

SEARCHING FOR NEW PHYSICS IN $b \rightarrow s\ell^+\ell^-$ 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.

July 16, 2015

ABSTRACT

Flavour Changing Neutral Currents are transitions between different quarks with the same charge such as $b \rightarrow s$ processes. These are forbidden at tree level in the Standard Model but can happen through loop electroweak 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 $b \rightarrow s\ell^+\ell^-$ decays are presented. First $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ 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, $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays are analysed measuring the ratio between the muon, $B^0 \rightarrow K^{*0}\mu^+\mu^-$, and electron, $B^0 \rightarrow K^{*0}e^+e^-$, 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, the concept of flavour and explains how rare decays can help us in the quest for beyond the SM physics. Chapter 2 describes the LHCb detector, which was used to collect the data analysed in this thesis. The rest is organised in two parts: Part I dealing with the analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays and Parts II describing the analysis of $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays. In Part I chapters from 3 to 7 treat the measurement of the differential branching fraction of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$, while chapters from 8 to 10 describe their angular analysis. Each of the two parts contains a brief theoretical introduction to the specific topic, a description of the data samples used and a description of the analysis which follows in both parts the a series of steps. First of all the selection process is described through which the interesting decay channels are isolated. Secondly, the yield of interesting signal events is separated from the remaining background by fitting the invariant mass distributions of the selected candidates. In the third step the efficiency of the detector is evaluated in order to be able to correct the raw yields. Finally, the systematic uncertainties are studied and the result determined.

DECLARATION OF AUTHORS CONTRIBUTION

I've done this and that bla bla... for the advice given throughout. The work in this part was also published and can be found at Ref. [1].

ACKNOWLEDGEMENTS

I thank everybody, evvvvvvvveverybody!

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

*Nel niente c'è una via che conduce
lontano dalla polvere del mondo.
(F. Bertossa)*

Contents

1	Introduction	1
1.1	Electromagnetic and weak interactions	4
1.2	Flavour and the CKM matrix	5
1.3	The puzzles of the SM	9
1.3.1	The flavour problem	10
1.4	Beyond the Standard Model	11
1.4.1	Flavour and BSM theories	11
1.5	Rare decays: a tool to search for new physics	13
1.5.1	Theoretical framework: the effective Hamiltonian	14
1.5.2	Operators	16
1.5.3	Phenomenology of $b \rightarrow s\ell^+\ell^-$ decays	18
1.5.4	Observables in $b \rightarrow s\ell^+\ell^-$ decays	19
1.6	Experimental status	20
1.6.1	Dimuon decays of b hadrons	20
1.6.2	Semileptonic $b \rightarrow s\ell^+\ell^-$ decays of b hadrons	22
1.6.3	Lepton Flavour Violation searches	23
2	The LHCb detector at the Large Hadron Collider	25
2.1	The Large Hadron Collider	25
2.2	The LHCb detector	27
2.3	The magnet	29
2.4	Tracking system	29
2.5	Calorimeters	32
2.5.1	Bremsstrahlung recovery for electrons	34
2.6	RICH	35
2.7	The muon system	36
2.8	Particle identification	37
2.8.1	PID calibration	39
2.9	Trigger and software	39
2.10	Kinematical fits	41
I	Differential branching fraction and angular analysis of the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay	43
3	The rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay	45
3.1	Analysis strategy and q^2 regions	47

3.2 Candidate types	48
3.3 Simulation	49
3.3.1 Decay Model	50
3.3.2 Kinematic re-weighting	51
3.3.3 Event type	52
4 Selection	54
4.1 Pre-selection	55
4.2 Neural Networks	56
4.3 MVA optimization	60
4.4 Trigger	62
4.5 Background from specific decays	63
5 Yield extraction	65
5.1 Fit description	65
5.2 Fit results	70
6 Efficiency	76
6.1 Geometric acceptance	77
6.2 Reconstruction and neural network efficiencies	78
6.3 Trigger efficiency	79
6.4 PID efficiency	80
6.5 Relative efficiencies	80
7 Systematic uncertainties and results	84
7.1 Systematic uncertainty on the yields	84
7.2 Systematic uncertainties on the efficiency determination	86
7.2.1 Effect of new physics on the decay model	86
7.2.2 Simulation statistics	87
7.2.3 Production polarisation and decay structure	88
7.2.4 Λ_b^0 lifetime	88
7.2.5 Downstream candidates reconstruction efficiency	88
7.2.6 Data-simulation discrepancies	90
7.3 Differential branching ratio extraction	90
8 Angular analysis	95
8.1 One-dimensional angular distributions	96
8.2 Multi-dimensional angular distributions	99
8.3 Angular resolution	101
9 Angular fit strategy	104
9.1 Feldman-cousins plug-in method	105
9.2 Modelling the angular distributions	107
9.3 Angular acceptance	107
9.4 Studies on a three-dimensional fit	109
10 Systematics uncertainties on angular observables and results	112

10.1 Angular correlations	112
10.2 Resolution	113
10.3 Efficiency description	114
10.4 Background parameterisation	116
10.5 Polarisation	117
10.6 J/ψ cross-check	118
10.7 Results	119
II The $R_{K^{*0}}$ analysis	123
11 Testing lepton flavour universality with R_H	125
11.0.1 Combining ratios	128
11.0.2 Experimental status	129
11.1 Analysis strategy	129
11.2 Choice of q^2 intervals	131
11.3 Data samples and simulation	131
11.3.1 Data-simulation corrections	132
12 Selection	135
12.1 Trigger and Stripping	136
12.2 PID	140
12.3 Peaking backgrounds	142
12.3.1 Charmonium vetoes	142
12.3.2 ϕ veto	142
12.3.3 $B^+ \rightarrow K^+ \ell^+ \ell^-$ plus a random pion	143
12.3.4 Λ_b decays	144
12.3.4.1 Other peaking backgrounds	144
12.3.5 Mis-reconstructed background	144
12.4 Multivariate analysis	145
12.5 MVA optimisation	149
13 Mass fits	152
13.1 Muon channels	153
13.2 Electron channels	154
13.2.1 Signal PDFs for the electron channels in the central- q^2 interval	157
13.2.2 Background PDFs for the electron channels in the central- q^2 interval	158
13.2.3 Summary of the fit to electron channels in the central- q^2 interval	160
13.2.4 Electron channels fits in the high- q^2 interval	161
13.3 Fit summary	165
14 Efficiency	167
14.1 Geometric efficiency	170
14.2 Reconstruction efficiency and bin migration	170
14.3 PID efficiency	172
14.4 Trigger efficiency	172
14.4.1 Electron channels	173

14.4.2 TISTOS cross-check	175
14.5 Neural Networks efficiency	176
15 Systematic uncertainties	178
16 Result extraction	179
16.1 $R_{J/\psi}$ sanity check	179
16.2 R_{K^*} result summary	180
16.3 Branching ratios and expectations	181
17 Conclusions	183
A Decay models	193
A.1 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution	193
A.2 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution	197
B Data-simulation comparison	198
C Systematic uncertainties on the efficiency calculation for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction analysis.	201
D Invariant mass fits to $B^0 \rightarrow K^{*0} e^+ e^-$ simulated candidates	204

1

CHAPTER 1

2

3

Introduction

4

5 The Standard Model of particle physics (SM) is a Quantum Field Theory (QFT)
6 describing strong and electroweak (EW) interactions. It was formulated in his cur-
7 rent form in the mid-70s and has been an extremely successful and predictive theory
8 since then. Almost all known phenomena from 1 eV up to several hundred GeV are
9 well described by the SM and experiments at the Large Hadron Collider (LHC) are
10 now probing the SM up to the TeV scale. As an example of the level of accuracy
11 of the SM in Tab. 1.1 are reported the predicted and measured values of the widths
12 of the Z and W bosons [2]. Finally, in 2013 the Higgs boson was observed, one
13 of the fundamental building blocks of the theory, which gives a solid basis to it by
14 introducing a mechanism that produces particles' masses [3]. Despite the success
15 of the SM, experimentally well established effects, like neutrino oscillations and the
16 presence of dark matter, are outside the reach of this theory. Furthermore, the
17 model does not include the description of gravity, which can be neglected at the
18 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.

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 $SU(3)_C \times SU(2)_W \times U(1)_Y$, 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 [4], 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	Rel. strength	Range (m)	Mediator mass (GeV/c^2)
Strong	g	1	∞	0
EM	γ	10^{-3}	∞	0
Weak	Z, W^\pm	10^{-16}	10^{-18}	$W^\pm = 80.399$ $Z_0 = 91.188$
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,

³⁷ so called, strong interaction, while leptons, which do not carry colour charge, are
³⁸ insensitive to it. For each particle a corresponding anti-particle exists with opposite
³⁹ quantum numbers. Finally, fermions are divided into three families having similar
⁴⁰ properties but different masses. This last structure embedded in the SM is also
⁴¹ called flavour structure and it will be the main tool used in this thesis, a more
⁴² 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

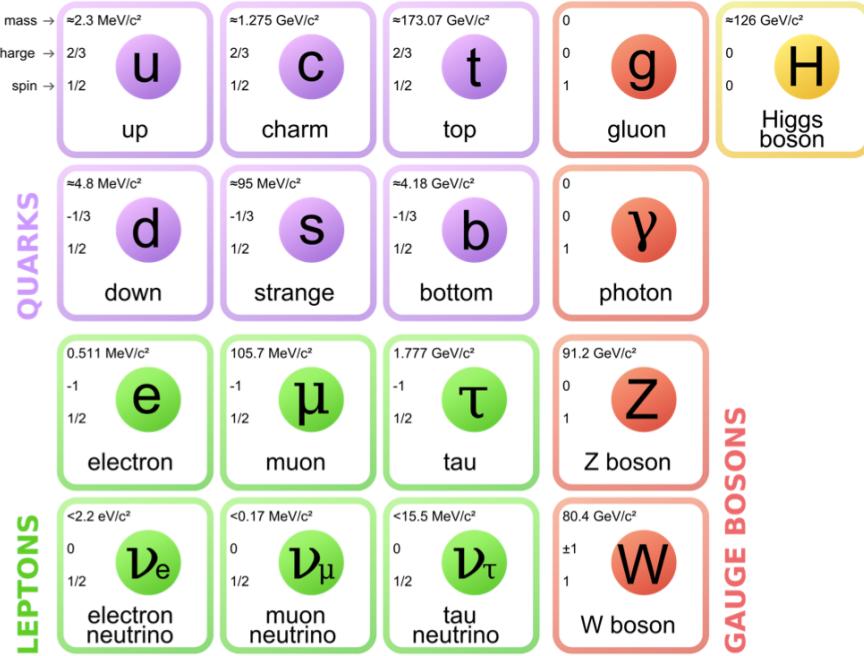


Figure 1.1: A scheme of the fundamental particles in the SM with their properties.

⁴³
⁴⁴ of the strong interaction quarks cannot be observed alone but are always combined
⁴⁵ with other quarks to form color singlets. Non-fundamental particles composed by
⁴⁶ quarks are called hadrons and can be divided in mesons, where the color singlet is
⁴⁷ achieved by the combination of a quark and its antiquark ($q \bar{q}$), and baryons formed
⁴⁸ by three quarks ($q q q$) of different colours. Recently, in 2014 and 2015 evidence for
⁴⁹ new states, formed by four and five quarks, was found [].

⁵⁰ 1.1 Electromagnetic and weak interactions

⁵¹ The Electromagnetic (EM) force is responsible for binding electrons and nuclei to-
⁵² gether in atoms. Its force carrier, the photon, is the gauge boson of the EM force.
⁵³ In the SM the photon must be massless...

⁵⁴ *write something here!*

⁵⁵ The weak interaction is responsible for the β decay of nuclei. Unlike the electromag-
⁵⁶ netic force, that affects only charged particles, all known fermions interact through
⁵⁷ the weak interaction. In the SM this interaction is mediated by the emission or ab-
⁵⁸ sorption of W^\pm and Z bosons. The electroweak symmetry is spontaneously broken
⁵⁹ by the Higgs field [5] and this causes the W^\pm and Z bosons to become massive (see
⁶⁰ Tab. 1.2) and consequently the weak force has a very short range. In fact using
⁶¹ Heisenberg's Principle together with Einstein's formula $\Delta E = mc^2$, which relates
⁶² mass and energy, and knowing that the maximum space that a particle can cover
⁶³ in a time Δt is $r = c\Delta t$, qualitatively $r \sim \hbar/mc$. In this picture the carriers of the
⁶⁴ weak force can travel $r \sim 2 \cdot 10^{-3}$ fm. The weak interaction is also the only one
⁶⁵ that violates parity-symmetry, which states that interactions are invariant under
⁶⁶ a reflection of all coordinates. This symmetry breaking arises from the fact that
⁶⁷ only left-handed fermions interact through the weak interaction as discovered by
⁶⁸ Wu in 1957 [6]. Similarly, the weak interaction is the only one that also breaks
⁶⁹ the CP symmetry, which combines parity transformations and charge conjugation.
⁷⁰ This is particularly interesting because all interactions are invariant under the CPT
⁷¹ transformation, which combines the CP transformation and time reversal, hence,
⁷² breaking CP the weak interaction must also be not invariant under time reversal.

⁷³ In 1968 Salam, Glashow and Weinberg unified the weak and electromagnetic forces in
⁷⁴ a single theory, with a single coupling constant [2]. The EW interactions are divided
⁷⁵ into charged currents (CC) and neutral currents (NC). In the first group, quarks
⁷⁶ and leptons interact with the W^\pm bosons, producing decays such as $\mu^+(\mu^-) \rightarrow$
⁷⁷ $e^+\nu_e\bar{\nu}_\mu(e^-\bar{\nu}_e\nu_\mu)$ and $n \rightarrow pe^-\bar{\nu}_e(\bar{p}e^+\nu_e)$. The study of these processes confirmed

78 that only the left-handed (right-handed) component of fermions (anti-fermions)
79 takes part in weak processes. The CC interactions have a peculiarity: they are
80 the only interactions in the SM that violate flavour conservation at tree level (see
81 next section), while any other interaction not conserving flavour has to happen
82 through loops. The second group of EW interactions, NC, corresponds to diagrams
83 mediated by a photon or a Z boson interacting with a fermion and its anti-fermion.

84 1.2 Flavour and the CKM matrix

85 “Flavour” in particle physics refers to the quark-lepton composition of a particle.
86 The introduction of flavour quantum numbers was motivated in order to explain
87 why some decays, although kinematically allowed, have never been observed. To all
88 leptons is assigned a quantum number $L_\ell = 1$ (where $\ell = e, \mu, \tau$), which in the SM is
89 conserved by all interactions. This conservation is experimentally well established;
90 for example decays like $\mu^- \rightarrow e^- \gamma$, which is kinematically possible, have never been
91 observed. This is explained by the fact that the lepton number in the initial and final
92 state are different and therefore lepton flavour would be violated. In the hadronic
93 sector particles carry flavour numbers described as follow:

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

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

101 Measuring branching fractions of weak decays like $\pi \rightarrow \mu\nu_\mu$ and $K \rightarrow \mu\nu_\mu$, corre-
102 sponding respectively to $ud \rightarrow \mu\nu_\mu$ and $us \rightarrow \mu\nu_\mu$ processes, suggested the existence

of more than one coupling constant for different quarks. Nicola Cabibbo [2], in order to preserve the universality of weak interactions, suggested that the branching fraction differences could arise from the fact that the doublets participating in the weak interactions are an admixture of the flavour 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.

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

Considering a 6 quark system one angle is not enough to describe a rotation but the mixing system 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 $(N - 1)^2$. 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. Eq. 1.2 reports a parametrisation

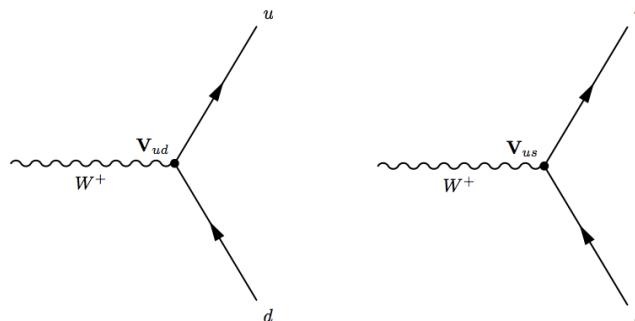


Figure 1.2: Feynman diagrams with CKM weights on weak interaction vertices

of the CKM matrix together with the most recent measured values of its elements [2]. In this parametrisation ρ , A , and λ are the real constants and η the imaginary phase;

in Eq. 1.3 are reported their relations with the 3 mixing angles.

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & A\lambda^2 & 1 \end{pmatrix} + O(\lambda^3) = \\ = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.0014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.00412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \quad (1.2)$$

109

$$\begin{aligned} \lambda &= \sin(\theta_{12}) = \sin(\theta_c) \\ A\lambda^2 &= \sin(\theta_{23}) \\ A\lambda^3(\rho - i\eta) &= \sin(\theta_{13})e^{i\delta} \end{aligned} \quad (1.3)$$

110 Figure 1.2 displays examples of CC processes together with the CKM elements
 111 associated with their vertices. It is interesting to note that the CKM matrix has a
 112 hierarchical form, namely elements on the diagonal are approximately 1 and become
 113 smaller and smaller going farther from the diagonal. This structure is not explained
 114 in the SM. Another feature to note is that, due to the unitarity of the matrix, the
 115 transformation has no effect on neutral interactions. In fact defining $q' = Vq$

$$\bar{q}'q' = \bar{q}V^*Vq = \bar{q}q. \quad (1.4)$$

116 As a result flavour-changing neutral currents are forbidden at tree level in the SM.
 117 The unitarity of the CKM matrix imposes constraints to its elements of the form:

$$\sum_i |V_{ik}|^2 = 1 \text{ and } \sum_k V_{ik}V_{jk}^* = 0. \quad (1.5)$$

118 These correspond to constraints to three complex numbers, which can be viewed
 119 as triangles in the (ρ, η) plane; these are called “unitarity triangles”. The most
 120 commonly used unitarity triangle arises from $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$. Figure 1.3
 121 shows a representation of such triangle together with a plot summarising the most
 122 up to date experimental constraints to its parameters [7]. 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.

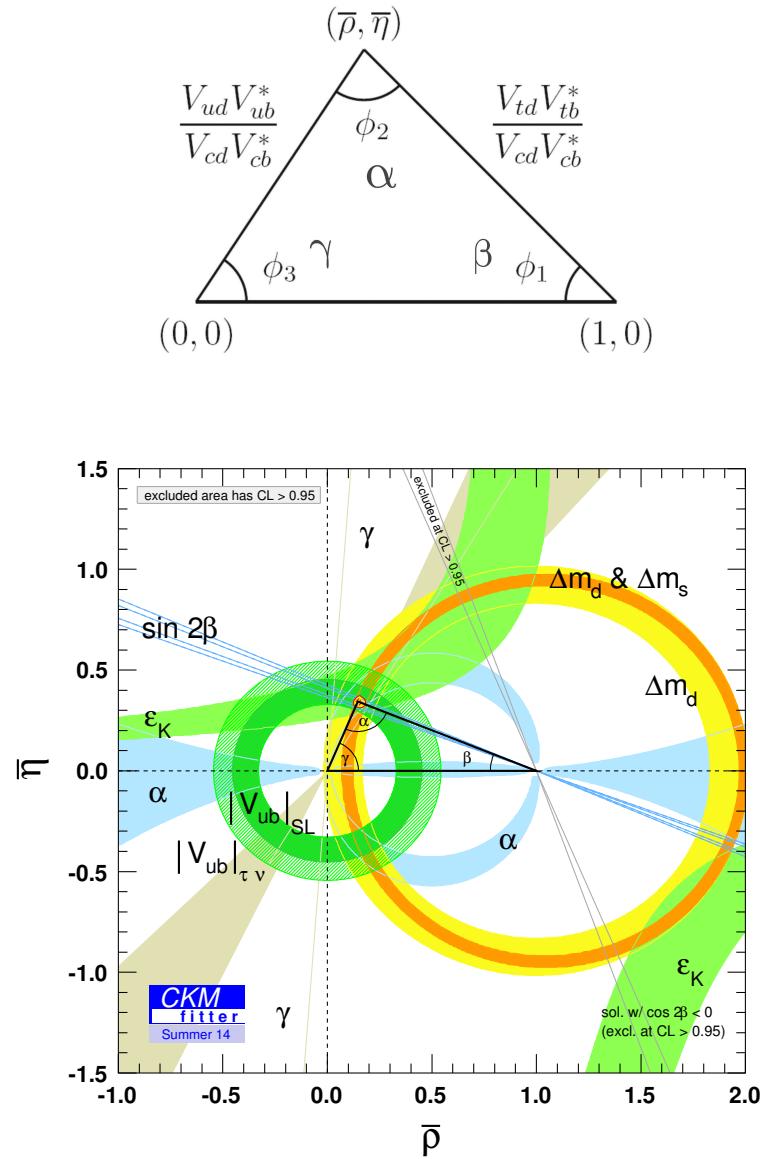


Figure 1.3: (top) A representation of the unitarity triangle and its parameters. (bottom) A summary of the most up to date measurements of the unitarity triangle parameters [7].

¹²⁷ 1.3 The puzzles of the SM

¹²⁸ Despite the confirmation of many predictions of the SM, the theory has several
¹²⁹ limitations and is unable to account for some well established experimental facts:

¹³⁰ • *Dark matter*: experimental evidence tells us that the content of visible matter
¹³¹ in the universe is not enough to account for the observed rotation of galaxies [8].
¹³² The most natural way to solve the problem is the hypothesis of a form of
¹³³ matter that interacts with the gravitational field but not with the other SM
¹³⁴ interactions.

¹³⁵ • *Matter-antimatter asymmetry*: a large asymmetry is observed between the
¹³⁶ quantity of matter and antimatter in the universe, $O(10^{-9})$. Assuming that
¹³⁷ both were equally created in the initial state of the universe, a condition such
¹³⁸ as the violation of the CP symmetry is necessary to account for such observed
¹³⁹ differences. However, the magnitude of CP violation predicted by the SM,
¹⁴⁰ $O(10^{-20})$, is not enough to explain the observed imbalance [9].

¹⁴¹ • *Gravity*: even though the gravitational force was the first to be discovered this
¹⁴² is not included in the SM. When introducing gravity in the framework of QFT
¹⁴³ the theory diverges. On the other hand gravity becomes irrelevant for small
¹⁴⁴ masses as those of particles and can be neglected in good approximation at the
¹⁴⁵ EW energy scale. Many attempts were made but there is not yet a consistent
¹⁴⁶ procedure to introduce gravity in the SM.

¹⁴⁷ • *Neutrino oscillation*: measurements regarding solar and atmospheric neutrinos
¹⁴⁸ as well as neutrinos from nuclear reactors established that neutrinos can
¹⁴⁹ change flavour while propagating in space. This is not predicted in the SM, in
¹⁵⁰ fact in the SM neutrinos are massless, while an oscillation requires a non zero
¹⁵¹ mass [10].

¹⁵² • *The hierarchy problem*: The mass of a scalar (spin 0) particle, such as the
¹⁵³ Higgs boson, suffers from quantum corrections due to the physics above a

¹⁵⁴ certain scale. As new physics can appear anywhere up to the Planck scale,
¹⁵⁵ $\sim 10^{19}$ GeV, at which gravity cannot be neglected any more, these corrections
¹⁵⁶ can be very large and it would require a high level of fine-tuning for them to
¹⁵⁷ cancel out and give such a small value as the one measured for the Higgs Mass,
¹⁵⁸ ~ 126 GeV/ c^2 [11, 12].

¹⁵⁹ In conclusion, even though the SM has been very successful in describing the prop-
¹⁶⁰ erties of the observed particles and their interactions so far, because of its many
¹⁶¹ puzzles, it is believed only to be part of a more general theory or only to be valid
¹⁶² up to a certain energy scale.

¹⁶³ 1.3.1 The flavour problem

¹⁶⁴ Flavour Changing Charged Currents (FCCC) that are mediated by the W^\pm bosons
¹⁶⁵ are the only sources of flavour changing interaction in the SM and, in particular, of
¹⁶⁶ generation changing interactions, where a quark or a lepton of a family transforms
¹⁶⁷ into one of another family. Another class of processes is the Flavour Changing
¹⁶⁸ Neutral Currents (FCNCs), e.g. transitions from a b quark with a charge of -1/3 to
¹⁶⁹ 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. In the SM there is no fundamental reason

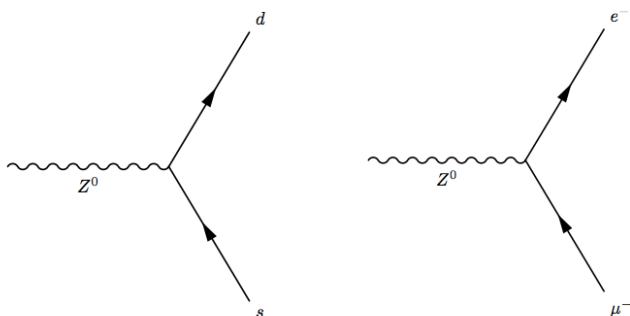


Figure 1.4: Feynman diagrams of FCNCs processes forbidden in the SM.

¹⁷⁰
¹⁷¹ why there cannot be FCNCs and, yet, they are experimentally observed to be highly
¹⁷² suppressed. On the other hand the observation of neutrino oscillation proves that

173 flavour is not an exact symmetry and is not always conserved. Furthermore, the
174 values of the terms of the CKM matrix and the PMNS matrix, which the mixing-
175 matrix, equivalent to the CKM, in the lepton sector, are not explained in the SM
176 and have to be measured experimentally. These open problems motivate searches
177 for flavour symmetries and deeper motivations for flavour conservation.

178 1.4 Beyond the Standard Model

179 From the last sections it is evident that, despite the great success of the SM, there
180 is a need to explore theories Beyond the SM (BSM). Among the most promising
181 approaches there are those invoking Super-Symmetry and extra-dimensions. In
182 Super-Symmetry new degrees of freedom are introduced to suppress the diverging
183 terms of the scalar mass. This theory assumes that for each fermion there is a
184 corresponding boson and, since bosons and fermions contribute with opposite sign
185 to the mass term, these would cancel out [13]. Supersymmetry also provides a
186 natural candidate for dark matter, the neutralino, which is a weakly interacting
187 stable particle. The idea to introduce extra-dimensions was triggered by the fact
188 that, normally, gravity is not relevant in particle physics but it would be natural if
189 all forces had similar strength. By adding extra dimensions to the normal 3 spatial
190 dimensions, one can restore the strength of gravity, as this could be dispersed by the
191 wider space available [14]. In all these approaches severe constraints to masses and
192 couplings must be imposed to maintain compatibility with the SM at the electroweak
193 scale.

194 1.4.1 Flavour and BSM theories

195 Most BSM theories predict processes violating flavour conservation. Therefore, the
196 observation or non-observation of these processes can give important information
197 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 those 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. These features can be assured by symmetry principles, which makes these types of models naturally compatible with the SM. Examples of such models include the MSSM which minimal flavour violation and the SM with one extra-dimension. A review of MFV models is presented in Refs. [15, 16]. The MFV paradigm provides a way to resolve the tension between expectation, driven by naturalness arguments, that NP should be at the TeV scale and limits on FCNC processes that point to much higher scales. A powerful test of MFV is provided by the study of ratios between $b \rightarrow d$ and $b \rightarrow s$ 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 [17], as this is a purely leptonic decay free from hadronic uncertainties. In the SM such ratios are approximately equal to $|V_{td}/V_{ts}| \sim 1/25$, 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 models valid up to the GUT scale one can consider simplified models, which typically start from the SM and incorporate 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 data 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. As for the SM penguins, this 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. [18]. Leptoquarks are bosonic particles that carry one quark and one lepton flavour quantum number. They can be spin 1 but they are commonly assumed to

228 be scalar particles. A tree level exchange of a leptoquark induces processes such as
 229 $b \rightarrow (s, d)\ell^+\ell^-$, and therefore can result in an enhancement of their decay rates with
 230 respect to the SM [19]. Leptoquarks would also provide a natural explanation for
 231 non-universal couplings to leptons.

232 1.5 Rare decays: a tool to search for new physics

233 In the Standard Model FCNC processes are forbidden at tree level but can occur
 234 through loops diagrams such as W box or penguin diagrams (see Fig. 1.5). The
 235 branching fractions of decays going through these processes are small, typically
 236 $\sim 10^{-6}$ or lower, and therefore they are called “rare decays”. Additional NP contri-
 237 butions to the virtual loops are not necessarily suppressed with respect to the SM
 238 component and this makes these decays very sensitive to new physics. This approach
 239 to new physics searches is interesting as new particles could be at a high mass scale
 240 not accessible at colliders but their effect could be observed in loops. Radiative and
 241 penguin decays are particularly interesting because they are theoretically well un-
 242 derstood, which allows precise comparisons with measurements. Furthermore, they
 243 provide a great quantity of observables that can be affected by NP, not only decay
 244 rates, but also CP asymmetries and angular observables such as forward-backward
 245 asymmetries. The joint analysis of different observables can help building a consist-
 tent picture and rule out specific models.

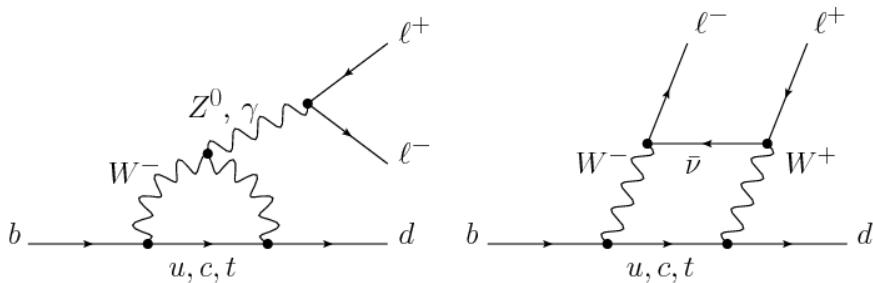


Figure 1.5: Loop Feynmann diagrams allowing $b \rightarrow d$ FCNC processes: penguin diagram (left) and W box (right).

²⁴⁷ 1.5.1 Theoretical framework: the effective Hamiltonian

²⁴⁸ Rare decays of b hadrons are governed by an interplay between weak and strong
²⁴⁹ interactions. The large masses of W , Z and top quark compared to that of the b quark
²⁵⁰ allow the construction of an effective theory that divides the problem of calculating
²⁵¹ weak decay amplitudes into two parts: “short-distance” and “long-distance” effects
²⁵² separated at an energy scale μ . The first part, dealing with short distance physics,
²⁵³ handles perturbative contributions due to energy scales above the b mass. The
²⁵⁴ second part typically deals with non-perturbative contributions. The classic example
²⁵⁵ of an effective theory is the Fermi theory of weak interactions which describes the
²⁵⁶ 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.

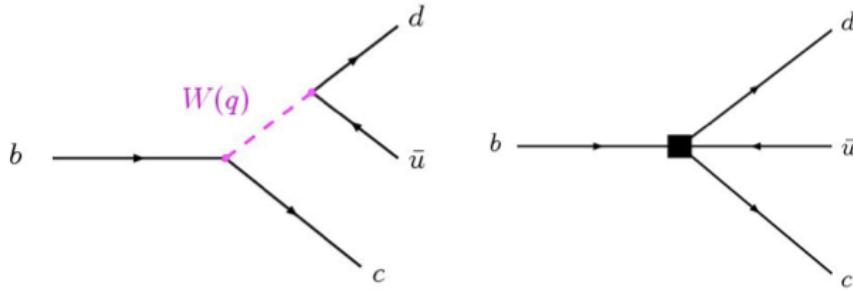


Figure 1.6: Example of a Fermi theory in which the full theory is divided between a short distance contribution, hidden in the vertex, and a long distance contribution.

²⁵⁷

²⁵⁸ The effective hamiltonian [20] relevant to $b \rightarrow s/d\gamma$ and $b \rightarrow s/d\ell^+\ell^-$ transitions
²⁵⁹ can be written as:

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

²⁶⁰ where G_F denotes the Fermi coupling constant and the λ constants are the CKM
²⁶¹ factors, $\lambda_q^t = V_{tb}V_{tq}^*$ and $\lambda_q^u = V_{ub}V_{uq}^*$. To obtain this formula the method of the
²⁶² Operator Product Expansion (OPE) [21] is used. This implements a summation
²⁶³ over all contributing operators weighted by corresponding constants called Wilson
²⁶⁴ coefficients. In this Hamiltonian the long-distance contributions are contained in
²⁶⁵ the operator matrix elements, \mathcal{O}_i , and the short-distance physics is described by the

266 Wilson Coefficients, C_i . Operators and coefficients are evaluated at the renormalization
267 scale μ . Any particle that contributes to the decay and has a mass greater than
268 the scale μ will affect the value of at least one of the Wilson coefficients, including
269 SM particles as the top quark. In $b \rightarrow s$ quark transitions, which are the main topic
270 of this thesis, the doubly Cabibbo-suppressed contributions proportional to λ_s^u can
271 be neglected.

272 In order to describe SM processes the effective theory must be matched with the
273 SM by requiring the equality between each term in effective theory and the full
274 theoretical calculation at a matching scale, typically the EW scale (μ_W). Then, using
275 the scale independence of the effective Hamiltonian, one can derive a renormalization
276 group equation for the Wilson Coefficients

$$\mu \frac{d}{d\mu} C_i(\mu) = \gamma_{ij} C_j(\mu), \quad (1.7)$$

277 where the matrix γ is the anomalous dimensions matrix of the operators \mathcal{O}_i . At
278 leading order the solution is given by [22]:

$$C_i(\mu) = \left[\frac{\alpha_s(\mu_W)}{\alpha_s(\mu)} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W) = \left[\frac{1}{1 + \beta_0 \frac{\alpha_s(\mu)}{4\pi} \ln \frac{\mu_W^2}{\mu^2}} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W), \quad (1.8)$$

279 where α_s is the strong coupling constant. In the SM, using $\mu_W = m_b$, the Wilson
280 Coefficients have values:

$$C_7^{SM} = -0.3, \quad C_9^{SM} = 4.2, \quad C_{10}^{SM} = -4.2. \quad (1.9)$$

281 New physics contributions appear in the Wilson Coefficients in the form of additive
282 factors:

$$C_i = C_i^{NP} + C_i^{SM}. \quad (1.10)$$

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

²⁸⁵ $I = B$ and $F = K^{*0}\mu^+\mu^-$) the decay amplitude can be calculated as

$$A(M \rightarrow F) = \langle M | \mathcal{H}_{eff} | F \rangle = \mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} \sum V_{CKM}^i C_i(\mu) \langle M | \mathcal{O}_i(\mu) | F \rangle \quad (1.11)$$

²⁸⁶ where $\langle M | \mathcal{O}_i(\mu) | F \rangle$ are the hadronic matrix elements also called “form factors”.

²⁸⁷ These can be evaluated using non perturbative methods such as lattice calculations.

²⁸⁸ However, due to the limitations of these methods, they represent the dominant

²⁸⁹ source of uncertainty in theoretical calculations.

²⁹⁰ 1.5.2 Operators

²⁹¹ Separating the left- and right-handed components the effective Hamiltonian is

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{\alpha_e}{4\pi} \sum_{i=1}^{10} [C_i \mathcal{O}_i + C'_i \mathcal{O}'_i]. \quad (1.12)$$

²⁹² The operators are that are relevant for radiative transitions and leptonic penguin

²⁹³ processes are the following [17]:

$$\begin{aligned} \mathcal{O}_7 &= \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_R b) F_{\mu\nu} & \mathcal{O}'_7 &= \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_L b) F_{\mu\nu} \\ \mathcal{O}_9 &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell) & \mathcal{O}'_9 &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu \ell) \\ \mathcal{O}_{10} &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell) & \mathcal{O}'_{10} &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu \gamma_5 \ell) \end{aligned} \quad (1.13)$$

²⁹⁴ where $P_{L/R} = (1 \mp \gamma_5)/2$ denote the left/right handed chiral projection, T^a are the

²⁹⁵ QCD generators and $F_{\mu\nu}$ is the electromagnetic field tensor. The \mathcal{O}' operators cor-

²⁹⁶ respond to right-handed coupling obtained by swapping P_R and P_L in the equations.

²⁹⁷ In the SM, as well as in MFV models where the flavour violation is entirely ruled by

²⁹⁸ the CKM matrix, the C' Wilson Coefficients are suppressed by the strange coupling,

²⁹⁹ $C'_i \sim (m_s/m_b)C_i$. The operator \mathcal{O}_7 relates to penguin diagrams that are mediated

³⁰⁰ via a photon. It represents the dominant contribution to the radiative $b \rightarrow s\gamma$

³⁰¹ transition and contributes to $b \rightarrow s\ell^+\ell^-$ processes when the virtual photon decays

³⁰² into a dilepton pair. The semileptonic \mathcal{O}_9 and \mathcal{O}_{10} correspond to penguin diagrams

³⁰³ mediated by a Z boson and W mediated box diagrams. These are the dominant
³⁰⁴ contributions in semileptonic $b \rightarrow s\ell^+\ell^-$ decays. The vertices corresponding to the
radiative and semileptonic operators are illustrated in Fig. 1.7

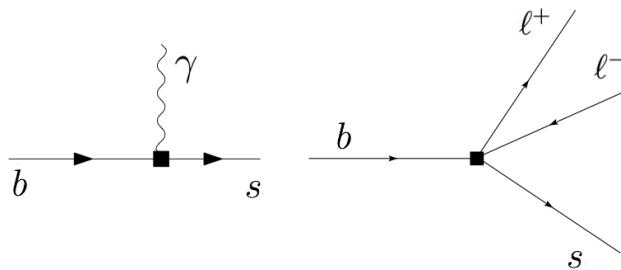


Figure 1.7: Interaction vertices corresponding to the radiative (left) and semileptonic (right) operators.

³⁰⁵

³⁰⁶ It is also common to express the semileptonic operators in a basis with left and right
³⁰⁷ projected leptons

$$\begin{aligned} \mathcal{O}_{LL} &= (\mathcal{O}_9 - \mathcal{O}_{10})/2 & \mathcal{O}_{LR} &= (\mathcal{O}_9 + \mathcal{O}_{10})/2 \\ \mathcal{O}_{RR} &= (\mathcal{O}'_9 - \mathcal{O}'_{10})/2 & \mathcal{O}'_{RL} &= (\mathcal{O}'_9 + \mathcal{O}'_{10})/2 \end{aligned} \quad (1.14)$$

³⁰⁸ where the Wilson Coefficients are also redefined as

$$\begin{aligned} C_{LL} &= C_9 - C_{10}, & C_{LR} &= C_9 + C_{10}, \\ C_{RR} &= C'_9 - C'_{10}, & C'_{RL} &= C'_9 + C_{10}. \end{aligned} \quad (1.15)$$

³⁰⁹ This basis is particularly useful in frameworks where BSM physics at a high mass
³¹⁰ scale respects the $SU(2)_W$ part of the SM gauge symmetry group. Finally, in the
³¹¹ picture presented in this section all operators were considered as universal with
³¹² respect of the flavour of the involved leptons. However, BSM models often contain
³¹³ sources of lepton flavour violation leading to a split of the same operators depending
³¹⁴ on the lepton considered: $C_i \rightarrow C_i^e, C_i^\mu, C_i^\tau$ and $\mathcal{O}_i \rightarrow \mathcal{O}_i^e, \mathcal{O}_i^\mu, \mathcal{O}_i^\tau$.

315 1.5.3 Phenomenology of $b \rightarrow s\ell^+\ell^-$ decays

316 Semileptonic b hadron decays are characterised by two kinematic regimes which
 317 are treated theoretically in different ways. Table 1.3 shows a scheme of the q^2
 318 spectrum. The ‘high q^2 ’ is the region of low hadron recoil, $q^2 > 15 \text{ GeV}^2/c^4$,
 319 and is characterised by the energy of the hadron being less than the energy scale
 320 of QCD interactions within the meson, $\Lambda_{QCD} \sim 1 \text{ GeV}$. In this region theoretical
 321 calculations of B meson decays can be simplified by working in the heavy quark
 322 limit, $m_b \rightarrow \infty$. In this limit a Heavy Quark Effective Theory (HQET) [23] can
 323 be constructed in which the heavy quark interacts only via ‘soft’ hadronic processes
 324 and an OPE in $1/m_b$ is valid.

325 The ‘low q^2 ’ is instead the region where the light spectator quark is energetic
 326 and cannot be neglected. Furthermore, the light quark interacts not only via ‘soft’
 327 hadronic processes, as in HQET, but also via the so-called ‘collinear’ hadronic pro-
 328 cesses. The boundary of this region can be set at $\sim 7 \text{ GeV}^2/c^4$ as this corresponds
 329 to the threshold for $c\bar{c}$ production, $(2m_c)^2$. In this region the hadronic interactions
 330 are handled by expanding in terms of the energy of the emitted energetic hadron,
 331 $1/E_h$, forming the so-called Soft-Collinear Effective Theory (SCET) [24]. In both
 332 regions decay rates can be predicted using the different methods and the biggest un-
 333 certainties come from the limited knowledge of hadronic transition matrix elements.
 334 The intermediate region is characterised by the presence of charmonium resonances,
 335 produced through tree level $b \rightarrow c\bar{c}s$ transitions and no precise theoretical calculation
 336 is available [25].

337 As can be seen in Fig. 1.8 the very low q^2 is characterised by a peak due to the
 338 virtual photon contribution, associated with C_7 . In the region $1 - 6 \text{ GeV}^2/c^4$ the
 339 interference between C_7 and C_9 becomes large, yielding sensitivity to NP in C_9 .

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

$q^2 = 0$	$E_{K^{*0}} >> \Lambda_{QCD}$	$q^2 \sim m_{J/\psi, \psi(2S)}^2$	$E_{K^{*0}} \sim \Lambda_{QCD}$	$q^2 = (m_B - m_K^{*0})^2$
max. recoil	large recoil (SCET)	$c\bar{c}$ resonances	low recoil (HQET)	zero recoil

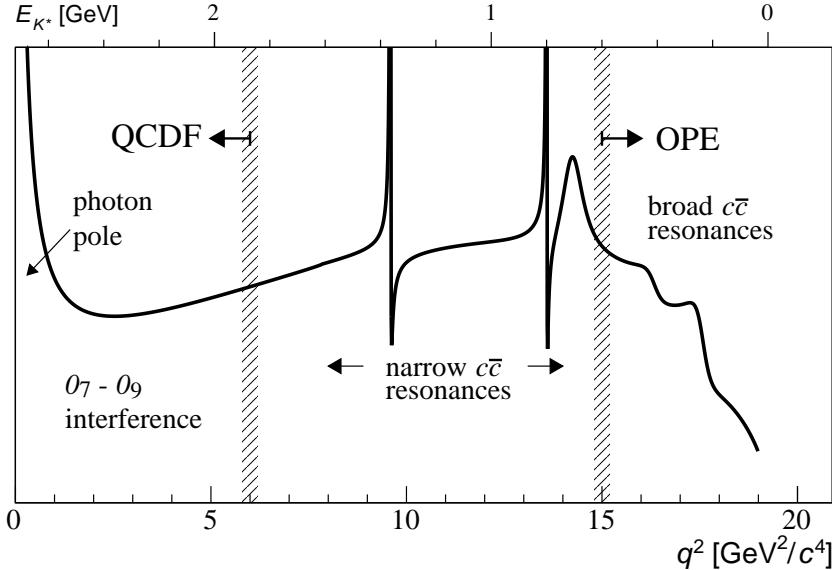


Figure 1.8: A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ process characterised by the photon pole at very low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2 .

340 The $7 - 15$ GeV^2/c^4 interval is dominated by the charmonium resonances, J/ψ and
 341 $\psi(2S)$. Although these decays can be experimentally vetoed in principle charmonia
 342 affect the entire q^2 space. Finally, at high q^2 broad charmonium resonances can
 343 contribute, like those observed by LHCb in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays [26].

344 1.5.4 Observables in $b \rightarrow s\ell^+\ell^-$ decays

345 Rare decays and especially semileptonic $b \rightarrow s\ell^+\ell^-$ processes offer a number of ob-
 346 servables which can be used to benchmark BSM models. The most direct effects
 347 appear in decay rates that can be enhanced by NP but the precision on these mea-
 348 surements is often limited by the uncertainty on form factor calculations. Therefore,
 349 it is important to also look for different observables. One important class of observ-
 350 ables are angular quantities that can often carry complementary information about
 351 NP with respect to branching ratio measurements. The most basic of these observ-
 352 able are forward-backward asymmetries that characterise the angular distribution
 353 of final particles. For the $B^0 \rightarrow K^* \mu^+ \mu^-$ decay combinations of observables have

³⁵⁴ been proposed that are independent of form factor uncertainties at leading order
³⁵⁵ order [17].

³⁵⁶ One way to build safe observables is to construct ratios between similar decays, in
³⁵⁷ which uncertainties due to the hadronization process cancel out. These observables
³⁵⁸ include the R_H ratios, between B^0 decays into electrons and muons, that are de-
³⁵⁹ scribed in detail in Ch. 11. It is also interesting to compare decays which go through
³⁶⁰ the same fundamental process but where the spectator quark has a different flavour.
³⁶¹ This is the case of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays, which are both
³⁶² $b \rightarrow s$ transitions where the spectator quark is an u quark in the first case and a d
³⁶³ quark in the second. The ratio of the branching fractions of these decays is called
³⁶⁴ isospin asymmetry.

³⁶⁵ 1.6 Experimental status

³⁶⁶ To set the background for the searches included in this thesis, this section reports a
³⁶⁷ brief review of recent results of NP searches involving rare decays or lepton flavour
³⁶⁸ violation. Among these, results recently obtained by the LHCb experiment show a
³⁶⁹ series of anomalies with respect to the SM that have the potential to yield to NP
³⁷⁰ scenarios.

³⁷¹ 1.6.1 Dimuon decays of b hadrons

Decays of B mesons into two muons have been recently studied at the LHCb and CMS experiments. These are two-body decays where the two muons are back to back in the hadron rest frame. The simple signatures of these decays make 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 $B^0 \rightarrow \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^-$ decays are exceedingly rare in the SM. First of all they are FCNCs that can only happen in loops and furthermore they

are CKM-suppressed. 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 [27]:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9} \text{ and} \quad (1.16)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}. \quad (1.17)$$

The uncertainties on these values mainly come from 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 [28]:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9} \text{ and} \quad (1.18)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}. \quad (1.19)$$

Both decays were previously unobserved and now the B_s^0 decay is observed 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 to 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.

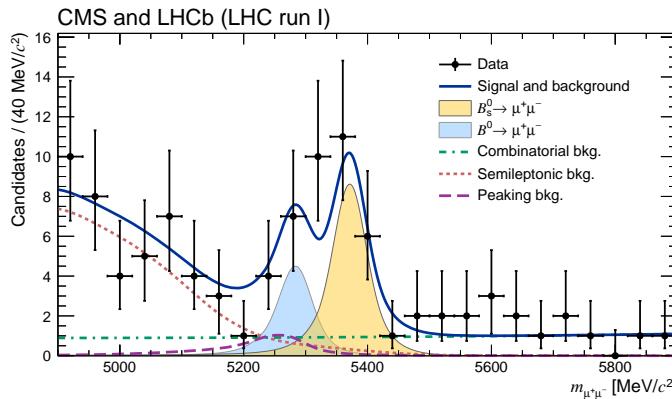


Figure 1.9: Dimuon invariant mass of B candidates showing peaks corresponding $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays [28].

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

³⁷⁸ At the LHC energies is now possible to collect large data sample of semileptonic
³⁷⁹ decays, especially those with a dimuon pair in the final state. Many branching
³⁸⁰ fractions of semileptonic B meson decays were recently measured at the LHCb ex-
³⁸¹ periment, including $B \rightarrow K\mu^+\mu^-$, $B \rightarrow K^{*0}\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ [29, 30, 31].
³⁸² Baryon decays were also studied at LHCb: including the branching fraction of the
³⁸³ rare $\Lambda_b \rightarrow \Lambda\mu^+\mu^-$ decay [1], which is described in this thesis. Unlike for pure lep-
³⁸⁴ tonic decays, SM predictions for semileptonic decays are affected by the knowledge
³⁸⁵ of hadronic form factors, which results in relatively large uncertainties, $\mathcal{O}(30\%)$. As
³⁸⁶ a result measurements are now typically more precise than predictions.

³⁸⁷ As described in Sec. 1.5.4 angular observables can be affected by new physics. Partic-
³⁸⁸ ular interest was risen by the measurement of a series of observables in $B \rightarrow K^{*0}\mu^+\mu^-$
³⁸⁹ decays, free from form factors uncertainties at leading order [32]. Most of the mea-
³⁹⁰ surements are found to be in agreement with SM predictions with the exception
³⁹¹ of the P'_5 observable, shown in Fig. 1.10, which presents a local 3.7σ deviation.
³⁹² Attempts to build a consistent picture point to a NP contribution to the Wilson Co-
³⁹³ efficient C_9 [33]. An angular analysis of $B^+ \rightarrow K^+\mu^+\mu^-$ decays was also performed,
³⁹⁴ where observables are found to be compatible with SM predictions [34].

³⁹⁵ Other observables for which the sensitivity to form factors effects is reduced are the
³⁹⁶ CP asymmetry between B and \bar{B} decays, \mathcal{A}_{CP} , and the isospin asymmetry between
³⁹⁷ B^0 ad B^+ decays, \mathcal{A}_{CP} . Due to the small numerical size of the corresponding CKM
³⁹⁸ elements CP asymmetries of $B^0 \rightarrow K^{(*)}\mu^+\mu^-$ decays are tiny in the SM, $O(10^{-3})$. In
³⁹⁹ BSM models new sources of CP violation can arise and therefore their measurement
⁴⁰⁰ is a powerful null test of the SM. The isospin asymmetry is not zero in the SM due
⁴⁰¹ to isospin breaking effects in the form factors. This is expected to be $\sim 1\%$ at low
⁴⁰² q^2 and grow up to $\sim 10\%$ as q^2 tends to zero. The LHCb experiment, using the

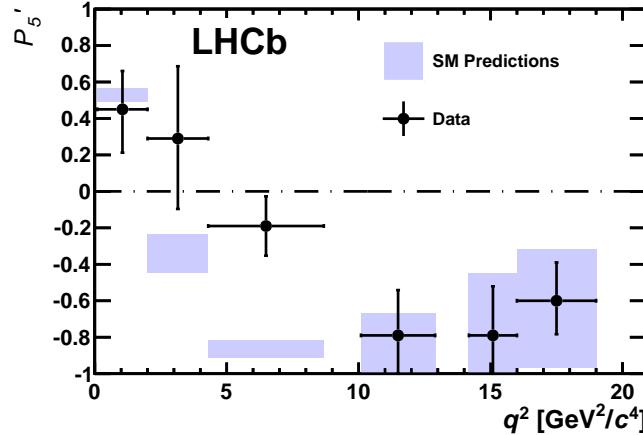


Figure 1.10: Measurement of the observable as a function of q^2 , showing a tension with SM predictions in the 2–6 GeV $^2/c^4$ region.

q^2 [GeV $^2/c^4$]	$B^0 \rightarrow K^+ \mu^+ \mu^-$		$B^0 \rightarrow K^{*0} \mu^+ \mu^-$	
	1.1–6	15.0–22.0	1.1–6	15.0–19.0
\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 $B^0 \rightarrow K^{(*)} \mu^+ \mu^-$ decays from the LHCb experiment.

403 full dataset collected in Run I, corresponding to an integrated luminosity of 3 fb $^{-1}$,
 404 measured both these asymmetries to be consistent with zero [29, 35], as reported in
 405 Tab. 1.4.

406 Recently, progress was made measuring also electron channels. The branching frac-
 407 tion of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay was measured to be $(3.1 \pm 1.3) \times 10^{-7}$ in the
 408 dilepton mass interval 30–1000 MeV/ c^2 [36]. Furthermore, for the first time angu-
 409 lar observables were measured for this decay and found to be consistent with SM
 410 predictions [37].

411 1.6.3 Lepton Flavour Violation searches

412 Several LFV searches are linked to rare decays as they involve small branching
 413 ratios in the SM that can be enhanced by new physics. They are therefore a natural
 414 place to look for NP. Lepton flavour conservation is well experimentally established

⁴¹⁵ measuring the branching ratios of decays of muons into electrons and no neutrinos
⁴¹⁶ but has no strong theoretical explanation in the context of the SM. In fact it is
⁴¹⁷ already observed that flavour is not conserved in neutrino oscillations.

⁴¹⁸ The best-studied decays violating lepton flavour are rare muon decays including
⁴¹⁹ $\mu^+ \rightarrow e^+\gamma$ and $\mu^+ \rightarrow e^+e^-e^+$. Since muons can be abundantly produced and
⁴²⁰ the final states are simple, these decays provide the best constraints to LFV. The
⁴²¹ present best upper limits are 1.2×10^{-11} for the radiative decay and 1.0×10^{-12}
⁴²² for $\mu^+ \rightarrow e^+e^-e^+$ obtained respectively by the MEGA [38] and SINDRUM [39] ex-
⁴²³ periments. Several LFV searches in the B sector have recently been performed at
⁴²⁴ the LHCb experiment including decays such as $B^0 \rightarrow e\mu$ [40] and τ decays such
⁴²⁵ as $\tau \rightarrow \mu^+\mu^-\mu^-$ [41]. None of these searches has found evidence of NP so far and
⁴²⁶ therefore they set limits, constraining the parameter space available for NP mod-
⁴²⁷ els. Figure 1.11 shows a summary of the best limits set at different times on LFV
⁴²⁸ searches [42].

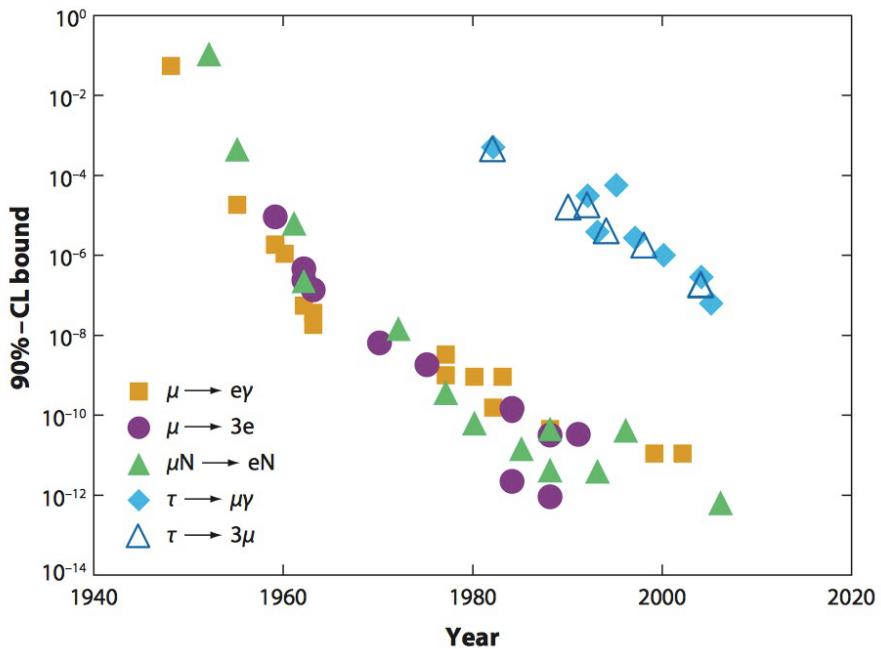


Figure 1.11: Summary of limits set in LFV searches as a function of time [42].

