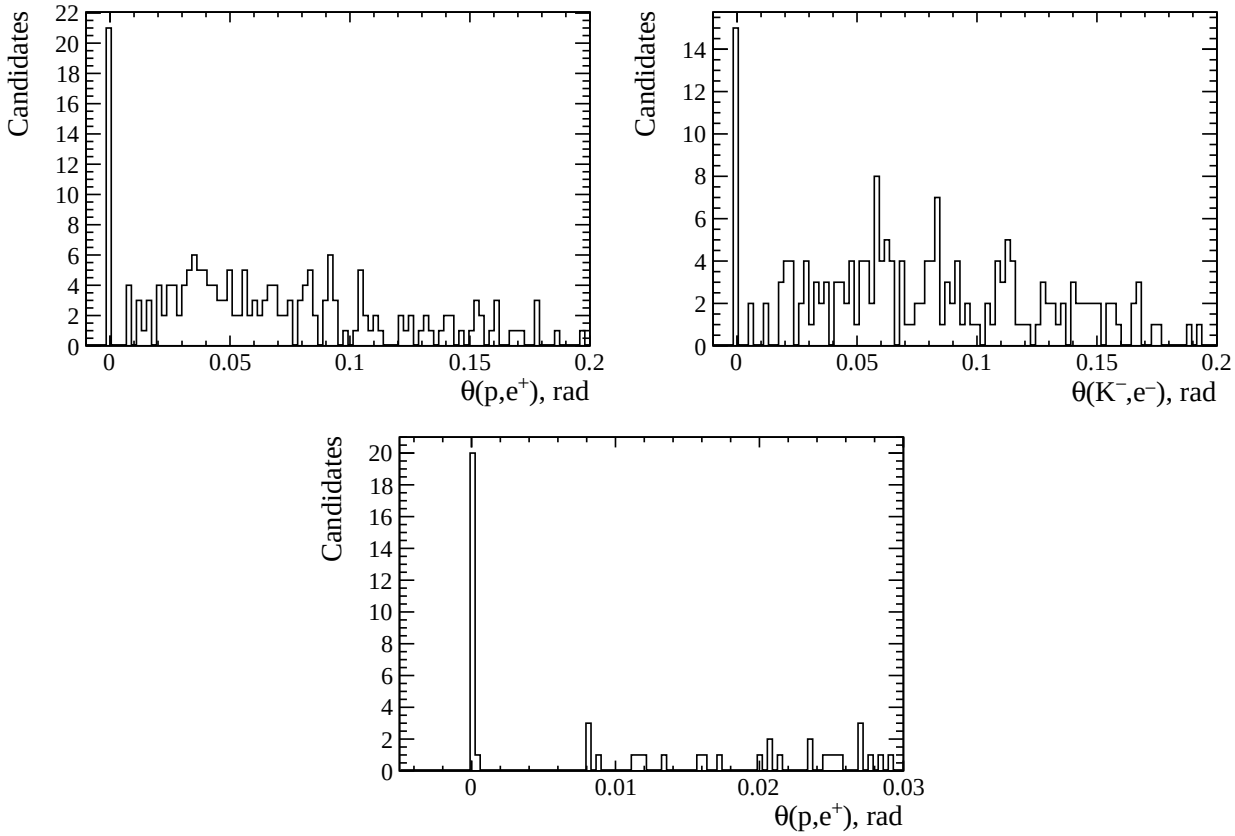


Figure 3.20: Angle between the hadron-lepton tracks of the same charge (top left: proton-electron, top right: kaon-electron), in the rare $\Lambda_b^0 \to pKe^+e^-$ dataset, after the complete selection applied except for the clone track veto. Bottom plot is the zoom of the top left plot to the region close to zero, justifying the cut value.

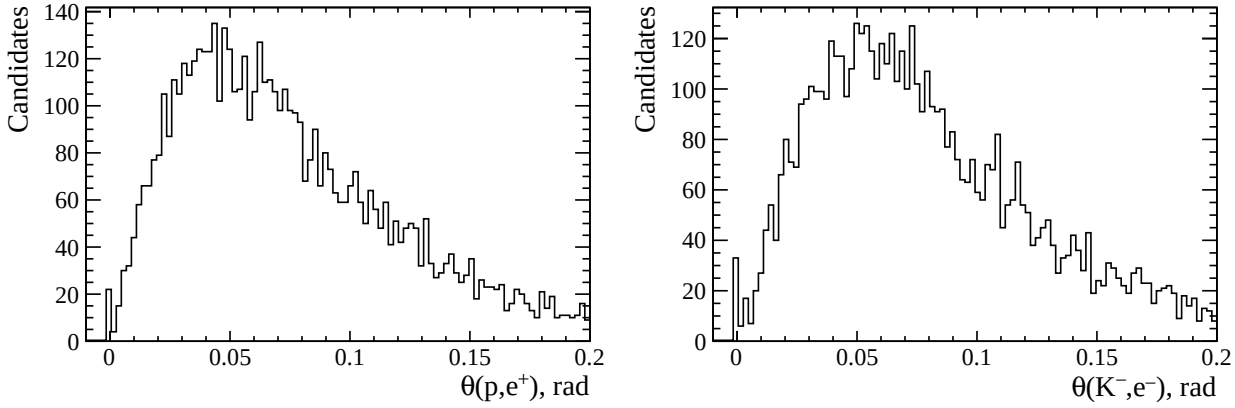


Figure 3.21: Angle between the hadron-lepton tracks of the same charge (left: proton-electron, right: kaon-electron), in the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ dataset, after the complete selection applied except for the clone track veto.

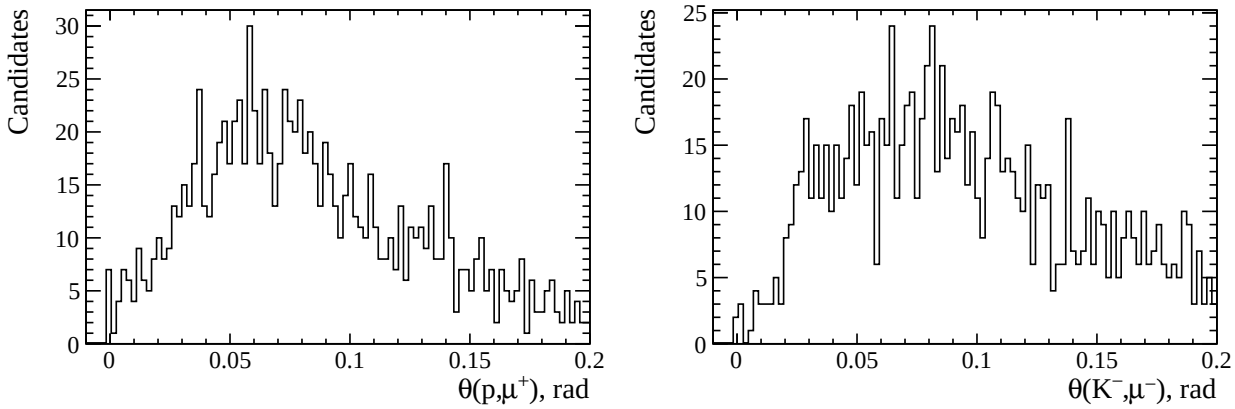


Figure 3.22: Angle between the hadron-lepton tracks of the same charge (left: proton-muon, right: kaon-muon), in the rare $\Lambda_b^0 \to pK\mu^+\mu^-$ dataset, after the complete selection applied except for the clone track veto.

3.4 Offline selection requirements

As a summary of the background studies performed above, the offline selection applied in order to minimise the impact of these backgrounds, is presented here. It is applied on top of the stripping requirements defined earlier. The acquired knowledge of the different background sources is used to develop the selection steps in order to reach the best sensitivity to the rare modes $\Lambda_b^0 \to pK\ell^+\ell^-$.

The summary of the offline selection requirements applied throughout this analysis, is presented in the Table 3.6 for Run I, and Table 3.7 for Run II. Further details and explanations of LHCb-specific terms are provided below.

3.4.1 Subdetector acceptance

For a reliable estimation of the detector performance, one has to make sure that the particles of interest fall into the acceptance of every relevant subdetector. This section describes all the subdetector acceptances requirements applied in this analysis.

In order to ensure that the corrections applied to the simulation properly take into account the differences of the response of the Ł0 triggers between data and simulation, electron tracks from the various simulated samples are required to be within the acceptance of the ECAL.

Additionally, some of the innermost ECAL cells are not read out during data taking, as shown in Fig. 3.23. In order to minimise the data-simulation disagreement, a veto of the innermost region of ECAL: $!(|x_{\mathrm{ECAL}}| < 363.6\mathrm{mm}\&|y_{\mathrm{ECAL}}| < 282.6)\mathrm{mm}$ is applied both in the simulation and data.

All final state tracks of the decays of interest, proton, kaon, muon, electron are required to be in the acceptance of RICH system. This is done in order to ensure the PID response for all tracks is valid.

Table 3.6: Offline selection requirements for Run I.

Type	Particle(s)	Requirement	Sample
Quality and acceptance	μ	`InAccMuon`	all
	e	`hasCalo`	all
	p, K, e, μ	`hasRich`	all
	p, K, e, μ	`GhostProb<0.3`	all
	e	$region_{ECAL} \geq 0$	all
	e	$!(\lvert x_{ECAL}\rvert < 363.6$ mm&&$\lvert y_{ECAL}\rvert < 282.6$ mm$)$	all
PID	p	$p > 10\,\text{GeV}/c$, $p_T > 400$ MeV/c	all
	K	$p > 2\,\text{GeV}/c$, $p_T > 250$ MeV/c	all
	e	$p > 3\,\text{GeV}/c$, $p_T > 500$ MeV/c	all
	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV/c	all
	p, K, e, μ	$p < 150$ GeV/c	all
	p	`MC12TuneV2_ProbNNp>0.2`, `MC12TuneV3_ProbNNk<0.8`, `MC12TuneV3_ProbNNpi<0.7`	all
	K	`MC12TuneV3_ProbNNk>0.2`, `MC12TuneV3_ProbNNp<0.8`	all
	μ	`MC12TuneV3_ProbNNmu>0.1`	all
	e	`MC12TuneV3_ProbNNe>0.1`	all
Mass windows	dilepton	$2900 < \text{m}(\mu^+\mu^-) < 3200$ MeV/c^2	J/ψ $(\mu^+\mu^-)$
		$6 < q^2 < 11$ GeV$^2/c^4$	J/ψ (e^+e^-)
		$3586 < \text{m}(\mu^+\mu^-) < 3786$ MeV/c^2	$\psi(\mu^+\mu^-)$
		$11 < q^2 < 15$ GeV$^2/c^4$	$\psi(e^+e^-)$
		$0.1 < q^2 < 6$ GeV$^2/c^4$	all rare
		$\text{m}(e^+e^-) < 20$ MeV/c^2	$\gamma(e^+e^-)$
Background vetoes	B^+	$m(K\ell^+\ell^-) < 5200$ MeV/c^2, $m(p\ell^+\ell^-)_{p\leftarrow K} < 5200$ MeV/c^2	all
	ϕ	$\lvert m(pK)_{p\leftarrow K} - 1020\rvert > 12$ MeV/c^2	all
	Λ_c^+	$m(pK\ell^+) > 2320$ MeV/c^2, $m(pK\ell^-)_{p\leftrightarrow K} > 2320$ MeV/c^2	all
	D^0	$\lvert m(K^-\ell^+)_{\ell\leftarrow\pi} - 1865\rvert > 20$ MeV/c^2	all rare
	swaps	$\lvert m(K^-\mu^+)_{K\leftarrow\mu} - 3097\rvert > 35$ MeV/c^2	rare $\mu\mu$
		$m(K^-e^+)_{K\leftarrow e} < 2900$ or > 3150 MeV/c^2	rare ee
	conversions	$m(K^-e^+)_{K\leftarrow e} > 10$ MeV/c^2, $m(pe^-)_{p\leftarrow e} > 10$ MeV/c^2	all ee
	clones	$\theta(K,\ell) > 0.5$ mrad, $\theta(p,\ell) > 0.5$ mrad	all
Combinatorial background	–	`BDT>0.8`	all $\mu\mu$
	–	`BDT>0.8`	all ee
HOP	–	$m_{HOP}(\Lambda_b^0) > (5265\,\text{MeV}/c^2 + 10 \times log(\chi^2_{FD}(\Lambda_b^0)))$	rare ee

Table 3.7: Offline selection requirements for Run II.

Type	Particle(s)	Requirement	Sample
Quality and acceptance	μ	`InAccMuon`	all
	e	`hasCalo`	all
	p, K, e, μ	`hasRich`	all
	p, K, e, μ	`GhostProb<0.3`	all
	e	$region_{ECAL} \geq 0$	all
	e	$!(\lvert x_{ECAL}\rvert < 363.6$ mm&&$\lvert y_{ECAL}\rvert < 282.6$ mm$)$	all
PID	p	$p > 10\,\text{GeV}/c$, $p_T > 1000$ MeV$/c$	all
	K	$p > 2\,\text{GeV}/c$, $p_T > 250$ MeV$/c$	all
	e	$p > 3\,\text{GeV}/c$, $p_T > 500$ MeV$/c$	all
	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV$/c$	all
	p, K, e, μ	$p < 150$ GeV$/c$	all
	p	`MC15TuneV1_ProbNNp>0.3`, `MC15TuneV1_ProbNNk<0.8`, `MC15TuneV1_ProbNNpi<0.7`	all
	K	`MC15TuneV1_ProbNNk>0.2`, `MC15TuneV1_ProbNNp<0.8`	all
	μ	`MC15TuneV1_ProbNNmu>0.1`	all
	e	`MC15TuneV1_ProbNNe>0.1`	all
Mass windows	dilepton	$2900 < \text{m}(\mu^+\mu^-) < 3200$ MeV$/c^2$	J/ψ $(\mu^+\mu^-)$
		$6 < q^2 < 11$ GeV$^2/c^4$	J/ψ (e^+e^-)
		$3586 < \text{m}(\mu^+\mu^-) < 3786$ MeV$/c^2$	$\psi(\mu^+\mu^-)$
		$11 < q^2 < 15$ GeV$^2/c^4$	$\psi(e^+e^-)$
		$0.1 < q^2 < 6$ GeV$^2/c^4$	all rare
		$\text{m}(e^+e^-) < 20$ MeV$/c^2$	$\gamma(e^+e^-)$
Background vetoes	B^+	$m(K\ell^+\ell^-) < 5200$ MeV$/c^2$, $m(p\ell^+\ell^-)_{p\leftarrow K} < 5200$ MeV$/c^2$	all
	ϕ	$\lvert m(pK)_{p\leftarrow K} - 1020\rvert > 12$ MeV$/c^2$	all
	Λ_c^+	$m(pK\ell^+) > 2320$ MeV$/c^2$, $m(pK\ell^-)_{p\leftrightarrow K} > 2320$ MeV$/c^2$	all
	D^0	$\lvert m(K^-\ell^+)_{\ell\leftarrow\pi} - 1865\rvert > 20$ MeV$/c^2$	all rare
	swaps	$\lvert m(K^-\mu^+)_{K\leftarrow\mu} - 3097\rvert > 35$ MeV$/c^2$	rare $\mu\mu$
		$m(K^-e^+)_{K\leftarrow e} < 2900$ or > 3150 MeV$/c^2$	rare ee
	conversions	$m(K^-e^+)_{K\leftarrow e} > 10$ MeV$/c^2$, $m(pe^-)_{p\leftarrow e} > 10$ MeV$/c^2$	all ee
	clones	$\theta(K,\ell) > 0.5$ mrad, $\theta(p,\ell) > 0.5$ mrad	all
Combinatorial background	–	`BDT>0.7`	all $\mu\mu$
	–	`BDT>0.7`	all ee
HOP	–	$m_{HOP}(\Lambda_b^0) > (5265\,\text{MeV}/c^2 + 10 \times log(\chi^2_{FD}(\Lambda_b^0)))$	rare ee

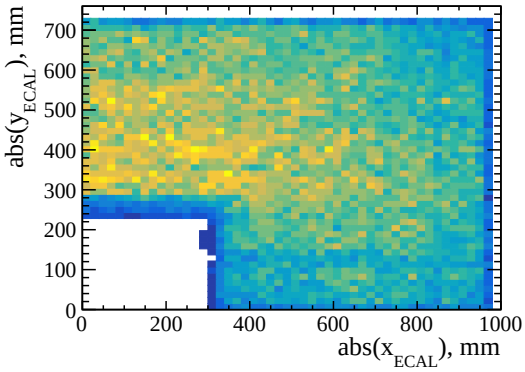


Figure 3.23: Occupancy of the innermost ECAL region in the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ data (Run I), with loose preselection applied. The blue cells in the down-left are not read out.

Muons are required to fall in the acceptance of the muon system. This is required to control properly both particle identification and trigger responses.

3.4.2 Particle identification requirements

As it has already been mentioned, there are several types of particle identification variables used in LHCb. The first one, `PIDX`, is the difference of the likelihoods of the track being identified as `X` and being identified as a pion. Another set of variables, `ProbNNX`, is the output of a neural network providing the probability of track being identified as `X`. The former variables are not limited in their values, while the latter are limited to the interval $[0, 1]$. For the `ProbNNX` variables, there are several alternative "tunes" of the neural network. There are separate tunes for Run I (labeled as "MC12") and Run II ("MC15") conditions. In all cases, the most optimal configuration for a given data taking period and particle species is used; these tunes are always specified in Tables 3.6-3.7. There are also some binary variables such as `isMuon`, returning 0 or 1.

In order to be able to reach a reasonable retention rate, loose PID requirements are already placed in the stripping lines (see Table 3.3). Tighter PID requirements are added on top of the stripping requirements, they are based on the misidentification background composition discussed above, and are different between Run I and II. This is due to differences in the global PID optimisation algorithms (so-called PID tunes) between Run I and Run II.

The usage of PID variables is strictly valid only in the kinematic regions covered by the calibration samples used to train the corresponding PID response. Furthermore, it is possible to precisely estimate the efficiency of these requirements only in these kinematic regions. The cuts are applied on the transverse momenta of the final state particles to match this requirement, as described in Tab. 3.6- 3.7. In addition, the momentum of the proton is required to be larger than 10 GeV/c being roughly the RICH threshold of creating a Cherenkov ring for kaons: below this value, the RICH cannot distinguish properly between protons and kaons.

Finally, the fake tracks (also called *ghost* tracks) are rejected applying the cut on the output of the dedicated neural net classifier, described in Ref. [109].

3.4.3 Multivariate selection against the combinatorial background

A separation of signal and background events can be performed in different ways. The most straightforward one is based on rectangular cuts on discriminating variables. Such approach is inefficient when these variables have complicated dependencies of each another. For this reason, so-called multivariate analysis techniques (MVA) are more powerful to separate signal and background than the cut-based methods. In order to further suppress the background and purify the data samples of the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays, MVA techniques described in Ref. [151] have been used. As it has already been mentioned in Sec. 3.3.1.a, the main purpose of the multivariate classifier in this analysis is to reject the combinatorial background.

The gradient boosting decision tree (BDTG) method was chosen as it combines high efficiency and simplicity, as well as high speed of the training phase. Several alternative classifiers have been tried, but their performance was found to be worse.

3.4.4 Overview of the BDTG algorithm

Multivariate algorithms used in high-energy physics are based on many different approaches, most widely used ones are Neutral Networks (NN), Multi-Layer Perceptrons (MLP) and Boosted Decision Trees (BDTs). Among them, BDTs are the most simple and fast in training, while being efficient enough for most of the use-cases. BDT classifiers are based on the machine learning techniques. The fundamental unit of a BDT is the Decision Tree [151, 152]. In a Decision Tree (Fig. 3.24), sequential rectangular cuts are applied on a provided set of variables, applying the optimal cut at each step. This allows to split the N-dimensional space, defined by the provided variables, into smaller partitions. In each partition, the optimised selection is applied, and the procedure repeats iteratively.

In order to overcome potential biases, instead of a single Decision Tree, an ensemble of trees ('forest') is usually considered. This procedure is called *boosting*. It allows to consider a collective boosted decision as an average of responses of each single tree. Events misclassified by the previous tree are taken with larger weight in the subsequent tree, allowing to improve the BDT response. The difference between various BDT algorithms is in the exact implementation of the boosting algorithm. The gradient boosting is more robust in presence of outliers or mislabeled data points [151].

Fundamental phases of the BDTG algorithm are:

- Training phase: the algorithm ”learns” how to discriminate signal from background, based on the provided signal and background proxy samples.
- Testing phase: the trained classifier is applied to statistically-independent signal and background samples, with respect to samples used for training. This allows to check whether the algorithm was trained properly, or was *overtrained*, *i.e.* learned to pick up insignificant statistical fluctuations.

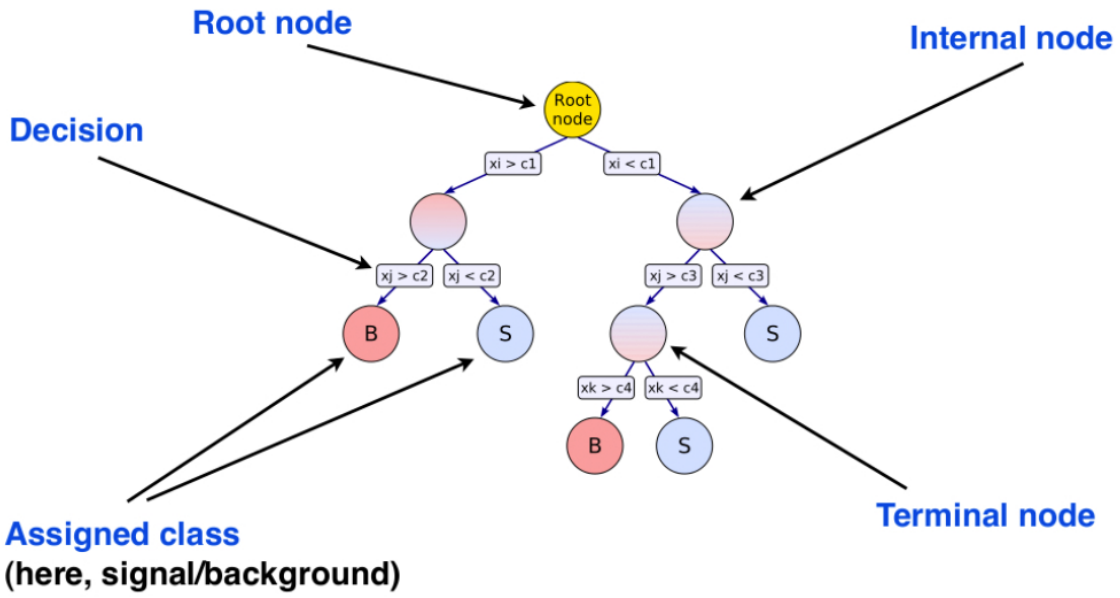


Figure 3.24: Structure of a single Decision Tree, allowing to classify events as signal (S) or background (B) by applying a set of rectangular cuts on the variables x_i. Taken from Ref. [152].

- The BDT classifier is applied to the data sample of interest, in a form of a new variable. The output of the classifier is a function of the variables used for training, and its value is distributed between -1, for "background-like", and 1, for "signal-like" events.

There are some important options which allow to tune the training phase of the BDT algorithm:

- Number of trees: the number of decision trees in the forest. Larger number of trees allows to improve the performance of the algorithm, but can also lead to the overtraining;
- Maximum depth: maximum depth of nodes allowed for each decision tree. Larger depth allows to improve the performance of the algorithm, but can also lead to the overtraining;
- Number of cuts: the number of steps during cut optimisation;
- Bagging fraction: the fraction of events to be used in each iteration;
- Shrinkage: learning rate for the gradient boost algorithm.

3.4.5 Training of the multivariate classifier

For the BDTG training, a signal and a background sample have to be provided, to be more precise – one set of samples for the muon mode and another one for the electron mode. As

a signal proxy, a set of simulated candidates, corresponding to the decay $\Lambda_b^0 \to pK\mu^+\mu^-$ for muons, and $\Lambda_b^0 \to pKe^+e^-$ for electrons, was used. A sample of data taken from the upper sideband of the signal, was used as a background proxy. The upper sideband was defined as $m(pKl^+l^-) > 5825$ MeV/c^2; and was limited from above by a stripping cut at $m(pKl^+l^-) = 6780$ MeV/c^2. The procedure was otherwise kept identical between the muon and electron channels. The number of events used for the training is about 15 000 signal candidates for each of the four classifiers, and about 4 500 background candidates for each of two muon classifiers, and about 2 500 background candidates for each of the two electron classifiers. The last two numbers are limited by the amount of data in the considered sideband region, and they limit the number of trees which can be used to train the classifier without overtraining.

Separate BDTs were trained for Run I and Run II to take into account differences in the kinematics, running conditions etc, however, the same BDT classifier was used for 2011 and 2012 datasets, given the little difference in their underlying kinematics. Therefore, four different classifiers are used in the current analysis (Run I/II, electron/muon modes).

The same exact BDT classifier and the cut value on its output are used throughout different q^2 regions (it is then validated in Fig. 3.29 that the BDT has a similar performance across the broad q^2 window). For instance, it is shown later in Sec. 3.6.3 that the BDT efficiency is very similar for the $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ decay modes, so cancels in the ratio of the two. If the similar cancellation is achieved for the electron modes (where the efficiencies are currently blind), the BDT efficiency will largely cancel in the double ratio, not bringing any potential biases to the measurement. It should be noted that the BDT classifiers are trained only in the q^2 region (0.1, 6) GeV$^2/c^4$ where the maximal sensitivity is required. The resonant region ($\Lambda_b^0 \to pKJ/\psi$) is not used for the training despite the large available amount of data due to the very different nature of background composition: the test classifier trained on the resonant mode had much worse performance on the non-resonant channel.

The output of the classifier provides a new variable that is then used to apply one of the selection requirements. Prior to the training, most of the selection requirements relying on detector acceptance, PID, q^2 range limited to (0.1, 6) GeV$^2/c^4$, *etc.*, described in Tab. 3.6-3.7, were applied to both signal and background samples in a consistent manner. This ensures the absence of mismatches in the kinematic cuts between signal and background proxy samples, which could be picked up by the BDT and reduce its performance. Also, both samples were required to pass the L0 (either of trigger categories used in this analysis) and HLT selections, and all the mass vetoes used throughout this analysis, were applied. This (together with the PID requirements) is particularly important, since the upper sideband originally contains not only the combinatorial background, but also certain misidentification and over-reconstructed backgrounds, which have quite different kinematics. Applying such requirements allows to access a rather pure sample of the combinatorial background.

The training setup is presented in the Table 3.8. The only difference in the setup of the classifiers is the number of trees: due to limited size of background samples, the usage of larger number of trees could lead to an over-training. This explains why the number of trees is smaller for 2012 than for 2016, and for electron than for muon samples.

Table 3.8: BDTG training setup.

	Run I μ	Run I e	Run II μ	Run II e
N_{trees}	300	220	350	300
max tree depth	4			
Shrinkage	0.11			
Bagging fraction	0.73			
N_{cuts}	17			

The PID, multiplicity and kinematic weights, as well as the reweighting of the pK spectrum (see Sec. 3.5), were applied to the simulation samples in order to correct for the data-simulation discrepancies in the latter. The reweighting of the decay model (pK spectrum) is especially important to apply, because the phase-space model used in the simulation significantly underestimates the fraction of low-mass pK contributions, leading to the BDT being trained for rather high pK region, where very few resonances are present in data.

A set of 21 variables (features) was used for the training. This set is already a result of an iterative removal of variables having low separation power, from the initial, much larger set. When building the classifier, *i.e.* identifying which variables should be used, it was observed that certain variables showed a high separation power, nevertheless it was decided to not take them into account. For example, the polar angle of the final state hadrons, although having a reasonable separation power, is not used in the training due to its correlation with the (pK) spectrum which is not modeled properly in the simulation (although the reweighting is applied, the precise decay model remains unknown).

The list of variables used includes:

- $\beta' = \frac{p(J/\psi)-p(p)-p(K)}{p(J/\psi)+p(p)+p(K)}$,
- $p_T(\Lambda_b^0)$, $p_T(\Lambda^*)$, $p_T(\Lambda^*) + p_T(dilepton)$,
- $\chi^2_{IP}(\Lambda_b^0)$, $\chi^2_{IP}(\Lambda^*)$,
- $DIRA(\Lambda_b^0)$,
- $\chi^2_{FD}(\Lambda_b^0)$, $\chi^2_{FD}(dilepton)$,
- $\chi^2_{DTF}/n_{dof}(\Lambda_b^0)$,
- $\chi^2_{vertex}(\Lambda_b^0)$, $\chi^2_{vertex}(\Lambda^*)$,
- $p_T^{min}(p, K)$, $p_T(p) + p_T K)$,
- $(\chi^2_{IP})^{min}(p, K)$, $\chi^2_{IP}(p) + \chi^2_{IP}(K)$,
- $p_T^{min}(\ell^+, \ell^-)$,

- $(\chi^2_{IP})^{min}(\ell^+, \ell^-)$, $\chi^2_{IP}(\ell^+) + \chi^2_{IP}(\ell^-)$,
- $p(p)$,
- $\eta(p) + \eta(K)$.

Some of these variables have significant correlations between themselves in the signal sample, however the correlations are much smaller for the background sample, and removal of one of these variables degraded significantly the BDTG performance.

A *k-fold* approach was adopted for the BDT training, with $k = 10$: 9/10 of the dataset were used to train the BDT applied afterwards to the remaining 1/10, and the procedure repeated 10 times.

The resulting ten receiver operational characteristic (ROC) curves are shown in Fig. 3.25. In addition, the classifier was checked to have no overtraining, by means of Kolmogorov-Smirnov (KS) test values. This test is implemented in the way that KS test values very close to zero (below 0.05) can be considered as a sign of overtraining (*i.e.*, training and testing distributions are different). An example of comparison of training and testing distributions is shown in Fig. 3.26, and no overtraining is seen, neither visually, nor by means of the KS test.

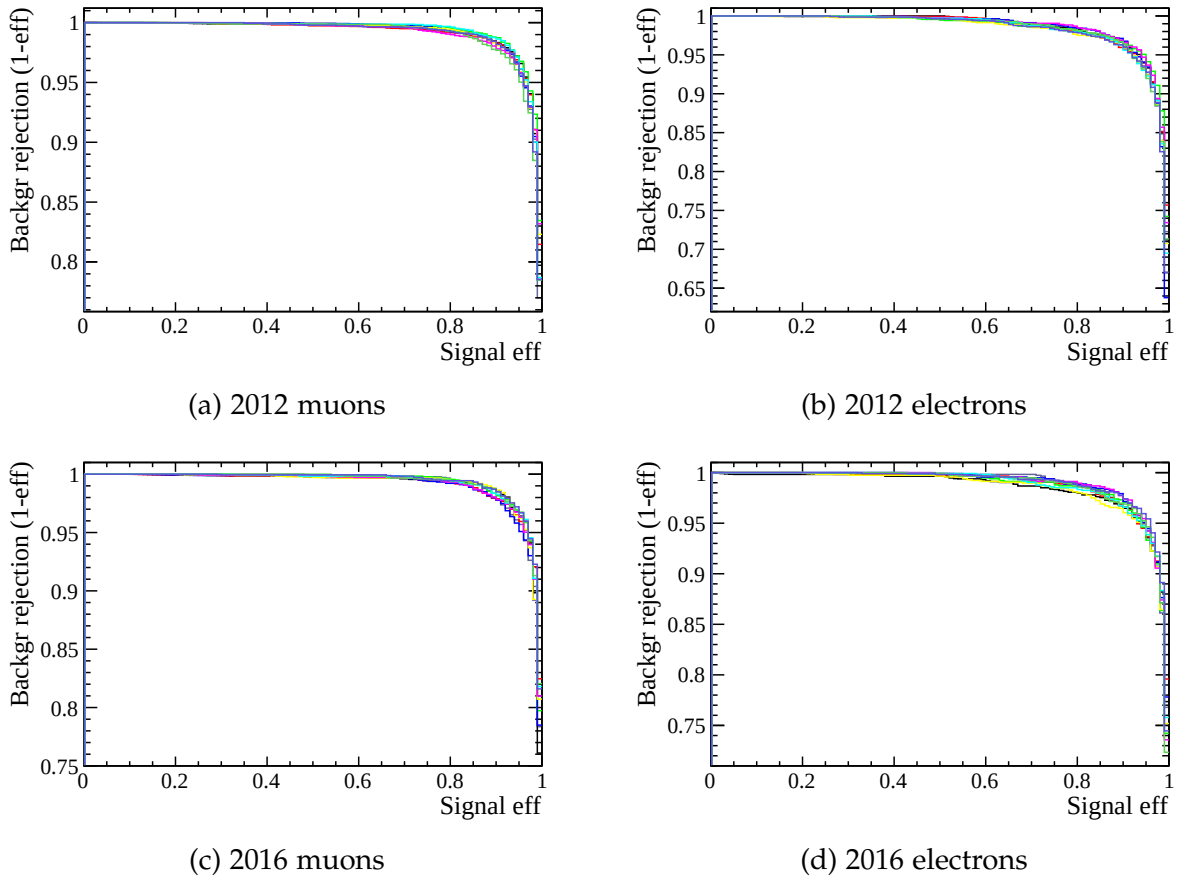
When using a large number of variables in the BDT classifier, one has to be careful and make sure not to pick up any discrepancies between the data and simulation resulting in a mismodeled BDT distribution. The most stringent test of the BDT validity is comparison of the BDT output variable between the data and simulation, shown in Figs. 3.51, 3.52 for the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ modes. The data distribution is accessed via the ${}_s\mathcal{P}lot$ technique (as it is later discussed in Sec. 3.5.3), while the simulation has the set of the discussed corrections (PID, multiplicity, kinematics) applied (Sec. 3.5). The fit to the data is performed with a loose BDT cut of -0.2 due to the large background at low BDT values making an ${}_s\mathcal{P}lot$ unreliable, so the validation works only above this value (which is anyway the region which is the most interesting for comparisons). No significant differences are seen between the data and simulation distributions. This also validates usage of a single BDT for 2011 and 2012 samples. A similar check is later performed for the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode in Sec. 3.10.6.a.

3.4.5.a Cut optimisation

The cut value of the BDT classifier output has to be optimised in order to achieve the optimal sensitivity to the rare $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ modes. No optimisation is performed for the normalisation or control modes, as the same exact cut value is used across q^2. For the cut optimisation, all ten folds considered at the training step are merged.

A figure of merit defined as $P = \frac{S}{\sqrt{S+B}}$, where S is the signal yield, and B is the background yield, was used to establish the optimal cut value on the BDT output. As it can be seen from the definition, it requires the knowledge of the absolute (expected) signal yield.

In practice, S is the expected signal yield at a given BDT cut value. As branching fractions of decays of interest $\Lambda_b^0 \to pK\ell^+\ell^-$ have never been measured, it is difficult to



(a) 2012 muons

(b) 2012 electrons

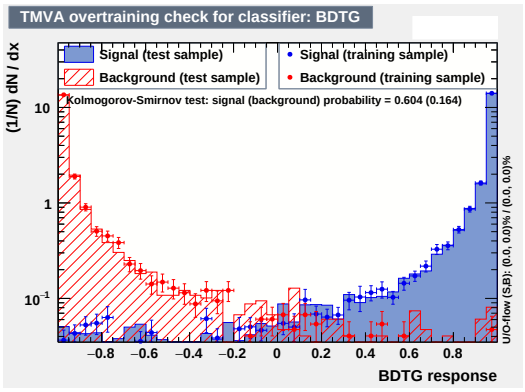
(c) 2016 muons

(d) 2016 electrons

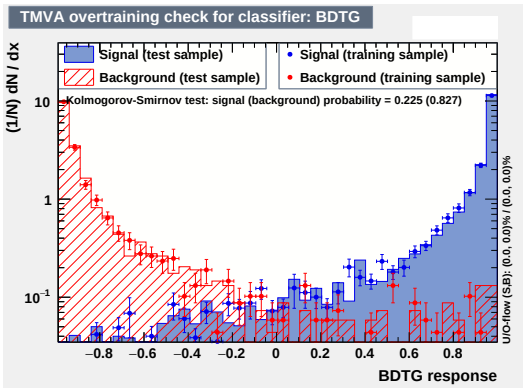
Figure 3.25: ROC curves for ten BDT folds, each shown in a different colour.

estimate the expected signal yield without relying on data. At the same time, taking the signal yield directly from data might bias the cut optimisation procedure. Therefore, the signal yield is estimated from the relevant signal simulation sample, scaled by a certain factor to match the expected signal yield in data. To compute this factor for the $\Lambda_b^0 \to pK\mu^+\mu^-$ mode, the signal yield in data at the starting point of the scanned BDT cut values range (BDT=0.5) was taken as a reference point. For the electron mode, the value taken in the muon mode was then scaled by the ratio of yields of $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ modes.

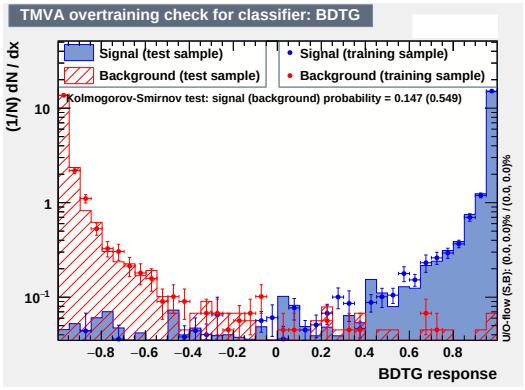
Then, at each point of the scan, a simple fit was performed to the data of the $\Lambda_b^0 \to pK\ell^+\ell^-$, involving only signal and exponential-shaped background. The signal yield in the fit was blinded and not looked at. The background shape was used to estimate the background yield under the signal peak, in a window of 5σ for the muon mode, and 3σ for the electron mode , where σ is a value of the mass resolution taken from the signal simulation. 5σ interval is required in the muon case due to small resolution and rather small background level, resulting in somewhat unreliable estimate coming from the narrow 3σ window. At the same time, for electrons, the mass resolution is much larger, so taking 5σ would result in a way too broad mass window. This obtained estimate of the



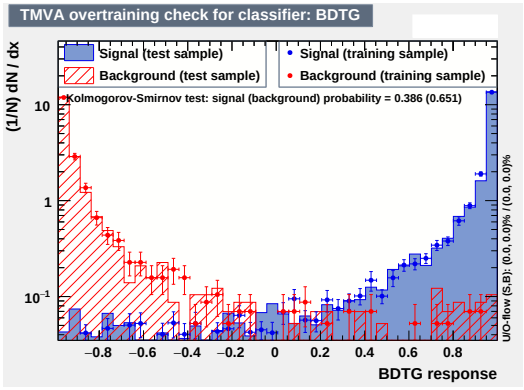
(a) 2012 muons



(b) 2012 electrons



(c) 2016 muons



(d) 2016 electrons

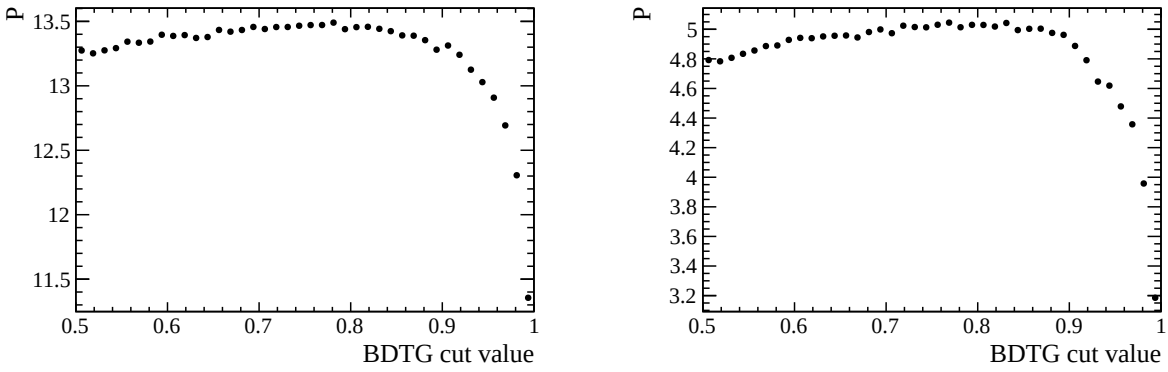
Figure 3.26: BDT distributions for signal and background proxy samples, compared between the training and testing samples (in logarithmic scale). No overtraining is visible.

background yield was taken as B in the figure of merit computation. The fit was always performed in a broader range than the one used to estimate the background yield, in order to ensure the fit stability. The fit convergence was checked at each step, and it was also controlled that the obtained values of B did never increase with increasing the cut value (that could have happened if one of the fits had failed).

The distributions of the FOM value (denoted as P) as a function of the BDT cut value, are shown in Fig. 3.27 for the Run I and II (muon and electron modes).

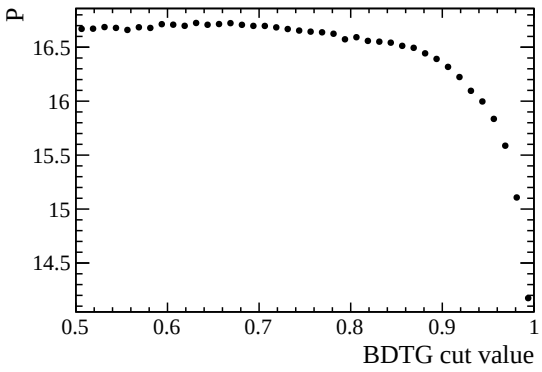
In Run I, the cut value of 0.8 was chosen for both the muon and the electron modes, providing the highest FOM values and keeping the signal efficiency high at the same time. At the same time, in Run II the FOM shape suggests a looser cut value, so the cut value of 0.7 was chosen as a default for both muon and electron modes.

It is also important to make sure that the BDT will not create fake peaks out of a flat background sample. To check that, the sample of background is taken from the simulation: taking the $\Lambda_b^0 \to pK\ell^+\ell^-$ simulation and requiring that none of the four tracks have Λ_b^0 as their mother, and not requiring that tracks have a correct identification, gives a sample behaving very close to the combinatorial background in the data. The trends of the mean

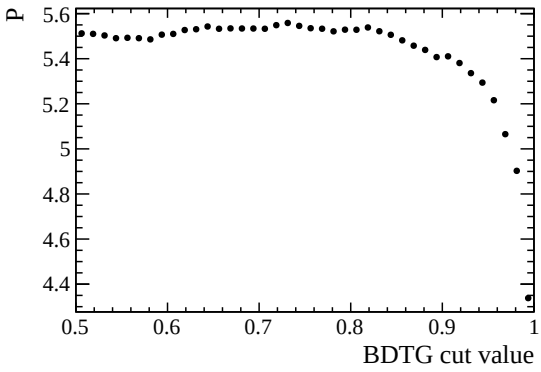


(a) Run I muons

(b) Run I electrons



(c) Run II muons



(d) Run II electrons

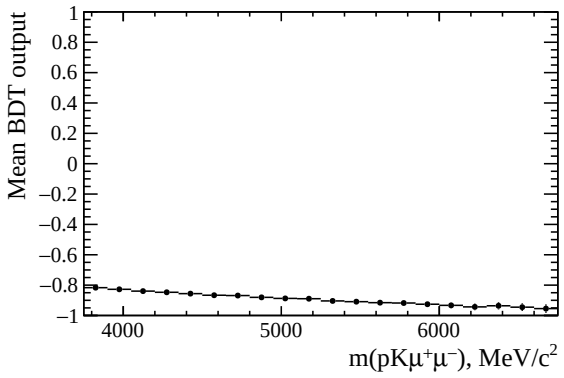
Figure 3.27: BDT output cut value optimisation: distributions of the FOM values as a function of the cut value. Only central values are shown, without uncertainties.

BDT value as a function of the invariant mass in these samples are shown in Fig. 3.28, and, even though the descending trend is seen, it has a very linear behaviour and does not create any fake peak around the known Λ_b^0 mass. We also check the BDT flatness as a function of the q^2 and $m(pK)$ in Fig. 3.29, using the signal simulation samples.

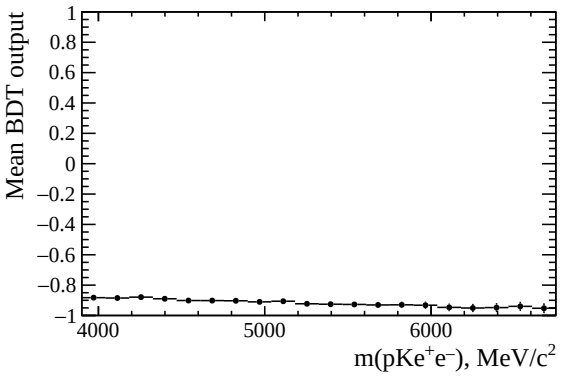
3.4.6 Momentum balance requirement: HOP

An additional suppression of partially reconstructed backgrounds that contaminate the electron channels, is provided by constraining the kinematics of the decay [153]. The vectorial sum of the momenta of the final state particles is not expected to have a component orthogonal to the flight direction of the Λ_b^0, defined by the primary and the decay vertices.

Fig. 3.30 presents schematically the quantities relevant for the definition of the so-called `HOP variable` α_{HOP} which accounts for the momentum balance in the event. The ratio between the transverse (with respect to the Λ_b^0 direction of flight) momenta of the Λ^* (*i.e.*, proton-kaon pair), and that of the dielectron pair, is used:

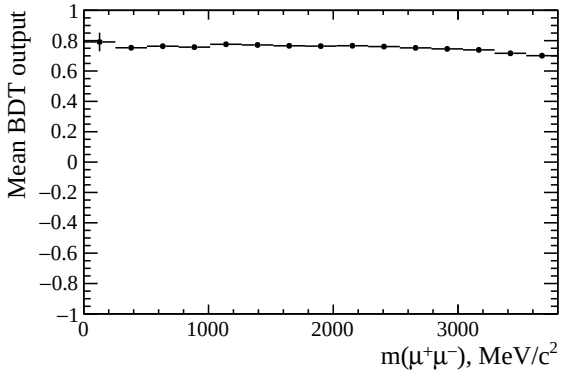


(a) Muons

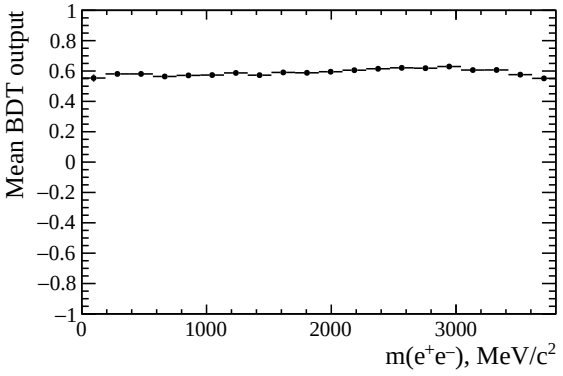


(b) Electrons

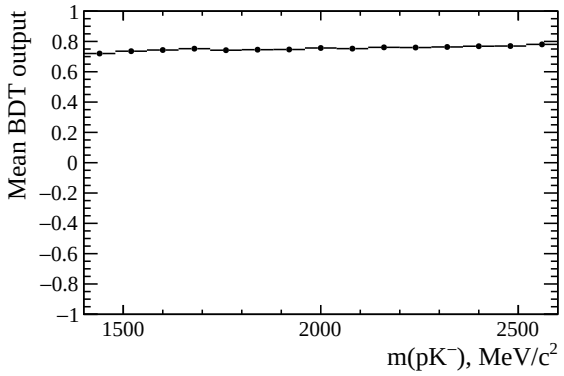
Figure 3.28: Mean BDT value as a function of invariant mass in a sample of background from the simulation. No peaking structures are seen.



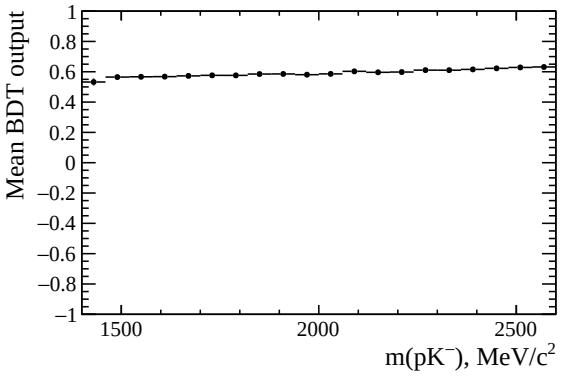
(a) Muons



(b) Electrons



(c) Muons



(d) Electrons

Figure 3.29: Mean BDT value as a function of (a-b) q^2 and (c-d) $m(pK)$, in the $\Lambda_b^0 \to pK\ell^+\ell^-$ signal simulation. No peaking structures are seen.

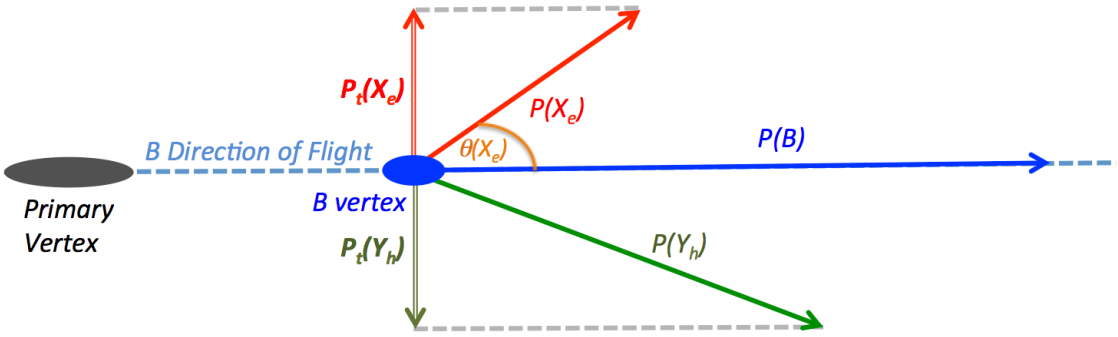


Figure 3.30: Sketch of the kinematics of a b-hadron decay, with hadronic Y_h and leptonic X_e parts. Taken from [153].

$$\alpha_{\mathrm{HOP}} = \frac{p_{\mathrm{T}}(\Lambda^*)}{p_{\mathrm{T}}(ee)}$$

When α_{HOP} deviates from one, some energy is missing in the final state. For signal events, the missing energy is most likely carried away by bremsstrahlung photons emitted by the electrons. Given that, the electron momentum can be corrected as:

$$p_{\mathrm{corr}}(ee) = \alpha_{\mathrm{HOP}} \cdot p(ee).$$

Since bremsstrahlung photons are emitted in the direction of flight of the electron, the same α_{HOP} correction is valid for the longitudinal component of the di-electron momentum. On the other hand, the missing particles in partially-reconstructed background events are not necessarily emitted in the direction of flight of the electrons, and therefore the α_{HOP} correction does not work in the same way. A similar argument applies to the combinatorial background.

The corrected momenta can be used to re-calculate the invariant mass of the Λ_b^0 candidate ($m(pKe^+e^-)_{\mathrm{HOP}}$). The resolution of $m(pKe^+e^-)_{\mathrm{HOP}}$ depends on the quality of the vertex reconstruction and on the Λ_b^0 decay time, and degrades as a function of q^2. This degradation happens because the average angle between the Λ_b^0 direction of flight and the dielectron system decreases with increasing the q^2, which affects the $m(pKe^+e^-)_{\mathrm{HOP}}$ resolution [153]. The effect of degrading the $m(pKe^+e^-)_{\mathrm{HOP}}$ resolution can be clearly seen from Fig. 3.32.

Figure 3.31 shows the dependence of the $\chi^2_{FD}(\Lambda_b^0)$ as a function of $m(pKe^+e^-)_{\mathrm{HOP}}$ in the considered q^2 region. Given the distributions for signal and background events, a two-dimensional cut is adopted:

$$m(pKe^+e^-)_{\mathrm{HOP}} > a_{\mathrm{HOP}} + b_{\mathrm{HOP}} \times log(\chi^2_{FD}(\Lambda_b^0))\,,$$

where the a_{HOP} and b_{HOP} coefficients are taken from the $R_{K^{*0}}$ analysis [2], with a_{HOP} being shifted by the mass difference of the Λ_b^0 and B^0. This results in the values $a_{\mathrm{HOP}} = 5265\,\mathrm{MeV}/c^2$ and $b_{\mathrm{HOP}} = 10\,\mathrm{MeV}/c^2$. Such an extrapolation is possible because the decays

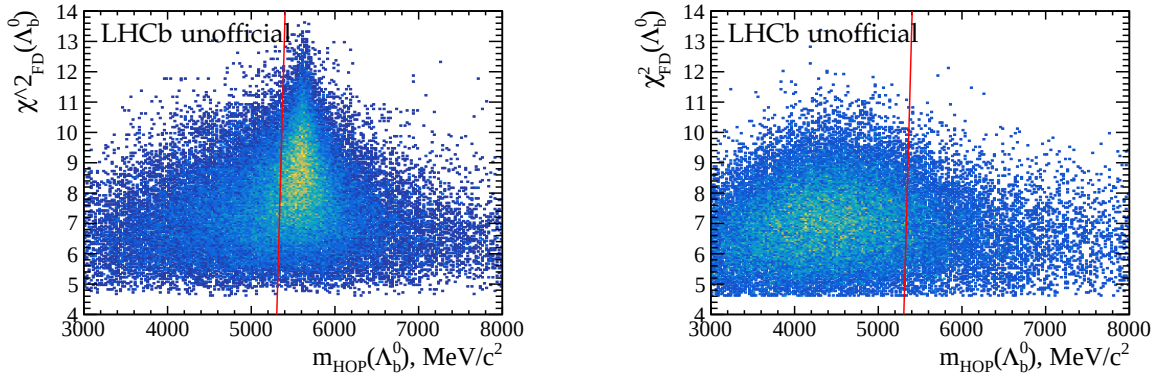


Figure 3.31: Distribution of the HOP mass as a function of the $\chi^2_{FD}(\Lambda_b^0)$ in the $0.1 < q^2 < 6$ GeV$^2/c^4$ region, for the (left) $\Lambda_b^0 \to pKe^+e^-$ signal simulation; and (right) simulation of the $\Lambda_b^0 \to pD^0\mu\nu_\mu$ semileptonic background. The red line depicts the applied cut.

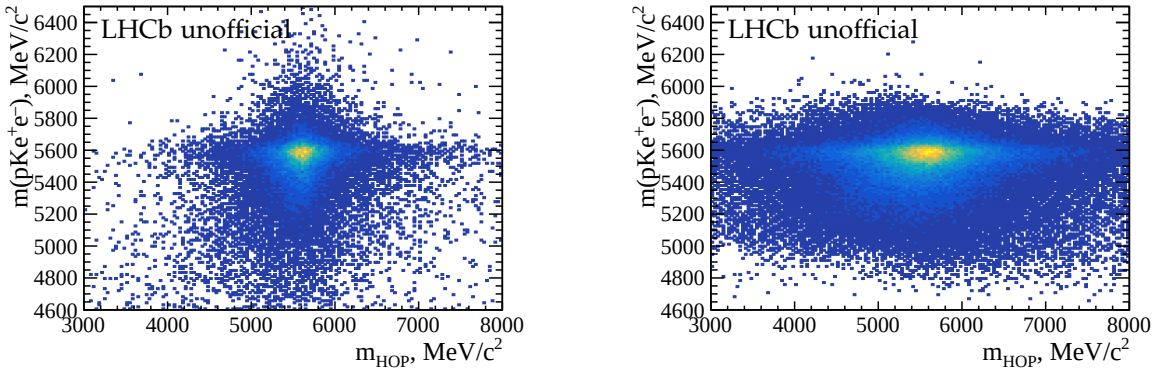


Figure 3.32: Relation between m_{HOP} and Λ_b^0 invariant mass in the simulation. Left: in the $\Lambda_b^0 \to pKe^+e^-$ decay in the $0.1 < q^2 < 6$ GeV$^2/c^4$; right: in the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ decay.

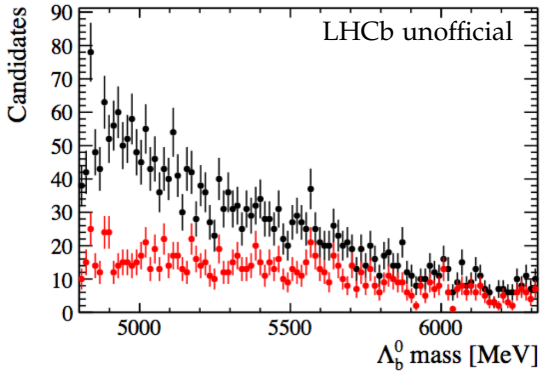


Figure 3.33: Effect on the invariant mass shape of the combinatorial background after applying the cut on the HOP mass. Run II data from the negative BDT output region is used for this test.

$B^0 \to K^{*0} e^+ e^-$ and $\Lambda_b^0 \to pKe^+e^-$ are quite similar from the kinematic point of view, so that the resolution in m_{HOP} is comparable for these two decay modes.

The HOP cut is not applied to the resonant decay modes (in the q^2 regions of J/ψ and $\psi(2S)$ resonances), due to the degradation of the HOP mass resolution with q^2, which is illustrated in Fig. 3.32. However, strictly speaking, this approach can introduce a bias in the double ratio, if the efficiency of the HOP requirement is mismodeled by the simulation, as the efficiency of the HOP cut does not cancel. Therefore, a systematic uncertainty is assigned, to account for a potential mismodeling of the HOP efficiency in the simulation (Sec. 3.10.2), and it is at the level of 2%.

It should be noted that the HOP mass does not correlate with the invariant mass, as it is shown in Fig. 3.32(left). Therefore, applying a cut on the HOP mass does not sculpt any smooth background and does not create any fake peaks. This is proven in Fig. 3.33 which is showing an effect of the HOP cut being applied to the sample of combinatorial background in the $\Lambda_b^0 \to pKe^+e^-$ dataset (extracted from the negative-BDT region). This check has been performed for both Run I and Run II data, in two alternative regions of the BDT output variable: $-1 < BDT < -0.8$ or $-0.8 < BDT < 0.4$, with a similar conclusion for all cases. Since in the invariant mass fits the shape of the combinatorial background is parametrised by an exponential of a free slope, this parametrisation holds either before or after applying the HOP cut.

3.4.7 Multiple candidates

Once the entire selection is applied, a search for candidates sharing the same event and run numbers is performed. Those candidates are called "multiple candidates". In data, when such candidates are found, only one is retained randomly for this analysis, and all the others are removed. This corresponds to a removal of about 2.6 (0.4) percent of candidates in the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$, 2.5 (0.6) percent in $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, 1 (1) percent in $\Lambda_b^0 \to pK\mu^+\mu^-$ and 1 (2) percent in $\Lambda_b^0 \to pKe^+e^-$ decay modes for Run I (Run II), in the $m(pK\ell^+\ell^-)$ mass range considered as the fit region. The Run II percentage of multiple candidates is in general lower than Run I, this is due to the tighter $p_T(p)$ cut and the more efficient PID selection.

In the signal simulation, the fully truth-matched events are considered, *i.e.* the events for which all reconstructed tracks match to all generated tracks. In addition, events with at most one mismatched track are allowed. Those can correspond, for example, to the cases of kaon decays in flight within the volume of the detector. If there is more than one candidate per event, the one which is kept is selected randomly, however, the preference is always given to the fully truth-matched events.

3.5 Calibration of the simulation

The LHCb simulation is known not to be a perfect reflection of data. Therefore the variables used for the measurement, in particular those used when estimating the efficiencies, have to be "corrected" for. In practice, this is done by applying a collection of *weights*, also called *corrections*, to the simulated spectra, in order to match them to the data ones. This section summarises the way these weights are computed, and how they are applied to the simulation.

The full chain of corrections is applied to the simulated samples of the signal and normalisation modes. It is also applied to the samples used for the BDT training (Section 3.4.3) and finally in the efficiency estimations (Section 3.6). As will be described in Section 3.7, backgrounds originating from the misidentification of tracks are modeled in the fits to the data using templates extracted from corresponding simulation samples. Dedicated particle identification corrections (Section 3.5.8) are also applied to these samples.

The corrections (except for the particle identification, which are discussed in Sec. 3.5.2) are computed using data and simulation samples of the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ decay modes, and consequently extrapolated to other decay modes of interest. These decay modes are chosen as they are the most abundant of those used in the current analysis.

Corrections are computed sequentially, each on top of the previous ones. The correction for the "physics", *i.e.* a proper model of the decay, being the first, and followed by corrections for the improper description of various quantities (event multiplicity, kinematics of the Λ_b^0, Λ_b^0 lifetime) at the generation step, and the improper simulation of the detector response (particle identification, Λ_b^0 decay time acceptance, and trigger response). In all the cases, where $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ or $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ data is exploited, almost complete selection is applied to this data, involving most notably the PID and a loose requirement on the preliminary BDT classifier output, in order to get a reasonably clean sample.

Corrections of the generated decay model, generated event multiplicity, generated Λ_b^0 kinematics and generated Λ_b^0 lifetime are applied both to the reconstructed- and generator-level simulation samples, and are assumed to be identical for electron and muon decay channels. Other corrections, which deal with describing the detector response (rather than event generation itself), are applied only to the reconstructed-level simulation samples, and are in principle different between different lepton channels.

3.5.1 Decay model

As it is shown in Ref. [4], the three-body decay $\Lambda_b^0 \to pKJ/\psi$ has a rich resonant structure in pK and pJ/ψ invariant mass distributions. No resonances are observed in the KJ/ψ invariant mass distribution. In the simulation samples of $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-, e^+e^-)$ decays, the proton and kaon in the final state are not produced via resonances, this type of event generation is called "phase-space" in the `EvtGen` package [154]. This can lead to important mismatches of the distributions of kinematic variables between data and

simulation. To account properly for the data-simulation differences emerging from an improper decay model, and those from improper detector simulation, the correction for the resonant structure of the hadronic part of the decay is applied before computing any other corrections.

To do so, a tool based on the current knowledge of the mixture of Λ^* resonances in the pK spectrum, and two pentaquark states in the pJ/ψ spectrum, and developed for the amplitude analysis of the $\Lambda_b^0 \to pKJ/\psi$ decay [4], was used. Effectively, it applies weights as a function of truth-level four-momenta of p, K and J/ψ, which can be visualised by a map shown in Fig. 3.34(a).

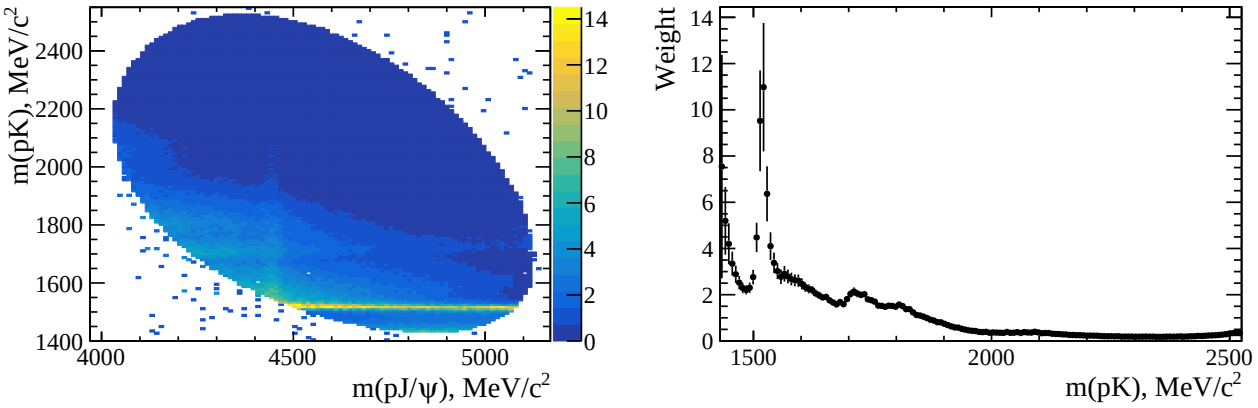


Figure 3.34: Map of corrections for the decay model of the $\Lambda_b^0 \to pKJ/\psi$ decay, a) as a function of its Dalitz plane, obtained using the dedicated tool; and b) one-dimensional weights as a function of $m(pK)$, as described in the text. Invariant mass is computed using truth-level four-momenta.

It is important to mention that this correction was measured on $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$ decays, and so far no equivalent exists for other $b \to c\bar{c}s$ transitions such as the $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ decay, or rare decays ($b \to s\ell^+\ell^-$ transition) such as $\Lambda_b^0 \to pK\ell^+\ell^-$.

An assumption is made that the composition of the pK spectrum is rather similar in all these transitions. This is a rather strong statement, but any bias related to the pK spectrum should largely cancel in a single ($r_{J/\psi}^{-1}$) or double (R_{pK}^{-1}) ratio. A systematic uncertainty is assigned to cope with this lack of information, as described in Sec. 3.8.4.

The pentaquark candidates were observed only in the (pJ/ψ) system, but not in the $(p\psi(2S))$. Even more, the pentaquark states decay via a strong transition, so are not expected to decay to non-resonant $(p\ \ell^+\ell^-)$ final state. Given these considerations, the tool described above cannot be directly used for channels other than $\Lambda_b^0 \to pKJ/\psi$. A dedicated set of one-dimensional weights is applied to the pK invariant mass only, computed using truth-level four-momenta of the proton and kaon particles. To do that, the decay model of the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ simulation is first corrected using the tool discussed above. Then, a comparison of the corrected and uncorrected distributions of $m(pK)$ in this simulation, allows to produce a set of binned one-dimensional weights presented in Fig. 3.34(b).

The effect of the correction applied to the phase-space pK spectrum is shown in

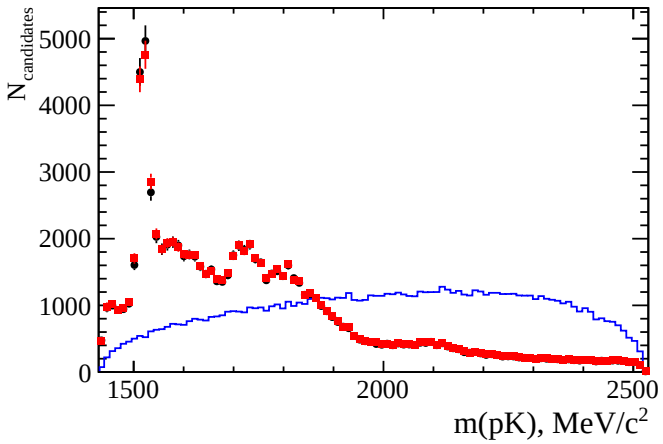


Figure 3.35: Invariant mass of pK system in the $\Lambda_b^0 \to pKJ/\psi$ simulation, computed using truth-level four-momenta, with the phase-space model (blue line), decay model corrected using the dedicated tool (black circles), and using the one-dimensional weights (red squares).

Fig. 3.35. It can be seen how different the corrected decay model is from the phase-space model: the latter is significantly underestimating the contribution of low-mass region, therefore biasing the kinematics of the decay. Also one can see that the one-dimensional weights in $m(pK)$ provide a sufficiently good description of the spectrum.

3.5.2 Particle identification

In Section 3.4 one can find the particle identification requirements applied to all the final state particles in the analysis. As it has been discussed in Sec. 2.8, the PID efficiency is estimated using the dedicated calibration samples, and is subsequently applied as a per-event weight to the simulation samples.

For protons, kaons and muons, the `PIDCalib` tool described in Sec. 2.8 is used. The calibration of the electron PID efficiencies, although relying on the same calibration samples and preselection requirements as those used by the `PIDCalib`, is performed using `fit-and-count` technique described in more detail in the Sec. 2.8.2.

3.5.2.a Calibration sample choice

The first important step before correcting the PID response in the simulation is the choice of the calibration sample, one important criterion is that the kinematic coverage of the calibration samples and the ones of the decays of interest overlap sufficiently.

The calibration samples available at the time of this analysis, for the Run I and Run II conditions of proton-proton collisions, have already been summarised in the Table 2.2.

To make sure that the calibration samples can be used for this analysis, the distributions of signal decay kinematics and the one of the calibration sample were compared for each particle, each sample, and both $\Lambda_b^0 \to pK\ell^+\ell^-$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$ decay modes. To do that, two-dimensional distribution of the particle momentum and pseudorapidity

(hereafter called "kinematic phase-space") is compared between calibration samples and simulation samples of decays of interest.

The event multiplicity cuts and subdetector acceptance requirements in the calibration samples are adjusted with those used in the offline selection.

For kaons, muons and electrons, default calibration samples have been used. For protons, the $\Lambda_c^+ \to pK\pi$ sample was used, as the default $\Lambda \to p\pi$ sample does not cover a part of the kinematic phase-space of $\Lambda_b^0 \to pK\ell^+\ell^-$ decays, and has an unnecessary correlation of the PID efficiency with the Λ decay time.

It should also be noted that various calibration samples have different selection cuts applied in order to purify the signal. For this analysis, the cuts on the (transverse) momentum of a considered particle are most important, because they define the kinematic range in which the PID efficiency can be obtained from these samples. These cuts (hereafter referred to as `fiducial`) have to be as well applied to all studied decays, and have already been discussed in Sec. 3.4.2 and Tables 3.6-3.7. The coverage of the kinematic phase-space by calibration samples was checked before and after applying these fiducial cuts, as illustrated (on an example of muons) in Fig. 3.36. It should be noted that in some cases, the efficiency of fiducial cuts is not close to one, however, these cuts are crucial to clean the signal from dangerous misidentification backgrounds, by avoiding the regions where the PID performance is not ideal.

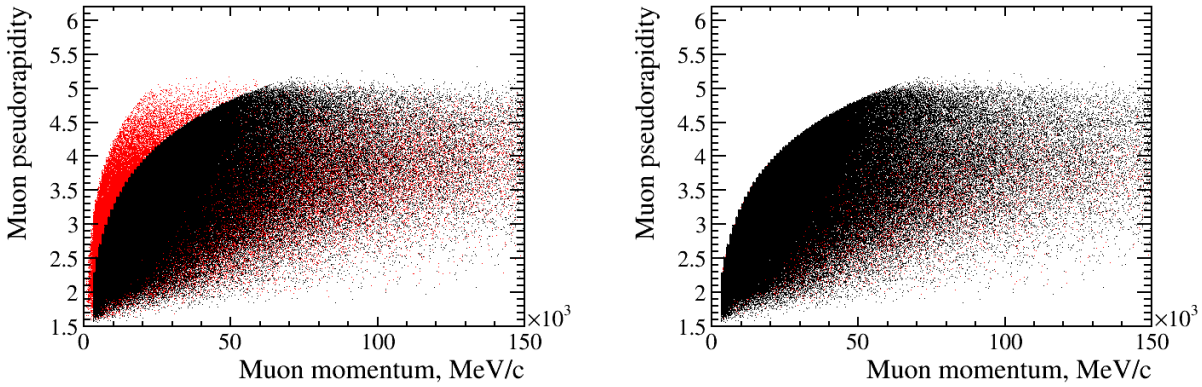


Figure 3.36: Comparison of the phase space covered by muons from the $\Lambda_b^0 \to pK\mu^+\mu^-$ simulation (red) and available calibration samples (black), before (left) and after (right) applying the fiducial cut. Regions which are red on this plot, are poorly covered by the calibration sample. It can be seen that the coverage is good after applying the fiducial cuts.

3.5.2.b Binning scheme choice

When computing the PID efficiency maps, it is important to pick the optimal binning scheme. The binning should be fine enough to describe sufficiently well the efficiency variations, and at the same time the statistics in each bin should be large enough to avoid significant fluctuations.

The binning variables considered were: track momentum, track pseudorapidity, and event multiplicity (represented by the number of tracks in the event). By default, the two-dimensional efficiency maps were studied (in track momentum versus track pseudorapidity), while the three-dimensional maps, also accounting for the minor multiplicity dependence, were used to estimate the systematic uncertainties. For electrons, the dependence of the PID efficiency on the number of bremsstrahlung photons attached to an electron, was considered in addition, because the electron PID is partially relying on the bremsstrahlung information. Therefore, electron maps are computed separately for the case of no bremsstrahlung photons attached (`HasBremAdded=0`) and the presence of such photons (`HasBremAdded=1`).

Bins are defined independently in different variables, and are required to be as small as necessary, such that the efficiency is nearly constant within a given bin, but as large as possible in order to achieve small uncertainties on the individual efficiencies. In each of the binning variables, N iso-populated bins were created and the cut efficiency was computed in these bins. N was set to be between 6 and 35 depending on the variable and abundance of the calibration sample. Then, these one-dimensional bins with the efficiency difference smaller than 5σ (3σ for electrons, due to the smaller statistics of the electron calibration sample), where σ is the uncertainty on the efficiency in a given one-dimensional bin, were merged. Resulting one-dimensional binning schemes were used to construct the final two-dimensional (and alternative three-dimensional) binning schemes, which were used to compute efficiency maps.

3.5.2.c PID calibration maps

The list of exact cuts and additional pre-requirements applied to the calibration samples in Run I and Run II, is shown in the Tables 3.9 and 3.10. It should be noted that in certain cases the applied cuts are tighter than the fiducial cuts in the calibration samples. This is mostly related to the fact that low-momentum protons have very low PID efficiency, so are removed in the offline selection. All these cuts are also applied to analysed data, as it have already been discussed in Sections 3.2.6 and 3.4. Requirements for electron samples are also included here, but are used in fit-and-count tool instead of the PIDCalib (see Sec. 2.8.2).

The resulting PID efficiency maps for protons, kaons and muons, for 2011, 2012 and 2016 conditions are shown in Fig. 3.37, 3.38, 3.39, respectively. For simplicity, only the maps for one magnet polarity are shown, although in reality all the maps are computed for both polarities and applied appropriately.

The PID maps for electrons, obtained using the fit-and-count tool described in Sec. 2.8.2, are presented in Fig. 3.40.

