Searching for new physics in $b \to s\ell^+\ell^-$ transitions at the LHCb experiment

L. Pescatore

Thesis submitted for the degree of
Doctor of Philosophy

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

February 8, 2016

ABSTRACT

Flavour Changing Neutral Currents are transitions between different quarks with the same charge such as $b \to s$ processes. These are forbidden at tree level in the Standard Model but can happen through loop diagrams, which causes the branching ratio of this type of decays to be small, typically $\sim 10^{-6}$ or less. Particles beyond the SM can contribute in the loops enhancing the branching fractions of these decays, which are therefore very sensitive new physics. In this work two analysis of semileptonic $b \to s\ell^+\ell^-$ decays are presented. First $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays are analysed to measure their branching fraction as a function of the dimuon invariant mass, q^2. Furthermore, an angular analysis of these decays is performed for the first time. Secondly, $B^0 \to K^{*0}\ell^+\ell^-$ decays are analysed measuring the ratio between the muon, $B^0 \to K^{*0}\mu^+\mu^-$, and electron, $B^0 \to K^{*0}e^+e^-$, channels, which is interesting as it is largely free from uncertainties due to the knowledge of the hadronic matrix elements. This thesis is organised in the following way. Chapter 1 introduces the Standard Model and the concept of flavour and explains how rare decays can help us in the quest for physics beyond the SM. Chapter 2 describes the LHCb detector, which was used to collect the data analysed in this thesis. This chapter also includes studies performed to validate the hadronic physics in LHCb simulation software. Chapter 3 deals with the measurement of the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, while Chapter 4 describes an angular analysis of these decays. Finally, Chapter 5 reports the measurement of the $R_{K^{*0}}$ ratio between the decay rates of the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays.

DECLARATION OF AUTHORS CONTRIBUTION

I am one of the main authors of the two analysis reported in Chapters 3, 4 and 5. For the analysis of the differential branching ratio of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay I collaborated with Michal Kreps, who took care of implementing the decay model to re-weight the simulation and provided a few of the simulated samples. Furthermore, I want to thank him for the advice given throughout. The work in this part was also published and can be found at Ref. [39]. For the $R_{K^{*0}}$ analysis, described in Ch. 5, I actively participated in most stages of the analysis collaborating with Simone Bifani. In particular I took care of the production of various simulated samples, participated in the definition of the selection, in the yields extraction and I provided a fit and data-reduction framework. Finally, as a service work for the LHCb experiment, I developed the tools used to perform the studies described in Sections 2.11–2.13. Furthermore, I contributed to the LHCb experiment in the role of "Monte Carlo liaison" for two years. This is a connection role between the physics analysis groups and the simulation team providing simulated samples vital for most analysis. Finally, I have taken several shifts in the control room during the Run II data taking in 2015, checking the smooth running of the detector.

ACKNOWLEDGEMENTS

I thank everybody, evvvvvvvverybody! First of all I would like to thank Nigel, who always supported me in these years and granted be many good opportunities. I think I could not hope for a better supervisor. A big thank you also to Simone, with whom I collaborated for the $R_{K^{*0}}$ analysis and from who taught me a lot. Thanks also to the people of the Birmingham LHCb group: Cristina, Jimmy and also to Michal, who adopted us, Birmingham students, for a while. A special 'thank you' goes to Pete, who shared with me this three years experience. I think it would have been a very different and less interesting experience without him. A 'thank you' also goes to the members of the LHCb collaboration and in particular of the Rare Decays Working Group.; in particular to the working group conveners Gaia, Tom and Marco and to Gloria, who patiently guided me through the depts of the LHCb software. Finally, I'm grateful to the Vincenzo and the LHCb Bologna group, who kindly hosted me for a few months and in particular to Umberto for all the wisdom he shared. I want also to thank the LTA folks, who where with me during the long period I spent at CERN and especially Mark and Lewis, adventure companions. And speaking about CERN people a great 'thank you' to Lorenzo, because when it's 1pm I always feel that I should be in front of the trays. Going now to who is always waiting for me in Italy when I go back, a big 'thank you' to my family for all their support and all the italian food they brought me while I was living abroad. Thank you my dad Orazio, my mum Paola and my sisters Giulia and Silvia. A big 'thank you' also to my friends Ivan, Enrico, Martina, Federico, Valentina, Letizia and all the others. And finally, last but not least, a giant thank you to Lucia, who is the engine of my life and to whom this thesis is dedicated.

A Lucia,
perché quando tutto perde di senso
tu sei il mio piccolo mondo felice.

Nec per se quemquam tempus sentire fatendumst
semotum ab rerum motu placidaque quiete.
(Lucrezio, De rerum natura, vv. 462-463) Nel niente c'è una via che conduce
lontano dalla polvere del mondo.
(F. Bertossa)

Contents

CHAPTER 1

Introduction

The Standard Model of particle physics (SM) is a Quantum Field Theory (QFT) describing strong and electroweak (EW) interactions. It was formulated in his current form in the mid-70s and has been an extremely successful and predictive theory since then. Almost all known phenomena from 1 eV up to several hundred GeV are well described by the SM and experiments at the Large Hadron Collider (LHC) are now probing the SM up to and above the TeV scale. As an example of the level of accuracy of the SM, Tab. 1.1 reports the predicted and measured values of the widths of the Z and W bosons [1]. Finally, in 2013 the Higgs boson was observed, one of the fundamental building blocks of the theory, which gives a solid basis to it by introducing a mechanism that produces particles' masses [2]. Despite the success of the SM, experimentally well established effects, like neutrino oscillations and the presence of dark matter, are outside the reach of this theory. Furthermore, the model does not include the description of gravity, which can be neglected at the EW energy scale. Therefore this motivates the search for New Physics (NP).

Table 1.1: Predicted and measured values of the decay widths of the Z and W bosons [1].

Quantity	Predicted	Measured
Γ_Z	2.4960 ± 0.0002 GeV	2.4952 ± 0.0023 GeV
Γ_W	2.0915 ± 0.0005 GeV	2.085 ± 0.042 GeV

The SM is based on the symmetry groups of strong, $SU(3)_C$, and electroweak, $SU(2)_W \times U(1)_Y$, interactions. The subscripts C, W and Y stand for colour charge, weak isospin and hyper-charge respectively. The Lagrangian describing the SM results from the application of the principle of invariance under the unitary group given by the product $SU(3)_C \otimes SU(2)_W \otimes U(1)_Y$, which reflects conservation laws such as the conservation of electric and strong charge. The model has then 26 free parameters, which are experimentally measured.

Particles included in the SM can be grouped under a few categories depending on their properties and ability to interact with each other. The first distinction is between fermions, half-integer spin particles, and bosons, integer spin particles. Fermions constitute the basic building blocks of matter, while bosons are the mediators of the interactions. Since the concept of bosonic mediators of interactions arises because of gauge symmetry [3], they are called "gauge bosons". The list of the

Table 1.2: Fundamental forces of nature together with their gauge bosons, relative strengths and range. Gravity is not included in the SM and the graviton is hypothetical at the current time.

Interaction	**Mediator**	**Strength**	**Range (m)**	**Mediator mass**
Strong	g	1	∞	0
EM	γ	10^{-3}	∞	0
Weak	Z, $W^{\pm}$	10^{-16}	10^{-18}	$W^{\pm} = 80.399$ GeV$/c^2$ $Z_0 = 91.188$ GeV$/c^2$
Gravity	g^0 (graviton?)	10^{-41}	∞	0

known interactions with their force carrier and properties is reported in Tab. 1.2. The matter of which we are made of is mainly composed of electrons and protons, which have spin 1/2; protons are in turn composed of u and d quarks, which again have spin 1/2. Among fermions one can then consider two smaller groups: quarks and leptons. Quarks carry colour charge and therefore can interact through the, so called, strong interaction, while leptons, which do not carry colour charge, are insensitive to it. For each particle a corresponding anti-particle exists with opposite quantum numbers. Finally, fermions are divided into three families having similar properties but different masses. This last structure embedded in the SM is also called "flavour structure" and it will be the main tool used in this thesis; a more detailed description of it is given in the next sections. A schematic view of the fundamental particles in the SM is shown in Fig. 1.1. Due to the asymptotic freedom

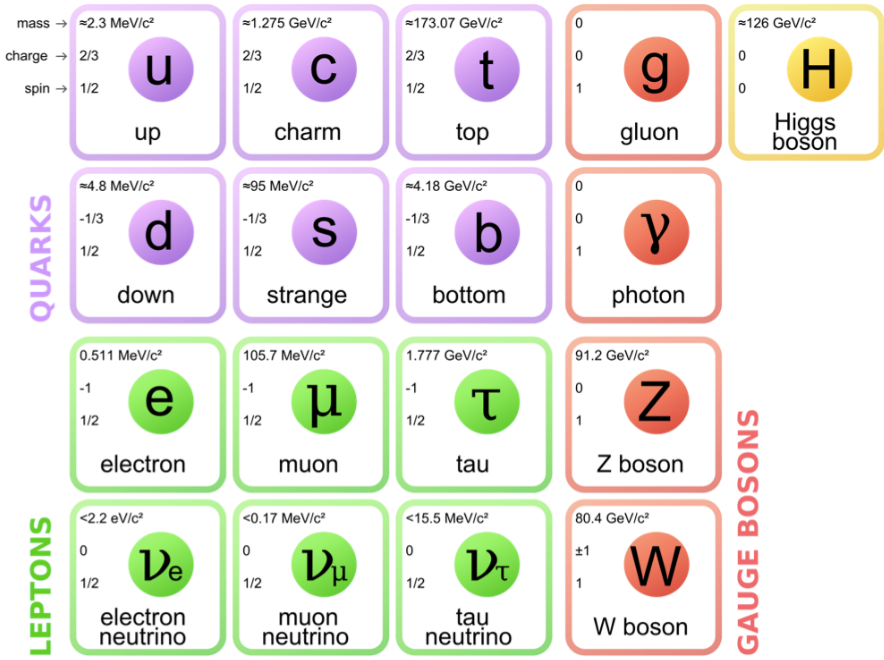

Figure 1.1: A scheme of the fundamental particles in the SM with their properties.

of the strong interaction quarks cannot be observed alone but are always combined with other quarks to form color singlets. Non-fundamental particles composed by quarks are called hadrons and can be divided in mesons, where the color singlet is achieved by the combination of a quark and its antiquark ($q\,\overline{q}$), and baryons formed by three quarks ($q\;q\;q$) of different colours. Recently, in 2014 and 2015 evidence for new states, formed by four and five quarks, was found [4, 5].

1.1 The electroweak interaction

The electromagnetic interaction is responsible for binding electrons and nuclei together in atoms and its mediator is the photon. The weak interaction is responsible for the β decay of nuclei and is mediated by the emission or absorption of $W^{\pm}$ and Z bosons. Unlike the electromagnetic force, that affects only charged particles, all known fermions interact through the weak interaction. The weak interaction is also the only one that violates the parity symmetry, which states that interactions are invariant under a reflection of all coordinates. This symmetry breaking arises from the fact that only left-handed fermions interact through the weak interaction as discovered by Wu in 1957 [6]. Similarly, the weak interaction is the only one that also breaks the CP symmetry, which combines parity transformations and charge conjugation. This is particularly interesting because all interactions are invariant under the CPT transformation, which combines the CP transformation and time reversal, hence, breaking CP the weak interaction must also be not invariant under time reversal. In 1968 Salam, Glashow and Weinberg unified the weak and electromagnetic forces in a single theory, where the coupling constants of the electromagnetic, e, and weak, g, interactions are linked by the weak mixing angle, θ_W by the relation $g \sin\theta_W = e$ [1]. The electroweak symmetry is spontaneously broken by the Higgs field [7] and this causes the $W^{\pm}$ and Z bosons to become massive (see Tab. 1.2) and consequently the weak force has a very short range. In fact using Heisenberg's Principle ($\Delta E \Delta t > \hbar$) together with Einstein's formula $\Delta E = mc^2$, which relates mass and energy, and knowing that the maximum space that a particle can cover in a time Δt is $r = c\Delta t$, qualitatively $r \sim \hbar/mc$. In this picture the carriers of the weak force can travel $r \sim 2 \cdot 10^{-3}$ fm. The photon must instead be massless in the theory, which accounts for the long range of the electromagnetic force. The EW interactions are divided into charged currents (CC) and neutral currents (NC). In the first group, quarks and leptons interact with the $W^{\pm}$ bosons, producing decays such as $\mu^+(\mu^-) \to e^+\nu_e\overline{\nu}_\mu(e^-\overline{\nu}_e\nu_\mu)$ and $n \to pe^-\overline{\nu}_e(\overline{p}e^+\nu_e)$. The study of these processes confirmed that only the left-handed (right-handed) component of fermions (anti-fermions) takes part in weak processes. The CC interactions have a peculiarity: they are the only interactions in the SM that violate flavour conservation at tree level (see next section), while any other interaction not conserving flavour has to happen through loops. The second group of EW interactions, NC, corresponds to diagrams mediated by a photon or a Z boson interacting with a fermion and its anti-fermion.

1.2 Flavour and the CKM matrix

"Flavour" in particle physics refers to the quark/lepton composition of a particle. The introduction of flavour quantum numbers was motivated in order to explain why some decays, although kinematically allowed, had never been observed. To all leptons is assigned a quantum number $L_\ell = 1$ (where $\ell = e, \mu, \tau$), which in the SM is conserved by all interactions. This conservation is experimentally well established; for example decays like $\mu^- \to e^-\gamma$ have never been observed. In the hadronic sector particles carry flavour numbers described as follow:

- *Isospin*: $I_3 = 1/2$ for the up quark and $I_3 = -1/2$ for the down quark;
- *Strangeness*: $S = -(n_s - \bar{n}_s)$, where n_s and $\bar{n}_s$ are the numbers of strange and anti-strange quarks respectively;
- *charmness, bottomness, topness*: in analogy to strangeness they are respectively defined as $C = -(n_c - \bar{n}_c)$, $B = -(n_b - \bar{n}_b)$, $T = -(n_t - \bar{n}_t)$.

As mentioned before, in the SM the only interaction violating flavour conservation is the weak interaction when mediated by $W^\pm$ bosons.

Measuring branching fractions of weak decays like $\pi \to \mu\nu_\mu$ and $K \to \mu\nu_\mu$, corresponding respectively to $ud \to \mu\nu_\mu$ and $us \to \mu\nu_\mu$ processes, suggested the existence of more than one coupling constant for different quarks. Nicola Cabibbo [1], in order to preserve the universality of weak interactions, suggested that the differences could arise from the fact that the doublets participating in the weak interactions are an admixture of the mass eigenstates. He therefore introduced the Cabibbo angle, θ_c, considering that eigenstates participating to the weak interaction are rotated with respect of the flavour eigenstates.

$$\begin{pmatrix} d_W \\ s_W \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} \cos\theta_c \cdot d + \sin\theta_c \cdot s \\ \cos\theta_c \cdot s - \sin\theta_c \cdot d \end{pmatrix} \tag{1.1}$$

Considering a 6 quark system one angle is not enough to describe a rotation but the mixing can be generalised using a 3×3 unitary matrix, which is called CKM matrix, from the names of Cabibbo, Kobayashi and Maskawa. The unitarity of the matrix is required to preserve the universality of the weak interaction. Theoretically, a $N \times N$ complex matrix depends on $2 \cdot N^2$ real parameters. Requiring unitarity ($AA^\dagger = A(A^*)^T = I$), the number of independent parameters left is

$$(N-1)^2 = \underbrace{\frac{1}{2}N(N-1)}_{\text{Number of mixing angles}} + \underbrace{\frac{1}{2}(N-1)(N-2)}_{\text{Number of complex phases}} . \tag{1.2}$$

Therefore a 3×3 matrix depends then on 4 real parameters, which can be divided in 3 real constants and one imaginary phase. The imaginary phase generates the CP-violation which was observed in weak interactions. Figure 1.2 displays examples of CC processes together with the CKM elements associated with their vertices. Equation 1.3 reports the most recent measured values of its elements [1] together

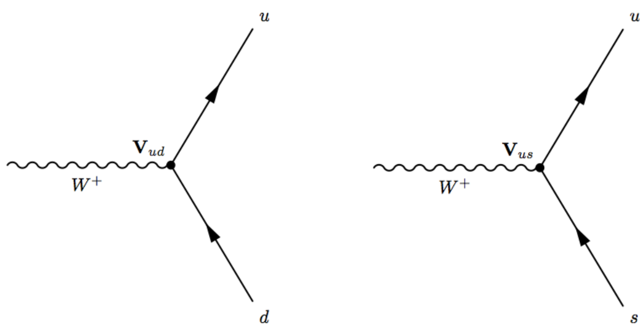

Figure 1.2: Feynman diagrams with CKM weights on weak interaction vertices

with the Wolfenstein parametrisation which highlights the hierarchical structure of the matrix. In fact elements on the diagonal, corresponding to transitions between quarks of the same generation, are approximately 1 and become smaller and smaller going farther from the diagonal. In the formula ρ, A, and λ are the real constants and η the imaginary phase and Eq. 1.4 shows their relations with the 3 mixing angles; terms further from the diagonal are proportional to higher powers of λ.

$$V_{CKM} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.0014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.00412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} = \\ = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \tag{1.3}$$

$$\begin{aligned} \lambda &= \sin(\theta_{12}) = \sin(\theta_c) \\ A\lambda^2 &= \sin(\theta_{23}) \\ A\lambda^3(\rho - i\eta) &= \sin(\theta_{13})e^{i\delta} \end{aligned} \tag{1.4}$$

The unitarity of the CKM matrix imposes constraints to its elements of the form:

$$\sum_i |V_{ik}|^2 = 1 \text{ and } \sum_k V_{ik}V^*_{jk} = 0. \tag{1.5}$$

These correspond to constraints to three complex numbers, which can be viewed as the sides of triangles in the (ρ, η) plane; these are called "unitarity triangles". The most commonly used unitarity triangle arises from $V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0$. Figure 1.3 shows a representation of such triangle together with a plot summarising the most up to date experimental constraints to its parameters [8]. Due to these unitarity constraints flavour-changing neutral currents are forbidden at tree level in the SM.

The precise measurement of the parameters of the CKM matrix is a powerful stability test of the SM and sets a solid base for new physics searches in the flavour sector. One of the main goals of the LHCb experiment is to precisely measure the angle γ, which is currently the least constrained by measurements.

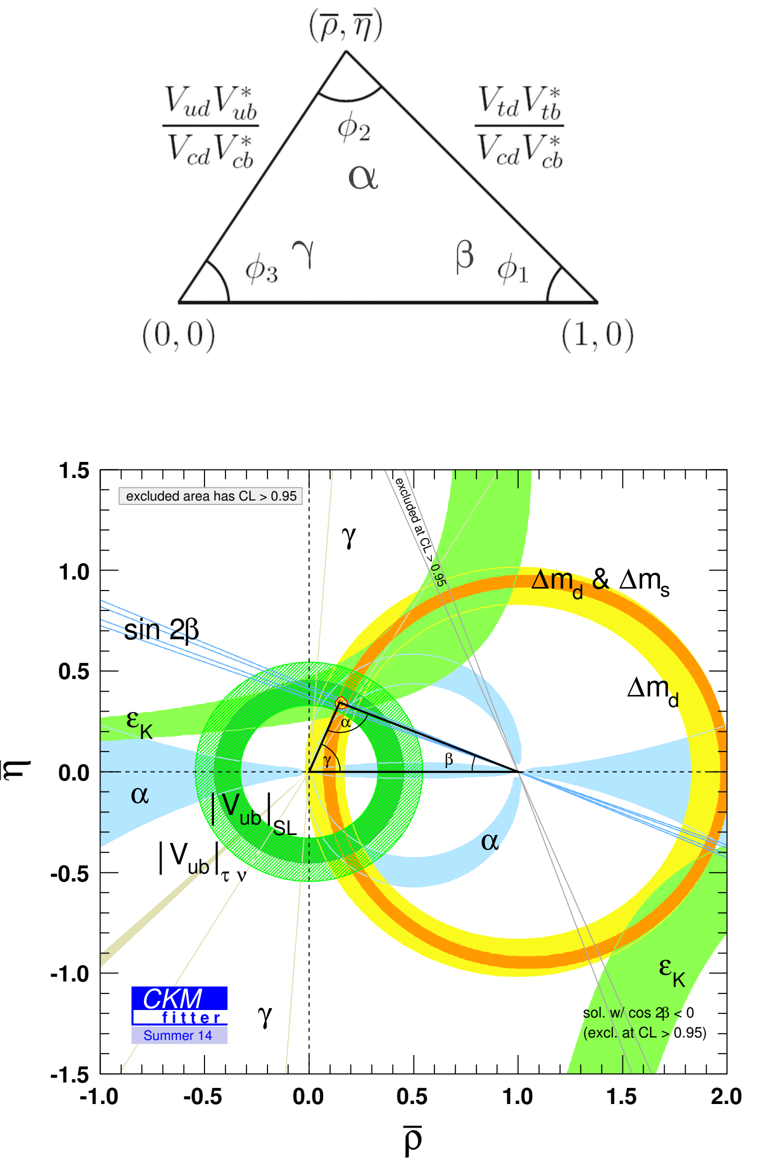

Figure 1.3: (top) A representation of the unitarity triangle and its parameters. (bottom) A summary of the most up to date measurements of the unitarity triangle parameters [8].

1.3 The puzzles of the SM

Despite the confirmation of many predictions of the SM, the theory has several limitations and is unable to account for some well established experimental facts:

- *Dark matter*: experimental evidence tells us that the content of visible matter in the universe is not enough to account for the observed rotation of galaxies [9]. The most natural way to solve the problem is the hypothesis of a form of matter that interacts with the gravitational field but not with the other SM interactions.

- *Matter-antimatter asymmetry*: a large asymmetry is observed between the quantity of matter and antimatter in the universe, $O(10^{-9})$. Assuming that both were equally created in the initial state of the universe, a condition such as the violation of the CP symmetry is necessary to account for the observed imbalance. However, the magnitude of CP violation predicted by the SM, $O(10^{-20})$, is not enough to explain the observed imbalance [10].

- *Gravity*: even though the gravitational force was the first to be discovered this is not included in the SM. When introducing gravity in the framework of QFT the theory diverges. On the other hand gravity becomes irrelevant for small masses as those of particles and can be neglected in good approximation at the EW energy scale. Many attempts were made but there is not yet a consistent procedure to introduce gravity in the SM.

- *Neutrino oscillation*: measurements regarding solar and atmospheric neutrinos as wells as neutrinos from nuclear reactors established that neutrinos can change flavour while propagating in space. This is not predicted in the SM, in fact in the SM neutrinos are massless, while an oscillation requires a non zero mass [11, 12, 13, 14].

- *The hierarchy problem*: the mass of a scalar (spin 0) particle, such as the Higgs boson, suffers from quantum corrections due to the physics at high energy scales. As new physics can appear anywhere up to the Planck scale, $\sim 10^{19}$ GeV, at which gravity cannot be neglected any more, these corrections can be very large and it would require a high level of fine-tuning for them to cancel out and give such a small value as the one measured for the Higgs Mass, ~ 126 GeV/c^2 [15, 16].

In conclusion, even though the SM has been very successful in describing the properties of the observed particles and their interactions so far, because of its many puzzles, it is believed only to be part of a more general theory or only to be valid up to a certain energy scale.

1.3.1 The flavour problem

Flavour Changing Charged Currents (FCCC) that are mediated by the $W^{\pm}$ bosons are the only sources of flavour changing interaction in the SM and, in particular, of generation changing interactions, where a quark or a lepton of a family transforms into one of an other family. Another class of processes is the Flavour Changing Neutral Currents (FCNCs), e.g. transitions from a b quark with a charge of -1/3 to a s or d quark with the same charge. Examples of FCNC transitions in the quark and lepton sector are shown in Fig. 1.4. FCNCs are experimentally observed to

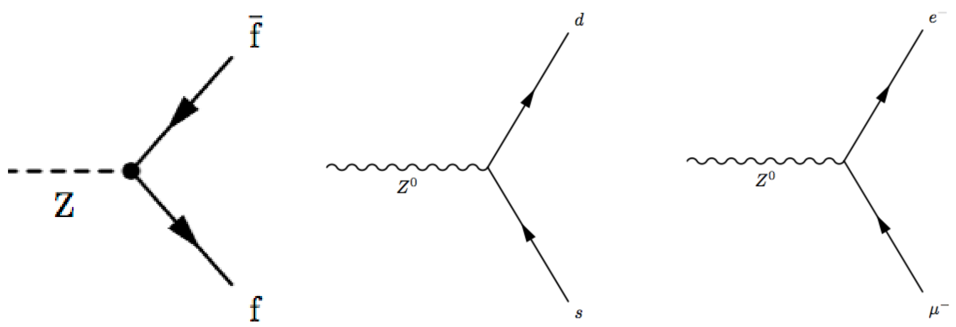

Figure 1.4: Feynman diagrams of (left) neutral current allowed in the SM, where$f$ represents any fermion, and (center-right) FCNCs processes forbidden in the SM.

be highly suppressed which derives from the unitarity of the CKM matrix, however there is no fundamental reason why there cannot be FCNCs at tree level. In fact the CKM matrix could be part of a larger matrix involving for example quark-lepton terms. This would introduce new sources of FCNCs but also allow for natural explanations of the equality of the proton and electron charges. On the other hand the observation of neutrino oscillation proves that flavour is not always conserved suggesting flavour structures beyond the SM. Furthermore, the values of the terms of the CKM matrix and the PMNS matrix [17, 18], which is the mixing-matrix equivalent to the CKM in the lepton sector, are not explained in the SM but have to be measured experimentally. These open problems motivate searches for flavour symmetries and deeper motivations for flavour conservation.

1.4 Beyond the Standard Model

From the previous sections it is evident that, despite the great success of the SM, there is a need to explore theories Beyond the SM (BSM). Among the most promising approaches there are those involving Super-Symmetry (SUSY) [19] and extra-dimensions [20]. In SUSY new degrees of freedom are introduced to suppress the diverging terms of the Higgs mass. This theory assumes that for each fermion there is a corresponding boson and, since bosons and fermions contribute with opposite sign to the mass term, these would naturally cancel out. Supersymmetry also provides a candidate for dark matter. In fact the lightest Super-Simmetric particle, the neutralino, which in R-parity [21] conserving variants of the theory must be stable, is a weakly interacting potentially heavy particle. The idea to introduce extra-dimensions was triggered by the fact that, normally, gravity is not relevant in particle physics but it would be natural if all forces had similar strength. By adding extra dimensions to the normal three spatial dimensions, one can restore the strength of gravity, as this could be dispersed by the wider space available. In all these approaches constraints to masses and couplings must be imposed to maintain compatibility with the SM at the electroweak scale and the existing experimental observations.

1.4.1 Flavour and BSM theories

Most BSM theories predict processes violating flavour conservation. Therefore, the observation or non-observation of these processes can give important information about new physics. BSM theories can be classified according to the amount of flavour violation they introduce. The first class of models to consider is that with Minimal Flavour Violation (MFV). These are models in which the only sources of flavour changing transitions are governed by the CKM matrix and the CKM phase is the only source of CP violation. This definition is driven by the fact that usually a solution of the hierarchy problem is expected at the TeV scale, while the very small amount of flavour violation observed in measurements seems to indicate that the SM would remain valid up to much higher energy scales. It is therefore assumed that new physics must respect flavour symmetry principles, which also makes these types of models naturally compatible with the SM. Examples of such models include the MSSM with minimal flavour violation and the SM with one extra-dimension. Reviews of MFV models are presented in Refs. [22, 23]. A powerful test of MFV is provided by the study of ratios between $b \to d$ and $b \to s$ transitions, because their hamiltonians share the same structure. One particularly important example is the ratio of B^0 and B_s^0 dimuon decay rates [24], as this is a purely leptonic decay free from hadronic uncertainties. In the SM such ratios are approximately equal to $|V_{td}/V_{td}| \sim 1/25$, only modified by phase space and hadronic matrix elements, while they can take very different values in non-MFV models.

In the quest for new physics an important role is also played by simplified models as an intermediate model building step. Instead of constructing thories valid up to the GUT scale one can consider simplified models, where the SM is extended by the addition of a new sector with a limited number of parameters. Such models are easier to constrain but can nevertheless point in the right direction to build more complete theories. The choice of the new sector to add can be driven by the need to explain existing tensions between measurements and SM predictions or by theoretical prejudice. Two models especially relevant when studying rare decays, which are the main topic of this thesis, are Z'-penguins and leptoquarks. A Z'-penguin is a FCNC process involving a neutral field arising from an extra U(1) gauge symmetry, for example $\mathrm{U}(1)_{B-L}$, where B and L are the baryon and lepton numbers. As for the SM penguins, the Z' field contributes in loops causing modifications of the effective couplings with respect to the SM. A survey of Z' models can be found in Ref. [25]. Leptoquarks are bosonic particles that carry both quark and lepton flavour quantum numbers, which for simplicity are commonly assumed to be scalar particles. A tree level exchange of a leptoquark induces processes such as $b \to (s,d)\ell^+\ell^-$, and therefore can result in an enhancement of their decay rates with respect to the SM [26]. Leptoquarks would also provide a natural explanation for non-universal couplings to leptons.

1.5 Rare decays: a tool to search for new physics

In the Standard Model FCNC processes are forbidden at tree level but can occur through loop diagrams such as penguin or W box diagrams (see Fig. 1.5). The branching fractions of decays going through these processes are small, typically $\sim 10^{-6}$ or lower, and therefore they are called “rare decays”. Additional contributions to the virtual loops are not necessarily suppressed with respect to the SM component and this makes these decays very sensitive to new physics. This approach to new physics searches is interesting as new particles could be at high mass scales that are not accessible via direct production at colliders but their effect could be observed in loops. Radiative and penguin decays are particularly interesting because they are theoretically well understood, which allows precise comparisons with measurements. Furthermore, they provide a large quantity of observables that can be affected by new physics, not only decay rates, but also CP asymmetries and angular observables such as forward-backward asymmetries. The joint analysis of different observables can help to build a consistent picture and rule out specific models.

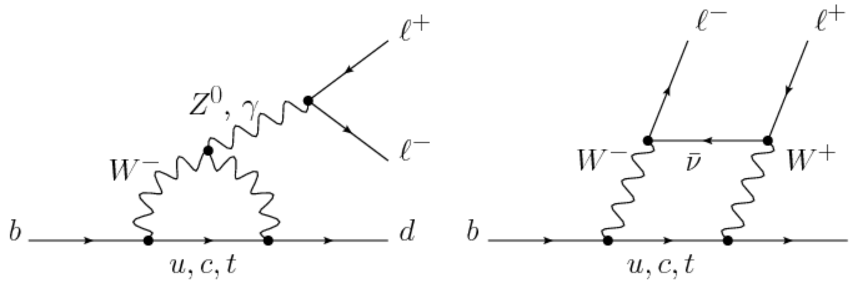

Figure 1.5: Loop Feynmann diagrams allowing $b \to d$ FCNC processes: penguin diagram (left) and W box (right).

1.5.1 Theoretical framework: the effective Hamiltonian

Rare decays of b hadrons are governed by an interplay between weak and strong interactions. The large masses of the W and Z bosons and top quark compared to that of the b quark allow the construction of an effective theory that divides the problem of calculating weak decay amplitudes into two parts: "short-distance" and "long-distance" effects separated at an energy scale μ. The first part, dealing with short distance physics, handles perturbative contributions due to energy scales above the b mass. The second part typically deals with non-perturbative contributions. A classic example of an effective theory is the Fermi theory of weak interactions which describes the β decay in terms of a four-fermion interaction, where the short distance physics is hidden into a point-like vertex as illustrated in Fig. 1.6.

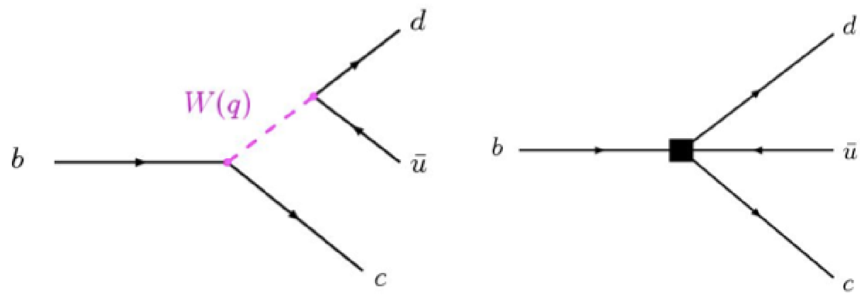

Figure 1.6: Example of a Fermi theory in which the full theory is divided between a short distance contribution, hidden in the vertex, and a long distance contribution.

The effective hamiltonian [27] relevant to $b \to s/d\gamma$ and $b \to s/d\ell^+\ell^-$ transitions can be written as:

$$\mathcal{H}_{eff} = \frac{-4G_F}{\sqrt{2}} \left[\lambda_q^t \sum C_i(\mu, M)\mathcal{O}_i(\mu) + \lambda_q^u \sum C_i(\mu, M)(\mathcal{O}_i(\mu) - \mathcal{O}_i^u(\mu)) \right], \quad (1.6)$$

where G_F denotes the Fermi coupling constant and the λ constants are the CKM factors, $\lambda_q^t = V_{tb}V_{tq}^*$ and $\lambda_q^u = V_{ub}V_{uq}^*$. In $b \to s$ quark transitions, which are the main topic of this thesis, the doubly Cabibbo-suppressed contributions can be neglected as $\lambda_s^u << \lambda_s^t$. To obtain this formula the Operator Product Expansion (OPE) [28] method is used, which implements a summation over all contributing operators weighted by corresponding constants called Wilson coefficients. In this Hamiltonian the long-distance contributions are described by the operators, $\mathcal{O}_i$, while the short-distance physics is encoded in the Wilson Coefficients, C_i. Operators and coefficients are evaluated at the renormalisation scale μ. Any particle that contributes to the decay and has a mass greater than the scale μ will affect the value of at least one of the Wilson coefficients, including SM particles as the top quark.

In order to describe SM processes the effective theory must be matched with the SM by requiring the equality between each term in effective theory and the full theoretical calculation at a matching scale, typically the EW scale (μ_W). Then, using the scale independence of the effective Hamiltonian, one can derive a renormalisation group equation for the Wilson Coefficients [29]. Taking into account only SM contributions and using $\mu_W = m_b$, the Wilson Coefficients have values:

$$C_7^{SM} = -0.3, \quad C_9^{SM} = 4.2, \quad C_{10}^{SM} = -4.2 \quad (1.7)$$

and new physics contributions appear in the Wilson Coefficients in the form of additive factors:

$$C_i = C_i^{NP} + C_i^{SM}. \quad (1.8)$$

The amplitudes of exclusive hadronic decays can be calculated as the expectation values of the effective Hamiltonian. Given an initial state I and a final state F

(e.g. $I = B$ and $F = K^{*0}\mu^+\mu^-$) the decay amplitude can be calculated as

$$\begin{aligned} A(I \to F) &= \langle I|\mathcal{H}_{eff}|F\rangle = \\ &= \frac{G_F}{\sqrt{2}} \sum V^i_{CKM} \underbrace{C_i(\mu)}_{\substack{\text{Perturbative} \\ \text{Includes new physics}}} \cdot \underbrace{\langle I|\mathcal{O}_i(\mu)|F\rangle}_{\substack{\text{Non-preturbative} \\ \text{Known physics}}} , \end{aligned} \tag{1.9}$$

where $\langle I|\mathcal{O}_i(\mu)|F\rangle$ are the hadronic matrix elements also called "form factors". These can be evaluated using non perturbative methods such as lattice calculations. However, due to the limitations of these methods, they represent the dominant source of uncertainty in theoretical calculations.

1.5.2 Operators

Separating the left- and right-handed components the effective Hamiltonian is

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb}V^*_{ts} \frac{\alpha_e}{4\pi} \sum_{i=1}^{10} \left[C_i\mathcal{O}_i + C'_i\mathcal{O}'_i\right]. \tag{1.10}$$

A complete basis is given by a set of 10 operators, where $\mathcal{O}_{1,2}$ are the tree level W operators; $\mathcal{O}_{3-6,8}$ are penguin diagrams mediated by gluons; and $\mathcal{O}_{7,9,10}$, which are the operators that are relevant for radiative and leptonic penguin processes are defined as [24]:

$$\begin{aligned} \mathcal{O}_7 &= \tfrac{m_b}{e}(\bar{s}\sigma^{\mu\nu}P_R b)F_{\mu\nu}, & \mathcal{O}'_7 &= \tfrac{m_b}{e}(\bar{s}\sigma^{\mu\nu}P_L b)F_{\mu\nu}, \\ \mathcal{O}_9 &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu\ell), & \mathcal{O}'_9 &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu\ell), \\ \mathcal{O}_{10} &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu\gamma_5\ell), & \mathcal{O}'_{10} &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu\gamma_5\ell), \end{aligned} \tag{1.11}$$

where $P_{L/R} = (1 \mp \gamma_5)/2$ denote the left- and right-handed chiral projections, T^a are the QCD generators and $F_{\mu\nu}$ is the electromagnetic field tensor. The $\mathcal{O}'$ operators correspond to right-handed coupling obtained by swapping P_R and P_L in the equations. In the SM, as well as in MFV models where the flavour violation is entirely ruled by the CKM matrix, the C' Wilson Coefficients are suppressed by the strange coupling, $C_i' \sim (m_s/m_b)C_i$. The operator $\mathcal{O}_7$ relates to penguin diagrams that are mediated via a photon. It represents the dominant contribution to the radiative $b \to s\gamma$ transition and contributes to $b \to s\ell^+\ell^-$ processes when the virtual photon decays into a dilepton pair. The semileptonic $\mathcal{O}_9$ and $\mathcal{O}_{10}$ correspond to penguin diagrams mediated by a Z boson and W mediated box diagrams. These are the dominant contributions in semileptonic $b \to s\ell^+\ell^-$ decays. The vertices corresponding to the radiative and semileptonic operators are illustrated in Fig. 1.7

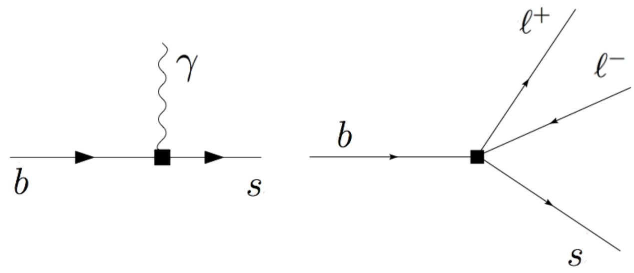

Figure 1.7: Interaction vertices corresponding to the radiative (left) and semileptonic (right) operators.

It is also common to express the semileptonic operators in a basis with left and right projected leptons

$$\begin{aligned} \mathcal{O}_{LL} &= (\mathcal{O}_9 - \mathcal{O}_{10})/2 \quad \mathcal{O}_{LR} = (\mathcal{O}_9 + \mathcal{O}_{10})/2 \\ \mathcal{O}_{RR} &= (\mathcal{O}_9' - \mathcal{O}_{10}')/2 \quad \mathcal{O}_{RL}' = (\mathcal{O}_9' + \mathcal{O}_{10})/2 \end{aligned} \tag{1.12}$$

where the Wilson Coefficients are redefined as

$$\begin{aligned} C_{LL} &= C_9 - C_{10}, \quad C_{LR} = C_9 + C_{10}, \\ C_{RR} &= C_9' - C_{10}', \quad C_{RL}' = C_9' + C_{10}. \end{aligned} \tag{1.13}$$

This basis is particularly useful in frameworks where BSM physics at a high mass scale respects the $\mathrm{SU}(2)_W$ part of the SM gauge symmetry group. Finally, in the picture presented in this section all operators were considered as universal with respect of the flavour of the involved leptons. However, BSM models often contain sources of lepton universality violation leading to a split of the same operators depending on the lepton considered: $C_i \to C_i^e, C_i^\mu, C_i^\tau$ and $\mathcal{O}_i \to \mathcal{O}_i^e, \mathcal{O}_i^\mu, \mathcal{O}_i^\tau$.

1.5.3 Phenomenology of $b \to s\ell^+\ell^-$ decays

Semileptonic b hadron decays are characterised by two kinematic regimes which are treated theoretically in different ways; Table 1.3 shows a scheme of the q^2 spectrum. The "high q^2" is the region of low hadron recoil, $q^2 > 15$ GeV$^2/c^4$, and is characterised by the energy of the hadron being less than the energy scale of QCD interactions within the meson, $\Lambda_{QCD} \sim 1$ GeV. In this region theoretical calculations of B meson decays can be simplified by working in the heavy quark limit, $m_b \to \infty$. In this limit a Heavy Quark Effective Theory (HQET) [30] can be constructed in which the heavy quark interacts only via 'soft' hadronic processes and an OPE in $1/m_b$ is valid. The "low q^2" region is where the light spectator quark is energetic and cannot be neglected. Furthermore, the light quark interacts not only via 'soft' hadronic processes, as in HQET, but also via the so-called 'collinear' hadronic processes. The boundary of this region can be set at ~ 7 GeV$^2/c^4$ which corresponds to the threshold for $c\bar{c}$ production, $(2m_c)^2$. In this region the hadronic interactions are handled by expanding in terms of the energy of the emitted energetic hadron, $1/E_h$, forming the so-called Soft-Collinear Effective Theory (SCET) [31]. In both regions decay rates can be predicted using the different methods and the biggest uncertainties come from the limited knowledge of hadronic transition matrix elements. The intermediate region is characterised by the presence of charmonium resonances, produced though tree level $b \to \bar{c}cs$ transitions and no precise theoretical calculation is available [32].

Table 1.3: A scheme of the q^2 spectrum.

q^2	$E_{K^{*0}}$	Regime	Valid theory
~ 0 GeV$^2/c^4$ < 6 GeV$^2/c^4$	$\sim m_B$ $>> \Lambda_{QCD}$	Max. recoil Large recoil	SCET
$q^2 \sim m^2_{J/\psi,\psi(2S)}$	~ 3 GeV	$c\bar{c}$ resonances	–
$q^2 > 15$ GeV$^2/c^4$ $q^2 = (m_B - m_K^{*0})^2$	$E_{K^{*0}} \sim \Lambda_{QCD}$ $E_{K^{*0}} \sim 0$	Low recoil Zero recoil	HQET

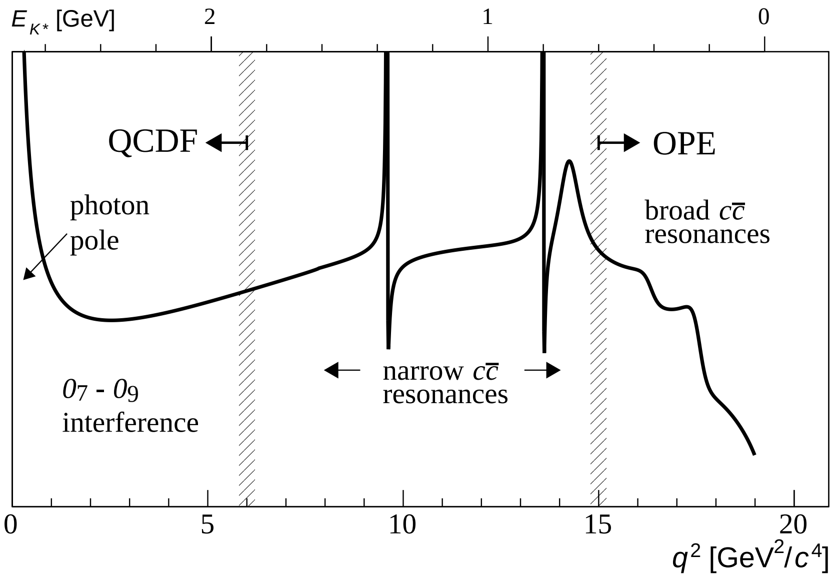

Figure 1.8: A typical q^2 spectrum of $b \to s\ell^+\ell^-$ process characterised by the photon pole at very low q^2, charmonium resonances at central q^2 and broad resonances at high q^2 [24].

As can be seen in Fig. 1.8 the very low q^2 region is characterised by a peak due to the virtual photon contribution, associated with C_7. In the region $1-6$ GeV$^2/c^4$ the interference between C_7 and C_9 becomes large, yielding sensitivity to new physics in C_9. The $7-15$ GeV$^2/c^4$ interval is dominated by the charmonium resonances, J/ψ and $\psi(2S)$. Although these decays can be experimentally vetoed in principle charmonia affect the entire q^2 space. Finally, at high q^2 broad charmonium resonances can contribute, like those observed by LHCb in $B^+ \to K^+\mu^+\mu^-$ decays [33].

1.5.4 Observables in $b \to s\ell^+\ell^-$ decays

Rare decays and especially semileptonic $b \to s\ell^+\ell^-$ processes offer a number of observables which can be used to study BSM models. The most direct effects appear in decay rates that can be enhanced by new physics but the precision on these measurements is often limited by uncertainties on the perturbative part of the calculations. Therefore, it is important to also look for different observables. One important class of observables are angular quantities that can often carry complementary information with respect to branching ratio measurements. The most basic of these observable are forward-backward asymmetries that characterise the angular distribution of final particles. For the $B^0 \to K^* \mu^+\mu^-$ decay combinations of observables have been proposed that are independent of form factor uncertainties at leading order order [24].

Another way to build safe observables is to construct ratios between similar decays, in which uncertainties due to the hadronisation process cancel out. These observables include the R_H ratios, between B^0 decays into electrons and muons, that are described in detail in Ch. 5. It is also interesting to compare decays which proceed via the same fundamental process but where the spectator quark has a different flavour. This is the case of $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^0_S\mu^+\mu^-$ decays, which are both $b \to s$ transitions where the spectator quark is an u quark in the first case and a d quark in the second. The normalised difference of the branching fractions of these decays is called isospin asymmetry.

1.6 Experimental status

To set the background for the analysis described in this thesis, this section reports a brief review of recent results of new physics searches involving rare decays or lepton flavour violation. Among these, results recently obtained by the LHCb experiment show a series of anomalies with respect to the SM that have the potential to yield to BSM scenarios.

1.6.1 Dimuon decays of b hadrons

Decays of B mesons into a pair of muons are two-body decays where the two muons are back to back in the hadron rest frame. The simple signatures of these decays makes them easy to study and the fact that they are unaffected by hadronic physics in the final state makes predictions very clean and precise. Therefore these are essential tests of the SM. The $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ decays are FCNCs that can only happen via loops and furthermore they are CKM-suppressed, which makes them particularly rare. In addition to that the decay of a pseudo-scalar B meson into two muons has a significant helicity suppression. The latest SM predictions for these decay rates are [34]:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.65 \pm 0.23) \times 10^{-9} \text{ and} \tag{1.14}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}. \tag{1.15}$$

The uncertainties on these values are dominated by the knowledge of the decay constants and CKM-elements. BSM models can produce significant enhancement to these decay rates. Furthermore, the measurement of their ratio is a stringent test of the MFV hypothesis. A combination of the LHCb and CMS results measured the values [35]:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9} \text{ and} \tag{1.16}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}. \tag{1.17}$$

Neither decay had been previously observed, while now the B_s^0 decay is observed

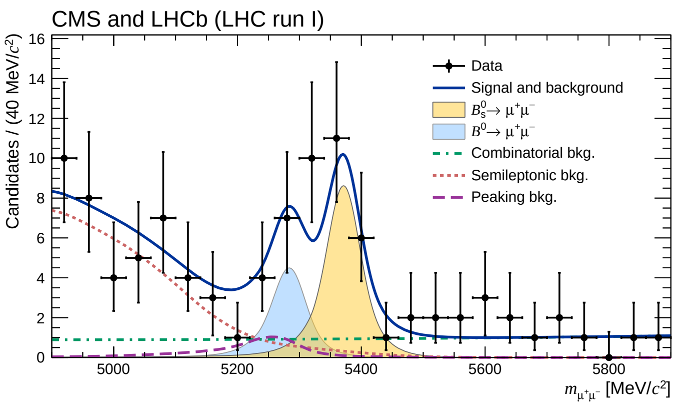

Figure 1.9: Dimuon invariant mass of B candidates showing peaks corresponding $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays [35].

with a significance of 6σ and evidence for the B^0 decay is found at 3σ significance level. The measured branching fractions are compatible with SM predictions within

2σ and put strong constraints on the available parameter-space for BSM theories. Figure 1.9 shows the fit the dimuon invariant mass of B meson candidates where the peaks of the two decays are visible.

1.6.2 Semileptonic $b \to s\ell^+\ell^-$ decays of b hadrons

At the LHC energies is possible to collect large data samples of semileptonic decays, especially those with muons in the final state. Many branching fractions of semileptonic B meson decays were recently measured at the LHCb experiment, including $B \to K\mu^+\mu^-$, $B \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ [36, 37, 38]. Baryon decays were also studied at LHCb: including the rare $\Lambda_b \to \Lambda\mu^+\mu^-$ decay [39], whose analysis is described in this thesis. In contrast to purely leptonic decays, SM predictions for semileptonic decays are affected by the knowledge of hadronic form factors, which results in relatively large uncertainties, $\mathcal{O}(30\%)$. As a result measurements are now typically more precise than predictions.

Among the measurements of angular observables that can be affected by new physics, particular interest was risen by the measurement of a set of observables in $B \to K^{*0}\mu^+\mu^-$ decays, free from form factors uncertainties at leading order [40]. Most of the measurements are found to be in agreement with SM predictions with the exception of the P_5' observable, shown in Fig. 1.10, which presents a local 3.7σ deviation. Attempts to build a consistent picture point to a new physics contribution to the Wilson Coefficient C_9 [41]. An angular analysis of $B^+ \to K^+\mu^+\mu^-$ decays was also performed, where observables are found to be compatible with SM predictions [42]. Other observables for which the sensitivity to form factors effects is reduced are the CP asymmetry between B and $\bar{B}$ decays, $\mathcal{A}_{CP}$, and the isospin asymmetry between B^0 and B^+ decays, $\mathcal{A}_{CP}$. Due to the small size of the corresponding CKM elements, CP asymmetries of $B^0 \to K^{(*)}\mu^+\mu^-$ decays are tiny in the SM, $O(10^{-3})$. In BSM models new sources of CP violation can arise and therefore $\mathcal{A}_{CP}$ measurements are a powerful null test of the SM. The isospin asymmetry is not zero in the SM due to isospin breaking effects in the form factors. This is expected to be $\sim 1\%$ at low

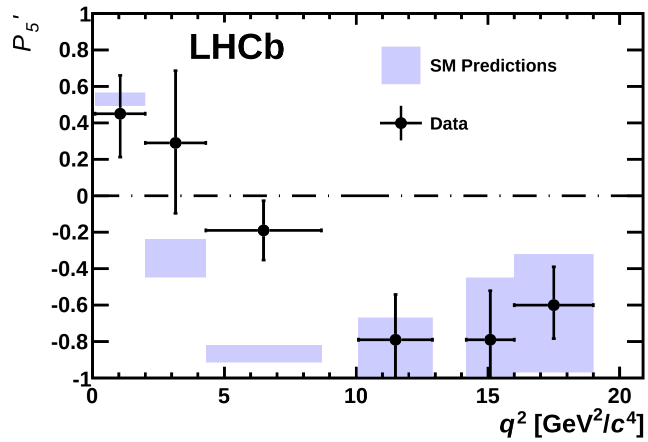

Figure 1.10: Measurement of the observable as a function of q^2, showing a tension with SM predictions in the 2–6 GeV^2/c^4 region [40].

Asymmetry	$B^0 \to K^+\mu^+\mu^-$		$B^0 \to K^{*0}\mu^+\mu^-$	
	1.1–6 $[\mathrm{GeV}^2/c^4]$	15.0–22.0 $[\mathrm{GeV}^2/c^4]$	1.1–6 $[\mathrm{GeV}^2/c^4]$	15.0–19.0 $[\mathrm{GeV}^2/c^4]$
$\mathcal{A}_{CP}$	0.004 ± 0.028	-0.005 ± 0.030	0.094 ± 0.047	-0.074 ± 0.044
$\mathcal{A}_I$	$-0.10^{+0.08}_{-0.09} \pm 0.02$	$-0.09 \pm 0.08 \pm 0.02$	$0.00^{+0.12}_{-0.10} \pm 0.02$	$0.06^{+0.10}_{-0.09} \pm 0.02$

Table 1.4: Measurement of CP and isospin asymmetry in $B^0 \to K^{*0}\mu^+\mu^-$ decays from the LHCb experiment [24].

q^2 and increase to $\sim 10\%$ as q^2 tends to zero. The LHCb experiment, using the full dataset collected in Run I, corresponding to an integrated luminosity of 3 fb^{-1} and $\sim 10^9$ B decays, measured both of these asymmetries to be consistent with zero [36, 43], as reported in Tab. 1.4. Recently, progress was also made measuring also electron channels. The branching fraction of the $B^0 \to K^{*0}e^+e^-$ decay was measured to be $(3.1 \pm 1.3) \times 10^{-7}$ in the dilepton mass interval 30–1000 MeV/c^2 [44]. Furthermore, for the first time angular observables were measured for this decay and found to be consistent with SM predictions [45].

1.6.3 Lepton Flavour Violation searches

Several Lepton Flavour Violation (LFV) searches are linked to rare decays as they involve small branching ratios in the SM that can be enhanced by BSM physics. Lepton flavour conservation is experimentally well-established measuring the branching ratios of decays of muons into electrons and no neutrinos, but has no strong theoretical explanation in the context of the SM. In fact it is already observed that flavour is not conserved in neutrino oscillations. The best-studied decays violating lepton flavour are rare muon decays including $\mu^+ \to e^+\gamma$ and $\mu^+ \to e^+e^-e^+$. Since muons can be abundantly produced and the final states are simple, these decays provide the best constraints to LFV. The present best upper limits are 1.2×10^{-11} for the radiative decay and 1.0×10^{-12} for $\mu^+ \to e^+e^-e^+$ obtained respectively by the MEGA [46] and SINDRUM [47] experiments. Several LFV searches in the B sector have recently been performed at the LHCb experiment including decays such as $B^0 \to e\mu$ [48] and τ decays such as $\tau \to \mu^+\mu^-\mu$ [49]. None of these searches has found evidence of new physics so far and therefore they set limits, constraining the parameter space available for BSM models. Figure 1.11 shows a summary of the best limits set at different times on LFV searches [50].

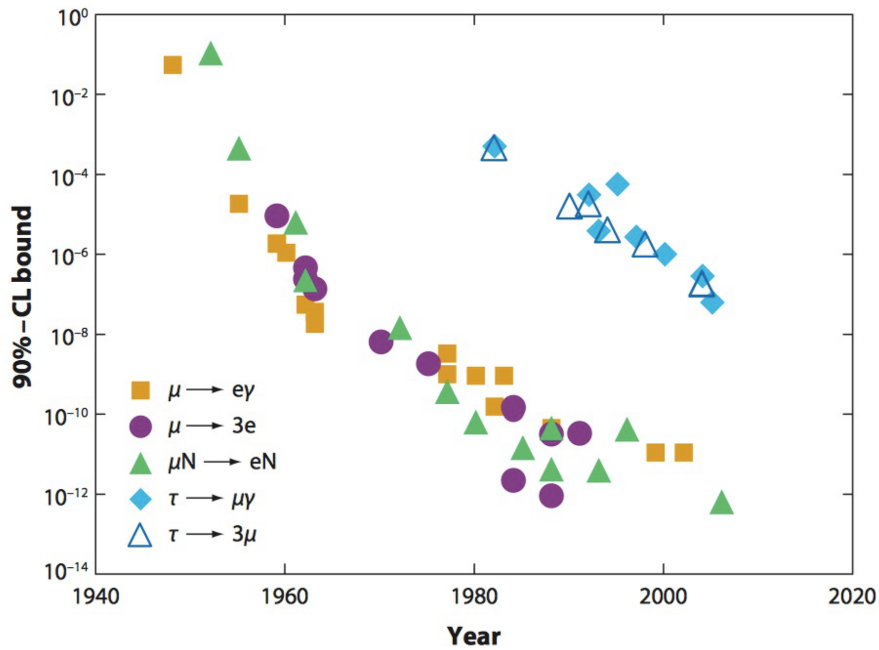

Figure 1.11: Summary of limits set in LFV searches as a function of time [50].

CHAPTER 2

The LHCb detector at the Large Hadron Collider

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [51] is a synchrotron particle accelerator with a circumference of 27 km located about 100 m underground at CERN in the surroundings of Geneva, Switzerland. Two proton beams circulate in opposite directions around the ring and cross each other in four points, in which particle detectors are placed. These include two general-purpose detectors, ATLAS and CMS, sitting on opposites sides of the ring and two smaller detectors, ALICE and LHCb that are designed to study specific topics (see Fig. 2.1).

Each beam consists of a series of proton bunches, up to a maximum of 2835. Each bunch consists of about 10^{11} protons and the bunch spacing is such that the nominal bunch crossing rate is 40 MHz. The beams are injected into pre-accelerators and then pass into the LHC through the CERN acceleration system shown in Fig. 2.1.

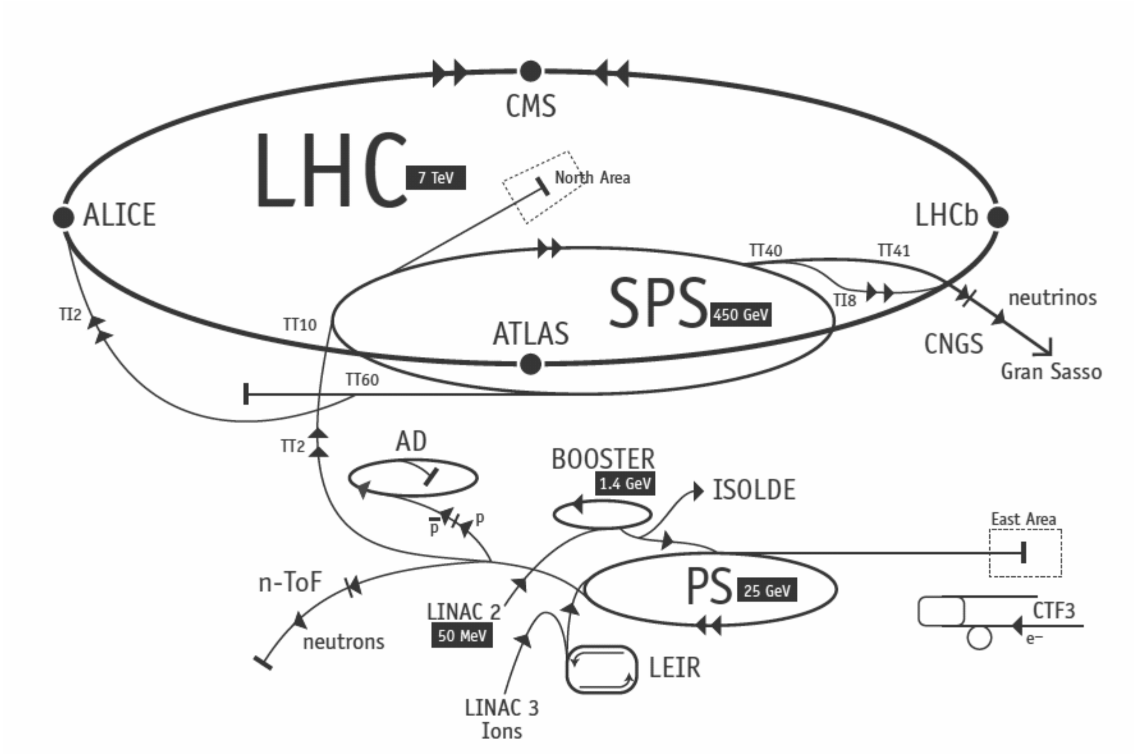

Figure 2.1: Scheme of CERN accelerators.

Protons are produced from hydrogen gas and are initially accelerated to an energy of 50 MeV in a linear accelerator (LINAC). Then they are injected into the Proton Synchrotron Booster (PSB), where they are boosted to an energy of 1.4 GeV, into the Proton Synchrotron (PS) to 25 GeV and into the Super Proton Synchrotron (SPS) to 450 GeV. Finally, protons enter into the LHC storage ring, where they are accelerated from injection energy to the final one by radio frequency (RF) cavities. The beams are steered around the ring by 8 T magnetic fields produced by 15 m long superconducting niobium-titanium dipole magnets and focused by quadrupole magnets. The LHC magnets use a design in which both proton beam pipes are contained in the same housing, allowing a common liquid helium cooling the system to be used. The LHC began colliding proton beams in "physics mode" in 2009 at a centre of mass energy of $\sqrt{s} = 900$ GeV and from April 2010 to November 2011 accelerated beams at $\sqrt{s} = 7$ TeV (3.5 TeV per proton beam) with a maximum instantaneous luminosity of $3 \cdot 10^{33}$ cm^{-2}s^{-1}, while in 2012 the energy was increased to 8 TeV. The LHC maximum design energy is 14 TeV, and its design luminosity is 10^{34} cm^{-2}s^{-1}. After a long shut down to upgrade and maintain the machine, a new run started in June 2015, in which protons are collided at a centre of mass energy of $\sqrt{s} = 13$ TeV. At this energy the total proton-proton cross-section is expected to be roughly 100 mb.

2.2 The LHCb detector

The LHCb detector [52] was designed to study decays of B and D mesons, mainly looking for CP-violating processes. In 2011, running at a centre of mass energy of 7 TeV, the cross-section for $b\bar{b}$ production was measured to be 284 ± 53 μb [53], while it will be $\sim$ 500 μb at the current LHC energy, 13 TeV. At these high energies, proton-proton interactions produce highly boosted virtual gluons which produce $b\bar{b}$ pairs at small angles, close to the beam pipe. For this reason the LHCb detector is designed to have a very forward angular coverage. The detector is fully instrumented from 10 mrad to 300 mrad, corresponding to an interval $2 < \eta < 5$, where η is the "pseudorapidity", a quantity defined as:

$$\eta = -\ln(\tan(\theta/2)), \tag{2.1}$$

where θ is the angle between a particle's momentum and the beam direction [1].

At LHCb's collision point the luminosity can be adjusted by displacing the beams from head on collisions while keeping the same crossing angle. This allows the experiment to keep an approximately constant instantaneous luminosity, compensating for the reduction in beam intensity due to extended operation periods. This also means that the average number of interactions per bunch crossing can be regulated, which is important because the detector efficiency, especially in detecting secondary vertices, decreases for events with an high number of primary vertices (PV). Reducing the particle occupancy through the detector also keeps radiation damage to a

[1]LHCb's reference system has the z axis in the direction of the beam, the x axis directed to the centre of the accelerator and y is directed upward. Then we define θ as the angle with the beam direction and ϕ as the position around the beam in the xy plane, taking $\phi = 0$ on the x axis. The origin, $(x, y, z) = (0, 0, 0)$, corresponds to the centre of the interaction area.

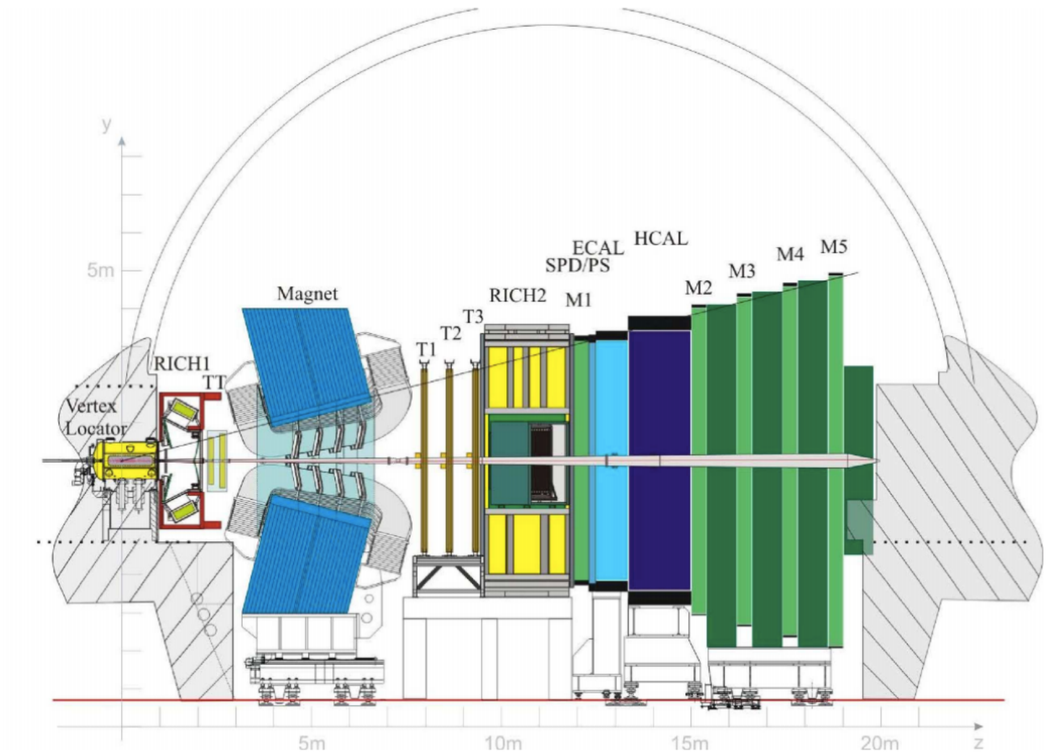

Figure 2.2: A side view of the LHCb detector [52].

minimum. Until the end of 201 the instantaneous luminosity was $3 \cdot 10^{32}$ cm^{-2}s^{-1}, corresponding to an average number of 1.5 PVs per bunch crossing and at the end of 2011 LHCb had collected an integrated luminosity of 1 fb^{-1}. In 2012 the luminosity was increased and a further 2 fb^{-1} of data were collected.