429

CHAPTER 2

430

431

The LHCb detector at the Large Hadron Collider

432

433

2.1 The Large Hadron Collider

434 The Large Hadron Collider (LHC) [43] is a circular particle accelerator with a cir-
435 cumference of 27 km located about 100 m underground at CERN in the surroundings
436 of Geneva, Switzerland. Two proton beams circulate in opposite directions around
437 the ring and cross each other at several points, in which particle detectors are placed.
438 These include two general-purpose detectors, ATLAS and CMS, siting on opposites
439 sides of the ring and the two smaller specialty detectors, ALICE and LHCb, are at
440 the interaction points to either side of ATLAS (see Fig. 2.1).

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

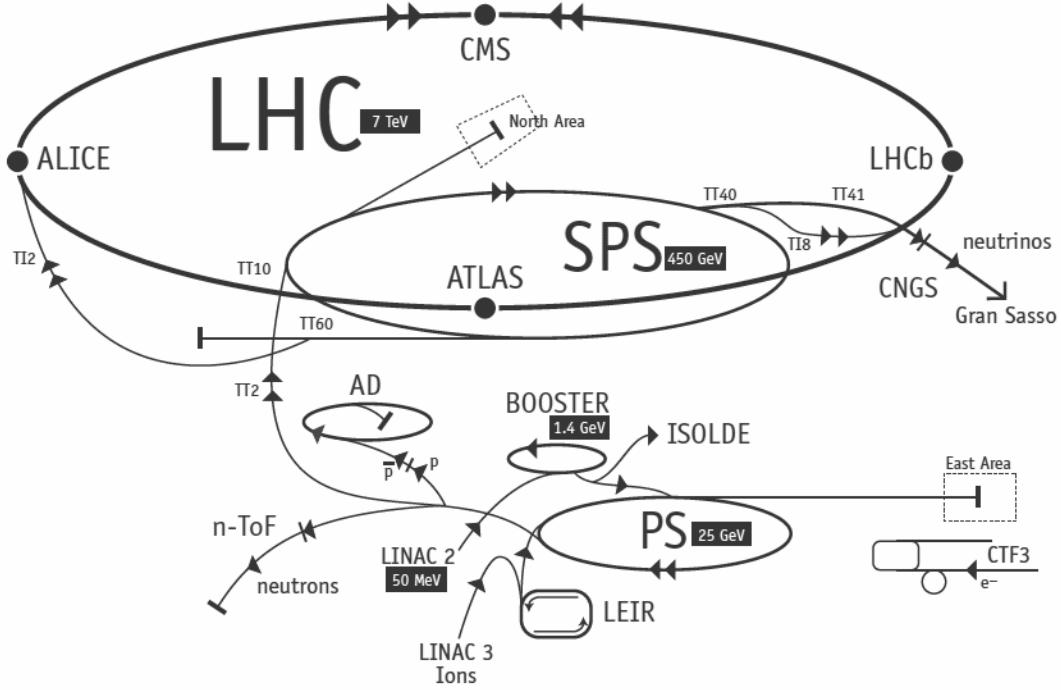


Figure 2.1: Scheme of CERN accelerators.

445 Protons are produced from hydrogen gas and are initially accelerated to the energy
 446 of 50 MeV in a linear accelerator (LINAC). Then they are injected into the Proton
 447 Synchrotron Booster (PSB), where they are boosted to an energy of 1.4 GeV, into
 448 the Proton Synchrotron (PS) to 25 GeV and into the Super Proton Synchrotron
 449 (SPS) to 450 GeV. Finally, protons enter into the LHC storage ring, where they are
 450 accelerated from injection energy to the final one by radio frequency (RF) cavities.
 451 The beams are steered around the ring by 8 T magnetic fields produced by 15 m
 452 long superconducting niobium-titanium dipole magnets and focused by quadrupole
 453 magnets. The LHC magnets use a design in which both proton beam pipes are
 454 contained in the same housing, allowing the same liquid helium to cool the system
 455 down for both. The LHC began colliding proton beams in physics mode in 2009
 456 at a center of mass energy of $\sqrt{s} = 900$ GeV and from April 2010 to November
 457 2011 accelerated beams at $\sqrt{s} = 7$ TeV (3.5 TeV per proton beam). At this energy
 458 it delivered over 5.7 fb^{-1} of collisions, with a maximum instantaneous luminosity
 459 of $3 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. The LHC maximum design energy is 14 TeV, and its design
 460 luminosity is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. After a long shut down to upgrade and maintain the

461 machine, a new run started in June 2015, in which protons are collided at a center
462 of mass energy of $\sqrt{s} = 13$ TeV. At this energy the total proton-proton cross section
463 is expected to be roughly 100 mb.

464

2.2 The LHCb detector

465 The LHCb detector [44] was built with the main purpose of studying the decays of B
466 and D mesons, looking in particular for CP-violating processes. In 2011, running at
467 a centre of mass energy of 7 TeV, the cross section for $b\bar{b}$ production was measured
468 to be $284 \pm 53 \mu b$ [45], while it will be $\sim 500 \mu b$ at the current LHC energy, 13 TeV.
469 At these high energies, proton-proton interactions produce highly boosted virtual
470 gluons which interact to produce $b\bar{b}$ pairs at small angles, close to the beam pipe. For
471 this reason the LHCb detector is designed to have a very forward angular coverage:
472 it is fully instrumented from approximately 10 mrad to 300 mrad, corresponding
473 to $2 < \eta < 5$, where η is the “pseudorapidity”, a quantity used in particle physics
474 defined as:

$$\eta = -\ln(\tan(\theta/2)), \quad (2.1)$$

475 where θ is the angle between a particle’s momentum and the beam direction ¹.

476 At the collision point of LHCb the luminosity can be adjusted by displacing the
477 beams from head on collisions while keeping the same crossing angle. This allows the
478 experiment to keep an approximately constant instantaneous luminosity. This also
479 means that the average number of interactions per bunch crossing can be limited as
480 the detector efficiency, especially in detecting secondary vertices, decreases for events
481 with an high number of primary vertices (PV). Reducing the particle occupancy
482 through the detector also keeps radiation damage to a minimum. Since the LHC
483 started colliding protons in November 2009 until the end of 2011, the instantaneous
484 luminosity was at an average of $3 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, with an average number of 1.5 PV

¹LHCb’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.

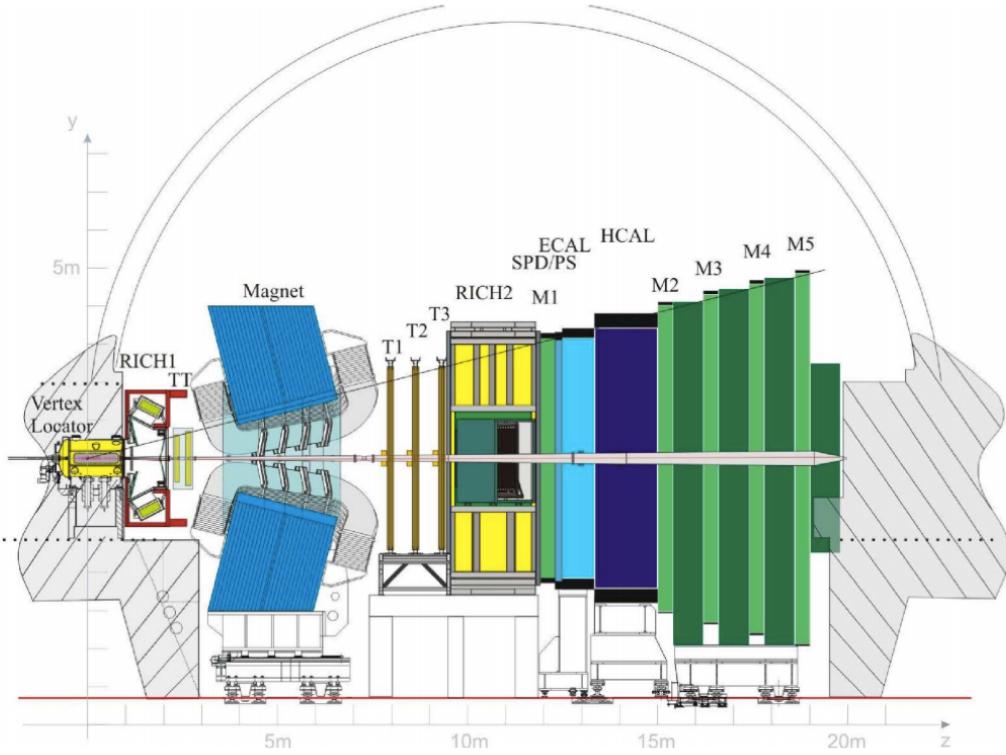


Figure 2.2: A side view of the LHCb detector [44].

per bunch crossing. At the end of 2011 LHCb had collected an integrated luminosity of 1 fb^{-1} ; in 2012 the luminosity was increased and 2 fb^{-1} more were collected.

Other B physics experiments, like BaBar at the Stanford Linear Accelerator (SLAC), Belle at KEK at J-PARC (Japan) and the Tevatron experiments at Fermilab have made accurate measurements in heavy flavour physics. All of these results have so far been consistent with the Standard Model predictions. However, some of the deviations from the Standard Model are expected to be very small, therefore LHCb has begun to make the most precise measurements in heavy flavour physics to test the consistency of the Standard Model 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, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream. Charged hadrons are identified using two Ring-Imaging Cherenkov detectors

499 (RICH) [46]. Photon, electron and hadron candidates are identified by a calorimeter
500 system consisting of a scintillating-pad, a pre-shower detector, an electromagnetic
501 calorimeter and a hadronic calorimeter. Muons are identified by a system composed
502 of alternating layers of iron and multi-wire proportional chambers [47]. A schematic
503 view of the detector is shown in Fig. 2.2 and more details on each sub-detector are
504 given in the following sections.

505 2.3 The magnet

506 Charged particle are bent horizontally in the magnetic field so that their momentum
507 can be measured from the curvature radius. The LHCb dipole magnet is comprised
508 of two coils supported by an iron yoke and is shaped to fit the LHCb angular
509 acceptance. Unlike the other LHC experiments, LHCb uses a warm magnet, so that
510 it can be ramped easily and the field can be reversed periodically. When the polarity
511 is flipped particles of a given sign are bent in the opposite direction. This method is
512 used to limit systematic uncertainties that can arise form different performances in
513 different areas of the detector and average out using data taken in both polarities.
514 In the magnet flows a 5.85 kA current that generates an integrated magnetic field
515 of 4 Tm for 10 m long tracks. In order to achieve the required momentum precision
516 the magnetic field must be mapped with a 10^{-4} precision. For this reason a grid
517 of 60 sensors is positioned inside the magnet and provides real time magnetic field
518 maps.

519 2.4 Tracking system

520 B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means
521 they travel about 1 cm before decaying and they form a displaced vertex. To study
522 specific decays, it is therefore important to be able to separate the particles pro-
523 duced at the primary pp vertex and at the B decay vertex. The tracking system

524 consists in the Vertex Locator (VeLo), and 4 tracking stations: the Tracker Turi-
 525 censis (TT), which are located before the magnet and the T1, T2 and T3 stations,
 526 located downstream of the magnet. The latter tree stations are in turn formed by
 527 two subsystems: the Inner Tracker (IT) close to the beam-line, where the particle
 528 density is greatest and the Outer Tracker (OT) covering the rest of the acceptance.

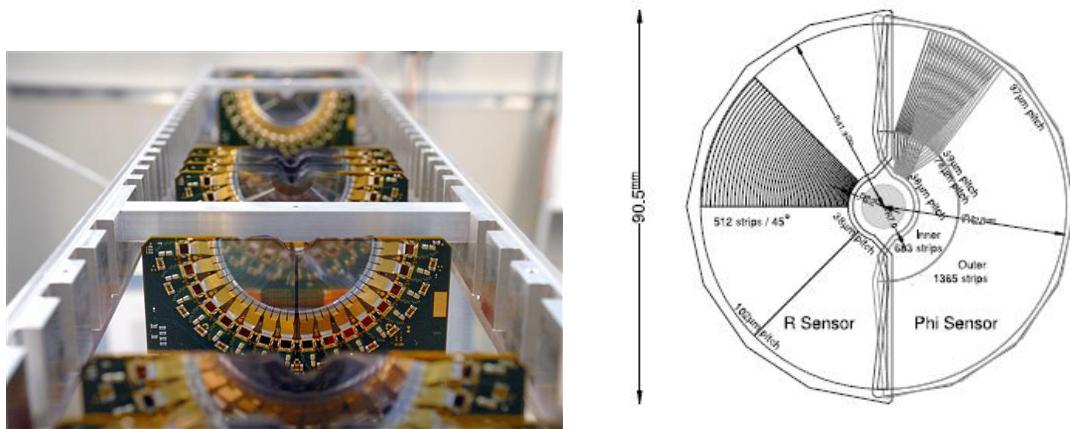


Figure 2.3: On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [44].

529

530 The VeLo accurately measures positions of tracks close to the interaction point so
 531 that production and decay vertices of bottom and charm hadrons can be recon-
 532 structed. The VeLo is composed by 21 staggered silicon modules which surround
 533 the beam axis and are positioned from $z = -18$ cm to $+80$ cm. It is able to detect
 534 particles within a pseudorapidity range $1.6 < \eta < 4.9$. The sensitive region of the
 535 VeLo starts at an inner diameter of only 8 mm from the beam axis. The VeLo is
 536 housed in its own vacuum vessel of thin aluminium foil, which protects the vacuum
 537 of the beam pipe from any outgassing. The silicon layers composing the VeLo consist
 538 of two modules each including two types of sensors: the ϕ -sensor, which measures
 539 the azimuthal position around the beam, and the R-sensor, which measures the ra-
 540 dial distance from the beam axis. A sketch of the VeLo sensor is shown in Fig. 2.3.
 541 The sensors are $300 \mu\text{m}$ thick, approximately semicircular and are positioned on
 542 either side of the beam axis. To ensure that they cover the full azimuthal angle the

543 right-side module is placed 1.5 cm behind the left-side module on the z-axis and
544 they overlap. There are two modules which cover the backward direction and are
545 used as a veto for multiple interactions, this is called the pileup system.

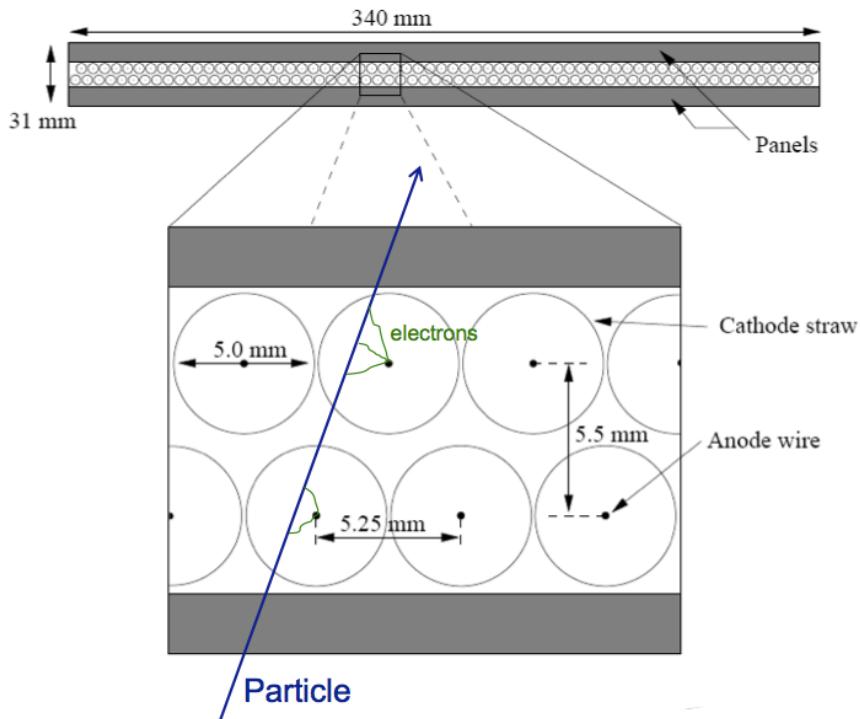


Figure 2.4: Sketch of the straw tubes which constitute the Outer Tracker layers [44].

546

547 The IT and TT both use silicon microstrips and together constitute the Silicon
548 Tracker (ST). Straw tubes are used in the OT, of which a sketch is shown in Fig. 2.4.
549 The IT requires an higher inner granularity because of the higher flux of particles
550 nearer the beam pipe, in fact it covers only 1.3% of the total area of IT plus OT
551 but it contains about 20% of the tracks. Each ST station has four detection layers,
552 the first and last being vertical, measuring the track position in x. The second
553 and third layer are rotated by an angle of +5 and -5 degrees, which allows the y-
554 coordinate to be measured. The TT is placed upstream of the magnet to allow the
555 reconstruction of tracks from low-momentum particles, which are swept out of the
556 downstream acceptance. Overall the tracking system provides a measurement of
557 momentum, p , with a relative uncertainty that varies from 0.4% at 5 GeV/c to 1.0%

558 at 200 GeV/ c . The impact parameter (IP), namely the minimum distance of a track
559 to a primary vertex, is measured with a resolution of $(15 + 29/p_T) \mu m$, where p_T
560 is the component of the momentum transverse to the beam, in GeV/ c . The z-axis
561 position of a PV reconstructed with 35–40 tracks can be measured to a precision
562 of roughly 50–60 μm . Accurate IP and vertex displacement measurements allow
563 LHCb to effectively distinguish between B meson decays and background processes.
564 In fact B mesons typically travel ~ 1 cm in the detector before decaying into lighter
565 particles, which tend to have high IP as the B decay imparts transverse momentum
566 to them.

567 2.5 Calorimeters

568 The main purpose of the calorimeter system is to determine the energy of particles
569 traversing the detector. A calorimeter is composed by layers of absorber and active
570 material. The absorber lets particles interact and produce a cascade of secondaries,
571 which multiply quickly and are detected by the active part. In LHCb the sensitive
572 material consists of scintillating layers, where the light detected is approximately
573 proportional to the number of deposited particles. Calibration is then used to trans-
574 late the signal into a measurement of deposited energy. The LHCb calorimeter sys-
575 tem consists of the Scintillator Pad Detector (SPD), the Pre-Shower Detector (PS)
576 as well as the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter
577 (HCAL). A sketch of the LHCb calorimeters is shown in Fig. 2.5 The information
578 from the calorimeter is also used to identify electron and hadrons. The most difficult
579 identification is that of electrons. The rejection of a high background of charged pi-
580 ons requires a longitudinal segmentation of the electromagnetic calorimeters. This
581 is provided by the PS detector added in front of the main electromagnetic calorime-
582 ter, ECAL. Electrons also have to be distinguished from high energy π^0 s. For this
583 purpose the SPD calorimeter, detecting charged particles, is located in front of the
584 PS and ECAL detectors. Figure 2.5 shows how the ratio of the energy detected
585 in the ECAL and the particle momentum allows the separation of electrons and

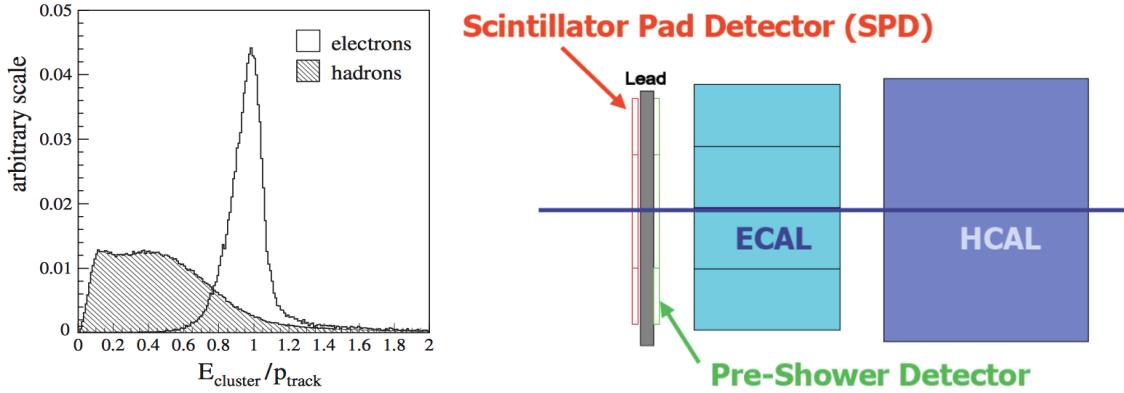


Figure 2.5: The ratio of the energy deposited in the ECAL and the particle momentum, which allows the separation between electrons and hadrons.

hadrons. All four detectors transmit scintillation light via wavelength-shifting fibres to photo-multiplier tubes (PMTs). The SPD/PS cells are read out with MAPMTs (Multi-anode PMTs) located outside the LHCb acceptance. Instead the ECAL and HCAL have individual MAPMTs located on the modules. All four detectors are segmented, which allows to associate the energy deposits to tracks in the tracking system. The segmentation of the cells varies according to the distance from the beam pipe.

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 [44]

$$\frac{\sigma_{\text{ECAL}}(E)}{E} = \frac{10\%}{\sqrt{E(\text{GeV})}} + 1\%, \quad (2.2)$$

which results in a mass resolution of $\sim 70 \text{ MeV}/c^2$ for B mesons for $\sim 8 \text{ GeV}/c^2$ for π^0 . The trigger requirements on the HCAL resolution do not depend on the containment of the hadron showers as much as for the ECAL, so, due to a limited space, its thickness is only 5.6 interaction lengths and its resolution

$$\frac{\sigma_{\text{HCAL}}(E)}{E} = \frac{69\%}{\sqrt{E(\text{GeV})}} + 9\%. \quad (2.3)$$

600 2.5.1 Bremsstrahlung recovery for electrons

601 Bremsstrahlung is an electromagnetic radiation produced by charged particles that
 602 decelerate or deviate. Typically electrons produce Bremsstrahlung when deflected by
 603 atomic nuclei. The probability of emitting bremsstrahlung radiation is proportional
 604 to the inverse of the squared mass of the particle ($1/m^2$) and therefore it is relevant
 only for electrons. At LHC energies, if electrons radiate after the magnet, the

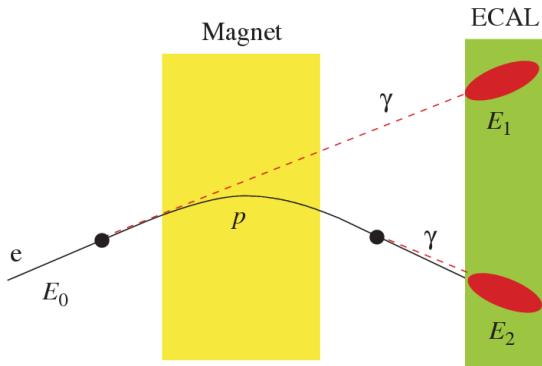


Figure 2.6: Schematic view of the bremsstrahlung recovery.

605
 606 photon will hit in the same calorimeter cells as the electron and the energy will be
 607 automatically recovered. However, if the photon is emitted before the magnet, the
 608 electron will be deflected by the magnetic field whereas the photon will continue
 609 on its initial trajectory, with its energy being deposited in a different part of the
 610 calorimeter. Missing this energy results in a poorer reconstructed invariant mass
 611 resolution, so it is desirable to recover these bremsstrahlung photons, when possible.
 612 A tool for bremsstrahlung recovery is available in the LHCb analysis software. This
 613 tool looks for other clusters in the calorimeter and, reconstructing the trajectory
 614 of the electron, checks if they may be associate with photons emitted by that.
 615 The photon energy is then added to the electron and its momentum recalculated.
 616 Figure 2.6 displays a schematic view of the process. For more information see
 617 Ref. [48].

618 2.6 RICH

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

$$\cos \theta = \frac{1}{\beta n}, \quad (2.4)$$

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

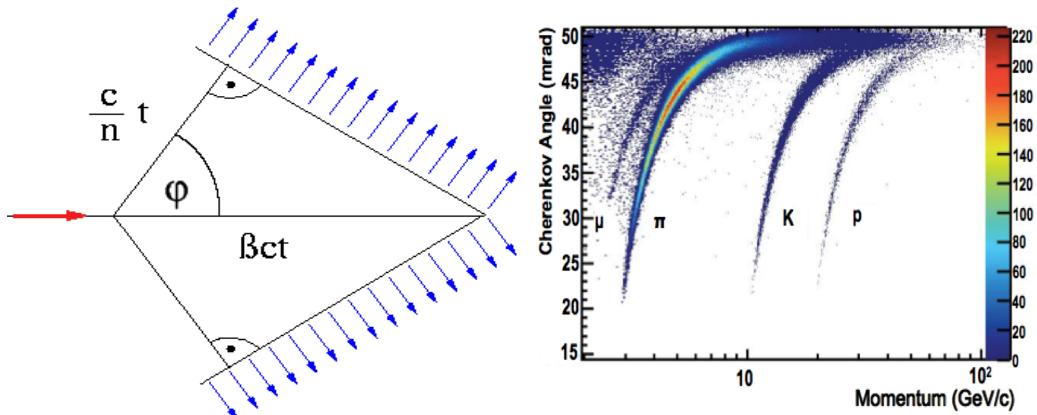


Figure 2.7: On the left a sketch of Cherenkov light emission and on the right the Cherenkov angle versus the particle momentum, where one can see that the study of the Cherenkov angle allows distinguish particles identities.

625

626 RICH 1 is situated before the magnet in order to cover a large angular accep-
 627 tance. Its purpose is to ensure particle identification over the momentum range
 628 $1 < p < 70 \text{ GeV}/c$. It uses two radiators: C_4F_{10} that covers the momentum range
 629 $5 - 70 \text{ GeV}/c$ and silica aerogel which covers $1 - 10 \text{ GeV}/c$. RICH 2 is positioned
 630 after the magnet and tracking stations. It identifies higher momentum particles
 631 from approximately $20 \text{ GeV}/c$ up to beyond $100 \text{ GeV}/c$ using CF_4 as a radiator.
 632 The Cherenkov light produced when charged particles travel through the radiators,
 633 is reflected and focused using flat and spherical mirrors, which are tilted so that the

ring image is reflected onto arrays of photo-detectors. The radius of the ring can be used to measure the opening angle of the Cherenkov cone because of the known geometry. The photo-detectors 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 the final state. 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 is formed by five stations (M1 - M5), the first one being located before the calorimeters in order to improve p_T measurements. A scheme of the muon system is shown in Fig. 2.8. The remaining four stations lay behind the HCAL and are separated by 1.2 m from each other, interleaved with 80 cm thick iron blocks, which absorb hadrons, electrons and photons to ensure that only muons reach the final muon station. 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 provide position measurements. Since there is a larger particle flux towards the beam pipe, the stations are divided into four concentric rectangular regions (R1-R4), their size increasing according to the ratio 1 : 2 : 4 : 8. This results in a similar channel 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 because they have to withstand a rate up to 500 kHz cm⁻² of charged particles. In these detectors particles traversing through the drift gap between the cathode and the first GEM foil produce ionisation electrons, which are then attracted by electric fields though all of the GEM foils and multiply. They

662 then drift into the anode inducing a signal on the pads. A gas mixture of Argon,
 CO_2 and CF_4 , is used to give a time resolution better than 3 ns.

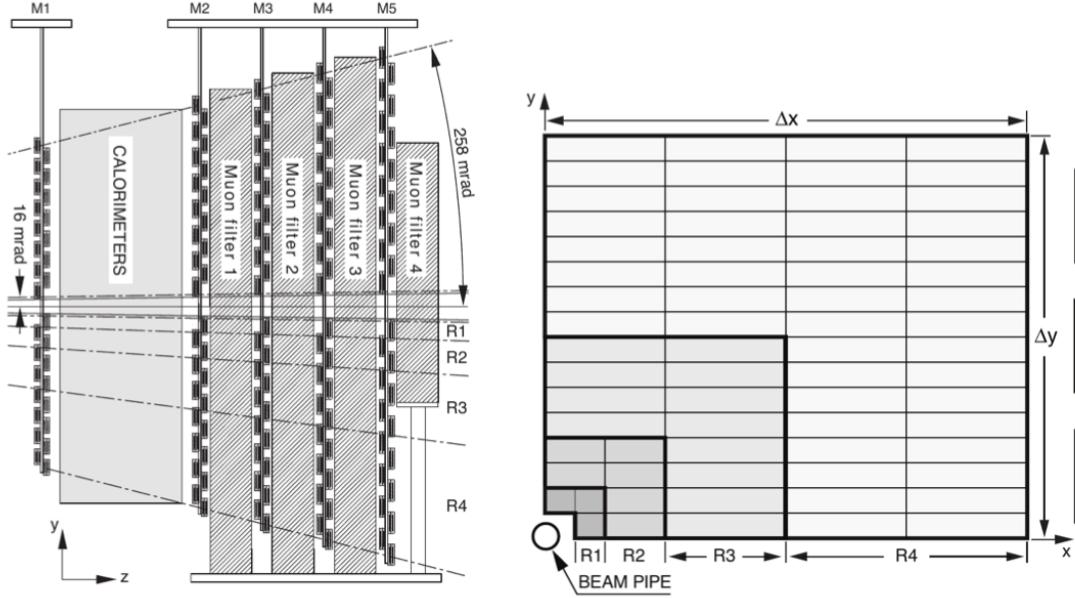


Figure 2.8: The LHCb muon system [44].

663

664 2.8 Particle identification

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

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

676 To quantify the quality of the ID the pion hypothesis is used as a reference point
 677 and the probability of a specific ID is given in terms of Log-Likelihood difference
 678 between the given ID hypothesis and the pion one. This variable is called Delta
 679 Log-Likelihood (DLL) and denoted with “PID”. For example:

$$\text{PID}_K = \text{DLL}_{K-\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi) \quad (2.5)$$

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

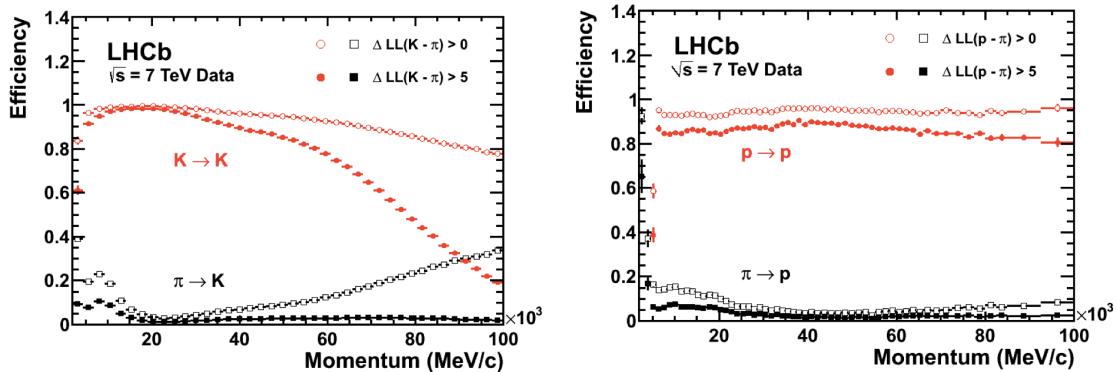


Figure 2.9: Particle Identification performances for kaons (left) and protons (right) as a function of the measured momentum of the particles.

688
 689 muons, which is quantified using two different variables: the DLL_μ and in addition
 690 the `isMuon` variable. This is a boolean variable determined by defining a ‘field of
 691 interest’ around an extrapolated track trajectory through the muon chambers. The
 692 variable is set to true if hits in multiple muon stations are found in the field of
 693 interest.

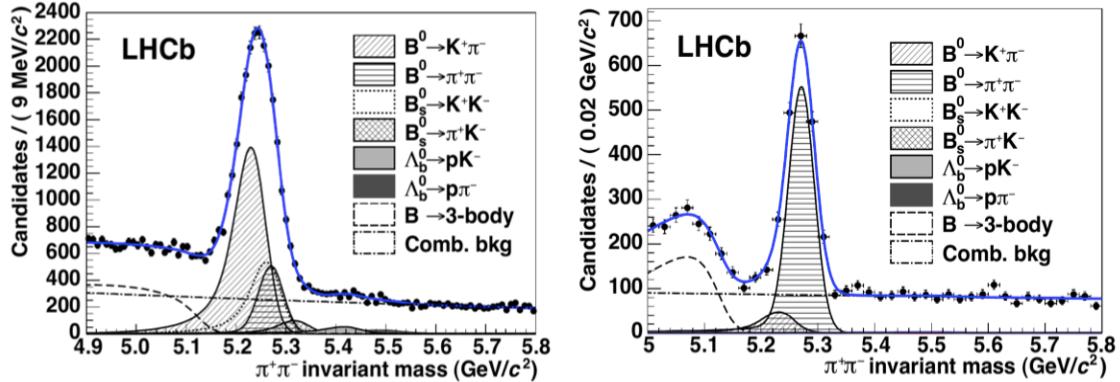


Figure 2.10: Invariant mass peak of the $B^0 \rightarrow \pi^+\pi^-$ decay before (left) and after (right) the application of PID requirements.

2.8.1 PID calibration

In order to be able to calculate efficiencies a data-driven method was developed. The calibration software is referred to as `PIDCalib` package [49]. This tool uses decays where final particles can be identified thanks to their kinematic properties. For example the $K_s^0 \rightarrow \pi^+\pi^-$ decay has a clear signature with a displaced vertex and can be easily singled out from other decays and used to test pion PID efficiency. The narrow peaks of the $J/\psi \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow e^+e^-$ decays allow to calibrate muon and electron efficiencies. 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, $\phi \rightarrow KK$ samples and $D^{*+} \rightarrow D(\rightarrow K^-\pi^+)\pi^+$ decays, where the D^{*+} is used to tag the decay, are used to test the kaon efficiency. Residual background is subtracted using the $\mathcal{P}\text{Plot}$ technique [50].

2.9 Trigger and software

The LHCb trigger system [51] 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 event reconstruction. To increase performances

711 the HLT is split again into stages, HLT1 and HLT2. The HLT1 phase happens in
 712 real time and saves data in local disks while the HLT2 phase uses the resources
 713 available during periods with no beam. The event selected by the HLT2 stage are
 714 then saved for offline analysis. The bunch crossing frequency is 40 MHz, which
 715 corresponds to an instantaneous luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ for LHCb. About
 716 15% of the total number of $b\bar{b}$ pairs produced will contain at least one B meson
 717 with all of its decay products within the detector acceptance. This rate needs to be
 718 reduced down to about 2 kHz so that the events can be written to disk. Figure 2.11
 shows a scheme of the trigger system.

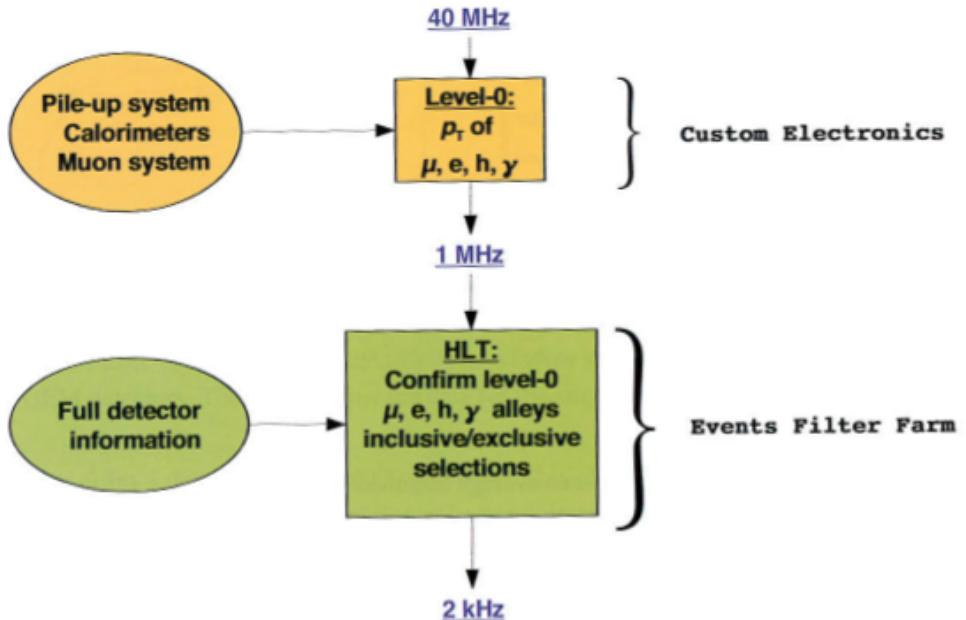


Figure 2.11: Scheme of the LHCb trigger system [44].

719

720 The L0 trigger reduces the rate of visible interactions from 10 MHz to 1 MHz.
 721 Due to the heavy mass of B mesons, they often produce particles with high p_T or
 722 E_T . Therefore the trigger selects events with large E_T deposits in the calorime-
 723 ter or high p_T muons. The event is classified as L0Muon if it was triggered due to
 724 information for the muon detector, while the information from the calorimeters is
 725 combined to divide the events in 5 categories: L0Photon, L0Electron, L0LocalPion,
 726 L0GlobalPion, L0Hadron. The PS detector information is converted to a photon
 727 flag (PS && !SPD) or an electron flag (PS && SPD). The “local” label of the L0Pion

728 trigger refers to π^0 reconstructed through their $\gamma\gamma$ decay, where the two photons fall
729 in the same ECAL element, they are labelled “global” otherwise. The first four
730 calorimeter triggers require energy clusters in the ECAL, while L0Hadron requires
731 clusters also in the HCAL. The HLT1 uses information from the VELO and trackers
732 performing a partial reconstruction of the event and reduces the rate to 2 kHz by
733 adding requirements of the IP and χ^2 of tracks. Finally, the HLT2 involves a full
734 reconstruction of the event and includes many “lines” designed to trigger specific
735 decay structures.

736 LHCb also developed an extended simulation software in order to reconstruct ef-
737 ficiencies and signal shapes. In the simulation, pp collisions are generated using
738 PYTHIA8 [52, 53] with a specific LHCb configuration [54]. Decays of hadronic par-
739 ticles are described by EVTGEN [55], and final state radiation is generated using
740 PHOTOS [56]. Finally, the interaction of the generated particles with the detec-
741 tor and its response are implemented using the GEANT4 toolkit [57] as described
742 in Ref. [58]. For this analysis in this thesis, the ROOT framework [59] is used to
743 analyse data and the RooFit package to perform maximum likelihood fits. A multi-
744 variate analysis is also performed based on the NeuroBayes package [60, 61], which
745 provides a framework for neural network training.

746 2.10 Kinematical fits

747 The resolution of key variables, such as the measured invariant mass of decaying
748 particles, can be improved by imposing constraints to the measured quantities to
749 remove redundant degrees of freedom. The four-momentum conservation can be
750 ensured at each vertex and the origin and decay vertices of a particle are related via
751 the momentum of the particle. Furthermore, additional constraints can be imposed
752 due to a particular decay hypothesis such as the known invariant masses of final and
753 intermediate particles. In order to do this the `DecayTreeFitter` tool was developed
754 by the BaBar experiment and later used by LHCb [62]. The algorithm takes a

complete decay chain and parameterises it in terms of vertex positions, decay lengths and momentum parameters. Then these parameters are fit simultaneously, taking into account the relevant constraints, including the information from photons. To perform the fit efficiently a Kalman filter is used. 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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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 likelihood of 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.

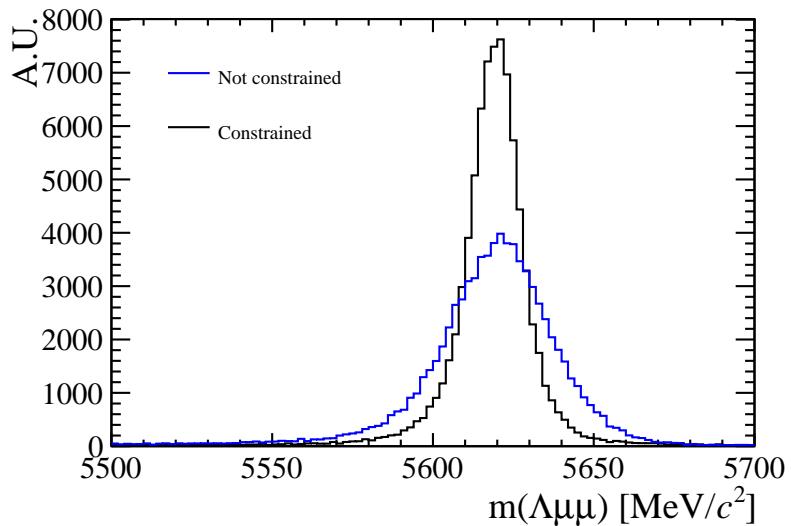


Figure 2.12: Invariant mass of the final daughters of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays calculated with and without constraints using the `DecayTreeFitter` tool.

765

Part I

766

Differential branching fraction and angular
analysis of the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay

767

768

CHAPTER 3

769

770

The rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay

771

772 The rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay is a FCNC process governed by the $b \rightarrow s\mu^+\mu^-$ quark
773 level transition. In the SM this decay proceeds through electroweak penguin and W
774 box diagrams (see Fig. 1.5). Since this process happens only through loop diagrams,
775 it is highly sensitive to new particles entering the loops. Moreover, as the final state
776 contains only a single long-lived hadron, the hadronic physics is easier to handle
777 than in fully hadronic decays.

778 Interest in Λ_b^0 baryon decays arises from two important facts. First of all, Λ_b^0 has
779 non-zero initial spin, thus is a potential to learn information about the helicity
780 structure of the underlying Hamiltonian, which cannot be extracted from the meson
781 decays [63, 64]. Secondly, the Λ_b^0 baryon can be considered in first approximation
782 as composed of an heavy quark and a light di-quark, therefore the hadronic physics
783 significantly differs from similar meson decays. This itself provides the possibility
784 to better understand and test the hadronic physics in the theory, which could yield
785 improved understanding and confidence also for the meson case.

786 With respect to B^0 decays going through the same transitions, such as $B^0 \rightarrow K^{*0}\mu^+\mu^-$,
 787 Λ_b^0 decays can provide independent confirmations of the results as they involve
 788 the same operators but different hadronic matrix elements. Furthermore, Λ de-
 789 cays weakly, which results in complementary constraints with respect to B^0 decays.
 790 Finally, the narrow width approximation, used in theoretical calculations is fully
 791 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
 792 $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays because the contribution from the non resonant channel
 793 $B^0 \rightarrow K\pi\mu^+\mu^-$ is unconstrained.

 794 Theoretical aspects of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays were considered by a number of
 795 authors both in the SM and in different new physics scenarios [65, 66, 67, 68, 69, 70,
 796 71, 72, 73, 74, 75]. All authors start from the same effective Hamiltonian already
 797 described in Sec. 1.5.1. However, form factors, describing hadronic physics, are
 798 not developed as well as for the meson case, partly because there are not as many
 799 experimental constraints, and this leads to a relatively large spread in predicted
 800 branching fractions. An interesting quantity to study is the differential branching
 801 fraction as function of q^2 . This still suffers from the knowledge of form factors,
 802 but, as different approaches to form factors calculations are applicable in different
 803 q^2 regions, this allows a more meaningful comparison to theory.

Experimentally, the decay $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ was observed for the first time in 2011 by
 the CDF collaboration [76], with a signal yield of 24 ± 5 events. Later this result
 was updated using the full CDF statistics [77]. Their result on full statistics yields

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+\mu^-) = [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}.$$
 CDF observed the signal only in the q^2 region above the square of the $\psi(2S)$ mass. Recently, the decay
 was observed also at LHCb [78] 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 $m_\psi^2(2S)$. The LHCb result for the branching fraction relative
 to the $J/\psi\Lambda$ decay, used as normalisation channel, is

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

and for absolute branching fraction

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

This part of the thesis describes the measurement of the differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay using 3 fb^{-1} of pp collisions collected by the LHCb experiment in 2011 and 2012. Furthermore, 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 $b \rightarrow s\ell^+\ell^-$ 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 . For this analysis two regions are defined: the “low q^2 ” region, below the J/ψ resonance ($q^2 < 8 \text{ GeV}^2/c^4$), where the signal is unobserved, and the “high q^2 ” region, above the J/ψ resonance ($q^2 > 11 \text{ GeV}^2/c^4$). The decay $\Lambda_b^0 \rightarrow J/\psi \Lambda$, where J/ψ decays into two muons, which has the same final state as the signal, is used as a normalisation channel and the branching fraction measurement is given in relative form to limit systematic uncertainties. In both cases the Λ decay mode into a pion and a proton, $\Lambda \rightarrow p\pi$, is used to reconstruct the decays. The rare and normalisation channels are naturally distinguished by the q^2 interval they fall into. The regions in which the rare channel is studied include:

- $0.1 < q^2 < 8 \text{ GeV}^2/c^4$, where the selection is optimised to observe the signal as explained in Sec. 4.3. The upper bound of this interval was chosen to be sufficiently far from the J/ψ radiative tail at low masses, that could contaminate the rare sample;
- $11 < q^2 < 12.5 \text{ GeV}^2/c^4$ in between two charmonium resonances and

826 $q^2 > 15 \text{ GeV}^2/c^4$, above $\psi(2S)$. In the latter two intervals the selection is
827 optimised to maximise the yield which is particularly important for a stable
828 angular analysis.

829 The above regions are then divided in smaller intervals, as much as the available
830 statistics allows, which results in $\sim 2 \text{ GeV}^2/c^4$ wide bins. The binning used is the
831 following

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

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

- 833 • 1.1-6.0 GeV^2/c^4 : this interval is theoretically clean since it is far from the
834 photon pole, which dominates at low q^2 , washing out the sensitivity to NP
835 contributions. The lower bound of this interval is chosen to exclude the possi-
836 ble contribution from the ϕ resonance, which appears at $1 \text{ GeV}^2/c^4$. The
837 upper bound of the interval is chosen to totally exclude a small contribution
838 from the J/ψ resonance that leaks below $8 \text{ GeV}^2/c^4$.
- 839 • 15.0-20.0 GeV^2/c^4 : this interval is the one that contains most of the statistics
840 and it is used as a natural cross check that the analysis in smaller bins is
841 stable.

842 3.2 Candidate types

843 This analysis deals with Λ baryons, which have a lifetime of $(2.632 \pm 0.020) \times 10^{-10} \text{ s}$ [2].
844 These are considered long-lived particles in particle physics terms and can travel into
845 the detector for several meters generating well distinguished secondary vertices. In
846 LHCb Λ baryons can be reconstructed from tracks with or without hits in the VELO
847 (see Sec. 2.4) and therefore define two candidate types as follows:

- 848 • **Long candidates:** built from tracks which have hits in the VELO, “long
849 tracks”. These candidates, also denoted as “LL”, are characterised by a better
850 momentum resolution thanks to the longer leverage arm available to their
851 tracks.
- 852 • **Downstream candidates:** built from tracks without hits in the VELO,
853 “downstream tracks”, also denoted as “DD”.

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

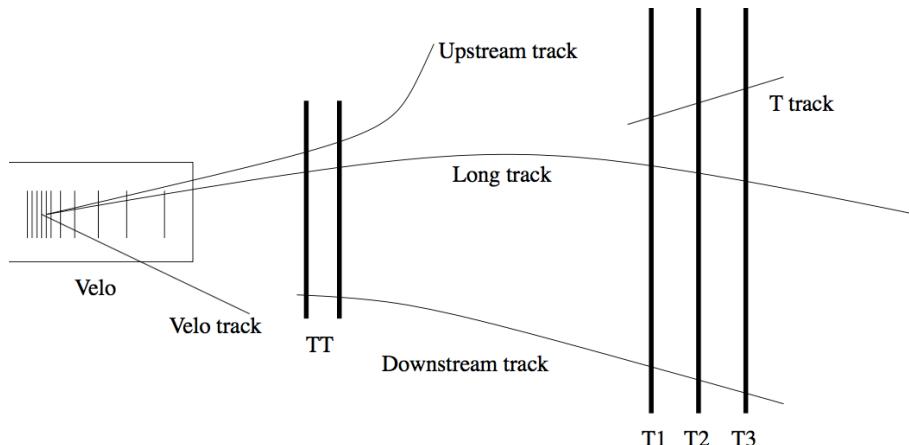


Figure 3.1: Representation of possible track types in LHCb. Candidates built from “long” and “downstream” tracks are used in this analysis.

859 3.3 Simulation

860 Samples of simulated events are needed in order to train the multivariate classifier
861 (see Sec. 4.2), calculate the selection efficiency and study possible background; in
862 particular for this analysis samples of ~ 2 millions $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and ~ 5 millions
863 $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulated events are used. Samples of simulated $B^0 \rightarrow J/\psi K_s^0$,

864 $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ and $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ events are also used to study backgrounds
865 from these decays. The events are generated using PYTHIA8, hadronic particle are
866 decayed using EVTGEN and GEANT4 is used to simulate the interaction of final
867 state particles with the detector. Simulated events are then reconstructed using the
868 same reconstruction version used for real data. The L0 hardware trigger is emulated
869 in the simulation, while for the software stage, Hlt1 and Hlt2 (see Sec. 2.9), the same
870 code can be used as for real data. Events are simulated using both 2011 and
871 2012 conditions in the same amount in which data is available. It is important that
872 the simulation gives an accurate description of the data, especially for quantitative
873 estimations, as the extraction of efficiencies. While the simulation gives a generally
874 good description of data some discrepancies remain. The next sections describe
875 corrections applied to the simulation in order to have a better description of data.
876 In App. B data distributions are compared with simulated ones for variables relevant
877 to this analysis.

878 3.3.1 Decay Model

879 Little is known about Λ_b^0 decays structure and therefore the simulation software
880 generates events according to the phase space given by the available kinematic. To
881 include a reasonably realistic q^2 dependence, the simulation is weighted using decay
882 amplitudes based on the predictions in Ref. [79]. Equations in this paper are for the
883 case of unpolarised Λ_b^0 production and for this analysis those are extended to include
884 polarisation. Details about the models used are in Appendix A.1. The value of the
885 Λ_b^0 production polarisation, P_b , used in the calculations is $P_b = 0.06$ as measured by
886 LHCb [80]. Figure 3.2 shows the phase space q^2 distribution and the one obtained
887 re-weighting the events. The latter can be qualitatively compared to the q^2 spectrum
888 of a generic $b \rightarrow s\ell^+\ell^-$ decay reported in Fig. 1.8. For the normalisation mode, the
889 decay model used is described in Appendix A.2, with amplitude magnitudes and
890 production polarisation taken from the measurement in Ref. [80]. Phases are not
891 yet measured and are all set to zero.

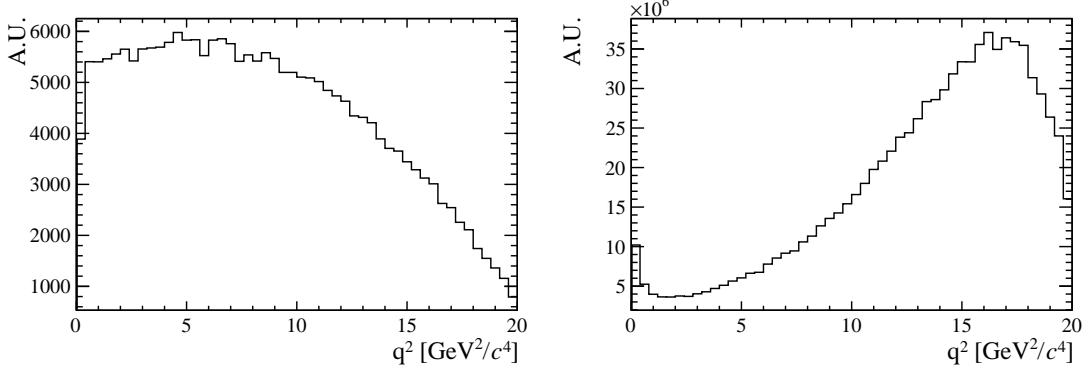


Figure 3.2: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulates events according to the phase space of the decay (left) and reweighted using the decay amplitudes (right).

3.3.2 Kinematic re-weighting

Small data-simulation differences are found in the kinematic properties of the mother particle, Λ_b^0 , which also affect the final state particles. The simulation is re-weighted by comparing the two-dimensional momentum and transverse momentum of Λ_b^0 between real and simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ which passed pre-selection. To do this a data sample as clean as possible is obtained selecting a narrow interval around J/ψ and Λ_b^0 peaks. Then the Λ_b^0 invariant mass is fitted to extract the amount of background under the peak. The background fraction with respect to the signal, $f_b = B/(S+B)$, is then used to statistically subtract the background from the kinematical distributions as described by the equation:

$$S(p, p_T) = T(p, p_T) - f_b \cdot B(p, p_T), \quad (3.2)$$

where $S(p, p_T)$ is the distribution of pure signal events, which we want to obtain, $T(p, p_T)$ 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 $B(p, p_T)$ 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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ simulated events and a weight, $w(p_{\Lambda_b^0}, p_{T\Lambda_b^0})$ is defined by taking the ratio of

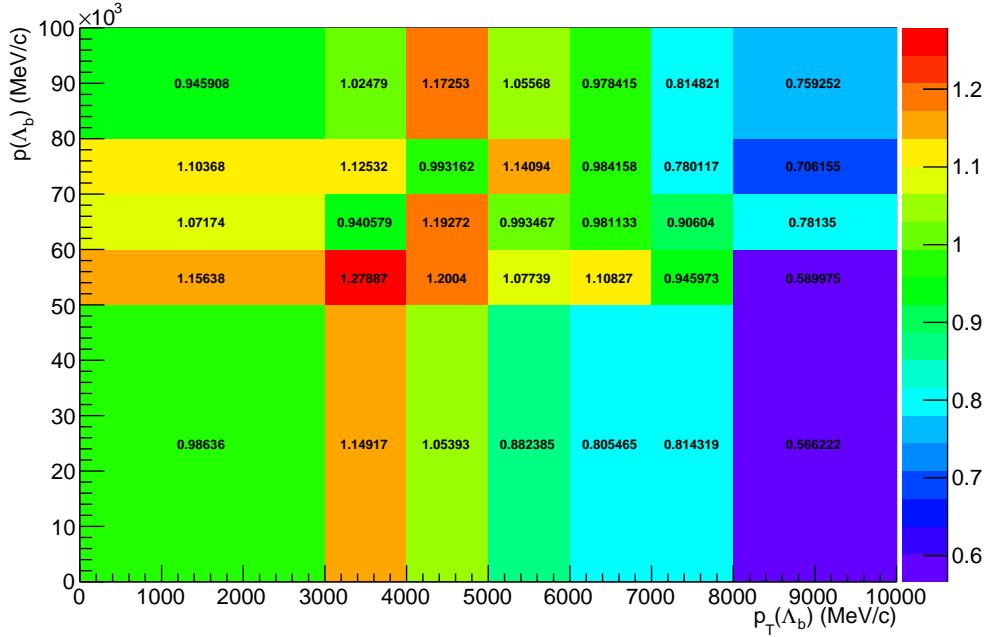


Figure 3.3: Weights used for the kinematical reweighting as a function of the momentum and transverse momentum of Λ_b^0 .

the two dimensional distributions. The result is shown in Fig. 3.3. 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 $\Lambda_b^0 p$ and p_T distributions 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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays which passed the full selection, $\sim 70\%$ of candidates are reconstructed from downstream tracks. On the contrary, in the simulation of the same decay, $\sim 75\%$ of candidates are reconstructed using downstream tracks. The fraction of downstream and long

923 tracks also varies as a function of q^2 and the biggest differences are found at low
924 q^2 . In order to deal with these differences all efficiencies are obtained separately for
925 downstream and long candidates and the analysis is done separately for the two
926 categories joining results at the end. It is therefore not required to correct the
927 simulation to reproduce the correct fraction of events in each category.