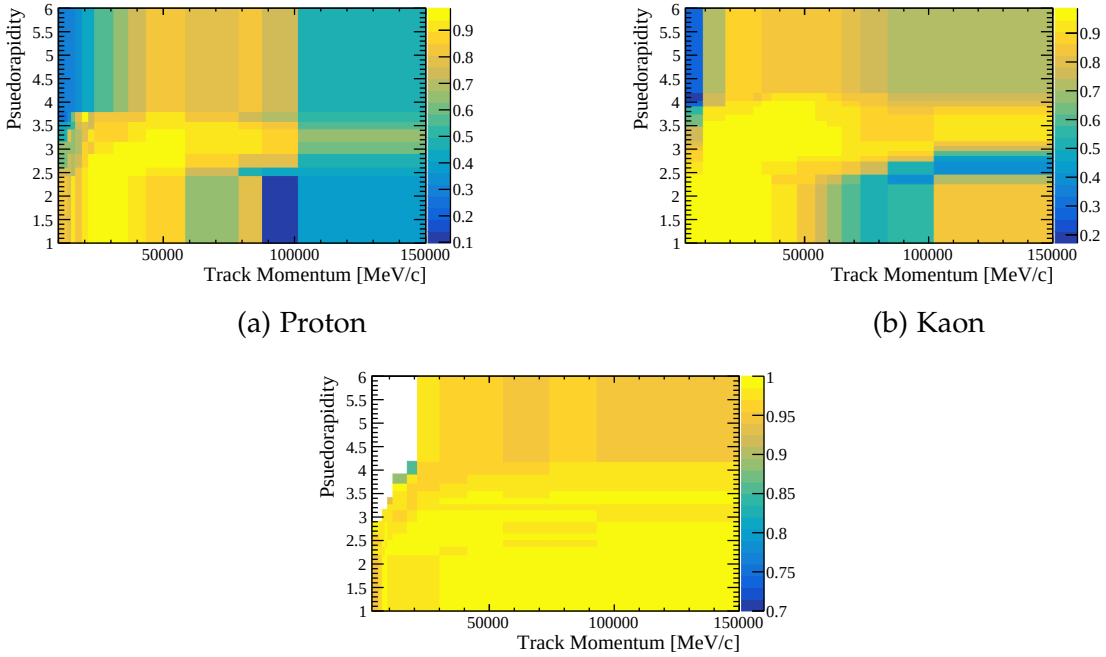
When applying the PID weights to the simulation, the following procedure was adopted. Each weight was applied per track, interpolating the corresponding map to the point of a given momentum and pseudorapidity. In case the interpolation fails (*e.g.* in the corners of the kinematic phase-space), the value of the closest bin was assigned as a weight. In exceptional cases of any unphysical values outside the range (0, 1) in the corners of the map, usually resulting from ${}_s\mathcal{P}lot$ features or divisions by zero in regions

Table 3.9: Requirements applied to the Run I calibration samples.

Particle	PID cut	Additional requirements
p	MC12TuneV2_ProbNNp>0.2 && MC12TuneV3_ProbNNK<0.8 && MC12TuneV3_ProbNNpi<0.7 && DLLp>-5.	HasRich==1.0 && nSPDHits<600 && PT>400 && P>10000
K	MC12TuneV3_ProbNNK>0.2 && MC12TuneV3_ProbNNp<0.8 && DLLK>-5.	HasRich==1.0 && nSPDHits<600 && PT>250 && P>2000
μ	MC12TuneV3_ProbNNmu>0.1 && IsMuon==1.	HasRich==1.0 && InMuonAcc==1.0 && PT>800 && P>3000 && nSPDHits<600
e (no brem)	MC12TuneV3_ProbNNe>0.1 && DLLe>0.	HasRich==1.0 && HasCalo==1.0 && PT>500 && P>3000 && nSPDHits<600 && HasBremAdded==0
e (has brem)	MC12TuneV3_ProbNNe>0.1 && DLLe>0.	HasRich==1.0 && HasCalo==1.0 && PT>500 && P>3000 && nSPDHits<600 && HasBremAdded==1

Table 3.10: Requirements applied to the Run II calibration samples.

Particle	PID cut	Additional requirements
p	Brunel_MC15TuneV1_ProbNNp>0.3 && Brunel_MC15TuneV1_ProbNNK<0.8 && Brunel_MC15TuneV1_ProbNNpi<0.7 && Brunel_DLLp>-5.	Brunel_HasRich==1.0 && Brunel_P>10000 && Brunel_PT>1000 && nSPDHits<450
K	Brunel_MC15TuneV1_ProbNNK>0.2 && Brunel_MC15TuneV1_ProbNNp<0.8 && Brunel_DLLK>-5.	Brunel_HasRich==1.0 && Brunel_P>2000 && Brunel_PT>400 && nSPDHits<450
μ	Brunel_MC15TuneV1_ProbNNmu>0.1 && Brunel_IsMuon==1.	Brunel_HasRich==1.0 && Brunel_InMuonAcc==1.0 && Brunel_P>3000 && nSPDHits<450
e (no brem)	Brunel_MC15TuneV1_ProbNNe>0.1 && Brunel_DLLe>0.	Brunel_HasRich && Brunel_HasCalo && Brunel_P>3000 && Brunel_PT>500 && nSPDHits<450 && Brunel_HasBremAdded==0
e (has brem)	Brunel_MC15TuneV1_ProbNNe>0.1 && Brunel_DLLe>0.	Brunel_HasRich && Brunel_HasCalo && Brunel_P>3000 && Brunel_PT>500 && nSPDHits<450 && Brunel_HasBremAdded==1



(a) Proton

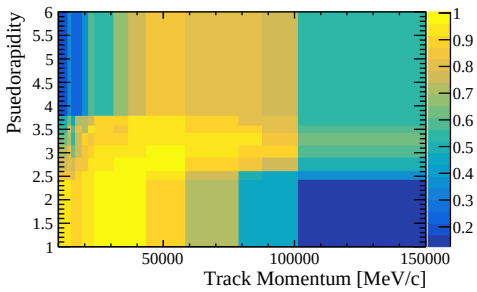
(b) Kaon

(c) Muon

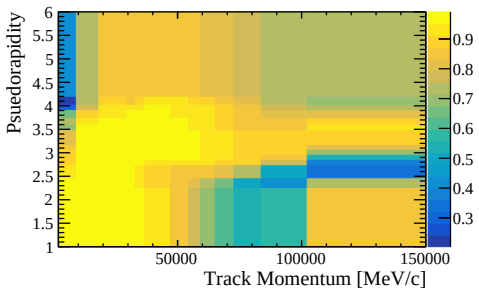
Figure 3.37: PID efficiency maps for 2011.

of low statistics, the weight was set to the closest physical value (either 0 or 1).

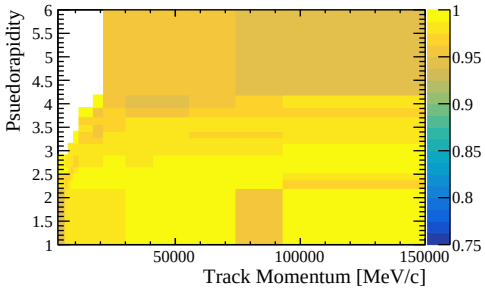
Consequently, the following strategy was adopted in this analysis. In the data, the complete PID selection was applied, as described in Tables 3.6-3.7. To measure the efficiency of htese PID requirements, signal simulation samples were used. In these simulation samples, PID cuts were not applied, but instead the PID weights were, by the procedure described above. Each per-event weight provides directly the PID efficiency for a given event. Therefore, reweighting the full simulation sample allows to evaluate the PID efficincy for the given decay mode. This procedure is a part of selection efficiency estimation described in Sec. 3.6.3.



(a) Proton

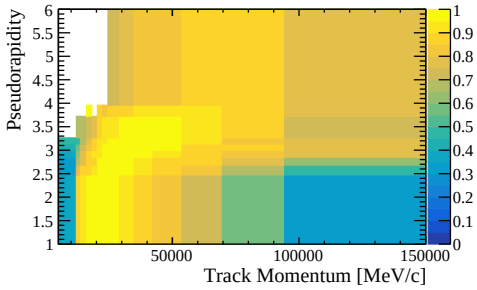


(b) Kaon

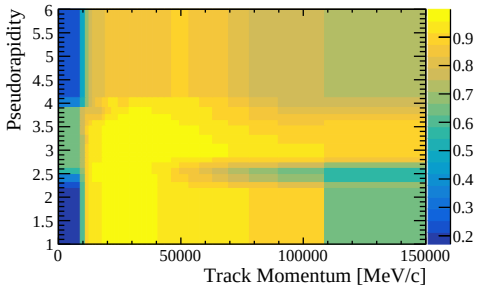


(c) Muon

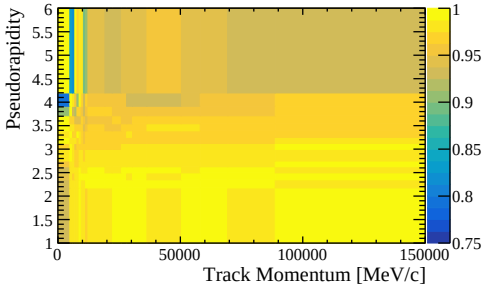
Figure 3.38: PID efficiency maps for 2012.



(a) Proton

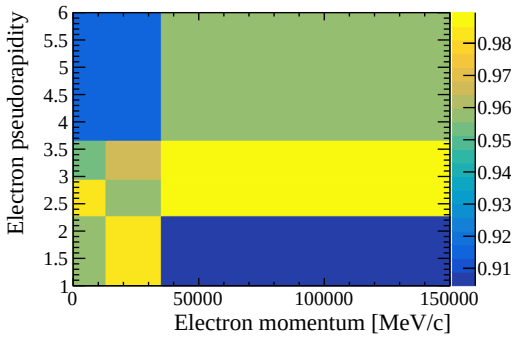


(b) Kaon

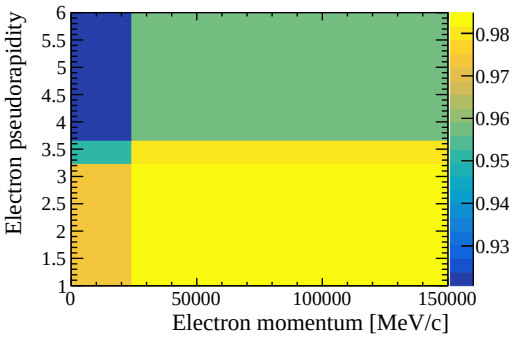


(c) Muon

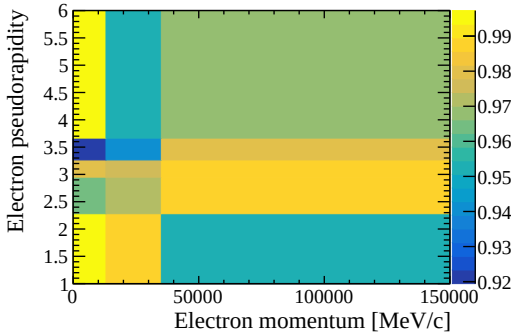
Figure 3.39: PID efficiency maps for 2016.



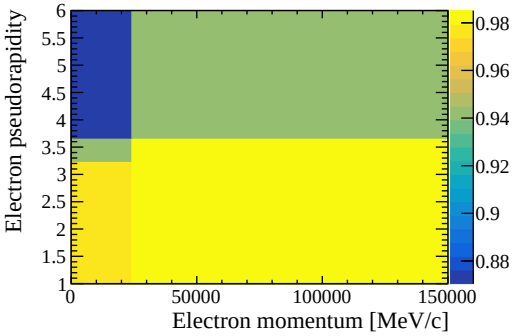
(a) 2011, no brem



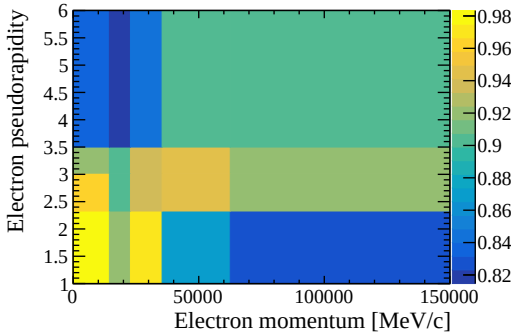
(b) 2011, has brem



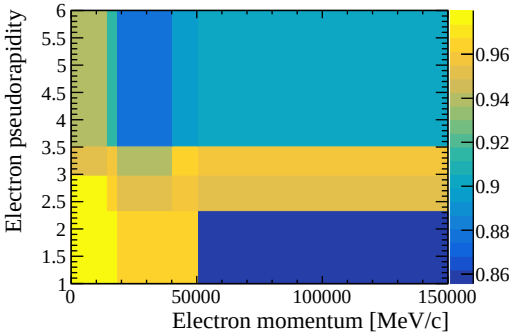
(c) 2012, no brem



(d) 2012, has brem



(e) 2016, no brem



(f) 2016, has brem

Figure 3.40: Electron PID efficiencies computed using the fit-and-count technique.

3.5.3 Event multiplicity

There are several variables which can be considered as a proxy to the event multiplicity distribution. The most common variables considered in various analyses are `nTracks` describing the number of tracks in the event, and `nSPDHits` describing the number of hits in the scintillating pad detector. Both of them are known to be mismodeled in the simulation up to different extents. The second one is important because the L0 trigger applies a cut on it; however it is known that its data-simulation discrepancy is much worse due to the improper model of the `M1` muon station, and in fact correcting for this leads to even worse discrepancy on other multiplicity proxies. Given that this analysis aims at measuring a double ratio, the efficiency of the SPD cut is expected to cancel, as it is further discussed in Sec. 3.8.2. In this analysis, the `nTracks` variable is used as a proxy for the event multiplicity, but several alternative choices, among them the `nSPDHits` variable, are used to estimate the systematic uncertainty related to the choice of the multiplicity proxy (see Sec. 3.8.4 and 3.10.2). The distribution of `nTracks` variable for 2016 data and simulation in shown in Fig. 3.41 for the $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$ data and simulation, and a disagreement can be seen between the two.

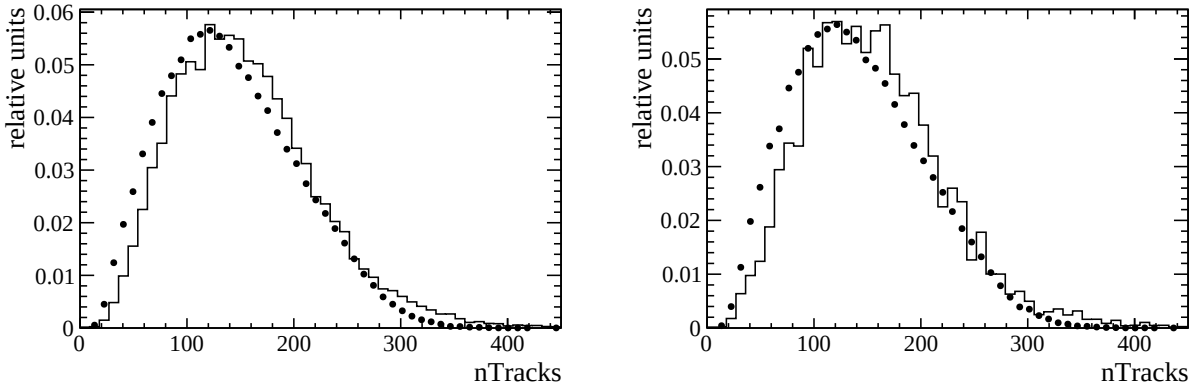


Figure 3.41: nTracks variable distribution for (left) $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ and (right) $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ channels in 2016, for an ${}_s\mathcal{P}lot$ of data (line) and simulation (dots)

Because no differences are expected between the muon and electron modes, the corrections to the simulation are computed using muon data and dedicated simulation samples with the largest statistical power.

To be able to compare data and simulation, several steps are required. First, a loose selection, including a loose BDT (see Sec. 3.4.3) requirement (BDT > -0.2) is applied to the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ data and simulation samples. Applying this preselection not only allows to obtain a clean Λ_b^0 invariant mass peak in data, but also ensures that the comparison between data and simulation is performed in the same kinematic region of interest. At the same time, the considered preselection is substantially looser than the selection used for the main part of the analysis, which not only allows to gain additional statistical power when computing corrections, but also to reduce correlations arising from reusing the same dataset when measuring signal yields. Given the fact that most of the

corrections rely on the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ decay mode, rare modes $\Lambda_b^0 \to pK\ell^+\ell^-$ are completely unaffected by such correlations.

Nominal corrections are computed, requiring both data and simulation candidates to pass the `L0Muon` trigger requirement, although no significant difference was observed if TIS events are kept in addition. The simulation is corrected for the decay model and PID response (PID weights are applied instead of the corresponding PID cuts), as described above. Then, a simple fit to the Λ_b^0 (with J/ψ mass constraint) invariant mass is performed on the data sample. Three fits corresponding to 2011, 2012 and 2016 datasets are shown in Fig. 3.42. The simplified fit model was used, including only the signal shape (described by a Bifurcated Crystal Ball function (defined in Eq. 2.7) and tuned on the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ simulation), and an exponential background shape of a free slope and yield. This fit is much simpler than the default fits used to extract signal yields in Sec. 3.7, as it runs over a sample with a loose selection (so, larger background), and separation of individual background components is difficult and not necessary for the scope of this task.

This signal shape was tuned using the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ simulation, with a selection applied which is identical to the one in the fitted dataset. The tail parameters were subsequently fixed in the fits to data, while the mean and the width were kept free.

After performing the fits, all the free parameters but the yields are fixed, and the ${}_s\mathcal{P}lot$ technique is used to subtract statistically the remaining background. This technique is valid, as multiplicity variables do not correlate with the invariant mass.

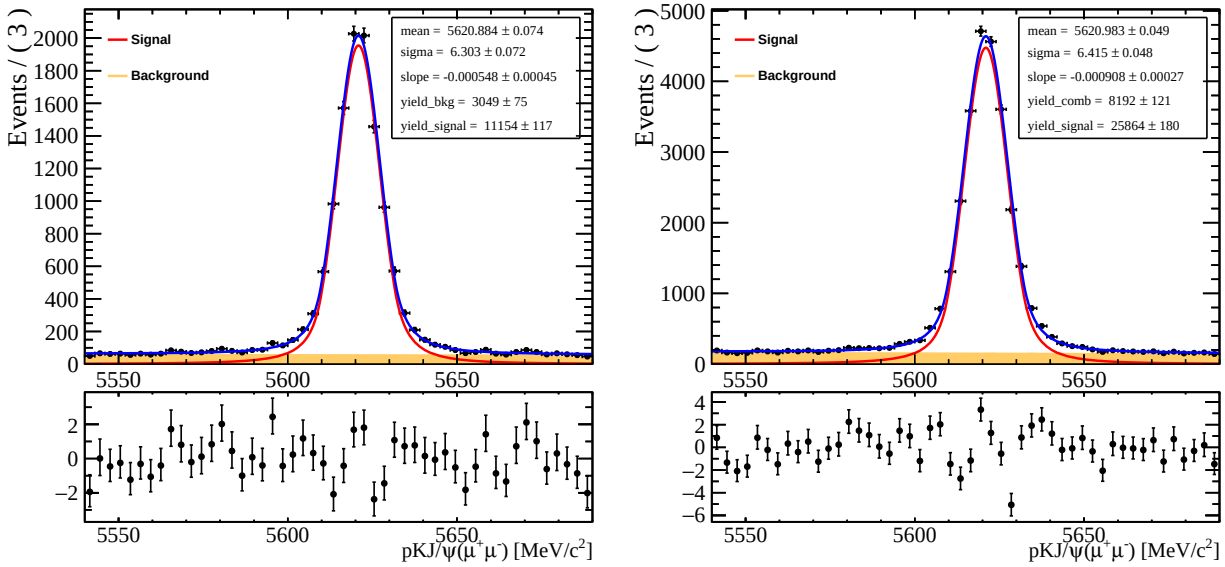
The corrections are computed using a `k-fold` approach, with $k = 10$, when the dataset (both data and simulation) is split into ten equal parts, and nine of them are used to compute the correction which is afterwards applied to the tenth one, and this process is repeated ten times. This means, the correction which has been computed using a given subsample, is never applied to this same subsample. Such approach allows to reduce potential correlations arising from reusing the same dataset twice.

For each dataset, an iso-populated binning in `nTracks` variable was created, with the number of bins dependent on the available statistics in the given dataset. The corrections obtained for 2011, 2012 and 2016 datasets, are shown in Fig. 3.43. Different colours represent ten folds. The weights for 2016 have a different trend from the Run I samples, this is partially due to a tighter cut on the number of SPD hits in the 2016 hardware trigger configuration, which reflects on the `nTracks` behaviour.

When applying them to the simulation, the weights are interpolated to the corresponding value of `nTracks`. Binned weights (as compared to interpolated ones) are used to assess the systematic uncertainties related to this procedure (Sec. 3.8.4).

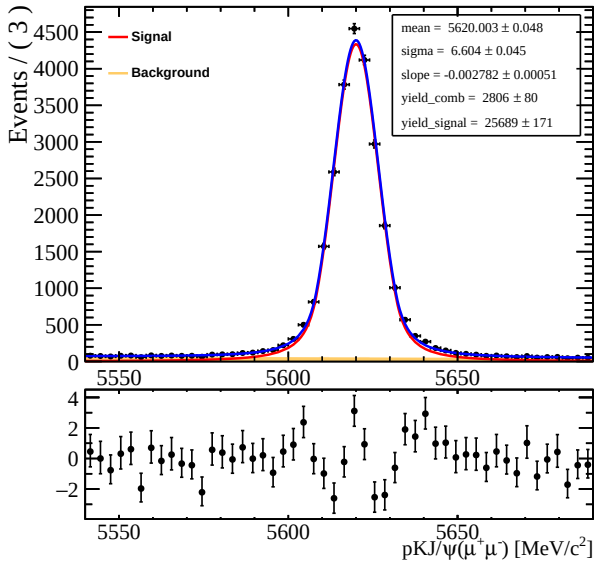
It can be seen that, in general, the relative difference between ten folds is comparable to the per-bin statistical uncertainty on the weights in each fold.

For the systematic uncertainty studies (Sec. 3.8.4), an alternative set of weights is computed, reweighting `nSPDHits`, `nPVs` or `nVeloTracks` variable as a multiplicity proxy, instead of `nTracks`. The case of absence of any multiplicity correction is also considered.



(a) 2011

(b) 2012



(c) 2016

Figure 3.42: Preliminary fits to the $pKJ/\psi\,(\mu^+\mu^-)$ invariant mass with J/ψ mass constraint applied, to different datasets; used for further background subtraction via ${}_s\mathcal{P}lot$ method in order to evaluate corrections to the simulation.

3.5.4 Generated event kinematics

After applying the decay model, PID and event multiplicity corrections to the simulation, the data-simulation agreement is being checked for the generated kinematics of the Λ_b^0. In particular, the simulation is known to mismodel the strong p_T dependence of the Λ_b^0

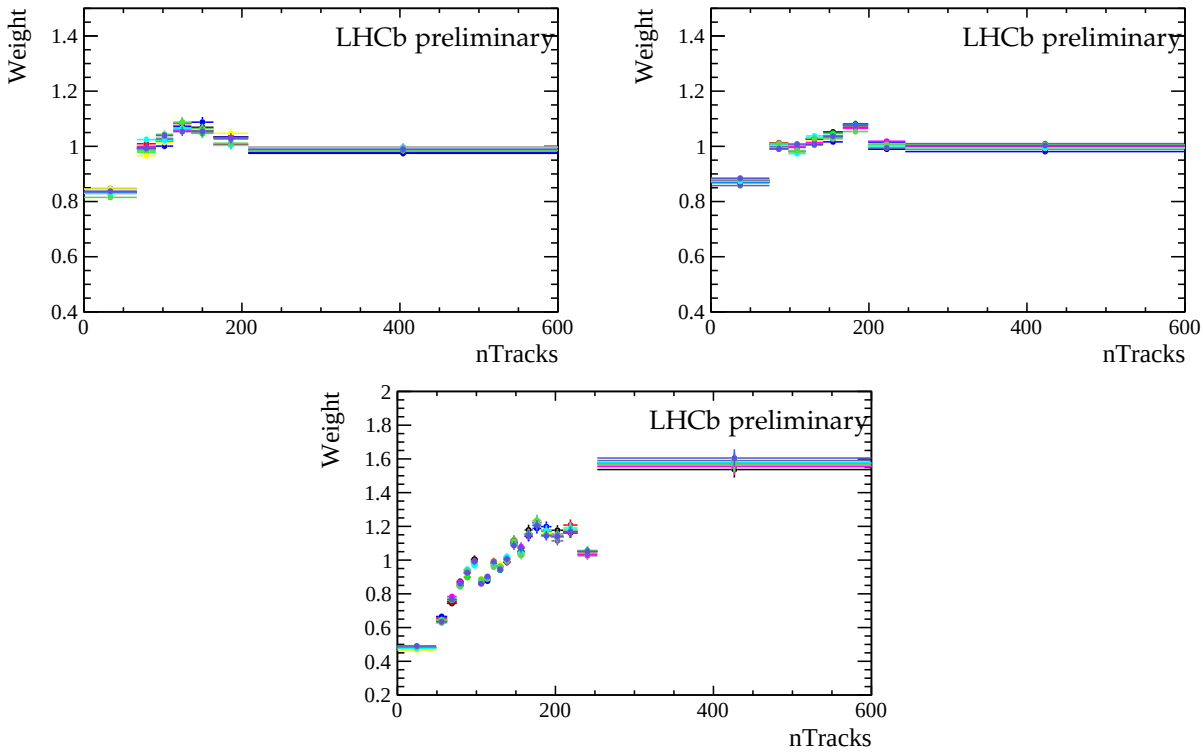


Figure 3.43: `nTracks` corrections for 2011, 2012 and 2016 years of data taking. Different colours represent 10 equivalent folds.

production in LHCb, as measured in Ref. [97]. This means that the corrections for the Λ_b^0 decays should not necessarily be similar to those computed for the decays of B mesons.

As a proxy, two variables are used: the transverse momentum p_T and the pseudorapidity η of the Λ_b^0. Since these two variables are correlated, the correction is computed in two dimensions simultaneously. As in the previous step, the simulation was compared to the ${}_s\mathcal{P}lot$ of the data in order to evaluate this correction. The ${}_s\mathcal{P}lot$ technique is still valid because in the $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$ decay mode the invariant mass does not correlate with the kinematic variables. The *k-fold* approach similar to the one discussed in the previous section was adopted when computing the weights. The number of iso-populated two-dimensional bins in which the correction was computed, depends on the size of the dataset: 100 bins for 2011, 120 for 2012, and 250 for 2016 data taking periods.

The resulting corrections for 2011, 2012 and 2016 samples are shown in Fig. 3.44. As it is impossible to visualise ten folds on the same two-dimensional plot, only one of ten folds is shown here.

Same as in the previous section, these corrections are computed using the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ dataset passing the `L0Muon` trigger requirement. However, a possible mismodeling of the trigger response in the simulation can be a potential source of bias for the kinematic corrections. To estimate this effect, two alternative sets of kinematic corrections are computed for systematic studies (Sec. 3.8.4): extracted either using the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ dataset passing the `L0Electron` trigger requirement, or using the

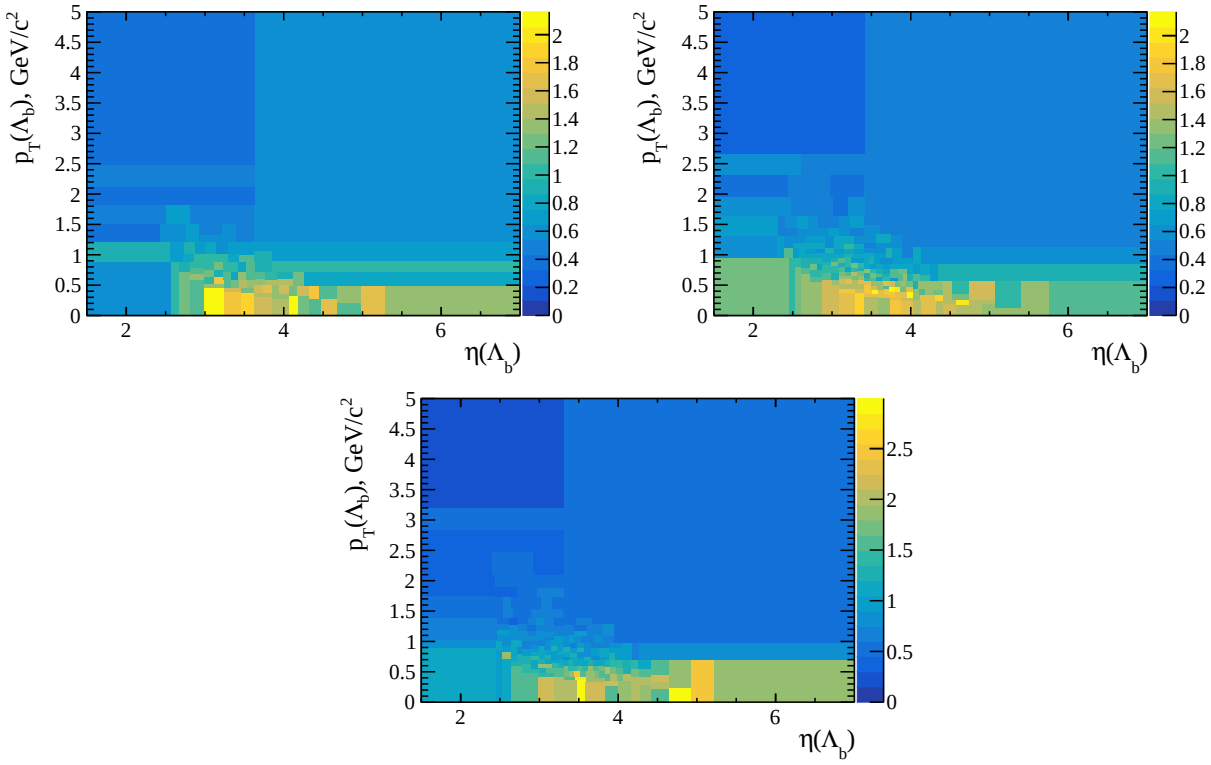


Figure 3.44: Kinematic weights for 2011, 2012 and 2016 years of data taking

$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ dataset passing the `L0TIS` requirement.

3.5.5 Decay time of Λ_b^0 and its acceptance

It is known that the value of the Λ_b^0 lifetime used in the simulation is slightly different from the most recent value reported in the PDG [6]. All the simulation samples used in this analysis were generated assuming $c\tau(\Lambda_b^0) = 0.4350$ mm, while the PDG reports $c\tau(\Lambda_b^0) = 0.4407$ mm. To correct for this discrepancy, per-event weights defined as:

$$e^{c\tau \times (1./0.4350 - 1./0.4407)} \tag{3.4}$$

were applied to all the generator- and reconstructed-level simulation samples, for all data taking periods and decay modes. In this correction, the generator-level values of the decay time were used, as the correction accounts only for the lifetime mismatch at the generation step.

However, the data-simulation comparison performed on Run I samples after applying the complete selection and reweighting procedure, still shows a discrepancy in the decay time distribution, mostly in the region of small decay time values, where a proper modeling of the decay time acceptance is essential. This is believed to happen due to the fact that some of the reconstruction and selection steps affecting the decay time acceptance (for instance, related to the tracking) are not ideally described by the simulation. To

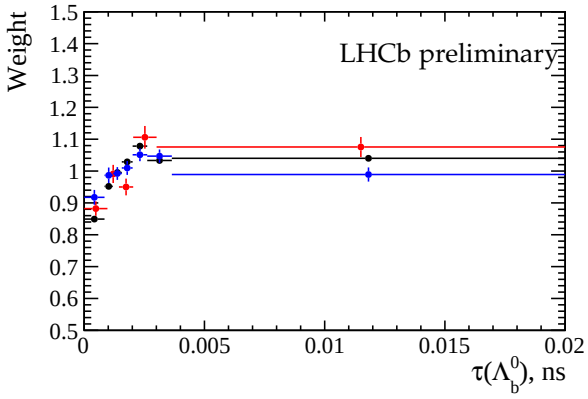
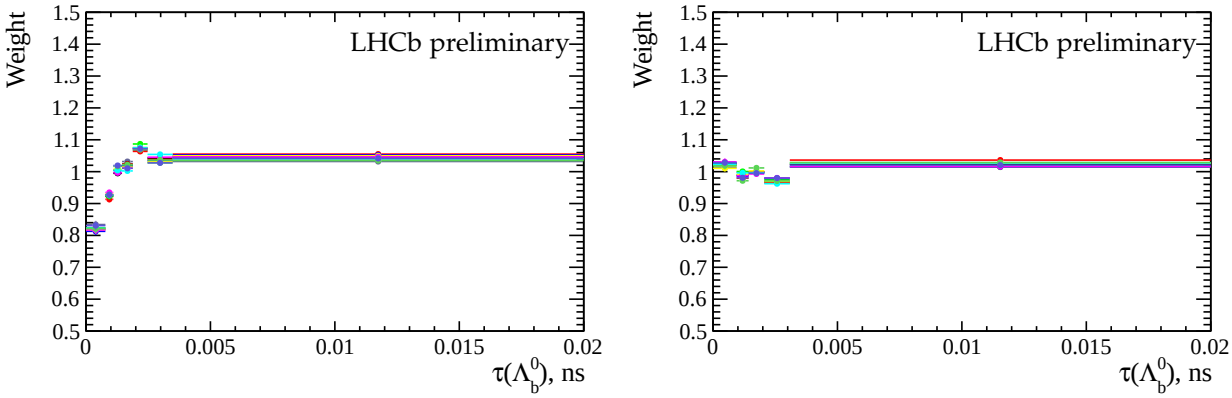


Figure 3.45: Λ_b^0 decay time weights for Run I, obtained from (black) L0M, (red) L0E and (blue) L0I trigger categories. For visual purposes, only one of ten folds is shown.



(a) Run I (b) Run II

Figure 3.46: a) Λ_b^0 decay time weights taken as default for Run I, and b) computed in the same way for Run II. The latter are compatible with unity. Different colours represent 10 equivalent folds.

correct for this effect, weights are computed in bins of the decay time, and applied to all the simulation samples. It is shown in Fig. 3.45 that the weights look consistent between various trigger categories, so the ones extracted from the most abundant category (L0M) are applied throughout. The behaviour of this correction is similar regardless of the fact whether an ${}_s\mathcal{P}lot$ or the tight mass cut is used to select the signal on the data.

The weights applied to the Run I simulation samples are shown in Fig. 3.46(a). Fig. 3.46(b) shows the absence of a significant trend or deviation from unity for the Run II weights, so for Run II the weights are not applied. This can probably be explained by improvements in the reconstruction in Run II.

The impact of (not) applying this correction on the single and double ratios is negligible (as all channels used are the decay modes of Λ_b^0), and is considered as a minor source of systematic uncertainty (Sec. 3.8.4).

3.5.6 Trigger

After correcting for the PID response, event multiplicity, generated Λ_b^0 kinematics, generated Λ_b^0 lifetime and the decay time acceptance, the trigger response is being corrected for. This is performed using the *TISTOS* method [155] described below, in two subsequent steps: L0 and HLT response. Since the L0 and HLT requirements are different between the electron and muon decay modes, these corrections are evaluated separately.

The TISTOS method is based on the usage of a sample triggered independently of the signal (TIS), as opposed to the sample triggered on signal (TOS). As in data there is no access to all events which have not passed the trigger, it is impossible to estimate the trigger efficiency directly. However, a well-tested assumption lying under the TISTOS method is that in a sample triggered independently of signal, there is a subsample which is at the same time triggered by signal ($TIS\&TOS$), and the fraction of this subsample is a good estimate of the trigger efficiency. To evaluate the corrections to the simulation, the following quantity is computed:

$$\epsilon_{TOS} = \frac{N_{TIS\&\&TOS}}{N_{TIS}}, \tag{3.5}$$

where N_{TIS} is the number of events (in the selected sample) triggered independently of the trigger line of interest (or, in the "TIS subsample"), and $N_{TIS\&\&TOS}$ is the number of events in the TIS subsample which were at the same time also triggered by the trigger line of interest (TOS). Exact definitions of the TOS and TIS subsamples differ depending on the studied trigger category, as discussed below.

The quantity ϵ_{TOS} is computed on both data and simulation, with the same selection applied to both, and all the previous corrections applied to the simulation. Given the possible dependence of ϵ_{TOS} on various kinematic variables, especially those on which a cut is applied by the trigger decision, the trend of ϵ_{TOS} in bins of these variables is studied. Finally, the correction to the simulation is defined as $\epsilon_{TOS}^{data}/\epsilon_{TOS}^{sim}$.

Similarly to the previous cases, a *k-fold* approach is adopted when computing the trigger corrections.

In the data, instead of an ${}_s\mathcal{P}lot$ technique, the cut around the J/ψ-constrained Λ_b^0 invariant mass is applied. This is to avoid issues related to the ${}_s\mathcal{P}lot$ performance for the electron modes (it is not strictly valid, as bremsstrahlung introduces a correlation between the invariant mass and some of the kinematic variables), but for consistency the same technique is also used for the muon modes. To clean up the remaining background, a loose requirement on the BDT output is applied. It should be noted here that the trigger correction is the only one where the electron dataset is used to extract default correction weights. In all other cases, we rely on the muon dataset which has no bremsstrahlung, and so the ${}_s\mathcal{P}lot$ technique is valid there. This explains why a slightly different approach was chosen here to extract a sample of data, with respect to what has been done for other correction steps.

3.5.6.a L0

In the scope of this section, all the trigger categories are considered as inclusive. This is because the TISTOS method is not directly applicable to the exclusive categories which have only TOS events but not TIS events.

For the `L0Muon` and `L0Electron` trigger categories, the TIS sample is defined as `Lb_L0GlobalDecision_TIS`. For the L0I category, it is respectively `Lb_L0GlobalDecision_TOS`. Some alternative choices are possible, for example, one could use events triggered by a hadron as a TIS sample for L0M and `L0Electron` categories, or events triggered by lepton for the L0I category. These alternatives are used to estimate the related systematic uncertainty, as discussed in Sec. 3.8.4.

All the corrections are binned in the variable on which the cut in the trigger is applied.

- `L0Muon`: the binning variable is the transverse momentum of a muon. A one-dimensional iso-populated binning is adopted, however at low values of p_T, several extra bins are added by hand to ensure the proper description of a region close to the trigger threshold. Corrections are computed per one muon (each of two muons is used in this procedure to increase the statistics), and applied to each of two. Maps of corrections for different years of data taking are shown in Fig. 3.47.

  In 2016, the tightest trigger configuration, having a threshold of about 36 ADC counts, where one ADC count corresponds to approximately 50 MeV/c, was emulated in the simulation. However, it was used in only about 35% of the dataset, while a looser trigger configuration with a threshold of 26 ADC counts was used in about 45% of the dataset. This explains the large correction at the low momentum, in the region below approximately $36 \times 50 = 1800$ MeV/c.

  The resulting reweighted per-event L0 decision is computed as

$$w_{L0} = \texttt{L0Decision}_{\mu 1} \times w_{\mu 1} + \texttt{L0Decision}_{\mu 2} \times w_{\mu 2} - \texttt{L0Decision}_{\mu 1} \times w_{\mu 1} \times \texttt{L0Decision}_{\mu 2} \times w_{\mu 2} \quad (3.6)$$

- `L0Electron`: The approach is similar to `L0Muon`, but the binning variable is the transverse energy of electron in the `ECAL`. In addition, the correction is evaluated separately for three regions of the ECAL (outer, middle, inner) which were described in Sec. 2.5. The resulting corrections are presented in Fig. 3.48. Variations in the behaviour of the weights between different years of data taking can be explained by the different L0 thresholds on the electron E_T in the ECAL. The resulting reweighted per-event L0 decision is computed in a similar way as for the `L0Muon`.

- `L0TIS`: For the `L0TIS` trigger response, the binning variable is the transverse momentum of the Λ_b^0. This implies that the correction is by construction applied per event, but not per track, like it was done in the two previous cases. This correction is evaluated in three regions of `nTracks` variable: `0-150-210-600` for Run I, and `0-110-180-600` for Run II, the difference between the two is due to a tighter

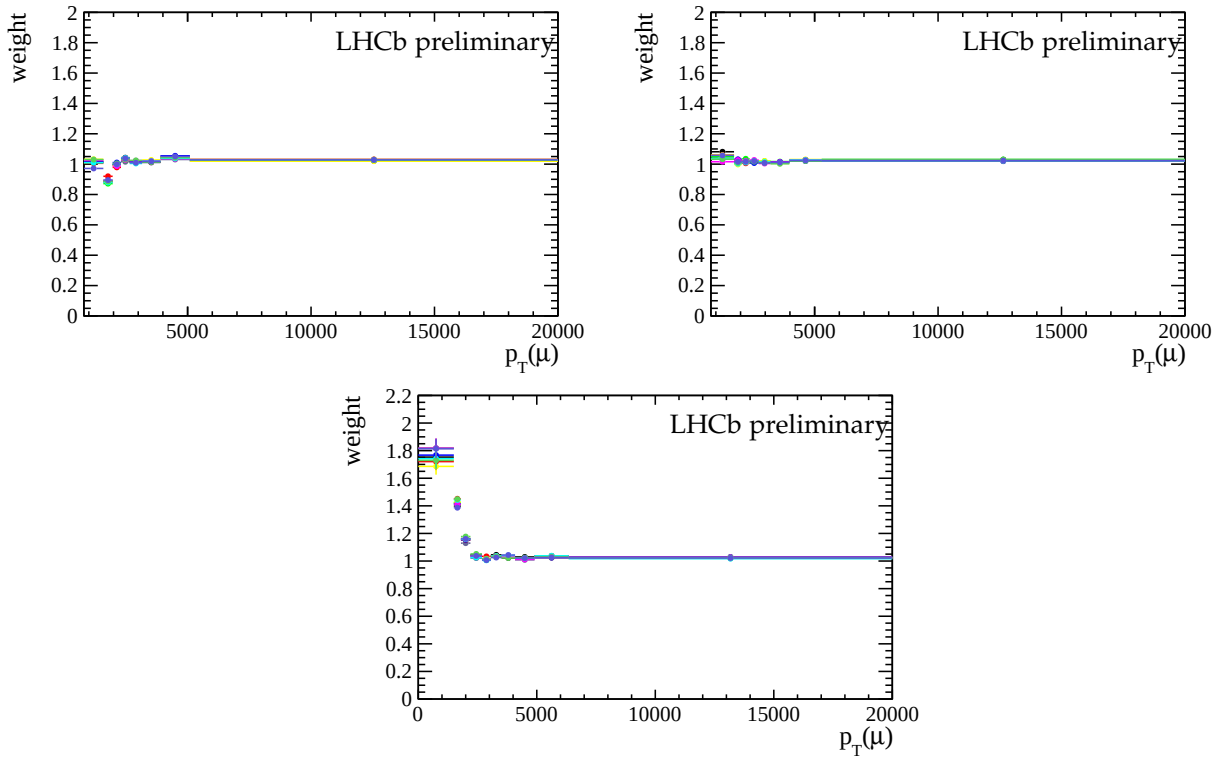


Figure 3.47: `L0Muon` weights for 2011, 2012 and 2016 datasets. Different colours represent 10 equivalent folds.

`nSPDHits` cut in Run II, which effectively removes events with larger `nTracks` values. Regions of different event multiplicity are considered, as the chance to trigger an event independently of signal changes with multiplicity. The resulting corrections are shown in Fig. 3.49.

It should be noted that the `L0TIS` trigger correction is computed using the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data and simulation samples, but is also applied to the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$. This is done because no difference is expected for this trigger correction between the two modes, while the muon mode provides a larger statistical power. This approach is validated in Fig 3.50, where four alternative methods are compared in order to extract the L0I trigger correction, and a reasonable agreement is observed, within uncertainties. The correction is extracted either from muon, or from electron dataset, and alternative definitions of the TIS sample (compared to the default `Lb_L0GlobalDecision_TOS`) are used. These alternative corrections are used to estimate systematic uncertainty related to the `L0Global (TIS)` trigger correction, by applying these alternative weights when computing signal efficiencies. To gain additional statistical power in the electron mode, no binning in `nTracks` is used for this test.

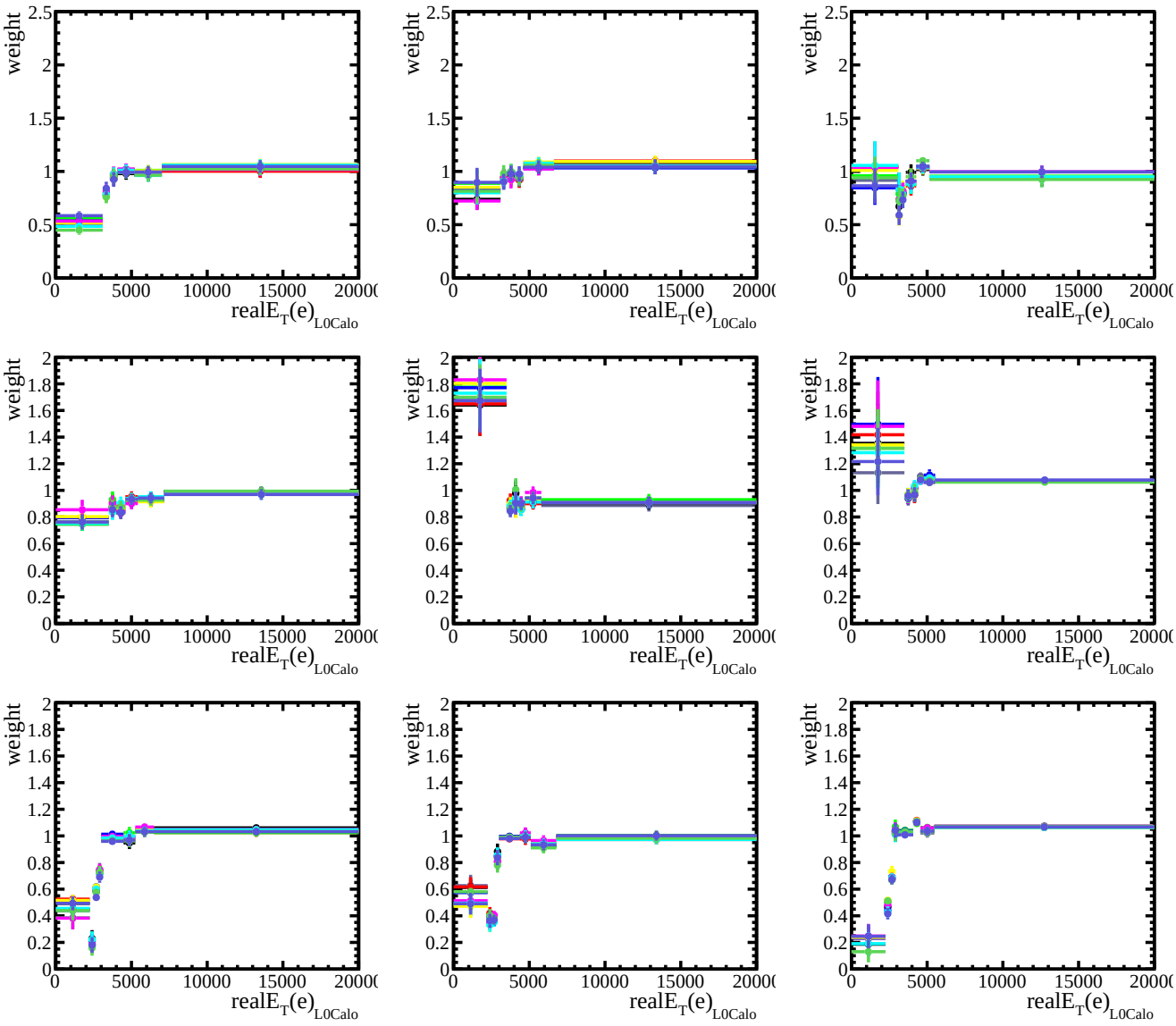


Figure 3.48: `L0Electron` weights for (from top to bottom) 2011, 2012 and 2016 datasets. From left to right, in three regions of ECAL – outer, middle, inner. Different colours represent 10 equivalent folds.

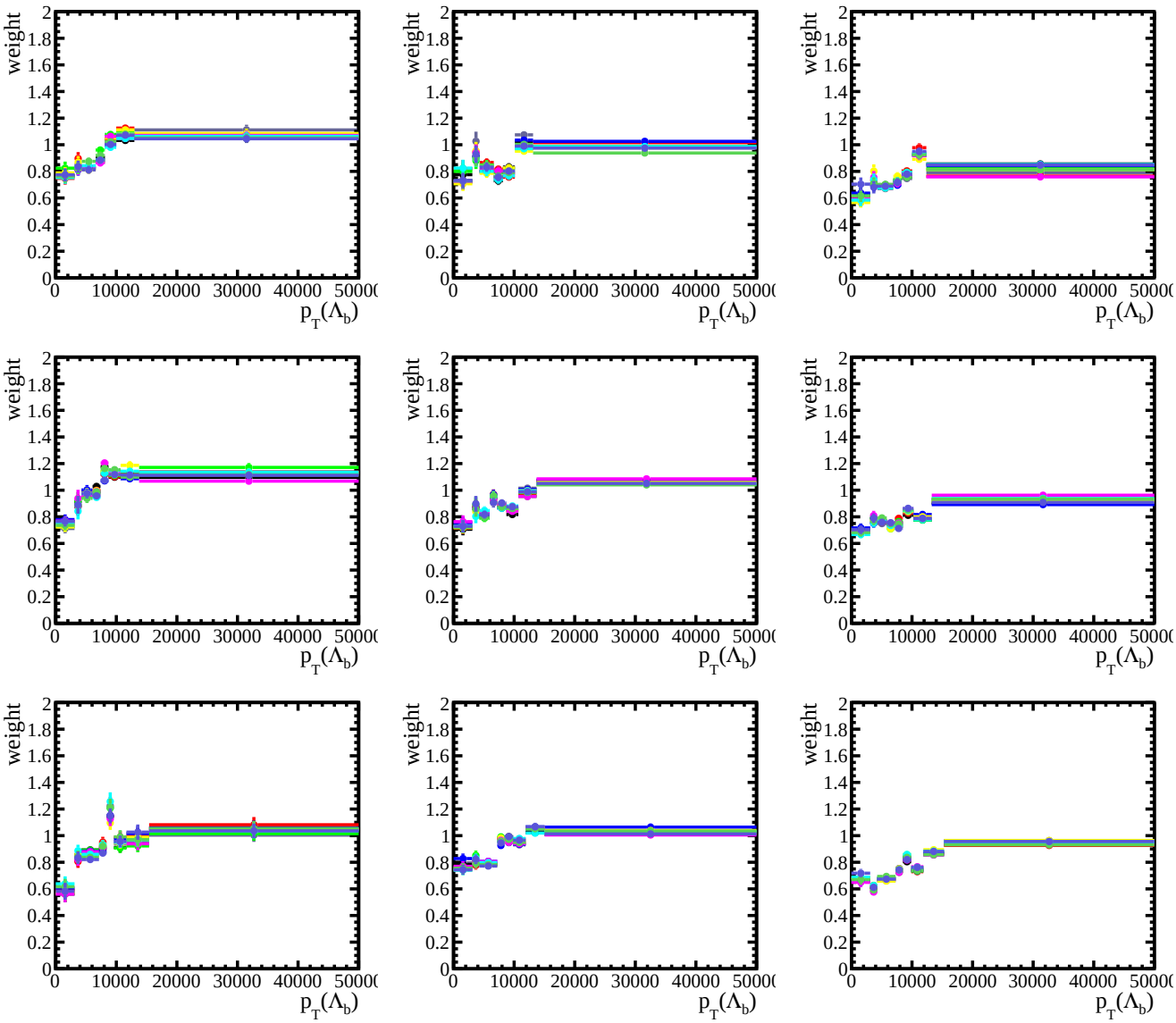
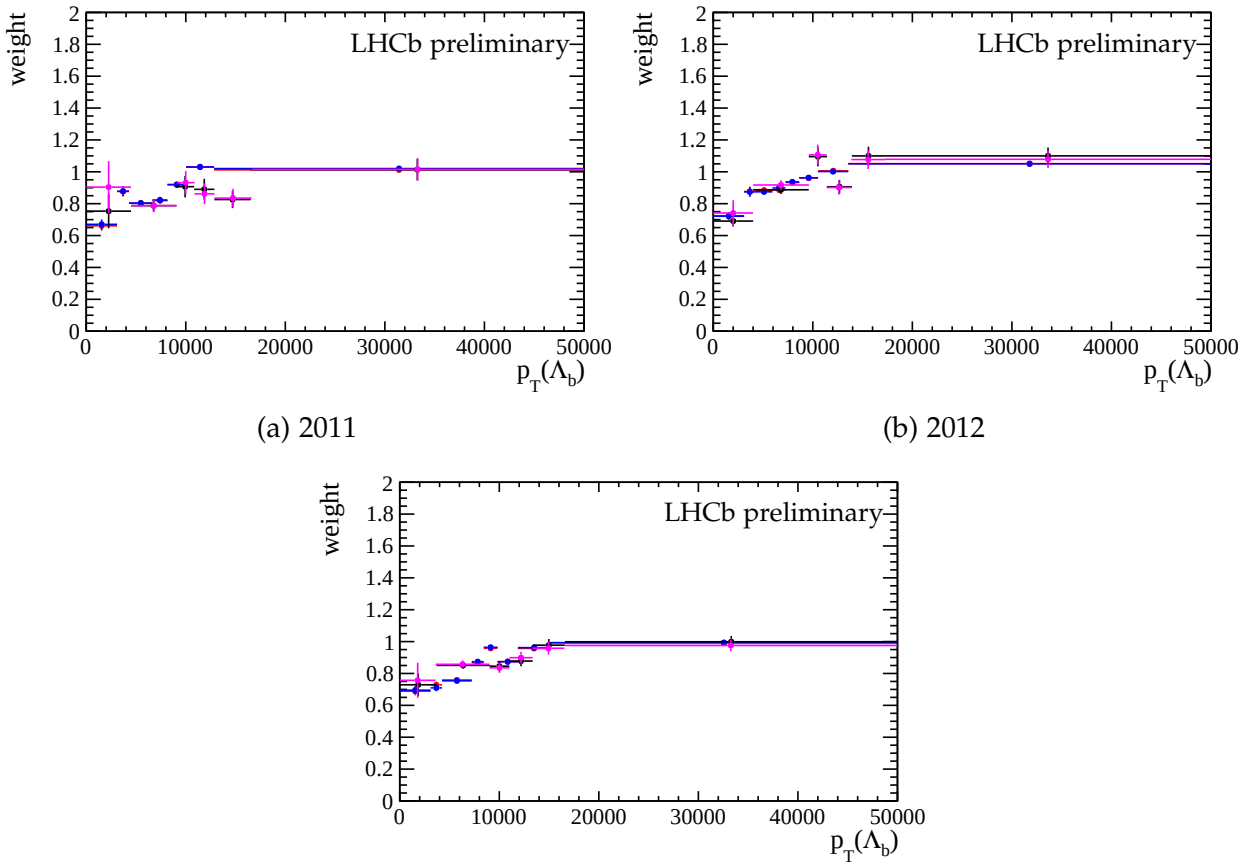


Figure 3.49: `L0Global (TIS)` weights for (from top to bottom) 2011, 2012 and 2016 datasets. From left to right, in three regions of `nTracks` values, as described in the text. Different colours represent 10 equivalent folds.



(a) 2011

(b) 2012

(c) 2016

Figure 3.50: Comparison of the L0I trigger correction computed via (blue) default procedure; (red) muon sample using `Lb_L0MuonDecision_TOS` as a TIS requirement; (magenta) electron sample using `Lb_L0ElectronDecision_TOS` as a TIS requirement; and (black) electron sample using `Lb_L0Global_TOS` as a TIS requirement.

3.5.6.b HLT

Differences between simulation and data in the HLT modeling have been found to be very small. No trend in kinematic variables have been observed. The integrated corrections are well compatible with one, and have large uncertainties. The central values, integrated over kinematics, are presented in the Table 3.11.

Table 3.11: Integrated HLT weights.

Run period	L0M	L0E	L0I
Run I	1.027 ± 0.012	0.971 ± 0.030	1.027 ± 0.038
Run II	0.996 ± 0.008	0.980 ± 0.015	0.998 ± 0.010

Given that, the HLT corrections are not applied to the simulation samples. Instead, a systematic uncertainty is assigned to account for possible mismodeling, taking the largest statistical uncertainty on the correction value (Sec. 3.8.4).

3.5.7 Residual differences

After applying the corrections described above, no other corrections are applied, however a check is performed for the data-simulation agreement in various variables. Most importantly, the data-simulation agreement is studied for the BDT output (above -0.2), and is found to be very good, which proves the absence of a need for residual corrections, as shown in Figures 3.51 and 3.52.

On top of that, the data-simulation agreement is studied in other relevant variables, and the corresponding plots for $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ modes are shown in Appendix C.

To assign a systematic uncertainty related to the agreement between the data and the simulation, some of the variables showing visible discrepancies are reweighted to check the stability of all the results, as it is discussed in Sec. 3.8.4.

3.5.8 Simulation samples of the specific backgrounds

When performing a fit to data (Section 3.7), it is important to make sure that the shapes of the specific backgrounds (described in Sec. 3.3) are described properly. Therefore, the corrections which can potentially impact the invariant mass shapes, are applied to the relevant simulation samples. Those include the reweighting of the PID response, and when applicable correcting for the wrong decay model. The PID response in the background samples is corrected using the `PIDCalib` maps computed assuming a certain misID hypothesis.

One of the most important background modes, the $B_s^0 \to K^+K^-J/\psi\,(\to \ell^+\ell^-)$, has a decay model with several resonances contributing to the K^+K^- spectrum [149]: ϕ (vetoed in this analysis), $f_2'(1525)$ and others. In order to correct the simulation which assumes a phase-space decay model, for the resonant model, an ${}_s\mathcal{P}lot$ is performed to the

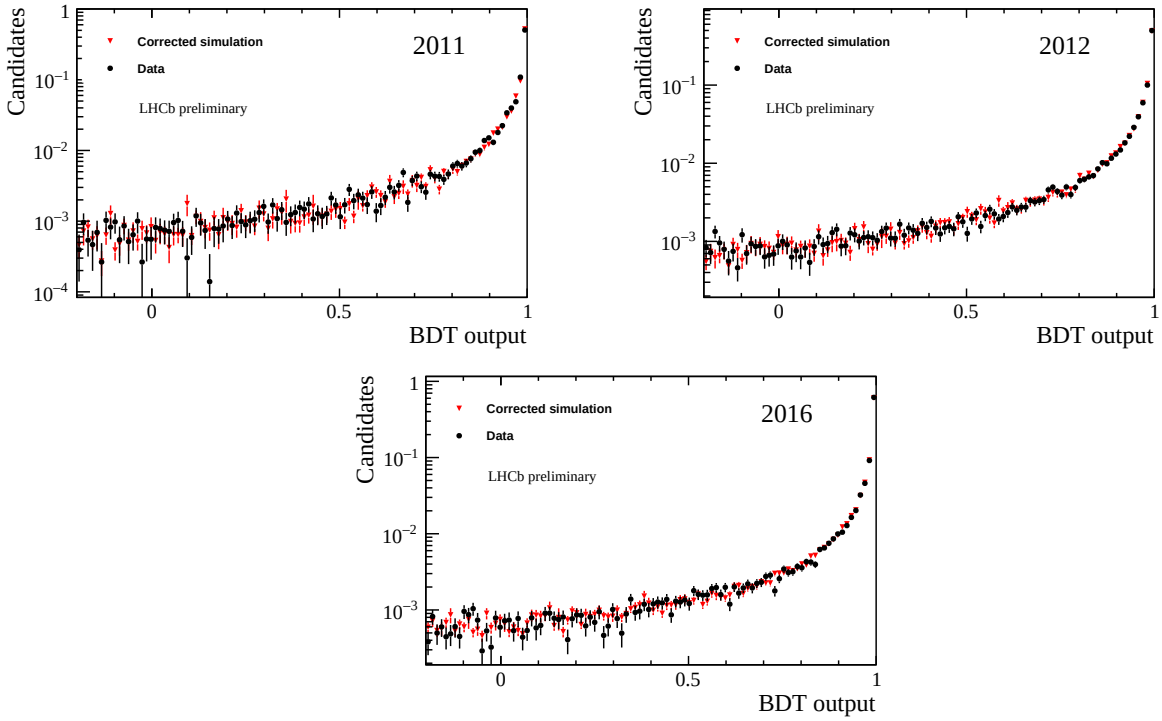


Figure 3.51: Comparison of the BDT output variable, between the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data (black circles) and simulation (red triangles), for different years of data taking.

$B_s^0 \to K^+K^-J/\psi(\to \mu^+\mu^-)$ data, extracting the proper spectrum (which is uncorrelated with the B_s^0 invariant mass) and reweighting the simulation according to it. The result of this correction can be seen in Fig. 3.53.

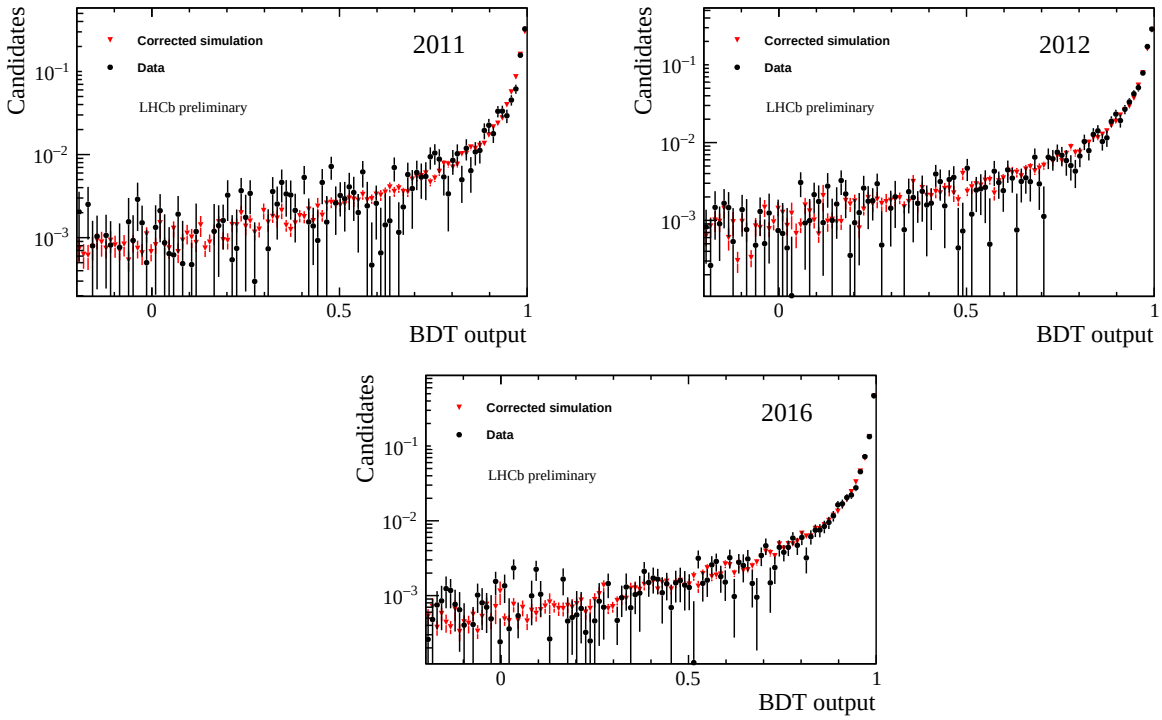


Figure 3.52: Comparison of the BDT output variable, between the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ data (black circles) and simulation (red triangles), for different years of data taking.

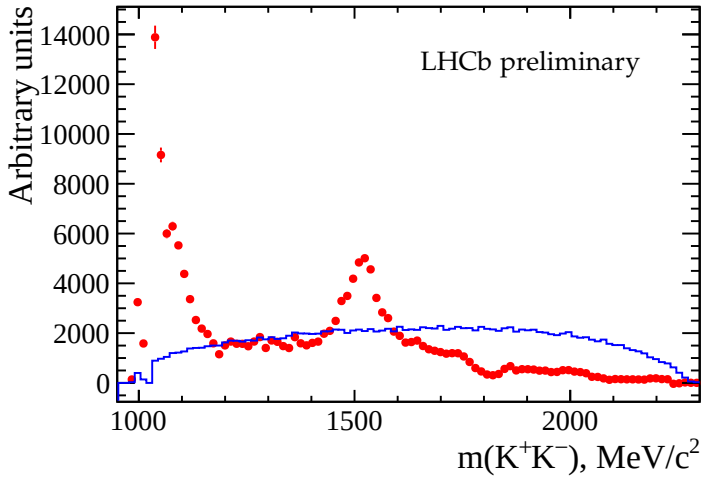


Figure 3.53: The $m(K^+K^-)$ spectrum in the simulation of the $B_s^0 \to K^+K^-J/\psi(\to \mu^+\mu^-)$ decay: (blue) – phase-space model; (red) – reweighted model. The $\phi(1020)$ resonance is vetoed.

3.6 Extraction of signal efficiencies

Signal efficiency describes the proportion of signal candidates which fall in the acceptance of the LHCb detector, and pass all the reconstruction and selection steps. The total selection efficiency is defined as the product of following efficiencies:

$$\epsilon_{tot} = \epsilon_{geom} \times \epsilon_{filt} \times \epsilon_{reco} \times \epsilon_{sel}, \tag{3.7}$$

where the first term represents the geometric efficiency, the second one is the filtering one, third – the reconstruction and the last one – the offline selection efficiency. Each of these terms is defined and described in the following sections. Throughout this section, each mentioned efficiency is computed on top of the previous ones.

3.6.1 Geometric efficiencies

The geometric efficiency ϵ_{geom} accounts for the geometric acceptance of the detector. It is taken as a single number for each data taking period and magnet polarity, from the official tables provided by the LHCb simulation group. A summary of the efficiencies used for Run I and II, is provided in Table 3.12. The efficiencies for two magnet polarities are averaged.

Table 3.12: Geometric efficiencies, %.

Channel	2011	2012	2016
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	16.23 ± 0.02	16.52 ± 0.02	17.41 ± 0.02
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	16.10 ± 0.02	16.48 ± 0.02	17.37 ± 0.02
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	16.52 ± 0.02	16.84 ± 0.02	17.74 ± 0.02
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	16.43 ± 0.02	16.78 ± 0.02	17.67 ± 0.02
$\Lambda_b^0 \to pK\mu^+\mu^-$	16.01 ± 0.03	16.33 ± 0.02	17.30 ± 0.03
$\Lambda_b^0 \to pKe^+e^-$	15.75 ± 0.03	16.08 ± 0.03	17.04 ± 0.03

3.6.2 Filtering efficiencies

The filtering efficiency ϵ_{filt} accounts for the fact that (most of) simulation samples were produced using the filtering script, which is applying the selection of the stripping line to the simulation sample. The idea behind using the filtering script is to save the disk space used for storing the simulated candidates. The PID requirements, which are present in the stripping, are removed from the filtering, so that their efficiency can be estimated more precisely using data-driven techniques (see Sec. 3.5.2). The efficiency of these PID requirements is quoted as a part of the total PID efficiency, so enters the selection efficiency described in the following section. Therefore, the filtering efficiency is not exactly equal to the stripping efficiency. The filtering efficiency is provided by the LHCb simulation

group. The summary of the efficiencies used for Run I and II, is provided in the Table 3.13. Efficiencies for the two magnet polarities are averaged. If an unfiltered production was used for a certain sample, the filtering efficiency is equal to 100%.

Table 3.13: Filtering efficiencies, %

Channel	2011	2012	2016
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	100	51.29	49.07
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	38.39	39.88	38.71
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	100	52.76	50.43
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	38.14	39.47	39.34
$\Lambda_b^0 \to pK\mu^+\mu^-$	100	43.74	46.81
$\Lambda_b^0 \to pKe^+e^-$	30.70	32.65	39.31

3.6.3 Reconstruction and offline selection efficiencies

The reconstruction efficiency is defined on simulation samples as a ratio between the number of reconstructed events (so, passing all the reconstruction and stripping algorithms, and truth-matched as a signal), and the number of generated events. It therefore includes by construction the stripping and the truth-matching efficiencies. The offline selection efficiency accounts for all the cuts applied in the offline selection, as well as the PID cuts which were 'removed' from the filtering/stripping. In practice, it is difficult to separate the two efficiencies, accounting properly for the multiple candidates, not fully truth-matched cases such as decays in flight, and corrections to the simulation. Therefore, in this analysis we combine these two efficiencies. All the corrections to the simulation are applied in this case, including the PID and trigger weights. The efficiency is taken as a ratio of the weighted number of events passing the selection requirements in the reconstructed-level sample, and the weighted (only for the decay model, event multiplicity, Λ_b^0 kinematics and the generated Λ_b^0 lifetime value) number of events in the generator-level sample:

$$\epsilon_{reco} \times \epsilon_{sel} = \frac{N_{reco-level}^{\text{reweighted, fully selected, truth-matched}}}{N_{gen-level}^{\text{reweighted}}} \tag{3.8}$$

The definition provided above holds for the normalisation and control modes, but not for the rare $\Lambda_b^0 \to pK\ell^+\ell^-$ decays. In this analysis, all the principal measurements – R_{pK}^{-1} and branching ratios of the decays $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$, are performed only in the region $0.1 < q^2 < 6$ GeV$^2/c^4$ and $m(pK) < 2600$ MeV$/c^2$. The efficiency of these cuts is not accounted for, and these cuts are also applied at the generator level. At the generator level, so-called true-q^2 is used, to account for the final state radiation effects for leptons (Sec. 3.1.2). The efficiency of the cut $m(pK) < 2600$ MeV$/c^2$ cannot be accounted for properly, as the pK spectrum is not described by the simulation (which has a phase

space model), while the cut in the data at the Stripping level does not allow to access this spectrum above the cut value. It should be noted that this problem does not affect the resonant modes, as their available phase-space stops before 2600 MeV/c^2.

Similarly, the efficiency of the q^2 cut cannot be accounted for properly, as the q^2 model in the simulation is very generic and does not exploit the correct form-factors, which are currently unknown for the $\Lambda_b^0 \to pK\ell^+\ell^-$ mode.

Therefore the precise definition of the efficiencies for the $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ decay modes is the following:

$$\epsilon_{reco} \times \epsilon_{sel} = \frac{N_{reco-level}^{\text{reweighted, fully selected, truth-matched}}[0.1 < q^2 < 6; m(pK) < 2600]}{N_{gen-level}^{\text{reweighted}}[0.1 < q^2_{\text{true}} < 6; m(pK)_{true} < 2600]} \tag{3.9}$$

For completeness, the reconstruction, and the step-by-step selection efficiencies are provided in Tables 3.14, 3.16, 3.15 and 3.17 for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$, $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ (in L0E! and L0I categories) and $\Lambda_b^0 \to pK\mu^+\mu^-$ decay modes, respectively. Each mentioned efficiency is computed on top of the previous ones, and all them are computed using a simulation sample which has already passed geometric and filtering selection requirements. In the computation of these step-by-step efficiencies, only fully matched decay chains are taken into account in the simulation. The $\Lambda_b^0 \to pKe^+e^-$ efficiencies are totally blinded and are not shown. The multiple candidates are not removed in these step-by-step computations, as their removal makes sense only after applying the complete selection.

One of the main challenges of this analysis is to accurately control all the efficiencies so no artificial asymmetries between muons and electrons are introduced. This is achieved by correcting the simulation in any aspect where it does not match the data distributions for the control modes, as explained in Sec. 3.5. To assess the effect of these corrections, the raw efficiencies without corrections have been checked and dominant differences were understood: as expected, the largest change after reweighting occurs in trigger and PID efficiencies, as well as efficiencies of kinematic cuts.

3.6.4 Total efficiencies

The values actually used in the computations are the total efficiencies, taken as a product of the generation, filtering, and, as one step, reconstruction & selection efficiencies. These final values[4] are presented in the Table 3.18.

[4]The efficiencies for the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode have been unblinded when unblinding the $r_{J/\psi}^{-1}$, and were blind before.

Table 3.14: Step-by-step reconstruction and selection efficiencies (in %) for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$.

Step	2011	2012	2016
reco	14.37±0.09	22.86±0.10	25.74±0.09
L0	84.73±0.25	82.81±0.18	81.51±0.13
HLT1	76.24±0.31	79.53±0.21	86.95±0.18
HLT2	88.01±0.27	91.33±0.17	94.08±0.14
q^2	98.21±0.11	98.02±0.08	97.94±0.06
Kinematics	62.2±0.4	65.30±0.29	47.89±0.24
PID	66.8±0.6	70.6±0.4	69.0±0.4
BDT	87.36±0.35	85.53±0.25	93.38±0.16
Mass vetoes	88.8±0.4	88.26±0.28	87.88±0.27
Clones	100.0± 0	99.9970±0.0020	99.9970±0.0020
Fit range	99.931±0.022	99.912±0.025	99.877±0.029
Total sel	17.96±0.20	20.51±0.16	17.69±0.12

Table 3.15: Step-by-step reconstruction and selection efficiencies (in %) for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ L0I.

Step	2011	2012	2016
reco	18.12±0.07	16.16±0.06	18.60±0.06
L0	27.13±0.17	24.82±0.17	27.66±0.13
HLT1	64.39±0.34	69.1±0.4	75.34±0.25
HLT2	70.4±0.4	70.6±0.4	92.73±0.17
q^2	93.05±0.23	92.52±0.27	91.08±0.17
Kinematics	75.2±0.4	74.1±0.5	56.21±0.32
PID	69.9±0.6	72.6±0.7	66.2±0.4
BDT	81.8±0.4	79.7±0.5	90.61±0.23
Mass vetoes	88.4±0.5	87.9±0.5	87.66±0.33
Clones	99.986±0.010	100.0± 0	99.992±0.004
Fit range	99.928±0.027	99.916±0.024	99.897±0.020
Total sel	4.35±0.06	4.22±0.06	5.20±0.05

Table 3.16: Step-by-step reconstruction and selection efficiencies (in %) for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ L0E!.