Experiments like BaBar at the Stanford Linear Accelerator (SLAC), Belle at KEK at J-PARC (Japan) and the Tevatron experiments at Fermilab have made measurements in heavy flavour physics which have so far been found to be consistent with the SM predictions. However, some of the deviations from the SM are expected to be very small. Therefore LHCb was designed to make the most precise measurements in heavy flavour physics to test the consistency of the SM and look for new physics.

The LHCb detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the *pp* interaction region, and a larger silicon-strip and drift tubes detectors located on both sides of a dipole magnet with a bending power of about 4 Tm. Charged hadrons are identified using two Ring-Imaging

Cherenkov detectors (RICH) [54]. Photon, electron and hadron candidates are identified by a calorimeter system and muons by a system composed of alternating layers of iron and multi-wire proportional chambers [55]. A schematic view of the detector is shown in Fig. 2.2 and more details on each sub-detector are given in the following sections.

2.3 The magnet

Charged particle trajectories are deflected horizontally in the magnetic field so that their momentum can be measured from the radius of curvature. The LHCb dipole magnet is composed of two coils supported by an iron yoke and is shaped to fit the LHCb angular acceptance. Unlike the other LHC experiments, LHCb uses a warm magnet which can be easily ramped allowing the field polarity to be inverted periodically. When the polarity is flipped, particles of a given sign are bent in the opposite direction. This method is used to limit systematic uncertainties that can arise due to performance variations in different areas of the detector and average out using data taken in both polarities. A current of 5.85 kA flows in the magnet generating an integrated magnetic field of 4 Tm for 10 m long tracks. In order to achieve the required momentum precision the magnetic field must be mapped with a 10^{-4} precision. For this reason a grid of 60 sensors is positioned inside the magnet and provides real time magnetic field maps.

2.4 Tracking system

B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means they travel about 1 cm before decaying to form a displaced vertex. To study specific decays, it is therefore important to be able to separate the particles produced at the primary *pp* vertex and at the B decay secondary vertex (SV). The tracking system consists of the Vertex Locator (VeLo), and 4 tracking stations: the Tracker Turicensis

(TT), which are located before the magnet and the T1, T2 and T3 stations, located after of the magnet. The latter three stations are in turn formed by two subsystems: the Inner Tracker (IT) close to the beam-line, where the particle density is greatest, and the Outer Tracker (OT) covering the rest of the acceptance.

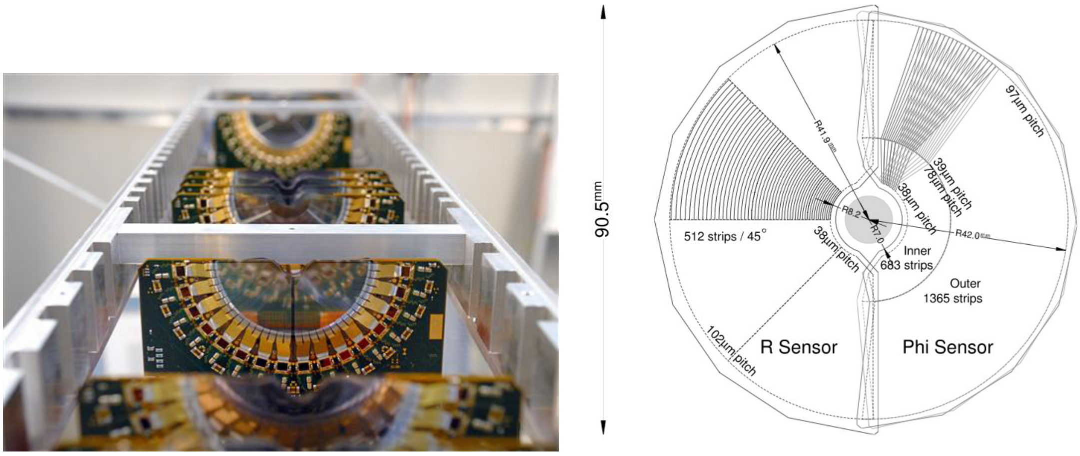

Figure 2.3: On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [52].

The VeLo accurately measures positions of tracks close to the interaction point which is essential to reconstruct production and decay vertices of bottom and charm hadrons. The VeLo is composed by 21 silicon modules that surround the beam axis and are positioned from $z = -18$ cm to $+80$ cm. The sensitive region of the VeLo starts at an inner diameter of only 8 mm from the beam axis and it is able to detect particles within a pseudorapidity range $1.6 < \eta < 4.9$. The VeLo is housed in its own vacuum vessel of thin aluminium foil, which protects the vacuum of the beam pipe from any outgassing. The silicon layers composing the VeLo consist of two modules each including two types of sensors: the ϕ-sensor, which measures the azimuthal position around the beam, and the R-sensor, which measures the radial distance from the beam axis. A sketch of the VeLo sensors is shown in Fig. 2.3. The sensors are 300 μm thick and to ensure that they cover the full azimuthal angle the right-side module is placed 1.5 cm behind the left-side module on the z-axis and they overlap. There are two modules which cover the backward direction and are used as a veto for multiple interactions; this is called the pileup system.

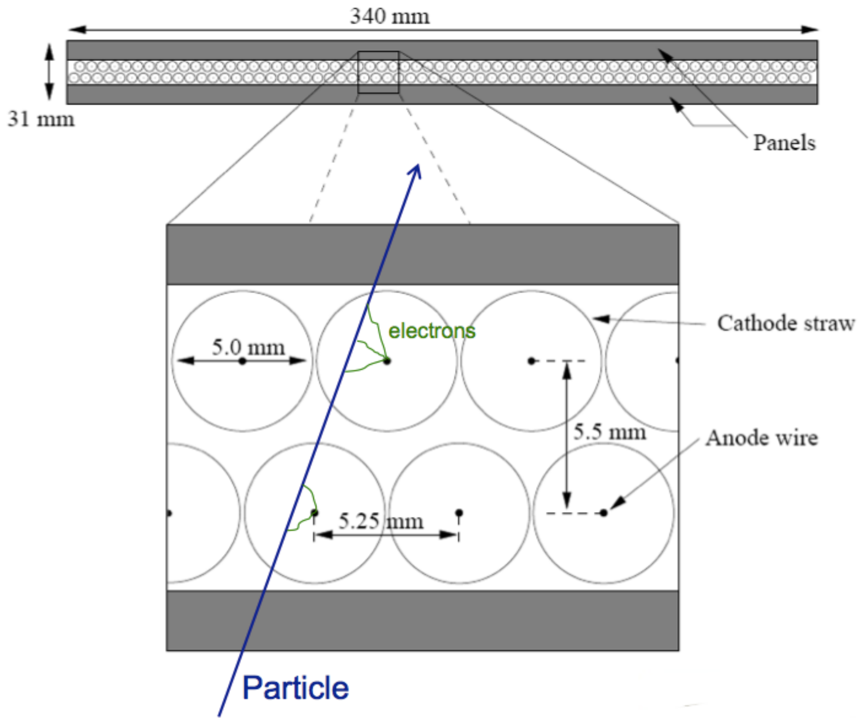

Figure 2.4: Sketch of the straw tubes which constitute the Outer Tracker layers [52].

The IT and TT both use silicon strips and together constitute the Silicon Tracker (ST). Straw tubes are instead used in the OT, of which a sketch is shown in Fig. 2.4. The IT requires a higher inner granularity because of the greater flux of particles close to the beam pipe. In fact, it covers only 1.3% of the total area of IT plus OT but it contains about 20% of the tracks. Each ST station has four detection layers: the first and last are vertical, measuring the track position in x, while the second and third layers are rotated by an angle of +5 and -5 degrees, which allows the measurement ofthe y coordinate. The TT is placed upstream of the magnet to allow the reconstruction of tracks from low-momentum particles, which are bent out of the downstream acceptance. Overall the tracking system provides a measurement of momentum, p, with a relative uncertainty that varies from 0.4% at 5 GeV$/c$ to 1.0% at 200 GeV$/c$. The impact parameter (IP), namely the minimum distance of a track to a primary vertex, is measured with a resolution of $(15 + 29/p_{\mathrm{T}})$ μm, where p_{T} is the component of the momentum transverse to the beam, in GeV$/c$. The z-axis position of a PV reconstructed with 35–40 tracks can be measured with a precision of roughly 50–60 μm. The decay products of B mesons tend to have high IP values because the B decay imparts transverse momentum to them. Therefore, accurate IP and vertex displacement measurements allow LHCb to distinguish effectively between B meson decays and background processes.

2.5 Calorimeters

In general the main purpose of a calorimeter system is to determine the energy of particles but in LHCb it is mostly used to help the identification electrons and hadrons. Sampling calorimeters, as those used in LHCb, are composed of layers of absorber and active material. Particles interact with the absorber layers and produce a cascade of secondaries, that multiply quickly and are detected by the active part, which is usually composed of scintillating layers. The light produced is detected by photo-multipliers (PMTs) and it is approximately proportional to the energy of the deposited particles. Calibration is then used to translate the signal into an energy measurement. The LHCb's calorimeter system consists of the Scintillator Pad Detector (SPD), the Pre-Shower Detector (PS) as well as the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). A sketch of the LHCb calorimeters is shown in Fig. 2.5. The SPD/PS cells are read out with PMTs located outside the LHCb acceptance, while the ECAL and HCAL have individual PMTs located on the modules. All four detectors are segmented, which allows the energy deposits to be associated to the tracks detected by the tracking system. The segmentation of the cells varies according to the distance from the beam pipe due to the different track density.

The most difficult identification in LHCb is that of electrons. The rejection of a high background of charged pions is achieved using a longitudinal segmentation of the

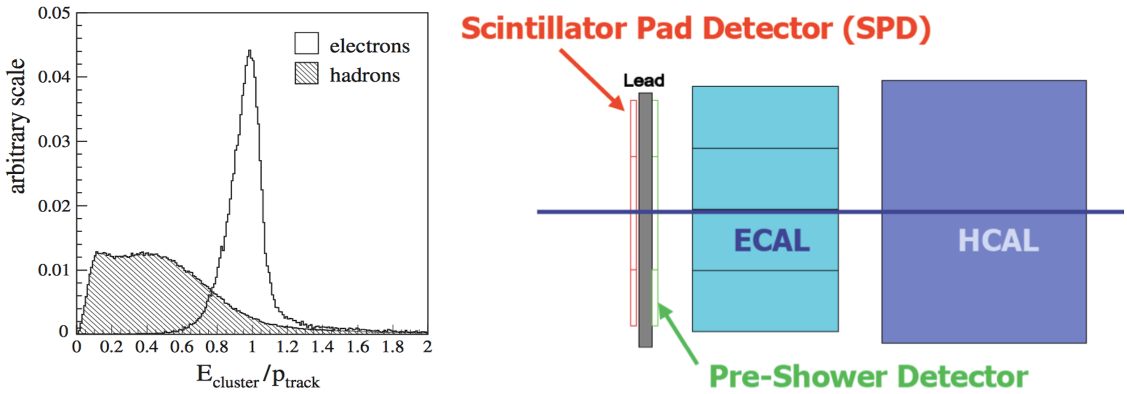

Figure 2.5: (left) The ratio of the energy deposited in the ECAL and the particle momentum, which allows the separation between electrons and hadrons [52]. (right) A schematic of the calorimeter system.

electromagnetic calorimeter which is provided by the PS detector added in front of the main electromagnetic calorimeter, ECAL. Electrons also have to be distinguished from high energy π^0s and photons. For this purpose the SPD calorimeter, detecting charged particles, is located in front of the PS and ECAL detectors. Figure 2.5 illustrates how the ratio between the energy detected in the ECAL and a particle's momentum allows the separation of electrons and hadrons.

The ECAL is formed by 66 lead layers (2 mm thick) separated by 4 mm thick plastic scintillator layers. In order to obtain the highest energy resolution the showers from high energy photons must be fully absorbed. For this reason the ECAL has a thickness of 25 radiation lengths and its resolution is measured to be $\sigma_{\mathrm{ECAL}}(E)/E = 10\%/\sqrt{E(\text{ GeV})} + 1\%$ [52], which results in a mass resolution of ~ 70 MeV/c^2 for B mesons and ~ 8 MeV/c^2 for π^0. The HCAL is mainly used for triggering and it is similar to the ECAL but with 4 mm thick scintillator layers and 16 mm thick absorber layers. The trigger requirements on the HCAL resolution do not depend on the containment of the hadron showers as much as for the ECAL, therefore, due to space limits, its thickness is only 5.6 interaction lengths and its resolution is given by $\sigma_{\mathrm{HCAL}}(E)/E = 69\%/\sqrt{E(\text{ GeV})} + 9\%$.

2.5.1 Bremsstrahlung recovery for electrons

Bremsstrahlung is an electromagnetic radiation produced by charged particles that undergo an acceleration. Typically electrons produce Bremsstrahlung when deflected by atomic nuclei. The probability of emitting bremsstrahlung radiation is proportional to the inverse of the squared mass of the particle ($1/m^2$) and therefore it is most relevant for electrons. At LHC energies, if electrons radiate after the magnet, the photon will hit the same calorimeter cell as the electron and the energy will be automatically recovered, as illustrated in Fig. 2.6. However, if the photon is emitted before the magnet, the electron will be deflected by the magnetic field whereas the photon will continue on its initial trajectory, with its energy being deposited in a different part of the calorimeter. Missing this energy results in a poorer reconstructed invariant mass resolution, so it is desirable to recover these bremsstrahlung photons. A tool for bremsstrahlung recovery is available in the LHCb analysis software. This tool looks for other clusters in the calorimeter and, reconstructing the trajectory of the electron, checks if they may be associated with photons emitted. The photon energy is then added to the electron and its momentum is recalculated. For more information see Ref. [56].

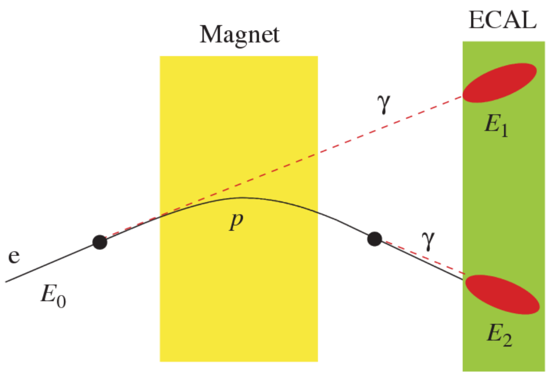

Figure 2.6: Schematic view of the bremsstrahlung recovery [52].

2.6 RICH

The two RICH detectors are a special feature of LHCb, as it is the only experiment at LHC using them. These detectors take advantage of the Cherenkov radiation produced by particles passing through a medium with speed higher than the speed of light in the medium. The Cherenkov light, as shown in Fig. 2.7, is produced in cones with a specific opening angle depending on the velocity of the particle. The relation between the angle and the particle velocity can be written as

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

where $\beta = v/c$ and n is the refraction index of the medium.

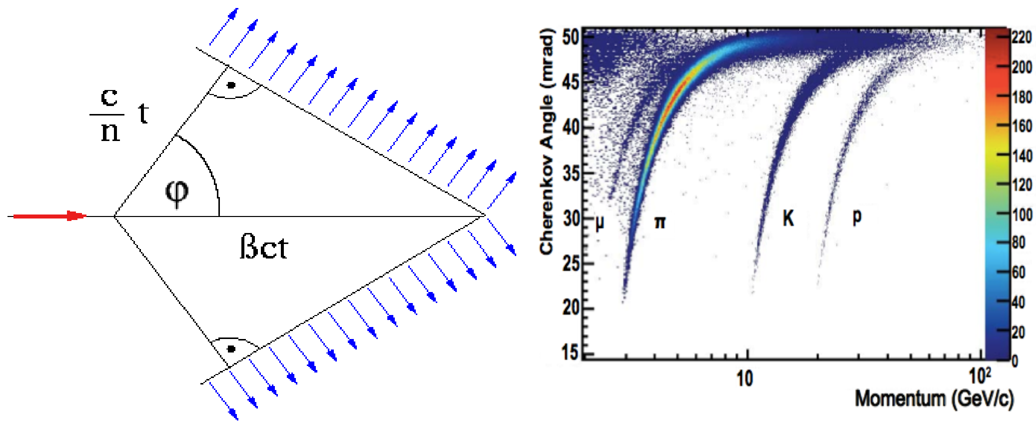

Figure 2.7: (left) A sketch of Cherenkov light emission and on the right the Cherenkov angle versus the particle momentum, where one can see that the study of the Cherenkov angle allows distinguish particles identities. (right) Measured Cherenkov angle as a function of particle momentum [52].

RICH 1 is located before the magnet in order to cover a larger angular acceptance. Its purpose is to ensure particle identification over the momentum range $1 < p < 70$ GeV$/c$. It uses two radiators: C_4F_{10} that covers the momentum range $5 - 70$ GeV$/c$ and silica aerogel which covers $1 - 10$ GeV$/c$. RICH 2 is positioned after the magnet and tracking stations and it identifies higher momentum particles from approximately 20 GeV$/c$ up to beyond 100 GeV$/c$ using CF_4 as a radiator. The Cherenkov light produced when charged particles travel through the radiators, is reflected and focussed using mirrors, which are tilted so that the ring image is reflected onto arrays of PMTs. The radius of the ring can be used to measure the opening angle of the Cherenkov cone because of the known geometry. The photodetectors are located outside of the LHCb acceptance in order to reduce the amount of material that the particles have to traverse. Pattern recognition algorithms are then used to reconstruct the Cherenkov rings.

2.7 The muon system

It is essential for many of the key physics analyses in LHCb to be able to identify muons in decay final states. Muons are the most penetrating particles that can be detected at LHC experiments, so the muon chambers are the farthest sub-detectors from the interaction point. The muon system consists of five stations (M1 - M5), the first one being located before the calorimeters in order to improve p_{T} measurements. The remaining four stations are behind the HCAL and are separated from each other and interleaved with 80 cm thick iron blocks, which absorb hadrons, electrons and photons to ensure that only muons reach the final muon station. A schematic of the muon system is shown in Fig. 2.8. Only muons with a minimum momentum of 10 GeV/c traverse all of the five stations and, for positive identification of a muon, the trigger requires a signal in each of them. Each station has a detection efficiency of at least 95% and the detectors also provide position measurements. Since there is a larger particle flux close to the beam pipe, the stations are divided into four concentric rectangular regions (R1-R4) with increasing cell size, which results in a similar occupancy over the four regions. All of the muon stations use Multi Wire Proportional Chambers (MWPC) except for the inner region of M1, where the particle flux is too high. In this region triple-GEM (Gas Electron Multiplier) detectors are used because of their better ageing properties as they have to withstand a rate up to 500 kHz cm^{-2} of charged particles. These detectors consist of three gas electron multiplier foils sandwiched between and anode and a cathode.

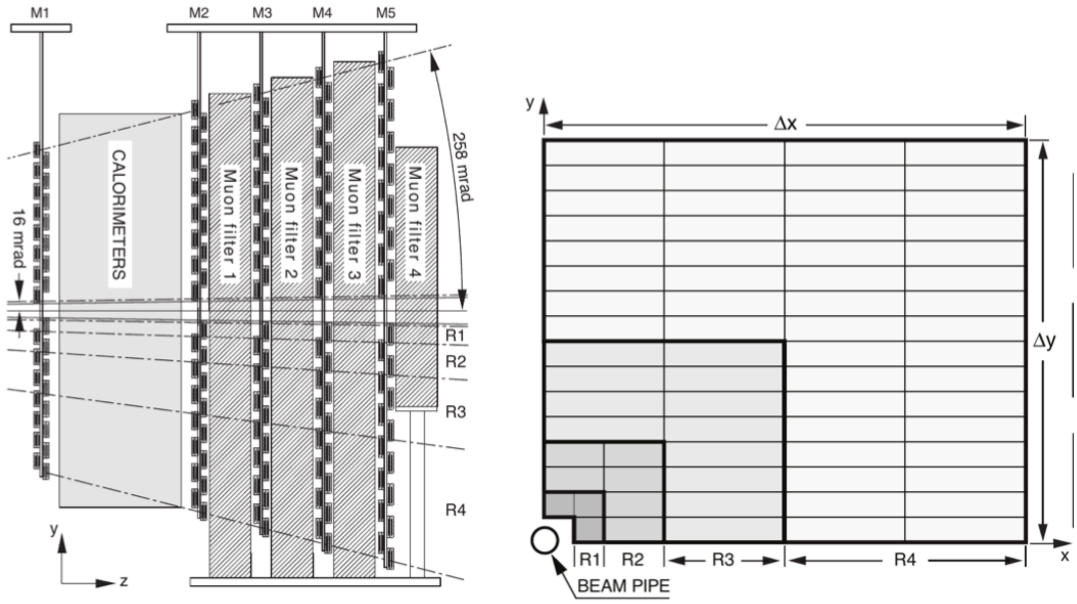

Figure 2.8: The LHCb muon system [52].

2.8 Particle identification

Particle identification (PID) is an important feature in LHCb and it is performed in various ways. The electromagnetic calorimeters can distinguish between pions and electron, the muon chambers identify muons and the RICH detectors can be used to identify heavier charged particles such as protons and kaons.

The RICH assigns an ID to a track calculating the global likelihood for the observed distribution of hits being consistent with the expected distribution from various ID hypotheses. The algorithm iterates through each track and recalculates the likelihood when the track PID hypothesis is changed to that of an electron, muon, kaon or proton. For electrons and muons additional information from the calorimeter and muon systems is also used. The hypothesis which maximises the likelihood is assigned to the track.

To quantify the quality of the ID the pion hypothesis is used as a reference point and the probability of a specific ID is given in terms of Log-Likelihood difference between the given ID hypothesis and the pion one. This variable is called Delta

Log-Likelihood (DLL) and denoted with “PID”. For example:

$$\mathtt{PID}_K = \mathrm{DLL}_{K-\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi) \tag{2.3}$$

quantifies the probability of a particle being a kaon rather than a pion. Figure 2.9 shows the efficiency for correctly identifying and mis-identifying kaons and protons as a function of the measured momentum of the particle. For kaons the efficiency drops at momenta below 10 GeV, where they fall below threshold for the gas radiators. The DLL cuts enable LHCb physics analyses to distinguish between kinematically similar decays with different final states, such as B^0 and B_s^0 mesons decaying into two hadrons. Figure 2.10 illustrates the power of particle identification, showing how the application of DLL cuts can be used to isolate $B^0 \to \pi^+\pi^-$ decays from other two-body B decays. The identification of muons is particularly important in

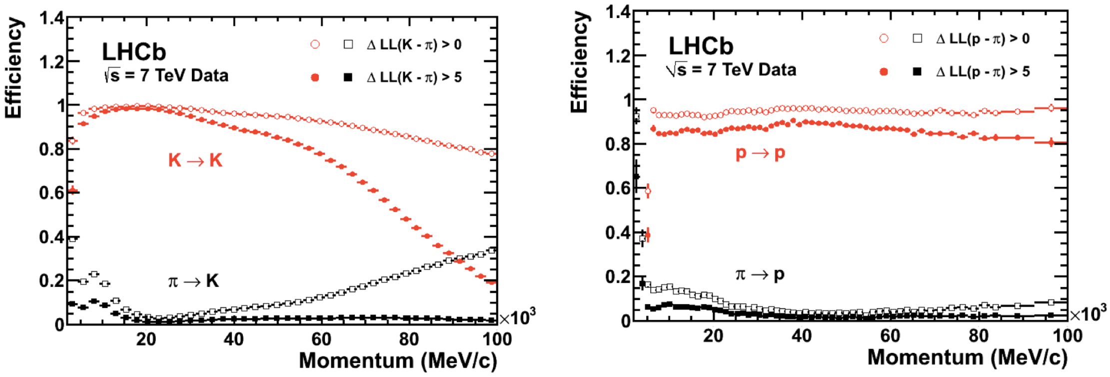

Figure 2.9: Particle Identification performances for kaons (left) and protons (right) as a function of the measured momentum of the particles [52].

LHCb and it is quantified using two variables: the DLLμ and the `isMuon` variable. The latter is a boolean variable determined by defining a ‘field of interest’ around a track trajectory extrapolated through the muon chambers. The variable is set to true if hits in multiple muon stations are found in the field of interest.

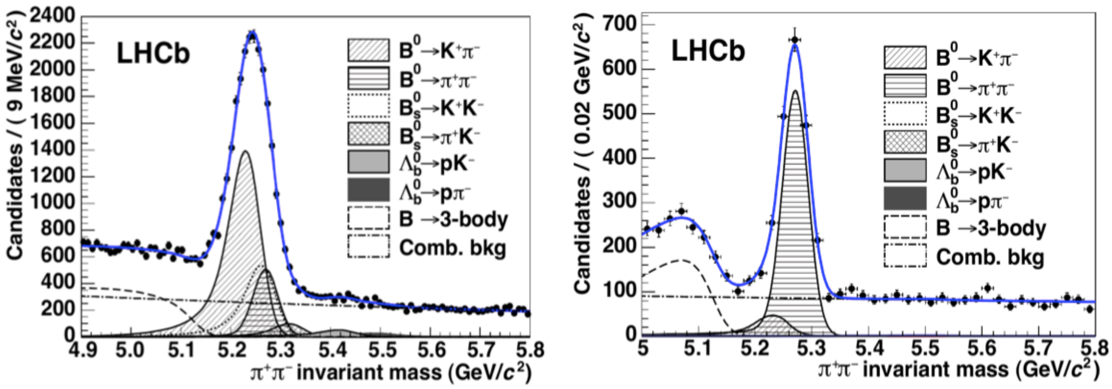

Figure 2.10: Invariant mass peak of the $B^0 \to \pi^+\pi^-$ decay before (left) and after (right) the application of PID requirements [57].

2.8.1 PID calibration

In order to be able to calculate detection efficiencies, a "data-driven" method was developed. The calibration software is referred to as `PIDCalib` package [57]. This tool uses decays where final particles can be identified thanks to their kinematic properties. For example the $K^0_S \to \pi^+\pi^-$ decay has a clear signature with a displaced vertex and can be easily singled out from other decays and used to test pion ID efficiency. The narrow peaks of the $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$ decays allow muon and electron efficiencies to be calibrated. A "tag-and-probe" method is used in this case, where only one of the two leptonic tracks is reconstructed requiring the correct identity and the other one is used to probe the PID efficiency. Finally, $\phi \to KK$ samples and $D^{*+} \to D(\to K^-\pi^+)\pi^+$ decays, where the D^{*+} is used to tag the decay, are used to test the kaon efficiency. In all cases the residual background is subtracted using the ${}_s\mathcal{P}$lot technique [58].

2.9 Trigger and software

The LHCb trigger system [59] consists of a hardware stage, L0, based on information from the calorimeters and muon system, followed by a software stage, the High-Level Trigger (HLT), which applies a full reconstruction of the events. To increase performance, the HLT is further split into two stages, HLT1 and HLT2. The HLT1 phase happens in real time and saves data in local disks while the HLT2 phase uses the resources available during periods with no beam. The event selected by the HLT2 stage are then saved for offline analysis. Figure 2.11 shows a scheme of the trigger system. The bunch crossing frequency is 40 MHz, which corresponds to an instantaneous luminosity of $2 \cdot 10^{32}$ cm^{-2}s^{-1} for LHCb. About 15% of the total number of $b\bar{b}$ pairs produced will contain at least one B meson with all of its decay products within the detector acceptance. This rate needs to be reduced to about 2 kHz at which the events can be written to disk.

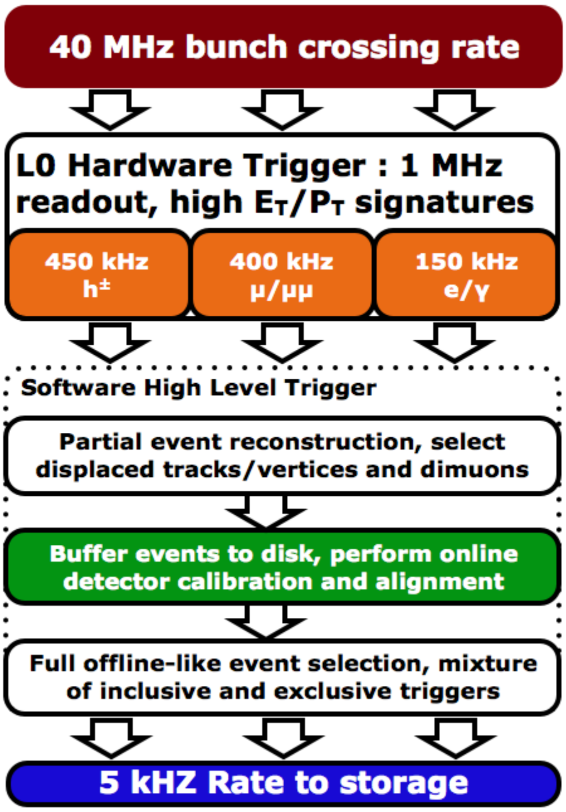

Figure 2.11: Scheme of the LHCb trigger system [52].

The L0 trigger reduces the rate of visible interactions from 10 MHz to 1 MHz. Due to the heavy mass of B mesons, they often produce particles with high energy and momentum. Therefore the trigger selects events with large E_{T} deposits in the calorimeter or high p_{T} muons. The event is classified as `L0Muon` if it was triggered due to information from the muon detector, while the information from the calorimeters is used to divide the events into five categories: `L0Photon`, `L0Electron`, `L0LocalPion`, `L0GlobalPion`, `L0Hadron`. The PS detector information is converted to a photon flag (`PS && !SPD`) or an electron flag (`PS && SPD`). The "local" label of the `L0Pion` trigger refers to π^0 reconstructed though their $\gamma\gamma$ decay, where the two photons fall in the same ECAL element, they are labelled "global" otherwise. The first four calorimeter triggers require energy clusters in the ECAL, while `L0Hadron` requires clusters also in the HCAL. The HLT1 uses information from the VELO and trackers performing a partial reconstruction of the event and reduces the rate to 2 kHz by adding requirements of the IP and χ^2 of tracks. Finally, the HLT2 involves a full reconstruction of the event and includes many "lines" designed to select specific decay structures.

LHCb also developed an extended simulation software in order to reconstruct efficiencies and signal shapes. In the simulation, pp collisions are generated using PYTHIA8 [60, 61] with a specific LHCb configuration [62]. Decays of hadronic particles are described by EVTGEN [63], and final state radiation is generated using PHOTOS [64]. Finally, the interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [65] as described in Ref. [66]. For this analysis in this thesis, the ROOT framework [67] is used to analyse data and the RooFit package to perform maximum likelihood fits. A multivariate analysis is also performed based on the NeuroBayes package [68, 69], which provides a framework for neural network training.

2.10 Constrained kinematic fits

The resolution of key variables, such as the measured invariant mass of decaying particles, can be improved by imposing constraints on the measured quantities to remove redundant degrees of freedom. The four-momentum conservation can be ensured at each vertex and the origin and decay vertices of a particle are related via the momentum of the particle. Furthermore, additional constraints can be imposed due to a particular decay hypothesis such as the known invariant masses of final and intermediate particles. In order to do this the `DecayTreeFitter` tool was developed by the BaBar experiment and later used by LHCb [70]. The algorithm takes a complete decay chain and parametrises it in terms of vertex positions, decay lengths and momentum parameters. These parameters are then fit simultaneously, taking into account the relevant constraints, including the information from photons. Figure 2.12 illustrates the effect of the application of the kinematical fit on the 4-body invariant mass of the final daughters of the $\Lambda_b^0 \to J/\psi\,\Lambda$ decay. The resolution in this case improves by over a factor of 2. Furthermore, the χ^2 from the kinematic fit can be used to quantify the compatibility with a specific decay structure, which helps to separate candidates where random particles from the event have been added to the decay tree, or where one or more particles is not reconstructed or mis-identified.

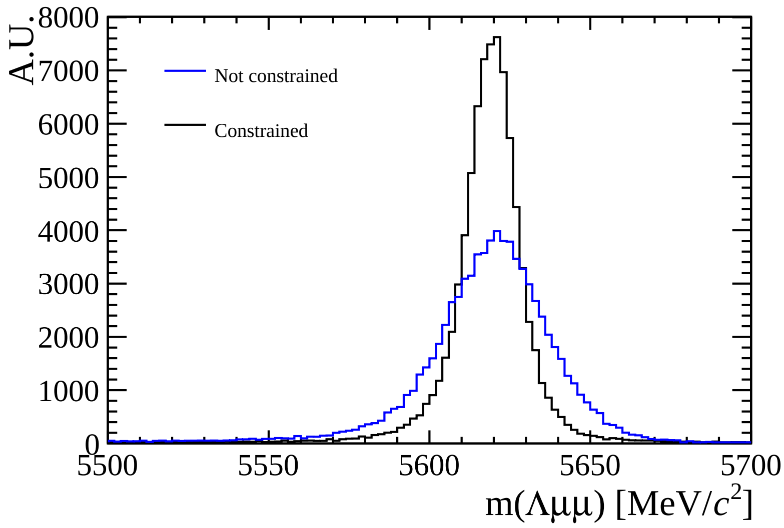

Figure 2.12: Invariant mass of the final daughters of simulated $\Lambda_b^0 \to J/\psi\,\Lambda$ decays calculated with and without constraints using the `DecayTreeFitter` tool.

2.11 Validation of hadronic processes in the simulation

Particle-antiparticle asymmetries are of major interest for LHCb and detection efficiencies are usually obtained from simulation. It is therefore important, in order to limit systematic uncertainties, to have a model that parametrises correctly the cross-sections of particles and antiparticles or at least their ratio.

The LHCb simulation software propagates particles though the detector using the GEANT4 toolkit [52]. This offers a variety of models for physics processes over a wide range of energies for both electromagnetic and strong interactions. Given a combination of projectile, target and energy there can be several models applicable with different reliability and computational costs. GEANT4 provides a number of pre-packaged physics lists each representing complete and consistent sets of models chosen to be appropriate for a given use case. In LHCb mainly two hadronic physics lists are considered:

- **LHEP** (Low and High Energy Parametrisation): based on a parametrised modelling of all hadronic interactions for all particles. This list combines the High Energy Parametrised model (HEP) and the low energy one (LEP). There is a sharp switch from the low to the high energy model at 25 GeV. The modelling of elastic scattering off a nucleus and of nuclear capture also proceeds via parametrised models.
- **FTFP_BERT**: includes the following models:
  - Bertini cascade model (BERT) [71], which simulates the intra-nuclear cascade, followed by pre-equilibrium and evaporation phases of the residual nucleus, for protons, neutrons, pions and kaons interaction with nuclei at kinetic energies below 9.9 GeV. The Bertini model produces more secondary neutrons and protons than the LEP model, yielding a better agreement with experiment data.
  - FTFP model, which implements high energy inelastic scattering of hadrons by nuclei using the FRITIOF model [72]. The change between the two models happens with a linear shift from BERT to FTF that starts at 4 GeV and ends at 5 GeV.

Figure 2.13 summarises the composition of the different models.

When two models overlap in an energy interval the choice of the model for each interaction is made using a random number: the probability to select each model varies linearly from 0 to 100% over the overlap range. Because of the differences of the two models in the overlap region, unphysical discontinuities can be produced as a function of energy.

2.11.1 Geometry and interaction probability

The results presented in the following sections are produced using the version v45r0 of the full LHCb framework for simulation, Gauss [66], interfaced to Geant4 v95r2p1. A simple geometry setup is used in order to be able to calculate in a clean way the interaction cross-sections in a specific material. This is constituted by a series of rectangular boxes filled with the most relevant materials for LHCb: Aluminium, Silicon and Beryllium. For each material three boxes are defined with different thicknesses (1mm, 10mm, 50mm). These values are chosen to be indicative of the amount of material present in the LHCb detector.

The simplest quantity available to extract the cross-section is the interaction probability (P_{int}), defined as:

$$P_{int} = \frac{N_{int}}{N_{tot}}, \tag{2.4}$$

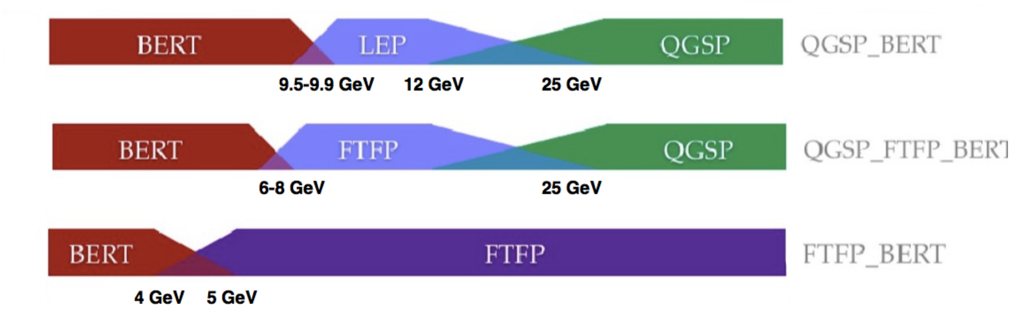

Figure 2.13: Diagram of LHEP, FTFP_BERT and QGSP_BERT models composition as a function of energy.

where N_{int} is the number of particles which interacted in the material and N_{tot} is the number of generated particles. As GEANT4 provides an ID for the end process of a particle (e.g. 121 for inelastic interaction, 111 for elastic, 201 for decay) it is possible to distinguish the inelastic and elastic probabilities of interaction and therefore cross-sections.

To compare simulation and data the cross-section and P_{int} are linked by the following formula valid for thin layers:

$$\sigma_{int} = \frac{A}{\rho N_A \Delta x} \cdot P_{int}, \tag{2.5}$$

where ρ is the density of the material and A is its mass number, Δx is the thickness of the considered layer and N_A is the Avogadro number.

2.11.2 PDG prediction

In the Review of Particle Physics [1] cross-sections of protons and neutrons are parametrised as:

$$\sigma_{tot}^{ab} = Z^{ab} + B^{ab}\log^2(s/s_M) + Y_1^{ab}(s_M/s)^{\eta_1} - Y_2^{ab}(s_M/s)^{\eta_2}, \tag{2.6}$$

$$\sigma_{tot}^{\bar{a}b} = Z^{ab} + B^{ab}\log^2(s/s_M) + Y_1^{ab}(s_M/s)^{\eta_1} + Y_2^{ab}(s_M/s)^{\eta_2}, \tag{2.7}$$

where $s_M = (m_a + m_b + M)^2$ and $B^{ab} = \lambda\pi(\frac{\hbar c}{M})^2$. Some of the constants in these equations are universal and valid for any kind of collision: M = 2.15, η_1 = 0.462, η_2 = 0.551, λ = 1 (for p, n and γ) and 1.63 (for d). The other ones are characteristic of each type of collision and are listed in Tab. 2.11.2. In these formulae the particle-antiparticle asymmetry arises from the last term which has opposite sign in the two equations. This term becomes less and less important with increasing energies. Therefore a net asymmetry is found at low energies, while the cross-sections tend to a common point at high energy and continue increasing logarithmically.

2.11.3 Validation results

This section reports particle and antiparticle cross-sections and their ratios compared, where available, with predictions and with data from the COMPASS experiment [73]. Figure 2.14 shows the probability of interaction for protons and anti-protons in 1mm of Aluminium using the FTFP_BERT and LHEP models compared with COMPASS data and Fig. 2.15 shows the ratios of $\sigma_{\bar{p}}^{tot}/\sigma_{p}^{tot}$ together with the PDG prediction. A difference of 40% is found between the two considered models for 1 GeV incoming anti-protons. This difference becomes negligible at higher energies. The discrepancies between the two physics lists for kaons and pions are of a few percents (2–3%) and usually constant with the energy. From the comparison with data and PDG predictions it can be qualitatively concluded that the FTFP_BERT model gives a better description of hadronic interactions at low energies, while both models give good results at high energy, above $\sim$ 10 GeV. The tool developed for this studies is not limited to cross-sections but can also give information on other simulated quantities. As an example, Fig. 2.11.3 shows a comparison between the types of particles generated in inelastic collisions of protons and anti-protons onto Aluminium using different models. Physics lists can give very different results, for example the LHEP model does not produce photons in inelastic collisions. However, it is difficult to use these quantities for validation as there is no data available for comparison.

Proj / Targ	Z^{ab}	Y_1^{ab}	Y_2^{ab}
$\bar{p}$,p / p	34.71	12.72	7.35
$\pi^{\pm}$ / p	19.02	9.22	1.75
$K^{\pm}$ / p	16.56	4.02	3.39
$K^{\pm}$ / n	16.49	3.44	1.82
$\bar{p}$,p / n	35.00	12.19	6.62

Table 2.1: Values for the constants Z^{ab}, Y_1^{ab} and Y_2^{ab} [1], which parametrise hadronic cross-sections.

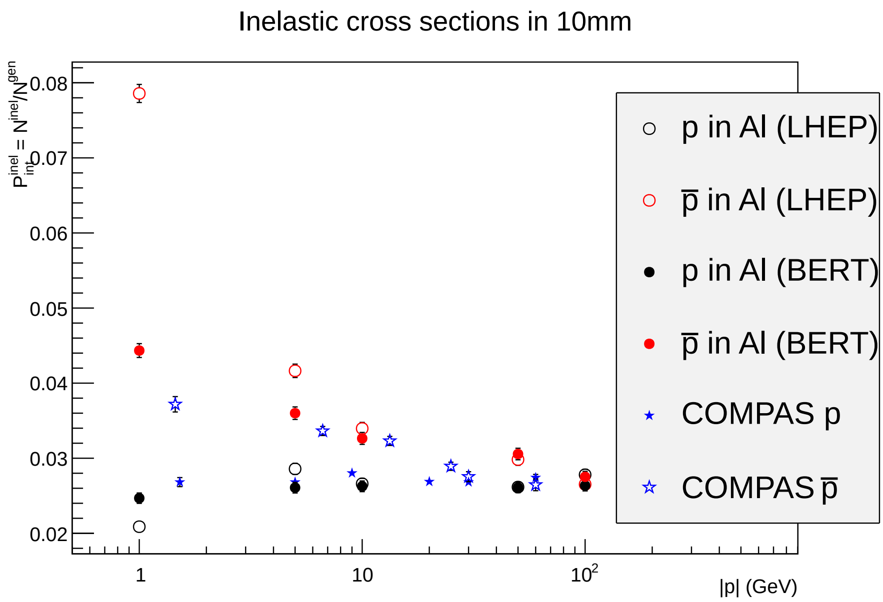

Figure 2.14: Probability of interaction for protons and anti-protons in Aluminium as a function of the projectile momentum. Two physics lists are used to generate events that can be compared with data from the COMPASS experiment.

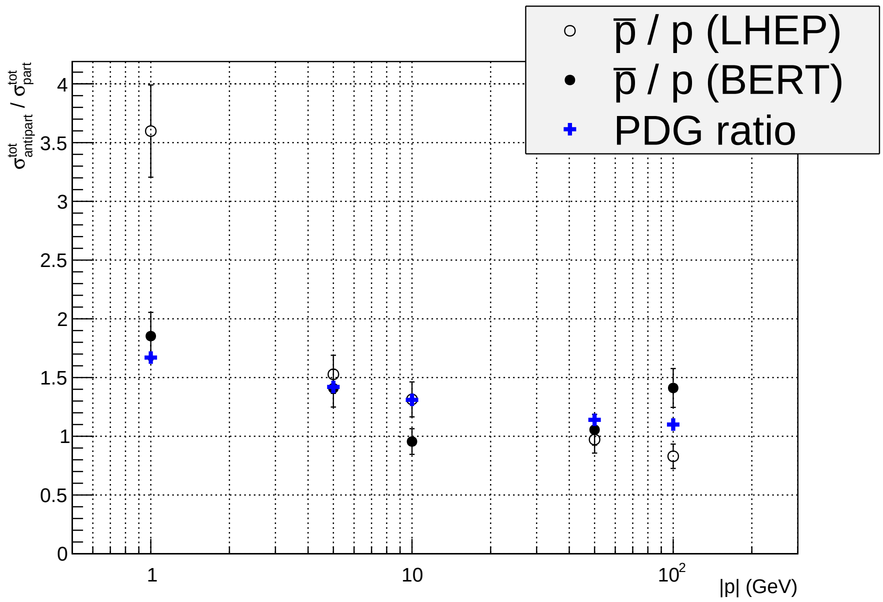

Figure 2.15: Ratio of antiproton over proton total interaction cross-section as a function of energy compared with PDG predictions.

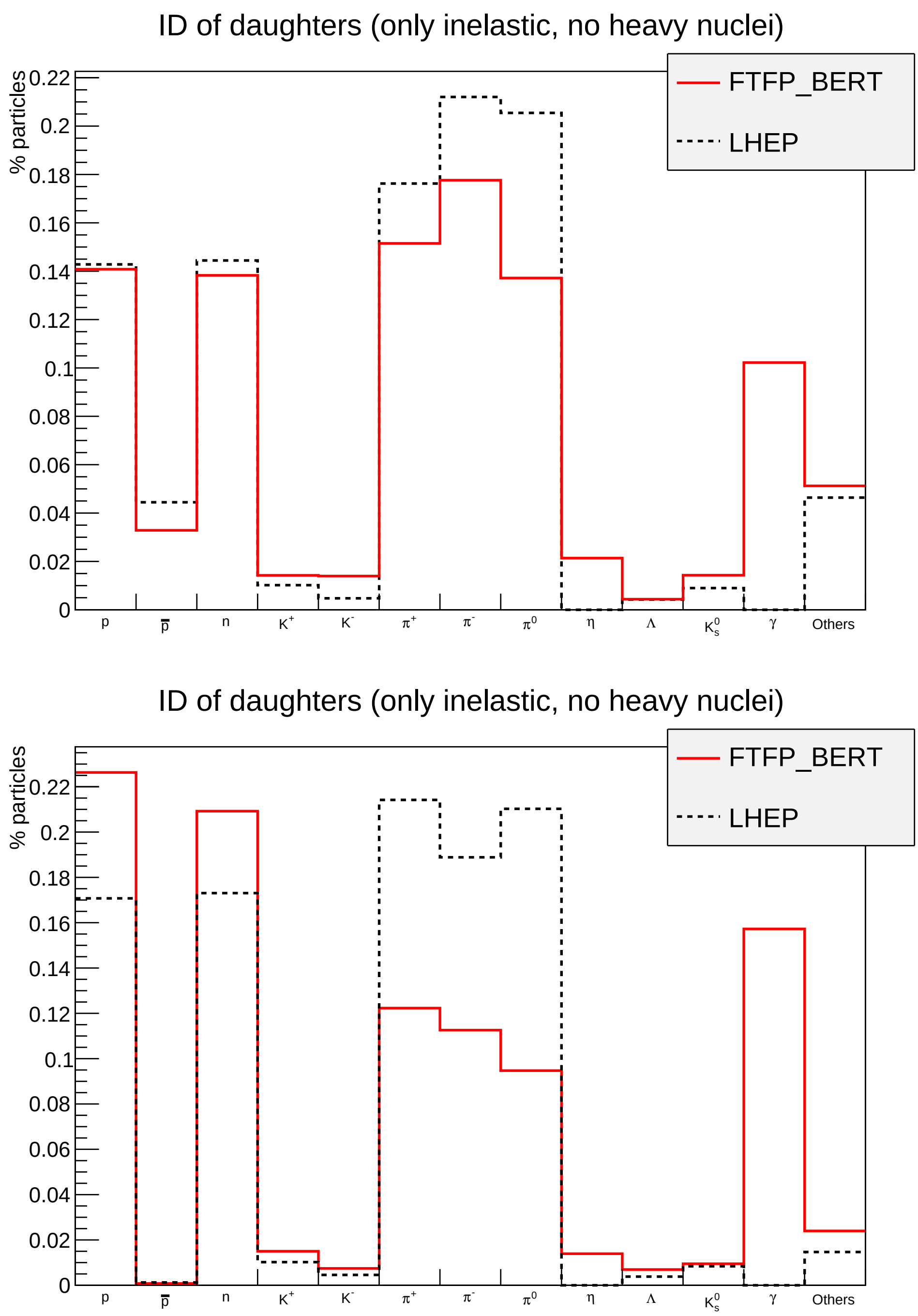

Figure 2.16: Composition of secondary particles produced in 100 GeV protons (top) and anti-protons (bottom) collisions in 1 mm Aluminium.

2.12 Material budget studies

It is important for many analysis to quantify the amount of material present in the detector, for example to estimate the amount of multiple scattering. In Geant4 particles are propagated in steps through the detector and for each step the framework analyses the geometry to understand in what material the particle is and modifies its trajectory accordingly. A tool was developed where neutrinos are used as probes to scan the detector summing the radiation length seen at each step up to a certain point.

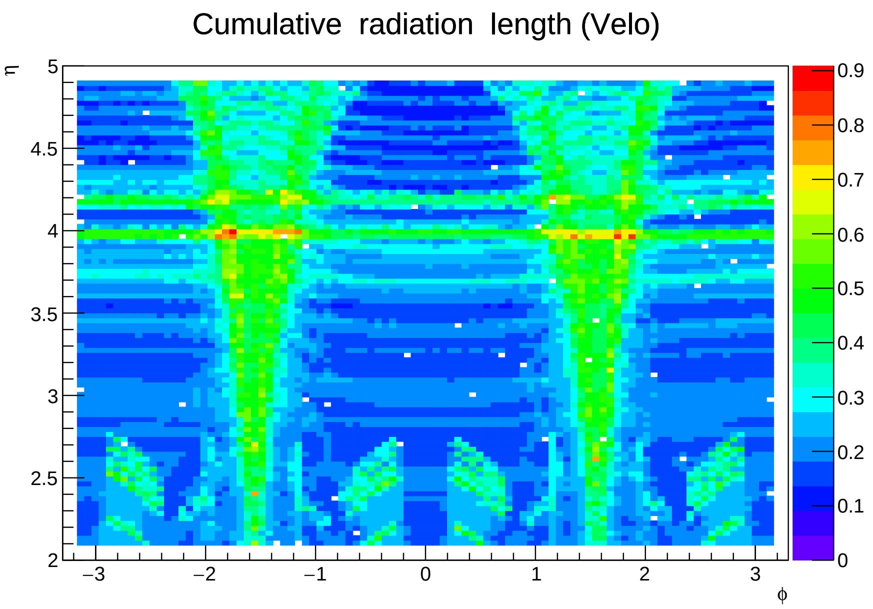

Figure 2.17: Map of cumulative radiation length traversed by a particle starting from the interaction point up to the end of the VeLo.

Neutrinos are used as they do not bend in magnetic field and do not interact with the detector to any appreciable extent. Thin air planes are inserted after each sub-detector. When these are traversed by the neutrinos, the information about the accumulated radiation and interaction length is saved. In this way it is possible to obtain maps of the detector, such as the one shown in Fig. 2.17. Using the tool developed for this study it is also possible to obtain the cumulative radiation and interaction lengths as a function of the position along the beam axis and the pseudorapidity. As an example Fig. 2.18 shows the average radiation length as a function of the distance from the interaction point. Furthermore, it is possible to displace the primary vertex from its position, normally set at the origin, in order to study how this translates into the amount of material traversed.

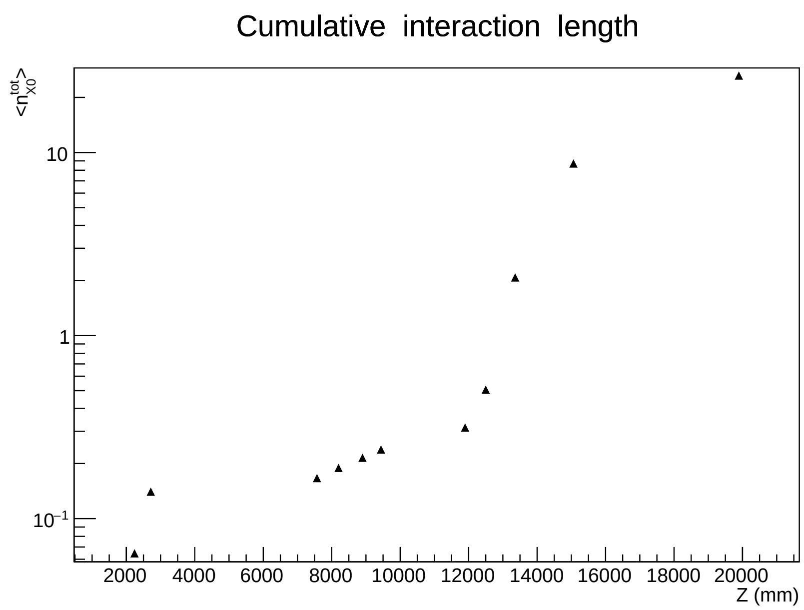

Figure 2.18: Average cumulative radiation length as a function of the horizontal distance from the interaction point. Each considered point corresponds to the end of a sub-detector: VeLo, RICH1, RICH2, tracking stations, ECAL and HCAL and muon detector.

2.13 Validation and material budget studies conclusions

The studies outlined in the previous two sections are based on tools which are now officially part of the LHCb simulation framework. These tools were used to validate the framework when passing from GEANT4 version 95 to version 96 and will continue to be used in the future. In particular a patch was provided by the GEANT4 team including improved kaon cross section. And it was verified these go in the right direction. Furthermore, the tools can be used by analyses sensitive to the quality of the simulation of particle and antiparticles cross section in order to study systematic effects and uncertainties.

CHAPTER 3

Differential branching fraction of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$

The rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay is a FCNC process governed by the $b \to s\mu^+\mu^-$ quark level transition. In the SM this decay proceeds only through loop diagrams, electroweak penguin and W box as discussed in Sec. 1.5 (see Fig. 1.5), and therefore it is highly sensitive to new particles entering the loops. Interest in Λ_b^0 baryon decays arises from two important facts. First of all, Λ_b^0 has non-zero initial spin, which allows us to learn information about the helicity structure of the underlying Hamiltonian that cannot be extracted from the meson decays [74, 75]. Secondly, the Λ_b^0 baryon can be considered in a first approximation as being composed of a heavy quark and a light di-quark, therefore the hadronic physics differs significantly from similar meson decays. This provides the possibility to better understand and test the hadronic physics in the theory, which could yield improved understanding that would be relevant also for the meson case.

With respect to B^0 decays going though the same transitions, such as $B^0 \to K^{*0}\mu^+\mu^-$, Λ_b^0 decays can provide independent confirmations of the results as they involve the same operators but different hadronic matrix elements. Furthermore, Λ baryons decays weakly, which results in complementary constraints with respect to B^0 decays. Finally, the narrow width approximation, used in theoretical calculations, is fully applicable in the Λ_b^0 case, which has $\Gamma_{\Lambda_b^0} \sim 2.5 \cdot 10^{-6}$ eV. This is not assured for $B^0 \to K^{*0}\mu^+\mu^-$ decays because the contribution from the non resonant channel $B^0 \to K\pi\mu^+\mu^-$ is unconstrained.

The theory of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays was considered by a number of authors both in the SM and in different new physics scenarios [76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86]. All authors start from the same effective Hamiltonian already described in Sec. 1.5.1. However, form factors, describing hadronic physics, are not developed as well as for the meson case because there are fewer experimental constraints. This leads to a relatively large spread in predicted branching fractions. For these reasons an interesting quantity to study is the differential branching fraction as a function of q^2. This still suffers from the knowledge of form factors but, as different approaches to form factors calculations are applicable in different q^2 regions, it allows a more meaningful comparison with theory.

Experimentally, the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ was observed for the first time in 2011 by the CDF collaboration [87], with a signal yield of 24±5 events and was later updated in preliminary form using their full statistics [88]. CDF observed the signal only in the q^2 region above the square of the $\psi(2S)$ mass. Their result on full statistics yields $\mathcal{B}(\Lambda_b^0 \to \Lambda\ \mu^+\mu^-\ ) = [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}$. Recently, the decay was also observed at LHCb [89] with a yield of 78 ± 12 signal events using 1 fb^{-1} of integrated luminosity collected in 2011. The signal was again found only in the high q^2 region, above $m^2_{\psi(2S)}$. The LHCb result for the branching fraction relative to the $J/\psi\Lambda$ decay, used as normalisation channel, is

$$\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)} = [1.54 \pm 0.30\ (\text{stat})\ \pm 0.20\ (\text{syst})\ \pm 0.02\ (\text{norm})] \times 10^{-3}$$

and for absolute branching fraction

$$\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-) = [0.96 \pm 0.16\ \text{(stat)}\ \pm 0.13\ \text{(syst)}\ \pm 0.21\ \text{(norm)}] \times 10^{-6}.$$

This chapter describes the measurement of the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay using 3 fb^{-1} of pp collisions collected by the LHCb experiment in 2011 and 2012. Furthermore, in the next chapter an angular analysis of these decays is performed for the first time, measuring observables including the forward-backward asymmetries in the leptonic and hadronic systems.

3.1 Analysis strategy and q^2 regions

A typical q^2 spectrum of $b \to s\ell^+\ell^-$ decays was shown in Fig. 1.8. This is characterised by the presence of the photon pole at low q^2 and the narrow peaks of the J/ψ and $\psi(2S)$ resonances at mid q^2. In the analysis $\Lambda_b^0 \to J/\psi\,\Lambda$ decays, where J/ψ decays into two muons have the same final state as the signal and are used as a normalisation channel. The rare and normalisation channels are naturally distinguished by the q^2 intervals in which they are reconstructed. The Λ decay mode into a pion and a proton, $\Lambda \to p\pi$, is always used to reconstruct the decays. The intervals in which the rare channel is studied are:

- $0.1 < q^2 < 8$ GeV$^2/c^4$, where the signal is unobserved and the selection is optimised to observe the signal as explained in Sec. 3.4.3. The upper bound of this interval is chosen to be sufficiently far from the J/ψ radiative tail at low masses, that could contaminate the rare sample;

- $11 < q^2 < 12.5$ GeV$^2/c^4$ in between two charmonium resonances, and

- $q^2 > 15$ GeV$^2/c^4$, above $\psi(2S)$.

The first interval is referred to as “low q^2 ” region, below the J/ψ resonance ($q^2 < 8$ GeV^2/c^4), and the other two as “high q^2 ” region, above the J/ψ resonance ($q^2 > 11$ GeV^2/c^4). The above regions are then sub-divided into smaller intervals, as the available statistics allows, which results in ~ 2 GeV^2/c^4 wide bins. The binning used is the following:

$$[0.1, 2.0, 4.0, 6.0, 8.0], J/\psi, [11.0, 12.5], \psi(2S), [15.0, 16.0, 18.0, 20.0]. \tag{3.1}$$

In addition the result is also provided in two integrated regions:

- 1.1-6.0 GeV^2/c^4: this interval is theoretically clean since it is far from the photon pole, which dominates at low q^2 washing out the sensitivity to new physics contributions. The lower bound of this interval it chosen to exclude the possible contribution from from the ϕ resonance, which appears at 1 GeV^2/c^4. The upper bound of the interval is chosen to totally exclude a small contribution from the J/ψ resonance that leaks below 8 GeV^2/c^4.
- 15.0-20.0 GeV^2/c^4: this interval is the one that contains most of the statistics and it is used as a natural cross check that the analysis in smaller bins is stable.

3.2 Candidate types

This analysis deals with Λ baryons, which have a lifetime of $(2.632 \pm 0.020) \times 10^{-10}$ s [1]. These are considered long-lived particles in particle physics terms and can travel into the detector for several metres generating well distinguished secondary vertices. In LHCb, Λ baryons can be reconstructed from tracks with or without hits in the VeLo (see Sec. 2.4) and therefore two candidates types are defined as follows:

- **Long candidates**: built from tracks which have hits in the VeLo, “long tracks”. These candidates, also denoted as “LL”, are characterised by a better momentum resolution thanks to the longer lever arm available to their tracks.

- **Downstream candidates**: built from tracks without hits in the VeLo, "downstream tracks", also denoted as "DD".

Figure 3.1 shows a depiction of the two types of candidates used in the analysis together with other possible track types in LHCb, which are not used in this analysis. As the long and downstream candidate categories are characterised by different resolution and different kinematic properties the analysis is performed separately on the two samples and the results are then combined.

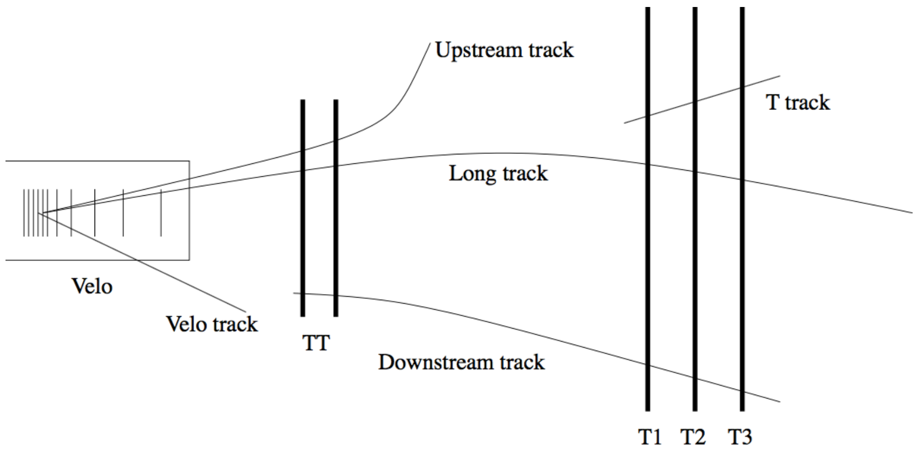

Figure 3.1: Representation of possible track types in LHCb. Candidates built from "long" and "downstream" tracks are used in this analysis [52].

3.3 Simulation

Samples of simulated events are needed in order to train the multivariate classifier (see Sec. 3.4.2), calculate the selection efficiency and study possible backgrounds; in particular for this analysis samples of ~ 2 millions $\Lambda_b^0 \to J/\psi\,\Lambda$ and ~ 5 millions $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulated events are used. Samples of simulated $B^0 \to J/\psi\,K_S^0$, $B^0 \to K_S^0\mu^+\mu^-$ and $B^+ \to \mu^+\mu^-K^{*+}$ events are also used to study backgrounds from these decays. The events are generated using Pythia8; hadronic particle are decayed using EvtGen and Geant4 is used to simulate the interaction of final state particles with the detector. Simulated events are then reconstructed by the same reconstruction software that is used for real data. The L0 hardware trigger is emulated in the simulation, while for the software stage, HLT (see Sec. 2.9), the same code can be used as for data. Events are simulated using both 2011 and 2012 beam and detector conditions in the same proportion as data are available. While the simulation gives a generally good description of data some discrepancies remain. However, it is important that the simulation gives an accurate description of the data, especially for quantitative estimations, e.g the extraction of efficiencies. The next sections describe corrections applied to the simulation in order to provide a better description of data. In Appendix B data distributions are compared with simulated ones for variables relevant to this analysis.

3.3.1 Decay Model

Little is known about Λ_b^0 decays structure and therefore the simulation software generates events according to the phase space given by the available kinematics. To include a reasonably realistic q^2 dependence, the simulation is weighted using decay amplitudes based on the predictions in Ref. [90]. Equations in this paper are for the case of unpolarised Λ_b^0 production and for this analysis those are extended to include polarisation. Details about the models used are in Appendix A.1. The value of the Λ_b^0 production polarisation, P_b, used in the calculations is $P_b = 0.06$ as measured by LHCb [91]. Figure 3.2 shows the phase space q^2 distribution and the one obtained by re-weighting the events. The latter can be qualitatively compared to the q^2 spectrum of a generic $b \to s\ell^+\ell^-$ decay shown in Fig. 1.8. For the normalisation mode, the decay model used is described in Appendix A.3, with amplitude magnitudes and production polarisation taken from the measurements in Ref. [91]. Phases are not yet measured and are all set to zero.