CHAPTER 4

928

929

Selection

930

932 The reconstruction of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ candidates begins with the
933 application of requirements on basic kinematic properties as the p_T of the final
934 particles and quality requirements for the track and vertices. The selection described
935 in this chapter aims to first form a dimuon candidate from two oppositely charged
936 muons. Then, in events containing a dimuon candidate, two oppositely charged
937 tracks are combined and retained as Λ candidate, if they form a good vertex which
938 is well separated from all primary vertices. In the final step, dimuon candidates are
939 combined with Λ candidates to form Λ_b^0 baryons and requirements are set on the
940 properties of this combination. The selection procedure is divided in a pre-selection,
941 where cuts are applied in order to be able to work with manageable samples and a
942 multivariate analysis which combines information from several variables.

₉₄₃ **4.1 Pre-selection**

₉₄₄ The full list of pre.selection cuts is reported in Tab. 4.1. In the table χ^2_{IP} is de-
₉₄₅ fined as the projected distance from a vertex divided by its uncertainty, for example
₉₄₆ the $\chi^2_{IP}(\text{primary}) > n$ requirement on Λ_b^0 means that the Λ_b^0 vertex is at least
₉₄₇ \sqrt{n} standard deviations away from the primary vertex. Another useful quantity
₉₄₈ used to remove combinatorial background is a pointing variable defined as the an-
₉₄₉ gle between the direction of the particle momentum and the flight direction from
₉₅₀ its mother vertex, called DIRA. This allows the selection of particles with well-
₉₅₁ defined primary vertices. Graphical representation of the χ^2_{IP} and DIRA variables
 are shown in Fig. 4.1. The variable χ^2_{FD} represents the flight distance with respect

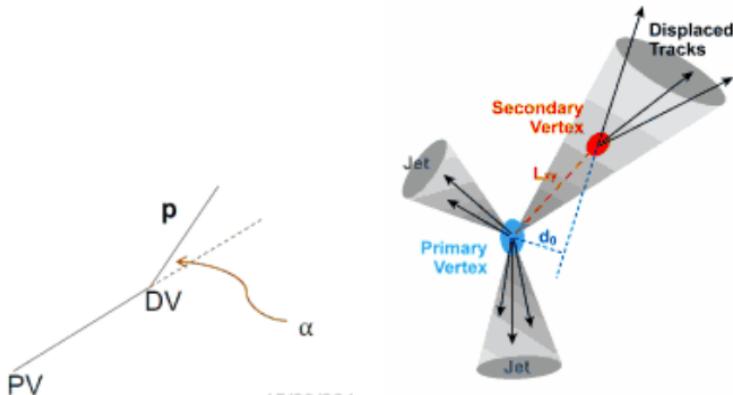


Figure 4.1: Graphical representation of the DIRA (left) and χ^2_{IP} (right) variables.

₉₅₂
₉₅₃ to the origin vertex divided by its uncertainty. The χ^2_{trk}/ndf and χ^2_{vtx}/ndf quanti-
₉₅₄ ties are the χ^2 from the fit to tracks and vertices, which are used to quantify their
₉₅₅ quality. The `GhostProb` quantity describes the probability of a track being fake.
₉₅₆ By construction, cutting at k , removes $(1 - k) \cdot 100\%$ of fake tracks. The `hasRich`,
₉₅₇ `hasCalo` and `isMuon` variables are binary indicators that the information from the
₉₅₈ RICH/calorimeter/muon detector is available for the track. Loose PID requirements
₉₅₉ are also applied in pre-selection to limit the size of the samples. Details about the
₉₆₀ quantification of the PID quality are given in Sec. 2.8. A large mass window is kept
₉₆₁ around the Λ_b^0 peak in order to be able to fit the sideband, to train the multivariate
₉₆₂ analysis and better constrain backgrounds. Rare candidates are selected by the q^2

963 region requirements described in Sec. 3.1, while resonant candidates are further con-
964 strained to have dimuon invariant mass in a $100 \text{ MeV}/c^2$ interval around the known
965 J/ψ mass [2].

Particle	Requirement
A_b^0	$4.6 < m(p\pi\mu\mu) < 7.0 \text{ GeV}/c^2$ $\text{DIRA} > 0.9999$ $\chi_{\text{IP}}^2 < 16.0$ $\chi_{\text{FD}}^2 > 121.0$ $\chi_{vtx}^2/\text{ndf} < 8.0$
A	$\chi_{vtx}^2/\text{ndf} < 30.0(25.0)$ $\text{Decay time} > 2 \text{ ps}$ $ m(p\pi) - m^P D G_A < 35(64) \text{ GeV}/c$
p/π	$p > 2 \text{ GeV}/c$ $p_T > 250 \text{ MeV}/c$ $\chi_{\text{IP}}^2 > 9(4)$
p (only long cand.)	hasRICH PIDp > -5
μ	isMuon $\chi_{trk}^2/\text{ndf} < 5$ GhostProb < 0.4 PIDμ > -3 $\chi_{\text{IP}}^2 > 9.0$
Dimuon	$\chi_{vtx}^2/\text{ndf} < 12.0$ $m(\mu\mu) < 7.1 \text{ GeV}/c^2$

Table 4.1: 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.

966 4.2 Neural Networks

967 The final selection is performed using a neural network (NN) classifier based on
968 the NeuroBayes package [60, 61]. The input to the neural network consists of 14
969 variables carrying information about the kinematics of the decay, the quality of
970 tracks and vertices and the PID of the muons. The list of inputs to the neural
971 network is reported in Tab. 4.2, together with information about the importance of
972 inputs. Variables related to A and its daughters are considered as different inputs
973 depending on the candidate type (long or downstream). This effectively corresponds

974 to making a separate training for the two categories. The graphical representation
975 of the correlation matrix is shown in Fig. 4.2, where the variable with ID= 1 is the
976 NN output and the IDs of the other variables can be found in Tab. 4.2.

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

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

999 Figure 4.3 shows distributions of neural network output for the signal and back-
1000 ground samples and purity, $P = N(\text{signal})/N(\text{background})$, as a function of the
1001 neural network output. The distributions from test samples are also overlaid in
1002 order to check for overtraining. The distributions follow the same shape but with

Table 4.2: Summary of inputs to the neural network in order of importance. Column “ID” lists the indices used for the correlation matrix (see Fig. 4.2). Column “adds” gives the significance added by a given input when it is added to the list of those ranked above. Finally, 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	ID	adds	only this	loss
$\Lambda_{DD} p_T$	15	143.11	143.11	29.20
χ^2_{DTF}	2	77.81	134.00	51.10
$\min(\chi^2_{IP} \mu)$	7	61.31	113.62	29.76
$\chi^2_{IP} \Lambda_b^0$	5	52.94	113.23	40.98
$\chi^2_{IP} \pi_{LL}$	16	20.29	60.72	12.82
$\min(\text{PID } \mu)$	8	17.91	59.11	13.44
$\tau_{\Lambda_b^0}$	3	16.24	35.36	11.24
$\Lambda_b^0 \text{DIRA}$	4	12.28	73.96	9.98
$\Lambda_{DD} \text{flight distance}$	14	9.47	86.75	11.24
$\chi^2_{IP} \Lambda_{DD}$	13	10.58	59.84	8.88
$\max(\chi^2_{IP} \mu)$	6	9.51	97.24	8.15
$\chi^2_{IP} \Lambda_{LL}$	10	7.31	54.27	10.32
$\max(\text{PID } \mu)$	9	6.99	69.33	6.87
$\pi_{LL} p_T$	18	6.13	47.03	7.12
$\Lambda_{LL} p_T$	12	5.58	49.64	5.86
$\chi^2_{IP} p_{LL}$	17	4.48	53.01	4.18
$\chi^2_{IP} p_{DD}$	20	3.43	55.09	3.31
$\Lambda_{LL} \text{flight distance}$	11	0.87	52.52	0.86
$p_{DD} p_T$	21	0.74	129.58	0.75
$\chi^2_{IP} \pi_{DD}$	19	0.24	70.43	0.24

1003 different fluctuations indicating no significant overtraining. In general it can be con-
 1004 cluded that the neural network is able to separate signal from background and the
 1005 training converged properly.

1006 It can happen that too much information is given to the classifier, which becomes
 1007 able to calculate the invariant mass of the candidates from the input variables.
 1008 This can generate fake peaks and it is therefore important to check for correlations
 1009 between the 4-body invariant mass and the NN output. Figure 4.4 reports the
 1010 average NN output value as a function of 4-body $m(K\pi\mu\mu)$ invariant mass for data
 1011 and simulation. The distributions are flat indicating that no significant correlation
 1012 is present.

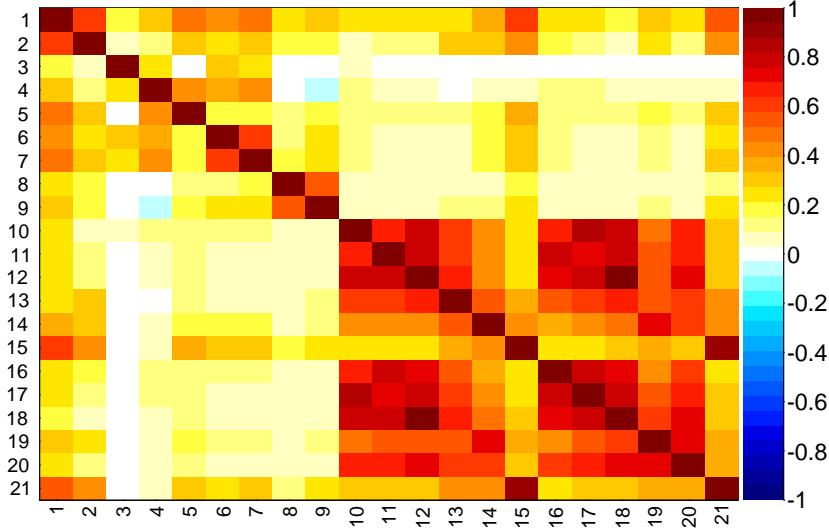


Figure 4.2: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of Tab. 4.2. Correlation is calculated using all events without distinguishing signal and background.

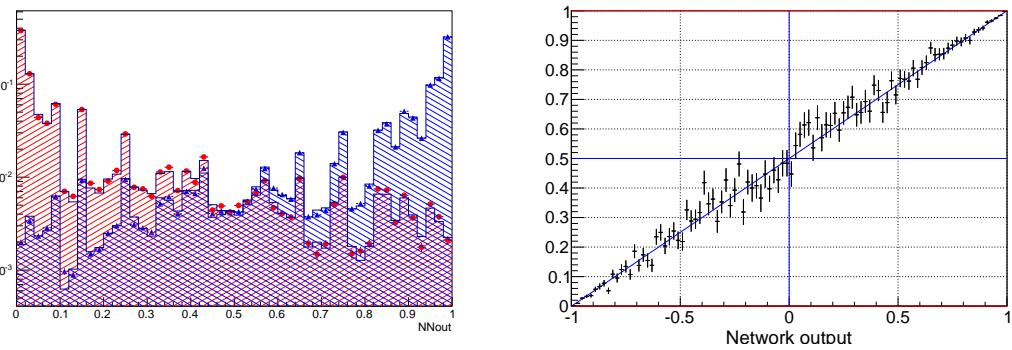


Figure 4.3: (left) NN output distribution for training (points) and test (stripes) samples, for signal and background events. (right) Purity as a function of neural network output.

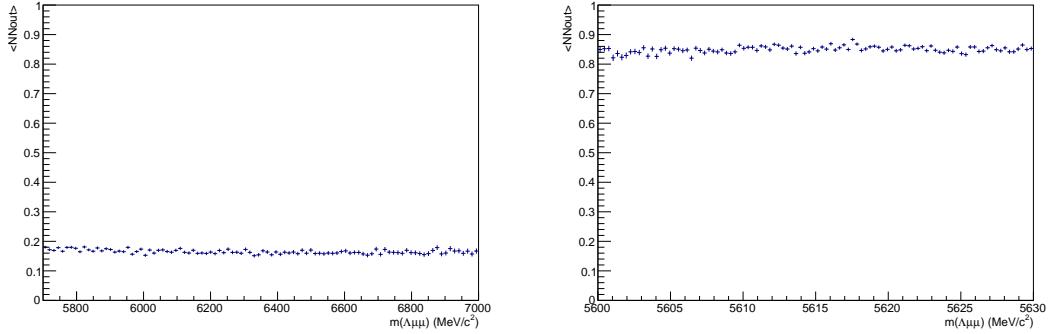


Figure 4.4: Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and simulated signal (right) events.

1013 4.3 MVA optimization

1014 In the high q^2 region, where the signal is already observed, the final requirement
 1015 on the neural network output is chosen maximising the significance, $N_S/\sqrt{N_S + N_B}$,
 1016 where N_S and N_B are the numbers of expected signal and background candidates
 1017 respectively. N_S is derived from simulation but, as an arbitrary number of events can
 1018 be generated, it needs to be normalised. To do this, the invariant mass distribution
 1019 of real $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates is fit after preselection (including all requirements but
 1020 MVA). This is possible as the peak of the resonant channel is already well visible
 1021 after preselection. Then the resonant yield is scaled by the ratio of between the
 1022 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ branching fractions as measured by LHCb on 2011
 1023 data

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) = 1.54 \times 10^{-3} \quad (4.1)$$

1024 and the $J/\psi \rightarrow \mu^+ \mu^-$ branching fraction. In summary:

$$N_S = N_{J/\psi} \cdot \frac{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}. \quad (4.2)$$

1025 The number of expected background events instead is derived fitting the data side-
 1026 band with an exponential and extrapolating under the singnal region.

1027 In the low q^2 region, where the signal is unobserved, the so called ‘‘Punzi figure-of-
 1028 merit’’, $N_S/(n_\sigma/2 + \sqrt{N_B})$, is maximised [81]. This figure-of-merit is considered to be

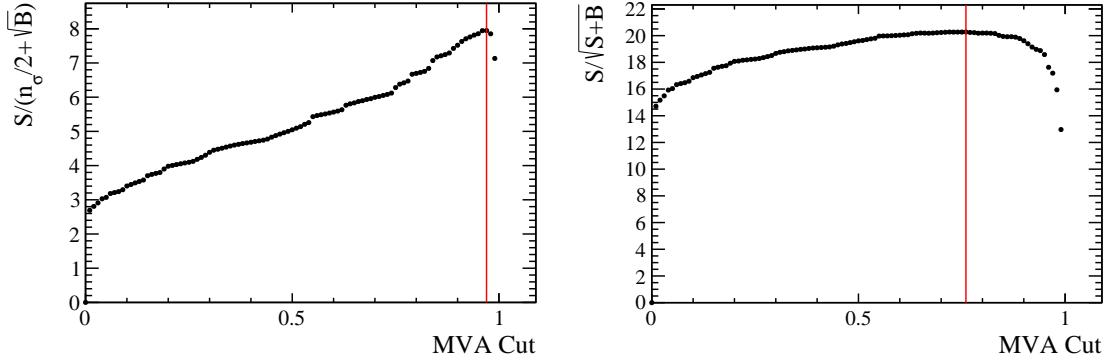


Figure 4.5: Dependence of the figure-of-merits on the neural network output requirement for the low q^2 (left) and high q^2 (right) regions. The vertical lines correspond to the chosen cuts.

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. 4.5, and curves of signal efficiency versus background rejection are shown in Fig. 4.6.

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 requirements the neural network retains approximately 96% (66 %) of downstream candidates and 97 % (82 %) of long candidates for the selection at high (low) q^2 , with respect to the pre-selected samples. After full selection $\sim 0.5\%$ of the events contain multiple candidates which are randomly rejected to keep 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.

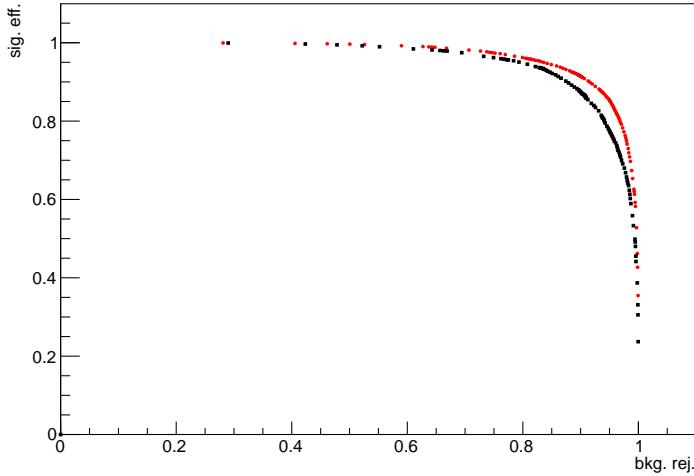


Figure 4.6: Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection. The optimal points on these curves are the closest ones to (1,1).

1045 4.4 Trigger

1046 Finally, specific trigger lines are selected, corresponding to events triggered by the
1047 muons of the reconstructed candidate. This is denoted as Trigger On Signal (TOS).
1048 The trigger lines used in the analysis are shown in Tab. 4.3. The logical *or* of
1049 lines on the same lever is required and the logical *and* and lined in different levels.
1050 The L0Muon trigger requires hits in the muon detector and triggers if a muon with
1051 $p_T > 1.5 \text{ GeV}/c$ is identified. L0Dimuon imposes the same requirement on the sum
1052 of the transverse momenta of two tracks. The Hlt1TrackAllL0 performs a partial
1053 reconstruction of the events and applies basic requirements on the IP, χ^2 and p_T
1054 of tracks; it triggers if the L0 decision is confirmed. Hlt1TrackMuon applies looser
1055 requirements but in addition requires the `isMuon` variable (see Sec. 2.8) to be true
1056 to limit the yield. Finally, at the Hlt2 level, a complete reconstruction is done and a
1057 multivariate analysis is used to identify decay structures. One of the main variables
1058 used at this stage is the Distance Of Closest Approach (DOCA), which is required
1059 to be less than 0.2 mm to form a 2-body object.

Table 4.3: 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

1060 4.5 Background from specific decays

1061 Candidates from other decays can be reconstructed as the decays of interest if par-
 1062 ticles are not reconstructed or mis-identified. A survey of possible backgrounds
 1063 concluded that the only physics background to take into account is coming from
 1064 misreconstructed decays of B^0 to K_s^0 with two muons, whether via J/ψ or not,
 1065 where the K_s^0 is reconstructed as a Λ with a $p \rightarrow \pi$ identity swap and $m(p\pi)$ in the
 1066 Λ mass window. The lack of background from other decays is mainly due to the par-
 1067 ticular topology of the Λ decay, which is long-lived and decays at a displaced vertex.
 1068 To study the effect of misreconstructed $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays
 1069 simulated samples are used. On data the $B^0 \rightarrow J/\psi K_s^0$ contribution is clearly visible
 1070 in the resonant channel mass distribution. This background is not suppressed with
 1071 specific cuts in this analysis as its mass shape is sufficiently distinct from Λ_b^0
 1072 signal, which allows to reliably model its contribution in the mass fits (see Sec. 5.1).
 1073 For the rare case a rough estimate of the K_s^0 background size is obtained using the
 1074 yield in the resonant channel rescaled by the measured ratio between the rare and
 1075 resonant branching fractions. Details are given in Sec. 5.1 and numbers of events
 1076 predicted are reported in Tab. 5.1. This contribution, although close to negligible
 1077 is again considered in the fit. A possible pollution due to $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ decays,
 1078 where the K^{*+} further decays into $K_s^0 \pi$ is also investigated using a dedicated sim-
 1079 ulated sample and found to be negligible. Finally, $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events radiating
 1080 photons from the final state, can escape the J/ψ veto and be reconstructed in the

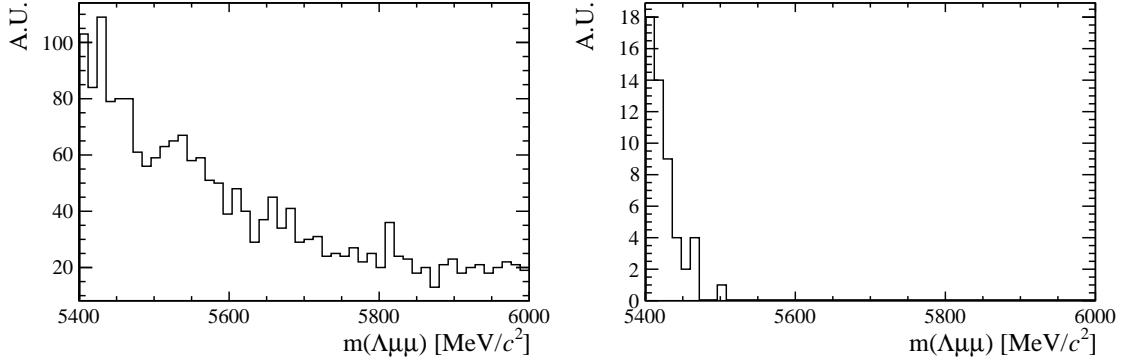


Figure 4.7: Invariant mass distributions of simulated $B^+ \rightarrow \mu^+\mu^-K^{*+}$ (left) and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ (right) candidates passing the full selection. Only $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed in $q^2 < 8$ GeV $^2/c^4$ are selected. Distributions are shown in the invariant mass range relevant for the analysis (see Sec. 5.1).

rare channel. Analysing simulated events it was found that the only contribution is in the closest q^2 interval to the J/ψ tail, $6 < q^2 < 8$ GeV $^2/c^4$. In this interval 1.3% of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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 4.7 shows the invariant mass distribution of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events falling into the rare q^2 region and the distribution of simulated $B^+ \rightarrow \mu^+\mu^-K^{*+}$ events misreconstructed and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays.

1091

CHAPTER 5

1092

1093

Yield extraction

1094

1095 5.1 Fit description

1096 Extended unbinned maximum likelihood fits are used to extract the yields of the
1097 rare and resonant channels. The likelihood has the form:

$$\mathcal{L} = e^{-(N_S + N_C + N_B)} \times \prod_{i=1}^N [N_S P_S(m_i) + N_C P_C(m_i) + N_B P_B(m_i)] \quad (5.1)$$

1098 where N_S , N_C and N_B are respectively the numbers of signal, combinatorial and
1099 K_S^0 background events and the $P_i(m_i)$ are the corresponding probability density
1100 functions (PDF). The fit variable is the 4-body $m(p\pi\mu\mu)$ invariant mass obtained
1101 from a kinematical fit of the full decay chain in which each particle is constrained
1102 to point to its assigned origin vertex and the invariant mass of the $p\pi$ system is
1103 constrained to be equal to the world average for the Λ baryon mass. In the resonant

1104 case a further constrain is used on the dimuon mass to be equal to the known J/ψ
1105 mass. This method allows to improve the mass resolution giving better defined
1106 peaks and therefore a more stable fit. For brevity, in the following these variables
1107 are simply referred to as “invariant mass”.

1108 For the resonant channel the signal is described as the sum of two Crystal Ball
1109 functions (CB) with common mean (m_0) and tail slope (n). This is also known
1110 as Double Crystal Ball (DCB) function. A single Crystal Ball [82] is a probability
1111 density function commonly used to model various processes involving energy loss.
1112 In particular it is used to model the radiative tail, which can be seen in many
1113 resonances’ peaks. This function consists of a Gaussian core and a power-law tail,
1114 below a certain threshold and has form

$$C(x; \alpha, n, \bar{x}, \sigma) = N \cdot \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma}\right) & \text{if } \frac{(x-\bar{x})}{\sigma} > \alpha, \\ A \left(B - \frac{(x-\bar{x})}{\sigma}\right)^{-n} & \text{if } \frac{(x-\bar{x})}{\sigma} < \alpha, \end{cases} \quad (5.2)$$

1115 where for normalisation and continuity

$$\begin{aligned} A &= \left(\frac{c}{|\alpha|}\right)^n \cdot \exp\left(-\frac{\alpha^2}{2}\right), \\ B &= \frac{n}{|\alpha|} - |\alpha|. \end{aligned} \quad (5.3)$$

1116 The full form of the PDF for the resonant channel is therefore:

$$P_S(m; m_0, \alpha_1, \alpha_2, f, n) = f \text{CB}(m; m_0, \sigma_1, \alpha_1, n) + (1-f) \text{CB}(m; m_0, \sigma_2, \alpha_2, n), \quad (5.4)$$

1117 where f is the relative fraction of candidates falling into the first CB function.

1118 As a first step simulated $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distributions are fitted using
1119 the signal PDF alone separately for long and downstream candidates. Figure 5.1
1120 shows distributions of candidates selected in the resonant sample with the fit function
1121 overlaid.

1122 In a second step the fit to the resonant channel data sample is performed. For this fit
1123 the tail slope parameter, “ n ”, which is highly correlated with α_1 and α_2 , is fixed to

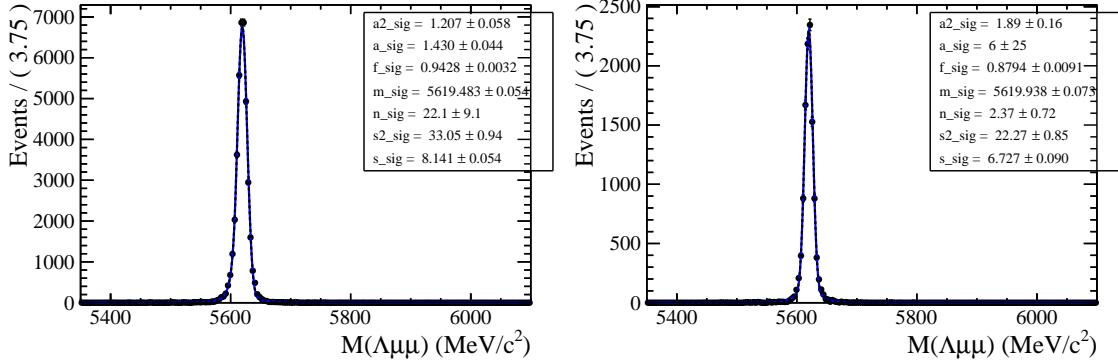


Figure 5.1: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ downstream (left) long (right) candidates. The points show simulated data and the blue line is the signal fit function.

the value found in the fit to simulated data. In this fit two background components are modelled: the combinatorial background, parameterized by an exponential and the background from $B^0 \rightarrow J/\psi K_s^0$ decays. The K_s^0 background is described using the shape obtained using a $B^0 \rightarrow J/\psi K_s^0$ simulated sample and applying to it the full selection. The invariant distribution of these events is fit with a DCBall function, which is then used to model the K_s^0 background in the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ fit. The fit to the simulated misreconstructed $B^0 \rightarrow J/\psi K_s^0$ events is reported in Fig. 5.2. When the K_s^0 shape is introduced in the fit to the data all 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 in account possible data-simulation differences, an horizontal shift is added and left floating (by adding a constant to the central value, m_0 of the DCB). In summary, the free parameters in the fit to the resonant $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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. Notice that all parameters of the fit to the long and downstream samples are independent.

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

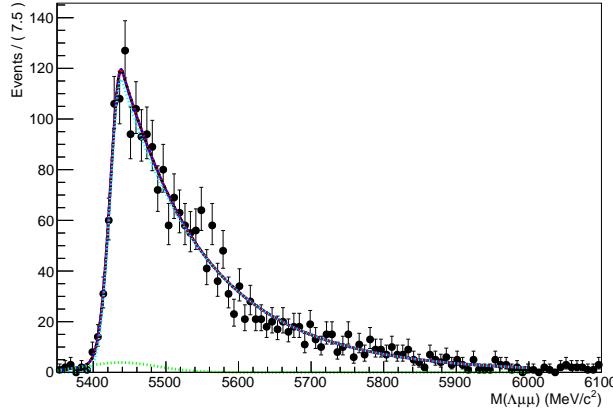


Figure 5.2: Invariant mass distribution of simulated $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted a Double Crystal Ball function.

to limit systematic uncertainties. In fact 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 on the normalisation channel. To account for possible differences, arising from a different resolution in different q^2 regions, a scale factor is multiplied to the width of the two gaussians cores at the signal DCB: $\sigma_1 \rightarrow c \cdot \sigma_1$ and $\sigma_2 \rightarrow c \cdot \sigma_2$, where the two scale factors are the same. This factors are fixed in the fit to data by fitting the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ simulated sample in each q^2 bin and comparing the widths with the ones found on the fit to the resonant simulated sample, namely

$$c = \sigma_{\mu^+ \mu^-}^{MC} / \sigma_{J/\psi}^{MC}. \quad (5.5)$$

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 fits on the DD and LL samples the parameters are always fixed to the corresponding J/ψ fit; in this analysis shape parameters are never shared between DD and LL fits.

The modelled background components are also in the rare case the combinatorial

Table 5.1: Predicted numbers of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ 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

background, described with an exponential function and the K_s^0 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 the 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 and for the downstream and long samples. The background component from $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ 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 $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events is small and does not allow to determine reliably the yield. Therefore this is fixed to the the yield of $B^0 \rightarrow J/\psi K_s^0$ decays, rescaling it by the expected ratio of branching fractions between the resonant and rare channels. The q^2 distribution of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ simulated events is then used to predict the yield as a function of q^2 . Table 5.1 reports the number of predicted $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in each q^2 interval obtained with the following formula:

$$N_{K_s^0 \mu^+ \mu^-}(q^2) = N_{J/\psi K_s^0} \frac{B(B^0 \rightarrow K_s^0 \mu^+ \mu^-)}{B(B^0 \rightarrow K_s^0 J/\psi)} \cdot \frac{1}{\epsilon_{rel}} \cdot B(J/\psi \rightarrow \mu^+ \mu^-) \frac{N(q^2)_{MC}}{N_{MC}^{tot}} \quad (5.6)$$

where $N(q^2)_{MC}$ is the number of simulated rare candidates falling in a q^2 interval after full selection and N_{MC}^{tot} is the total number of simulated events. The $K_s^0 \mu^+ \mu^-$ contribution is then completely taken out to study systematic uncertainties as described in Sec. 7.

1180 The fit on the rare sample is performed simultaneously on the LL and DD candidate
1181 categories. Therefore the two separate yields are not separately floating but are
1182 parameterised as a function of a common branching fraction using the following
1183 formula:

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2}{\mathcal{B}(J/\psi\Lambda)} \right] \cdot N(J/\psi\Lambda)_k \cdot \varepsilon_k^{rel} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}, \quad (5.7)$$

1184 where $k = (\text{LL}, \text{DD})$, Δq^2 is width of the q^2 bin and the only free parameter is the
1185 relative branching fraction ratio of the rare over J/ψ channels. For the branching
1186 fraction of the $J/\psi \rightarrow \mu^+\mu^-$ decay the value reported in the PDG book, $(5.93 \pm$
1187 $0.06) \cdot 10^{-2}$ [2] is used and ε^{rel} corresponds to the relative efficiency between the
1188 rare and resonant channels obtained in Sec. 6. In this formula the efficiencies and
1189 the normalisation yield appear as constants, namely $N(\Lambda\mu^+\mu^-)_k = C_k \cdot \mathcal{B}^{rel}$. These
1190 constants are then varied in order to obtain systematic uncertainties on the final
1191 result as described in Sec. 7.

1192 5.2 Fit results

1193 Figures 5.3 and 5.4 show fitted invariant mass distributions for the normalisation
1194 channel, selected with the high q^2 and low q^2 requirements respectively. Table 5.2
1195 reports measured yields of $\Lambda_b^0 \rightarrow J/\psi\Lambda$ candidates found using the low and high q^2
1196 selections. Values for the signal shape parameters are shown on Fig. 5.3. Fits to
1197 the rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ samples are shown in Fig. 5.5 for the integrated $15 < q^2 < 20$
1198 and $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ q^2 intervals, while fitted invariant mass distribution
1199 in all other considered q^2 intervals are in Figs. 5.6 and 5.7 for downstream and
1200 long candidates respectively. The yields of rare candidates obtained from the fit are
1201 reported in Tab. 5.3 together with their significances. Most candidates are found
1202 in the downstream sample comprising $\sim 80\%$ of the total yield. Note that, since
1203 the fit is simultaneous on DD and LL candidates, the yields in the two categories
1204 yields are not parameters free to float independently in the fit but are correlated

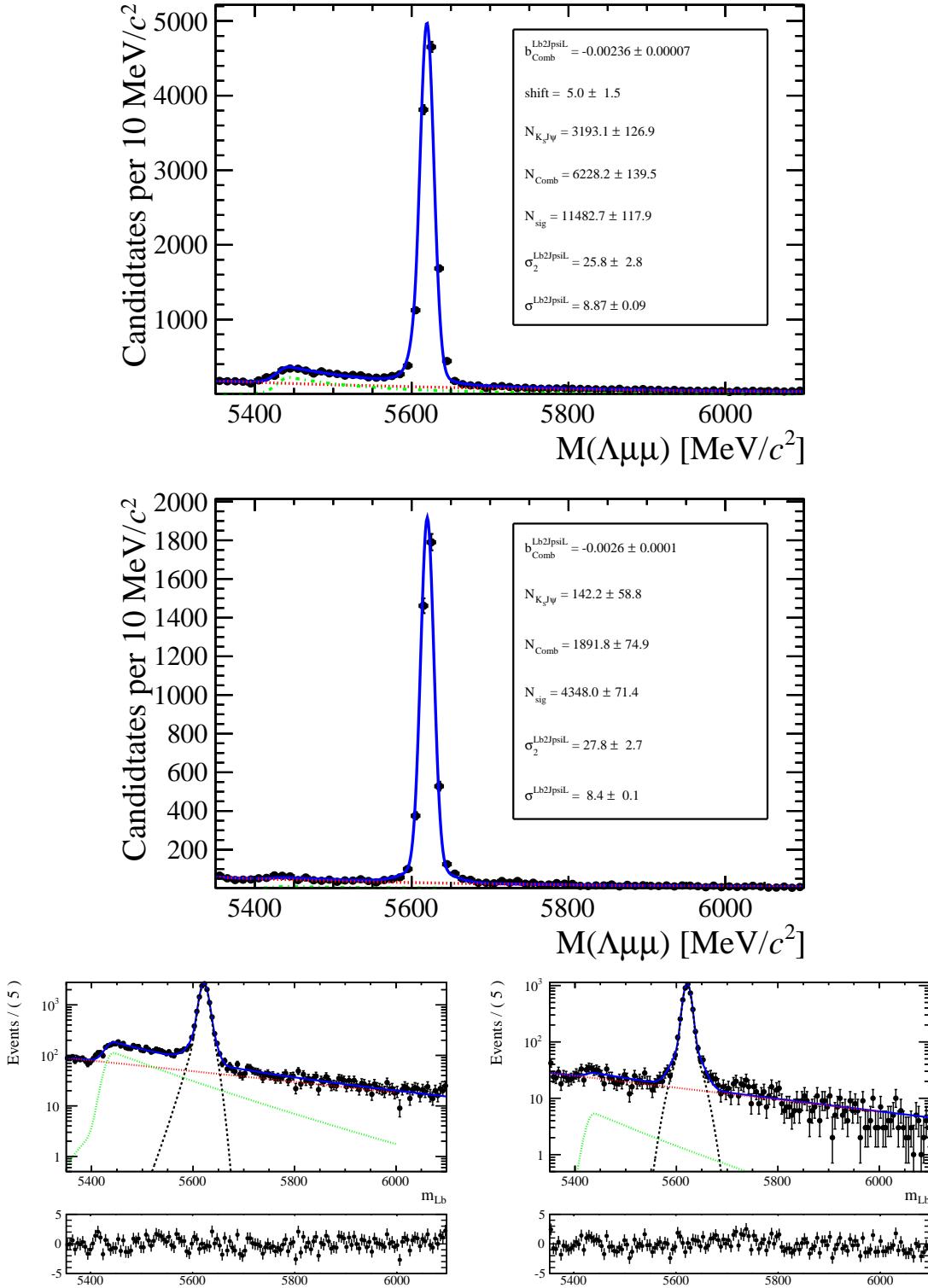


Figure 5.3: Invariant mass distributions of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ downstream (top) and long (middle) candidates selected with high q^2 requirements. Bottom plots are the same as the upper ones but shown in logarithmic scale. Black points show data. The blue solid line represents the total fit function, the black dashed line the signal, the red dashed line the combinatorial background and the green dashed line the $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ background.

Table 5.2: Number of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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 5.3: Signal yields (N_S) obtained from the mass fit to $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ 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

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.

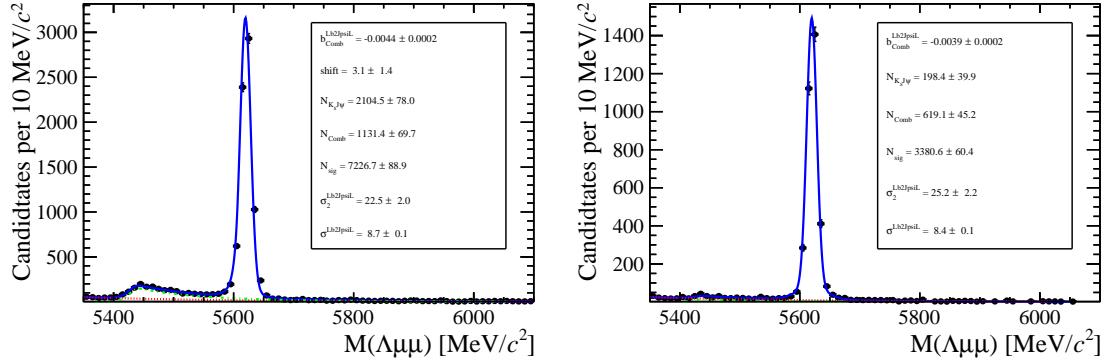


Figure 5.4: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ for downstream (left) and long (right) candidates selected with low q^2 requirements.

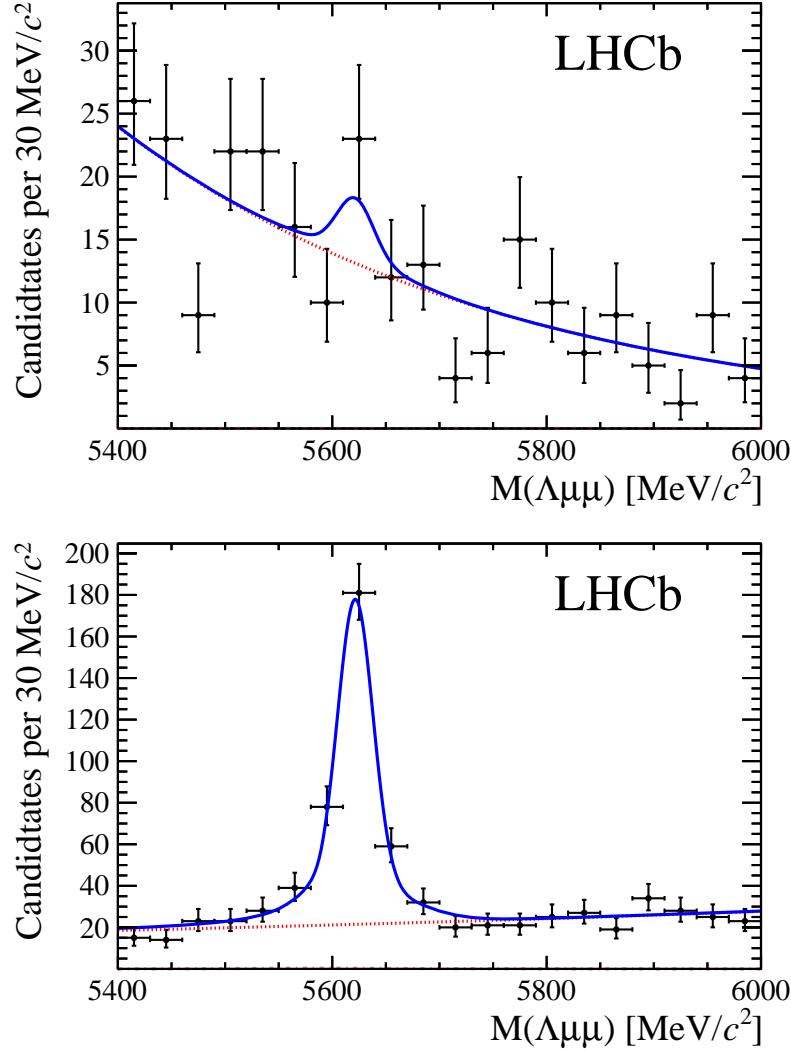


Figure 5.5: Invariant mass distributions of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the integrated 0.1–6.0 GeV $^2/c^4$ (top) and 15–20 GeV $^2/c^4$ (bottom) q^2 intervals. Points show data combining downstream and long candidates together. The blue solid line represents the total fit function and the dashed red line the combinatorial background.

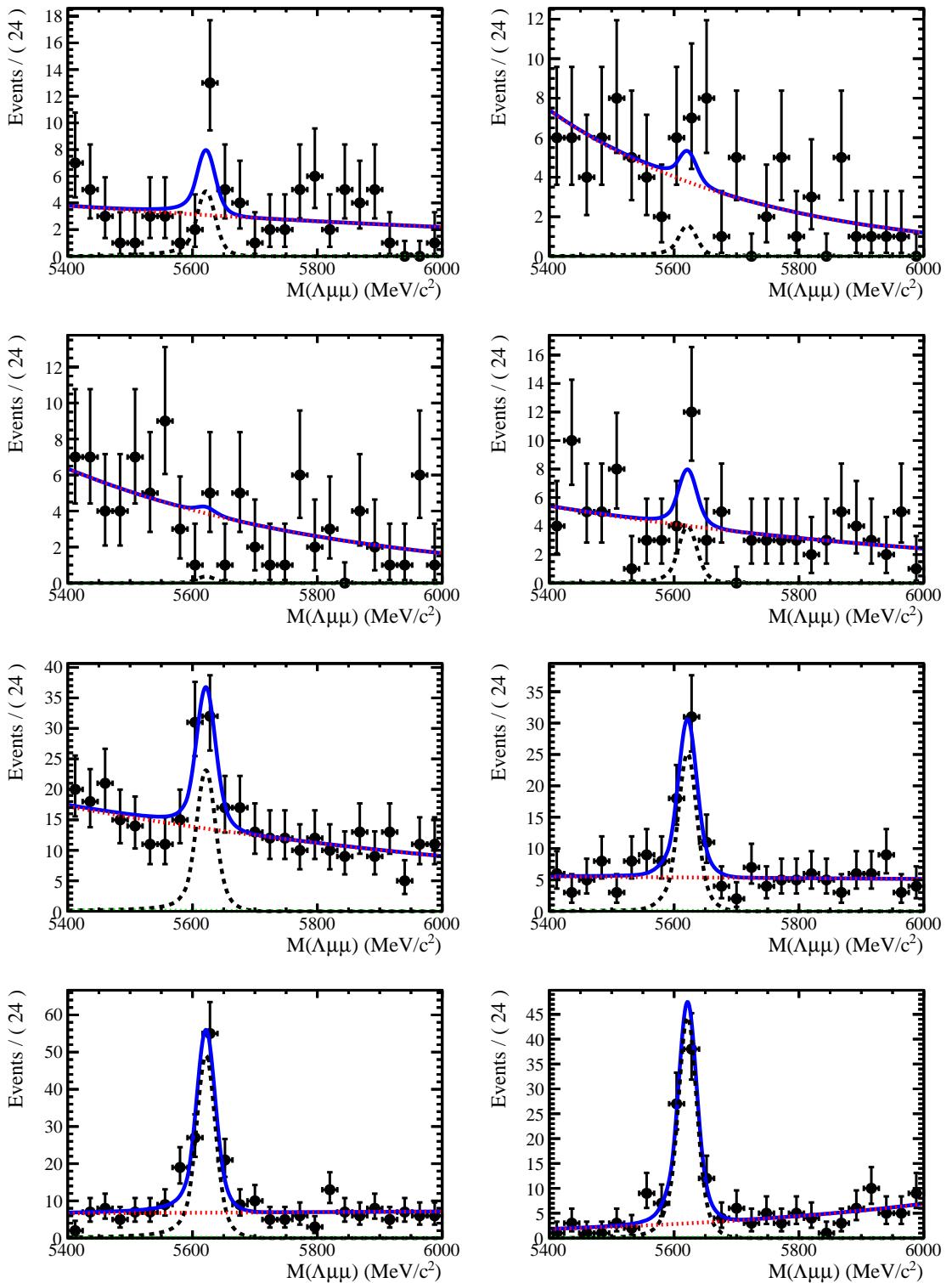


Figure 5.6: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for downstream candidates.

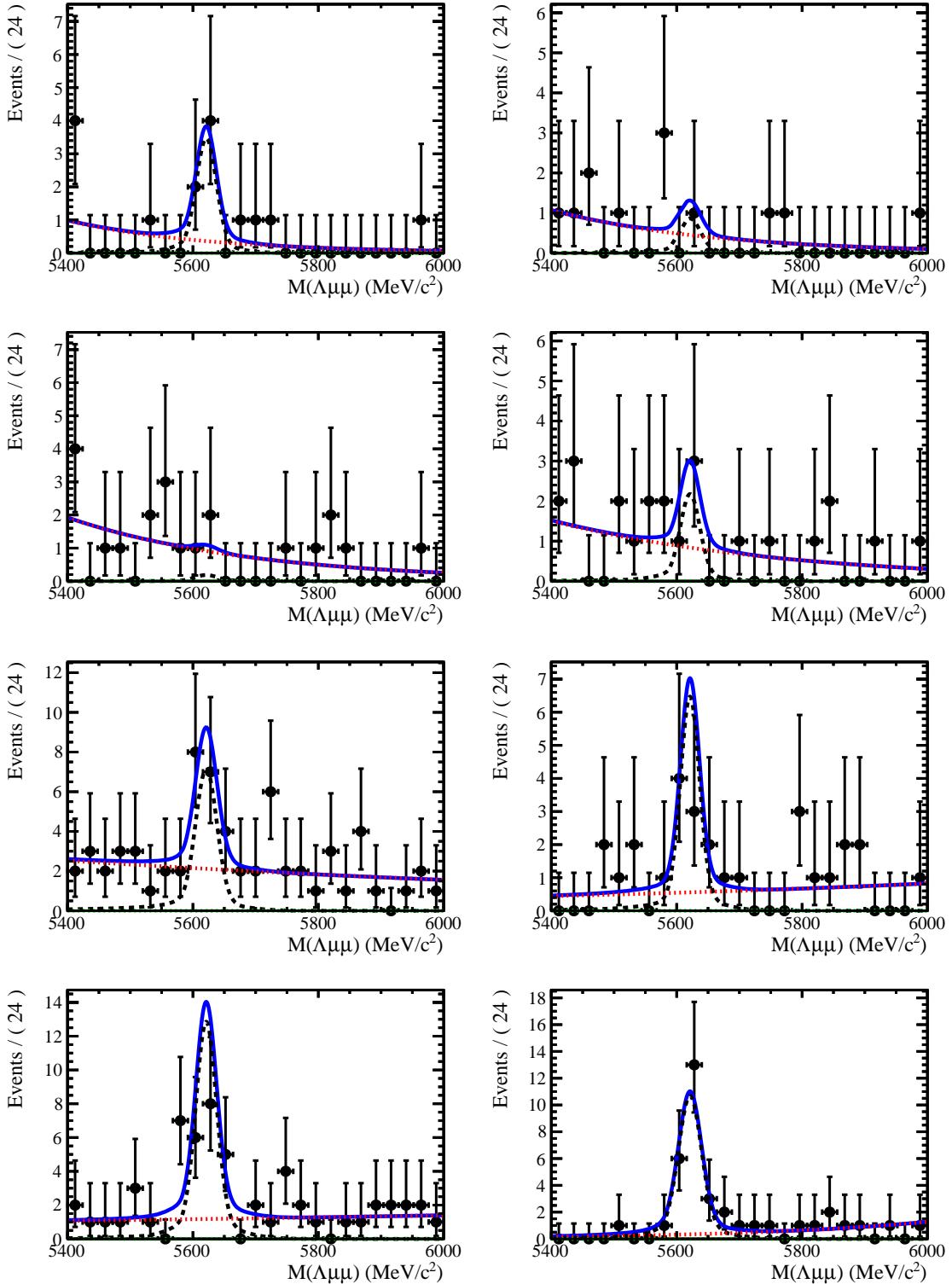


Figure 5.7: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for long candidates.

1209

CHAPTER 6

1210

1211

Efficiency

1212

1213 The efficiency for each of the decays is calculated according to the formula

$$\varepsilon^{tot} = \varepsilon(Geom)\varepsilon(Det|Geom)\varepsilon(Reco|Det)\epsilon(MVA|Reco)\varepsilon(Trig|MVA). \quad (6.1)$$

1214 In this expression the first term gives the efficiency to have final state particles in
1215 the LHCb acceptance. The second term handles the possibility of Λ escaping the
1216 detector or interacting with it and therefore never decaying into $p\pi$. This term
1217 is referred to as “detection” efficiency. The third term carries information about
1218 the reconstruction and pre-selection efficiencies, which are kept together given that
1219 boundaries between them are completely artificial. The fourth part deals with the
1220 efficiency of the Neural Network for those events which passed the pre-selection.
1221 Finally, the last term handles the trigger efficiency. Most of the efficiency com-
1222 ponents are evaluated using the simulated samples described in Sec. 3.3. Only the
1223 efficiency of PID requirement for the proton (see Tab. 4.1) is separately derived with

1224 a data–driven method because the simulation does not provide a good description of
1225 PID variables. For complete information, all absolute efficiencies for the two decays
1226 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ are separately listed in the next sections. However,
1227 for the analysis itself only relative efficiency, $\varepsilon(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-)/\varepsilon(\Lambda_b^0 \rightarrow J/\psi \Lambda)$, is
1228 used.

1229 6.1 Geometric acceptance

1230 In order to save disk space and time, only events are simulated, in which the
1231 final muons are in the detector acceptance and therefore can be reconstructed.
1232 This corresponds to a requirement for each of the muons to be in an interval
1233 $10 < \theta < 400$ mrad, where θ is the angle between the muon momentum and the
1234 beam line. The efficiency of this requirement is obtained by using a separate simu-
1235 lated sample, where events are generated in the full space. In Tab. 6.1 the efficiencies
due to the geometrical acceptance are listed in bins of q^2 for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays.

Table 6.1: Absolute geometrical acceptance in bins of q^2 derived from generator level simulated samples. Uncertainties are statistical only.

q^2 [GeV $^2/c^4$]	Geom. acc.
0.1–2.0	0.2359 ± 0.0008
2.0–4.0	0.2098 ± 0.0007
4.0–6.0	0.2008 ± 0.0007
6.0–8.0	0.1960 ± 0.0008
11.0–12.5	0.1897 ± 0.0010
15.0–16.0	0.1896 ± 0.0015
16.0–18.0	0.1872 ± 0.0012
18.0–20.0	0.1870 ± 0.0016
1.1–6.0	0.2072 ± 0.0005
15.0–20.0	0.1876 ± 0.0008

6.2 Reconstruction and neural network efficiencies

The efficiency to reconstruct the decays together with the pre-selection requirements is evaluated from simulated data. Table 6.2 reports values of reconstruction efficiency in bins of q^2 for long and downstream candidates. In the table the 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 efficiency for this to happen is what is called “Detection” efficiency. The “Reconstruction and Stripping” 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. 4.1, which is kept separate because PID variables are not well described by the simulation and therefore a data-driven method is used instead (see Sec. 6.4).

Table 6.2: Absolute detection and reconstruction plus stripping efficiencies. Reconstruction efficiency is given separately for DD and LL candidates. Uncertainties are statistical only.

q^2 [GeV $^2/c^4$]	Detection	Reco and pre-sel (DD)	Reco and pre-sel (LL)
0.1–2.0	0.8793 ± 0.0005	0.0519 ± 0.0006	0.0194 ± 0.0004
2.0–4.0	0.8850 ± 0.0004	0.0664 ± 0.0006	0.0195 ± 0.0004
4.0–6.0	0.8902 ± 0.0004	0.0717 ± 0.0007	0.0209 ± 0.0004
6.0–8.0	0.8962 ± 0.0005	0.0756 ± 0.0007	0.0212 ± 0.0004
11.0–12.5	0.9084 ± 0.0006	0.0799 ± 0.0009	0.0221 ± 0.0005
15.0–16.0	0.9187 ± 0.0009	0.0736 ± 0.0012	0.0179 ± 0.0007
16.0–18.0	0.9247 ± 0.0007	0.0696 ± 0.0010	0.0169 ± 0.0005
18.0–20.0	0.9318 ± 0.0009	0.0600 ± 0.0011	0.0136 ± 0.0006
1.1–6.0	0.8868 ± 0.0003	0.0684 ± 0.00041	0.0202 ± 0.0002
15.0–20.0	0.9260 ± 0.0005	0.0669 ± 0.00063	0.0159 ± 0.0003

The NN selection efficiency is again evaluated from simulated samples. Results are shown in Tab. 6.3 in bins 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.