Step	2011	2012	2016
reco	18.12±0.07	16.16±0.06	18.60±0.06
L0	19.71±0.14	14.59±0.13	24.94±0.13
HLT1	81.21±0.32	81.7±0.4	89.54±0.18
HLT2	85.71±0.33	89.7±0.4	96.59±0.11
q^2	94.00±0.21	93.79±0.26	92.88±0.14
Kinematics	72.0±0.4	69.9±0.5	54.63±0.30
PID	67.2±0.6	71.6±0.7	65.2±0.4
BDT	79.1±0.4	75.6±0.6	89.81±0.23
Mass vetoes	89.0±0.4	88.1±0.5	87.99±0.31
Clones	100.0± 0	99.992±0.006	99.991±0.006
Fit range	99.86±0.04	99.86±0.06	99.887±0.019
Total sel	4.39±0.06	3.34±0.05	5.63±0.05

Table 3.17: Step-by-step reconstruction and selection efficiencies (in %) for $\Lambda_b^0 \to pK\mu^+\mu^-$.

Step	2011	2012	2016
reco	14.46±0.15	26.25±0.13	27.27±0.15
L0	71.2±0.5	68.17±0.26	68.00±0.29
HLT1	76.9±0.6	80.15±0.29	86.85±0.33
HLT2	89.1±0.5	91.17±0.23	93.36±0.28
Kinematics	53.3±0.8	55.2±0.4	43.0±0.4
PID	69.2±1.2	73.1±0.6	70.6±0.7
BDT	87.5±0.7	87.70±0.34	92.4±0.4
Vetoes	89.9±0.8	89.0±0.4	89.8±0.5
Clones	100.0± 0	99.999±0.001	99.998±0.001
Fit range	99.31±0.15	99.15±0.12	99.19±0.10
Total sel	14.05±0.33	15.54±0.17	13.76±0.19

Table 3.18: Total efficiencies (including geometric, filtering, reconstruction, selection), all the numbers displayed in the Table are to be multiplied by $\times 10^{-3}$.

Channel	Run I, L0I	Run I, L0E!	2016, L0I	2016, L0E!
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	4.031		3.910	
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	0.455	0.391	0.608	0.655
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	3.581		3.282	
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	0.362	0.431	0.489	0.632
$\Lambda_b^0 \to pK\mu^+\mu^-$	3.017		3.070	
$\Lambda_b^0 \to pKe^+e^-$	blind	blind	blind	blind

3.7 Invariant mass fits

In order to extract the yields of various decay modes of interest throughout this work, unbinned maximum likelihood fits to the $pK\ell^+\ell^-$ invariant mass distributions are performed. Simulation samples of signal and background modes are used to model the invariant mass shape of each component. The fit is constructed relying on the knowledge of the background composition, acquired in Sec. 3.3. In this section, the fit models which describe the invariant mass distributions of the normalisation, control and signal modes, are outlined in detail.

It is assumed throughout the analysis that the simulation (once the reweighting procedure is performed as described above) describes the invariant mass shapes of the $pK\ell^+\ell^-$ spectra properly. Possible imperfections are accounted for in form of dedicated systematic uncertainties, these studies will be discussed in the following sections.

On top of the description of the signal shapes, the purpose of this section is to give a snapshot of the mass shapes of the dominant background modes which will be encountered later on.

3.7.1 $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$ resonant modes

A striking feature of using resonant modes such as $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$, rises from an additional constraint in the kinematics of the leptons due to the fact that they originate from a specific resonance, a J/ψ meson in this case. This means that applying a J/ψ mass constraint can significantly improve the invariant mass resolution. As it can be seen from Fig. 3.54, the improvement is especially important for the electron mode. In the harsh background environment, improving the mass resolution will have an impact on the uncertainties of the fitted yields. This makes the fits with the J/ψ mass constraint our default choice. Difference between these nominal fits and the fits without the J/ψ mass constraint is then considered as a source of systematic uncertainty. The J/ψ mass constraint in this analysis is implemented using the `DecayTreeFitter` algorithm [156].

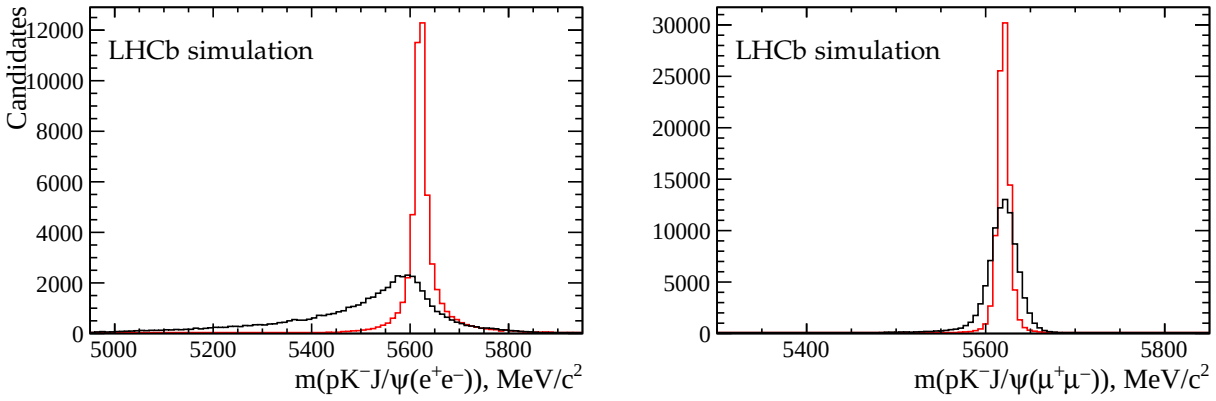


Figure 3.54: Invariant mass shapes of the (left) $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ and (right) $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ signal simulation, with (red) and without (black) the J/ψ mass constraint applied.

The models used for these fits are described in this section, while the fits without constraint are shown in the following section, a comparison between them is one of the important cross-checks of the fit stability. The fit is constructed relying on the knowledge of the background composition, acquired in Sec. 3.3.

3.7.1.a Signal model

Simulation samples of the relevant decay modes are used to describe the shapes of the signal modes. In order to account for differences in the kinematics of the decays, for each channel different shapes will be adopted for different trigger categories.

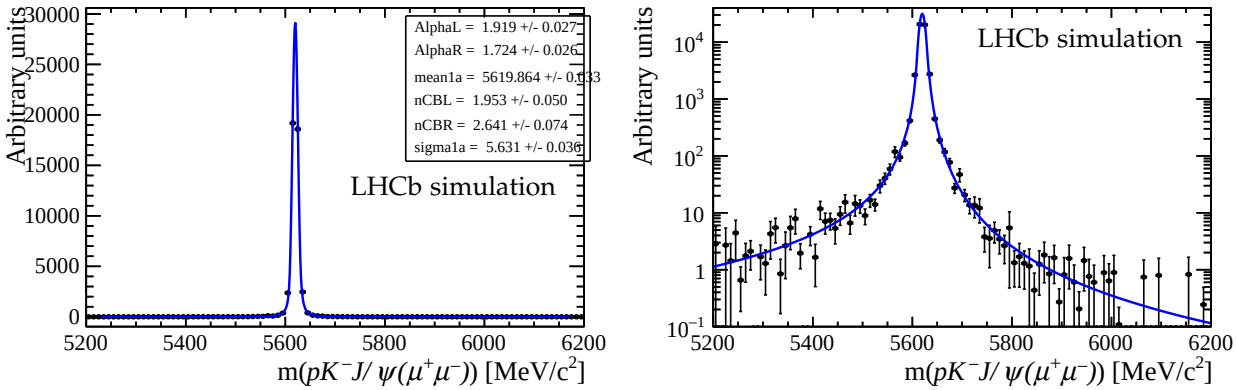


Figure 3.55: Invariant mass shapes of the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ signal in the simulation, with the J/ψ mass constraint applied; in linear and logarithmic scales.

For the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ control mode, one Bifurcated Crystal Ball function (defined in Eq. 2.7) is sufficient to describe the signal PDF.

For what concerns the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ fits where the J/ψ mass constraint is applied, the sum of two Bifurcated Crystal Ball functions is used to describe the signal invariant mass distribution, as it was observed that one single function is not sufficient to describe properly the tails of the signal distribution. It should be noted that, the tails in the electron mode have different properties than those in the muon one due to bremsstrahlung effects and their recovery procedure. The two Bifurcated Crystal Ball functions have no shared parameters, and all the shape parameters are free in the fit to the simulation. When fitting to the data, all the tail parameters are fixed, as well as the difference between the two means, and the ratio of the two widths. One mean and one width are kept free.

The fitted invariant mass shapes are shown in Fig. 3.55 for the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode, and Fig. 3.56 for the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode.

3.7.1.b Background model

To describe the shapes of the dominant background modes introduced in Sec. 3.3, dedicated simulation samples were used. The only exception concerns the combinatorial

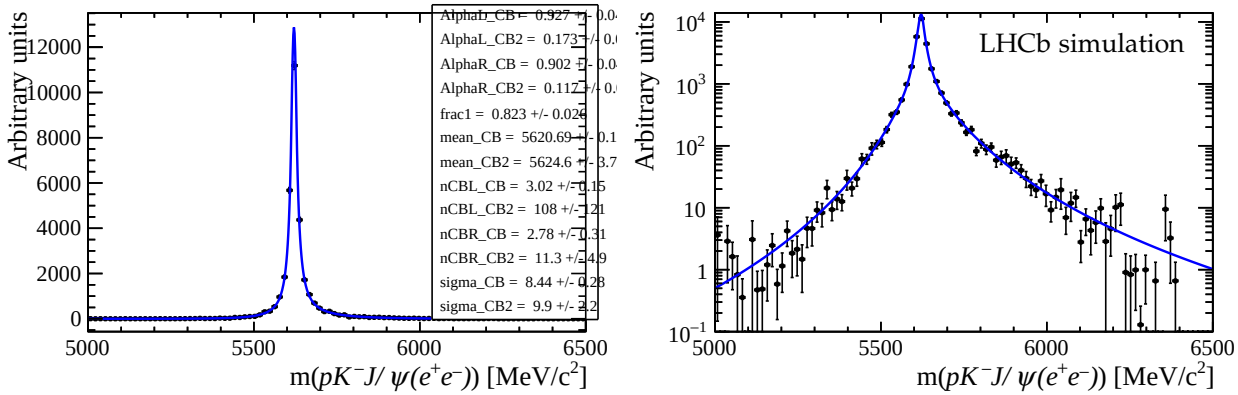


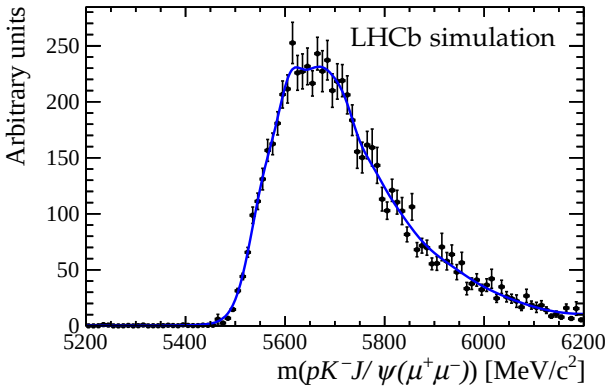
Figure 3.56: Invariant mass shapes of the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ signal, with the J/ψ mass constraint applied; in linear and logarithmic scales.

background which is described by an exponential distribution where the value of the slope is fitted directly in data. The shapes of the specific background modes are fitted with the kernel estimation method (`RooKeysPdf`) [146]. A particular care is needed to monitor the behaviour of the `RooKeysPdf` at the edges of the defined range: the fits to the simulation samples are always performed in a wider range than the final fit to the data.

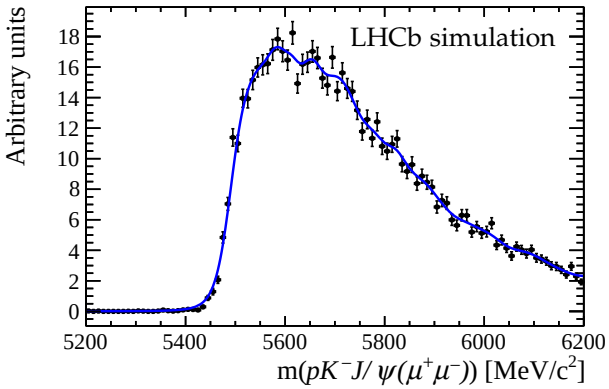
The invariant mass shapes of signal and specific background modes are evaluated separately for each trigger category and data-taking period. As the difference between the categories and runs is small, for simplicity, only one example of each shape will be shown throughout this section.

3.7.1.c Fits to the data

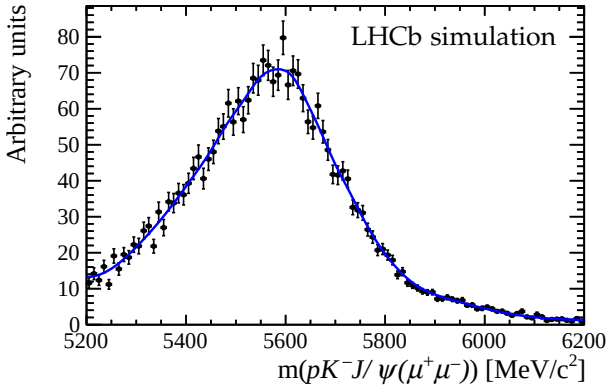
The default fit to the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data includes the following components: signal with free mean and width, combinatorial background described by an exponential shape with a free slope, and three `RooKeysPdf`-modeled misidentification backgrounds for $B_s^0 \to K^+K^-J/\psi(\to \mu^+\mu^-)$, $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$, signal with the proton-kaon ID swap. The yields of the signal, as well as of the combinatorial and of the $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ components are free parameters of the fit. The yields of the hadron-ID swap component is related to the signal yields by a factor estimated from data independently before performing the final fit (the procedure is explained below and in Appendix D). This ratio is also cross-checked on the simulation with the PID correction applied, and estimated as equal to 4.5% in Run I and 0.6% in Run II. Finally, since the B_s^0 and B^0 components have rather similar shapes, and that toy studies show a correlation between their yields, it was decided to fix their ratio using information from data. To do so, the number of relevant events is estimated under the B_s^0 and B^0 signal peaks, as described in Appendix D, (under the appropriate mass hypothesis, while the Λ_b^0 signal candidates are vetoed applying a veto of 35 MeV/c^2 around the J/ψ-constrained Λ_b^0 invariant mass in the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$



(a) $B_s^0 \to K^+K^-J/\psi\,(\to \mu^+\mu^-)$



(b) $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$



(c) $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ with hadrons ID swap

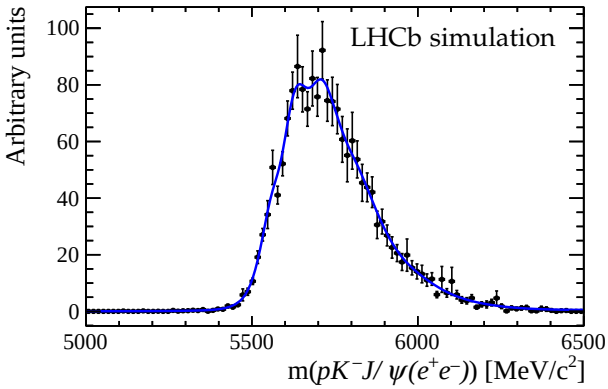
Figure 3.57: Invariant mass shapes of the specific backgrounds to the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ mode, with J/ψ mass constraint applied, in the L0M category for the Run I conditions.

mode). Their ratio ($\frac{N(B_s^0)}{N(B^0)}$=2.3 in Run I, 0.7 in Run II) is taken as a factor relating the yields of these components into the fit. The differences between Run I and Run II originate from the different PID versions of the tuning, as well as the stronger proton p_T cut, resulting into the stronger suppression of misidentified kaons in the Run II. This factor is assumed to be preserved between muon and electron modes, as decay modes with J/ψ in the final state are lepton-universal. Dedicated systematic studies are performed where this ratio is allowed to float, as described in Sec. 3.8.4.

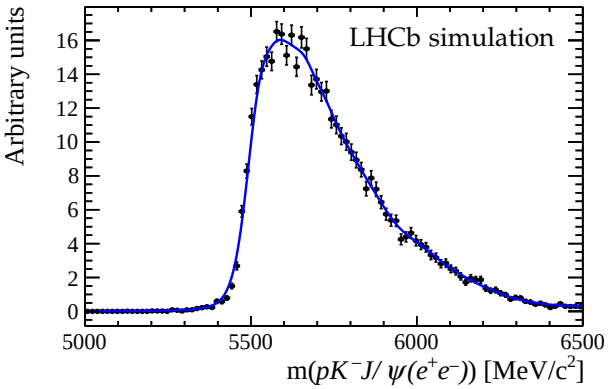
The summary of the fit model (free and derived parameters, invariant mass shapes) is provided in Table 3.19. It should also be noted that for all invariant mass fits presented in this manuscript, the final values of all free fit parameters are shown in the legend.

The exact same fit model (but with electrons instead of muons in the final state) is adopted as default for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode (with the J/ψ mass constraint applied), as no conceptual differences are expected in the background composition between the two.

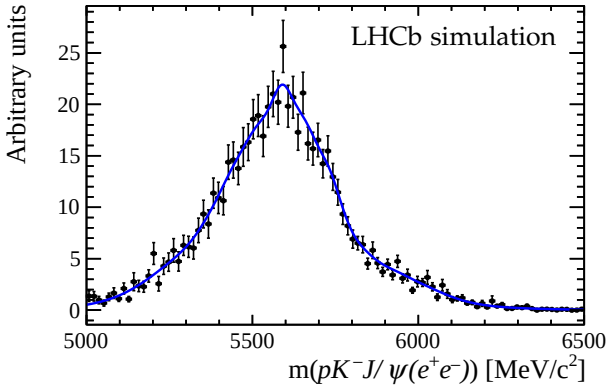
The fit quality is validated by means of the χ^2 per degree of freedom (χ^2/ndf) for each



(a) $B_s^0 \to K^+K^-J/\psi\,(\to e^+e^-)$



(b) $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$



(c) $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ with hadrons ID swap

Figure 3.58: Invariant mass shapes of the specific backgrounds to the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, with J/ψ mass constraint applied, in the L0I category for the Run I conditions.

fit, and making sure these values are reasonably close to unity. They were found to be: for $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ channel, 1.6 in Run I and 1.1 in Run II fit; for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ channel, between 0.6 and 0.8 for all four datasets.

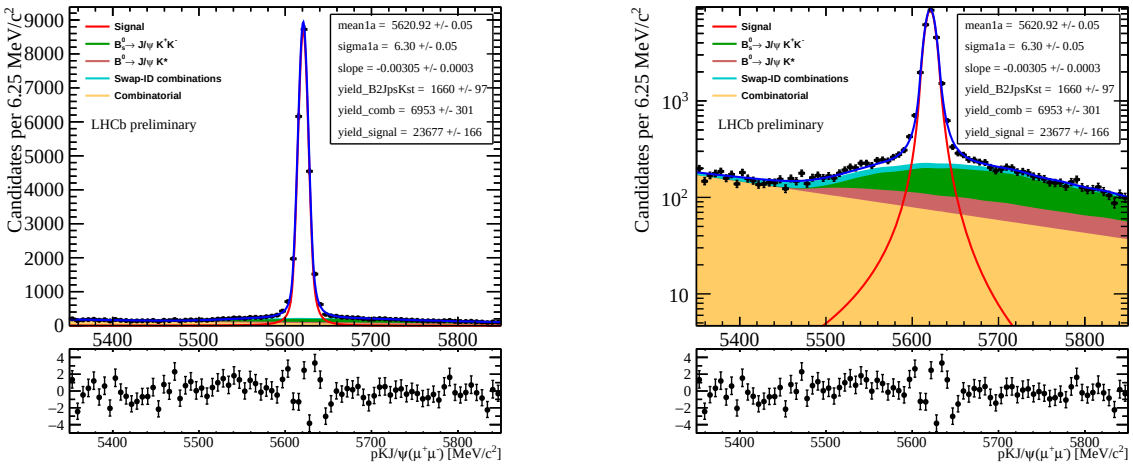


Figure 3.59: Fit to the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data, with the J/ψ mass constraint applied, for the Run I data taking period, in the L0M category, in linear and logarithmic scales.

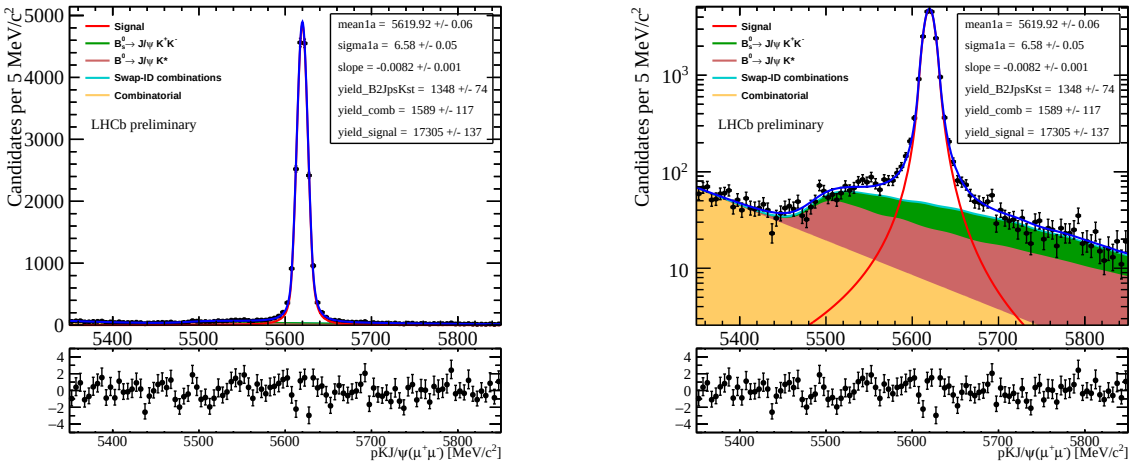


Figure 3.60: Fit to the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data, with the J/ψ mass constraint applied, for the Run II data taking period, in the L0M category, in linear and logarithmic scales

3.7.1.d Fit validation

Pseudo-experiments are performed, generating a pseudo-dataset according to the fit model developed earlier, and subsequently fitting to it by the same model. The generated values of the fit parameters are chosen to be close to the ones obtained in the default fits to the data. The fit stability is studied looking at the pull distributions, quantifying how close they are to Gaussian distributions centered at zero and having a width equal to unity. After generating tens of thousands pseudo-experiments, it can be observed that no significant bias is seen in the signal yields and most of other free fit parameters, as it is shown in Figs. 3.63 and 3.64.

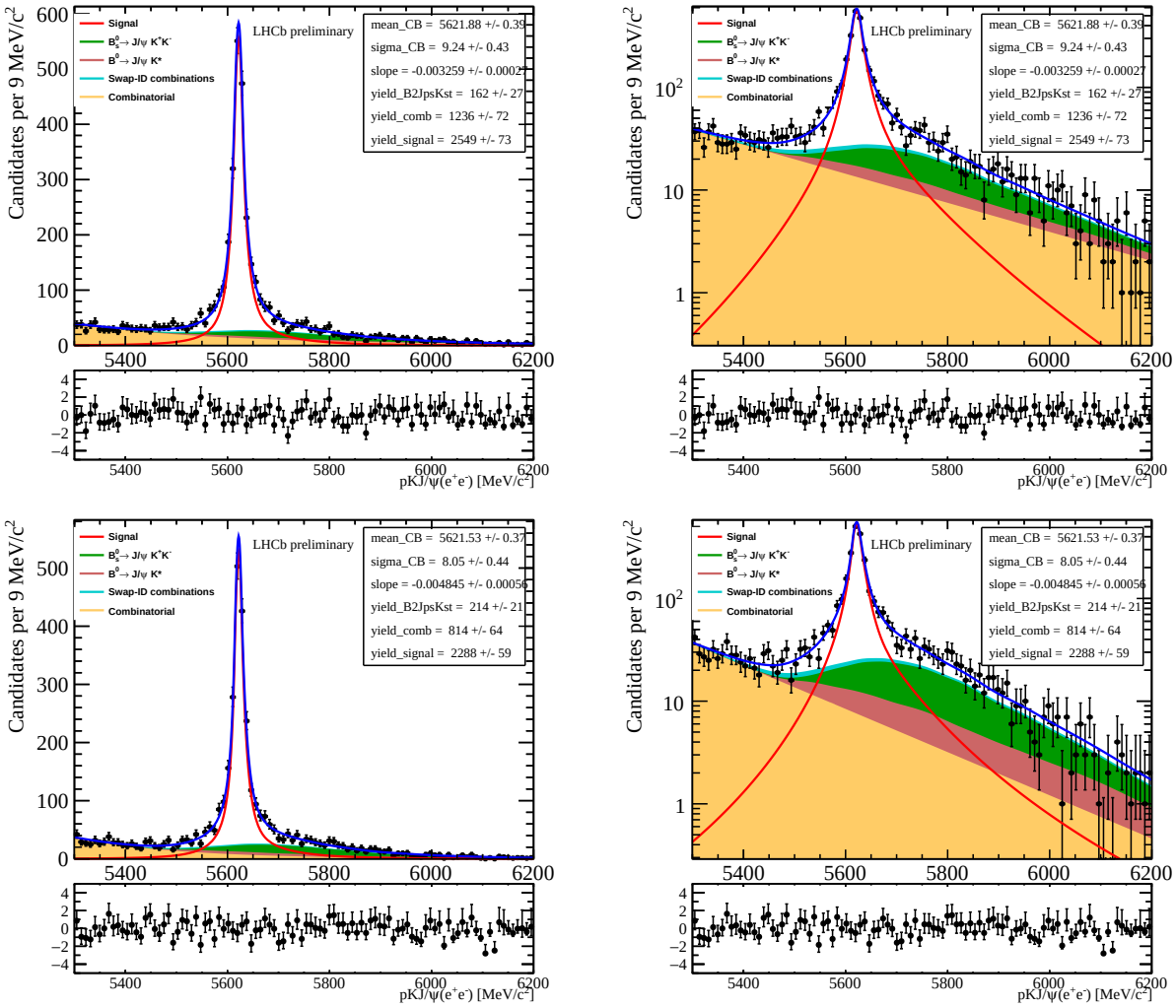


Figure 3.61: Fit to the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ data, with the J/ψ mass constraint applied, for the Run I data taking period, in the (top) L0I category, (bottom) L0E! category, in linear and logarithmic scales

It can be concluded that the default fits to the normalisation modes behave in a stable manner, showing no significant bias and a good coverage of uncertainties.

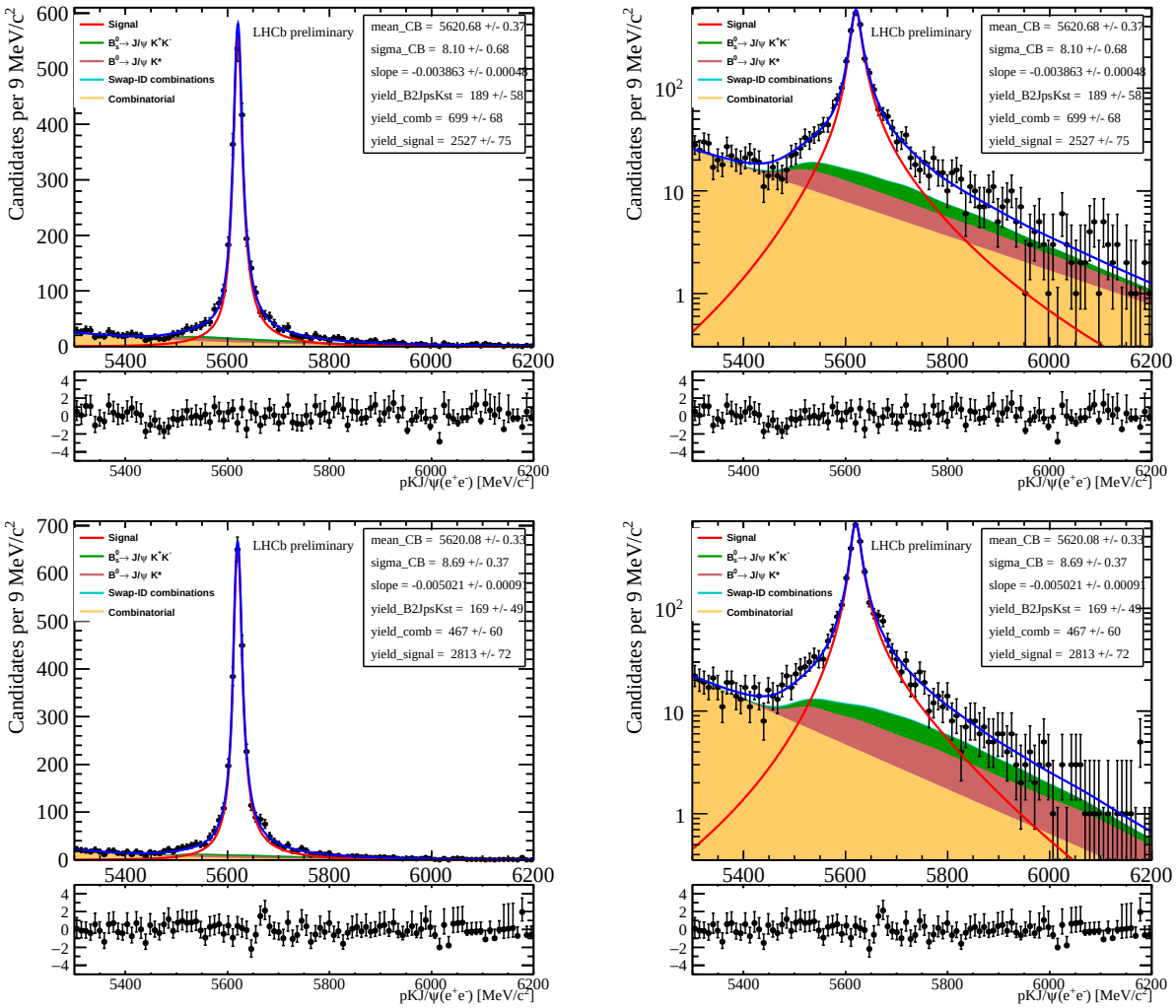


Figure 3.62: Fit to the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ data, with the J/ψ mass constraint applied, for the Run II data taking period, in the (top) L0I category, (bottom) L0E! category, in linear and logarithmic scales

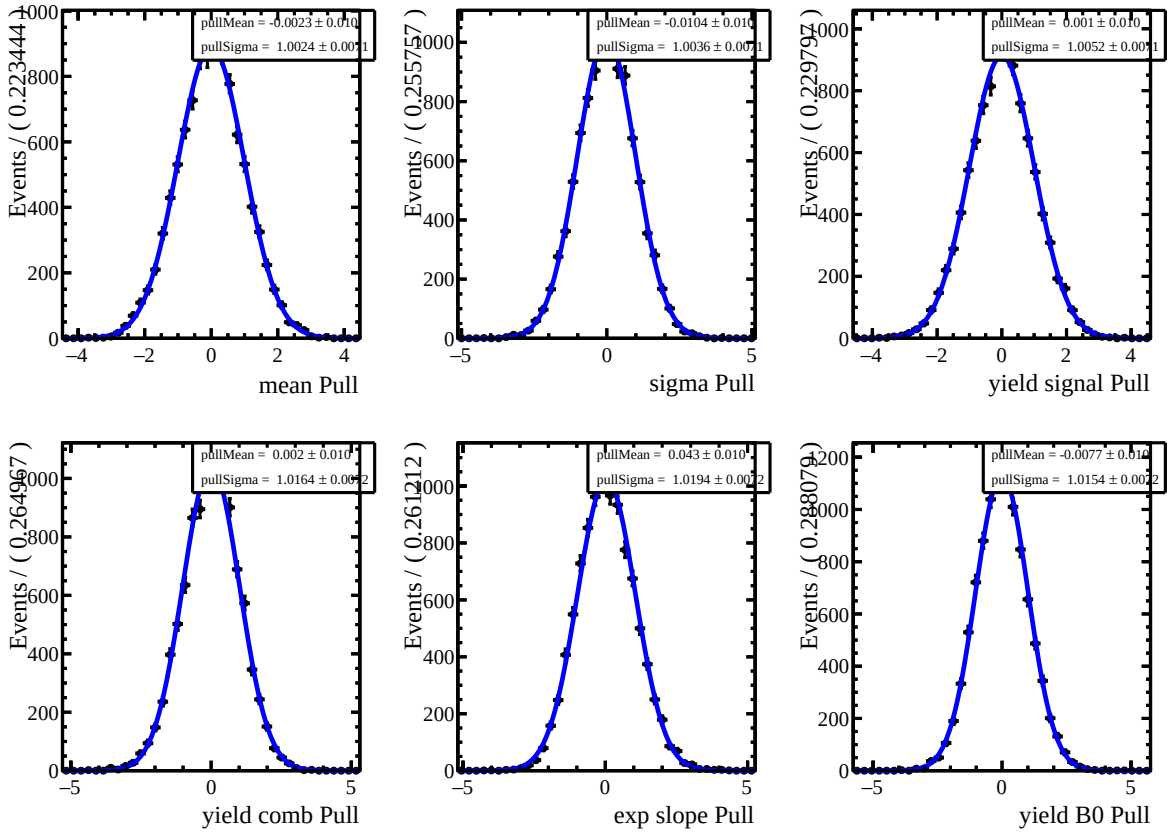


Figure 3.63: Pull distributions of the free fit parameters for the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ fit, based on 100000 pseudo-experiments. The investigated observables are (from left to right, top to bottom) the mean of the signal distribution, width of the signal distribution the signal yields, the combinatorial background yields, the slope of the exponential background, the $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ yields.

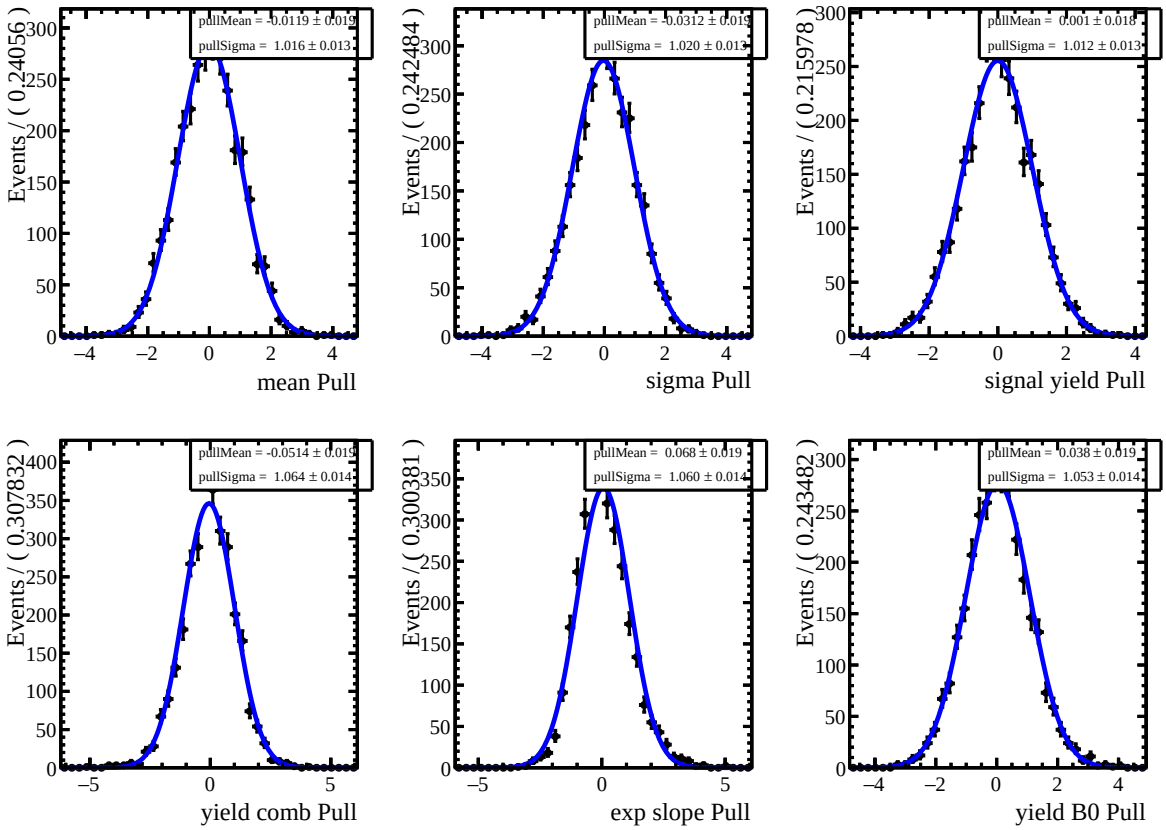


Figure 3.64: Pull distributions of the free fit parameters for the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ fit, based on 30000 pseudo-experiments. The investigated observables are (from left to right, top to bottom) the mean of the signal distribution, the width of the signal distribution, the signal yield, yield of the combinatorial background, the slope of the exponential background, the $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ yield.

3.7.2 $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$

The fits to $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ decays are performed with a $\psi(2S)$ mass constraint. It should be noted as already mentioned in the introduction that relative branching fraction of the $\Lambda_b^0 \to pK\psi(2S)$ decay (6.6×10^{-5}) is significantly lower than those of the similar decays of the B^0 and B_s^0 mesons which are backgrounds (*e.g.*, $\mathcal{B}$ ($B^0 \to \psi(2S)K^{*0}$)=5.9×10^{-4}).

It should also be noted that for the electron modes, candidates coming from the leakage of J/ψ events, as described in the Sec. 3.3.4.a, populate the upper sideband when applying the $\psi(2S)$ mass constraint. In order to clean up the sideband, these events are vetoed applying a cut on the Λ_b^0 invariant mass with J/ψ mass constraint $m(pK(\ell^+\ell^-)_{J/\psi}) < 5500\,\text{MeV}/c^2$. The efficiency of this cut is taken into account when estimating $R_{\psi(2S)}$.

3.7.2.a Signal model

A Bifurcated Crystal Ball function is used to describe the signal shape in the muon mode, and is shown in Fig. 3.65 (left). For the electron mode, same as in the J/ψ case, the one Bifurcated CB is not sufficient to fit the tails properly, so the sum of two Bifurcated CB functions is used instead, this shape is shown in Fig. 3.65 (right).

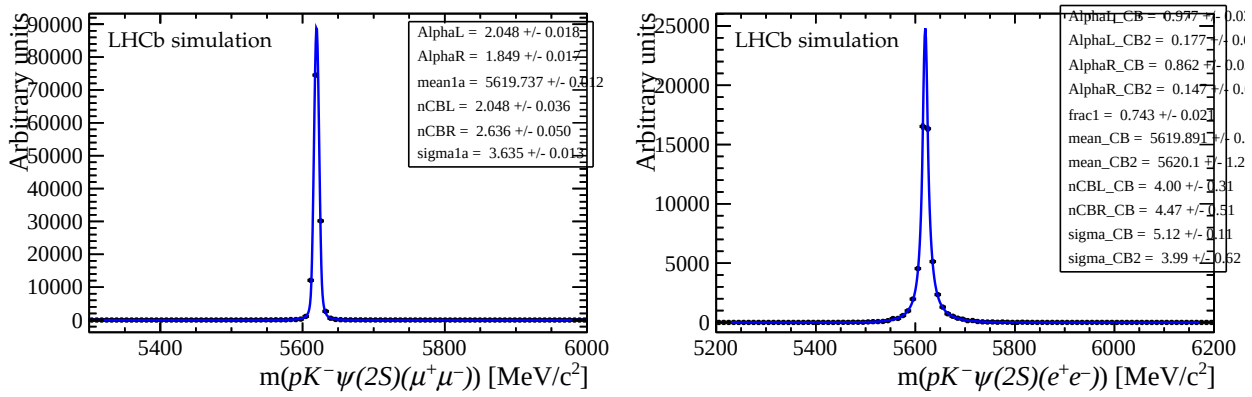


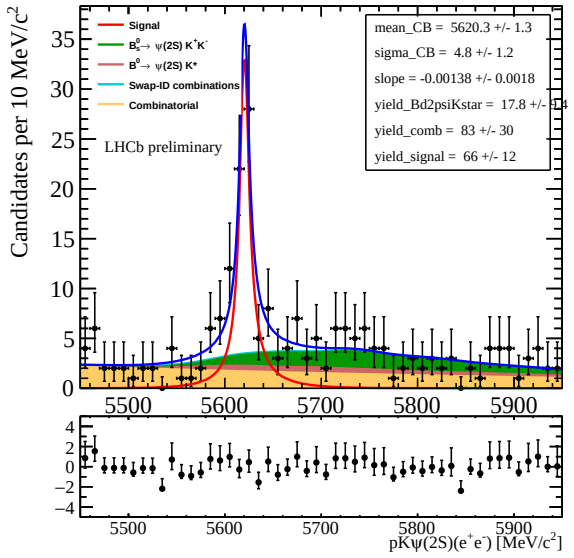
Figure 3.65: Invariant mass shape of the (left) $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$ and (right) $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$ signal, both with the $\psi(2S)$ mass constraint applied.

3.7.2.b Background model

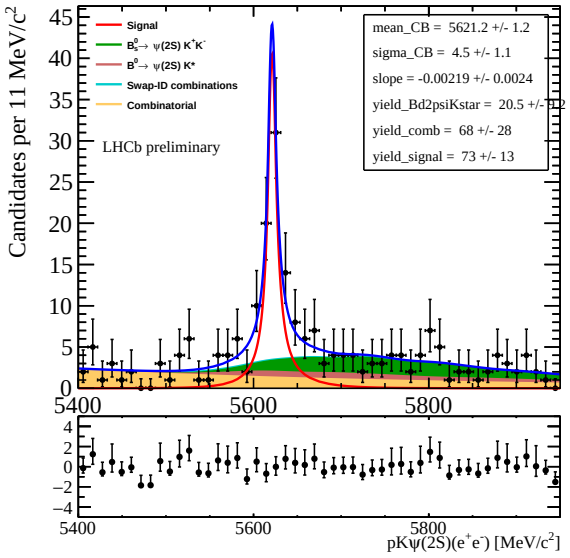
The following background modes are considered as important and are included into the fit: the combinatorial background (modeled as an exponential, the proton misidentification modes (decays of B_s^0, B^0), and the hadron ID swap. For simplicity, their invariant mass distributions are not shown here, as they are quite similar to the ones in the relevant J/ψ channels. With the current statistics, no need for including other background modes is seen.

3.7.2.c Fits to the data

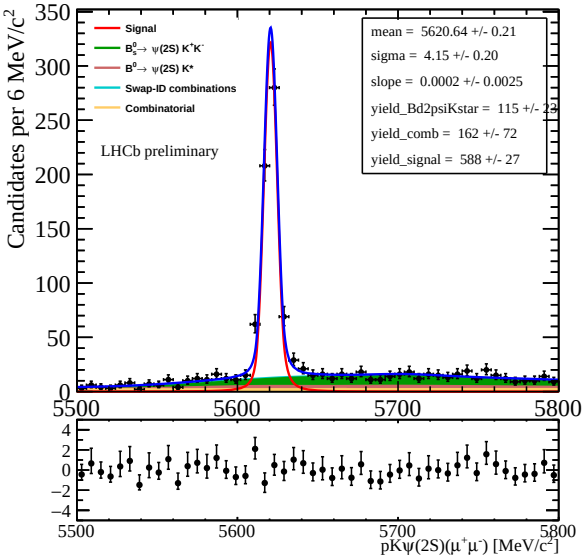
The following fit parameters are free: the mean and width of the signal shape; the slope of the combinatorial background; the yields of signal, combinatorial and B^0 components. The B_s^0 component has yields which are related to the one of B^0 via the same factor as in the fits in the J/ψ region, which is an approximation. The fit model is summarized in Tab. 3.19.



(a) $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$, L0I category

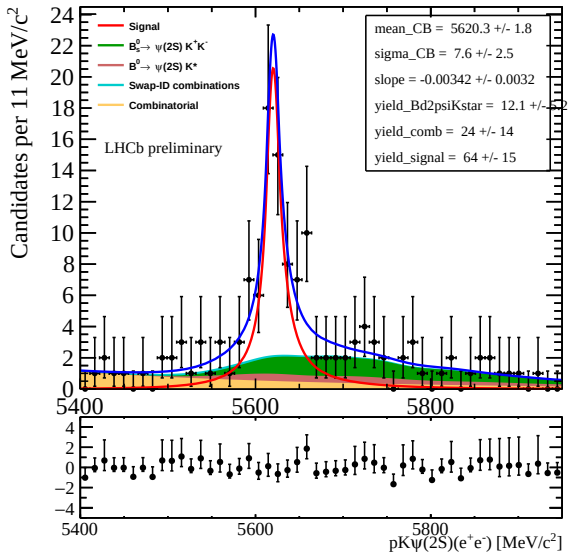


(b) $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$, L0E! category

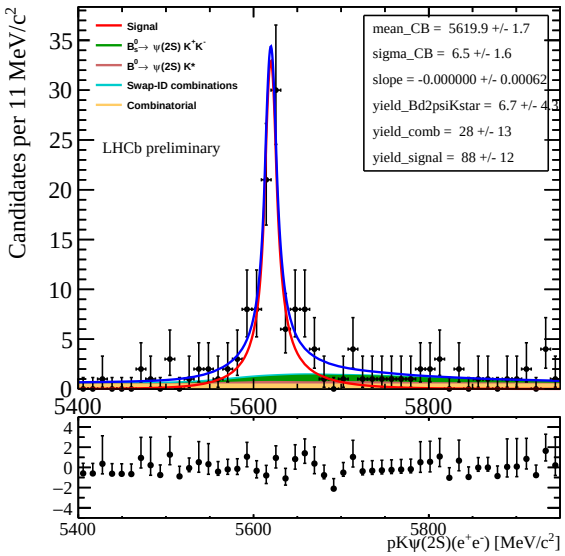


(c) $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$, L0M category

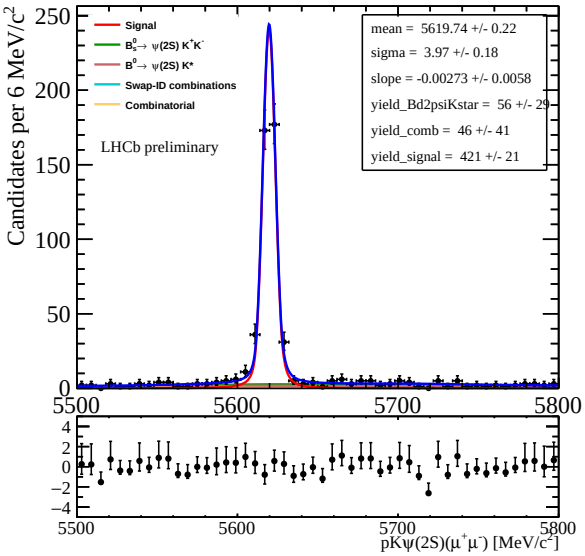
Figure 3.66: Fits to $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$ and $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$ data, with $\psi(2S)$ mass constraint applied, for the Run I conditions.



(a) $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$, L0I category



(b) $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$, L0E! category



(c) $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$, L0M category

Figure 3.67: Fits to $\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$ and $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$ data, with $\psi(2S)$ mass constraint applied, for the Run II conditions.

Decay	Signal	Comb.	Hadronic misidentification
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	Bifur. CB	exp.	$B^0 \to K^{*0}J/\psi$ $B_s^0 \to K^+K^-J/\psi$ pK-swap
Free parameters	yields, mean width	yields, slope	Y. of $B^0 \to K^{*0}J/\psi$
Fixed parameters	tails		(Y. of $B_s^0 \to K^+K^-J/\psi$)/(Y. of $B^0 \to K^{*0}J/\psi$) (Y. of pK-swap)/(Y. of $\Lambda_b^0 \to pKJ/\psi$)
Derived parameters			Y. of $B_s^0 \to K^+K^-J/\psi$, Y. of pK-swap
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	sum of 2 Bifur. CB	exp.	$B^0 \to K^{*0}J/\psi$ $B_s^0 \to K^+K^-J/\psi$ pK-swap
Free parameters	yields, mean	yields, slope	Y. of $B^0 \to K^{*0}J/\psi$
Fixed parameters	tails, diff of two means, ratio of two widths		(Y. of $B_s^0 \to K^+K^-J/\psi$)/(Y. of $B^0 \to K^{*0}J/\psi$) (Y. of pK-swap)/(Y. of $\Lambda_b^0 \to pKJ/\psi$)
Derived parameters			Y. of $B_s^0 \to K^+K^-J/\psi$, Y. of pK-swap
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	Bifur. CB	exp.	$B^0 \to K^{*0}\psi(2S)$ $B_s^0 \to K^+K^-\psi(2S)$ pK-swap
Free parameters	yields, mean	yields, slope	Y. of $B^0 \to K^{*0}\psi(2S)$
Fixed parameters	tails, width		(Y. of $B_s^0 \to K^+K^-\psi(2S)$)/(Y. of $B^0 \to K^{*0}\psi(2S)$) (Y. of pK-swap)/(Y. of $\Lambda_b^0 \to pK\psi(2S)$)
Derived parameters			Y. of $B_s^0 \to K^+K^-J/\psi$, Y. of pK-swap
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	Bifur CB	exp.	$B^0 \to K^{*0}\psi(2S)$ $B_s^0 \to K^+K^-\psi(2S)$ pK-swap
Free parameters	yields, mean, width	yields, slope	Y. of $B^0 \to K^{*0}\psi(2S)$
Fixed parameters	tails, diff of two means, ratio of two widths		(Y. of $B_s^0 \to K^+K^-\psi(2S)$/(Y. of $B^0 \to K^{*0}\psi(2S)$) (Y. of pK-swap)/(Y. of $\Lambda_b^0 \to pK\psi(2S)$)
Derived parameters			Y. of $B_s^0 \to K^+K^-J/\psi$, Y. of pK-swap

Table 3.19: Summary of the nominal fit configurations to invariant mass distributions used on data for the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ and $\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-)$ resonant modes. All the hadron misindentified backgrounds are modelled with `RooKeysPdf`, their invariant mass shapes are always taken from simulation. "Y." is used to refer to the yields of each mode.

3.7.3 Rare mode $\Lambda_b^0 \to pK\mu^+\mu^-$

For the non-resonant decay modes, we do not have an additional handle in a form of the dilepton invariant mass constraint. Therefore, a typical mass resolution in these modes is larger than in the default fits to the resonant modes.

3.7.3.a Signal model

To describe the signal distribution, a Bifurcated Crystal Ball function is used. The fit to the simulation is presented in Fig. 3.68.

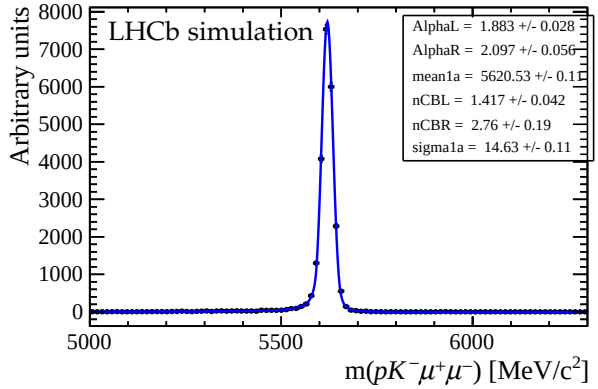
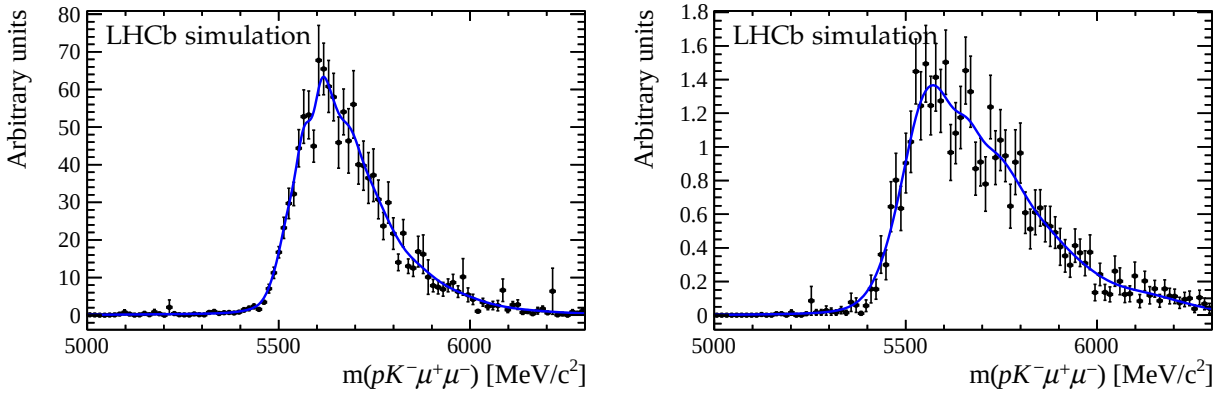


Figure 3.68: Invariant mass shape of the $\Lambda_b^0 \to pK\mu^+\mu^-$ signal.

3.7.3.b Background model

The important background components within the adopted fit range are the misidentification backgrounds. The component with proton-kaon ID swap is omitted as its expected yields are very small.



(a) $B_s^0 \to K^+K^-\mu^+\mu^-$ (b) $B^0 \to K^{*0}\mu^+\mu^-$

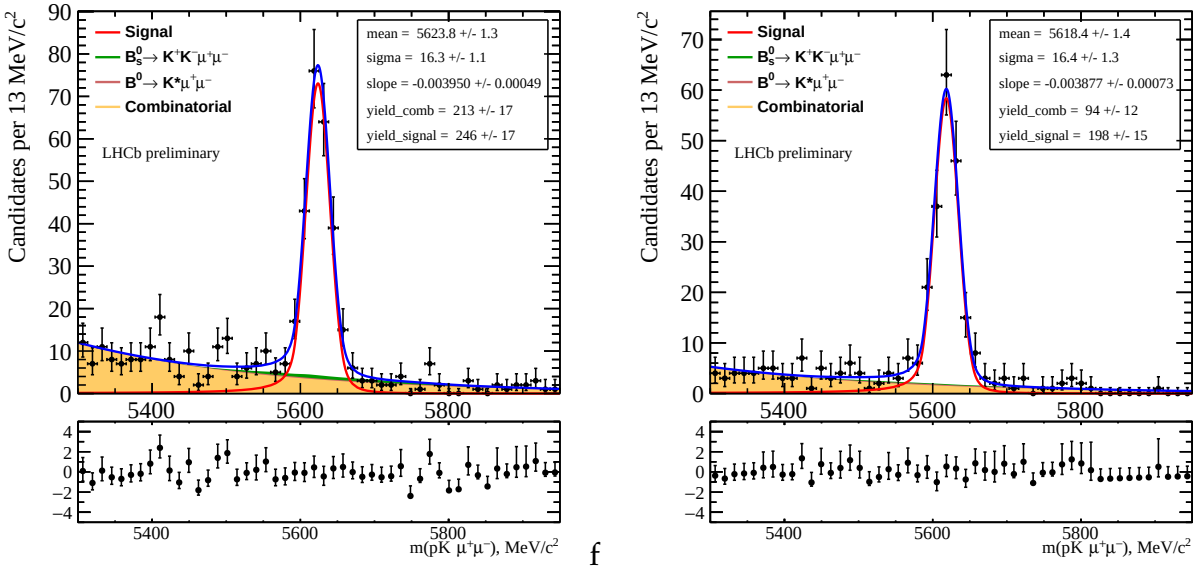
Figure 3.69: Invariant mass shapes of the specific backgrounds to the $\Lambda_b^0 \to pK\mu^+\mu^-$ mode.

3.7.3.c Fits to the data

The yields of the misidentification components ($B_s^0 \to K^+K^-\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$) are small, therefore they have to be constrained in the fit. The branching fraction of the $B^0 \to K^{*0}\mu^+\mu^-$ decay is known [6], this allows to rely on the $B^0 \to K^{*0}\mu^+\mu^-$ simulation in order to estimate the expected background yield. This estimate results[5] in a ratio of yields of the $B^0 \to K^{*0}\mu^+\mu^-$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ components to be about 0.027 in both Run I and Run II. A systematic uncertainty is then assigned for this assumption, by lifting this constraint in the toy studies (Sec. 3.10.1).

Similar estimate cannot be obtained for the $B_s^0 \to K^+K^-\mu^+\mu^-$ decay, as its branching fraction has not been measured outside the (vetoed in this analysis) $\phi(1020)$ region. Therefore, the best assumption which can be made is that the ratio of the B_s^0 and B^0 channels is the same as in the J/ψ window. As in the previous case, a systematic uncertainty is then assigned for this assumption, performing the toy studies with different hypotheses for this value (Sec. 3.10.1).

The fit model is summarized in Tab. 3.21. The fits to $\Lambda_b^0 \to pK\mu^+\mu^-$ data are shown in Fig. 3.70 for Run I and Run II datasets, and the fitted values of all free fit parameters are shown in the legend of each fit.



f

Figure 3.70: Fits to $\Lambda_b^0 \to pK\mu^+\mu^-$ data, for the Run I (left) and Run II (right) conditions, in the L0M category.

3.7.3.d Fit studies using pseudo-experiments

In this study, the nominal Run I yields and fit parameters were used. Pull results are presented in Fig. 3.71. The fit exhibits reasonable behaviour and a good coverage of the

[5] Assuming the value of the $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$ measured in this analysis. Therefore, this is a somewhat iterative procedure.

uncertainty on the signal yield.

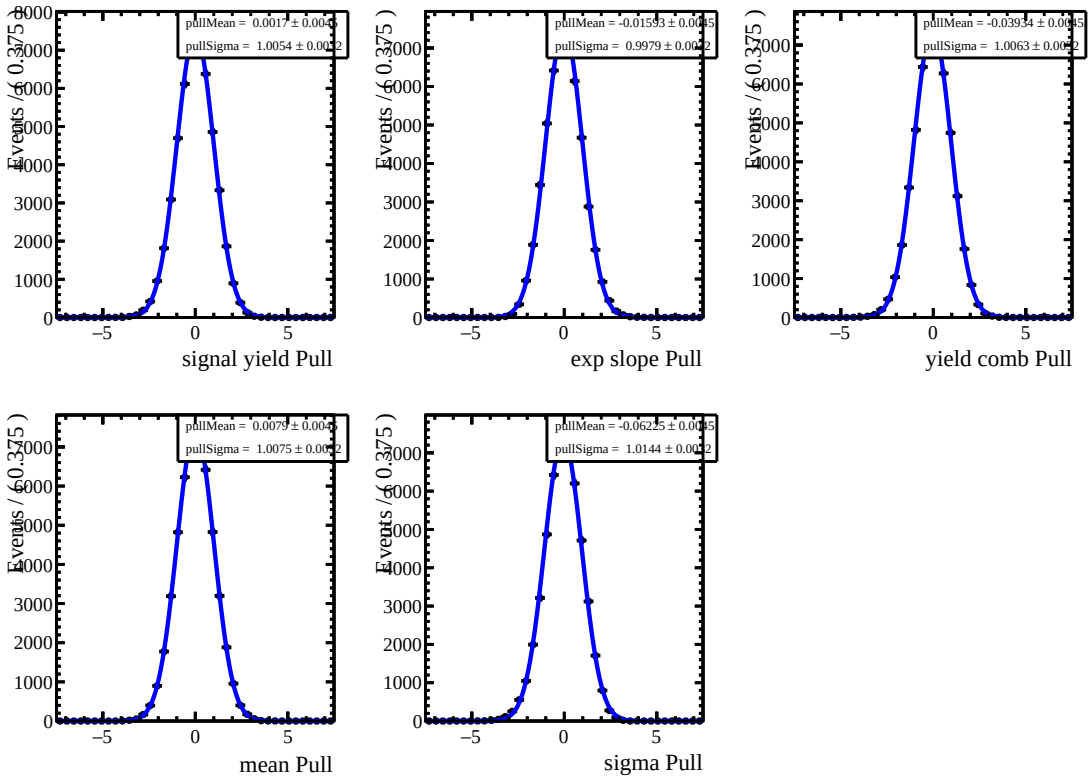


Figure 3.71: Pull distributions of the 50000 pseudo-experiments based on the fit model for the $\Lambda_b^0 \to pK\mu^+\mu^-$ channel. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential background, the combinatorial background yield, the mean of the signal distribution, as well as the width of the signal distribution.

3.7.4 Rare mode $\Lambda_b^0 \to pKe^+e^-$

3.7.4.a Signal model

For the electron mode, three bremsstrahlung categories having very different shapes are considered, depending on the number of bremsstrahlung photons attached to the dielectron pair: category 0 (no photons), 1 (one photon) and 2 (more than one photon). As category 0 has no photons attached, it has only a lower mass tail but the upper tail is tiny. The other two categories have both lower and upper tails. This difference is accounted for when choosing the PDF to describe the signal shape: the invariant mass shape is described by a Crystal Ball function (category 0), or the sum of two Bifurcated Crystal Ball functions (1 and 2), as it is shown in Fig. 3.72. The Crystal Ball function [157] consists of a Gaussian core with a power-law tail below a certain threshold, and it is given by:

$$CB(x,\alpha,n\mu_B,\sigma) = N \times \begin{cases} A \times (B - \frac{x-\mu_B}{\sigma})^{-n}, \text{ for } \frac{x-\mu_B}{\sigma} \leq -|\alpha| \\ \exp\left(-\frac{(x-\mu_B)^2}{2\sigma^2}\right), \text{ for } -\frac{x-\mu_B}{\sigma} \leq -|\alpha| \end{cases} \tag{3.10}$$

where:

$$A = (\tfrac{n}{|\alpha|})^n \times \exp\left(-\tfrac{|\alpha|^2}{2}\right), \quad B = \tfrac{n}{|\alpha|} - |\alpha|$$

and N is a normalisation factor. All the tail parameters are kept free when fitting to the simulation, but then frozen in the fit to data. The three categories are not obliged to share the same mean and width: these parameters are free for the category 0, and for the other two categories they are shifted by a free difference (mean) or ratio (width) with respect to the category 0. When fitting to the data, these parameters of difference of the means values (ratio of widths), are frozen. The relative ratio between categories is fixed from the simulation.

3.7.4.b Background model

The invariant mass shapes of the important background contributions are presented in Fig. 3.73. They include misidentified backgrounds, but also some types of backgrounds which are not important in other decay modes. Due to a very broad fit range, a partially-reconstructed background with a lost π^0 has to be taken into account. In addition, the so-called J/ψ leakage, *i.e.* candidates of $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ decay "leaking" into the q^2 window of interest due to bremsstrahlung (Sec. 3.3.4.a), is also important at the lower edge of the fit range.

The partially reconstructed background with a lost π^0 arises from decays of the type $\Lambda_b^0 \to pKe^+e^-\pi^0$. Full MC samples were only generated with a phase-space decay model, however the invariant pKe^+e^- mass shape is actually sensitive to the presence of hadronic resonances in this decay, since the energy carried away by the non-reconstructed π^0 can vary significantly. This is studied in `RapidSim` by generating samples with the phase-space model, with an intermediate K^{*-} resonance decaying to $K^-\pi^0$ and with an intermediate Δ resonance going to $p\pi^0$. The pKe^+e^- invariant mass distribution for these three models as generated by `RapidSim` is shown in Fig. 3.74 (left), where the signal shape as obtained in `RapidSim` is also shown for comparison. The kinematic selection, background vetoes and HOP requirement have been applied to these samples. Significant differences are observed for the phase-space and the model with containing hadronic resonances. The more realistic model, which includes a K^{*-} intermediate resonance, is used for the nominal fit, while the other two models are used to compute the systematic uncertainty associated to the shape of this background. It is also checked that the change in shape is negligible if the pK^{*-} originates from a Λ^* resonance[6] or when a physics model (VSS) instead of phase-space is used for the K^{*-} decay. The invariant mass shape obtained from `RapidSim` cannot be directly used in the fit to data due to imprecise modeling of detector effects in this fast simulation, most notably the bremsstrahlung emission and recovery. The physics model is instead implemented using the full phase-space simulation, by re-weighting it with the

[6]Note that Δ^+K^- final state is unlikely to originate from Λ^* resonances due to isospin suppression.

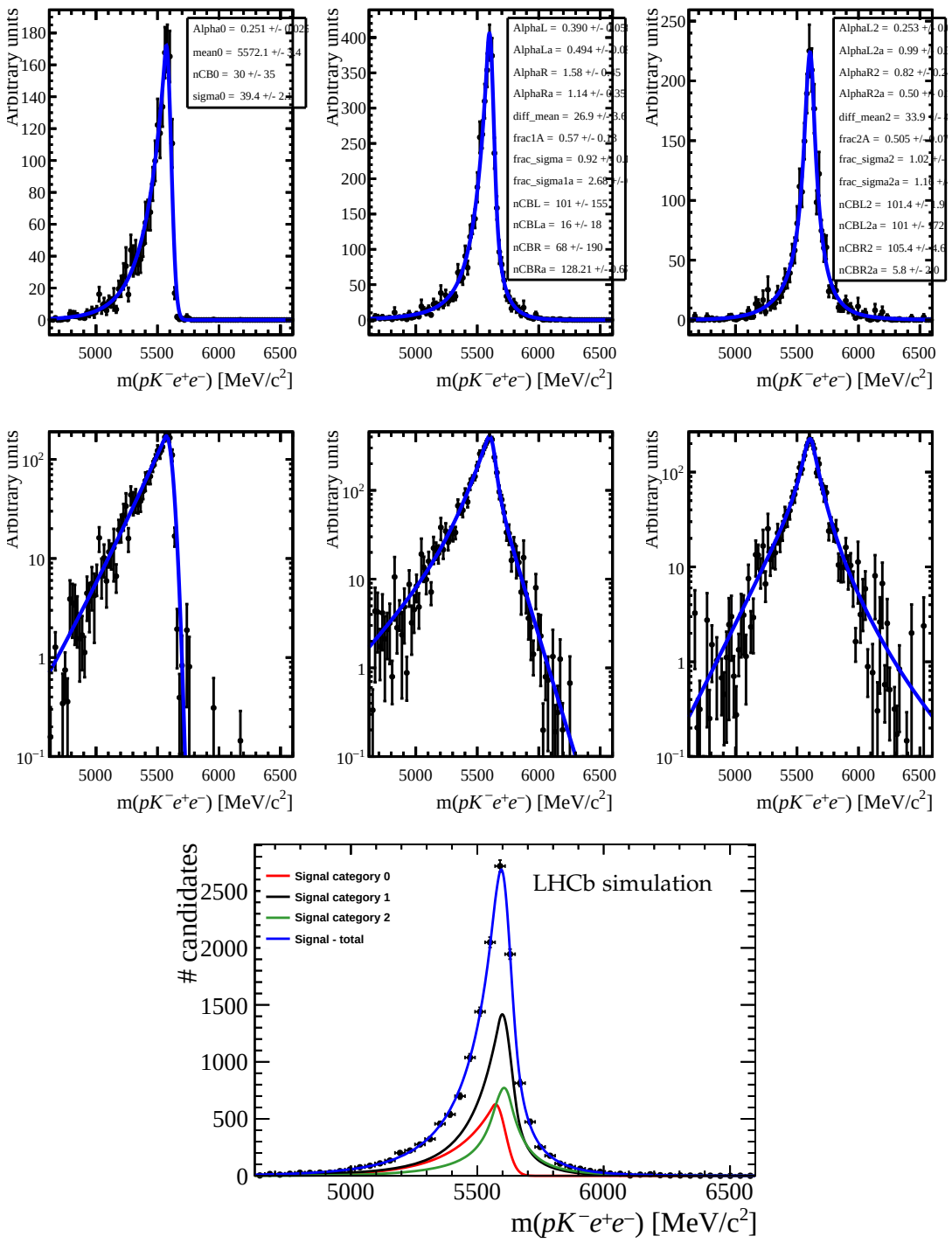
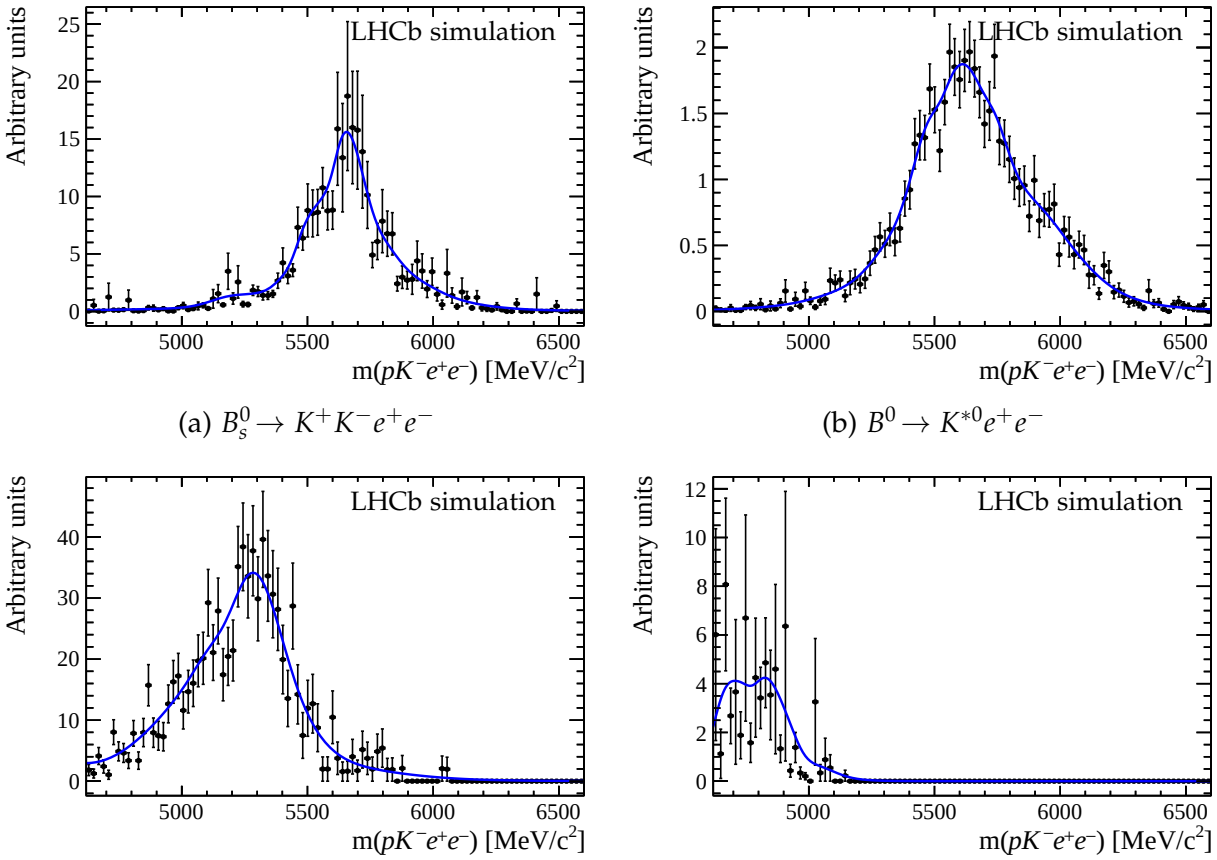


Figure 3.72: Fits to the $\Lambda_b^0 \to pKe^+e^-$ signal simulation. Three bremsstrahlung categories in (top) linear and (middle) logarithmic scales. Bottom: final signal shape created by merging three categories.



(a) $B_s^0 \to K^+K^-e^+e^-$ (b) $B^0 \to K^{*0}e^+e^-$

(c) Partially reconstructed (lost π^0) (d) J/ψ leakage

Figure 3.73: Invariant mass shapes of the specific backgrounds to the $\Lambda_b^0 \to pKe^+e^-$ mode.

ratio between the model including a given resonance and the one using plain phase-space in `RapidSim`. The weights used to do so are shown in Fig. 3.74 (right) for the K^{*-} case.

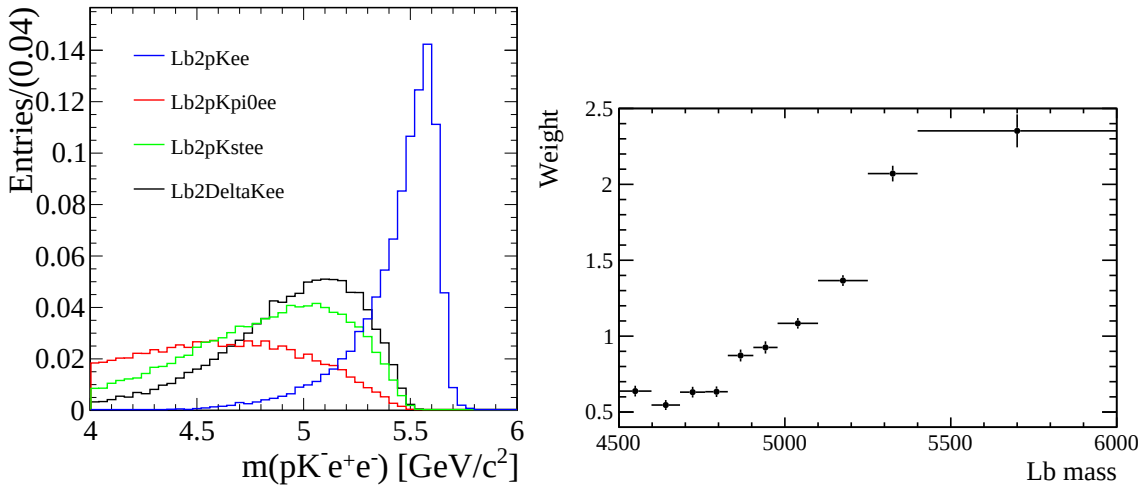


Figure 3.74: Left: Invariant mass shapes of the partially reconstructed background with a lost π^0, for different decay models generated in *RapidSim*. The signal shape is also included for reference. Right: weights used to reweight full simulation sample in order to match the $\Lambda_b^0 \to pK^{*-}e^+e^-$ decay model.