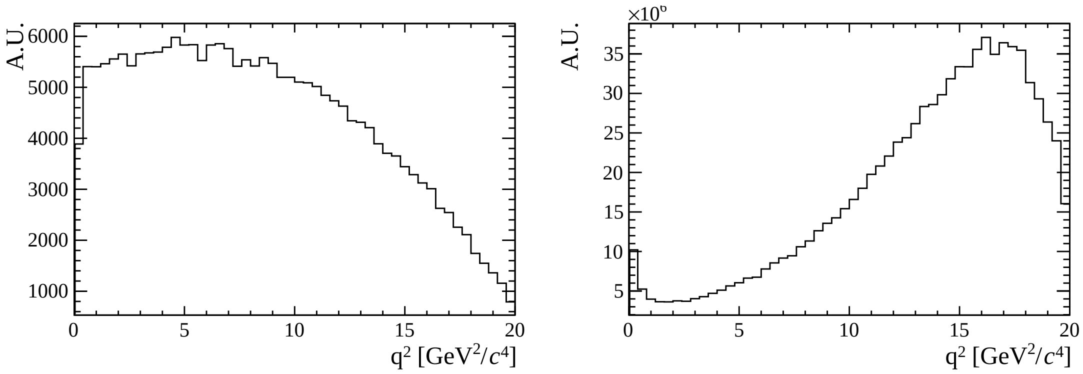

Figure 3.2: The q^2 spectrum of $\mathit{\Lambda}_b^0 \to \mathit{\Lambda}\mu^+\mu^-$ simulates events according to the phase space of the decay (left) and re-weighted using the decay amplitudes (right).

3.3.2 Kinematic re-weighting

Small data-simulation differences are found in the kinematic properties of the mother particle, $\mathit{\Lambda}_b^0$, which also affect the final state particles. The simulation is re-weighted by comparing the momentum and transverse momentum of $\mathit{\Lambda}_b^0$ in real and simulated $\mathit{\Lambda}_b^0 \to J/\psi\,\mathit{\Lambda}$ candidates that satisfy the pre-selection (see Sec. 3.4). To do this a high purity data sample is obtained by selecting a narrow interval around the J/ψ and $\mathit{\Lambda}_b^0$ peaks; this contains about 400.000 candidates. Then the $\mathit{\Lambda}_b^0$ invariant mass distribution is fitted to extract the amount of background under the peak. The background fraction, $f_b = B/(S+B)$, is then used to subtract statistically the background from the kinematical distributions as described by the equation:

$$S(p, p_{\mathrm{T}}) = T(p, p_{\mathrm{T}}) - f_b \cdot B(p, p_{\mathrm{T}}), \tag{3.2}$$

where $S(p, p_{\mathrm{T}})$ is the distribution of pure signal events, which we want to obtain, $T(p, p_{\mathrm{T}})$ is the total distribution of signal plus background, namely the distribution of all events in the signal interval, $5605 < m(p\pi\mu^+\mu^-) < 5635$ MeV$/c^2$, and $B(p, p_{\mathrm{T}})$ is the pure background distribution obtained using events from the upper sideband, $m(p\pi\mu^+\mu^-) > 5800$ MeV$/c^2$.

After obtaining the signal distributions from data, these are compared with $\mathit{\Lambda}_b^0 \to J/\psi\,\mathit{\Lambda}$ simulated events and a weight, $w(p_{\mathit{\Lambda}_b^0}, p_{\mathrm{T}\,\mathit{\Lambda}_b^0})$ is defined by taking the ratio of the two

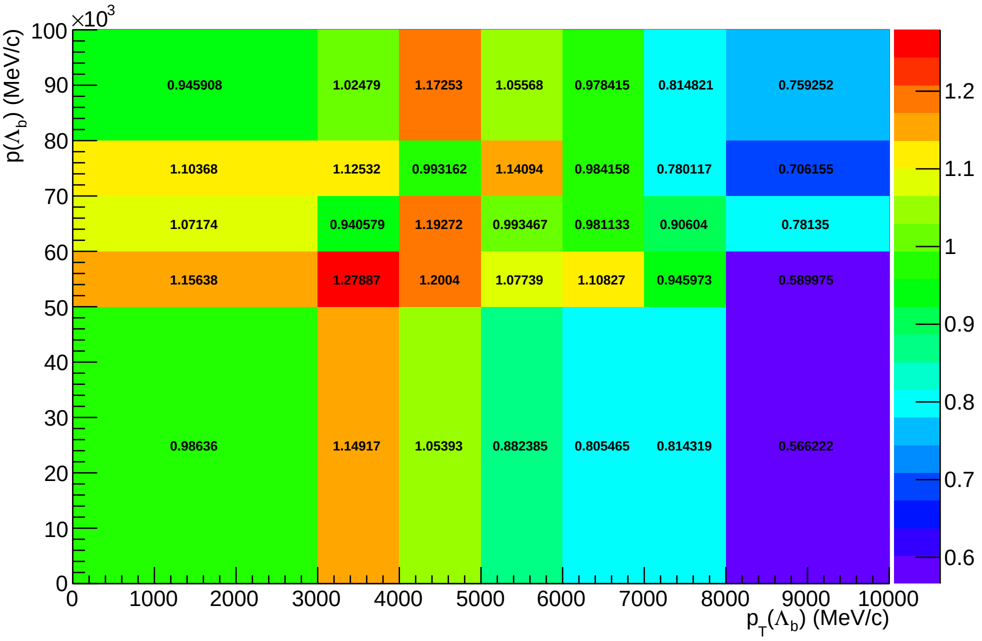

Figure 3.3: Weights used for the kinematical re-weighting as a function of the momentum and transverse momentum of Λ_b^0.

dimensional (p, p_{T}) distributions. The result is shown in Fig. 3.3, while Appendix B reports distributions of sideband subtracted data in the signal and sideband regions together with weighted and unweighted simulated events. In these plots the momentum and p_{T} distributions of Λ_b^0 match by construction but the re-weighting also improves the agreement between the kinematical distributions of all final particles. Small differences remain due to the finite binning used for the weights calculation. Quality variables, such as the χ^2 of tracks and vertices, show little dependence on the kinematics and are relatively unaffected by the weighting procedure.

3.3.3 Event type

The fraction of Λ baryons reconstructed from long tracks and downstream tracks does not fully agree between data and simulation. For $\Lambda_b^0 \to J/\psi\,\Lambda$ decays passing the full selection, $\sim 70\%$ of candidates are reconstructed from downstream tracks, compared with $\sim 75\%$ in the simulation. The fraction of downstream and long tracks also varies as a function of q^2 and the biggest differences are found at low

q^2. In order to deal with these differences all efficiencies are obtained separately for downstream and long candidates and the analysis is carried out separately for the two categories, results are then joined to ensure the best use of the available information. It is therefore not necessary to correct the simulation to reproduce the correct fraction of events in each category.

3.4 Selection

This section described the requirements applied to reconstruct $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates The selection procedure is divided into two steps: a pre-selection, where cuts are applied in order to be able to work with manageable datasets and a multivariate analysis (MVA) which combines information from several variables. At first good tracks are selected using requirements on basic kinematic properties, as the p_{T} of the final particles, and quality requirements, as the track χ^2. The selection then aims to first form a dimuon candidate from two oppositely changed muons. Then, in events containing a dimuon candidate, two oppositely charged tracks are combined and retained as Λ candidate if they form a good vertex which is well separated from all primary vertices. Finally the dimuon and Λ candidates are combined to form Λ_b^0 baryons and requirements are set on the properties of this combination.

3.4.1 Pre-selection

The full list of pre-selection cuts is reported in Tab. 3.1. In the table χ^2_{IP} is defined as the projected distance from a vertex divided by its uncertainty, for example the $\chi^2_{\mathrm{IP}}(primary) > n$ requirement on Λ_b^0 means that the Λ_b^0 vertex must be at least $\sqrt{n}$ standard deviations away from the primary vertex. Another quantity, especially useful to remove combinatorial background, is a pointing variable called DIRA defined as the cosine of the angle between the direction of a particle's momentum and the flight direction from its mother vertex. Requiring a DIRA close to unity corresponds to the selection of particles with well-defined origin vertices. Graphical representation of the χ^2_{IP} and DIRA variables are shown in Fig. 3.4. The variable

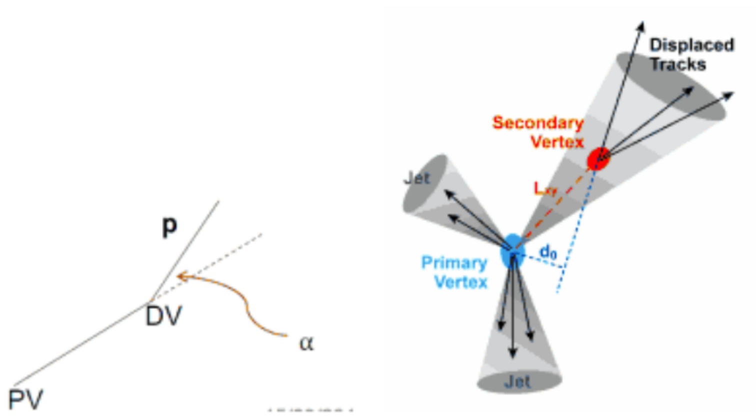

Figure 3.4: Graphical representation of the DIRA (left) and χ^2_{IP} (right) variables.

χ^2_{FD} represents the flight distance of a particle from its origin vertex divided by its uncertainty. The χ^2_{trk}/ndf and χ^2_{vtx}/ndf quantities are the χ^2 from the fit to tracks and vertices, which are used to quantify their quality. The `GhostProb` quantity describes the probability of a track being fake. By construction, cutting at k, removes $(1-k)\cdot 100\%$ of fake tracks. The `hasRich`, `hasCalo` and `isMuon` variables are binary indicators that the information from the RICH/calorimeter/muon detector is available for the track. Loose PID requirements on the proton are also applied in pre-selection. Details about the quantification of the PID quality are given in Sec. 2.8. A large mass window is kept around the Λ_b^0 peak in order to be able to fit the sideband, to train the multivariate analysis and to better constrain backgrounds. Rare candidates are selected by the q^2 region requirements described in Sec. 3.1, while resonant candidates are further constrained to have dimuon invariant mass in a 100 MeV/c^2 interval around the known J/ψ mass [1].

3.4.2 Neural Networks

The final selection is performed using a neural network (NN) classifier based on the NeuroBayes package [68, 69]. The input to the neural network consists of 14 variables

Particle	Requirement
Λ_b^0	$4.6 < m(p\pi\mu\mu) < 7.0$ GeV/c^2
	`DIRA`> 0.9999
	$\chi^2_{\text{IP}} < 16.0$
	$\chi^2_{\text{FD}} > 121.0$
	$\chi^2_{vtx}/\text{ndf} < 8.0$
Λ	$\chi^2_{vtx}/\text{ndf} < 30.0(25.0)$
	Decay time > 2 ps
	$\lvert m(p\pi) - m_{\Lambda}^{\text{PDG}}\rvert < 35(64)$ GeV/c
p/π	$p > 2$ GeV/c
	$p_{\text{T}} > 250$ MeV/c
	$\chi^2_{\text{IP}} > 9(4)$
p (only long cand.)	`hasRICH`
	`PIDp`> -5
μ	`isMuon`
	$\chi^2_{trk}/\text{ndf} < 5$
	`GhostProb`< 0.4
	`PID`$\mu > -3$
	$\chi^2_{\text{IP}} > 9.0$
Dimuon	$\chi^2_{vtx}/\text{ndf} < 12.0$
	$m(\mu\mu) < 7.1$ GeV/c^2

Table 3.1: Summary of pre-selection requirements. Where two values are given, the main one applies to long candidates and the one in parenthesis to downstream candidates.

carrying information about the kinematics of the decay, the quality of tracks and vertices and the PID of the muons. The list of the 10 most significant inputs is reported in Tab. 3.2, together with information about the importance of each input. Variables related to Λ and its daughters are considered as different inputs depending on the candidate type (long or downstream). This effectively corresponds to making a separate training for the two categories.

The NN is trained using representative samples of signal and background. A sample of simulated $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ events is used as a proxy for the signal, while for the background a representative sample is given by candidates in the upper $m(p\pi\mu\mu)$ invariant mass sideband. Only the upper sideband, $m(p\pi\mu\mu) > 6$ GeV/c^2, is used since it contains only combinatorial background, while the lower sideband may contain partially reconstructed and misreconstructed candidates. In the q^2 spectrum of background samples the J/ψ and $\psi(2S)$ peaks are still present indicating that charmonium resonances are often combined with other random tracks. These candidates do not give a good description of purely combinatorial background and, in order to avoid biases, they are removed from the training sample by rejecting events in a 100 MeV$/c^2$ interval around the nominal J/ψ and $\psi(2S)$ masses [1]. A total of 30000 total events is used for the training from each sample. This corresponds $\sim$ 50% of the available sideband data sample and $\sim$ 20% of the simulated sample. The full simulated sample is not used as the same sample will also be used to study efficiencies. For reproducibility the events are sampled uniformly.

The single most important variable used for downstream candidates is the transverse momentum of Λ, which allows to reject random combination of tracks as these have preferentially low p_{T}. For long candidates instead the best variable is the χ^2 from a kinematic fit that constrains the decay products of the Λ_b^0, the Λ and the dimuon, to originate from their respective vertices performed using the `DecayTreeFitter` tool (see Sec. 2.10). Other variables that contribute significantly are the χ^2_{IP} of Λ_b^0, Λ and muons, the separation between the Λ_b^0 and Λ vertices and, finally, the muon PID.

Figure 3.5 shows distributions of neural network output for the signal and background samples and purity, $P = N_{\mathrm{sig}}/N_{\mathrm{bkg}}$, as a function of the neural network output. The distributions from test samples are also overlaid in order to check for overtraining. The distributions follow the same shape but with different fluctuations indicating no significant overtraining. In general it can be concluded that the neural network is able to separate signal from background and the training converged properly. It can happen that too much information is given to the classifier, which becomes able to calculate the invariant mass of the candidates from the input variables. This can generate fake peaks and it is therefore important to check for correlations between the 4-body invariant mass and the NN output. Figure 3.6 reports the average neural network output as a function of the 4-body $m(p\pi\mu\mu)$ invariant mass for data and simulation. The distributions are flat indicating that no significant correlation is present.

Table 3.2: Summary of the 10 most significant inputs to the neural network in order of importance. Column "adds" gives the significance added by a given input when it is added to the list of those ranked above. Column "only this" provides the power of a given input alone and "loss" shows how much information is lost when removing only a given input.

Input	**adds**	**only this**	**loss**
Λ DD p_{T}	143.11	143.11	29.20
χ^2_{DTF}	77.81	134.00	51.10
$\min(\chi^2_{\mathrm{IP}}\ \mu)$	61.31	113.62	29.76
$\chi^2_{\mathrm{IP}}\ \Lambda_b^0$	52.94	113.23	40.98
$\chi^2_{\mathrm{IP}}\ \pi_{\mathrm{LL}}$	20.29	60.72	12.82
min(PID μ)	17.91	59.11	13.44
$\tau_{\Lambda_b^0}$	16.24	35.36	11.24
Λ_b^0 `DIRA`	12.28	73.96	9.98
Λ DD flight distance	9.47	86.75	11.24
$\chi^2_{\mathrm{IP}}\ \Lambda$ DD	10.58	59.84	8.88

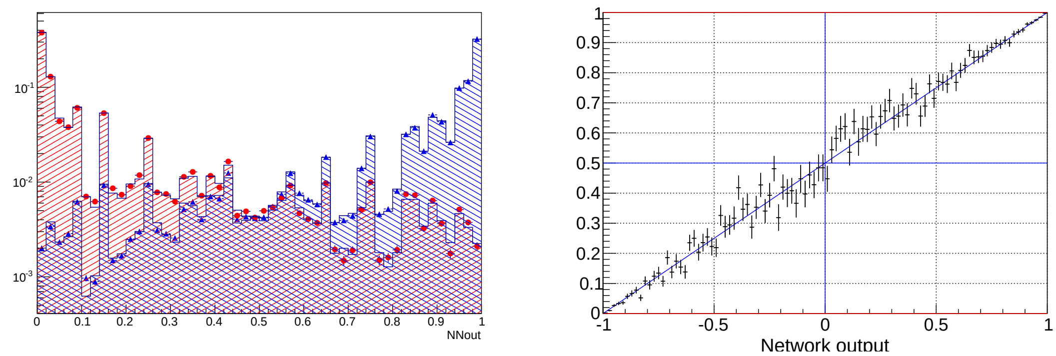

Figure 3.5: (left) Neural network output distribution for training (points) and test (stripes) samples, for signal and background events. (right) Purity as a function of neural network output.

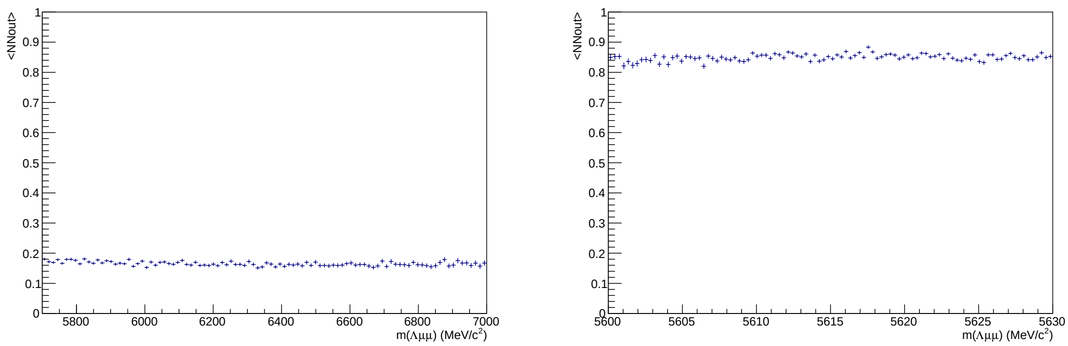

Figure 3.6: Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and simulated signal (right) events.

3.4.3 MVA optimisation

In the high q^2 region, where the signal is already observed, the requirement on the neural network output is chosen maximising the significance, $N_{\mathrm{S}}/\sqrt{N_{\mathrm{S}}+N_{\mathrm{B}}}$, where N_{S} and N_{B} are the numbers of expected signal and background candidates respectively. N_{S} is derived from simulation but, as an arbitrary number of events can be generated, it needs to be normalised. To do this, the invariant mass distribution of real $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates is fit after pre-selection (including all requirements but MVA). This is possible as the peak of the resonant channel is already well visible before the MVA cut. The resonant yield is then scaled by the ratio of between the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ branching fractions as measured by LHCb on 2011 data

$$\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda) = 1.54 \times 10^{-3} \tag{3.3}$$

and by the $J/\psi \to \mu^+\mu^-$ branching fraction. In summary:

$$N_{\mathrm{S}} = N_{J/\psi} \cdot \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda)\cdot\mathcal{B}(J/\psi \to \mu^+\mu^-)}. \tag{3.4}$$

The number of expected background events instead is derived fitting the data side-band with an exponential and extrapolating under the signal region.

In the low q^2 region, where the signal is unobserved, the so called "Punzi figure-of-merit", $N_{\mathrm{S}}/(n_\sigma/2+\sqrt{N_{\mathrm{B}}})$, is maximised [92]. This figure-of-merit is considered to be optimal for discovery and the parameter n_σ corresponds to the number of expected standard deviations of significance, in this analysis $n_\sigma = 3$ is used. Moreover, the Punzi shape does not depend on the relative normalisation between signal and background, which is important since the signal is still unobserved at low q^2 and the existing predictions vary significantly for this region. The dependence of the figure-of-merit for both q^2 regions is shown in Fig. 3.7, and curves of signal efficiency versus background rejection are shown in Fig. 3.8.

For final selection the neural network output is required to be larger than 0.76 for candidates in the high q^2 region and 0.97 for the low q^2 ones. Using these

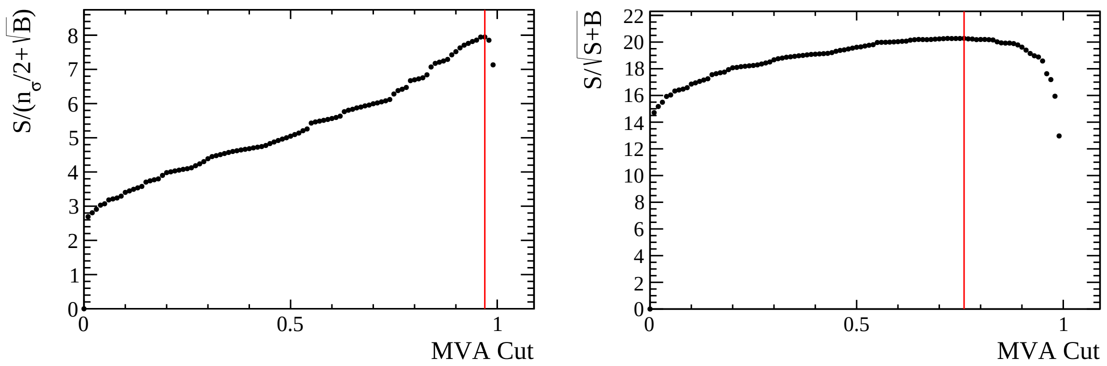

Figure 3.7: Dependence of the figure-of-merits on the neural network output requirement for the low q^2 (left) and high q^2 (right) regions. The vertical lines correspond to the chosen cuts.

requirements the neural network retains approximately 96% (66 %) of downstream candidates and 97 % (82 %) of long candidates for the high (low) q^2 selection, with respect to the pre-selected samples. After full selection $\sim$ 0.5% of the events contain multiple candidates which are randomly rejected keeping only one candidate per event. To normalise the branching ratio measurement J/ψ events are selected using the low and high q^2 requirements to normalise respectively low and high q^2 intervals.

3.4.4 Trigger

Finally, specific trigger lines are selected, corresponding to events triggered by muons which formed the reconstructed candidate. This is denoted as Trigger On Signal (TOS). The trigger lines used in the analysis are listed in Tab. 3.3. The logical *or* of the lines on the same lever is required and the logical *and* of those on different levels. The `L0Muon` trigger requires hits in the muon detector and triggers if a muon with $p_{\mathrm{T}} > 1.5$ GeV/c is identified. `L0Dimuon` imposes the same requirement on the sum of the transverse momenta of two tracks. The `Hlt1TrackAllL0` performs a partial reconstruction of the events and applies basic requirements on the IP, χ^2 and p_{T} of tracks; it triggers if the L0 decision is confirmed. `Hlt1TrackMuon` applies looser requirements but in addition requires the `isMuon` variable (see Sec. 2.8) to be

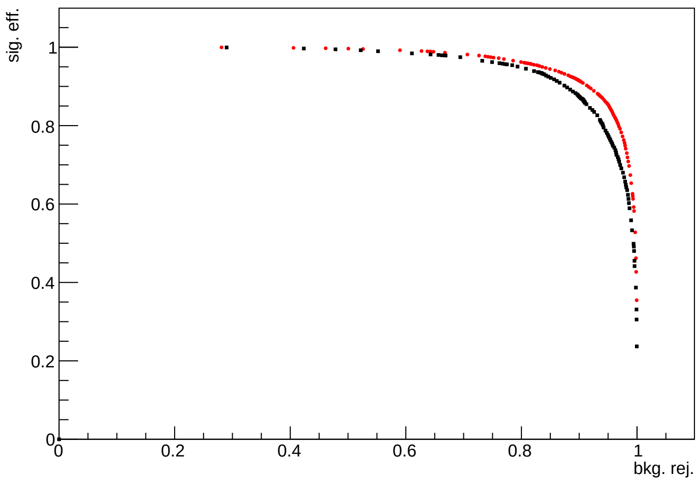

Figure 3.8: Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection. The optimal points on these curves are the closest ones to (1,1).

true to limit the yield. Finally, at the Hlt2 level, a complete reconstruction is done and a multivariate analysis is used to identify decay structures. One of the main variables used at this stage is the Distance Of Closest Approach (DOCA), which is required to be less than 0.2 mm to form a 2-body object.

Table 3.3: Summary of trigger lines which candidates have to pass at various trigger levels. Trigger is always required to be due to tracks of the candidate itself.

Trigger Level	**Lines**
L0	`L0Muon`
	`L0DiMuon`
Hlt1	`Hlt1TrackAllL0`
	`Hlt1TrackMuon`
Hlt2	`Hlt2Topo[2-4]BodyBBDT`
	`Hlt2TopoMu[2-4]BodyBBDT`
	`Hlt2SingleMuon`
	`Hlt2DiMuonDetached`

3.4.5 Background from specific decays

Candidates from other decays can be reconstructed as the decays of interest if particles are not reconstructed of mis-identified. A survey of possible backgrounds concluded that the only physics background to take into account comes from misreconstructed decays of B^0 to K^0_S with two muons in the final state, whether via J/ψ or not, where the K^0_S is reconstructed as a Λ with a $p \to \pi$ identity swap. The lack of background from other decays is mainly due to the particular topology of the Λ decay, which is long-lived and decays at a displaced vertex. To study the effect of misreconstructed $B^0 \to J/\psi K^0_S$ and $B^0 \to K^0_S\mu^+\mu^-$ decays simulated samples are used. On data the $B^0 \to J/\psi K^0_S$ contribution is clearly visible in the resonant channel mass distribution. This background is not suppressed with specific cuts in this analysis as its mass shape is sufficiently distinct the from Λ_b^0 signal and its contribution can be reliably modelled in the mass fits (see Sec. 3.5.1). For the rare case a rough estimate of the K^0_S background size is obtained using the yield in the resonant channel rescaled by the measured ratio between the rare and resonant branching fractions. Details are given in Sec. 3.5.1 and numbers of events predicted are reported in Tab. 3.4. This contribution, although close to negligible is again considered in the fit. A possible pollution due to $B^+ \to \mu^+\mu^- K^{*+}$ decays, where the K^{*+} further decays into $K^0_S\pi$ is also investigated using a dedicated simulated sample and found to be negligible. Finally, $\Lambda_b^0 \to J/\psi\ \Lambda$ events radiating photons from the final state, can escape the J/ψ veto and be reconstructed in the rare channel sample. Analysing simulated events it was found that the only contribution is in the closest q^2 interval to the J/ψ tail, $6 < q^2 < 8\ \mathrm{GeV}^2/c^4$. In this interval 1.3% of the $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates are reconstructed but only 0.06% fall into the 4-body invariant mass window used for the fits. This corresponds to ~ 6 events, 4 of which in the downstream category. Given the low yield and that these events do not peak under the signal but show a decaying distribution at the edge of the fit mass window, this background is considered as absorbed in the combinatorial background. Figure 3.9 shows the invariant mass distribution of simulated $\Lambda_b^0 \to J/\psi\,\Lambda$ events falling into the rare q^2 region and the distribution of simulated $B^+ \to \mu^+\mu^- K^{*+}$ events mis-reconstructed as $\Lambda_b^0 \to J/\psi\,\Lambda$ decays.

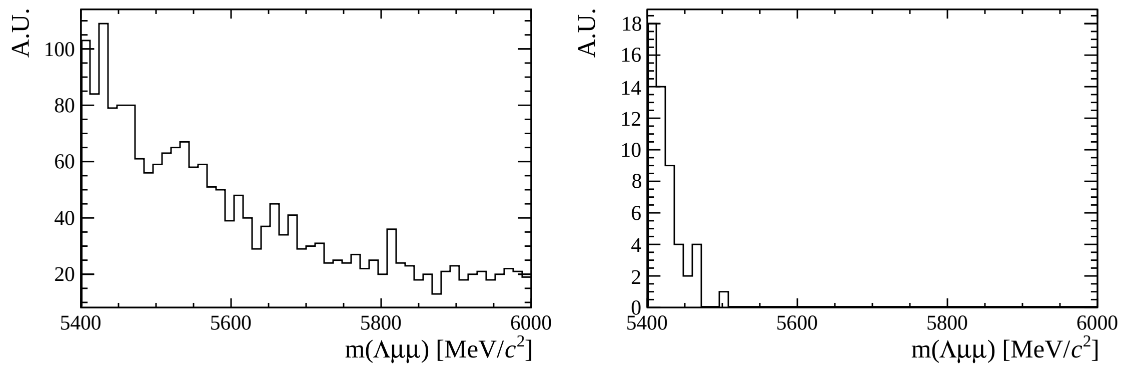

Figure 3.9: Invariant mass distributions of simulated $B^+ \to \mu^+\mu^- K^{*+}$ (left) and $\Lambda_b^0 \to J/\psi\ \Lambda$ (right) candidates passing the full selection. Only $\Lambda_b^0 \to J/\psi\ \Lambda$ candidates reconstructed in $q^2 < 8\ \mathrm{GeV}^2/c^4$ are selected. Distributions are shown in the invariant mass range relevant for the analysis (see Sec. 3.5.1).

3.5 Yield extraction

Extended unbinned maximum likelihood fits are used to extract the yields of the rare and resonant channels. The likelihood has the form:

$$\mathcal{L} = e^{-(N_\mathrm{S}+N_\mathrm{C}+N_\mathrm{B})} \times \prod_{i=1}^{N} \left[N_\mathrm{S} P_\mathrm{S}(m_i) + N_\mathrm{C} P_\mathrm{C}(m_i) + N_\mathrm{B} P_\mathrm{B}(m_i) \right] \tag{3.5}$$

where N_S, N_C and N_B are respectively the numbers of signal, combinatorial and K^0_S background events and the $P_i(m_i)$ are the corresponding probability density functions (PDF). The fit variable is the 4-body $m(p\pi\mu\mu)$ invariant mass obtained from a kinematical fit of the full decay chain in which each particle is constrained to point to its assigned origin vertex and the invariant mass of the $p\pi$ system is constrained to be equal to the world average for the Λ baryon mass. In the resonant case a further constrain is used on the dimuon mass to be equal to the known J/ψ mass. This method allows to improve the mass resolution giving better defined peaks and therefore a more stable fit. For brevity, in the following these variables are simply referred to as “invariant mass”.

3.5.1 Fit description

The fit is performed though the following steps:

- simulated distributions are fit to extract initial parameters;
- the resonant data sample is fitted;
- the rare sample is fitted fixing some parameters to those obtained in the previous cases.

In the first step simulated $\Lambda_b^0 \to J/\psi\,\Lambda$ distributions are fitted using the signal PDF alone. This is done separately for long and downstream candidates. Figure 3.10 shows distributions of candidates selected in the resonant sample with the fit function overlaid. The signal is described as the sum of two Crystal Ball functions (CB) with common mean (m_0) and tail slope (n). This is also known as Double Crystal Ball (DCB) function. A single Crystal Ball [93] is a probability density function commonly used to model processes involving energy loss. In particular it is used to describe resonances' peaks with radiative tails. This function consists of a Gaussian core and a power-law tail below a certain threshold and has form

$$C(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} exp\left(-\frac{(x-\bar{x})^2}{2\sigma}\right) & \text{if } \frac{(x-\bar{x})}{\sigma} > \alpha, \\ A\left(B - \frac{(x-\bar{x})}{\sigma}\right)^{-n} & \text{if } \frac{(x-\bar{x})}{\sigma} < \alpha, \end{cases} \tag{3.6}$$

where for normalisation and continuity

$$\begin{aligned} A &= \left(\frac{c}{|\alpha|}\right))^n \cdot exp(-\frac{\alpha^2}{2}), \\ B &= \frac{n}{|\alpha|} - |\alpha|. \end{aligned} \tag{3.7}$$

The full PDF for the resonant channel is therefore:

$$P_{\mathrm{S}}(m;m_0,\alpha_1,\alpha_2,f,n) = f\mathrm{CB}(m;m_0,\sigma_1,\alpha_1,n) + (1-f)\mathrm{CB}(m;m_0,\sigma_2,\alpha_2,n), \tag{3.8}$$

where f is the relative fraction of candidates falling into the first CB function.

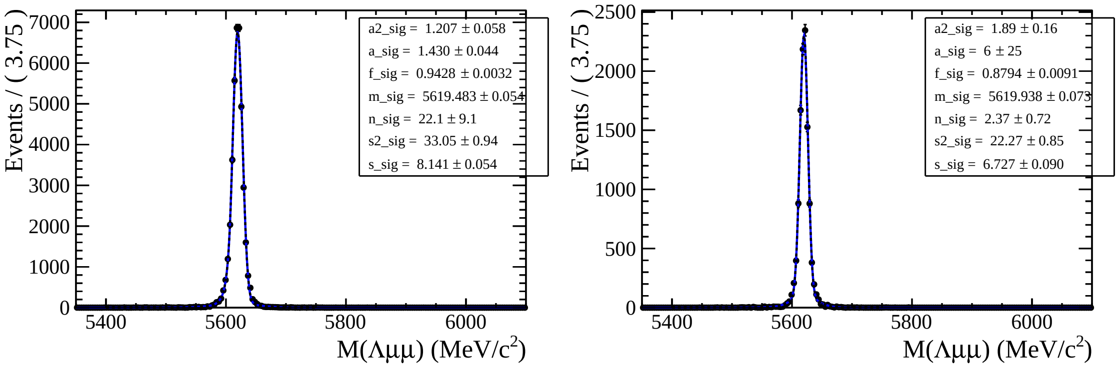

Figure 3.10: Invariant mass distribution of $\Lambda_b^0 \to \Lambda J/\psi$ downstream (left) long (right) candidates. The points show simulated data and the blue line is the signal fit function.

In a second step the fit to the resonant channel data sample is performed. For this fit the tail slope parameter, "n", which is highly correlated with α_1 and α_2, is fixed to the value found in the fit to simulated data. In this fit two background components are modelled: the combinatorial background, parameterized with an exponential and the background from $B^0 \to J/\psi K_S^0$ decays. The shape used to describe the K_S^0 background is obtained from a $B^0 \to J/\psi K_S^0$ simulated sample to which the full selection is applied. The invariant distribution of these events is fit with a DCB function, which is then used to model the K_S^0 background in the $\Lambda_b^0 \to J/\psi \Lambda$ fit. The fit to the simulated $B^0 \to J/\psi K_S^0$ events is reported in Fig. 3.11. When the K_S^0 shape is introduced in the fit to the data all its parameters are fixed. This is particularly important when fitting long candidates, where the K_S^0 peak is less evident, which does not allow to constrain many parameters. On the other hand, in order to take into account possible data-simulation differences, an horizontal shift is added and left floating (by adding a constant to the central value of the DCB, $m_0 \to m_0 + m'$). In summary, the free parameters in the fit to the resonant $\Lambda_b^0 \to J/\psi \Lambda$ sample are the yields of the signal and the combinatorial and K_S^0 backgrounds, the slope of the exponential and the horizontal shift of the K_S^0 shape. Note that all parameters of the fit to the long and downstream samples are independent.

Finally, the rare $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ data sample is fit. In this case the fit to the long and downstream samples is performed simultaneously to obtain a more stable con-

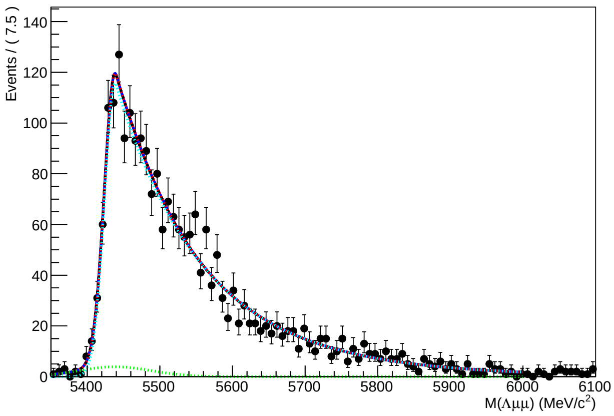

Figure 3.11: Invariant mass distribution of simulated $B^0 \to J/\psi K^0_{\mathrm{s}}$ events after full selection fitted a Double Crystal Ball function.

vergence. In this fit the signal is modelled with the same shape used in the resonant case as there is no physical reason why they should be different. This method is also useful to limit systematic uncertainties as the result will be given as a ratio between rare and resonant quantities. However, the low statistics for the rare sample does not allow to constrain many parameters. Therefore, all parameters of the signal shape are fixed to the ones derived from the fit to the normalisation channel. However, to account for possible differences, arising from a different resolution in different q^2 regions, a scale factor is multiplied to the widths of the two gaussian cores of the signal DCB: $\sigma_1 \to c \cdot \sigma_1$ and $\sigma_2 \to c \cdot \sigma_2$, where the two scale factors are the same. This factors are fixed in the fit to data by fitting rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulated events in each q^2 bin and comparing the widths with the ones found on the fit to the resonant simulated sample, namely

$$c = \sigma^{MC}_{\mu^+\mu^-} / \sigma^{MC}_{J/\psi}. \tag{3.9}$$

Values obtained are ~ 1.9 for downstream candidates and ~ 2.3 for long candidates, corresponding to the fact that in the resonant case a further constrain on the dimuon mass is used, which improves the resolution by a factor of ~ 2. The dependence of the scaling factor on q^2 is found to be small. For the fits on the long and downstream samples the parameters are always fixed to the corresponding J/ψ fit; in this analysis

Table 3.4: Predicted numbers of $B^0 \to K^0_S\mu^+\mu^-$ events in each considered q^2 interval.

q^2 **interval** [GeV^2/c^4]	**Downstream**	**Long**
0.1–2.0	0.9	0.1
2.0–4.0	0.9	0.1
4.0–6.0	0.8	0.1
6.0–8.0	1.1	0.1
11.0–12.5	1.9	0.2
15.0–16.0	1.1	0.1
16.0–18.0	2.0	0.2
18.0–20.0	1.1	0.1
1.1–6.0	2.1	0.1
15.0–20.0	4.2	0.5

shape parameters are never shared between the two candidate categories.

Also in the rare case the modelled background components are the combinatorial background, described with an exponential function and the K^0_S background. The slope of the background is visibly different depending on the q^2 interval. This is partly due to the fact that at high q^2 the combinatorial changes slope because of a kinematical limit at low 4-body masses imposed by the q^2 requirements. The exponential slopes are therefore left as independent parameters in each q^2 interval. The background component from $B^0 \to K^0_S\mu^+\mu^-$ decays is modelled using the same shapes used for the resonant channel. However, in this case the horizontal shift is fixed to what found for the resonant channel. The expected amount of misreconstructed $B^0 \to K^0_S\mu^+\mu^-$ events is small and does not allow to determine reliably the yield. Therefore this is fixed to the yield of $B^0 \to J/\psi\, K^0_S$ decays rescaled by the expected ratio of branching fractions between the resonant and rare channels. The q^2 distribution of $B^0 \to K^0_S\mu^+\mu^-$ simulated events is used to predict the yield as a function of q^2. Table 3.4 reports the number of predicted $B^0 \to K^0_S\mu^+\mu^-$ events in each q^2 interval obtained with the following formula:

$$N_{K^0_S\mu^+\mu^-}(q^2) = N_{J/\psi\, K^0_S} \frac{B(B^0 \to K^0_S\mu^+\mu^-)}{B(B^0 \to K^0_S J/\psi)} \cdot \frac{1}{\epsilon_{rel}} \cdot B(J/\psi \to \mu^+\mu^-) \frac{N(q^2)_{MC}}{N^{tot}_{MC}} \quad (3.10)$$

where $N(q^2)_{MC}$ is the number of simulated rare candidates falling in a q^2 interval after full selection and N^{tot}_{MC} is the total number of simulated events.

As the fit on the rare sample is performed simultaneously on long and downstream candidates, their two yields are not free to vary separately but are parameterised as a function of the common branching fraction using the following formula:

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{\mathrm{d}\mathcal{B}(\Lambda\mu^+\mu^-)/\mathrm{d}q^2}{\mathcal{B}(J/\psi\,\Lambda)}\right] \cdot N(J/\psi\,\Lambda)_k \cdot \varepsilon_k^{\mathrm{rel}} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \to \mu^+\mu^-)}, \tag{3.11}$$

where $k =$(LL,DD), Δq^2 is the width of the q^2 interval and the only free parameter is the relative branching fraction ratio of the rare over J/ψ channels. For the branching fraction of the $J/\psi \to \mu^+\mu^-$ decay the value reported in the PDG book, $(5.93 \pm 0.06) \cdot 10^{-2}$ [1] is used and ε^{rel} corresponds to the relative efficiency between the rare and resonant channels obtained in Sec. 3.6. In this formula the efficiencies and the normalisation yield appear as constants, namely $N(\Lambda\mu^+\mu^-)_k = C_k \cdot \mathcal{B}^{rel}$.

3.5.2 Fit results

Figures 3.12 and 3.13 show fitted invariant mass distributions for the normalisation channel, selected with the high q^2 and low q^2 requirements respectively. Table 3.5 reports the measured yields of $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates found using the low and high q^2 selections. Values for the signal shape parameters are shown on Fig. 3.12. Fits to the rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ samples are shown in Fig. 3.14 for the integrated $15 < q^2 < 20$ and $1.1 < q^2 < 6.0$ GeV$^2/c^4$ q^2 intervals, while fitted invariant mass distribution in all other considered q^2 intervals are in Figs. 3.15 and 3.16 for downstream and long candidates respectively. The yields of rare candidates obtained from the fit are listed in Tab. 3.6 together with their significances. Most candidates are found in the downstream sample, which comprises $\sim 80\,\%$ of the total yield. Note that, since the fit is simultaneous to the two candidate categories, their yields are not parameters free to float independently in the fit but are correlated via the branching ratio. The statistical significance of the observed signal yields is evaluated as $\sqrt{2\Delta\ln\mathcal{L}}$, where $\Delta\ln\mathcal{L}$ is the change in the logarithm of the likelihood function when the signal component is excluded from the fit, relative to the nominal fit in which it is present.

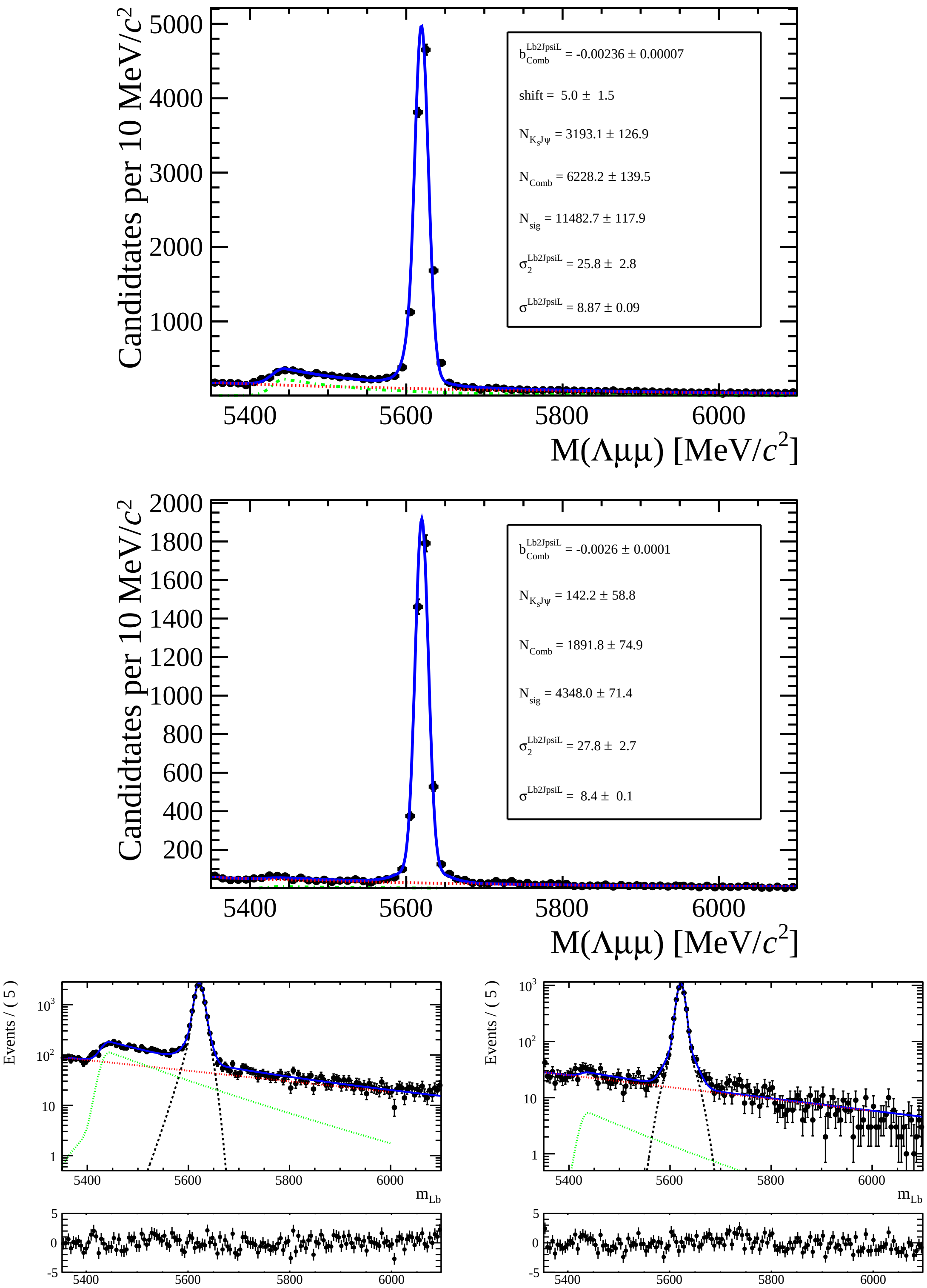

Figure 3.12: Invariant mass distributions of $\Lambda_b^0 \to J/\psi\,\Lambda$ downstream (top) and long (middle) candidates selected with high q^2 requirements. Bottom plots are the same as the upper ones but shown in logarithmic scale. Black points show data. The blue solid line represents the total fit function, the black dashed line the signal, the red dashed line the combinatorial background and the green dashed line the $B^0 \to K_S^0\mu^+\mu^-$ background.

Table 3.5: Number of $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates in the long and downstream categories found using the for low- and high-q^2 requirements. Uncertainties shown are statistical only.

Selection	**Long**	**Downstream**
high-q^2	4313 ± 70	$11\,497 \pm 123$
low-q^2	3363 ± 59	7225 ± 89

Table 3.6: Signal yields (N_S) obtained from the mass fit to $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ candidates in each q^2 interval together with their statistical significances. The $8-11$ and $12.5-15$ GeV$^2/c^4$ q^2 intervals are excluded from the study as they are dominated by decays via charmonium resonances.

q^2 **interval** [GeV$^2/c^4$]	**DD**	**LL**	**Tot. yield**	**Significance**
0.1 – 2.0	6.9 ± 2.2	9.1 ± 3.0	16.0 ± 5.3	4.4
2.0 – 4.0	1.8 ± 1.7	3.0 ± 2.8	4.8 ± 4.7	1.2
4.0 – 6.0	0.4 ± 0.9	0.6 ± 1.4	0.9 ± 2.3	0.5
6.0 – 8.0	4.3 ± 2.0	7.2 ± 3.3	11.4 ± 5.3	2.7
11.0 – 12.5	14.6 ± 2.9	42.8 ± 8.5	60 ± 12	6.5
15.0 – 16.0	13.5 ± 2.2	43.5 ± 7.2	57 ± 9	8.7
16.0 – 18.0	28.6 ± 3.3	88.8 ± 10.1	118 ± 13	13
18.0 – 20.0	22.4 ± 2.6	78.0 ± 8.9	100 ± 11	14
1.1 – 6.0	3.6 ± 2.4	5.7 ± 3.8	9.4 ± 6.3	1.7
15.0 – 20.0	64.6 ± 4.7	209.6 ± 15.3	276 ± 20	21

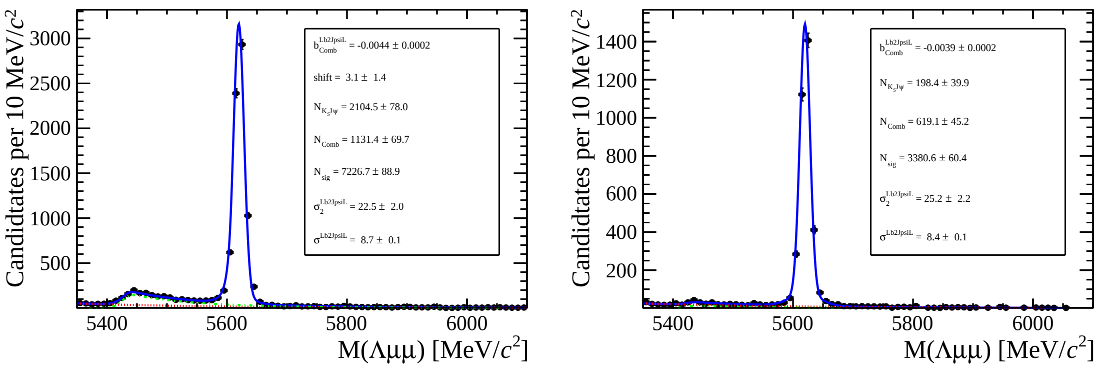

Figure 3.13: Invariant mass distribution of $\Lambda_b^0 \to J/\psi\,\Lambda$ for downstream (left) and long (right) candidates selected with low q^2 requirements.

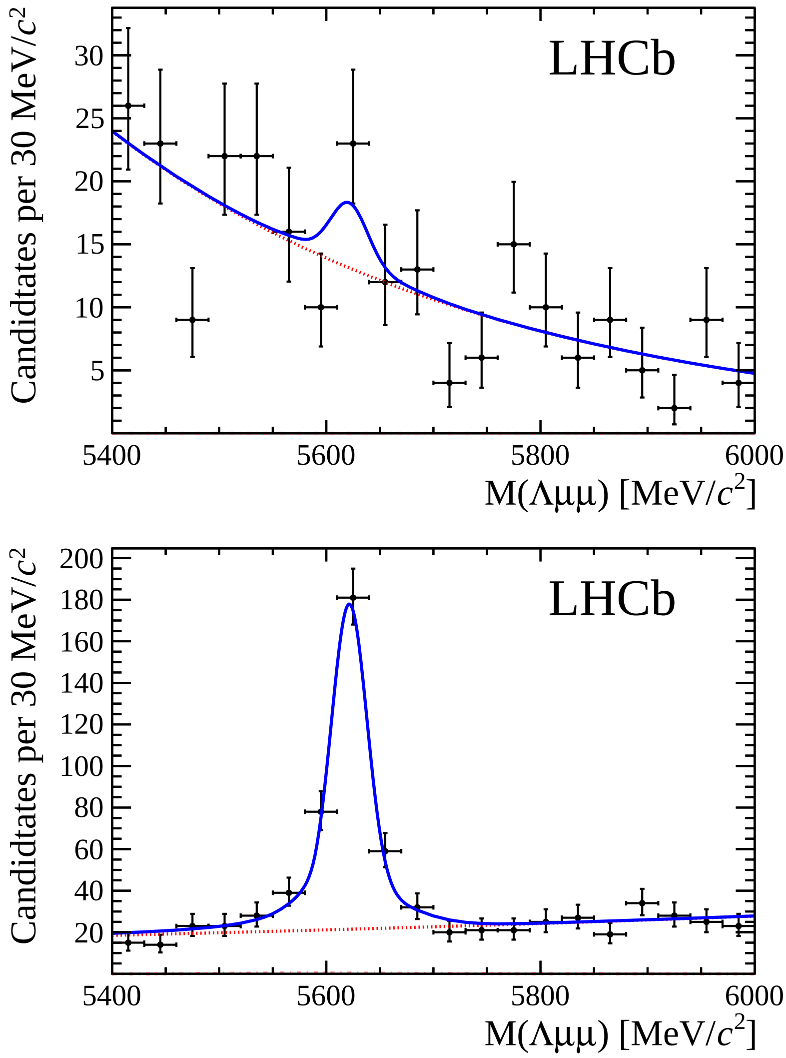

Figure 3.14: Invariant mass distributions of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ candidates in the integrated 0.1–6.0 GeV^2/c^4 (top) and 15–20 GeV^2/c^4 (bottom) q^2 intervals. Points show data combining downstream and long candidates together. The blue solid line represents the total fit function and the dashed red line the combinatorial background.

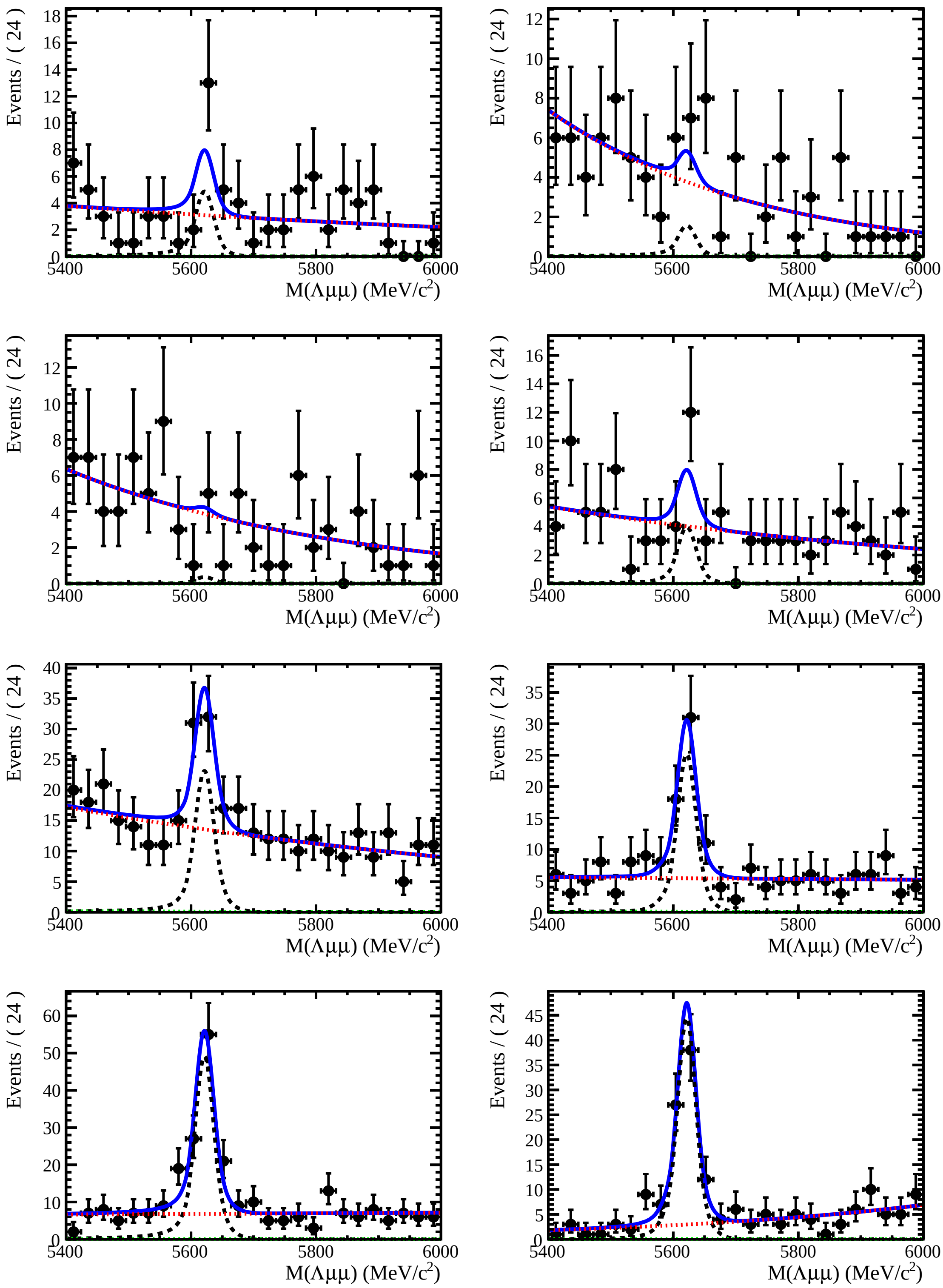

Figure 3.15: Invariant mass distributions of rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for downstream candidates.

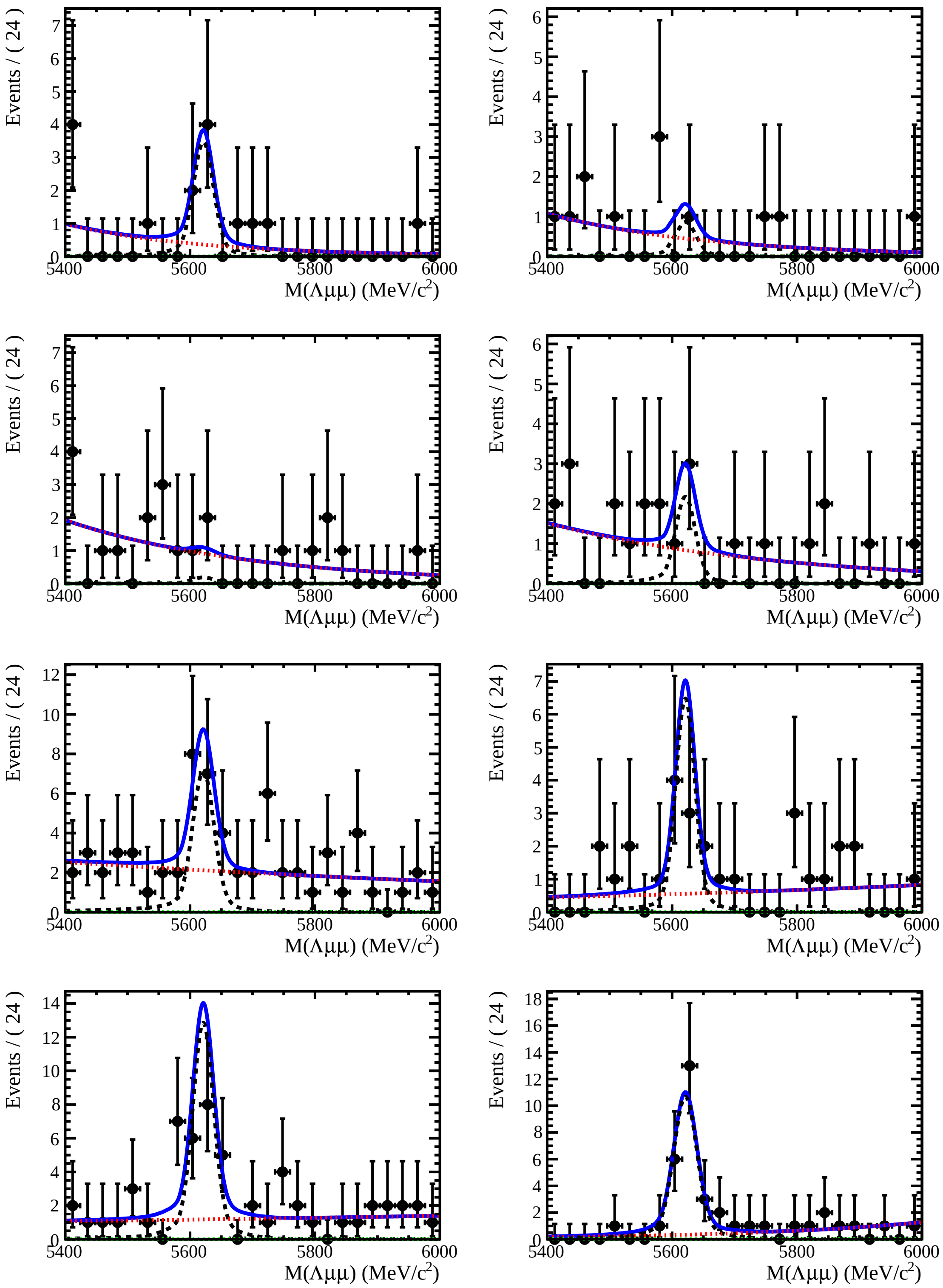

Figure 3.16: Invariant mass distributions of rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for long candidates.

3.6 Efficiency

The selection efficiency is calculated for each decay according to the formula

$$\varepsilon^{tot} = \varepsilon(Geom)\varepsilon(Det|Geom)\varepsilon(Reco|Det)\epsilon(MVA|Reco)\varepsilon(Trig|MVA). \qquad (3.12)$$

In this expression the first term gives the efficiency to have final state particles in the LHCb acceptance. The second term handles the possibility of Λ escaping the detector or interacting with it and therefore never decaying into $p\pi$; this term is referred to as "detection" efficiency. The third term carries information about the reconstruction and pre-selection efficiencies, which are kept together given that boundaries between them are completely artificial. The fourth part deals with the efficiency of the Neural Network for those events which passed the pre–selection. Finally, the last term handles the trigger efficiency for events which are accepted by the full selection. Most of the efficiency components are evaluated using the simulated samples described in Sec. 3.3. Only the efficiency of the PID requirement for the proton (see Tab. 3.1) is separately derived with a data–driven method because the simulation does not provide a good description of PID variables. For complete information, all absolute efficiencies for the two decays $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ are separately listed in the next subsections. However, for the analysis itself only the relative efficiency, $\varepsilon(\Lambda_b^0 \to \Lambda\ \mu^+\mu^-\ )/\varepsilon(\Lambda_b^0 \to J/\psi\,\Lambda)$, is used.

3.6.1 Geometric acceptance

In order to save disk space and time only events are simulated in which the final muons are in the detector acceptance and therefore can be reconstructed. This corresponds to a requirement for each of the muons to be in an interval $10 < \theta < 400$ mrad, where θ is the angle between the muon momentum and the beam line. The efficiency of this requirement is obtained by using a separate simulated sample, where events are generated in the full space. The geometric efficiency varies between 18% at high-q^2 and 20% at low-q^2; Fig. 3.17 shows the dependence of this efficiency as a function of q^2.

3.6.2 Reconstruction and neural network efficiencies

The efficiency to reconstruct the decays together with the pre-selection requirements is evaluated from simulated data. The reconstruction efficiency is subdivided in "Detection" and "Reconstruction and pre-selection" efficiencies. In fact, since Λ is a long lived particle, there is a non-negligible probability that it interacts in the detector or escapes from it and therefore never decays in proton and pion. The reconstruction efficiency includes the efficiency of for the tracks to produce observable signatures and the efficiency for candidates to pass the pre-selection requirements. This component does not include the efficiency of the PID cut that appears in Tab. 3.1, which is kept separate because PID variables are not well described by the simulation. The detection efficiency varies between 88% at high-q^2 and 20% at low-q^2 while the reconstruction efficiency for downstream candidates is almost flat at 6.6% and for long candidates it varies from 1.6% at high-q^2 to 2.0% at low-q^2. Fig. 3.17 shows the dependence of these efficiencies as a function of q^2. The NN selection efficiency is again evaluated from simulated samples and it is observed to vary from 58% to 84% for downstream candidates and from 77% to 92% for log candidates. Fig. 3.17 shows the dependence of this efficiency as a function of q^2. The sudden jump in efficiency at ~ 9 GeV$/c^2$ is due to the fact that a different figure-of-merit is used to optimise the NN cut in the low and high q^2 regions, which results in different efficiencies.

3.6.3 Trigger efficiency

The trigger efficiency is again calculated using a simulated sample and it varies between 61% and 84% for downstream candidates and from 65% to 85% for long candidates. Fig. 3.17 shows the dependence of this efficiency as a function of q^2. Using the resonant channel it is possible to crosscheck on data the efficiency obtained using the simulation. In LHCb triggered events can fall in two categories: events triggered by a track which is part of a signal candidate, Trigger On Signal (TOS), or by other tracks in the event, Trigger Independent of Signal (TIS). As the TIS and TOS categories are not exclusive the TIS sample provides a control sample which can be used to obtain the efficiency for TOS trigger. This is calculated with the formula:

$$\varepsilon_{\mathrm{TOS}} = \frac{\text{TIS and TOS}}{\text{TIS}}. \tag{3.13}$$

As data contains background the numbers of signal candidates in the "TIS" and "TIS && TOS" categories are not just determined by counting events but from a fit to the 4-body invariant mass, $m(p\pi\mu\mu)$. This procedure takes the name of TISTOS method. Using the data–driven method an efficiency of $(70 \pm 5)\%$ is obtained, while this is calculated to be $(73.33 \pm 0.02)\%$ using the simulation. Results are therefore compatible within 1σ.

3.6.4 PID efficiency

For long tracks a PID requirement on protons (`PIDp`> -5) is applied. The simulation is known not to describe particle ID variables well and therefore a data-driven method is used to obtain this efficiency component. This is done using the `PIDCalib` package (see Sec. 2.8.1), which uses as calibrations samples decays where particles can be identified due to their kinematic properties. In the case of protons a sample of Λ particles is used, where the proton can be identified because it always has the highest momentum. The package allows to divide the phase space in bins of variables relevant for PID performances; in this analysis momentum and pseudorapidity are used. Using the calibration sample the efficiency is derived in each two-dimensional bin. Finally, to take into account that the decay channel under study could have different kinematical distributions than the calibration sample these efficiency tables are used to re-weight the simulation. The PID efficiency varies from 97.3% at low-q^2 to 98.2% at high-q^2.

Table 3.7: Absolute efficiency values for $\Lambda_b^0 \to J/\psi\,\Lambda$. Uncertainties are statistical only.