Table 6.3: Neural network selection efficiency. Uncertainties are statistical only.

q^2 [GeV $^2/c^4$]	NN eff. (DD)	NN eff. (LL)
0.1–2.0	0.623 ± 0.008	0.813 ± 0.011
2.0–4.0	0.583 ± 0.007	0.757 ± 0.011
4.0–6.0	0.584 ± 0.007	0.776 ± 0.011
6.0–8.0	0.588 ± 0.007	0.778 ± 0.011
11.0–12.5	0.888 ± 0.005	0.944 ± 0.007
15.0–16.0	0.882 ± 0.007	0.929 ± 0.012
16.0–18.0	0.847 ± 0.007	0.928 ± 0.009
18.0–20.0	0.831 ± 0.009	0.889 ± 0.016
1.1–6.0	0.584 ± 0.005	0.772 ± 0.007
15.0–20.0	0.849 ± 0.005	0.917 ± 0.007

₁₂₅₅ 6.3 Trigger efficiency

₁₂₅₆ The trigger efficiency is again calculated on a simulated sample for events which
₁₂₅₇ are accepted by the full selection. Using the resonant channel it is possible to
₁₂₅₈ crosscheck on data the efficiency obtained using the simulation. In LHCb triggered
₁₂₅₉ events can fall in two categories: events triggered by a track which is part of a
₁₂₆₀ signal candidate, Trigger On Signal (TOS), or by other tracks in the event, Trigger
₁₂₆₁ Independent of Signal (TIS). As the TIS and TOS categories are not exclusive the
₁₂₆₂ TIS sample provides a control sample which can be used to obtain the efficiency for
₁₂₆₃ TOS trigger. This is calculated with the formula:

$$\varepsilon_{TOS} = \frac{\text{TOS and TIS}}{\text{TIS}}. \quad (6.2)$$

₁₂₆₄ As data contains background the numbers of signal candidates in the “TIS” and
₁₂₆₅ “Tis&&Tos” categories are determined from a fit to the 4-body invariant mass,
₁₂₆₆ $m(p\pi\mu\mu)$. This procedure takes the name of TISTOS method. Using the data-
₁₂₆₇ driven method an efficiency of $(70 \pm 5)\%$ is obtained, while this is calculated to be
₁₂₆₈ $(73.33 \pm 0.02)\%$ using the simulation. Results are therefore compatible within 1σ .

Table 6.4: Absolute trigger efficiencies for selected events as determined from the simulation separately for LL and DD events.

q^2 [GeV $^2/c^4$]	Trigger eff. (DD)	Trigger eff. (LL)
0.1–2.0	0.560 ± 0.008	0.577 ± 0.012
2.0–4.0	0.606 ± 0.006	0.651 ± 0.010
4.0–6.0	0.623 ± 0.006	0.674 ± 0.010
6.0–8.0	0.669 ± 0.006	0.706 ± 0.010
11.0–12.5	0.744 ± 0.006	0.738 ± 0.011
15.0–16.0	0.818 ± 0.008	0.826 ± 0.015
16.0–18.0	0.836 ± 0.006	0.860 ± 0.011
18.0–20.0	0.857 ± 0.008	0.863 ± 0.015
1.1–6.0	0.610 ± 0.004	0.653 ± 0.007
15.0–20.0	0.839 ± 0.004	0.853 ± 0.008

1269 6.4 PID efficiency

1270 For long tracks a PID cut on protons ($\text{PIDp} > -5$) is used. The simulation is known
 1271 not to describe particle ID variables well and therefore a data-driven method is
 1272 used to obtain this efficiency component. This is done using the `PIDCalib` package
 1273 (see Sec. 2.8.1), which uses as calibrations samples decays where particles can be
 1274 identified due to their kinematic properties. In the case of protons a sample of
 1275 Λ particles is used, where the proton can be identified because it always has the
 1276 highest momentum. The package allows to divide the phase space in bins of variables
 1277 relevant for PID performances, in this analysis momentum and pseudorapidity are
 1278 used. Using the calibration sample the efficiency is derived in each two-dimensional
 1279 bin. To take in account that the decay channel under study could have different
 1280 kinematical distributions than the calibration sample these efficiency tables are used
 1281 to re-weight the simulation. Absolute PID efficiencies are listed in Tab. 6.5 in bins
 1282 of q^2 .

1283 6.5 Relative efficiencies

1284 In the previous sections absolute efficiencies values were given for the rare channel
 1285 in different q^2 intervals. Figure 6.1 contains a summary of those values in these

Table 6.5: Absolute PID efficiencies in q^2 bins

q^2 [GeV $^2/c^4$]	PID efficiency
0.1–2.0	97.32 ± 0.012
2.0–4.0	97.42 ± 0.012
4.0–6.0	97.59 ± 0.011
6.0–8.0	97.70 ± 0.010
11.0–12.5	98.04 ± 0.009
15.0–16.0	98.31 ± 0.006
16.0–18.0	98.10 ± 0.005
18.0–20.0	98.11 ± 0.001
1.1–6.0	97.49 ± 0.007
15.0–20.0	98.17 ± 0.003
J/ψ	97.89 ± 0.005

Table 6.6: Absolute efficiency values for $\Lambda_b^0 \rightarrow J/\psi \Lambda$. Uncertainties are statistical only.

Efficiency	Downstream	Long
Geometric	0.1818 ± 0.0003	
Detection	0.9017 ± 0.0003	
Reconstruction	0.0724 ± 0.0004	0.0203 ± 0.0002
MVA	0.882 ± 0.002	0.942 ± 0.002
Trigger	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

1286 tables in graphical form. This section reports the corresponding relative efficiencies
 1287 with respect to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ channel, which will be used to correct the yields and
 1288 obtain the differential branching fraction. Table 6.6 reports the absolute efficiency
 1289 values for the J/ψ channel used to derive the relative efficiencies. Relative geometric,
 1290 detection and PID efficiencies are listed in Tab. 6.7. In Tabs. 6.9 and 6.8 relative
 1291 reconstruction, trigger and NN efficiencies are listed separately for downstream and
 1292 long candidates. Since these three components are obtained from the same simulated
 1293 sample their statistical errors are correlated. Therefore the total of the three is also
 1294 reported as a single efficiency and labeled “Full Selection” in the table. Finally,
 1295 Tab. 7.3 reports the total of all relative efficiencies, which will be then used to correct
 1296 the raw yields and calculate the differential branching fraction. Uncertainties reflect
 1297 the statistics of both rare and resonant samples, while systematic uncertainties are
 1298 discussed in next chapter.

Table 6.7: Relative geometric, detection and PID relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ 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 6.8: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for long events. Uncertainties reflect the statistics of both samples.

q^2 [GeV $^2/c^4$]	Reco and strip	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 6.9: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for downstream events. Uncertainties reflect the statistics of both samples.

q^2 [GeV $^2/c^4$]	Reco and strip	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

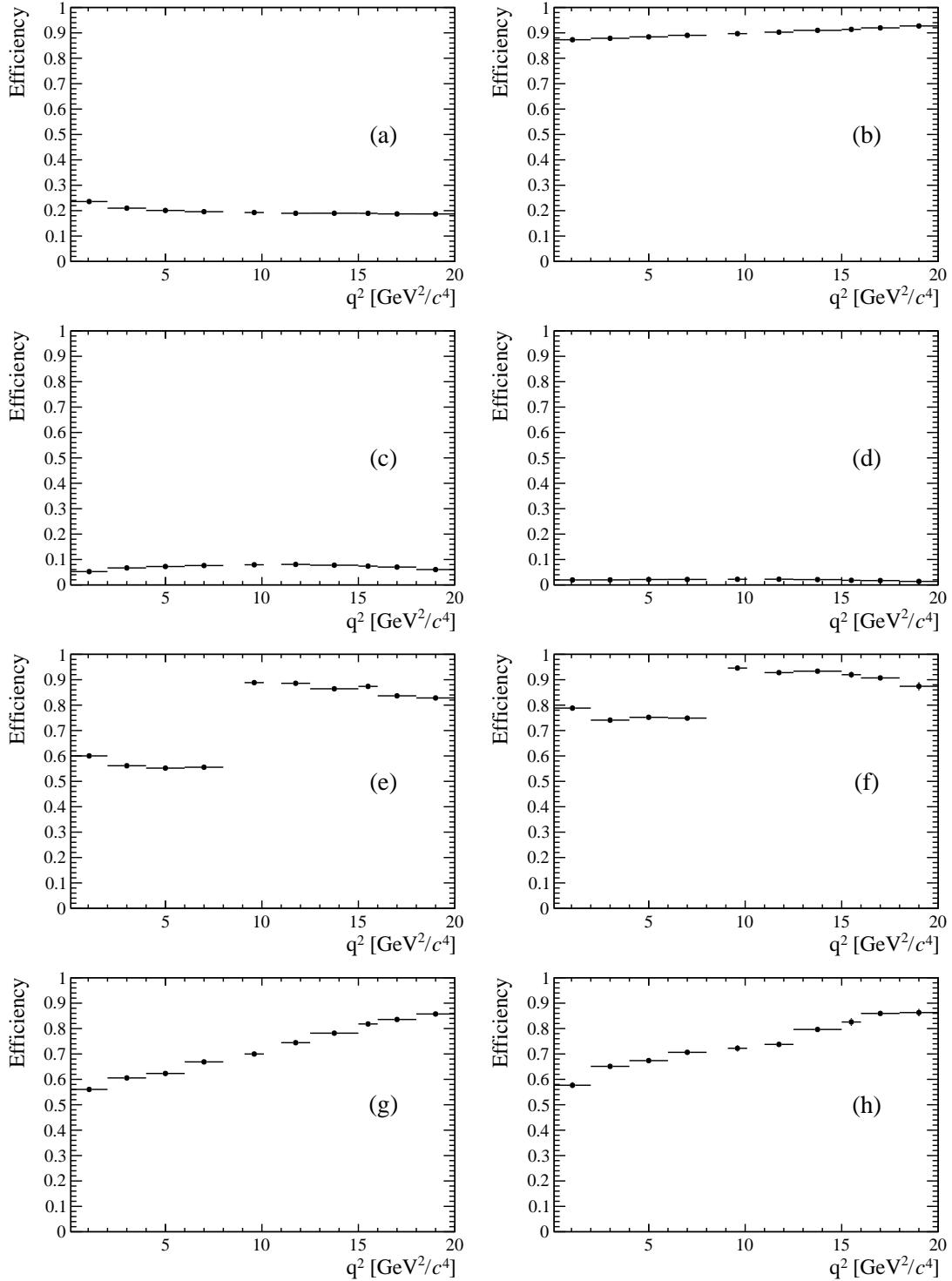


Figure 6.1: Absolute efficiencies as a function of q^2 : geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) candidates, NN efficiency for DD (e) and LL (f) and trigger efficiency for DD (g) and LL (h).

1299

CHAPTER 7

1300

1301

Systematic uncertainties and results

1302

1303 In this chapter the main sources of systematic uncertainty considered are described,
1304 and finally the results of the angular analysis are reported.

1305 7.1 Systematic uncertainty on the yields

1306 The choice of a specific PDF to model the invariant mass distribution could result
1307 in a bias. To asses the effect of the signal PDF choice a number of models are tried
1308 on the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ data sample in order to understand which ones are plausible.
1309 Table 7.1 reports the χ^2 and relative probabilities obtained using different models
1310 including: the default model (a DCB function), a simple Gaussian function, a single
1311 Crystal Ball function and the sum of two Gaussians. The only two models that give
1312 a reasonable p-value are the default DCB and the sum of two Gaussian functions
1313 (DG). As 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

$$b = \left(\frac{N_{\ell\ell}^{DCB}}{N_{J/\psi}^{DCB}} - \frac{N_{\ell\ell}^{DG}}{N_{J/\psi}^{DG}} \right) / \frac{N_{\ell\ell}^{DCB}}{N_{J/\psi}^{DCB}} \quad (7.1)$$

where $N_{\ell\ell}^{model}$ and $N_{J/\psi}^{model}$ are the numbers of rare and resonant candidates observed using a specific model. The distribution of biases have approximately gaussian shape. Finally, 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 float in the second model in order to asses at the same time the systematic due to the parameters constraint.

Table 7.1: χ^2 , NDF, p-values and number of signal events obtained fitting $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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

1326

For the background PDF systematic the rare channel is re-fit leaving the yield of K_s^0 component floating, which is fixed to the predicted value in the default fit. The same procedure as for the signal PDF is applied. Results are reported in Tab. 7.2. The most affected bin 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 $B^+ \rightarrow K^{*+}(K_s^0\pi^+)\mu^+\mu^-$ decays is added in 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 7.2: Values of systematic uncertainties due to the choice of signal and background shapes in bins of q^2 .

1335

1336 7.2 Systematic uncertainties on the efficiency determination

1337 Systematic uncertainties in the efficiency determination are due to the limited knowl-
 1338 edge of the decay properties such as the Λ_b^0 lifetime and production polarisation. The
 1339 uncertainties are directly calculated on the relative efficiencies as these are the ones
 1340 that are actually used in the analysis. It should be noted that not all sources con-
 1341 tribute to each part of the efficiency. For brevity, this section only reports estimates
 1342 of the systematic uncertainties obtained while the full information is contained in
 1343 Appendix C.

1344 7.2.1 Effect of new physics on the decay model

1345 New physics could affect the decay model modifying the Wilson Coefficients by
 1346 adding contributions to the C_7 and C_9 coefficients. This would result in a modifi-
 1347 cation of the q^2 spectrum and therefore of the efficiency. To asses this systematic
 1348 Wilson Coefficients are modified by adding a NP component ($C_i \rightarrow C_i + C_i^{\text{NP}}$). Fig-
 1349 ure 7.1 shows q^2 spectra obtained weighting the simulation for a model embedding

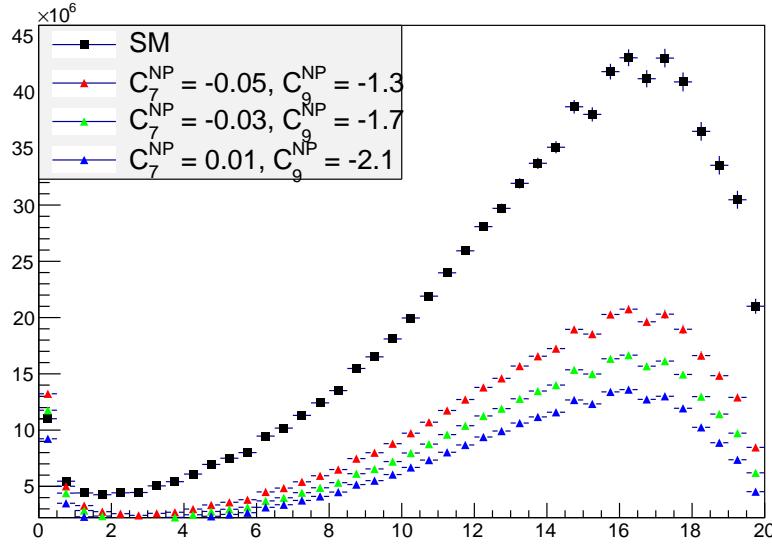


Figure 7.1: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

1350 the default and 3 modified sets of wilson coefficients. Used values used are reported
 1351 on top of each plot and are inspired to maintain compatibility with the recent LHCb
 1352 result for the P'_5 observable [33]. The biggest effect is observed in the very low q^2 ,
 1353 below $2 \text{ GeV}^2/c^4$, where the efficiency can change up to 7%, while it changes 3-4 %
 1354 between 3 and 4 GeV^2/c^4 and 2-3 % in the rest of the spectrum. This analysis is
 1355 performed under the hypothesis that the decays are described by a the SM. There-
 1356 fore, this values are given in order to provide the full information but are not added
 1357 as systematic uncertainties.

1358 7.2.2 Simulation statistics

1359 The limited statistics of the simulated samples used to determine efficiencies is
 1360 considered as a source of systematic uncertainty. While it is not the dominant
 1361 source, its size does not allow to completely neglect it. When reporting relative
 1362 efficiency values the statistical uncertainty due to the rare and resonant channels is
 1363 always considered.

¹³⁶⁴ 7.2.3 Production polarisation and decay structure

¹³⁶⁵ One of the main unknown, which affects the determination of the efficiencies, is
¹³⁶⁶ the angular structure of the decays. And, connected to it, also the production
¹³⁶⁷ polarisation, which is a parameter of the model. To assess the systematic uncertainty
¹³⁶⁸ due to the knowledge of the production polarisation for $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays the
¹³⁶⁹ polarisation parameter in the model is varied within one standard deviation from
¹³⁷⁰ the central value of the most recent LHCb measurement, $P_b = 0.06 \pm 0.09$ [80].
¹³⁷¹ The full difference observed is taken as systematic uncertainty. To assess systematic
¹³⁷² uncertainty due to the decay structure an alternative set of form factors is used based
¹³⁷³ on lattice QCD calculation [83]. Details of this are explained in Appendix A.1. The
¹³⁷⁴ two models are compared and the full difference is taken as systematic uncertainty.
¹³⁷⁵ In total this results in an uncertainty of $\sim 1.3\%$ for long candidates and $\sim 0.6\%$
¹³⁷⁶ for downstream candidates, mostly coming from the knowledge of the production
¹³⁷⁷ polarisation.

¹³⁷⁸ 7.2.4 Λ_b^0 lifetime

¹³⁷⁹ The Λ_b^0 lifetime is known only with limited precision. For evaluation of the efficien-
¹³⁸⁰ cies the world average value, 1.482 ps^{-1} [84] is used. To evaluate the systematic
¹³⁸¹ uncertainty, this is varied within one standard deviation from the measured value.
¹³⁸² Only the case where both signal and normalisation channel are varied in same di-
¹³⁸³ rection are considered. The larger difference with the default lifetime case is taken
¹³⁸⁴ as systematic uncertainty, which is found to range from $\sim 0.4\%$ at low q^2 to $\sim 0.1\%$
¹³⁸⁵ at high q^2 .

¹³⁸⁶ 7.2.5 Downstream candidates reconstruction efficiency

¹³⁸⁷ Other analysis in LHCb using particles reconstructed with downstream tracks showed
¹³⁸⁸ that the efficiency for these candidates is not well simulated. For example, Fig. 7.2

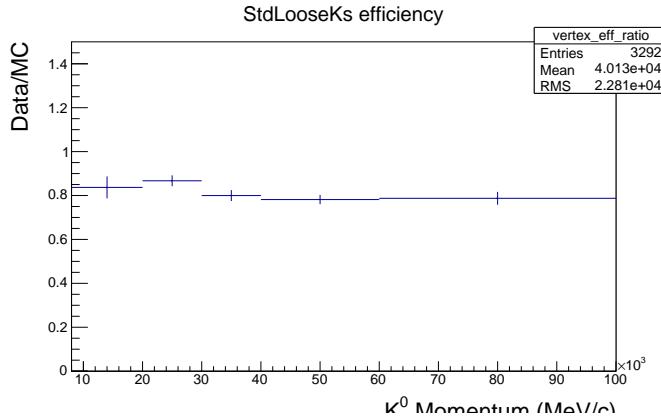


Figure 7.2: Ratio of reconstruction efficiency in Data and MC found using K_S events [85].

1389 shows the ratio between the reconstruction efficiency for downstream candidates in
1390 data and simulation found analysing K_S^0 events [85]. This effect is not yet fully
1391 understood and is currently under study. It seems to be mainly due to a poor
1392 simulation of the vertexing efficiency for downstream tracks.

1393 This effect is dealt with in two steps. Firstly, the analysis is performed separately for
1394 downstream and long candidates. Since efficiencies are also calculated separately,
1395 the effect should mostly cancel in the ratio between the rare and resonant channels.
1396 In a second step a systematic uncertainty is assigned for downstream candidates
1397 only. To do this the simulation is re-weighted by the efficiency ratio between data
1398 and simulation found for K_S as a function of momentum shown in Fig. 7.2. Then
1399 efficiencies obtained using the weighted and unweighted simulation are compared
1400 and the full difference is taken as systematic uncertainty. As the discrepancy shows
1401 little dependence on momentum, dependencies due to the different momentum dis-
1402 tributions of Λ and K_S^0 are assumed to be negligible. This results in an extra 0.4%
1403 systematic uncertainty at low q^2 and 1.2% at high q^2 , only for downstream candi-
1404 dates.

1405 7.2.6 Data-simulation discrepancies

1406 The simulation used to extract efficiency is re-weighted as described in Sec. 3.3.2.
1407 The influence of this procedure on the efficiency determination is checked by com-
1408 paring values obtained with and without re-weighting. The effect is negligible with
1409 respect to other systematics considered.

1410 7.3 Differential branching ratio extraction

1411 In this chapter the differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay is
1412 calculated relative to the $\Lambda_b^0 \rightarrow J/\psi\Lambda$ channel as a function of q^2 . The values are
1413 directly obtained from the fit to the rare sample by parameterising the downstream
1414 and long yields with the following formula:

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2}{\mathcal{B}(J/\psi\Lambda)} \right] \cdot N(J/\psi\Lambda)_k \cdot \varepsilon_k^{\text{rel}} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}, \quad (7.2)$$

1415 where $k = (\text{LL}, \text{DD})$, Δq^2 is width of the q^2 bin and the only free parameter is the
1416 relative branching fraction ratio. For the $J/\psi \rightarrow \mu^+\mu^-$ branching ratio the value
1417 reported in the PDG book, $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$ [2], is used.
1418 Table 7.3 summarises the total relative efficiencies, ε^{rel} , for downstream and long
1419 candidates together with their correlated and uncorrelated uncertainties, where the
1420 correlation is intended between the downstream and long samples. On the table
1421 the uncorrelated uncertainty corresponds to the total systematic uncertainty on the
1422 efficiency determination. The correlated uncertainty is given in percent form since
1423 it can be applied to either downstream, long candidates or their combination. This
1424 includes the PDF systematic described in Sec. 7.1 and the systematic due to the
1425 uncertainty on the $J/\psi \rightarrow \mu^+\mu^-$ branching fraction.

1426 Figure 7.3 shows the branching fraction obtained by fitting the downstream and
1427 long samples independently, while the combined result, obtained fitting both sam-
1428 ples simultaneously, is shown in Fig. 7.4. Values are also listed in Tab. 7.4, where

Table 7.3: Absolute values of the total relative efficiency and the absolute value of the uncorrelated uncertainty, together with relative values of the correlated uncertainty.

q^2 interval [GeV $^2/c^4$]	Eff. (DD)	σ_{uncorr}^{DD}	Eff. (LL)	σ_{uncorr}^{LL}	Correlated err.
0.1–2.0	0.694	0.058	1.136	0.066	1.012%
2.0–4.0	0.693	0.027	0.907	0.047	2.697%
4.0–6.0	0.699	0.018	0.964	0.044	2.697%
6.0–8.0	0.733	0.020	0.953	0.048	2.697%
11.0–12.5	1.254	0.032	1.140	0.057	3.356%
15.0–16.0	1.260	0.035	1.035	0.060	2.977%
16.0–18.0	1.163	0.029	0.997	0.048	1.727%
18.0–20.0	1.023	0.027	0.782	0.040	2.697%
1.1–6.0	0.696	0.032	0.950	0.058	1.012%
15.0–20.0	1.132	0.014	0.927	0.031	1.423%

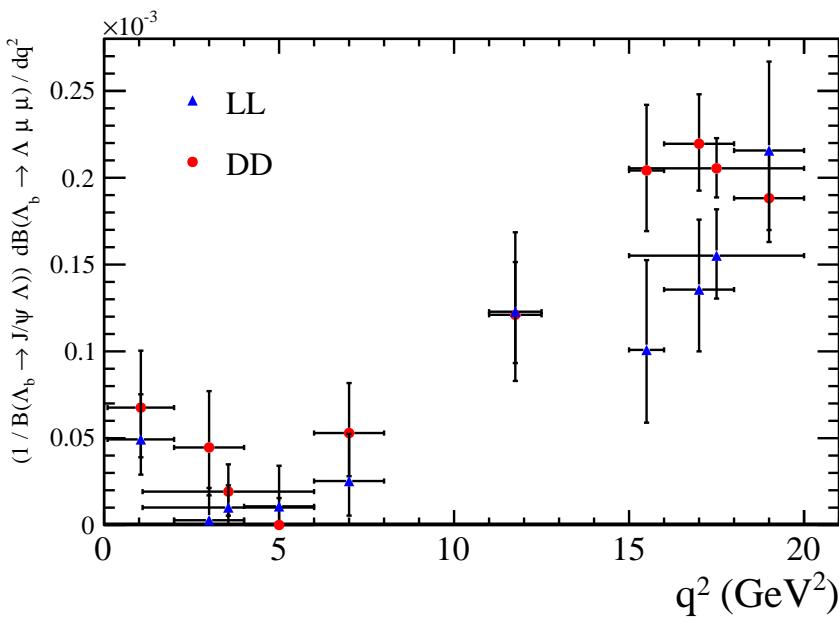


Figure 7.3: Measured values of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction relative to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay as a function of q^2 obtained fitting the downstream and long samples independently. Error bars represent the total statistical and systematic uncertainty.

the statistical uncertainty on the rare channel (stat) and the total systematic uncertainty (stat) are shown separately. The statistical uncertainty is calculated using the MINOS tool, which returns 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 translate into a different branching fraction and the full difference with respect to the default fit is taken as systematic uncertainty in each direction.

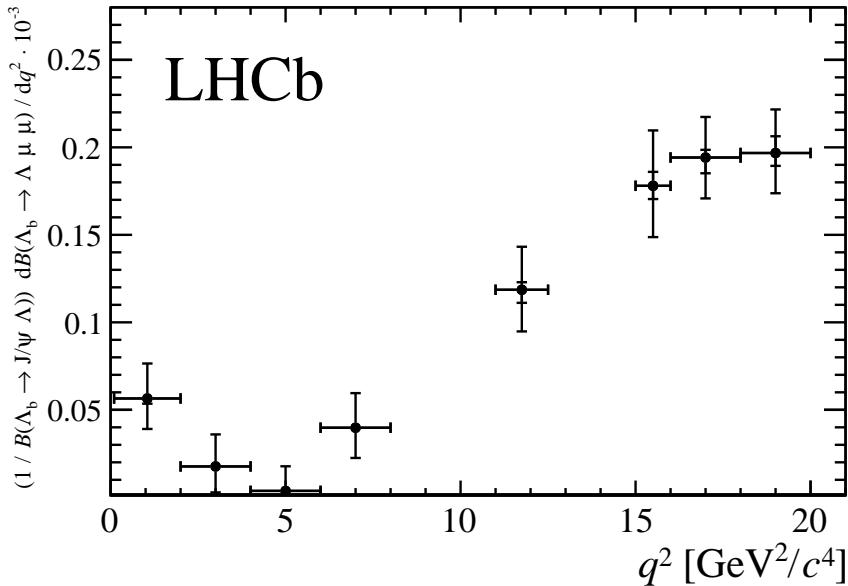


Figure 7.4: Branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay normalised to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ mode. The inner error bar represents the systematic uncertainty and the outer error bar includes the statistical uncertainty.

Finally, values for the absolute branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay are obtained by multiplying the relative branching fraction by the absolute branching fraction of the normalisation channel, $\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) = (6.3 \pm 1.3) \times 10^{-4}$ [2]. Values are shown in Fig. 7.5 and summarised in Tab. 7.5, 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. [83].

Evidence for the signal is found for the first time in the q^2 region between the charmonium resonances and in the interval $0.1 < q^2 < 2.0$ GeV^2/c^4 , where an increased yield is expected due to the proximity of the photon pole. The uncertainty on the

Table 7.4: Differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay relative to $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays, where the uncertainties are statistical and systematic, respectively.

q^2 interval [GeV^2/c^4]	$\frac{d\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-)/dq^2}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda)} \cdot 10^{-3} [(\text{GeV}^2/c^4)^{-1}]$		
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

¹⁴⁴⁵ absolute branching fraction is dominated by the precision with which the branch-
¹⁴⁴⁶ ing fraction for the normalisation channel is known, while the uncertainty on the
¹⁴⁴⁷ relative branching fraction it is dominated by the size of the data sample available.
¹⁴⁴⁸ The data are consistent with the theoretical predictions in the high- q^2 region but lie
¹⁴⁴⁹ below the predictions in the low- q^2 region.

Table 7.5: Measured differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay, where the uncertainties are statistical, systematic and due to the uncertainty on the normalisation mode, $\Lambda_b^0 \rightarrow J/\psi \Lambda$, respectively.

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

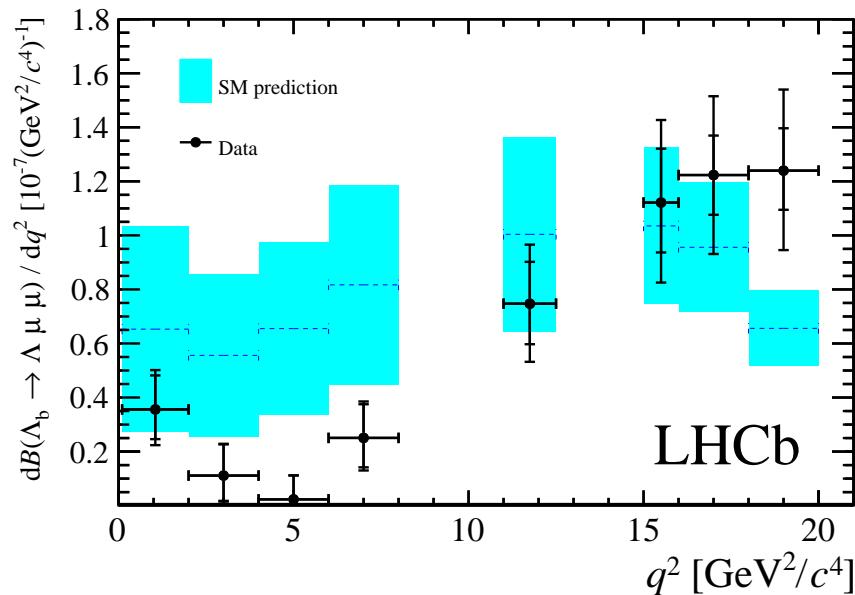


Figure 7.5: Measured $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction as a function of q^2 with the SM predictions [83] superimposed. The inner error bars on data points represent the total uncertainty on the relative branching fraction (statistical and systematic); the outer error bar also includes the uncertainties from the branching fraction of the normalisation mode.

CHAPTER 8

Angular analysis

Neglecting Λ_b^0 production polarisation, the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ 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 chapters and are defined in Fig. 8.1, where θ_ℓ is the angle between the positive (negative) muon direction and the dimuon system direction in the Λ_b^0 ($\bar{\Lambda}_b^0$) rest frame, and θ_h is defined 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. The following chapters describe a measurement of two forward-backward asymmetries in the leptonic, A_{FB}^ℓ , and in the hadronic, A_{FB}^h , systems. These forward-backward asymmetries are defined as

$$A_{\text{FB}}^i(q^2) = \frac{\int_0^1 \frac{d^2\Gamma}{dq^2 d\cos\theta_i} d\cos\theta_i - \int_{-1}^0 \frac{d^2\Gamma}{dq^2 d\cos\theta_i} d\cos\theta_i}{d\Gamma/dq^2}, \quad (8.1)$$

¹⁴⁵⁴ where $d^2\Gamma/dq^2 d\cos\theta_i$ is the two-dimensional differential rate and $d\Gamma/dq^2$ is rate
¹⁴⁵⁵ integrated over the angles.

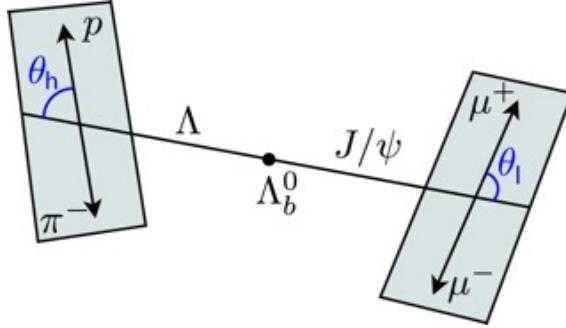


Figure 8.1: Graphical representation of the angles for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay.

¹⁴⁵⁶ The A_{FB}^ℓ observable was also measured by LHCb in $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays, going
¹⁴⁵⁷ through the same quark traditions as $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays. Instead the hadronic
¹⁴⁵⁸ asymmetry, A_{FB}^h , is interesting only in the Λ_b^0 case as it is zero by definition in B^0
¹⁴⁵⁹ decays, where K^{*0} decays strongly.

¹⁴⁶⁰ 8.1 One-dimensional angular distributions

¹⁴⁶¹ This section describes the derivation of the functional form of the angular distribu-
¹⁴⁶² tions as a function of the $\cos\theta_\ell$ and $\cos\theta_h$, which are used to measure the observables.
¹⁴⁶³ The content of this section is based on the calculations in Ref. [79].

¹⁴⁶⁴ For unpolarised Λ_b^0 production, integrating over the three angles the differential
¹⁴⁶⁵ branching fraction is given in Eq. V of Ref. [79] as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2} = \frac{v^2}{2} \cdot \left(U^{V+A} + L^{V+A} \right) + \frac{2m_\ell^2}{q^2} \cdot \frac{3}{2} \cdot \left(U^V + L^V + S^A \right), \quad (8.2)$$

¹⁴⁶⁶ and the lepton helicity angle θ_ℓ differential distribution, given in Eq. 15, has the

¹⁴⁶⁷ form

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_\ell} &= v^2 \cdot \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) \cdot \frac{1}{2} U^{V+A} + \frac{3}{4} \sin^2 \theta_\ell \cdot \frac{1}{2} L^{V+A} \right] \\ &- v \cdot \frac{3}{4} \cos \theta_\ell \cdot P^{VA} + \frac{2m_\ell^2}{q^2} \cdot \frac{3}{4} \cdot \left[U^V + L^V + S^A \right]. \end{aligned} \quad (8.3)$$

¹⁴⁶⁸ In these formulas m_ℓ is the mass of the lepton and $v = \sqrt{1 - 4m_\ell^2/q^2}$, U denotes
¹⁴⁶⁹ the unpolarised-transverse contributions, L the longitudinal contributions and S
¹⁴⁷⁰ the scalar contribution. The apices V and A represent respectively vector and axial-
¹⁴⁷¹ vector currents, with $X^{V+A} = X^V + X^A$. The authors of Ref. [79] define then the
¹⁴⁷² lepton-side forward-backward asymmetry as

$$A_{FB}^\ell(q^2) = -\frac{3}{2} \frac{v \cdot P^{VA}}{v^2 \cdot (U^{V+A} + L^{V+A}) + \frac{2m_\ell^2}{q^2} \cdot 3 \cdot (U^V + L^V + S^A)}. \quad (8.4)$$

Using these results as a starting point one can rewrite Eq. 8.3 as

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_\ell} &= \frac{3}{8} \frac{d\Gamma}{dq^2} (1 + \cos^2 \theta_\ell) U^{V+A} + \frac{d\Gamma}{dq^2} A_{FB}^\ell \cos \theta_\ell + \frac{3}{8} \sin^2 \theta_\ell v^2 (L^{V+A}) \\ &\quad (U^V + L^V + S^A) \frac{3m_\ell^2}{q^2} \left(\frac{1}{8} - \frac{3}{8} \cos^2 \theta_\ell \right) \end{aligned} \quad (8.5)$$

¹⁴⁷³ For this analysis the massless leptons limit, $m_\ell \rightarrow 0$, is used, which is a good
¹⁴⁷⁴ approximation except at very low q^2 . In the massless limit the differential rates
¹⁴⁷⁵ simplify to

$$\frac{d\Gamma}{dq^2} = \frac{v^2}{2} \cdot \left(U^{V+A} + L^{V+A} \right) \quad (8.6)$$

and

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{3}{8} \frac{d\Gamma}{dq^2} (1 + \cos^2 \theta_\ell) U^{V+A} + \frac{d\Gamma}{dq^2} A_{FB}^\ell \cos \theta_\ell + \frac{3}{8} v^2 \sin^2 \theta_\ell (L^{V+A}). \quad (8.7)$$

Equations 8.6 and 8.7 can be then combined to achieve the form

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) \frac{U^{V+A}}{U^{V+A} + L^{V+A}} + A_{\text{FB}}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell \frac{L^{V+A}}{U^{V+A} + L^{V+A}} \right]. \quad (8.8)$$

¹⁴⁷⁶ The amplitude combination in the last term can be viewed as ratio between longitudinal and sum of longitudinal and unpolarized transverse contributions and therefore
¹⁴⁷⁷ one can define the longitudinal fraction
¹⁴⁷⁸

$$f_L = \frac{L^{V+A}}{U^{V+A} + L^{V+A}}, \quad (8.9)$$

which leads to the distribution used in the analysis

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) (1 - f_L) + A_{\text{FB}}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell f_L \right]. \quad (8.10)$$

¹⁴⁷⁹ Using the same steps the proton helicity distribution is given in Ref. [79] as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda(\rightarrow p\pi^-)\ell^+\ell^-)}{dq^2 d \cos \theta_h} = \text{Br}(\Lambda \rightarrow p\pi^-) \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+\ell^-)}{dq^2} \left(\frac{1}{2} + A_{\text{FB}}^h \cos \theta_h \right), \quad (8.11)$$

¹⁴⁸⁰ and A_{FB}^h is defined as

$$A_{\text{FB}}^h = \frac{1}{2} \alpha_A P_z^A(q^2), \quad (8.12)$$

¹⁴⁸¹ where $P_z^A(q^2)$ is the polarisation of the daughter baryon, Λ , and $\alpha_A = 0.642 \pm 0.013$ [2]
¹⁴⁸² is the Λ decay asymmetry parameter.

¹⁴⁸³ These expressions assume that Λ_b^0 is produced unpolarised, which is in agreement
¹⁴⁸⁴ with the recent LHCb measurement of the production polarisation [86]. Possible
¹⁴⁸⁵ effects due to a non zero production polarisation are investigated as systematic
¹⁴⁸⁶ uncertainties (see Sec. 10.5).

¹⁴⁸⁷ 8.2 Multi-dimensional angular distributions

To incorporate effects of production polarisation this was introduced in the equations. In the modified version the angle θ is defined as the polar angle between the Λ direction in the Λ_b^0 rest frame with respect to $\hat{n} = \hat{p}_{inc} \times \hat{p}_{\Lambda_b^0}$, where \hat{p}_{inc} represents the direction of the incoming proton. This angle is sensitive to the production polarisation through the spin-density matrix in Eq. A.3. Integrating over all the angles but θ_ℓ results in the same distribution as in the unpolarised case (Eq. 8.3). Therefore, in the case of uniform efficiency, the lepton side forward-backward asymmetry, A_{FB}^ℓ , is unaffected by the production polarisation. To estimate 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)

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d\cos\theta d\cos\theta_\ell} = \frac{d\Gamma}{dq^2} \left\{ \frac{3}{8} (1 + \cos^2\theta_\ell) (1 - f_L) + A_{FB}^\ell \cos\theta_\ell + \frac{3}{4} \sin^2\theta_\ell f_L + P_b \cos\theta \left[-\frac{3}{4} \sin\theta_\ell^2 O_{Lp} + \frac{3}{8} (1 + \cos\theta_\ell^2) O_P - \frac{3}{8} \cos\theta_\ell O_{UVA} \right] \right\}, \quad (8.13)$$

where three more observables are defined

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

¹⁴⁸⁸ In the massless leptons approximation two of these quantities are related to the hadron side forward-backward asymmetry as

$$\frac{1}{2} \alpha_\Lambda (O_P + O_{Lp}) = A_{FB}^h. \quad (8.14)$$

$q^2 [GeV^2/c^2]$	A_{FB}^ℓ	P_z^A	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 8.1: Prediction for angular observables entering two-dimensional angular distributions. Prediction is based on covariant quark model form factors from Ref. [79].

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_{FB}^h , is also unaffected by the production polarisation in case of uniform efficiency. The differential distribution in θ and θ_h has the form

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d(\cos \theta) d(\cos \theta_h)} = \frac{d\Gamma}{dq^2} [1 + 2A_{FB}^h \cos \theta_h + P_b (O_P - O_{Lp}) \cos \theta + \alpha_A P_b (1 - 2f_L) \cos \theta \cos \theta_h]. \quad (8.15)$$

In order to use these distributions, expectations for the three additional observables, which do not enter one-dimensional distributions, are needed. Expectations are calculated using form factors and numerical inputs from Ref. [79] and are shown in Tab. 8.1.

For completeness, the differential distribution in $\cos \theta_\ell$ and $\cos \theta_h$ has the form

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_h d \cos \theta_\ell} = & \frac{3}{8} + \frac{6}{16} \cos^2 \theta_\ell (1 - f_L) - \frac{3}{16} \cos^2 \theta_\ell f_L + A_{FB}^l \cos \theta_\ell + \\ & \left(\frac{3}{2} A_{FB}^h - \frac{3}{8} \alpha_A O_P \right) \cos \theta_h - \frac{3}{2} A_{FB}^h \cos^2 \theta_\ell \cos \theta_h - \frac{3}{16} f_L + \\ & \frac{9}{16} f_L \sin^2 \theta_\ell + \frac{9}{8} \alpha_A \cos^2 \theta_\ell \cos \theta_h O_P - \\ & \frac{3}{2} \alpha_A \cos \theta_\ell \cos \theta_h O_{UVA}. \end{aligned} \quad (8.16)$$

¹⁴⁹⁴

8.3 Angular resolution

¹⁴⁹⁵ This section describes a study of the angular resolution done in order to achieve a
¹⁴⁹⁶ better understanding of detector and reconstruction effects. This will be then used to
¹⁴⁹⁷ study systematic uncertainties (see Sec. 10.5). The study is done by analysing simu-
¹⁴⁹⁸ lated events and comparing generated and reconstructed quantities. Figure 8.2 con-
¹⁴⁹⁹ tains plots of the difference between true and measured angular observables ($\cos \theta_\ell$
¹⁵⁰⁰ and $\cos \theta_h$) as a function of the observable itself. These are centred at zero indicat-
¹⁵⁰¹ ing no bias in the measurement. In Fig. 8.2 the same difference is shown also as a
¹⁵⁰² function of q^2 showing again no bias. The spread of these distributions around the
¹⁵⁰³ central value is an estimate of the angular resolution. Taking vertical slices of the
¹⁵⁰⁴ distributions in Fig. 8.2 one obtains approximately gaussian distributions centred at
¹⁵⁰⁵ zero. These distributions are fit with a single gaussian and its width is interpreted
¹⁵⁰⁶ as angular resolution. Table 8.2 reports the average resolutions for the two angular
¹⁵⁰⁷ variables separately for the long and downstream candidates. As expected candi-
¹⁵⁰⁸ dates built from long tracks are characterised by a better resolution due to a better
¹⁵⁰⁹ momentum and vertex resolution. Figure 8.3 shows response matrices, representing
the correlation between reconstructed and generated angular observables.

Table 8.2: Average angular resolutions integrated over the full interval and the full available q^2 .

Observable	DD	LL
$\cos \theta_\ell$	0.015	0.010
$\cos \theta_h$	0.066	0.014

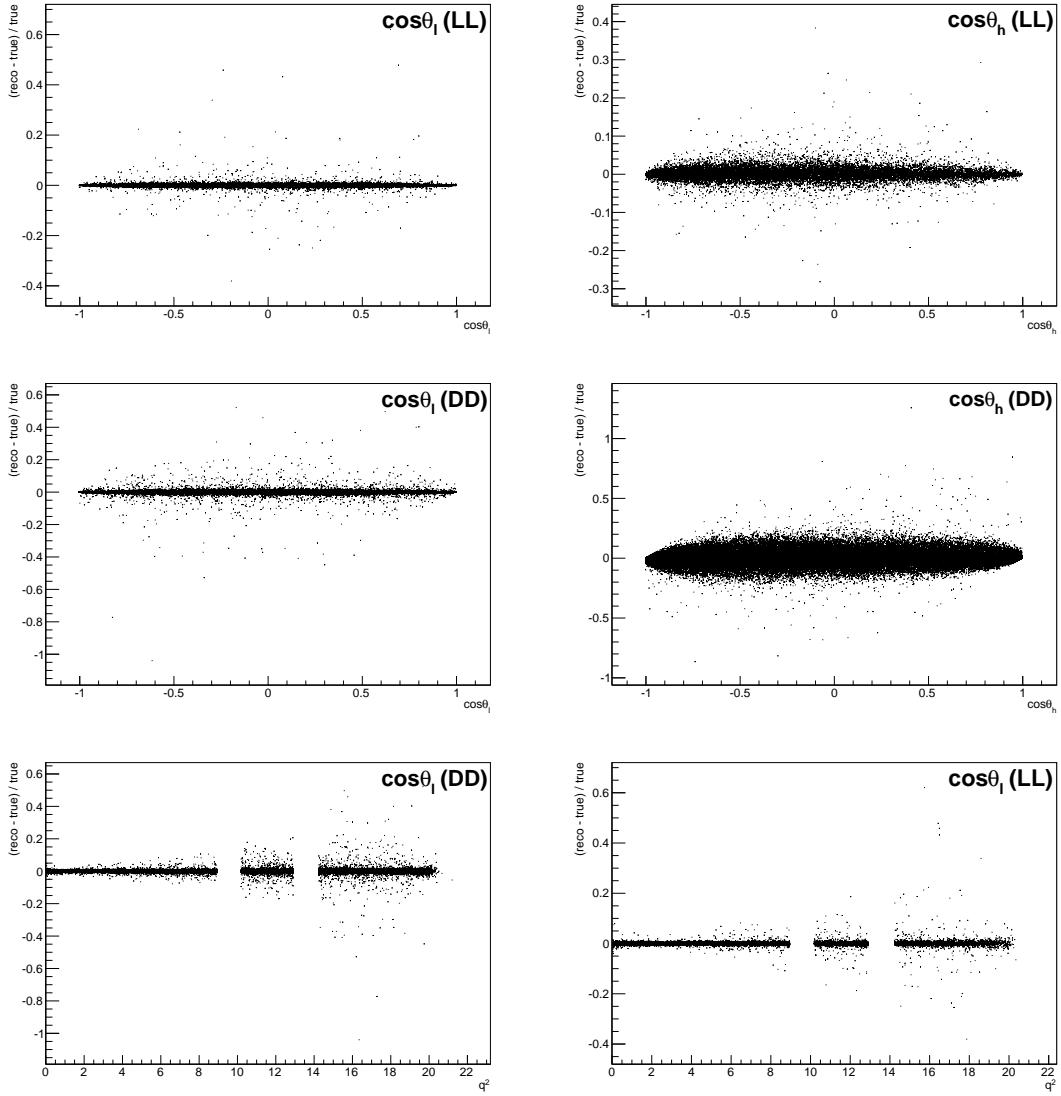


Figure 8.2: Difference of between generated and reconstructed angular observables as a function of the observables themselves for long (top) and downstream (bottom) candidates and as a function of q^2 for long (bottom left) and downstream (bottom right) candidates. As the plots are made using fully selected rare samples the bottom plots present empty bands corresponding to the charmonium vetoes.

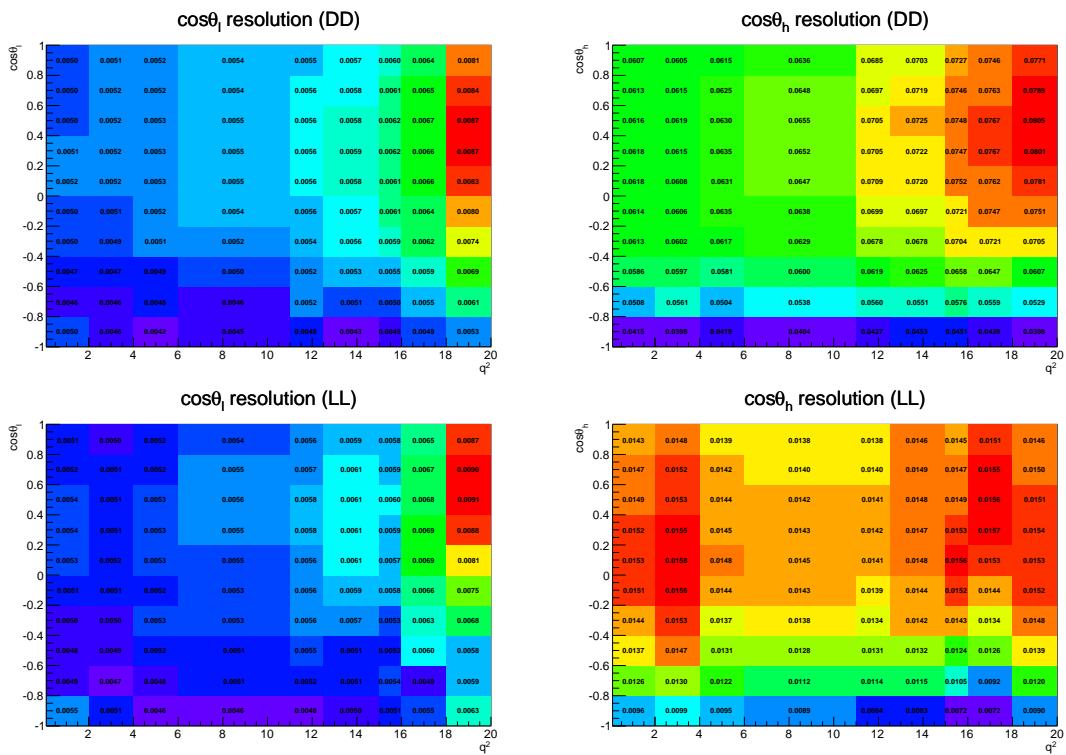


Figure 8.3: Angular resolution for $\cos\theta_\ell$ (left plots) and $\cos\theta_h$ (right plots) as a function of the angular variables and q^2 for downstream (upper plots) and long (lower plots) candidates. White bands correspond to the J/ψ and $\psi(2S)$ resonances which are excluded from the study.

CHAPTER 9

1511

1512

1513

Angular fit strategy

1514

1515 This chapter describes the fitting technique applied to extract the angular observ-
1516 ables. There are physical limits to the values of the parameters of interests: A_{FB}^h is
1517 limited in the $[-0.5, 0.5]$ interval and for the f_L and A_{FB}^ℓ parameters the physical re-
1518 gion, given by $|A_{\text{FB}}^\ell| < 3/4(f_L - 1)$, is the triangle shown in Fig. 9.1. If the measured
1519 value is close to the border the fit does not always converge. For this reason a "brute
1520 force" fitting technique is applied. Fit parameters are divided into two categories:
1521 parameters of interest (PoIs), A_{FB}^ℓ , A_{FB}^h and f_L and all other parameters, which are
1522 referred to as "nuisances". The value of the Log-Likelihood, $\log \mathcal{L}$, of the fit model
1523 with respect to data is evaluated in a grid of points in the PoIs allowed area to find
1524 the function minimum. A first coarse scan finds a candidate minimum and then the
1525 procedure is reiterated two more times in finer intervals around it. For each point
1526 all the nuisances are fitted using a maximum likelihood fit. Using this method the
1527 fit is therefore constrained inside the physical region, if the best log-likelihood is
1528 found to be outside it, the point at the boundary is chosen as the best fit.

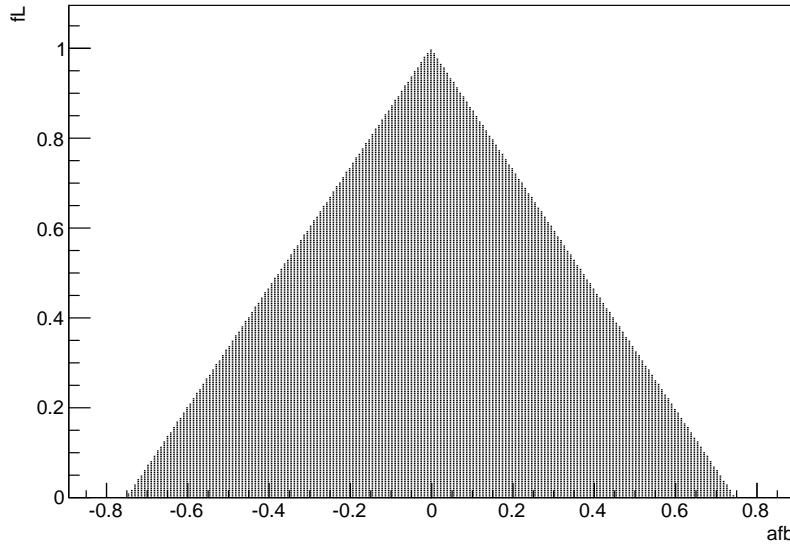


Figure 9.1: The physical (A_{FB}^{ℓ}, f_L) parameter space. The dark region corresponds to points where the PDF is positive in the whole $[-1, 1]$ interval.

¹⁵²⁹ 9.1 Feldman-cousins plug-in method

¹⁵³⁰ Physical boundaries of the parameter space could result in a wrong estimation of
¹⁵³¹ the uncertainties, especially if the measured value is close to the border. To deal
¹⁵³² with this effects the likelihood-ordering method [87] is used to estimate uncertain-
¹⁵³³ ties in this analysis and nuisance parameters are accounted for using the plug-in
¹⁵³⁴ method [88]. This is a unified method to calculate confidence intervals and up-
¹⁵³⁵ per/lower limits, based on simulated experiments and has the advantage of having
¹⁵³⁶ a well defined frequentist coverage.

¹⁵³⁷ The method is constituted by the following steps:

- ¹⁵³⁸ 1. fit real data distributions with all parameters free;
- ¹⁵³⁹ 2. fit real data fixing the PoIs to a value of choice and keeping nuisance param-
¹⁵⁴⁰ eters free;
- ¹⁵⁴¹ 3. generate simulated samples following the distribution given by the fit model,
¹⁵⁴² where all nuisance parameters are taken from the fit in point 2 and PoIs are

1543 fixed to the same value used in point 2;

1544 4. repeat the two fits made on data on each simulated sample: fit with all pa-
1545 rameters free and with fixed PoIs;

1546 5. extract the value of the Log-Likelihoods at the minimum for all cases;

1547 6. calculate the percentage of simulated experiments in which the ratio $\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free}$
1548 is bigger than in data.

1549 7. repeat the procedure for many values of the PoIs scanning around the best fit
1550 values.

1551 The confidence interval at $k\%$ is given by the points where the free-to-fixed likelihood
1552 ratio is bigger in data than simulation for $(1 - k)\%$ of times. As an example, in
1553 Fig. 9.2 are reported the p-values obtained with the plug-in method for A_{FB}^{ℓ} and
1554 f_L . For the analysis the two-dimensional allowed region is scanned giving a grid of
p-values, which translated into two-dimensional confidence regions.