3.7.4.c Fits to the data

The fit to $\Lambda_b^0 \to pKe^+e^-$ data includes the following components and free parameters. The signal shape has free yields; the mean fixed from the simulation; the width fixed from the simulation and scaled by a resolution factor of 1.15 obtained as an average of the scaling factors in the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode, in order to take into account the data-simulation discrepancies. Background components include a combinatorial background described by an exponential (which has a free slope and yield), a partially-reconstructed background with a lost π^0 (free yield), and the J/ψ leakage component. The yield of this J/ψ-leakage component is estimated separately in each trigger category and run period, from the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ yield in the fit to the normalisation mode, and a ratio of relevant efficiencies:

$$N_{leak} = N_{\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)} \times \frac{\varepsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-), q^2 = (0.1, 6))}{\varepsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-), q^2 = (6, 11))}, \tag{3.11}$$

where $\varepsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-), q^2 = (6,11))$ is computed using the corrected $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ simulation with the selection corresponding to the J/ψ window, while $\varepsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, q^2 =(0.1,6)) – using the same $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ simulation but with the selection corresponding to the $\Lambda_b^0 \to pKe^+e^-$ channel. The values of the estimated N_{leak} are provided in the Tab. 3.20, and then fixed in the fit.

The semileptonic $\Lambda_b^0 \to pD^0e\nu_e$ background is not modelled in this fit, because a) it is located further away from the signal region, b) is heavily suppressed by the HOP cut removing more than 85% of it, c) the simulation statistics is too small to get a precise invariant mass shape, and d) it interplays largely with the combinatorial background.

Table 3.20: Estimated yields of the J/ψ-leakage, per trigger category and dataset.

Run I, L0I	Run I, L0E!	2016, L0I	2016, L0E!
9.6 ± 1.1	6.9 ± 0.9	11.6 ± 0.7	10.7 ± 0.7

The strategy of dealing with the misidentification backgrounds is identical to the one used for the muon mode (Sec. 3.7.3.c), with the relations between the yields fixed to the same values.

The fit model is summarized in Tab. 3.21.

The signal window is kept blind in the region (5200, 5800) MeV/c^2, which is covering about 90% of signal events according to the simulation. However, the full fit framework is run, with all the components included. The total fit, the signal component and the misidentification background components (related to the signal by known fractions of yields) are not plotted, as well as the data points in the signal window; and their yields are excluded from the printout.

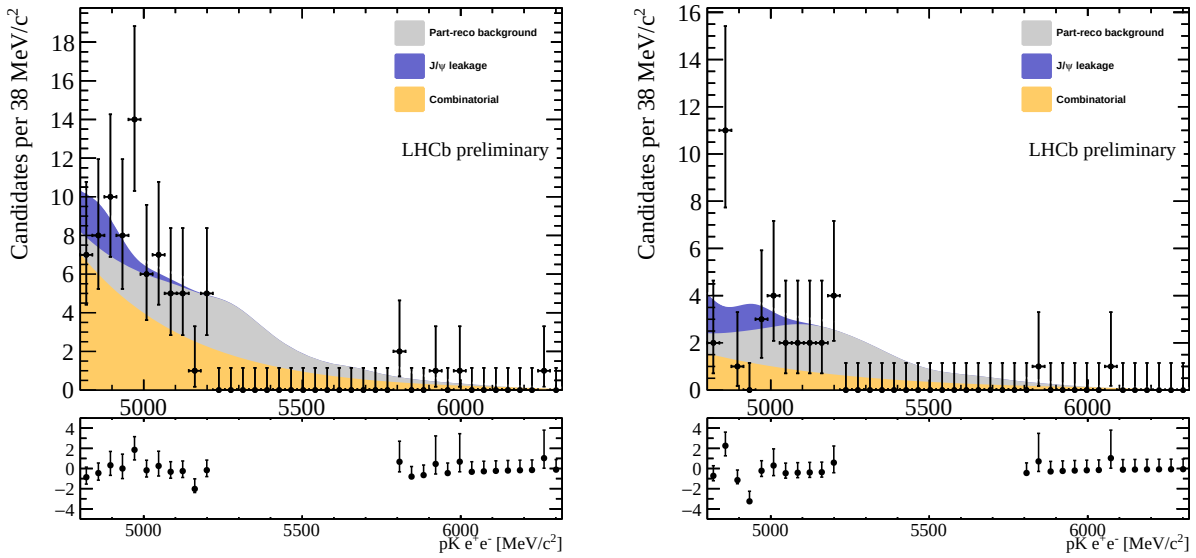


Figure 3.75: Blinded fit to the Run I $\Lambda_b^0 \to pKe^+e^-$ data in the (left) L0I and (right) L0E! trigger category.

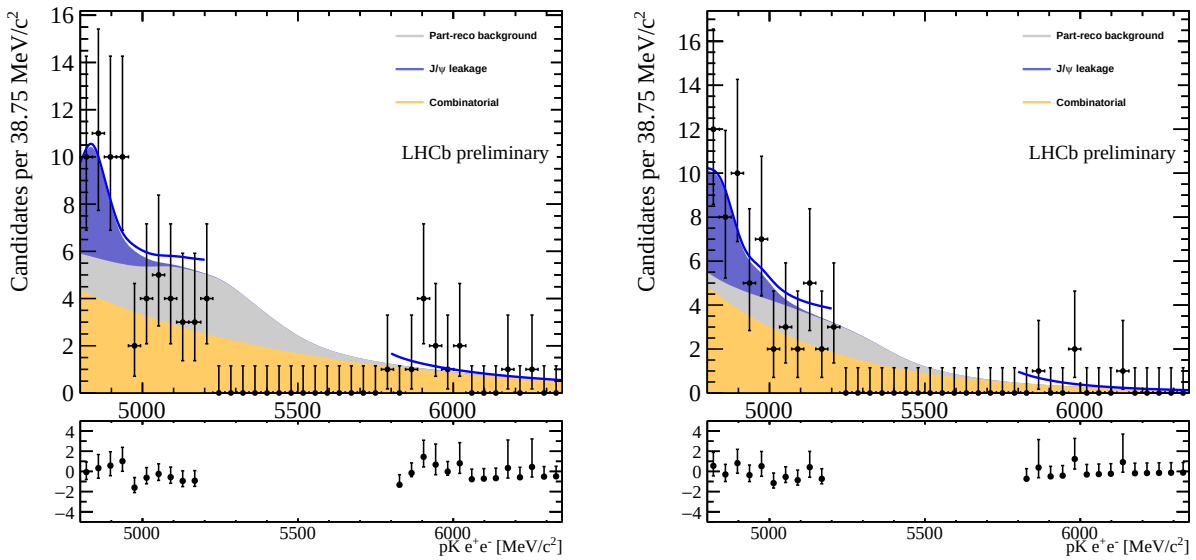


Figure 3.76: Blinded fit to the Run II $\Lambda_b^0 \to pKe^+e^-$ data in the (left) L0I and (right) L0E! trigger category.

3.7.4.d Fit studies using pseudo-experiments

To have an idea of how the unblinded fit will behave, a toy dataset is generated according to the fit model, and then the same model is then fitted to it. At the generation stage, the background yields are obtained from the blinded fit. The signal yields can only be "guessed" by scaling the $\Lambda_b^0 \to pK\mu^+\mu^-$ signal yield by the ratio of $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ yields (it is impossible to predict precisely the expected yields under the hypothesis R_{pK}^{-1} =1, as the efficiencies of the $\Lambda_b^0 \to pKe^+e^-$ mode are blind as well). An example of such a fit for each trigger category is shown in Figs. 3.77, 3.78.

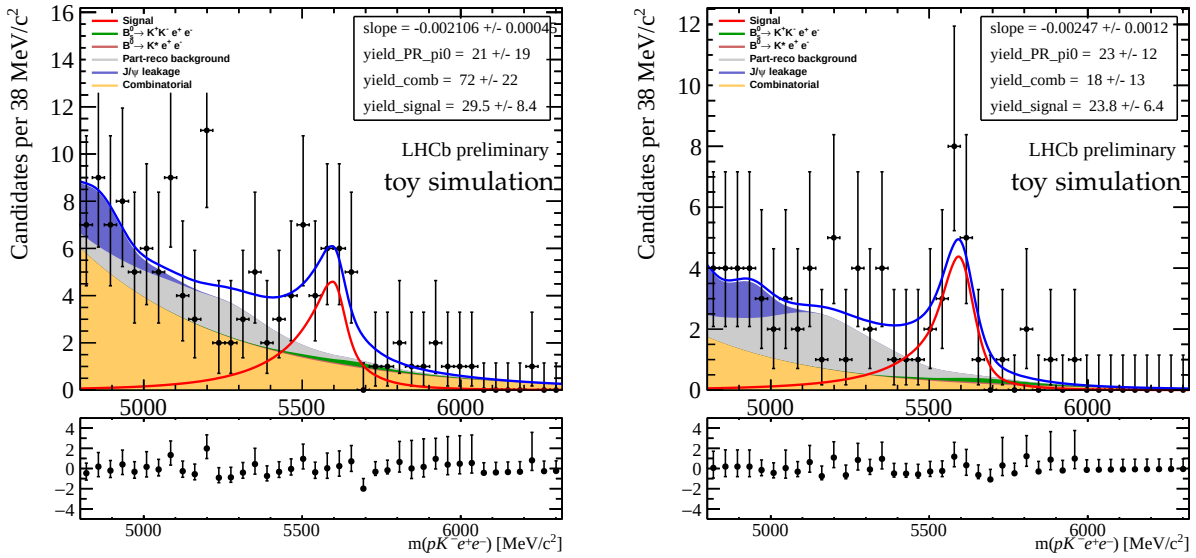


Figure 3.77: Example fits to the pseudo-data generated according to the Run I fit model in the (left) L0I and (right) L0E! trigger category.

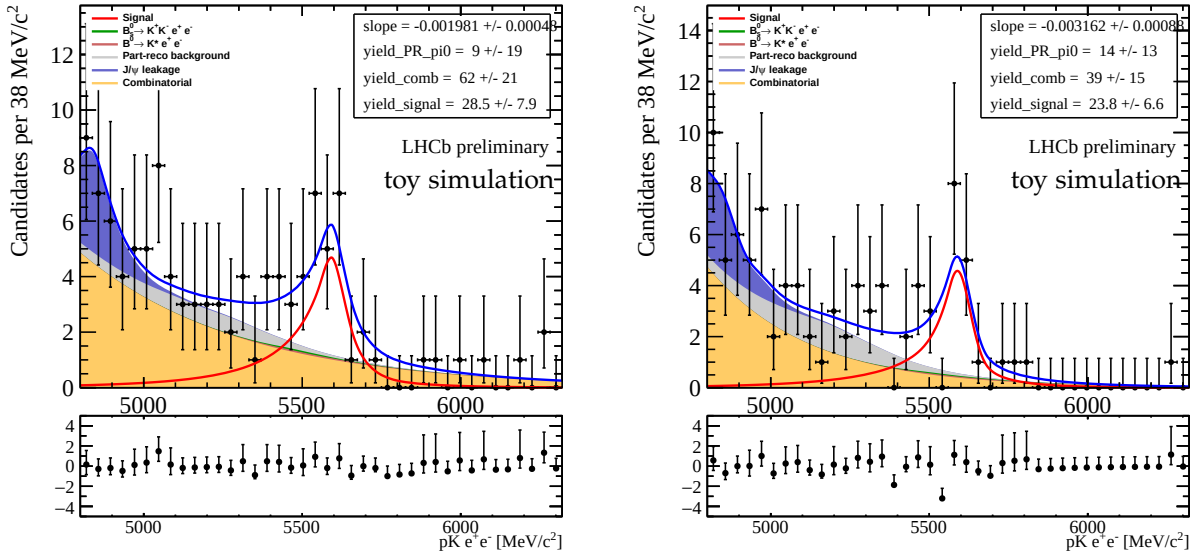


Figure 3.78: Example fits to the pseudo-data generated according to the Run II fit model in the (left) L0I and (right) L0E! trigger category.

However, given that the efficiencies are blinded, and the value of R_{pK}^{-1} is unknown, it is useful to consider several signal yield hypotheses, and validate the fit behaviour. Studies have been performed with signal yields hypotheses of 20, 35 and 50 candidates (Figs. 3.79–3.81). It can be seen that the signal yields have a very stable and rather unbiased behaviour, showing a good coverage of uncertainties. This is not always the case for the

background components. It can be seen that the pulls of the combinatorial yield have a tail to the left, resulting in a tail to the right in the pulls of partially reconstructed background yields. This is explained by the fact that the statistics in the upper sideband is tiny, making it very difficult for the fit to estimate the slope of an exponential in certain cases. This results in a distortion seen in the distribution of the slope pulls, and biases the yields of the combinatorial background. As long as this issue does not have a visible impact on the signal yields behaviour (as the signal pull looks good), it is not considered as a significant problem.

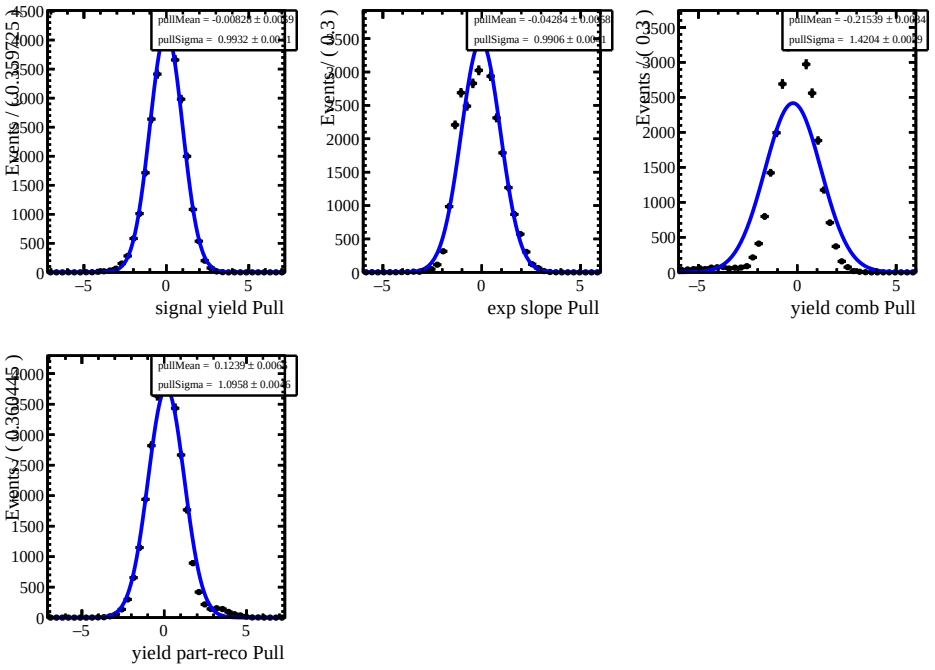


Figure 3.79: Pulls of the free fit parameters, with 20 signal candidates used as a hypothesis. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential shape of the combinatorial background, the yields of the combinatorial background and the yields of the partially reconstructed background.

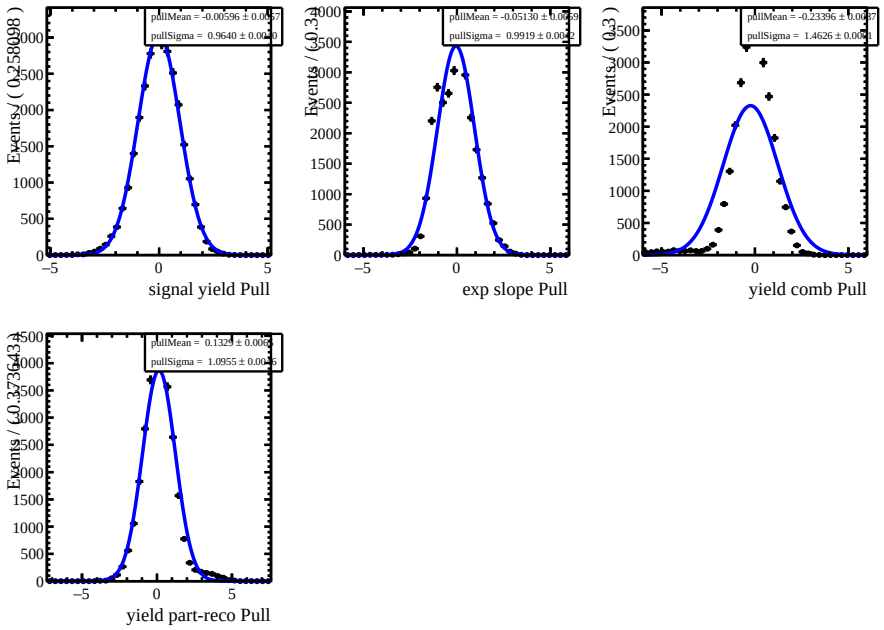


Figure 3.80: Pulls of the free fit parameters, with 35 signal candidates used as a hypothesis. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential shape of the combinatorial background, the yields of the combinatorial background and the yields of the partially reconstructed background.

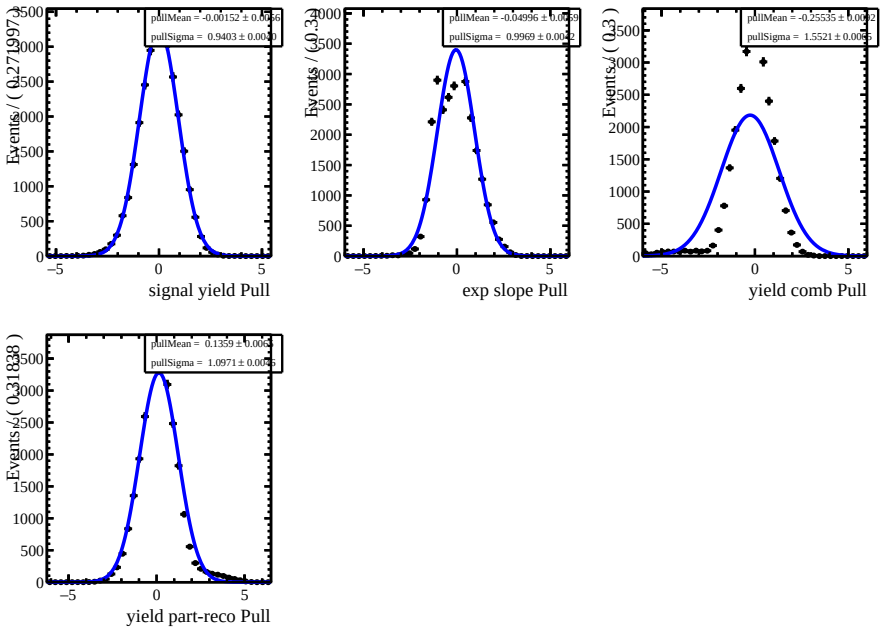


Figure 3.81: Pulls of the free fit parameters, with 50 signal candidates used as a hypothesis. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential shape of the combinatorial background, the yields of the combinatorial background and the yields of the partially reconstructed background.

Decay	Signal	Comb.	Hadronic misidentification	Other
$\Lambda_b^0 \to pK\mu^+\mu^-$	CB	exp.	$B^0 \to K^{*0}\mu^+\mu^-$ $B_s^0 \to K^+K^-\mu^+\mu^-$	
Free parameters	yields, width, mean	yields, slope		
Fixed parameters	tails		(Y. of $B^0 \to K^{*0}\mu^+\mu^-$)/(Y. of $\Lambda_b^0 \to pK\mu^+\mu^-$) (Y. of $B_s^0 \to K^+K^-\mu^+\mu^-$)/(Y. of $B^0 \to K^{*0}\mu^+\mu^-$)	
Derived parameters			Y. of $B^0 \to K^{*0}\mu^+\mu^-$ Y. of $B_s^0 \to K^+K^-\mu^+\mu^-$	
$\Lambda_b^0 \to pKe^+e^-$	CB + 2 Bifur. CB	exp.		PR , J/ψ-leakage
Free parameters	yields	yields, slope		Y. of PR
Constrained param's				Y. of J/ψ-leakage
Fixed parameters	tails, means widths, fractions		(Y. of $B^0 \to K^{*0}e^+e^-$)/(Y. of $\Lambda_b^0 \to pKe^+e^-$) (Y. of $B_s^0 \to K^+K^-e^+e^-$)/(Y. of $B^0 \to K^{*0}e^+e^-$)	
Derived parameters			Y. of $B^0 \to K^{*0}e^+e^-$ Y. of $B_s^0 \to K^+K^-e^+e^-$	

Table 3.21: Summary of the nominal fit configurations to invariant mass distributions used on data for the rare $\Lambda_b^0 \to pK\ell^+\ell^-$ decay modes. All the hadron misindentified and partially reconstructed backgrounds are modelled with `RooKeysPdf`, their invariant mass shapes are always taken from simulation. "Y." is used to refer to the yields of each mode. "PR" refers to partially-reconstructed background $\Lambda_b^0 \to pKe^+e^-\pi^0$.

3.8 $r^{-1}_{J/\psi}$ cross-check

Although R^{-1}_{pK} is measured as a double ratio, which implies a large cancellation of numerous systematic effects and also requires only a relative efficiency control, it can be considered as important to reassure ourselves and have a strong control of the absolute efficiencies. This is crucial, given that the differences in the detection and reconstruction of the muons and the electrons in the apparatus. So, as a first step, one needs to ensure a good control of the relative total efficiencies for muon and electron modes, in various regions of the phase-space. These cross-checks are presented in the first part of this section. The second part is dedicated to the extraction of the main results of this work.

One stringent test of the control on efficiencies is the so-called $r^{-1}_{J/\psi}$ ratio, defined by the Equation 3.3.

As it has been already discussed in Sec. 3.1, the idea behind this check is the observed universality of J/ψ decays to dilepton final states. Consequently, this ratio is assumed to be strictly equal to unity, independently of the region of the phase-space where it is measured. It should not show significant trends as a function of any kinematic, geometric or other variables.

In practice, Equation 3.3 can be rewritten in terms of measured numbers of candidates (denoted as N) and efficiencies (ϵ), so it is in fact the ratio of efficiency-corrected yields:

$$r^{-1}_{J/\psi} = \frac{N(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))}{N(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} \times \frac{\epsilon(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}{\epsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))} \tag{3.12}$$

As it has already been mentioned in the introduction regarding the construction of R^{-1}_{pK}, in this analysis the overall statistical uncertainties on the observable of interest are driven by the limited yields of the electron mode. Therefore to ensure that the uncertainties are Gaussian (symmetric), it was decided to invert the measured ratio. Hence, what is measured is $r^{-1}_{J/\psi}$.

The following tests related to the measurement of $r^{-1}_{J/\psi}$ are performed:

- Comparison of the central values (integrated over the whole phase-space) of $r^{-1}_{J/\psi}$ between different data-sets and trigger categories, as shown in the Section 3.8.3;
- Trends of $r^{-1}_{J/\psi}$ as a function of various variables, as a check of the flatness, presented in the Section 3.8.6;

One can also exploit the less abundant $\psi(2S)$ charmonium resonance as a check of the efficiency control in a q^2 region different than the J/ψ one. To do so, a double ratio $R^{-1}_{\psi(2S)}$ is measured (Sec.3.8.8), it is constructed in an analogous to the R^{-1}_{pK} observable:

$$R^{-1}_{\psi(2S)} = \frac{BR(\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-))}{BR(\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-))} \times \frac{BR(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}{BR(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))} \tag{3.13}$$

Finally, one can also investigate the branching ratio of $\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$, and check whether it is consistent between the data-sets (Sec. 3.8.8), and with the published measurement [87].

3.8.1 A note on the blinding of the cross-checks

As it has been discussed in Sec. 3.1.5, the values of $r_{J/\psi}^{-1}$ tests are kept blind in order to avoid potential biases when evaluating the corrections to the simulation and efficiencies. This is achieved by scaling the final value by an unknown randomly generated factor, which is kept constant across various data-taking periods and trigger categories, to nevertheless allow for comparisons and consistency tests. As a consequence, as the signal yields are not blind for the $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ modes, it was decided not to look at the efficiency-corrected yields for the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode. So, the efficiencies of the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode are blinded. This leaves the freedom to compare the relative efficiencies of the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ in different trigger categories, which is an important cross-check of the control over trigger efficiencies.

The $R_{\psi(2S)}^{-1}$ test is not blinded, as it is computed by means of a double ratio, so it is less sensitive to potential biases of efficiencies. In addition, it is statistically limited.

The branching ratio of the decay $\Lambda_b^0 \to pK\mu^+\mu^-$ was not measured before, but is not kept blind in this analysis since the decay was already observed by a different group of analysts. This is not the case for the electron mode which is kept blind at both fit and efficiency levels.

3.8.2 A note on the `nSPDHits` cut efficiencies

According to Equation 3.12, one has to know the absolute efficiencies of both $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ channels, or, at least, their ratio. This knowledge can be achieved for most of the selection requirements using the simulation samples or data-driven techniques. However, in certain cases it is impossible to estimate this efficiency precisely.

In the L0 trigger, as well as in the Stripping, a cut on the `nSPDHits` variable is applied. This variable, describing the number of hits in the SPD subdetector, is a proxy for the multiplicity in the event. It is used in the L0 trigger due to the fast response of the SPD detector, allowing to take quick decisions. The limitation of the event multiplicity is required, as the processing time of high-multiplicity events is larger than acceptable in the software trigger. Unfortunately, the `nSPDHits` variable is known to be poorly modeled in the simulation, so the efficiency of the cut applied in the trigger cannot be accurately estimated directly from the simulation. Also, data cannot be used to reweight the simulation above the cut value, as only the data which passed the L0 trigger is saved to disk. Any indirect methods to estimate the efficiency of this cut will lead to unnecessary large uncertainties. As a consequence, the efficiency of this cut is not taken into account in the current analysis.

In the double ratios, the efficiency of this cut is expected to fully cancel, as the distributions of the nSPDHits variable do not depend on the q^2 region. For the same reason, it is also expected to cancel in the measurements of the branching ratios of the rare modes, when normalising them to the resonant channels. However, when it comes to computing ratios such as $r_{J/\psi}^{-1}$, the efficiency of the cut on nSPDHits is not expected to cancel completely in the ratio since there are some small differences between the muon and electrons modes in the nSPDHits distributions as shown in Figure. 3.82. This difference is seen in both data and simulation. The origin of this behaviour is probably related to the secondary activity, created by electrons in the calorimeter.

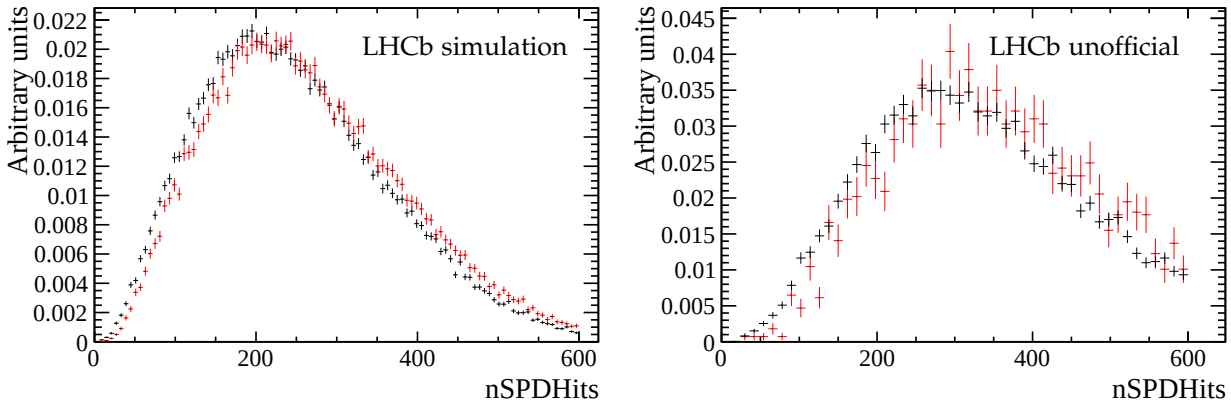


Figure 3.82: Comparison of the nSPDHits distributions between the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ (black) and $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ (red) channels in Run I, in the (left) simulation not corrected by event multiplicity; and (right) data with a full selection applied and a ± 35 MeV/c^2 cut around the J/ψ-constrained Λ_b^0 mass.

A first rough estimate can be obtained by applying a tighter nSPDHits < 450 cut and estimating its efficiency on top of the existing one nSPDHits < 600 in Run I datasets. In the simulation, the efficiency of the cut nSPDHits < 450 is 0.942 ± 0.003 for the muon channel, and 0.925 ± 0.004 for the electron channel; while in the data it is 0.82 ± 0.01 for the muon channel and 0.79 ± 0.02 for the electron channel. It can be seen that on average, this gives about two percent difference in the nSPDHits cut efficiency between the electron and muon modes, but, as it was explained above, it is impossible to estimate this difference precisely. The implication of this observation is that the expected value of the $r_{J/\psi}^{-1}$, having this efficiency unaccounted, is rather about 0.98, than strictly 1. Being unable to correct precisely for this effect, we assign a systematic uncertainty of two percent on the $r_{J/\psi}^{-1}$ measurement in each trigger category (Sec. 3.8.4). This uncertainty is assigned only to the $r_{J/\psi}^{-1}$ measurement.

3.8.3 Integrated $r^{-1}_{J/\psi}$ test

The test is performed separately for each of the trigger categories, and each data taking period Run I/Run II. For what concerns the muon mode, only one trigger category (L0M) is used. For the electron mode, two default categories are used (L0I and L0E!), and in addition for the cross-check the inclusive L0E is considered. This check is motivated by the fact that both default categories use the L0TIS trigger correction, however a comparison to the inclusive L0E allows to make sure there is no significant issues with this trigger correction. The notation in this section represents the trigger category of the electron mode, as the muon one is always the same.

The procedure described below is identical between the datasets, with the only difference coming in selection requirements and weights applied.

1. The $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ generator-level simulation sample, with the relevant weights applied (see Sec. 3.5), is used to extract the number of simulated events at the generator stage. No cuts are applied to this sample: this sample is used to estimate the initial number of events before applying selection requirements. However, only events which pass in reconstructed-level sample the requirement $\texttt{nSPDHits} < 600(450)$ in Run I (Run II), are kept in the generator-level sample, due to the reasons discussed in Sec. 3.8.2.

2. The $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ reconstructed-level simulation sample, passed through the complete chain of selection requirements, and having all the relevant weights applied, is used to extract the number of simulated events after the final selection.

3. The number of reconstructed-level events (step 2) is divided by the number of generated-level events (step 1), and multiplied by the values of generating and filtering efficiencies, to get the total efficiency.

4. For Run I, the 2011 and 2012 datasets have to be merged. To do so, their simulation samples are merged, having a dedicated weight applied ensuring the correct ratio of integrated luminosity (1:2) between the two datasets. In addition, the ratio of Λ_b^0 production cross-sections at 8 and 7 TeV is considered as 1.23, as per LHCb measurement [86]. Relevant values of generation and filtering efficiencies are passed to each of the two. Steps 1 and 2 are performed with already merged generator-level and reconstruction-level samples.

5. The simulation samples of the specific backgrounds, with the final selection applied, are used to extract their invariant mass shapes (see Sec. 3.7 here and after).

6. The $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ simulation sample is used to extract the signal invariant mass shape.

7. The fit is performed to the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data, with the final selection applied. The signal yield and its uncertainty are extracted from the fit.

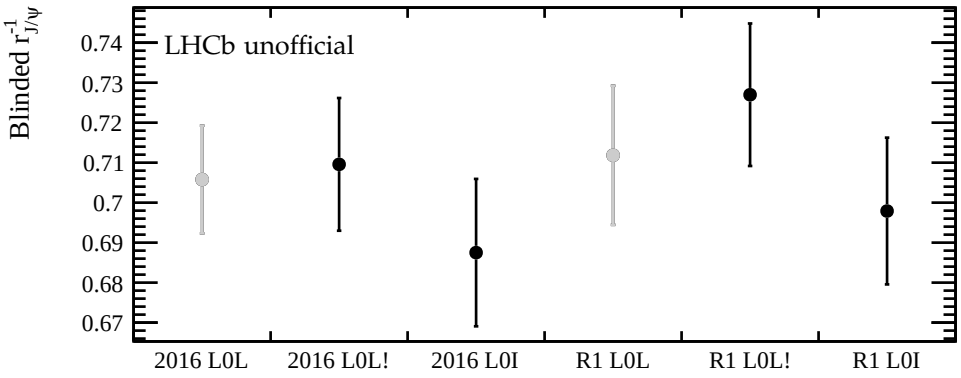


Figure 3.83: Comparison of blinded $r_{J/\psi}^{-1}$ values in different electron trigger categories and datasets. The blinding factor is identical between the datasets. Black points denote the trigger categories used for this analysis, grey points – used only as a cross-check.

8. The signal yield is divided by the total efficiency (step 3) to obtain the efficiency-corrected yield.

9. All the previous steps are repeated for the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode.

10. Efficiency-corrected yield of the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ is divided by the efficiency-corrected $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ yield to get the value of $r_{J/\psi}^{-1}$.

11. The value of $r_{J/\psi}^{-1}$ is multiplied by a blinding factor and printed out.

Comparisons of $r_{J/\psi}^{-1}$ between trigger categories and datasets are shown in Fig. 3.83 and in the Table 3.22. Uncertainties are only those coming from the fit result (*i.e.* statistical). The Table 3.22 also features the values of $r_{J/\psi}^{-1}$ if no corrections are applied to the simulation samples, and the relative difference in percents. This allows to understand the importance of these corrections for a single ratio. In particular, the trigger and PID corrections affect significantly the corresponding efficiencies. At the same time, the kinematic reweighting and decay model weights affect the efficiency of the kinematic cuts.

The systematic uncertainties on $r_{J/\psi}^{-1}$ are discussed in the Sec. 3.8.4.

3.8.4 $r_{J/\psi}^{-1}$: systematic uncertainties

Let's assume the nominal value of a given observable R computed by the standard procedure is $\overline{R}$. To estimate a systematic uncertainty on R due to a certain source of potential bias s, the value of R is computed in n alternative ways, resulting in values R_i ($1 < i \leq n$). The systematic uncertainty on R due to this source of bias is then estimated as the square root of variance on R:

Table 3.22: Comparison of blinded $r_{J/\psi}^{-1}$ values in different electron trigger categories and datasets. Uncertainties provided are statistical only.

Category	Value with nominal corrections	Value without corrections
Run I, L0E/L0M	0.712 ± 0.017	0.485 ± 0.006 (-31.8%)
Run I, L0E!/L0M	0.727 ± 0.018	0.505 ± 0.007 (-30.4%)
Run I, L0I/L0M	0.698 ± 0.018	0.547 ± 0.013 (-21.5%)
Run II, L0E/L0M	0.706 ± 0.013	0.492 ± 0.005 (-30.3%)
Run II, L0E!/L0M	0.710 ± 0.017	0.522 ± 0.008 (-26.4%)
Run II, L0I/L0M	0.688 ± 0.018	0.526 ± 0.010 (-23.7%)

$$\sigma_s^R = \left(\frac{1}{n} \sum_{i=1}^{n} \left(R_i - \overline{R} \right)^2 \right)^{\frac{1}{2}} \tag{3.14}$$

The list below covers what is considered to be the dominant sources of systematic uncertainties, as well as the procedure used to estimate each one of them. The resulting values are then summarised in Table 3.23.

- **Decay model**: To assign this uncertainty, the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ candidates selected in data (to be precise, the ${}_s\mathcal{P}lot$ distribution described in Sec. 3.5.3) is used to extract the one-dimensional $m(pK)$ spectrum from data. This spectrum is then used to reweight the $m(pK)$ distribution in the simulation at both reconstructed- and generator-level. The difference in $r_{J/\psi}^{-1}$ between the default method and this one, is assigned as a systematic uncertainty. It should be noted that the default method is based on an amplitude model developed on the Run I data, and does not necessary describe the Run II dataset very well. Consequently, the related uncertainty is smaller in Run I than in Run II.

- **Statistics of samples used to extract the corrections**: The corrections discussed in Section 3.5 and applied to the simulation, have uncertainties related to the size of the data and simulation samples used for their extraction, as well as the size of the PID calibration samples. To account for these statistical uncertainties, the `bootstrapping` technique is used: all the data and simulation candidates are assigned a Poisson-distributed weight of unity mean; all the corrections and efficiencies are re-evaluated. The PID efficiencies are oscillated around their mean assuming Gaussian uncertainties. This procedure is repeated a 100 times, resulting in a 100 alternative efficiency values. These efficiencies are then combined with the nominal yields (the fits to the data are not re-performed) to obtain a 100 alternative $r_{J/\psi}^{-1}$ values. It should be noted that this uncertainty on $r_{J/\psi}^{-1}$ is correlated with the statistical uncertainty. This systematic uncertainty is always larger in the L0E! category than in the L0I one, as the L0Electron correction is the only one which is extracted from the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ dataset which has a smaller statistical power.

- **Binning of the corrections**: Binned weights are used instead of interpolated ones in order to estimate the uncertainty related to the binning of the weights.The effect is larger in Run I as a smaller datasets (when considering 2011 and 2012 separately) were available to compute these corrections.

- **Multiplicity proxy**: As it has already been discussed, the `nTracks` variable is chosen as a default proxy for the multiplicity. Four additional choices are considered to estimate the relevant systematic uncertainty: a) the `nSPDHits` variable, b) the `nVeloTracks` variable, c) the `nPVs` variable, d) no multiplicity correction at all.

- **PID procedure:**

  - **Proton**: The three-dimensional binning (in $p:\eta$:`nTracks`) is compared to the default two-dimensional binning used in Section 3.5.2, and the difference in the $r^{-1}_{J/\psi}$ value is taken as a systematic uncertainty. No uncertainty is assigned due to the usage of an ${}_s\mathcal{P}lot$ method since it is expected to cancel in the ratio.
  - **Kaon**: Three-dimensional binning (in $p:\eta$:`nTracks`) is compared to the default two-dimensional binning, and the difference in the $r^{-1}_{J/\psi}$ value is taken as a systematic uncertainty. No uncertainty is assigned due to the usage of an ${}_s\mathcal{P}lot$ method since it is expected to cancel in the ratio.
  - **Muon**: A systematic uncertainty of 0.1% is assigned per each muon, accounting for an ${}_s\mathcal{P}lot$ procedure used to extract the weights, following the recommendations of the LHCb PID group. Two alternative binning schemes are compared.
  - **Electron**: The maximal difference in the efficiency obtained by the default fit-and-count method, and ${}_s\mathcal{P}lot$ approach, amounting in 2 percent, is assigned as a systematic uncertainty (see Fig. 2.25). Two alternative binning schemes are compared.

- **Trigger category bias on kinematic correction**: The kinematics weights (Sec. 3.5.4) are by default computed using the L0M trigger category. As the kinematic correction is performed before the trigger one, the effect of data-simulation disagreement in the trigger description can introduce a little bias in the kinematic weights. To assign a related uncertainty, the corrections are recomputed using either L0I (on the $\Lambda^0_b \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset) or L0E (on the $\Lambda^0_b \to pKJ/\psi\,(\to e^+e^-)$ dataset) categories, and applied in a symmetrical manner to both electron and muon modes, to both generator- and reconstructed-level simulation samples. Two resulting $r^{-1}_{J/\psi}$ values are used to estimate the associated systematic uncertainty. The effect seems to be larger in Run II, that can be explained by the fact that L0Muon trigger corrections (Sec. 3.5.6.a) are very close to one for Run I, but can be quite large for Run II.

- **L0 Trigger**: As discussed in Sec. 3.5.6.a, the TISTOS method is used to evaluate the L0 trigger corrections to the simulation. To estimate the associated systematic uncertainty, the "TIS tag requirement" can be changed, and an impact on the $r^{-1}_{J/\psi}$ value studied. This is done in the following way for each of the trigger categories:

- **L0Muon**: The "TIS requirement" used for the TISTOS method when computing the weights, is changed from `L0Global_TIS` to `L0Hadron_TOS` on the proton or kaon tracks;
- **L0Electron**: The "TIS requirement" is changed from `L0Global_TIS` to `L0Hadron_TOS` on the proton or kaon tracks;
- **L0TIS**: On one hand, the "TIS requirement" is changed from `L0Global_TOS` to `L0Muon_TOS` (and the relevant weights are evaluated using the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset). On the other hand, the "TIS requirement" is changed from `L0Global_TOS` to `L0Electron_TOS` (and weights are evaluated on the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ sample).

These alternative weights for the L0TIS case were already compared in Fig. 3.50, and a reasonable agreement within uncertainties is observed between four methods. To gain additional statistical power in the electron mode, no binning in `nTracks` is used.

- **Decay time corrections**: in Run I, additional corrections were applied to the decay time distribution (Sec. 3.5.5). The effect of not applying these small corrections, on $r_{J/\psi}^{-1}$ value, is considered as a systematic uncertainty. In Run II, such corrections were not applied, and applying them is leading to a difference smaller than 0.01% in $r_{J/\psi}^{-1}$, therefore no uncertainty is applied.

- **Residual corrections** (or their absence): As described in Sec. 3.5.7, there are only minor residual data-MC differences in the important variables. To estimate a related uncertainty, a correction is applied for the proton p_T and χ^2_{IP} , and the resulting difference in the $r_{J/\psi}^{-1}$ value is assigned as the systematic uncertainty. It can be seen that the effect is larger in Run II. This is correlated with the fact that the decay model used to reweight the simulation is data-driven, and has been developed using (statistically-limited) Run I dataset. It is expected that this same model is not describing that well the statistically independent Run II dataset, and this effect propagates to the proton p_T distribution.

- **HLT**: The maximal (between the muon and electron sides) statistical uncertainty on the HLT correction (Sec. 3.5.6.b) is taken as the systematic uncertainty.

- **MVA**: As the data-MC agreement in the BDT variable is good (Sec. 3.5.7), no uncertainty is assigned.

- q^2 **window mass cuts**: The efficiency of the J/ψ mass window requirement in the electron sample might be mismodeled by the simulation. This effect is independent of the hadronic structure and so is similar to the one which has been observed in published R_K and $R_{K^{*0}}$ analyses. No dedicated study was performed so far, because it is difficult to obtain a clean signal sample of dielectron candidates from the data (and given that the statistics is not sufficient for a precise study), but by analogy with the $R_{K^{*0}}$ analysis, a 2% systematic uncertainty is assigned.

- **Fits**: The systematic uncertainties associated with the fit modeling, described in Section 3.7, can be related to the following procedures:
  - J/ψ **mass constraint**: The difference in the $r_{J/\psi}^{-1}$ values, obtained through invariant mass fits with the J/ψ mass constraint (Sec. 3.7.1) and without it, is assigned as a systematic uncertainty.
  - **Background model**: All the fixed ratios of background yields are allowed to float. The partially reconstructed $\Lambda_b^0 \to J/\psi\, pK\pi^0$ background component, omitted in the default fit, is allowed. The fit range is increased to accommodate it. The difference in the $r_{J/\psi}^{-1}$ computed using such fits, with respect to the the default values, is taken as the systematic uncertainty.

A summary of the values of each source of uncertainty is provided in Table 3.23.

Table 3.23: Systematic uncertainties on $r_{J/\psi}^{-1}$, %.

Type	Run I L0I	Run I L0E!	Run II L0I	Run II L0E!
Decay model	0.41	1.63	2.03	3.95
Corrections (statistics)	1.20	2.26	1.22	1.82
Corrections (binning)	2.94	3.22	0.48	0.35
Multiplicity proxy	3.25	1.62	3.76	4.37
Trigger bias (kin.corr.)	2.11	1.65	4.16	3.95
PID procedure	2.93	2.02	2.02	2.03
L0 procedure	0.75	3.20	1.98	1.53
HLT	3.80	3.00	1.00	1.50
Decay time corrections	0.17	0.37	0	0
Residual corrections	0.43	0.45	2.03	1.18
q^2 window	2	2	2	2
Fit (J/ψ constraint)	1.88	2.80	1.64	0.80
Fit (background model)	1.85	0.15	1.61	0.30
nSPDHits shape	2	2	2	2
Total systematic	8.00	7.94	7.98	8.52

3.8.5 Averaging the datasets

To get the average value of $r_{J/\psi}^{-1}$, we perform a simultaneous fit to the Run I and Run II $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ datasets, in the L0I and L0E! categories for electrons and L0M category for muons. The muon yields and $r_{J/\psi}^{-1}$ are free parameters of the fit, as well as the background parameters which were kept free in the per-category fits. The $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ yields are expressed in terms of the free parameters and the ratio of efficiencies. The latter is blinded by the same shift as in the previous section. In addition, the uncorrelated systematic uncertainties (discussed in detail in Sec. 3.8.4), are plugged

into the fit as Gaussian constraints on the efficiency ratios. Correlated uncertainties are added to the final result. We do not split the statistical and systematic uncertainties as they are correlated.

The resulting blinded value of $r_{J/\psi}^{-1}$, obtained from the simultaneous fit, is:

$$r_{J/\psi}^{-1}(\text{blind}) = 0.704 \pm 0.035, \tag{3.15}$$

where the total uncertainty is provided. It should be noted that the uncertainty is also scaled by the same blinding factor as the central value.

3.8.6 $r_{J/\psi}^{-1}$ as a function of important variables

In this section, the 'flatness' of the $r_{J/\psi}^{-1}$ in bins of various kinematic and geometric variables, important for this analysis, is checked. It should be noted that this check can only be performed if the variable has the same meaning for the electron and muon mode.

The procedure used to evaluate $r_{J/\psi}^{-1}$ in this section is similar to the one described above. Before proceeding through all the steps, an iso-populated binning (five bins) is computed in a variable of interest, using the $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ simulation sample with the full selection applied. In each bin, all the steps are repeated. All the fit shapes are recomputed as they depend on the kinematic requirements and the fits are re-performed.

In some corner cases, the fits may not behave in the best possible way, this is due to a few following reasons:

a) a lack of statistics in the simulation samples in a given bin might cause the invariant mass shapes to be slightly mismodeled;

b) the description of the fit model could be imperfect in some regions of phase-space, because of variations of the relative background composition;

c) the global assumption of the exponential shape for the combinatorial background might fail in certain regions of the phase space,

d) the signal statistics itself may not be sufficient, although this effect should be minimised by the choice the iso-populated binning.

This means some fluctuations are expected around the flat $r_{J/\psi}^{-1}$ behaviour even if the efficiencies are under the best control.

The $r_{J/\psi}^{-1}$ trends for Run I and Run II, L0I and L0E! (overlaid) are presented in Figs. 3.84-3.85. No significant trends are seen in any of the relevant variables.

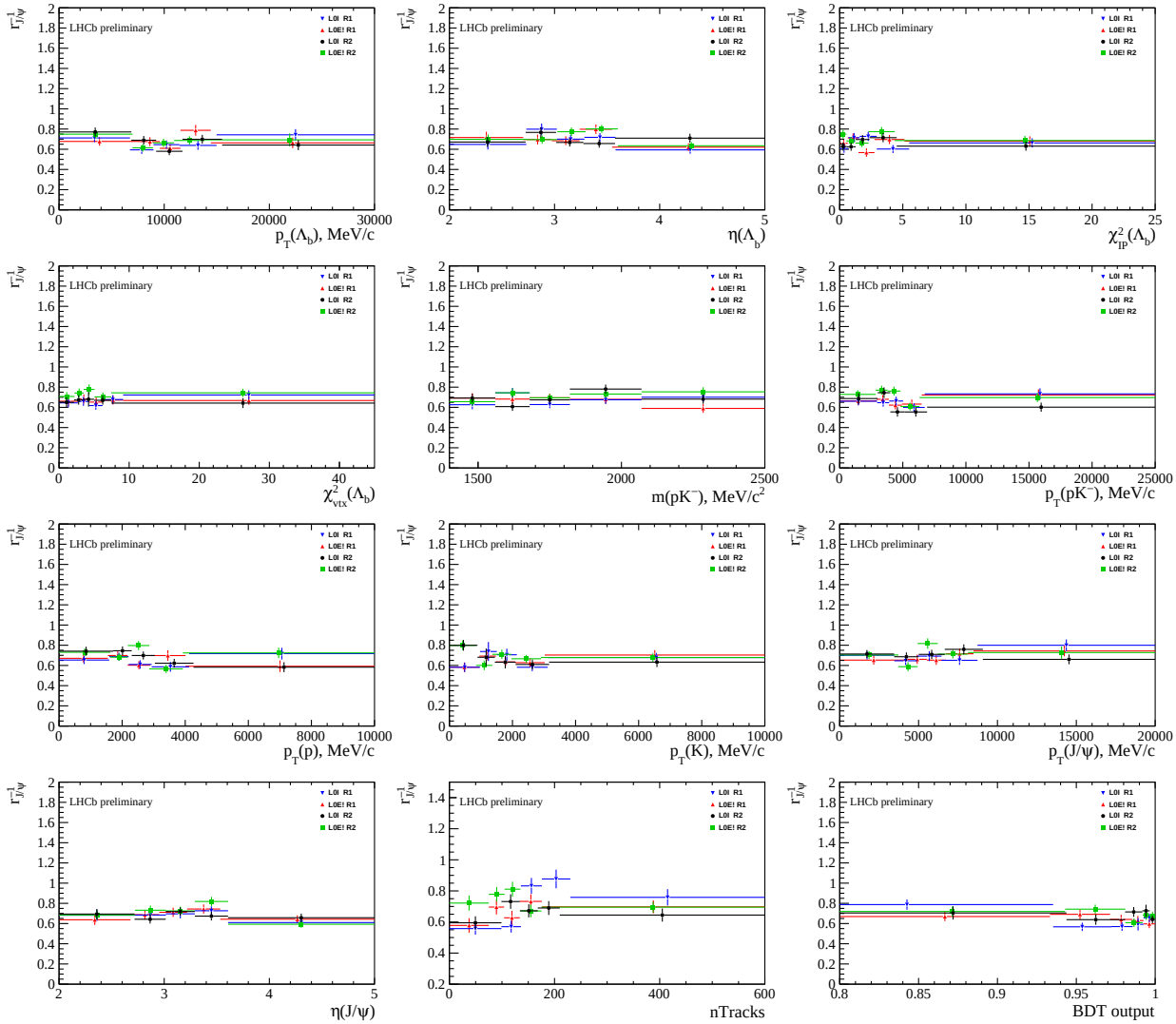


Figure 3.84: $r_{J/\psi}^{-1}$ trends in bins of various variables, for four different datasets. Central values are blinded.

3.8.7 Unblinding of $r_{J/\psi}^{-1}$

As it has been discussed above, the value of $r_{J/\psi}^{-1}$ has been blinded by scaling it with a random factor. After getting the green light to unblind by the review committee, this factor was found to be 0.731, which means the blinded result, obtained from the simultaneous fit to the data, and shown in Eq. 3.15, translates in the following result:

$$r_{J/\psi}^{-1} = 0.962 \pm 0.048, \tag{3.16}$$

where the total uncertainty is provided. This result falls within less than one standard deviation from unity.

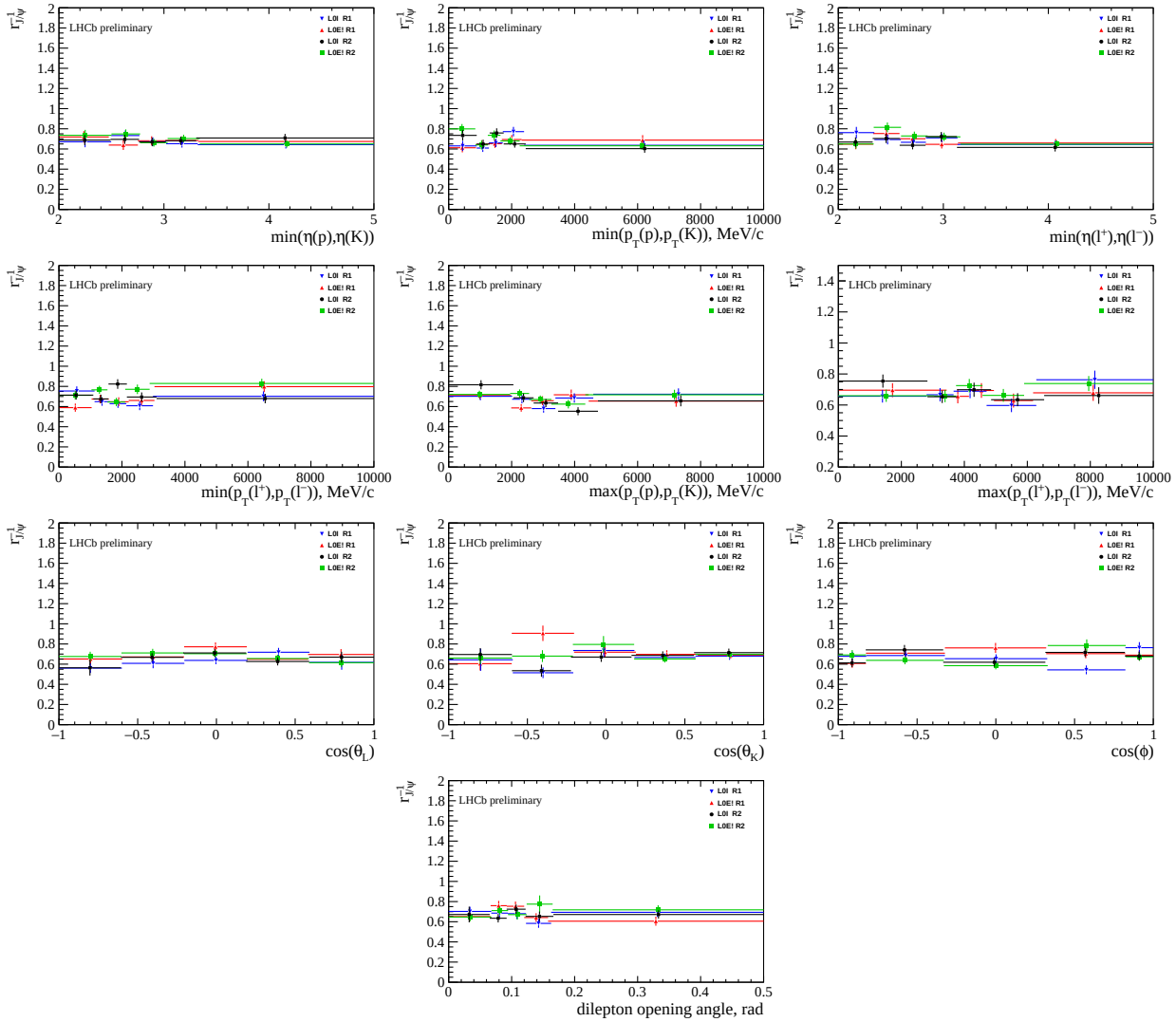


Figure 3.85: $r_{J/\psi}^{-1}$ trends in bins of various variables, for four different datasets. Central values are blinded.

Fig. 3.86 features the comparison of $r_{J/\psi}^{-1}$ values between the different datasets of interest. It can be seen that within the estimated uncertainties, all the values of $r_{J/\psi}^{-1}$ are in good agreement with unity. However, the central values are always slightly below unity – this can be explained by several systematic effects which are taken into account when assigning the systematic uncertainties (Sec. 3.8.4). For example, the unaccounted efficiency of `nSPDHits` cut which is applied in the hardware trigger, differs by about 2% between muon and electron modes, this "pushes" the expected $r_{J/\psi}^{-1}$ value down, as it was already explained in Sec. 3.8.2. The imprecise modeling of the efficiency of dielectron mass window requirement on the J/ψ meson, is also associated to a 2% systematic uncertainty as discussed in Sec. 3.8.4, this also "pushes" the expected value of $r_{J/\psi}^{-1}$ value down. Finally,

it should be noted that these effects are correlated between datasets.

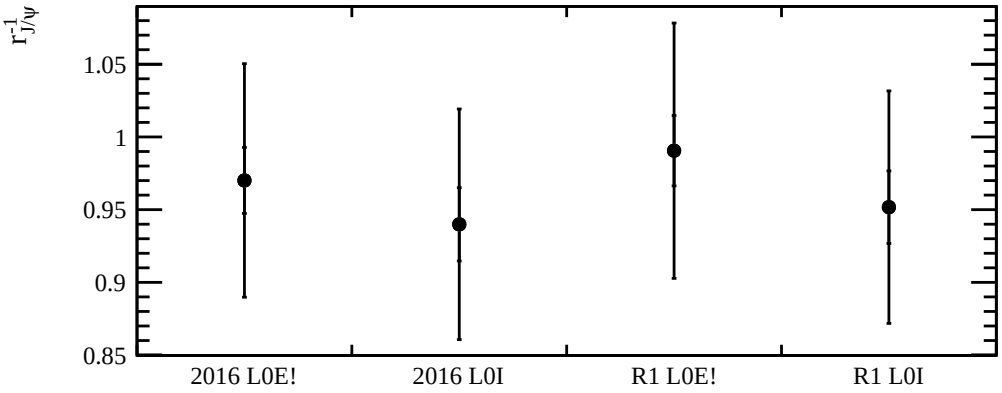


Figure 3.86: Comparison of unblinded $r^{-1}_{J/\psi}$ values in different trigger categories and datasets, with total and statistical-only uncertainties displayed.

As a conclusion, $r^{-1}_{J/\psi}$ is found to agree with unity, and systematic effects studied and assigned to this value are under control.

3.8.8 $R^{-1}_{\psi(2S)}$ cross-check

Similarly to R^{-1}_{pK}, $R^{-1}_{\psi(2S)}$ is constructed as a double ratio, and is computed in two steps: first, single ratios $\mathcal{B}\,(\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-))/\mathcal{B}\,(\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-))$ are computed for muon and electron modes, and then the double ratio is obtained dividing the two. The procedure for dealing with the efficiencies and fits is similar to the one described for the integrated $r_{J/\psi}$ measurement. To obtain $\mathcal{B}\,(\Lambda_b^0 \to pK\psi(2S)(\to \ell^+\ell^-))/\mathcal{B}\,(\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-))$, one has to know $\mathcal{B}(\psi(2S) \to \ell^+\ell^-)$ and $\mathcal{B}(J/\psi \to \ell^+\ell^-)$. For these branching fractions, lepton universality is assumed, and the most precise value among dimuon and dielectron decays is taken. Results are presented in Fig. 3.87.

In the Fig. 3.88, one can see the comparison of single ratio of branching fractions $\mathcal{B}(\Lambda_b^0 \to pK\psi(2S))/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi\,)$ in different datasets. In addition, the reported LHCb result based on the Run I dataset, exploiting also the $\Lambda_b^0 \to pK\psi(2S)$ decay with both $\psi(2S) \to \mu^+\mu^-$ and $\psi(2S) \to J/\psi\,\pi\pi$ decays from Ref. [87], is shown as a blue band. It can be seen that our results are consistent with, but slightly higher than the published result, nevertheless, the double ratio is consistent with unity within one standard deviation.

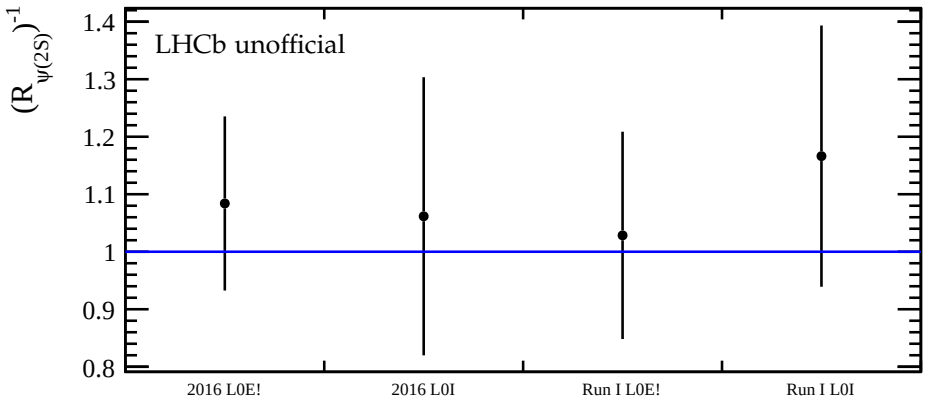


Figure 3.87: Comparison of the $R^{-1}_{\psi(2S)}$ values in different trigger categories and datasets.

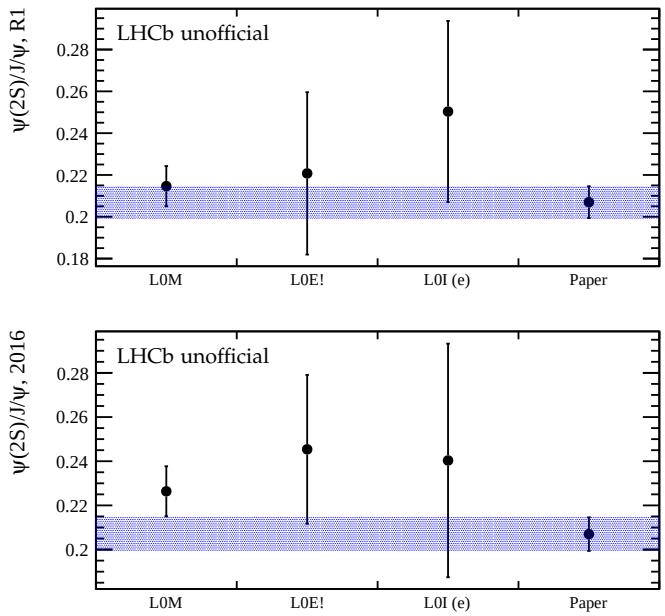


Figure 3.88: Comparison of the $\mathcal{B}(\Lambda_b^0 \to pK\psi(2S))/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ values in different trigger categories and datasets (top – Run I, bottom – Run II).

3.9 Towards the observation of the decay $\Lambda_b^0 \to pKe^+e^-$

In order to establish the observation of the $\Lambda_b^0 \to pKe^+e^-$ decay mode, one can merge all the electron datasets and perform a fit to the $\Lambda_b^0 \to pKe^+e^-$ invariant mass. Such an approach allows to validate the background model on a higher-statistics dataset, and isolate any potential issues with the fit model prior to unblinding the per-category fits, used in the following section to measure the R_{pK}^{-1} ratio. This also simplifies the computation of the signal significance. As a conclusion, in this section both Run I and II $\Lambda_b^0 \to pKe^+e^-$ datasets, as well as both trigger categories, are merged.

3.9.1 Invariant mass fit (blinded)

The fit model used towards the observation of the $\Lambda_b^0 \to pKe^+e^-$ decay is very similar to the one introduced in Sec. 3.7.4. The signal shape is 'frozen' by fixing its mean to a value obtained from the simulation, and its width to a value also obtained from the simulation but scaled by the factor of 1.15, as it was already discussed. The free fit parameters include the yields of the signal, of the combinatorial background and of the partially reconstructed background. The yields of the J/ψ leakage component is Gaussian-constrained to the estimate of an expected yield derived from the knowledge of the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ in data, and the ratio of efficiencies.

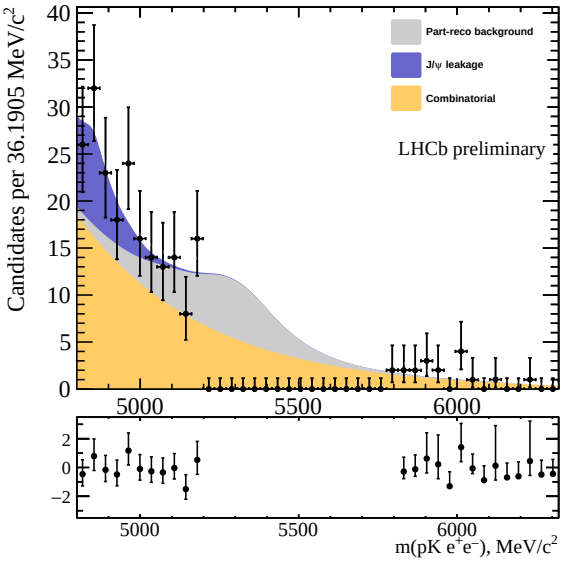


Figure 3.89: Blinded fit to the $\Lambda_b^0 \to pKe^+e^-$ invariant mass in data, for the merged datasets.

3.9.2 Pseudo-experiments with various signal yield hypotheses

Before unblinding the $\Lambda_b^0 \to pKe^+e^-$ invariant mass distribution, it is important to asses the behaviour of the final fit under various signal yield hypotheses, and to check the fit

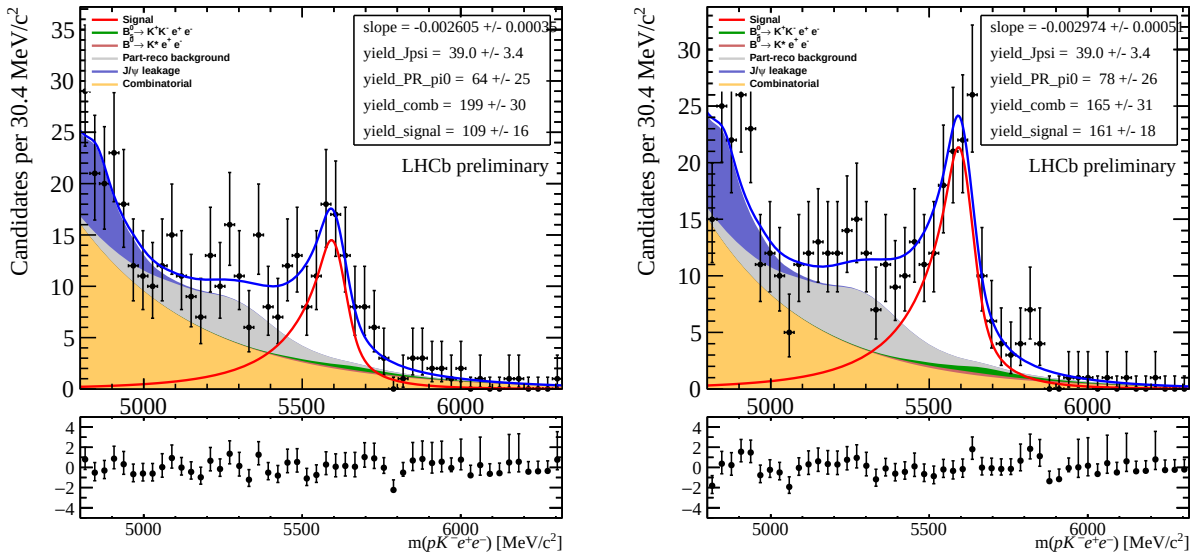


Figure 3.90: Example fit to the pseudo-data containing (left) 100 and (right) 150 generated signal candidates.

stability under such hypotheses.

Pseudo-experiments are generated according to this model, adding some amount of signal candidates close to the expected signal yield (100 or 150), as it is shown in Fig. 3.90. The output of the blinded fit shown in Fig. 3.89 is used to estimate the yields of specific backgrounds in data, and these yields are then generated for these toy studies. The pull distributions of the free fit parameters were checked, in general they confirm a reasonable fit behaviour. An example of the toy study for the case of 100 signal events is presented in Fig. 3.91.

In addition, the correlation matrix between free fit parameters has been computed by averaging the correlation matrices of 100 pseudo-experiments (with a signal yield hypothesis of 100 events). It is shown in Fig. 3.92. It can be seen that there is a significant anti-correlation between the yields of combinatorial and partially reconstructed backgrounds. In other words, the fit is not always able to distinguish the two. However, what is reassuring is that the correlation of the background yields with the yield of the signal component is much smaller.

3.9.3 Test of the fit stability with respect to the background model

While a lot of effort has gone into understanding in details the background composition of the $\Lambda_b^0 \to pKe^+e^-$ data, one can assume that background model used in this fit is still approximate, and might not be a perfect reflection of the background data. In order to estimate potential misbehaviours, the strongest test that one can perform is to validate the fit using pseudo-experiments. 100 000 toy datasets are generated according to the default fit model, including all the pieces discussed earlier, *i.e.* the signal, misID, combinatorial,

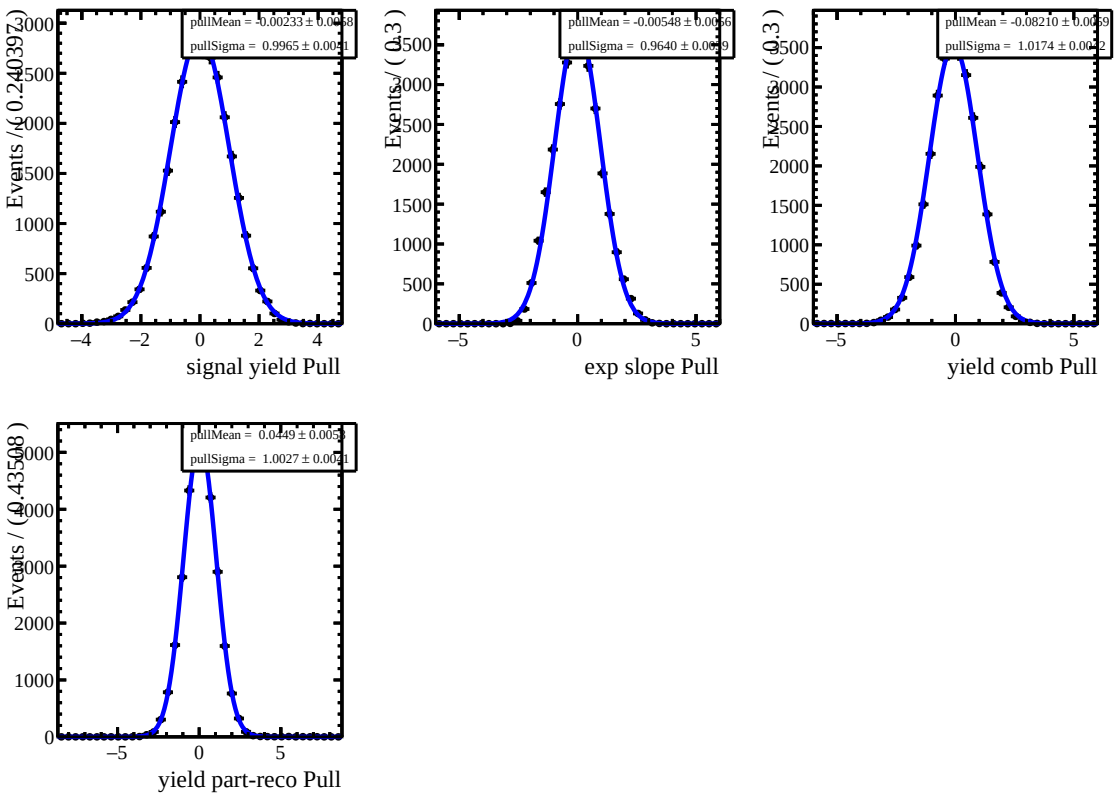


Figure 3.91: Pull distributions for the fit to the pseudo-data containing 100 generated $\Lambda_b^0 \to pKe^+e^-$ signal candidates, using 100000 pseudo-experiments. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential shape of the combinatorial background, the yields of the combinatorial background, and the yields of the partially reconstructed background.

partially reconstructed backgrounds, and the J/ψ-leakage component. Subsequently, each dataset is fitted using the simplest fit model that can be constructed, which includes only the signal and an exponential-like background shape ("combinatorial"). This fit model has only three free parameters: the yields of the signal and the background components, and the slope of the exponential. The pull distributions of these three parameters are presented in Fig. 3.93. It can be seen that even with such a drastic simplification of the fit model, only the yields of the background components are biased. However, the important conclusion from this study is that the pull for the signal yield is at the level of 0.06σ (6% of the statistical uncertainty). Similar checks have been performed with alternative fit models, such as those including the partially reconstructed background, but not J/ψ leakage, or vice versa. The largest bias in the signal yield observed in these studies was 7.5% of the statistical uncertainty. This conclusion is reassuring for our fit model, as even in case it is not completely correct, the signal yield is fairly robust against this. The background components correlate significantly between themselves, however, their correlation with the signal yield is rather small.

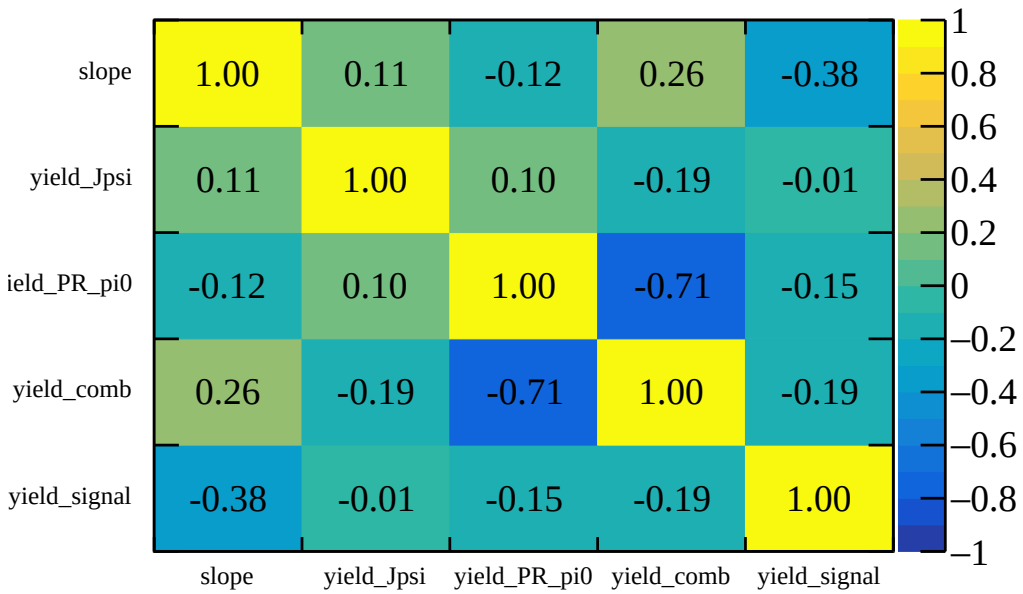


Figure 3.92: Correlation matrix between free fit parameters (in the following order: slope of the combinatorial, yield of the J/ψ leakage component, yield of partially reconstructed background, yield of combinatorial background, and signal yield), obtained using 100 pseudoexperiments.