Efficiency	**Downstream**	**Long**
$\varepsilon(PID)$	0.1818 ± 0.0003	
$\varepsilon(Det)$	0.9017 ± 0.0003	
$\varepsilon(Reco)$	0.0724 ± 0.0004	0.0203 ± 0.0002
$\varepsilon(PID)$	–	97.89 ± 0.005
$\varepsilon(MVA)$	0.882 ± 0.002	0.942 ± 0.002
$\varepsilon(Trig)$	0.697 ± 0.003	0.734 ± 0.005
Full Selection	0.0445 ± 0.0003	0.0140 ± 0.0002
Total	0.00729 ± 0.00005	0.00230 ± 0.00003

3.6.5 Relative efficiencies

In the previous sections absolute efficiencies values were given for the rare channel in different q^2 intervals. Figure 3.17 contains a summary of those values in these tables in graphical form. This section reports the corresponding relative efficiencies with respect to the $\Lambda_b^0 \to J/\psi\,\Lambda$ channel, which will be used to correct the yields and obtain the differential branching fraction. Table 3.7 reports the absolute efficiency values for the J/ψ channel used to derive the relative efficiencies. Relative geometric, detection and PID efficiencies are listed in Tab. 3.8, while Tabs. 3.10 and 3.9 report relative reconstruction, trigger and NN efficiencies separately for downstream and long candidates. Since the latter three components are obtained from the same simulated sample their statistical errors are correlated. Therefore the total of the three is also reported as a single efficiency and labeled "Full Selection". Finally, Tab. 3.13 reports the total of all relative efficiencies, which will be then used to correct the raw yields and calculate the differential branching fraction. Uncertainties reflect the statistics of both rare and resonant samples, while systematic uncertainties are discussed in next sections.

Table 3.8: Relative geometric, detection and PID relative efficiencies between $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ decays. Uncertainties reflect the statistics of both samples.

q^2 [GeV$^2/c^4$]	Geometric	Detection	PID
0.1 – 2.0	1.2976 ± 0.0050	0.9751 ± 0.0006	0.99418 ± 0.00013
2.0 – 4.0	1.1541 ± 0.0043	0.9814 ± 0.0005	0.99523 ± 0.00013
4.0 – 6.0	1.1043 ± 0.0044	0.9872 ± 0.0006	0.99699 ± 0.00012
6.0 – 8.0	1.0778 ± 0.0045	0.9939 ± 0.0006	0.99805 ± 0.00011
11.0 – 12.5	1.0431 ± 0.0058	1.0074 ± 0.0007	1.00151 ± 0.00010
15.0 – 16.0	1.0426 ± 0.0084	1.0188 ± 0.0010	1.00431 ± 0.00008
16.0 – 18.0	1.0296 ± 0.0068	1.0255 ± 0.0008	1.00215 ± 0.00008
18.0 – 20.0	1.0288 ± 0.0087	1.0333 ± 0.0010	1.00226 ± 0.00005
1.1 – 6.0	1.1396 ± 0.0031	0.9835 ± 0.0004	0.99589 ± 0.00009
15.0 – 20.0	1.0320 ± 0.0048	1.0269 ± 0.0006	1.00281 ± 0.00006

Table 3.9: Relative efficiencies between $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ decays for long events. Uncertainties reflect the statistics of both samples.

q^2 [GeV$^2/c^4$]	Recoscruction	MVA	Trigger	Full Selection
0.1 – 2.0	0.96 ± 0.02	0.863 ± 0.012	0.79 ± 0.02	0.65 ± 0.02
2.0 – 4.0	0.97 ± 0.02	0.803 ± 0.012	0.89 ± 0.02	0.69 ± 0.02
4.0 – 6.0	1.04 ± 0.02	0.824 ± 0.012	0.92 ± 0.02	0.79 ± 0.02
6.0 – 8.0	1.05 ± 0.02	0.825 ± 0.012	0.96 ± 0.02	0.84 ± 0.02
11.0 – 12.5	1.10 ± 0.03	1.002 ± 0.008	1.01 ± 0.02	1.10 ± 0.03
15.0 – 16.0	0.89 ± 0.03	0.987 ± 0.013	1.13 ± 0.02	0.98 ± 0.04
16.0 – 18.0	0.84 ± 0.03	0.985 ± 0.010	1.17 ± 0.02	0.97 ± 0.03
18.0 – 20.0	0.67 ± 0.03	0.944 ± 0.017	1.18 ± 0.02	0.75 ± 0.04
1.1 – 6.0	1.00 ± 0.02	0.820 ± 0.008	0.89 ± 0.01	0.73 ± 0.02
15.0 – 20.0	0.78 ± 0.02	0.973 ± 0.008	1.16 ± 0.01	0.89 ± 0.02

Table 3.10: Relative efficiencies between $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,\Lambda$ decays for downstream events. Uncertainties reflect the statistics of both samples.

q^2 [GeV$^2/c^4$]	Reconstruction	MVA	Trigger	Full Selection
0.1 – 2.0	0.721 ± 0.009	0.706 ± 0.010	0.805 ± 0.011	0.410 ± 0.009
2.0 – 4.0	0.920 ± 0.010	0.661 ± 0.008	0.870 ± 0.010	0.529 ± 0.010
4.0 – 6.0	0.997 ± 0.010	0.662 ± 0.008	0.895 ± 0.010	0.590 ± 0.011
6.0 – 8.0	1.050 ± 0.011	0.665 ± 0.008	0.960 ± 0.010	0.671 ± 0.012
11.0 – 12.5	1.112 ± 0.014	1.007 ± 0.006	1.069 ± 0.009	1.197 ± 0.019
15.0 – 16.0	1.019 ± 0.018	1.000 ± 0.009	1.175 ± 0.012	1.197 ± 0.026
16.0 – 18.0	0.968 ± 0.014	0.961 ± 0.008	1.200 ± 0.010	1.115 ± 0.020
18.0 – 20.0	0.832 ± 0.016	0.943 ± 0.010	1.231 ± 0.012	0.966 ± 0.023
1.1 – 6.0	0.950 ± 0.007	0.663 ± 0.005	0.876 ± 0.007	0.551 ± 0.007
15.0 – 20.0	0.929 ± 0.010	0.963 ± 0.005	1.204 ± 0.007	1.077 ± 0.014

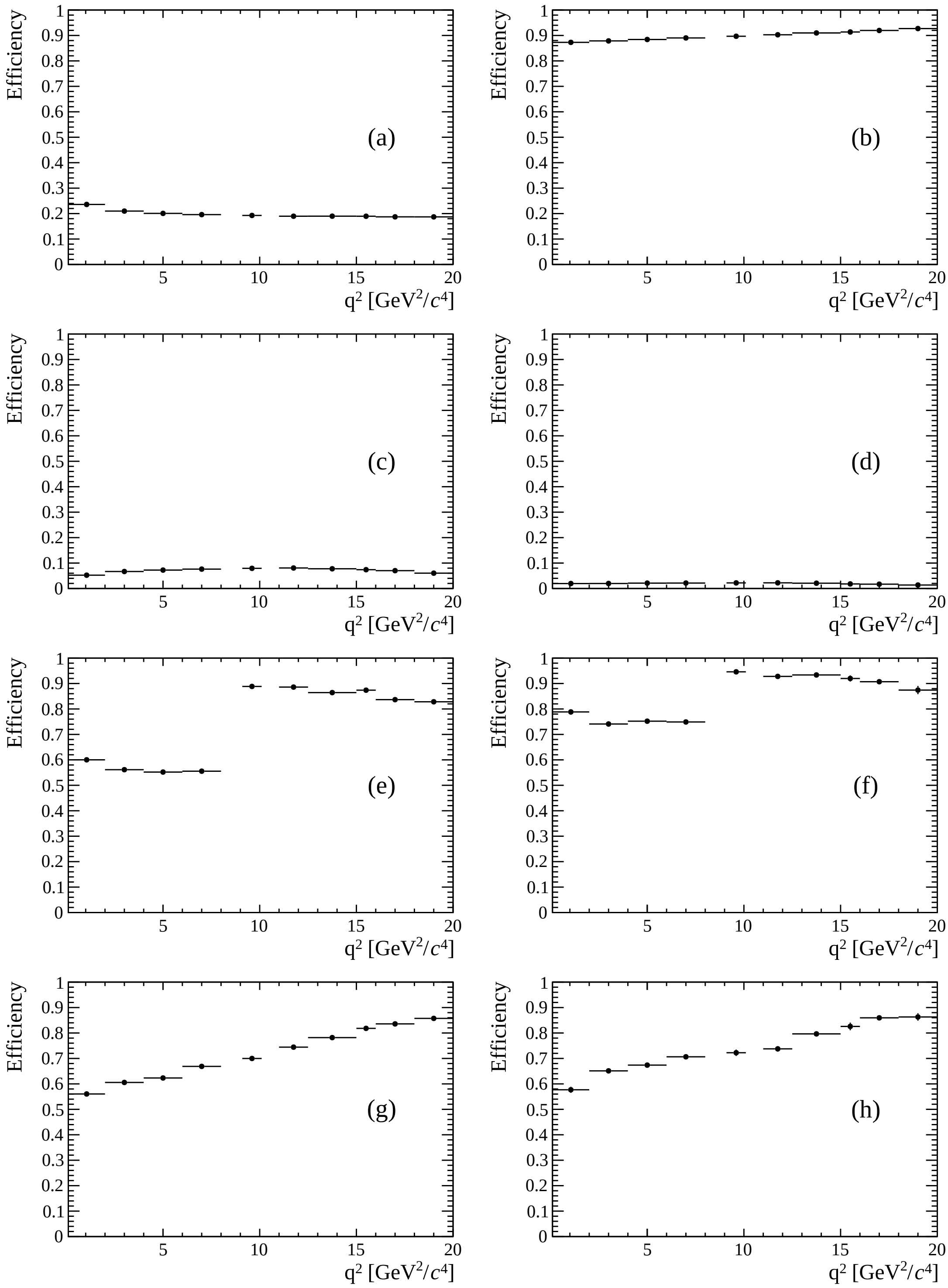

Figure 3.17: Absolute efficiencies as a function of q^2: geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) candidates, NN efficiency for DD (e) and LL (f) and trigger efficiency for DD (g) and LL (h).

3.7 Systematic uncertainties

This section describes the main considered sources of systematic uncertainty.

3.7.1 Systematic uncertainty on the yields

The choice of specific PDFs to model the invariant mass distribution could result in a bias. To asses the effect of the signal PDF choice as a first step a number of models are tried on the $\Lambda_b^0 \to J/\psi\,\Lambda$ data sample to understand which ones are plausible. Table 3.11 reports the χ^2 and relative probabilities obtained using different models including: the default model (a DCB function), a simple Gaussian function, a single Crystal Ball function and the sum of two Gaussians. The only two models that give a reasonable p-value are the default DCB and the sum of two Gaussian functions (DG). In a second step simulated experiments are generated and fit with the two chosen models. Events are generated according to a density function given by the default model fitted on data separately for each q^2 interval. In this way, for each q^2 interval, a specific shape is reproduced including the background level and slope. Furthermore, a number of events comparable to the one found in data is generated. For each experiment a normalised bias is calculated as

$$b = \left(\frac{N_{\ell\ell}^{DCB}}{N_{J/\psi}^{DCB}} - \frac{N_{\ell\ell}^{DG}}{N_{J/\psi}^{DG}} \right) / \frac{N_{\ell\ell}^{DCB}}{N_{J/\psi}^{DCB}} \tag{3.14}$$

where $N_{\ell\ell}^{model}$ and $N_{J/\psi}^{model}$ are the numbers of rare and resonant candidates observed using a specific model. The average bias over 1000 pseudo-experiments is taken as systematic uncertainty. Note that in each case the rare and normalisation channels are fit with the same signal model and, while for the default case the rare parameters are fixed to what found for the resonant channel, they are left free to vary in the second model in order to asses at the same time the systematic due to the parameters constraints.

Table 3.11: χ^2, NDF, p-values and number of signal events obtained fitting $\Lambda_b^0 \to J/\psi\ \Lambda$ data using different models.

Model	χ^2/ndf	NDF	**p-value**	N_{evts}
DCB (default)	1.0	187	0.51	9965.4
Gauss	1.8	193	~ 0	9615.7
Double Gauss	1.1	191	0.45	9882.4
CB	1.5	191	~ 0	9802.4

For the background PDF systematic the rare channel is re-fit leaving the yield of the K^0_S component free to vary; this is instead fixed to the predicted value in the default fit. The same procedure as for the signal PDF is applied. Results are reported in Tab. 3.12. The most affected q^2 interval is the one in the middle of the charmonium resonances, where a combination of lower statistics and higher background leaves more freedom to the signal shape. Finally, a background component for $B^+ \to K^{*+}(K^0_S\pi^+)\mu^+\mu^-$ decays is added to the fit, modelled using the distribution of simulated events after full selection. No significant bias is found for this component.

q^2 [GeV$^2/c^4$]	**Sig. PDF bias (%)**	**Bkg. PDF bias (%)**	**Tot. sys. (%)**
0.1 – 2.0	3.2	1.1	3.4
2.0 – 4.0	2.9	2.4	3.8
4.0 – 6.0	4.6	4.8	6.6
6.0 – 8.0	1.2	1.7	2.0
11.0 – 12.5	2.6	1.8	3.2
15.0 – 16.0	1.3	2.5	2.8
16.0 – 18.0	0.6	1.3	1.4
18.0 – 20.0	1.7	1.8	2.5
1.1 – 6.0	0.1	4.2	4.2
15.0 – 20.0	1.0	0.2	1.1

Table 3.12: Values of systematic uncertainties due to the choice of signal and background shapes in bins of q^2.

3.7.2 Systematic uncertainties on the efficiency determination

Systematic uncertainties in the efficiency determination are due to the limited knowledge of the decay properties such as the Λ_b^0 lifetime and production polarisation. The uncertainties are directly calculated on the relative efficiencies as these are the ones that are actually used in the analysis. It should be noted that not all sources contribute to each part of the efficiency. For brevity, this section only reports estimates of the systematic uncertainties obtained while the full information is contained in Appendix C.

3.7.2.1 Effect of new physics on the decay model

New physics could affect the decay model by adding contributions to the C_7 and C_9 Wilson Coefficients. This would result in a modification of the q^2 spectrum and therefore of the efficiency. To asses this systematic the Wilson Coefficients are modified by adding a new physics component ($C_i \to C_i + C_i^{\text{NP}}$). Figure 3.18 shows q^2 spectra obtained weighting the simulation for a model embedding the default and 3 modified sets of Wilson Coefficients. The used values, reported on top of each plot, are inspired to maintain compatibility with the recent LHCb result about the P_5' observable [41]. The biggest effect is observed in the very low q^2, below 2 GeV^2/c^4, where the efficiency can change up to 7%, while it changes 3-4 % between 3 and 4 GeV^2/c^4 and 2-3 % in the rest of the spectrum. As this analysis is performed under the hypothesis that the decays are described by a the SM, these values are given in order to provide the full information but are not added as systematic uncertainties.

3.7.2.2 Simulation statistics

The limited statistics of the simulated samples used to determine efficiencies is considered as a source of systematic uncertainty. While it is not the dominant source, its size does not allow to completely neglect it. When reporting relative

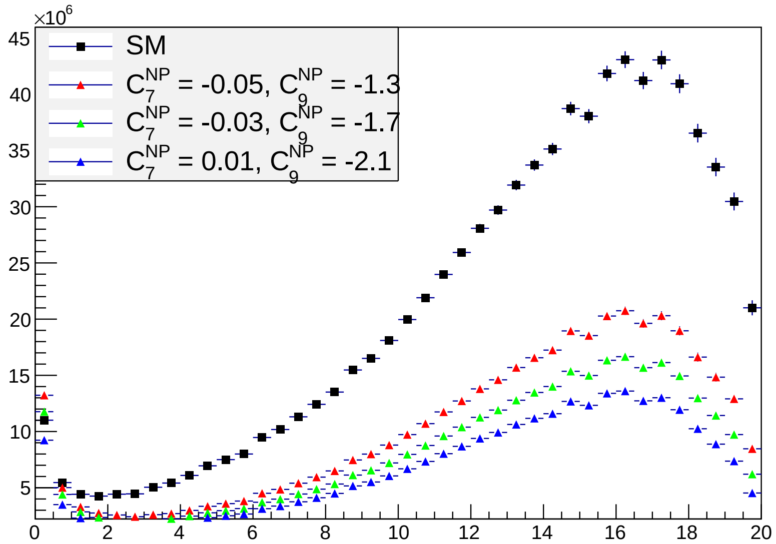

Figure 3.18: The q^2 spectrum of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

efficiency values the statistical uncertainty due to the rare and resonant channels is always considered.

3.7.2.3 Production polarisation and decay structure

One of the main unknown, which affects the determination of the efficiencies, is the angular structure of the decays. And, connected to it, also the production polarisation, which is a parameter of the model. To assess the systematic uncertainty due to the knowledge of the production polarisation for $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays the polarisation parameter in the model is varied within one standard deviation from the central value of the most recent LHCb measurement, $P_b = 0.06 \pm 0.09$ [91]. The full observed difference is taken as systematic uncertainty. To assess the systematic uncertainty due to the decay structure an alternative set of form factors is used based on lattice QCD calculation [94]. Details of this are explained in Appendix A.1. The two models are compared and the full difference is taken as systematic uncertainty. In total this results in an uncertainty of $\sim 1.3\%$ for long candidates and $\sim 0.6\%$

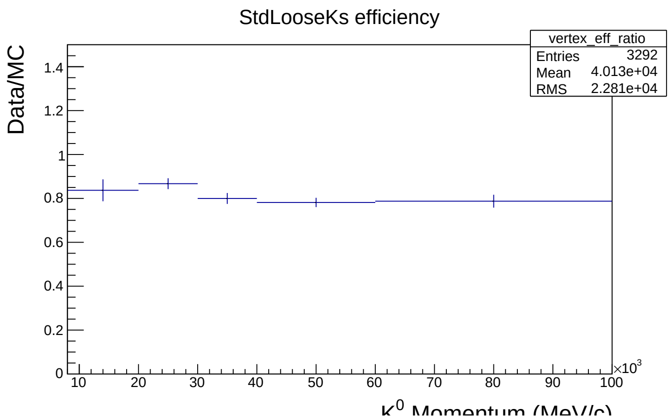

Figure 3.19: Ratio of reconstruction efficiency in Data and MC found using K_S events [96].

for downstream candidates, mostly coming from the knowledge of the production polarisation.

3.7.2.4 Λ_b^0 lifetime

The Λ_b^0 lifetime is known with limited precision. For evaluation of the efficiencies the world average value, 1.482 ps^{-1} [95] is used. To evaluate the systematic uncertainty, this is varied within one standard deviation from the measured value. Only the case where both signal and normalisation channel are varied in same direction are considered. The larger difference with the default lifetime case is taken as systematic uncertainty, which is found to range from $\sim 0.4\%$ at low q^2 to $\sim 0.1\%$ at high q^2.

3.7.2.5 Downstream candidates reconstruction efficiency

Other analysis in LHCb using particles reconstructed with downstream tracks showed that the efficiency for these candidates is not well simulated. For example, Fig. 3.19 shows the ratio between the reconstruction efficiency for downstream candidates in data and simulation found analysing K_S^0 events [96]. This effect is not yet fully understood and is currently under study. It seems to be mainly due to a poor simulation of the vertexing efficiency for downstream tracks. This effect is dealt with in two steps. Firstly, the analysis is performed separately for downstream and long candidates. Since efficiencies are also calculated separately, the effect mostly cancels in the ratio between the rare and resonant channels. In a second step a systematic uncertainty is assigned for downstream candidates only re-weighting the simulation by the efficiency ratio between data and simulation found for K_S as a function of momentum (see Fig. 3.19). The efficiencies obtained using the weighted and unweighted simulation are compared and the full difference is taken as systematic uncertainty. As the discrepancy shows little dependence on momentum, dependencies due to the different momentum distributions of Λ and K^0_S are assumed to be negligible. This results in an extra 0.4% systematic uncertainty at low q^2 and 1.2% at high q^2, only for downstream candidates.

3.7.2.6 Data-simulation discrepancies

The simulation used to calculate the efficiency is re-weighted as described in Sec. 3.3.2. The influence of this procedure on the efficiency determination is checked by comparing values obtained with and without re-weighting. The effect is negligible with respect to other systematics considered.

3.8 Differential branching ratio extraction

In this section the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay is calculated relative to the $\Lambda_b^0 \to J/\psi\,\Lambda$ channel as a function of q^2. The values are directly obtained from the fit to the rare sample by parameterising the downstream and long yields with the following formula:

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{\mathrm{d}\mathcal{B}(\Lambda\mu^+\mu^-)/\mathrm{d}q^2}{\mathcal{B}(J/\psi\,\Lambda)}\right] \cdot N(J/\psi\,\Lambda)_k \cdot \varepsilon_k^{\mathrm{rel}} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \to \mu^+\mu^-)}, \qquad (3.15)$$

where $k =$(LL,DD), Δq^2 is the width of the q^2 interval, $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$ [1] and the only free parameter is the relative branching fraction ratio.

Table 3.13: Absolute values of the total relative efficiency and the absolute value of the uncorrelated uncertainty (σ^k_{uncorr}), together with relative values of the correlated uncertainty (σ_{corr}).

q^2 [GeV$^2/c^4$]	Eff. (DD)	σ^{DD}_{uncorr}	Eff. (LL)	σ^{LL}_{uncorr}	σ_{corr}
0.1 – 2.0	0.694	0.058	1.136	0.066	1.0%
2.0 – 4.0	0.693	0.027	0.907	0.047	2.7%
4.0 – 6.0	0.699	0.018	0.964	0.044	2.7%
6.0 – 8.0	0.733	0.020	0.953	0.048	2.7%
11.0 – 12.5	1.254	0.032	1.140	0.057	3.4%
15.0 – 16.0	1.260	0.035	1.035	0.060	3.0%
16.0 – 18.0	1.163	0.029	0.997	0.048	1.7%
18.0 – 20.0	1.023	0.027	0.782	0.040	2.7%
1.1 – 6.0	0.696	0.032	0.950	0.058	1.0%
15.0 – 20.0	1.132	0.014	0.927	0.031	1.4%

Table 3.13 summarises the total relative efficiencies, ε^{rel}, for downstream and long candidates together with their correlated and uncorrelated uncertainties, where the correlation is intended between the downstream and long samples. On the table the uncorrelated uncertainty corresponds to the total systematic uncertainty on the efficiency determination. The correlated uncertainty is given in percent form since it can be applied to either downstream, long candidates or their combination. This includes the PDF systematic described in Sec. 3.7.1 and the systematic due to the uncertainty on the $J/\psi \to \mu^+\mu^-$ branching fraction.

Figure 3.20 shows the branching fraction obtained by fitting the downstream and long samples independently, while the combined result, obtained fitting both samples simultaneously, is shown in Fig. 3.21. Values are also listed in Tab. 3.14, where the statistical uncertainty on the rare channel and the total systematic uncertainty are shown separately. The statistical uncertainty is calculated using the MINOS application of the MINUIT package [97], which provides an asymmetric interval. The normalisation and systematic uncertainties are evaluated by pushing the efficiencies and normalisation yields up and down by one standard deviation and re-performing the fit. The different efficiencies used traslate into a different branching fraction and the full difference with respect to the default fit is taken as systematic uncertainty in each direction.

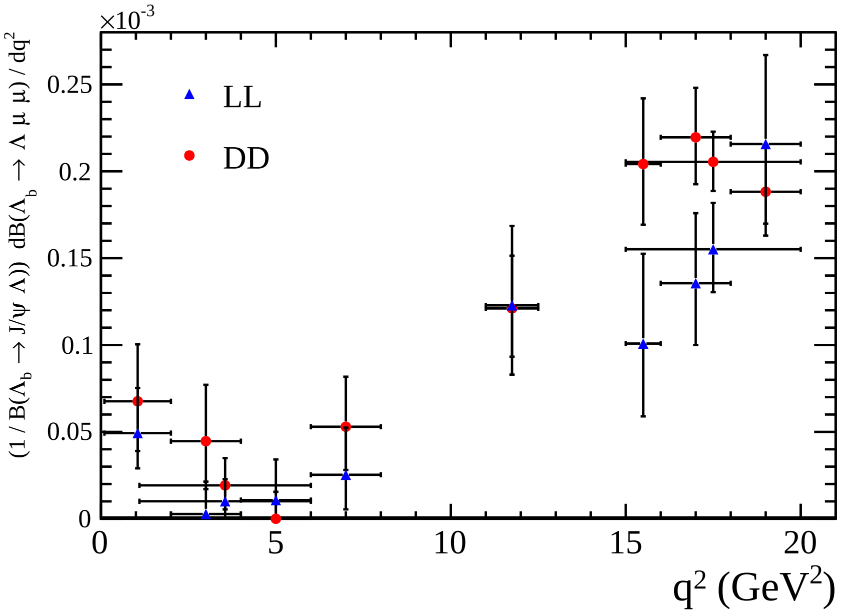

Figure 3.20: Measured values of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ branching fraction relative to the $\Lambda_b^0 \to J/\psi\ \Lambda$ decay as a function of q^2 obtained fitting the downstream and long samples independently. Error bars represent the total statistical and systematic uncertainty.

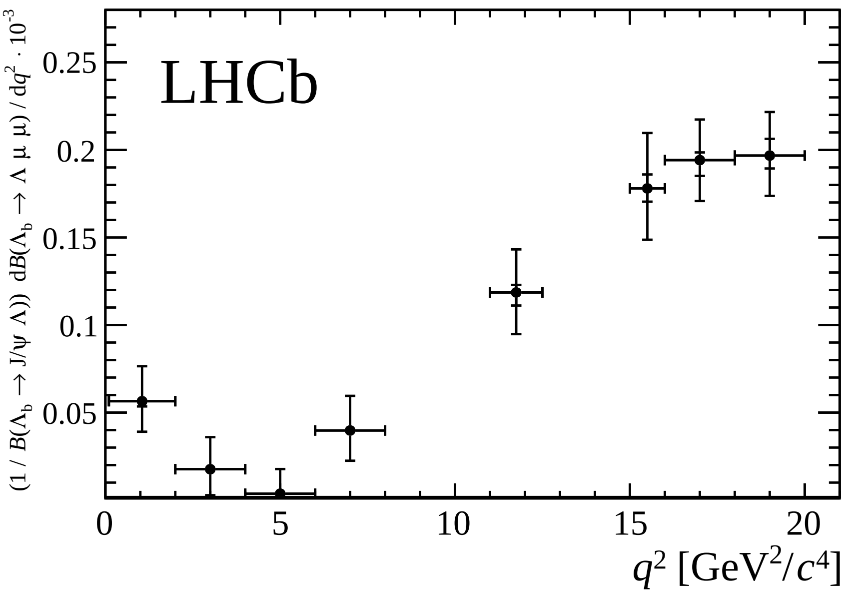

Figure 3.21: Branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay normalised to the $\Lambda_b^0 \to J/\psi\,\Lambda$ mode. The inner error bar represents the systematic uncertainty and the outer error bar includes the statistical uncertainty.

Table 3.14: Differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay relative to $\Lambda_b^0 \to J/\psi\,\Lambda$ decays, where the uncertainties are statistical and systematic, respectively.

q^2 interval [GeV^2/c^4]	$\frac{\mathrm{d}\mathcal{B}(\Lambda_b^0\to\Lambda\mu^+\mu^-)/\mathrm{d}q^2}{\mathcal{B}(\Lambda_b^0\to J/\psi\,\Lambda)} \cdot 10^{-3}[(\mathrm{GeV}^2/c^4)^{-1}]$
0.1 – 2.0	$0.56\ ^{+0.20}_{-0.17}\ ^{+0.03}_{-0.03}$
2.0 – 4.0	$0.18\ ^{+0.18}_{-0.15}\ ^{+0.01}_{-0.01}$
4.0 – 6.0	$0.04\ ^{+0.14}_{-0.04}\ ^{+0.01}_{-0.01}$
6.0 – 8.0	$0.40\ ^{+0.20}_{-0.17}\ ^{+0.01}_{-0.02}$
11.0 – 12.5	$1.19\ ^{+0.24}_{-0.23}\ ^{+0.04}_{-0.07}$
15.0 – 16.0	$1.78\ ^{+0.31}_{-0.28}\ ^{+0.08}_{-0.08}$
16.0 – 18.0	$1.94\ ^{+0.23}_{-0.22}\ ^{+0.04}_{-0.09}$
18.0 – 20.0	$1.97\ ^{+0.23}_{-0.22}\ ^{+0.10}_{-0.07}$
1.1–6.0	$0.14\ ^{+0.10}_{-0.09}\ ^{+0.01}_{-0.01}$
15.0–20.0	$1.90\ ^{+0.14}_{-0.14}\ ^{+0.04}_{-0.06}$

Finally, values for the absolute branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay are obtained by multiplying the relative branching fraction by the absolute branching fraction of the normalisation channel, $\mathcal{B}(\Lambda_b^0 \to J/\psi\,\Lambda) = (6.3 \pm 1.3) \times 10^{-4}$ [1]. Values are shown in Fig. 3.22 and summarised in Tab. 3.15, where the uncertainty due to the knowledge of the normalisation channel (norm), which is correlated across q^2, is shown separately. The SM predictions on the plot are obtained from Ref. [94].

Evidence for the signal is found for the first time in the q^2 region between the charmonium resonances and in the interval $0.1 < q^2 < 2.0\ \mathrm{GeV}^2/c^4$, where an increased yield is expected due to the proximity of the photon pole. The uncertainty on the absolute branching fraction is dominated by the precision with which the branching fraction of the normalisation channel is known, while the uncertainty on the relative branching fraction is dominated by the size of the available data sample. The data are consistent with the theoretical predictions in the high-q^2 region but lie below the predictions in the low-q^2 region.

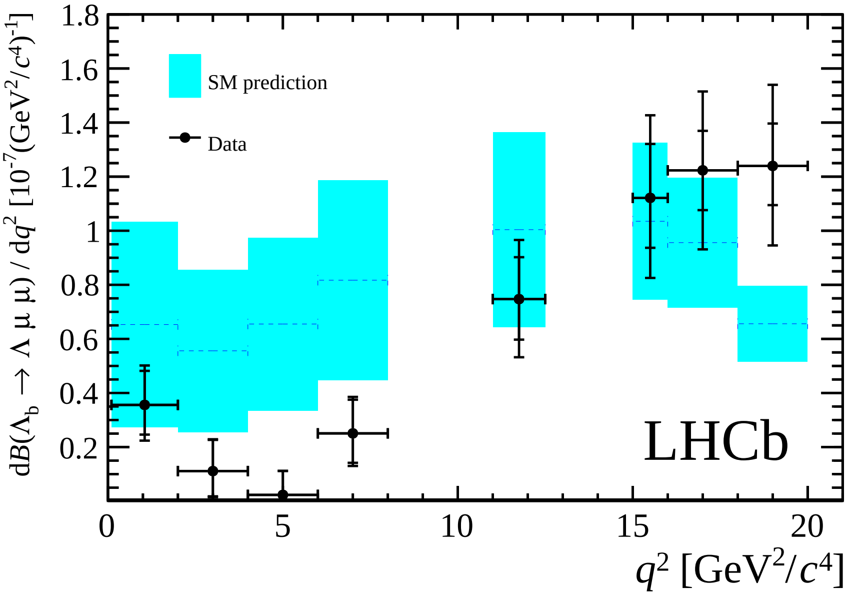

Figure 3.22: Measured $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ branching fraction as a function of q^2 with the SM predictions [94] superimposed. The inner error bars on data points represent the total uncertainty on the relative branching fraction (statistical and systematic); the outer error bar also includes the uncertainties from the branching fraction of the normalisation mode.

Table 3.15: Measured differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, where the uncertainties are statistical, systematic and due to the uncertainty on the normalisation mode, $\Lambda_b^0 \to J/\psi\,\Lambda$, respectively.

q^2 interval [GeV$^2/c^4$]	$\mathrm{d}\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-)/\mathrm{d}q^2 \cdot 10^{-7}[(\mathrm{GeV}^2/c^4)^{-1}]$
0.1 – 2.0	$0.36\ ^{+0.12}_{-0.11}\ ^{+0.02}_{-0.02} \pm 0.07$
2.0 – 4.0	$0.11\ ^{+0.12}_{-0.09}\ ^{+0.01}_{-0.01} \pm 0.02$
4.0 – 6.0	$0.02\ ^{+0.09}_{-0.00}\ ^{+0.01}_{-0.01} \pm 0.01$
6.0 – 8.0	$0.25\ ^{+0.12}_{-0.11}\ ^{+0.01}_{-0.01} \pm 0.05$
11.0 – 12.5	$0.75\ ^{+0.15}_{-0.14}\ ^{+0.03}_{-0.05} \pm 0.15$
15.0 – 16.0	$1.12\ ^{+0.19}_{-0.18}\ ^{+0.05}_{-0.05} \pm 0.23$
16.0 – 18.0	$1.22\ ^{+0.14}_{-0.14}\ ^{+0.03}_{-0.06} \pm 0.25$
18.0 – 20.0	$1.24\ ^{+0.14}_{-0.14}\ ^{+0.06}_{-0.05} \pm 0.26$
1.1 – 6.0	$0.09\ ^{+0.06}_{-0.05}\ ^{+0.01}_{-0.01} \pm 0.02$
15.0 – 20.0	$1.20\ ^{+0.09}_{-0.09}\ ^{+0.02}_{-0.04} \pm 0.25$

CHAPTER 4

Angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays

---

Neglecting Λ_b^0 production polarisation, the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay angular distributions can be described as a function of three angles and q^2. The first two angles are the ones which are relevant for the analysis in this chapter and are defined in Fig. 4.1, where θ_ℓ is the angle between the positive (negative) muon direction and the dimuon system direction in the Λ_b^0 ($\overline{\Lambda}{}_b^0$) rest frame, and θ_h is defined as the angle between the proton and the Λ baryon directions, also in the Λ_b^0 rest frame. The third angle is the angle between the dimuon and Λ decay planes, which is integrated over in this analysis. This chapter describes a measurement of two forward-backward asymmetries in the leptonic ($A_{\rm FB}^\ell$) and in the hadronic ($A_{\rm FB}^h$) systems. These forward-backward asymmetries are defined as

$$A_{\rm FB}^i(q^2) = \frac{\int_0^1 \frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_i}\mathrm{d}\cos\theta_i - \int_{-1}^0 \frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_i}\mathrm{d}\cos\theta_i}{\mathrm{d}\Gamma/\mathrm{d}q^2}, \tag{4.1}$$

where $\mathrm{d}^2\Gamma/\mathrm{d}q^2\,\mathrm{d}\cos\theta_i$ is the two-dimensional differential rate and $d\Gamma/dq^2$ is rate integrated over the angles.

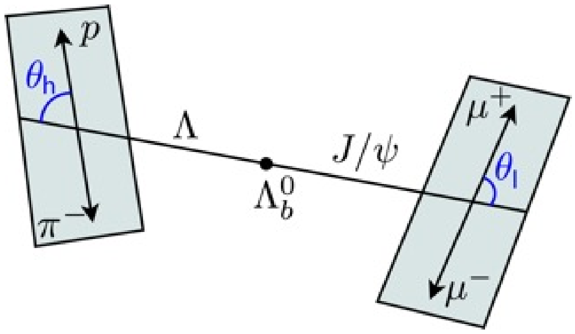

Figure 4.1: Graphical representation of the angles for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.

The A_{FB}^{ℓ} observable was also measured by LHCb in $B^0 \to K^{*0}\mu^+\mu^-$ decays which are going through the same quark level transition as $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays. Instead the hadronic asymmetry, A_{FB}^{h}, is interesting only in the Λ_b^0 case as it is zero by definition in B^0 decays, where K^{*0} decays strongly.

4.1 One-dimensional angular distributions

This section describes the derivation of the functional form of the differential distributions as a function of $\cos\theta_\ell$ and $\cos\theta_h$, which are used to measure the observables. The content of this section is based on the calculations in Ref. [90].

For unpolarised Λ_b^0 production, integrating over the three angles the differential branching fraction is given in Eq. 11 of Ref. [90] as

$$\frac{d\Gamma(\Lambda_b \to \Lambda\,\ell^+\ell^-)}{dq^2} = \frac{v^2}{2}\cdot\left(U^{V+A} + L^{V+A}\right) + \frac{2m_\ell^2}{q^2}\cdot\frac{3}{2}\cdot\left(U^V + L^V + S^A\right), \quad (4.2)$$

and the lepton helicity angle differential distribution, given in Eq. 15, has the form

$$\begin{aligned}\frac{d\Gamma(\Lambda_b \to \Lambda\,\ell^+\ell^-)}{dq^2 d\cos\theta_\ell} \; &= \; v^2\cdot\left[\frac{3}{8}\,(1+\cos^2\theta_\ell)\cdot\frac{1}{2}U^{V+A} \; + \; \frac{3}{4}\,\sin^2\theta_\ell\cdot\frac{1}{2}L^{V+A}\right] \\ &- \; v\cdot\frac{3}{4}\cos\theta_\ell\cdot P^{VA} \; + \; \frac{2m_\ell^2}{q^2}\cdot\frac{3}{4}\cdot\left[U^V + L^V + S^A\right]. \quad (4.3)\end{aligned}$$

In these formulas m_ℓ is the mass of the lepton and $v = \sqrt{1 - 4m_\ell^2/q^2}$; U denotes the unpolarised-transverse contributions, L the longitudinal contributions and S the scalar contribution. The apices V and A represent respectively vector and axial-vector currents, with $X^{V+A} = X^V + X^A$. The authors of Ref. [90] define then the lepton-side forward-backward asymmetry as

$$A^{\ell}_{\mathrm{FB}}(q^2) = -\frac{3}{2}\,\frac{v \cdot P^{VA}}{v^2 \cdot \left(U^{V+A} + L^{V+A}\right) + \frac{2m_\ell^2}{q^2} \cdot 3 \cdot \left(U^V + L^V + S^A\right)}\,. \tag{4.4}$$

For this analysis the massless leptons limit, $m_\ell \to 0$, is used, which is a good approximation except at very low q^2. Combining the previous equations ad taking the massless limit the differential rates simplify to

$$\frac{d\Gamma}{dq^2} = \frac{v^2}{2} \cdot \left(U^{V+A} + L^{V+A}\right) \tag{4.5}$$

and

$$\frac{d\Gamma}{dq^2 d\cos\theta_\ell} = \frac{v^2}{2}\left[\frac{3}{8}\left(1 + \cos^2\theta_\ell\right) U^{V+A} + A^{\ell}_{\mathrm{FB}} \cos\theta_\ell (U^{V+A} + L^{V+A}) + \frac{3}{4}\sin^2\theta_\ell \left(L^{V+A}\right)\right]. \tag{4.6}$$

Equations 4.5 and 4.6 can be then combined to achieve the form

$$\begin{aligned}\frac{d\Gamma}{dq^2 d\cos\theta_\ell} = \frac{d\Gamma}{dq^2}\Bigg[&\frac{3}{8}\left(1 + \cos^2\theta_\ell\right)\frac{U^{V+A}}{U^{V+A} + L^{V+A}} + A^{\ell}_{\mathrm{FB}}\cos\theta_\ell + \\ &\frac{3}{4}\sin^2\theta_\ell \frac{L^{V+A}}{U^{V+A} + L^{V+A}}\Bigg].\end{aligned} \tag{4.7}$$

The amplitude combination in the last term can be viewed as ratio between longitudinal and sum of longitudinal and unpolarised transverse contributions and therefore one can define the longitudinal fraction

$$f_{\mathrm{L}} = \frac{L^{V+A}}{U^{V+A} + L^{V+A}}, \tag{4.8}$$

which leads to the functional form used in the analysis

$$\frac{d\Gamma}{dq^2 d\cos\theta_\ell} = \frac{d\Gamma}{dq^2}\left[\frac{3}{8}\left(1+\cos^2\theta_\ell\right)(1-f_{\rm L}) + A^\ell_{\rm FB}\cos\theta_\ell + \frac{3}{4}\sin^2\theta_\ell f_{\rm L}\right]. \tag{4.9}$$

Using the same steps the proton helicity distribution is given in Ref. [90] as

$$\frac{d\Gamma(\Lambda_b\to\Lambda(\to p\pi^-)\ell^+\ell^-)}{dq^2\, d\cos\theta_h} = {\rm Br}(\Lambda\to p\pi^-)\frac{d\Gamma(\Lambda_b\to\Lambda\,\ell^+\ell^-)}{dq^2}\left(\frac{1}{2}+A^h_{\rm FB}\cos\theta_h\right), \tag{4.10}$$

and $A^h_{\rm FB}$ is defined as

$$A^h_{\rm FB} = \frac{1}{2}\alpha_\Lambda P^\Lambda_z(q^2), \tag{4.11}$$

where $P^\Lambda_z(q^2)$ is the polarisation of the daughter baryon, Λ, and $\alpha_\Lambda = 0.642\pm0.013$ [1] is the Λ decay asymmetry parameter.

These expressions assume that Λ^0_b is produced unpolarised, which is in agreement with the recent LHCb measurement [98]. Possible effects due to a non zero production polarisation are investigated as systematic uncertainties (see Sec. 4.5.5).

4.2 Multi-dimensional angular distributions

To incorporate effects of production polarisation this was introduced in the equations. In the modified version an angle θ is defined as the angle between the Λ direction in the Λ^0_b rest frame with respect to $\hat{n} = \hat{p}_{inc}\times\hat{p}_{\Lambda^0_b}$, where $\hat{p}_{inc}$ represents the direction of the incoming proton. This angle is sensitive to the production polarisation through the spin-density matrix. Integrating over all the angles but θ_ℓ results in the same distribution as in the unpolarised case (Eq. 4.3). Therefore, in the case of uniform efficiency, the lepton side forward-backward asymmetry, $A^\ell_{\rm FB}$, is unaffected by the production polarisation. To be able to estimate the effect of the production polarisation in the case of non-uniform efficiency, the differential distribution in θ and θ_ℓ is also derived, which in the massless leptons limit becomes (up to a constant multiplicative factor)

$$
\begin{aligned}
\frac{d\Gamma(\Lambda_b \to \Lambda\,\ell^+\ell^-)}{dq^2 d\cos\theta d\cos\theta_\ell} =& \frac{d\Gamma}{dq^2}\left\{\frac{3}{8}\left(1+\cos^2\theta_\ell\right)(1-f_{\rm L}) + A^\ell_{\rm FB}\cos\theta_\ell + \frac{3}{4}\sin^2\theta_\ell f_{\rm L} + \right.\\
& P_b\cos\theta\left[-\frac{3}{4}\sin\theta_\ell^2 O_{Lp} + \frac{3}{8}\left(1+\cos\theta_\ell^2\right)O_P\right.\\
& \left.\left.-\frac{3}{8}\cos\theta_\ell O_{UVA}\right]\right\}\,,
\end{aligned}
\tag{4.12}
$$

where three more observables are defined

$$
\begin{aligned}
O_{Lp} =& \frac{L_P^V + L_P^A}{U^{V+A} + L^{V+A}},\\
O_P =& \frac{P^V + P^A}{U^{V+A} + L^{V+A}},\\
O_{UVA} =& \frac{U^{VA}}{U^{V+A} + L^{V+A}}.
\end{aligned}
$$

1575 In the massless leptons approximation two of these quantities are related to the
1576 hadron side forward-backward asymmetry as

$$
\frac{1}{2}\alpha_\Lambda\left(O_P + O_{Lp}\right) = A^h_{\rm FB}\,. \tag{4.13}
$$

Following the same steps as for the lepton case, after integrating over all the angles but θ_h one finds that the hadron side, $A^h_{\rm FB}$, is also unaffected by the production polarisation in case of uniform efficiency. The differential distribution in θ and θ_h has the form

$$
\begin{aligned}
\frac{d\Gamma(\Lambda_b \to \Lambda\,\ell^+\ell^-)}{dq^2 d(\cos\theta) d(\cos\theta_h)} = \frac{d\Gamma}{dq^2}&\left[1 + 2A^h_{\rm FB}\cos\theta_h + P_b\left(O_P - O_{Lp}\right)\cos\theta\right.\\
&\left.+\alpha_\Lambda P_b\left(1-2f_{\rm L}\right)\cos\theta\cos\theta_h\right].
\end{aligned}
\tag{4.14}
$$

1577 In order to use these distributions, expectations for the three additional observables,
1578 which do not enter one-dimensional distributions, are needed. Expectations are
1579 calculated using form factors and numerical inputs from Ref. [90] and are listed in
1580 Appendix A.1 in Tab. A.1.

For completeness, the differential distribution in $\cos\theta_\ell$ andcos θ_h has the form

$$
\begin{aligned}
\frac{d\Gamma(\Lambda_b \to \Lambda\,\ell^+\ell^-)}{dq^2 d\cos\theta_h d\cos\theta_\ell} =& \frac{3}{8} + \frac{6}{16}\cos^2\theta_\ell(1-f_{\mathrm{L}}) - \frac{3}{16}\cos^2\theta_\ell f_{\mathrm{L}} + A^l_{FB}\cos\theta_\ell + \\
& \left(\frac{3}{2}A^h_{\mathrm{FB}} - \frac{3}{8}\alpha_\Lambda O_P\right)\cos\theta_h - \frac{3}{2}A^h_{\mathrm{FB}}\cos^2\theta_\ell\cos\theta_h - \frac{3}{16}f_{\mathrm{L}} + \\
& \frac{9}{16}f_{\mathrm{L}}\sin^2\theta_\ell + \frac{9}{8}\alpha_\Lambda\cos^2\theta_\ell\cos\theta_h O_P - \\
& \frac{3}{2}\alpha_\Lambda\cos\theta_\ell\cos\theta_h O_{UVA}.
\end{aligned} \tag{4.15}
$$

4.3 Angular resolution

This section describes a study of the angular resolution done in order to achieve a better understanding of detector and reconstruction effects. This will be then used to study systematic uncertainties (see Sec. 4.5.5). The study is performed by analysing simulated events and comparing generated and reconstructed quantities. Figure 4.2 shows plots of the difference between true and measured angular observables ($\cos\theta_\ell$ and $\cos\theta_h$) as a function of the observable itself. These are centred at zero indicating no bias in the measurement. Figure 4.3 shows the angular resolution in two-dimensional bins of q^2 and angular observables. In Fig. 4.2 the same differ-

Table 4.1: Average angular resolutions integrated over the full interval and the full available q^2.

Observable	**Downstream**	**Long**
$\cos\theta_\ell$	0.015	0.010
$\cos\theta_h$	0.066	0.014

ence is shown also as a function of q^2 revealing again no bias. The spread of these distributions around the central value can be takes as an estimate of the angular resolution. Taking vertical slices of the plots in Fig. 4.2 one obtains approximately gaussian distributions centred at zero. These are fit with a single gaussian and its width is interpreted as the angular resolution. Table 4.1 reports the average resolutions for the two angular observables separately for long and downstream candidates. As expected candidates built from long tracks are characterised by a better angular

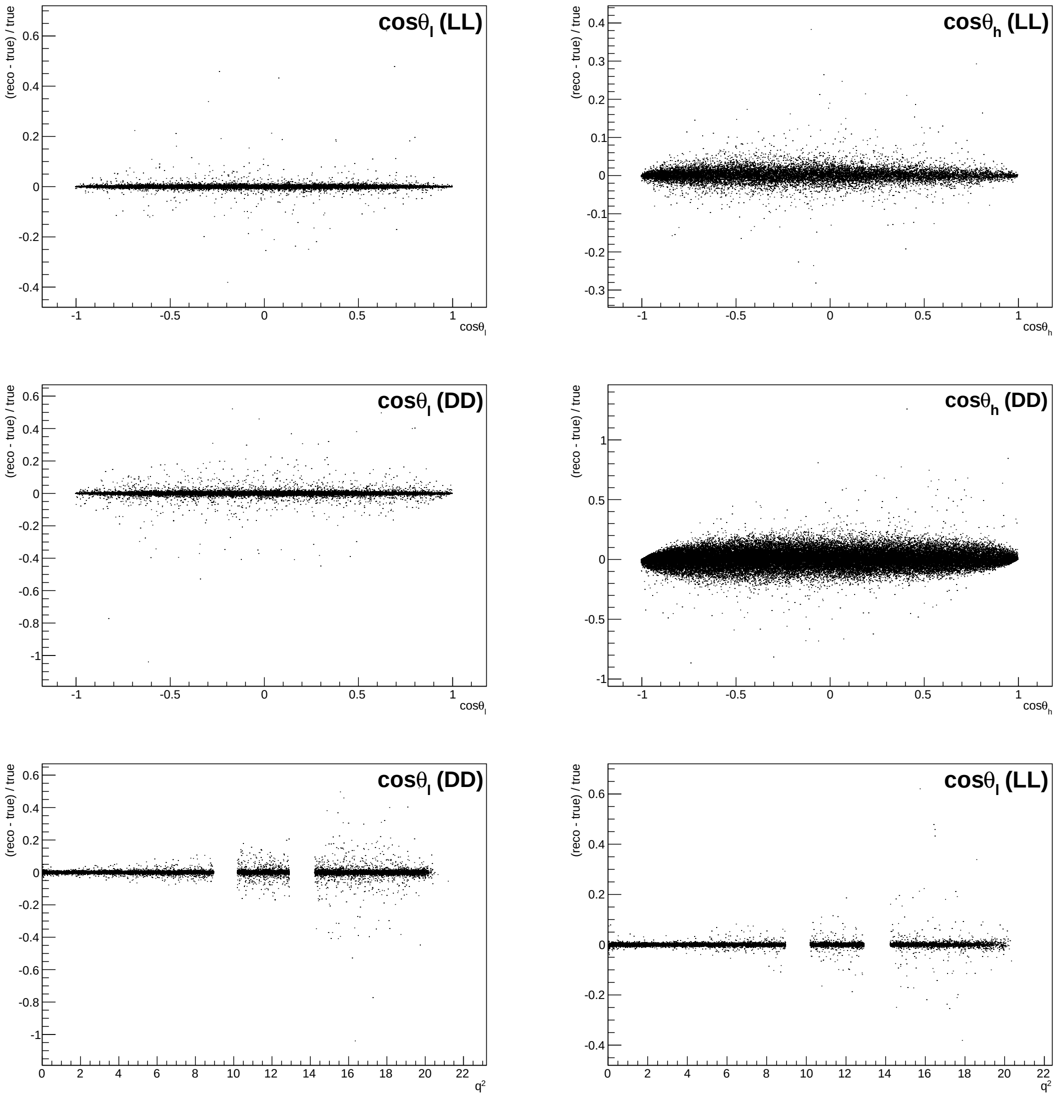

Figure 4.2: Difference of between generated and reconstructed angular observables as a function of the observables themselves for long (top) and downstream (bottom) candidates and as a function of q^2 for long (bottom left) and downstream (bottom right) candidates. As the plots are made using fully selected rare samples the bottom plots present empty bands corresponding to the charmonium vetoes.

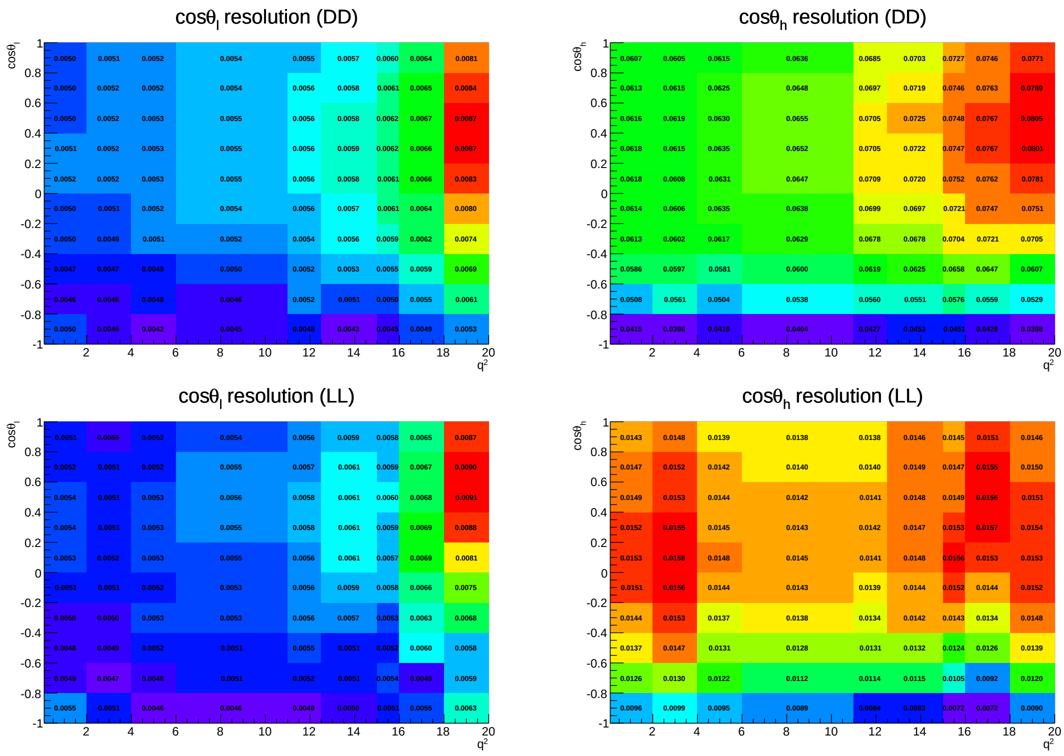

Figure 4.3: Angular resolution for $\cos\theta_\ell$ (left plots) and $\cos\theta_h$ (right plots) as a function of the angular observables and q^2 for downstream (upper plots) and long (lower plots) candidates. White bands correspond to the J/ψ and $\psi(2S)$ resonances which are excluded from the study.

resolution due to a better momentum and vertex resolution.

4.4 Fit strategy

There are physical limits to the values of the parameters of interests: $A^h_{\rm FB}$ is limited in the [-0.5,0.5] interval and for the $f_{\rm L}$ and $A^\ell_{\rm FB}$ parameters the physical region, given by $|A^\ell_{\rm FB}| < 3/4(f_{\rm L} - 1)$, is the triangle shown in Fig. 4.4. If the measured value is close to the border the fit does not always converge. Therefore a "brute force" fitting technique is applied. For this purpose fit parameters are divided into two categories: parameters of interest (PoIs), $A^\ell_{\rm FB}$, $A^h_{\rm FB}$ and $f_{\rm L}$ and all other parameters, which are referred to as "nuisances". The value of the Log-Likelihood ($\log\mathcal{L}$) of the fit model with respect to data is evaluated in a grid of points in the PoIs allowed area to find the function minimum. A first coarse scan finds a candidate minimum and then the procedure is reiterated two more times in finer intervals around it. For each point all the nuisances are fitted using a maximum likelihood fit. Using this method the best fit point is therefore constrained inside the physical region. If the minimum of the log-likelihood is found to be outside it, the closest point on the boundary is chosen as the best fit.

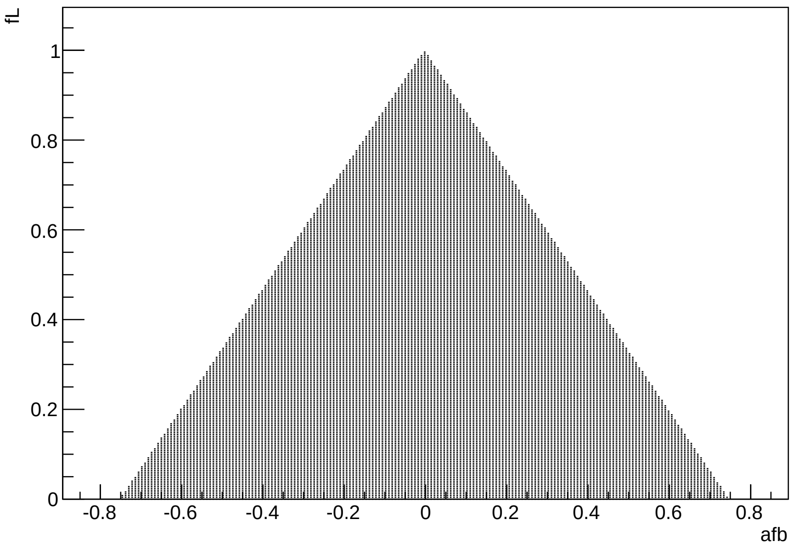

Figure 4.4: The physical (A^{ℓ}_{FB},f_{L}) parameter space. The dark region corresponds to points where the PDF is positive in the whole $[-1, 1]$ interval.

4.4.1 Feldman-cousins plug-in method

Physical boundaries of the parameter space could result in a wrong estimation of the uncertainties, especially if the measured value is close to the border. To deal with this effect in this analysis the likelihood-ordering method [99] is used to estimate uncertainties and nuisance parameters are accounted for using the plug-in method [100]. This is a unified method to calculate confidence intervals and upper/lower limits, based on simulated experiments and has the advantage of having a well defined frequentist coverage.

The method is constituted by the following steps:

1. fit real data distributions with all parameters free;
2. fit real data fixing the PoIs to a value of choice while keeping nuisance parameters free;
3. generate simulated samples following the distribution given by the fit model, where all nuisance parameters are taken from the fit in point 2 and PoIs are fixed to the same value used in point 2;
4. repeat the two fits made on data (points 1 and 2) on each simulated sample: fit with all parameters free and with fixed PoIs;
5. extract the value of the Log-Likelihoods at the minimum for all cases;
6. calculate the percentage of simulated experiments in which the free-to-fixed likelihood ratio is bigger than in data: $\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free} > (\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free})_{data}$;
7. repeat the procedure for many values of the PoIs scanning around the best fit point.

The confidence interval at $k\%$ is given by the points where the free-to-fixed likelihood ratio is bigger in data than simulation for $(1-k)\%$ of times. As an example, Fig. 4.5 shows the p-values obtained with the plug-in method for $A_{\rm FB}^h$ and $f_{\rm L}$. A two-dimensional region can also be scanned giving a grid of p-values, which translates into two-dimensional confidence regions.

4.4.2 Modelling the angular distributions

The observables are extracted from fits to one-dimensional angular distributions. The PDFs used to model the data are defined as

$$P^k(\cos\theta_{\ell/h}) = (1-f_b)P_S(\cos\theta_{\ell/h}) \times \varepsilon^k(\cos\theta_{\ell/h}) + f_b P_B^k(\cos\theta_{\ell/h}), \tag{4.16}$$

where k =(LL,DD). The signal function is composed by a theoretical shape (P_S) given by Eq. 4.10 and 4.9, which is multiplied by an acceptance function ε described

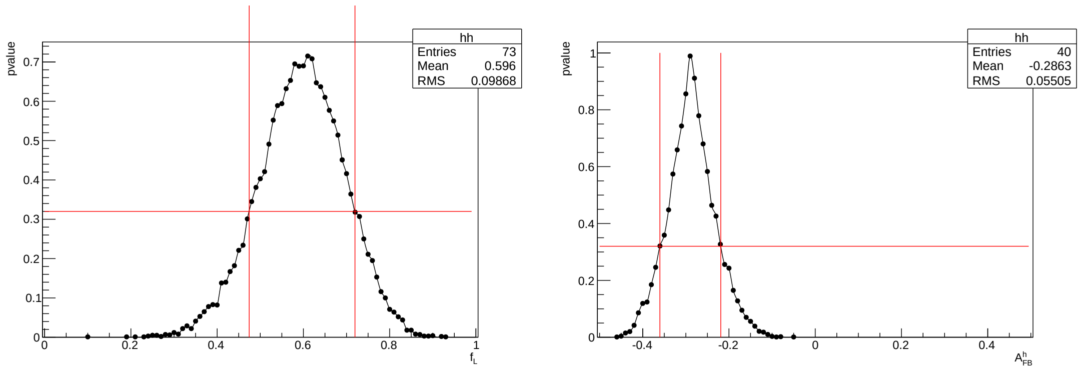

Figure 4.5: Dependence of the p-value from the values of the angular observables f_{L} (left) and A_{FB}^h (right) in simulated experiments. The red lines mark the points at p-value 32% corresponding to a 68% CL.

in Sec. 4.4.3. The background function, P_B, is parameterised with a linear function times the efficiency shape: $P_B^k(\cos\theta_{\ell/h}) = (cx + q) \times \varepsilon^k(\cos\theta_{\ell/h})$. The free parameter of this model is fixed by fitting candidates in the sideband which contains only background. Finally, f_b is the background fraction: $f_b = B/(S+B)$. To limit systematic effects due to the background parameterisation the fit is performed in a restricted mass region around the Λ_b^0 mass peak dominated by the signal: $5580 < m(\Lambda\mu^+\mu^-) < 5660$ MeV/c^2 ("signal region"). The background fraction, f_b, is obtained by looking at the 4-body $m(p\pi\mu\mu)$ invariant mass distribution in a wider interval and fitting it to extract the fraction of background in the signal region. In the fit to the angular distributions this is then gaussian constrained to the obtained value. Figure 4.6 shows the background distributions in the sideband, $m(p\pi\mu^+\mu^-) > 5700$ MeV/c^2, for the high q^2 integrated interval with overlaid the background function. Note that a different acceptance shape is used for downstream and long events and for each q^2 interval. In summary the only free fit parameter in each of the final fits to data is the forward-backward asymmetry (and f_{L} in the leptonic case).

4.4.3 Angular acceptance

Selection requirements on the minimum momentum of the muons may distort the $\cos\theta_\ell$ distribution by removing candidates with extreme values of $\cos\theta_\ell$. Similarly,

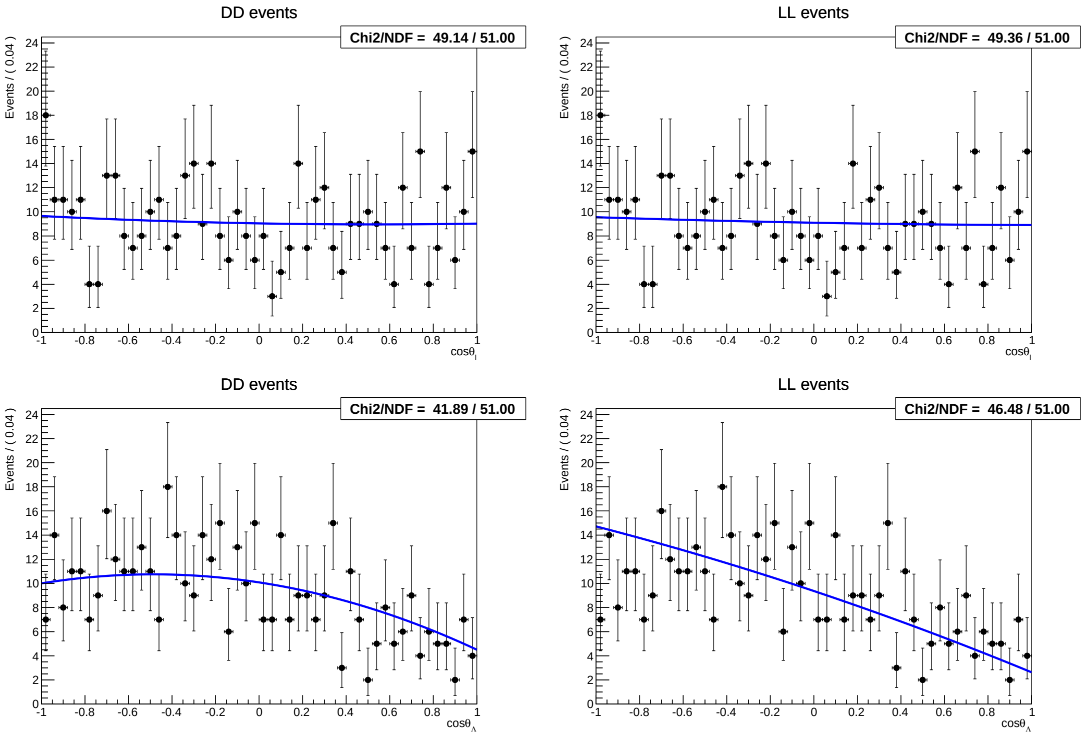

Figure 4.6: Background distribution as a function of $\cos\theta_\ell$ (top) and $\cos\theta_h$ (bottom) for downstream (left) and long (right) candidates in the 15-20 GeV^2/c^4 q^2 interval.

the impact parameter requirements affect $\cos\theta_h$ as very forward hadrons tend to have smaller impact parameter values. While in principle one could take this into account by an additional weight, to minimise the distortion of the uncertainties estimate, the efficiency function is incorporated in the fit model. The angular efficiency is parametrised using a second-order polynomial and determined separately for downstream and long candidates by fitting simulated events, using an independent set of parameters obtained for each q^2 interval. These parameters are then fixed for the fits to data. Using polynomial functions allows to calculate the PDF normalisation analytically. Figure 4.7 shows the acceptance as a function of $\cos\theta_h$ and $\cos\theta_\ell$ for the 15.0–20.0 integrated q^2 interval obtained using a $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulated sample. For the lepton side, even though the efficiency is symmetric by construction, all parameters are left free to float, namely it is not constrained to be symmetric.