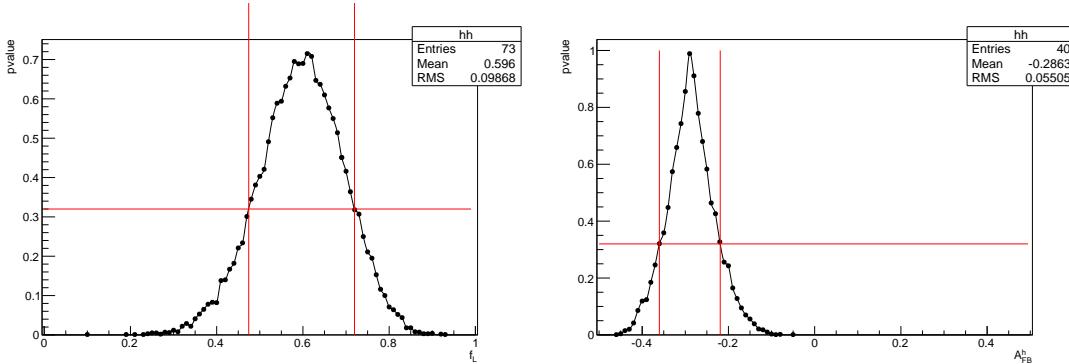


Figure 9.2: Dependence of the p-value from the values of the angular observables f_L (left) and A_{FB}^h (right) in simulated experiments. The red lines mark the points at p-value 32% corresponding to a 68% CL.

1556 9.2 Modelling the angular distributions

1557 The observables are extracted from fits to one-dimensional angular distributions.

1558 The PDFs used to model the data are defined as

$$P^k(\cos \theta_{\ell/h}) = f_b P_S(\cos \theta_{\ell/h}) \times \varepsilon^k(\cos \theta_{\ell/h}) + (1 - f_b) P_B^k(\cos \theta_{\ell/h}), \quad (9.1)$$

1559 where $k = (\text{LL}, \text{DD})$, P_S is the signal function composed by a theoretical shape given
1560 by Eq. 8.11 and 8.10, which is multiplied by an acceptance function ε described
1561 in Sec. 9.3 and P_B is a background component. To limit systematic effects due to
1562 the background parameterisation, the fit is performed in a restricted mass region
1563 around the peak: $5580 < m(\Lambda\mu^+\mu^-) < 5660 \text{ MeV}/c^2$ (“signal region”), which is
1564 dominated by the signal. The background fraction, f_b , is obtained by looking at
1565 the 4-body $m(p\pi\mu\mu)$ invariant mass distribution in a wider interval and fitting it to
1566 extract the fraction of background in the signal region. In the fit to the angular dis-
1567 tributions this is then gaussian constrained to the obtained value. The background
1568 shape is parameterised with a linear function times the efficiency shape. A different
1569 efficiency shape is used for downstream and long events and for each q^2 interval.
1570 The free parameter of this model is fitted on sideband candidates which contain
1571 only background and fixed for the fit to the signal region. Figure 9.3 reports the
1572 background distributions in the sideband, $m(p\pi\mu^+\mu^-) > 5700 \text{ MeV}/c^2$, for the high
1573 q^2 integrated interval with overlaid the background function. In summary the only
1574 fit parameter in the total fit function is the forward-backward asymmetry (and f_L
1575 in the leptonic case).

1576 9.3 Angular acceptance

1577 Selection requirements on the minimum momentum of the muons may distort the
1578 $\cos \theta_\ell$ distribution by removing candidates with extreme values of $\cos \theta_\ell$. Similarly,
1579 the impact parameter requirements affect $\cos \theta_h$ as very forward hadrons tend to

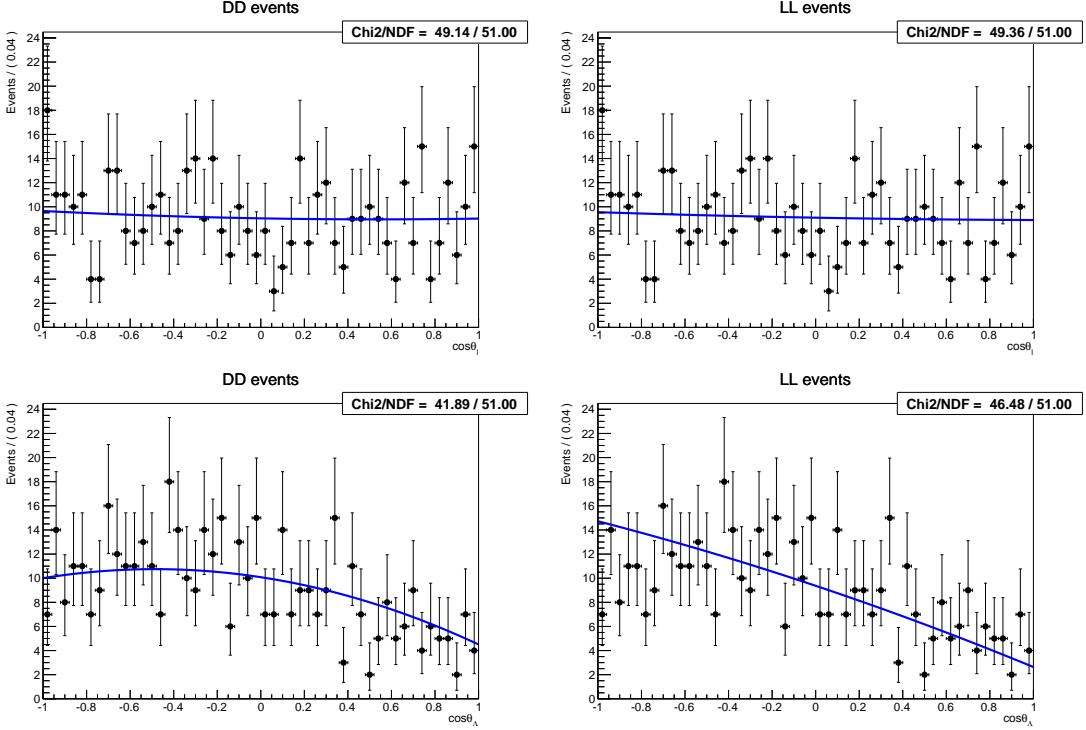


Figure 9.3: Background distribution as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

have smaller impact parameter values. While in principle one could take it 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, with an independent set of parameters obtained for each q^2 interval. These parameters are fixed for the fits to data. Using polynomial functions allows to calculate the PDF normalisation analytically. Figure 9.4 shows total efficiency as a function of $\cos \theta_h$ and $\cos \theta_\ell$ in the $15.0\text{--}20.0$ integrated q^2 interval obtained using a $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ 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.

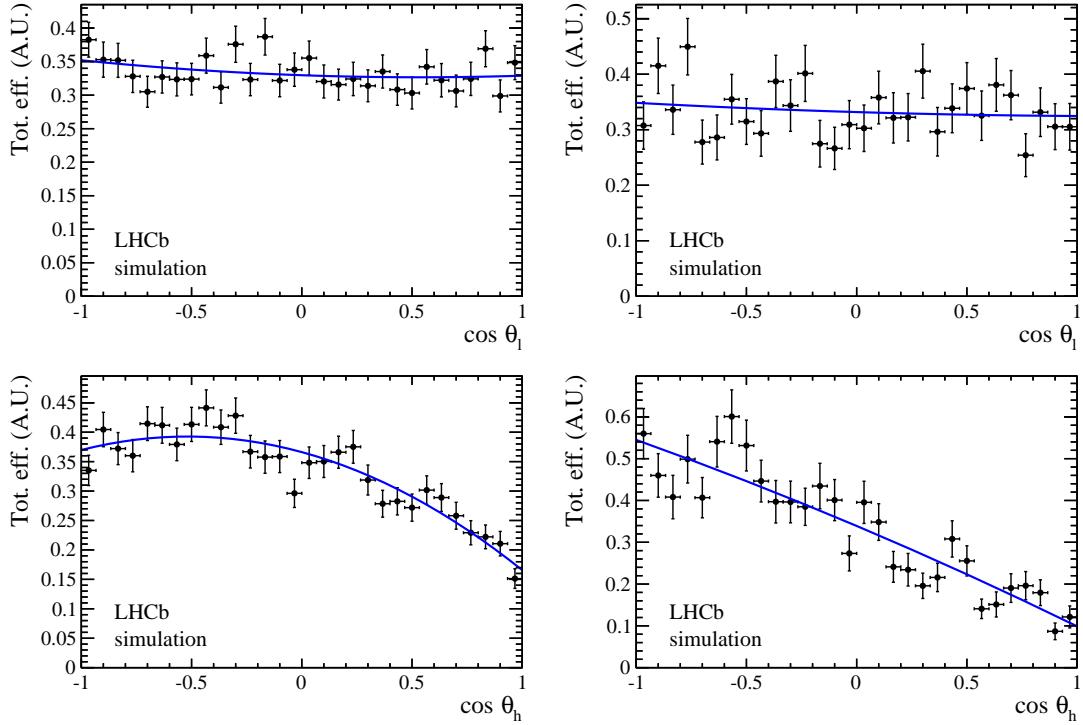


Figure 9.4: Efficiency as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

1591 9.4 Studies on a three-dimensional fit

1592 One other way of extracting the angular observables would be to fit at the same
 1593 time both angles and also the invariant mass distribution in order to have a better
 1594 handle on the level of background. In this case one can use more of the information
 1595 available. On the other hand it is necessary to use a larger mass window including
 1596 more background and this method involves more parameters to fit. In the 1D case
 1597 the free parameters are the two parameters of interest (A_{FB}^ℓ and f_L) for the lepton
 1598 case and one (A_{FB}^h) for the hadron one. For the 3D case the free parameters are the
 1599 three parameters of interest plus two background fractions and the two exponential
 1600 slopes for the invariant mass background. A high number of free parameters is
 1601 difficult to constrain with the very limited statistics available. Furthermore, to take
 1602 correctly into account correlations 3 more observables enter the fit. To check which
 1603 method gives the best sensitivity 500 pseudo-experiments are generated. Events are
 1604 generated in a 3D $(\cos \theta_\ell, \cos \theta_h, m)$ space using shapes taken from the fit on real

Table 9.1: RMS values for toy experiments on the extraction of the three parameters of interest with the 1D or 3D fitting methods.

q^2 [GeV $^2/c^4$]	Fit type	A_{FB}^h	A_{FB}^ℓ	f_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

1605 data. The generated values of the parameters of interest are $A_{FB}^\ell = 0$, $f_L = 0.7$ and
 1606 $A_{FB}^h = -0.37$. These are data-like values inspired to a preliminary measurement in
 1607 the highest statistics interval. The overall statistics and the fraction of background
 1608 events in the mass window are generated using the values found from the 1D fit
 1609 on data. Each pseudo-experiment is fitted with both methods and Fig. 9.5 reports
 1610 distributions of parameters of interest obtained from the fit in the 1D and 3D cases.
 1611 The RMS of these distributions can be taken as a measure of the sensitivity of each
 1612 method. In Tab. 9.1 RMSs from both methods can be compared. For all parameters
 1613 of interest the 1D fit method gives a smaller RMS, hence a better sensitivity.

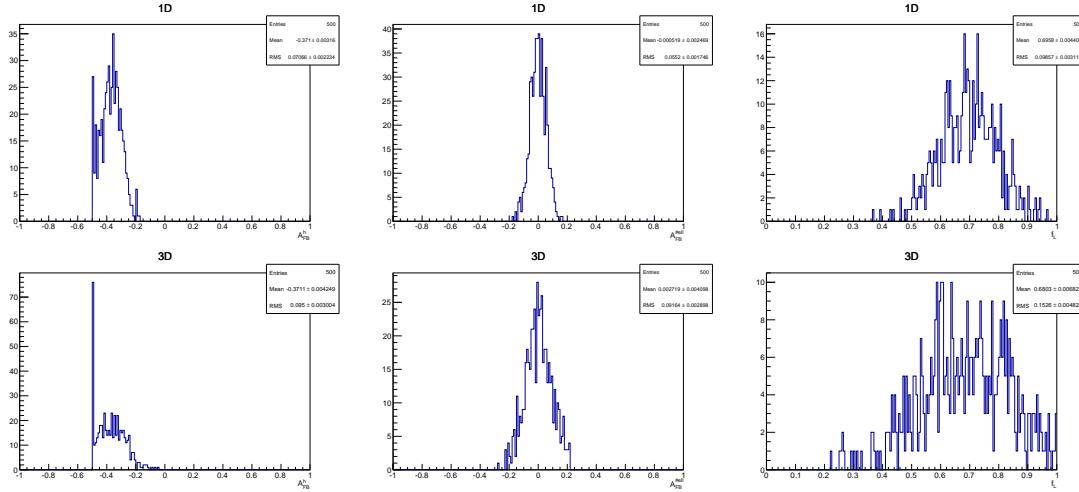


Figure 9.5: Distribution of observed parameters of interest over 500 pseudo-experiments obtained using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what is observed in the 15–20 q^2 interval.

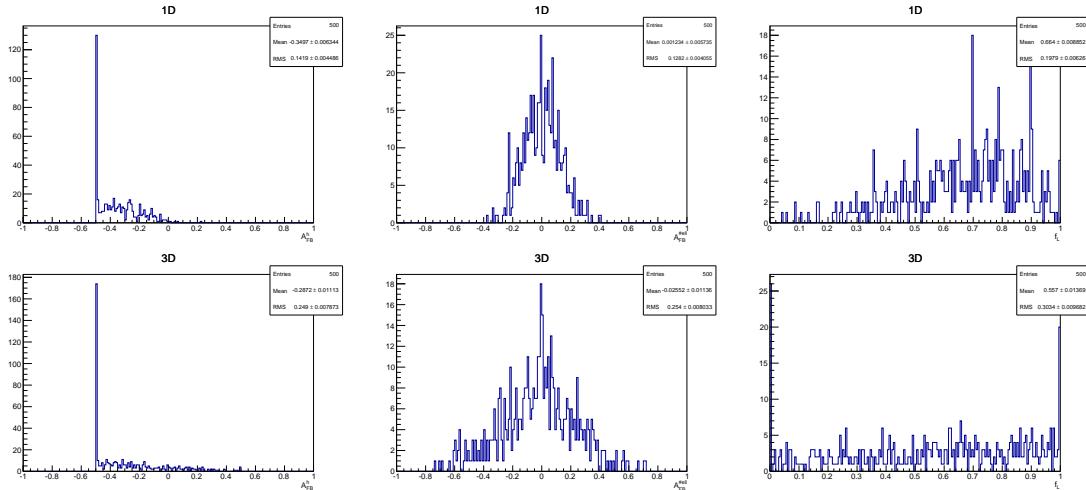


Figure 9.6: Distribution of observed parameters of interest over 500 pseudo-experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the 11–12.5 q^2 interval.

1614

CHAPTER 10

1615

1616 Systematics uncertainties on angular observables and results

1617

1618 The following sections describe the five main sources of systematic uncertainties
1619 that are considered for the angular observables measurement and, finally, results
1620 are reported in Sec. 10.7. Results are derived only for q^2 intervals where the signal
1621 significance, shown in Tab. 5.3, is above 3 standard deviations. This includes all
1622 q^2 intervals above the J/ψ resonance and the lowest q^2 interval, where an increased
1623 yield is due to the presence of the photon pole.

1624 10.1 Angular correlations

1625 The angular efficiency is non-flat as a function of $\cos \theta_\ell$ and $\cos \theta_h$. Therefore, while
1626 integrating the full angular distribution, terms that cancel with perfect efficiency
1627 may remain and generate a bias in the final result. In order to deal with this effect
1628 simulated events are generated in a two-dimensional $(\cos \theta_\ell, \cos \theta_h)$ space according

1629 to the theoretical distribution described by Eq. 8.16 multiplied by two-dimensional
 1630 efficiency histograms obtained from simulation and reported in Fig. 10.1. Then,
 1631 one-dimensional projections are taken and fit using the default one-dimensional ef-
 1632 ficiency functions. Figure 10.2 shows the distribution of observed deviations from
 1633 the generated value, $\Delta x = x_{true} - x_{measured}$. Since the mean of these distributions
 1634 is non-zero by more than 3σ , they are taken as systematic uncertainties.

1635 10.2 Resolution

1636 The angular resolution could bias the observables measurement generating an asym-
 1637 metric migration of events. This is especially important in the $\cos \theta_h$ case, because
 1638 this has worse resolution and a considerably asymmetric distribution. Simulated
 1639 experiments are used to asses this systematic. Events are generated according to
 1640 the measured distributions (including efficiencies). The generated events are then
 1641 smeared by the angular resolution (gaussian smearing). To be conservative the case
 1642 with biggest angular resolution (downstream candidates), is always used. Finally,
 1643 the smeared and not-smeared distributions are fitted with the same PDF. The av-
 1644 erage deviation from the default values are reported in Tab. 10.1 as a function of q^2
 1645 and assigned as systematic uncertainties.

Table 10.1: 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	ΔA_{FB}^ℓ	Δf_L	ΔA_{FB}^h
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

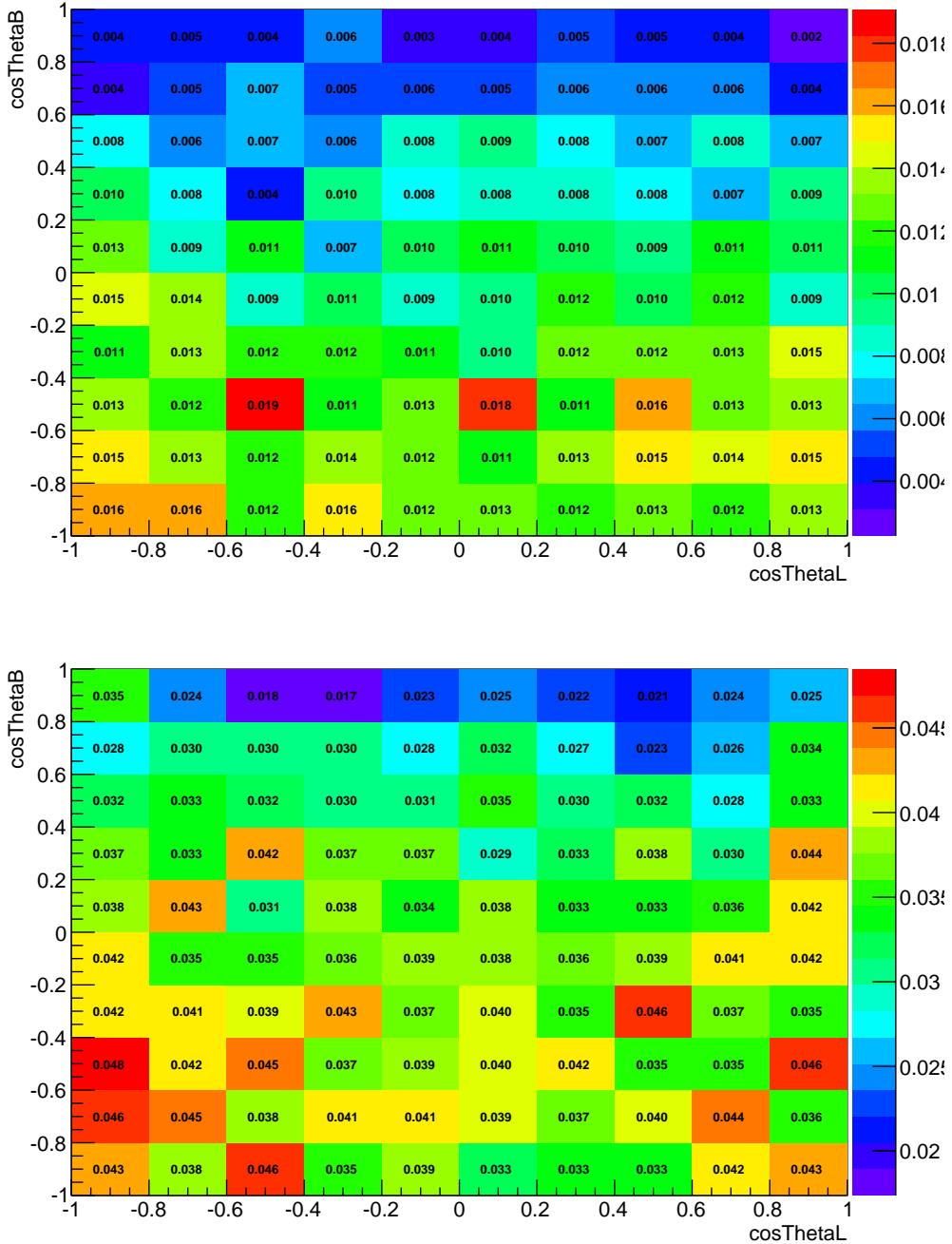


Figure 10.1: Angular acceptance as a function of $\cos \theta_\ell$ and $\cos \theta_h$ for long (top) and downstream (bottom) candidates, integrated over the full available q^2 range.

10.3 Efficiency description

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

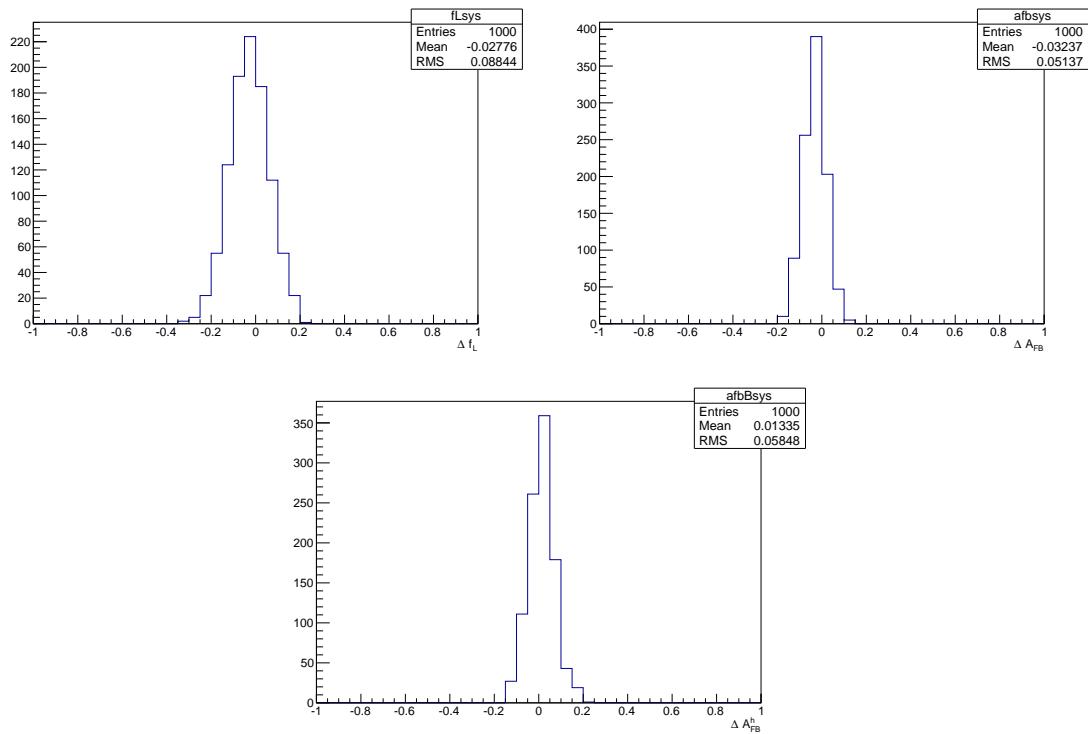


Figure 10.2: Deviations of the observables' values obtained fitting simulated events generated with a 2D distribution multiplied by a 2D efficiency and fitting 1D projections with respect to generated values. For f_L (top left), A_{FB}^ℓ (top right) and A_{FB}^h (bottom).

1649 kinematic re-weighting described in Se. 3.3.2 is removed from the simulation and
1650 the efficiency is determined again. Simulated events are then fit using the same the-
1651 oretical PDF and multiplied by the efficiency function obtained with and without
1652 kinematical weights. As in the previous cases the average bias is taken as system-
1653 atic uncertainty. Results are shown in Tab. 10.2. Furthermore, for the effect of
1654 the limited simulated statistics is taken into account and added to the systematic
uncertainty.

Table 10.2: 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_{FB}^h	A_{FB}^ℓ	f_L
0.1–2.0	0.0093	0.0020	0.0440
11.0–12.5	0.0069	0.0069	0.0027
15.0–16.0	0.0109	0.0018	0.0046
16.0–18.0	0.0159	0.0012	0.0043
18.0–20.0	0.0148	0.0030	0.0017
15.0–20.0	0.0138	0.0002	0.0046

Table 10.3: 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_{FB}^ℓ	f_L	A_{FB}^h
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

1655

1656 10.4 Background parameterisation

1657 There is a certain degree of arbitrariness in the choice of a parameterisation for the
1658 background, especially in q^2 intervals with low statistics. To assess possible biases
1659 due to the PDF choice, simulated experiments are generated using the shapes from
1660 data fits and the same statistics as observed in data for each q^2 interval. Each
1661 pseudo-experiment is fit with two models: the default one, a “line times efficiency”

1662 function and the efficiency function alone, corresponding to the assumption that
 1663 background distributions are originally flat and only modified by the interaction
 1664 with the detector. The average bias with respect to the default model is taken as
 1665 systematic uncertainty. Results are reported in Tab. 10.4.

Table 10.4: Values of systematic uncertainties due to the choice of background parameterisation in bins of q^2 .

q^2 [GeV $^2/c^4$]	A_{FB}^ℓ	f_L	A_{FB}^h
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

1666

1667 10.5 Polarisation

1668 To study the effect of a non-zero A_b^0 production polarisation simulated events are
 1669 generated using the distributions given by Eqs. 8.13 and 8.15 as a function of the
 1670 angle under study ($\cos \theta_\ell$ or $\cos \theta_h$) and $\cos \theta$, defined in Sec. 8.2, which is sensitive to
 1671 polarisation. Similarly to the procedure used for the branching ratio measurement,
 1672 events are generated using values of the polarisation corresponding to $\pm\sigma$ from the
 1673 LHCb measurement [80]. In the theoretical functions $\cos \theta$ is always odd therefore
 1674 with perfect efficiency it always drops out by integrating over $\cos \theta$. Therefore the
 1675 generated distributions are multiplied by the two-dimensional efficiency function.
 1676 No significant bias is found.

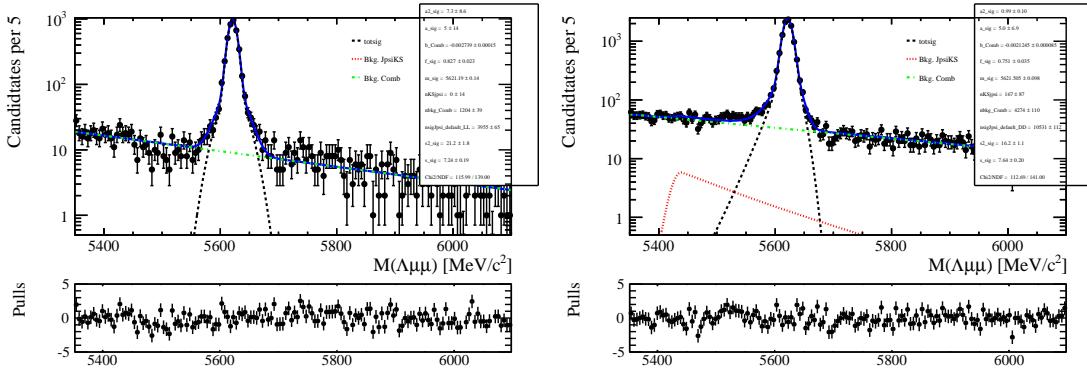


Figure 10.3: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ long (left) and downstream (right) candidates with an extra proton PID cut to remove K_s^0 background.

1677 10.6 J/ψ cross-check

1678 To cross-check the fitting procedure this is applied on the high statistics $\Lambda_b^0 \rightarrow J/\psi \Lambda$
 1679 sample. For this purpose events are selected with an additional requirement on the
 1680 proton PID, $\text{PID}_p > 10$. This is needed to reduce the $B^0 \rightarrow K_s^0 J/\psi$ background,
 1681 which is particularly important for the hadronic side fit, since the K_s^0 events are
 1682 not distributed in a flat way in the $\cos \theta_h$ variable and would therefore bias the fit.
 1683 Figure 10.3 shows the invariant mass distributions after this requirement is applied,
 1684 which can be compared with the ones in Fig. 5.3. After the PID cut there are 0.2%
 1685 of K_s^0 events left in the downstream sample and a fraction compatible with zero in
 1686 the long sample. The signal model is the same used for the rare case and described
 1687 in Sec. 9.2. For the background instead the higher statistics allows to leave more
 1688 freedom to the fit. Therefore a second-order Chebyschev polynomial is used, where
 1689 the two parameters are free to float. As for the rare case the background fractions are
 1690 gaussian-constrained to what found in the invariant mass fit. Figures 10.4 and 10.5
 1691 show fitted angular distributions for the J/ψ channel. The measured values of the
 1692 observables are $A_{\text{FB}}^\ell = -0.002^{+0.011}_{-0.011}$, $A_{\text{FB}}^h = -0.402^{+0.010}_{-0.009}$ and $f_L = 0.485^{+0.019}_{-0.020}$,
 1693 where the uncertainties are 68% Feldman Cousins confidence intervals. The lepton
 1694 side asymmetry as expected is measured to be zero.

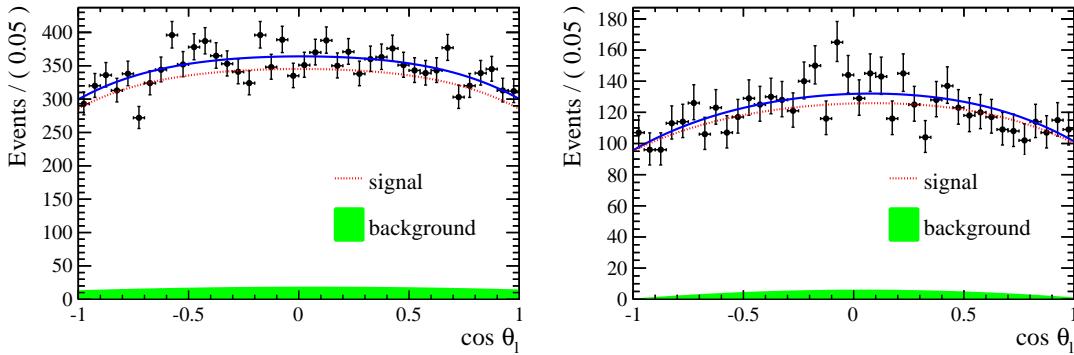


Figure 10.4: Fitted angular distribution as a function of $\cos \theta_\ell$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

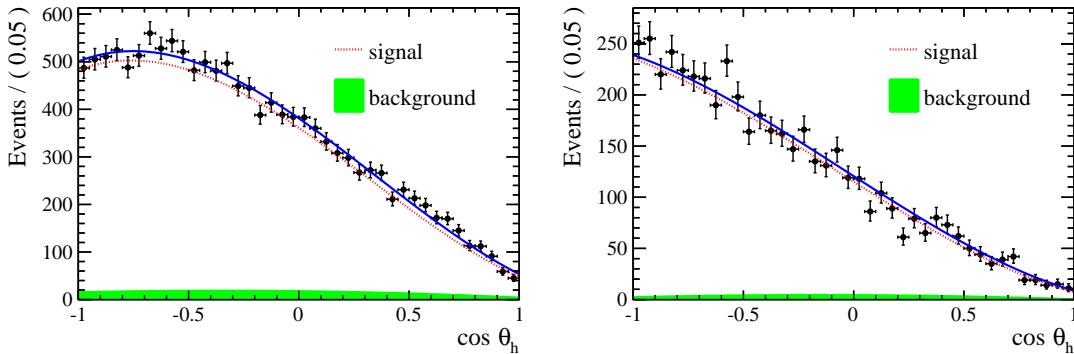


Figure 10.5: Fitted angular distribution as a function of $\cos \theta_h$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

1695 10.7 Results

1696 Figures 10.6 and 10.7 show fits to the angular distributions for the 15-20 $\text{GeV}^2/c^4 q^2$
 1697 interval and Tab. 10.5 reports measured values of A_{FB}^ℓ , A_{FB}^h and f_L . The asymmetries
 1698 are also shown in Fig. 10.8 together with SM predictions obtained from Ref. [83].
 1699 The statistical uncertainties on these tables are obtained using the likelihood-ratio
 1700 ordering method described in Sec. 9.1, where only one of the two observables at
 1701 a time is treated as the parameter of interest. In Fig. 10.9 the statistical uncer-
 1702 tainties on A_{FB}^ℓ and f_L are also reported as two-dimensional 68 % confidence level
 1703 (CL) regions, where the likelihood-ratio ordering method is applied by varying both
 1704 observables and therefore taking correlations into account. Total systematic uncer-
 1705 tainties correspond to the square root sum of the single considered sources.

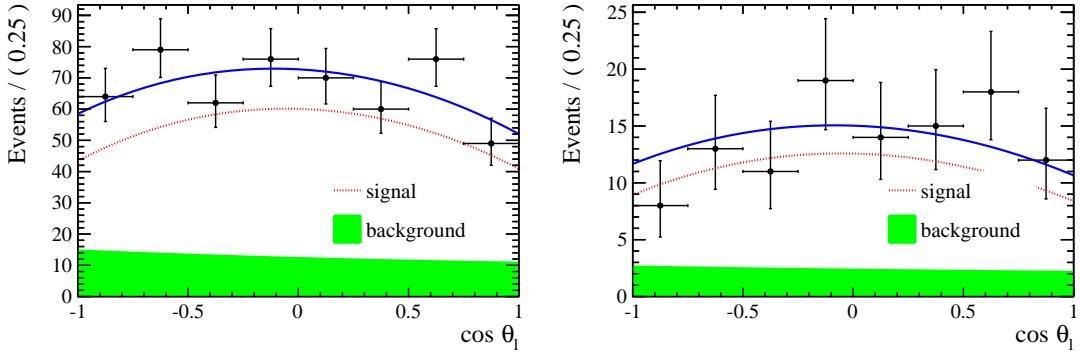


Figure 10.6: Fitted angular distributions as a function of $\cos \theta_\ell$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

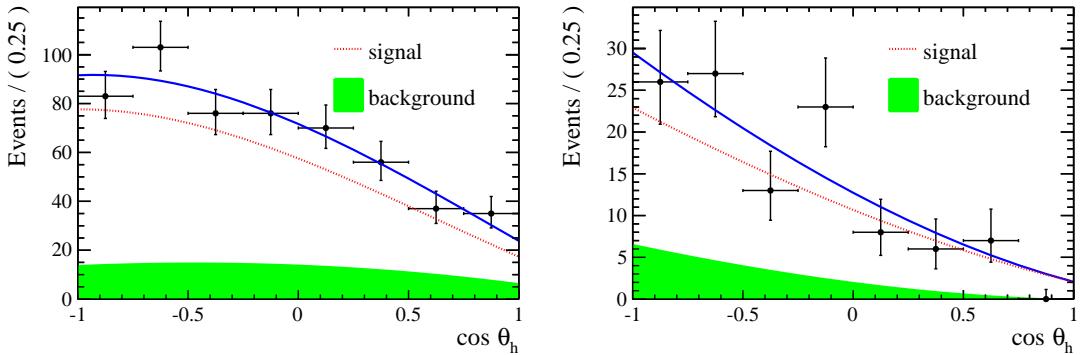


Figure 10.7: Fitted angular distributions as a function of $\cos \theta_h$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

Table 10.5: 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_{FB}^ℓ	f_L	A_{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	$-0.05^{+0.09}_{-0.09} \pm 0.03$	$0.61^{+0.11}_{-0.14} \pm 0.03$	$-0.29^{+0.07}_{-0.07} \pm 0.03$

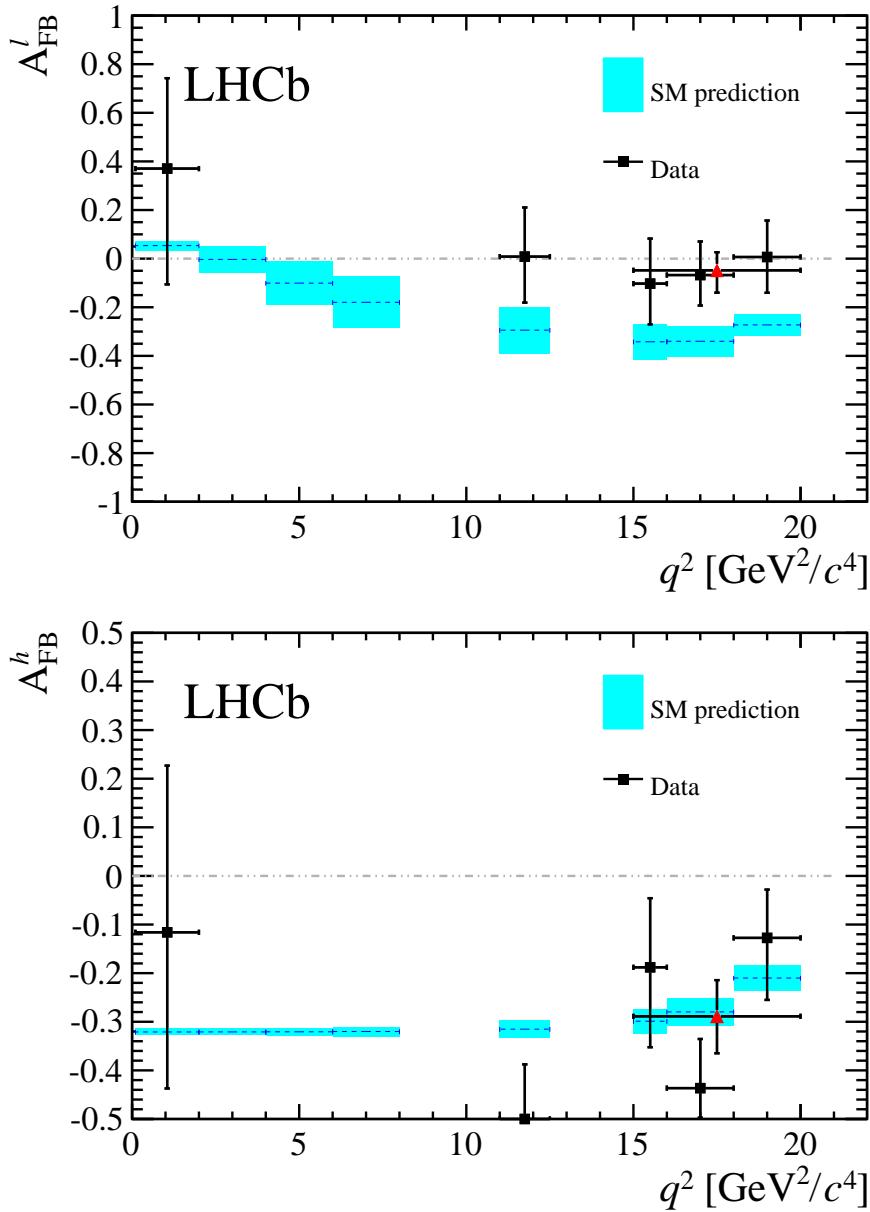


Figure 10.8: Measured values of the leptonic (top) and the hadronic (bottom) forward-backward asymmetries in bins of q^2 . Data points are only shown for q^2 intervals where a statistically significant signal yield is found, see text for details. The (red) triangle represents the values for the $15 < q^2 < 20$ GeV^2/c^4 interval. Standard Model predictions are obtained from Ref. [89].

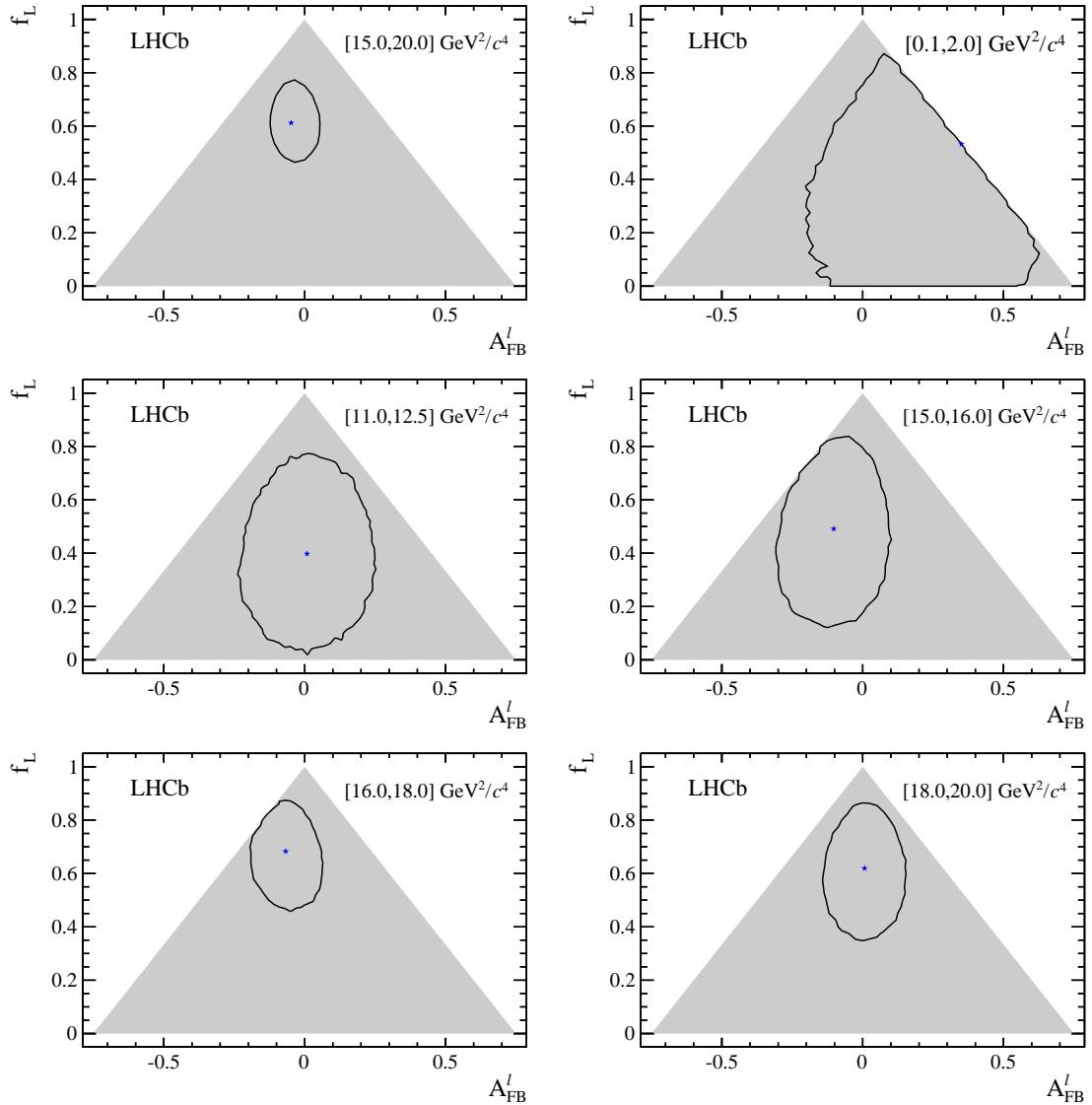


Figure 10.9: Two-dimensional 68 % CL regions (black) as a function of A_{FB}^{ℓ} and f_L . The shaded areas represent the regions in which the PDF is positive over the complete $\cos \theta_{\ell}$ range. The best fit points are indicated by the (blue) stars.

1706

Part II

1707

The $R_{K^{*0}}$ analysis

1708

CHAPTER 11

1709

1710

Testing lepton flavour universality with R_H

1711

1712 Lepton Favour Universality (LFU) is the equality of the weak coupling constants for
1713 all leptons. FCNCs processes, forbidden in the SM at tree level and happening only
1714 at loop level, are an ideal laboratory for studying LFU as NP in the loops could
1715 break the flavour symmetry.

1716 In this work, decays of $b \rightarrow s\mu^+\mu^- (e^+e^-)$ type, are considered to test LFU between
1717 electrons and muons using penguin decays. In particular, B^0 meson semileptonical
1718 decays to $B^0 \rightarrow K^{*0}\ell^+\ell^-$ are considered. Figure 11.1 shows the possible Feyn-
1719 man diagrams producing such decays while Fig. 11.2 illustrates how the Feynman
1720 diagrams of these processes may include new particles. A series of recent LHCb mea-
1721 surements [17] points to a tension with SM predictions, which make these processes
1722 very interesting to better understand the nature of the discrepancy.

1723 In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger pro-
1724 posed the measurement of the R_H ratio [90], defined in Eq. 11.1, where H can be an

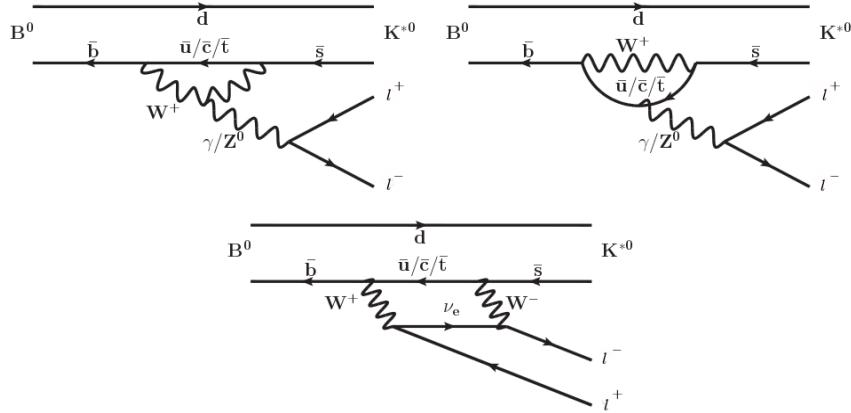


Figure 11.1: Loop diagrams of the $B^0 \rightarrow K^{(*)0} \ell^+ \ell^-$ process.

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

$$R_H = \frac{\int_{4m_\mu^2}^{m_b} \frac{d\mathcal{B}(B^0 \rightarrow H \mu^+ \mu^-)}{dq^2} dq^2}{\int_{4m_\mu^2}^{m_b} \frac{d\mathcal{B}(B^0 \rightarrow He^+ e^-)}{dq^2} dq^2} \quad (11.1)$$

In this quantity the decay width is integrated over the squared dilepton invariant mass, q^2 , from $q_{min}^2 = 4m_\mu^2$, which is the threshold for the $\mu\mu$ process, up to $q_{max}^2 = m_b^2$.

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 which happen 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 happen. The luminosity measurement is usually a source of systematic uncertainty, however 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_μ^2/m_b^2 Hiller and Kruger calculate that in the SM and in the full q^2

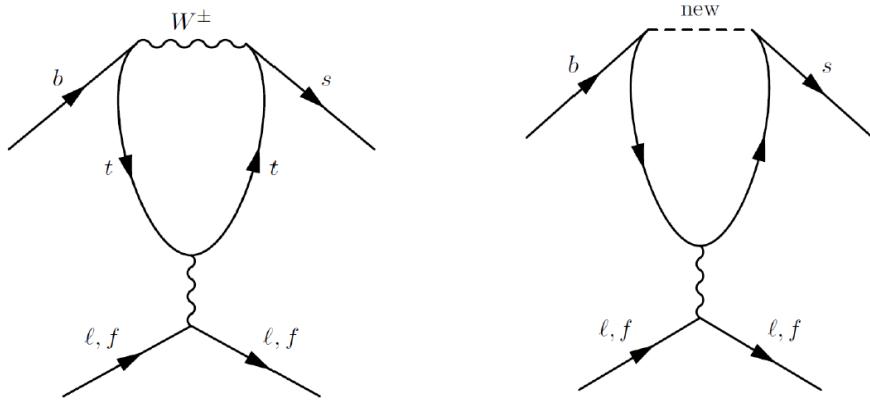


Figure 11.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

range [90]:

$$R_{X_s} = 0.987 \pm 0.006 \quad (11.2)$$

$$R_K = 1.0000 \pm 0.0001 \quad (11.3)$$

$$R_{K^{*0}} = 0.991 \pm 0.002 \quad (11.4)$$

$$(11.5)$$

¹⁷³⁷ under the assumptions that:

- ¹⁷³⁸ • right-handed currents are negligible;
- ¹⁷³⁹ • (pseudo-)scalar couplings are proportional to the lepton mass;
- ¹⁷⁴⁰ • there are no CP-violating phases beyond the SM.

¹⁷⁴¹ The measurement of the R_H ratios is of particular interest after the recent measure-
¹⁷⁴² ment of the branching ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay [28], where no evidence of NP
¹⁷⁴³ was found. In fact the quantities $(R_H - 1)$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ remain proportional
¹⁷⁴⁴ with

$$\frac{R_H - 1}{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)} \sim 2 \cdot 10^{-5} \quad (11.6)$$

¹⁷⁴⁵ A joint measurement of these two quantities can give much information and constrain
¹⁷⁴⁶ MFV models. If $R_H = 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ is close to the SM prediction as it is

1747 measured to be this will allow to put strong constraints on extensions of the SM. If
1748 instead $R_H > 1$ and the equation above is not verified, this would mean that one of
1749 the assumptions listed above are not verified, which can happen in some extensions of
1750 the SM as Super-Symmetric models with broken R-parity. A series of recent LHCb
1751 measurements [17] shows tensions with SM predictions, which makes it interesting
1752 to further investigate these processes.

1753 11.0.1 Combining ratios

1754 The full power of the R_H ratios in understanding new physics scenarios comes from
1755 their combinations. In Ref. [91] Hiller and Schmaltz propose the measurement of
1756 the double ratios, $X_H = R_H/R_K$, which not only can test LFU but also allow to
1757 disentangle the kind of new physics that lies behind. These ratios are in fact sensitive
1758 to FCNCs of right-handed currents. Furthermore, in Ref. [91] the study is extended
1759 to B_s^0 decays such as $B_s^0 \rightarrow \phi\ell^+\ell^-$ or $B_s^0 \rightarrow \eta\ell^+\ell^-$.

1760 Parity and Lorentz invariance require that the Wilson Coefficients with left-handed
1761 chirality (C) and their right-handed counterparts (C') appear in the decay amplitude
1762 of exclusive decays in determined combinations, e.g.

$$\begin{aligned} C + C' : & K, K_{\perp}^*, \dots \\ C - C' : & K_0(1430), K_{0,\parallel}^*, \dots \end{aligned} \tag{11.7}$$

1763 where the labels for the K^* meson represent its longitudinal (0), parallel (\parallel) and
1764 perpendicular (\perp) transversity components. The C contributions are universal to
1765 all decays and therefore X_H double ratios are sensitive to right-handed currents. In
1766 fact the R_H ratios can be expressed in terms of their deviation from unity as

$$\begin{aligned} R_K &\simeq 1 + \Delta_+ \\ R_{K_0(1430)} &\simeq 1 + \Delta_- \\ R_K^* &\simeq 1 + p(\Delta_- - \Delta_+) + \Delta_+ \end{aligned} \tag{11.8}$$

where the Δ_{\pm} quantities are combinations of Wilson coefficients described in Eq. 10 of Ref. [91] and the parameter p is the polarisation of K^* that in Ref. [91] 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} .

11.0.2 Experimental status

The R_K and $R_{K^{*0}}$ ratios have already been measured at the B-factories [92, 93], and the R_K ratio has been also recently measured at LHCb [94] 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. 11.1. 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 11.1: Experimental status of the $R_{K^{(*)}}$ measurements.

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

1783

11.1 Analysis strategy

The aim of this analysis is to measure the $R_{K^{*0}}$ ratio using pp collision data collected by the LHCb detector in 2011 and 2012, corresponding to a total of 3 fb^{-1} of

1787 integrated luminosity. The $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$, “rare channels”,
1788 are reconstructed with the K^{*0} decaying into a kaon and a pion with opposite
1789 charges.

1790 The analysis has to separate signal candidates from background candidates which
1791 have similar observed properties. The selection presented in Sec. 12 aims to max-
1792 imise the yield while minimising the background contamination. Two types of back-
1793 grounds are identified: “peaking background” and “combinatorial background”. The
1794 first comes from the mis-reconstruction of other decays or from partially recon-
1795 structed events. This type of background, because its specific kinematic properties,
1796 usually peaks in some variable, such as the invariant mass of all final particles,
1797 therefore we can remove these events by removing the peak. The combinatorial
1798 background instead comes from the random combination of particles and can be
1799 lowered selecting events with good-quality tracks and vertices.

1800 To further reduce systematic uncertainties the measurement is performed as a dou-
1801 ble ratio as shown in Eq. 11.9 where decays reaching the same final states as the
1802 rare channels via a J/ψ resonance, $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$, also referred as “char-
1803 monium” or “resonant” channels, are used as control samples. These decays are
1804 distinguished from the rare channels using the invariant mass of the dilepton pair.

1805 In Sec. 14 the efficiency of selecting and reconstructing each channel is extracted
1806 and, finally, in Sec. 16 the R_{K^*} ratio defined is built as the double ratio of rare and
1807 resonant channels:

$$R_{K^{*0}} = \frac{N_{B^0 \rightarrow K^{*0}\mu^+\mu^-}}{N_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}} \cdot \frac{N_{B^0 \rightarrow K^{*0}J/\psi e^+e^-}}{N_{B^0 \rightarrow K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}}{\varepsilon_{B^0 \rightarrow K^{*0}\mu^+\mu^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}e^+e^-}}{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow e^+e^-}} \quad (11.9)$$

1808 As NP is expected not to affect charmonium resonances the ratio of the J/ψ channels
1809 is 1 and therefore $R'_{K^{*0}} = R_{K^{*0}} \times R_{J/\psi} = R_{K^{*0}}$. On the other hand using the relative
1810 efficiencies between the rare and resonant channels allows to cancel out many effects

1811 resulting in a better control of systematic uncertainties.

1812 For brevity, the rare channels will also be denoted as “ $\ell\ell$ ”, or specifically “ee” and
1813 “ $\mu\mu$ ”, and the resonant channels as “ $J/\psi(\ell\ell)$ ”, or “ $J/\psi(ee)$ ” and “ $J/\psi(\mu\mu)$ ”.

1814 11.2 Choice of q^2 intervals

1815 Two q^2 intervals are considered in this work:

- 1816 • the “central- q^2 ” region, $[1.1,6.0]$ GeV^2/c^4 ;
- 1817 • the “high- q^2 ” region, above 15 GeV^2/c^4 .

1818 The central- q^2 region is the most interesting place to look for new physics. In fact,
1819 at low q^2 , below 1 GeV^2/c^4 the photon pole dominates leaving little space for NP
1820 to be found [17]. The upper bound of this interval is set at 1.1 GeV^2/c^4 , in order
1821 entirely include the contribution from $\phi \rightarrow \ell^+\ell^-$ decays, that can dilute new physics
1822 effects, into the low q^2 interval. The upper bound of the central interval is chosen
1823 to be sufficiently far away from the J/ψ radiative tail, where predictions cannot be
1824 cleanly extracted. The 6 – 15 GeV^2/c^4 region is characterised by the presence of the
1825 narrow peaks of the J/ψ and $\psi(2S)$ resonances. The lower bound of the high- q^2
1826 region, where the signal in the electron channel is still unobserved, is chosen to be
1827 sufficiently far from the $\psi(2S)$ resonance. Rare and resonant channels are selected
1828 depending on which q^2 interval they fall in (for details see Sec. 12).

1829 11.3 Data samples and simulation

1830 This analysis is based on a data set corresponding to 3 fb^{-1} of integrated luminosity
1831 collected by the LHCb detector in 2011 and 2012. In order to study the background
1832 properties, determine efficiencies and to train the multivariate analysis simulated

1833 events are used. After the hard interactions are generated with Pythia8 hadronic
 1834 particles are decayed using EvtGen and, finally, propagated into the detector us-
 1835 ing Geant4 and reconstructed with the same software used for data. Samples are
 1836 generated with both 2011 and 2012, magnet up and down conditions and are com-
 1837 bined in the right proportions, according to the luminosity registered on data. The
 1838 next section describes the corrections applied to the simulation to obtain a better
 1839 description of data.

1840 11.3.1 Data-simulation corrections

1841 Since the multivariate classifier training (see Sec. 12.4) and the calculation of most
 1842 of the efficiency components (see Sec. 14) are obtained from the study of simulated
 1843 events it is important to verify that the simulation is a reliable reproduction the
 1844 data. In particular it is important to match data and Monte Carlo in the kinemat-
 1845 ics of the final particles and the occupancy of the detector. The kinematics of the
 1846 decays is characterised by the transverse momentum spectrum of the B^0 . Discrep-
 1847 ancies in this distribution cause also the spectra of the final particles to differ from
 1848 data and affect the efficiency determination as its value often depends on the mo-
 1849 mentum distribution of final particles. The occupancy of the detector is correlated
 1850 to the invariant mass shape of the signal because the addition of energy clusters in
 1851 the electromagnetic calorimeter, affects the electron momenta for bremsstrahlung
 1852 photons emitted before the magnet. The hit multiplicity in the SPD detector is a
 1853 proxy for the detector occupancy.

1854 Since it is important that these quantities are well modelled, the simulation is
 1855 reweighted so that the distributions in data and simulation match for these vari-
 1856 ables. This can be done using resonant $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ events, for which
 1857 the signal peak is already visible in data after pre-selection (see Sec. 12). However,
 1858 the data still includes a high level of background and distributions cannot be directly
 1859 compared. The $s\mathcal{P}$ lot technique [50] is used to statistically subtract the background
 1860 from data and obtain pure signal distributions using the invariant mass as control

variable. Fig. 11.3 shows fits to the 4-body invariant mass of candidates after pre-selection, done in order to estimate the signal density. Data and simulation are then compared and the ratio between the distributions is used to re-weight the Monte Carlo. 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 multiplicities only. Distributions of B^0 transverse momentum and SPD multiplicities are reported in Fig. 11.4 and ratios of these distribution, which are used to re-weight the simulation, are reported in Fig. 11.5. The weights for the SPD multiplicity are calculated separately for 2011 and 2012 events, because distributions are significantly different in the two years. Binnings for these distributions are chosen to have approximately the same number of events in each bin to limit fluctuations. Further corrections are made re-weighting simulated events for PID efficiency using the `PIDCalib` package as described in Sec. 14.3 and, finally, ee samples are also reweighted for L0 trigger efficiency as described in Sec. 14.4.

Weights are always applied throughout unless specified.

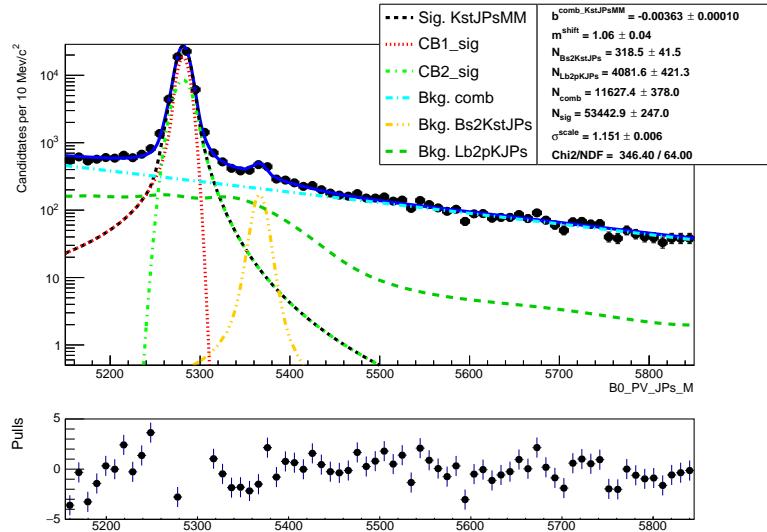


Figure 11.3: Fitted 4-body invariant mass distributions of muonic resonant candidates.

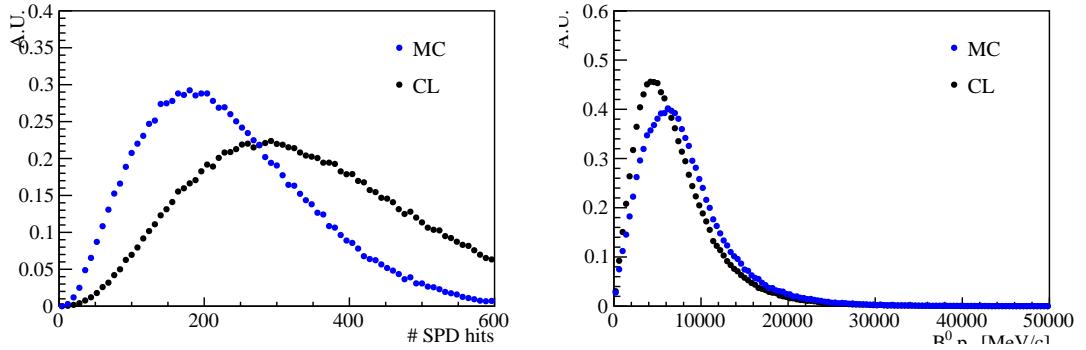


Figure 11.4: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

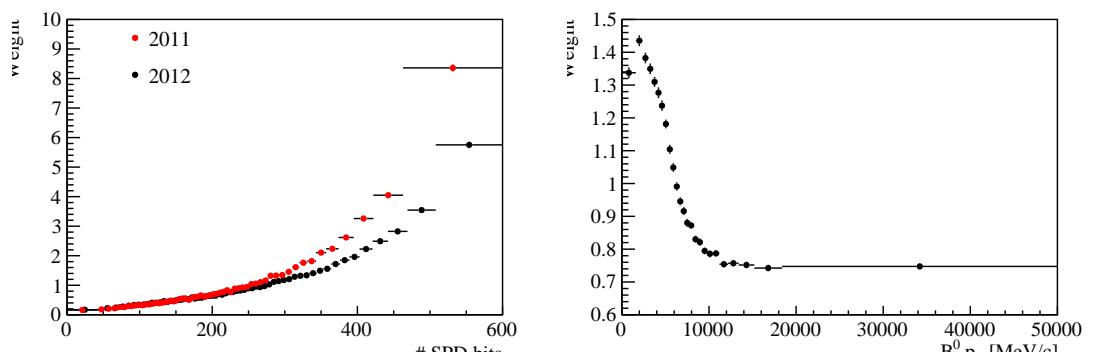


Figure 11.5: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

1876

CHAPTER 12

1877

1878

Selection

1879

1880 The selection process, described in the following subsections, is divided into several
1881 steps. First of all events have to fall into the detector acceptance, produce hits
1882 and be selected on the basis of quality features, such as χ^2 of tracks and vertices,
1883 this stage is called “stripping”. Secondly it is required that some specific trigger
1884 lines were fired by the events. After the trigger and stripping requirements, cuts
1885 are applied to remove backgrounds from specific decays. These first three steps are
1886 referred to as “pre-selection”.

1887 The next step consists in the application of particle identification (PID) conditions
1888 which remove a good part of misreconstructed background and clear the way for
1889 the last step where a neural network is used to remove combinatorial background.

1890 In order to minimise systematic uncertainties the same selection requirements are
1891 applied to the rare signal candidates and on their relative charmonium channel, a
1892 part from the q^2 cuts which serve to distinguish them. In order to identify the
1893 $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ channel a dilepton mass interval of 100 MeV/ c^2 around the

nominal J/ψ peak [2] is selected. For the electron resonant channel it is not possible to use a narrow cut on the q^2 and 4-body $m(K\pi e^+e^-)$ invariant mass distributions are characterised by a long radiative tail at low masses due to bremsstrahlung radiation. Furthermore, a cut in q^2 also distorts the 4-body mass distribution at low masses and it is important to be able to fit a wide mass range to constrain backgrounds. For these reasons the interval to select $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates is chosen to go as low as possible without overlapping with the rare channel interval. The electron resonant channel is therefore selected in the interval [6,11] GeV^2/c^4 . Figure 12.1 shows two-dimensional distributions of q^2 versus the 4-body $m(K\pi\ell^+\ell^-)$ invariant mass for events which pass the full selection. On these plots horizontal bands can be seen at the q^2 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.

12.1 Trigger and Stripping

Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported in Tab. 12.1, 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 to the L0 candidates and triggers if the L0 decision is confirmed; finally, Hlt2Topo[2,3]BodyBBDT uses a reconstruction of the event and a neural network trained on events with a specific topology in order to detect decays.

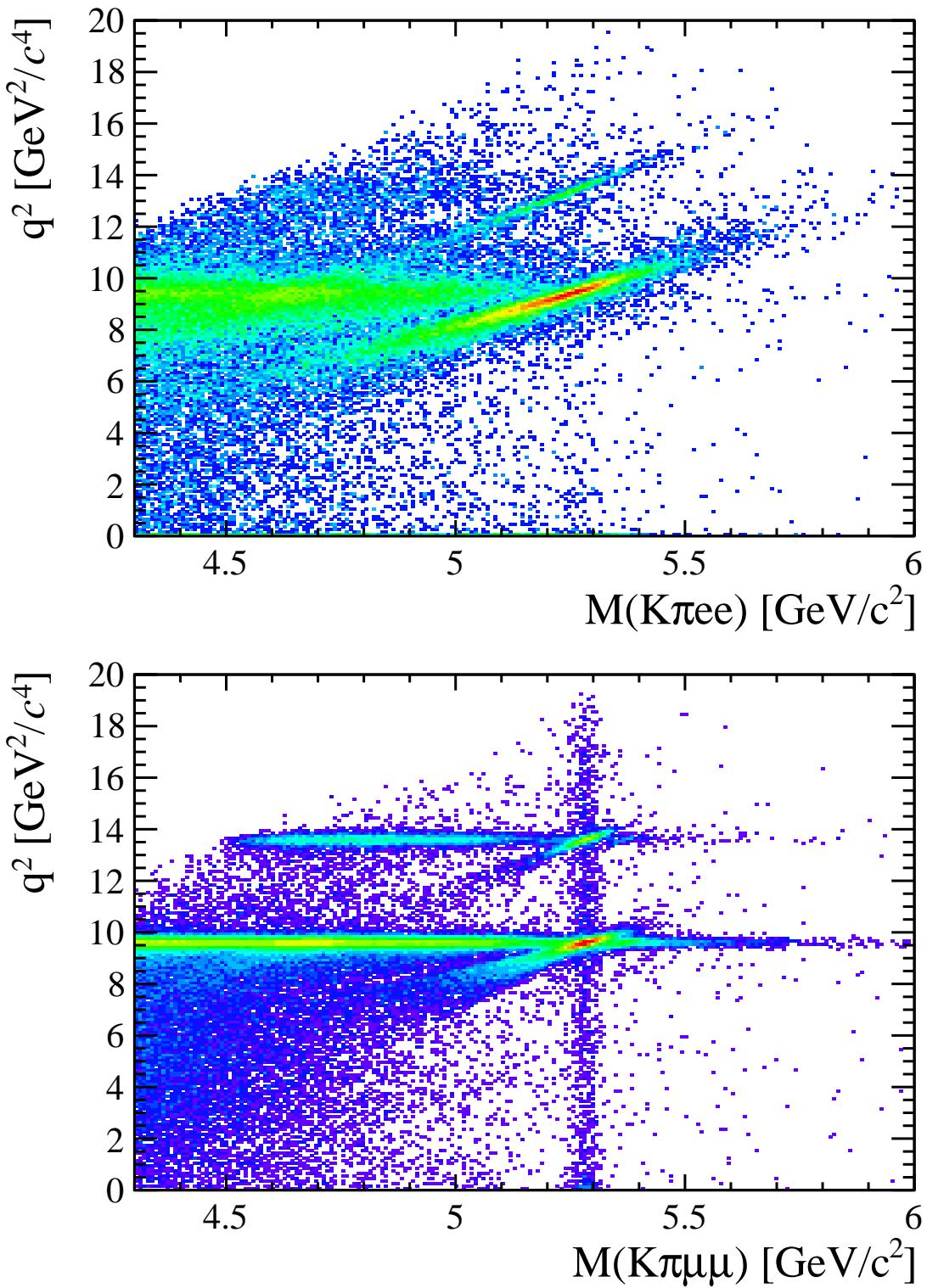


Figure 12.1: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electron (top) and muonic (bottom) channels in 2012 data.

Table 12.1: 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

1921 For the electron channels the L0 lines have different properties, therefore the analysis
1922 is performed separately for three categories of events, depending on the L0 trigger
1923 that fired them. These categories are defined to be exclusive in the following way:

1924 • Events triggered by at least one of the electrons in the signal candidate (L0E):
1925 `L0Electron_TOS`

1926 • Events triggered by at least one of the hadrons in the signal candidate and
1927 not by L0Electron (L0H):

1928 `L0Hadron_TOS && !L0Electron_TOS`

1929 • Events triggered by particles not in the signal candidate (Trigger Independent
1930 of Signal, TIS) and not by the previous cases (L0I):

1931 `L0_TIS && !(L0Electron_TOS || L0Hadron_TOS)`

1932 The majority of the selected events falls in the L0Electron category. The L0Hadron
1933 category is more efficient at low q^2 were the K^{*0} has more momentum.

1934 Candidates are then required to pass the kinematic and quality cuts summarised in
1935 Tab. 12.2. The meaning of variables in the table was already explained in Sec. 4.

1936 Loose PID cuts are applied in preselection to limit the size of the samples, while
1937 tighter cuts are applied in a second stage. A large mass window is kept around
1938 the B^0 peak in order to be able to fit the sideband and to train the multivariate

Table 12.2: Summary of stripping requirements used for the central and high q^2 regions.

Particle	Requirements
π	$\chi_{\text{IP}}^2(\text{primary}) > 9$
K	$\text{PID}_K > -5$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>hasRICH</code>
K^{*0}	$p_{\text{T}} > 500 \text{ MeV}/c$ $ m - m_{K^{*0}}^{\text{PDG}} < 100 \text{ MeV}/c^2$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ Origin vertex $\chi^2/\text{ndf} < 25$
μ	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>isMuon</code>
e	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>hasCalo</code> $\text{PID}_e > 0$
Dilepton	$m_{\ell\ell} < 5500 \text{ MeV}/c^2$ End vertex $\chi^2/\text{ndf} < 9$ Origin vertex χ^2 separation > 16
B^0	$\text{DIRA} > 0.9995$ End vertex $\chi^2/\text{ndf} < 9$ $\chi_{\text{IP}}^2(\text{primary}) < 25$ Primary vertex χ^2 separation > 100

1939 analysis and constrain backgrounds. Track-quality and vertex quality cuts are also
1940 applied using the $\chi^2_{track}/\text{ndf}$, `GhostProb`, and χ^2_{vtx}/ndf variables. The `GhostProb`
1941 quantity describes the probability of a track being fake. By construction cutting at
1942 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. For details about the definition of
1943 the variables used see Ref. [95].

1944 12.2 PID

1945 After preselection there still are high levels of misreconstructed background. In par-
1946 ticular, as the ID of kaons and pions are not constrained, the samples still contain
1947 both ID combinations for most candidates, therefore tighter PID cuts are applied.
1948 In the LHCb analysis framework the particle identification probability can be quan-
1949 tified using the “`ProbNN`” variables [96]. These variables are the output of a Neural
1950 Network which takes as input information from the calorimeters, the RICH detec-
1951 tors and the muon system. Unlike the DLL variables these are bounded from 0 to 1
1952 and can be therefore directly be interpreted as probabilities. For example `ProbNNk`
1953 is the probability for a reconstructed particle to be a kaon. Two tunes of the `ProbNN`
1954 variables, labelled V2 and V3, are available. Tune V3 was shown to be optimal for
1955 positive ID, while tune V3 was found to be optimal for background rejection and
therefore it is used to quantify the mis-ID probability.

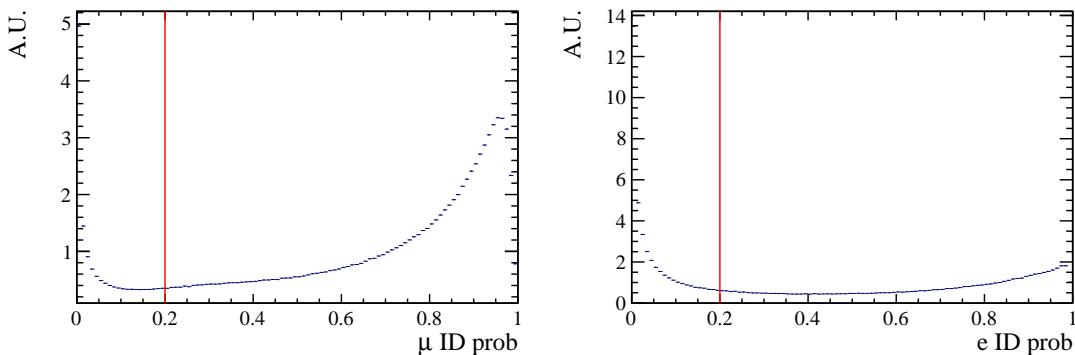


Figure 12.2: Correct ID probability distributions for muons (left) and electron (right) in 2012 data.

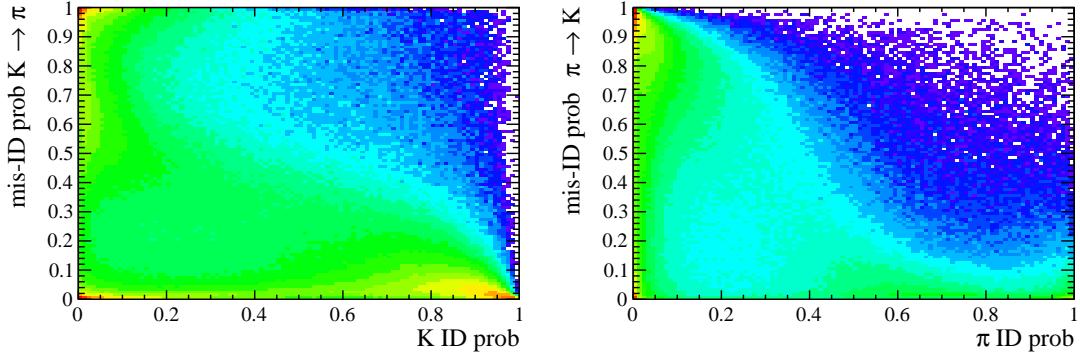


Figure 12.3: On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis show the mis-ID probability.

1957 Figure 12.2 shows distributions of the correct ID variables in the 2012 data sample
 1958 while Fig. 12.3 shows in a two-dimensional plane the probability of correct identifi-
 1959 cation and mis-identification of kaons and pions. These plots are characterised by
 1960 clear peak at maximal ID probability and minimal mis-ID probability, corresponding
 1961 to particles to which is possible to assign a well defined identification.

1962 In order to maximise the power of the PID cuts probabilities of correct ID and
 1963 mis-ID are combined using the following cuts:

$$\begin{aligned}
 \pi &\rightarrow \text{ProbNNpi-V3} \times (1 - \text{ProbNNk-V2}) \times (1 - \text{ProbNNp-V2}) > 0.1 \\
 K &\rightarrow \text{ProbNNk-V3} \times (1 - \text{ProbNNp-V2}) > 0.05 \\
 \mu &\rightarrow \min(\text{ProbNNmu-V3}, \text{ProbNNmu-V3}) > 0.2 \\
 e &\rightarrow \min(\text{ProbNNe-V3}, \text{ProbNNe-V3}) > 0.2
 \end{aligned}$$

1964
 1965 In the first formula, for example, **ProbNNpi** is the probability of correctly identifying
 1966 the pion as a pion, while **ProbNNk** is the probability of mistaking it for a kaon. There-
 1967 fore by maximising the quantity “**ProbNNpi** × (1 - **ProbNNk**)”, one can maximise
 1968 the correct ID probability and minimise at the same time the mis-ID probability.

1969 12.3 Peaking backgrounds

1970 Cuts are applied in order to remove background sources due to specific decays.
 1971 These types of backgrounds usually peak in some variable because of their mass or
 1972 distinctive kinematic properties and therefore they can be removed without signifi-
 1973 cant signal efficiency loss. In the following sections are described the main sources
 1974 of peaking background.

1975 12.3.1 Charmonium vetoes

1976 Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2 . The choice of q^2 binning
 1977 described in Sec. 11.2 constitutes a natural veto for these decays. Simulated events
 1978 were used to check if resonant events leak inside the q^2 intervals chosen for the rare
 1979 channel analysis. For the muonic channels the leakage is negligible as the peaks are
 1980 sharper due to a better resolution and muons emit fewer bremsstrahlung photons, re-
 1981 sulting in shorter radiative tails. The electronic channels are instead characterised by
 1982 a worse resolution and at the same time electrons can radiate several bremsstrahlung
 1983 photons, yielding long tails at low q^2 . Analysing Monte Carlo events it was found
 1984 that 1.3–2% (depending on the trigger category) of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candi-
 1985 dates leak into the $1.1 < q^2 < 6$ GeV^2/c^4 interval and 1.8% of $\psi(2S)$ events leak
 1986 above 15 GeV^2/c^4 . The contribution from these candidates is modelled in the fit.

1987 12.3.2 ϕ veto

1988 It can happen that a kaon from the decay $B_s \rightarrow \phi\ell^+\ell^-$, where the ϕ decays in two
 1989 kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as a
 1990 K^{*0} . This results in a candidate with a value of $m(K\pi)$ that is less than the nominal
 1991 K^{*0} mass but still high enough to pass the selection requirements. In Fig. 12.4 is
 1992 reported the plot of $m(K\pi)$ versus $m(K\pi\ell\ell)$, where kaon mass hypothesis is assigned
 1993 to the pion. A peak can clearly be seen around the ϕ mass (1020 MeV/c^2). To remove

1994 this background only candidates with $m_{K(\pi \rightarrow K)} > 1040 \text{ MeV}/c^2$ are selected. This
 1995 results in a 98% background rejection while keeping a 99% signal efficiency. The ϕ
 1996 could also constitute a background when it decays into two leptons but the branching
 1997 ratio of this decay is small compared to the one into kaons and this contribution is
 1998 taken into account by the choice of the q^2 intervals.

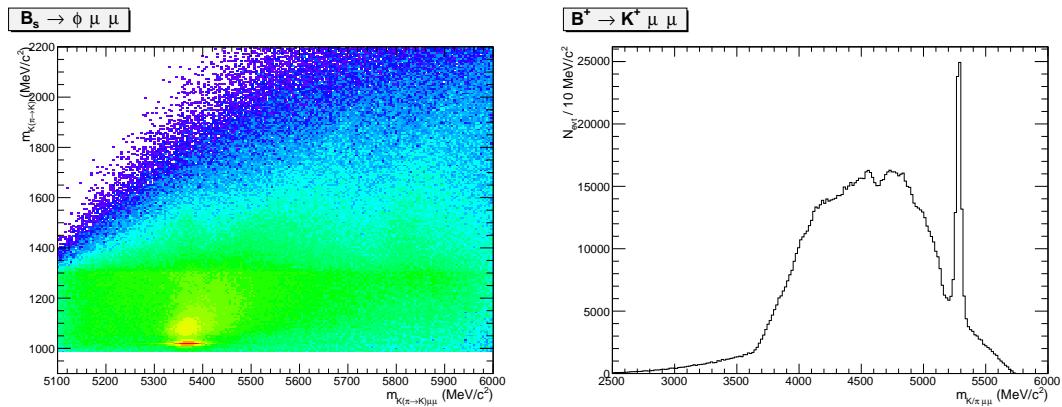


Figure 12.4: (left) Distribution of 2011 data events as a function of the variables $(m_{K(\pi \rightarrow K)})$ and $(m_{K(\pi \rightarrow K)\mu\mu})$, where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. (right) The invariant mass distribution of the three-body system $(K\mu\mu)$, where the peak due to the $B^+ \rightarrow K^+\mu^+\mu^-$ decay is visible.

1999

2000 12.3.3 $B^+ \rightarrow K^+\ell^+\ell^-$ plus a random pion

2001 Some $B^+ \rightarrow K^+\ell^+\ell^-$ decays can contaminate the upper B^0 mass sideband if they
 2002 are combined with a soft pion from somewhere else in the event and therefore re-
 2003 constructed as a B^0 decay. The same can also happen with a kaon misidentified
 2004 as a pion combined with an other kaon in the event. In Fig. 12.4 the invariant
 2005 mass distribution of the three-body $K\mu^+\mu^-$ system, $m(K\mu\mu)$, is shown. This is
 2006 characterised by a narrow peak at the B^+ mass. Since these candidates have
 2007 $m(K\pi\ell\ell) > 5380 \text{ MeV}/c^2$ there is no contribution under the B^0 peak, but they
 2008 can cause problems when using sidebands events to train the neural network. An
 2009 effective veto for this decay was found to be $\max(m_{K\ell\ell}, m_{(K \rightarrow \pi)\ell\ell}) < 5100 \text{ MeV}/c^2$,
 2010 which results in 95% background rejection while keeping 99% signal efficiency.

2011 12.3.4 Λ_b decays

2012 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays are unlikely to be reconstructed as $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ because the
 2013 Λ is long-lived and decays further in the detector with a separate vertex. Simulated
 2014 events were used to check how many candidates fall into the B^0 samples, which
 2015 results to be negligible. The $\Lambda_b^0 \rightarrow J/\psi pK$ decay can instead contribute more easily
 2016 since the $m(pK)$ is above the Λ threshold and therefore they must come from Λ^*
 2017 resonances, which are not long-lived. This background is already reduced using
 2018 PID but a non-negligible contribution is still expected in the $\mu\mu$ sample, which is
 2019 modelled in the fit.

2020 12.3.4.1 Other peaking backgrounds

2021 A possible background could come from $B^0 \rightarrow K^{*0}\gamma$ decays where the photon con-
 2022 verts into two electrons while traversing the detector. In LHCb, around 40% of
 2023 photons convert before the calorimeter, but only a small fraction of these, $\sim 10\%$,
 2024 are reconstructed. Furthermore these events fall into a q^2 region well below the
 2025 intervals considered in these analysis and their contribution is therefore negligible.
 2026 Similar decays are also $B^0 \rightarrow K^{*0}\eta$ and $B^0 \rightarrow K^{*0}\pi^0$ where η and the pion decay
 2027 into two photons. Once again the contribution from these decays falls well below
 2028 the considered q^2 intervals. Finally, a potentially dangerous background could come
 2029 from events where the identity of the kaon and the pion are swapped as these candi-
 2030 dates peak under the signal. Their contribution is found to be small, 0.5%, however
 2031 the effect of their modelling into the fit is taken into account in the systematic
 2032 uncertainties.

2033 12.3.5 Mis-reconstructed background

2034 A source of mis-reconstructed background is due to cascade decays with a B^0 de-
 2035 caying semileptonically into a D meson which also decays semileptonically, e.g.

2036 $B^0 \rightarrow D^- \ell^+ \bar{\nu}_\ell$ followed by $D^- \rightarrow K^{*0} \ell^- \nu_\ell$. The candidates built from these de-
2037 cays tend to have a low 4-body invariant mass as two or more particles are not
2038 reconstructed.

2039 In order to remove this background in the muonic channels, the 4-body $m(K\pi\mu^+\mu^-)$
2040 invariant mass is recalculated with a kinematical fit using the `DecayTreeFitter`
2041 package. In the resonant case this includes a constraint of the dilepton mass to
2042 be the J/ψ nominal mass and in both rare and resonant cases each particles is
2043 constrained to point to its origin vertex. This constraint has the effect of pushing the
2044 misreconstructed events far from the B^0 peak. Therefore, to avoid this background,
2045 it is sufficient to limit the analysis to 4-body invariant masses above $5150 \text{ MeV}/c^2$.

2046 In the electron case it is instead important to fit a wider mass window to correctly
2047 constrain the background therefore one cannot eliminate this mis-reconstructed
2048 background which is then modelled in the fit (for details see Sec. 13.2.2).

2049 12.4 Multivariate analysis

2050 The final selection is performed using a Neural Network classifier (NN) based on the
2051 `NeuroBayes` package [60, 61]. The multivariate analysis is intended to remove some
2052 combinatorial background and obtain a clearer signal peak.

2053 For the final selection in the central and high q^2 intervals a Neural Network clas-
2054 sifier (NN) is used based on the `NeuroBayes` package [60, 61]. Representative
2055 samples of the signal and background are needed to train the classifier. For the
2056 signal, fully reconstructed $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$ simulated events
2057 can be used. A sample representative of the background can be obtained taking
2058 real data events in the upper B^0 sideband: $m(K\pi\mu^+\mu^-) > 5400 \text{ MeV}/c^2$ and
2059 $m(K\pi e^+ e^-) > 5600 \text{ MeV}/c^2$. The lower sideband is not used in the training as it
2060 contains a significant fraction of mis-reconstructed background. All pre-selection
2061 cuts are applied to the background samples used for the training. As L0 and PID

Particle	Variables
B^0	χ_{DTF}^2/ndf [1], DIRA [19], χ_{FD}^2 [15], χ_{vtx}^2/ndf [12], χ_{IP}^2 [14], p_T [7]
K^*	χ_{FD}^2 [21], χ_{vtx}^2/ndf [11], χ_{IP}^2 [2], p_T [5]
Dilepton	χ_{FD}^2 [17], χ_{vtx}^2/ndf [13], χ_{IP}^2 [20], p_T [6]
e	χ_{IP}^2 [3][4], p_T [9][10]
μ	χ_{IP}^2 [14][15], p_T [9][10]
K	χ_{IP}^2 [18], p_T [16]
π	χ_{IP}^2 [22], p_T [8]

Table 12.3: Variables used as inputs for the NN training. Next to each variable the ID number in brackets provides the index reported in the correlation matrices shown in Fig. 12.5.

variables are not well described these cuts are not applied in the Monte Carlo samples but their effect is taken into account by the event weight. To train the classifier 50% of the sideband events was used, keeping the other 50% for testing. For the signal sample a number of Monte Carlo events was used equal to the number available for the background sample.

The input to the NN consists of 22 variables containing information about the kinematic of the decays and the quality of tracks and vertices. All the variables used are listed in Tab. 12.3. The graphical representation of the correlation matrices are shown in Fig. 12.5, in these figures the variable with ID = 1 is the NN output and the other IDs are reported in Tab. 12.3. 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_{IP}^2 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. The list the 10 most important variables is reported in Tab. 12.4, together with information on the relative importance of each input. The meaning of the column headings in this table was already explained in Sec. 4.2.

Figure 12.6 shows neural network output distributions for signal and background. On this plot distributions from test samples are also overlaid in order to check for overtraining. The distributions follow the same shape but with different fluctuations so we conclude that we have no significant overtraining. In general we conclude that

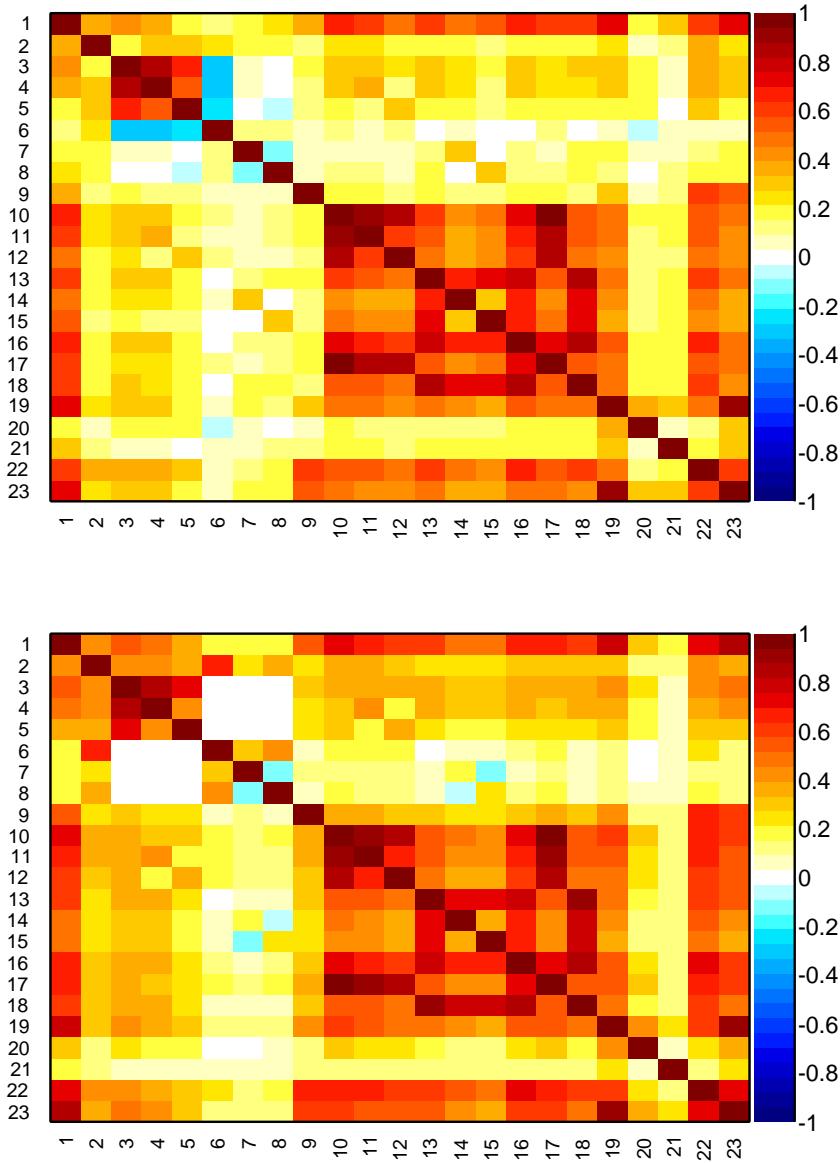


Figure 12.5: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of Tab. 12.4. Correlation is calculated using all events without distinguishing signal and background.

Table 12.4: Summary of inputs to the neural network in order of importance. The 10 most discriminating variables are shown. Column “adds” gives correlation significance added by given input when adding it to list of those ranked above, “only this” provides power of given input alone and “loss” shows how much information is lost when removing only given input. Decay Tree Fit is performed using DecayTreeFitter tool on whole decay chain with constraining tracks to appropriate vertex topology and the $m(p\pi)$ invariant mass to the PDG value.

Muons				Electrons			
Input	Adds	Only this	Loss	Input	Adds	Only this	Loss
$B^0 \chi_{DTF}^2/\text{ndf}$	80.44	80.44	13.14	$B^0 \chi_{DTF}^2/\text{ndf}$	28.70	28.70	3.94
$K^* \chi_{IP}^2$	22.26	67.58	3.48	$K^* \chi_{IP}^2$	12.71	25.11	1.57
$B^0 \text{DIRA}$	10.58	71.24	3.95	$e_2 \chi_{IP}^2$	6.56	20.19	3.30
$K^* p_T$	9.16	49.13	2.07	$e_1 \chi_{IP}^2$	5.54	19.66	2.60
$J/\psi \chi_{IP}^2$	6.58	56.15	1.35	$K^* p_T$	3.74	15.35	3.14
$B^0 p_T$	6.00	41.42	4.39	$J/\psi p_T$	4.81	5.55	3.18
$\mu_1 p_T$	2.96	15.85	3.79	$B^0 p_T$	2.78	13.01	2.20
$\mu_2 p_T$	2.73	15.04	3.46	πp_T	3.08	7.93	1.83
$J/\psi p_T$	3.06	16.41	2.84	$e_2 p_T$	2.35	9.81	2.74
$K^* \chi_{vtx}^2/\text{ndf}$	2.41	28.14	2.38	$e_1 p_T$	2.15	8.04	2.28

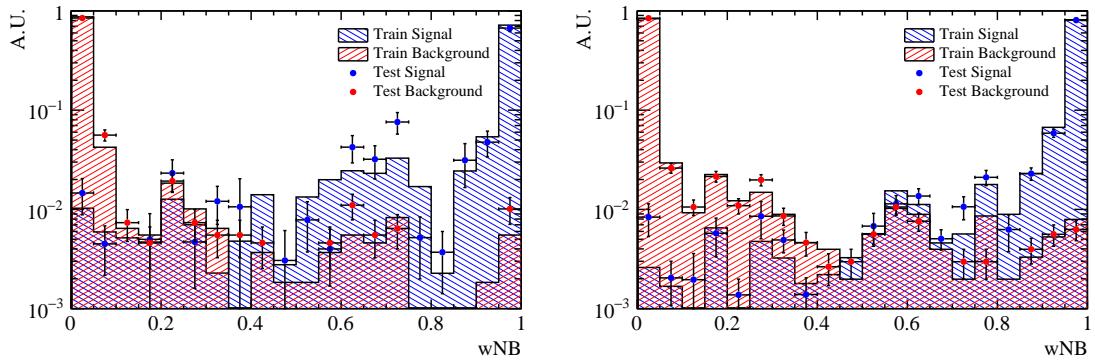


Figure 12.6: NN output distributions for training (solid) and test (stripes) samples, for simulated signal and data sideband events. For the electron (left) and muon (right) training.

2083 the neural network is able to separate signal from background and that the training
2084 converged properly.

2085 It can happen that too much information is given to the classifier, which becomes
2086 able to calculate the invariant mass of the candidates from its inputs. This could
2087 generate fake peaks and it is therefore important to check for correlations between
2088 the B^0 mass and the NN output. Fig 12.7 reports plots of the average NN output
2089 as a function of the B^0 mass on sideband data and simulated signal events. The
2090 distributions are flat showing that no significant correlation is present.

2091 12.5 MVA optimisation

2092 In order to optimise the cut on our neural network output the expected signal
2093 significance, $N_S/\sqrt{N_S + N_B}$, was maximised. In this formula N_S is number of rare
2094 signal events and N_B the number of background events.

2095 The number of signal events accepted for a given NN output cut is determined
2096 exploiting the resonant channel and simulation. First, as an arbitrary number of
2097 events can be simulated, this has to be rescaled to the expected yield. This is done
2098 by fitting $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ events after pre-selection, including all selection
2099 cuts except MVA. The resonant yield is then scaled down by the expected ratio
2100 between the rare and the resonant channels. The number of background events is
2101 instead derived by fitting the combinatorial background in the sideband with an
2102 exponential function and extrapolating the fit function below the signal peak.

2103 The dependence of the figure-of-merit for both the electron and muon trainings are
2104 shown in Fig.12.8, where the red line indicate the chosen cut: 0.75 for both samples.
2105 Curves of signal efficiency versus background rejection are shown in Fig. 12.8. Using
2106 the described MVA cuts the signal efficiency is $\sim 91\%$ for the muon channels and
2107 $\sim 84\%$ for the electron channels (for more details see Sec. 14), while the background
2108 rejections is $\sim 98\%$ on both samples.

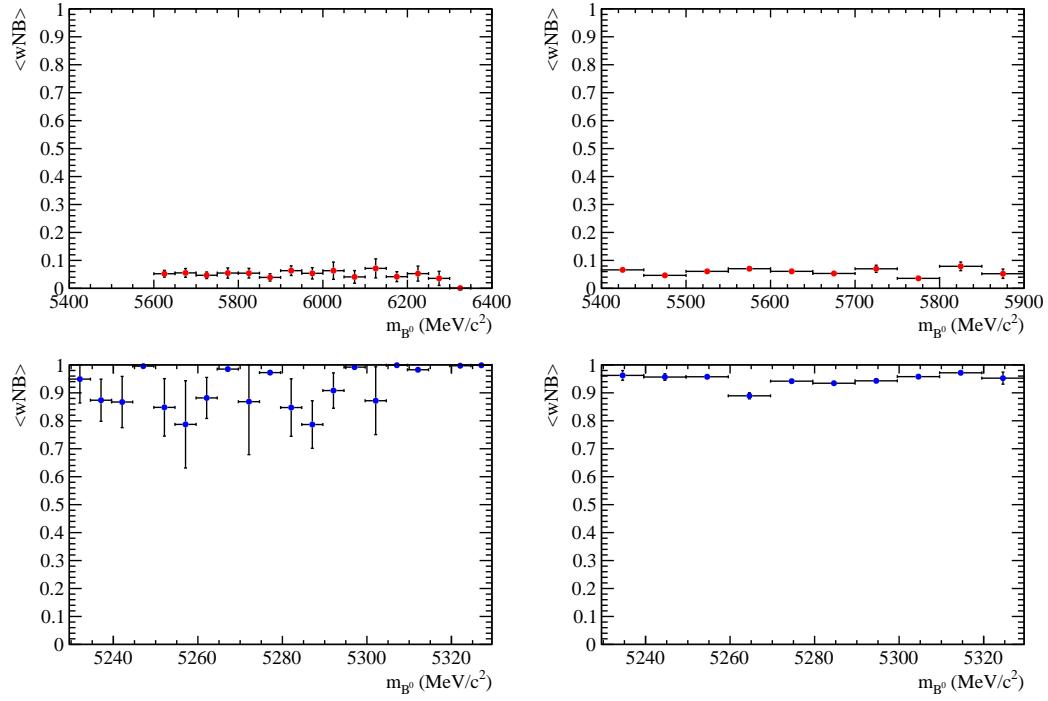


Figure 12.7: Average value of NN output as a function of B^0 mass for data sideband (top) and simulated signal (bottom) events for the electron (left) and muon (right) training.

- ²¹⁰⁹ After full selection about $\sim 3\%$ of events still contain multiple candidates which are
- ²¹¹⁰ removed at random keeping only a single candidate per event.

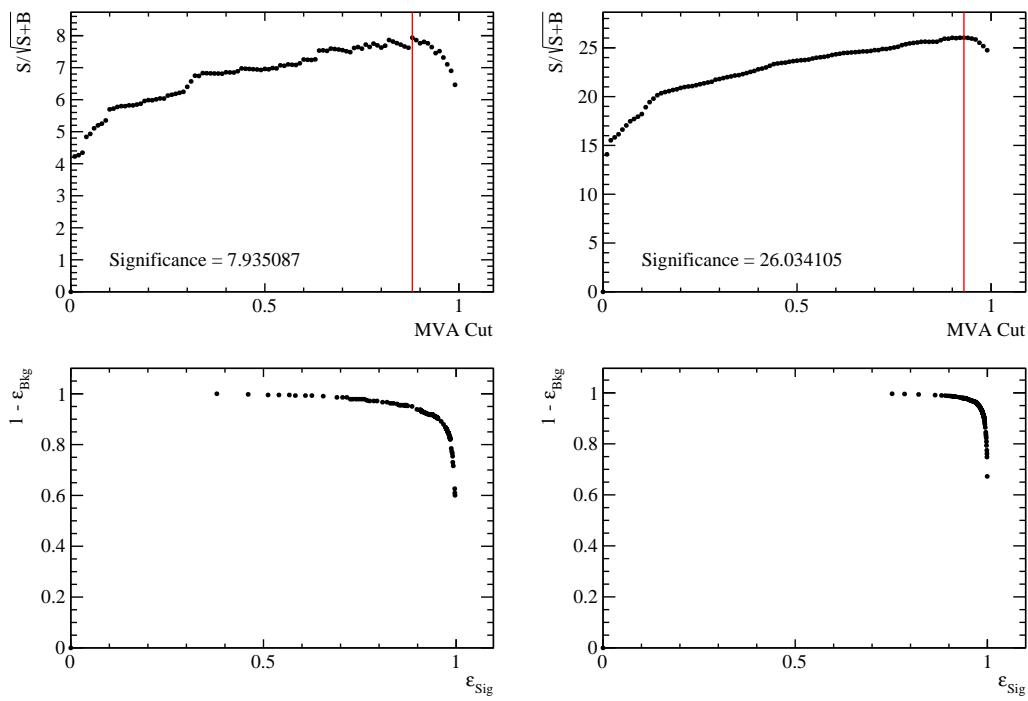


Figure 12.8: (top) Dependence of figure-of-merit on the requirement on neural network output. Vertical lines corresponds to the chosen cuts. (bottom) Signal efficiency versus the background rejection. Plots correspond to the electron (left) and muons (right) samples.

CHAPTER 13

2111

2112

2113

Mass fits

2114

2115 In order to extract the signal yields a simultaneous unbinned maximum likelihood
2116 fit to the 4-body invariant masses, $m(K\pi\ell\ell)$, is performed on the rare and resonant
2117 samples. The following subsections contain a description of the line shapes used to
2118 model the signal and background components in each sample. The simultaneous fit
2119 allows to share parameters between the two e.g. those describing data-simulation
2120 differences. The yields of the rare channels are parameterised as a function of the
2121 corresponding J/ψ yields as

$$N_{\ell\ell} = N_{J/\psi} \cdot \varepsilon^{\text{rel}} \cdot R_{\ell\ell}, \quad (13.1)$$

2122 where ε^{rel} is the relative efficiency between the rare and resonant channels given in
2123 Tab. 14.4 and $R_{\ell\ell}$ corresponds to the efficiency corrected ratio of the raw rare and
2124 resonant yields:

$$R_{\ell\ell} = \frac{N_{\ell\ell}/\varepsilon^{\ell\ell}}{N_{J/\psi}/\varepsilon^{J/\psi}}. \quad (13.2)$$

2125 The two ratios, R_{ee} and $R_{\mu\mu}$, are then used to determine the $R_{K^{*0}}$ quantity, as
2126 described in Sec. 16.

2127 13.1 Muon channels

2128 For the rare and resonant $\mu^+\mu^-$ channels the fitted variable is the $m(K\pi\mu\mu)$ in-
2129 variant mass coming from a kinematic fit where all vertices are required to point
2130 to their mother particle. In the resonant case it is beneficial to also constrain the
2131 the dimuon mass to the known J/ψ mass. The effect of the kinematical fit is to
2132 improve the mass resolution by roughly a factor of 2, which results a more stable
2133 fit. Furthermore, misreconstructed events are pushed away from the B^0 peak, which
2134 allows to use a wider mass window to better constrain the combinatorial background
2135 slope. The mass spectrum is fitted in the range 5150–5800 MeV/ c^2 with the lower
2136 limit of the mass range chosen to exclude partially reconstructed background. As it
2137 is not needed to model misreconstructed backgrounds in the fit this also eliminates
2138 systematic uncertainties associated with the knowledge of its shape.

2139 The PDF chosen to describe the signal in both the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and its relative
2140 J/ψ channel is a Double Crystal Ball function already described in Sec. 5.1 and also
2141 in this case the mean value (m_0) is kept in common:

2142 As a first step simulated distributions are fit using the signal model to extract
2143 parameters to be constrained in the fit to data. The fitted MC distribution for the
2144 resonant channel is reported in Fig. 13.1.

2145 For the fit to real data the signal parameters are fixed to the ones found for the
2146 simulated samples. In order to account for possible data-simulation discrepancies
2147 a scale factor is multiplied to the widths and a shift is added to the masses. In

2148 summary the PDFs used for the signal fits on data are defined as

$$\begin{aligned} P(m; c, m') = & f^* \cdot C(m; \alpha_1^*, n_1^*, c \cdot \sigma_1^*, m_0^* + m') \\ & + (1 - f^*) \cdot C(m; \alpha_2^*, n_2^*, c \cdot \sigma_2^*, m_0^* + m') \end{aligned} \quad (13.3)$$

2149 where the free parameters are the width scale factor, c , and the mass shift, m' ,
2150 which are common between the rare and resonant samples. The other parameters,
2151 denoted with $*$, are taken from the fit to simulated events, separately for the rare and
2152 resonant samples and are fixed in the fit on data. The parameter f^* in the formula
2153 is the relative fraction of candidates falling in the first Crystal Ball function.

2154 To model the combinatorial background an exponential function was used. This is
2155 the only background component for the rare channel. In the normalisation channel
2156 fit the $B_s^0 \rightarrow K^{*0} J/\psi$ background is described using the same PDF used for the sig-
2157 nal but a different central value, m , which is set at the B_s^0 nominal mass [2]. Finally,
2158 a $\Lambda_b^0 \rightarrow J/\psi pK$ background component is modelled using simulated $\Lambda_b^0 \rightarrow J/\psi pK$
2159 events to which the full $B^0 \rightarrow K^{*0} J/\psi$ selection is applied. The invariant mass dis-
2160 tribution of these candidates is a broad shape under the signal peak. The simulated
2161 distribution is smoothed using a kernel estimation method (using the `RooKeysPdf`
2162 class of the `RooFit` package [97]).

2163 In summary the free parameters in the simultaneous fit to rare and resonant $\mu^+ \mu^-$
2164 data samples are: the signal and background yields, the combinatorial background
2165 slopes, the widths scale, c , and the the mass shift, m' .

2166 Fig. 13.2 reports fits to real data distributions for the rare and resonant $\mu^+ \mu^-$
2167 channels. Values of fitted parameters are reported on the plots.

2168 13.2 Electron channels

2169 In the electron case the variable we fit is the $m(K\pi e^+ e^-)$ invariant mass coming
2170 from the kinematic fit where all vertices are required to point to their mother par-

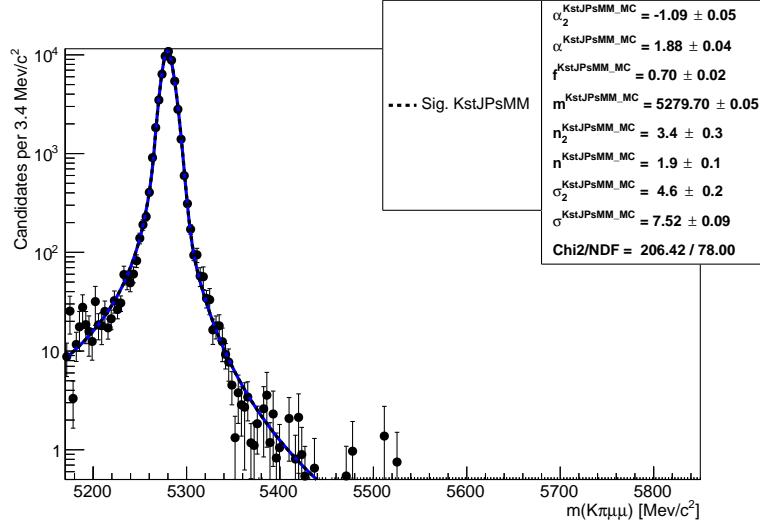


Figure 13.1: Fitted $m(K\pi\mu\mu)$ mass spectrum for $K^{*0}J/\psi$ simulated events.

ticle. In contrast to the muon channel, the constraint to the dilepton mass to the nominal J/ψ nominal value is not applied. In fact, due to the longer bremsstrahlung tail, the J/ψ mass constraint distorts the invariant mass distribution and makes it is hard to model it. Furthermore, mis-reconstructed background enters in the rare channel sample and its amount can be constrained by exploiting the higher statistics resonant channel, but this implies the usage of the same variable in both fits. In order to better constrain the parameters modelling the radiative tail and the mis-reconstructed backgrounds a wide mass window is used [4500,5800] MeV/c^2 . 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.

In the electron case the invariant mass distributions are different depending on which hardware trigger was used and especially how many bremsstrahlung photons were recovered. Therefore our sample is divided in 3 trigger categories, as described in Sec. 12.1, and 3 bremsstrahlung categories defined as:

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

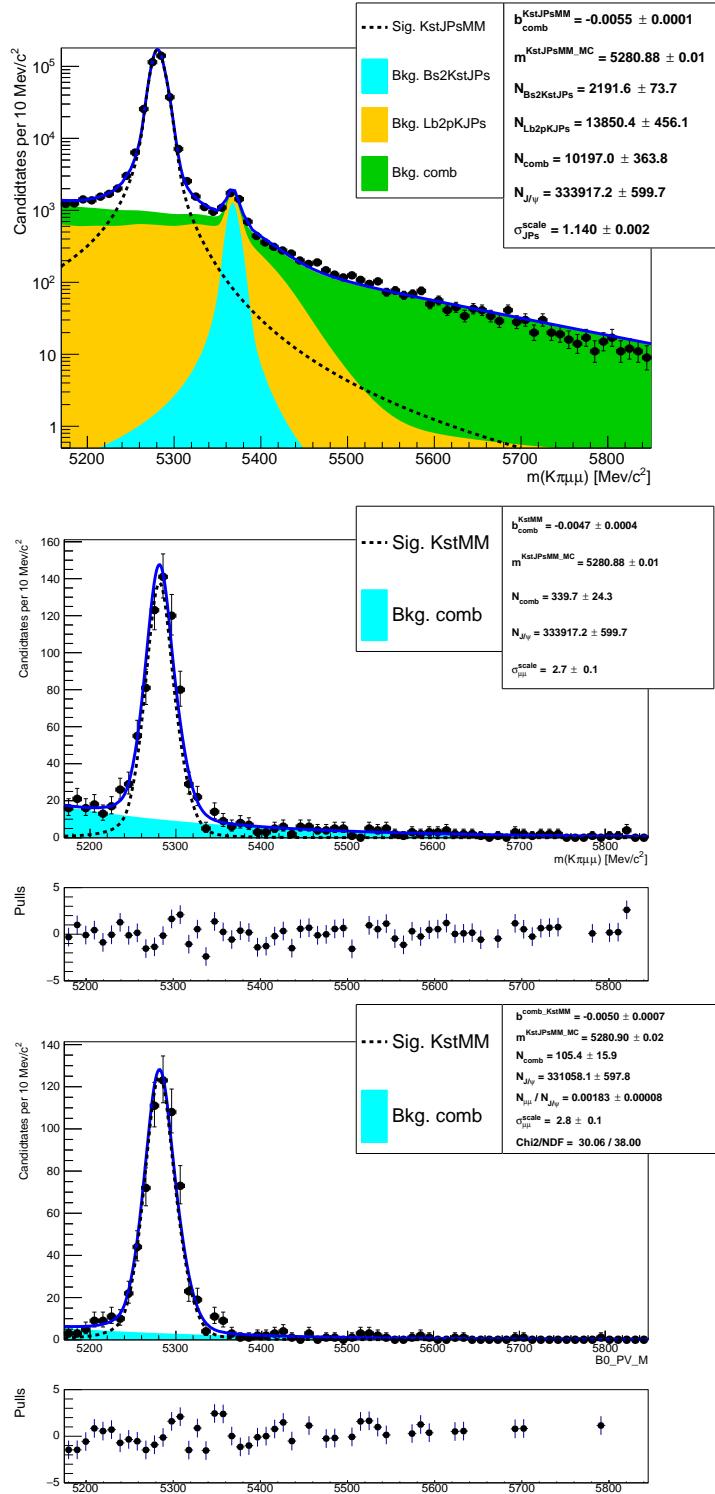


Figure 13.2: Fitted $m(K\pi\mu\mu)$ invariant mass distribution for $K^{*0}J/\psi$ candidates (top) and for rare candidates in the central (bottom left) and high (bottom right) q^2 intervals. Dashed black lines represent the signal PDFs and filled shapes the background components.

- 2188 The three samples, divided by trigger, are fitted simultaneously. This allows a better
2189 use of statistics as the simultaneous fit gathers information from the three categories
2190 at the same time and is more stable. Furthermore using this method the results for
2191 the three categories are naturally combined in a single R_{ee} ratio.
- 2192 The PDFs used to fit the invariant mass distributions are described in the next
2193 section.

2194 13.2.1 Signal PDFs for the electron channels in the central- q^2 interval

2195 As for the muonic channel simulated events are fitted at first to constrain the shapes
2196 for the subsequent fit on data. The signal PDFs are built using the following method:

- 2197 • Simulated $B^0 \rightarrow K^{*0}J/\psi(ee)$ and $B^0 \rightarrow K^{*0}ee$ events are divided in each
2198 trigger and bremsstrahlung category and an independent fit is performed to
2199 each sample.
- 2200 • For each trigger category a PDF is built as the sum of the three PDFs for each
2201 bremsstrahlung category.

$$P(m)^{\text{trg}} = f_{0\gamma}^{\text{trg}} P_{0\gamma}^{\text{trg}}(m) + f_{1\gamma}^{\text{trg}} P_{1\gamma}^{\text{trg}}(m) + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) P_{2\gamma}^{\text{trg}}(m). \quad (13.4)$$

2202 where the $P(x)_{n\gamma}^{\text{trg}}$ functions are the chosen PDFs for each trigger and bremsstrahlung
2203 category and the $f_{n\gamma}^{\text{trg}}$ parameters are the relative fractions of events falling in
2204 each category.

- 2205 • Most parameters are fixed (details later) and the combined PDF, $P(m)$, is
2206 used to fit real data divided only in trigger categories.

2207 The 0γ category is characterised by a better resolution and a sharp tail on the right-
2208 hand side and it is fitted with a simple Crystal Ball function (CB). Instead the 1γ
2209 and 2γ samples are modelled using the sum of a Crystal Ball and a Gaussian func-
2210 tions (CBG) with all parameters independent. When the combined PDF, $P(m)^{\text{trg}}$,

Table 13.1: Percentages of events with 0, 1 and 2 emitted photons in the three trigger categories, extracted from simulated events.

Trigger	0γ	1γ	2γ
J/ψ			
L0E	28.3 %	50.5 %	21.2 %
L0H	18.1 %	51.0 %	30.9 %
L0I	25.1 %	52.0 %	22.9 %
$1-6 \text{ GeV}^2/c^4$			
L0E	30.1 %	50.2 %	19.7 %
L0H	23.1 %	51.7 %	25.2 %
L0I	28.5 %	50.8 %	20.7 %

is built all parameters are fixed leaving one global mass shift and one scale factor for the widths free to vary, as done for the muonic samples.

Finally, combining the three bremsstrahlung PDFs one needs to specify in which fractions they contribute to the total. These fractions have been shown to be in good agreement between resonant data and Monte Carlo and therefore they are fixed to the values found on simulation, separately for the normalisation channel and each q^2 bin. In Tab. 13.1 are reported percentages of events with 0, 1 and 2 recovered photons in the three trigger categories.

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

$$P(m; c, m')^{\text{trg}} = f_{0\gamma}^{\text{trg}} \text{CB}_{0\gamma}^{\text{trg}}(m; c, m') + f_{1\gamma}^{\text{trg}} \text{CBG}_{1\gamma}^{\text{trg}}(m; c, m') + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) \text{CBG}_{2\gamma}^{\text{trg}}(m; c, m') \quad (13.5)$$

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

13.2.2 Background PDFs for the electron channels in the central- q^2 interval

In the fit to the resonant sample three background components are modelled: combinatorial background, and misreconstructed background coming from the hadronic

and the leptonic systems. The combinatorial is described with an exponential function.

The misreconstructed background is split in two categories, that involving higher hadronic resonances and that coming from higher $c\bar{c}$ resonances. These backgrounds are modelled using inclusive $B^0 \rightarrow J/\psi X$ simulated samples to which the full selection is applied. The distributions for the hadronic (leptonic) background are defined selecting candidates where the K^{*0} (dimuon) is not a direct daughter of the B^0 . The invariant mass distributions of these events, shown in Fig. 13.3, are smoothed using a kernel estimation method and their yields are left floating in the fit. Given the low statistics available, the same shape is used for the three trigger categories.

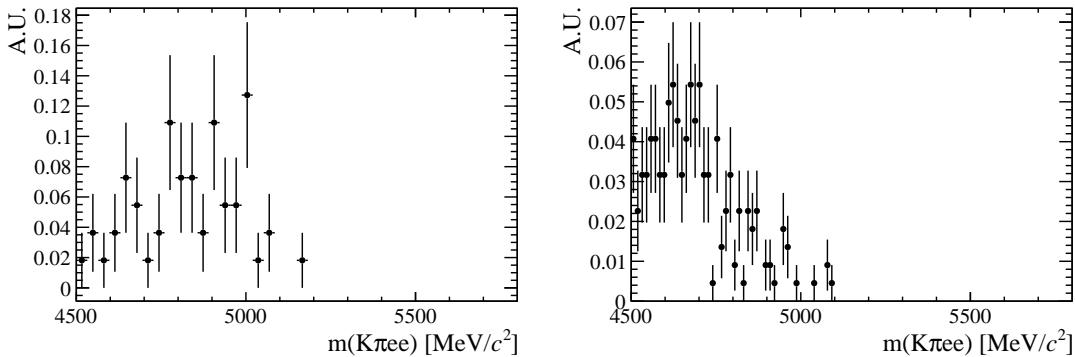


Figure 13.3: Simulated distributions of misreconstructed background events falling into the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ sample coming from the hadronic (left) and leptonic (right) systems.

In the fit for the rare sample in the central- q^2 interval the modelled backgrounds are: combinatorial background, again modelled with an exponential; misreconstructed background coming from the hadronic system and the leakage of the J/ψ radiative tail into the lower q^2 interval. The shape for the misreconstructed component is obtained from simulated distributions similarly to what described for the resonant channel. However, as there are no inclusive samples for the rare case, a sample including higher K^* resonances is generated, including $K_1^+(1400)$ and $K_2^+(1460)$. The yield of this component is not floating independently but its relative proportion with respect to the signal yield is constrained to be the same as in the resonant

sample, namely:

$$N_{\ell\ell}^{mis-reco} = N_{ee} \cdot k = N_{ee} \cdot \frac{N_{J/\psi}^{mis-reco}}{N_{J/\psi}}. \quad (13.6)$$

Notice that, as the fit is simultaneous for the rare and resonant samples, this fraction is not fixed in the fit but floats using information from both samples.

The shape to describe the J/ψ tail leakage is obtained using simulated $B^0 \rightarrow J/\psi K^{*0}$ candidates and selecting those falling in q^2 below $6 \text{ GeV}^2/c^4$. The 4-body invariant mass distribution of these events is reported in Fig. 13.4. The yield of this component again is not floating independently but it is linked to the yield found in the resonant fit as follows

$$N_{\ell\ell}^{leak} = N_{J/\psi} \cdot k^{MC} = N_{J/\psi} \cdot \frac{N_{leak}^{MC}}{N_{J/\psi}^{MC}} \quad (13.7)$$

where k is the ratio between $N_{J/\psi}^{MC}$, the number of J/ψ events that fall into the $J/\psi q^2$ window ($6-11 \text{ GeV}^2/c^4$) in the simulation and N_{leak}^{MC} , the number of J/ψ events leaking below $6 \text{ GeV}^2/c^4$ in the simulation. In this case k is previously extracted from simulated events and fixed in the fit on data.

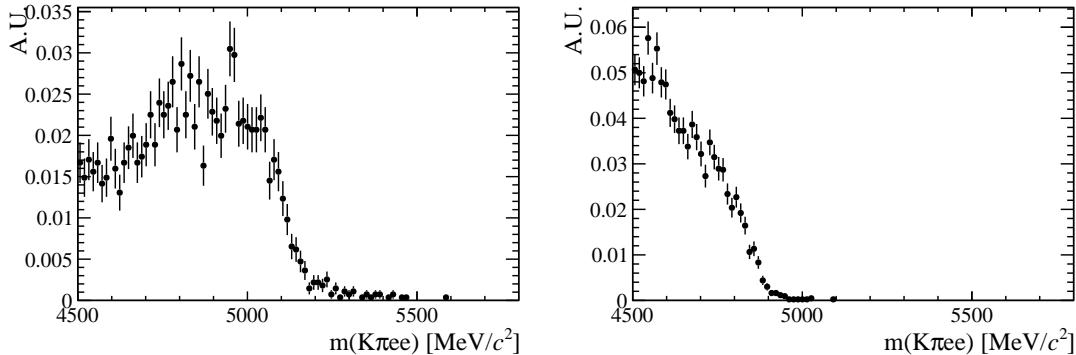


Figure 13.4: (left) Simulated 4-body invariant mass distributions for events involving higher K^* states and passing out full selection. (right) Simulated invariant mass distribution of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ events leaking into the central- q^2 interval.

13.2.3 Summary of the fit to electron channels in the central- q^2 interval

In summary in the resonant fit on data the floating parameters are the yields of all the components in the resonant channel, a common R_{ee} ratio, the combinatorial

Table 13.2: Percentages of events with 0, 1 and 2 emitted photons in the three trigger categories, extracted from simulated events.

Sample	0γ	1γ	2γ
$\psi(2S)$ (L0E)	25.7 %	52.1 %	22.2 %
15–20 GeV^2/c^4 (L0E)	20.7 %	51.7 %	27.6 %

background yield in the rare sample, one scale factor c , one mass shift m_0 and the combinatorial background slopes.

Fits on simulated $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates are shown in Appendix D. Figure 13.5 and 13.6 present fits on real $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}e^+e^-$ candidates (central- q^2 interval) in the three trigger categories. Values of fitted parameters are reported on the plots.

13.2.4 Electron channels fits in the high- q^2 interval

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 $4500 < m(K\pi ee) < 6000 \text{ MeV}/c^2$. In the L0TIS category, even if the yield is bigger a clear peak cannot be seen, therefore only L0Electron triggered events are used in this region.

The signal PDF is described in the same way as for the central bin. Simulated events are divided in three bremsstrahlung categories and fitted using the same PDFs described in Sec. 13.2.1. While the signal tail parameters are similar for the J/ψ and central- q^2 samples in the case of the high- q^2 interval it is particularly important to keep them independent. In fact, as can be seen in Fig. 13.7, the invariant mass distributions are significantly different for the two intervals. The fractions of 0, 1 and 2 γ components used to build the total PDF are also in this case taken from simulated events are are reported in Tab. 13.2.

The background components, as for the central- q^2 interval, include a combinatorial background and a misreconstructed background coming from the hadronic system.

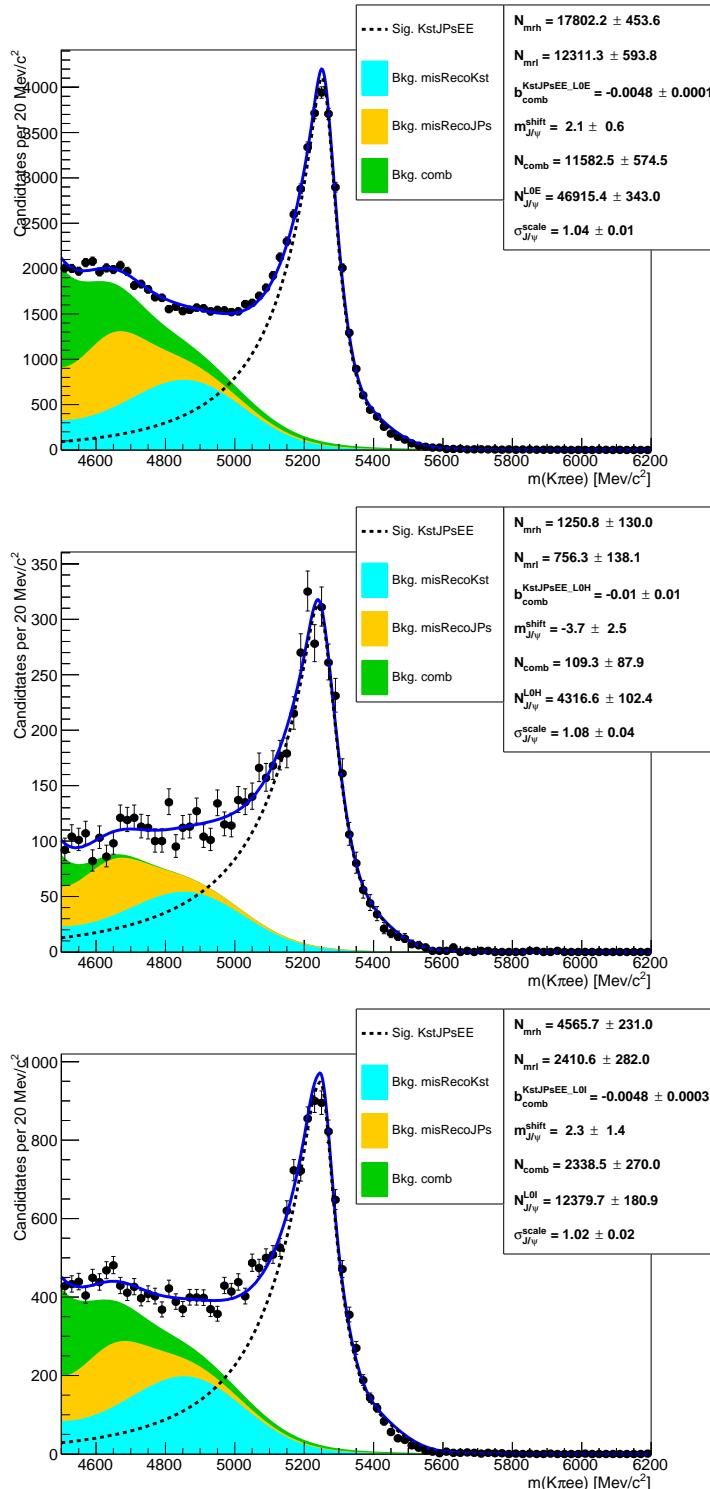


Figure 13.5: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

Furthermore there is a leakage due to the $\psi(2S)$ resonance, that is wide enough to contribute in q^2 above 15 GeV^2/c^4 . The combinatorial background is modelled using

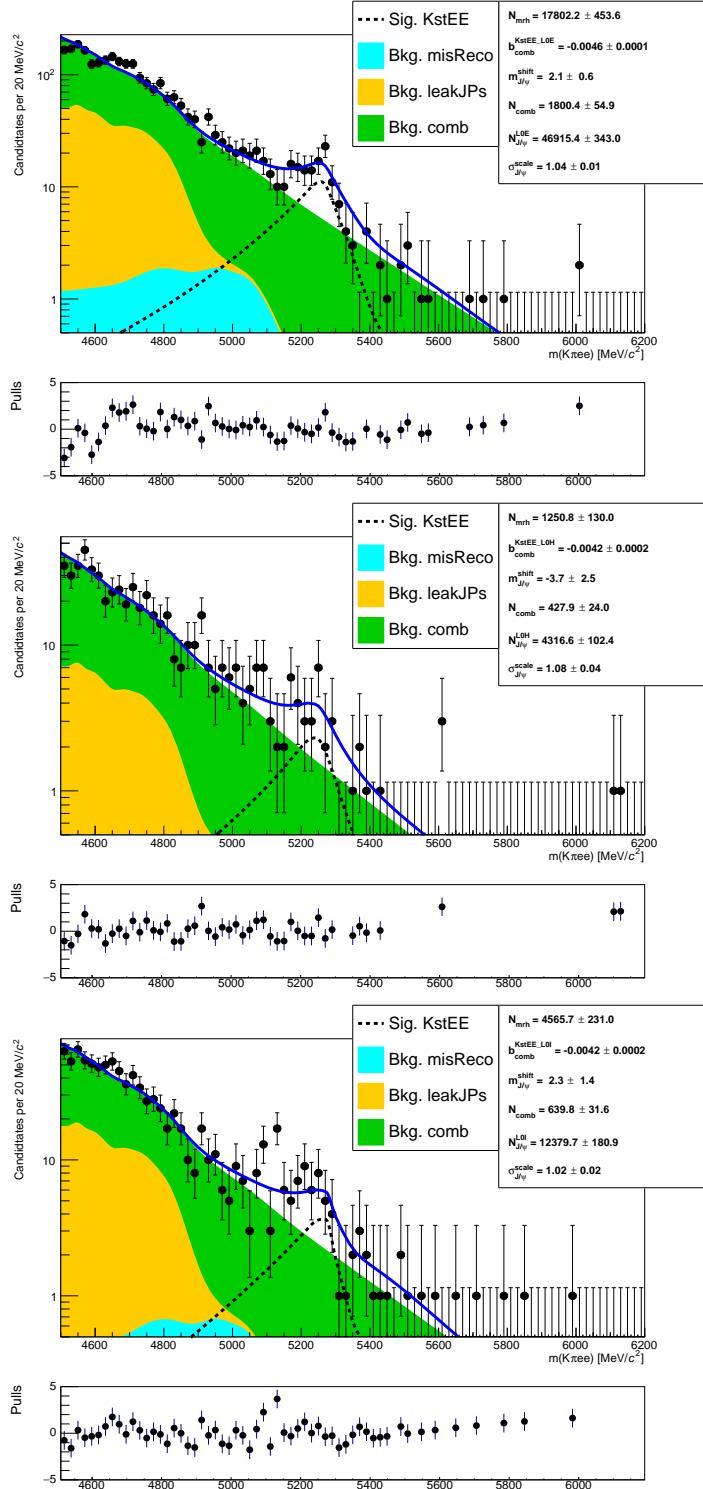


Figure 13.6: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} e^+ e^-$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

2284 the shape obtained by reversing the NN output cut. In Fig. 13.8 are shown invariant
 2285 mass distributions for different NN output cuts using the electron and muon high- q^2

samples. The shapes are normalised to the same integral. This plot shows that the shape is similar for the two samples and as a function of the NN output cut, which supports the hypothesis that these distributions contain mostly combinatorial background and that its shape has no strong dependency on the NN output cut. The shape is therefore obtained from the muon sample, which has higher statistics, selecting events with NN output < 0.1 . The distribution is then smoothed using the `RooKeysPdf` method.

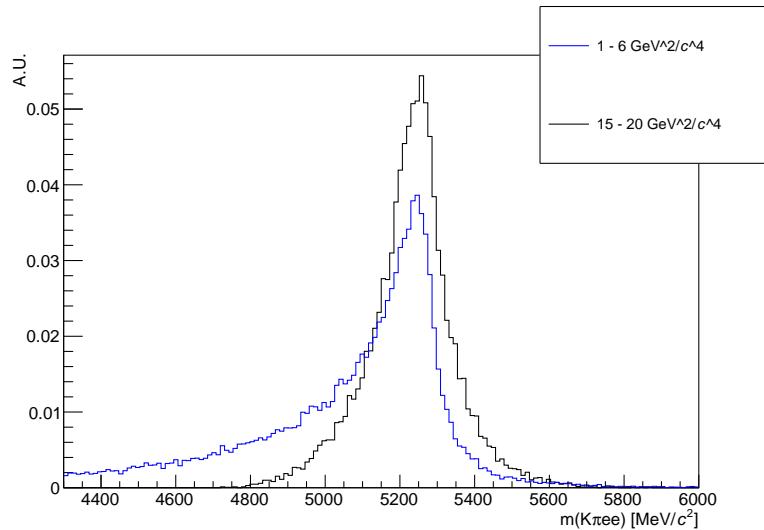


Figure 13.7: Simulated invariant mass of the $K\pi ee$ system in the $1.1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.

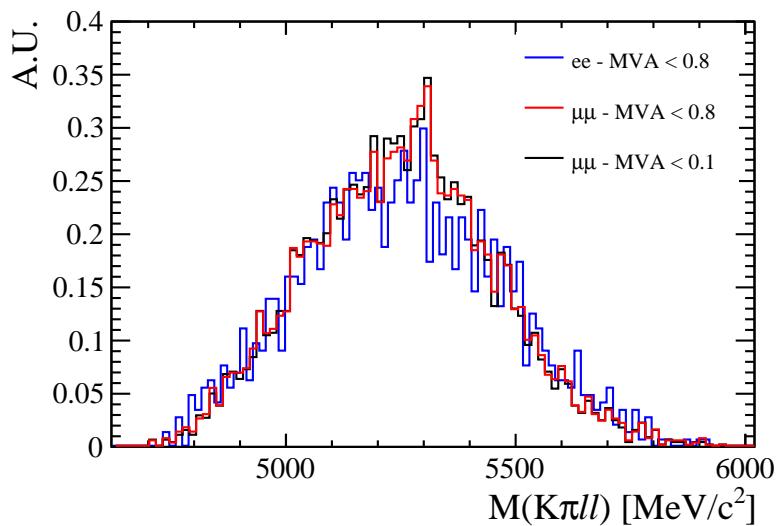


Figure 13.8: Invariant mass distribution of candidates selected by reversed NN output cuts.

2293 The misreconstructed component is modelled in the same way described for the
2294 central- q^2 interval. However, in this case, its yield is not constrained to the resonant
2295 channel.

2296 The $\psi(2S)$ leakage component is modelled from $B^0 \rightarrow K^*(\psi(2S) \rightarrow e^+e^-)$ simulated
2297 events with the same method used for the J/ψ leakage in the central- q^2 interval. The
2298 yield of this component is fixed to the yield of $\psi(2S)$ as

$$N_{\ell\ell}^{leak} = N_{\psi(2S)} \cdot k^{MC} = N_{\psi(2S)} \cdot \frac{N_{leak}^{MC}}{N_{\psi(2S)}^{MC}}. \quad (13.8)$$

2299 In order to do this the $\psi(2S)$ yield, $N_{\psi(2S)}$, is obtained from a fit to the $\psi(2S)$
2300 invariant mass peak. Since we are only interested in the $\psi(2S)$ yield we fit the
2301 $m(K\pi ee)$ obtained from a kinematic fit where the dimuon mass is constrained to
2302 the known $\psi(2S)$ mass. This allows to eliminate the misreconstructed background
2303 form the fit mass window and use a simple model composed by a signal component
2304 and a combinatorial background component. The signal is described with a Double
2305 Crystal Ball function with opposite tails already described the Λ_b^0 fits (see Sec. 5.1),
2306 and the combinatorial background is described with an exponential. The fit to the
2307 $\psi(2S)$ peak is reported in Fig. 13.9 together with the fit to the $B^0 \rightarrow K^*e^+e^-$
2308 candidates in the high- q^2 interval.

2309 13.3 Fit summary

2310 In Tab. 13.3 are reported raw yields obtained from the fits described in the previous
2311 sections. The values for the rare channels are not directly floating in the fits but as
2312 described in Sec. 13 they are parameterised as a function of the number of resonant
2313 events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare branching
2314 fractions. Measured values of these ratios are reported in Tab. 16.2.

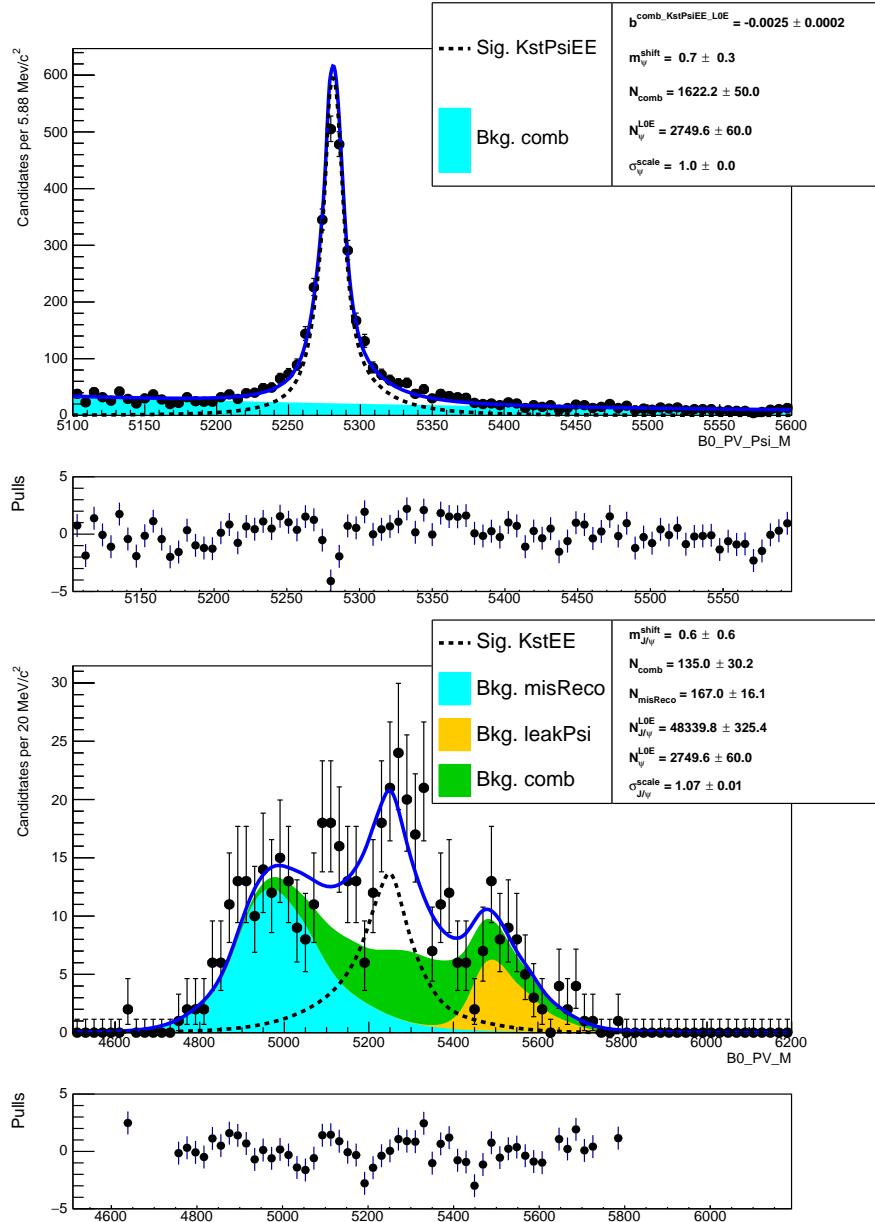


Figure 13.9: Fitted $m(K\pi ee)$ invariant mass distribution in the $\psi(2S)$ interval, $11 < q^2 < 15$ GeV²/c⁴ and in the high- q^2 interval.

Sample	1–6 GeV ² /c ⁴	15–20 GeV ² /c ⁴	J/ψ
$\mu\mu$	625.38 ± 29.60	606.87 ± 27.56	333917.20 ± 599.73
ee LOE	131.77 ± 18.06	132.28 ± 27.92	48103.10 ± 329.77
ee L0H	32.50 ± 4.50	—	4439.51 ± 98.38
ee L0I	48.53 ± 6.68	—	12683.18 ± 174.25

Table 13.3: Raw yields of events found fitting invariant mass distributions of the rare and resonant events.

2315

CHAPTER 14

2316

2317

Efficiency

2318

2319 The efficiency for each of the decay channels is calculated according to the formula

$$\varepsilon^{tot} = \varepsilon(geom)\varepsilon(reco|geom)\varepsilon(PID|reco)\varepsilon(trig|PID)\varepsilon(MVA|trig). \quad (14.1)$$

2320 In this expression the first term is the efficiency to have final state particles in the
2321 LHCb detector acceptance. The second term carries information on reconstruction
2322 and stripping efficiency (we keep these together given that boundaries between them
2323 are completely artificial). The third part corresponds to the efficiency of the PID
2324 requirements. The fourth term handles the trigger efficiency for those events which
2325 are selected by the preselection process. Finally, the latter term deals with the
2326 efficiency of the NN classifier. Reconstruction, trigger and MVA efficiencies are
2327 evaluated on simulated data with the trigger efficiency for $B^0 \rightarrow K^* J/\psi$ being
2328 cross-checked using the data-driven TISTOS method as described in Sec. 6.3. The
2329 PID efficiency is calculated with a data-driven method as described in Sec. 14.3.

2330 All absolute efficiencies for the muon and electron rare channels are separately listed
2331 in Tab. 14.3 for the central and high q^2 intervals and in Tab. 14.1 for the resonant
2332 channels. However for the analysis itself only efficiencies relative to the resonant
2333 channels are used in order to limit systematic uncertainties.

2334 Tab. 14.3 reports relative efficiencies between the rare and resonant channels, $\varepsilon(B^0 \rightarrow$
2335 $K^{*0}\ell^+\ell^-)/\varepsilon(B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-))$. Finally, in Tab. 14.2 are listed ratios of rel-
2336 ative efficiencies for the ee and $\mu\mu$ channels, $[ee/(J/\psi \rightarrow ee)]/[\mu\mu/(J/\psi \rightarrow \mu\mu)]$.

Table 14.1: Absolute efficiencies for the resonant ee and $\mu\mu$ channels.

ε	$\mu\mu$	ee		
		L0E	L0H	LOI
Geom	0.1598 ± 0.0005		0.1589 ± 0.0005	
Reco	0.0947 ± 0.0001		0.0603 ± 0.0001	
PID	0.8148 ± 0.0000		0.8222 ± 0.0000	
Trig	0.7620 ± 0.0005	0.1976 ± 0.0005	0.0166 ± 0.0002	0.0720 ± 0.0003
MVA	0.8958 ± 0.0004	0.8586 ± 0.0007	0.8974 ± 0.0006	0.8260 ± 0.0017
Tot	0.0084 ± 0.0000	0.0013 ± 0.0000	0.0001 ± 0.0000	0.0005 ± 0.0000

Table 14.2: Double ratios of efficiencies $(\varepsilon^{ee}/\varepsilon^{J/\psi \rightarrow ee})/(\varepsilon^{\mu^+\mu^-}/\varepsilon^{J/\psi \rightarrow \mu^+\mu^-})$ in the cen-
 tral and high q^2 intervals.

ε	1–6 GeV^2/c^4			15–20 GeV^2/c^4
	L0E	L0H	LOI	L0E
q^2		0.697 ± 0.007		0.770 ± 0.010
Geom		1.022 ± 0.012		1.022 ± 0.012
Reco		0.931 ± 0.006		0.443 ± 0.443
PID		0.982 ± 0.000		0.970 ± 0.000
Trig	0.890 ± 0.008	2.449 ± 0.049	1.239 ± 0.016	1.416 ± 0.014
MVA	0.969 ± 0.003	0.941 ± 0.003	0.972 ± 0.006	1.061 ± 0.006
Tot	1.155 ± 0.023	3.086 ± 0.083	1.613 ± 0.037	0.856 ± 0.020

Table 14.3: Absolute efficiencies for the rare ee and $\mu\mu$ channels in the central and high q^2 intervals.

ε	$\mu\mu$		ee		15–20 GeV^2/c^4
	1–6 GeV^2/c^4	15–20 GeV^2/c^4	L0E	L0H	
q^2	0.2142 ± 0.0015	0.1552 ± 0.0013		0.1493 ± 0.0012	0.1196 ± 0.0011
Geom	0.1630 ± 0.0014	0.1630 ± 0.0014		0.1657 ± 0.0012	0.1657 ± 0.0012
Reco	0.0177 ± 0.0001	0.0110 ± 0.0001		0.0105 ± 0.0000	0.0031 ± 0.0000
PID	0.7824 ± 0.0002	0.8420 ± 0.0001		0.7750 ± 0.0001	0.8239 ± 0.0001
Trig	0.6852 ± 0.0029	0.8456 ± 0.0028	0.1581 ± 0.0012	0.0366 ± 0.0006	0.0802 ± 0.0009
MVA	0.9097 ± 0.0022	0.8298 ± 0.0032	0.8447 ± 0.0021	0.8571 ± 0.0020	0.8156 ± 0.0046
Tot	0.0066 ± 0.0001	0.0068 ± 0.0001	0.0012 ± 0.0000	0.0003 ± 0.0000	0.0006 ± 0.0000

Table 14.4: Relative efficiencies, $\varepsilon^{rel} = \varepsilon^\ell / \varepsilon^{J/\psi}$, for the ee and $\mu\mu$ channels in the central and high q^2 intervals.

ε	1–6 GeV^2/c^4		15–20 GeV^2/c^4		15–20 GeV^2/c^4
	$\mu\mu$	L0E	ee	L0H	
Geom	1.0200 ± 0.0091		1.0429 ± 0.0084		1.0200 ± 0.0091
Reco	0.1873 ± 0.0011		0.1743 ± 0.0006		0.1159 ± 0.0009
PID	0.9602 ± 0.0002		0.9425 ± 0.0001		1.0334 ± 0.0001
Trig	0.8993 ± 0.0038	0.8002 ± 0.0065	2.2025 ± 0.0434	1.1138 ± 0.0136	1.1098 ± 0.0037
MVA	1.0154 ± 0.0025	0.9838 ± 0.0025	0.9551 ± 0.0023	0.9874 ± 0.0060	0.9262 ± 0.0036
Tot	0.7821 ± 0.0108	0.9033 ± 0.0132	2.4136 ± 0.0558	1.2618 ± 0.0228	0.8092 ± 0.0126

2337 14.1 Geometric efficiency

2338 The simulated samples used contain the requirement that daughters are in the LHCb
 2339 detector acceptance. This corresponds to the requirement for each of the final par-
 2340 ticles to have polar angle θ between 10 and 400 mrad. The efficiency of this cuts is
 2341 obtained using a generator level simulated sample.

2342 14.2 Reconstruction efficiency and bin migration

2343 The reconstruction efficiency is here defined as the efficiency to reconstruct each
 2344 decay channel given that its daughters are into the geometrical acceptance of the
 2345 detector. This includes both the probability that a particle generates observable sig-
 2346 natures and the efficiency of all the preselection cuts described in Sec. 12, including
 2347 those done to remove peaking backgrounds. The efficiency of the PID cuts is kept
 2348 separate as it is known to be not well simulated and there are reliable data-driven
 2349 methods which can be used to extract it (see Sec. 14.3).

2350 It can happen that events generated in a q^2 interval are reconstructed in a different
 2351 one, this is referred to as “bin migration”. Two different effects can cause bin
 2352 migration. First of all, as the resolution of real detectors is not perfect, events
 2353 close to the edges of the considered intervals can fall on the wrong side of the
 2354 edge. This effect is only important in case of non-flat true distributions, as the
 2355 amount of bin migration in the two directions is different. The second possible
 2356 source of bin migration are systematic effects due, for example, to the presence
 2357 of bremsstrahlung photons that cannot be recovered. It is particularly important
 2358 to take into account the bin migration in the electron channels case because more
 2359 photons are radiated from the final state. Figure 14.1 reports the response matrix for
 2360 simulated $B^0 \rightarrow K^{*0} e^+ e^-$ events, which shows the correlation between reconstructed
 2361 and generated q^2 . In the ideal case of perfect resolution and no bias this plot would
 2362 look like a diagonal line and in the case of no bias its slope would be 1. Table 14.5

2363 reports net amounts of bin migration, M_{net} , in the considered q^2 intervals defined
2364 as:

$$M_{net} = N(\text{in} \rightarrow \text{in}) + N(\text{out} \rightarrow \text{in}) - N(\text{in} \rightarrow \text{out}) \quad (14.2)$$

2365 where $N(\text{in} \rightarrow \text{in})$ is the number of candidates that are generated and reconstructed
2366 inside the considered interval, $N(\text{out} \rightarrow \text{in})$ the number of candidates that are
2367 generated outside the interval but reconstructed inside and $N(\text{in} \rightarrow \text{out})$ the number
2368 of candidates generated inside that fall outside.

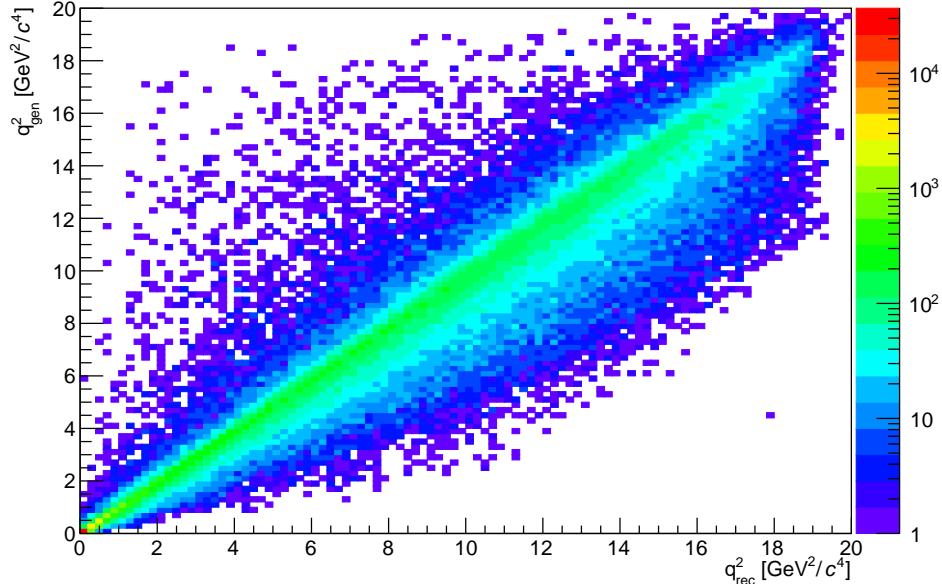


Figure 14.1: Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^* ee$ events.

Table 14.5: Net bin migration amounts (M_{net}) in the considered q^2 intervals. Positive values indicate “net in”, negative values “net out”.

Sample	1–6 GeV^2/c^4	15–20 GeV^2/c^4	J/ψ
$\mu\mu$	-0.0018 ± 0.0002	0.0042 ± 0.0003	-0.0012 ± 0.0000
ee	0.0834 ± 0.0013	-0.4469 ± 0.0091	-0.0258 ± 0.0003

2369 The reconstruction efficiency is calculated comparing generated to reconstructed
2370 samples and therefore already includes bin migration effects. Nevertheless, it is
2371 useful to single out this component to better assess the corresponding systematic
2372 uncertainties.

2373 14.3 PID efficiency

2374 The Monte Carlo is known not to reliably describe particle ID variables and therefore
 2375 a data-driven method is used to obtain this efficiency component. This is done
 2376 using the `PIDCalib` described in Sec. 2.8.1. Furthermore, the same method is used
 2377 to weight the MC in order to extract MVA and trigger efficiencies. The package
 2378 `PIDCalib` allows to divide the phase-space in bins and obtain a data-driven efficiency
 2379 for each bin. For this analysis the phase-space is divided in equi-populated bins
 2380 of momentum and pseudorapidity of the particle under study. Figure 14.2 shows
 2381 performance tables for pions, kaons, muons and electrons.

2382 The decay channel under study generally has different kinematical distributions than
 2383 the calibration sample. Therefore, once the efficiency table is obtained for each
 2384 particle, the total efficiency for each candidate is calculated as the product of the
 2385 four final particles efficiencies. $\varepsilon^{ev} = \varepsilon_K \cdot \varepsilon_\pi \cdot \varepsilon_{\ell_1} \cdot \varepsilon_{\ell_2}$. Finally, the total efficiency is
 2386 found by averaging over all simulated events.

$$\varepsilon_{PID} = \frac{1}{N} \sum_i^N \varepsilon_K(p_K^i, \eta_K^i) \cdot \varepsilon_\pi(p_\pi^i, \eta_\pi^i) \cdot \varepsilon_\ell(p_{\ell_1}^i, \eta_{\ell_1}^i) \cdot \varepsilon_K(p_{\ell_2}^i, \eta_{\ell_2}^i) \quad (14.3)$$

2387 14.4 Trigger efficiency

2388 While the muon channels the trigger efficiency is calculated using simulated events,
 2389 for the electron channels a combination of simulation and data-driven methods is
 2390 used. The efficiency of the software stage, Hlt, is always obtained from simulation,
 2391 while the efficiency of the hardware stage, L0, is obtained using a data-driven method
 2392 as described in the next section. For both muon and electronc channels it is possible
 2393 to use the resonant sample to crosscheck the efficiency obtained using the simulation
 2394 as explained in Sec. 14.4.2.

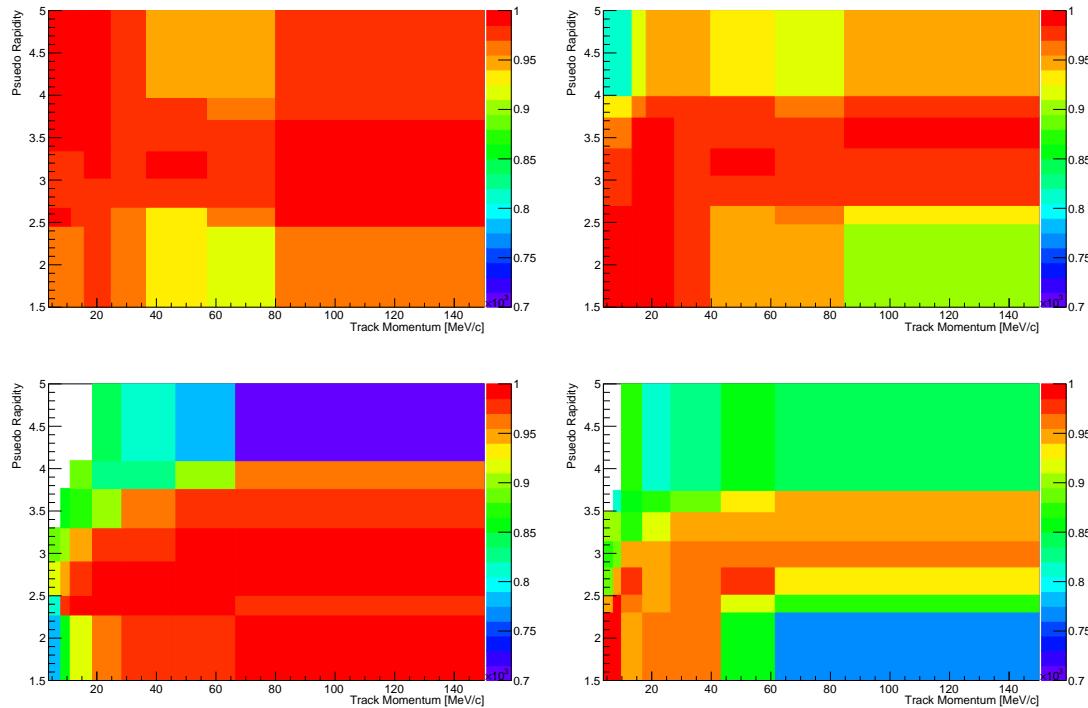


Figure 14.2: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

2395 14.4.1 Electron channels

2396 For the electron channels data is fitted separately in three trigger categories: L0Electron,
2397 L0Hadon and L0TIS. Therefore we need to extract the efficiency separately for each
2398 category.

2399 While the Hlt (1 and 2) efficiency is still computed using simulated events, the
2400 L0Electron and L0Hadron efficiencies cannot be modelled with the Monte Carlo.
2401 The discrepancy between data and simulation is mainly due to the ageing of the
2402 calorimeters, on which the decision of these triggers relies. The ageing is not simu-
2403 lated in the Monte Carlo and affects the L0 trigger efficiency which, therefore, must
2404 be calibrated using data driven-methods. Tables of efficiencies are obtained applying
2405 the TIS-TOS method to a calibration sample.

2406 For each trigger category these tables contain efficiency as a function of p_T of the
2407 considered particle and are given for different calorimeter regions as these have

²⁴⁰⁸ different properties (e.g. cell size) due to the different position with respect to the
²⁴⁰⁹ beam line. Regions considered are inner and outer HCAL, and inner, middle and
²⁴¹⁰ outer ECAL. Figure 14.3 shows data-driven efficiencies for the L0Electron trigger in
the three ECAL regions.

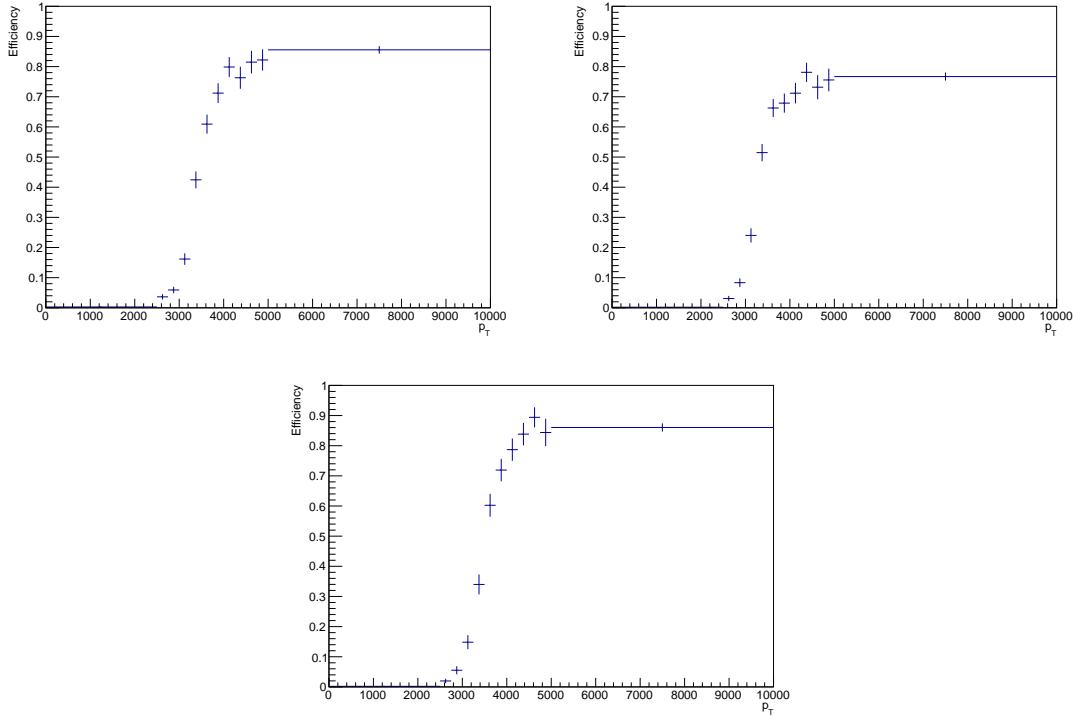


Figure 14.3: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

²⁴¹¹

²⁴¹² The probability of L0Electron trigger is calculated for each event as $P_{L0E} = \varepsilon(e^+) +$
²⁴¹³ $\varepsilon(e^-) - \varepsilon(e^+)\varepsilon(e^-)$. Similarly, the L0Hadron trigger probability is $P_{L0H} = \varepsilon(\pi) +$
²⁴¹⁴ $\varepsilon(K) - \varepsilon(\pi)\varepsilon(K)$. For L0TIS a probability $P_{L0I} = (1 - P_{L0E}) \cdot (1 - P_{L0H})$ is calculated.
²⁴¹⁵ Note that the probability of TIS trigger is defined to be independent of the signal
²⁴¹⁶ and therefore must be the same in the rare and resonant channels and cancel in
²⁴¹⁷ their ratio.

²⁴¹⁸ Then event by event efficiencies for the three trigger categories are defined to be
²⁴¹⁹ exclusive in the following way:

- ²⁴²⁰ • 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 event triggered. Note that in this case ε^{L0I} does not correspond to the efficiency of L0TIS trigger.

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

$$\varepsilon^{\text{trg}} = \frac{1}{N} \sum_i^N \varepsilon^{\text{trg}}(p_T^i) \quad (14.4)$$

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

14.4.2 TISTOS cross-check

The efficiency obtained using the simulation is crosschecked on resonant data using the TISTOS method already described in Sec. 6.3. For this purpose the logical *and of* of `LOGlobal_TIS`, `Hlt1Phys_TIS` and `Hlt2Phys_TIS` are used as control sample (TIS). As data also contains non negligible amounts of background the `DecayTreeFitter` tool is used to constrain the J/ψ mass to its known value and a narrow interval around the peak, dominated by the signal, is selected. Finally, to remove residual background in the data sample the `sPlot` method is used. Results are shown in Tab. 14.6, where the efficiency obtained using the TISTOS method is compared between data and simulation. Efficiencies obtained on simulation and data are found to be in agreement for the muon channel while they show deviations in the electron channels. Especially for the L0I category, for which the procedure explained in Sec. ?? does not ensure a correct calibration, a significant discrepancy is found. The table also reports a correction factor obtained according to the formula

$$f = 1 + \frac{\varepsilon_{\text{date}}^{\text{TISTOS}} - \varepsilon_{\text{MC}}^{\text{TISTOS}}}{\varepsilon_{\text{MC}}^{\text{TISTOS}}} \quad (14.5)$$

which is used to correct the resonant yields as described in Sec. ??.

2444 No correction is applied for the extraction of the $R_{K^{*0}}$ ratio because these discrepan-
 2445 cies do not affect the ratio between the rare and resonant samples. To check this the
 2446 efficiency obtained using the TISTOS method is plot as a function of the kinematic
 2447 quantity relevant for each trigger category: the maximum p_T of the electrons for
 2448 L0E, the maximum p_T of p and π for L0H, and the maximum p_T of all the final
 2449 particles for L0I. Results are shown in Fig. 14.4. The rare and resonant simulated
 2450 samples are used to find the average efficiency in the two cases. This corresponds
 2451 to making an average of the efficiency weighted by the kinematical distributions in
 2452 the two samples. This procedure is applied using the efficiency obtained using the
 2453 TISTOS method and data and simulation and the two are compared. The ratios
 2454 between the relative efficiencies, $(\varepsilon_{\ell\ell}^{data}/\varepsilon_{J/\psi}^{data})/(\varepsilon_{\ell\ell}^{sim}/\varepsilon_{J/\psi}^{sim})$, are found to be ??? for
 2455 L0E, ??? for L0H and ??? for L0I, which are all compatible with one.

Table 14.6: Trigger efficiencies obtained using the TISTOS method on simulated and real $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell\ell)$ decays.

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

2456 14.5 Neural Networks efficiency

2457 The NN efficiency is again evaluated from fully weighted Monte Carlo samples. For
 2458 the electron channels it is obtained separately for each trigger category.

2459 In order to cross check that this efficiency component is extracted correctly one
 2460 can compare the efficiency obtained using $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ events and rare
 2461 $B^0 \rightarrow K^{*0}\ell^+\ell^-$ events in the same q^2 region selected for the resonant case. The ratio
 2462 between the two should be close to 1 with small deviations due the fact that the bin
 2463 width is finite and the events are distributed differently inside the bin. This ratio is
 2464 found to be 0.997 ± 0.004 for the $\mu\mu$ channels and on average 0.981 ± 0.005 for the

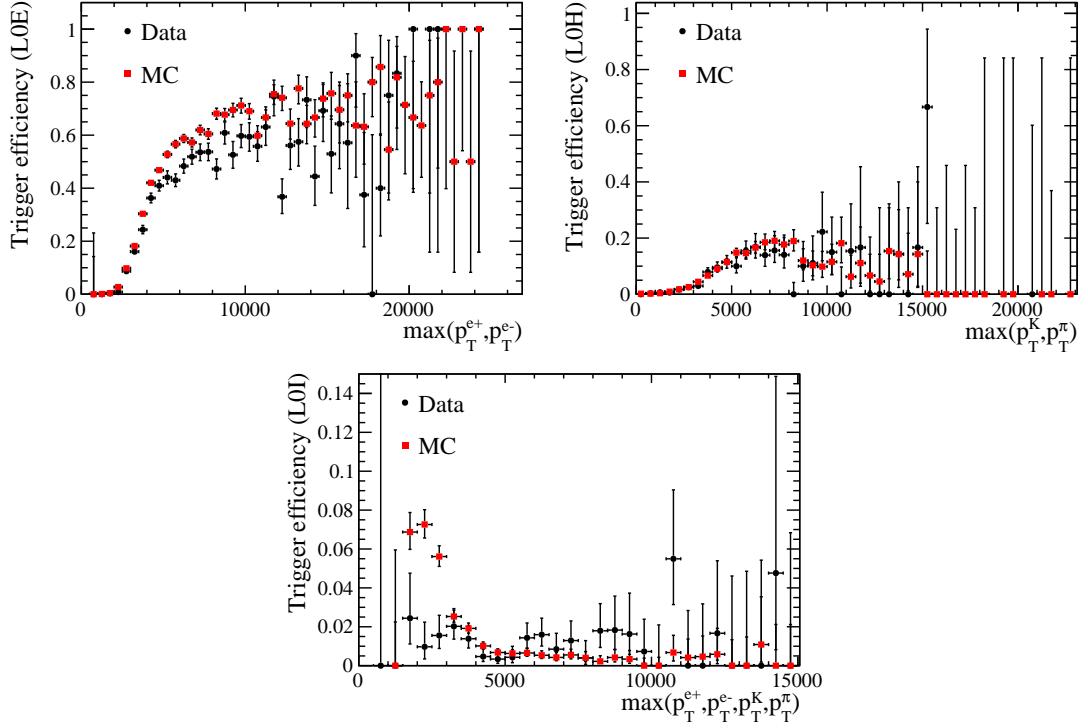


Figure 14.4: Trigger efficiency obtained applying the TISTOS method on $B^0 \rightarrow J/\psi K^*$ candidates as a function of the maximum p_T of the electrons for the L0E category (top left), the maximum p_T of p and π for L0H (top right) and the maximum p_T of all the final particles for L0I (bottom).

²⁴⁶⁵ ee channels. Values for the electron channels show a small deviation from one due
²⁴⁶⁶ to the very large q^2 interval used to select the resonant channel ($[6,11]$ GeV^2/c^4).

2467

CHAPTER 15

2468

2469

Systematic uncertainties

2470

CHAPTER 16

2471

2472

2473

Result extraction

2474

2475 In the following sections are presented the final results of this analysis together with
2476 the description of sanity checks performed to verify the stability of the methods
2477 used.

2478 16.1 $R_{J/\psi}$ sanity check

2479 In order to cross-check the analysis procedure, the ratio between the measured
2480 branching ratio of the electron and muonic resonant channels is calculated:

$$R_{J/\psi} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-))} = \frac{\varepsilon_{J/\psi(\mu\mu)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)}}{\varepsilon_{J/\psi(ee)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-)}}. \quad (16.1)$$

2481 One could also look at the single absolute branching fractions but this ratio is a
2482 better defined sanity check for our analysis. In fact the absolute branching fractions

²⁴⁸³ can be calculated from the raw yields as

$$\mathcal{B} = \mathcal{L} \cdot \sigma_{b\bar{b}} \cdot f_d \cdot \varepsilon \cdot N_{raw} \quad (16.2)$$

²⁴⁸⁴ where \mathcal{L} is the luminosity, $\sigma_{b\bar{b}}$ is the cross section for $b\bar{b}$ production and f_d is the
²⁴⁸⁵ fragmentation fraction, the probability for a b quark to produce a B^0 meson. All
²⁴⁸⁶ this quantities come with large uncertainties but they cancel in the $R_{J/\psi}$ ratio.

²⁴⁸⁷ Measured values of the $R_{J/\psi}$ ratio are reported in Tab. 16.1, where the error shown is
²⁴⁸⁸ statistical only. Notice that on these values systematic uncertainties, which cancel
²⁴⁸⁹ doing the ratio between the rare and resonant channels with same leptonic final
²⁴⁹⁰ state, do not cancel.

Table 16.1: Fully corrected measured values of the ratio $\mathcal{B}(B^0 \rightarrow K^* J/\psi \rightarrow \mu\mu)/\mathcal{B}(B^0 \rightarrow K^* J/\psi \rightarrow ee)$ in the three electron trigger categories.

$R_{J/\psi}$ (L0E)	1.01618 ± 0.00896
$R_{J/\psi}$ (L0H)	0.94276 ± 0.02315

²⁴⁹¹ 16.2 R_{K^*} result summary

²⁴⁹² The ratio R_{K^*} is extracted by dividing the R_{ee} and $R_{\mu\mu}$ parameters described in
²⁴⁹³ Sec. 13. These ratios are floating in the fit but they can also be built from the yields
²⁴⁹⁴ in Tab. 13.3 and the efficiencies in Tab. 14.2. In summary the definition of the R_{K^*}
²⁴⁹⁵ ratio is the following :

$$R_{K^*} = \frac{R_{ee}}{R_{\mu\mu}} = \frac{N_{ee}}{N_{J/\psi(ee)}} \cdot \frac{N_{J/\psi(\mu\mu)}}{N_{\mu\mu}} \cdot \frac{\varepsilon_{J/\psi(ee)}}{\varepsilon_{ee}} \cdot \frac{\varepsilon_{\mu\mu}}{\varepsilon_{J/\psi(\mu\mu)}}. \quad (16.3)$$

²⁴⁹⁶ As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three
²⁴⁹⁷ electron categories its value is already a combination of the three samples. Results
²⁴⁹⁸ are shown in Tab. 16.2.

Table 16.2: Measured values of R_{ee} , $R_{\mu\mu}$ and R_{K^*} ratios.

Ratio	1–6 GeV^2/c^4	15–20 GeV^2/c^4
R_{ee}	0.00303 ± 0.00042	0.00395 ± 0.00083
$R_{\mu\mu}$	0.00187 ± 0.00009	0.00182 ± 0.00008
R_{K^*}	0.61756 ± 0.08950	0.45980 ± 0.09919

2499 16.3 Branching ratios and expectations

2500 Multiplying the ratios R_{ee} and $R_{\mu\mu}$ by the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ [2]
 2501 branching ratios one can obtain absolute branching ratios for the rare channels:

$$\begin{aligned} \mathcal{B}(B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)) &= \mathcal{B}(B^0 \rightarrow K^*J/\psi) \times \mathcal{B}(B^0 \rightarrow \ell\ell) \\ &= (1.32 \pm 0.06)10^{-3} \times (5.96 \pm 0.03)10^{-2} = (7.87 \pm 0.36) \times 10^{-5} \end{aligned} \quad (16.4)$$

2502 Table 16.3 reports absolute branching ratio values for the rare channels in the considered q^2 intervals, where the errors are statistical only.

Table 16.3: Measured absolute branching ratio of the rare $\mu\mu$ and ee channels in the central and high q^2 regions. Errors shown are statistical only.

Channel	1–6 GeV^2/c^4	15–20 GeV^2/c^4
ee	$(2.37 \pm 0.34) \times 10^{-7}$	$(3.09 \pm 0.67) \times 10^{-7}$
$\mu\mu$	$(1.47 \pm 0.10) \times 10^{-7}$	$(1.42 \pm 0.09) \times 10^{-7}$

2503

2504 The results for the central q^2 bin can be compared also with SM predictions obtained
 2505 from Ref. [98]. This paper reports branching predicted ratios in the $1 < q^2 < 6$
 2506 GeV^2/c^4 interval for the rare electronic and muonic channels. These are rescaled
 2507 to the range $1.1 < q^2 < 6 \text{ GeV}^2/c^4$ using the simulation. Finally, the measured
 2508 value of the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ decay is used. The predicted ratio
 2509 is found to be 0.75 ± 0.14 , which is in agreement with our measurement within one
 2510 standard deviation. In Tab. 16.4 are also reported observed and expected ratios of
 2511 rare over resonant raw numbers of candidates ($N_{\ell\ell}/N_{J/\psi}$). In this table the observed
 2512 ratios are simply obtained dividing the rare and resonant yields in Tab. 13.3 and
 2513 the expected ones are obtained using the predicted rare channel branching ratios
 2514 and the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ branching ratio, rescaled by the relative

²⁵¹⁵ efficiencies in Tab. 14.4.

Table 16.4: Expected and observed ratios of raw event yields, $N_{\ell\ell}/N_{J/\psi}$.

Sample	Expected	Observed	Obs / exp ratio
$\mu\mu$	0.0026 ± 0.0003	0.0019 ± 0.0001	0.7265 ± 0.1012
ee (L0E)	0.0027 ± 0.0004	0.0027 ± 0.0004	1.0220 ± 0.1939
ee (L0H)	0.0072 ± 0.0009	0.0073 ± 0.0010	1.0220 ± 0.1970
ee (L0I)	0.0037 ± 0.0005	0.0038 ± 0.0005	1.0220 ± 0.1951

2516

CHAPTER 17

2517

2518

Conclusions

2519

2520

2521

REFERENCES

2522

- 2523 [1] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and*
 2524 *angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays*, *JHEP* **1506** (2015) 115,
 2525 [[arXiv:1503.07138](#)].
- 2526 [2] **Particle Data Group** Collaboration, K. Olive et al., *Review of Particle*
 2527 *Physics, Chin.Phys.* **C38** (2014) 090001.
- 2528 [3] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the*
 2529 *Weinberg-Salam Theory*, *Phys.Rev.* **D20** (1979) 2619–2625.
- 2530 [4] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961)
 2531 579–588.
- 2532 [5] F. Strocchi, *Spontaneous Symmetry Breaking in Local Gauge Quantum Field*
 2533 *Theory: The Higgs Mechanism*, *Commun.Math.Phys.* **56** (1977) 57.
- 2534 [6] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental*
 2535 *Test of Parity Conservation in Beta Decay*, *Phys.Rev.* **105** (1957) 1413–1414.
- 2536 [7] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, et al.,
 2537 *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$*
 2538 *New Physics*, *Phys.Rev.* **D91** (2015), no. 7 073007, [[arXiv:1501.05013](#)].
- 2539 [8] F. Zwicky, *Spectral displacement of extra galactic nebulae*, *Helv.Phys.Acta* **6**
 2540 (1933) 110–127.
- 2541 [9] M. Gavela and Hernandez, *Standard model CP violation and baryon*
 2542 *asymmetry*, *Mod.Phys.Lett.* **A9** (1994) 795–810, [[hep-ph/9312215](#)].
- 2543 [10] M. Maltoni, *Status of three-neutrino oscillations*, *PoS EPS-HEP2011* (2011)
 2544 090.

- 2545 [11] J. L. Feng, *Naturalness and the Status of Supersymmetry*,
 2546 *Ann.Rev.Nucl.Part.Sci.* **63** (2013) 351–382, [[arXiv:1302.6587](#)].
- 2547 [12] **ATLAS** Collaboration, G. Aad et al., *Observation of a new particle in the*
 2548 *search for the Standard Model Higgs boson with the ATLAS detector at the*
 2549 *LHC*, *Phys. Lett.* **B716** (2012) 1–29, [[arXiv:1207.7214](#)].
- 2550 [13] P. Fayet and S. Ferrara, *Supersymmetry*, *Phys.Rept.* **32** (1977) 249–334.
- 2551 [14] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra*
 2552 *dimension*, *Phys.Rev.Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- 2553 [15] G. Isidori and D. M. Straub, *Minimal Flavour Violation and Beyond*,
 2554 *Eur.Phys.J.* **C72** (2012) 2103, [[arXiv:1202.0464](#)].
- 2555 [16] A. J. Buras, *Minimal flavor violation*, *Acta Phys.Polon.* **B34** (2003)
 2556 5615–5668, [[hep-ph/0310208](#)].
- 2557 [17] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*,
 2558 *Ann.Rev.Nucl.Part.Sci.* **65** (2015) 8007, [[arXiv:1501.03309](#)].
- 2559 [18] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, *Can we reach*
 2560 *the Zeptouniverse with rare K and B_{s,d} decays?*, *JHEP* **1411** (2014) 121,
 2561 [[arXiv:1408.0728](#)].
- 2562 [19] G. Hiller and M. Schmaltz, *R_K and future b → sℓℓ physics beyond the standard*
 2563 *model opportunities*, *Phys.Rev.* **D90** (2014) 054014, [[arXiv:1408.1627](#)].
- 2564 [20] K. G. Chetyrkin, M. Misiak, and M. Munz, *Weak radiative B meson decay*
 2565 *beyond leading logarithms*, *Phys.Lett.* **B400** (1997) 206–219,
 2566 [[hep-ph/9612313](#)].
- 2567 [21] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond*
 2568 *leading logarithms*, *Rev.Mod.Phys.* **68** (1996) 1125–1144, [[hep-ph/9512380](#)].
- 2569 [22] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*,
 2570 [[hep-ph/9806471](#)].
- 2571 [23] M. Della Morte, J. Heitger, H. Simma, and R. Sommer, *Non-perturbative*
 2572 *Heavy Quark Effective Theory: An application to semi-leptonic B-decays*,
 2573 *Nucl.Part.Phys.Proc.* **261–262** (2015) 368–377, [[arXiv:1501.03328](#)].
- 2574 [24] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An Effective field*
 2575 *theory for collinear and soft gluons: Heavy to light decays*, *Phys.Rev.* **D63**
 2576 (2001) 114020, [[hep-ph/0011336](#)].
- 2577 [25] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop*
 2578 *effect in B → K^(*)ℓ⁺ℓ[−] and B → K^{*}γ*, *JHEP* **1009** (2010) 089,
 2579 [[arXiv:1006.4945](#)].

- [26] **LHCb** Collaboration, R. Aaij et al., *Observation of a resonance in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays at low recoil*, *Phys. Rev. Lett.* **111** (2013) 112003, [[arXiv:1307.7595](#)].
- [27] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, et al., *$B_{s,d} \rightarrow l + l -$ in the Standard Model with Reduced Theoretical Uncertainty*, *Phys. Rev. Lett.* **112** (2014) 101801, [[arXiv:1311.0903](#)].
- [28] **CMS, LHCb** Collaboration, V. Khachatryan et al., *Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data*, *Nature* **522** (2015) 68–72, [[arXiv:1411.4413](#)].
- [29] **LHCb** Collaboration, R. Aaij et al., *Differential branching fractions and isospin asymmetry of $B \rightarrow K^{(*)} \mu^+ \mu^+$ decays*, *JHEP* **06** (2014) 133, [[arXiv:1403.8044](#)].
- [30] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , *JHEP* **07** (2013) 084, [[arXiv:1305.2168](#)].
- [31] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , *JHEP* **08** (2013) 131, [[arXiv:1304.6325](#)].
- [32] **LHCb** Collaboration, R. Aaij et al., *Measurement of form-factor-independent observables in the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , *Phys. Rev. Lett.* **111** (2013) 191801, [[arXiv:1308.1707](#)].
- [33] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \rightarrow K^* \mu^+ \mu^-$ Anomaly*, *Phys. Rev.* **D88** (2013), no. 7 074002, [[arXiv:1307.5683](#)].
- [34] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of charged and neutral $B \rightarrow K \mu^+ \mu^-$ decays*, *JHEP* **05** (2014) 082, [[arXiv:1403.8045](#)].
- [35] **LHCb** Collaboration, R. Aaij et al., *Measurement of CP asymmetries in the decays $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^+ \rightarrow K^+ \mu^+ \mu^-$* , *JHEP* **1409** (2014) 177, [[arXiv:1408.0978](#)].
- [36] **LHCb** Collaboration, R. Aaij et al., *Measurement of the $B^0 \rightarrow K^{*0} e^+ e^-$ branching fraction at low dilepton mass*, *JHEP* **05** (2013) 159, [[arXiv:1304.3035](#)].
- [37] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay in the low- q^2 region*, *JHEP* **04** (Jan, 2015) 064. 18 p.
- [38] **MEGA** Collaboration, M. Ahmed et al., *Search for the lepton family number nonconserving decay $\mu u^+ \rightarrow e^+ \gamma$* , *Phys. Rev.* **D65** (2002) 112002, [[hep-ex/0111030](#)].
- [39] **SINDRUM** Collaboration, U. Bellgardt et al., *Search for the Decay $\mu u^+ \rightarrow e^+ e^+ e^-$* , *Nucl. Phys.* **B299** (1988) 1.

- [40] **LHCb** Collaboration, R. Aaij et al., *Search for the lepton-flavour-violating decays $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow e^\pm \mu^\mp$* , *Phys. Rev. Lett.* **111** (2013) 141801, [[arXiv:1307.4889](#)].
- [41] **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](#)].
- [42] W. J. Marciano, T. Mori, and J. M. Roney, *Charged Lepton Flavor Violation Experiments*, *Ann.Rev.Nucl.Part.Sci.* **58** (2008) 315–341.
- [43] L. Evans, *The LHC machine*, *PoS EPS-HEP2009* (2009) 004.
- [44] **LHCb** Collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.
- [45] **LHCb** Collaboration, R. Aaij et al., *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, *Phys.Lett.* **B694** (2010) 209–216, [[arXiv:1009.2731](#)].
- [46] M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J.* **C73** (2013) 2431, [[arXiv:1211.6759](#)].
- [47] A. A. Alves Jr. et al., *Performance of the LHCb muon system*, *JINST* **8** (2013) P02022, [[arXiv:1211.1346](#)].
- [48] **LHCb** Collaboration, R. e. a. Aaij, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030.
- [49] **LHCb** Collaboration, R. e. a. Aaij, *LHCb Detector Performance*, *Int. J. Mod. Phys. A* **30** (Dec, 2014) 1530022. 82 p.
- [50] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl.Instrum.Meth.* **A555** (2005) 356–369, [[physics/0402083](#)].
- [51] R. Aaij et al., *The LHCb trigger and its performance in 2011*, *JINST* **8** (2013) P04022, [[arXiv:1211.3055](#)].
- [52] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].
- [53] 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](#)].
- [54] 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.
- [55] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

- [56] 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](#)].
- [57] **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.
- [58] M. Clemencic et al., *The LHCb simulation application, GAUSS: design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023.
- [59] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, *Conf.Proc. C000917* (2000) 11–42.
- [60] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, *Nucl.Instrum.Meth.* **A559** (2006) 190–194.
- [61] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, [physics/0402093](#).
- [62] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, *Nucl.Instrum.Meth.* **A552** (2005) 566–575, [[physics/0503191](#)].
- [63] 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](#)].
- [64] T. Mannel and S. Recksiegel, *Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \rightarrow \Lambda \gamma$* , *J.Phys.* **G24** (1998) 979–990, [[hep-ph/9701399](#)].
- [65] M. J. Aslam, Y.-M. Wang, and C.-D. Lu, *Exclusive semileptonic decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in supersymmetric theories*, *Phys.Rev.* **D78** (2008) 114032, [[arXiv:0808.2113](#)].
- [66] Y.-m. Wang, Y. Li, and C.-D. Lu, *Rare Decays of $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in the Light-cone Sum Rules*, *Eur.Phys.J.* **C59** (2009) 861–882, [[arXiv:0804.0648](#)].
- [67] C.-S. Huang and H.-G. Yan, *Exclusive rare decays of heavy baryons to light baryons: $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Rev.* **D59** (1999) 114022, [[hep-ph/9811303](#)].
- [68] C.-H. Chen and C. Geng, *Rare $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decays with polarized lambda*, *Phys.Rev.* **D63** (2001) 114024, [[hep-ph/0101171](#)].
- [69] C.-H. Chen and C. Geng, *Baryonic rare decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Rev.* **D64** (2001) 074001, [[hep-ph/0106193](#)].
- [70] C.-H. Chen and C. Geng, *Lepton asymmetries in heavy baryon decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Lett.* **B516** (2001) 327–336, [[hep-ph/0101201](#)].

- [71] F. Zolfagharpour and V. Bashiry, *Double Lepton Polarization in $\Lambda_b^0 \rightarrow \Lambda l^+ l^-$ Decay in the Standard Model with Fourth Generations Scenario*, *Nucl.Phys.* **B796** (2008) 294–319, [[arXiv:0707.4337](#)].
- [72] 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](#)].
- [73] T. Aliev, K. Azizi, and M. Savci, *Analysis of the $\Lambda_b^0 \rightarrow \Lambda l^+ l^-$ decay in QCD*, *Phys.Rev. D81* (2010) 056006, [[arXiv:1001.0227](#)].
- [74] R. Mohanta and A. Giri, *Fourth generation effect on Λ_b decays*, *Phys.Rev. D82* (2010) 094022, [[arXiv:1010.1152](#)].
- [75] S. Sahoo, C. Das, and L. Maharana, *Effect of both Z and Z' -mediated flavor-changing neutral currents on the baryonic rare decay $\Lambda_b^0 \rightarrow \Lambda l^+ l^-$* , *Int.J.Mod.Phys. A24* (2009) 6223–6235, [[arXiv:1112.4563](#)].
- [76] **CDF Collaboration** Collaboration, T. Aaltonen et al., *Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$* , *Phys.Rev.Lett. 107* (2011) 201802, [[arXiv:1107.3753](#)].
- [77] **CDF** Collaboration, S. Behari, *CDF results on $b \rightarrow s \mu \mu$ decays*, [arXiv:1301.2244](#).
- [78] **LHCb** Collaboration, R. Aaij et al., *Measurement of the differential branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$* , *Phys. Lett. B725* (2013) 25, [[arXiv:1306.2577](#)].
- [79] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, *Rare baryon decays $\Lambda_b \rightarrow \Lambda l^+ l^- (l = e, \mu, \tau)$ and $\Lambda_b \rightarrow \Lambda \gamma$: differential and total rates, lepton- and hadron-side forward-backward asymmetries*, *Phys.Rev. D87* (2013) 074031, [[arXiv:1301.3737](#)].
- [80] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys.Lett. B724* (2013) 27, [[arXiv:1302.5578](#)].
- [81] 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](#).
- [82] 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.
- [83] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, *$\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ form factors and differential branching fraction from lattice QCD*, *Phys. Rev. D87* (2013), no. 7 074502, [[arXiv:1212.4827](#)].
- [84] **LHCb** Collaboration, R. Aaij et al., *Precision measurement of the Λ_b^0 baryon lifetime*, *Phys.Rev.Lett. 111* (2013) 102003, [[arXiv:1307.2476](#)].

- [85] T. Blake, S. Coquereau, M. Chrzaszcz, S. Cunliffe, C. Parkinson, K. Petridis, and M. Tresch, *The $B_0 \rightarrow K_0^* \mu\mu$ selection using $3fb^{-1}$ of LHCb data*, Tech. Rep. LHCb-INT-2013-058, CERN-LHCb-INT-2013-058, CERN, Geneva, Nov, 2013.
- [86] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Lett. B* **724** (2013) 27, [[arXiv:1302.5578](#)].
- [87] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys.Rev. D* **57** (1998) 3873–3889, [[physics/9711021](#)].
- [88] T. M. Karbach, *Feldman-Cousins Confidence Levels - Toy MC Method*, [arXiv:1109.0714](#).
- [89] S. Meinel, *Flavor physics with Λ_b baryons*, *PoS LATTICE2013* (2014) 024, [[arXiv:1401.2685](#)].
- [90] G. Hiller and F. Kruger, *More model independent analysis of $b \rightarrow s$ processes*, *Phys.Rev. D* **69** (2004) 074020, [[hep-ph/0310219](#)].
- [91] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \rightarrow s\ell\ell$* , *JHEP* **1502** (2015) 055, [[arXiv:1411.4773](#)].
- [92] **BaBar** Collaboration, J. Lees et al., *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \rightarrow K^{(*)} l^+ l^-$* , *Phys.Rev. D* **86** (2012) 032012, [[arXiv:1204.3933](#)].
- [93] **BELLE** Collaboration, J.-T. Wei et al., *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \rightarrow K^{(*)} l^+ l^-$* , *Phys.Rev.Lett.* **103** (2009) 171801, [[arXiv:0904.0770](#)].
- [94] **LHCb** Collaboration, R. Aaij et al., *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, [arXiv:1406.6482](#). to appear in Phys. Rev. Lett.
- [95] “Lhcb loki twiki.” <https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters>. Accessed: 2015-09-30.
- [96] “Probnn presentation at ppts meeting.” <https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf>. Accessed: 2015-09-30.
- [97] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf C0303241* (2003) MOLT007, [[physics/0306116](#)].
- [98] A. Ali, E. Lunghi, C. Greub, and G. Hiller, *Improved model independent analysis of semileptonic and radiative rare B decays*, *Phys.Rev. D* **66** (2002) 034002, [[hep-ph/0112300](#)].

- ²⁷⁶⁴ [99] J. Hrivnac, R. Lednický, 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](#)].
²⁷⁶⁵
²⁷⁶⁶

APPENDIX A

2767

2768

2769

Decay models

2770

A.1 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution

2771 The q^2 and angular dependancies of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays are modelled based on
2772 Ref. [79], where the angular distribution for unpolarised Λ_b^0 production is defined as

$$\begin{aligned} W(\theta_\ell, \theta_B, \chi) &\propto \sum_{\lambda_1, \lambda_2, \lambda_j, \lambda'_j, J, J', m, m', \lambda_\Lambda, \lambda'_\Lambda, \lambda_p} h_{\lambda_1 \lambda_2}^m(J) h_{\lambda_1 \lambda_2}^{m'}(J') e^{i(\lambda_j - \lambda'_j)\chi} \\ &\times \delta_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} \delta_{J, J'} d_{\lambda_j, \lambda_1 - \lambda_2}^J(\theta_\ell) d_{\lambda'_j, \lambda_1 - \lambda_2}^{J'}(\theta_\ell) H_{\lambda_\Lambda \lambda_j}^m(J) H_{\lambda'_\Lambda \lambda'_j}^{m'\dagger}(J') \\ &\times d_{\lambda_\Lambda \lambda_p}^{1/2}(\theta_B) d_{\lambda'_\Lambda \lambda_p}^{1/2}(\theta_B) h_{\lambda_p 0}^B h_{\lambda_p 0}^{B\dagger}. \end{aligned} \quad (\text{A.1})$$

2773 In this formula θ_ℓ and θ_B correspond to the lepton and proton helicity angles, χ
2774 is angle between dimuon and Λ decay planes (for unpolarised production we are
2775 sensitive only to difference in azimuthal angles), $d_{i,j}^J$ are Wigner d-functions and h ,
2776 h^B and H are helicity amplitudes for virtual dimuon, Λ and Λ_b^0 decays. The sum
2777 runs over all possible helicities with the dimuon being allowed in spin 0 and 1 states
2778 (J and J'). The m and m' indices run over the vector and axial-vector current
2779 contributions.

2780 The production polarisation is introduced by removing $e^{i(\lambda_j - \lambda'_j)\chi}$ from the expression,
2781 swapping small Wigner d-functions $d_{i,j}^J$ to the corresponding capital ones $D_{i,j}^J$ which
2782 are related as

$$D_{i,j}^J(\theta, \phi) = d_{i,j}^J(\theta) e^{i\phi(i-j)} \quad (\text{A.2})$$

²⁷⁸⁴ and substitute spin density matrix for $\delta_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} \delta_{JJ'}$. The spin density matrix
²⁷⁸⁵ itself is given by

$$\rho_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} = \frac{1}{2} \begin{pmatrix} 1 + P_b \cos \theta & P_b \sin \theta \\ P_b \sin \theta & 1 - P_b \cos \theta \end{pmatrix}. \quad (\text{A.3})$$

²⁷⁸⁶ Those changes lead to the formula

$$\begin{aligned} W(\theta\ell, \theta_B, \chi) \propto & \sum_{\lambda_1, \lambda_2, \lambda_j, \lambda'_j, J, J', m, m', \lambda_\Lambda, \lambda'_\Lambda, \lambda_p} h_{\lambda_1 \lambda_2}^m(J) h_{\lambda_1 \lambda_2}^{m'}(J') \\ & \times \rho_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} D_{\lambda_j, \lambda_1 - \lambda_2}^J(\theta\ell, \phi_L) D_{\lambda'_j, \lambda_1 - \lambda_2}^{J'}(\theta\ell, \phi_L) H_{\lambda_\Lambda \lambda_j}^m(J) H_{\lambda'_\Lambda \lambda'_j}^{m'\dagger}(J') \\ & \times D_{\lambda_\Lambda \lambda_p}^{1/2}(\theta_B, \phi_B) D_{\lambda'_\Lambda \lambda_p}^{1/2}(\theta_B, \phi_B) h_{\lambda_p 0}^B h_{\lambda_p 0}^{B\dagger}. \end{aligned} \quad (\text{A.4})$$

²⁷⁸⁷ The lepton amplitudes come directly from Ref. [79], eq. 3. The Λ decay amplitudes
²⁷⁸⁸ are related to the Λ decay asymmetry parameter as

$$\alpha_\Lambda = \frac{|h_{\frac{1}{2}0}^B|^2 - |h_{-\frac{1}{2}0}^B|^2}{|h_{\frac{1}{2}0}^B|^2 + |h_{-\frac{1}{2}0}^B|^2}. \quad (\text{A.5})$$

²⁷⁸⁹ Finally, the Λ_b^0 decay amplitudes receive contributions from vector and axial-vector
²⁷⁹⁰ currents and can be written as

$$H_{\lambda_2, \lambda_j}^m = H_{\lambda_2, \lambda_j}^{Vm} - H_{\lambda_2, \lambda_j}^{Am}. \quad (\text{A.6})$$

Finally, the remaining amplitudes are expressed in terms of form factors (Ref. [79],
eq. C6) as

$$\begin{aligned} H_{\frac{1}{2}t}^{Vm} &= \sqrt{\frac{Q_+}{q^2}} \left(M_- F_1^{Vm} + \frac{q^2}{M_1} F_3^{Vm} \right), \\ H_{\frac{1}{2}1}^{Vm} &= \sqrt{2Q_-} \left(F_1^{Vm} + \frac{M_+}{M_1} F_2^{Vm} \right), \\ H_{\frac{1}{2}0}^{Vm} &= \sqrt{\frac{Q_-}{q^2}} \left(M_+ F_1^{Vm} + \frac{q^2}{M_1} F_2^{Vm} \right), \\ H_{\frac{1}{2}t}^{Am} &= \sqrt{\frac{Q_-}{q^2}} \left(M_+ F_1^{Am} - \frac{q^2}{M_1} F_3^{Am} \right), \\ H_{\frac{1}{2}1}^{Am} &= \sqrt{2Q_+} \left(F_1^{Am} - \frac{M_-}{M_1} F_2^{Am} \right), \\ H_{\frac{1}{2}0}^{Am} &= \sqrt{\frac{Q_+}{q^2}} \left(M_- F_1^{Am} - \frac{q^2}{M_1} F_2^{Am} \right), \end{aligned} \quad (\text{A.7})$$

²⁷⁹¹ where $M_\pm = M_1 \pm M_2$, $Q_\pm = M_\pm^2 - q^2$. The form factors F are expressed in

2792 terms of dimensionless quantities in eqs. C8 and C9 in Ref. [79]. In our actual
2793 implementation form factors calculated in the covariant quark model [79] are used
2794 and for the numerical values of the Wilson coefficients Ref. [79] 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. [83]. The form factors relations are found by comparing eqs. 66 and 68 in Ref. [79] to eq. 51 in Ref. [83]. 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

$$\begin{aligned} c_\gamma &= 1 - \frac{\alpha_s(\mu^2)}{\pi} \left[\frac{4}{3} + \ln\left(\frac{\mu}{m_b}\right) \right], \\ c_v &= \frac{2}{3} \frac{\alpha_s(\mu^2)}{\pi}, \\ c_\sigma &= 1 - \frac{\alpha_s(\mu^2)}{\pi} \left[\frac{4}{3} + \frac{5}{3} \ln\left(\frac{\mu}{m_b}\right) \right]. \end{aligned} \quad (\text{A.8})$$

2795 In the calculations $\mu = m_b$ is used. For the strong coupling constant, we start
2796 from the world average value at the Z mass, $\alpha_s(m_Z^2) = 0.1185 \pm 0.0006$ [2], and we
2797 translate it to the scale m_b^2 by

$$\alpha_s(\mu^2) = \frac{\alpha_s(m_Z^2)}{1 + \frac{\alpha_s(m_Z^2)}{12\pi} (33 - 2n_f) \ln\left(\frac{\mu^2}{m_Z^2}\right)}, \quad (\text{A.9})$$

2798 where $n_f = 5$. The LQCD form factors F_1^L and F_2^L can be then taken directly from
2799 Ref. [83] and plugged into the code implementing the calculation from Ref. [79].

i	f_{1i}	f_{2i}	F_i
0	$a_+a_+^* + a_-a_-^* + b_+b_+^* + b_-b_-^*$	1	1
1	$a_+a_+^* - a_-a_-^* + b_+b_+^* - b_-b_-^*$	P_b	$\cos \theta$
2	$a_+a_+^* - a_-a_-^* - b_+b_+^* + b_-b_-^*$	α_Λ	$\cos \theta_1$
3	$a_+a_+^* + a_-a_-^* - b_+b_+^* - b_-b_-^*$	$P_b\alpha_\Lambda$	$\cos \theta \cos \theta_1$
4	$-a_+a_+^* - a_-a_-^* + \frac{1}{2}b_+b_+^* + \frac{1}{2}b_-b_-^*$	1	$d_{00}^2(\theta_2)$
5	$-a_+a_+^* + a_-a_-^* + \frac{1}{2}b_+b_+^* - \frac{1}{2}b_-b_-^*$	P_b	$d_{00}^2(\theta_2) \cos \theta$
6	$-a_+a_+^* + a_-a_-^* - \frac{1}{2}b_+b_+^* + \frac{1}{2}b_-b_-^*$	α_Λ	$d_{00}^2(\theta_2) \cos \theta_1$
7	$-a_+a_+^* - a_-a_-^* - \frac{1}{2}b_+b_+^* - \frac{1}{2}b_-b_-^*$	$P_b\alpha_\Lambda$	$d_{00}^2(\theta_2) \cos \theta \cos \theta_1$
8	$-3Re(a_+a_-^*)$	$P_b\alpha_\Lambda$	$\sin \theta \sin \theta_1 \sin^2 \theta_2 \cos \phi_1$
9	$3Im(a_+a_-^*)$	$P_b\alpha_\Lambda$	$\sin \theta \sin \theta_1 \sin^2 \theta_2 \sin \phi_1$
10	$-\frac{3}{2}Re(b_-b_+^*)$	$P_b\alpha_\Lambda$	$\sin \theta \sin \theta_1 \sin^2 \theta_2 \cos(\phi_1 + 2\phi_2)$
11	$\frac{3}{2}Im(b_-b_+^*)$	$P_b\alpha_\Lambda$	$\sin \theta \sin \theta_1 \sin^2 \theta_2 \sin(\phi_1 + 2\phi_2)$
12	$-\frac{3}{\sqrt{2}}Re(b_-a_+^* + a_-b_+^*)$	$P_b\alpha_\Lambda$	$\sin \theta \cos \theta_1 \sin \theta_2 \cos \theta_2 \cos \phi_2$
13	$\frac{3}{\sqrt{2}}Im(b_-a_+^* + a_-b_+^*)$	$P_b\alpha_\Lambda$	$\sin \theta \cos \theta_1 \sin \theta_2 \cos \theta_2 \sin \phi_2$
14	$-\frac{3}{\sqrt{2}}Re(b_-a_-^* + a_+b_+^*)$	$P_b\alpha_\Lambda$	$\cos \theta \sin \theta_1 \sin \theta_2 \cos \theta_2 \cos(\phi_1 + \phi_2)$
15	$\frac{3}{\sqrt{2}}Im(b_-a_-^* + a_+b_+^*)$	$P_b\alpha_\Lambda$	$\cos \theta \sin \theta_1 \sin \theta_2 \cos \theta_2 \sin(\phi_1 + \phi_2)$
16	$\frac{3}{\sqrt{2}}Re(a_-b_+^* - b_-a_+^*)$	P_b	$\sin \theta \sin \theta_2 \cos \theta_2 \cos \phi_2$
17	$-\frac{3}{\sqrt{2}}Im(a_-b_+^* - b_-a_+^*)$	P_b	$\sin \theta \sin \theta_2 \cos \theta_2 \sin \phi_2$
18	$\frac{3}{\sqrt{2}}Re(b_-a_-^* - a_+b_+^*)$	α_Λ	$\sin \theta_1 \sin \theta_2 \cos \theta_2 \cos(\phi_1 + \phi_2)$
19	$-\frac{3}{\sqrt{2}}Im(b_-a_-^* - a_+b_+^*)$	α_Λ	$\sin \theta_1 \sin \theta_2 \cos \theta_2 \sin(\phi_1 + \phi_2)$

Table A.1: Different terms describing angular distributions of $\Lambda_B^0 \rightarrow J/\psi \Lambda$ decays by eq. A.10.

2800 **A.2 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution**

2801 The angular distribution of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay is modelled using Ref. [99]. The
2802 differential rate is written as

$$w(\Omega, \Omega_1, \Omega_2) = \frac{1}{(4\pi)} \sum_{i=0}^{3^{i=19}} f_{1i} f_{2i}(P_b, \alpha_\Lambda) F_i(\theta, \theta_1, \theta_2, \phi_1, \phi_2), \quad (\text{A.10})$$

2803 where f_{1i} , f_{2i} and F_i are listed in Tab. A.1. The expression uses four observables
2804 (angles) and depends on four complex amplitudes a_+ , a_- , b_+ , b_- and two real valued
2805 parameters for the production polarisation, P_b , and the Λ decay asymmetry, α_Λ . The
2806 angle θ is the angle of the Λ momentum in Λ_b^0 rest frame with respect to the vector
2807 $\vec{n} = \frac{\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}}{|\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}|}$, where \vec{p}_{inc} and $\vec{p}_{\Lambda_b^0}$ are the momenta of incident proton and Λ_b^0 in the
2808 center of mass system. The angles θ_1 and ϕ_1 are polar and azimuthal angle of the
2809 proton coming from the Λ decay in the Λ rest frame with axis defined as $z_1 \uparrow\uparrow \vec{p}_\Lambda$,
2810 $y_1 \uparrow\uparrow \vec{n} \times \vec{p}_\Lambda$. Finally, the angles θ_2 and ϕ_2 are the angles of the momenta of the
2811 muons in J/ψ rest frame with axes defined as $z_2 \uparrow\uparrow \vec{p}_{J/\psi}$, $y_2 \uparrow\uparrow \vec{n} \times \vec{p}_{J/\psi}$.

2812 The distribution depends on the Λ decay asymmetry parameter, α_Λ , the production
2813 polarisation P_b and four complex amplitudes. The α_Λ is measured to be 0.642 ± 0.013
2814 for Λ . The production polarisation P_b and magnitudes of a_+ , a_- , b_+ and b_- are
2815 measured in Ref. [86]. Phases are not measured therefore, as default all phases are
2816 set to zero and then they are randomly varied to calculate the systematic uncertainty.

2817

APPENDIX B

2818

2819

Data-simulation comparison

2820

This appendix reports a comparison between distributions in data and simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events. 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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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.

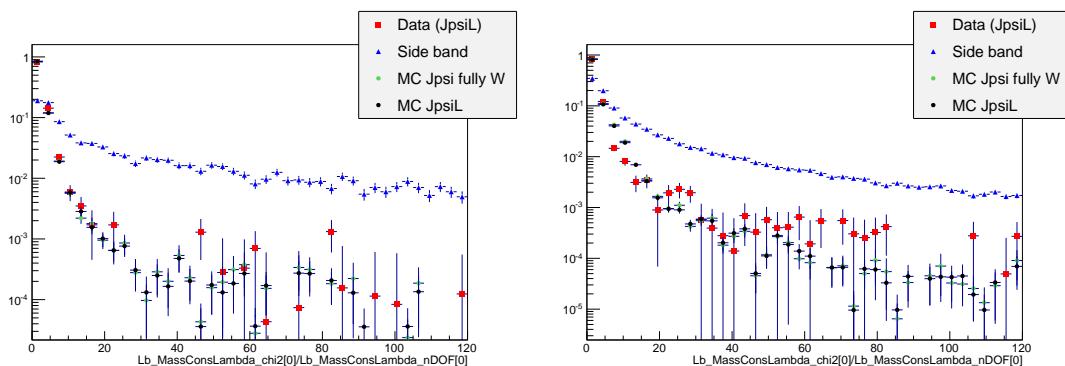


Figure B.1: Distributions of χ^2/NdF of the kinematic fit in data and simulation for LL (left) and DD (right) events.

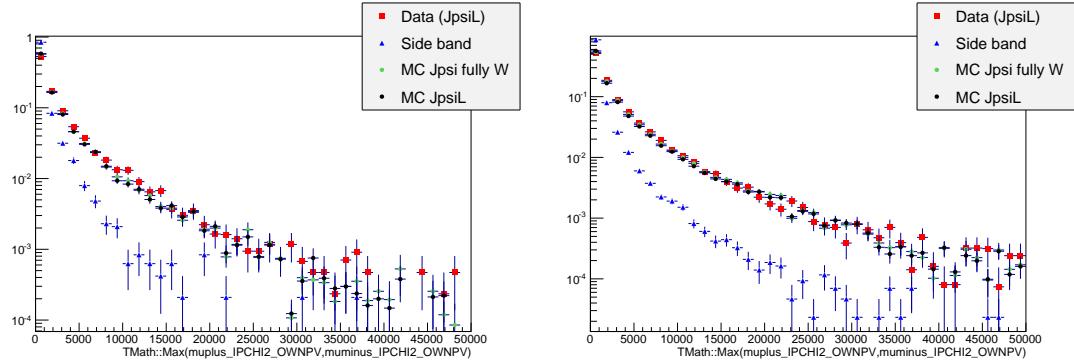


Figure B.2: Distributions of maximum muon $IP\chi^2$ variable in data and simulation for LL (left) and DD (right) events.

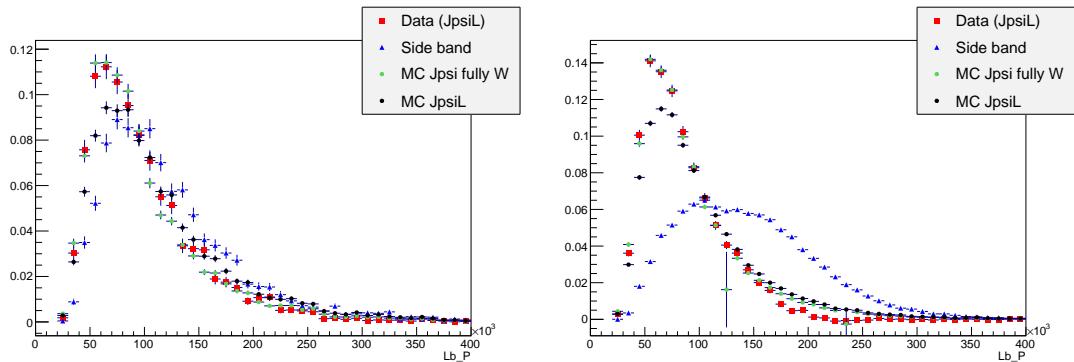


Figure B.3: Distributions of Λ_b^0 momentum variable in data and simulation for LL (left) and DD (right) events.

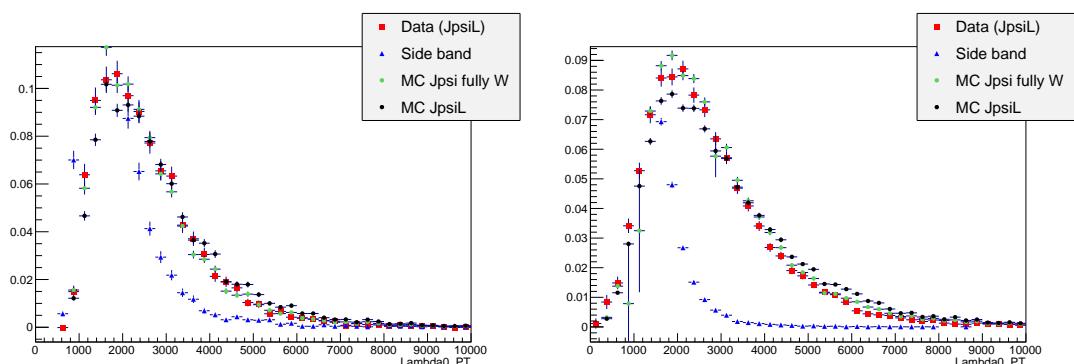


Figure B.4: Distributions of Λ transverse momentum variable in MC, data signal and data background for LL (left) and DD (right) events.

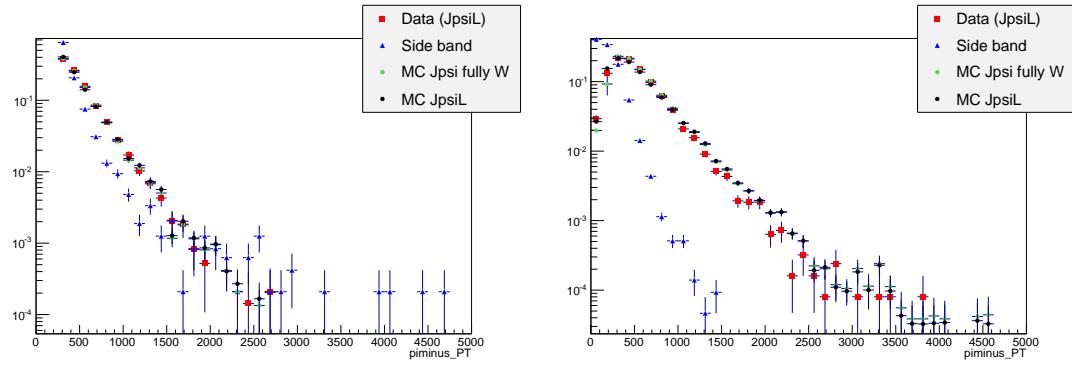


Figure B.5: Distributions of pion transverse momentum variable in data and simulation for LL (left) and DD (right) events.

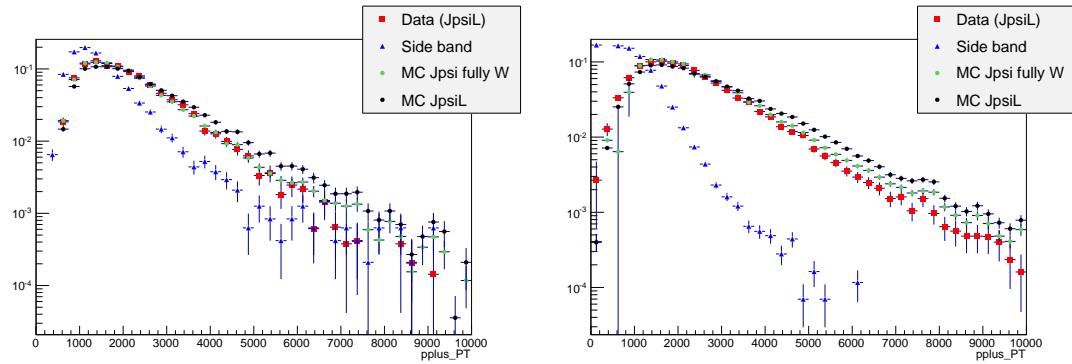


Figure B.6: Distributions of proton transverse momentum variable in data and simulation for LL (left) and DD (right) events.

2830

APPENDIX C

2831

2832 **Systematic uncertainties on the efficiency calculation for the**
 2833 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ **branching fraction analysis.**

2834

2835 This appendix reports systematic uncertainties on absolute and relative efficiencies
 2836 for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ 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.

q^2 [GeV $^2/c^4$]	Downstream			Long		
	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.

q^2 [GeV $^2/c^4$]	Downstream			Long		
	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.

q^2 [GeV $^2/c^4$]	Downstream			Long		
	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.

2837

APPENDIX D

2838

2839 **Invariant mass fits to $B^0 \rightarrow K^{*0}e^+e^-$ simulated candidates**

2840

2841 This appendix contains fits to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}e^+e^-$ sim-
2842 ulated candidates used to constrain parameters in the fit to data.

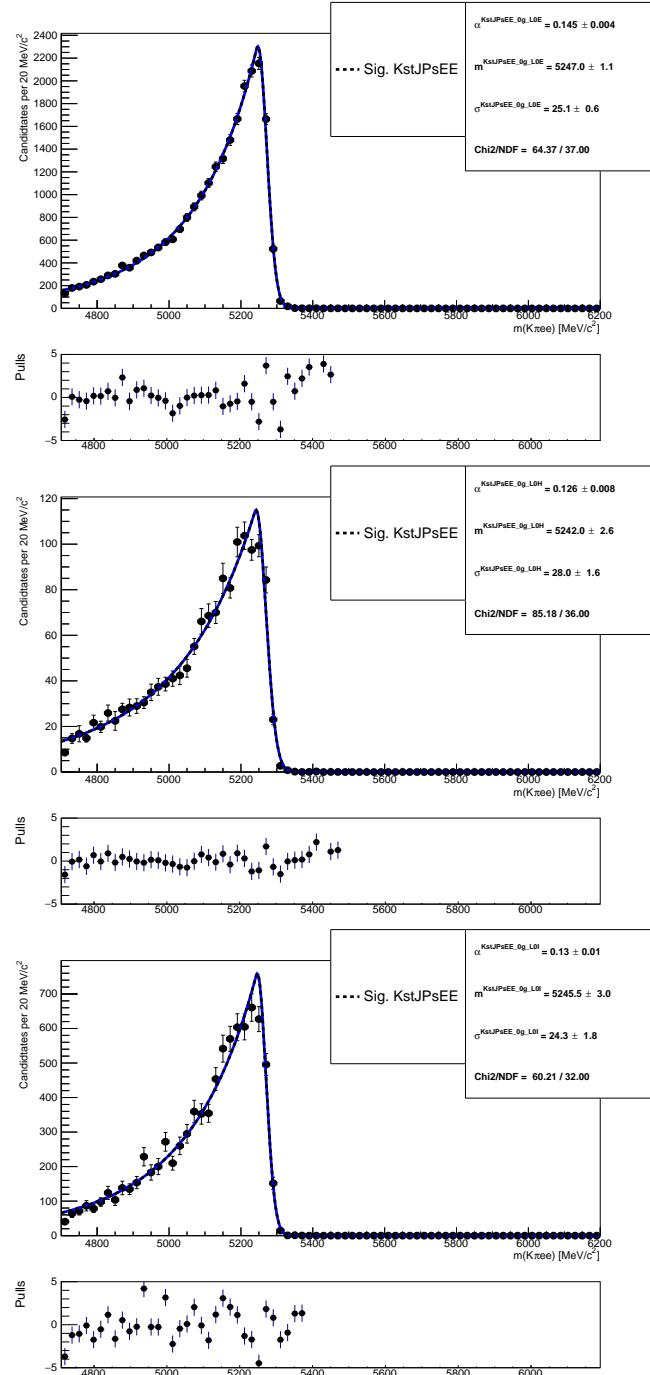


Figure D.1: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and no photon emitted.

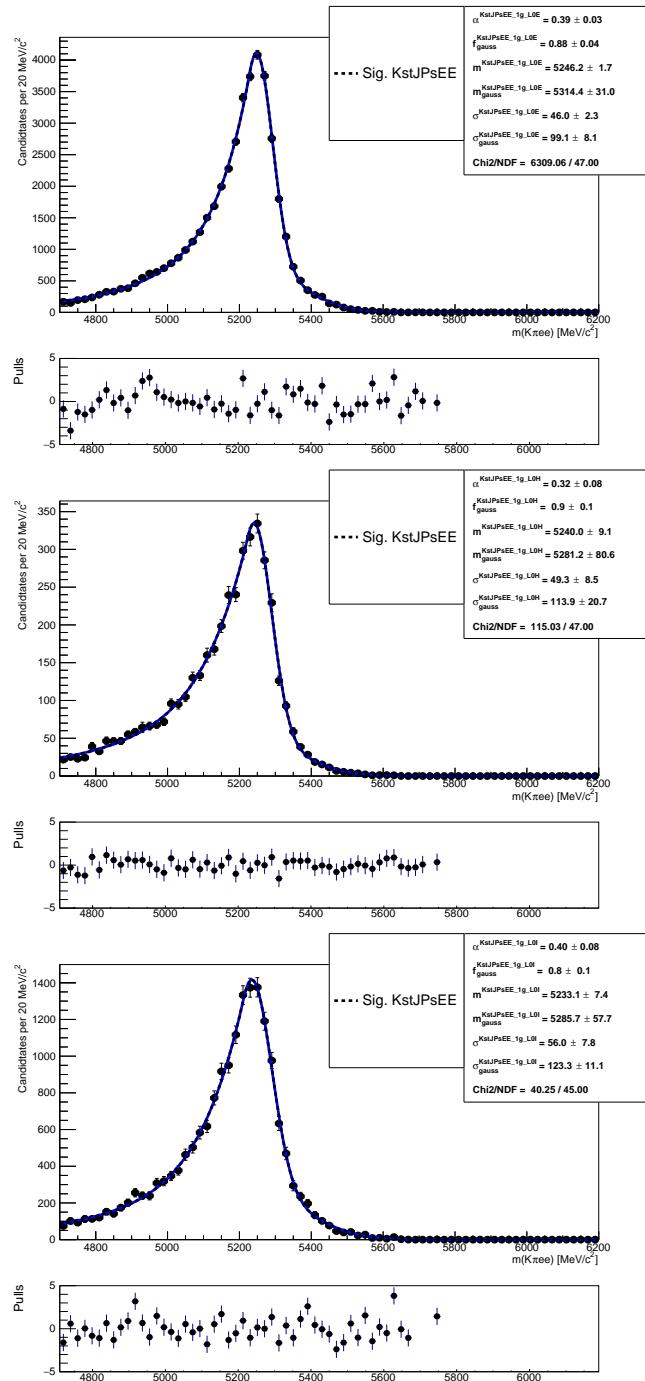


Figure D.2: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and one photon emitted.

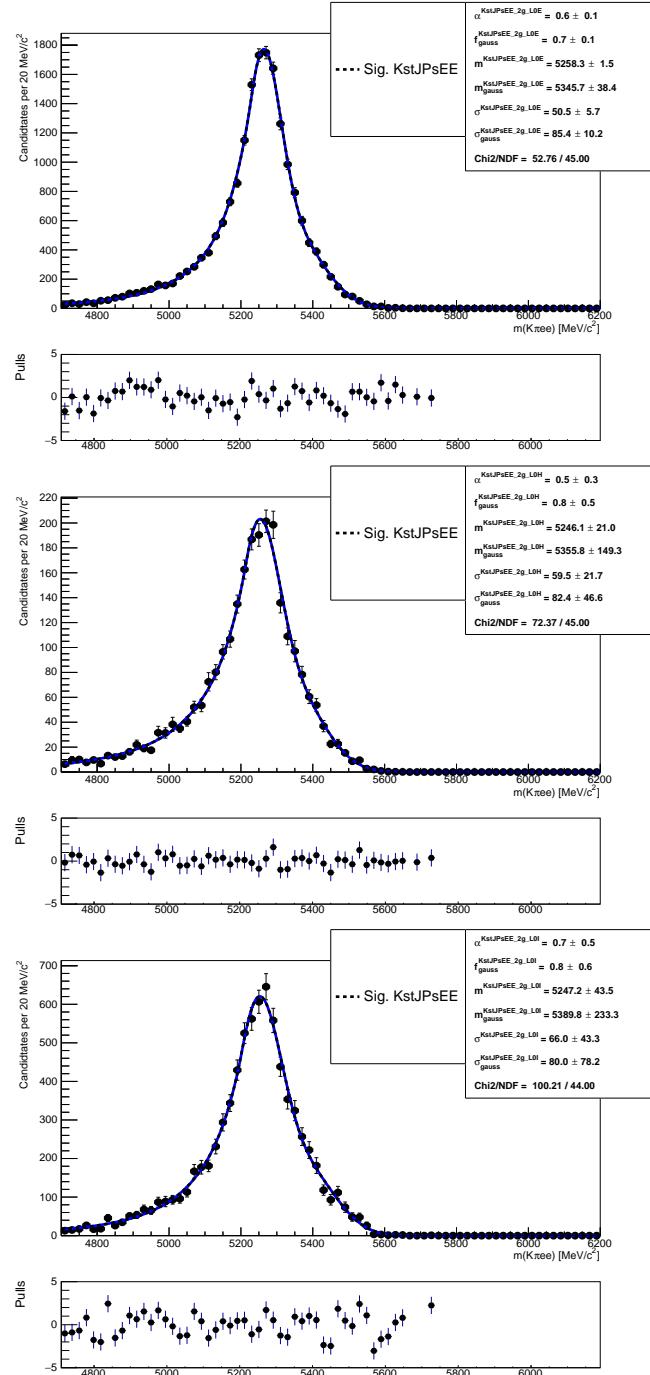


Figure D.3: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and two photons emitted.