This test will be repeated after unblinding the $\Lambda_b^0 \to pKe^+e^-$ fit, with realistic signal yields.

3.9.4 Expected signal significance

Once the $\Lambda_b^0 \to pKe^+e^-$ fit will be unblinded, the probability that the observed $\Lambda_b^0 \to pKe^+e^-$ signal peak is a statistical fluctuation has to be quantified.

The fit model with the signal yield constrained to the one returned by the default fit, and the same model but with signal yield constrained to null, are separately fitted to the dataset, and the log-likelihoods are recorded. A test statistic is defined as twice the difference of the log-likelihoods between the two models. Assuming the probability distribution of the test statistic follows a chi-squared distribution, one can compute the p-value corresponding to the probability of this difference.

In case the obtained significance is relatively low, the obtained result can be validated using pseudo-experiments: the background-only PDF resulting from the nominal fit can be used to generate a large number of pseudo-experiments, and searching for fluctuations mimicking the given signal yield. However, if the obtained significance is much larger than five standard deviations, this exercise becomes extremely computationally intensive due to the need to generate an extremely large number ($> \mathcal{O}(10^{10})$) of pseudo-experiments.

The expected signal significance has been estimated by pseudo-experiments, using the `RooDLLSignificanceMCSModule` tool. In each pseudo-experiment, the default model is

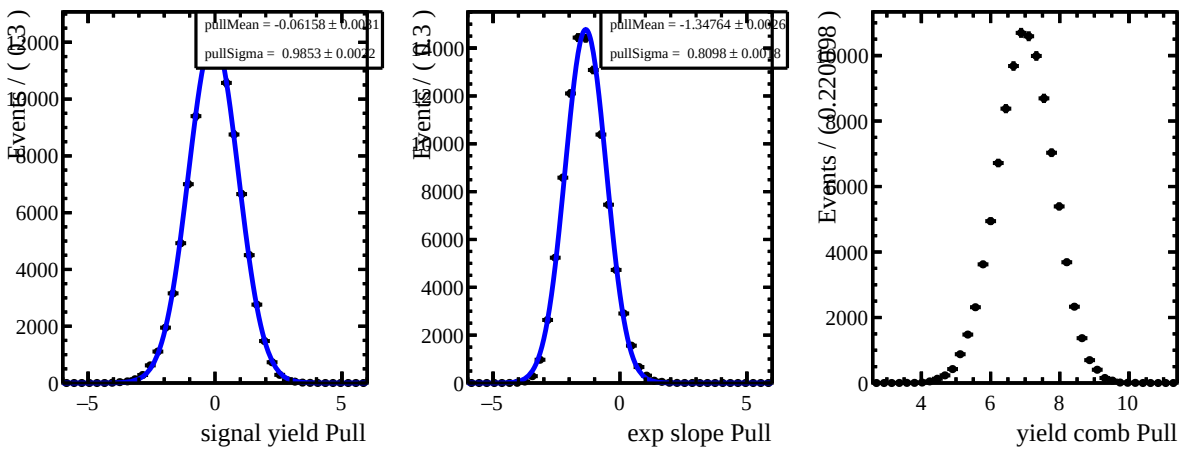


Figure 3.93: Pull distributions for a stringent test of the fit stability, described in the text, using 100 000 pseudo-experiments. The investigated observables are (from left to right) the signal yield, the slope of the exponential shape of the combinatorial background and the yields of the combinatorial background. It can be seen that the signal pull is only at the level of 0.06σ. The Gaussian fit is missing on the right plot due to the limitations of the `RooMCStudy` fit framework: it allows for the values of the mean not larger than five standard deviations.

used to generate events (assuming signal yield of 100 events). As it can be seen from the distribution of signal significance for 2000 pseudo-experiments presented in Fig. 3.94, the expected signal significance (in case signal yield is 100 events) is centered at about eight standard deviations.

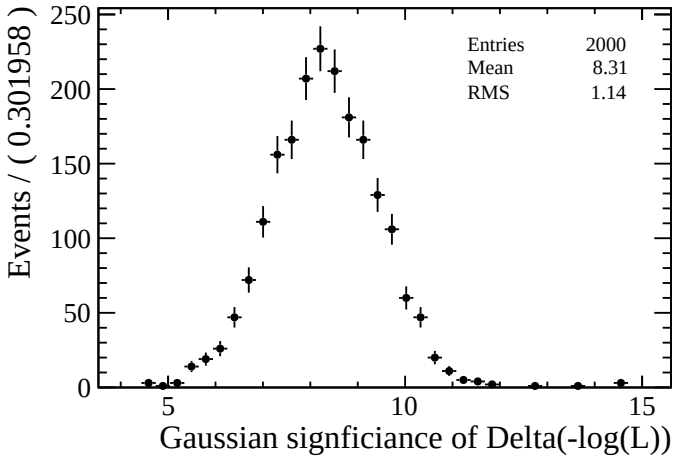


Figure 3.94: Distribution of signal significance for 2000 pseudo-experiments, generated assuming signal yield of 100 events.

3.9.5 Unblinded fit to $\Lambda_b^0 \to pKe^+e^-$ data

The unblinded fit to the merged $\Lambda_b^0 \to pKe^+e^-$ dataset is presented in Fig. 3.95. The signal is observed with the yield of 122 ± 17 candidates, which has a statistical precision of about 13.9%. Accounting for the systematic uncertainties discussed above, the signal yield is found to be $122 \pm 17(\text{stat}) \pm 6(\text{syst})$ candidates.

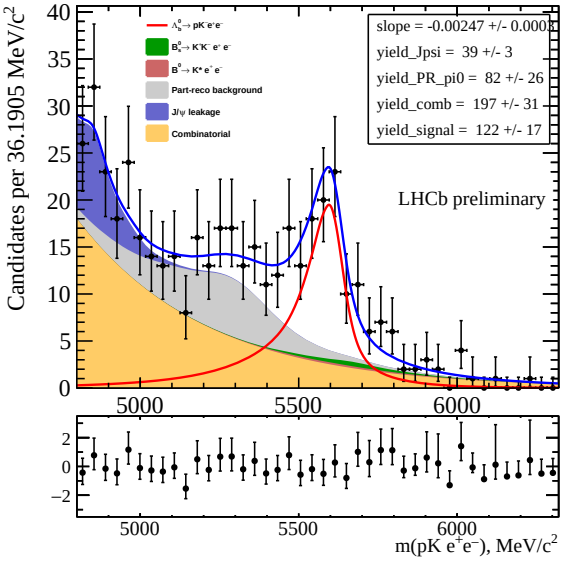


Figure 3.95: Unblinded fit to the $\Lambda_b^0 \to pKe^+e^-$ invariant mass in data, for the merged datasets.

3.9.5.a Signal significance

Following the procedure described in Sec. 3.9.4, the signal significance is estimated to be 8.9σ. This result has been cross-checked using pseudo-experiments (as described in Sec. 3.9.4), generating the amount of each fit component equal to the yield returned by the fit shown in Fig. 3.95. It can be seen that the obtained result is close to the expected signal significance.

To account for the fit model systematics, the profile likelihood of the $\Lambda_b^0 \to pKe^+e^-$ signal yield is smeared by the Gaussian of a width equal to the $\Lambda_b^0 \to pKe^+e^-$ fit model systematic uncertainty (5.1%), which includes the uncertainties due to the model of partially reconstructed background, signal model, and models of other backgrounds, as described in Sec. 3.10.2. The signal significance is computed from the difference of log-likelihood values at the best fit point and the point of signal yield equal to zero. As a consequence of taking systematic effects into account, the global significance lowers to 7.7 standard deviations.

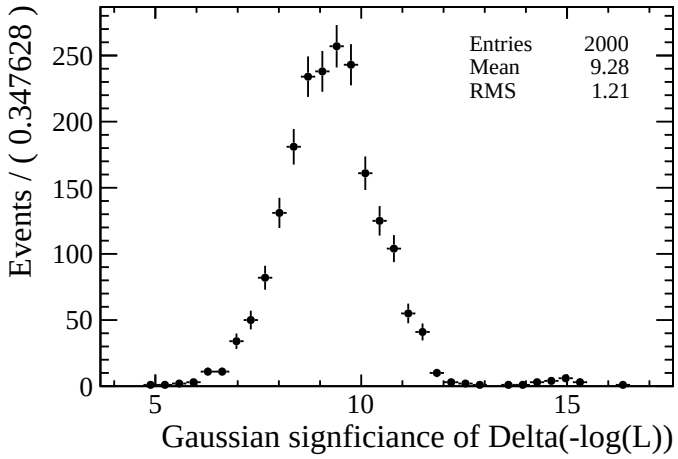


Figure 3.96: Distribution of signal significance for 2000 pseudo-experiments, generated according to the fit result presented in Fig. 3.95.

3.9.5.b Correlation matrix

Figure 3.97 features the correlation matrix between the free fit parameters, obtained from the fit to pKe^+e^- data. It is similar to the one obtained using pseudo-experiments (see Fig. 3.92).

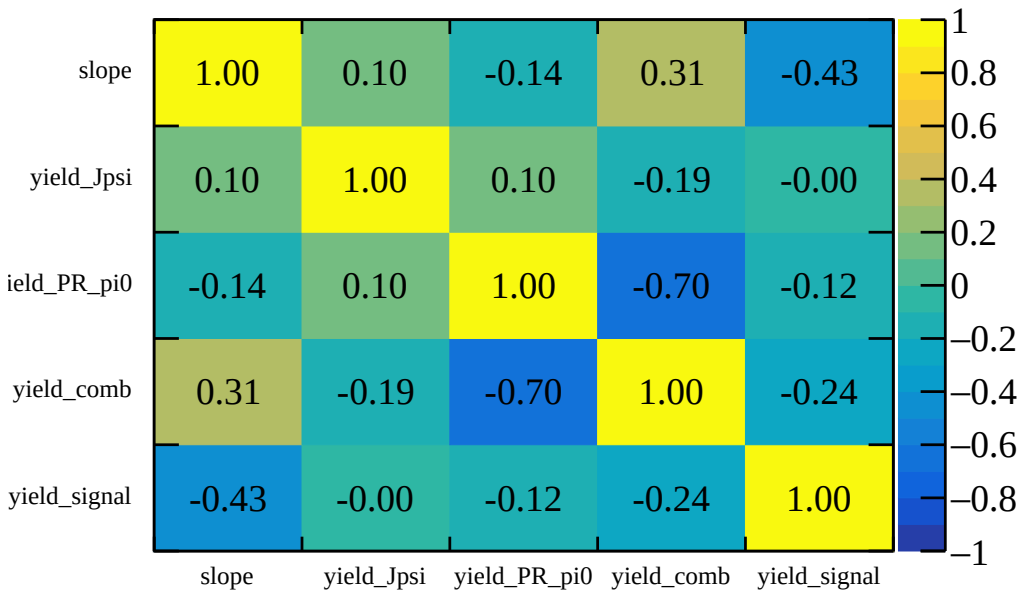


Figure 3.97: Correlation matrix between free fit parameters (in the following order: slope of the combinatorial, yield of the J/ψ leakage component, yield of partially reconstructed background, yield of combinatorial background, and signal yield), in the fit to data presented in Fig. 3.95.

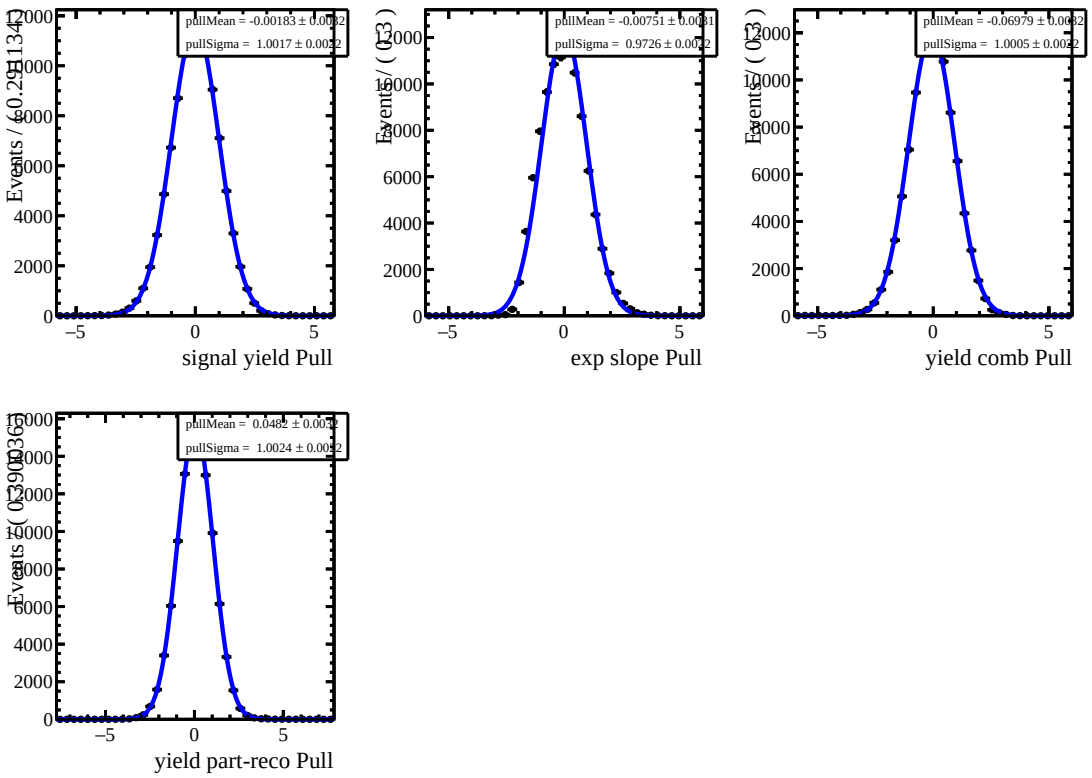


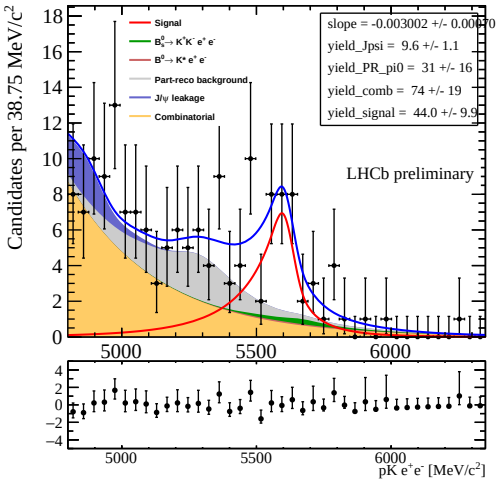
Figure 3.98: Pull distributions for the fit to the pseudo-data, using 100 000 pseudo-experiments generated according to the fit result presented in Fig. 3.95. The investigated observables are (from left to right, top to bottom) the signal yield, the slope of the exponential shape of the combinatorial background, the yields of the combinatorial background, and the yields of the partially reconstructed background.

3.9.5.c Fit validation with pseudo-experiments

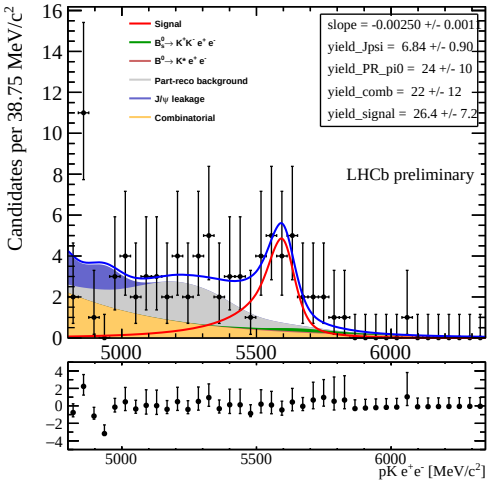
The stability of the fit for the values of the free parameters, returned by the unblinded fit (Fig. 3.95), is validated using 100 000 pseudo-experiments. As it can be seen from Fig. 3.98, the pull distribution of the signal yield exhibits a Gaussian behaviour, with a mean pull which is consistent with zero, and a good coverage of the statistical uncertainties.

3.9.5.d Fits per trigger category and run period

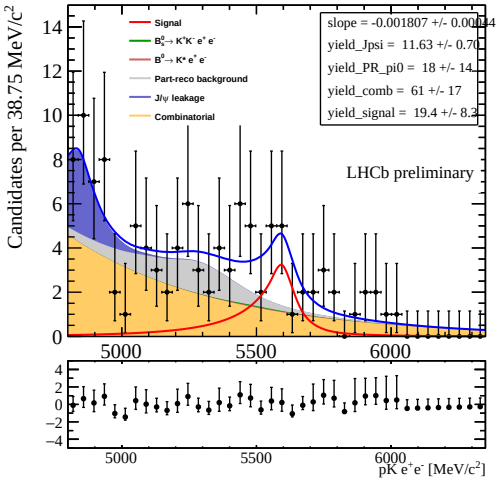
Blinded fits to $\Lambda_b^0 \to pKe^+e^-$ data have already been presented in Sec. 3.7.4.c. In Fig. 3.99, the same fits are presented unblinded.



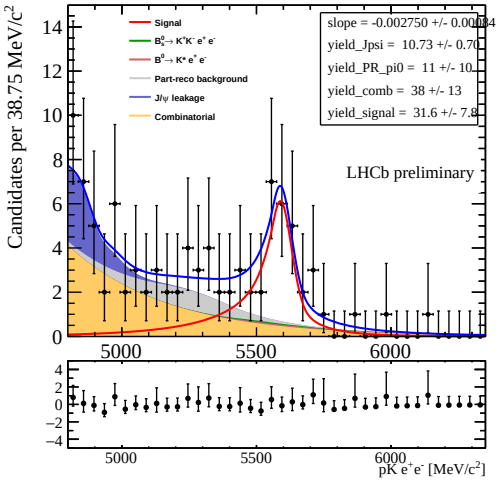
(a) Run 1 L0I



(b) Run 1 L0E!



(c) Run 2 L0I



(d) Run 2 L0E!

Figure 3.99: Unblinded fits to $\Lambda_b^0 \to pKe^+e^-$ data, split per trigger category and dataset.

3.10 Extraction of R_{pK}^{-1} and $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$

In order to extract the measurements of interest, a simultaneous fit is performed to both the $\Lambda_b^0 \to pKe^+e^-$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ decay modes, using Run I and Run II datasets, in both trigger categories (*i.e.* two muon and four electron datasets). The yields and efficiencies of the normalisation $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ and $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ decay modes are extracted beforehand, and are combined with the efficiencies of the $\Lambda_b^0 \to pK\ell^+\ell^-$ modes into a single coefficient, which enters the expression defining R_{pK}^{-1} in the fit model. The fit has R_{pK}^{-1} and the ratio $\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}$ as free parameters, and the $\mathcal{B}(\Lambda_b^0 \to pKe^+e^-)$ is derived from these two later on. In addition, in each dataset, the yield and the slope of the combinatorial background, as well as the yield of partially reconstructed background (electron mode only), are kept free. In the muon datasets, the mean and the width of the signal shape are free as well in the fit. This results in 22 free fit parameters. Other background yields (such as the J/ψ leakage and the misID backgrounds) are constrained as described in the previous sections (Sec. 3.7.3.c for muons, Sec. 3.7.4 for electrons).

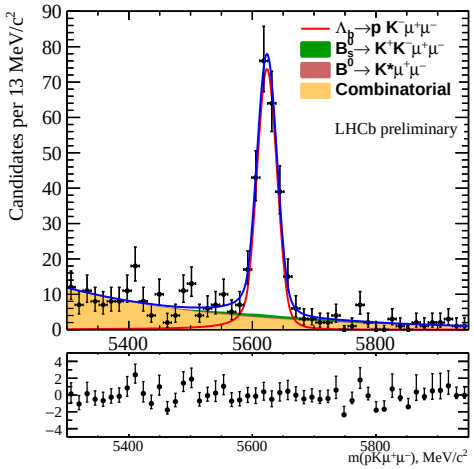
To validate this procedure, the following test was performed: data was used for the Run I and Run II $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ datasets. (Before unblinding of $\Lambda_b^0 \to pKe^+e^-$ data, toy datasets were generated for the $\Lambda_b^0 \to pKe^+e^-$ mode, according to the background model acquired in the blinded fits to the data, as described in the Sec. 3.7.4.d.) The $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ datasets were first fitted separately, in order to extract the fitted signal yields. Then, "fake" efficiency ratios were computed as a ratio of fitted $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ yields. Plugging in these "fake" efficiency ratios into the simultaneous fit, one expects the value of R_{pK}^{-1} to be strictly equal to unity in each category by construction, and therefore also in the result of this simultaneous fit. The obtained result (1.00 ± 0.15) is indeed in an excellent agreement with unity. Since in this section all the efficiency ratios are fixed to their values in the fit, the quoted uncertainty is a purely statistical uncertainty. To get the total uncertainties (as it will be discussed later in Sec. 3.10.4), the uncorrelated systematic uncertainties will be used as Gaussian constraints on efficiency ratios, while the correlated ones will be added to the final result and used to smear the final log-likelihood.

The results of this fit are shown in Fig. 3.100, and the profile log-likelihood in Fig 3.101. The table of fit parameters is shown in Tab. 3.24.

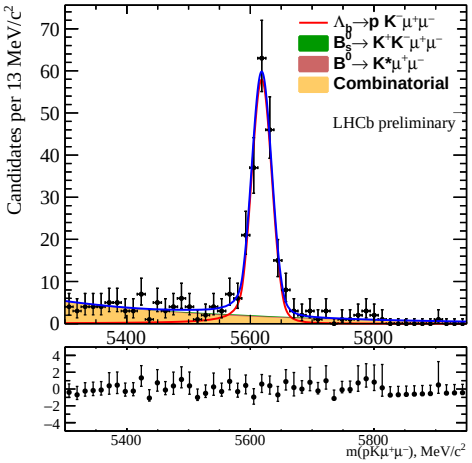
A few useful conclusions can be drawn from this study. First, given that the muon dataset and efficiencies are not blind, this fit directly returns:

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} = (8.45^{+0.45}_{-0.43}) \times 10^{-4}, \tag{3.17}$$

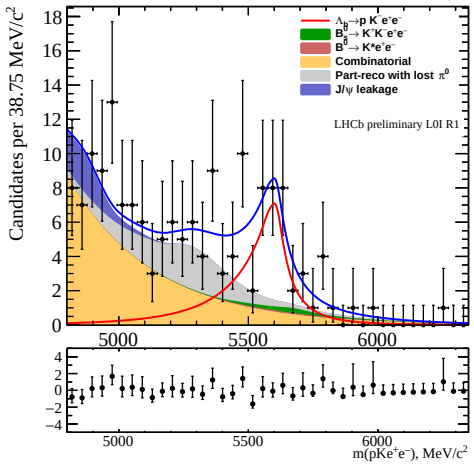
where the uncertainty is statistical only, leading to about 5% statistical precision on this branching fraction measurement. The log-likelihood profile of this measurement is presented in Fig. 3.102. One should keep in mind that this measurement is restricted to the region $m(pK) < 2600$ MeV/c^2 and $0.1 < q^2 < 6$ GeV$^2/c^4$.



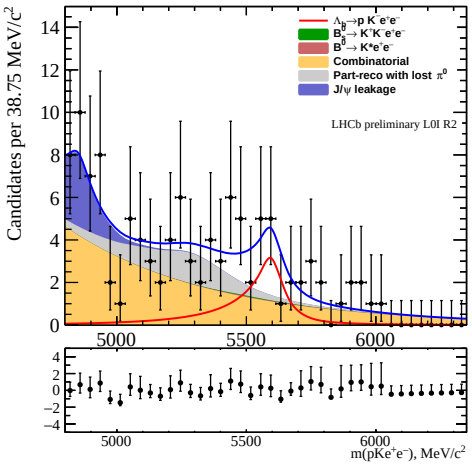
(a) $\Lambda_b^0 \to pK\mu^+\mu^-$ Run 1 L0M



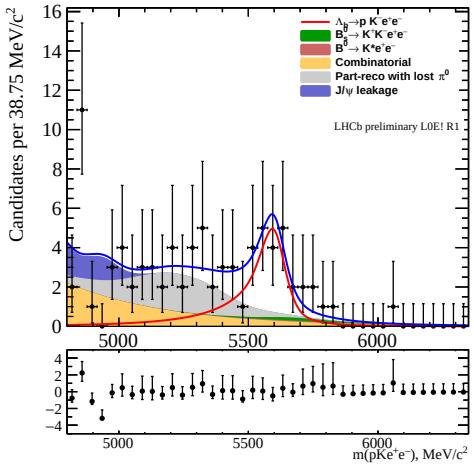
(b) $\Lambda_b^0 \to pK\mu^+\mu^-$ Run 2 L0M



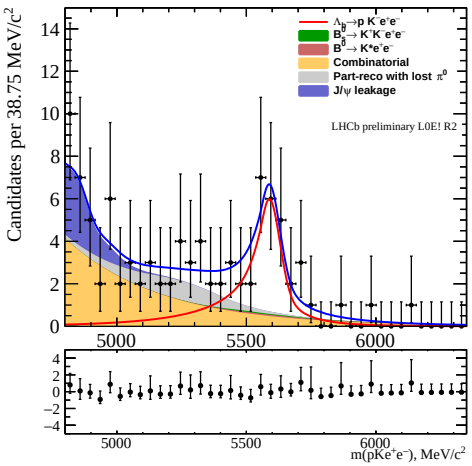
(c) $\Lambda_b^0 \to pKe^+e^-$ Run 1 L0I



(d) $\Lambda_b^0 \to pKe^+e^-$ Run 2 L0I



(e) $\Lambda_b^0 \to pKe^+e^-$ Run 1 L0E!



(f) $\Lambda_b^0 \to pKe^+e^-$ Run 2 L0E!

Figure 3.100: Result of the simultaneous fit (with ”fake” efficiency ratios).

NO.	NAME	VALUE	ERROR	NEGATIVE	POSITIVE	
1	BR_pKmm/pKJpsi		8.45431e-04	4.38786e-05	-4.33011e-05	4.45603e-05
2	1/RpK		9.99428e-01	1.43938e-01	-1.38841e-01	1.49895e-01
3	mean_mm_R1_L0M		5.62375e+03	1.27176e+00	-1.27281e+00	1.27458e+00
4	mean_mm_R2_L0M		5.61842e+03	1.38046e+00	-1.38671e+00	1.39315e+00
5	sigma_mm_R1_L0M		1.63097e+01	1.13776e+00	-1.09591e+00	1.17130e+00
6	sigma_mm_R2_L0M		1.62826e+01	1.27287e+00	-1.28648e+00	1.30559e+00
7	slope_ee_R1_L0Eex		-2.54724e-03	1.07021e-03		9.15068e-04
8	slope_ee_R1_L0I		-3.02353e-03	6.67761e-04	-8.96692e-04	5.62918e-04
9	slope_ee_R2_L0Eex		-2.72022e-03	7.74181e-04	-1.06893e-03	6.77619e-04
10	slope_ee_R2_L0I		-1.80229e-03	4.09367e-04	-4.27038e-04	4.09817e-04
11	slope_mm_R1_L0M		-3.98070e-03	4.90357e-04	-5.04973e-04	4.77850e-04
12	slope_mm_R2_L0M		-3.83612e-03	7.11620e-04	-7.41749e-04	6.87370e-04
13	yield_Jpsi_ee_R1_L0Eex		6.83462e+00	8.99437e-01	-8.97613e-01	9.01477e-01
14	yield_Jpsi_ee_R1_L0I		9.56973e+00	1.09856e+00	-1.09853e+00	1.09887e+00
15	yield_Jpsi_ee_R2_L0Eex		1.07244e+01	6.99308e-01	-6.99775e-01	6.98970e-01
16	yield_Jpsi_ee_R2_L0I		1.16295e+01	6.98818e-01	-6.98445e-01	6.99299e-01
17	yield_PR_pi0_ee_R1_L0Eex		2.43858e+01	1.02812e+01	-1.00914e+01	1.05395e+01
18	yield_PR_pi0_ee_R1_L0I		3.01209e+01	1.63021e+01	-1.59514e+01	1.68432e+01
19	yield_PR_pi0_ee_R2_L0Eex		1.08771e+01	1.03111e+01	-1.01140e+01	1.05877e+01
20	yield_PR_pi0_ee_R2_L0I		1.79914e+01	1.38127e+01	-1.34552e+01	1.42798e+01
21	yield_comb_ee_R1_L0Eex		2.22139e+01	1.14000e+01	-1.80837e+01	1.21339e+01
22	yield_comb_ee_R1_L0I		7.40394e+01	1.85188e+01	-1.81396e+01	1.91386e+01
23	yield_comb_ee_R2_L0Eex		3.82090e+01	1.30229e+01	-1.23799e+01	1.37395e+01
24	yield_comb_ee_R2_L0I		6.17369e+01	1.64501e+01	-1.59945e+01	1.69909e+01
25	yield_comb_mm_R1_L0M		2.11519e+02	1.69821e+01	-1.65808e+01	1.73934e+01
26	yield_comb_mm_R2_L0M		9.50497e+01	1.17807e+01	-1.13883e+01	1.21944e+01

Table 3.24: Result from `Minuit` of the fit shown in Fig. 3.100. The third column shows the fitted values of each fit parameter. The fourth column shows symmetric uncertainties, while asymmetric uncertainties are shown in the last two columns.

The R_{pK}^{-1} (with fake efficiencies constructing it to be equal to unity) is found to be:

$$R_{pK}^{-1} = 1.00^{+0.15}_{-0.14}(\text{stat}), \tag{3.18}$$

which means we reach about 15% statistical precision on the LU ratio.

Secondly, one can compare the shape of the profile likelihood for the R_{pK}^{-1} and R_{pK} observables, validating our choice to "invert" the definition of the lepton universality ratio and put the electron mode in the numerator. This comparison is presented in Fig. 3.101, and it is clearly visible that our choice to measure R_{pK}^{-1} provides us with a much more parabolic likelihood shape, than the R_{pK} definition where electron mode is in the denominator.

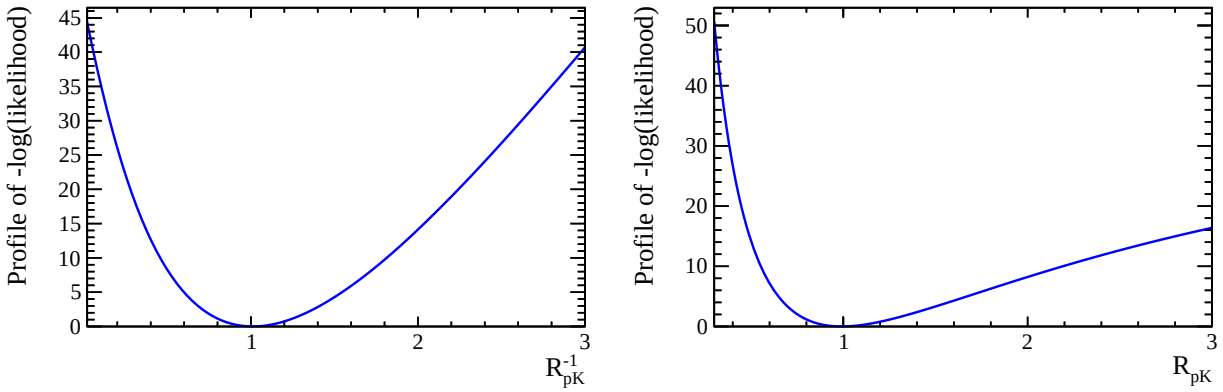


Figure 3.101: Left: log-likelihood profile for R_{pK}^{-1} in the fit shown in Fig. 3.100. Right: log-likelihood profile for the R_{pK}, if it is considered as a free parameter instead of R_{pK}^{-1}.

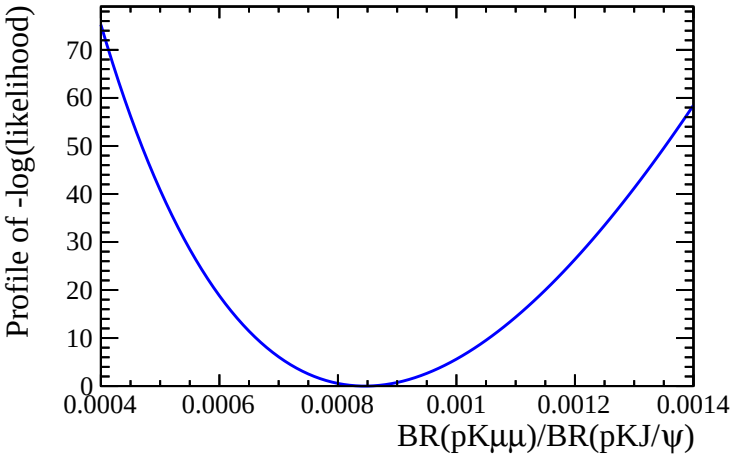


Figure 3.102: Log-likelihood profile for the ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi$, in the fit shown in Fig. 3.100.

3.10.1 Systematic uncertainty on $\mathcal{B}$ ($\Lambda_b^0 \to pK\mu^+\mu^-$)

The global strategy of estimating the systematic uncertainties is similar to the one adopted in Sec. 3.8.4. Since the branching ratio of the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay is evaluated exploiting the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ decay as a normalisation, certain uncertainties related to the specific final state will cancel. However, the uncertainties due to the imprecise knowledge of the branching fraction of the normalisation mode should be taken into account.

The list of sources of systematic uncertainty considered is presented below, and is in general similar to the one discussed above.

- **Decay model**: the default $m(pK)$ weights described in Section 3.5.1 are extracted using the normalisation $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode. However, the $m(pK)$ spectrum is not necessarily the same in the non-resonant mode. To assign a related systematic uncertainty, alternative $m(pK)$ weights are extracted from the non-resonant $\Lambda_b^0 \to pK\mu^+\mu^-$ data, and applied to the reconstructed-level and generator-level $\Lambda_b^0 \to pK\mu^+\mu^-$ simulation. The spectrum exploited here is shown in Fig. 3.110, and the correction weights are presented in Fig. 3.103(left). In order to gain additional statistical power, the Run I and Run II datasets are merged for this study. The difference in the value of the branching fraction, obtained by the default and this methods, is assigned as a systematic uncertainty accounting for an imprecise knowledge of the decay model.

- q^2 **distribution**: The shape of the q^2 distribution might be mismodeled in simulation, if for example wrong assumptions are made on the decay model or on the form-factors. As the simulation used in this analysis has a phase-space model, it certainly does not describe properly the presence of the so-called photon pole at low-q^2 for instance. In the absence of theoretical predictions for the $\Lambda_b^0 \to pK\mu^+\mu^-$, the $m(\mu^+\mu^-)$ shape is extracted directly from the $\Lambda_b^0 \to pK\mu^+\mu^-$ data (by the procedure described above for the $m(pK)$). Then, the simulation is reweighted to match this distribution applying the weights presented in Fig. 3.103(right). The difference in branching fraction computed using this and the default methods, is assigned as a systematic uncertainty. It should be noted that the used q^2 weights have very large statistical uncertainties, which reflect in the value of systematic uncertainty. Therefore, three alternative binning schemes in q^2 have been considered for this correction (9, 11 and 13 bins).

- **Helicity angles**: Even after correcting for the $m(pK)$ spectrum or q^2 model, there can remain some residual discrepancies in the distributions of the helicity angles $\cos\theta_K$, $\cos\theta_l$. Those discrepancies can originate, for example, from the spin effects unaccounted for in the one-dimensional reweighting. As it can be seen in Appendix E, there are quite significant efficiency trends in these variables. Therefore, we apply a systematic uncertainty for any potential biases. To do that, the correction on angles $\cos\theta_K$, $\cos\theta_l$ is extracted using $\Lambda_b^0 \to pK\mu^+\mu^-$ data and simulation (in a similar manner to the q^2 correction described above), and the $\Lambda_b^0 \to pK\mu^+\mu^-$ simulation is reweighted accordingly. The difference in branching fraction computed

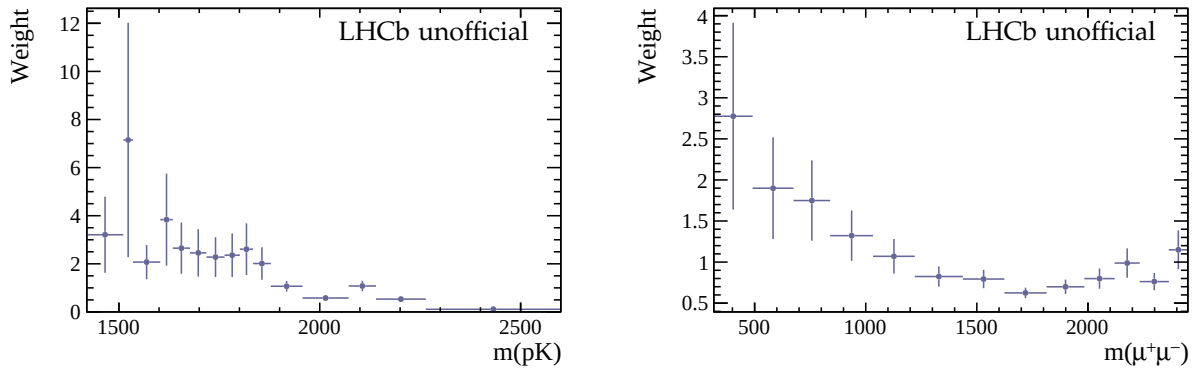


Figure 3.103: Correction to the $\Lambda_b^0 \to pK\mu^+\mu^-$ simulation for (left) $m(pK)$ and (right) $m(\mu^+\mu^-)$ distributions, extracted using the $\Lambda_b^0 \to pK\mu^+\mu^-$ data.

using reweighted and the default simulation samples, is assigned as a systematic uncertainty (averaging uncertainties coming from two angles). The third helicity angle ϕ is not considered, as the efficiency is flat in it (see Appendix E).

- **Size of samples used to extract corrections**: The bootstrapping technique is used the way described above, in order to assess the limited size of the samples used to extract the corrections to the simulations described in Sec. 3.5.

- **Binning of the corrections**: Binned weights are used instead of interpolated ones.

- **Multiplicity proxy**: The `nTracks` variable is chosen as a default proxy for the event multiplicity. Four additional choices are considered to estimate the relevant systematic uncertainty: a) the `nSPDHits` variable, b) the `nVeloTracks` variable, c) the `nPVs` variable, d) no multiplicity correction at all. It should be noted that when reweighting the `nSPDHits` variable, this correction is not applied to the generator-level simulation sample, as it is known that the major reason for the data-simulation disagreement in this variable comes from the improper modeling of detector (*i.e.* reconstruction) effects.

- **PID: procedure**: As the final state is identical between the signal and normalisation modes, the uncertainties related to the PID procedure such as usage of an ${}_s\mathcal{P}lot$ technique (see Sec. 3.5.2), should cancel in the ratio. However, due to the slightly different kinematics of the final state, some effects can be seen. To estimate their size, alternative binning schemes are used.

- **Trigger category bias**: The kinematics weights (Sec. 3.5.4) are by default computed using the L0M trigger category. To assign a related uncertainty, the corrections are recomputed using the L0I category on the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ dataset, and applied in a symmetrical manner to both signal and normalisation modes, to both generator- and reconstructed-level simulation samples. The resulting alternative value of branching fraction is used to estimate the associated systematic uncertainty.

- **L0Muon trigger**: The "TIS requirement" used for the TISTOS method when computing the weights(see Sec. 3.5.6.a), is changed from `L0Global_TIS` to `L0Hadron_TOS` on the proton or kaon tracks.

- **Decay time corrections**: in Run I, additional corrections were applied to the decay time distribution (Sec. 3.5.5). In Run II, such corrections were not applied. Effect of absence/presence of these corrections on the value of the BR is considered as a systematic uncertainty.

- **Residual corrections** (or their absence). As described in Sec. 3.5.7, there are only minor residual data-simulation differences in the important variables. To estimate a related uncertainty, the correction is applied for the proton p_T and χ^2_{IP}, and the difference in the BR value is assigned as an uncertainty.

- **HLT**: Given that the HLT weights do not exhibit any dependence on the kinematics, the related uncertainty should fully cancel. No uncertainty is assigned.

- **MVA**: As the data-MC agreement in the BDT variable is good (Sec. 3.5.7), no uncertainty is assigned.

- **Fit models**: The systematic uncertainties associated with the fit modeling, described in Section 3.7, can be related to the following procedures:

  - **Resonant mode**: The same strategy as described in the $r^{-1}_{J/\psi}$ systematics section.
  - **Rare mode**:
    * The fractions of yields of misidentification backgrounds are allowed to float. Obtained values of these fractions are used to generate 10 000 toy datasets. Fitting to these toy datasets with the default fit model allows to estimate the bias on the signal yield. Given the large uncertainty on these fractions as returned by the fit, several input values of them have been explored (as well as the option of absence of any mis-indentified backgrounds), and the largest bias on the signal yield is taken as a systematic uncertainty.
    * The partially reconstructed background components (lost-π^0 and semileptonic), omitted in the default fit, are allowed. The returned yields are consistent with zero, but have large uncertainties. These value of uncertainty are taken as as input to generate 10 000 toy datasets, which are then fitted with the default fit model. This allows to determine the bias on the signal yield.
    * The impact of increasing or decreasing the smearing parameter of the `RooKeysPdf` method, when producing the background PDFs, is quantified using 10 000 pseudo-experiments. Several input values between 0.8 and 2.0 of the smoothing parameter have been explored, and the largest bias is taken as a systematic uncertainty.

A summary of the values of each source of uncertainty is provided in Table 3.25.

3.10.1.a External input

In this analysis, one of the observables which is extracted from the nominal simultaneous fit is the ratio of branching fractions $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$. Therefore, in order to quote in the final results the absolute value of $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$ obtained from this work, the $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ is used as an external input. The Particle Data Group reports: $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi = 3.17 \pm 0.04 \pm 0.07 \pm 0.34^{+0.45}_{-0.28}) \times 10^{-4}$ [6], which is based on the LHCb measurement [86].

Here, the first uncertainty is statistical, the second one is systematic, the third one is due to the knowledge of the normalisation mode branching fraction, and the last one due to the knowledge of the hadronisation fractions $f_{\Lambda_b^0}/f_d$. As it can be seen, one of the largest contribution to the error budget comes from the imprecise knowledge of the Λ_b^0 production, and of $f_{\Lambda_b^0}/f_d$ dependence on the Λ_b^0 kinematics. In the present analysis, as both the signal and normalisation modes are decays of the Λ_b^0 baryon, this uncertainty does not enter in the ratio of branching fractions, but it is nevertheless relevant for the absolute value of the branching fraction, and will be quoted separately from other contributions.

The published measurement of $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ was performed with Run I LHCb data only. Nevertheless, given that our measurement was performed with a different selection (including trigger and stripping requirements), no potential correlations are taken into account.

3.10.2 Systematic uncertainties on the double ratio R_{pK}^{-1}

The sources of systematic uncertainty considered for the measurement R_{pK}^{-1} are similar to the ones discussed above. The usage of a double ratio implies the cancellation of certain systematic effects, thus reducing the overall size of the uncertainties.

- **Decay model**: the default $m(pK)$ weights described in Section 3.5.1 are extracted using the normalisation $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode. However, the $m(pK)$ spectrum is not necessarily identical to the non-resonant mode. To assign a related systematic uncertainty to account for this possibility, alternative $m(pK)$ weights are extracted from the non-resonant $\Lambda_b^0 \to pK\mu^+\mu^-$ data, and applied to the reconstructed-level and generator-level $\Lambda_b^0 \to pK\ell^+\ell^-$ simulation. The spectrum exploited here is shown in Fig. 3.110, and the correction weights are presented in Fig. 3.103(left). In order to gain additional statistical power, the Run I and Run II datasets are merged for this study. The difference in the branching fraction value, obtained by the default and this methods, is assigned as a systematic uncertainty accounting for an imprecise knowledge of the decay model.

- q^2 **distribution**: The shape of the q^2 distribution might be mismodeled by the simulation, if the wrong assumptions are made on the decay model or form-factors. As the simulation used in this analysis has a phase-space model, it certainly does not

Table 3.25: Systematic uncertainties on the ratio of branching fractions $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$, %.

Type	Run I L0M	Run II L0M
$m(pK)$ model	1.20	1.41
q^2 model	0.99	0.24
Helicity angles	2.01	1.74
Total decay model	2.54	2.25
Corrections (statistics)	1.34	1.40
Corrections (binning)	0.71	1.90
Multiplicity proxy	0.46	1.31
Trigger bias	0.54	0.51
PID procedure	0.37	0.03
L0 procedure	0.17	0.63
Decay time corrections	0.59	0.04
Residual corrections	1.65	1.62
Total corrections	2.46	3.25
Fit (normalisation)	0.50	1.20
Fit (background model)	1.43	
Total systematic	3.85	4.37
Statistical uncertainty	5.2	

describe properly the presence of the so-called photon pole at low-q^2 for instance. As it is shown in Appendix E, the selection efficiency can have a mild q^2 dependence for electron mode. Therefore, an uncertainty is applied to account for this systematic effect. In the absence of theoretical predictions for the $\Lambda_b^0 \to pK\ell^+\ell^-$, the q^2 shape is extracted from the $\Lambda_b^0 \to pK\mu^+\mu^-$ data (by the procedure described above for the $m(pK)$), and the generated value of q^2 in both reconstructed- and generator-level simulation is reweighted to match this distribution. Typical weights are presented in Fig. 3.103 (right). It should be noted that the used q^2 weights have very large uncertainties themselves, which reflect in the value of systematic uncertainty. Therefore, three alternative binning schemes in q^2 have been considered for this correction (9, 11 and 13 bins).

- **Helicity angles**: Even after correcting for the $m(pK)$ spectrum or q^2 model, there can remain some residual discrepancies in the distributions of helicity angles $\cos\theta_K$, $\cos\theta_l$. Those discrepancies can come, for example, from the spin effects unaccounted for in one-dimensional reweighting. As it can be seen in Appendix E, there are quite significant efficiency trends in these variables. However, the trend is usually similar for the muon and electron modes, which implies a partial cancellation of systematic effects. To apply a systematic uncertainty for potential biases, the correction on angles $\cos\theta_K$, $\cos\theta_l$ is extracted using $\Lambda_b^0 \to pK\mu^+\mu^-$ data and simulation (as it

was discussed for the $\mathcal{B}$ ($\Lambda_b^0 \to pK\mu^+\mu^-$) systematics). Both $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ simulation samples are reweighted with the same set of weights. The difference in the double ratio of efficiencies, computed using reweighted and the default simulation samples, is assigned as a systematic uncertainty (averaging uncertainties coming from two angles). The third helicity angle ϕ is not considered, as efficiency is flat in it (see Appendix E).

- **Statistics of the samples used to extract corrections**: The bootstrapping technique is used the way described above, in order to assess the limited size of the samples used to extract the corrections to the simulations.

- **Binning of the corrections**: Binned weights (kinematics, trigger, `nTracks`) are used instead of interpolated ones.

- **Multiplicity proxy**: The similar strategy is adopted as for the $r_{J/\psi}^{-1}$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ systematics, discussed above.

- **PID procedure**: As the final state is identical between the signal and normalisation modes, the uncertainties related to the PID procedure such as the usage of an ${}_s\mathcal{P}lot$ technique (see Sec. 3.5.2), should cancel in the double ratio. However, due to the slightly different kinematics of the final state, some effects can be seen. To estimate their size, alternative binning schemes are used.

- **Trigger category bias**: The kinematics weights (Sec. 3.5.4) are by default computed using the L0M trigger category. To assign a related uncertainty, the similar strategy is adopted as for the case of $r_{J/\psi}^{-1}$ systematics.

- **L0 trigger**: To assign a related uncertainty, the similar strategy is adopted as for the case of $r_{J/\psi}^{-1}$ systematics.

- **Decay time corrections**: Effect of absence/presence of these corrections on the value of the BR is considered as a systematic uncertainty, in the way already discussed above.

- **Residual corrections** (or their absence). As described in Sec. 3.5.7, there are only minor residual data-simulation differences in the important variables. To estimate a related uncertainty, the correction is applied for the proton p_T or proton χ^2_{IP} (on all four reco-level simulation samples entering the computation of R_{pK}^{-1}), and the maximal difference in the double ratio of efficiencies is assigned as an uncertainty.

- **HLT**: Given that the HLT weights do not exhibit any dependence on the kinematics, the related uncertainty should fully cancel in a double ratio. No uncertainty is assigned.

- **Fit models**: The systematic uncertainties associated with the fit modeling, described in Section 3.7, can be related to the following procedures:

– **Resonant modes**: The same strategy as described in the $r_{J/\psi}^{-1}$ systematics uncertainties section. We apply the quadratic sum of two systematics effects from Table 3.23 as an uncertainty.

– **Rare mode** $\Lambda_b^0 \to pK\mu^+\mu^-$: the same approach is used as in the systematics on $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$.

– **Rare mode** $\Lambda_b^0 \to pKe^+e^-$:

  * The mean of the signal distribution (by default, fixed to the value obtained from the fit to the simulation, 5573 MeV/c^2) is allowed to float in the fit to the merged datasets. The resulting value from the fit to the data (5578 ± 12 MeV) is in a good agreement with a value from the fit to the simulation, therefore no uncertainty is applied to account for a possible bias.
  * The signal width (by default fixed to the value from simulation times the resolution factor 1.15, see Sec. 3.7.4) is increased or decreased by 5% (which is a conservative uncertainty on the resolution factor. The resulting bias on the signal yield is quantified using 10 000 pseudo-experiments.
  * An effect of increasing or decreasing the smearing parameter of the `RooKeysPdf` method, when producing the background PDFs, is quantified using 10 000 pseudo-experiments. Several input values between 0.8 and 2.0 of the smoothing parameter have been explored, and the largest bias is taken as a systematic uncertainty.
  * The semileptonic background component, omitted in the default fit, is included in the fit to the merged dataset. The returned yield is consistent with zero, but has a large uncertainty. This value of the uncertainty is taken as as an input to generate 10 000 toy datasets, which are then later fitted with the default fit model. This allows to determine the bias on the signal yield.
  * The dominant systematic uncertainty of this analysis comes from an imprecise knowledge of the composition of the partially reconstructed background $\Lambda_b^0 \to pKe^+e^-\pi^0$ hence the invariant mass shape of the component. By default, this decay is assumed to proceed through an intermediate K^{*-} resonance. Alternative hypothesis considered include an intermediate Δ resonance, an intermediate Λ^* resonance, a phase-space hypothesis without resonant structures, and a hypothesis of absence of this background. The reweighting procedure has been detailed in Sec. 3.7.4.b. First, the default fit (Sec. 3.9) is performed to the merged dataset, under each of the part-reco hypotheses. This allows to extract the values of free fit parameters for each of them. These values are used as an input to series of 10 000 pseudo-experiments, which are generated for each of the alternative models, and fit with the nominal one. The largest bias observed in the $\Lambda_b^0 \to pKe^+e^-$ yield, 4.28%, comes from the case where a phase-space model is used. This largest deviation is assigned as a systematic uncertainty. This effect is evaluated in

the full Run I + Run II sample with both trigger categories and thus the same fully correlated value is assigned to all the categories.

* The impact of not assuming LU in the peaking backgrounds: by default, we assume that the proportion of misID backgrounds is similar between the muon and the electron modes. However, if LU were to be violated in these decay modes, then this assumption does not hold anymore. An extreme scenario is the one where we have no misID background in the muon mode, and nominal misID yields in the electron mode. This is equivalent to the study which already has been performed for the $\Lambda_b^0 \to pK\mu^+\mu^-$ fit systematics evaluation, where it was assumed that there is no misID backgrounds, while we include them into the fit with fixed proportions. This systematic uncertainty was found to be at the sub-percent level (0.75%) in that extreme case, and is expected to be much smaller for more realistic scenarios. This is not surprising, as the misID yields are very small. Therefore, no dedicated systematic is added to account for this effect in R_{pK}^{-1} evaluation.

In the table of systematics shown below, all the uncertainties related to the $\Lambda_b^0 \to pK\ell^+\ell^-$ signal and background models (except for the dominant one, related to the partially reconstructed backgound model), are added in quadrature and grouped together.

Besides the sources of uncertainties described above, the following additional sources, related to the specificity of the $\Lambda_b^0 \to pKe^+e^-$ decay mode, have been considered:

- **Bin migration**: due to the energy loss suffered by electrons and the relatively large resolution on their reconstructed momentum, candidates with q^2_{true} in the rare decay region can have a value of q^2_{reco} outside this range and vice versa. This effect is accounted for in the reconstruction efficiency, since the requirement is both applied in the generator level (where q^2_{true} is used) and reconstruction level samples (q^2_{reco} is used). However, potential mismodelling of the q^2 distribution in the simulation samples can change the proportion of events migrating in and out the region of interest. Two separate effects are studied: the effect of a potential underestimation of the q^2_{reco} resolution in simulation, and the effect of the physics model used to generate the simulation sample. The first effect is studied by comparing the J/ψ mass resolution in simulation and data. A small shift in the mean of the distribution, and a slightly larger resolution are observed in data with respect to simulation. These parameters are used to smear the dielectron mass in the $\Lambda_b^0 \to pKe^+e^-$ simulation sample, and the bin migration is re-computed using the smeared mass instead of the reconstructed one. The largest difference between such a smeared model and the default one is found to be 1.1%. To evaluate the second effect, simulated $B^0 \to K^{*0}e^+e^-$ candidates, generated with the `BTOSLLBALL` physical model from the `EvtGen` package [154], are used. In order to study the bin migration effect, one needs a model that can be applied to a large q^2 region, much broader than the region used

for the analysis. The most important migration is coming from the regions just above and just below the q^2 window of interest. Since in data the region above 6 GeV$^2/c^4$ is dominated by the J/ψ pole, one cannot directly study this effect from s-weighted data (effect of J/ψ leakage is already taken into account in the invariant mass fit). As the q^2 distribution of $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates shown in Fig. 3.3 is similar to the q^2 model observed in the $B^0 \to K^{*0}e^+e^-$ decay, relevant simulated sample can help to estimate the size of the effect. Correction weights are obtained as the ratio of the q^2_{true} distribution in generator level $B^0 \to K^{*0}e^+e^-$ and $\Lambda_b^0 \to pKe^+e^-$ samples, and are applied to reconstructed MC $\Lambda_b^0 \to pKe^+e^-$ events, to correct their q^2_{true} distribution. The bin migration is re-computed taking into account this correction, and maximal difference with respect to the default model is found to be about 2%, with large uncertainties.

Consequently, a 2% systematic uncertainty is assigned, taking into account both effects. This systematic uncertainty is applied as 100% correlated between years and trigger categories. It has been checked that the effect is similar for L0I and L0E! trigger categories, as well as for Run I and Run II.

- **HOP cut efficiency**: the possible data-simulation disagreement in the HOP resolution is studied exploiting the resonant mode. Correction weights are computed as the ratio of the α_{HOP} distribution in simulation and s-weighted $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ data candidates. The corrections are then applied as per event weights to the rare mode simulation, and the relative difference on the HOP cut efficiency computed when applying or not these corrections is assigned as systematic uncertainty. The relative variation on the efficiency when including or not the weights is computed without looking at the absolute HOP cut efficiency, since it is still blinded. The procedure is performed independently for each year and trigger category. Since this effect arises from the imperfect description of the HOP resolution in MC, it is considered to be fully correlated between different years and categories. All values of uncertainty are consistent with zero, but within quite large errors (1%). Therefore, a half of the error (0.5%) is assigned as a conservative systematic uncertainty on our knowledge of the HOP cut efficiency.

The summary of the values of each source of uncertainty is provided in Table 3.26. It can be seen that in most of the cases, the uncertainties are smaller than those on the single ratio $r_{J/\psi}^{-1}$.

3.10.3 Systematic uncertainties on $\mathcal{B}(\Lambda_b^0 \to pKe^+e^-)$

The branching fraction of the decay mode $\Lambda_b^0 \to pKe^+e^-$ will be extracted as a derived parameter from the simultaneous fit, which returns both $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$ and R_{pK}^{-1} observables. Consequently, when all the uncertainties on these two free parameters are plugged into the fit as Gaussian constraints on the relevant ratios of efficiencies, the fit also returns the uncertainty on the derived $\mathcal{B}(\Lambda_b^0 \to pKe^+e^-)$.

Table 3.26: Systematic uncertainties on R_{pK}^{-1}, %.

Type	Run I L0I	Run I L0E!	Run II L0I	Run II L0E!
$m(pK)$ model	0.85	0.84	0.36	1.02
q^2 model	0.64	0.80	1.30	0.44
Helicity angles	1.46	1.57	0.99	1.68
Total decay model	1.81	1.95	1.67	2.01
Corrections (statistics)	2.12	2.64	2.86	2.51
Corrections (binning)	1.03	1.27	0.83	0.37
Multiplicity proxy	2.24	1.29	0.49	1.88
Trigger bias	0.09	1.39	1.61	0.33
PID procedure	0.05	0.45	0.04	0.15
L0 procedure	0.65	0.14	1.20	0.18
Decay time corrections	0.75	0.44	0.05	0.01
Residual corrections	0.13	0.11	0.18	0.23
Total corrections	3.40	3.55	3.63	3.19
Fit (normalisation)	2.64	2.80	2.30	0.86
Fit (part-reco model)	4.28			
Fit (signal and background model)	2.87			
HOP cut efficiency	0.5			
Bin migration	2			
Total systematic	7.25	7.42	7.21	6.77

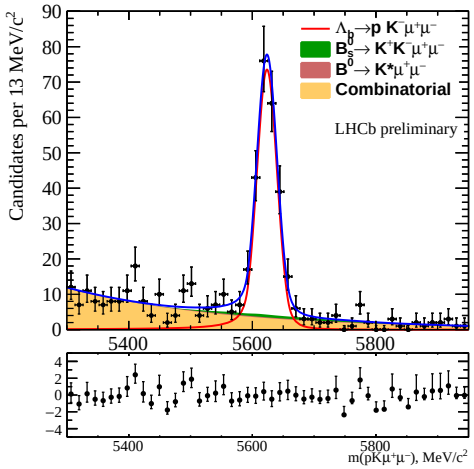
3.10.4 Including the systematic uncertainties in the results

The systematic uncertainties on the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, if they are uncorrelated between datasets, are accounted for as Gaussian constraints on the ratio of efficiencies $(\Lambda_b^0 \to pK\mu^+\mu^- / \Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-))$ in the simultaneous fit presented in the previous section. The fit is run again, and the resulting uncertainty is a combination of the statistical and systematic uncertainties. As these two are almost uncorrelated, we can separate them. After that, the correlated systematic uncertainties are introduced. The profile likelihood of $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-))$ is smeared by a Gaussian reflecting the correlated uncertainties.

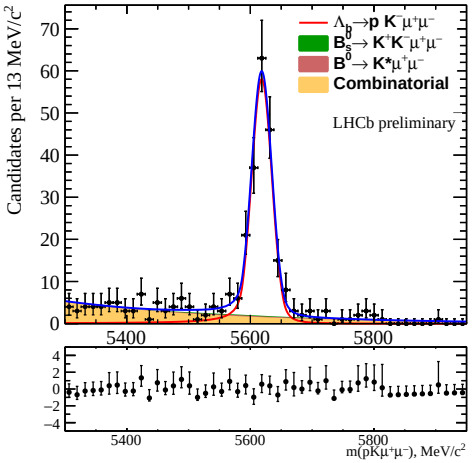
A similar strategy is adopted for the systematic uncertainty on the R_{pK}^{-1} observable. In this case, the uncertainties enter as Gaussian constraints on the double ratio of efficiencies.

The fit result in this configuration is presented below (Tab. 3.27, Fig. 3.104). It should be noted that the number of free fit parameters is now larger, as the ratios of efficiencies are not fixed anymore but Gaussian-constrained. It should also be noted that the table of free fit parameters below does not represent the last step of inclusion of correlated systematic uncertainties, as this is performed independently afterwards. This one is however added in the results quoted below.

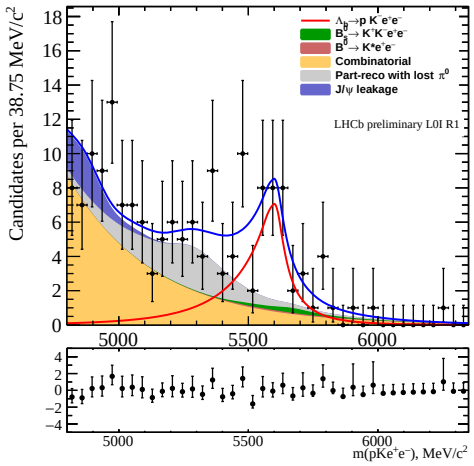
This leads to the following result, accounting for the statistical and systematic uncer-



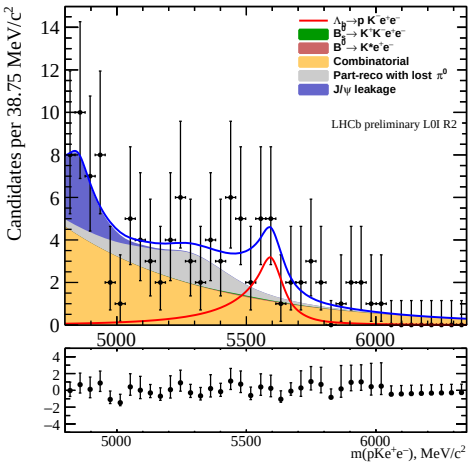
(a) $\Lambda_b^0 \to pK\mu^+\mu^-$ Run 1 L0M



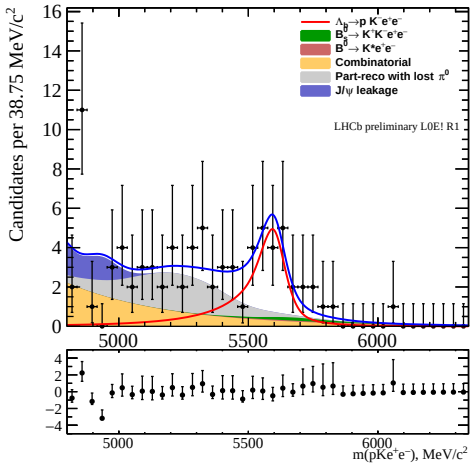
(b) $\Lambda_b^0 \to pK\mu^+\mu^-$ Run 2 L0M



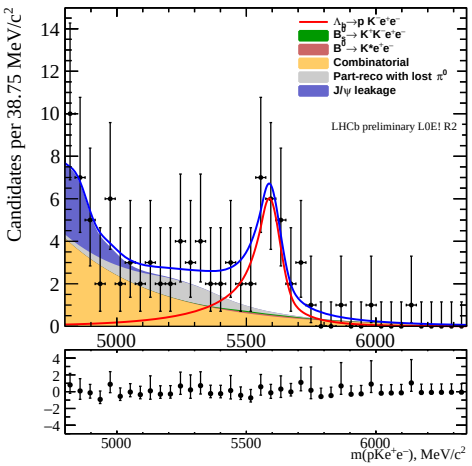
(c) $\Lambda_b^0 \to pKe^+e^-$ Run 1 L0I



(d) $\Lambda_b^0 \to pKe^+e^-$ Run 2 L0I



(e) $\Lambda_b^0 \to pKe^+e^-$ Run 1 L0E!



(f) $\Lambda_b^0 \to pKe^+e^-$ Run 2 L0E!

Figure 3.104: Result of the simultaneous fit (with ”fake” efficiency ratios) with Gaussian constraints on efficiencies.

NO.	NAME	VALUE	ERROR	NEGATIVE	POSITIVE	
1	BR_pKmm/pKJpsi	8.45329e-04	4.94404e-05	-4.83819e-05	5.06644e-05	
2	Ceff_pKmm_R1	3.00225e-03	1.02274e-04	-1.02020e-04	1.02517e-04	
3	Ceff_pKmm_R2	3.09017e-03	1.17709e-04	-1.17383e-04	1.18019e-04	
4	1/RpK	1.00027e+00	1.46103e-01	-1.40602e-01	1.52535e-01	
5	cL0Eex_R1	9.32260e+00	4.53590e-01	-4.53406e-01	4.53987e-01	
6	cL0Eex_R2	6.25476e+00	2.40253e-01	-2.40248e-01	2.40309e-01	
7	cL0I_R1	5.60056e+00	2.57772e-01	-2.57490e-01	2.58121e-01	
8	cL0I_R2	1.01929e+01	4.67664e-01	-4.67672e-01	4.67897e-01	
9	mean_mm_R1_L0M	5.62375e+03	1.27100e+00	-1.27052e+00	1.27676e+00	
10	mean_mm_R2_L0M	5.61842e+03	1.38155e+00	-1.38590e+00	1.39341e+00	
11	sigma_mm_R1_L0M	1.62945e+01	1.13695e+00	-1.09541e+00	1.17410e+00	
12	sigma_mm_R2_L0M	1.62994e+01	1.27798e+00	-1.28807e+00	1.30754e+00	
13	slope_ee_R1_L0Eex	-2.53524e-03	1.06379e-03		9.09069e-04	
14	slope_ee_R1_L0I	-3.01308e-03	6.63436e-04	-8.99085e-04	5.57811e-04	
15	slope_ee_R2_L0Eex	-2.73119e-03	7.82427e-04	-1.09278e-03	6.83113e-04	
16	slope_ee_R2_L0I	-1.80275e-03	4.10442e-04	-4.30826e-04	4.08433e-04	
17	slope_mm_R1_L0M	-3.97269e-03	4.90186e-04	-5.05749e-04	4.76916e-04	
18	slope_mm_R2_L0M	-3.84468e-03	7.14780e-04	-7.46741e-04	6.88954e-04	
19	yield_Jpsi_ee_R1_L0Eex	6.83615e+00	8.99439e-01	-8.99163e-01	8.99876e-01	
20	yield_Jpsi_ee_R1_L0I	9.56942e+00	1.09858e+00	-1.09814e+00	1.09922e+00	
21	yield_Jpsi_ee_R2_L0Eex	1.07238e+01	6.99318e-01	-6.99237e-01	6.99498e-01	
22	yield_Jpsi_ee_R2_L0I	1.16296e+01	6.98830e-01	-6.98429e-01	6.99300e-01	
23	yield_PR_pi0_ee_R1_L0Eex	2.44148e+01	1.02821e+01	-1.01061e+01	1.05321e+01	
24	yield_PR_pi0_ee_R1_L0I	3.01207e+01	1.62810e+01	-1.59066e+01	1.68925e+01	
25	yield_PR_pi0_ee_R2_L0Eex	1.08237e+01	1.03027e+01	-1.00902e+01	1.06061e+01	
26	yield_PR_pi0_ee_R2_L0I	1.79637e+01	1.38195e+01	-1.34804e+01	1.42553e+01	
27	yield_comb_ee_R1_L0Eex	2.22466e+01	1.14030e+01	-1.80976e+01	1.21708e+01	
28	yield_comb_ee_R1_L0I	7.42061e+01	1.85070e+01	-1.82348e+01	1.90712e+01	
29	yield_comb_ee_R2_L0Eex	3.81361e+01	1.30239e+01	-1.24054e+01	1.37224e+01	
30	yield_comb_ee_R2_L0I	6.16446e+01	1.64559e+01	-1.59571e+01	1.70263e+01	
31	yield_comb_mm_R1_L0M	2.11888e+02	1.70472e+01	-1.66885e+01	1.74119e+01	
32	yield_comb_mm_R2_L0M	9.47969e+01	1.18171e+01	-1.14192e+01	1.22324e+01	

Table 3.27: Output from `Minuit` of the fit (with "fake" efficiency ratios) shown in Fig. 3.104.

tainties:

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} = (8.45^{+0.45}_{-0.43}(\text{stat}) \pm 0.26(\text{syst})) \times 10^{-4}; \tag{3.19}$$

This result can be translated into the absolute branching fraction, using the value of $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ from the PDG $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi) = (3.17 \pm 0.04(\text{stat}) \pm 0.07(\text{syst}) \pm 0.34(\text{norm})^{+0.45}_{-0.28}(f_{\Lambda_b^0})) \times 10^{-4}$, where the first uncertainty is statistical, second systematic, third – due to the knowledge of the normalisation mode, and the fourth – due to the

knowledge of the Λ_b^0 production ratio. We group the first to third uncertainties, adding them in quadrature, as an uncertainty related to the knowledge of $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$. The fourth uncertainty is treated separately.

$$\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-) = (2.68 \pm 0.14(\text{stat}) \pm 0.08(\text{syst}) \pm 0.30(\text{norm})^{+0.38}_{-0.24}(f_{\Lambda_b^0})) \times 10^{-7}; \tag{3.20}$$

where the first uncertainty is statistical, second systematic, third – due to the knowledge of $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$, and the fourth – due to the knowledge of the Λ_b^0 production ratio.

The R_{pK}^{-1} (with fake efficiencies constructing it to be equal to unity) is found to be:

$$R_{pK}^{-1} = 1.00^{+0.15}_{-0.14}(\text{stat}) \pm 0.06(\text{syst}). \tag{3.21}$$

As these results have been obtained using unblinded datasets, they reflect the expected sensitivity on observables of interest.

The profile log-likelihood with this configuration is presented in Fig. 3.105 as a red curve. For comparison, the same likelihood without systematic uncertainties is overlaid as a blue curve.

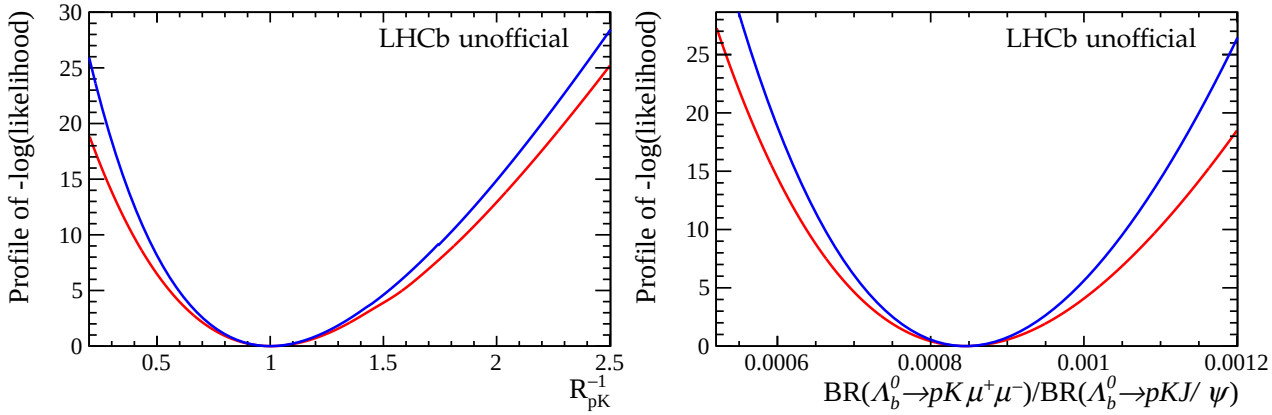


Figure 3.105: Red curve: Log-likelihood profile for the (left) R_{pK}^{-1} and (right) ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi$, in the fit (with "fake" efficiency ratios) shown in Fig. 3.104, including all the systematic uncertainties. Blue curve: same profiles, but without systematic uncertainties (reproduced from Fig. 3.101- 3.102) .

It should also be noted that in the configuration presented above, we force fake efficiencies in the way that in each trigger category, R_{pK}^{-1} is strictly equal to one. However, this does not account for possible statistical fluctuations. To study whether the fit converges properly in more realistic case, the fake efficiencies for each trigger category were randomly offset by two standard deviations (about 30%) either up or down. This mimics the case of 2σ fluctuations on the electron yields. Two efficiencies are scaled up and two – down, so that the resulting central value stays close to unity, but this does not imply that this central value is strictly 1 anymore. It has been checked that the fit anyway

converges properly, and that the uncertainty on R_{pK}^{-1} increases only mildly under these conditions.

3.10.5 Fit studies using pseudo-experiments

It is very important to check whether the simultaneous fit produces an unbiased pull distribution of the R_{pK}^{-1} parameter. To do so, 8000 pseudo-experiments were ran based on the fit model presented above, and the pull distribution of R_{pK}^{-1} is shown in Fig. 3.106, while the pull distribution of the ratios of $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$ and $\mathcal{B}(\Lambda_b^0 \to pKJ/\psi)$ is presented in Fig. 3.107. It can be seen that there is no significant bias on the R_{pK}^{-1} parameter and the ratio of BR's, and the coverage of uncertainties is good. The behaviour of other fit parameters approximately matches the one presented in the Sections 3.7.3.d and 3.7.4.d.