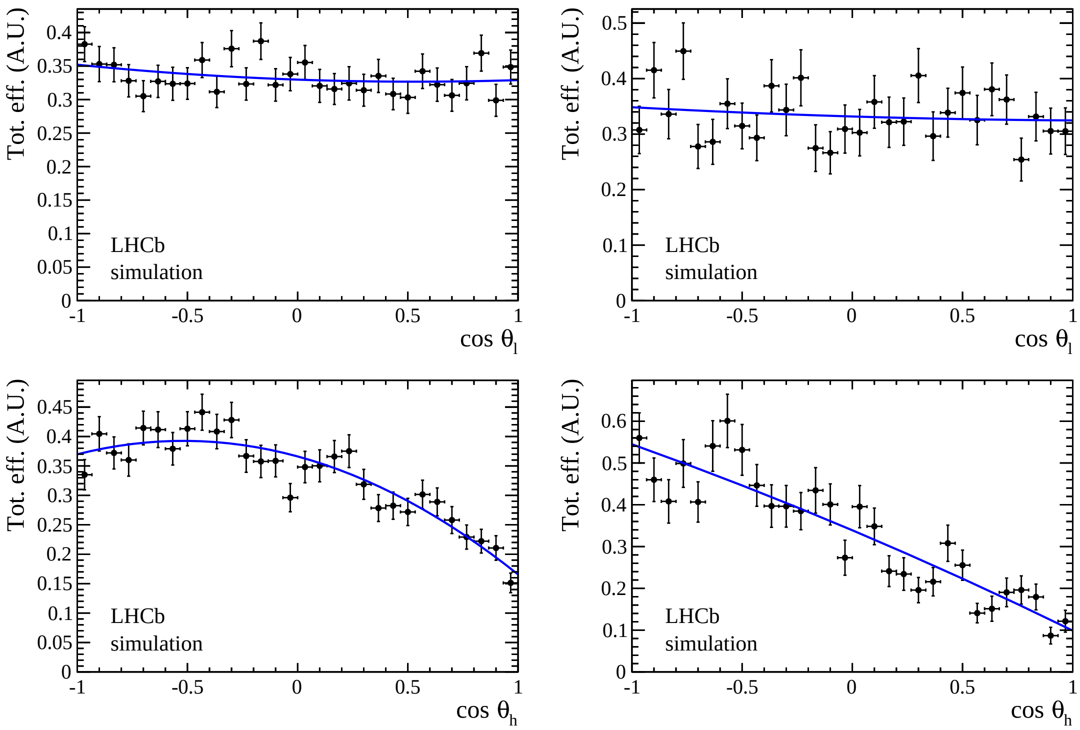

Figure 4.7: Efficiency as a function of $\cos\theta_\ell$ (top) and $\cos\theta_h$ (bottom) for downstream (left) and long (right) candidates in the 15–20 GeV^2/c^4 q^2 interval.

4.4.4 Studies on a three-dimensional fit

One other way of extracting the angular observables would be to fit at the same time both angles and also the invariant mass distribution in order to have a better handle on the level of background. In this case one can use more of the information available. On the other hand it is necessary to use a larger mass window including more background and this method involves more parameters to fit. In the 1D case the free parameters are the two parameters of interest ($A_{\rm FB}^{\ell}$ and $f_{\rm L}$) for the lepton case and one ($A_{\rm FB}^{h}$) for the hadron one. For the 3D case in addition to the three PoIs there are two background fractions and the two exponential slopes for the invariant mass background. Furthermore, to take correctly into account correlations three more observables enter the fit (see Eq. 4.12). As an high number of free parameters is difficult to constrain with the very limited statistics available, pseudo-experiments are used to check which method gives the best sensitivity. Events are generated in a 3D ($\cos\theta_\ell$,$\cos\theta_h$,$m_{p\pi\mu\mu}$) space. The generated values of the PoIs are $A_{\rm FB}^{\ell} = 0$, $f_{\rm L} = 0.7$ and $A_{\rm FB}^{h} = -0.37$, which are data-like values inspired to a preliminary measurement in the highest statistics interval. The overall statistics and the fraction of background events in the mass window are generated to be data-like using information from the preliminary fit to data. Each pseudo-experiment is fitted with both methods and Fig. 4.8 reports distributions of parameters of interest obtained from the fit in the 1D and 3D cases. The RMS of these distributions can be taken as a measure of the sensitivity of each method. Table 4.2 lists the RMSs obtained from both methods; for all parameters of interest the 1D fit method gives a smaller RMS, hence a better sensitivity.

Table 4.2: RMS values for toy experiments on the extraction of the three parameters of interests with the 1D or 3D fitting methods.

q^2 [GeV^2/c^4]	**Fit type**	$A_{\rm FB}^{h}$	$A_{\rm FB}^{\ell}$	$f_{\rm L}$
15.0–20.0	1D	0.070	0.055	0.099
	3D	0.092	0.095	0.153
11.0–12.5	1D	0.142	0.128	0.198
	3D	0.249	0.254	0.303

4.5 Systematics uncertainties on angular observables

The following section describes the five main sources of systematic uncertainties that are considered for the angular observables measurement and, finally, results are reported in Sec. 4.7. Results are derived only for q^2 intervals where the signal significance, shown in Tab. 3.6, is above 3 standard deviations. This includes all q^2 intervals above the J/ψ resonance and the lowest q^2 interval, where an increased yield is due to the presence of the photon pole.

4.5.1 Angular correlations

The angular acceptance is non-flat as a function of $\cos\theta_\ell$ and $\cos\theta_h$. Therefore, while integrating the full angular distribution, terms that cancel with perfect efficiency

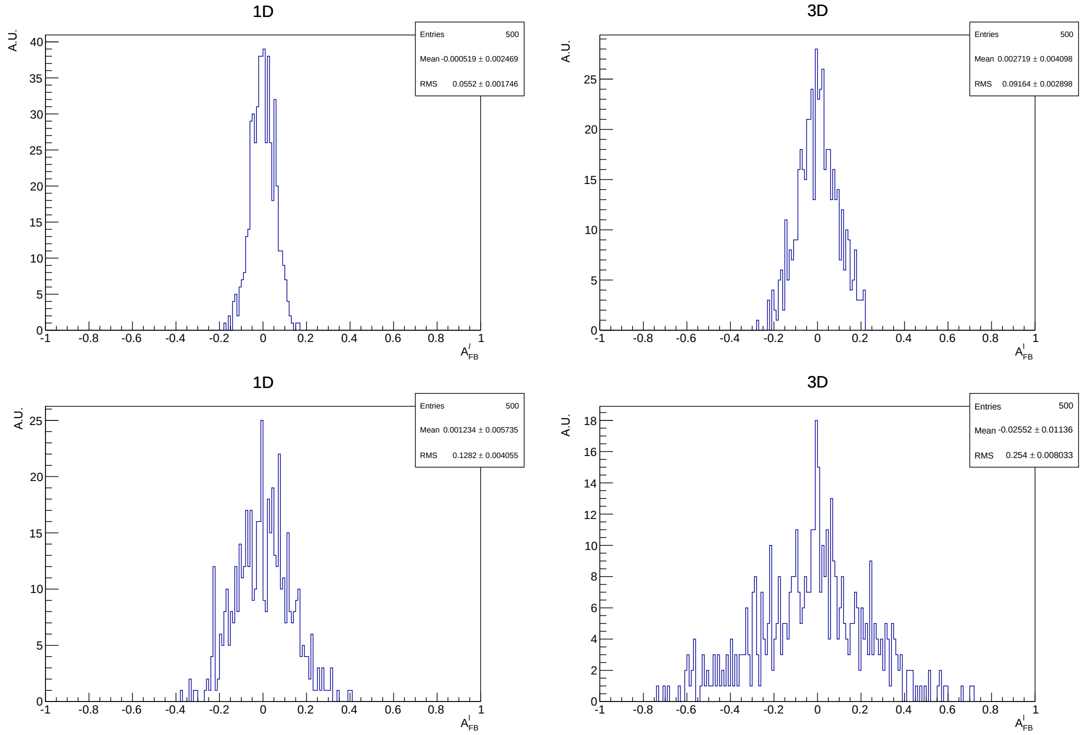

Figure 4.8: Values of the $A_{\rm FB}^{\ell}$ parameter observed over pseudo-experiments with input $A_{\rm FB}^{\ell} = 0$ using the 1D fit method (left) and the 3D one (right). Events are generated with parameters and statistics corresponding to what is observed in the highest statistics interval (top), 15–20 GeV$^2/c^4$, and in the lowest statistics one, 11–12.5 GeV$^2/c^4$.

may remain and generate a bias in the final result. In order to deal with this effect simulated events are generated in a two-dimensional ($\cos\theta_\ell$,$\cos\theta_h$) space according to the theoretical distribution described by Eq. 4.15 multiplied by the two-dimensional efficiency function obtained from simulation. Then, one-dimensional projections are taken and fit using the default one-dimensional efficiency functions. Figure 4.9 shows the distribution of observed deviations from the generated value, $\Delta x = x_{true} - x_{measured}$. Since the mean of these distributions is non-zero by more then 3σ, they are taken as a systematic uncertainties.

4.5.2 Resolution

The angular resolution could bias the observables measurement generating an asymmetric migration of events. This is especially important in the $\cos\theta_h$ case, because

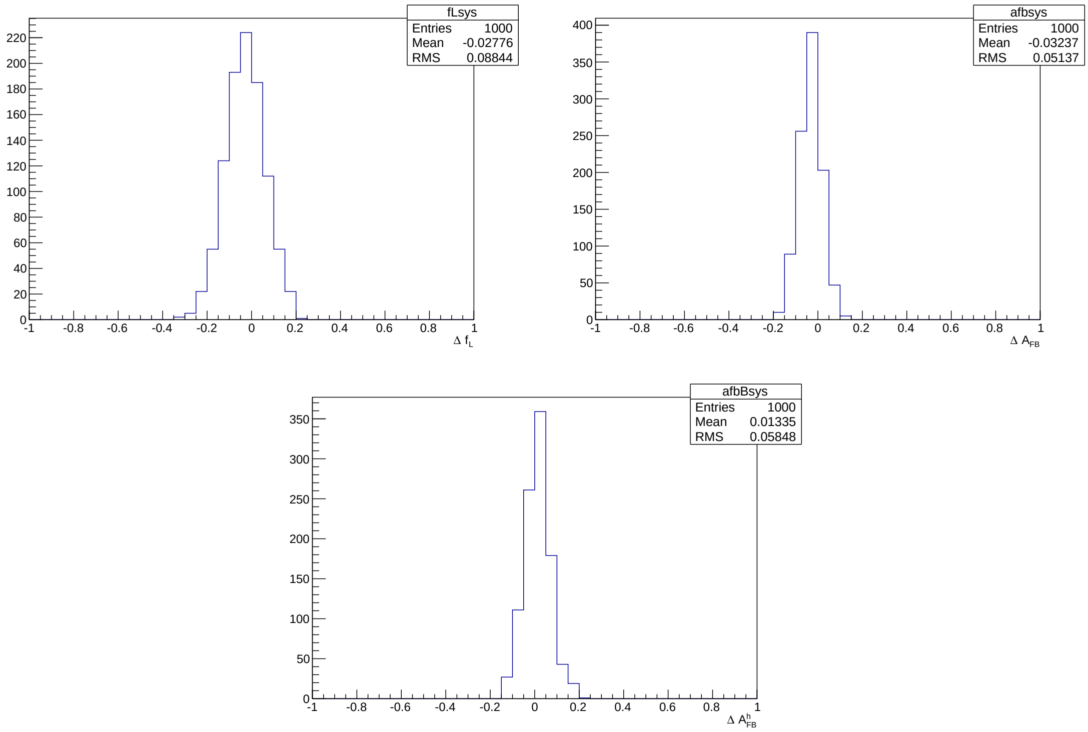

Figure 4.9: Deviations of the observables' values obtained fitting simulated events generated with a 2D distribution multiplied by a 2D efficiency and fitting 1D projections with respect to generated values. For f_{L} (top left), A^{ℓ}_{FB} (top right) and A^{h}_{FB} (bottom).

it has worse resolution and a considerably asymmetric distribution. Simulated experiments are used to asses this systematic. Events are generated according to the measured distributions including efficiencies. The generated events are then smeared by the angular resolution (gaussian smearing). To be conservative the case with biggest angular resolution, downstream candidates, is always used. Finally, the smeared and not-smeared distributions are fit with the same PDF. The average deviation from the default values are reported in Tab. 4.3 as a function of q^2 and assigned as systematic uncertainties.

4.5.3 Efficiency description

An imprecise determination of the reconstruction and selection efficiency can introduce an extra oddity and therefore bias the measurement. To asses this effect

Table 4.3: Values of simulated $\cos\theta_\ell$ and $\cos\theta_\Lambda$ resolutions and systematic uncertainties on angular observables due to the finite resolution in bins of q^2.

q^2 [GeV$^2/c^4$]	σ_ℓ	σ_Λ	$\Delta A^\ell_{\rm FB}$	$\Delta f_{\rm L}$	$\Delta A^h_{\rm FB}$
0.1 – 2.0	0.0051	0.061	0.0011	-0.0022	-0.007
11.0 – 12.5	0.0055	0.067	0.0016	-0.0051	-0.013
15.0 – 16.0	0.0059	0.070	0.0006	-0.0054	-0.010
16.0 – 18.0	0.0064	0.070	0.0014	-0.0077	-0.010
18.0 – 20.0	0.0081	0.074	0.0014	-0.0062	-0.010
15.0 – 20.0	0.0066	0.072	0.0013	-0.0076	-0.011

the kinematic re-weighting described in Sec. 3.3.2 is removed from the simulation and the efficiency is determined again. Simulated events are then fit using the same theoretical PDF but multiplied by the efficiency function obtained with and without kinematical weights. As in the previous cases the average bias is taken as systematic uncertainty; results are shown in Tab. 4.4. Furthermore, the effect of the limited simulated statistics is taken into account and added to the systematic uncertainty.

Table 4.4: Values systematic uncertainties due to limited knowledge of the efficiency function on the three angular observables in bins of q^2

q^2 [GeV$^2/c^4$]	$A^\ell_{\rm FB}$	$f_{\rm L}$	$A^h_{\rm FB}$
0.1 – 2.0	0.0020	0.0440	0.0093
11.0 – 12.5	0.0069	0.0027	0.0069
15.0 – 16.0	0.0018	0.0046	0.0109
16.0 – 18.0	0.0012	0.0043	0.0159
18.0 – 20.0	0.0030	0.0017	0.0148
15.0 – 20.0	0.0002	0.0046	0.0138

Table 4.5: Values of systematic uncertainties due to the statistics of the simulated samples on the three angular observables in bins of q^2.

q^2 [GeV$^2/c^4$]	$A^\ell_{\rm FB}$	$f_{\rm L}$	$A^h_{\rm FB}$
0.1 – 2.0	0.00151	0.00170	0.00213
11.0 – 12.5	0.00121	0.00154	0.00196
15.0 – 16.0	0.00004	0.00017	0.00103
16.0 – 18.0	0.00065	0.00246	0.00417
18.0 – 20.0	0.00023	0.00372	0.00162
15.0 – 20.0	0.00039	0.00091	0.00137

4.5.4 Background parameterisation

There is a certain degree of arbitrariety in the choice of a parameterisation for the background, especially for q^2 intervals with low statistics. To asses possible biases due to the PDF choice, simulated experiments are generated using the shapes from data fits and the same statistics as observed in data for each q^2 interval. Each pseudo-experiment is fit with two models: the default one, a “line times efficiency” function and the efficiency function alone, corresponding to the assumption that background distributions are originally flat and only modified by the interaction with the detector. The average bias with respect to the default model is taken as systematic uncertainty; results are reported in Tab. 4.6.

Table 4.6: Values of systematic uncertainties due to the choice of background parameterisation in bins of q^2.

q^2 [GeV^2/c^4]	A^{ℓ}_{FB}	f_{L}	A^{h}_{FB}
0.1 – 2.0	0.003	0.049	0.053
11.0 – 12.5	0.045	0.034	0.035
15.0 – 16.0	0.010	0.038	0.026
16.0 – 18.0	0.026	0.036	0.022
18.0 – 20.0	0.011	0.031	0.025
15.0 – 20.0	0.007	0.014	0.017

4.5.5 Polarisation

To study the effect of a non-zero Λ_b^0 production polarisation simulated events are generated using the distributions given by Eqs. 4.12 and 4.14 as a function of the angle under study ($\cos\theta_\ell$ or $\cos\theta_h$) and $\cos\theta$, defined in Sec. 4.2, which is sensitive to polarisation. Similarly to the procedure used for the branching ratio measurement, events are generated using values of the polarisation corresponding to $\pm\sigma$ from the LHCb measurement [91]. In the theoretical functions $\cos\theta$ is always odd therefore with perfect efficiency it always drops out by integrating over $\cos\theta$. Therefore

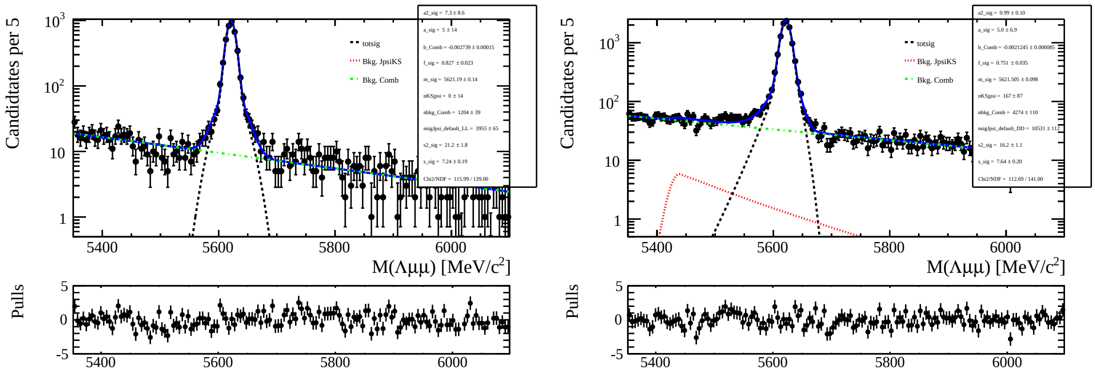

Figure 4.10: Invariant mass distribution of $\Lambda_b^0 \to J/\psi\ \Lambda$ long (left) and downstream (right) candidates with an extra proton PID cut to remove K_S^0 background.

the generated distributions are also multiplied by the two-dimensional efficiency function. No significant bias is found.

4.6 J/ψ cross-check

To cross-check the fitting procedure this is applied on the high statistics $\Lambda_b^0 \to J/\psi\,\Lambda$ sample. For this purpose events are selected with an additional requirement on the proton PID, `PIDp`> 10. This is needed to reduce the $B^0 \to K_S^0 J/\psi$ background, which is particularly important for the hadronic side fit, since the K_S^0 events are not distributed in a flat way in the $\cos\theta_h$ variable and would therefore bias the fit. Figure 4.10 shows the invariant mass distributions after this requirement is applies, which can be compared with the ones in Fig. 3.12. After the PID cut there are 0.2% of K_S^0 events left in the downstream sample and a fraction compatible with zero in the long sample. The signal model used for this fit is the same used for the rare case and described in Sec. 4.4.2. For the background instead the higher statistics allows to leave more freedom to the fit. Therefore a second-order Chebyschev polynomial is used, where the two parameters are free to vary. As for the rare case the background fractions are gaussian-constrained to what found from the invariant mass fit. Figures 4.11 and 4.12 show fitted angular distributions for the J/ψ channel. The measured values of the observables are $A_{\rm FB}^{\ell} = -0.002^{+0.011}_{-0.011}$, $A_{\rm FB}^{h} = -0.402^{+0.010}_{-0.009}$ and $f_{\rm L} = 0.485^{+0.019}_{-0.020}$, where the uncertainties are 68% Feldman

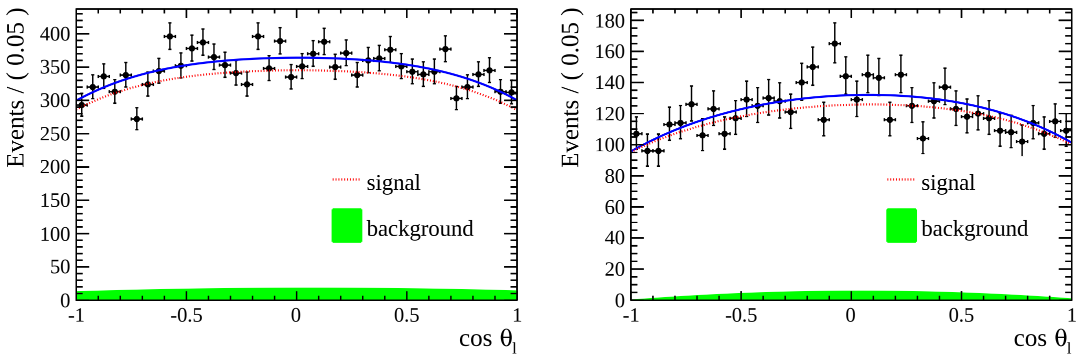

Figure 4.11: Fitted angular distribution as a function of $\cos\theta_\ell$ for $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

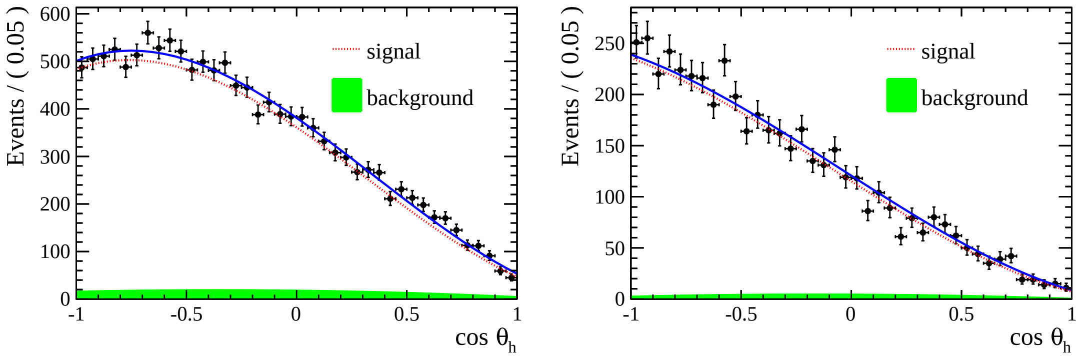

Figure 4.12: Fitted angular distribution as a function of $\cos\theta_h$ for $\Lambda_b^0 \to J/\psi\,\Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

Cousins confidence intervals. The lepton side asymmetry as expected is measured to be zero.

4.7 Results

Figures 4.13 and 4.14 show fits to the angular distributions for the 15-20 GeV^2/c^4 q^2 interval and Tab. 4.7 reports measured values of $A^\ell_{\rm FB}$, $A^h_{\rm FB}$ and $f_{\rm L}$. The asymmetries are also shown in Fig. 4.15 together with SM predictions obtained from Ref. [94]. The statistical uncertainties on these tables are obtained using the likelihood-ratio ordering method described in Sec. 4.4.1, where only one of the two observables is treated as the PoI at a time. The statistical uncertainties on $A^\ell_{\rm FB}$ and $f_{\rm L}$ are also reported in Fig. 4.16 as two-dimensional 68 % confidence level (CL) regions, where

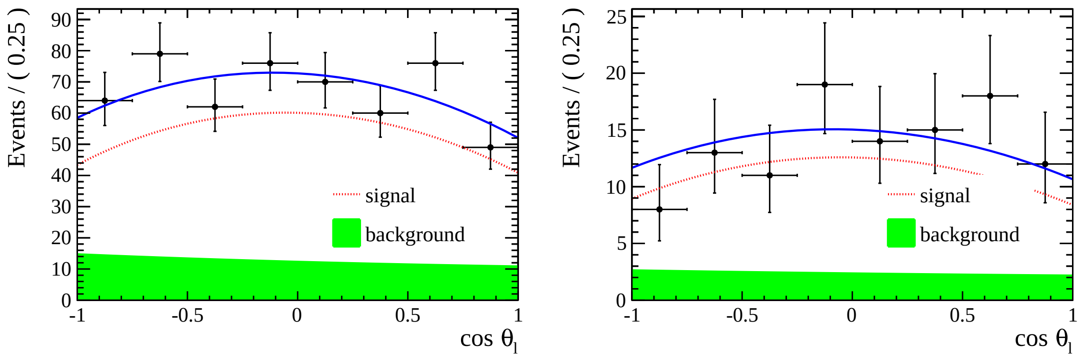

Figure 4.13: Fitted angular distributions as a function of $\cos\theta_\ell$ for downstream (left) and long (right) candidates in the 15–20 GeV$^2/c^4$ q^2 interval.

the likelihood-ratio ordering method is applied by varying both observables and therefore taking correlations into account. Total systematic uncertainties correspond to the square root sum of the single considered sources.

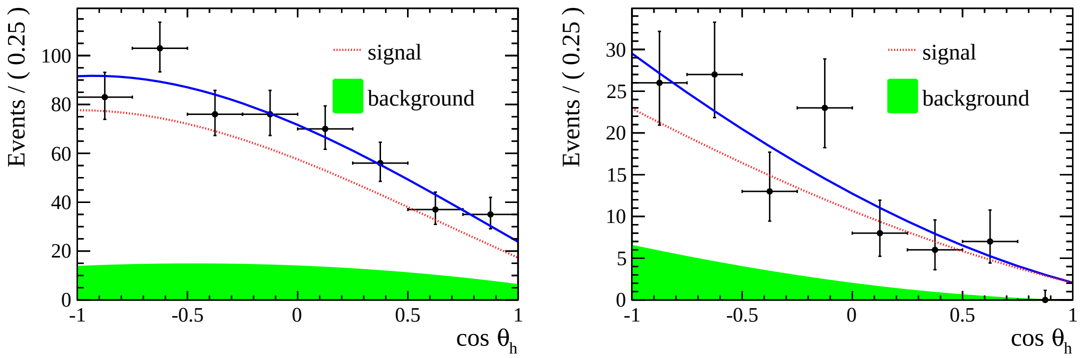

Figure 4.14: Fitted angular distributions as a function of $\cos\theta_h$ for downstream (left) and long (right) candidates in the 15–20 GeV$^2/c^4$ q^2 interval.

Table 4.7: Measured values of leptonic and hadronic angular observables, where the first uncertainties are statistical and the second systematic.

q^2 interval [GeV$^2/c^4$]	$A_{\rm FB}^{\ell}$	$f_{\rm L}$	$A_{\rm FB}^{h}$
0.1 – 2.0	$0.37\,^{+0.37}_{-0.48} \pm 0.03$	$0.56\,^{+0.23}_{-0.56} \pm 0.08$	$-0.12\,^{+0.31}_{-0.28} \pm 0.15$
11.0 – 12.5	$0.01\,^{+0.19}_{-0.18} \pm 0.06$	$0.40\,^{+0.37}_{-0.36} \pm 0.06$	$-0.50\,^{+0.10}_{-0.00} \pm 0.04$
15.0 – 16.0	$-0.10\,^{+0.18}_{-0.16} \pm 0.03$	$0.49\,^{+0.30}_{-0.30} \pm 0.05$	$-0.19\,^{+0.14}_{-0.16} \pm 0.03$
16.0 – 18.0	$-0.07\,^{+0.13}_{-0.12} \pm 0.04$	$0.68\,^{+0.15}_{-0.21} \pm 0.05$	$-0.44\,^{+0.10}_{-0.05} \pm 0.03$
18.0 – 20.0	$0.01\,^{+0.15}_{-0.14} \pm 0.04$	$0.62\,^{+0.24}_{-0.27} \pm 0.04$	$-0.13\,^{+0.09}_{-0.12} \pm 0.03$
15.0 – 20.0	$-0.05\,^{+0.09}_{-0.09} \pm 0.03$	$0.61\,^{+0.11}_{-0.14} \pm 0.03$	$-0.29\,^{+0.07}_{-0.07} \pm 0.03$

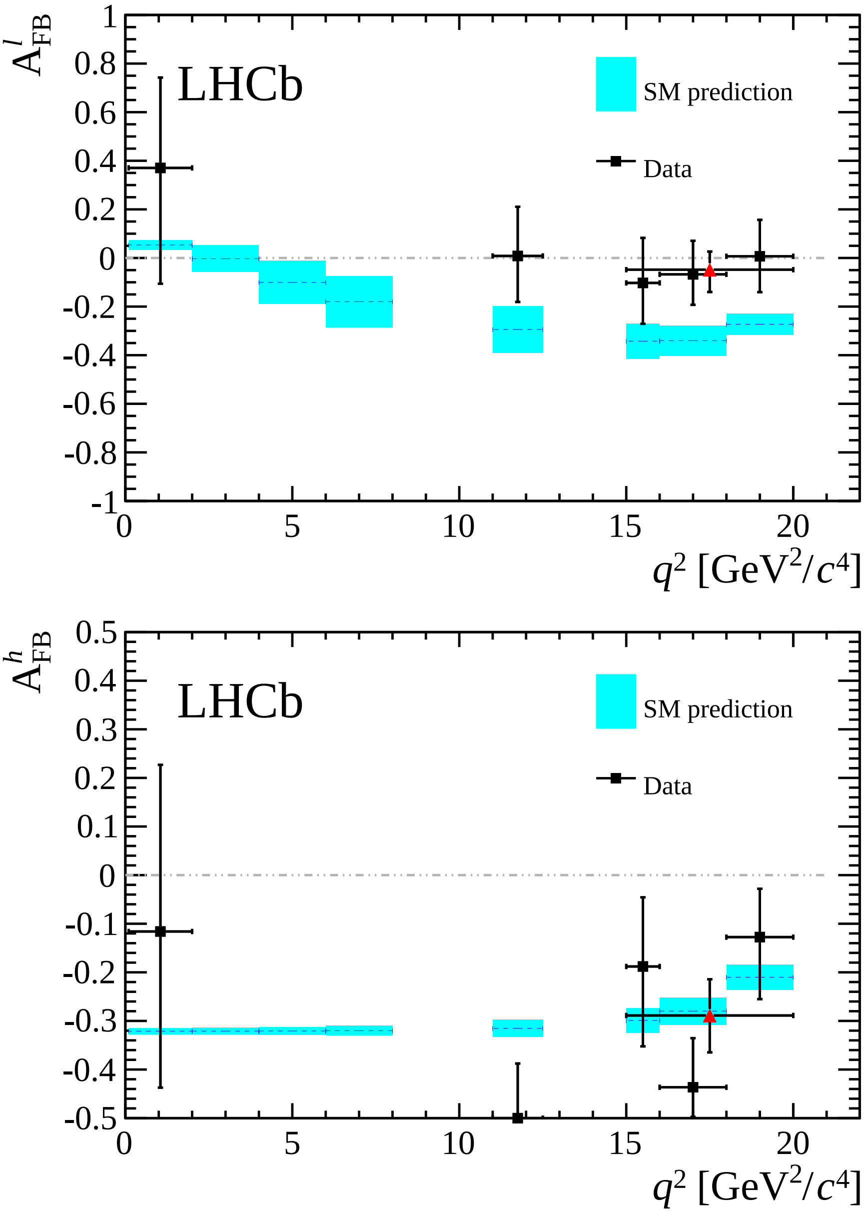

Figure 4.15: Measured values of the leptonic (top) and the hadronic (bottom) forward-backward asymmetries in bins of q^2. Data points are only shown for q^2 intervals where a statistically significant signal yield is found, see text for details. The (red) triangle represents the values for the $15 < q^2 < 20$ GeV^2/c^4 interval. Standard Model predictions are obtained from Ref. [101].

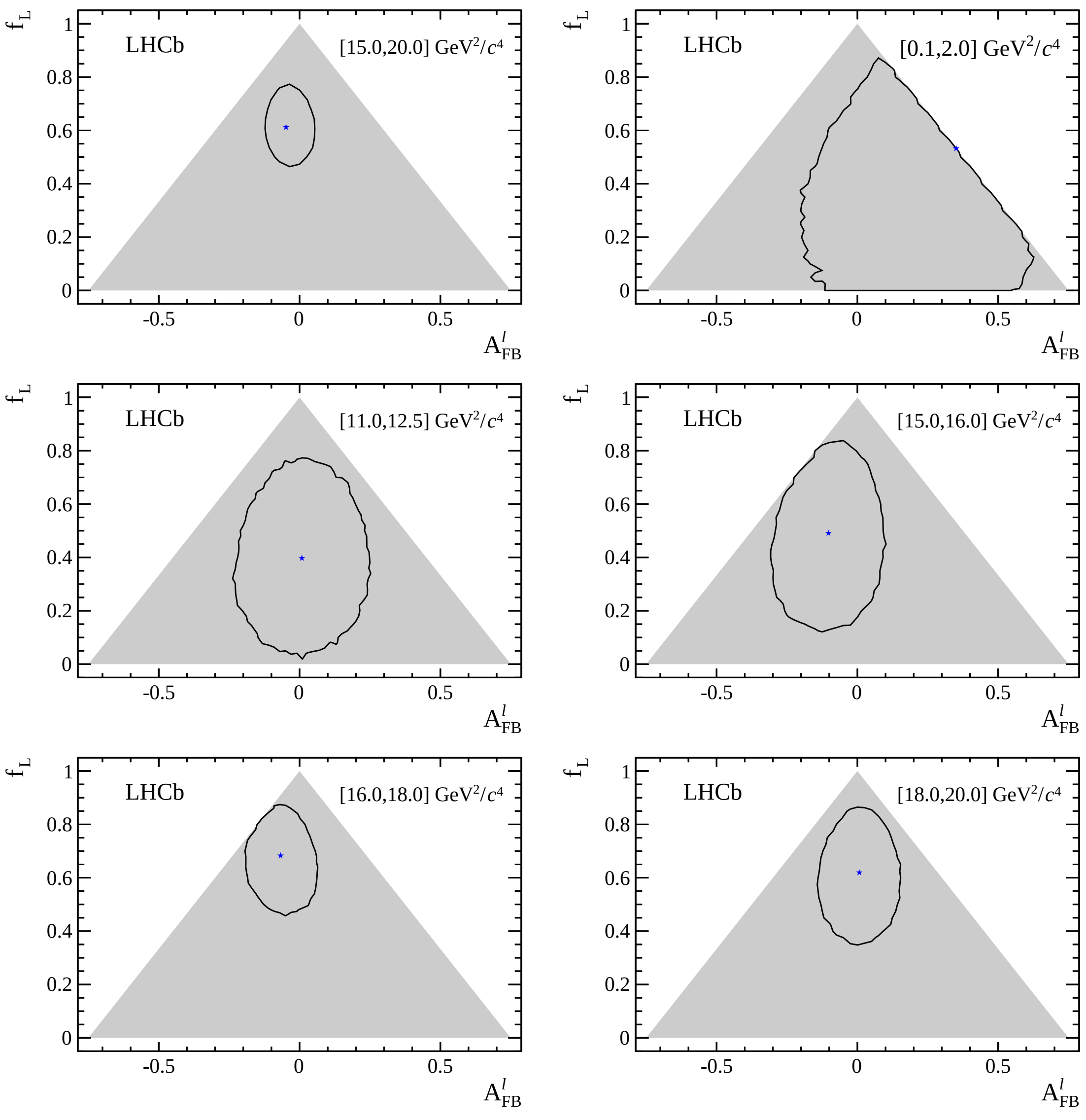

Figure 4.16: Two-dimensional 68 % CL regions (black) as a function of A^{ℓ}_{FB} and f_{L}. The shaded areas represent the regions in which the PDF is positive over the complete $\cos\theta_\ell$ range. The best fit points are indicated by the (blue) stars.

CHAPTER 5

Testing lepton flavour universality with $R_{K^{*0}}$

Lepton Favour Universality (LFU) is the equality of the weak coupling constants for all leptons. FCNC processes, which are forbidden in the SM at tree level and happen only at loop level, are an ideal laboratory to study LFU as new physics in the loops could break the flavour symmetry.

In this work $b \to s\mu^+\mu^-(e^+e^-)$ decays are studied to test LFU between electrons and muons using penguin decays. In particular, the B^0 meson semileptonic decays $B^0 \to K^{*0}\ell^+\ell^-$ are considered. Figure 5.1 shows the possible Feynman diagrams producing such decays while Fig. 5.2 illustrates how these Feynman diagrams may include new particles. A series of recent LHCb measurements [24] points to a tension with SM predictions, which make these processes very interesting to better understand the nature of the discrepancy.

In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger proposed the measurement of the R_H ratio [102], defined in Eq. 5.1, where H can be an

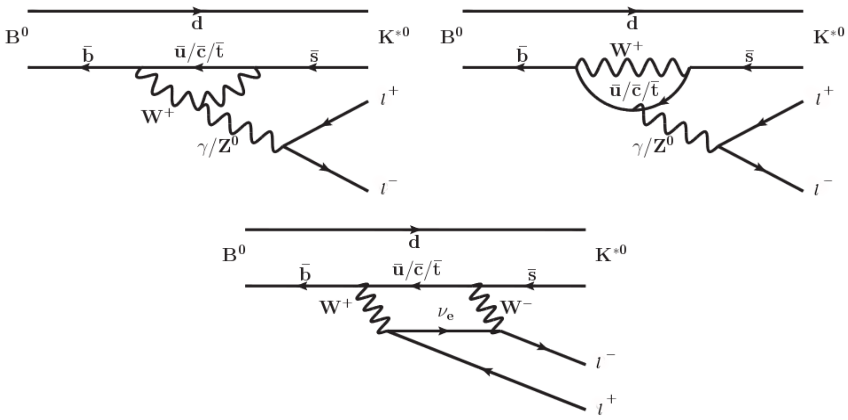

Figure 5.1: Loop diagrams of the $B^0 \to K^{(*0)}\ell^+\ell^-$ process.

inclusive state containing an s quark (X_s) or an s-quark resonance like K or K^{*0}.

$$R_H = \frac{\int_{4m_\mu^2}^{m_b} \frac{d\mathcal{B}(B^0 \to H\mu^+\mu^-)}{dq^2}}{\int_{4m_\mu^2}^{m_b} \frac{d\mathcal{B}(B^0 \to He^+e^-)}{dq^2}} dq^2 \tag{5.1}$$

In this quantity the differential branching ratio is integrated over the squared dilepton invariant mass, q^2, from $q^2_{min} = 4m^2_\mu$, which is the threshold for the $\mu\mu$ process, up to $q^2_{max} = m^2_b$.

The advantage of using ratios of branching fractions as observables is that, in the theoretical prediction, hadronic uncertainties cancel out. Furthermore, experimentally, some of the systematic uncertainties on the ratios are reduced giving a better measurement. For example, what is measured is the number of $\mu\mu$ and ee decays happening in a certain period of time. Then, the luminosity, $\mathcal{L}$, is used to obtain a cross section, σ, using $R = \mathcal{L}\sigma$, where R is the rate at which the decays occur. The luminosity measurement is usually a source of systematic uncertainty, but it appears on both sides of the ratio and therefore cancels out.

Since the SM does not distinguish between lepton flavours, the predicted value of the ratio is $R_H = 1$, under the assumption of massless leptons. Taking into account effects of order m^2_μ/m^2_b Hiller and Kruger calculate that in the SM and in the full q^2

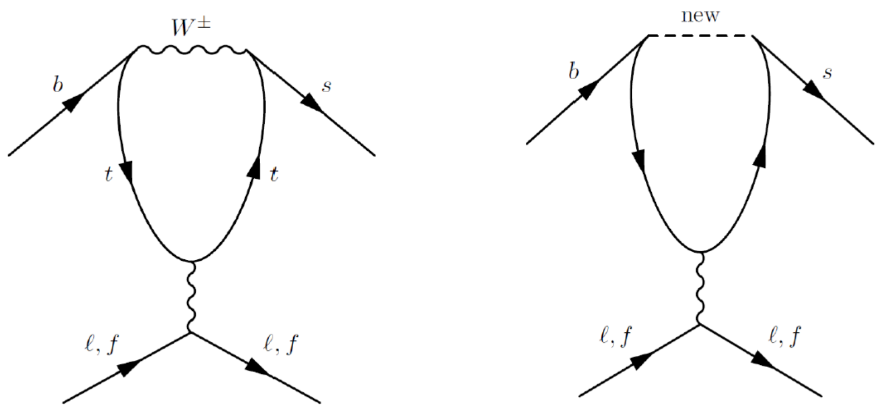

Figure 5.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

range [102]:

$$R_{X_s} = 0.987 \pm 0.006, \tag{5.2}$$
$$R_K = 1.0000 \pm 0.0001, \tag{5.3}$$
$$R_{K^{*0}} = 0.991 \pm 0.002; \tag{5.4}$$
$$\tag{5.5}$$

under the assumptions that:

- right-handed currents are negligible;
- (pseudo-)scalar couplings are proportional to the lepton mass;
- there are no CP-violating phases beyond the SM.

The measurement of the R_H ratios is of particular interest after the recent measurement of the branching ratio of the $B_s^0 \to \mu^+\mu^-$ decay [35], where no evidence of new physics was found. In fact the quantities $(R_H - 1)$ and $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ remain proportional with

$$\frac{R_H - 1}{\mathcal{B}(B_s^0 \to \mu^+\mu^-)} \sim 2 \cdot 10^{-5}. \tag{5.6}$$

A joint measurement of this two quantities can give much information and constrain MFV models. If $R_H = 1$ and $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is close to the SM prediction as it is measured to be this will allow to put strong constraints on extensions of the SM. If instead $R_H > 1$ and the equation above is not verified, this would mean that one of the assumptions listed above are not verified, which can happen is some extensions of the SM as Super-Symmetric models with broken R-parity. A series of recent LHCb measurements [24] shows tensions with SM predictions, which makes it interesting to further investigate these processes.

5.1 Combining ratios

The full power of the R_H ratios in understanding new physics scenarios comes from their combinations. In Ref. [103] Hiller and Schmaltz propose the measurement of the double ratios, $X_H = R_H/R_K$, which not only can test LFU but also allow to disentangle the kind of new physics that lies behind. These ratios are in fact sensitive to FCNCs of right-handed currents. Furthermore, in Ref. [103] the study is extended to B_s^0 decays such as $B_s^0 \to \phi\ell^+\ell^-$ or $B_s^0 \to \eta\ell^+\ell^-$.

Parity and Lorentz invariance require that the Wilson Coefficients with left-handed chirality (C) and their right-handed counterparts (C') appear in the decay amplitude of exclusive decays in determined combinations, e.g.

$$\begin{aligned} C + C' : &\quad K, K^*_\perp, ... \\ C - C' : &\quad K_0(1430), K^*_{0,\|}, ... \end{aligned} \tag{5.7}$$

where the labels for the K^* meson represent its longitudinal (0), parallel ($\|$) and perpendicular ($\perp$) transversity components. The C contributions are universal to all decays and therefore X_H double ratios are sensitive to right-handed currents. In fact the R_H ratios can be expressed in terms of their deviation from unity as

$$\begin{aligned} R_K &\simeq 1 + \Delta_+, \\ R_{K_0(1430)} &\simeq 1 + \Delta_-, \\ R_K^* &\simeq 1 + p(\Delta_- - \Delta_+) + \Delta_+, \end{aligned} \tag{5.8}$$

where the $\Delta_\pm$ quantities are combinations of Wilson coefficients described in Eq. 10 of Ref. [103] and the parameter p is the polarisation of K^* that in Ref. [103] is determined to be close to 1 simplifying the formula to $R_{K^*} \simeq 1 + \Delta_-$. In particular one can observe the following correlations:

- $R_K < 1$, as it is measured to be, and $X_{K^*} > 1$ points to dominant BSM contributions into C_{LR} (see definition in Sec. 1.5.2);
- a SM like $R_K \sim 1$ together with $X_{K^*} \neq 1$ requires BSM with $C_{LL} + C_{RL} \simeq 0$;
- $R_K \neq 1$ and $X_{K^*} \simeq 1$ corresponds to new physics in C_{LL}.

5.2 Experimental status

The R_K and $R_{K^{*0}}$ ratios have already been measured at the B-factories [104, 105], and the R_K ratio has been recently measured also at LHCb [106] in the $1 < q^2 < 6$ GeV$^2/c^4$ q^2 interval, which represents the most precise measurement to date. This measurement manifests a 2.6σ deviation from the SM prediction. The current experimental status is summarised in Tab. 5.1. By profiting of the large dataset collected during Run-I, the LHCb experiment is expected to reduce the uncertainty on $R_{K^{*0}}$ by at least a factor of 2 with respect to the B-factories.

Table 5.1: Experimental status of the $R_{K^{(*)}}$ measurements.

Ratio	Belle	BaBar	LHCb
R_K	$1.06 \pm 0.48 \pm 0.05$	$1.38^{+0.39+0.06}_{-0.41-0.07}$	$0.745^{+0.090}_{-0.074} \pm 0.036$
$R_{K^{*0}}$	$0.93 \pm 0.46 \pm 0.12$	$0.98^{+0.30+0.08}_{-0.31-0.08}$	—

5.3 Analysis strategy

The aim of the analysis in this chapter is to measure the $R_{K^{*0}}$ ratio using pp collision data collected by the LHCb detector in 2011 and 2012, corresponding to 3 fb^{-1} of integrated luminosity. The $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$, "rare channels", are reconstructed via the K^{*0} decay into a kaon and a pion with opposite charges.

The analysis has to separate signal candidates from background candidates which have similar observed properties. The selection presented in Sec. 5.6 aims to maximise the yield while minimising the background contamination. Two types of backgrounds are identified: "peaking background" and "combinatorial background". The first comes from the mis-reconstruction of other decays or from partially reconstructed events. This type of background, because its specific kinematic properties, usually peaks in some variable, such as the invariant mass of all final particles. Therefore these candidates can be removed using specific cuts. The combinatorial background instead comes from the random combination of particles and can be lowered selecting events with good-quality tracks and vertices.

To further reduce the systematic uncertainties the measurement is performed as the double ratio

$$R_{K^{*0}} = \frac{N_{B^0\to K^{*0}\mu^+\mu^-}}{N_{B^0\to K^{*0}J/\psi\to\mu^+\mu^-}} \cdot \frac{N_{B^0\to K^{*0}J/\psi\to e^+e^-}}{N_{B^0\to K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0\to K^{*0}J/\psi\to\mu^+\mu^-}}{\varepsilon_{B^0\to K^{*0}\mu^+\mu^-}} \cdot \frac{\varepsilon_{B^0\to K^{*0}e^+e^-}}{\varepsilon_{B^0\to K^{*0}J/\psi\to e^+e^-}}, \tag{5.9}$$

where decays reaching the same final states as the rare channels via a J/ψ resonance, $B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$, also referred to as "charmonium" or "resonant" channels, are used as control samples. These decays are distinguished from the rare channels using the invariant mass of the dilepton pair.

As new physics is expected not to affect charmonium resonances the ratio of the J/ψ channels is 1 and therefore $R'_{K^{*0}} = R_{K^{*0}} \times R_{J/\psi} = R_{K^{*0}}$. On the other hand using the relative efficiencies between the rare and resonant channels allows to cancel out many effects resulting in a better control of systematic uncertainties. For brevity, the rare channels will also be denoted as "$\ell\ell$", or specifically "ee" and "$\mu\mu$", and the resonant channels as "$J/\psi\,(\ell\ell)$", or "$J/\psi\,(ee)$" and "$J/\psi\,(\mu\mu)$".

5.4 Dilepton invariant mass intervals

Three q^2 intervals are considered in this work:

- the "low-q^2 " region between 0.0004 and 1.1 GeV^2/c^4, where the $b \to s\ell^+\ell^-$ process is dominated by the photon pole;
- the "central-q^2 " region, [1.1,6.0] GeV^2/c^4;
- the "high-q^2 "region, above 15 GeV^2/c^4.

The central-q^2 region is the most interesting place to look for new physics. In fact, at low q^2, below 1 GeV^2/c^4 the photon pole dominates leaving little space for new physics to be found 1.5.3. The choice of the lower limit of the low-q^2 bin is driven by the need to reject the background due to the $B^0 \to K^{*0}\gamma$ decay where the photon converts into electrons in the material. The lower bound of the central interval is set at 1.1 GeV^2/c^4, in order to exclude the contribution from $\phi \to \ell^+\ell^-$ decays, that can dilute new physics effects. The upper bound of the central interval is chosen to be sufficiently far away from the J/ψ radiative tail, where predictions cannot be cleanly obtained. The 6–15 GeV^2/c^4 region is characterised by the presence of the narrow peaks of the J/ψ and $\psi(2S)$ resonances. The lower bound of the high-q^2 region, where the signal in the electron channel is still unobserved, is chosen to be sufficiently far from the $\psi(2S)$ resonance. Rare and normalisation channels are selected depending on the q^2 interval they fall into (for details see Sec. 5.6).

5.4.1 Control channels

Beyond the normalisation channels, $B^0 \to K^{*0}(J/\psi \to e^+e^-/\mu^+\mu^-)$, extra-control channels are used to perform cross-checks and better constrain some of the background components in the electron fit. In particular, $B^0 \to K^{*0}(\gamma \to e^+e^-)$, also denoted as "$\gamma(ee)$", and $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$, also denoted as "$\psi(2S)(ee)$". All the normalisation and control channels are distinguished depending on the q^2 interval they fall into (for details see Sec. 5.6).

5.5 Data samples and simulation

Simulated samples are used to study the background properties, determine efficiencies and to train the multivariate analysis. The hard interactions are generated with Pythia8 hadronic particles are decayed using EvtGen and, finally, propagated into the detector using Geant4 and reconstructed with the same software used for data. Samples are generated with both 2011 and 2012, magnet up and down conditions and are combined in the right proportions, according to the luminosity registered on data. The next section describes the corrections applied to the simulation to obtain a better description of data.

5.5.1 Data-simulation corrections

Since the multivariate classifier training (see Sec. 5.6.6) and the calculation of most of the efficiency components (see Sec. 5.9) are obtained from the study of simulated events it is important to verify that the simulation provides a reliable reproduction the data. In particular it is important to match data and Monte Carlo in the kinematics of the final particles and the occupancy of the detector. The kinematics of the decays is characterised by the transverse momentum spectrum of the B^0. Discrepancies in this distribution cause also the spectra of the final particles to differ from data and affect the efficiency determination as its value often depends on the momentum of the final particles. The occupancy of the detector is relevant as it is correlated to the invariant mass shape of the signal because of the addition of energy clusters in the electromagnetic calorimeter, which affects the electron's momenta especially when bremsstrahlung photons emitted before the magnet. The hit multiplicity in the SPD detector is a proxy for the detector occupancy.

Since it is important that these quantities are well modelled, the simulation is reweighted so that the distributions in data and simulation match for these variables. The weight is calculated using resonant $B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$ candidates, for which the signal peak is already visible in data after pre-selection (see Sec. 5.6). However, the data still includes a high level of background and distributions cannot be directly compared. The ${}_s\mathcal{P}$lot technique [58] is used to statistically subtract the background from data and obtain pure signal distributions using the invariant mass as control variable. Figure 5.3 shows fits to the 4-body invariant mass of candidates after pre-selection. Data and simulation are then compared and the ratio between the two distributions is used to re-weight the simulation. The discrepancy in the SPD hits multiplicity is solved as a first step and then the B^0 transverse momentum distributions are compared between data and simulation reweighted for the SPD multiplicity only. Distributions of B^0 transverse momentum and SPD multiplicity are reported in Fig. 5.4 and ratios of these distribution, which are used to re-weight the simulation, are reported in Fig. 5.5. The weights for the SPD multiplicity are calculated separately for 2011 and 2012 events, because distributions are significantly different in the two years. The binnings for these distributions are chosen to have approximately the same number of events in each bin to limit fluctuations. Further corrections are made by re-weighting the simulation for PID efficiency using the `PIDCalib` package as described in Sec. 5.9.3 and, finally, ee samples are also reweighted for L0 trigger efficiency as described in Sec. 5.9.4. Weights are always applied throughout unless specified.

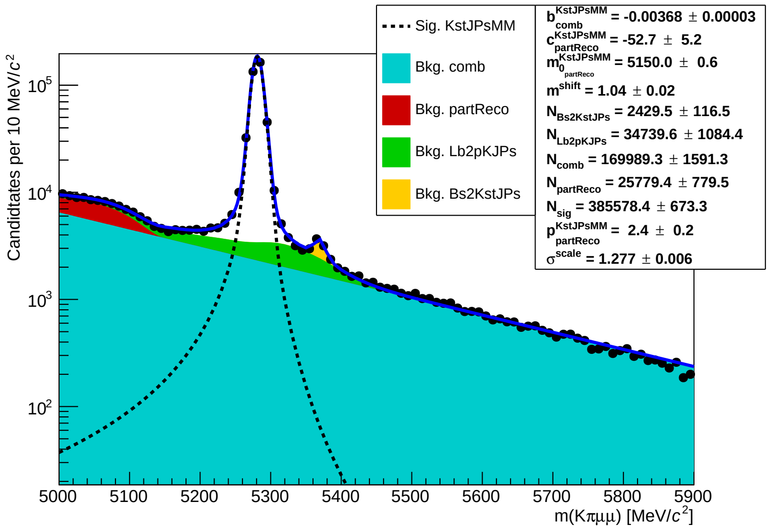

Figure 5.3: Fitted 4-body invariant mass distributions of muonic resonant candidates.

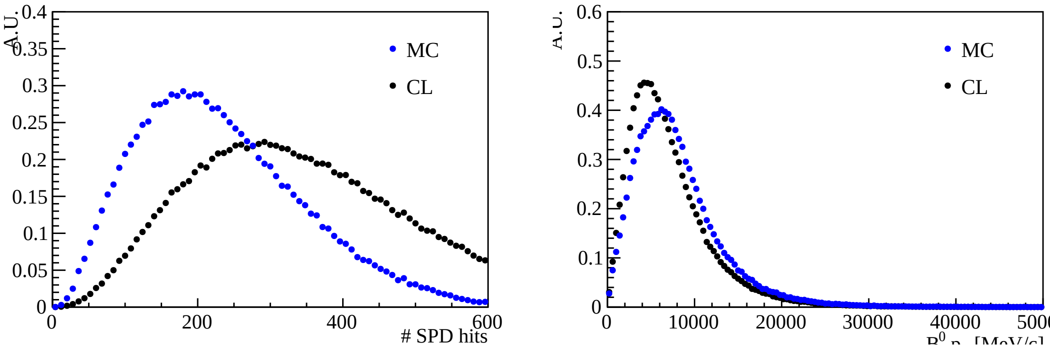

Figure 5.4: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

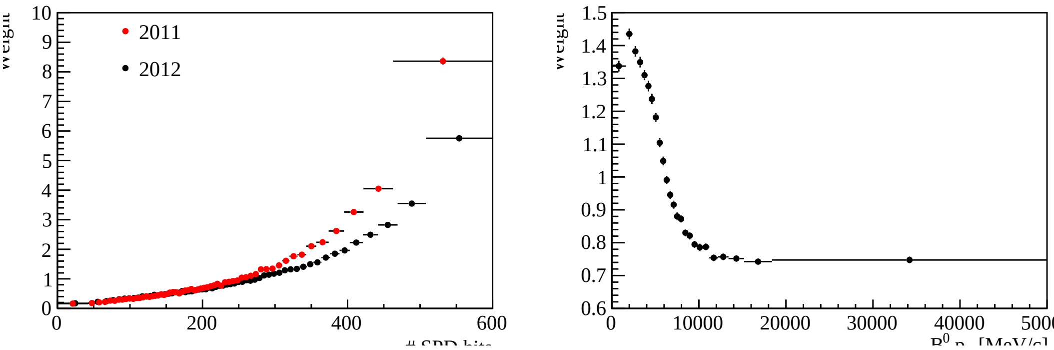

Figure 5.5: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

5.6 Selection

The selection process, described in this section, is divided into several steps:

- first of all candidates have to fall into the detector acceptance, produce hits and be selected on the basis of quality features, such as χ^2 of tracks and vertices and basic kinematic cuts. This stage is called "stripping". Furthermore, it is required that the events are triggered by specific trigger lines and cuts are applied to remove backgrounds from specific decays. All these first three steps are referred to as "pre-selection";
- secondly, particle identification requirements are applied to remove part of misreconstructed background and clear the way for the last step;
- in the final step a neural network is used to remove combinatorial background. Furthermore, for the electron channels, which are more challenging, the kinematic structure of the decays is also used to improve the samples purity.

To identify the $B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ candidates a dilepton mass interval of 100 MeV/c^2 around the nominal J/ψ peak [1] is selected. On the other hand it is not possible to use a narrow interval around $J/\psi(ee)$ mass peak as the invariant mass distribution is characterised by a long radiative tail at low masses due to bremsstrahlung radiation. Furthermore, a requirement in q^2 distorts the 4-body $m(K\pi ee)$ mass distribution which is not advisable as is important to be able to fit a wide mass range to constrain the backgrounds. For these reasons the interval to select $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates is chosen to go as low as possible without overlapping with the rare channel interval. Candidates are therefore identified as $J/\psi(ee)$ if they fall in the q^2 interval $6 < q^2 < 11$ GeV$^2/c^4$. Similarly, candidates are identified as $\psi(2S)(ee)$ is they fall into $11 < q^2 < 15$ GeV$^2/c^4$ and $\gamma(ee)$ if they fall into $q^2 < 0.004$ GeV$^2/c^4$. Table 5.2 summarises the requirements to distinguish sample from different channels. Figure 5.6 shows two-dimensional distributions of q^2 versus the 4-body invariant mass for candidates passing the full selection. Horizontal

Table 5.2: Summary of the channel categories.

Type	**Sample**	$\boldsymbol{q^2}$
	$B^0 \to K^{*0}\mu^+\mu^-$ (low)	$0.0004 < q^2 < 1.1\,\mathrm{GeV}^2/c^4$
$\mu\mu$	$B^0 \to K^{*0}\mu^+\mu^-$ (central)	$1.1 < q^2 < 6\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}\mu^+\mu^-$ (high)	$q^2 > 15\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ $(m(K\pi\mu\mu))$	$\lvert m_{\,\mathrm{mm}} - m_{J/\psi}^{PDG}\rvert < 100$ MeV$/c^2$
	$B^0 \to K^{*0}e^+e^-$ (low)	$0.0004 < q^2 < 1.1\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}e^+e^-$ (central)	$1.1 < q^2 < 6\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}e^+e^-$ (high)	$q^2 > 15\,\mathrm{GeV}^2/c^4$
ee	$B^0 \to K^{*0}(J/\psi \to e^+e^-)$ $(m(K\pi ee))$	$6 < q^2 < 11\,\mathrm{GeV}^2/c^4$
	Control samples	
	$B^0 \to K^{*0}(\gamma \to e^+e^-)$ $(m(K\pi ee))$	$q^2 < 0.0004\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}(J/\psi \to e^+e^-)$ $(m(K\pi ee))$	$6 < q^2 < 11\,\mathrm{GeV}^2/c^4$
	$B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ $(m(K\pi ee))$	$11 < q^2 < 15\,\mathrm{GeV}^2/c^4$

bands can be clearly seen at q^2 values corresponding to the J/ψ and $\psi(2S)$ resonances. On the plot for muons it is also evident a vertical band which corresponds to rare decay of interest.

5.6.1 Trigger and Stripping

Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported in Tab. 5.3, where the logical *and* of L0, HLT1 and HLT2 lines is required and the logical *or* of the lines on the same level. The candidates are required to be triggered-on-signal (TOS) for most of the stages, namely it is required for the particle which triggered to be one of the particles used to build the signal candidates. Only for `L0Global`, used in the electron case, we require a trigger-independent-of-signal (TIS), this is aimed to collect all the possible statistics for the electron channels, which are the most challenging. The `L0Muon` trigger requires hits in the muon detector, while `L0Electron` and `L0Hadron` use information from the calorimeters; `HLT1TrackAllL0` adds information from the trackers and triggers if the L0 decision is confirmed; finally, `HLT2Topo[2,3]BodyBBDT` uses a full reconstruction of the event and a neural network trained on events with a specific topology in order to detect specific decay

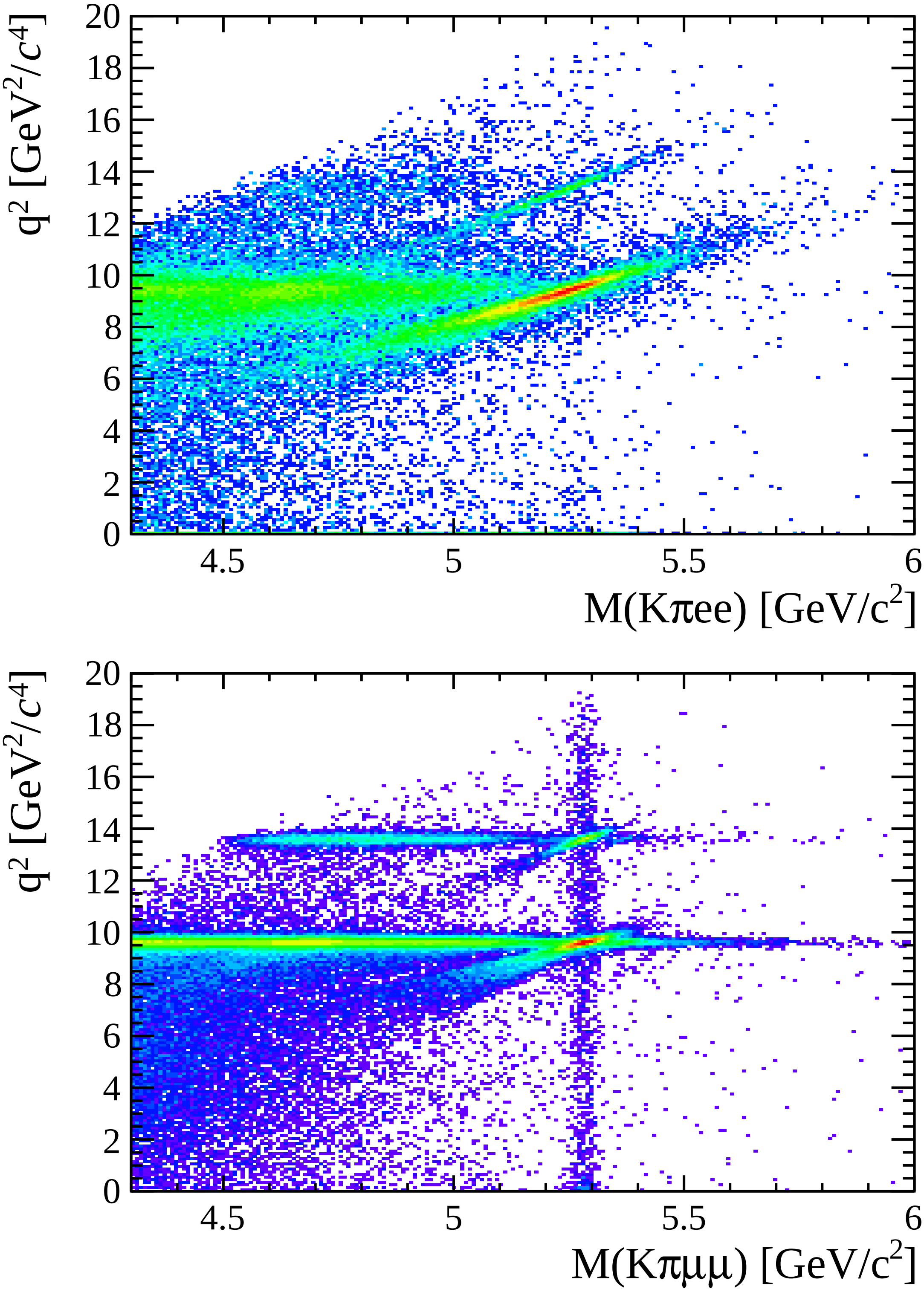

Figure 5.6: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electron (top) and muonic (bottom) channels in 2012 data.

structures.

Table 5.3: Summary of the trigger lines used to select the $\mu\mu$ and the ee channels. Where not explicitly indicated, the lines are required to be TOS.

$\mu\mu$ **candidates**	ee **candidates**
L0Muon	L0Electron L0Hadron L0Global (TIS)
Hlt1TrackAllL0 Hlt1TrackMuon	Hlt1TrackAllL0
Hlt2Topo[2,4]BodyBBDT Hlt2TopoMu[2,4]BodyBBDT Hlt2DiMuonDetachedDecision	Hlt2Topo[2,4]BodyBBDT Hlt2TopoE[2,4]BodyBBDT

For the electron channels the L0 lines have different properties, therefore the analysis is performed separately for three categories of events, depending on the L0 trigger that fired them. These categories are defined to be exclusive in the following way:

- **L0E**: events triggered by at least one of the electrons in the signal candidate (`L0Electron_TOS`);
- **L0H**: events triggered by at least one of the hadrons in the signal candidate and not by L0Electron (`L0Hadron_TOS && !L0Electron_TOS`);
- **L0I**: events triggered by particles not in the signal candidate and not by the previous cases (`L0_TIS && !(L0Electron_TOS || L0Hadron_TOS)`).

The majority of the selected events falls in the L0E category, while the L0H category is more efficient at low q^2 were the K^{*0} has more momentum. Because L0I is defined to be independent of the signal candidate, the corresponding signal efficiency is the same in the rare and resonant cases and cancels out in their ratio.

Candidates are then required to pass the kinematic and quality cuts summarised in Tab. 5.4. The meaning of the variables in the table was already explained in Sec. 3.4. Loose PID cuts are applied in preselection to limit the size of the samples, while tighter cuts are applied in a second stage. A large mass window is kept

Table 5.4: Summary of stripping requirements.

Particle	**Requirements**
π	$\chi^2_{\mathrm{IP}}(primary) > 9$
K	$\texttt{PID}_K > -5$
	$\chi^2_{\mathrm{IP}}(primary) > 9$
	`hasRICH`
K^{*0}	$p_{\mathrm{T}} > 500$ MeV/c
	$\lvert m - m^{PDG}_{K^{*0}} \rvert < 300$ MeV/c^2
	$\chi^2_{\mathrm{IP}}(primary) > 9$
	Origin vertex χ_2/ndf < 25
μ	$p_{\mathrm{T}} > 300$ MeV/c
	$\chi^2_{\mathrm{IP}}(primary) > 9$
	`isMuon`
e	$p_{\mathrm{T}} > 300$ MeV/c
	$\chi^2_{\mathrm{IP}}(primary) > 9$
	`hasCalo`
	$PID_e > 0$
$\ell\ell$	$m < 5500$ MeV/c^2
	End vertex χ^2/ndf < 9
	Origin vertex χ^2 separation > 16
B^0	`DIRA` > 0.9995
	End vertex χ^2/ndf < 9
	$\chi^2_{\mathrm{IP}}(primary) < 25$
	Primary vertex χ^2 separation > 100

around the B^0 peak in order to be able to use the sideband to train the multivariate analysis and to constrain the backgrounds. Track and vertex quality cuts are also applied using the χ^2_{track}/ndf, `GhostProb`, and χ^2_{vtx}/ndf variables. The `GhostProb` quantity describes the probability of a track being fake. By construction cutting at 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. For details about the definition of the variables used see Ref. [107].

5.6.2 PID

After preselection there still are high levels of misreconstructed background. In particular, as the ID of kaons and pions are not constrained, the samples still contain both ID combinations for most candidates, therefore tighter PID cuts are applied.

In the LHCb analysis framework the particle identification probability can be quantified using the "`ProbNN`" variables [108]. These variables are the output of a neural network which takes as input information from the calorimeters, the RICH detectors the muon system and the tracking system. Unlike the DLL variables (see Sec. 2.8) the `ProbNN` are bound from 0 to 1 and can be therefore directly be interpreted as probabilities. For example `ProbNNk` is the probability for a reconstructed particle to be a kaon. Figure 5.7 shows distributions of the correct ID variables in the

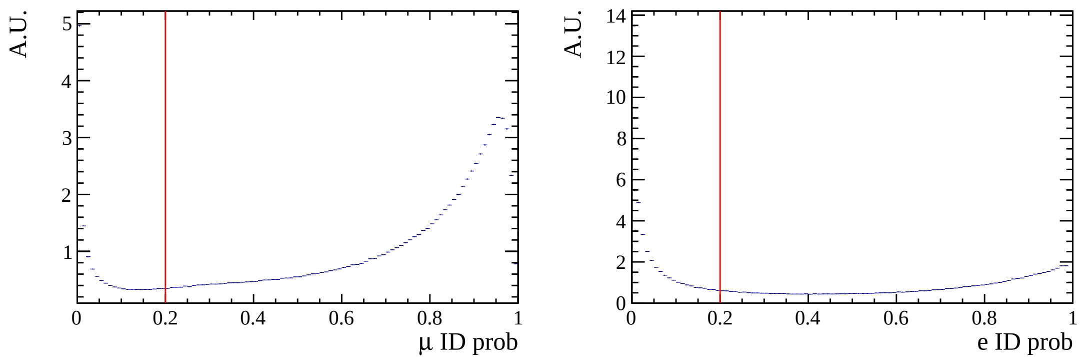

Figure 5.7: Correct ID probability distributions for muons (left) and electron (right) in 2012 data.

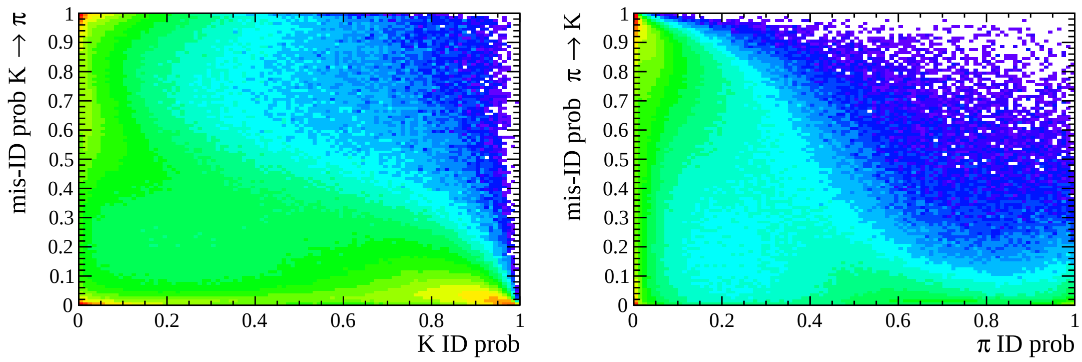

Figure 5.8: On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis show the mis-ID probability.

2012 data sample while Fig. 5.8 shows in a two-dimensional plane the probability of correct identification and mis-identification of kaons and pions. These plots are characterised by clear peak at maximal ID probability and minimal mis-ID probability, corresponding to particles to which a well defined identification can be assigned. In order to maximise the power of the PID requirements probabilities of correct ID and mis-ID are combined using the following cuts:

$$
\begin{aligned}
\pi &\rightarrow \texttt{ProbNNpi} \times (1 - \texttt{ProbNNk}) \times (1 - \texttt{ProbNNp}) > 0.1 \\
K &\rightarrow \texttt{ProbNNk} \times (1 - \texttt{ProbNNp}) > 0.05 \\
\mu &\rightarrow \min(\ \texttt{ProbNNmu},\ \texttt{ProbNNmu}\ ) > 0.2 \\
e &\rightarrow \min(\ \texttt{ProbNNe},\ \texttt{ProbNNe}\ ) > 0.2
\end{aligned}
$$

In the first formula, for example, `ProbNNpi` is the probability of correctly identifying the pion as a pion, while `ProbNNk` is the probability of mistaking it for a kaon. Therefore by maximising the quantity "`ProbNNpi` × (1 - `ProbNNk`)", one can maximise the correct ID probability and minimise at the same time the mis-ID probability.

5.6.3 Peaking backgrounds

Backgrounds due to specific decays usually peak in some variable because of their distinctive kinematic properties and therefore they can be removed without significant signal efficiency loss. The following sections describe the main sources of peaking background. The same cuts are applied to the muon and electron channels, unless specified.

5.6.3.1 Charmonium vetoes

Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2. The choice of q^2 binning described in Sec. 5.4 constitutes a natural veto for these decays. Simulated events were used to check if resonant candidates leak inside the q^2 intervals chosen for the rare channel analysis. For the muonic channels the leakage is negligible as the peaks are sharper due to a better resolution and muons emit fewer bremsstrahlung photons, resulting in shorter radiative tails. The electronic channels are instead characterised by a worse resolution and at the same time electrons can radiate several bremsstrahlung photons, yielding long tails at low q^2. Analysing Monte Carlo events it was found that 1.3–2% (depending on the trigger category) of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candidates leak into the $1.1 < q^2 < 6\ \mathrm{GeV}^2/c^4$ interval and 1.8% of $\psi(2S)$ events leak above 15 GeV$^2/c^4$. The contribution from these candidates is modelled in the fit.

5.6.3.2 ϕ veto

It can happen that a kaon from the decay $B_s \to \phi\ell^+\ell^-$, where the ϕ decays in two kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as a K^{*0}. This results in a candidate with a value of $m(K\pi)$ that is less than the nominal K^{*0} mass but still high enough to pass the selection requirements. Figure 5.9 shows the plot of $m(K\pi)$ versus $m(K\pi\ell\ell)$, where the kaon mass hypothesis is assigned to the pion. A peak can clearly be seen around the ϕ mass (1020 MeV$/c^2$). To remove this background only candidates with $m_{K(\pi\to K)} > 1040$ MeV$/c^2$) are selected. This results in a 98% background rejection while keeping a 99% signal efficiency. The ϕ could also constitute a background when it decays into two leptons but the branching ratio of this decay is small compared to the one into kaons and this contribution is taken into account by the choice of the q^2 intervals (see Sec. 5.4).

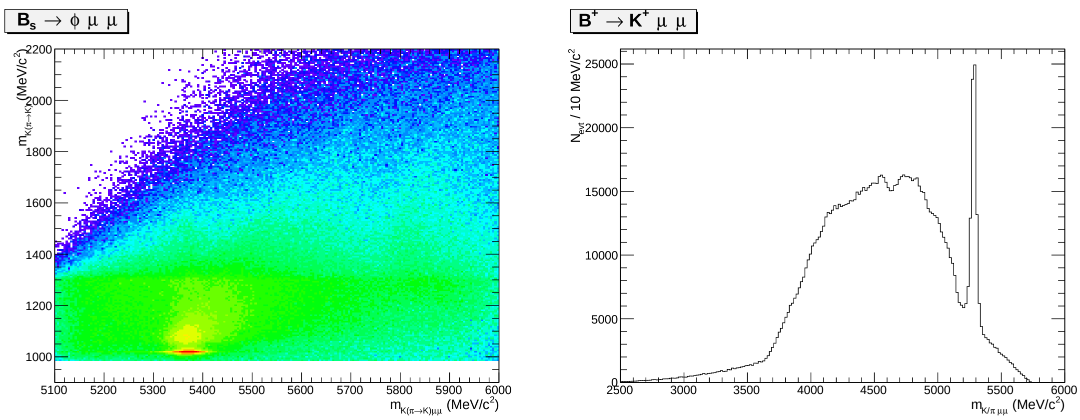

Figure 5.9: (left) Distribution of 2011 data events as a function of the variables $(m_{K(\pi\to K)})$ and $(m_{K(\pi\to K)\mu\mu})$, where $\pi \to K$ means that the kaon mass is given to the pions too. (right) The invariant mass distribution of the three-body system $(K\mu\mu)$, where the peak due to the $B^+ \to K^+\mu^+\mu^-$ decay is visible.

5.6.3.3 $B^+ \to K^+\ell^+\ell^-$ plus a random pion

$B^+ \to K^+\ell^+\ell^-$ decays can contaminate the upper B^0 mass sideband if they are combined with a soft pion from somewhere else in the event and therefore reconstructed as a B^0 decay. Similarly a kaon can be mis-identified as a pion and combined with an other kaon in the event. Figure 5.9 shows the invariant mass distribution of the three-body $K\mu^+\mu^-$ system, $m(K\mu\mu)$. This is characterised by a narrow peak at the B^+ mass. Since these candidates have $m(K\pi\ell\ell) > 5380$ MeV$/c^2$ there is no contribution under the B^0 peak, but they can cause problems when using sidebands events to train the neural network. An effective veto for this decay was found to be $\max(m_{K\ell\ell}, m_{(K\to\pi)\ell\ell}) < 5100$ MeV$/c^2$, which results in 95% background rejection while keeping 99% signal efficiency.

5.6.3.4 Λ_b decays

$\mathit{\Lambda}_b^0 \to J/\psi\,\mathit{\Lambda}$ decays are unlikely to be reconstructed as $B^0 \to K^{*0}\ell^+\ell^-$ because the $\mathit{\Lambda}$ is long-lived and decays further in the detector with a separate vertex. However, simulated events were used to check how many candidates fall into the B^0 samples, which results to be negligible. The $\mathit{\Lambda}_b^0 \to J/\psi pK$ decay, when the proton is mis-identified, can instead contribute more easily since the $m(pK)$ is above the $\mathit{\Lambda}$ threshold and therefore they must come from $\mathit{\Lambda}^*$ resonances, which are not long-lived. This background is already reduced by the PID requirements but a non-negligible contribution is still expected, which is modelled in the fit.

5.6.3.5 $B^0 \to (D^- \to Ke^-\overline{\nu})e^+\nu$

The $B^0 \to D^-e^+\nu$ decay, where the D^- in turn decays semileptonically to $K^{*0}e^-\nu$ has the same final particles as the $B^0 \to K^{*0}e^+e^-$ decay plus two neutrinos which are not reconstructed. This decay has a branching ratio four orders of magnitude larger than $B^0 \to K^{*0}e^+e^-$ in the low-q^2 region and it may pass the selection

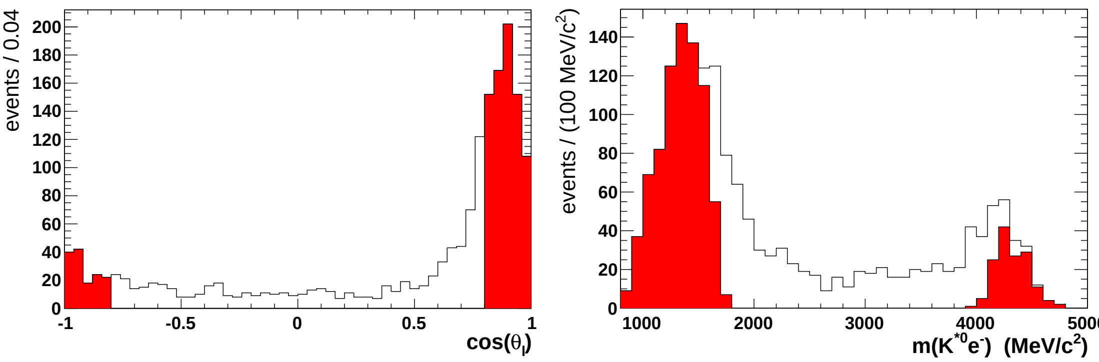

Figure 5.10: Distribution of (left) $\cos(\theta_\ell)$ and of (right) the $m(K^{*0}e)$ invariant mass, where the $B^0 \to (D^- \to Ke^-\overline{\nu})e^+\nu$ background is selected by requiring $m(K^{*0}e^+e^-) < 4800\ \mathrm{MeV}/c^2$. The red distribution corresponds to events with $|\cos(\theta_\ell)| > 0.8$.

requirements when the two neutrinos carry a low momentum. To lower the level of this background the angle θ_ℓ is used, which is defined as the angle between the direction of the e^+ (e^-) in the di-electron rest frame and the direction of the di-electron in the B^0 ($\overline{B}^0$) rest frame. Low momentum neutrinos demand the D^- and the e^+ to be almost back-to-back in the B^0 rest frame giving the e^+ a relatively large energy compared to the e^-. As a consequence, the direction of the e^+ is close to the direction of the di-electron pair, thus the θ_ℓ angle is close to 0. This explains why the distribution of background selected in data with an invariant mass cut of $m(K^{*0}ee) < 4800\ \mathrm{MeV}/c^2$is asymmetric towards higher $\cos(\theta_\ell)$ values as it can be seen in Fig. 5.10(*left*). The cut is chosen to be $|\cos(\theta_\ell)| < 0.8$, and is not applied in the high-q^2 bin as the variable looses its discriminating power.

In the muon channels the background from $B^0 \to (D^- \to K\mu^-\overline{\nu})\mu^+\nu$ decays is suppressed by the choice of the fitting range.

5.6.3.6 $B^0 \to K^{*0}(\gamma \to e^+e^-)$

For the low-q^2 region, a potentially dangerous peaking background is due to the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ decay followed by a conversion of the photon in the detector. The branching fraction of $B^0 \to K^{*0}(\gamma \to e^+e^-)$ has been measured to be $\mathcal{B}$= $(4.33\pm$

$0.15) \times 10^{-5}$ and when the photon converts to an electron and a positron has similar characteristics to $B^0 \to K^{*0} e^+ e^-$. In LHCb around 40% of the photons convert before the calorimeter. Although only a small fraction of these, $\sim 10\%$, converts in the VELO and are reconstructed as long tracks, the resulting B^0 mass should peak under that of the signal, making it a dangerous background. To veto this signal-like background an effective veto is in the reconstructed invariant mass window for the e^+e^-pair that was chosen above 20 MeV$/c^2$. Furthermore, the e^+e^-pair from $B^0 \to K^{*0}(\gamma \to e^+e^-)$ has a vertex at the point of conversion of the photon, but it may still be reconstructed as originating from the B^0 decay when the e^+e^-vertex is determined with a large error. Therefore a requirement is applied on the uncertainty of the reconstructed z-coordinate of the e^+e^-pair: $\sigma_z(e^+e^-) < 30$mm. iSimulation is used to predict the contamination from $B^0 \to K^{*0}(\gamma \to e^+e^-)$ events in the signal region which is found to be $(3.2 \pm 1.6)\%$.

5.6.3.7 Other peaking backgrounds

Contamination from $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$ where η and the pion decay into two photons was considered and found to be small. Furthermore, a potentially dangerous background could come from events where the identity of the kaon and the pion are swapped as these candidates peak under the signal. Their contribution is found to be small, 0.5%, however the effect of their modelling into the fit is taken into account in the systematic uncertainties. Finally, charmonium decays where the identity of the kaon, or the pion, and one of the muons are swapped is a potentially dangerous background. These decays are rejected by requiring that the hadron-μ invariant mass $m((h \to \mu)\mu)$, where the muon mass hypothesis is assigned to the hadron, is not compatible with a J/ψ ($\psi(2S)$) resonance: $|m((h \to \mu)\mu)?m_{J/\psi,(\psi(2S))}| > 60$ MeV$/c^2$.

5.6.4 Mis-reconstructed background

Mis-reconstructed candidates are defined as decays where one or more particles in the final state are not reconstructed, resulting in $m(K\pi\ell\ell)$ values smaller than m_{B^0}, but with tails that can still contaminate the signal peak. Sources of mis-reconstructed background are decays involving higher hadronic states such as $B^0 \to (Y \to K\pi X)(J/\psi \to e^+e^-)$, where X represents one or more not reconstructed particles. The Y state can be a K^* resonance as well as D mesons that decay semileptonically (*e.g.* $B^0 \to D^-\ell^+\bar{\nu}_\ell$ followed by $D^- \to K^{*0}\ell^-\nu_\ell$). In case of the $J/\psi\,(ee)$ channel, an additional source of mis-reconstructed background are decays of higher $c\bar{c}$ resonances, $B^0 \to (K^{*0} \to K\pi)(Y \to (J/\psi \to e^+e^-)X)$. To reject this backgrounds in the mm channels the 4-body invariant mass $m(K\pi\mu\mu)$ is recalculated using `DecayTreeFitter` with a vertex constraint. For the resonant case this also includes a J/ψ mass constraint to the dilepton pair. By using this procedure mis-reconstructed events are pushed towards low masses, resulting in no contamination above 5150 MeV/c^2. To correctly model the long radiative tail of the ee and $J/\psi\,(ee)$ mass shapes, a fit region that extends down to 4500 MeV/$c^2$is used. As a consequence, no mass constraint to the dilepton pair is applied, as this could bias the 4-body mass distribution, and the mis-reconstructed background is modelled in the fit (for details see Sec. 5.8.2.2).

5.6.5 Bremsstrahlung corrected mass

An additional handle against backgrounds that contaminate the ee channels is provided the analysis of the kinematics of the decay. In fact for the B^0 daughters the momentum component orthogonal to the flight direction of the B^0 meson should cancel out. The flight direction is defined using the primary and the decay vertices and sketch is shown in Fig. 5.11.

The ratio between the p_{T} of the K^{*0} and the di-electron pair can be used to check

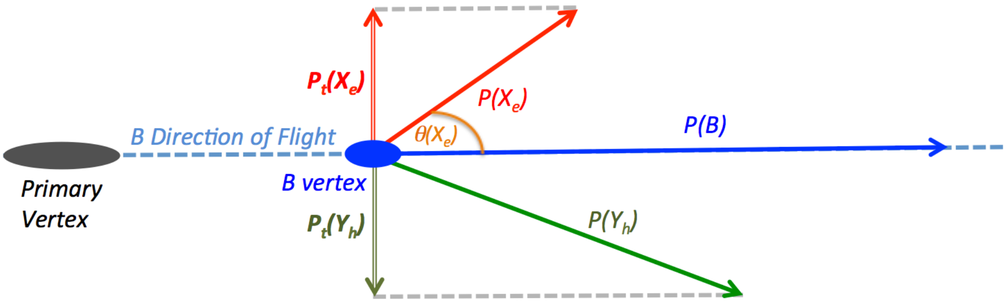

Figure 5.11: Schema of the kinematic of a $B \to Y_h X_e$ decay, highlighting the quantities relevant for the definition of the bremsstrahlung correction factor, α.

this hypothesis

$$\alpha = \frac{p_{\mathrm{T}}(K^*)}{p_{\mathrm{T}}(e^+e^-)}\,.$$

When α 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. Therefore we can use α to correct the electron momentum as

$$p_{\mathrm{corr}}(e^+e^-) = \alpha \times p(e^+e^-)\,.$$

Since bremsstrahlung photons are emitted in the same direction of the electron, the same α correction can be applied to the longitudinal component of the di-electron momentum. In contrast, the missing particles in partially-reconstructed background candidates are not necessarily emitted in the direction of the electrons, and therefore the α correction does not work properly. A similar argument applies to the combinatorial background.

The corrected momenta can be used to re-calculate the invariant mass of the B^0 candidate, which in the following will be called Bremsstrahlung Corrected Mass (m_{BCM}). The resolution of m_{BCM} depends on the quality of the vertex reconstruction and on the B^0 lifetime, and degrades as a function of q^2. Figure 5.12 shows the dependence of the B^0 χ^2_{FD} (flight distance χ^2) as a function of m_{BCM} in the considered q^2 regions.

As the correction does not work properly for backgrounds this leads the candidates to

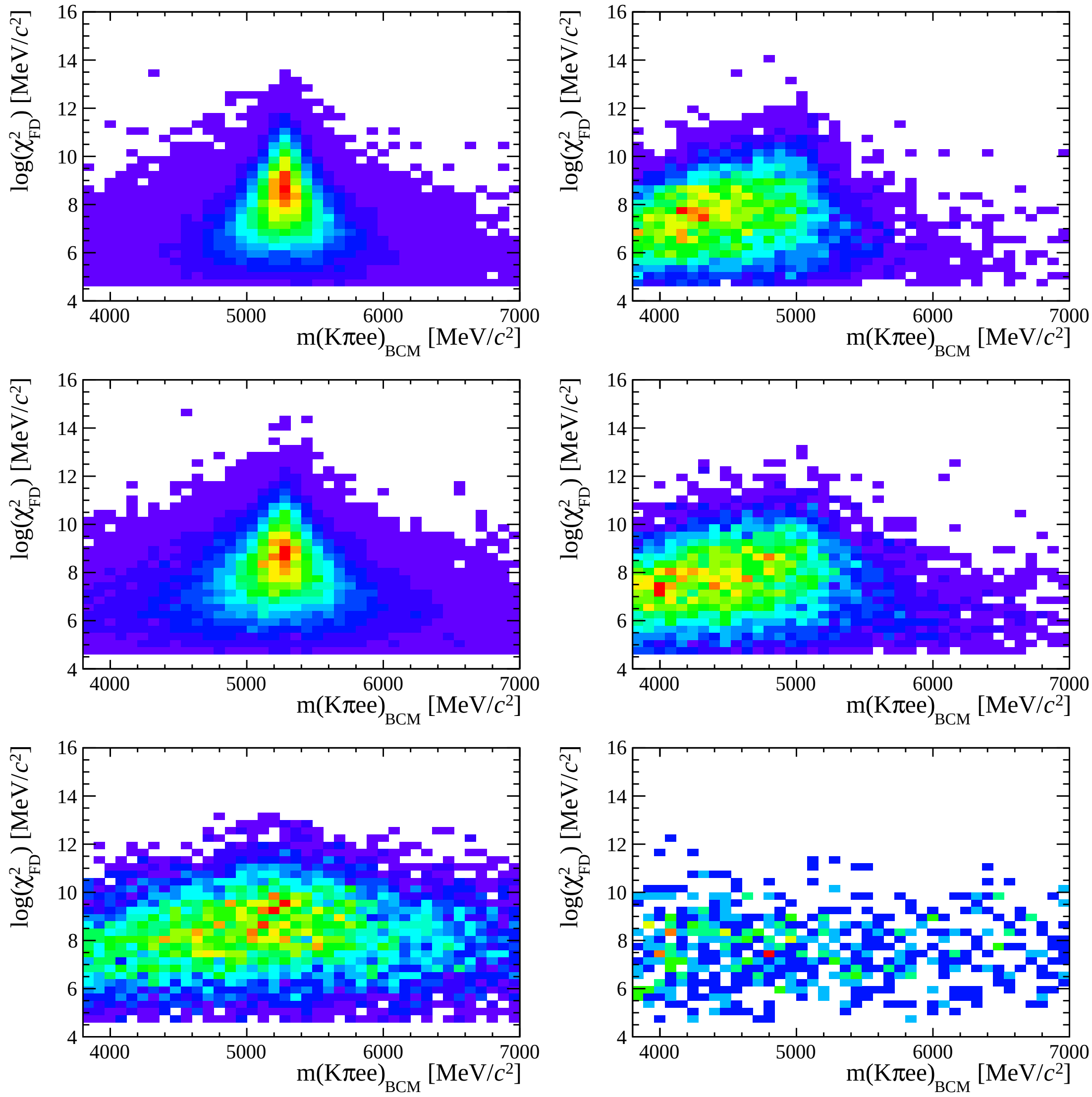

Figure 5.12: Two-dimensional distribution of $\chi^2_{\rm FD}$ *vs.* $m_{\rm BCM}$ for (left) $B^0 \to K^{*0}e^+e^-$ signal and (right) partially-reconstructed background. From top to bottom the low-, central- and high-q^2 intervals.

spread out making $m_{\rm BCM}$ a discriminating variable between signal and background shapes. A two-dimensional cut is adopted

$$m_{\rm BCM} > a_{\rm BCM} + b_{\rm BCM} \cdot \log(\chi^2_{\rm FD})$$

where the $a_{\rm BCM}$ and $b_{\rm BCM}$ coefficients are optimised as described in Sec. 5.6.7.

No cut is applied at high-q^2 nor on the muon channels for which the bremsstrahlung radiation is negligible.

5.6.6 Multivariate analysis

The final selection is performed using a Neural Network classifier (NN) based on the NeuroBayes package [68, 69]. The multivariate analysis is intended to remove some combinatorial background and obtain a clearer signal peak. In order to avoid biases, a k-fold approach is adopted to train and optimise the classifier, using $k = 10$. This method consists in dividing the samples in k equally sized subsamples; k classifiers are then trained and optimised each on $(k-1)$ samples and applied to the kth one. This approach ensures that a classifier is never applied to the events used for its training. Each classifier is trained on half of the events included in the $(k-1)$ samples and optimised using the other half, which ensures that events used for training are not used for optimisation.

Samples:

Representative samples of the signal and background are needed to train the classifier. For the signal, fully reconstructed $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ simulated events can be used. Instead a sample representative of the background can be obtained using real data candidates in the upper B^0 sideband: $m(K\pi\mu\mu) > 5400$ MeV/c^2 and $m(K\pi ee) > 5600$ MeV/c^2. The lower sideband is not used in the training as it contains a significant fraction of mis-reconstructed background. All pre-selection cuts are applied to the background samples used for the training. As L0 and PID variables are not well described in simulation these cuts are not applied to the simulation but their effect is taken into account by the event weight. An approximately equal number of signal and background events is used for the training which corresponds to about 1000 events for the electron case and 10,000 for the muon one.

Training:

The the neural-network input consists of 24 variables containing information about the kinematic of the decays and the quality of tracks and vertices. All the variables used are listed in Tab. 5.5. In these figures the variable with ID = 1 is the

Table 5.5: List of variables used as inputs for the neural-network training.

Particle	**Variables**
B^0	p_{T}, χ^2_{IP}, χ^2_{FD}, $\chi^2_{vtx}/\mathrm{ndf}$, `DIRA`, $\chi^2_{\mathrm{DTF}}/\mathrm{ndf}$
K^{*0}	p_{T}, χ^2_{IP}, χ^2_{FD}, $\chi^2_{vtx}/\mathrm{ndf}$, `DIRA`
h	$min, max(p_{\mathrm{T},K}, p_{\mathrm{T},\pi})$, $min, max(\chi^2_{\mathrm{IP},K}, \chi^2_{\mathrm{IP},\pi})$
$\ell\ell$	p_{T}, χ^2_{IP}, χ^2_{FD}, $\chi^2_{vtx}/\mathrm{ndf}$, `DIRA`
ℓ	$min, max(p_{\mathrm{T},\ell^+}, p_{\mathrm{T},\ell^-})$, $min, max(\chi^2_{\mathrm{IP},\ell^+}, \chi^2_{\mathrm{IP},\ell^-})$

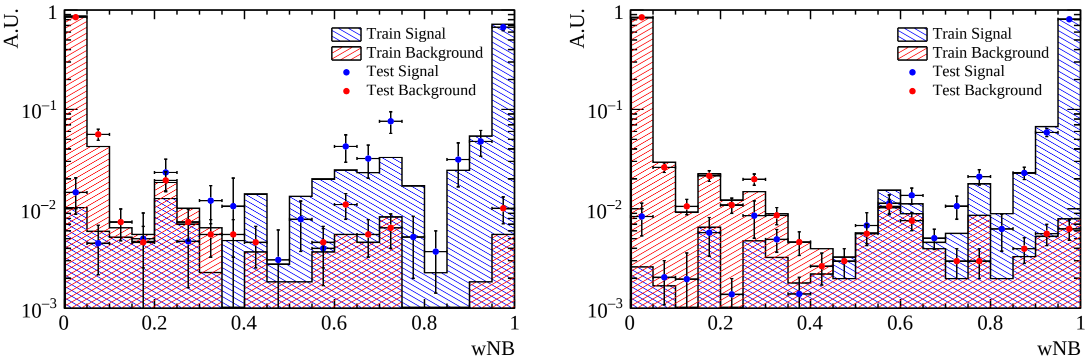

Figure 5.13: NN output distributions for training (solid) and test (stripes) samples, for simulated signal and data sideband events. For the electron (left) and muon (right) training.

neural-network output and the other IDs are reported in Tab. 5.5. The single most discriminating variable used is the χ^2 of a kinematic fit that constrains the decay product of the B^0, the K^{*0} and the dimuon, to originate from their respective vertices. Other variables that contribute significantly are the χ^2_{IP} of J/ψ and K^{*0}, the transverse momentum of the B^0 and the pointing direction (`DIRA`) of the reconstructed B^0 to the primary vertex.

Figure 5.13 shows neural network output distributions for signal and background. On this plot the distributions from the test samples are also overlaid in order to check for overtraining. The distributions follow the same shape but with different fluctuations indicating no significant overtraining. In general it can be concluded that the neural network is able to separate signal from background and that the training converged properly.

It can happen that too much information is given to the classifier, which becomes able to calculate the invariant mass of the candidates from its inputs. This could

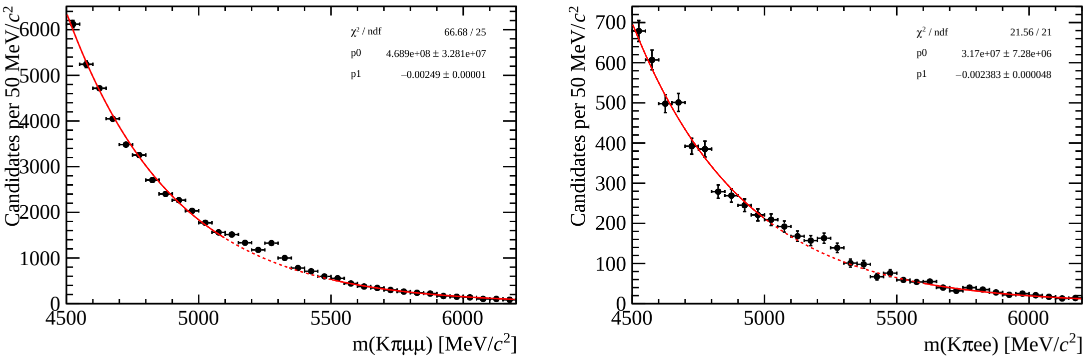

Figure 5.14: Fit to the data sidebands performed to estimate the amount of residual background in the signal mass window for (left) muons and (right) electrons. The region corresponding to the dashed line is excluded from the fit.

generate fake peaks and it is therefore important to check for correlations between the B^0 mass and the neural-network output. Figure 5.15 reports plots of the average neural-network output as a function of the B^0 mass on sideband data and simulated signal events. The distributions are flat showing that no significant correlation is present.

5.6.7 Optimisation

In order to optimise the requirements on the m_{BCM} and the neural network output the expected signal significance, $N_{\mathrm{S}}/\sqrt{N_{\mathrm{S}}+N_{\mathrm{B}}}$, is maximised, where N_{S} (N_{B}) is number of rare signal (background) candidates. When the BCM requirement is applied, the optimisation is performed in a three-dimensional space (t_{MVA}, a_{BCM}, b_{BCM}) where t_{MVA} is the neural-network output threshold below which a candidate is considered background, and a_{BCM} and b_{BCM} are the parameters of the BCM cut described in Sec. 5.6.5. Otherwise, only the MVA cut is optimised (for all muons samples and the high-q^2 electron sample).

The number of signal events accepted by a given requirement is determined using a data-driven method. Firstly, $B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$ candidates selected with all the requirements except for the MVA, and BCM when applicable, cut are fitted to determine the total yield. This number is then scaled by the ratio of the signal to

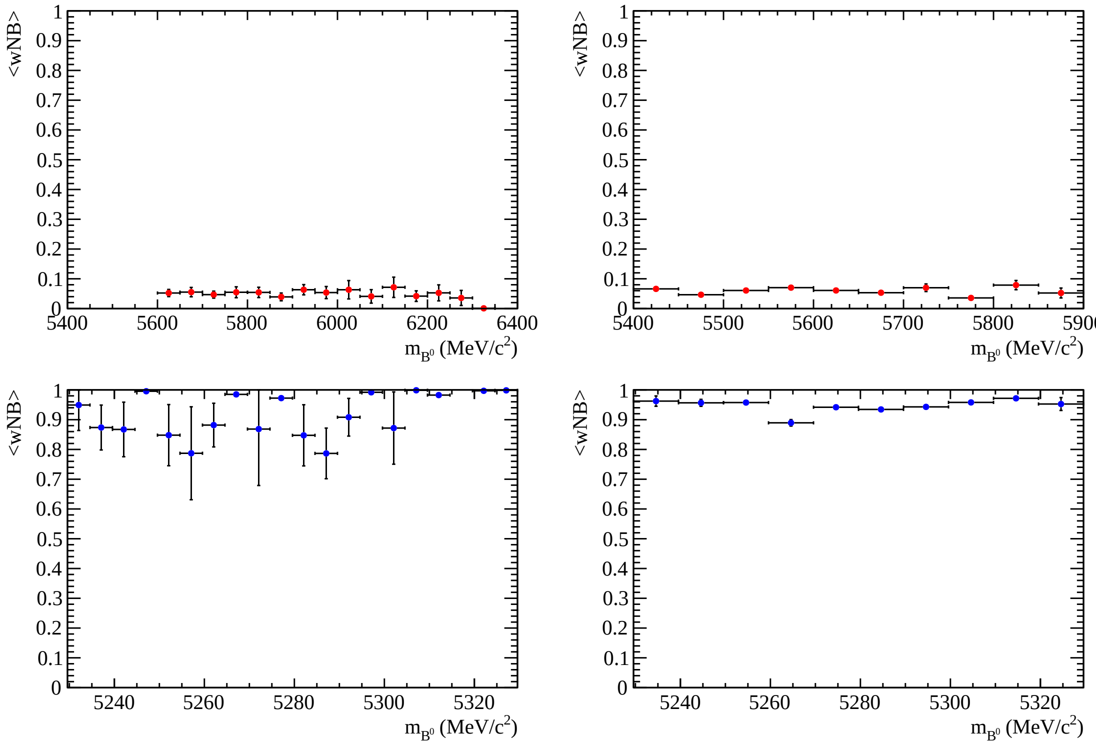

Figure 5.15: Average value of neural-network output as a function of B^0 mass for data sideband (top) and simulated signal (bottom) events for the electron (left) and muon (right) training.

$B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$ branching fractions and the efficiency ratio as a function of the cut

$$N_{\mathrm{S}} = N_{J/\psi\,(\ell\ell)} \cdot \frac{\mathcal{B}(\mathrm{S})}{\mathcal{B}(B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-))} \cdot \frac{\varepsilon_{\mathrm{S}}}{\varepsilon_{J/\psi\,(\ell\ell)}}\,.$$

The number of background events is also derived from data by fitting the background in the lower- and upper-mass sidebands with an exponential function, and extrapolating the residual yield in the signal region (Fig. 5.14). Because the background shape changes as a function of the requirement that is being optimised, the sidebands are refitted for each considered cut value.

The cut optimisation is performed in a signal mass window of ± 100 MeV$/c^2$ around the nominal B^0 mass for muons, and between 5000 and 5400 MeV$/c^2$ for electrons. The average result of the k-fold optimisations is taken as the nominal requirement. The variation of the signal and background efficiency, signal purity and figure-of-merit as a function of the neural-network output requirement for the central-q^2 is

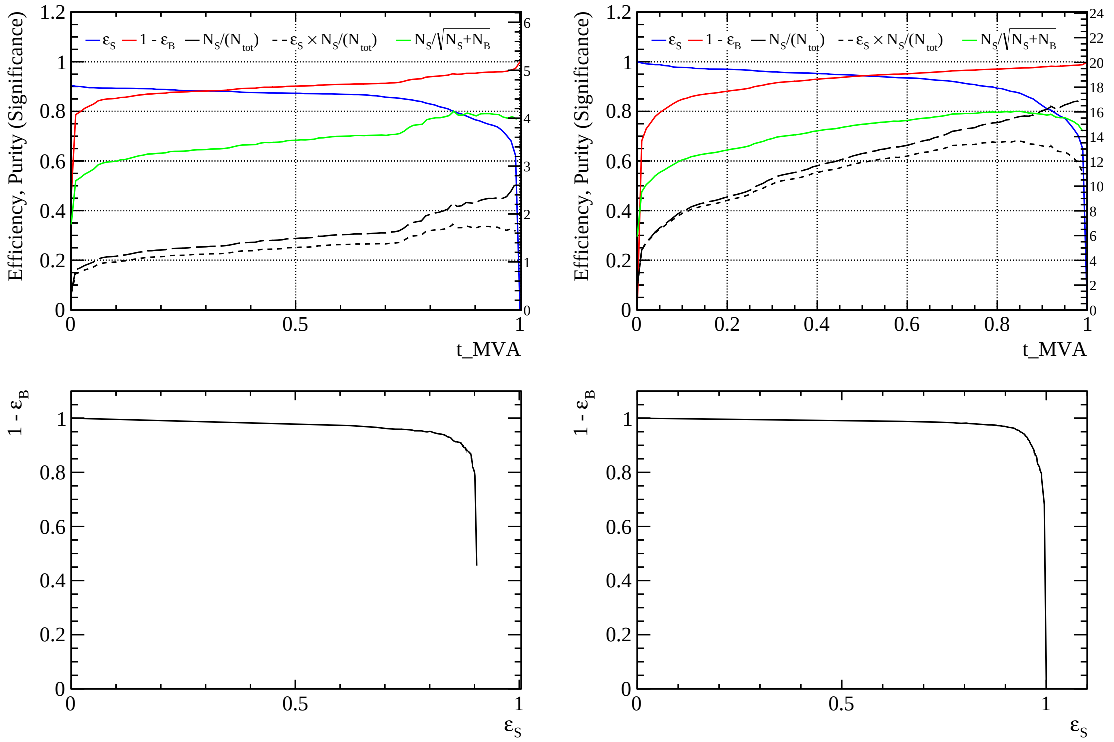

Figure 5.16: (top) Dependence of figure-of-merit on the requirement on neural network output. (bottom) Signal efficiency versus background rejection. Plots correspond to the electron (left) and muons (right) samples.

shown in Fig. 5.16 together with curves of the background rejection as a function of the signal efficiency. After full selection about $\sim 3\%$ of events still contain multiple candidates which are removed at random keeping only a single candidate per event.

5.7 Selection summary

Table 5.6 summarises the requirements applied for each cut after stripping.

5.8 Mass fits

The signal yields are extracted using a simultaneous unbinned maximum likelihood fit to the 4-body invariant mass, $m(K\pi\ell\ell)$, of the rare and normalisation

Table 5.6: Summary of the selection requirements. The last column indicates to which q^2 intervals the requirement is applied.

Type		**Requirement**	$\boldsymbol{q^2}$
Quality	All tracks	$\chi^2/\mathrm{ndf} < 3$	all
		$\texttt{GhostProb} < 0.4$	all
ID	K^{*0}	$\lvert m(K\pi) - m^{PDG}_{K^{*0}}\rvert < 100$ MeV$/c^2$	all
PID	K	$\texttt{ProbNNk} \cdot (1 - \texttt{ProbNNp}) > 0.05$	all
	π	$\texttt{ProbNNpi} \cdot (1 - \texttt{ProbNNk}) \cdot (1 - \texttt{ProbNNp}) > 0.1$	all
	μ	$\texttt{ProbNNmu} > 0.2$	all mm
	e	$\texttt{ProbNNe} > 0.2$	all e^+e^-
BKG	Swap	$\lvert m((h\to\mu)\mu) - m^{PDG}_{J/\psi,(\psi(2S))}\rvert > 60$ MeV$/c^2$	all
	$B^+ \to K^+\ell^+\ell^-$	$max(m(K\ell\ell), m((\pi\to K)\ell\ell)) < 5.1$ GeV$/c^2$	all
	$B^0_s \to \phi\ell^+\ell^-$	$m(K(\pi\to K)) > 1040$ MeV$/c^2$	all
	$B^0 \to D^- e^+\nu$	$\lvert\cos\theta_\ell\rvert < 0.8$	except ee high-
	$B^0 \to K^{*0}\gamma$	$\sigma_z(e^+e^-) < 30$ mm	except $\gamma(ee)$
	Comb	$\texttt{NNout} > 0.68$	$\mu\mu$ low-
		$\texttt{NNout} > 0.64$	ee low-
		$\texttt{NNout} > 0.85$	$\mu\mu$ central-
		$\texttt{NNout} > 0.97$	ee central-
		$\texttt{NNout} > 0.40$	$\mu\mu$ high-
		$\texttt{NNout} > 0.93$	ee high-
		$\texttt{NNout} > 0.06$	$J/\psi\,(\mu\mu)$
		$\texttt{NNout} > 0.20$	$J/\psi\,(ee)$
		$\texttt{NNout} > 0.16$	$\gamma(ee)$
		$\texttt{NNout} > 0.68$	$\psi(2S)(ee)$
	Comb, part-reco	$m_{\mathrm{BCM}} > 4680 + 31 \cdot \log(\chi^2_{\mathrm{FD}})$	ee low-
		$m_{\mathrm{BCM}} > 4437 + 64 \cdot \log(\chi^2_{\mathrm{FD}})$	ee central-
		$m_{\mathrm{BCM}} > 3380 + 140 \cdot \log(\chi^2_{\mathrm{FD}})$	$\gamma(ee)$

samples. The simultaneous fit allows to share parameters e.g. those describing data-simulation differences. The yields of the rare channels are parameterised as a function of the corresponding J/ψ yields as

$$N_{\ell\ell}(r_{\ell\ell}, N_{J/\psi}) = N_{J/\psi} \cdot \varepsilon^{\mathrm{rel}} \cdot r_{\ell\ell}, \tag{5.10}$$

where $\varepsilon^{\mathrm{rel}}$ is the relative efficiency between the rare and resonant channels (given in Tab. 5.10). Consequently, $r_{\ell\ell}$ corresponds to the efficiency corrected ratio of the raw rare and resonant yields:

$$R_{\ell\ell} = \frac{N_{\ell\ell}/\varepsilon^{\ell\ell}}{N_{J/\psi}/\varepsilon^{J/\psi(\ell\ell)}}. \tag{5.11}$$

The two ratios, R_{ee} and $R_{\mu\mu}$, are then used to determine the $R_{K^{*0}}$ quantity, as described in Sec. 5.11. The following subsections contain a description of the line shapes used to model the signal and background components in each sample.

5.8.1 Muon channels

For the rare and resonant $\mu\mu$ channels the fitted variable is the $m(K\pi\mu\mu)$ invariant mass coming from a kinematic fit where all vertices are required to point to their mother particle. In the resonant case it is beneficial to also constrain the the dimuon mass to the known J/ψ mass. The effect of the kinematical constraint is to improve the mass resolution by roughly a factor of 2, which results in a more stable fit. Furthermore, mis-reconstructed background candidates are pushed away from the B^0 peak, which allows to use a wider mass window to better constrain the combinatorial background slope. The mass spectrum is fitted in the range 5150–5800 MeV/c^2 with the lower limit chosen to totally exclude partially reconstructed background. As it is not needed to model partially reconstructed backgrounds in the fit this also eliminates the systematic uncertainties associated with the knowledge of their shape.

5.8.1.1 $B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ PDF

The signal PDF adopted to describe the reconstructed $m(K\pi\mu\mu)$ invariant mass of $B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ candidates is the sum of a Double Crystal Ball [93] (DCB) function with opposite-side tails and a Gaussian function with a common mean, μ:

$$\begin{aligned}\mathcal{P}_{\rm sig}(m|\vec{\lambda}) = \;& f_{\rm CB1}\cdot\mathcal{P}_{\rm CB}(m|\mu,\sigma_1,\alpha_1,n_1) + \\ & f_{\rm CB2}\cdot\mathcal{P}_{\rm CB}(m|\mu,\sigma_2,\alpha_2,n_2) + (1-f_{\rm CB1}-f_{\rm CB2})\cdot\mathcal{P}_{\rm Gauss}(m|\mu,\sigma_3)\,,\end{aligned}$$

where $f_{\text{CB}i}$ is the relative fraction of candidates falling in the i^{th} Crystal Ball function, σ_i is the width, α_i and n_i are the parameters controlling the power law tail of each CB, and σ_3 is the width of the Gaussian function.

As a first step, the parameters of the signal PDF are extracted by fitting the $m(K\pi\mu\mu)$ distribution on $B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ simulation and fixed for the fit to the data. Figure E.1 shows the fitted simulated distribution for the normalisation channel, while fits or the rare channel in the three q^2 bins are reported in Appendix E. In order to account for possible discrepancies in the invariant mass distribution between data and simulation, the mass is allowed to shift, $\mu \to \mu + m'$, and the widths are allowed to scale, $\sigma_i \to c \cdot \sigma_i$, where the scale factor c is common between the three σs.

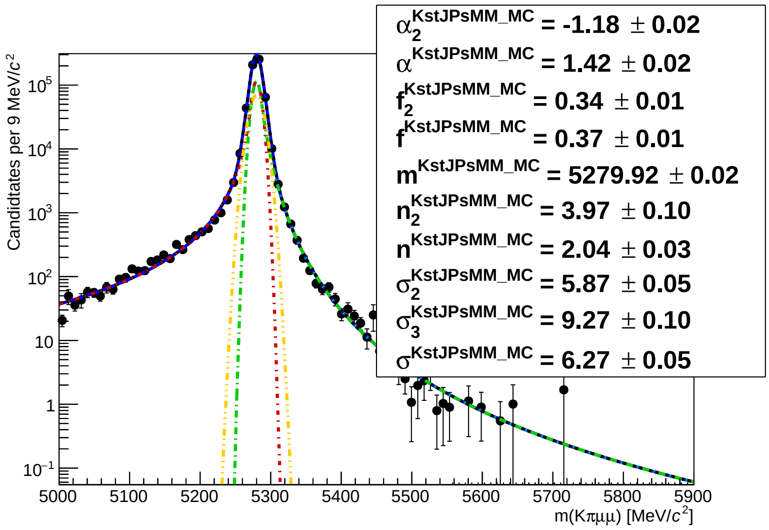

Figure 5.17: Fitted $m(K\pi\mu\mu)$ mass spectrum for $K^{*0}J/\psi$ simulated events.

In summary, the signal PDF for the $J/\psi\,(\mu\mu)$ channel fit on data is defined as

$$\mathcal{P}_{\text{J}/\psi\,(\mu\mu)}(m|m',c) = f_{\text{CB1}}\cdot\mathcal{P}_{\text{CB}}(m|m',c)+f_{\text{CB2}}\cdot\mathcal{P}_{\text{CB}}(m|m',c)+(1-f_{\text{CB1}}-f_{\text{CB2}})\cdot\mathcal{P}_{\text{Gauss}}(m|m',c)\,.$$

where the only free parameters are the mass shift, m' and the width scale factor, c.

The following backgrounds are considered:

- *Combinatorial*: modelled with an exponential function;
- $\Lambda_b^0 \to pK(J/\psi \to \mu^+\mu^-)$: described using simulated events to which the full selection selection and weights for the pK Dalitz plot are applied; this distribution has a broad shape under the signal peak and is smoothed using the `RooKeysPdf` class of the RooFit [109] package;
- $B_s^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$: described using the same PDF adopted for the signal, but a different central value, μ, which is set at the B_s^0 nominal mass. The same shift m' is used as for the signal.

5.8.1.2 $B^0 \to K^{*0}\mu^+\mu^-$ PDF

The signal PDF adopted to describe the reconstructed 4-body invariant mass of the $B^0 \to K^{*0}\mu^+\mu^-$ candidates is a DCB function with opposite-side tails with a common mean, μ. The parameters of the PDF are fixed to values obtained by fitting simulated candidates, separately in each q^2 interval. As for the charmonium channel, the mass is allowed to shift and the widths are allowed to scale with a common factor:

$$\mathcal{P}_{\mathrm{mm,q^2}}(m|m'_{q^2}, c_{q^2}) = f_{\mathrm{core,q^2}} \cdot \mathcal{P}_{\mathrm{CB}}(m|m'_{q^2}, c_{q^2}) + (1 - f_{\mathrm{core,q^2}}) \cdot \mathcal{P}_{\mathrm{CB}}(m|m'_{q^2}, c_{q^2}).$$

where $f_{\mathrm{core,q^2}}$ is the relative fraction of candidates falling in the first Crystal Ball function, m'_{q^2} is the mass shift and c_{q^2} is the width scale. The subscript "q^2" indicates that independent parameters are used for each q^2 interval.

The background is described by an exponential function in all the three q^2 bins.

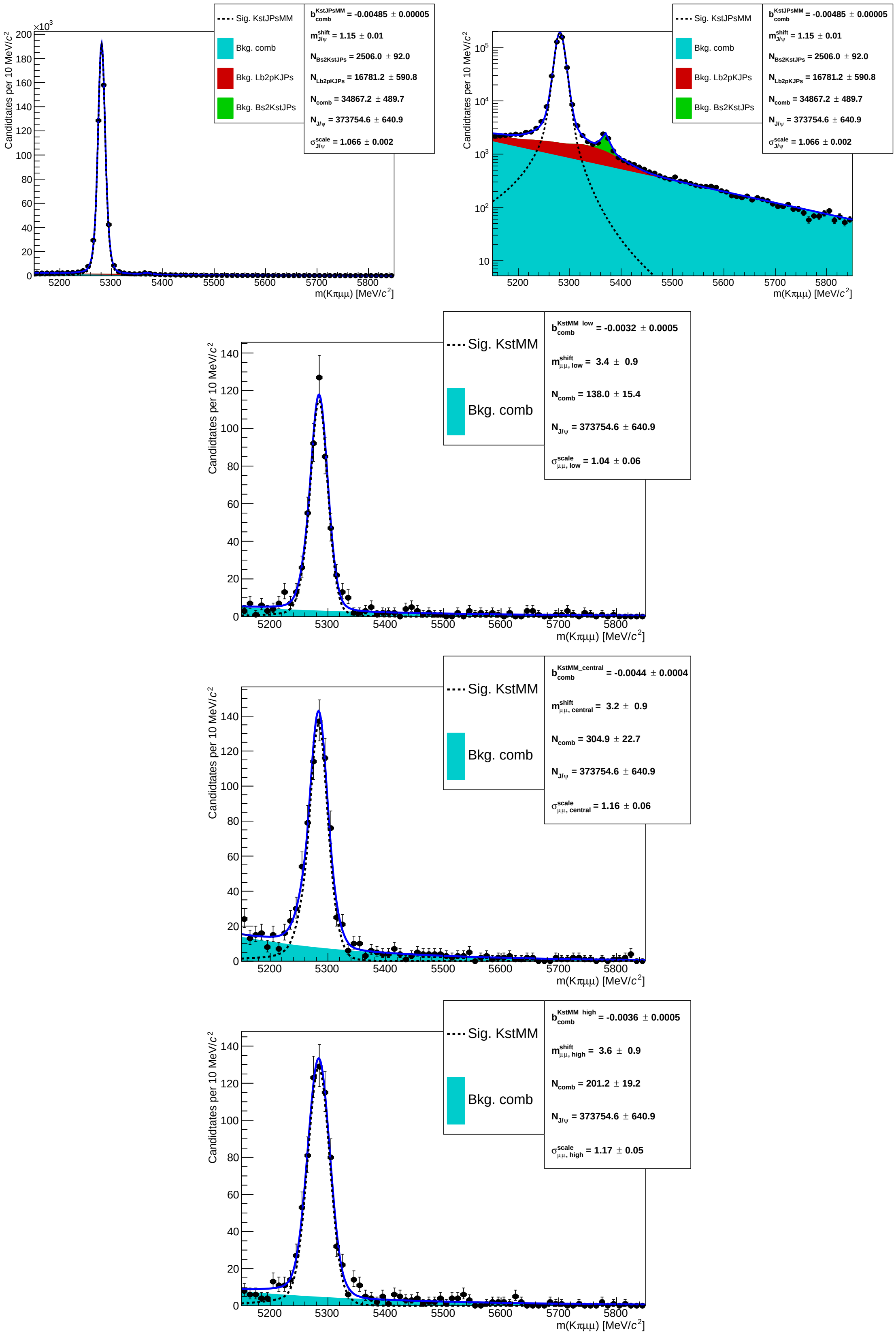

Figure 5.18: From top to bottom fitted $m(K\pi\mu\mu)$ invariant mass distributions for $K^{*0}J/\psi$ candidates and for rare candidates in the low-, central- and high-q^2 intervals. Dashed black lines represent the signal PDFs and filled shapes the background components.

5.8.1.3 Summary

In summary, the free parameters of the simultaneous fit to the $J/\psi\,(\mu\mu)$ and mm candidates are the signal and background yields, the combinatorial background slopes, the mass shifts and the width scales. Figure 5.18 shows the results of the fit to the rare and resonant $\mu\mu$ candidates. Values of the fitted parameters are reported on the plots.

5.8.2 Electron channels

The reconstructed invariant mass of the B^0 depends on which L0 line triggered the event. For this reason, a simultaneous fit to the 4-body invariant mass of the $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ and $B^0 \to K^{*0}e^+e^-$ channels in the three trigger categories is performed. In each trigger category, the $J/\psi\,(ee)$ and ee yields are extracted from the following signal channel categories:

- $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ (with a J/ψ mass constraint);
- $B^0 \to K^{*0}e^+e^-$.

Extra control channels are fit simutanously:

- $B^0 \to K^{*0}(\gamma \to e^+e^-)$ to constrain the yield of partially-reconstructed background in the low-q^2 and the leakage of $B^0 \to K^{*0}\gamma$ into the low-q^2;
- $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ (without the J/ψ mass constraint) to constrain the leakage to $B^0 \to K^{*0}e^+e^-$ in the central-q^2 and the parameters that model residual data-simulation discrepancies;
- $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ (with a $\psi(2S)$ mass constraint) to constrain the leakage to lower and higher q^2 values.

When fitting the variable without a J/ψ mass constraint it is important to fit a wider mass range to better constrain the parameters modelling the radiative tail and the backgrounds; a mass window [4500,6200] MeV/c^2 is used. The lower limit is given by the point in which the q^2 cut (at 6 GeV$^2/c^4$ to separate the rare and resonant channels) starts to affect the 4-body invariant mass distribution.

The invariant mass distributions are different depending on the trigger category and also on the number of bremsstrahlung photons recovered. Therefore, our samples are divided in three trigger categories, as described in Sec. 5.6.1, and three bremsstrahlung categories defined as:

- 0γ: candidates with no photon emitted
- 1γ: candidates with one photon by either of the electrons
- 2γ: candidates with ono photon emitted by each electron

All samples are fitted simultaneously, which allows a better use of the available statistics as the simultaneous fit gathers information from the three categories at the same time. Furthermore, using this method the results for the three categories are naturally combined in a single r_{ee} ratio. The PDFs used to fit the invariant mass distributions are described in the next subsections.

5.8.2.1 Signal PDFs for the electron channels

As for the muon channels, simulated candidates are fitted fist to constrain the shape parameters for the subsequent fit to data. The signal PDFs are built using the following method:

- Simulated $B^0 \to K^{*0}J/\psi(ee)$ and $B^0 \to K^{*0}ee$ events are divided in each trigger and bremsstrahlung category and an independent fit is performed to each sample. A different fit is also performed for the central, J/ψ and high q^2

Table 5.7: Percentages of events with 0, 1 and 2 emitted photons in the three trigger categories, obtained from simulated events.

Trigger	**0γ (%)**	**1γ (%)**	**2γ (%)**
$B^0 \to K^{*0}e^+e^-$ low-q^2			
L0E	34.2	56.0	9.8
L0H	27.8	58.1	14.2
L0I	31.7	56.9	11.4
$B^0 \to K^{*0}e^+e^-$ central-q^2			
L0E	29.2	50.0	20.8
L0H	23.6	50.5	26.0
L0I	28.5	49.9	21.6
$B^0 \to K^{*0}e^+e^-$ high-q^2			
L0E	20.6	51.2	28.2
L0I	10.0	53.8	36.2
$B^0 \to K^{*0}(\gamma \to e^+e^-)$			
L0E	40.4	59.6	–
L0H	32.2	67.8	–
L0I	39.3	60.7	–
$B^0 \to K^{*0}(J/\psi \to e^+e^-)$			
L0E	29.0	50.1	20.8
L0H	18.9	51.3	29.8
L0I	26.9	51.7	21.4
$B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$			
L0E	27.2	51.3	21.5
L0H	17.4	51.5	31.2
L0I	22.0	55.0	23.0

intervals. In the case of the high-q^2 interval it is particularly important to keep signal tail parameters independent from J/ψ channel ones because, as can be seen in Fig. 5.19, the invariant mass distributions are significantly different for the two intervals.

- For each trigger category a PDF is built as the sum of the three PDFs of the bremsstrahlung categories:

$$\mathcal{P}^{\mathrm{L0}}(m) = f_{0\gamma}^{\mathrm{L0}}\mathcal{P}_{0\gamma}^{\mathrm{L0}}(m) + f_{1\gamma}^{\mathrm{L0}}\mathcal{P}_{1\gamma}^{\mathrm{L0}}(m) + (1 - f_{0\gamma}^{\mathrm{L0}} - f_{1\gamma}^{\mathrm{L0}})\mathcal{P}_{2\gamma}^{\mathrm{L0}}(m), \qquad (5.12)$$

where the $\mathcal{P}(m)_{n\gamma}^{L0}$ functions are the chosen PDFs for the trigger and bremsstrahlung categories and the $f_{n\gamma}^{L0}$ parameters are the relative fractions of events falling in each category.

- Most parameters are fixed (details later) and the combined PDF, $P(m)$, is used to fit real data divided only in trigger categories.

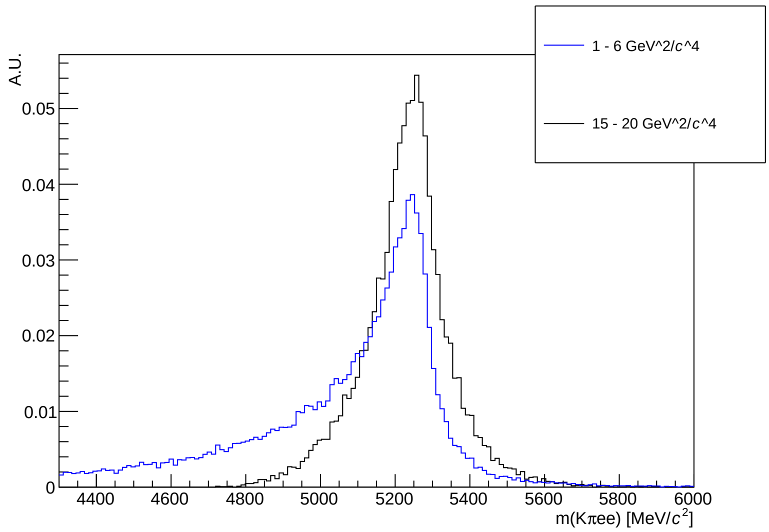

Figure 5.19: Simulated invariant mass of the $K\pi ee$ system in the $1.1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.

The distribution of the $m(K\pi ee)$ mass in the 0γ category is characterised by a sharp tail on the righthand side and is described with a Crystal Ball function (CB), while the 1γ and 2γ categories are modelled using the sum of a Crystal Ball and a Gaussian function (CBG) with independent parameters. In all the bremsstrahlung categories the distribution of the 4-body invariant mass with a mass with the J/ψ mass constraint is modelled using the sum of a DCB and a Gaussian functions as done in the muon fit. To account for possible data-simulation discrepancies, the mass (widths) of each trigger PDF is allowed to shift (scale), similarly to the muon channels. However, due to the larger background contamination these parameters are shared between the rare and the $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ control sample (no J/ψ mass constraint). The tail parameters are similar between the $J/\psi\,(ee)$ and the central-q^2 but this is not the case at high-q^2, as can be seen in Fig. 5.19, due to the migration of candidates in the tail to lower reconstructed q^2. For this reason the initial parameters for each candidate type are obtained fitting a simulated sample of the same candidate type.

The $f^{L0}_{n\gamma}$ fractions have been shown to be in good agreement between resonant data and simulation and therefore they are fixed to the simulated values, separately for the normalisation channel and each q^2 interval. Table 5.7 lists the percentages of candidates with 0, 1 and 2 recovered photons for each trigger category.

In summary the signal PDF for the fit on data is defined as:

$$\begin{aligned}\mathcal{P}_{sig}(m;c,m')^{\rm trg} &= f^{\rm L0}_{0\gamma}\mathcal{P}^{\rm L0}_{0\gamma}(m;c,m') \\ &+f^{\rm L0}_{1\gamma}\mathcal{P}^{\rm L0}_{1\gamma}(m;c,m') + (1-f^{\rm L0}_{0\gamma}-f^{\rm L0}_{1\gamma})\mathcal{P}^{\rm L0}_{2\gamma}(m;c,m')\end{aligned} \tag{5.13}$$

where the free parameters are: c, the scaling factor for the widths, and m', the mass shift.

5.8.2.2 Background PDFs for the electron channels

This section reports the background components considered for each fitted sample.

$B^0 \to K^{*0}e^+e^-$ low-q^2

- *Combinatorial*: described using an exponential function; the yield and slope parameters are free to vary in the fit;

- *Partially-reconstructed* (hadronic): the shape is obtained from a $K_1^+(1270)$ simulated samples as in the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ case; the fraction of partially-reconstructed candidates with respect to signal ones is expected to be very similar to that in $B^0 \to K^{*0}(\gamma \to e^+e^-)$ and therefore the normalisation is fixed as:
$$N^{\rm part-reco}_{e^+e^-,\,{\rm low}} = N_{e^+e^-} \cdot \frac{N^{\rm part-reco}_{\gamma(ee)}}{N_{\gamma(ee)}},$$
where $N^{\rm part-reco(hadronic)}_{\gamma(ee)}/N_{\gamma(ee)}$ is the fraction of the hadronic partially-reconstructed background relative to the signal yield in the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ channel;

- $B^0 \to K^{*0}\gamma$ *leakage*: the leakage from the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ decay in the low-q^2 region is modelled using a simulated candidates that pass the low-q^2 requirements: the distribution is smoothed using a `RooKeysPdf`; the normalisation is fixed to the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ yield, $N_{\gamma(ee)}$ as:

$$N^{\text{leak}}_{e^+e^-,\,\text{low}} = N_{\gamma(ee)} \cdot f^{\text{leak, MC}}_{\gamma(ee)},$$

where $f^{\text{leak, MC}}_{\gamma(ee)}$ is the fraction of $\gamma(ee)$ simulated candidates which leaks in the low-q^2 region.

$B^0 \to K^{*0}e^+e^-$ central-q^2

- *Combinatorial*: described using an exponential function; the yield and slope parameters are free to vary in the fit.
- *Partially-reconstructed* (hadronic): modelled in the same way as described for the low-q^2 but in this case the normalisation is left free to vary.
- $B^0 \to K^{*0}J/\psi$ *leakage*: the leakage from the J/ψ radiative tail into the central-q^2 interval is modelled by selecting simulated $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates which pass the central-q^2 requirements and smoothing the distributions with kernel estimation method. The normalisation is fixed to the $B^0_s \to K^{*0}(J/\psi \to e^+e^-)$ yield, $N_{J/\psi\, ee}$, as:

$$N^{\text{leak}}_{e^+e^-,\,\text{central}} = N_{J/\psi\,(ee)} \cdot f^{\text{leak, MC}}_{J/\psi\,(ee)},$$

where $f^{\text{leak, MC}}_{J/\psi\,(ee)}$ is the fraction of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ simulated events reconstructed in the central-q^2 interval.

$B^0 \to K^{*0}e^+e^-$ high-q^2

- *Combinatorial*: modelled using a shape obtained by reversing the NN output cut on data, which has the effect of selecting background candidates instead of signal ones. Figure 5.21 shows the invariant mass distributions for different anti-cuts on the electron and muon samples at high-q^2. The shapes are very similar between the two samples and as a function of the cut value. In order to have a larger statistics, the shape is taken from the muon sample with a tight NN output anti-cut at 0.1 and smoothed with a `RooKeysPdf`;

- *Partially-reconstructed* (hadronic): the hadronic mis-reconstructed background is modelled in the same way as in the central-q^2 interval. The normalisation is left free to vary in the fit;

- $B^0 \to K^{*0}\psi(2S)$ *leakage*: the leakage from the $\psi(2S)$ radiative tail is modelled using simulated $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ events in the high-q^2 region. The normalisation is fixed to the $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ yield, $N_{\psi(2S)(ee)}$ as:

$$N^{\rm leak}_{e^+e^-,\,{\rm high}} = N_{\psi(2S)(ee)} \cdot f^{\rm leak,\,MC}_{\psi(2S)(ee)},$$

  where $f^{\rm leak,\,MC}_{\psi(2S)(ee)}$ is the fraction of $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ simulated candidates leaking in the high-q^2 interval.

$B^0 \to K^{*0}\gamma$

- *Combinatorial*: described using an exponential function; the yield and slope parameters are free to vary in the fit;

- *Partially-reconstructed* (hadronic): the shape is obtained from simulation similarly to the $J/\psi\,(ee)$ mode. However, as there are no inclusive samples for the rare channel, a sample including higher K^{*0} resonances, such as $K_1^+(1270)$, which is the dominant component, is used.; the yield is left free to vary;

- $B^0 \to K^{*0}e^+e^-$ *leakage*: as the $K^{*0}\gamma$ was added to the low-q^2 also the low-q^2 leakage is added to $K^*\gamma$. The yield is constrained to the N^{low}_{ee} yield.

$B^0 \to K^{*0}J/\psi$ and $B^0 \to K^{*0}\psi(2S)$

The following backgrounds are considered for the fits to the invariant mass of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates (with the J/ψ mass constraint):

- *Combinatorial*: described using an exponential function. The yield and slope parameters are free to vary in the fit;

- $\Lambda_b^0 \to pK(J/\psi \to e^+e^-)$: described using simulated events to which the full selection is applied. This distribution has a broad shape under the signal peak and is smoothed using a `RooKeysPdf`. The normalisation is constrained to the $\Lambda_b^0 \to pK(J/\psi \to \mu^+\mu^-)$ yield returned by the $\mu\mu$ fit after correcting for efficiency differences between final states with muons and electrons.

- $B_s^0 \to K^{*0}(J/\psi \to e^+e^-)$: described using the same PDF adopted for the signal, but a different central value, m_0, which is set at the B_s^0 nominal mass. The normalisation is constrained to the $B^0 \to K^{*0}(J/\psi \to \mu^+\mu^-)$ yield returned by the $\mu\mu$ fit after correcting for efficiency differences between final states with muons and electrons;

The J/ψ mass constraint has the effect of pushing the partially-reconstructed background away from the peak outside the fit window. Therefore it does not need to be modelled. Instead the fit model for the variable without J/ψ mass constraint, used as control sample, includes further components:

- *Partially-reconstructed*: the partially-reconstructed background from both higher hadronic and $c\bar{c}$ resonances is modelled using inclusive $B^0 \to J/\psi X$ simulated events to which the full selection is applied. The invariant mass distributions of these candidates, shown in Fig. 5.20, is smoothed using a kernel estimation method and the yield is left free to vary;

- $B^0 \to K^{*0}\psi(2S)$ *leakage*: the leakage from the $\psi(2S)$ radiative tail into the J/ψ interval is modelled by selecting simulated $\psi(2S) \to e^+e^-$ which pass the requirements for J/ψ candidates. The normalisation is fixed to the $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ yield, $N_{\psi(2S)(ee)}$, as:

$$N^{\text{leak}}_{J/\psi(ee)} = N_{\psi(2S)(ee)} \cdot f^{\text{leak, MC}}_{\psi(2S)(ee)},$$

where $f^{\text{leak, MC}}_{\psi(2S)(ee)}$ is the fraction of $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ simulated events reconstructed in the J/ψ interval.

For the fit to $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$, which includes a $\psi(2S)$ mass constraint, only the combinatorial background is considered and described using an exponential function.

5.8.2.3 Summary of the fit to the electron samples

In summary, the free parameters in the fit to data are:

- the $B^0 \to K^{*0}(J/\psi \to e^+e^-)$, $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ and $B^0 \to K^{*0}(\gamma \to e^+e^-)$ yields in each trigger category;
- the r_{ee} ratio common to all trigger categories; one for the low, one for the central- and one for the high-q^2 region;
- one mass shift, m', and one width scale factor, c, for the signal PDF common between $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ and $B^0 \to K^{*0}e^+e^-$, but different for the three trigger categories and for $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ and $B^0 \to K^{*0}(\gamma \to e^+e^-)$;
- the yield and slope, when applicable (e.g. no slope at high-q^2), of the combinatorial background in each trigger category and for each channel;
- the yield of the backgrounds when not fixed as described in the previous section.

Fits to simulated $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates are shown in Appendix E, while fits to real candidates are shown in Fig. 5.22 for the normalisation channel, in Fig. 5.23 for the rare channel and in Fig. 5.24 for the control channels. For simplicity the latter two figures show the sub of the three trigger categories, while the separate plots are reported in Appendix F, where fitted parameters are also reported on the

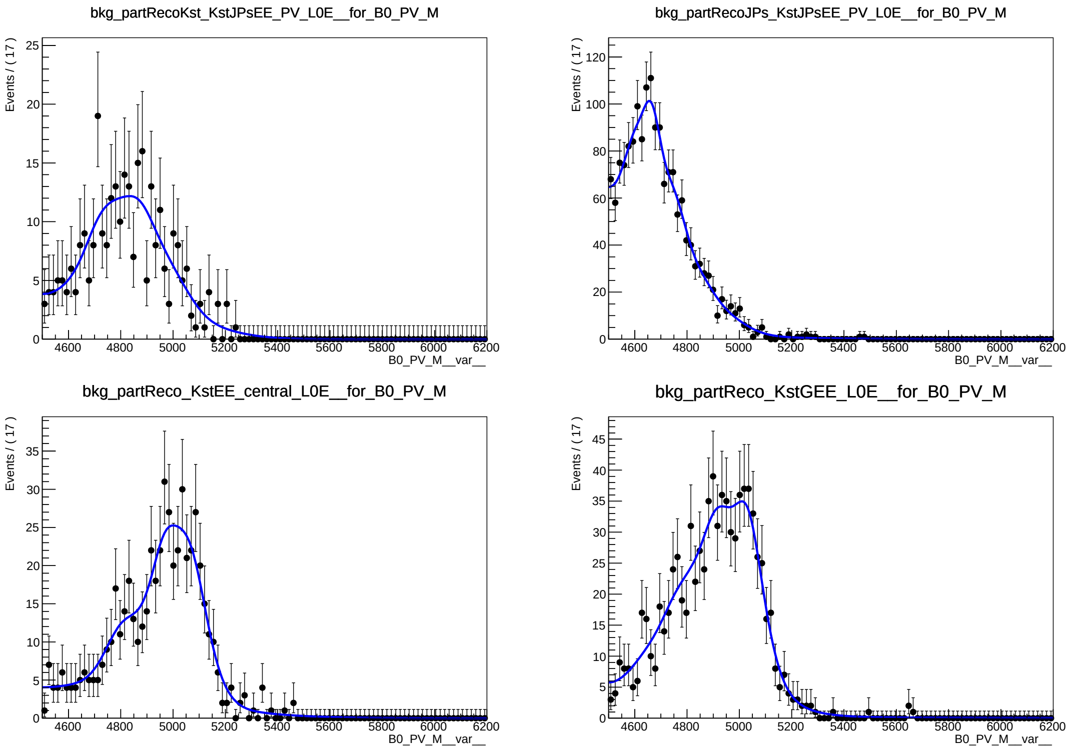

Figure 5.20: Distributions of the $m(K\pi ee)$ invariant mass for the (top left) hadronic and (top right) leptonic mis-reconstructed background to $B^0 \to K^{*0}(J/\psi \to e^+e^-)$. Distributions of the $m(K\pi ee)$ invariant mass for decays involving higher K^{*0} resonances in the (bottom left) central-q^2 interval and (bottom right) the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ interval.

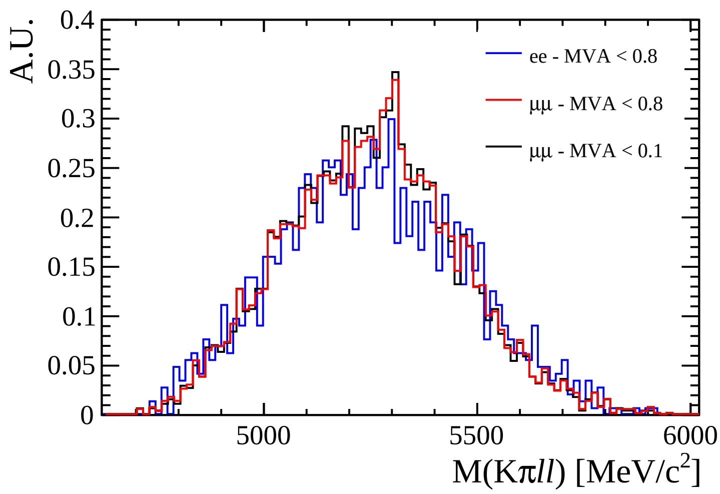

Figure 5.21: Distributions of the $m(K\pi\ell\ell)$ invariant mass for $B^0 \to K^{*0}\ell^+\ell^-$ candidates selected with a reversed cuts on the NN output.

Table 5.8: Raw yields of events found fitting invariant mass distributions of the rare and resonant events.

Sample	**1.1–6 GeV^2/c^4**	**15–20 GeV^2/c^4**	J/ψ
$\mu\mu$	626 ± 30	605 ± 27	333113 ± 604
ee L0E	132 ± 17	137 ± 27	48601 ± 326
ee L0H	31.7 ± 4.2	–	4324 ± 94
ee L0I	49.6 ± 6.5	–	12791 ± 172

plots. In the high-q^2 interval, above 15 GeV$^2/c^4$, the efficiency for the L0Hadron trigger becomes very low as the K^* has very low momentum. In this region only 9 candidates are found spread in the interval $4500 < m(K\pi ee) < 6000$ MeV/c^2. Therefore only L0E and L0I triggered events are fitted in this region.

5.8.3 Event yields

Table 5.8 reports raw yields obtained from the fits described in the previous subsections. The values for the rare channels are not directly floating in the fits but, as described in Sec. 5.8, they are parameterised as a function of the number of resonant events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare branching fractions. Measured values of these ratios are reported in Tab. 5.16.

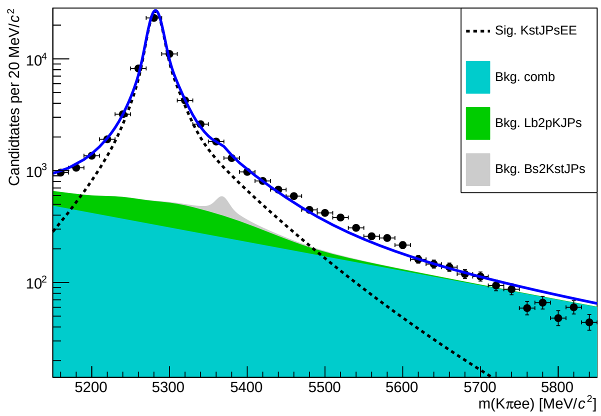

Figure 5.22: Fit to the mass constrained $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates. The dashed black line (shaded shapes) represents the signal (background) PDF.

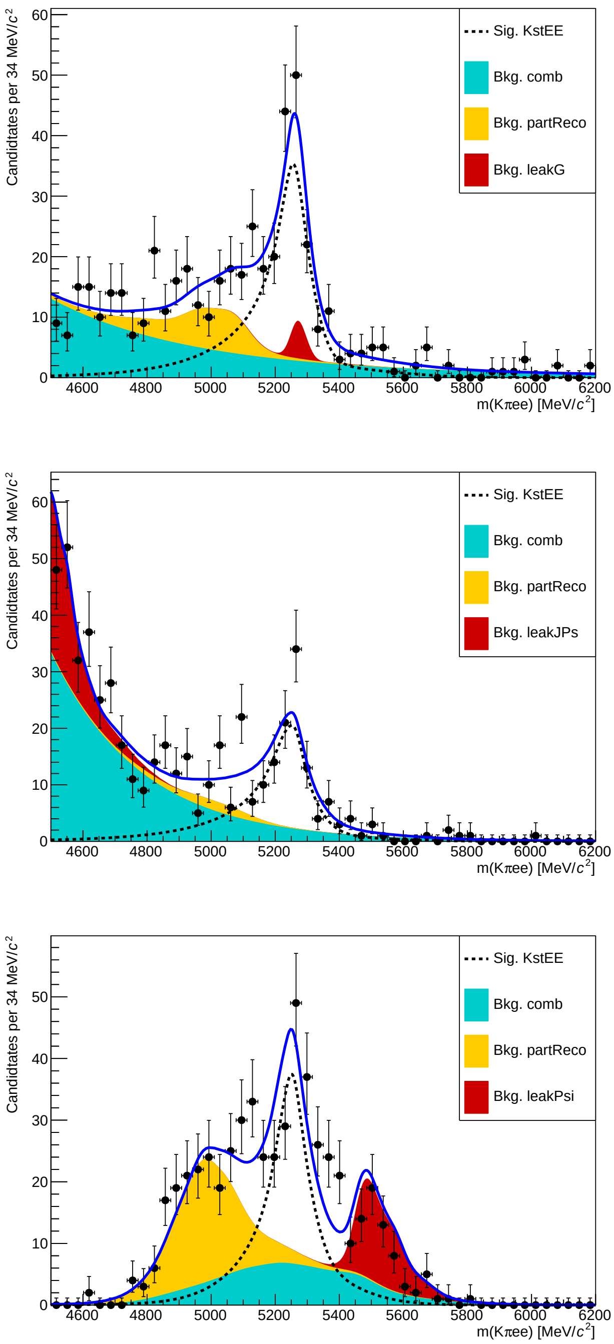

Figure 5.23: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}e^+e^-$ candidates. From top to bottom for the low-, central- and high-q^2 intervals. The dashed black line (shaded shapes) represents the signal (background) PDF.

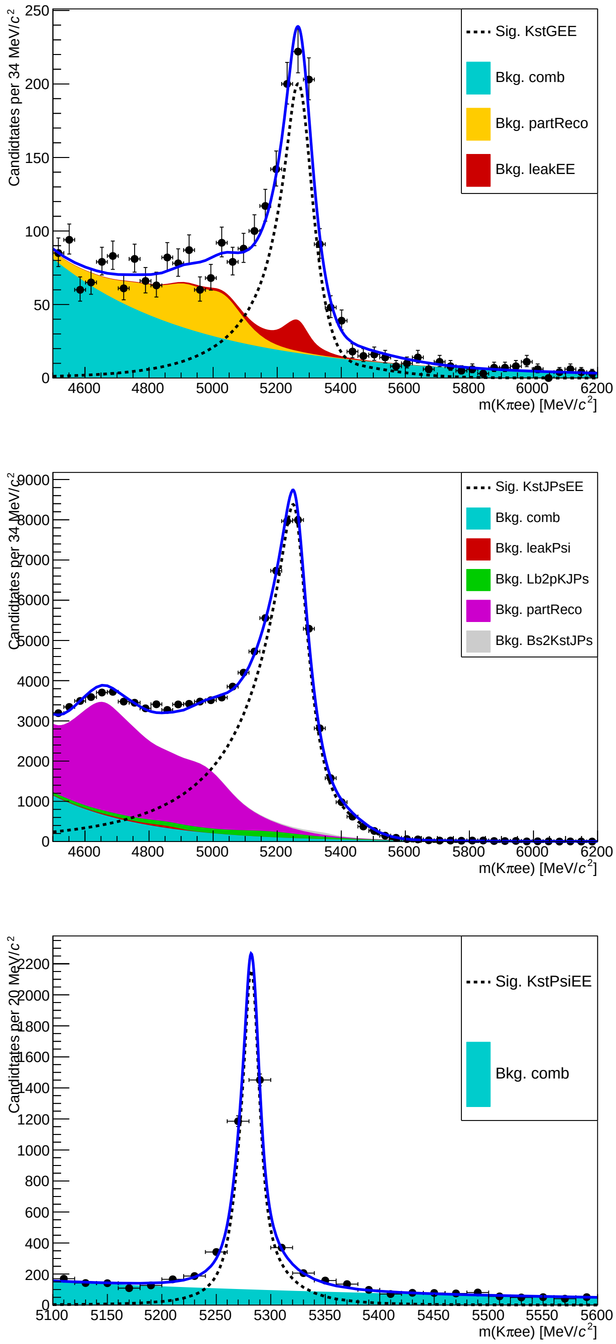

Figure 5.24: Fit to the $m(K\pi ee)$ invariant mass of control channel candidates. From top to bottom: invariant mass distribution without mass constraint of $B^0 \to K^{*0}(\gamma \to e^+e^-)$ and $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates and mass constrained mass of $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ candidates. The dashed black line (shaded shapes) represents the signal (background) PDF.

5.9 Efficiency

The efficiency for each of the decay channels is calculated according to the formula

$$\varepsilon^{tot} = \varepsilon^{\text{geom}} \cdot \varepsilon^{\text{reco|geom}} \cdot \varepsilon^{\text{PID|reco}} \cdot \varepsilon^{\text{trig|PID}} \cdot \varepsilon^{\text{MVA|trig}} \cdot \varepsilon^{\text{BCM|MVA}}$$

, where the first term is the efficiency to have final state particles in the LHCb detector acceptance; the second term ($\varepsilon^{\text{reco|geom}}$) carries information on reconstruction and stripping efficiency; the third ($\varepsilon^{\text{PID|reco}}$) corresponds to the efficiency of the PID requirements; the fourth ($\varepsilon^{\text{trig|PID}}$) handles the trigger efficiency for those events which are selected by the pre-selection process; and, finally, the latter term deals with the efficiency of the neural network classifier. Reconstruction, trigger, MVA and BCM efficiencies are evaluated using simulated data samples with the trigger efficiency for $B^0 \to K^* J/\psi$ being cross-checked using the data-driven TIS-TOS method as described in Sec. 3.6.3. The PID efficiency is calculated with a data-driven method as described in Sec. 5.9.3.

Absolute efficiencies for the muon and electron normalisation channel are reported in Tab. 5.9 and relative efficiencies between the rare and resonant channel, $\varepsilon(\ell\ell)/\varepsilon(J/\psi\,(\ell\ell))$, are listed in Tab. 5.10; these are the efficiencies which are used in the fit.

Table 5.9: Absolute efficiencies for the resonant ee and $\mu\mu$ channels.

ε	$\mu\mu$	ee		
		L0E	L0H	L0I
$\varepsilon^{\text{geom}}$	0.1598 ± 0.0005	0.1589 ± 0.0005		
$\varepsilon^{\text{reco\|geom}}$	0.0947 ± 0.0001	0.0603 ± 0.0001		
$\varepsilon^{\text{PID\|reco}}$	0.8148 ± 0.0000	0.8222 ± 0.0000		
$\varepsilon^{\text{trig\|PID}}$	0.7511 ± 0.0005	0.1939 ± 0.0005	0.0163 ± 0.0002	0.0707 ± 0.0003
$\varepsilon^{\text{MVA\|trig}}$	0.8944 ± 0.0004	0.8597 ± 0.0007	0.8983 ± 0.0006	0.8276 ± 0.0017
$\varepsilon^{\text{Total}}$	0.0083 ± 0.0000	0.0013 ± 0.0000	0.0001 ± 0.0000	0.0005 ± 0.0000

Table 5.10: Relative efficiencies, $\varepsilon^{rel} = \varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$, for the ee and $\mu\mu$ channels in the central and high q^2 intervals.

$\boldsymbol{\varepsilon}$	$\boldsymbol{\mu\mu}$	$\boldsymbol{ee}$		
		L0E	L0H	L0I
		low-q^2		
$\varepsilon^{\text{geom}}$	1.0200 ± 0.0091	1.0429 ± 0.0084		
$\varepsilon^{\text{reco}\|\text{geom}}$	0.1309 ± 0.0010	0.1961 ± 0.0007		
$\varepsilon^{\text{PID}\|\text{reco}}$	0.9861 ± 0.0003	0.9718 ± 0.0001		
$\varepsilon^{\text{trig}\|\text{PID}}$	0.8103 ± 0.0048	0.6478 ± 0.0058	2.5556 ± 0.0455	1.2748 ± 0.0139
$\varepsilon^{\text{MVA}\|\text{trig}}$	0.9528 ± 0.0024	0.9568 ± 0.0014	0.9570 ± 0.0013	0.9463 ± 0.0030
$\varepsilon^{\text{BCM}\|\text{MVA}}$	–	0.9394 ± 0.0014	0.9492 ± 0.0013	0.9590 ± 0.0023
ε^{tot}	0.7810 ± 0.0168	0.5809 ± 0.0097	2.2685 ± 0.0514	1.1073 ± 0.0200
		central-q^2		
$\varepsilon^{\text{geom}}$	1.0200 ± 0.0091	1.0429 ± 0.0084		
$\varepsilon^{\text{reco}\|\text{geom}}$	0.1891 ± 0.0012	0.1580 ± 0.0006		
$\varepsilon^{\text{PID}\|\text{reco}}$	0.9784 ± 0.0002	0.9672 ± 0.0001		
$\varepsilon^{\text{trig}\|\text{PID}}$	0.8925 ± 0.0038	0.7909 ± 0.0069	2.1344 ± 0.0439	1.1208 ± 0.0141
$\varepsilon^{\text{MVA}\|\text{trig}}$	0.9068 ± 0.0024	0.8397 ± 0.0024	0.8512 ± 0.0022	0.7946 ± 0.0054
$\varepsilon^{\text{BCM}\|\text{MVA}}$	–	0.8960 ± 0.0020	0.8978 ± 0.0020	0.9283 ± 0.0037
ε^{tot}	0.7171 ± 0.0124	0.8145 ± 0.0157	2.2235 ± 0.0595	1.0542 ± 0.0236
		high-q^2		
$\varepsilon^{\text{geom}}$	1.0200 ± 0.0091	1.0429 ± 0.0084		
$\varepsilon^{\text{reco}\|\text{geom}}$	0.1172 ± 0.0009	0.0530 ± 0.0003		
$\varepsilon^{\text{PID}\|\text{reco}}$	1.0286 ± 0.0001	1.0113 ± 0.0002		
$\varepsilon^{\text{trig}\|\text{PID}}$	1.1122 ± 0.0038	1.5639 ± 0.0148	–	0.8090 ± 0.0195
$\varepsilon^{\text{MVA}\|\text{trig}}$	0.8986 ± 0.0027	0.8228 ± 0.0036	–	0.7201 ± 0.0115
ε^{tot}	0.7843 ± 0.0155	0.6063 ± 0.0131	–	0.2745 ± 0.0095

5.9.1 Geometric efficiency

In order to save disk space, simulated samples only contain decays with final daughters in the LHCb detector acceptance, which can therefore be reconstructed. This corresponds to the requirement for each of the final particles to have polar angle θ between 10 and 400 mrad. The efficiency of this cuts is obtained using a generator level simulated sample.

5.9.2 Reconstruction efficiency and bin migration

The reconstruction efficiency is here defined as the efficiency to reconstruct each decay channel given that its daughters are into the geometrical acceptance of the detector. This includes both the probability that the final particles generate observable signatures and the efficiency of all the pre-selection requirements described in Sec. 5.6, including those done to remove peaking backgrounds. The efficiency of the PID cuts is kept separate as it is known to be not well simulated and there are reliable data-driven methods which can be used to extract it (see Sec. 5.9.3).

5.9.2.1 Bin migration

It can happen that events generated in a q^2 interval are reconstructed in a different one, this is referred to as "bin migration" and can be due to two different effects. First of all, as the resolution of real detectors is not perfect, events close to the edges of the considered intervals can fall on the wrong side of the edge. This effect is only important in case of non-flat true distributions, as the amount of bin migration in the two directions is different. The second possible source of bin migration are systematic effects due, for example, to the presence of bremsstrahlung photons that cannot be recovered. It is particularly important to take into account the bin migration in the electron channels case because more photons are radiated from the final state and the mass resolution is worse. Figure 5.25 shows the response matrix for simulated

$B^0 \to K^{*0}e^+e^-$ events, which represents the correlation between reconstructed and generated q^2. In the ideal case of perfect resolution this plot would look like a diagonal line and in case no bias is present its slope would be 1. Table 5.11 lists the net amounts of bin migration, M_{net}, in the considered q^2 intervals defined as:

$$M_{net} = N(\text{in} \to \text{in}) + N(\text{out} \to \text{in}) - N(\text{in} \to \text{out}) \tag{5.14}$$

where $N(\text{in} \to \text{in})$ is the number of candidates that are generated and reconstructed inside the considered interval, $N(\text{out} \to \text{in})$ the number of candidates that are generated outside the interval but reconstructed inside and $N(\text{in} \to \text{out})$ the number of candidates generated inside that fall outside. The reconstruction efficiency is calculated comparing generated to reconstructed samples and therefore already includes bin migration effects. Nevertheless, it is useful to single out this component to better asses the corresponding systematic uncertainty.

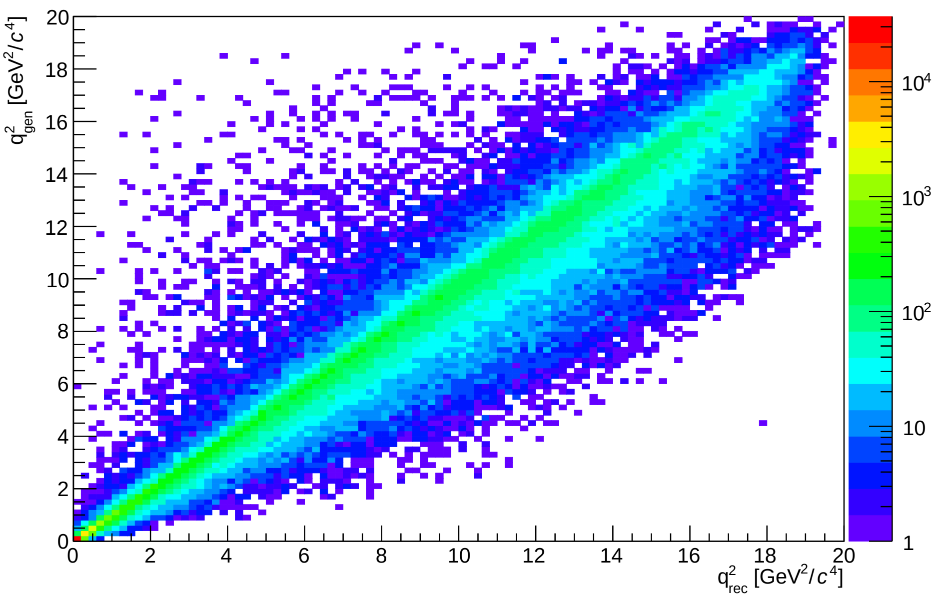

Figure 5.25: Generated versus reconstructed q^2 in simulated $B^0 \to K^* e^+e^-$ events.

5.9.3 PID efficiency

The simulation is known not to reliably describe particle ID variables and therefore a data-driven method is used to obtain this efficiency component. This is done using the `PIDCalib` package described in Sec. 2.8.1. Furthermore, the same method is used to weight the simulation in order to calculate the MVA and trigger efficiencies. The package `PIDCalib` allows to divide the phase-space in intervals of quantities relevant for the determination of the PID efficiency and obtain a data-driven efficiency for each interval. For this analysis the phase-space is divided in equi-populated bins of momentum and pseudorapidity of the particle under study. Figure 5.26 shows performance tables for pions, kaons, muons and electrons. Once the efficiency tables are obtained for each particle, the total efficiency is calculated for each candidate as the product of the four final particles efficiencies. $\varepsilon^{ev} = \varepsilon_K \cdot \varepsilon_\pi \cdot \varepsilon_{\ell_1} \cdot \varepsilon_{\ell_2}$. Finally, as the decay channel under study generally has different kinematical distributions than the calibration sample, the total efficiency is found by averaging over simulated events.

$$\varepsilon_{PID} = \frac{1}{N}\sum_{i}^{N} \varepsilon_K(p_K^i, \eta_K^i) \cdot \varepsilon_\pi(p_\pi^i, \eta_\pi^i) \cdot \varepsilon_\ell(p_{\ell_1}^i, \eta_{\ell_1}^i) \cdot \varepsilon_K(p_{\ell_2}^i, \eta_{\ell_2}^i) \tag{5.15}$$

5.9.4 Trigger efficiency

While the trigger efficiency for the muon channels is calculated using simulated events, for the electron channels a combination of simulation and data-driven methods is used. The efficiency of the software stage, HLT, is always obtained from

Table 5.11: Net bin migration amounts (M_{net}) in the considered q^2 intervals. Positive values indicate "net in", negative values "net out".

Sample	**low-q^2**	**central-q^2**	**J/ψ**	**high-q^2**
$\mu\mu$	0.0002 ± 0.0001	-0.0021 ± 0.0003	0.0032 ± 0.0004	-0.0012 ± 0.0000
ee	0.0268 ± 0.0005	0.0663 ± 0.0009	-0.4277 ± 0.0048	-0.0445 ± 0.0003

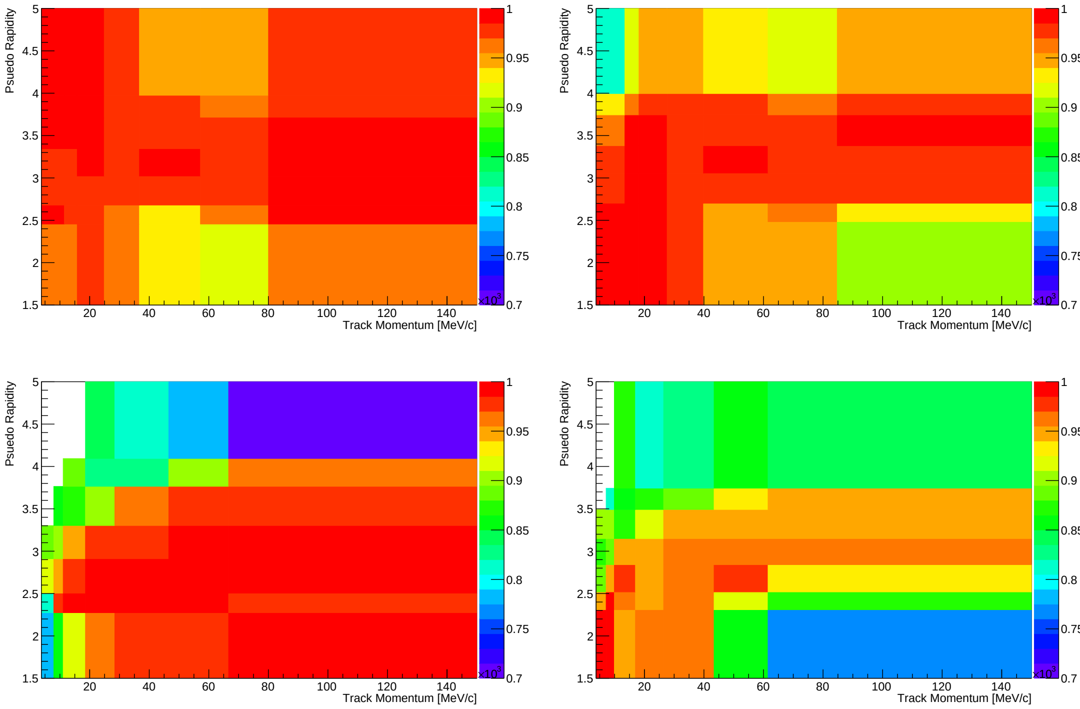

Figure 5.26: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

simulation, while the efficiency of the hardware stage, L0, is obtained using a data-driven method as described in the next subsection. For both muon and electron channels it is possible to use the resonant sample to cross-check the efficiency obtained using the simulation, as explained in Sec. 5.9.4.2.

5.9.4.1 Electron triggers

For the electron channels data is fitted separately in three trigger categories: L0E, L0H and L0I. Therefore the efficiency is calculated separately for each category. While the HLT (1 and 2) efficiency is always derived using simulated events, the L0Electron and L0Hadron efficiencies cannot be reliably modelled in simulation. In fact data-simulation discrepancies are caused by the ageing of the calorimeters, which is not simulated in the Monte Carlo. The ageing modifies the response of the calorimeters with time, which affects the L0 trigger efficiency. Therefore this must be calibrated using data driven-methods.

Tables of efficiencies are obtained applying the TISTOS method to a calibration sample. For each trigger category these tables contain the efficiency as a function of p_{T} of the considered particle and are given for different calorimeter regions as these have different properties (e.g. cell size) due to the different position with respect to the beam line. The considered regions are the inner and outer HCAL, and the inner, middle and outer ECAL. Figure 5.27 shows data-driven efficiencies for the L0Electron trigger in the three ECAL regions.

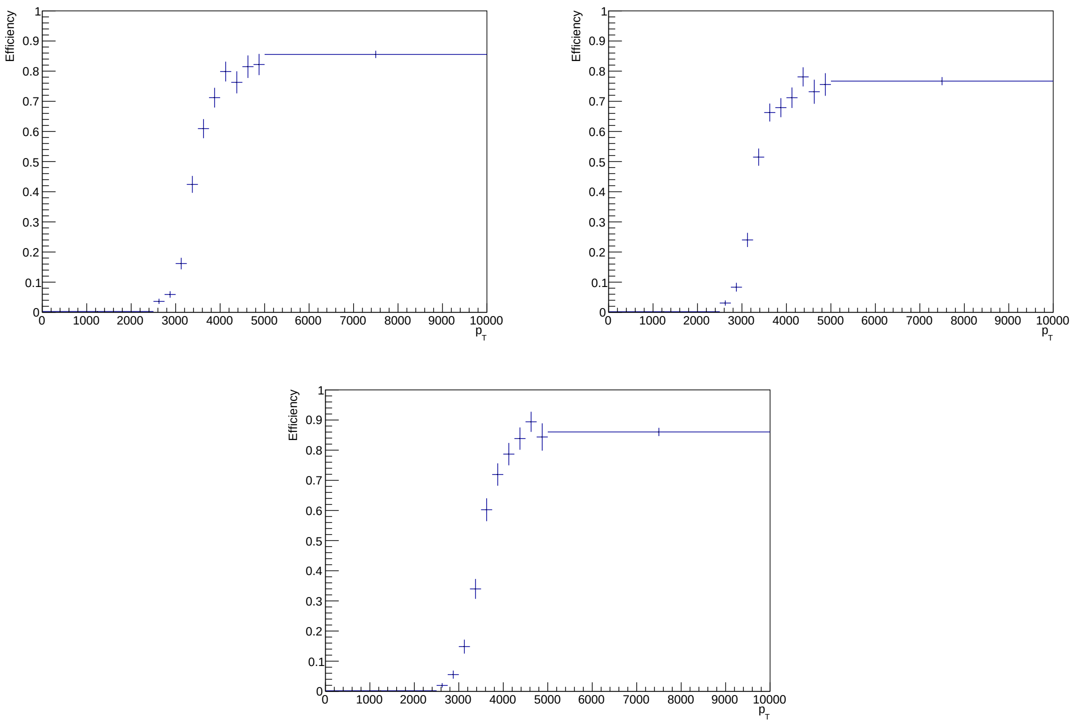

Figure 5.27: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

The probabilities of an event being triggered by L0Electron or L0Hadron are calculated for each candidate as:

$$P_{L0E} = \varepsilon(e^+) + \varepsilon(e^-) - \varepsilon(e^+)\varepsilon(e^-) \text{ and } P_{L0H} = \varepsilon(\pi) + \varepsilon(K) - \varepsilon(\pi)\varepsilon(K).$$

The probability of TIS trigger is defined to be independent of the signal and therefore must be the same in the rare and resonant channels and cancel in their ratio.

Then event by event efficiencies for the three trigger categories are defined to be exclusive in the following way:

- L0E: $\varepsilon^{L0E} = P_{L0E}$, namely the probability that at least one electron triggered;
- L0H: $\varepsilon^{L0H} = P_{L0H} \cdot (1 - P_{L0E})$, namely the probability that at least one hadron triggered but none of the electrons;
- L0I: $\varepsilon^{L0I} = (1 - P_{L0H}) \cdot (1 - P_{L0E})$, namely the probability that neither the hadrons or the electrons in the candidate triggered. Note that in this case ε^{L0I} does not correspond to the efficiency of TIS trigger but to the probability that the event does not fall into the L0E or L0H categories.

Finally, as in the PID case, the total efficiency is found averaging over all events of a simulated sample:

$$\varepsilon^{\text{trg}} = \frac{1}{N} \sum_{i}^{N} \varepsilon^{\text{trg}}(p_{\text{T}}{}^{i}) \tag{5.16}$$

where "trg" is a label indicating the trigger category under consideration.

5.9.4.2 TISTOS cross-check

The efficiency obtained using the simulation is cross-checked applying the TISTOS method, already described in Sec. 3.6.3, to resonant data. For this purpose a sample of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates triggered independent-of-signal (TIS) is used as control sample. As data also contains non negligible amounts of background a narrow interval around the peak, dominated by the signal, is selected and the $_s\mathcal{P}lot$ method is used to remove residual background in the data sample. Results are shown in Tab. 5.12, where the efficiency obtained using the TISTOS method is compared between data and simulation. These are found to be in agreement for the muon channel, while they show deviations in the electron channels. In particular a significant discrepancy is found, for the L0I category, for which the procedure explained in Sec. 5.9.4.1 does not ensure a correct calibration. The table also reports

Table 5.12: Trigger efficiencies obtained using the TISTOS method on simulated and real $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays.

Sample	**MC**	**Data**	**Correction factor**
$J/\psi \to \mu\mu$	0.797 ± 0.002	0.803 ± 0.004	1.0073
$J/\psi \to ee$ L0E	0.268 ± 0.002	0.255 ± 0.004	0.9536
$J/\psi \to ee$ L0H	0.028 ± 0.001	0.026 ± 0.002	0.9269
$J/\psi \to ee$ L0I	0.017 ± 0.001	0.011 ± 0.001	0.6760

a correction factor obtained according to the formula

$$f = 1 + \frac{\varepsilon_{data}^{\mathrm{TISTOS}} - \varepsilon_{MC}^{\mathrm{TISTOS}}}{\varepsilon_{MC}^{\mathrm{TISTOS}}}, \tag{5.17}$$

which can be used to correct the absolute resonant yields. On the other hand, even though discrepancies are present, they should cancel out in the ratio between the rare and J/ψ channels; only the residual discrepancy on this ratio is relevant for the measurement of $R_{K^{*0}}$. In order to check if discrepancies cancel out we need to obtain a data-driven efficiency also for the rare channels. To do this the TISTOS efficiency obtained on $B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$ candidates must be reweighted for the difference in the kinematics between the rare and resonant channels. This is done by determining the TISTOS efficiency as function of the maximum p_{T} of the particles that fired L0 (the leptons for `L0Electron` and `L0Muon`, the kaon and the pion for `L0Hadron`, and all final state particles for `L0Global`). Results are shown in Fig. 5.28 and used to re-weight the distribution of rare simulated candidates. The ratios $\varepsilon_{\ell\ell}^{tistos}/\varepsilon_{J/\psi}^{tistos}$ obtained using the data-driven method and simulation are compared and found to be fully compatible. This means that, even though the TISTOS correction has an effect on the absolute efficiency of each channel, this becomes negligible on their ratio. Therefore, no correction due to this effect is applied for the calculation of the $R_{K^{*0}}$ ratio.

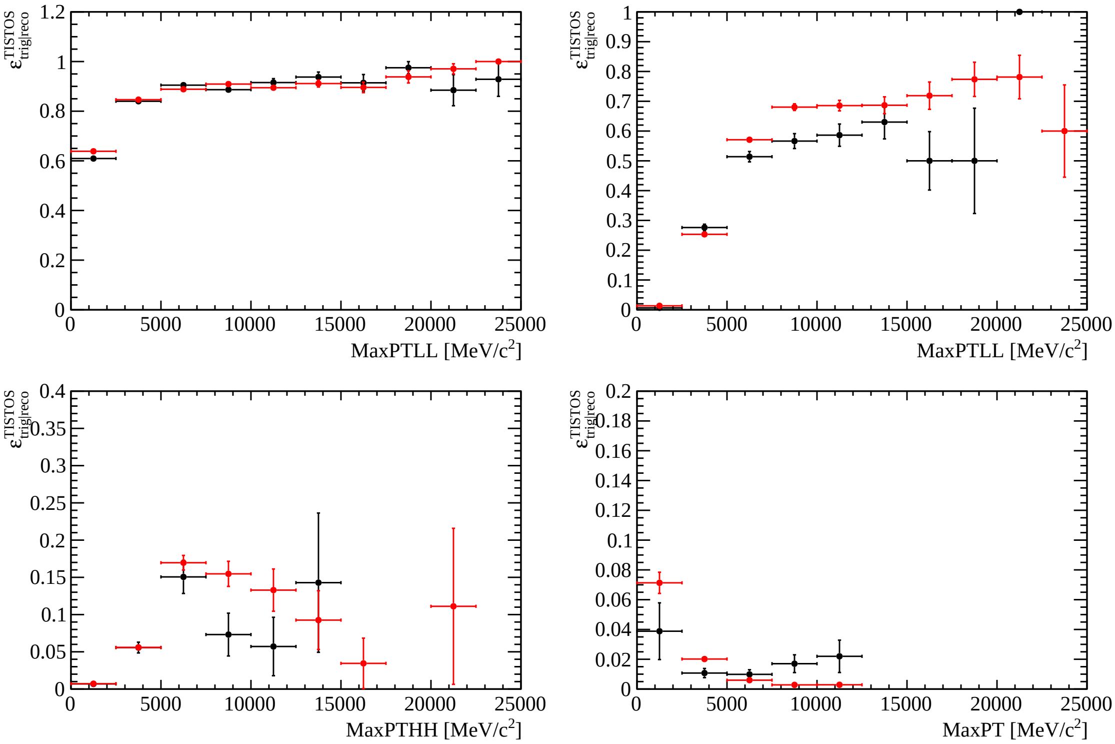

Figure 5.28: Trigger efficiency obtained applying the TISTOS method on $B^0 \to J/\psi\, K^*$ candidates as a function of the maximum p_{T} of the two muons (top left) of the two electrons for the L0E category (top right), the maximum p_{T} of p and π for L0H (bottom left) and the maximum p_{T} of all the final particles for L0I (bottom right).

5.9.5 Neural networks and BCM efficiencies

The the neural network and BCM efficiencies are evaluated from fully weighted simulated samples, and separately for each trigger category for the electron channels. In order to check for biases one can compare the efficiency obtained using $B^0 \to K^{*0}(J/\psi \to \ell^+\ell^-)$ events and rare $B^0 \to K^{*0}\ell^+\ell^-$ events in the same q^2 region selected for the resonant case. The ratio between the two should be close to unity with small deviations due the fact that the q^2 interval width is finite and the events are distributed differently inside the interval. This ratio is found to be 0.997 ± 0.004 for the $\mu\mu$ channels and on average 0.981 ± 0.005 for the ee channels. Values for the electron channels show a small deviation from one due to the very large q^2 interval used to select the resonant channel (6–11 GeV$^2/c^4$).

Table 5.13: Summary of the relative percent systematic uncertainties on $R_{K^{*0}}$.

Source	**low-q^2 (%)**	**central-q^2 (%)**	**high-q^2 (%)**
Signal shape	1.65	1.10	2.92
Bremsstrahlung categories	0.04	0.06	0.37
Swap	0.30	0.12	0.13
$\Lambda_b^0 \to pK\ell^+\ell^-$	0.25	0.28	0.77
Partially-reconstructed	0.11	4.13	0.10
Combinatorial	0.00	0.02	8.02
J/ψ leakage	0.06	0.01	0.10
$\psi(2S)$ leakage	0.03	0.01	2.00
`RooKeysPdf` ($\rho = 1.1$)	0.11	0.28	0.14
`RooKeysPdf` ($\rho = 1.3$)	0.10	0.24	0.49
Efficiency	0.65	0.74	0.83
TISTOS	2.47	2.30	2.80
Bin migration	0.69	1.43	1.19

5.10 Systematic uncertainties

This section describes the main sources of systematic uncertainties considered. Other sources, which would matter in measurements of absolute quantities, cancel in the ratio between the rare and resonant channels. A list of the systematic uncertainties that are considered and their effect on the $R_{K^{*0}}$ ratio is summarised in Tab. 5.13. The total uncertainty is evaluated by summing in quadrature the single components.

5.10.1 Choice of signal and background PDFs

There is a certain arbitrariety in the choice of PDFs to model signal and background contributions in the invariant mass fits, which could translate in a bias on the final result. The systematic uncertainty due to the parameterisation of line shapes is studied in the following ways.

For the signal PDF:

- *Shape*: in the electron channels the PDF is changed from a Crystal Ball and Gaussian to a Double Crystal Ball. Modifying the PDF has a negligible effect in the muon modes, while it affects the electron ones. Furthermore the data-simulation discrepancy parameters ($m^{'}$ and c) are constrained using the $B^0 \to K^{*0}(\gamma \to e^+e^-)$ sample instead of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$.
- *Bremsstrahlung categories*: gaussian constraints are applied to the relative fractions of the bremsstrahlung categories, instead of fixing them to the values observed on simulation. This yields a $\sim$ % systematic on $R_{K^{*0}}$ in the central- and high-q^2 region.

For the background PDFs:

- *Swaps*: a component that describes candidates where the particle identities are swapped is added both to the muon and electron resonant fits, and constrained to the number of candidates expected from simulation. This amounts to a $\sim$ % variation on $R_{K^{*0}}$ in the central- and high-q^2 region.
- $\Lambda_b^0 \to pK J/\psi (\to e^+e^-)$: the normalisation is left free to vary. This results in a $\sim$ % variation on $R_{K^{*0}}$ in the central- and high-q^2 region.
- *Partially-reconstructed*: the yield of the mis-reconstructed background to $B^0 \to K^{*0}e^+e^-$ is left free to vary in the fit. This only applies to the central-q^2 interval as this contribution is already free to vary in the high-q^2 range. This yields a $\sim$ % systematic on $R_{K^{*0}}$.
- *Combinatorial*: the PDF at high-q^2 is changed from an exponential (anti-MVA cut) to an anti-MVA cut (exponential) for the $\mu\mu$ (ee) mode. This amounts to a $\sim$ % variation on $R_{K^{*0}}$ in the central- and high-q^2 region.
- $\Lambda_b^0 \to pK\ell^+\ell^-$: this background is added to the fit to the rare channel and returns zero yield for both the muon and the electron samples. Therefore this yields no systematic uncertainty.

- *Leakage*: gaussian constraints are applied to the amounts of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ leakage in the central-q^2 region and to the $B^0 \to K^{*0}(\psi(2S) \to e^+e^-)$ leakage in the high-q^2 region, which are fixed in the default fit. This results in a $\sim$ % variation on $R_{K^{*0}}$ in the central- and high-q^2 region.

Furthermore in all case where a simulated sample is used and smoothed to obtain a PDF the kernel of the density estimation is varied by ± 0.1 from the value used in the nominal fit.

5.10.2 Efficiency determinations

The statistical uncertainty on the efficiency determinations is taken as the corresponding systematic uncertainty. The correlation among the electron trigger categories is taken into account (e.g. L0E and L0H are anti-correlated). A further source of systematic uncertainty associated to the trigger efficiency is estimated using the data-simulation differences observed in Sec. 5.9.4.2. Ratios of efficiencies for the rare to resonant decays are found to be compatible between the electron and muon modes, indicating that the effect on $R_{K^{*0}}$ is negligible, but the statistical precision on the determinations is taken as an extra systematic uncertainty.

5.10.3 Bin migration

The determination of the reconstruction efficiency is affected by the knowledge of the amount of bin migration as explained in Sec. 5.9.2. This amount depends on the shape of the q^2 distribution, which in turn depends on the simulated $B^0 \to K^{*0}e^+e^-$ decay model. In order to asses this systematic, simulated samples are generated using different models corresponding to different form factors [110, 111]. The q^2 distributions obtained using each model are compared with the ones obtained using the default one [112]. Figure 5.29 shows normalised ratios between these q^2 distributions and the default one, which are used to re-weight the simulation.

The amount of bin migration is recalculated using the simulation reweighted to reproduce each model; Table 5.14 lists the percent variations obtained. The largest difference between two values is taken as systematic uncertainty. This results in a $\sim 5\%$ uncertainty for the central-q^2 interval and $\sim 11\%$ for the high-q^2 one, which represent in both channel the biggest systematic uncertainty.

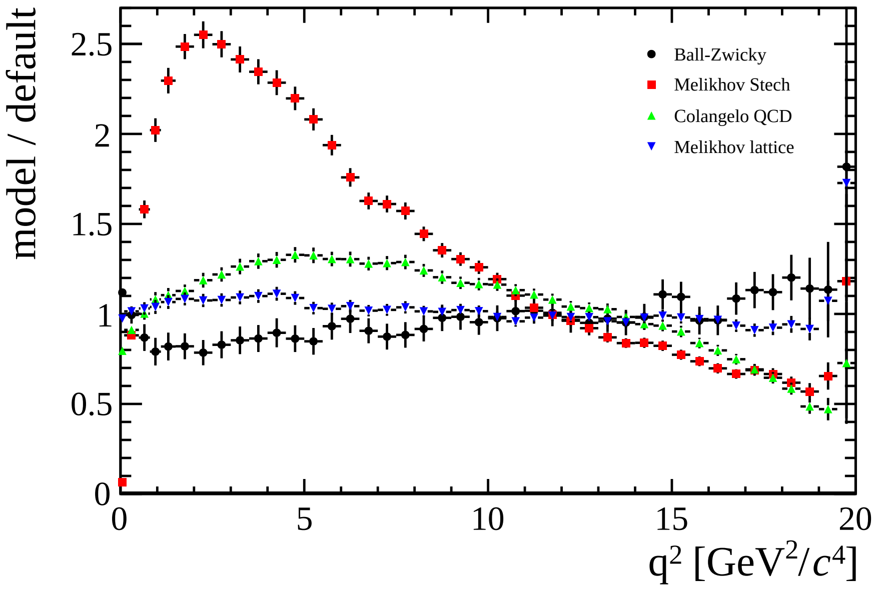

Figure 5.29: Ratios between the q^2 distributions obtained using different form factors models with respect to the default model.

Table 5.14: Percent variation on the bin migration amount obtained using different form factors models.

Model	low-q^2	central-q^2	central-q^2
Ball-Zwicky (6)	-0.3	1.0	0.2
Colangelo 2pt QCD (3)	0.4	0.4	0.8
Melikhov lattice (4)	0.1	-0.4	-0.4

5.11 Result extraction

This section presents the final results of this analysis together with the description of sanity checks performed to verify the stability of the methods used.

5.11.1 $R_{J/\psi}$ sanity check

In order to cross-check the analysis procedure, the ratio between the measured branching ratio of the electron and muon resonant channels is calculated:

$$r_{J/\psi} = \frac{\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-))}{\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to e^+e^-))} = \frac{\varepsilon_{J/\psi\,(\mu\mu)} \cdot N_{B^0 \to K^{*0} J/\psi\,(\to e^+e^-)}}{\varepsilon_{J/\psi\,(ee)} \cdot N_{B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)}}. \tag{5.18}$$

Unlike absolute branching fractions calculations, the determination of $R_{J/\psi}$ represents a better sanity test as it is not affected by uncertainties due to the knowledge of the amount of collected luminosity, $\mathcal{L}$, or the fragmentation fraction, f_d, the probability for a b quark to produce a B^0 meson. These quantities come with large uncertainties but they cancel in the $r_{J/\psi}$ ratio.

Measured values of the $R_{J/\psi}$ ratio are reported in Tab. 5.15, where the error shown is statistical only. For this purpose the trigger efficiencies are corrected using the factors obtained in Sec. 5.9.4.2. Note that systematic uncertainties, which cancel when doing the ratio between the rare and resonant channels with same leptonic final state, do not cancel in this case. A reasonable agreement with unity is found.

Table 5.15: Fully corrected measured values of the ratio $R_{J/\psi}$ in the three electron trigger categories.

Trigger	$r_{J/\psi}$
L0E	$1.073 \pm 0.009 \pm 0.021$
L0H	$1.016 \pm 0.024 \pm 0.068$
L0I	$1.150 \pm 0.017 \pm 0.148$

5.11.2 $\mathcal{B}(B^0 \to K^{*0}\gamma)$ sanity check

As a further check, the $B^0 \to K^{*0}\gamma$ branching fraction is determined using the ratio

$$r_\gamma = \frac{\mathcal{B}(B^0 \to K^{*0}\gamma)}{\mathcal{B}(B^0 \to K^{*0}(J/\psi \to e^+e^-))} = \frac{N_{\gamma(ee)}}{N_{J/\psi\,(ee)}} \cdot \frac{\varepsilon_{J/\psi\,(ee)}}{\varepsilon_{\gamma(ee)}} = 0.053 \pm 0.002\,.$$

Table 5.16: Measured values of r_{ee}, $r_{\mu\mu}$ and $R_{K^{*0}}$ ratios.

Ratio	low-q^2	central-q^2	high-q^2
R_{ee}			
$R_{\mu\mu}$			
$R_{K^{*0}}$		blind	blind

The measured value is

$$\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.18 \pm 0.26) \times 10^{-5}\,,$$

where the uncertainty is only statistical. A reasonable agreement with the PDG [1] value, $(4.33 \pm 0.15) \times 10^{-5}$, is found.

5.11.3 $R_{K^{*0}}$ result summary

The ratio $R_{K^{*0}}$ is extracted by dividing the r_{ee} and $r_{\mu\mu}$ parameters described in Sec. 5.8. These ratios are direct parameters of the fit but they can also be built from the yields in Tab. 5.8 and the efficiencies in Tab. ??. In summary the definition of the $R_{K^{*0}}$ ratio is the following:

$$R_{K^{*0}} = \frac{r_{ee}}{r_{\mu\mu}} = \frac{N_{ee}}{N_{J/\psi\,(ee)}} \cdot \frac{N_{J/\psi\,(\mu\mu)}}{N_{\mu\mu}} \cdot \frac{\varepsilon_{J/\psi\,(ee)}}{\varepsilon_{ee}} \cdot \frac{\varepsilon_{\mu\mu}}{\varepsilon_{J/\psi\,(\mu\mu)}}. \tag{5.19}$$

As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three electron categories its value is already a combination of the three samples. Results are shown in Tab. 5.16.

CHAPTER 6

Conclusions

In this work rare decays were analysed in order to look for hints of new physics using data recorded by the LHCb detector at centre-of-mass energies of 7 and 8 TeV and corresponding to a total integrated luminosity of 3.0 fb^{-1}. First, a measurement of the differential branching fraction of the rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay was performed together with the first measurement of angular observables for this decay. Evidence for the signal was found for the first time in the q^2 region below the square of the J/ψ mass in in the $0.1 < q^2 < 2.0$ GeV^2/c^4 interval, where an enhanced yield is expected due to the vicinity of the photon pole. Due to a larger data sample and a better control of systematic, the uncertainty of the measurement in the $15 < q^2 < 20$ GeV^2/c^4 interval are reduced by approximately a factor of three with respect to the previous LHCb measurements. The branching fraction measurements are compatible with SM predictions in the high-q^2 region, above the square of the J/ψ mass, and lie below the predictions in the low-q^2 region. In the angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays two forward-backward asymmetries, in the dimuon and

$p\pi$ systems, were measured. The measurements of the A_{FB}^{h} observable are in good agreement with the SM predictions while for the A_{FB}^{ℓ} observable measurements are consistently above the SM predictions.

Secondly lepton flavour universality was tested bla bla bla ...

REFERENCES

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

[2] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, *Phys.Rev.* **D20** (1979) 2619–2625.

[3] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961) 579–588.

[4] **LHCb** Collaboration, R. Aaij et al., *Observation of the resonant character of the $Z(4430)^-$ state*, *Phys. Rev. Lett.* **112** (2014), no. 22 222002, [arXiv:1404.1903].

[5] **LHCb** Collaboration, R. Aaij et al., *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda\_b^0 \to J/\psi K^- p$ decays*, arXiv:1507.03414.

[6] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, *Phys.Rev.* **105** (1957) 1413–1414.

[7] F. Strocchi, *Spontaneous Symmetry Breaking in Local Gauge Quantum Field Theory: The Higgs Mechanism*, *Commun.Math.Phys.* **56** (1977) 57.

[8] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, et al., *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, *Phys.Rev.* **D91** (2015), no. 7 073007, [arXiv:1501.05013].

[9] F. Zwicky, *Spectral displacement of extra galactic nebulae*, *Helv.Phys.Acta* **6** (1933) 110–127.

[10] M. Gavela and Hernandez, *Standard model CP violation and baryon asymmetry*, *Mod.Phys.Lett.* **A9** (1994) 795–810, [hep-ph/9312215].

[11] M. Maltoni, *Status of three-neutrino oscillations*, *PoS* **EPS-HEP2011** (2011) 090.

[12] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, *Astrophys. J.* **496** (1998) 505–526.

[13] **Super-Kamiokande** Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [hep-ex/9807003].

[14] **KamLAND** Collaboration, K. Eguchi et al., *First results from KamLAND: Evidence for reactor anti-neutrino disappearance*, *Phys. Rev. Lett.* **90** (2003) 021802, [hep-ex/0212021].

[15] J. L. Feng, *Naturalness and the Status of Supersymmetry*, *Ann.Rev.Nucl.Part.Sci.* **63** (2013) 351–382, [arXiv:1302.6587].

[16] **ATLAS** Collaboration, 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) 1–29, [arXiv:1207.7214].

[17] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, *Sov. Phys. JETP* **26** (1968) 984–988. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[18] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870–880.

[19] P. Fayet and S. Ferrara, *Supersymmetry*, *Phys.Rept.* **32** (1977) 249–334.

[20] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys.Rev.Lett.* **83** (1999) 3370–3373, [hep-ph/9905221].

[21] J. R. Ellis, G. Gelmini, C. Jarlskog, G. G. Ross, and J. W. F. Valle, *Phenomenology of Supersymmetry with Broken R-Parity*, *Phys. Lett.* **B150** (1985) 142.

[22] G. Isidori and D. M. Straub, *Minimal Flavour Violation and Beyond*, *Eur.Phys.J.* **C72** (2012) 2103, [arXiv:1202.0464].

[23] A. J. Buras, *Minimal flavor violation*, *Acta Phys.Polon.* **B34** (2003) 5615–5668, [hep-ph/0310208].

[24] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*, *Ann.Rev.Nucl.Part.Sci.* **65** (2015) 8007, [arXiv:1501.03309].

[25] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, *Can we reach the Zeptouniverse with rare K and $B_{s,d}$ decays?*, *JHEP* **1411** (2014) 121, [arXiv:1408.0728].

[26] G. Hiller and M. Schmaltz, *R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities*, *Phys.Rev.* **D90** (2014) 054014, [arXiv:1408.1627].

[27] K. G. Chetyrkin, M. Misiak, and M. Munz, *Weak radiative B meson decay beyond leading logarithms*, *Phys.Lett.* **B400** (1997) 206–219, [hep-ph/9612313].

[28] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev.Mod.Phys.* **68** (1996) 1125–1144, [hep-ph/9512380].

[29] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, hep-ph/9806471.

[30] M. Della Morte, J. Heitger, H. Simma, and R. Sommer, *Non-perturbative Heavy Quark Effective Theory: An application to semi-leptonic B-decays*, *Nucl.Part.Phys.Proc.* **261-262** (2015) 368–377, [arXiv:1501.03328].

[31] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, *Phys.Rev.* **D63** (2001) 114020, [hep-ph/0011336].

[32] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \to K^{(*)}\ell^+\ell^-$ and $B \to K^*\gamma$*, *JHEP* **1009** (2010) 089, [arXiv:1006.4945].

[33] **LHCb** Collaboration, R. Aaij et al., *Observation of a resonance in $B^+ \to K^+\mu^+\mu^-$ decays at low recoil*, *Phys. Rev. Lett.* **111** (2013) 112003, [arXiv:1307.7595].

[34] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, et al., *$B_{s,d} \to l + l-$ in the Standard Model with Reduced Theoretical Uncertainty*, *Phys.Rev.Lett.* **112** (2014) 101801, [arXiv:1311.0903].

[35] **CMS, LHCb** Collaboration, V. Khachatryan et al., *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, *Nature* **522** (2015) 68–72, [arXiv:1411.4413].

[36] **LHCb** Collaboration, R. Aaij et al., *Differential branching fractions and isospin asymmetry of $B \to K^{(*)}\mu^+\mu^+$ decays*, *JHEP* **06** (2014) 133, [arXiv:1403.8044].

[37] **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].

[38] **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].

[39] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, *JHEP* **1506** (2015) 115, [arXiv:1503.07138].

[40] **LHCb** Collaboration, R. Aaij et al., *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, *Phys. Rev. Lett.* **111** (2013) 191801, [arXiv:1308.1707].

[41] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \to K^*\mu^+\mu^-$ Anomaly*, *Phys.Rev.* **D88** (2013), no. 7 074002, [arXiv:1307.5683].

[42] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of charged and neutral $B \to K\mu^+\mu^-$ decays*, *JHEP* **05** (2014) 082, [arXiv:1403.8045].

[43] **LHCb** Collaboration, R. Aaij et al., *Measurement of CP asymmetries in the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$*, *JHEP* **1409** (2014) 177, [arXiv:1408.0978].

[44] **LHCb** Collaboration, R. Aaij et al., *Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction at low dilepton mass*, *JHEP* **05** (2013) 159, [arXiv:1304.3035].

[45] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay in the low-q^2 region*, *JHEP* **04** (Jan, 2015) 064. 18 p.

[46] **MEGA** Collaboration, M. Ahmed et al., *Search for the lepton family number nonconserving decay $mu^+ \to e^+\gamma$*, *Phys.Rev.* **D65** (2002) 112002, [hep-ex/0111030].

[47] **SINDRUM** Collaboration, U. Bellgardt et al., *Search for the Decay $mu^+ \to e^+e^+e^-$*, *Nucl.Phys.* **B299** (1988) 1.

[48] **LHCb** Collaboration, R. Aaij et al., *Search for the lepton-flavour-violating decays $B_s^0 \to e^\pm\mu^\mp$ and $B^0 \to e^\pm\mu^\mp$*, *Phys. Rev. Lett.* **111** (2013) 141801, [arXiv:1307.4889].

[49] **LHCb** Collaboration, R. Aaij et al., *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, *Phys. Lett.* **B724** (2013) 36, [arXiv:1304.4518].

[50] W. J. Marciano, T. Mori, and J. M. Roney, *Charged Lepton Flavor Violation Experiments*, *Ann.Rev.Nucl.Part.Sci.* **58** (2008) 315–341.

[51] L. Evans, *The LHC machine*, *PoS* **EPS-HEP2009** (2009) 004.

[52] **LHCb** Collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.

[53] **LHCb** Collaboration, R. Aaij et al., *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, *Phys.Lett.* **B694** (2010) 209–216, [`arXiv:1009.2731`].

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

[55] A. A. Alves Jr. et al., *Performance of the LHCb muon system*, *JINST* **8** (2013) P02022, [`arXiv:1211.1346`].

[56] **LHCb** Collaboration, R. e. a. Aaij, *LHCb technical design report: Reoptimized detector design and performance*, `CERN-LHCC-2003-030`.

[57] **LHCb** Collaboration, R. e. a. Aaij, *LHCb Detector Performance*, *Int. J. Mod. Phys. A* **30** (Dec, 2014) 1530022. 82 p.

[58] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl.Instrum.Meth.* **A555** (2005) 356–369, [`physics/0402083`].

[59] R. Aaij et al., *The LHCb trigger and its performance in 2011*, *JINST* **8** (2013) P04022, [`arXiv:1211.3055`].

[60] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [`hep-ph/0603175`].

[61] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [`arXiv:0710.3820`].

[62] I. Belyaev et al., *Handling of the generation of primary events in* GAUSS*, the LHCb simulation framework*, *Nuclear Science Symposium Conference Record (NSS/MIC)* **IEEE** (2010) 1155.

[63] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[64] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur.Phys.J.* **C45** (2006) 97–107, [`hep-ph/0506026`].

[65] **Geant4 collaboration** Collaboration, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, et al., *Geant4 developments and applications*, *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

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

[67] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, *Conf.Proc.* **C000917** (2000) 11–42.

[68] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, *Nucl.Instrum.Meth.* **A559** (2006) 190–194.

[69] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, physics/0402093.

[70] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, *Nucl.Instrum.Meth.* **A552** (2005) 566–575, [physics/0503191].

[71] H. W. Bertini, *Low-Energy Intranuclear Cascade Calculation*, *Phys. Rev.* **131** (1963) 1801–1821.

[72] B. Andersson, G. Gustafson, and H. Pi, *The FRITIOF model for very high-energy hadronic collisions*, *Z. Phys.* **C57** (1993) 485–494.

[73] **COMPASS** Collaboration, P. Abbon et al., *The COMPASS experiment at CERN*, *Nucl. Instrum. Meth.* **A577** (2007) 455–518, [hep-ex/0703049].

[74] G. Hiller, M. Knecht, F. Legger, and T. Schietinger, *Photon polarization from helicity suppression in radiative decays of polarized Lambda(b) to spin-3/2 baryons*, *Phys.Lett.* **B649** (2007) 152–158, [hep-ph/0702191].

[75] T. Mannel and S. Recksiegel, *Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \to \Lambda\,\gamma$*, *J.Phys.* **G24** (1998) 979–990, [hep-ph/9701399].

[76] M. J. Aslam, Y.-M. Wang, and C.-D. Lu, *Exclusive semileptonic decays of $\Lambda_b^0 \to \Lambda\, l^+l^-$ in supersymmetric theories*, *Phys.Rev.* **D78** (2008) 114032, [arXiv:0808.2113].

[77] Y.-m. Wang, Y. Li, and C.-D. Lu, *Rare Decays of $\Lambda_b^0 \to \Lambda\,\gamma$ and $\Lambda_b^0 \to \Lambda\, l^+l^-$ in the Light-cone Sum Rules*, *Eur.Phys.J.* **C59** (2009) 861–882, [arXiv:0804.0648].

[78] C.-S. Huang and H.-G. Yan, *Exclusive rare decays of heavy baryons to light baryons: $\Lambda_b^0 \to \Lambda\,\gamma$ and $\Lambda_b^0 \to \Lambda\, l^+l^-$* , *Phys.Rev.* **D59** (1999) 114022, [hep-ph/9811303].

[79] C.-H. Chen and C. Geng, *Rare $\Lambda_b^0 \to \Lambda\, l^+l^-$ decays with polarized lambda*, *Phys.Rev.* **D63** (2001) 114024, [hep-ph/0101171].

[80] C.-H. Chen and C. Geng, *Baryonic rare decays of $\Lambda_b^0 \to \Lambda\, l^+l^-$*, *Phys.Rev.* **D64** (2001) 074001, [hep-ph/0106193].

[81] C.-H. Chen and C. Geng, *Lepton asymmetries in heavy baryon decays of $\Lambda_b^0 \to \Lambda\, l^+l^-$*, *Phys.Lett.* **B516** (2001) 327–336, [hep-ph/0101201].

[82] F. Zolfagharpour and V. Bashiry, *Double Lepton Polarization in $\Lambda_b^0 \to \Lambda l^+l^-$ Decay in the Standard Model with Fourth Generations Scenario*, *Nucl.Phys.* **B796** (2008) 294–319, [arXiv:0707.4337].

[83] L. Mott and W. Roberts, *Rare dileptonic decays of Λ_b^0 in a quark model*, *Int.J.Mod.Phys.* **A27** (2012) 1250016, [arXiv:1108.6129].

[84] T. Aliev, K. Azizi, and M. Savci, *Analysis of the $\Lambda_b^0 \to \Lambda l^+ l^-$ decay in QCD*, *Phys.Rev.* **D81** (2010) 056006, [arXiv:1001.0227].

[85] R. Mohanta and A. Giri, *Fourth generation effect on Λ_b decays*, *Phys.Rev.* **D82** (2010) 094022, [arXiv:1010.1152].

[86] S. Sahoo, C. Das, and L. Maharana, *Effect of both Z and Z'-mediated flavor-changing neutral currents on the baryonic rare decay $\Lambda_b^0 \to \Lambda l^+ l^-$*, *Int.J.Mod.Phys.* **A24** (2009) 6223–6235, [arXiv:1112.4563].

[87] **CDF Collaboration** Collaboration, T. Aaltonen et al., *Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \to \Lambda\mu^+\mu^-$*, *Phys.Rev.Lett.* **107** (2011) 201802, [arXiv:1107.3753].

[88] **CDF** Collaboration, S. Behari, *CDF results on $b \to s\mu\mu$ decays*, arXiv:1301.2244.

[89] **LHCb** Collaboration, R. Aaij et al., *Measurement of the differential branching fraction of the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$*, *Phys. Lett.* **B725** (2013) 25, [arXiv:1306.2577].

[90] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, *Rare baryon decays $\Lambda_b \to \Lambda l^+ l^- (l = e, \mu, \tau)$ and $\Lambda_b \to \Lambda\gamma$: differential and total rates, lepton- and hadron-side forward-backward asymmetries*, *Phys.Rev.* **D87** (2013) 074031, [arXiv:1301.3737].

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

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

[93] 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.

[94] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, *$\Lambda_b^0 \to \Lambda\mu^+\mu^-$ form factors and differential branching fraction from lattice QCD*, *Phys. Rev.* **D87** (2013), no. 7 074502, [arXiv:1212.4827].

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

[96] T. Blake, S. Coquereau, M. Chrzaszcz, S. Cunliffe, C. Parkinson, K. Petridis, and M. Tresch, *The $B_0 \to K_0^* \mu\mu$ selection using 3fb^{-1} of LHCb data*, Tech. Rep. LHCb-INT-2013-058. CERN-LHCb-INT-2013-058, CERN, Geneva, Nov, 2013.

[97] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343–367.

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

[99] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys.Rev.* **D57** (1998) 3873–3889, [physics/9711021].

[100] T. M. Karbach, *Feldman-Cousins Confidence Levels - Toy MC Method*, arXiv:1109.0714.

[101] S. Meinel, *Flavor physics with Λ_b baryons*, *PoS* **LATTICE2013** (2014) 024, [arXiv:1401.2685].

[102] G. Hiller and F. Kruger, *More model independent analysis of $b \to s$ processes*, *Phys.Rev.* **D69** (2004) 074020, [hep-ph/0310219].

[103] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \to s\ell\ell$*, *JHEP* **1502** (2015) 055, [arXiv:1411.4773].

[104] **BaBar Collaboration** Collaboration, J. 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].

[105] **BELLE Collaboration** Collaboration, J.-T. Wei et al., *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)}l^+l^-$*, *Phys.Rev.Lett.* **103** (2009) 171801, [arXiv:0904.0770].

[106] **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].

[107] "Lhcb loki twiki." https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters. Accessed: 2015-09-30.

[108] "Probnn presentation at ppts meeting." https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf. Accessed: 2015-09-30.

[109] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf* **C0303241** (2003) MOLT007, [physics/0306116].

[110] P. Ball and R. Zwicky, *New results on $B \to \pi, K, \eta$ decay form factors from light-cone sum rules*, *Phys. Rev.* **D71** (2005) 014015, [hep-ph/0406232].

[111] D. Melikhov and B. Stech, *Weak form-factors for heavy meson decays: An Update*, *Phys. Rev.* **D62** (2000) 014006, [hep-ph/0001113].

[112] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *A Comparative study of the decays $B \to (K, K^{*)}\ell^+\ell^-$ in standard model and supersymmetric theories*, *Phys. Rev.* **D61** (2000) 074024, [hep-ph/9910221].

[113] J. Hrivnac, R. Lednicky, and M. Smizanska, *Feasibility of beauty baryon polarization measurement in $\Lambda^0 J/\psi$ decay channel by ATLAS LHC*, *J.Phys.G* **G21** (1995) 629–638, [hep-ph/9405231].

[114] W. Detmold and S. Meinel, *$\Lambda_b \to \Lambda\ell^+\ell^-$ form factors, differential branching fraction, and angular observables from lattice QCD with relativistic b quarks*, arXiv:1602.01399.

APPENDIX A

Decay models

A.1 $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ distribution

The q^2 and angular dependancies of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays are modelled based on Ref. [90], where the angular distribution for unpolarised Λ_b^0 production is defined as

$$\begin{aligned} W(\theta\ell, \theta_B, \chi) \quad \propto & \sum_{\lambda_1,\lambda_2,\lambda_j,\lambda'_j,J,J',m,m',\lambda_\Lambda,\lambda'_\Lambda,\lambda_p} h^m_{\lambda_1\lambda_2}(J) h^{m'}_{\lambda_1\lambda_2}(J') e^{i(\lambda_j-\lambda'_j)\chi} \\ & \times \ \delta_{\lambda_j-\lambda_\Lambda,\lambda'_j-\lambda'_\Lambda} \delta_{JJ'} d^J_{\lambda_j,\lambda_1-\lambda_2}(\theta\ell) d^{J'}_{\lambda'_j,\lambda_1-\lambda_2}(\theta\ell) H^m_{\lambda_\Lambda\lambda_j}(J) H^{m'\dagger}_{\lambda'_\Lambda\lambda'_j}(J') \\ & \times \ d^{1/2}_{\lambda_\Lambda\lambda_p}(\theta_B) d^{1/2}_{\lambda'_\Lambda\lambda_p}(\theta_B) h^B_{\lambda_p 0} h^{B\,\dagger}_{\lambda_p 0}\,. \end{aligned} \tag{A.1}$$

In this formula θ_ℓ and θ_B correspond to the lepton and proton helicity angles, χ is angle between dimuon and Λ decay planes (for unpolarised production we are sensitive only to difference in azimuthal angles), $d^J_{i,j}$ are Wigner d-functions and h, h^B and H are helicity amplitudes for virtual dimuon, Λ and Λ_b^0 decays. The sum runs over all possible helicities with the dimuon being allowed in spin 0 and 1 states (J and J'). The m and m' indices run over the vector and axial-vector current contributions.

The production polarisation is introduced by removing $e^{i(\lambda_j-\lambda'_j)\chi}$ from the expression, swapping small Wigner d-functions $d^J_{i,j}$ to the corresponding capital ones $D^J_{i,j}$ which are related as

$$D^J_{i,j}(\theta,\phi) = d^J_{i,j}(\theta) e^{i\phi(i-j)} \tag{A.2}$$

and substitute spin density matrix for $\delta_{\lambda_j-\lambda_\Lambda,\lambda'_j-\lambda'_\Lambda}\delta_{JJ'}$. The spin density matrix itself is given by

$$\rho_{\lambda_j-\lambda_\Lambda,\lambda'_j-\lambda'_\Lambda} = \frac{1}{2}\begin{pmatrix} 1+P_b\cos\theta & P_b\sin\theta \\ P_b\sin\theta & 1-P_b\cos\theta \end{pmatrix}. \tag{A.3}$$

Those changes lead to the formula

$$\begin{aligned} W(\theta\ell,\theta_B,\chi) \;\propto & \sum_{\lambda_1,\lambda_2,\lambda_j,\lambda'_j,J,J',m,m',\lambda_\Lambda,\lambda'_\Lambda,\lambda_p} h^m_{\lambda_1\lambda_2}(J)h^{m'}_{\lambda_1\lambda_2}(J') \\ \times\; & \rho_{\lambda_j-\lambda_\Lambda,\lambda'_j-\lambda'_\Lambda} D^J_{\lambda_j,\lambda_1-\lambda_2}(\theta\ell,\phi_L)D^{J'}_{\lambda'_j,\lambda_1-\lambda_2}(\theta\ell,\phi_L)H^m_{\lambda_\Lambda\lambda_j}(J)H^{m'\dagger}_{\lambda'_\Lambda\lambda'_j}(J') \\ \times\; & D^{1/2}_{\lambda_\Lambda\lambda_p}(\theta_B,\phi_B)D^{1/2}_{\lambda'_\Lambda\lambda_p}(\theta_B,\phi_B)h^B_{\lambda_p0}h^{B\,\dagger}_{\lambda_p0}. \end{aligned} \tag{A.4}$$

The lepton amplitudes come directly from Ref. [90], eq. 3. The Λ decay amplitudes are related to the Λ decay asymmetry parameter as

$$\alpha_\Lambda = \frac{|h^B_{\frac{1}{2}0}|^2 - |h^B_{-\frac{1}{2}0}|^2}{|h^B_{\frac{1}{2}0}|^2 + |h^B_{-\frac{1}{2}0}|^2}. \tag{A.5}$$

Finally, the Λ_b^0 decay amplitudes receive contributions from vector and axial-vector currents and can be written as

$$H^m_{\lambda_2,\lambda_j} = H^{Vm}_{\lambda_2,\lambda_j} - H^{Am}_{\lambda_2,\lambda_j}. \tag{A.6}$$

Finally, the remaining amplitudes are expressed in terms of form factors (Ref. [90], eq. C6) as

$$\begin{aligned} H^{Vm}_{\frac{1}{2}t} &= \sqrt{\frac{Q_+}{q^2}}\left(M_-\,F^{Vm}_1 + \frac{q^2}{M_1}\,F^{Vm}_3\right), \\ H^{Vm}_{\frac{1}{2}1} &= \sqrt{2Q_-}\left(F^{Vm}_1 + \frac{M_+}{M_1}\,F^{Vm}_2\right), \\ H^{Vm}_{\frac{1}{2}0} &= \sqrt{\frac{Q_-}{q^2}}\left(M_+\,F^{Vm}_1 + \frac{q^2}{M_1}\,F^{Vm}_2\right), \\ H^{Am}_{\frac{1}{2}t} &= \sqrt{\frac{Q_-}{q^2}}\left(M_+\,F^{Am}_1 - \frac{q^2}{M_1}\,F^{Am}_3\right), \\ H^{Am}_{\frac{1}{2}1} &= \sqrt{2Q_+}\left(F^{Am}_1 - \frac{M_-}{M_1}\,F^{Am}_2\right), \\ H^{Am}_{\frac{1}{2}0} &= \sqrt{\frac{Q_+}{q^2}}\left(M_-\,F^{Am}_1 - \frac{q^2}{M_1}\,F^{Am}_2\right), \end{aligned} \tag{A.7}$$

where $M_\pm = M_1 \pm M_2$, $Q_\pm = M_\pm^2 - q^2$. The form factors F are expressed in terms of dimensioneless quantities in eqs. C8 and C9 in Ref. [90]. In our actual implementation form factors calculated in the covariant quark model [90] are used and for the numerical values of the Wilson coefficients Ref. [90] is used.

To assess effect of different form factors on efficiency calculations, an alternative set of form factors is implemented, based on the LQCD calculation from Ref. [94]. The form factors relations are found by comparing eqs. 66 and 68 in Ref. [90] to eq. 51 in Ref. [94]. Denoting LQCD form factors by F_i^L and dimensionless covariant quark model ones by f_i^{XX} we have

$$\begin{aligned}
f_1^V &= c_\gamma(F_1^L + F_2^L), \\
f_2^V &= -2c_\gamma F_2^L, \\
f_3^V &= c_v(F_1^L + F_2^L), \\
f_1^A &= c_\gamma(F_1^L - F_2^L), \\
f_2^A &= -2c_\gamma F_2^L, \\
f_3^A &= -c_v(F_1^L - F_2^L), \\
f_1^{TV} &= c_\sigma F_2^L, \\
f_2^{TV} &= -c_\sigma F_1^L, \\
f_1^{TA} &= c_\sigma F_2^L, \\
f_2^{TA} &= -c_\sigma F_1^L,
\end{aligned}$$

where

$$\begin{aligned}
c_\gamma &= 1 - \frac{\alpha_s(\mu^2)}{\pi}\left[\frac{4}{3} + \ln\left(\frac{\mu}{m_b}\right)\right], \\
c_v &= \frac{2}{3}\frac{\alpha_s(\mu^2)}{\pi}, \\
c_\sigma &= 1 - \frac{\alpha_s(\mu^2)}{\pi}\left[\frac{4}{3} + \frac{5}{3}\ln\left(\frac{\mu}{m_b}\right)\right].
\end{aligned} \tag{A.8}$$

In the calculations $\mu = m_b$ is used. For the strong coupling constant, we start from the world average value at the Z mass, $\alpha_s(m_Z^2) = 0.1185 \pm 0.0006$ [1], and we translate it to the scale m_b^2 by

$$\alpha_s(\mu^2) = \frac{\alpha_s(m_Z^2)}{1 + \frac{\alpha_s(m_Z^2)}{12\pi}(33 - 2n_f)\ln\left(\frac{\mu^2}{m_Z^2}\right)}, \tag{A.9}$$

where $n_f = 5$. The LQCD form factors F_1^L and F_2^L can be then taken directly from Ref. [94] and plugged into the code implementing the calculation from Ref. [90].

A.2 Bi-dimensional distribution parameters

Expectations values for parameters in the bi-dimensional angular distribution for the $\Lambda_b^0 \to \Lambda\,\mu^+\mu^-$ decay calculated using form factors and numerical inputs from Ref. [90].

q^2 $[GeV^2/c^2]$	A_{FB}^{ℓ}	P_z^{Λ}	f_L	O_P	O_{Lp}	O_{UVA}
0.1 – 2.0	0.082	-0.9998	0.537	-0.463	-0.537	0.055
2.0 – 4.0	-0.032	-0.9996	0.858	-0.142	-0.857	-0.021
4.0 – 6.0	-0.153	-0.9991	0.752	-0.247	-0.752	-0.102
V.0 – VA.5	-0.348	-0.9834	0.508	-0.478	-0.505	-0.239
15.0 – 16.0	-0.384	-0.9374	0.428	-0.524	-0.413	-0.280
16.0 – 18.0	-0.377	-0.8807	0.399	-0.513	-0.368	-0.294
18.0 – 20.0	-0.297	-0.6640	0.361	-0.404	-0.260	-0.314
1.0 – 6.0	-0.040	-0.9994	0.830	-0.170	-0.830	-0.027
15.0 – 20.0	-0.339	-0.7830	0.385	-0.461	-0.3A	-0.302

Table A.1: Prediction for angular observables entering two-dimensional angular distributions. Prediction is based on covariant quark model form factors from Ref. [90].

A.3 $\Lambda_b^0 \to J/\psi\,\Lambda$ distribution

The angular distribution of the $\Lambda_b^0 \to J/\psi\;\Lambda$ decay is modelled using Ref. [113]. The differential rate is written as

$$w(\Omega, \Omega_1, \Omega_2) = \frac{1}{(4\pi)}^{3} \sum_{i=0}^{i=19} f_{1i} f_{2i}(P_b, \alpha_\Lambda) F_i(\theta, \theta_1, \theta_2, \phi_1, \phi_2), \tag{A.10}$$

where f_{1i}, f_{2i} and F_i are listed in Tab. **??**. The expression uses four observables (angles) and depends on four complex amplitudes a_+, a_-, b_+, b_- and two real valued parameters for the production polarisation, P_b, and the Λ decay asymmetry, α_Λ. The angle θ is the angle of the Λ momentum in Λ_b^0 rest frame with respect to the vector $\vec{n} = \frac{\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}}{|\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}|}$, where $\vec{p_{inc}}$ and $\vec{p}_{\Lambda_b^0}$ are the momenta of incident proton and Λ_b^0 in the center of mass system. The angles θ_1 and ϕ_1 are polar and azimuthal angle of the proton coming from the Λ decay in the Λ rest frame with axis defined as $z_1 \uparrow\uparrow \vec{p}_\Lambda$, $y_1 \uparrow\uparrow \vec{n} \times \vec{p}_\Lambda$. Finally, the angles θ_2 and ϕ_2 are the angles of the momenta of the muons in J/ψ rest frame with axes defined as $z_2 \uparrow\uparrow \vec{p}_{J/\psi}$, $y_2 \uparrow\uparrow \vec{n} \times \vec{p}_{J/\psi}$.

The distribution depends on the Λ decay asymmetry parameter, α_Λ, the production polarisation P_b and four complex amplitudes. The α_Λ is measured to be 0.642 ± 0.013 for Λ. The production polarisation P_b and magnitudes of a_+, a_-, b_+ and b_- are measured in Ref. [98]. Phases are not measured therefore, as default all phases are set to zero and then they are randomly varied to calculate the systematic uncertainty.

APPENDIX B

Data-simulation comparison

This appendix reports a comparison between distributions in data and simulated $\Lambda_b^0 \to J/\psi\ \Lambda$ evets. In the plots what is labeled as "Data" is real data in a 20 MeV interval around the Λ_b^0 mass, where a sideband subtraction technique to remove background. "Side" is real data for masses above 6 GeV containing mostly combinatorial background. These can be compared to the previous sample to see which variables differ the most. "MC" corresponds to Pythia8 $\Lambda_b^0 \to J/\psi\ \Lambda$ simulated events. Finally, the label "MC fully W" refers to the same simulated sample but weighted for the Λ_b^0 and Λ kinematics (Sec. 3.3.2) and the decay model (Sec. 3.3.1). Distributions are shown separately for long and downstream events.

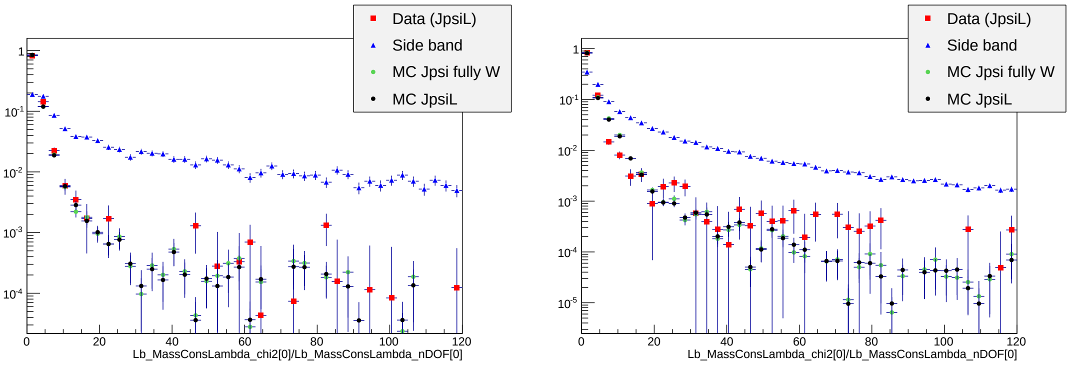

Figure B.1: Distributions of χ^2/NdF of the kinematic fit in data and simulation for LL (left) and DD (right) events.

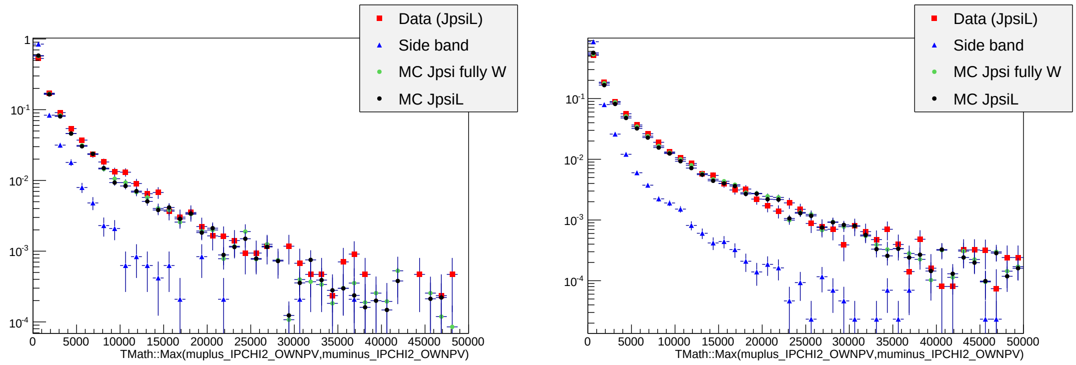

Figure B.2: Distributions of maximum muon $IP\chi^2$ variable in data and simulation for LL (left) and DD (right) events.

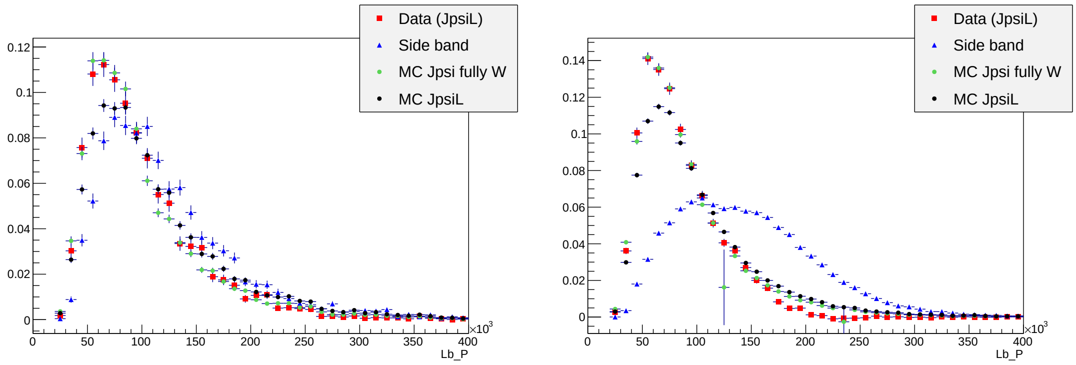

Figure B.3: Distributions of Λ_b^0 momentum variable in data and simulation for LL (left) and DD (right) events.

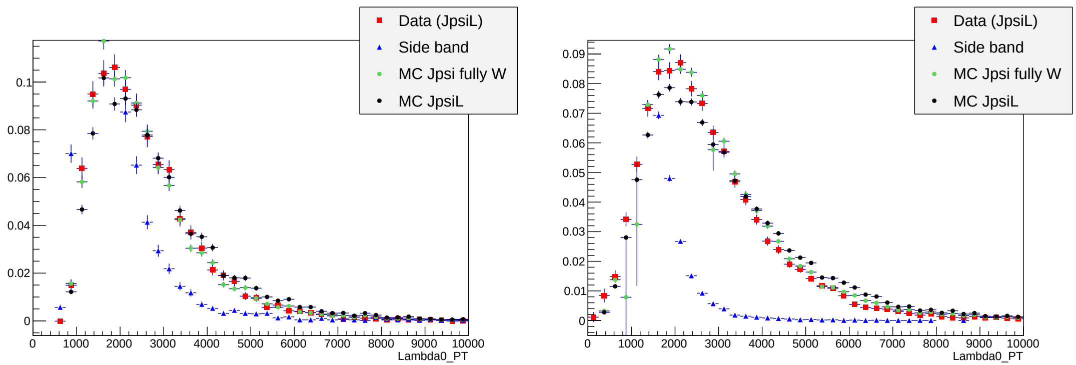

Figure B.4: Distributions of Λ transverse momentum variable in MC, data signal and data background for LL (left) and DD (right) events.

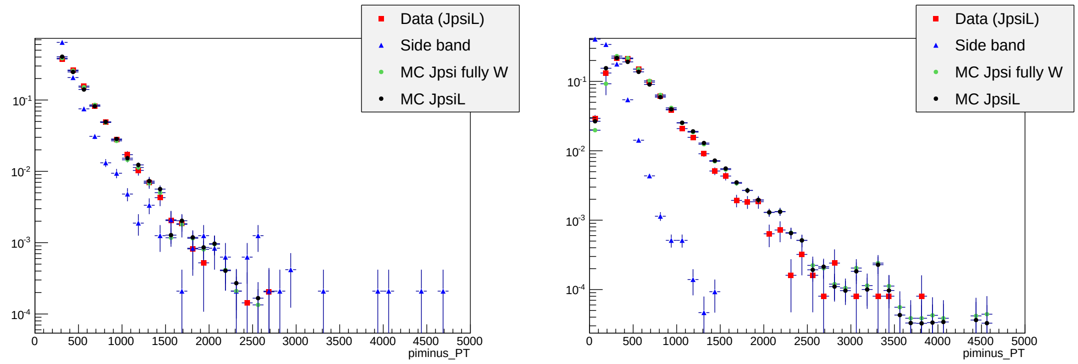

Figure B.5: Distributions of pion transverse momentum variable in data and simulation for LL (left) and DD (right) events.

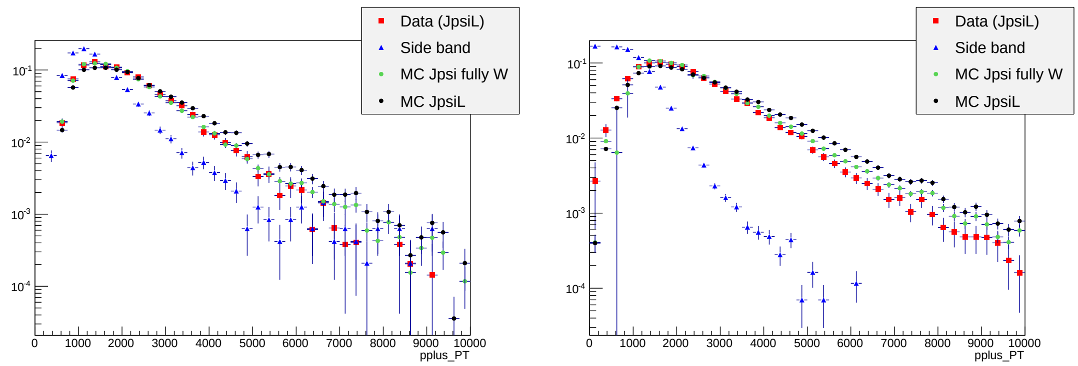

Figure B.6: Distributions of proton transverse momentum variable in data and simulation for LL (left) and DD (right) events.

APPENDIX C

Systematic uncertainties on the efficiency calculation for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ branching fraction analysis.

This appendix reports systematic uncertainties on absolute and relative efficiencies for the $\Lambda_b^0 \to \Lambda\, \mu^+\mu^-$ branching fraction analysis.

q^2 [GeV^2/c^4]	Lifetime	Decay Model	Polarisation
0.1-2.0	0.003%	0.059%	0.145%
2.0-4.0	0.007%	0.156%	0.145%
4.0-6.0	0.002%	0.156%	0.144%
6.0-8.0	0.003%	0.080%	0.144%
11.0-12.5	0.012%	0.101%	0.144%
15.0-16.0	0.007%	0.050%	0.144%
16.0-18.0	0.002%	0.059%	0.145%
18.0-20.0	0.009%	0.016%	0.145%
1.1-6.0	0.005%	0.651%	0.144%
15.0-20.0	0.007%	0.088%	0.144%

Table C.1: Absolute values of systematic uncertainties on relative geometric efficiency.

q^2 [GeV^2/c^4]	Lifetime	Decay Model	Polarisation
0.1-2.0	0.007%	0.004%	0.008%
2.0-4.0	0.006%	0.001%	0.009%
4.0-6.0	0.009%	0.003%	0.008%
6.0-8.0	0.008%	0.005%	0.008%
11.0-12.5	0.010%	0.005%	0.009%
15.0-16.0	0.004%	0.006%	0.008%
16.0-18.0	0.003%	0.010%	0.010%
18.0-20.0	0.004%	0.011%	0.008%
1.1-6.0	0.009%	0.043%	0.010%
15.0-20.0	0.005%	0.072%	0.009%

Table C.2: Absolute values of systematic uncertainties on relative detection efficiency.

	Downstream			Long		
q^2 [GeV^2/c^4]	Lifetime	Model	Polarisation	Lifetime	Model	Polarisation
0.1-2.0	0.350%	0.234%	0.463%	0.066%	0.264%	1.081%
2.0-4.0	0.170%	0.640%	0.488%	0.005%	0.953%	1.088%
4.0-6.0	0.073%	0.514%	0.465%	0.052%	1.607%	1.087%
6.0-8.0	0.054%	0.298%	0.458%	0.011%	1.517%	1.075%
11.0-12.5	0.043%	0.030%	0.469%	0.025%	0.187%	1.080%
15.0-16.0	0.078%	0.499%	0.462%	0.030%	0.110%	1.082%
16.0-18.0	0.100%	0.215%	0.477%	0.021%	0.412%	1.078%
18.0-20.0	0.130%	0.044%	0.471%	0.034%	0.216%	1.079%
1.1-6.0	0.137%	0.279%	0.460%	0.025%	0.656%	1.078%
15.0-20.0	0.107%	0.511%	0.460%	0.016%	0.742%	1.077%

Table C.3: Absolute values of systematic uncertainties on relative reconstruction efficiency for long and downstream candidates.

	Downstream			Long		
q^2 [GeV^2/c^4]	Lifetime	Model	Polarisation	Lifetime	Model	Polarisation
0.1-2.0	0.038%	0.226%	0.070%	0.003%	0.061%	0.117%
2.0-4.0	0.009%	0.091%	0.034%	0.020%	0.072%	0.076%
4.0-6.0	0.028%	0.162%	0.058%	0.018%	0.165%	0.040%
6.0-8.0	0.005%	0.080%	0.075%	0.041%	0.035%	0.053%
11.0-12.5	0.002%	0.207%	0.079%	0.002%	0.148%	0.076%
15.0-16.0	0.036%	0.094%	0.035%	0.022%	0.021%	0.089%
16.0-18.0	0.023%	0.027%	0.029%	0.023%	0.003%	0.031%
18.0-20.0	0.017%	0.145%	0.034%	0.008%	0.199%	0.063%
1.1-6.0	0.024%	0.215%	0.029%	0.012%	0.733%	0.051%
15.0-20.0	0.025%	0.220%	0.031%	0.004%	0.108%	0.029%

Table C.4: Absolute values of systematic uncertainties on relative trigger efficiency for long and downstream candidates.

	Downstream			Long		
q^2 [GeV^2/c^4]	Lifetime	Model	Polarisation	Lifetime	Model	Polarisation
0.1-2.0	0.022%	0.019%	0.025%	0.060%	0.106%	0.072%
2.0-4.0	0.127%	0.267%	0.017%	0.095%	0.002%	0.031%
4.0-6.0	0.116%	0.106%	0.045%	0.081%	0.139%	0.119%
6.0-8.0	0.111%	0.186%	0.020%	0.085%	0.387%	0.047%
11.0-12.5	0.008%	0.056%	0.017%	0.057%	0.030%	0.027%
15.0-16.0	0.002%	0.004%	0.066%	0.070%	0.124%	0.023%
16.0-18.0	0.024%	0.088%	0.027%	0.068%	0.105%	0.023%
18.0-20.0	0.031%	0.050%	0.027%	0.180%	0.506%	0.077%
1.1-6.0	0.118%	0.164%	0.037%	0.080%	0.183%	0.058%
15.0-20.0	0.001%	0.125%	0.037%	0.102%	0.541%	0.034%

Table C.5: Absolute values of systematic uncertainties on relative MVA efficiency for long and downstream candidates.

q^2 [GeV^2/c^4]	Reconstruction	Trigger	MVA
0.1-2.0	0.612%	0.250%	0.173%
2.0-4.0	0.515%	0.246%	0.223%
4.0-6.0	0.408%	0.180%	0.272%
6.0-8.0	0.412%	0.090%	0.218%
11.0-12.5	0.175%	0.047%	0.103%
15.0-16.0	0.962%	0.010%	0.141%
16.0-18.0	1.173%	0.037%	0.103%
18.0-20.0	1.557%	0.050%	0.122%
1.1-6.0	0.475%	0.220%	0.246%
15.0-20.0	1.254%	0.040%	0.083%

Table C.6: Values of DD vertexing systematic uncertainties on relative reconstruction, trigger and MVA efficiencies for downstream candidates.

APPENDIX D

Improved predictions for $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ observables.

The publication of the results included in this thesis triggered interest in the theory community, which produced improved lattice calculations and predictions [114]. Results with the new predictions overlaid are reported in Appendix ??. The latest predictions for the $f_{\rm L}$ and $A_{\rm FB}^h$ observables continue to agree with the measurements while a 3.3σ local tension is found for $A_{\rm FB}^{\ell}$ at high q^2. The branching fraction measurement is consistent with predictions at low q^2 while it is found to be above the theoretical value at high q^2. This section reports the measured quantities with the new predictions overlaid as reported in Ref. [114].

	Prediction	**Measurement**
$\langle \mathrm{d}\mathcal{B}/\mathrm{d}q^2\rangle_{[15,20]}$	0.756 ± 0.070	1.20 ± 0.27
$\langle F_L\rangle_{[15,20]}$	0.409 ± 0.013	$0.61^{+0.11}_{-0.14}$
$\langle A_{\rm FB}^{\ell}\rangle_{[15,20]}$	-0.350 ± 0.013	-0.05 ± 0.09
$\langle A_{\rm FB}^{\Lambda}\rangle_{[15,20]}$	-0.2710 ± 0.0092	-0.29 ± 0.08

Table D.1: Comparison of predictions for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ observables with the LHCb data presented in this thesis in the interval [15,20] GeV^2/c^4, where the measurement is most precise.

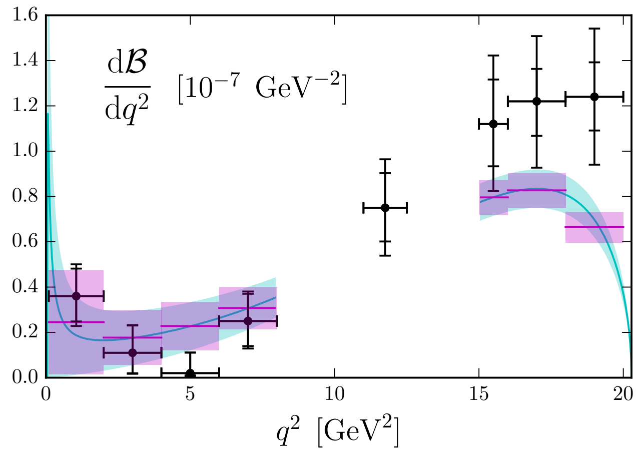

Figure D.1: Measurement of the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay as a function of q^2 already presented in Ch. 3 with improved Standard Model predictions from Ref. [114] overlaid.

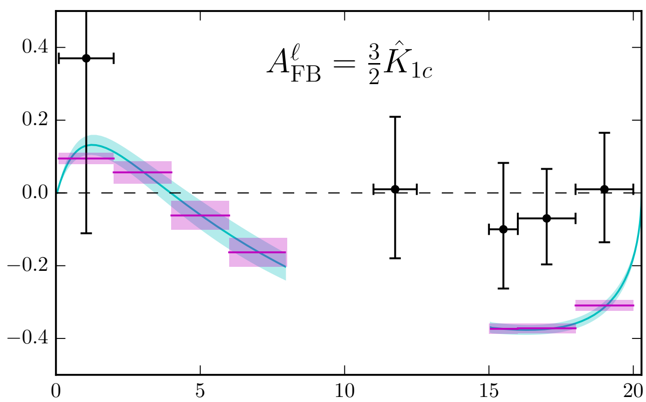

Figure D.2: Measurement of the lepton side forward-backward asymmetry, $A_{\rm FB}^{\ell}$, as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [114] overlaid.

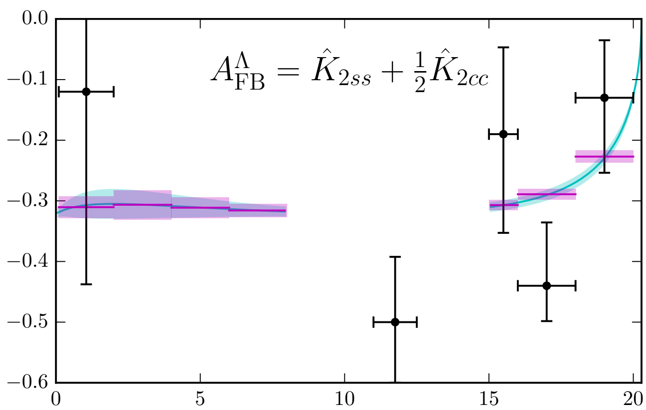

Figure D.3: Measurement of the hadron side forward-backward asymmetry, $A_{\rm FB}^{h}$, as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [114] overlaid.

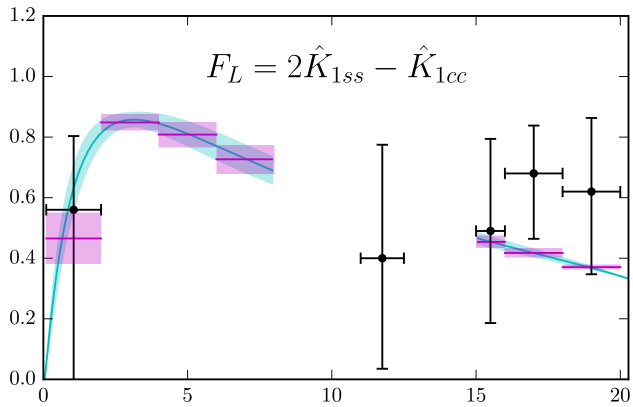

Figure D.4: Measurement of the fraction of longitudinally polarised dimuons, f_{L}, as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [114] overlaid.

APPENDIX E

Invariant mass fits to $B^0 \to K^{*0}\ell^+\ell^-$ simulated candidates

This appendix contains fits to the $m(K\pi\mu\mu)$ and $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}\ell^+\ell^-$ simulated candidates used to constrain parameters in the fit to data.

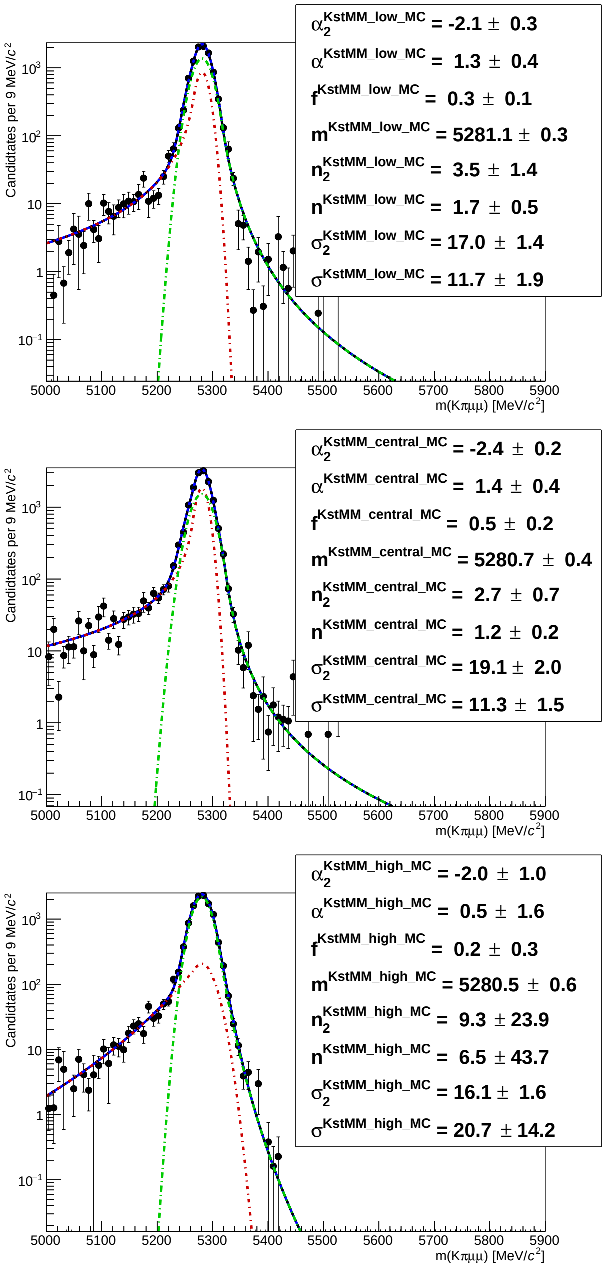

Figure E.1: Fitted $m(K\pi\mu\mu)$ mass spectrum for simulated events in the low (top), central (medium) and high (bottom) q^2 intervals.

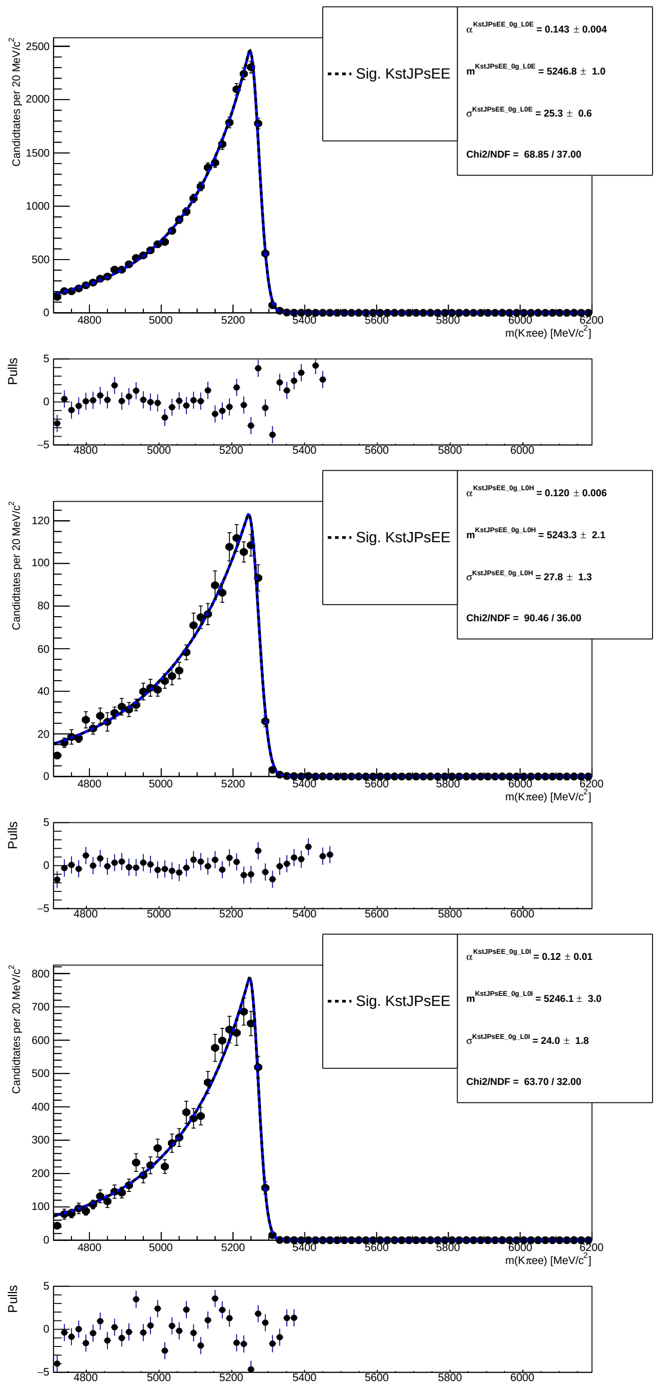

Figure E.2: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \to K^{*0}J/\psi(J/\psi \to ee)$ simulated events in the three trigger categories and no photon emitted.

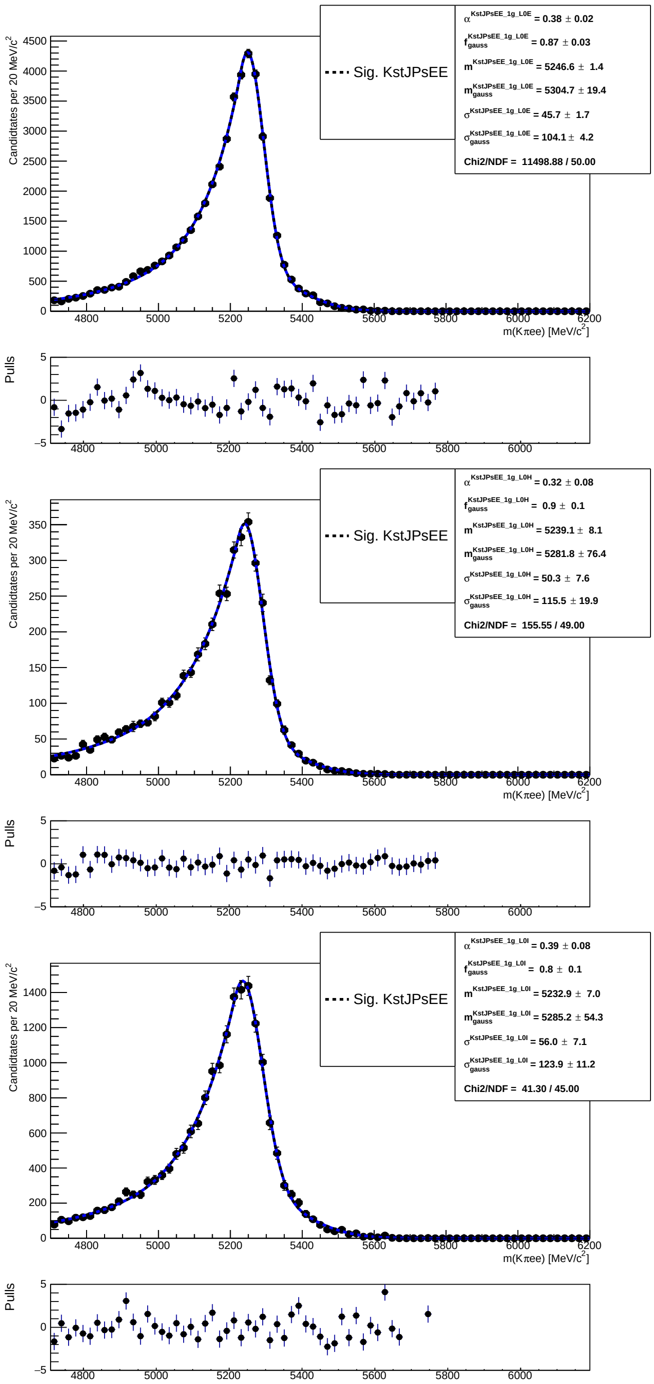

Figure E.3: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \to K^{*0}J/\psi(J/\psi \to ee)$ simulated events in the three trigger categories and one photon emitted.

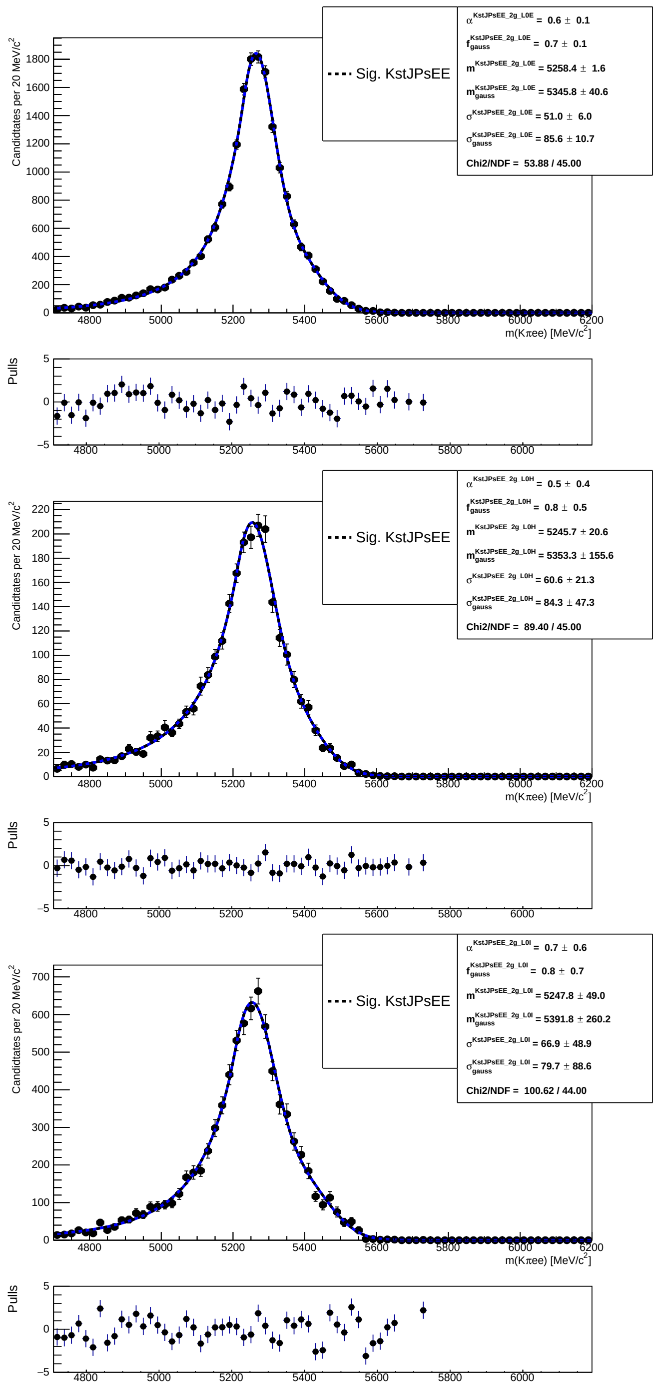

Figure E.4: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \to K^{*0}J/\psi(J/\psi \to ee)$ simulated events in the three trigger categories and two photons emitted.

APPENDIX F

Invariant mass fits to $B^0 \to K^{*0}e^+e^-$ candidates divided in trigger categories

This appendix contains fits to the $m(K\pi ee)$ invariant mass of rare and control channel candidates separately in the tree trigger categories. Each trigger category is always fit with its own PDF but in the main text only their sum is shown for simplicity.

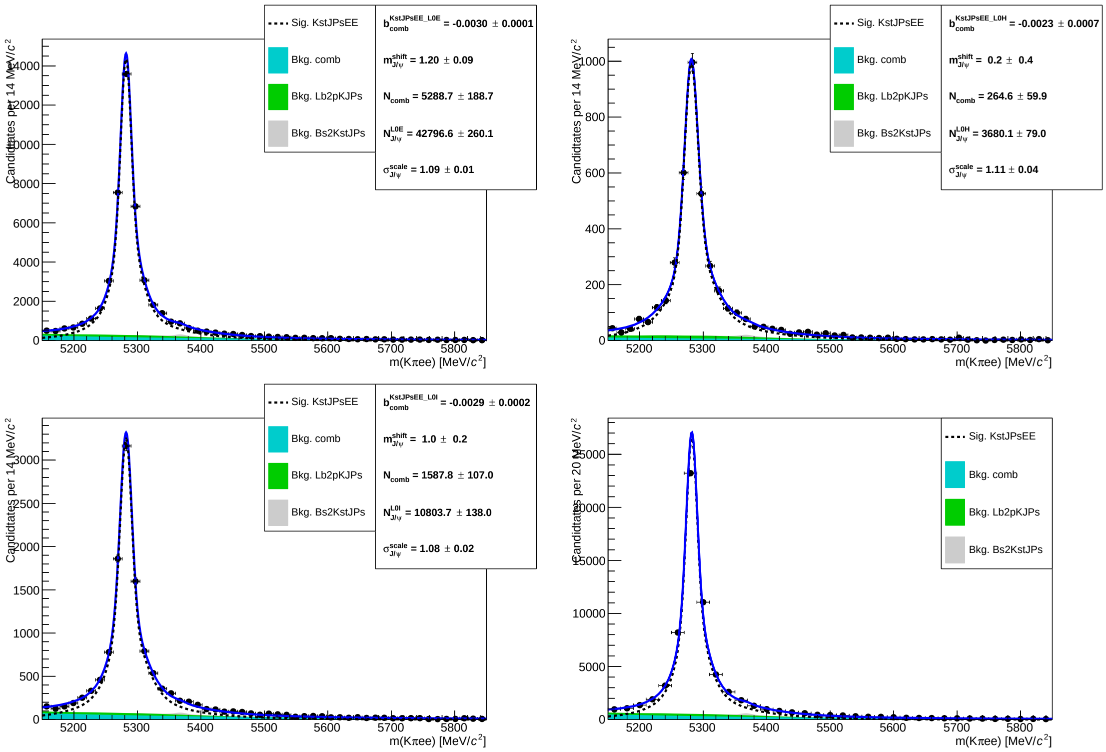

Figure F.1: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}(J/\psi \to e^+e^-)$ candidates in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

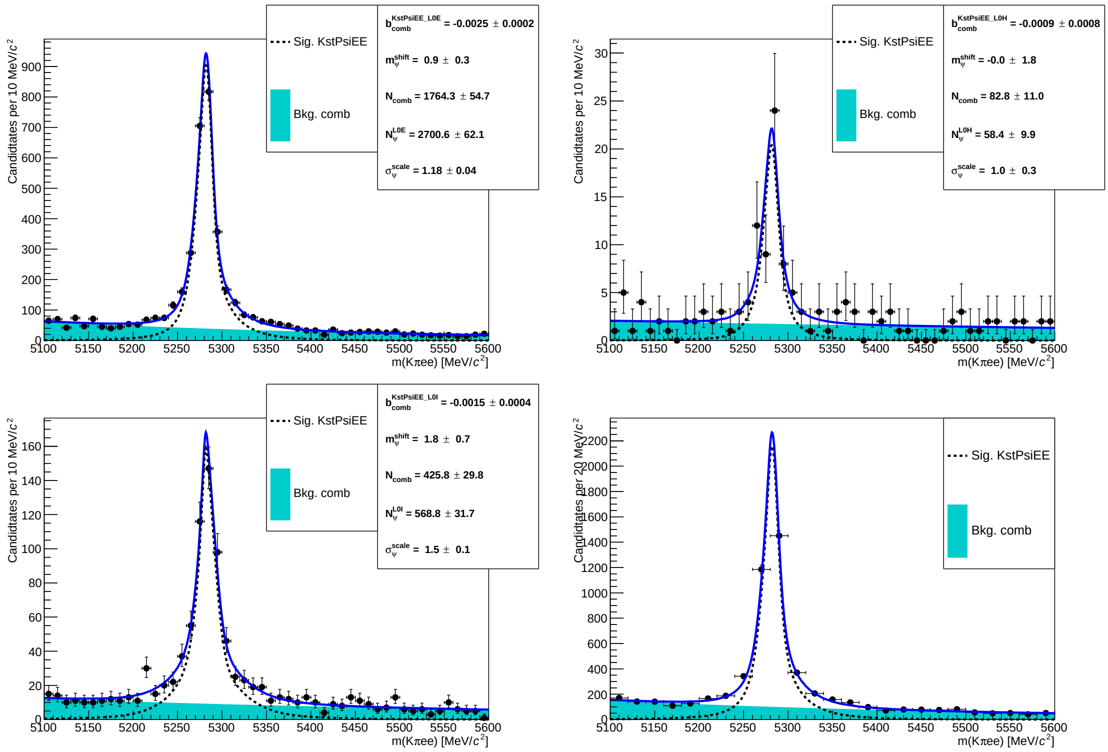

Figure F.2: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ candidates in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

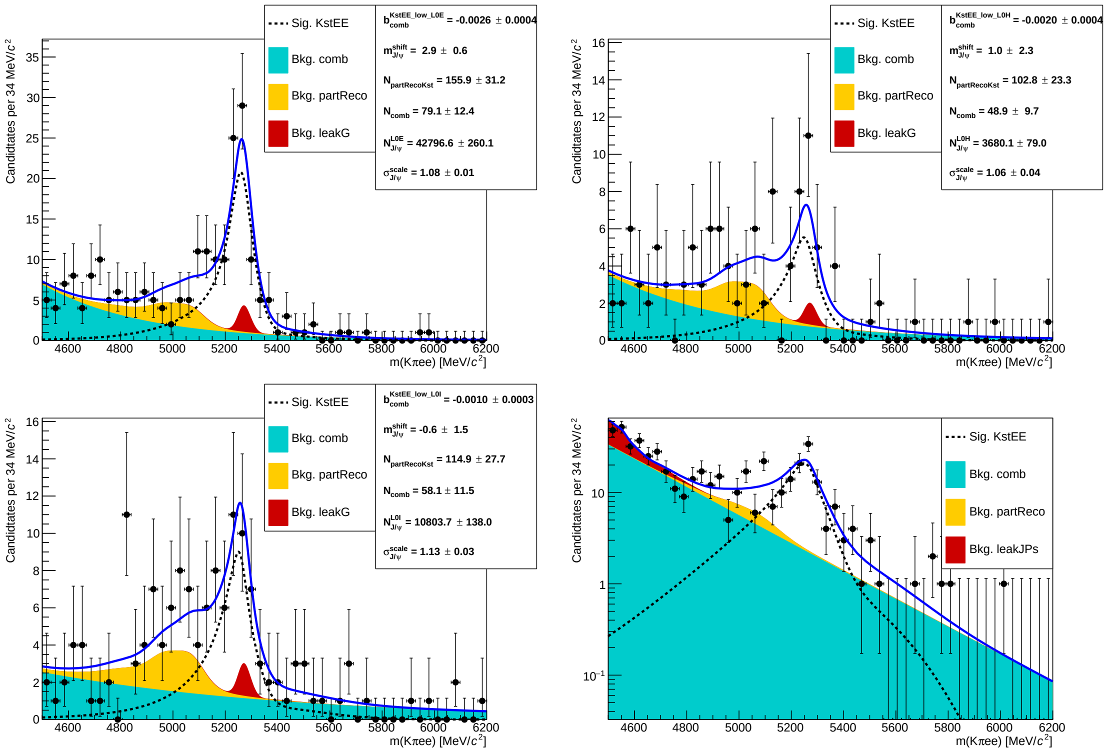

Figure F.3: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}e^+e^-$ candidates at low-q^2 in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

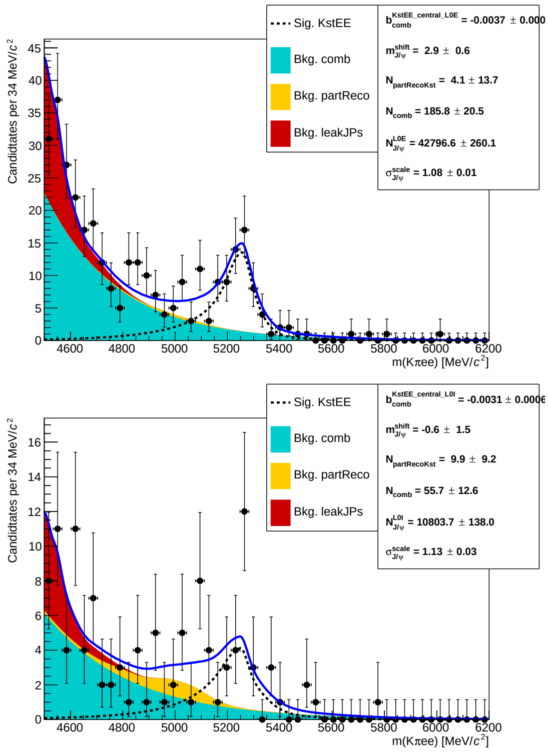

Figure F.4: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}e^+e^-$ candidates at central-q^2 in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

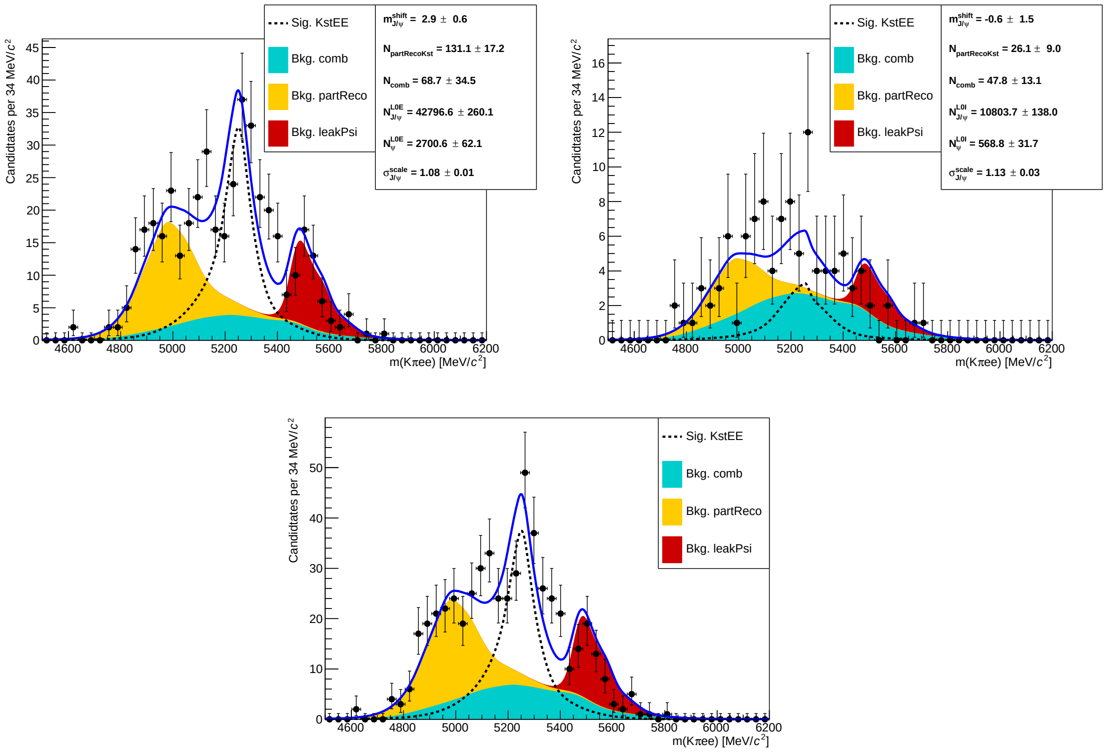

Figure F.5: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \to K^{*0}e^+e^-$ candidates at high-q^2 in the L0E and L0I trigger categories (top) separately, and (bottom) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.