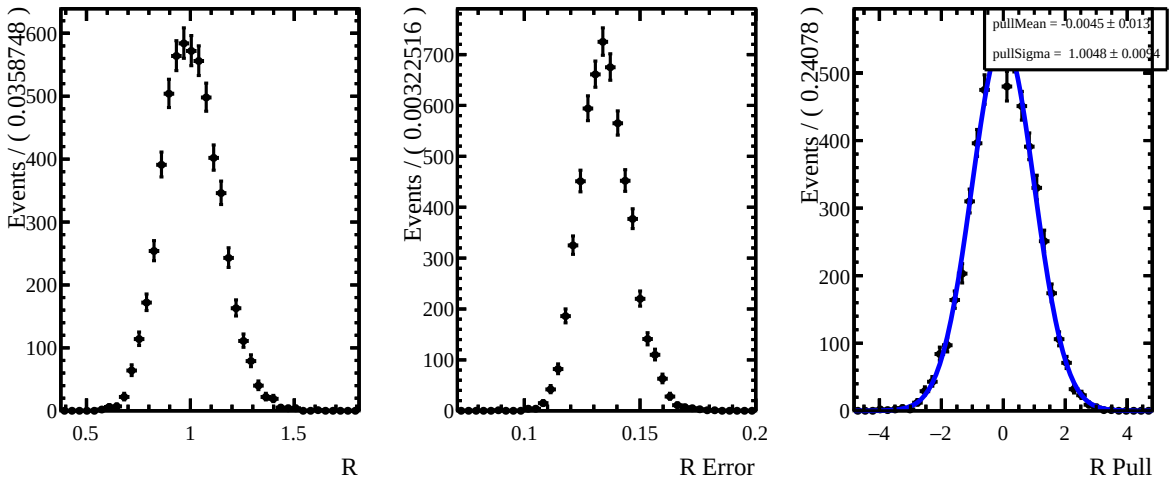


Figure 3.106: Pull distribution of R_{pK}^{-1}, as well as distributions of R_{pK}^{-1} and ΔR_{pK}^{-1} using 8000 pseudo-experiments based on the model presented in Fig. 3.100.

The values of R_{pK}^{-1} can also be extracted split per trigger category and run period. To do so, the simultaneous fit is run only on the relevant datasets (data taking period/trigger category). This fit configuration will be used for the cross-checks and tests of compatibility between the datasets. These configurations have also been validated with toys, and no significant bias was found.

3.10.6 Branching fraction of the rare decay $\Lambda_b^0 \to pK\mu^+\mu^-$ per dataset

This section does not cover the actual measurement of the branching fraction, which is performed in the simultaneous fit described above. Instead, it only presents a cross-check of the agreement of the values of the branching fraction, extracted using only muon mode information, between the trigger categories and datasets.

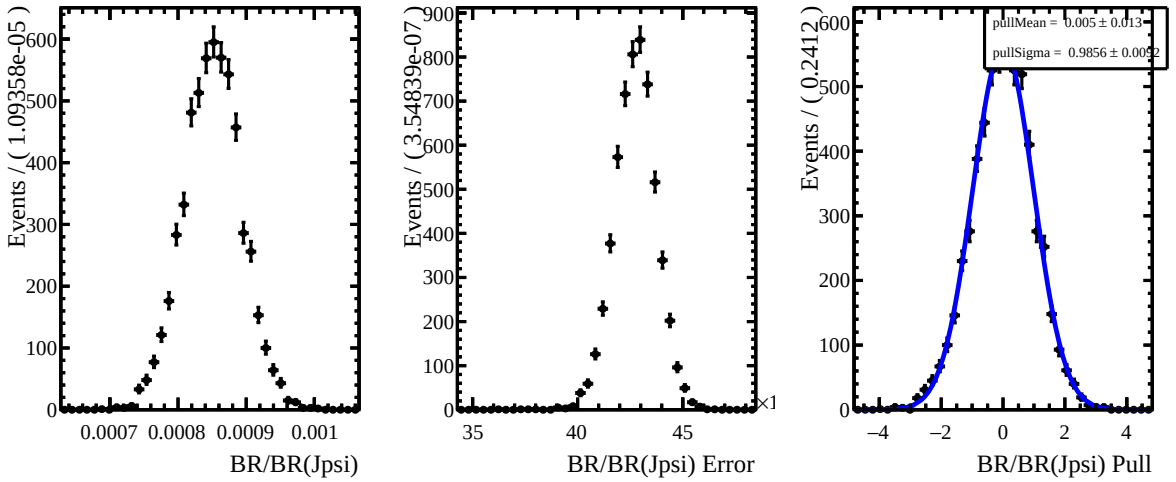


Figure 3.107: Pull distribution of the ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi$, as well as distributions of this ratio and its uncertainty, using 8000 pseudo-experiments based on the model presented in Fig. 3.100.

Figure 3.108 and Table 3.28 present a comparison of the ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$, including the L0M and L0I datasets. It should be noted that the L0M and L0I categories discussed in this section, are both inclusive, *i.e.* not mutually exclusive. It should also be noted that the statistical precision is much more limited in the L0I category, with the yields of about a quarter of those in the L0M category. Nevertheless, all the results are in a good agreement within one standard deviation.

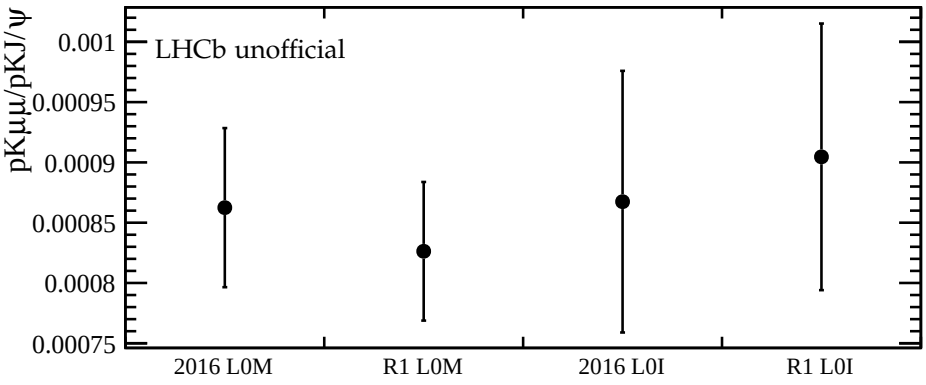


Figure 3.108: Comparison of the ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ in Run I and Run II. Both L0M and L0I categories are presented for comparison.

Table 3.28: Comparison of the ratio of branching fractions of $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ in Run I and Run II.

Category	Value
Run I, L0M	$(8.26 \pm 0.57) \times 10^{-4}$
Run II, L0M	$(8.62 \pm 0.66) \times 10^{-4}$
Run I, L0I	$(9.05 \pm 1.11) \times 10^{-4}$
Run II, L0I	$(8.67 \pm 1.09) \times 10^{-4}$

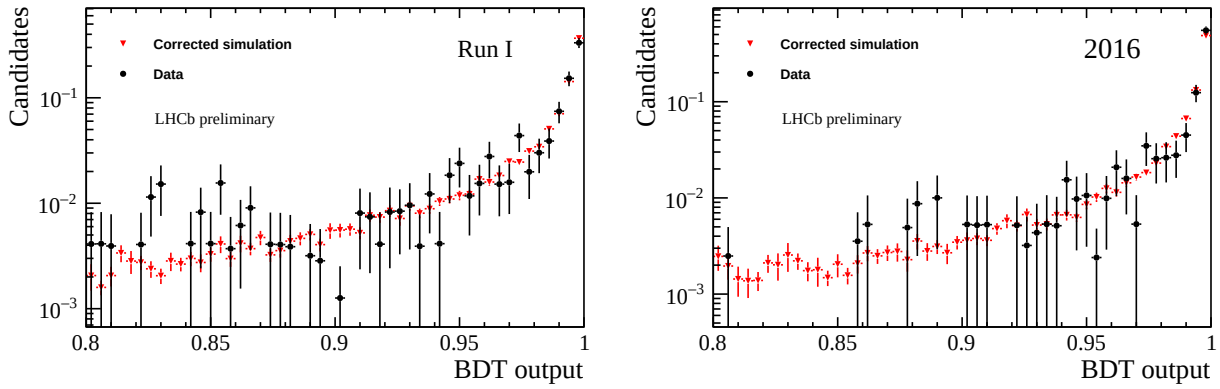


Figure 3.109: Data-simulation agreement in the BDT output variable in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay, for the (left) Run I and (right) Run II datasets.

3.10.6.a The BDT output variable distribution in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay

The data-simulation comparison in the BDT output variable is performed separately for Run I and Run II datasets, exploiting the default selection. Only the region of the BDT output values larger than 0.8 is used. An ${}_s\mathcal{P}lot$ technique was exploited in order to perform the background subtraction. The simulation has all default corrections applied. The comparison is presented in Fig. 3.109, and, accounting for a small statistics in data, a reasonable agreement can be seen.

3.10.7 The $m(pK)$ spectrum in the $\Lambda_b^0 \to pK\ell^+\ell^-$ decays

To extract the (pK) spectrum in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay, a fit was performed to the combined 2011-2012-2016 dataset. Both L0M and L0I categories were included, in order to gain additional statistics. An ${}_s\mathcal{P}lot$ technique was exploited in order to perform the background subtraction. The spectrum, not corrected by efficiency, is presented in Fig. 3.110 (left). The selection efficiency is shown in App. E.1, and it can be seen that it is almost flat in $m(pK)$, within uncertainties, which means the spectrum shown here represents properly the resonant structure of the decay.

For the $\Lambda_b^0 \to pKe^+e^-$ decay mode, an ${}_s\mathcal{P}lot$ has been performed relying on the fit presented in Sec. 3.9.5. It is assumed to be valid for extracting the $m(pK)$ spectrum as it

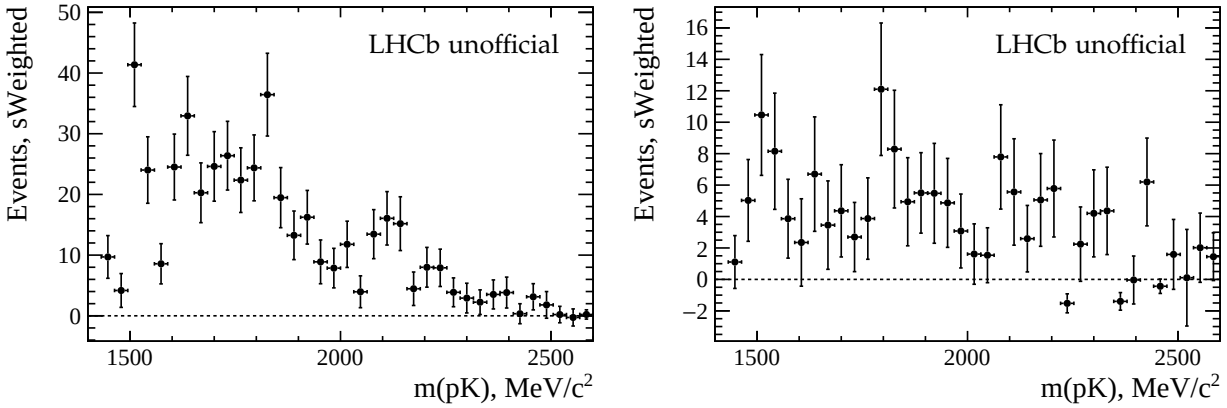


Figure 3.110: Background subtracted, proton-kaon invariant mass spectrum in the (left) $\Lambda_b^0 \to pK\mu^+\mu^-$ and (right) $\Lambda_b^0 \to pKe^+e^-$ decay modes (no efficiency corrections are applied), plotted with the same binning. The negative values on the vertical axis arise from negative sWeights.

does not correlate with the pKe^+e^- invariant mass, even though the background level is rather high, which complicates the statistical separation of signal. The spectrum, not corrected by efficiency, is presented in Fig. 3.110 (right). The selection efficiency is shown in App. E.2, and it can be seen that in all datasets this efficiency is almost flat in $m(pK)$, same as in the muon mode.

It can be seen that the spectrum has a complex resonant structure, while the relative abundance of different Λ^* resonances is somewhat different from the one observed in the $\Lambda_b^0 \to pKJ/\psi$ decay (Fig. 1.12). In addition to the prominent $\Lambda^*(1520)$ resonance, other structures can be seen in the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, which can be attributed to such states as $\Lambda^*(1690)$, $\Lambda^*(1820)$, $\Lambda^*(2110)$ etc. Some evidence for the $\Lambda^*(1520)$, $\Lambda^*(1820)$ and $\Lambda^*(2110)$ states can also be seen in the $\Lambda_b^0 \to pKe^+e^-$ channel, even though this check is statistically limited and statistical fluctuations may arise. One has to collect a much larger dataset and perform an amplitude analysis in order to identify structures observed in the spectrum as certain resonances.

A similar check can be performed for the q^2 distributions in $\Lambda_b^0 \to pK\ell^+\ell^-$ decay modes, as shown in Fig. 3.111. However, two caveats have to be taken into account when comparing the two. First, in this case the ${}_s\mathcal{P}lot$ technique is not strictly valid for the $\Lambda_b^0 \to pKe^+e^-$ decay mode, due to a significant correlation between the q^2 and $m(pKe^+e^-)$ arising due to bremsstrahlung radiation. Therefore, this check can only provide a glimpse on the underlying distribution. Second, the efficiency trends in q^2 are not flat for the electron modes (Appendix E.2), so ideally one has to correct for the efficiency. The latter exercise is not straightforward as the proper combination of four independent datasets with different efficiency distributions has to be performed. Even without this correction, the prominent photon pole can be observed in both distributions.

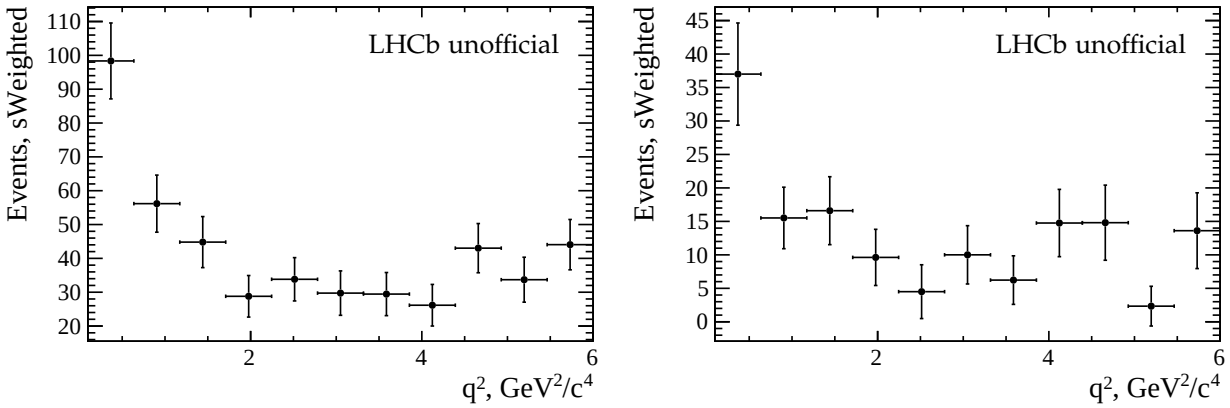


Figure 3.111: Background subtracted, q^2 distribution in the (left) $\Lambda_b^0 \to pK\mu^+\mu^-$ and (right) $\Lambda_b^0 \to pKe^+e^-$ decay modes (no efficiency corrections are applied), plotted with the same binning. The prominent photon pole can be observed in both distributions.

3.11 Conclusions and prospects

The test of lepton universality R_{pK}^{-1}, performed on the part of collected LHCb dataset, allows to obtain a statistical uncertainty of the order 15%. In addition, the branching ratio of the rare decay $\Lambda_b^0 \to pK\mu^+\mu^-$ is measured in the region $0.1 < q^2 < 6$ GeV$^2/c^4$ and $m(pK) < 2600$ MeV$/c^2$, with a 5.3% statistical precision. The decay mode $\Lambda_b^0 \to pKe^+e^-$ is observed for the first time, with the statistical significance of more than eight standard deviations. At the moment of writing this thesis, this analysis was in review within the LHCb Collaboration. The step-by-step strategy for unblinding the analysis is followed.

After concluding the current analysis, a possible continuation would be to perform a legacy measurement using the entire Run I and Run II dataset collected by the LHCb experiment. Besides the improvements in the statistical precision, and potential optimisations in the analysis procedure, one can consider splitting the measurement in two q^2 bins. Alternatively, the dataset can be split in $m(pK)$ bins, allowing to probe separately regions dominated by certain resonances. This kind of measurement can be especially interesting to perform in the coming years, after a larger dataset is collected. It would allow to directly compare the measured results to the theoretical predictions for a given resonance, and to study whether the potential NP effects are spin-dependent. Additional knowledge can be gained by studying rare decays of other b baryons such as Ξ_b^-.

Finally, the observation of the $\Lambda_b^0 \to pKe^+e^-$ decay would open a window to the angular analysis of this decay mode, or the measurement of its differential branching fraction. This would provide an additional set of observables to test the lepton universality or other NP effects.

CHAPTER 4

First observation and study of the decay $\Xi_b^0 \to pKJ/\psi$

4.1 Introduction and strategy

In this chapter, the first observation of the decay $\Xi_b^0 \to pKJ/\psi$ is presented. This decay is Cabibbo-suppressed with respect to the channel $\Lambda_b^0 \to pKJ/\psi$, used as a normalisation mode in this analysis. It can be considered as a baryonic analog of the decay $B_s^0 \to K^{*0}J/\psi$, observed by the LHCb Collaboration [158, 159], and being Cabibbo-suppressed with respect to the decay $B^0 \to K^{*0}J/\psi$. This suppression happens due to the CKM matrix element V_{cd} entering instead of V_{cs}.

Feynman diagram describing the decay $\Xi_b^0 \to pKJ/\psi$ is presented in Fig. 4.1.

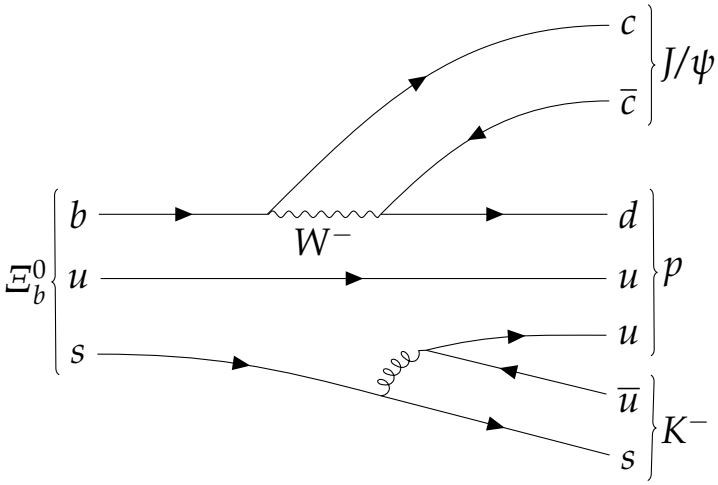


Figure 4.1: Feynman diagram of the decay $\Xi_b^0 \to pKJ/\psi$.

As it has already been discussed in Sec. 1.4, multiple resonances, such as

$\Lambda^*(1520), \Lambda^*(1670)$ *etc.*, populate the pK spectrum of the $\Lambda_b^0 \to pKJ/\psi$, but also $\Xi_b^0 \to pKJ/\psi$ decays. This rich structure has been characterised by the LHCb Collaboration in the $\Lambda_b^0 \to pKJ/\psi$ decays [4]. In addition, the structures in the pJ/ψ spectrum were observed, consistent with the pentaquark states. The decay $\Xi_b^0 \to pKJ/\psi$ is an alternative channel with the same exact final state, allowing for searches for exotic contributions in the Ξ_b^0 decays. This decay has a slightly larger phase space due to $m(\Xi_b^0) > m(\Lambda_b^0)$, and has somewhat different dynamics due to the fact that the s-quark is already present in the original baryon. In particular, the decay $\Xi_b^0 \to pKJ/\psi$ violates strong isospin, this might have an important impact on the resonant structure.

In order to minimise the systematic uncertainties on the measurement of interest, the following ratio is measured:

$$R = \frac{BR(\Xi_b^0 \to pK^- J/\psi\,(\to \mu^+\mu^-))}{BR(\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-))} \times \frac{f_{\Xi_b^0}}{f_{\Lambda_b^0}} \tag{4.1}$$

where $f_{\Xi_b^0}$ and $f_{\Lambda_b^0}$ are the hadronisation fractions of b quark to the corresponding b baryons. The first of them has never been measured, that is the reason why they enter the equation explicitly.

It should be noted that this analysis was born as a byproduct of the R_{pK}^{-1} lepton universality analysis presented in the previous chapter. When studying the background composition to the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ channel, the $\Xi_b^0 \to pK^- J/\psi\,(\to \mu^+\mu^-)$ signal was observed, which resulted in the current analysis. As a consequence, much of the analysis strategy, including the choice of the datasets exploited, data-simulation corrections *etc.*, were inherited from the original analysis. On the other hand, the selection was readjusted in order to get a better sensitivity to the $\Xi_b^0 \to pK^- J/\psi\,(\to \mu^+\mu^-)$ decay, as described in the Sec. 4.2.

The $\Xi_b^0 \to pKJ/\psi$ decay will be referred to as `signal` mode while the $\Lambda_b^0 \to pKJ/\psi$ will be referred to as `normalisation` mode.

4.2 Offline selection

In this section, the offline selection requirements applied to the data in order to get a reasonable signal versus background separation, are described. The acquired knowledge of the different background sources previously described in the Section 3.3, is used throughout the selection steps in order to optimise the cuts and to build a multivariate classifier to reach the best sensitivity for the measurement of interest.

The summary of the offline selection requirements applied throughout this analysis, is presented in the Table 4.1 for Run I, and Table 4.2 for Run II. Further details are provided below.

Table 4.1: Offline selection requirements for Run I.

Type	Particle(s)	Requirement
Quality and acceptance	μ	`InAccMuon`
	p, K, μ	`hasRich`
	p, K, μ	`GhostProb<0.3`
PID	p	$p > 10\,\text{GeV}/c$, $p_T > 1$ GeV/c
	K	$p > 2\,\text{GeV}/c$, $p_T > 250$ MeV/c
	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV/c
	p, K, μ	$p < 150$ GeV/c
	p	`MC12TuneV2_ProbNNp>0.5`, `MC12TuneV3_ProbNNk<0.8`, `MC12TuneV3_ProbNNpi<0.7`
	K	`MC12TuneV3_ProbNNk>0.4`, `MC12TuneV3_ProbNNp<0.8`
	μ	`MC12TuneV3_ProbNNmu>0.1`
Mass windows	dimuon	$3047 < \text{m}(\mu^+\mu^-) < 3147$ MeV/c^2
Background vetoes	B^+	$m(KJ/\psi) < 5200$ MeV/c^2, $m(pJ/\psi)_{p\leftarrow K} < 5200$ MeV/c^2
	ϕ	$\lvert m(pK)_{p\leftarrow K} - 1020\rvert > 12$
	B_s^0	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftarrow K} - 5366\rvert > 20$
	B^0	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftarrow \pi} - 5279\rvert > 20$
	Λ_b^0 swap	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftrightarrow K} - 5619\rvert > 8$
Combinatorial bkg	–	`BDT>0.2`
Clone tracks	–	$\theta_{p\mu^+} > 0.5$ mrad, $\theta_{K^-\mu^-} > 0.5$ mrad

4.2.1 Subdetector acceptance

All final state tracks of the decays of interest – protons, kaons and muons, are required to be in the acceptance of the RICH system. This is done in order to ensure that the PID response for all tracks is valid. Muons are also required to fall in the acceptance of the muon system. This is required to control properly both PID and trigger responses.

4.2.2 Particle identification requirements

As it has already been discussed, loose PID requirements are already placed in the stripping line (Table 3.3). Tighter PID requirements are added on top of the stripping, they are based on the misID background composition discussed above, and are different between Run I and II. This is due to differences in the global PID optimisation algorithms (PID tunes). The requirements are tighter than those used for the lepton universality analysis, as the small Ξ_b^0 signal is located in the region contaminated by the misidentification backgrounds.

Table 4.2: Offline selection requirements for Run II.

Type	Particle(s)	Requirement
Quality and acceptance	μ	`InAccMuon`
	p, K, μ	`hasRich`
	p, K, μ	`GhostProb<0.3`
PID	p	$p > 10\,\text{GeV}/c$, $p_T > 1$ GeV/c
	K	$p > 2\,\text{GeV}/c$, $p_T > 250$ MeV/c
	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV/c
	p, K, μ	$p < 150$ GeV/c
	p	`MC15TuneV1_ProbNNp>0.5`, `MC15TuneV1_ProbNNk<0.8`, `MC15TuneV1_ProbNNpi<0.7`
	K	`MC15TuneV1_ProbNNk>0.4`, `MC15TuneV1_ProbNNp<0.8`
	μ	`MC15TuneV1_ProbNNmu>0.1`
Mass windows	dimuon	$3047 < \text{m}(\mu^+\mu^-) < 3147$ MeV/c^2
Background vetoes	B^+	$m(KJ/\psi) < 5200$ MeV/c^2, $m(pJ/\psi)_{p\leftarrow K} < 5200$ MeV/c^2
	ϕ	$\lvert m(pK)_{p\leftarrow K} - 1020\rvert > 12$
	B_s^0	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftarrow K} - 5366\rvert > 20$
	B^0	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftarrow \pi} - 5279\rvert > 20$
	Λ_b^0 swap	$\lvert m(pKJ/\psi)^{J/\psi}_{p\leftrightarrow K} - 5619\rvert > 8$
Combinatorial bkg	–	`BDT>0.2`
Clone tracks	–	$\theta_{p\mu^+} > 0.5$ mrad, $\theta_{K^-\mu^-} > 0.5$ mrad

As it has already been discussed, the usage of PID variables is valid only in the kinematical regions covered by the calibration samples used to train the corresponding PID response. The required cuts for Run I and II are different for the proton (250 and 1000 MeV/c), but for consistency we apply the same (tighter) cut for all datasets, which also allows to get a cleaner data sample. The following cuts are applied on the transverse momenta of the final state particles to match this requirement: $p_T(p) > 1000$ MeV/c, $p_T(K) > 250$ MeV/c and $p(K) > 2$ GeV/c, $p_T(\mu) > 800$ MeV/c. In addition, the momentum of the proton is required to be larger than 10 GeV/c.

4.2.3 Mass vetoes

Most of the misidentification and over-reconstructed backgrounds are located in the upper sideband of the Λ_b^0 signal – exactly where the Ξ_b^0 signal is located. This makes the usage of mass vetoes necessary for the dominant background contributions, in order to improve the signal-to-background ratio for the Ξ_b^0 invariant mass peak.

It should be noted that broad vetoes are rather dangerous as they are likely to distort

the shape of the combinatorial background, making the invariant mass modeling to a required degree of precision a complicated task. In order to overcome this limitation, a J/ψ mass constraint is applied wherever possible, in order to significantly improve the mass resolution of the vetoed particle. For the most prominent backgrounds, a 3σ veto is applied (where σ is a typical value of the mass resolution), while for the minor background modes a 1σ veto is considered to be sufficient.

4.2.3.a B^+ meson

The cut $m(K\ell^+\ell^-) < 5200$ MeV/c^2 is applied to remove the $B^+ \to K\ell^+\ell^-$ candidates combined with a random proton. In addition, the cut $m(p\ell^+\ell^-)_{p\leftarrow K} < 5200$ MeV/c^2 removes similar candidates, but with the kaon misidentified as a proton.

4.2.3.b $\phi(1020)$ meson

The cut $|(m(pK)_{p\leftarrow K} - 1020| > 12$ MeV/c^2 allows to remove the misidentified $\phi \to K^+K^-$ candidates combined with dilepton pair (mostly coming from B_s^0 decays). This is a very powerful requirement suppressing not only fully-reconstructed (discussed below), but also partially reconstructed B_s^0 decays with a ϕ meson in the final state.

4.2.3.c B_s^0 meson

The ϕ veto discussed in the previous section does not account for all the possible misidentification background contributions coming from the B_s^0 meson decays, leaving room for the decays into excited K^+K^- states such as $f_2'(1525)$. This remains the major of all background contributions in the Ξ_b^0 region, so it is essential to suppress it by means of a mass veto. The cut $|m(pKJ/\psi)^{J/\psi}_{p\leftarrow K} - 5366| > 20$ MeV/c^2 (where superscript J/ψ means applying the J/ψ mass constraint) is applied. An effect of this requirement can be seen in Fig. 4.2(a). One can see that the efficiency of this requirement is not the same for the Λ_b^0 and Ξ_b^0 decay modes due to a different phase space: the veto is crossing close to the 'center of gravity' of the distribution of $\Lambda_b^0 \to pKJ/\psi$ events, however, is far in the tail for the $\Xi_b^0 \to pKJ/\psi$ events. Same stands for the vetoes discussed below.

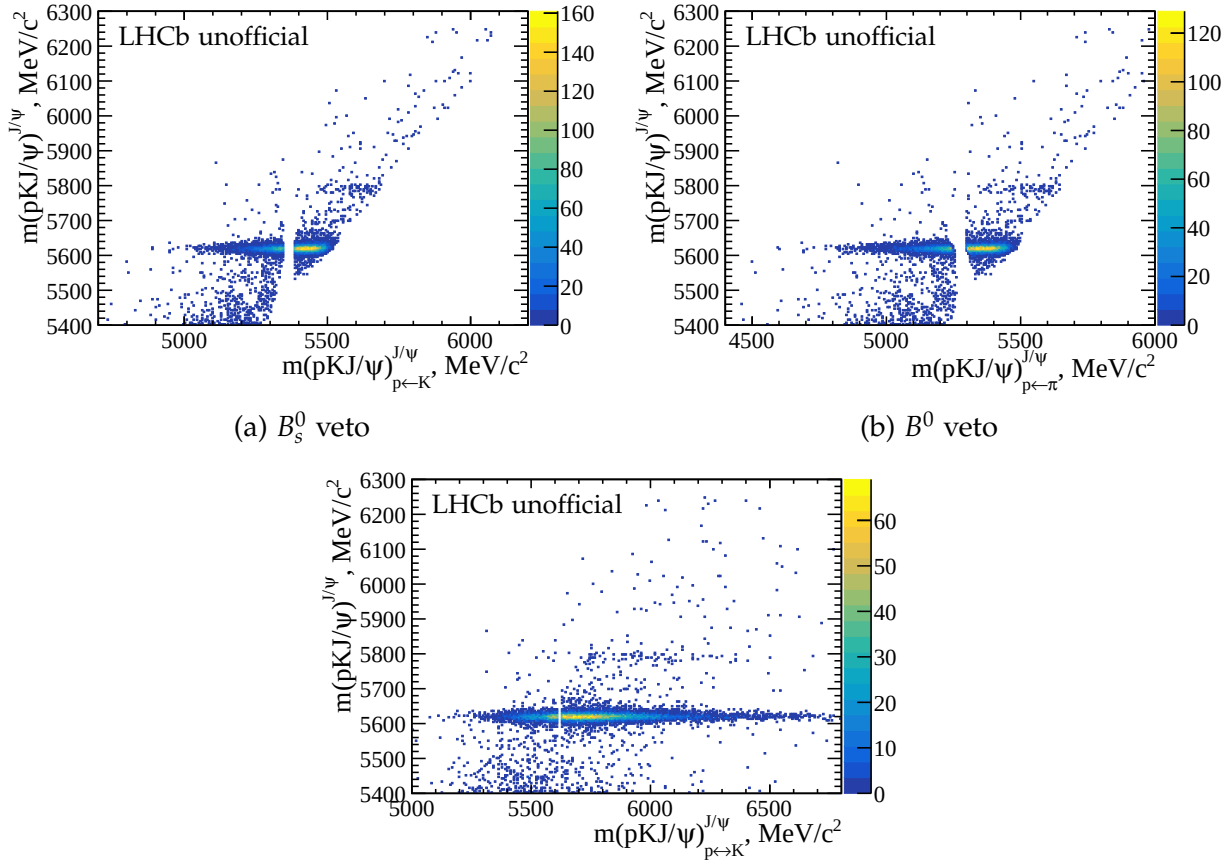
4.2.3.d B^0 meson

Misidentified decays of $B^0 \to J/\psi K\pi$ are the second largest source of background in the Ξ_b^0 mass region. The cut $|m(pKJ/\psi)^{J/\psi}_{p\leftarrow\pi} - 5279| > 20$ MeV/c^2 is applied. An effect of this requirement can be seen in Fig. 4.2(b).

4.2.3.e ID swap of the Λ_b^0 signal

Decays of $\Lambda_b^0 \to pKJ/\psi$, with the swapped identities of the proton and kaon, represent a minor source of the background. The cut $|m(pKJ/\psi)^{J/\psi}_{p\leftrightarrow K} - 5619| > 8$ MeV/c^2 is applied.

An effect of applying this requirement on data can be seen in Fig. 4.2(c).



(a) B_s^0 veto

(b) B^0 veto

(c) Λ_b^0 ID swap veto

Figure 4.2: Effect of the mass vetoes on the misidentified b hadrons, shown as a function of the Λ_b^0 invariant mass and the $m(pKJ/\psi)$ with a corresponding misidentification hypothesis. J/ψ mass constrain is applied throughout. The prominent horizontal band corresponds to the $\Lambda_b^0 \to pKJ/\psi$ decay, while the less prominent one above it – to the $\Xi_b^0 \to pKJ/\psi$ decay. The complete selection is applied in these plots.

In addition, it has to be checked that these vetoes do not create a fake peak or a step-like behavior in the signal region, in the sample composed of combinatorial background. To do so, the sample of data from the negative BDT output region was taken as a proxy for the background sample. A comparison of its invariant mass shape before and after applying the mass vetoes can be seen in Fig. 4.3. Indeed, no fake peak or step-like behavior is seen in the region where Λ_b^0 or Ξ_b^0 mass peaks are located. However, the invariant mass shape changes its slope below 5500 MeV/c^2. This is considered to be not problematic, as this region is characterised by the presence of partially reconstructed backgrounds, which behave in a similar manner, so would account for this slope change in the invariant mass fit.

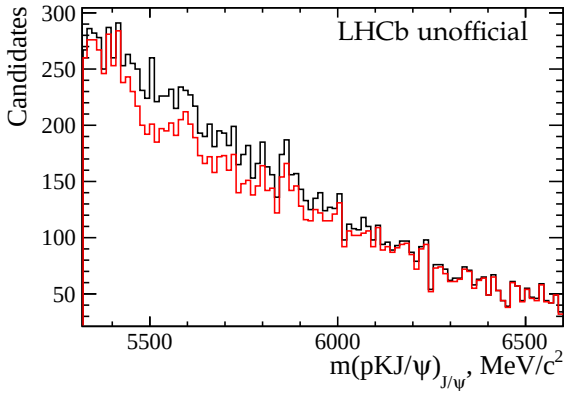


Figure 4.3: Effect of the mass vetoes on the smooth background. Black: original distribution; red: distribution after the background vetoes applied. No fake peak is created in the region above 5500 MeV/c^2.

4.2.3.f Clone tracks

The angle between two tracks of a same charge is required to be larger than 0.5 mrad. This requirement has almost 100% efficiency on signal, and allows to suppress the clone tracks.

4.2.4 Multivariate selection

Similarly to what was described for the LU analysis, the main purpose of the multivariate classifier is to reject the combinatorial background. The gradient boosting decision tree (BDTG) method was chosen for the training.

4.2.4.a Training of the multivariate classifier

Separate BDTs were trained for Run I and Run II to take into account differences in the kinematics, running conditions etc, however, the same BDT was used for 2011 and 2012 given the little difference in the underlying kinematics. Therefore, two different classifiers are used in the current analysis (Run I/II).

As a signal proxy, a set of truth-matched simulated candidates, corresponding to the $\Xi_b^0 \to pKJ/\psi$ decay mode, was used. A sample of the data from the upper sideband of the signal was used as a background proxy. The upper sideband was defined as $m(pKl^+l^-) > 5825$ MeV/c^2; and was limited from above by a stripping cut at approximately $m(pKl^+l^-) = 6780$ MeV/c^2.

The output of the classifier provides a new variable that is then used to apply one of the final selection cuts, before performing the fit to data in order to measure the observables of interest. This implies that, prior to the training, the selection on detector acceptance, PID, J/ψ invariant mass window, *etc.*, were applied to both signal and background samples in a consistent manner. This ensures the absence of mismatches in the kinematic cuts between signal and background proxy samples, which could be picked up by the BDT

and reduce its performance. Also, both samples were required to pass the L0 and HLT selections.

The training setup is presented in the Table 4.3.

Table 4.3: BDTG training setup.

	2012 and 2016
N_{trees}	500
max tree depth	3
Shrinkage	0.11
Bagging fraction	0.73
N_{cuts}	17

The PID, multiplicity and kinematical weights, as well as the reweighting of the pK spectrum (see Section 3.5), were applied to the simulation samples in order to correct for the data-simulation discrepancies in the latter.

A set of 15 variables was used for the training. Similarly to the lepton universality analysis, this set is already a result of an iterative removal of variables having low separation power, from the initial, much larger set.

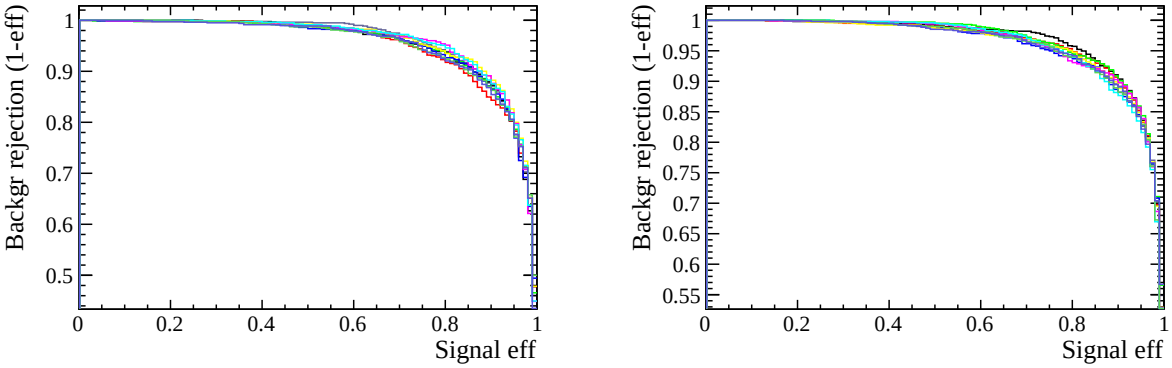
The list of variables used includes:

- $\beta' = \frac{p(J/\psi)-p(p)-p(K)}{p(J/\psi)+p(p)+p(K)}$,
- $p_T(\Lambda_b^0)$, $p_T(\Lambda^*)$, $p_T(\Lambda^*) - p_T(J/\psi)$,
- $\chi^2_{IP}(\Lambda_b^0)$, $\chi^2_{IP}(\Lambda^*)$,
- $DIRA(\Lambda_b^0)$,
- $\chi^2_{FD}(J/\psi)$,
- $\chi^2_{DTF}/n_{dof}(\Lambda_b^0)$,
- $\chi^2_{vertex}(J/\psi)$,
- $(\chi^2_{IP})^{min}(p, K)$, $(\chi^2_{IP})^{min}(\ell^+, \ell^-)$,
- $p(p)$,
- $\eta(p) + \eta(K)$,
- $\eta(J/\psi)$.

Some of these variables have significant correlations between themselves in the signal sample, however the correlations are much smaller for the background sample, and removal of one of these variables degraded significantly the BDTG performance.

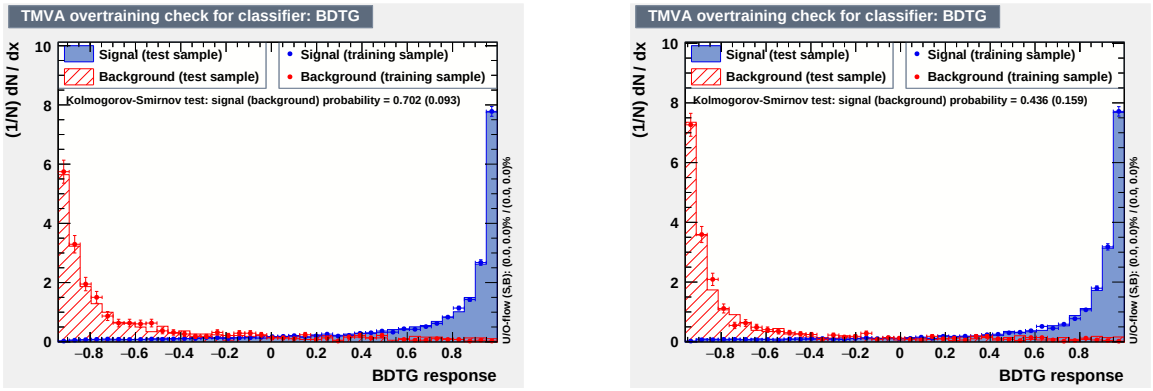
A *k-fold* approach was adopted for the BDT training, with $k = 10$: 9/10 of the dataset were used to train the BDT applied afterwards to the remaining 1/10, and the procedure repeated 10 times.

The resulting ten receiver operational characteristic (ROC) curves are shown in Fig. 4.4. The little differences between them can be assigned to the limited size of the training samples. In addition, the classifier was checked to have no overtraining, by means of Kolmogorov-Smirnov (KS) test values. An example of comparison of training and testing distributions is shown in Fig. 4.5.



(a) 2012 (b) 2016

Figure 4.4: ROC curves for ten BDT folds, each shown in a different colour.

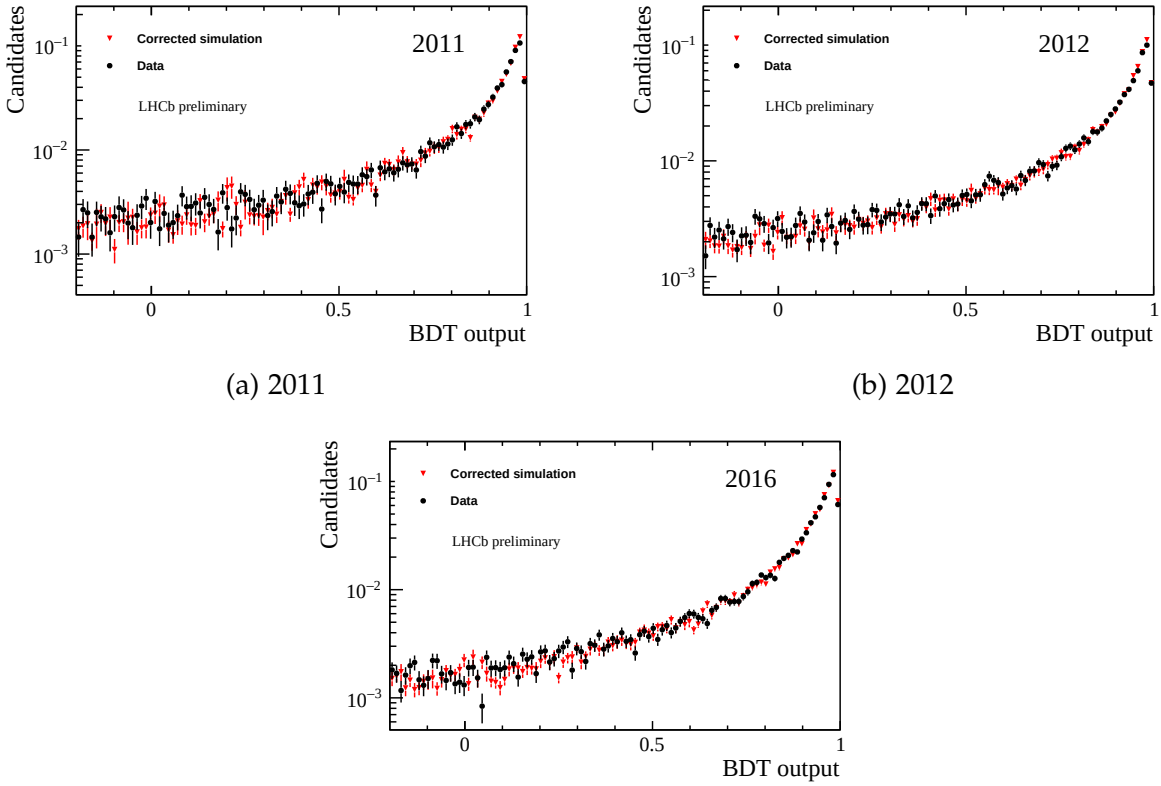


(a) 2012 (b) 2016

Figure 4.5: BDT response distributions for signal and background proxy samples, compared between the training and testing samples.

When using a large number of variables in the BDT classifier, one has to be careful and make sure not to pick up any discrepancies between the data and simulation resulting in a mismodeled BDT distribution. The most stringent test of the BDT validity is comparison of the BDT output variable between the data and simulation, shown in Fig. 4.6. The data

distribution is accessed via the ${}_s\mathcal{P}lot$ technique, while the simulation has the set of the discussed corrections (PID, multiplicity, kinematics *etc.*) applied. The fit to the data is performed with a loose BDT cut of -0.2: this cut enters in the first step of the preselection applied due to the large background at low BDT values making an ${}_s\mathcal{P}lot$ unreliable.



(a) 2011

(b) 2012

(c) 2016

Figure 4.6: Comparison of the BDT output variable, between the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ data (black circles) and simulation (red triangles), for different years of data taking.

No significant differences are seen between the data and simulation distributions. This also validates usage of a single BDT for 2011 and 2012 samples.

The BDT cut value was not explicitly optimised, but a working point was required to provide 90% background rejection. This corresponds to a cut value of 0.2, and gives about 85% signal efficiency. Tighter BDT values are not optimal due to a small signal statistics.

It is also important to make sure that the BDT will not create fake peaks out of a flat background sample. To check that, the sample of combinatorial background is taken from the simulation, following the same procedure as described for the lepton universality analysis. The trends of the mean BDT value as a function of the invariant mass in these samples are shown in Fig. 4.7, and, even though the descending trend is seen, it has a very linear behavior and does not create any fake peak around the known Λ_b^0 or Ξ_b^0 mass.

Finally, another test proving the BDT validity is a comparison of its efficiency for the $\Lambda_b^0 \to pKJ/\psi$ and $\Xi_b^0 \to pKJ/\psi$ decays. It can be seen that the BDT efficiency is within

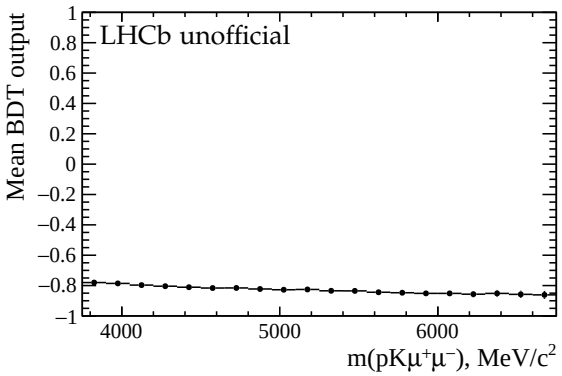


Figure 4.7: Mean BDT value as a function of invariant mass in a sample of combinatorial background. No peaking structures are seen.

1% difference (which might come from a slightly different kinematics) between the two modes, as expected.

Table 4.4: BDT cut efficiency for $\Lambda_b^0 \to pKJ/\psi$ and $\Xi_b^0 \to pKJ/\psi$ modes, %.

Year	$\Lambda_b^0 \to pKJ/\psi$	$\Xi_b^0 \to pKJ/\psi$
2011	83.7	84.4
2012	82.9	83.6
2016	87.7	88.3

4.2.5 Multiple candidates

The strategy is similar to the one adopted in the lepton universality analysis. In data, about 0.45% (0.42%) of candidates are removed for Run I (Run II), in the entire $m(pKJ/\psi)$ mass range, which is a negligibly small number.

4.3 Simulation samples and their calibration

Dedicated $\Xi_b^0 \to pK^-J/\psi\,(\to \mu^+\mu^-)$ simulation samples have been used for this analysis. In addition, the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ simulation samples have been reused from the R_{pK}^{-1} analysis.

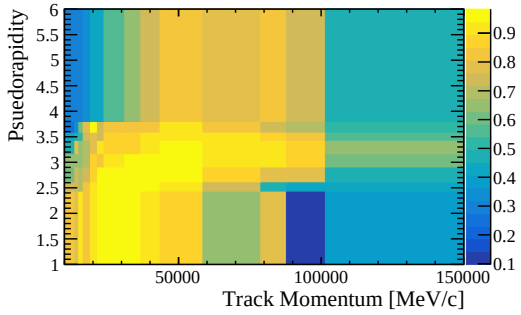
The calibration of the simulation has been inherited from the R_{pK}^{-1} analysis, and is described in Sec. 3.5. The corrections (except for the particle identification, which are discussed in the Section 4.3.1) are computed using data and simulation samples of the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ decay mode as being the most abundant in the current analysis, and consequently extrapolated to the signal channel.

Strictly speaking, the corrections computed using the Λ_b^0 decay channel do not necessarily represent well the Ξ_b^0 kinematics. However, in one previous LHCb analysis [44] it has already been shown that the kinematical distributions of Λ_b^0 and Ξ_b^0 are very similar in data, using the abundant decay channels $\Xi_b^0 \to \Xi_c^+ \pi^-$ and $\Lambda_b^0 \to \Lambda_c^+ \pi^-$. It is then concluded in a recent analysis measuring the production rate of Ξ_b^- baryons [160] that applying the Λ_b^0 kinematical weights to the Ξ_b^- channel results in a good data-simulation agreement. Given that in our case the final state is identical, possible biases are even further minimised. Consequently, the corrections computed using the $\Lambda_b^0 \to pKJ/\psi$ channel are applied throughout this analysis to the $\Xi_b^0 \to pKJ/\psi$ mode, except for the cases discussed below.

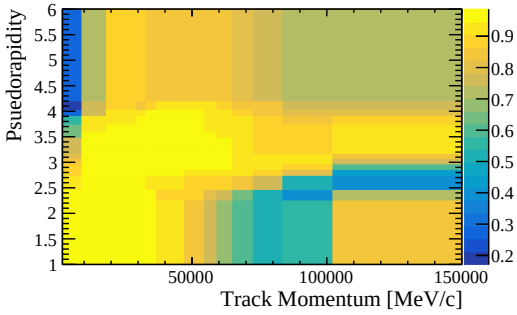
4.3.1 PID calibration maps

The PID requirements applied in this analysis are tighter than the ones in the lepton universality analysis. Therefore, the PID calibration maps have to be recomputed.

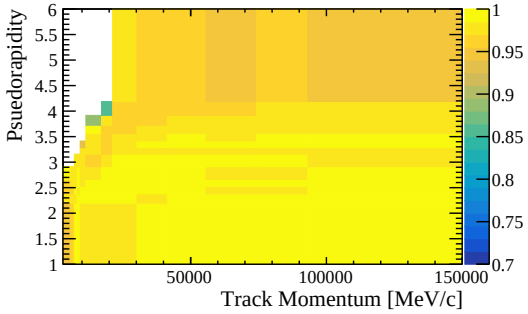
The resulting maps for 2011, 2012 and 2016 conditions are shown in Fig. 4.8, 4.9, 4.10. For simplicity, only the maps for MagDown conditions are shown, although in practice all the maps are computed for both polarities and applied appropriately.



(a) Proton

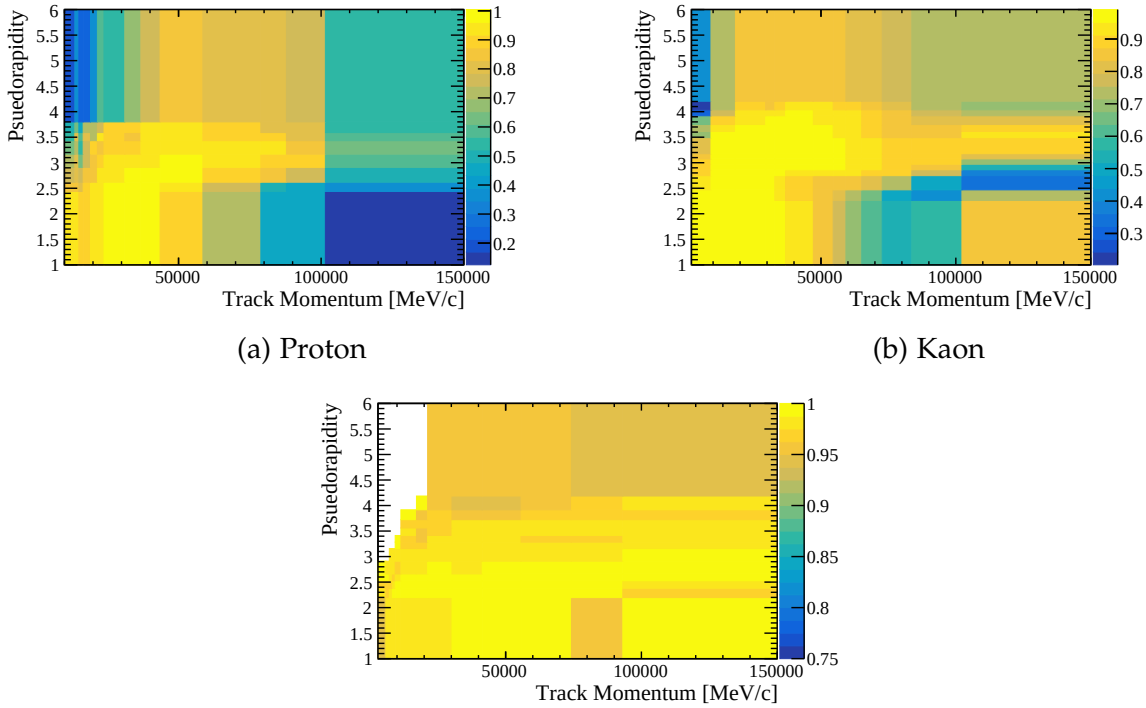


(b) Kaon



(c) Muon

Figure 4.8: PID efficiency maps for 2011 MagDown.



(a) Proton

(b) Kaon

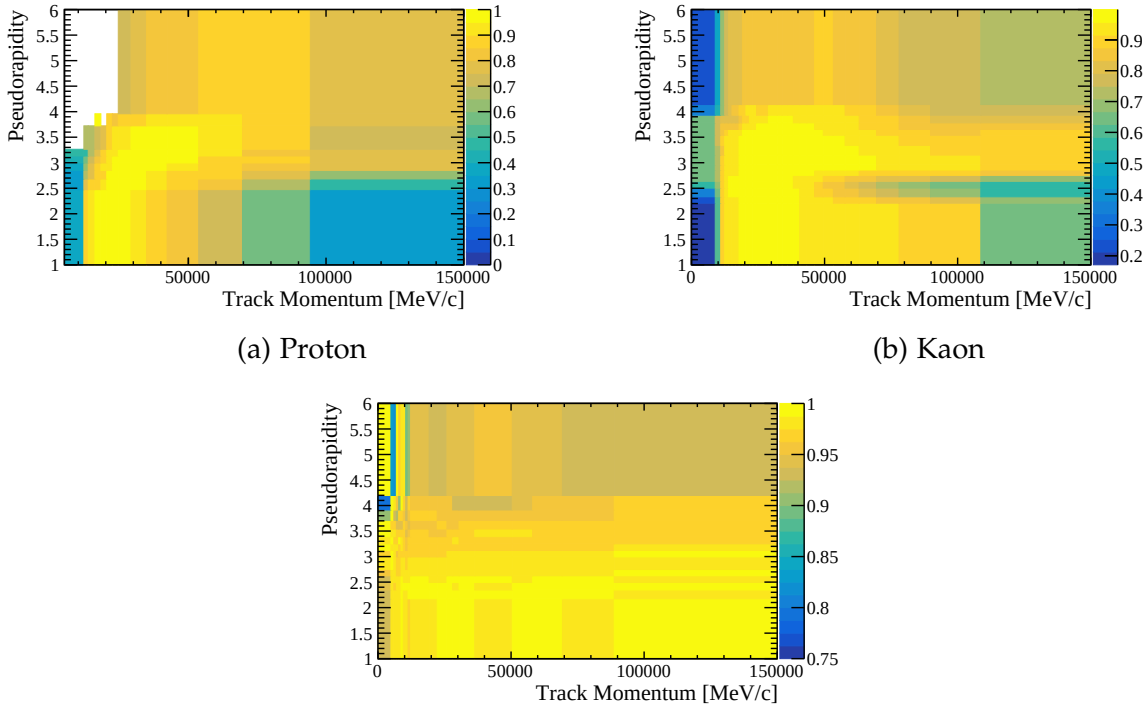
(c) Muon

Figure 4.9: PID efficiency maps for 2012 MagDown.

4.3.2 Decay time and its acceptance

As it has been discussed in Sec. 3.5.5, the generated Λ_b^0 lifetime is different from the one reported in Ref. [6]. This problem does not appear for the Ξ_b^0 simulation samples, as the generated value of the lifetime ($\tau(\Xi_b^0) = 1.49$ ps) is very close to the one reported in the PDG ($\tau(\Xi_b^0) = 1.479 \pm 0.031$ ps).

As it has also been discussed in Sec. 3.5.5, there is a residual data-simulation disagreement in Run I samples in the Λ_b^0 decay time distribution, mostly in the region of small decay time values, where the proper modeling of the decay time acceptance is essential. This is believed to happen due to the fact that some of the reconstruction and selection steps affecting the decay time acceptance (for instance, related to the tracking and subsequent selections) are not ideally described by the simulation. This correction is applied for the Λ_b^0 and Ξ_b^0 channels in Run I, as is believed to deal with the lifetime acceptance rather than with a given particle. Impact of (not) applying this correction to both Λ_b^0 and Ξ_b^0 decay channels is negligible, and is considered as a minor source of systematic uncertainty.



(a) Proton

(b) Kaon

(c) Muon

Figure 4.10: PID efficiency maps for 2016 MagDown.

4.4 Efficiencies

Given the closeness of the masses of Λ_b^0 and Ξ_b^0 baryons and their properties, most of the efficiencies are expected to be very similar between the two decay modes. This is however not the case for the mass vetoes, as it was already shown in Section 4.2 and Fig. 4.2. Therefore, it is important to keep track of efficiencies for each decay mode.

The total selection efficiency is defined as a product of geometric, filtering, reconstruction and selection efficiencies, as it has been already described by Eq. 3.7. For simplicity, only the total efficiencies are discussed in this section.

4.4.1 Total efficiencies

The values actually used in the computations are the total efficiencies, and they are evaluated according to Eq. 3.7. These final values are presented in the Table 4.5.

Table 4.5: Total efficiencies (including geometric, filtering, reconstruction, selection, as well as accounting for the fit range), 10^{-3}. The ratio of efficiencies between the signal and normalisation channel is also presented.

Channel	Run I	Run II
$\Lambda_b^0 \to pKJ/\psi$	2.50	3.67
$\Xi_b^0 \to pKJ/\psi$	3.14	4.49
Ratio Ξ_b^0/Λ_b^0	1.26	1.22

4.5 Invariant mass fits

In this section, the fits used for the extraction of the yields of the signal and normalisation modes in data, are shown.

A striking feature of using resonant modes such as $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ and $\Xi_b^0 \to pK^-J/\psi(\to \mu^+\mu^-)$ is the additional constraint on the lepton kinematics coming from the fact two leptons are originating from the specific J/ψ resonance. This means that applying a J/ψ mass constraint can significantly improve the resolution on the invariant mass distribution. Given the small yield of the $\Xi_b^0 \to pKJ/\psi$ mode, as well as the fact that it is located on the upper tail of the $\Lambda_b^0 \to pKJ/\psi$ invariant mass peak, the J/ψ mass constraint becomes an essential tool in this analysis.

The fit model is built according to the knowledge on the background composition, acquired in the Section 3.3.

4.5.1 Signal and normalisation modes: invariant mass shapes

Simulation samples of the relevant decay modes are used to describe the shapes of the signal and normalisation modes. The model used to describe the invariant mass shape depends on the channel. For the $\Xi_b^0 \to pKJ/\psi$ mode, the Bifurcated Crystal Ball function is used. For the more abundant $\Lambda_b^0 \to pKJ/\psi$ mode, the Hypatia function [161] was used, providing a slightly better description of the tails, however behaving significantly slower in the fits. Improving the description of the upper mass tail of the $\Lambda_b^0 \to pKJ/\psi$ channel is important, as the $\Xi_b^0 \to pKJ/\psi$ signal is located on this tail. Fig. 4.11 presents the invariant mass shape used for the $\Xi_b^0 \to pKJ/\psi$ signal mode, and Fig. 4.12 presents the shape used for the normalisation $\Lambda_b^0 \to pKJ/\psi$ mode.

4.5.2 Background model

To describe the shapes of specific background modes, dedicated simulation samples are used, and their shape are fitted with the `RooKeysPdf` kernel estimation technique [146]. The procedure is similar to the one used in the lepton universality analysis. The background simulation samples used have the selection applied identical to the one used for the final fit to the data.

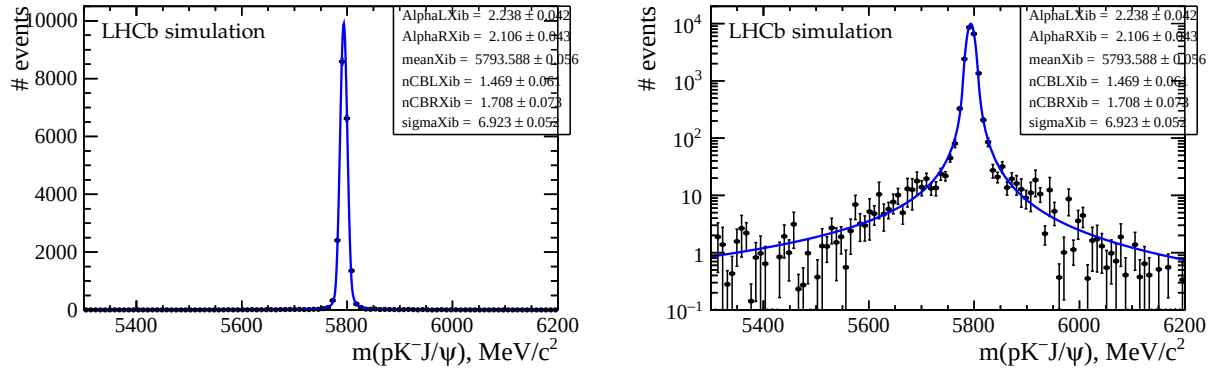


Figure 4.11: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the $\Xi_b^0 \to pKJ/\psi$ simulation, in linear and logarithmic scales.

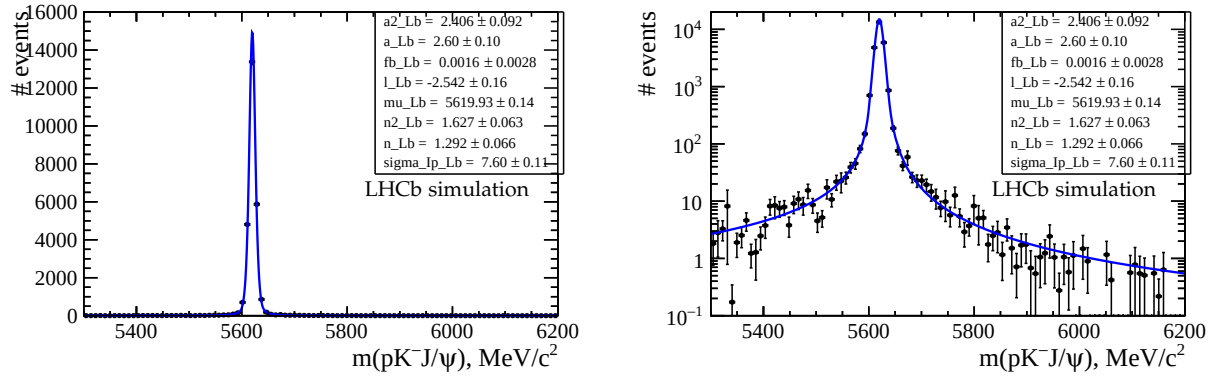


Figure 4.12: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the $\Lambda_b^0 \to pKJ/\psi$ simulation, in linear and logarithmic scales.

The selection applied in this analysis is supposed to suppress the most significant specific background contributions. The default fits include only the partially-reconstructed $\Lambda_b^0 \to pKJ/\psi\,\pi^0$ decay, with the π^0 not reconstructed, and the relevant invariant mass shapes are presented in Fig. 4.13. Other major specific background components (such as misidentification modes) are explicitly vetoed (see Sec. 4.2). It should be noted that the $\Xi_b^0 \to pKJ/\psi\,\pi^0$ partially-reconstructed background can also be present in the dataset. However, as the fraction of the partially reconstructed $\Lambda_b^0 \to pKJ/\psi\,\pi^0$ background is at a percent level with respect to the fully-reconstructed $\Lambda_b^0 \to pKJ/\psi$ (as it will be shown in the following section), the partially-reconstructed contribution for Ξ_b^0 is expected to be negligibly small. Therefore, it is not considered in the fits to data.

Finally, the combinatorial background is fitted with an exponential shape of a free slope.

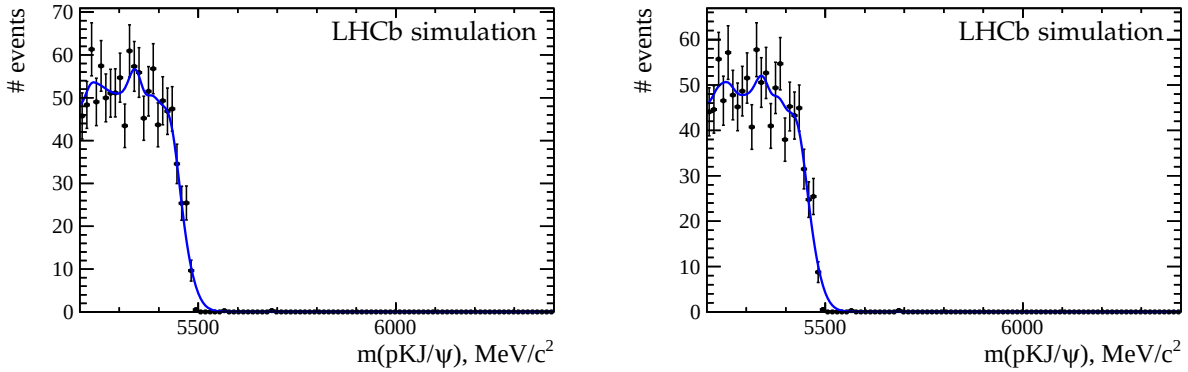


Figure 4.13: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the $\Lambda_b^0 \to pKJ/\psi\pi^0$ decay in the Run I (left) and Run II (right) simulation.

4.5.3 Fits to data

An unbinned maximum likelihood fit is used to extract the yields of the signal and the normalisation modes. The ratio of $\Xi_b^0 \to pKJ/\psi$ and $\Lambda_b^0 \to pKJ/\psi$ yields enters as a free parameter in the fit, allowing to extract it directly. At the same time, the yields of the $\Lambda_b^0 \to pKJ/\psi$ normalisation component, combinatorial and partially reconstructed backgrounds are also kept free. The mean and width of the $\Lambda_b^0 \to pKJ/\psi$ component are kept free, while those of the $\Xi_b^0 \to pKJ/\psi$ component are constrained: the mean is linked to the mean of the $\Lambda_b^0 \to pKJ/\psi$ via the mass difference taken from the PDG; the width is defined as the width of $\Lambda_b^0 \to pKJ/\psi$ times the factor taken from the simulation. Slope of the combinatorial background is kept free as well.

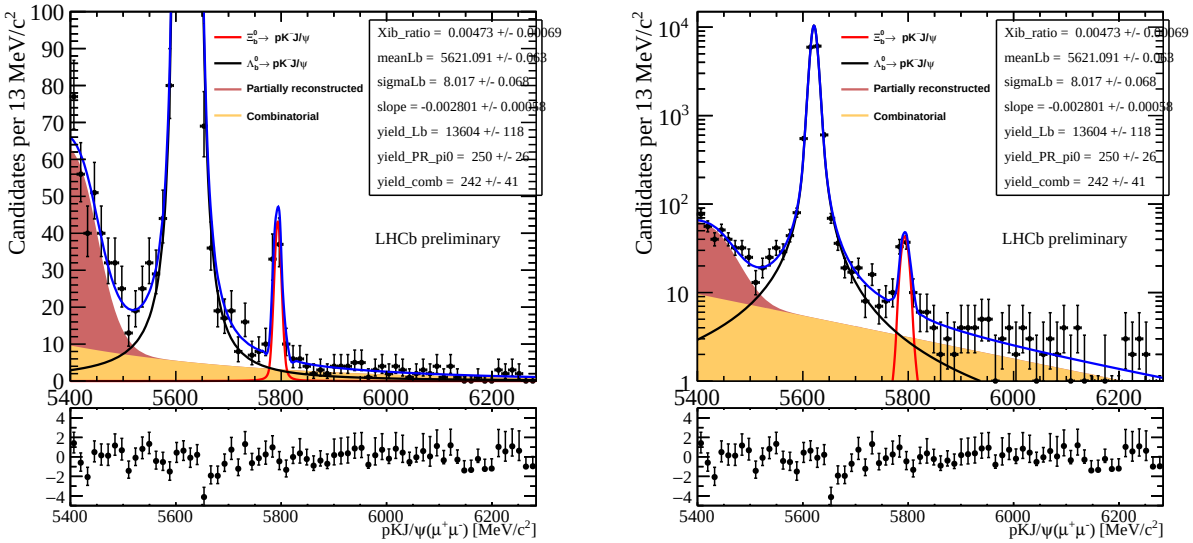


Figure 4.14: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the Run I data, in linear and logarithmic scales.

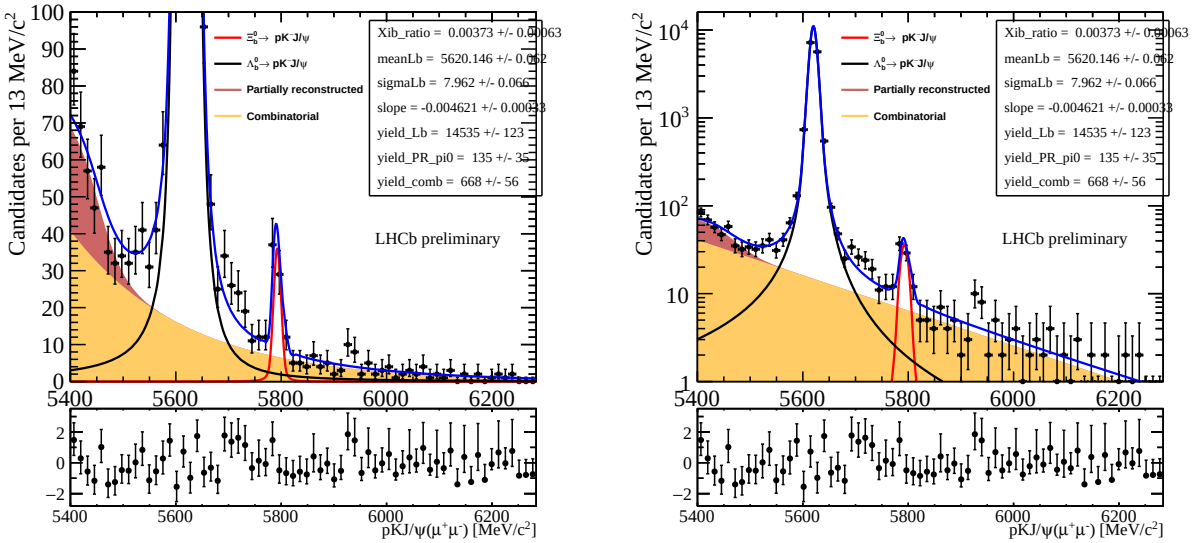


Figure 4.15: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the Run II data, in linear and logarithmic scales.

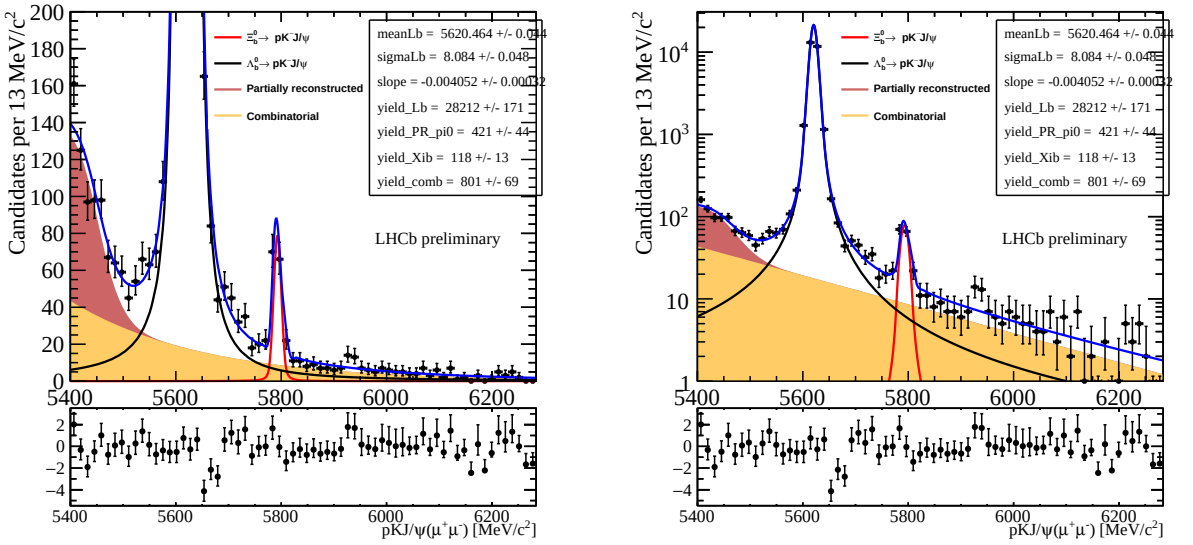


Figure 4.16: The fit to the $m(pKJ/\psi)$ invariant mass distribution in the combined Run I and Run II datasets, in linear and logarithmic scales.

Figures 4.14-4.15 present fits split by run period. Fig. 4.16 presents the total fit to the merged Run I and Run II datasets. The local significance of the $\Xi_b^0 \to pKJ/\psi$ peak, estimated by the Wilks' theorem from the latter fit, exceeds ten standard deviations.

4.5.4 Fit validation with pseudo-experiments

1000 pseudo-experiments were generated according to the fit model described above. A larger amount of (unbinned) pseudo-experiments is difficult to handle, as the Hypatia PDF is extremely slow in both generation and fit due to numerical computation of the integral. An example of such a toy fit is shown in Fig. 4.17. Pull distributions of the free fit parameters have been studied, and no significant bias was seen in any parameter, and the coverage of uncertainties is good as well.

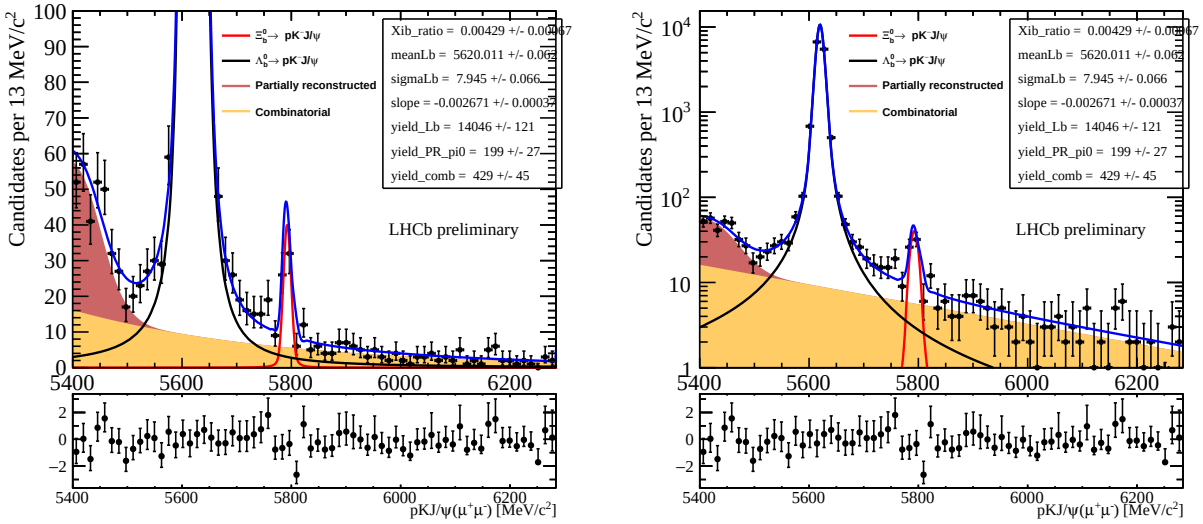


Figure 4.17: Example of a toy fit.

4.5.5 ${}_s\mathcal{P}lot$ of the $\Xi_b^0 \to pKJ/\psi$ signal

The fit presented in Fig. 4.16, is then used to perform an ${}_s\mathcal{P}lot$ in order to extract the background-subtracted $\Xi_b^0 \to pKJ/\psi$ signal from data. This can be used in order to study the Dalitz plane of the decay, as presented in Fig. 4.18. The variables studied here do not correlate with the pKJ/ψ invariant mass (especially when the J/ψ mass constraint applied), therefore, the ${}_s\mathcal{P}lot$ technique is valid. As it is expected from the analogous study in the $\Lambda_b^0 \to pKJ/\psi$ decay, the phase-space model does not describe properly the $m(pK)$ distribution. Instead, the prominent resonant structures are seen at low $m(pK)$ values, while the signal contribution above $2\,\mathrm{GeV}/c^2$ is very small (this is reassuring as it is unlikely that many signal events are located in the region above 2600 MeV/c^2). It should be noted that the relative proportion of various resonances in the $\Xi_b^0 \to pKJ/\psi$ decay is somewhat different from the one in the $\Lambda_b^0 \to pKJ/\psi$ decay, where the $\Lambda^*(1520)$ is clearly dominating the landscape. The structures seen here may be corresponding to the $\Lambda^*(1520)$, $\Lambda^*(1670)$ and other Λ^* resonances. An important difference with respect to the $\Lambda_b^0 \to pKJ/\psi$ decay is that the Σ^* resonances can contribute here as well. In the Λ_b^0 decays, the transition to the Σ^* states is isospin-suppressed. However, in the Ξ_b^0

decays both transitions to Λ^* and Σ^* are isospin-suppressed, making it possible to see both of them in equal proportions. This means, without an amplitude analysis it would be difficult to assign either $\Lambda^*(1670)$, $\Sigma^*(1660)$ or $\Sigma^*(1670)$ hypothesis to the structure around 1670 MeV/c^2. It should be noted that in addition to the above mentioned states reported in the PDG, the Belle experiment claims another state in the same mass range, seen in the Dalitz plane of the decay $\Lambda_c^+ \to pK^-\pi^+$ (see Refs. [162,163]). Larger statistics would be required to isolate and identify various states entering $m(pK)$ in the $\Xi_b^0 \to pKJ/\psi$ decay. No conclusion can be made with the current statistics from the distributions of $m(pJ/\psi)$ and $m(KJ/\psi)$.

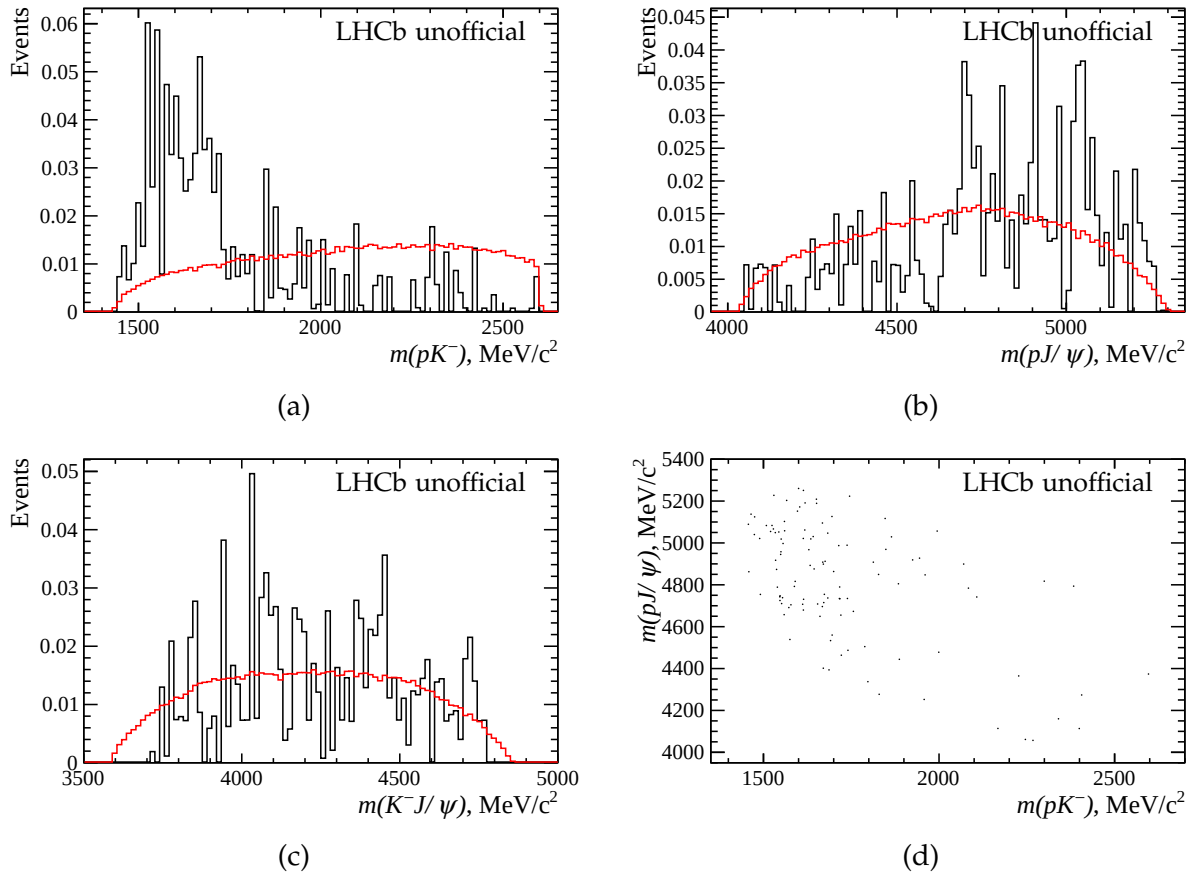


Figure 4.18: (a-c): Two-body invariant mass distributions in the $\Xi_b^0 \to pKJ/\psi$ decay, in the (red) simulation generated under the phase-space model, and (black) data, normalised to the unit area. (d): Dalitz plane of the decay in the data.

4.6 Combination of fit results and efficiencies

This section presents preliminary results, without an estimate of the systematic uncertainty, on the ratio of interest R defined in the Equation 4.1. Results are presented in Tab. 4.6 and

Fig. 4.19.

It can be seen that the Run II result is about 1.3 standard deviations smaller than the Run I result. This is somewhat counter-intuitive as the relative Ξ_b^0 production is expected to be about the same, or slightly larger at the Run II conditions (as it has been observed for the Ξ_b^- baryon [160]). However, a fluctuation at this level is not to be excluded. It should be noted that the Run I and Run II results cannot be directly merged, as the relative production of the Λ_b^0 and Ξ_b^0 can be different at 7-8 and 13 TeV, and has never been measured.

Table 4.6: Comparison of the R values in Run I and Run II datasets.

Run I	Run II
$(3.77 \pm 0.55) \times 10^{-3}$	$(3.05 \pm 0.51) \times 10^{-3}$

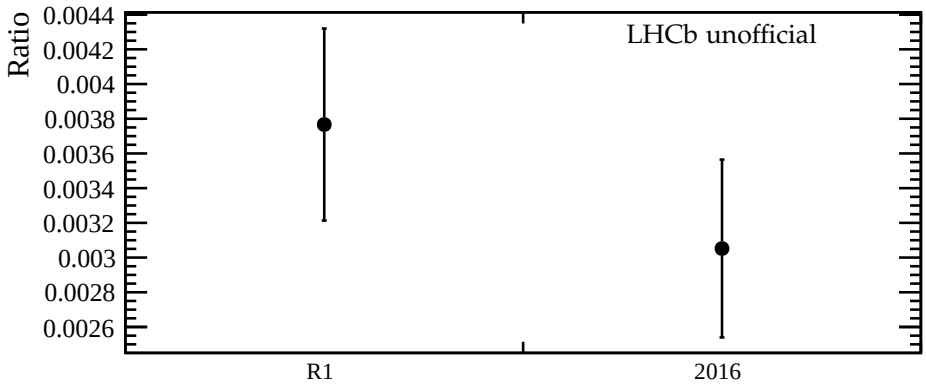


Figure 4.19: Comparison of the R values in Run I and Run II datasets.

4.6.1 Cross-checks

To understand the origin of the Run I – Run II difference, the results are split by magnet polarity in Fig. 4.20. In addition, the ratio of Ξ_b^0 and Λ_b^0 signal yields is shown in Fig. 4.21. It can be clearly seen that the MagUp results are in a good agreement between Run I and Run II, while the MagDown results are somewhat higher, especially in Run I. Given that this difference is entirely driven by the fitted value of signal yields, it is likely to be a statistical fluctuation.

4.7 Implications and future prospects

This work presents the first observation of the decay $\Xi_b^0 \to pKJ/\psi$, which is the first decay of Ξ_b^0 with a J/ψ in the final state. A more detailed study of this decay mode can be

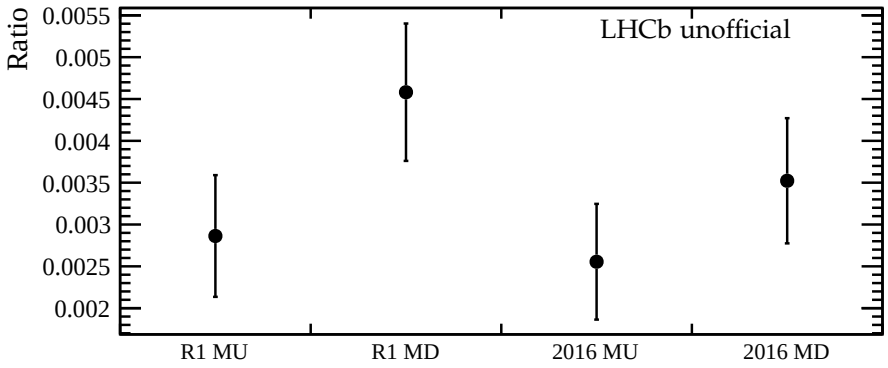


Figure 4.20: Comparison of the R values in Run I and Run II datasets, per magnet polarity.

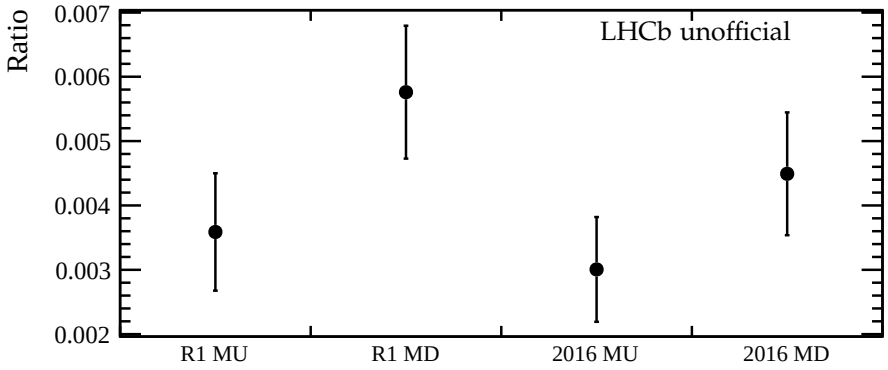


Figure 4.21: Comparison of the ratio of Ξ_b^0 and Λ_b^0 yields in Run I and Run II datasets, per magnet polarity.

performed with a larger dataset. In particular, one can double the amount of data by adding the 2015, 2017 and 2018 datasets which were not considered in the current work due to the lack of simulation samples. This would allow to perform a direct measurement of the mass difference of Λ_b^0 and Ξ_b^0 baryons, with the precision competitive to the value obtained through the $\Lambda_b^0, \Xi_b^0 \to D^0 pK$ channels, or to the indirect value provided by the PDG and obtained through the channels $\Xi_b^0 \to \Xi_c^+ \pi^-$ and $\Lambda_b^0 \to \Lambda_c^+ \pi^-$. The pKJ/ψ final state offers a smaller phase-space than other channels, resulting in a smaller systematic uncertainty related to the momentum scale uncertainty. With even larger dataset collected by the LHCb experiment in the future, the study of resonant structures, and the search for exotic contributions to the $\Xi_b^0 \to pKJ/\psi$ decay channel will become possible.

Conclusions

Rare $b \to s\ell^+\ell^-$ transitions provide an excellent laboratory for precision tests of the Standard Model. In the SM, they occur only at the loop level, which makes them sensitive to potential New Physics effects. Hints of deviations from the theoretical predictions have been reported recently in a number of observables which describe the $b \to s\ell^+\ell^-$ transitions, such as differential branching fractions, angular parameters, or lepton universality tests. The latter measurements test whether the processes involving different kinds of leptons in the final state – electrons or muons – have the same rate, as predicted by the SM. Lepton universality has been tested to a sub-percent precision in such tree-level decay modes as $Z \to \ell^+\ell^-$ or $J/\psi \to \ell^+\ell^-$. However, precise tests have not been performed until recently in suppressed $b \to s\ell^+\ell^-$ transitions. Two results have been reported by the LHCb Collaboration recently, testing lepton universality in decays $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$. These ratios, called R_K and $R_{K^{*0}}$, were found to be within 2.5 standard deviations from unity [1,2]. To achieve a definite conclusion on whether lepton universality holds in the $b \to s\ell^+\ell^-$ decays, two approaches can be taken: analysing a larger amount of data, and exploring additional decay modes.

This thesis presents the first test of lepton universality in the decays of b-baryons: to date, only their meson counterparts have been studied. The decay modes $\Lambda_b^0 \to pK\ell^+\ell^-$ have been chosen for this test, as being the most accessible experimentally. The mode $\Lambda_b^0 \to pK\mu^+\mu^-$ has been observed for the first time by the LHCb experiment [3], but its branching fraction has not been measured. The decay mode $\Lambda_b^0 \to pKe^+e^-$ has not been observed before. Therefore, prior to performing the lepton universality test, it is crucial to perform the first observation of the dielectron decay mode.

This thesis presents in detail all the milestones of the analysis. The LHCb dataset, collected in 2011 (amounting to 1 fb^{-1}), 2012 (2 fb^{-1}) and 2016 (1.7 fb^{-1}), at 7, 8 and 13 TeV centre-of-mass energy, respectively, is used. The analysis is performed in the dilepton invariant mass squared range $0.1 < q^2 < 6$ GeV^2/c^4, and dihadron invariant mass $m(pK) < 2600$ MeV/c^2. A dedicated set of selection requirements is developed in order to obtain clean signals of rare decays, and suppress all the dominant background modes. In particular, this includes particle identification requirements and mass vetoes dedicated

to suppress misidentified contributions. Dedicated multivariate classifiers are developed to suppress the background consisting of random track combinations. The efficiency of these selection requirements is estimated relying on dedicated simulation samples, which are calibrated using control datasets. Systematic uncertainties are estimated to account for imperfections of simulation, lack of knowledge on background composition, and other constraints which might impact the estimate of the signal yield.

The main difficulty of the analysis is due to the specificities of the electron reconstruction and the electron trigger at LHCb. Being light particles, electrons emit bremsstrahlung photons when interacting with the material of the detector. These photons are then recovered by dedicated algorithms, however, this recovery is not perfect, which leads to the degradation in invariant mass resolution. In addition, triggering on electrons is not as efficient at hardware stage, compared to triggering on muons, due to high occupancy of the electromagnetic calorimeter. This motivates the usage of two exclusive trigger categories for the electron dataset in the current analysis: events triggered independently of the signal candidate, and events triggered on the electrons from the signal candidate (but not entering the first category).

The calibration of simulation is verified on the control mode $\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$, profiting from the well-tested lepton universality in J/ψ decays to two leptons. To do that, the ratio $r_{J/\psi}^{-1}$ is explored. It is defined as the ratio of branching fractions of $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ and $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ decays, is expected therefore to be equal to unity. This test is performed blindly: the actual value of $r_{J/\psi}^{-1}$ was not looked at, until the entire machinery was prepared, and all cross-checks were performed. In particular, the absence of $r_{J/\psi}^{-1}$ dependence on any kinematic or geometrical variables is checked. Finally, the $r_{J/\psi}^{-1}$ value is unblinded and is found to be in a good agreement with unity:

$$r_{J/\psi}^{-1} = 0.962 \pm 0.048.$$

Another milestone achieved in this work is the first observation of the $\Lambda_b^0 \to pKe^+e^-$ decay mode. All the selection requirements, needed for this observation, have been prepared in a blind manner, without looking at the signal region. The invariant mass fit was prepared, accounting for the most important backgrounds, and the fit framework was first tested with pseudo-experiments. Numerous systematic studies have been performed, in order to quantify potential biases on the signal yield, coming from the limited knowledge of background composition. For example, the largest systematic effect comes from the contribution of partially reconstructed decay $\Lambda_b^0 \to pKe^+e^-\pi^0$, which has an unknown decay model. Upon unblinding, the decay mode $\Lambda_b^0 \to pKe^+e^-$ has been observed with the yield of $122 \pm 17(\text{stat}) \pm 6(\text{syst})$ candidates. The observed signal has a significance above seven standard deviations, accounting for statistical and systematic effects.

The observed signal yield in the electron mode is large enough to perform the test of lepton universality R_{pK}^{-1}. In order to cancel some of systematic uncertainties related to specific final state, the R_{pK}^{-1} is measured as a double ratio:

$$R_{pK}^{-1} = \frac{\mathcal{B}(\Lambda_b^0 \to pKe^+e^-)}{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)} \times \frac{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))}.$$

The measurement is performed by means of a simultaneous fit to $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to pKe^+e^-$ datasets in all trigger categories. Yields of $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ modes, as well as ratios of efficiencies, enter in the fit as external parameters. This allows to get R_{pK}^{-1} directly as a free fit parameter. Another important free parameter is the ratio $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))$, which comes as a standalone result of this work. The fit stability has been tested with pseudo-experiments. Systematic uncertainties are split in the two categories: those which are uncorrelated between the datasets, and those which are fully correlated. The first set of uncertainties enters as Gaussian constraints on the ratios of efficiencies. The second set of uncertainties is used to smear the final log-likelihood. All the datasets entering the fit are unblinded, however, the $\Lambda_b^0 \to pKe^+e^-$ efficiencies were still blind at the moment of writing this document. This was done in the way that the central value of R_{pK}^{-1} is scaled to be equal to one.

As a result, the two free fit parameters have the following values:

$$R_{pK}^{-1}(\text{blind}) = 1.00^{+0.15}_{-0.14}(\text{stat}) \pm 0.06(\text{syst});$$

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} = (8.45^{+0.45}_{-0.43}(\text{stat}) \pm 0.26(\text{syst})) \times 10^{-4}.$$

Both measurements are performed in range $0.1 < q^2 < 6$ GeV$^2/c^4, m(pK) < 2600$ MeV$/c^2$ only.

As it can be seen from these results, the lepton universality test R_{pK}^{-1} can be performed with a precision of about 16% (accounting for both statistical and systematic uncertainties). The first measurement of $\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))}$ has a precision of approximately 6%.

At the moment of presenting these results, the analysis was in the review by the LHCb Collaboration, and final unblinding was planned to happen on a short timescale. This result will provide a first test of lepton universality in the baryon sector, and will serve as an additional null test of the SM. This measurement can motivate the further theoretical and experimental interest in rare baryon decays.

The final chapter of this thesis is dedicated to a separate analysis: the first observation of the decay mode $\Xi_b^0 \to pKJ/\psi$. This decay mode has the same final state as the $\Lambda_b^0 \to pKJ/\psi$, where pentaquark candidates were discovered by the LHCb Colaboration [4,5]. However, the yield of this channel is suppressed by two factors: Cabibbo suppression, and smaller relative production rate of the Ξ_b^0 baryon compared to the Λ_b^0 baryon. Therefore, a dedicated selection had to be developed in order to achieve the best sensitivity to this suppressed signal. As a result, the $\Xi_b^0 \to pKJ/\psi$ decay mode is observed with 118 ± 13 signal candidates, with a statistical significance exceeding ten standard deviations. The branching fraction of this decay mode is measured relatively to the normalisation $\Lambda_b^0 \to pKJ/\psi$ mode. It is expressed through the ratio of hadronisation fractions to Λ_b^0

and Ξ_b^0 baryons, which is not measured to date, and can a priori be different at different energies. The resulting values are:

$$\frac{\mathcal{B}(\Xi_b^0 \to pK^- J/\psi(\to \mu^+\mu^-))}{\mathcal{B}(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))} \times \frac{f_{\Xi_b^0}}{f_{\Lambda_b^0}} = \begin{cases} (3.77 \pm 0.55) \times 10^{-3}, \text{ at 7-8 TeV;} \\ (3.05 \pm 0.51) \times 10^{-3}, \text{ at 13 TeV.} \end{cases}$$

Appendices

A List of variables used throughout the thesis

A decay of the type $A \to b, c...$ is considered, where $b, c, ...$ are tracks, and A is a short-lived particle candidate formed by the combination of these tracks. The proton-proton collision point, where the candidate A was created, is referred to as the *primary vertex* (PV). The decay vertex of A, defined by the crossing point of the tracks it decays into, is referred to as *secondary vertex* or decay vertex (DV).

p_T	Transverse (with respect to the beam axis) momentum;
$x_{ECAL}(b), y_{ECAL}(b)$	x and y coordinates of the track b when it reaches the ECAL;
$region_{ECAL}(b)$	ECAL region: outer=0, middle=1, inner=2; out of acceptance < 0;
$\chi^2_{DV \leftrightarrow PV}(A)$	χ^2-separation from the related PV, *i.e.* a difference in χ^2 obtained from two vertex fits: 1) assuming that all tracks come from the same vertex, 2) assuming two separate vertices (the primary and decay vertex);
`DIRA(A)`	Angle between the vector of momentum of a candidate A and the vector that links the PV to the DV of A;
$\chi^2_{IP}(A)$	The difference in χ^2 of a given PV when reconstructed with and without the candidate A;
$\chi^2_{FD}(A)$	Flight distance of the considered candidate A divided by its uncertainty;
$\chi^2_{vtx}(A)$	χ^2 of the fit of the decay vertex of candidate A;
$\chi^2_{vtx}(A)$/ndf	χ^2 per degree of freedom of the fit of the decay vertex of A;
$PID_\beta(b)$	Log-likelihood difference between the hypotheses that the track b is a particle of type β ($\beta = p, K, \mu, e$), and that b is a pion, as estimated using inputs from the RICH detectors and calorimeters;
ProbNN$\beta(b)$	Probability of the track b being a particle of type β ($\beta = p, K, \pi, \mu, e, d$), as estimated from information coming from all sub-detectors combined using a neural network;
`isMuon(b)`	Boolean variable that indicates whether b is compatible with being a muon, based on information from the muon system and calorimeters;

GhostProb(b)		Probability of b being a fake track coming from a random combination of hits in the tracking system ("ghost");
$\chi^2_{track}(b)$		χ^2 of the b track fit;
$\chi^2_{DTF}(A)$		χ^2 of the `DecayTreeFitter` fit to the A decay chain;
`nSPDHits`		Number of hits recorded in the scintillating pad detector;
`nTracks`		Number of (good) tracks recorded in the event;
`nVeloTracks`		Number of VELO tracks in the event;
`nPVs`		Number of primary vertices in the event;
`InAccMuon`(b)	or	Requirement of the track b being in acceptance of the muon system;
`hasMuon`(b)		
`hasCalo`(b),		Requirement of the track b having the information from the calorimeter
`hasRich`(b)		or RICH detectors;
$m(bc)_{b\leftarrow z}$		Two-body invariant mass of the (bc) system, where particle b has a mass hypothesis of the particle z;
$m(bc)_{b\leftrightarrow c}$		Two-body invariant mass of the (bc) system, where particles b and c have mass hypotheses swapped;
$\theta(b,c)$		Angle between the tracks b and c.

B Techniques to visualise misidentification contributions

There are several tests which can be performed in order to determine the presence of backgrounds due to the misidentification of a hadrons in $\Lambda_b^0 \to pK\ell^+\ell^-$ and $\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$ decays.

B.1 Two-body decays

One can start from two-body decays of a kind $A \to B + C$. In the scope of this analysis, this can be $\Lambda^* \to pK$ decay which has to be separated from various misidentification modes such as $\phi(1020) \to KK$. However, for completeness, let's consider the most general example of a two-body decay of a heavy particle with mass M, to two lighter particles of masses m_1 and m_2. Assume the parent particle was moving with a velocity β. The angle between the centre-of-mass momentum of a daughter particle, and the direction of flight of the parent particle, is θ. In a two-body decay, the energy-momentum conservation defines the center-of-mass momenta of daughter particles:

$$M = \sqrt{p_{cm}^2 + m_1^2} + \sqrt{p_{cm}^2 + m_2^2} \tag{2}$$

while center-of-mass energy for each daughter particle is $E_{cm,i}^2 = m_i^2 + p_{cm}^2$ $(i = 1, 2)$. The lab-momenta of the daughter particles can be expressed using Lorentz transformations. It is interesting to consider separately the longitudinal and transverse momenta of daughter particles. In the center-of-mass frame, they are defined as

$$p_L^{cm} = \pm p_{cm} \times \cos\theta \tag{3}$$

$$p_T^{cm} = p_{cm} \times \sin\theta \tag{4}$$

They transform the following way:

$$p_L = \gamma p_L^{cm} + \gamma\beta E^{cm} \tag{5}$$

$$p_T = p_T^{cm} \tag{6}$$

From Eq. 2, one can express the daughter momentum in the center-of-mass frame:

$$p_{cm} = \frac{1}{4M^2}(M^4 + m_1^4 + m_2^4 - 2m_1^2M^2 - 2m_2^2M^2 - 2m_1^2m_2^2) \tag{7}$$

Finally, this allows to express the longitudinal momenta:

$$p_L^i = \gamma p_{cm}\cos\theta + \gamma\beta E_i^{cm}, \tag{8}$$

where E_i^{cm} $(i = 1, 2)$ are the center-of-mass energies of daughter particles:

$$E_1^{cm} = \frac{1}{2M}(M^2 + m_1^2 - m_2^2) \tag{9}$$

$$E_2^{cm} = \frac{1}{2M}(M^2 - m_1^2 + m_2^2) \tag{10}$$

The asymmetry of longitudinal momenta of the daughter particles

$$\alpha = \frac{p_L^1 - p_L^2}{p_L^1 + p_L^2} = \frac{2p_{cm}}{\beta M}\cos\theta + \frac{m_1^2 - m_2^2}{M^2} \tag{11}$$

Denoting $\alpha_0 = \frac{m_1^2 - m_2^2}{M^2}$, and $r_\alpha = \frac{2p_{cm}}{M}$, and assuming the relativistic limit $\beta \to 1$, we get

$$\frac{\alpha - \alpha_0}{r_\alpha} = \cos\theta \tag{12}$$

Using the relation $p_T = p_{cm}\sin\theta$, we get

$$\frac{(\alpha - \alpha_0)^2}{r_\alpha^2} + \frac{p_T^2}{p_{cm}^2} = 1 \tag{13}$$

which is an equation of ellipse in the α, p_T plane.

As a conclusion, in a plane of daughter momentum asymmetry and the transverse momentum, a two-body decay defines an ellipse with a center $(\alpha_0, 0) = \left(\frac{m_1^2 - m_2^2}{M^2}, 0\right)$ and radii $(r_\alpha, r_{p_T}) = \left(\frac{2p_{cm}}{M}, p_{cm}\right)$. This ellipse is centered at $(0,0)$ if two daughter particles have identical mass, otherwise the shift from zero in the α plane is proportional to the squared mass asymmetry. The size of the ellipse is larger when there is a larger phase-space available for the decay, resulting in larger center-of-mass momentum.

This useful relation can be used to separate different two-body decays using the so-called Armenteros-Podolanski plot which is a plot of α, p_T for a given two-body decay [148]. It is useful to visualise different two-body decay modes present in the dataset. Fig. 22 features the transverse momentum of the proton with respect to the Λ^*, as a function of the asymmetry of longitudinal (with respect to Λ^*) momenta of proton and kaon. This plot is prepared using the 2012 $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ dataset, with loose preselection requirements applied. In the plot, several arc-like structures can be seen. The most prominent one, symmetric around zero in x-axis, is due to the narrow $\phi(1020)$ resonance decaying to two kaons. other symmetric structures around zero are due to higher meson states decaying to K^+K^-, notably $f_2'(1525)$. The broad arc displaced to the left originates from $K^*(892)$ resonance decaying to a pion and a kaon (so having negative mass asymmetry between decay products). Finally those arcs which are displaced to the right correspond to various Λ^* resonances decaying to a proton and a kaon, and having positive mass asymmetry between decay products. The only prominent of them corresponding to the narrowest $\Lambda^*(1520)$ resonance, while more broad resonances are not clearly visible in this plot.

B.2 Three-body decays

A somewhat different approach can be used for multi-body decays. Let's consider the invariant mass of the particle computed via its n-body decay to particles $i = 1...n$ which

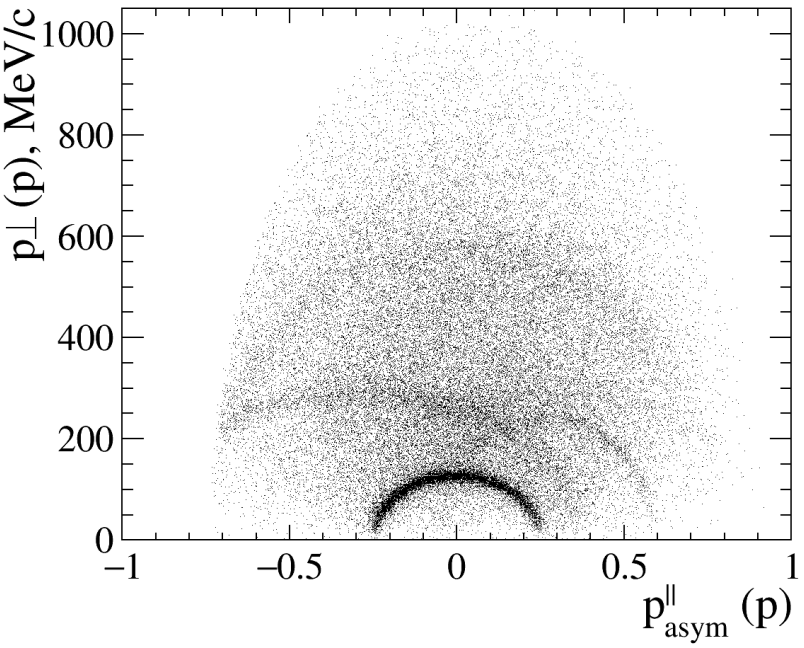


Figure 22: Armenteros-Podolanski plot for the two-body $\Lambda^* \to pK$ decay.

have masses m_i:

$$M_m^2 = \left(\sum_{i=1\ldots n} \sqrt{p_i^2 + m_i^2} \right)^2 - p_i p^i \tag{14}$$

Now, let's also consider the case when the mass hypotheses of daughter particles are not identified properly, and are μ_i:

$$M_\mu^2 = \left(\sum_{i=1\ldots n} \sqrt{p_i^2 + \mu_i^2} \right)^2 - p_i p^i \tag{15}$$

Taking the difference $M_\mu^2 - M_m^2$ and, assuming $\frac{m_i}{p_i} < 1$ and $\frac{\mu_i}{p_i} < 1$, one can expand for these parameters:

$$
\begin{aligned}
M_m^2 - M_\mu^2 &= \left(\sum_{i=1\ldots n}\sqrt{p_i^2+m_i^2}\right)^2 - \left(\sum_{i=1\ldots n}\sqrt{p_i^2+\mu_i^2}\right)^2 \\
&= \sum_{i=1\ldots n}(m_i^2-\mu_i^2) + 2\sum_{i=1\ldots n}\sum_{j=1\ldots n, j>i}\left(\sqrt{(p_i^2+m_i^2)(p_j^2+m_j^2)} - \sqrt{(p_i^2+\mu_i^2)(p_j^2+\mu_j^2)}\right) \\
&\approx \sum_{i=1\ldots n}(m_i^2-\mu_i^2) + \sum_{i=1\ldots n}\sum_{j=1\ldots n, j>i} p_i p_j \left(\left(\frac{m_i}{p_i}\right)^2\right) + \left(\frac{m_j}{p_j}\right)^2 - \left(\frac{\mu_i}{p_i}\right)^2 - \left(\frac{\mu_j}{p_j}\right)^2 \\
&= \sum_{i=1\ldots n}(m_i^2-\mu_i^2) + \sum_{i=1\ldots n}\sum_{j=1\ldots n, j>i}\frac{p_j}{p_i}\left(m_j^2\frac{p_i^2}{p_j^2} + m_i^2 - \mu_j^2\frac{p_i^2}{p_j^2} - \mu_i^2\right) \\
&\approx \sum_{i=1\ldots n}(m_i^2-\mu_i^2) + \sum_{i=1\ldots n}\sum_{j=1\ldots n, j>i}\frac{p_j}{p_i}(m_i^2-\mu_i^2) \qquad (16)
\end{aligned}
$$

Now, let's consider the case of only one misidentified track. In that case, only one mass difference is non-zero, $m_1^2 - \mu_1^2$. Consequently,

$$
M_m^2 - M_\mu^2 = (m_1^2-\mu_1^2)\left(1 + \sum_{j=2\ldots n}\frac{p_j}{p_1}\right) \qquad (17)
$$

Building an analogy to the two-body case, one can re-express this result in terms of the parameter which can be considered as the momentum asymmetry, defined as

$$
\beta = \frac{\left(\sum_{j=2\ldots n} p_j - p_1\right)}{\sum_{j=1\ldots n} p_j}. \qquad (18)
$$

The expression 17 then becomes

$$
M_m^2 - M_\mu^2 = (m_1^2-\mu_1^2)\frac{2}{1-\beta}. \qquad (19)
$$

For a three-body decay such as $\Lambda_b^0 \to pKJ/\psi$, with one misidentified particle (assume it's a proton), one has from Eq. 17:

$$
M_{\Lambda_b^0}^2 - M_\mu^2 = (m_p^2-\mu_p^2)\left(1 + \frac{p_K}{p_p} + \frac{p_{J/\psi}}{p_p}\right) = (m_p^2-\mu_p^2)\frac{2}{1-\beta}, \qquad (20)
$$

where the momentum asymmetry is

$$
\beta = \frac{p_K + p_{J/\psi} - p_p}{p_K + p_{J/\psi} + p_p}. \qquad (21)
$$

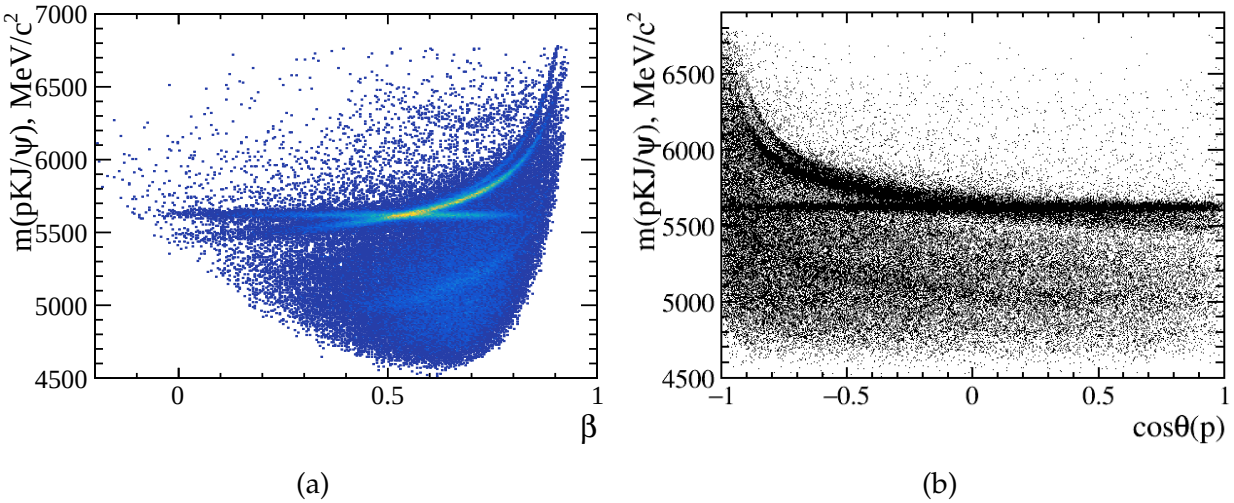


Figure 23: Two tests used to study the presence of misidentification backgrounds in the three-body decay $\Lambda_b^0 \to pKJ/\psi$. Prominent horizontal bands represent the signal, and arc-like structures correspond to misidentification backgrounds.

This shows that the invariant mass under another mass hypothesis, M_μ, is completely determined by the mass under proper hypothesis M_μ, the mass of the misidentified particle μ, and the momentum asymmetry β. This allows to visualise contributing misidentification backgrounds as curves in $(M_{pKJ/\psi}, \beta)$ space. Such a plot for our dataset is presented in Fig. 23(a). The flat horizontal distribution corresponds to the signal $\Lambda_b^0 \to pKJ/\psi$, while two prominent overlapping curves crossing the signal band correspond to two misidentifications (kaon or pion as a proton). Positions of these curves can be predicted mathematically from the Eq. 19. The arc at the very upper sideband corresponds to what is referred to "over-reconstructed" background described in Sec. 3.3.1.e, and the arc in the lower sideband is due to the misidentified partially reconstructed background.

A similar exercise can be performed, replacing the momentum asymmetry by the cosine of the angle between the proton momentum in the center-of-mass frame, and mother momentum in the lab frame. Such a plot for our dataset is shown in Fig. 23(b). A horizontal distribution corresponds to the signal, and two prominent arc-like structures crossing it are due to the misidentification backgrounds. The faint arc in the lower sideband is due to the misidentified partially reconstructed background.

Both plots show that the abundance of the misidentification backgrounds is of the same order, or larger, than the signal yield. This means, a careful selection is needed in order to suppress these contributions.

C Comparisons of data and corrected simulation

C.1 2012 $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$, L0M category

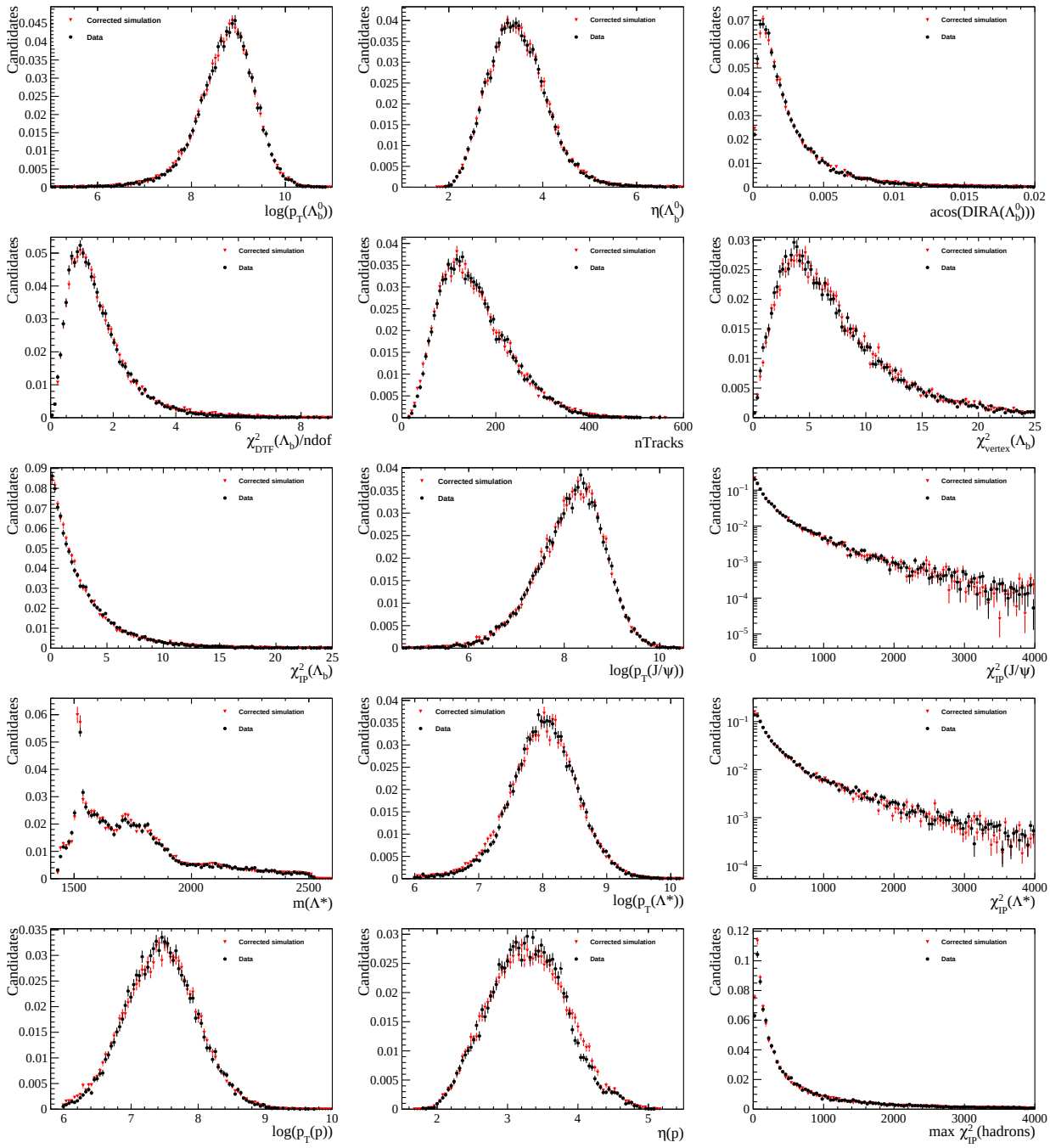


Figure 24: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode, with nominal corrections applied, in the L0M category for the 2012 conditions (part I).

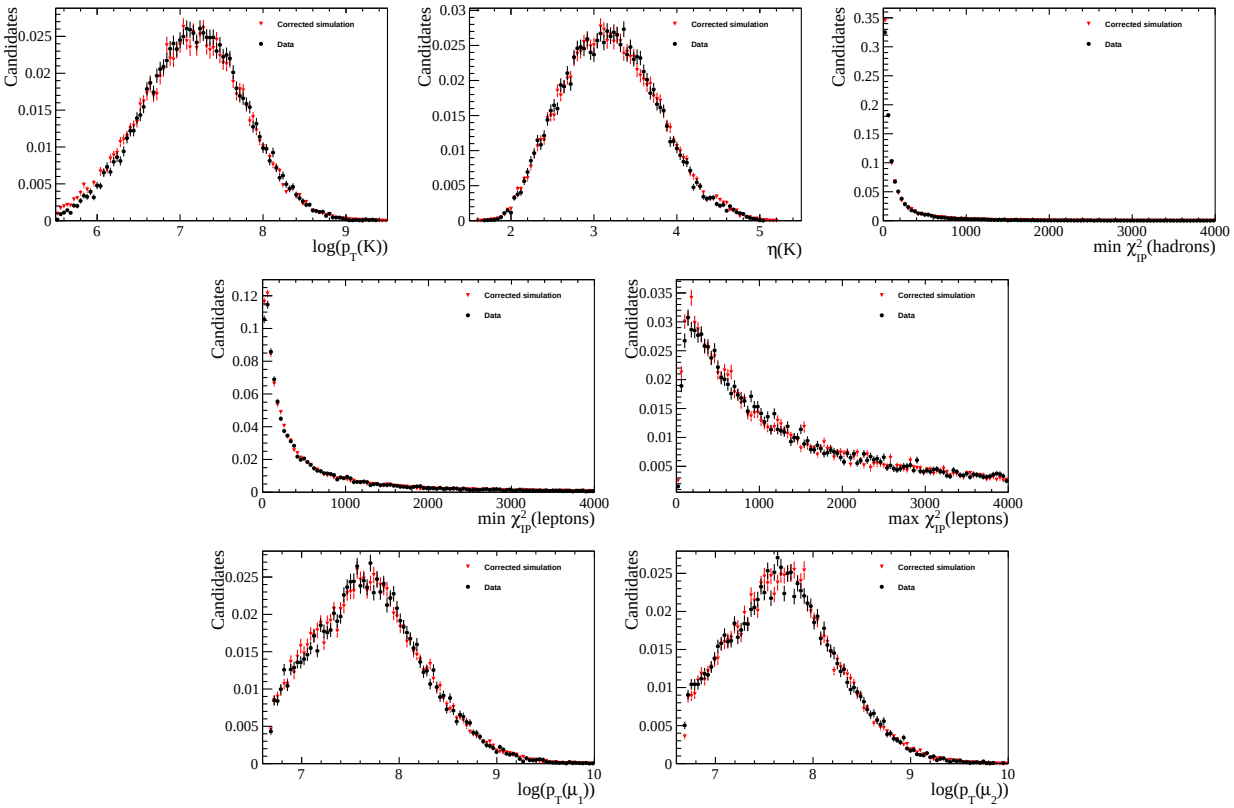


Figure 25: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode, with nominal corrections applied, in the L0M category for the 2012 conditions (part II).

C.2 2012 $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, L0E category

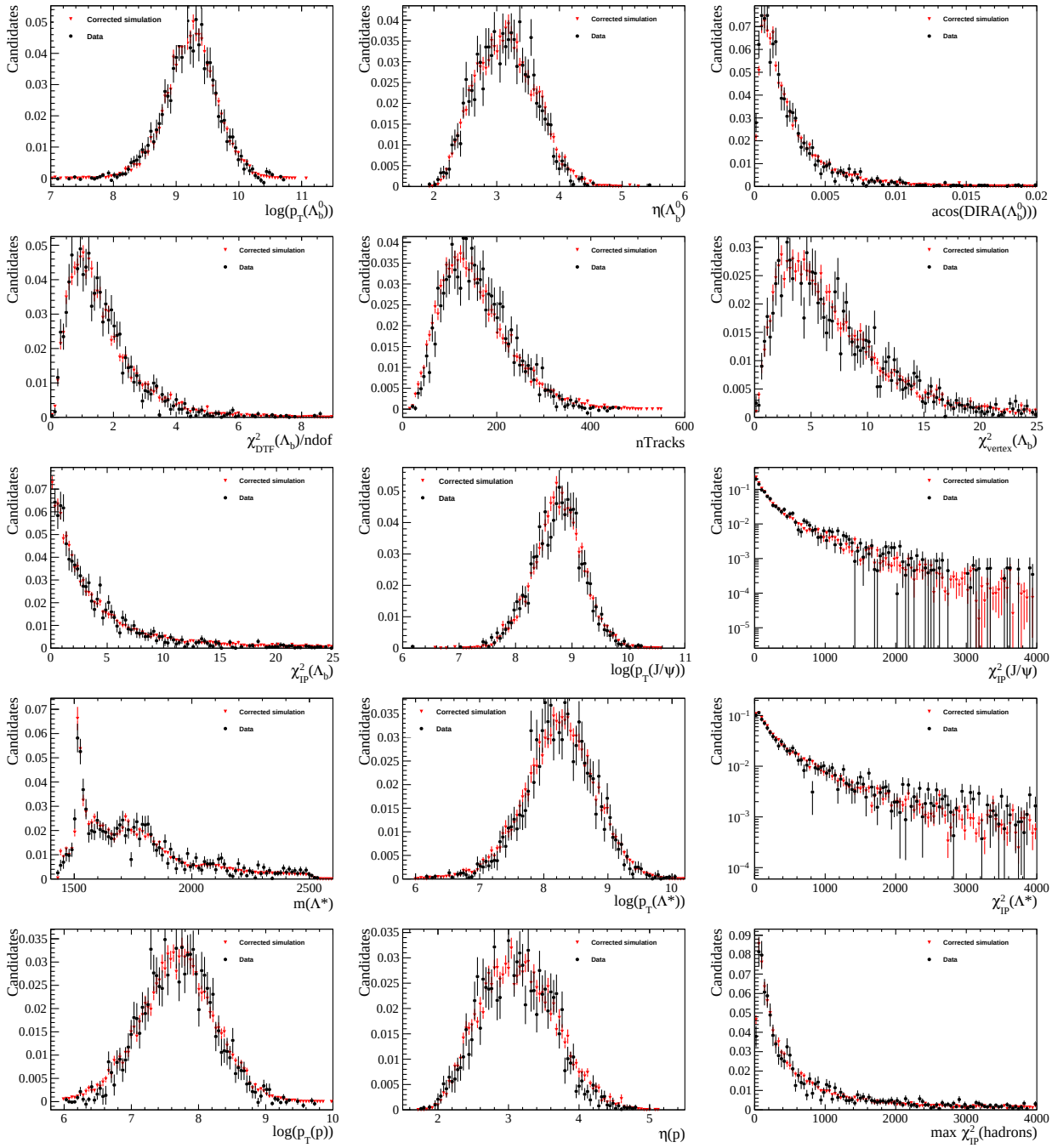


Figure 26: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode, with nominal corrections applied, in the L0E category for the 2012 conditions (part I).

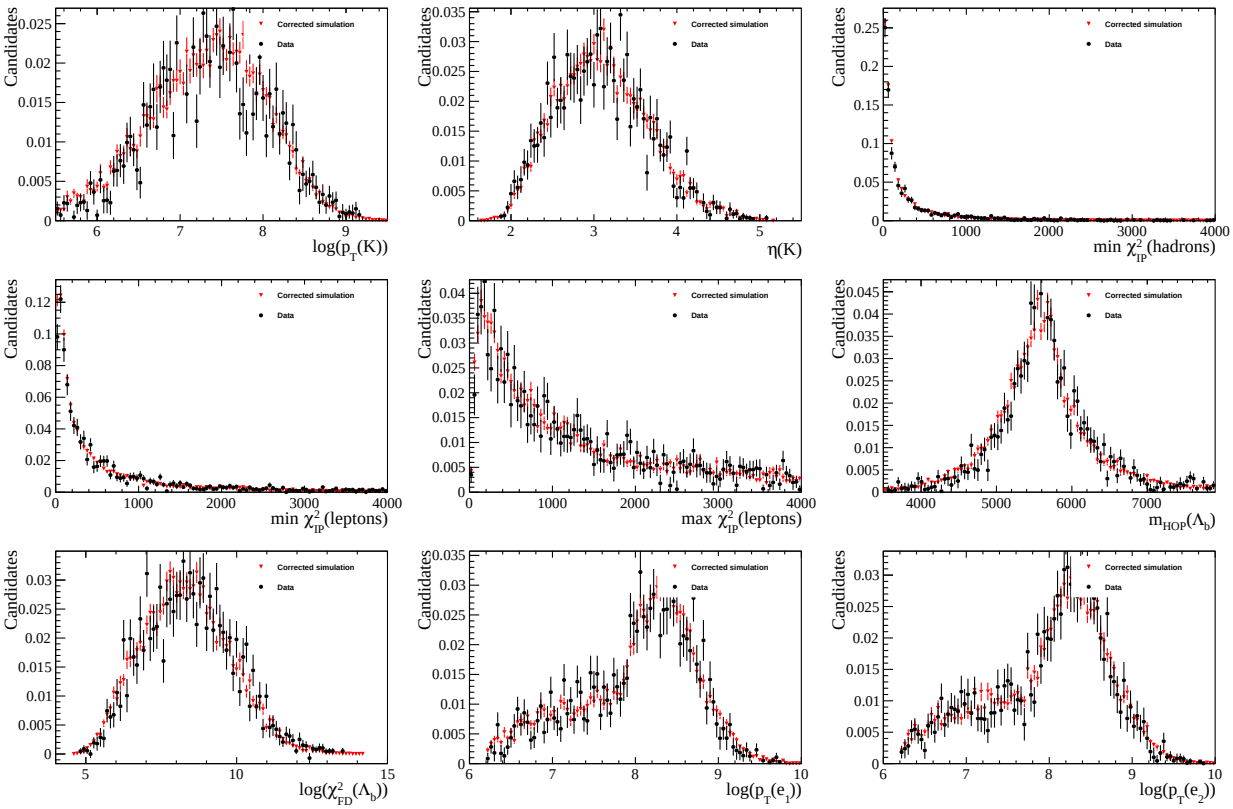


Figure 27: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, with nominal corrections applied, in the L0E category for the 2012 conditions (part II).

C.3 2012 $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, L0I category

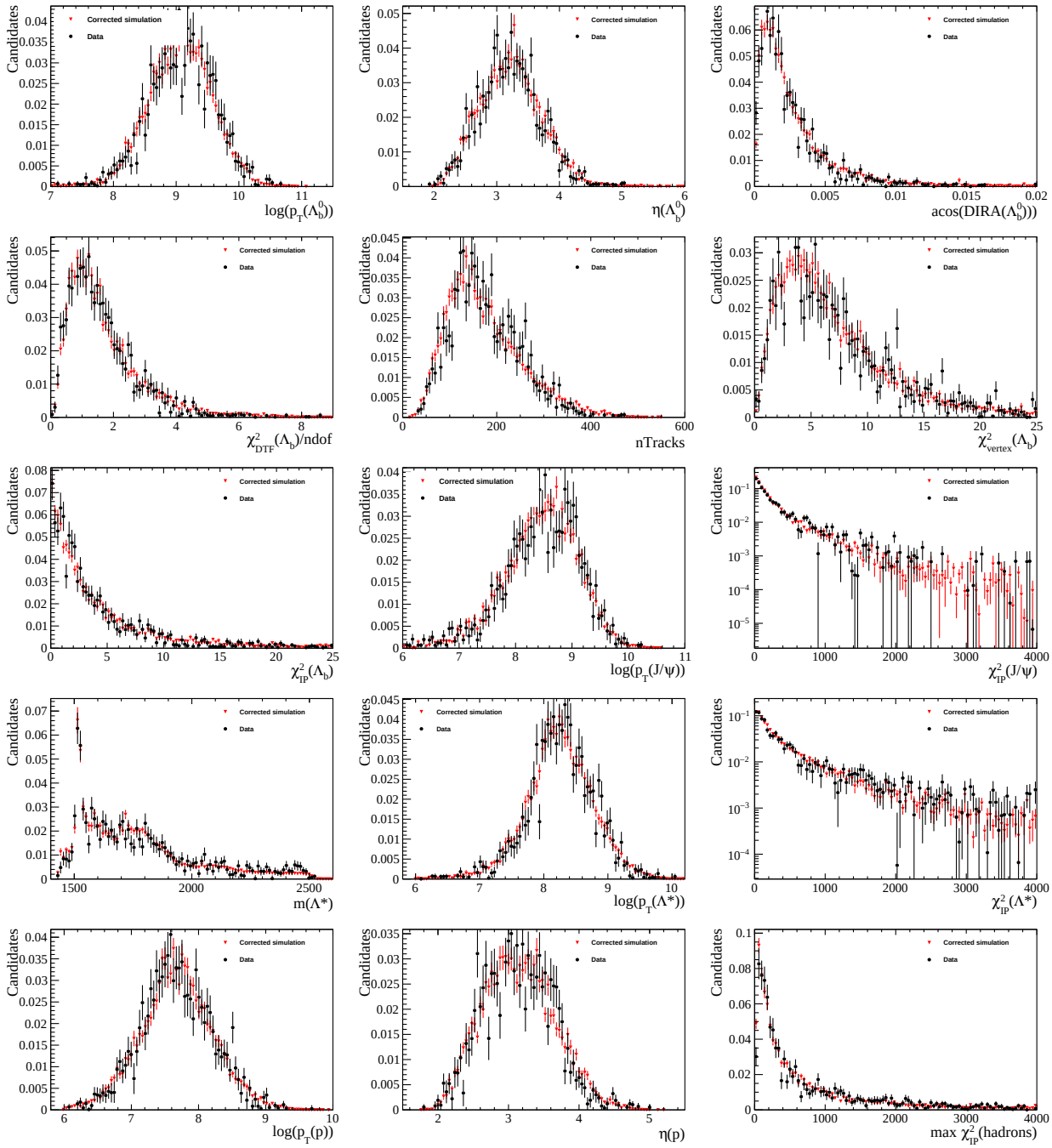


Figure 28: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode, with nominal corrections applied, in the L0I category for the 2012 conditions (part I).

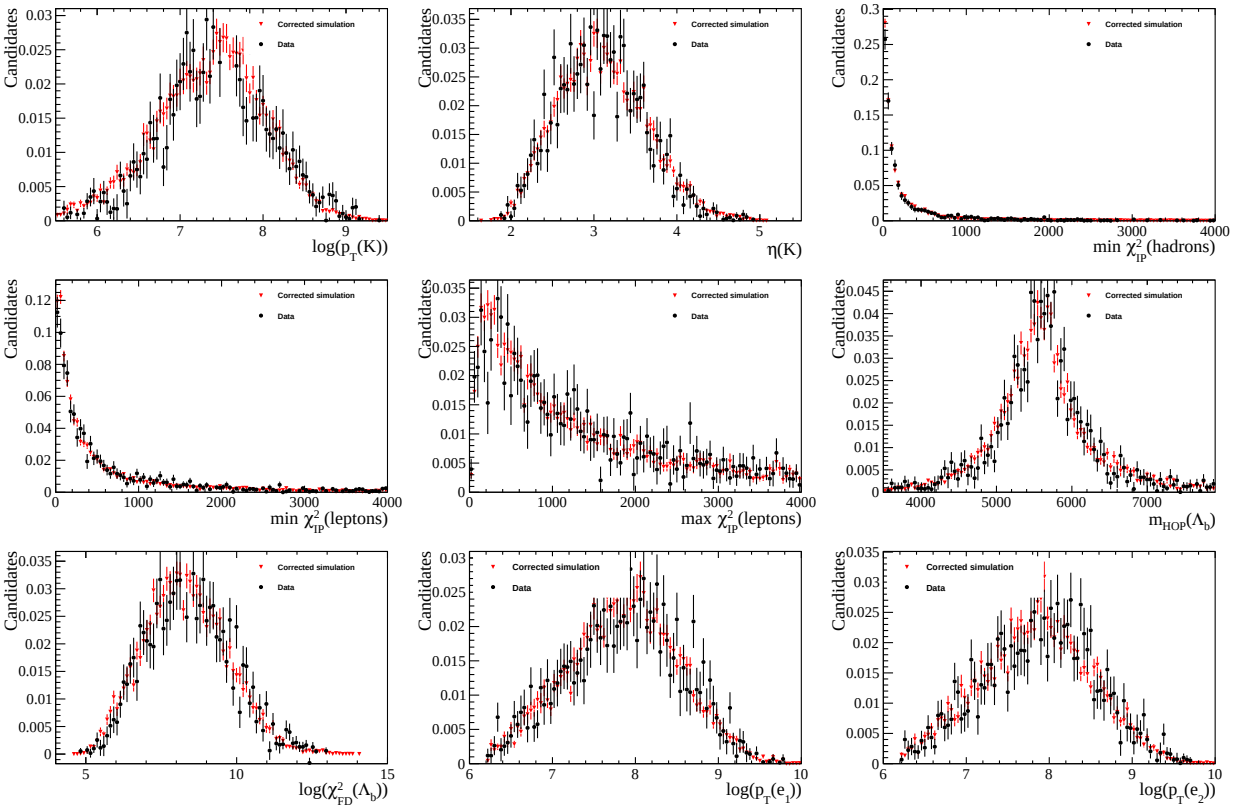


Figure 29: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, with nominal corrections applied, in the L0I category for the 2012 conditions (part II).

C.4 2016 $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$, L0M category

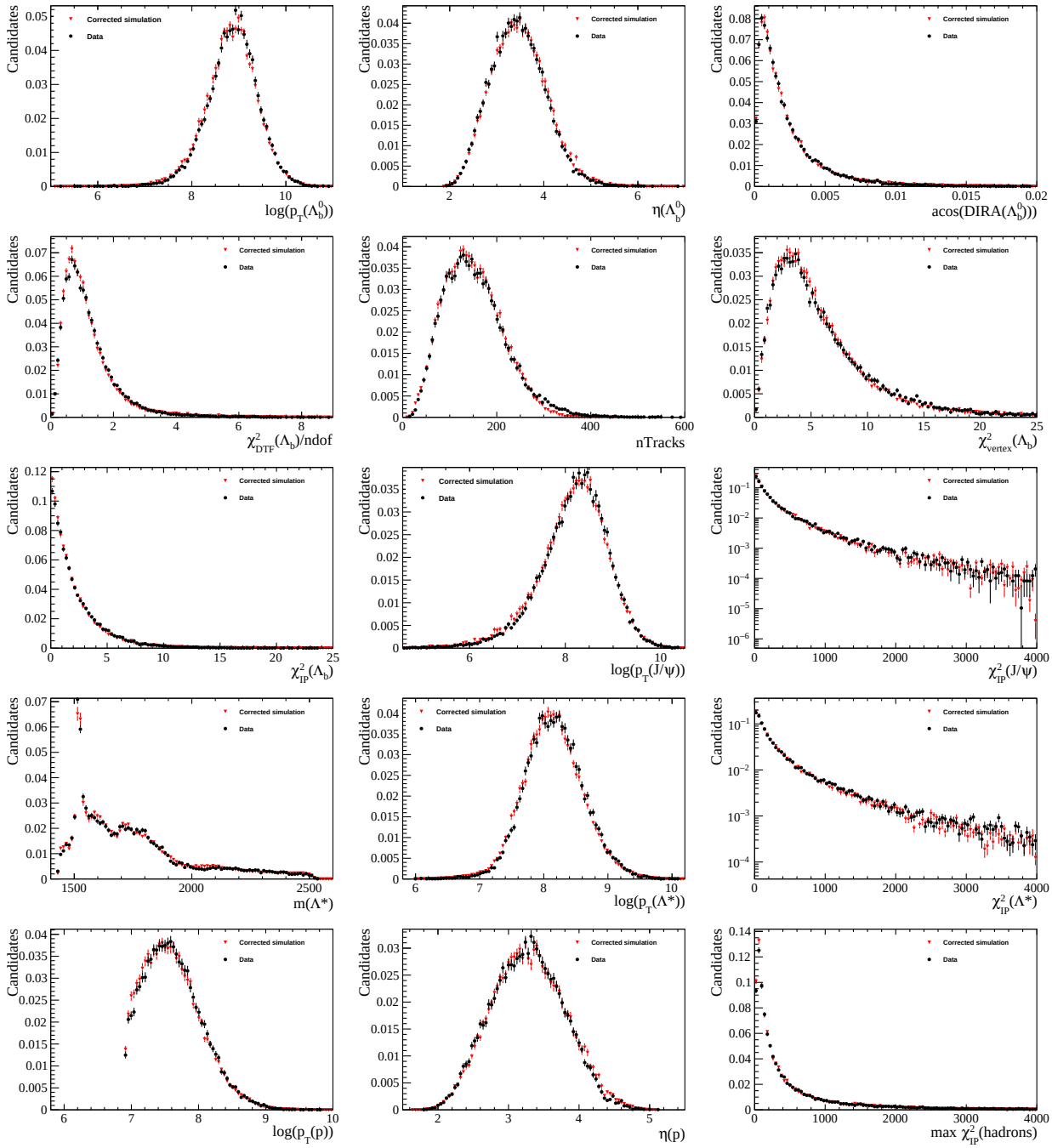


Figure 30: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode, with nominal corrections applied, in the L0M category for the 2016 conditions (part I).

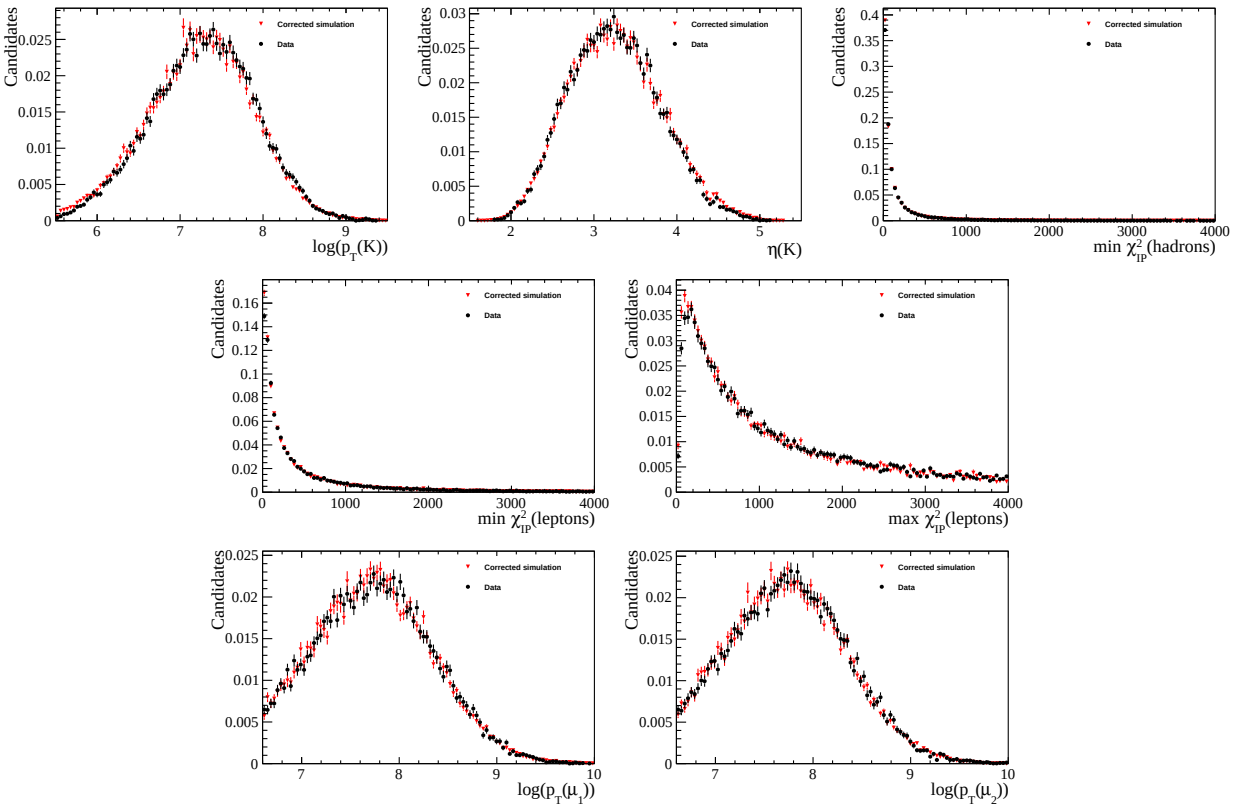


Figure 31: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ mode, with nominal corrections applied, in the L0M category for the 2016 conditions (part II).

C.5 2016 $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, L0E category

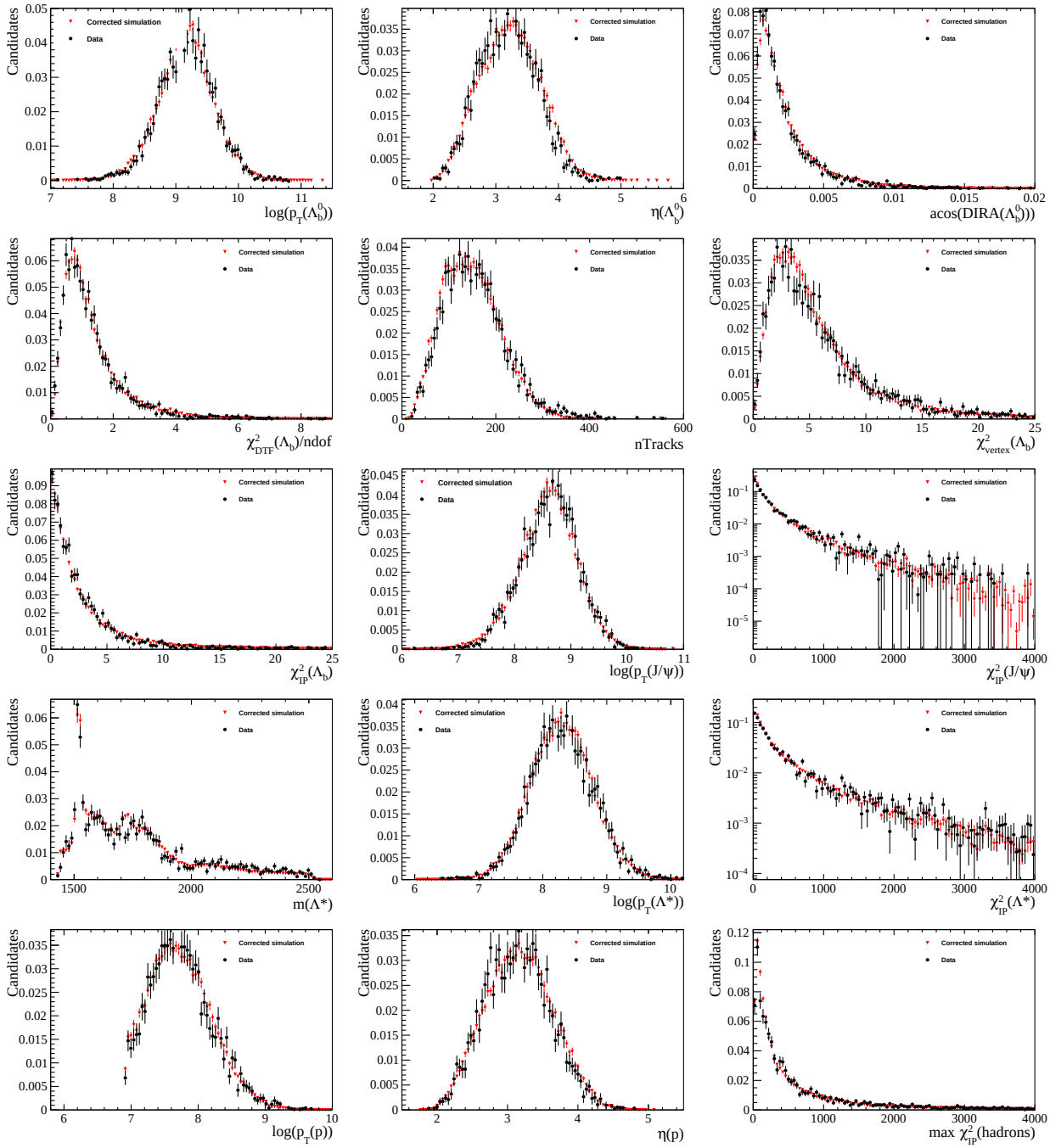


Figure 32: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode, with nominal corrections applied, in the L0E category for the 2016 conditions (part I).

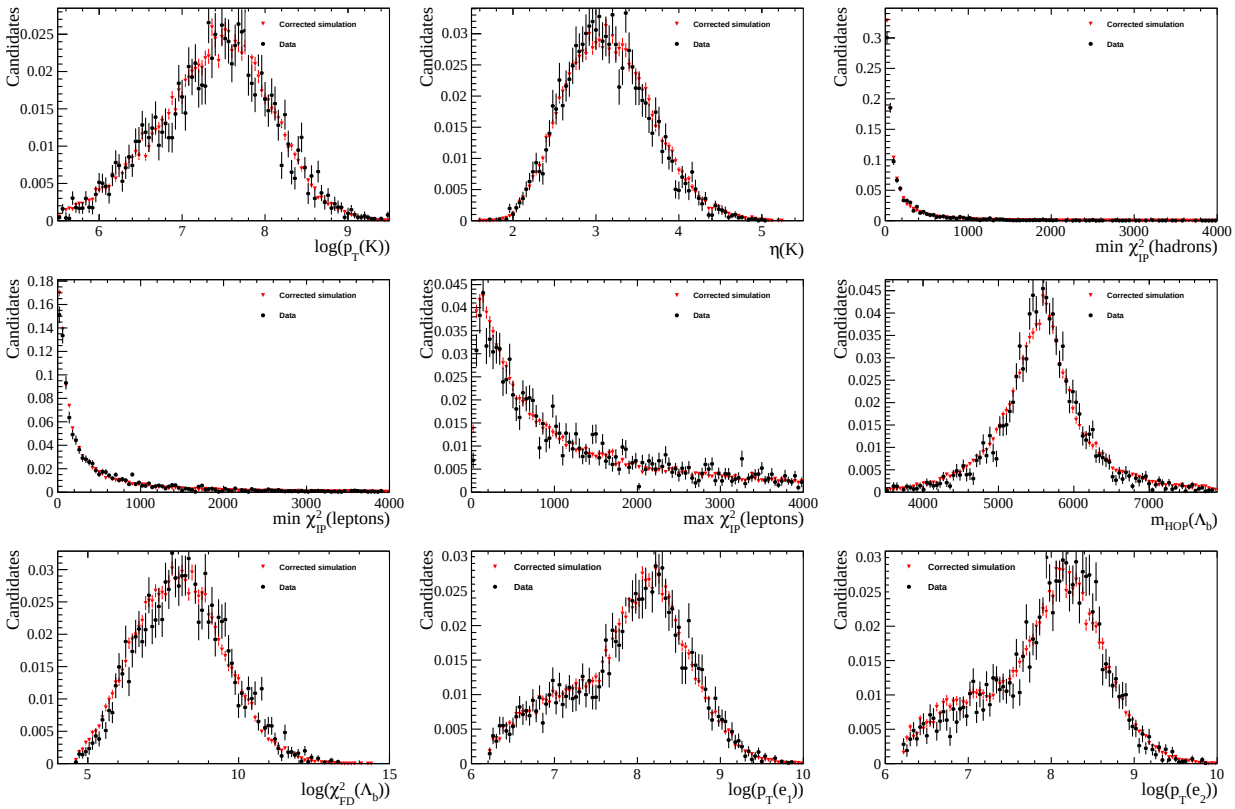


Figure 33: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, with nominal corrections applied, in the L0E category for the 2016 conditions (part II).

C.6 2016 $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$, L0I category

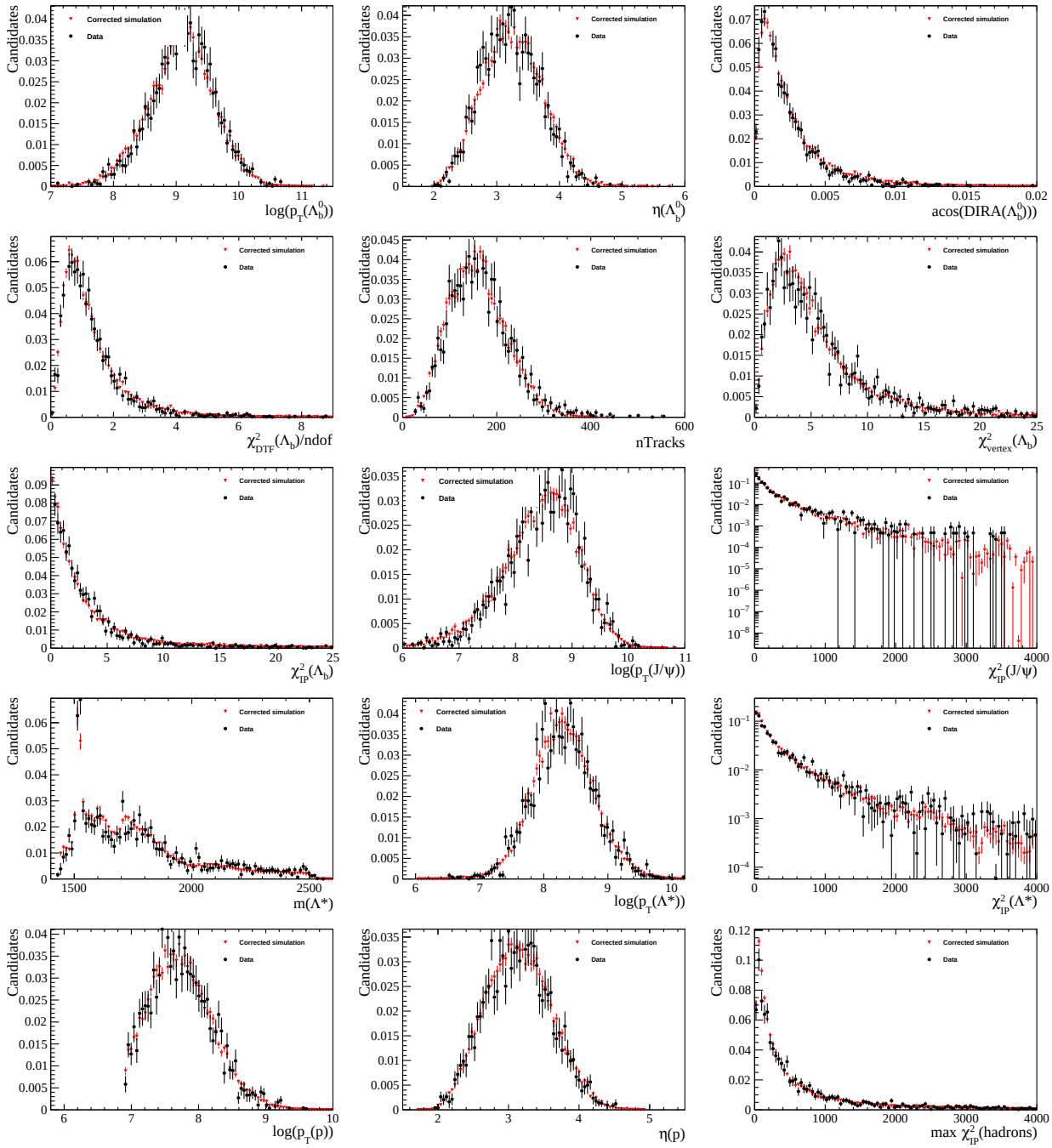


Figure 34: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ mode, with nominal corrections applied, in the L0I category for the 2016 conditions (part I).

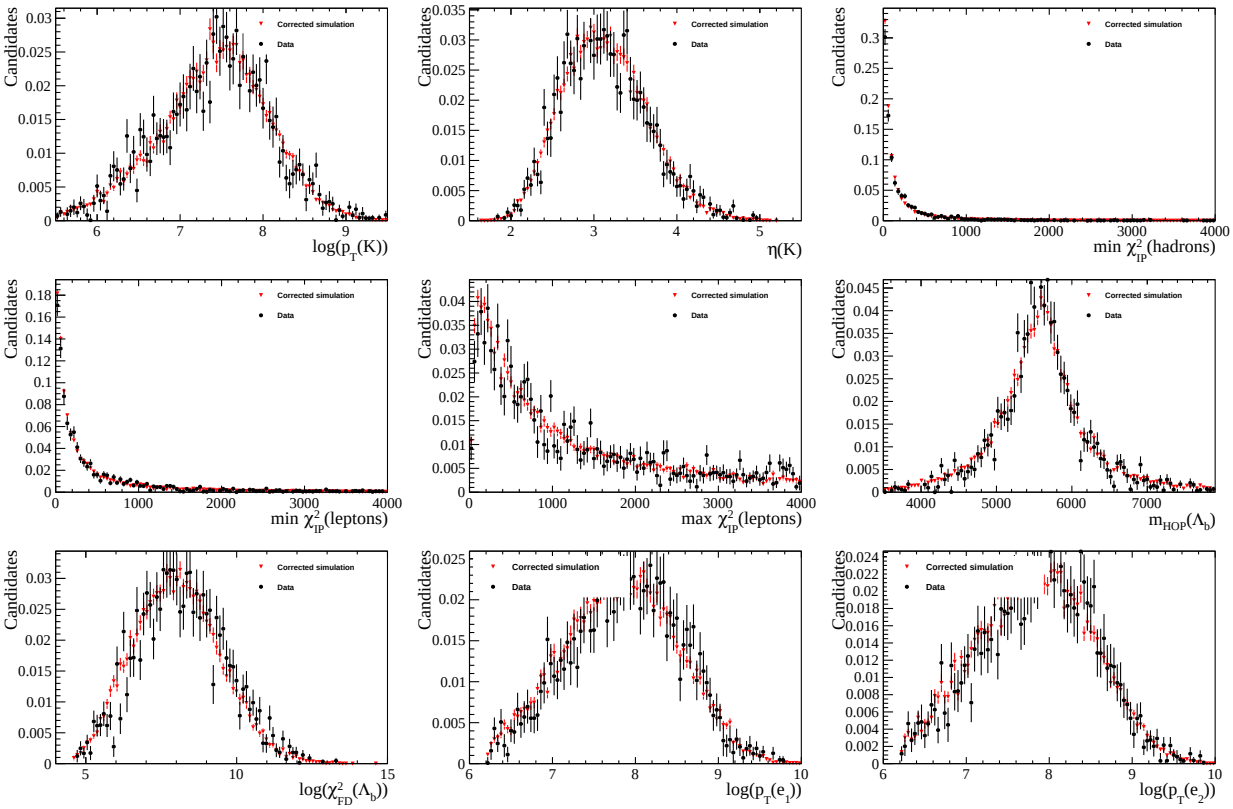
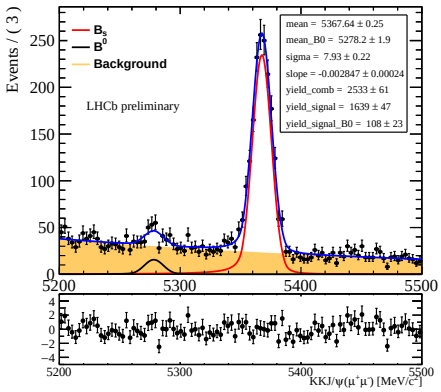


Figure 35: Data-simulation comparisons for $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ mode, with nominal corrections applied, in the L0I category for the 2016 conditions (part II).

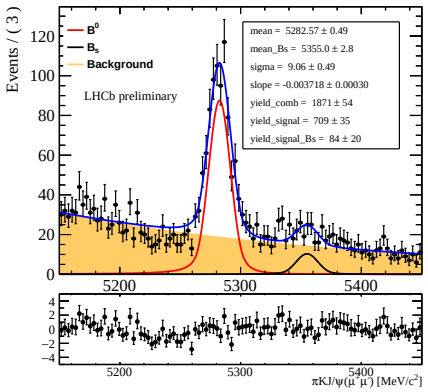
D Constraining background yields in fits to the data

The ratios of yields of peaking backgrounds are estimated using data. The result of this study is used as a constraint in the invariant mass fits.

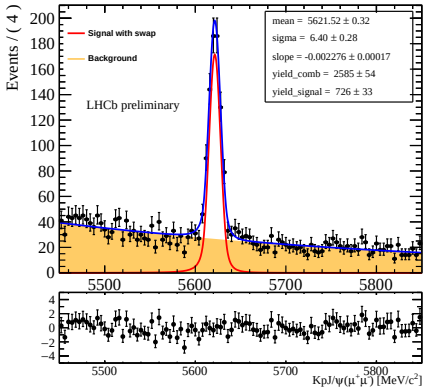
Invariant mass fits to the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ data in Run I are presented in Fig. 36, and in Run II – in Fig. 37. Four mass hypotheses are shown, with (a) proton being a kaon, (b) proton being a pion, (c) proton-kaon ID swap, (d) the reference plot without misidentifications. In each case, a $\pm 30\,\mathrm{MeV}/c^2$ veto is applied around the relevant mass peak in the other two other mass projections (for example $\Lambda_b^0 \to pKJ/\psi$ and $B^0 \to K\pi J/\psi$ when studying the KK projection, *etc.*), this is done in order to get a smooth background shape without additional bumps and make the estimate of the corresponding yields as clean as possible. For the hadron swap (c), three projections (pK, KK, πK) have a veto applied. The signal channel fit is also presented as a reference, with the KK, πK and swap components vetoed. Such a consistent approach allows to compare the yields of various misID components between themselves and to the signal yield, assuming that the veto efficiencies are of a similar order between the components.



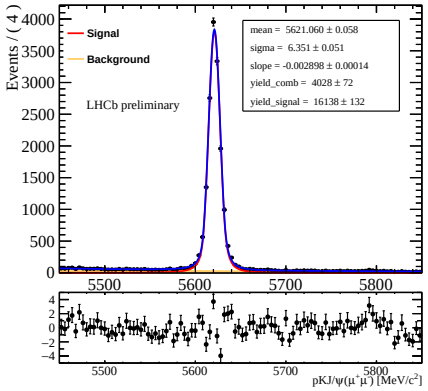
(a) $p \leftarrow K$ reflection



(b) $p \leftarrow \pi$ reflection

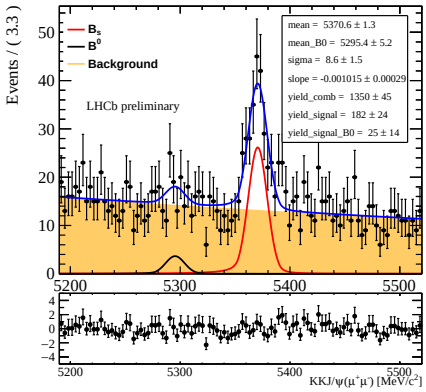


(c) $p \leftrightarrow K$ swap

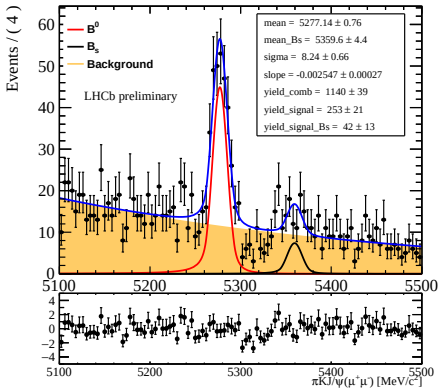


(d) Signal channel

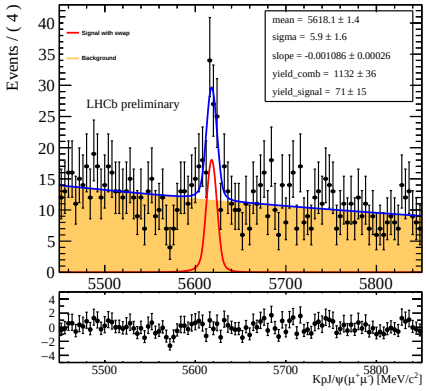
Figure 36: Fits to the misID mass projections in the $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ channel, in Run I data.



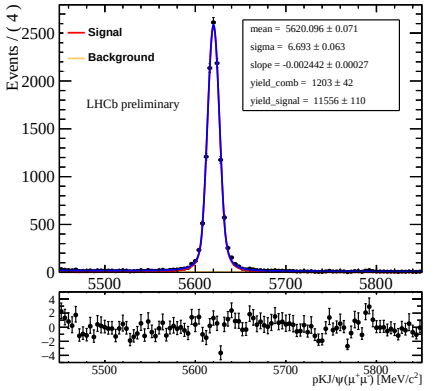
(a) $p \leftarrow K$ reflection



(b) $p \leftarrow \pi$ reflection



(c) $p \leftrightarrow K$ swap



(d) Signal channel

Figure 37: Fits to the misID mass projections in the $\Lambda_b^0 \to pKJ/\psi\,(\to \mu^+\mu^-)$ channel, in Run II data.

It can be seen that the background proportions are very different between Run I and Run II, owing to a tighter proton p_T requirement and a different PID response. The results are summarised in the Table 8.

Table 8: Ratios of misID backgrounds estimated from data.

Run period	$\frac{B_s^0 \to K^+K^-J/\psi}{B^0 \to K^{*0}J/\psi}$	$\frac{pK\text{swap}}{\text{signal}}$
Run I	2.31 ± 0.13	0.045 ± 0.002
Run II	0.72 ± 0.11	0.006 ± 0.001

E Trends of $\Lambda_b^0 \to pK\ell^+\ell^-$ selection efficiencies

In this chapter, trends of relative *selection* efficiency are shown in bins of variables related to the decay model. To avoid unblinding absolute values of efficiencies, all trends are normalised to unity.

It can be seen that efficiencies are usually flat in bins of $m(pK)$ and the helicity angle ϕ. Regarding the dependence on q^2, it is flat for the muon mode, and has a descending trend for the electron mode due to the HOP cut efficiency decreasing with q^2. Significant trends are observed in bins of other angles of the decay. The strong dependence on the dilepton opening angle is present, which is the strongest for L0E! trigger category.

E.1 $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode

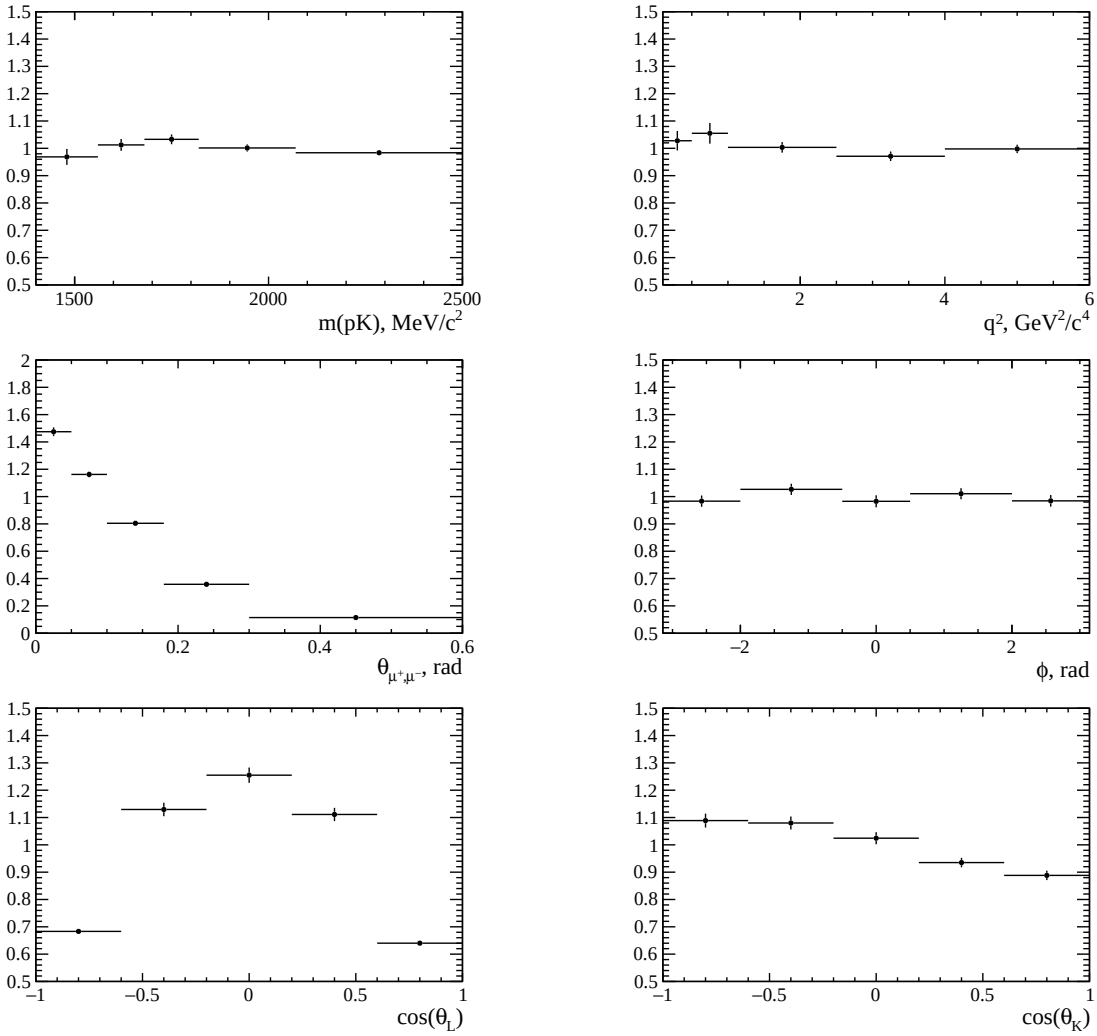


Figure 38: Selection efficiencies of the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, as a function of important decay parameters, in Run I simulation.

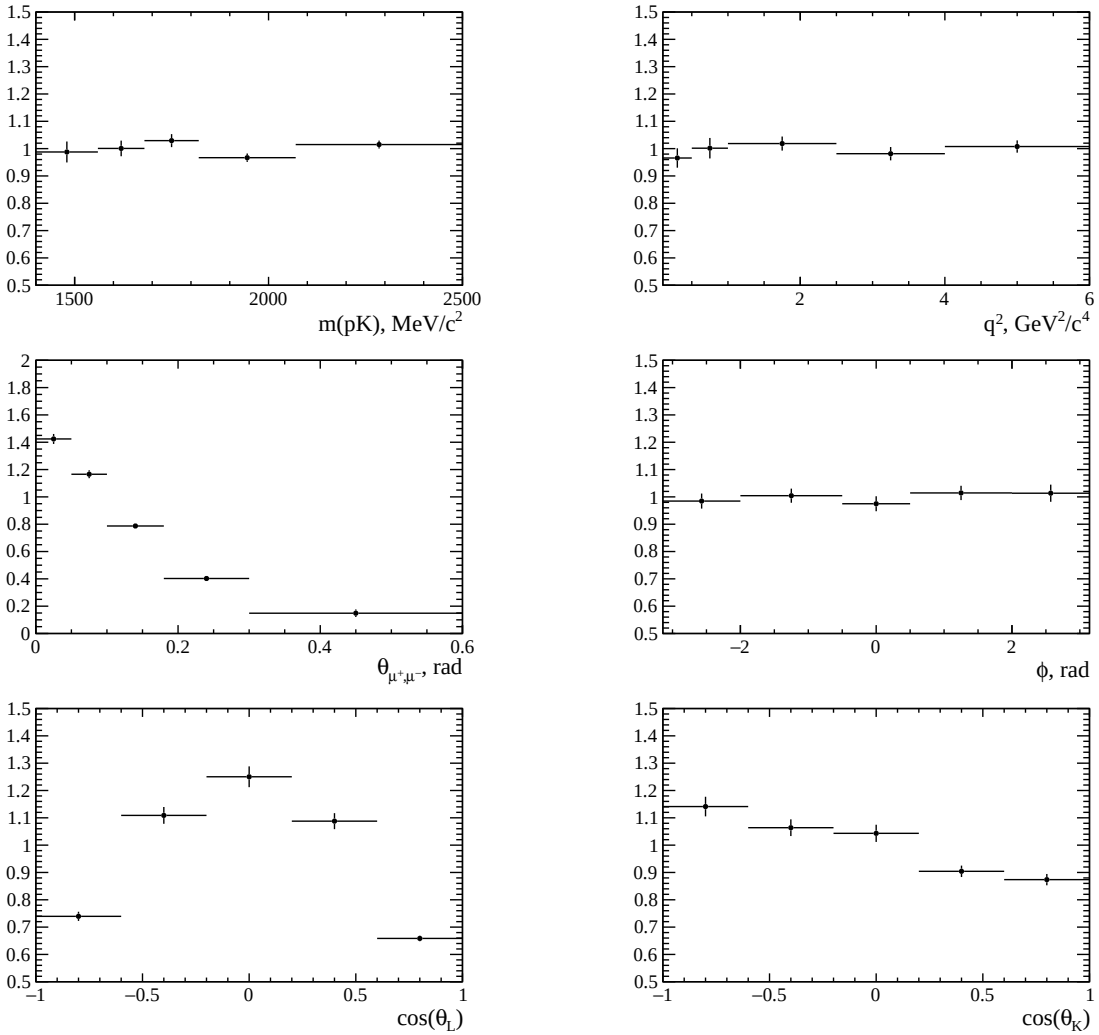


Figure 39: Selection efficiencies of the $\Lambda_b^0 \to pK\mu^+\mu^-$ decay mode, as a function of important decay parameters, in Run II simulation.

E.2 $\Lambda_b^0 \to pKe^+e^-$ decay mode

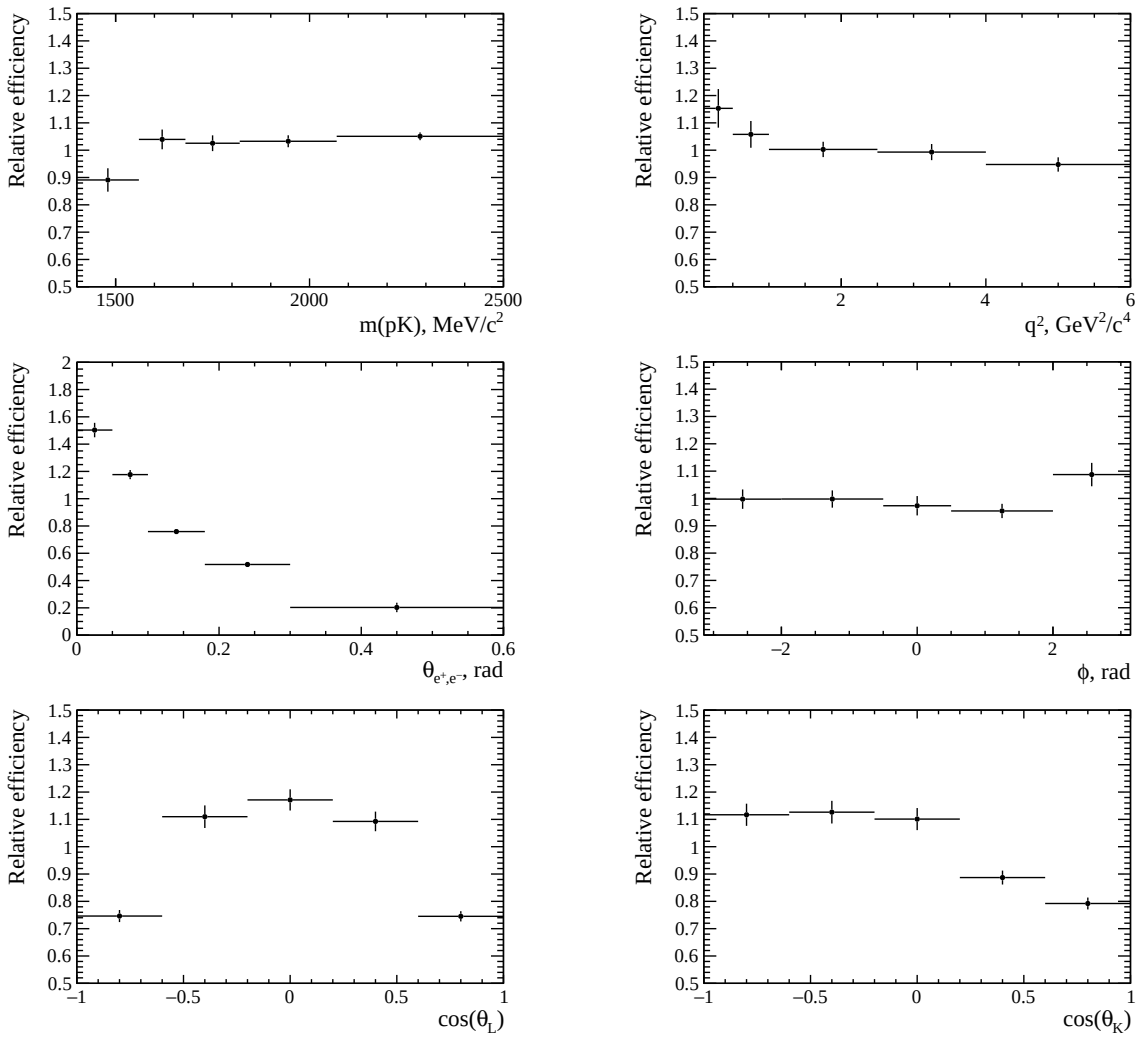


Figure 40: Selection efficiencies of the $\Lambda_b^0 \to pKe^+e^-$ decay mode in the L0I trigger category, as a function of important decay parameters, in Run I simulation.

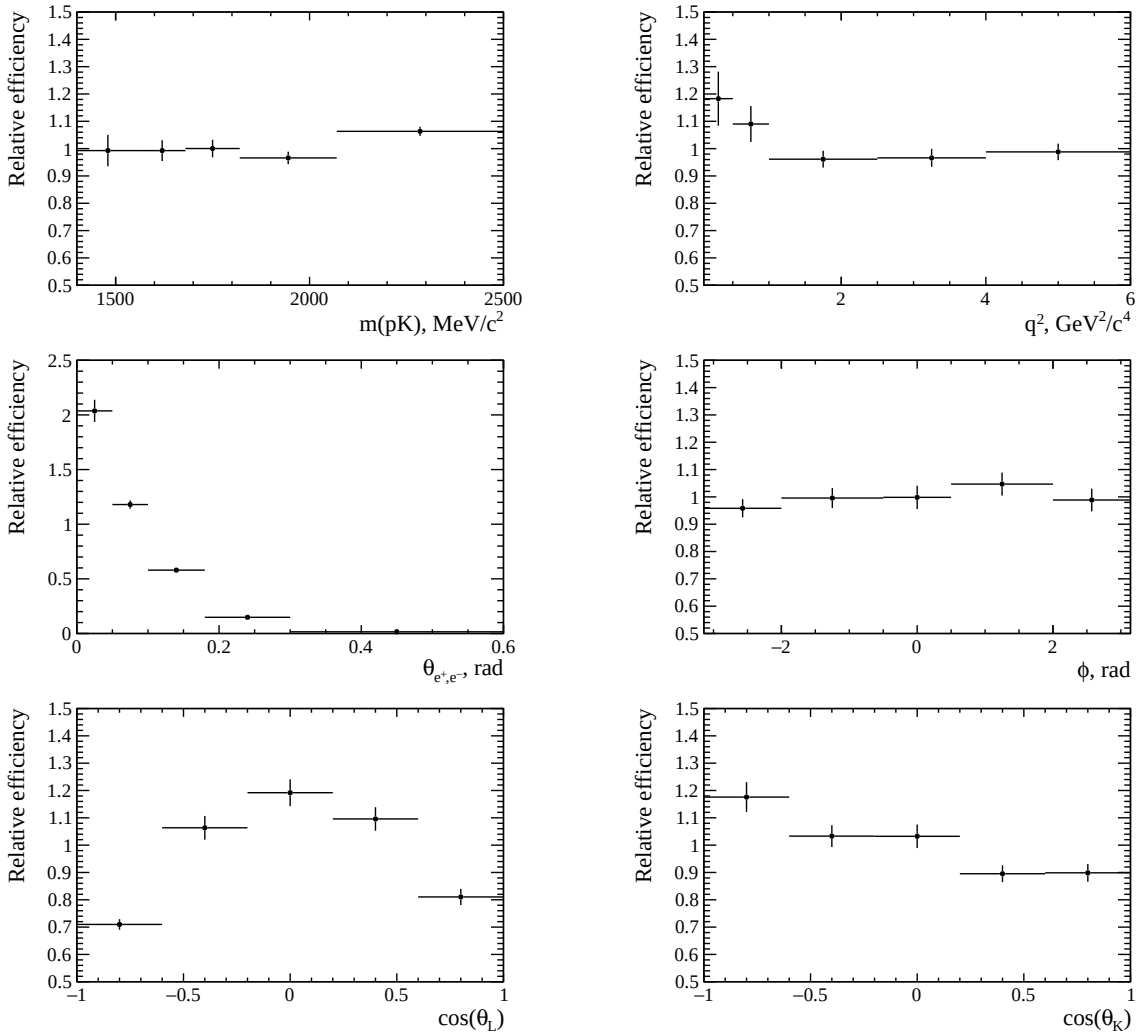


Figure 41: Selection efficiencies of the $\Lambda_b^0 \to pKe^+e^-$ decay mode in the L0E! trigger category, as a function of important decay parameters, in Run I simulation.

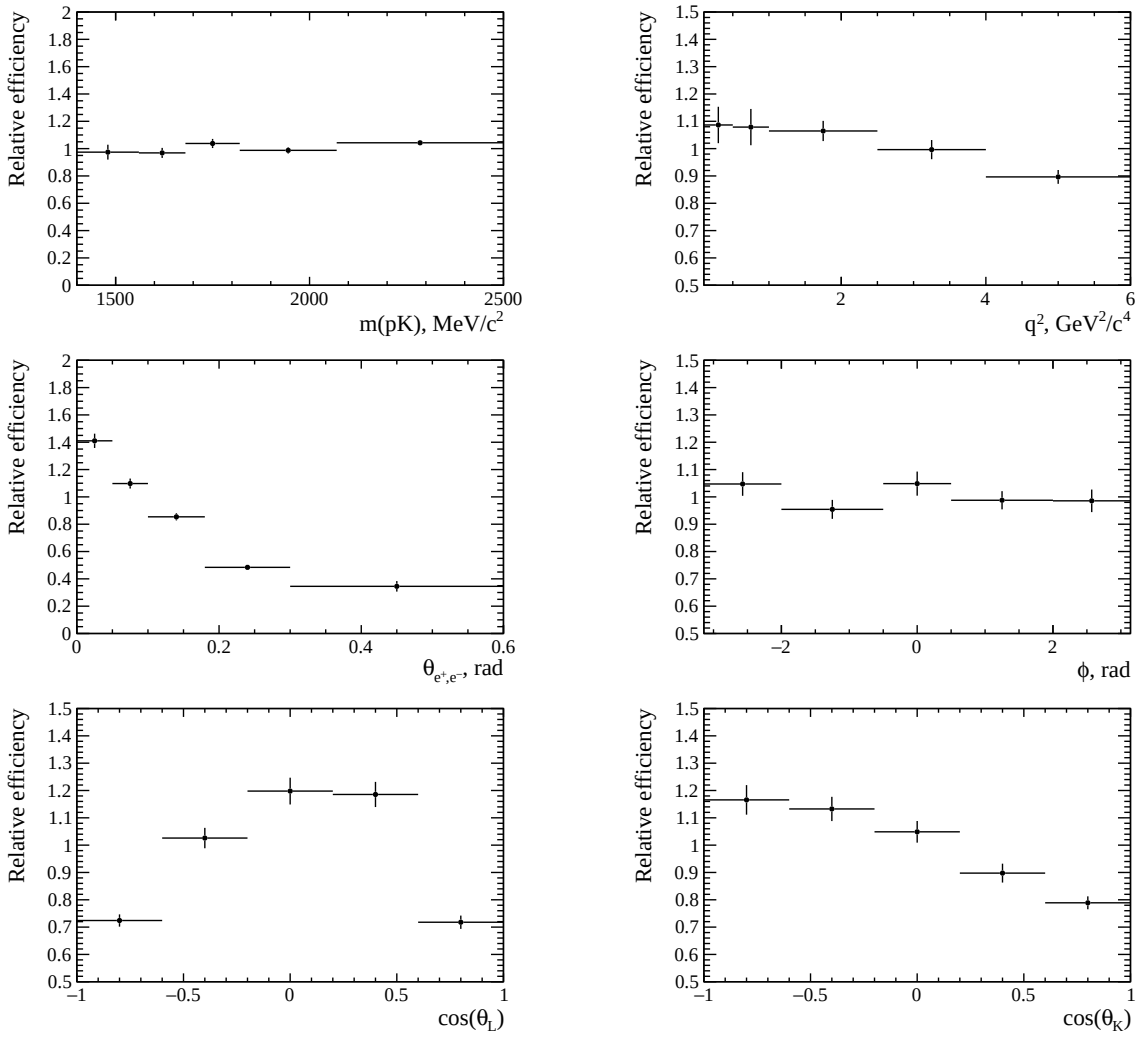


Figure 42: Selection efficiencies of the $\Lambda_b^0 \to pKe^+e^-$ decay mode in the L0I trigger category, as a function of important decay parameters, in Run II simulation.

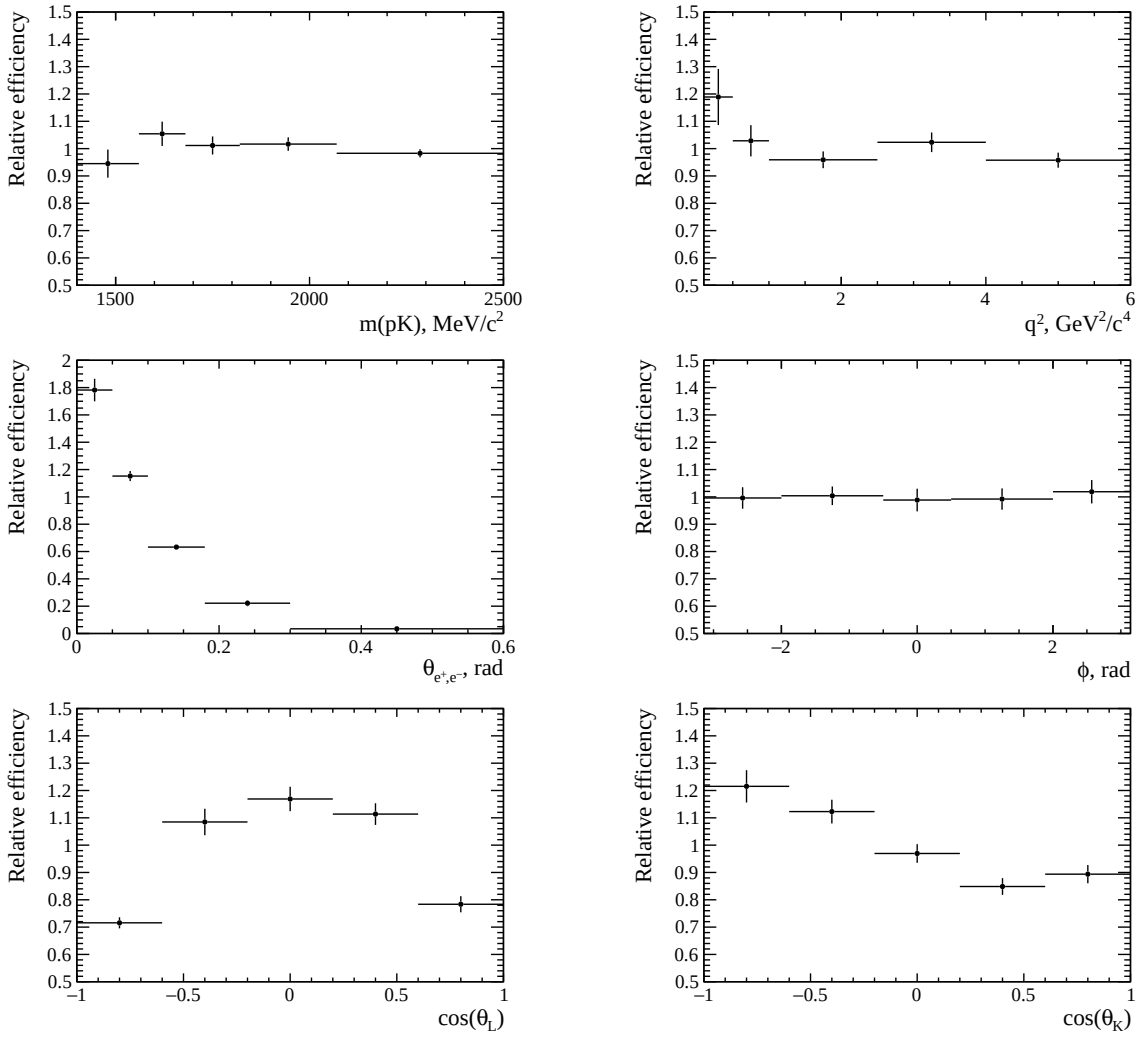


Figure 43: Selection efficiencies of the $\Lambda_b^0 \to pKe^+e^-$ decay mode in the L0E! trigger category, as a function of important decay parameters, in Run II simulation.

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, arXiv:1903.09252. (Cited on pages 2, 29, 33, and 243.)

[2] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, arXiv:1705.05802. (Cited on pages 2, 29, 33, 117, and 243.)

[3] LHCb collaboration, R. Aaij *et al.*, *Observation of the decay $\Lambda_b^0 \to pK^-\mu^+\mu^-$ and search for CP violation*, JHEP **06** (2017) 108, arXiv:1703.00256. (Cited on pages 2, 33, 35, 36, and 243.)

[4] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi pK^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414. (Cited on pages 5, 22, 33, 34, 75, 76, 120, 121, 222, and 245.)

[5] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow $P_c(4312)^+$ state and of two-peak structure of the $P_c(4450)^+$*, arXiv:1904.03947, submitted to Phys. Rev. Lett. (Cited on pages 5, 22, 76, and 245.)

[6] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001, and 2019 update. (Cited on pages 9, 12, 15, 20, 21, 23, 26, 36, 71, 91, 133, 164, 206, and 233.)

[7] J. Zupan, *Introduction to flavour physics*, in *2018 European School of High-Energy Physics (ESHEP2018) Maratea, Italy, June 20-July 3, 2018*, 2019. arXiv:1903.05062. (Cited on page 11.)

[8] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214. (Cited on pages 12 and 40.)

[9] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235. (Cited on pages 12 and 40.)

[10] S. Bifani, S. Descotes-Genon, A. Romero Vidal, and M.-H. Schune, *Review of Lepton Universality tests in B decays*, J. Phys. **G46** (2019), no. 2 023001, arXiv:1809.06229. (Cited on pages 15, 25, and 28.)

[11] PiENu, A. Aguilar-Arevalo *et al.*, *Improved Measurement of the $\pi \to e\nu$ Branching Ratio*, Phys. Rev. Lett. **115** (2015), no. 7 071801, arXiv:1506.05845. (Cited on page 16.)

[12] V. Cirigliano and I. Rosell, *Two-loop effective theory analysis of $\pi(K) \to e\bar{\nu}_e[\gamma]$ branching ratios*, Phys. Rev. Lett. **99** (2007) 231801, arXiv:0707.3439. (Cited on page 16.)

[13] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531, [648(1963)]. (Cited on page 16.)

[14] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652. (Cited on page 16.)

[15] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945. (Cited on page 17.)

[16] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass, $K^+ \to \pi^+\nu\bar{\nu}$, $B^0_{(s)}$ - $\overline{B^0_{(s)}}$ mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433, arXiv:hep-ph/9403384. (Cited on page 17.)

[17] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K^0_2 Meson*, Phys. Rev. Lett. **13** (1964) 138. (Cited on page 17.)

[18] KTeV Collaboration, A. Alavi-Harati *et al.*, *Observation of Direct CP Violation in $K_{S,L} \to \pi\pi$ Decays*, Phys. Rev. Lett. **83** (1999) 22. (Cited on page 17.)

[19] Belle Collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802. (Cited on page 17.)

[20] BABAR Collaboration, B. Aubert *et al.*, *Observation of CP Violation in the B^0 Meson System*, Phys. Rev. Lett. **87** (2001) 091801. (Cited on page 17.)

[21] Belle, K. Abe *et al.*, *Time-dependent CP asymmetries in $b \to s\bar{q}q$ transitions and $\sin 2\phi_1$ in $B^0 \to J/\psi K^0$ decays with 386 million B anti-B pairs*, in *Lepton and photon interactions at high energies. Proceedings, 22nd International Symposium, LP 2005, Uppsala, Sweden, June 30-July 5, 2005*, 2005. arXiv:hep-ex/0507037. (Cited on page 17.)

[22] BaBar, B. Aubert *et al.*, *Improved measurement of CP asymmetries in $B^0 \to (c\bar{c})K^{(*)0}$ decays*, Phys. Rev. Lett. **94** (2005) 161803, arXiv:hep-ex/0408127. (Cited on page 17.)

[23] BaBar, B. Aubert *et al.*, *Evidence for Direct CP Violation from Dalitz-plot analysis of $B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$*, Phys. Rev. **D78** (2008) 012004, arXiv:0803.4451. (Cited on page 17.)

[24] Belle, A. Garmash *et al.*, *Evidence for large direct CP violation in $B^{\pm} \to \rho(770)^0 K^{\pm}$ from analysis of the three-body charmless $B^{\pm} \to K^{\pm}\pi^{\pm}\pi^{\mp}$ decay*, Phys. Rev. Lett. **96** (2006) 251803, arXiv:hep-ex/0512066. (Cited on page 17.)

[25] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013) 221601, arXiv:1304.6173. (Cited on page 17.)

[26] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in charm decays*, arXiv:1903.08726, submitted to Phys. Rev. Lett. (Cited on page 17.)

[27] T. Gershon and V. V. Gligorov, *CP violation in the B system*, Rept. Prog. Phys. **80** (2017), no. 4 046201, arXiv:1607.06746. (Cited on page 17.)

[28] R. N. Cahn, *The eighteen arbitrary parameters of the standard model in your everyday life*, Rev. Mod. Phys. **68** (1996) 951. (Cited on page 17.)

[29] U. Husemann, *Top-Quark Physics: Status and Prospects*, Prog. Part. Nucl. Phys. **95** (2017) 48, arXiv:1704.01356. (Cited on page 18.)

[30] D. Flamm and F. Schoberl, *INTRODUCTION TO THE QUARK MODEL OF ELEMENTARY PARTICLES. VOL. 1. QUANTUM NUMBERS, GAUGE THEORIES AND HADRON SPECTROSCOPY*, 1982. (Cited on page 19.)

[31] V. V. Anisovich, M. N. Kobrinsky, J. Nyiri, and Yu. M. Shabelski, *QUARK MODEL AND HIGH-ENERGY COLLISIONS*, 1986. (Cited on page 19.)

[32] A. Seiden, *Particle physics: A comprehensive introduction*, 2004. (Cited on page 20.)

[33] A. Hosaka and H. Toki, *Quarks, baryons and chiral symmetry*, 2001. (Cited on page 20.)

[34] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214. (Cited on pages 20 and 22.)

[35] L. Maiani, A. D. Polosa, and V. Riquer, *The New Pentaquarks in the Diquark Model*, Phys. Lett. **B749** (2015) 289, arXiv:1507.04980. (Cited on page 22.)

[36] R. Ghosh, A. Bhattacharya, and B. Chakrabarti, *A study on P_c^* (4380) and P_c^* in the quasi particle diquark model*, arXiv:1508.00356arXiv:1508.00356, [Phys. Part. Nucl. Lett.14,no.4,550(2017)]. (Cited on page 22.)

[37] M. Karliner and J. L. Rosner, *New Exotic Meson and Baryon Resonances from Doubly-Heavy Hadronic Molecules*, Phys. Rev. Lett. **115** (2015), no. 12 122001, arXiv:1506.06386. (Cited on page 22.)

[38] J. He, *$\overline{D}\Sigma_c^*$ and $\overline{D}^*\Sigma_c$ interactions and the LHCb hidden-charmed pentaquarks*, Phys. Lett. **B753** (2016) 547, arXiv:1507.05200. (Cited on page 22.)

[39] U.-G. Meißner and J. A. Oller, *Testing the $\chi_{c1}p$ composite nature of the $P_c(4450)$*, Phys. Lett. **B751** (2015) 59, arXiv:1507.07478. (Cited on page 22.)

[40] LHCb collaboration, R. Aaij *et al.*, *Observation of two resonances in the $\Lambda_b^0\pi^\pm$ systems and precise measurement of $\Sigma_b^\pm$ and $\Sigma_b^{*\pm}$ properties*, Phys. Rev. Lett. **122** (2019) 012001, arXiv:1809.07752. (Cited on page 22.)

[41] M. Karliner, B. Keren-Zur, H. J. Lipkin, and J. L. Rosner, *The Quark Model and b Baryons*, Annals Phys. **324** (2009) 2, arXiv:0804.1575. (Cited on page 22.)

[42] UA1, C. Albajar *et al.*, *First observation of the beauty baryon Λ_b^0 in the decay channel $\Lambda_b^0 \to J/\psi\Lambda$ at the CERN proton - anti-proton collider*, Phys. Lett. **B273** (1991) 540, [249(1992)]. (Cited on page 22.)

[43] CDF, T. Aaltonen *et al.*, *Observation of the Ξ_b^0 Baryon*, Phys. Rev. Lett. **107** (2011) 102001, arXiv:1107.4015. (Cited on page 22.)

[44] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^0 baryon*, Phys. Rev. Lett. **113** (2014) 032001, arXiv:1405.7223. (Cited on pages 22 and 232.)

[45] LHCb collaboration, R. Aaij *et al.*, *Study of beauty baryon decays to D^0ph^- and $\Lambda_c^+h^-$ final states*, Phys. Rev. **D89** (2014) 032001, arXiv:1311.4823. (Cited on page 22.)

[46] LHCb collaboration, R. Aaij *et al.*, *Measurement of branching fractions of charmless four-body Λ_b^0 and Ξ_b^0 decays*, JHEP **02** (2018) 098, arXiv:1711.05490. (Cited on page 22.)

[47] LHCb collaboration, R. Aaij *et al.*, *Search for CP violation using triple product asymmetries in $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$, and $\Xi_b^0 \to pK^-K^-\pi^+$ decays*, JHEP **08** (2018) 039, arXiv:1805.03941. (Cited on pages 22 and 95.)

[48] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210** (2017) 103, arXiv:1601.05437. (Cited on page 24.)

[49] A. Buras and M. Lindner, *Heavy Flavors II*, 1992. (Cited on page 24.)

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

[51] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, in *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2*, pp. 281–539, 1998. arXiv:hep-ph/9806471. (Cited on page 24.)

[52] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, arXiv:1503.05534. (Cited on page 25.)

[53] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias, *Hadronic uncertainties in $B \to K^* \mu^+ \mu^-$: a state-of-the-art analysis*, JHEP **04** (2017) 016, arXiv:1701.08672. (Cited on page 25.)

[54] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^* \mu^+ \mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044. (Cited on pages 27 and 28.)

[55] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \to \phi \mu^+ \mu^-$*, JHEP **07** (2013) 084, arXiv:1305.2168. (Cited on page 27.)

[56] LHCb collaboration, R. Aaij *et al.*, *Measurement of the S-wave fraction in $B^0 \to K^+ \pi^- \mu^+ \mu^-$ decays and the $B^0 \to K^*(892)^0 \mu^+ \mu^-$ differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, arXiv:1606.04731. (Cited on pages 27 and 28.)

[57] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decays*, JHEP **06** (2015) 115, Erratum ibid. **09** (2018) 145, arXiv:1503.07138. (Cited on pages 27 and 37.)

[58] G. Hiller and F. Kruger, *More model-independent analysis of $b \to s$ processes*, Phys. Rev. **D69** (2004) 074020, arXiv:hep-ph/0310219. (Cited on page 28.)

[59] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}*, Eur. Phys. J. **C76** (2016), no. 8 440, arXiv:1605.07633. (Cited on pages 28, 29, and 73.)

[60] B. Capdevila *et al.*, *Patterns of New Physics in $b \to s \ell^+ \ell^-$ transitions in the light of recent data*, JHEP **01** (2018) 093, arXiv:1704.05340. (Cited on page 29.)

[61] L.-S. Geng *et al.*, *Towards the discovery of new physics with lepton-universality ratios of $b \to s \ell \ell$ decays*, Phys. Rev. **D96** (2017), no. 9 093006, arXiv:1704.05446. (Cited on page 29.)

[62] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *Status of the $B \to K^* \mu^+ \mu^-$ anomaly after Moriond 2017*, Eur. Phys. J. **C77** (2017), no. 6 377, arXiv:1703.09189. (Cited on page 29.)

[63] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482. (Cited on pages 29 and 33.)

[64] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)} l^+ l^-$*, Phys. Rev. **D86** (2012) 032012, arXiv:1204.3933. (Cited on page 29.)

[65] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)}\ell^+\ell^-$*, Phys. Rev. Lett. **103** (2009) 171801, arXiv:0904.0770. (Cited on page 29.)

[66] Belle, A. Abdesselam *et al.*, *Test of lepton flavor universality in $B \to K^*\ell^+\ell^-$ decays at Belle*, arXiv:1904.02440. (Cited on page 29.)

[67] A. Abdesselam *et al.*, *Test of lepton flavor universality in $B \to K\ell^+\ell^-$ decays*, arXiv:1908.01848. (Cited on page 29.)

[68] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, JHEP **08** (2013) 131, arXiv:1304.6325. (Cited on pages 30 and 31.)

[69] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of $B \to K^*\mu^+\mu^-$ at large recoil*, JHEP **01** (2013) 048, arXiv:1207.2753. (Cited on page 30.)

[70] J. Matias, F. Mescia, M. Ramon, and J. Virto, *Complete Anatomy of $\bar{B}_d \to \bar{K}^{*0}(\to K\pi)l^+l^-$ and its angular distribution*, JHEP **04** (2012) 104, arXiv:1202.4266. (Cited on page 30.)

[71] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442. (Cited on pages 30 and 31.)

[72] ATLAS, M. Aaboud *et al.*, *Angular analysis of $B^0_d \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2018) 047, arXiv:1805.04000. (Cited on pages 30 and 31.)

[73] CMS, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, arXiv:1710.02846. (Cited on pages 30 and 31.)

[74] Belle, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017), no. 11 111801, arXiv:1612.05014. (Cited on pages 30, 31, and 32.)

[75] B. Capdevila, S. Descotes-Genon, J. Matias, and J. Virto, *Assessing lepton-flavour non-universality from $B \to K^*\ell\ell$ angular analyses*, JHEP **10** (2016) 075, arXiv:1605.03156. (Cited on page 32.)

[76] M. Ciuchini *et al.*, *New Physics in $b \to s\ell^+\ell^-$ confronts new data on Lepton Universality*, arXiv:1903.09632. (Cited on page 32.)

[77] A. K. Alok, A. Dighe, S. Gangal, and D. Kumar, *Continuing search for new physics in $b \to s\mu\mu$ decays: two operators at a time*, arXiv:1903.09617. (Cited on page 32.)

[78] M. Algueró *et al.*, *Emerging patterns of New Physics with and without Lepton Flavour Universal contributions*, arXiv:1903.09578. (Cited on page 32.)

[79] J. Aebischer *et al.*, *B-decay discrepancies after Moriond 2019*, arXiv:1903.10434. (Cited on page 32.)

[80] A. Arbey *et al.*, *Update on the $b \to s$ anomalies*, arXiv:1904.08399. (Cited on page 32.)

[81] A. Angelescu, D. Bečirević, D. A. Faroughy, and O. Sumensari, *Closing the window on single leptoquark solutions to the B-physics anomalies*, JHEP **10** (2018) 183, arXiv:1808.08179. (Cited on page 32.)

[82] C. Cornella, J. Fuentes-Martin, and G. Isidori, *Revisiting the vector leptoquark explanation of the B-physics anomalies*, arXiv:1903.11517. (Cited on page 32.)

[83] S. Meinel and G. Rendon, *Lattice QCD calculation of form factors for $\Lambda_b \to \Lambda(1520)\ell^+\ell^-$ decays*, PoS **LATTICE2016** (2016) 299, arXiv:1608.08110. (Cited on page 33.)

[84] S. Descotes-Genon and M. Novoa Brunet, *Angular analysis of the rare decay $\Lambda_b \to \Lambda(1520)(\to NK)\ell^+\ell^-$*, arXiv:1903.00448. (Cited on page 33.)

[85] 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. (Cited on pages 35 and 37.)

[86] 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 $\Lambda_b^0 \to J/\psi\, pK^-$ branching fraction*, Chin. Phys. C **40** (2016) 011001, arXiv:1509.00292. (Cited on pages 36, 179, and 206.)

[87] LHCb collaboration, R. Aaij *et al.*, *Observation of $\Lambda_b^0 \to \psi(2S)pK^-$ and $\Lambda_b^0 \to J/\psi\,\pi^+\pi^- pK^-$ decays and a measurement of the Λ_b^0 baryon mass*, JHEP **05** (2016) 132, arXiv:1603.06961. (Cited on pages 36, 177, and 188.)

[88] LHCb collaboration, R. Aaij *et al.*, *Observation of the suppressed decay $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$*, JHEP **04** (2017) 029, arXiv:1701.08705. (Cited on page 36.)

[89] LHCb collaboration, R. Aaij *et al.*, *Angular moments of the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$*, JHEP **09** (2018) 146, arXiv:1808.00264. (Cited on page 37.)

[90] LHCb collaboration, R. Aaij *et al.*, *Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes*, JHEP **04** (2014) 114, arXiv:1402.2554. (Cited on page 37.)

[91] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008), no. LHCb-DP-2008-001 S08005. (Cited on pages 39, 42, 45, and 49.)

[92] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1. (Cited on page 39.)

[93] J.-P. Tock *et al.*, *The Second LHC Long Shutdown (LS2) for the Superconducting Magnets*, in *Proceedings, 9th International Particle Accelerator Conference (IPAC 2018): Vancouver, BC Canada, April 29-May 4, 2018*, p. MOPMF056, 2018. doi: 10.18429/JACoW-IPAC2018-MOPMF056. (Cited on page 39.)

[94] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, , General Photo. (Cited on page 40.)

[95] LHCb collaboration, C. Elsässer, *$b\bar{b}$ production angle plots*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html. (Cited on page 41.)

[96] 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. (Cited on page 42.)

[97] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, arXiv:1902.06794, submitted to Phys. Rev. Lett. (Cited on pages 42 and 132.)

[98] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140. (Cited on page 42.)

[99] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s}$ =7 TeV*, JHEP **02** (2015) 029, arXiv:1412.2500. (Cited on page 42.)

[100] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s}$ =13 TeV*, JHEP **06** (2018) 100, arXiv:1803.10974. (Cited on page 42.)

[101] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. (Cited on pages 42, 45, 49, and 60.)

[102] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352. (Cited on pages 42, 43, 44, 46, 50, and 60.)

[103] LHCb collaboration, LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. (Cited on page 42.)

[104] LHCb Collaboration, *LHCb Performance Plots*, https://lbgroups.cern.ch/online/OperationsPlots/index.htm. (Cited on page 44.)

[105] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251. (Cited on pages 46 and 49.)

[106] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, Tech. Rep. LHCb-PUB-2011-001. CERN-LHCb-PUB-2011-001, CERN, Geneva, Jan, 2011. LHCb. (Cited on page 46.)

[107] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, Tech. Rep. LHCb-2007-015. CERN-LHCb-2007-015, CERN, Geneva, May, 2007. (Cited on page 46.)

[108] W. Hulsbergen, *The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, 0810.22410810.2241. (Cited on page 47.)

[109] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, Tech. Rep. LHCb-PUB-2017-011. CERN-LHCb-PUB-2017-011, CERN, Geneva, Mar, 2017. (Cited on pages 47 and 108.)

[110] A. Perieanu, *A Fast Algorithm to Identify and Remove Clone Tracks*, Tech. Rep. LHCb-2008-020. CERN-LHCb-2008-020, CERN, Geneva, May, 2008. (Cited on page 47.)

[111] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, Tech. Rep. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, CERN, Geneva, Jan, 2008. (Cited on page 47.)

[112] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. (Cited on page 47.)

[113] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808. (Cited on pages 47 and 49.)

[114] LHCb Collaboration, *Approved VELO conference plots*, https://lbtwiki.cern.ch/bin/view/VELO/VELOConferencePlots. (Cited on page 47.)

[115] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021. (Cited on page 48.)

[116] M. Alexander *et al.*, *Mapping the material in the LHCb vertex locator using secondary hadronic interactions*, JINST **13** (2018) P06008, arXiv:1803.07466. (Cited on page 48.)

[117] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, arXiv:1812.10790. (Cited on pages 48, 49, and 50.)

[118] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. (Cited on page 48.)

[119] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. (Cited on page 48.)

[120] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893. (Cited on page 48.)

[121] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, arXiv:1708.00819. (Cited on page 48.)

[122] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. (Cited on page 51.)

[123] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759. (Cited on pages 51 and 52.)

[124] LHCb, D. Derkach *et al.*, *Machine-Learning-based global particle-identification algorithms at the LHCb experiment*, J. Phys. : Conf. Ser. **1085** (2018), no. 4 042038. 5 p. (Cited on page 51.)

[125] LHCb Collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000. (Cited on page 52.)

[126] P. Robbe, *Generators, Calorimeter Trigger and J/ψ production at LHCb*, habilitation à diriger des recherches, Université Paris Sud - Paris XI, Mar., 2012. (Cited on pages 53 and 57.)

[127] M. Borsato, M.-H. Schune, and N. Arnaud, *Study of the $B^0 \to K^{*0}e^+e^-$ decay with the LHCb detector and development of a novel concept of PID detector: the Focusing DIRC*, 2015. Presented 08 Sep 2015. (Cited on page 53.)

[128] P. Perret and X. Vilasis-Cardona, *Performance of the LHCb calorimeters during the period 2010-2012*, J. Phys. : Conf. Ser. **587** (2014) 012012. 6 p, see LHCb-TALK-2014-236. (Cited on page 54.)

[129] V. V. Gligorov, *Conceptualization, implementation, and commissioning of real-time analysis in the High Level Trigger of the LHCb experiment*, arXiv:1806.10912. (Cited on page 54.)

[130] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001. (Cited on page 54.)

[131] LHCb Collaboration, LHCb Collaboration, *LHCb muon system: second addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2005. Submitted on 9 Apr 2005. (Cited on page 54.)

[132] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346. (Cited on pages 54 and 55.)

[133] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249. (Cited on page 54.)

[134] F. Dordei, *LHCb detector and trigger performance in Run II*, EPJ Web Conf. **164** (2017) 01016. (Cited on page 55.)

[135] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055. (Cited on pages 56 and 57.)

[136] LHCb HLT project, J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544. (Cited on pages 56, 57, and 58.)

[137] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in run 2 of the LHC*, Journal of Instrumentation **14** (2019) P04013. (Cited on pages 56, 57, and 80.)

[138] *LHCb trigger plots and diagrams for conferences*, https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbTriggerConferenceDiagramsPlots. (Cited on page 56.)

[139] O. Callot, M. Ferro-Luzzi, and P. Perret, *Using the SPD multiplicity in the Level-0 trigger*, Tech. Rep. LHCb-2003-022, CERN, Geneva, Apr, 2003. (Cited on page 57.)

[140] D. A. Berninghoff, J. Albrecht, and V. Gligorov, *Bremsstrahlung Recovery of Electrons using Multivariate Methods*, LHCb-INT-2016-018. (Cited on pages 61 and 62.)

[141] O. Lupton, L. Anderlini, B. Sciascia, and V. Gligorov, *Calibration samples for particle identification at LHCb in Run 2*, Tech. Rep. LHCb-PUB-2016-005. CERN-LHCb-PUB-2016-005, CERN, Geneva, Mar, 2016. (Cited on page 62.)

[142] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2018) 1, arXiv:1803.00824. (Cited on page 62.)

[143] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021. (Cited on page 63.)

[144] A. Poluektov, *Correction of simulated paritcle identification response in LHCb using kernel density estimation*, LHCb-INT-2017-007. (Cited on page 63.)

[145] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083. (Cited on pages 63 and 64.)

[146] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116, [,186(2003)]. (Cited on pages 65, 151, and 235.)

[147] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861. (Cited on page 80.)

[148] J. Podolanski and R. Armenteros, *III. Analysis of V-events*, Philosophical Magazine **45** (1954) 13. (Cited on pages 84 and 250.)

[149] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis and branching fraction measurement of* $\bar{B}^0_s \to J/\psi K^+ K^-$, Phys. Rev. **D87** (2013) 072004, arXiv:1302.1213. (Cited on pages 87 and 141.)

[150] LHCb collaboration, R. Aaij *et al.*, *Observation of excited Λ_b^0 baryons*, Phys. Rev. Lett. **109** (2012) 172003, arXiv:1205.3452. (Cited on page 102.)

[151] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039. (Cited on page 108.)

[152] R. Quagliani, P. Robbe, and J. Rademacker, *Study of double charm B decays with the LHCb experiment at CERN and track reconstruction for the LHCb upgrade*, Oct, 2017. Presented 06 Oct 2017. (Cited on pages 108 and 109.)

[153] M. Borsato, F. Polci, and M.-H. Schune, *HOP an additional tool for decays involving electrons*, LHCb-INT-2015-037. (Cited on pages 115 and 117.)

[154] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152. (Cited on pages 120 and 210.)

[155] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014. (Cited on page 135.)

[156] W. D. Hulsbergen, *Decay chain fitting with a kalman filter*, physics/05physics/05. (Cited on page 149.)

[157] 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. (Cited on page 165.)

[158] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to J/\psi \bar{K}^{*0}$ branching fraction and angular amplitudes*, Phys. Rev. **D86** (2012) 071102(R), arXiv:1208.0738. (Cited on page 221.)

[159] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters and polarisation fractions in $B_s^0 \to J/\psi \bar{K}^{*0}$ decays*, JHEP **11** (2015) 082, arXiv:1509.00400. (Cited on page 221.)

[160] LHCb collaboration, R. Aaij *et al.*, *Measurement of the mass and production rate of Ξ_b^- baryons*, Phys. Rev. **D99** (2019) 052006, arXiv:1901.07075. (Cited on pages 232 and 241.)

[161] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000. (Cited on page 235.)

[162] Belle, S. B. Yang *et al.*, *First Observation of Doubly Cabibbo-Suppressed Decay of a Charmed Baryon: $\Lambda_c^+ \to pK^+\pi^-$*, Phys. Rev. Lett. **117** (2016), no. 1 011801, arXiv:1512.07366. (Cited on page 240.)

[163] C. Shen, *Recent Belle results on singly charmed baryon and Belle II prospects of baryons*, https://indico.in2p3.fr/event/17415/contributions/63275/attachments/49000/62074/01-shencp.pdf. (Cited on page 240.)

Acknowledgements

> There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery.
>
> Enrico Fermi

Over the past years, there have been a large number of people who have impacted this work in a direct or indirect way. It was a great experience to work in the LAL LHCb group. This is the group where one can find an expert for any question (not only in physics, but also in random general knowledge), and I very much appreciated being surrounded by such great scientists.

First, I would like to thank Marie-Hélène for guiding me into the world of flavour physics throughout the last five years. Back in the time, you managed to convince me that LHCb is a great collaboration to work in. Now, I am confident that this was a great choice, and I am grateful to you for your advises. It is remarkable that with your experience, you always managed to find answers to my weird questions and problems.

Special thanks to Yasmine, it was fantastic to work with you as a supervisor. I certainly appreciated the amount of freedom which I had during these years, which allowed me to learn so many things not always related to the core topic of my thesis, and gain a lot of experience. At the same time, you managed to teach me a plenty of things, and certainly contributed significantly to my formation as a physicist. I should separately acknowledge the amount of time which you dedicated to me, in particular when managing to reply my nasty messages in the middle of the night. If one day they establish a prize for the best PhD advisor, I would definitely nominate you!

I should also mention Carla, who contributed significantly to the analysis during the last year of my PhD, especially when I had to focus on writing this document. Thank you for the useful discussions and for sharing your experience. (As a sidenote, I am impressed by the amount of your time you dedicate to work!)

I would also like to thank my referees, who agreed to review my work and this manuscript. Special thanks to Marcella Bona and Ulrik Egede for reading it carefully.

Your deep questions and accurate remarks not only helped to improve the quality of this document, but also pushed me to rethink deeper some steps of the analysis. Thanks to Gino Isidori, Achille Stocchi and Fabien Cavalier who agreed to serve as examiners at my thesis defense. I would also like to express my gratitude to all the LHCb colleagues who reviewed or commented on the analysis in each step of its development. I am afraid I will not manage to mention everyone, but we received a lot of comments from Simone, Christoph, Eluned, Marco, Paula and Rolf. Each comment was valuable, and allowed to improve the robustness of the work presented in this manuscript. Thanks also to Patrick for his feedback on the analysis note. In addition, I would also like to mention the people who were working on previous lepton universality measurements in LHCb. Your experience was invaluable, and without your work, this thesis would have not been possible. Thanks to Simone for the useful discussions on the analysis procedure. Thanks to Thibaud and Paula for your answers to my questions, and so well-documented analysis note. Finally, thanks to Sébastien and Martín for their interest in rare Λ_b^0 decays, and for providing theoretical insights on the measurements presented in this thesis.

Huge thanks to all the members (current and former) of the LAL LHCb group, for creating a kind working environment, and helping me throughout these years. Thanks to Partick and Jacques for serving as perfect examples of ideal physicists, with impressive knowledge in literally every field. Thanks to João for all his deep questions about physics, which forced me to search for the answers and learn plenty of ideas. Thanks to Frédéric for tolerating me kindly in his office, and for all the help. I should also mention Victor who shared the office with me for two years. You helped me a lot with understanding how the things work (when I was just a fresh PhD student), and were always there to answer my questions or to share some thoughts. Thanks also to all the students who shared the office with me during the last year – in particular, Valeriia. A special mention to Sergey for his advises and help throughout the years. Thanks to Andrii for sharing with me the experience of being an Ukrainian PhD student at LAL (and a master student before that), and fighting together with all the paperwork. Special thanks for those long discussions we had from time to time, from which I learned a lot about physics, and could share some ideas.

Thanks to Fabrice and Elisabeth for the nice discussions, it was a pleasure to be there to teach you something, or to learn something from you. I hope this manuscript will be useful for you when you start preparing yours. I also hope your PhD goes well and you will have bright results to show! Elisabeth, I hope you will travel to many awesome destinations (take your time, it turns quite difficult to find the time for travel in the last year of PhD). Thanks to Jibo and Guy, I have learned a lot about flavour physics from you, your experience and stories during lunchtime. I should also mention Davide who arrived during the period when I was extremely focused on finalising this work, so we probably did not have enough time to discuss. Thanks also to the LLR colleagues (Émilie, Frédéric, Vladik and Felipe) for the discussions during our occasional meetings.

Thanks to the people who were a part of the LAL LHCb group, but, unfortunately, left before I finished. Special mention to Michael for the generous discussions and insights. You are a great physicist and you deserve a bright future. Thanks to Francesco who was there with curious facts and stories, and to Yanxi who was always doing some cool things

which impressed me (remember that Ξ_{cc} peak during your talk at the group meeting?).

I should mention all the people at CERN and other institutions, who helped me at any point. Thanks to all people who were teaching me things, or helping me to teach things to other people. Thanks to the starterkit and startertalk friends (I should notably mention Giulio), I learned so much from all of you, and got so much help. Thanks to Martino and Alex who were helping me with the calibration of electron identification. Thanks to the past LAL PhD students (especially Martino, Renato and Victor): your thesis manuscripts served as a source of knowledge and inspiration. Thanks to my family and friends who were always there to support me and to help. Sorry for complaining too much when something did not work. Thanks to anyone who I forgot to mention.

Special thanks to the Swiss Alps for the incredible nature and impressive hiking routes. These getaways were crucial to keep me in shape during the work on the thesis.

Finally, thanks to you (whoever you are) for reading this!

université PARIS-SACLAY | ÉCOLE DOCTORALE PHENIICS

Titre: Études de désintegrations rares de baryons beaux et test de l'universalité du couplage aux leptons avec LHCb

Mots clés: Physique de la saveur, LHCb, physique du B, universalité leptonique, baryons, désintégrations rares

Résumé: Les courants neutres changeant la saveur $b \to s\ell^+\ell^-$ ne sont autorisés dans le Modèle Standard (MS) qu'au niveau des boucles. Par conséquent, ils sont sensibles aux éventuels effets de la Nouvelle Physique (NP) au-delà du MS. Dans le MS, les transitions $b \to se^+e^-$ et $b \to s\mu^+\mu^-$ ont la même probabilité, cette propriété est appelée l'universalité du couplage aux leptons (UL). Tester l'UL dans les désintégrations de hadrons beaux a récemment été un domaine prometteur pour les recherches de NP. Des tensions par rapport aux prédictions du MS, de l'ordre de 2,5 déviations standards, ont été observées dans les désintégrations rares $B \to K^{(*)}\ell^+\ell^-$. À ce jour, les tests de l'UL ont été effectués uniquement sur des désintégrations de mesons beaux.
Cette thèse présente un premier test de la UL dans la désintégration rare du baryon $\Lambda_b^0 \to pK\ell^+\ell^-$, en utilisant les données collectées par l'expérience LHCb au CERN en 2011, 2012 et 2016. La sensibilité statistique attendue est d'environ 15%, tandis que la valeur centrale reste aveuglée. La première mesure du rapport de branchement de la désintégration $\Lambda_b^0 \to pK\mu^+\mu^-$ est effectuée avec une précision statistique de 5,3%. En outre, cette thèse présente la première observation de la désintégration $\Lambda_b^0 \to pKe^+e^-$, ainsi que la mesure de son rapport de branchement. Enfin, la désintégration $\Xi_b^0 \to pKJ/\psi$ est observée et étudiée pour la première fois.

Title: Study of rare b-baryon decays and test of lepton universality at LHCb

Keywords: Flavor Physics, LHCb, B physics, Lepton Universality, Baryons, Rare Decay

Abstract: Flavor-changing neutral-current $b \to s\ell^+\ell^-$ transitions are forbidden at tree level in the Standard Model (SM), and can only occur at loop level. Therefore, they are sensitive to possible New Physics (NP) effects beyond the SM. In the SM, the transitions $b \to se^+e^-$ and $b \to s\mu^+\mu^-$ have the same probability to happen, this property is called lepton universality (LU). Probing LU in b-hadron decays has been recently a promising area for NP searches. Tensions with respect to the SM predictions at the level of about 2.5 standard deviations have been observed in rare $B \to K^{(*)}\ell^+\ell^-$ decays. To date, tests of LU have been performed only in decays of b-mesons.
This thesis presents a first test of the LU in rare b-baryon decays $\Lambda_b^0 \to pK\ell^+\ell^-$, using the data collected by the LHCb experiment at CERN, during 2011, 2012 and 2016 data-taking periods. The expected statistical sensitivity is about 15%, while the central value remains blind at the moment. The first measurement of the branching fraction of the decay $\Lambda_b^0 \to pK\mu^+\mu^-$ is performed with the statistical precision of 5.3%. In addition, this thesis presents the first observation of the $\Lambda_b^0 \to pKe^+e^-$ decay, and the measurement of its branching fraction. Finally, a suppressed decay $\Xi_b^0 \to pKJ/\psi$ is observed and studied for the first time.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery