

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.

June 26, 2015

ABSTRACT

Flavour Changing Neutral Currents (FCNCs) 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 (SM) but can happen through loop level 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 these the branching fractions of these decays, which are therefore very sensitive new physics. In this work two analysis of semileptonic $b \rightarrow s\ell\ell$ transitions are presented. First al all $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ 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 ?? 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 ?? 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. Each of these 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 though which the interesting decay channels are isolated. Secondly, after selection, the yield of interesting signal events is separated for the remaining background fitting the invariant mass distributions of the selected events. In the third step the efficiency of the detector is extracted 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 was one of the main authors of the two analysis reported in Parts I and II of this thesis. For the analysis of the differential branching ratio of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay I collaborated with Michal Kreps, who took care of implementing the decay model to re-weight the simulation and provided a few of the simulated samples. Furthermore, I want to thank him for the advice given throughout. The work in this part was also published and can be found at Ref. [1]. For the $R_{K^{*0}}$ analysis I actively participated in all stages of the analysis collaborating with Simone Bifani. Further more I'd like to thank Nigel Watson for the editorial advice and the general support provided for both the analysis. Furthermore I must thank you the LHCb collaboration though which I obtained the data used in this work and that collectively participated to the analysis reviewing the work. I contributed to the collaboration in three ways. First of all, for two years, I have been "Monte Carlo liaison", for the Rare Decays Working Group. This is a connection role between the physics analysis groups and the simulation team providing simulated samples vital for most analysis. In second place I have given my contribution to the LHCb simulation project, Gauss, by developing tools to validated the reliability of the simulation. Finally, I have taken several shifts in the control room during the Run II data taking in 2015, checking the smooth running of the detector.

ACKNOWLEDGEMENTS

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

v

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

Motto latino

Contents

1	Introduction	1
1.1	Electromagnetic and weak interactions	3
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	17
1.5.4	Observables in $b \rightarrow s\ell^+\ell^-$ decays	18
1.6	Experimental status	19
1.6.1	Dimuon decays of b hadrons	19
1.6.2	Semileptonic $b \rightarrow s\ell^+\ell^-$ decays of b hadrons	20
1.6.3	Lepton Flavour Violation searches	22
2	The LHCb detector at the Large Hadron Collider	24
2.1	The Large Hadron Collider	24
2.2	The LHCb detector	26
2.3	The magnet	28
2.4	Tracking system	28
2.5	Particle identification	31
2.5.1	Calorimeters	31
2.5.1.1	Bremsstrahlung recovery for electrons	33
2.5.2	RICH	34
2.6	The muon system	35
2.7	Trigger and software	36
I	Differential branching fraction and angular analysis of the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay	39
3	Introduction	40
3.1	Analysis strategy and q^2 regions	42
3.2	Candidate types	43
3.3	Simulation	44

3.3.1	Decay Model	45
3.3.2	Kinematic re-weighting	46
3.3.3	Event type	47
4	Selection	49
4.1	Neural Networks	51
4.2	MVA optimization	57
4.3	Trigger	58
4.4	Background from specific decays	59
5	Yield extraction	61
5.1	Fit description	61
5.2	Fit results	67
6	Efficiency	72
6.1	Geometric acceptance	73
6.2	Reconstruction and neural network efficiencies	74
6.3	Trigger efficiency	75
6.4	PID efficiency	76
6.5	Relative efficiencies	76
7	Systematic uncertainties	81
7.1	Yields	81
7.2	Efficiencies	83
7.2.1	Effect of new physics on the decay model	83
7.2.2	Simulation statistics	84
7.2.3	Production polarisation and decay structure	85
7.2.4	A_b^0 lifetime	85
7.2.5	Downstream candidates reconstruction efficiency	85
7.2.6	Data-simulation discrepancies	86
8	Differential branching ratio extraction	88
9	Angular analysis	93
9.1	One-dimensional angular distributions	94
9.2	Multi-dimensional angular distributions	96
9.3	Angular resolution	98
10	Angular fit strategy	101
10.1	Feldman-cousins plug-in method	102
10.2	Modelling the angular distributions	104
10.3	Angular acceptance	104
10.4	Studies on a three-dimensional fit	106
11	Systematics uncertainties on angular observables and results	110
11.1	Non-flat angular efficiency	110
11.1.1	Resolution	111

11.1.2 Angular efficiency	111
11.1.3 Background parameterisation	114
11.1.4 Polarisation	115
11.2 J/ψ cross-check	115
11.3 Results	117
II The $R_{K^{*0}}$ analysis	120
12 Testing lepton flavour universality with R_H	121
12.0.1 Combining ratios	124
12.0.2 Experimental status	125
12.1 Analysis strategy	126
12.2 Choice of q^2 intervals	127
12.3 Data samples and simulation	128
12.3.1 Data-simulation corrections	128
13 Selection	131
13.1 Trigger and Stripping	132
13.2 PID	135
13.3 Peaking backgrounds	137
13.3.1 Charmonium vetoes	137
13.3.2 ϕ veto	138
13.3.3 $B^+ \rightarrow K^+ \ell^+ \ell^-$ plus a random pion	138
13.3.4 Λ_b decays	139
13.3.4.1 Other peaking backgrounds	140
13.3.5 Mis-reconstructed background	140
13.4 Multivariate analysis	141
13.5 MVA optimisation	146
14 Mass fits	148
14.1 Muonic channels	149
14.2 Electron channels	150
14.2.1 Signal PDFs for the electron channels in the central q^2 interval	153
14.2.2 Background PDFs for the electron channels in the central q^2 interval	154
14.2.3 Summary of the fit to electron channels in the central q^2 interval	156
14.2.4 Electron channels fits in the high q^2 interval	158
14.3 Fit summary	161
15 Efficiency	163
15.1 Geometric efficiency	166
15.2 Reconstruction efficiency and bin migration	166
15.3 PID efficiency	168
15.4 Trigger efficiency	169
15.4.1 muon channels	169
15.4.2 Electron channels	170
15.5 Neural Networks efficiency	172

16 Systematic uncertainties	173
17 Result extraction	174
17.1 $R_{J/\psi}$ sanity check	174
17.2 R_{K^*} result summary	175
17.3 Branching ratios and expectations	176
A Decay models	185
A.1 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution	185
A.2 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution	187
B Monte Carlo and data comparison	190
C Systematic uncertainties on the efficiency calculation for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction analysis.	193
D Invariant mass fits to $B^0 \rightarrow K^{*0} e^+ e^-$ simulated candidates	196

¹

CHAPTER 1

²

³

Introduction

⁴

⁵ The Standard Model of Particle Physics (SM) is a quantum field theory (QFT) de-
⁶ scribing strong and electroweak (EW) interactions. It was formulated in its current
⁷ 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 almost 200 GeV are well
⁹ described by the SM and experiments at the Large Hadron Collider (LHC) are now
¹⁰ probing the SM up to the TeV scale. Finally, in 2013 we were able to observe the
¹¹ Higgs boson, one of the fundamental building blocks of the theory, giving a solid
¹² theoretical basis to it. However, experimentally well established effects, like neu-
¹³ trino oscillations and the presence of dark matter, are outside the reach of the SM.
¹⁴ Furthermore, the model does not include the description of gravity. Therefore this
¹⁵ motivates the search for New Physics (NP).

¹⁶ The SM is based on the symmetry groups of strong ($SU(3)_C$) and electroweak in-
¹⁷ teractions ($SU(2)_W \times U(1)_Y$). The subscripts C, W and Y stand for colour charge,
¹⁸ weak isospin and hyper-charge. The Lagrangian describing the SM results from

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

Table 1.1: 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.

¹⁹ the application of the principle of invariance under the unitary product group
²⁰ $SU(3) \times SU(2) \times U(1)$, which reflects conservation laws such as the conservation
²¹ of electric and strong charge. The model had then 26 free parameters which are
²² experimentally measured.

²³ Particles included in the SM can be grouped under a few categories depending
²⁴ on their properties and ability to interact with each other. First of all we can
²⁵ distinguish between fermions (half-integer spin particles) and bosons (integer spin
²⁶ particles). Fermions constitute the basic building blocks of matter, while bosons
²⁷ are the mediators of the interaction between them. Since in the SM the concept
²⁸ of bosonic mediators of interactions arises because of gauge symmetry [2], they are
²⁹ called “gauge bosons”. The list of the known interactions with their force carrier
³⁰ and properties is reported in Tab. 1.1. The matter of which we are made is mainly
³¹ composed of electrons and protons, which have spin 1/2; protons are then composed
³² of u and d quarks, which again have spin 1/2. Among fermions one can then consider
³³ two smaller groups: quarks and leptons. Quarks carry colour charge and therefore
³⁴ can interact through the, so called, strong interaction, while leptons, which do not
³⁵ carry colour charge, are insensitive to it. For each particle exists a corresponding
³⁶ anti-particle with opposite quantum numbers. Finally, fermions are divided into
³⁷ three families having similar properties but different masses. This last structure
³⁸ embedded in the SM is also called flavour structure and it will be the main tool
³⁹ used in this thesis, a more detailed description of it is given in the next sections. A
⁴⁰ schematic view of the fundamental particles in the SM is shown in Fig. 1.1. Due to
⁴¹ the asymptotic freedom of the strong interaction quarks cannot be observed alone

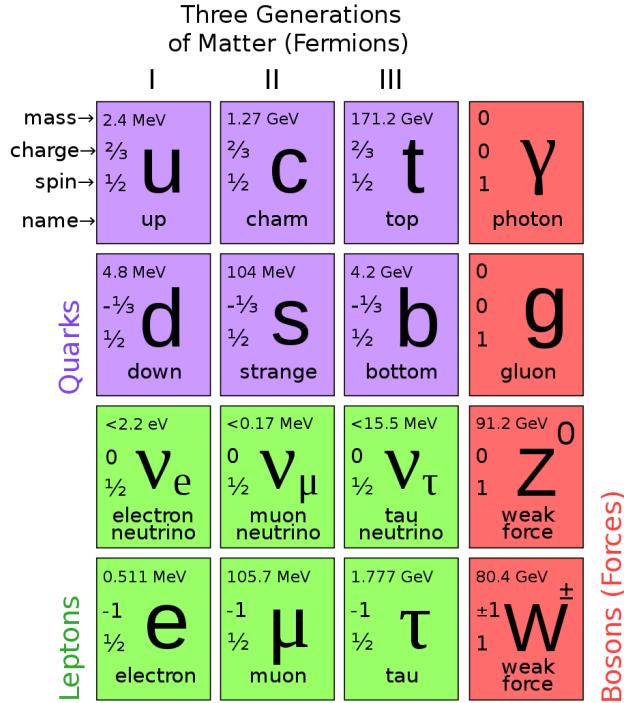


Figure 1.1: Diagram of SM particles with their properties [?].

⁴² but are always combined with other quarks to form color singlets. Non-fundamental
⁴³ particles composed by quarks are called hadrons and can be divided in mesons,
⁴⁴ where the color singlet is achieved by the combination of a quark and its antiquark
⁴⁵ ($q \bar{q}$), and baryons formed by three quarks ($q q q$).

⁴⁶ 1.1 Electromagnetic and weak interactions

⁴⁷ The Electromagnetic (EM) force is responsible for binding electrons and nuclei to-
⁴⁸ gether in atoms. Its force carrier, the photon, is the gauge boson of the EM force.
⁴⁹ In the SM the photon must be massless, which also sets the range of the EM force
⁵⁰ to infinity, since this is proportional to the inverse of the mediator mass. In fact
⁵¹ Heisenberg's Uncertainty Principle tells us that $\Delta E \Delta t > \hbar$, namely virtual particles
⁵² of energy ΔE are allowed to exist for time intervals inferior to Δt . Thus, since
⁵³ particles can move at most at the speed of light, $c = 299792458 \text{ ms}^{-1}$ [3], this also
⁵⁴ sets a relation between the length of time and space in which a virtual photons can

55 exist ($\Delta t > \hbar/(mc^2)$). As virtual photons can be very close to the mass shell, this
56 results in a very long lifetime. The EM force has therefore an infinite range.

57 The weak interaction is responsible for the β decay of nuclei. Unlike the electromagnetic
58 force that affects only charges particles, all known fermions interact through
59 the weak interaction. In the Standard Model this interaction is mediated by the
60 emission or absorption of W^\pm and Z bosons. The electroweak symmetry is sponta-
61 neously broken by the Higgs field [4] and this causes the W^\pm and Z bosons to become
62 massive (see Tab. ??). For this reason the weak force has a very short range. Using
63 Heisenberg’s Principle together with Einstein’s formula $\Delta E = mc^2$, which relates
64 mass and energy, and knowing that the maximum space that a particle can cover
65 in a time Δt is $r = c\Delta t$, qualitatively $r \sim \hbar/mc$. In this picture the carriers of the
66 weak force can travel $r \sim 2 \cdot 10^{-3}$ fm. The weak interaction is also the only one
67 that violates parity-symmetry, which states that interactions are invariant under a
68 reflection of all coordinates. This symmetry breaking arises from the fact that only
69 left-handed fermions interact through the weak interaction. The first experiment
70 showing this was made by Wu in 1957 [5]. Similarly, the weak interaction is the
71 only one that also breaks the CP symmetry, which combines parity transformations
72 and “charge conjugation”. This is particularly interesting because all interactions
73 are invariant under the CPT transformation, which combines the CP transforma-
74 tion and time reversal, hence, breaking CP the weak interaction must also be not
75 invariant under time reversal.

76 In 1968 Salam, Glasow and Weinberg unified the weak and electromagnetic forces
77 in a single theory, having a single coupling constant [3]. The EW interactions
78 are divided into charged currents (CC) and neutral currents (NC). In the first
79 group, quarks and leptons interact with the W^\pm bosons, producing decays such
80 as $\mu^+(\mu^-) \rightarrow e^+\nu_e\bar{\nu}_\mu(e^-\bar{\nu}_e\nu_\mu)$ and $n \rightarrow pe^-\bar{\nu}_e(\bar{p}e^+\nu_e)$. The study of these processes
81 confirmed that only the left-handed (right-handed) component of fermions (anti-
82 fermions) takes part in weak processes. The CC interactions have a peculiarity:
83 they are the only interactions in the SM that violate flavour conservation at tree
84 level (see next section), while any other interaction not conserving flavour has to

85 happen through loops. The second group of EW interactions, NC, corresponds di-
86 agrams mediated by the photon or the Z boson interacting with a fermion and its
87 anti-fermion.

88 1.2 Flavour and the CKM matrix

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

97 In the hadronic sector particles carry flavour numbers described as follow:

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

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

105 Measuring branching fractions of weak decays like $\pi \rightarrow \mu\nu_\mu$ and $K \rightarrow \mu\nu_\mu$, corre-
106 sponding respectively to $ud \rightarrow \mu\nu_\mu$ and $us \rightarrow \mu\nu_\mu$ processes, suggested the existence
107 of more than one coupling constant for different quarks. Cabibbo [3], in order to
108 preserve the universality of weak interactions, suggested that the branching fraction

¹⁰⁹ differences could arise from the fact that the doublets participating in the weak in-
¹¹⁰ teractions are an admixture of the flavour eigenstates. He therefore introduced the
¹¹¹ Cabibbo angle, θ_c , considering that eigenstates participating to the weak interaction
¹¹² 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 preserved the universality of the weak interaction. Theoretically, a $N \times N$ complex matrix is dependent 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. In Eq. 1.2 is reported a parametrisation of the CKM matrix together with the most recent measured values of its terms [3]. In this parametrisation ρ , A , and λ are the real constants and η the imaginary phase; in Eq. 1.5 are reported their relations with the 3 mixing angles.

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

$$\lambda = \sin(\theta_{12}) = \sin(\theta_c) \quad (1.3)$$

$$A\lambda^2 = \sin(\theta_{23}) \quad (1.4)$$

$$A\lambda^3(\rho - i\eta) = \sin(\theta_{13})e^{i\delta} \quad (1.5)$$

Figure 1.2 displays examples of CC processes together with the CKM elements

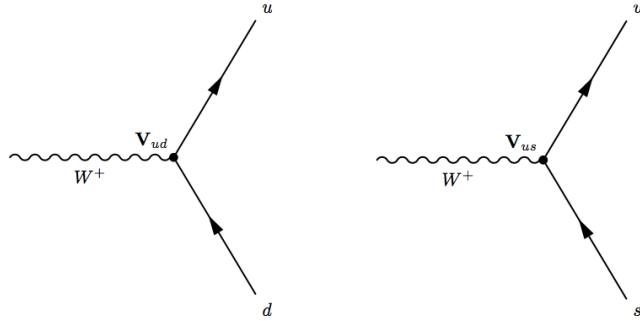


Figure 1.2: Feynman diagrams with CKM weights on weak interaction vertices

113

114 associated with their vertices. It is interesting to note that the CKM matrix has a
115 hierarchical form, namely elements on the diagonal are approximately 1 and become
116 smaller and smaller going farther from the diagonal. This structure is not explained
117 in the SM. Another feature to note is that, due to the unitarity of the matrix, the
118 transformation have no effect on neutral interactions. In fact defining $q' = Vq$

$$\bar{q}'q' = \bar{q}V^*Vq = \bar{q}q. \quad (1.6)$$

119 As a result flavour-changing neutral currents are forbidden at tree level in the SM.

120 As mentioned, the CKM matrix has to be unitary to preserve probability and this
121 imposes constraints to its terms of the form:

$$\sum_i |V_{ik}|^2 = 1 \text{ and } \sum_k V_{ik}V_{jk}^* = 0. \quad (1.7)$$

122 These correspond to a constraint to three complex numbers, which can be viewed
123 triangles in the (ρ, η) plane and are called “unitarity triangles”. The most commonly

¹²⁴ used unitarity triangle arises from

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (1.8)$$

¹²⁵ Figure 1.3 shows a representation of such triangle together with a plot summarising
¹²⁶ the most up to date experimental constraints to its parameters [6]. The precise
¹²⁷ measurement of the parameters of the CKM matrix is a powerful stability test of
¹²⁸ the standard model and sets a solid base for new physics searches in the flavour
¹²⁹ sector. One of the main goals of the LHCb experiment is to precisely measure the
angle γ , which is currently the least constrained from measurements.

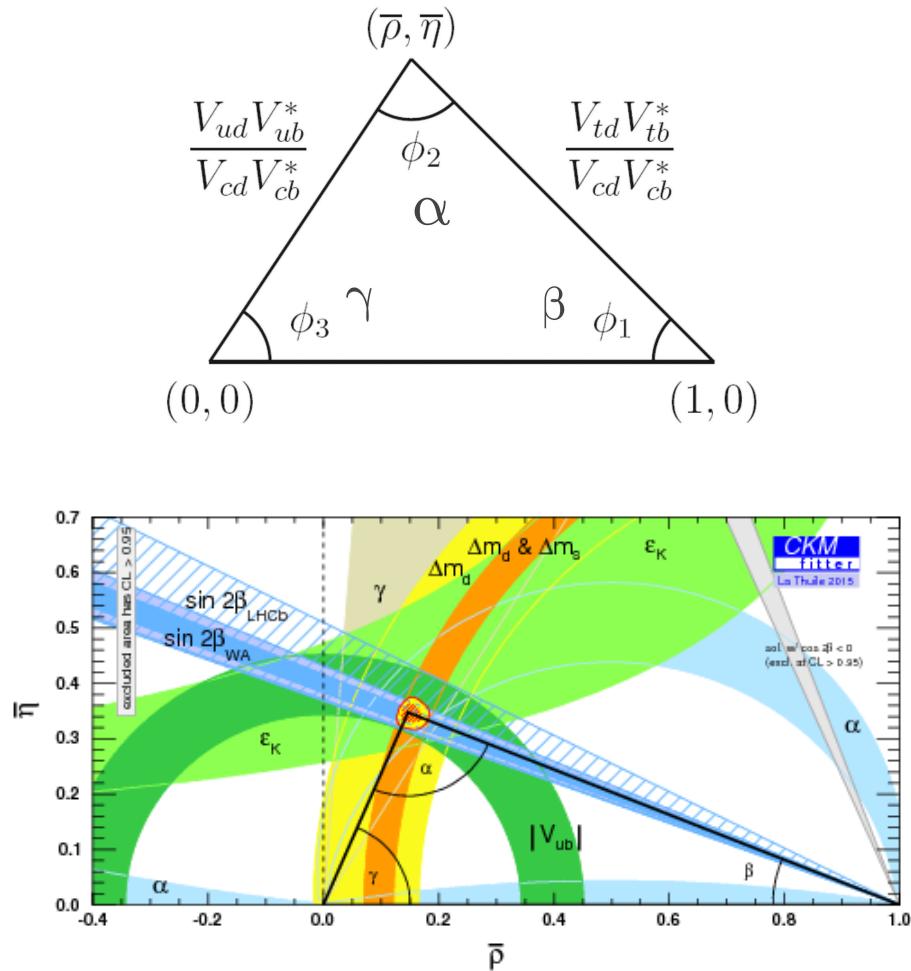


Figure 1.3: (top) A representation of the unitarity triangle and its parameters. (bottom) A summary of the most up to date measurements of the unitarity triangle parameters [6].

₁₃₁ 1.3 The puzzles of the SM

₁₃₂ Despite the confirmation of many predictions of the SM, the theory has several
₁₃₃ limitations and is unable to account for some well established experimental facts:

₁₃₄ • *Dark matter*: experimental evidence tells us that the content of visible matter
₁₃₅ in the universe is not enough to account for the observed rotation of galaxies [7].
₁₃₆ The more natural way to solve the problem is the hypothesis of a form of matter
₁₃₇ that interacts with the gravitational field but not with the interaction of the
₁₃₈ SM.

₁₃₉ • *Matter-antimatter asymmetry*: a large asymmetry is observed between the
₁₄₀ quantity of matter and antimatter in the universe. Assuming that both were
₁₄₁ equally created in the initial state of the universe, a condition such as the
₁₄₂ violation of the CP symmetry is necessary to account for such observed dif-
₁₄₃ ferences. However, the magnitude of CP violation predicted by the SM is not
₁₄₄ enough to explain the observed imbalance [8].

₁₄₅ • *Gravity*: even though the gravitational force was the first to be discovered this
₁₄₆ is not included in the SM. In fact when introducing gravity in the framework
₁₄₇ of QFT the theory diverges. On the other hand gravity becomes irrelevant for
₁₄₈ small masses as those of particles and can be neglected in good approximation.
₁₄₉ Many attempts were made but there is not yet a consistent procedure to
₁₅₀ introduce gravity in the SM.

₁₅₁ • *Neutrino oscillation*: measurements regarding solar and atmospheric neutrinos
₁₅₂ as wells as neutrinos from nuclear reactors established that neutrinos can
₁₅₃ change flavour while propagating in space. This is not predicted in the SM, in
₁₅₄ fact in the SM neutrinos are massless, while an oscillation requires a non zero
₁₅₅ mass [9].

₁₅₆ • *The hierarchy problem*: The mass of a scalar (spin 0) particle, such as the
₁₅₇ Higgs boson, suffers from quantum corrections due to the physics above a

158 certain scale. As new physics can appear anywhere up to the Planck scale,
159 $\sim 10^{19}$ GeV, these corrections can be very large and it would require a high
160 level of fine-tuning for them to cancel out and give such a small value as the
161 one measured for the Higgs Mass, ~ 126 GeV/ c^2 [10].

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

166 1.3.1 The flavour problem

167 Flavour Changing Charged Currents (FCCC) that are mediated by the W^\pm bosons
168 are the only sources of flavour changing interaction in the SM and, in particular, of
169 generation changing interactions, where a quark or a lepton of a family transforms
170 into one of another family. Another class of processes is the Flavour Changing
171 Neutral Currents (FCNCs), e.g. transitions from a b quark with charge of 1/3 to
172 a s or d with a charge of +2/3. Examples of FCNC transitions in the quark and
lepton sector are shown in Fig. 1.4. In the SM there is no fundamental reason why

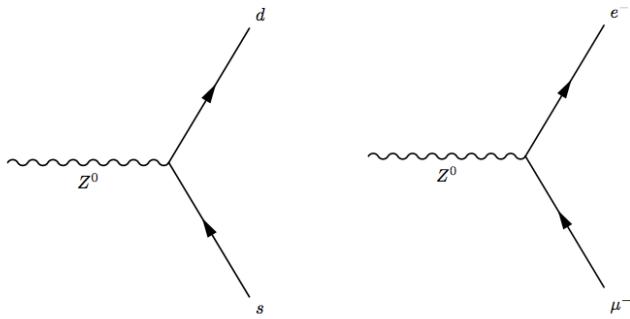


Figure 1.4: Feynman diagrams of FCNCs processes forbidden in the SM.

173

174 there cannot be FCNCs and, yet, they are experimentally observed to be highly
175 suppressed. On the other hand the observation of neutrino oscillation proves that
176 flavour is not an exact symmetry and is not always conserved. Furthermore, the

177 values of the terms of the CKM matrix and the PMNS matrix, which the mixing-
178 matrix, equivalent to the CKM, in the lepton sector, are not explained in the SM
179 and have to be measured experimentally. These open problems motivate searches
180 for flavour symmetries and deeper motivations for flavour conservation.

181 **1.4 Beyond the Standard Model**

182 From the last sections it is evident that, despite the great success of the SM, there
183 is a need to explore new theories. Among the most promising approaches there are
184 are those invoking Super-Symmetry and extra-dimensions In Super-Symmetry new
185 degrees of freedom are introduced to suppress the diverging terms of the scalar mass.
186 This theory assumes that for each fermion there is a corresponding boson and, since
187 bosons and fermions contribute with opposite sign to the mass term, these would
188 cancel out [11]. Supersymmetry also provides a natural candidate for dark matter,
189 the neutralino, which is a weakly interacting stable particle. The idea to introduce
190 extra-dimensions was triggered by the fact that normally gravity is not relevant in
191 particle physics but it would be natural if all forces had similar strength. By adding
192 extra dimensions to the normal 3 spatial dimensions, one can restore the strength
193 of gravity, as this could be dispersed by the wider space available [12]. In all these
194 approaches severe constraints to masses and couplings must be imposed to maintain
195 compatibility with the SM at the electroweak scale.

196 **1.4.1 Flavour and BSM theories**

197 Most BSM theories predict processes violating flavour conservation. Therefore, the
198 observation or non-observation of these processes can give important information
199 about new physics. BSM theories can be classified according to the amount of flavour
200 violation they introduce. The first class of models to consider is the Minimal Flavour
201 Violation (MFV). These are models in which the only sources of flavour changing

202 transitions are governed by the CKM matrix with the CKM phase being the only
 203 source of CP violation. These features can be assured by symmetry principles and
 204 these types of models are naturally compatible with the SM. Examples of such
 205 models include the MSSM which minimal flavour violation and the SM with one
 206 extra-dimension. A review of MFV models is presented in Refs. [13, ?]. The MFV
 207 paradigm provides a way to resolve the tension between expectation, driven by
 208 naturalness arguments, that NP should be at the TeVscales and limits on FCNC
 209 processes that point to much higher scales. A powerful test of MFV is provided by
 210 the study of ratios between $b \rightarrow d$ and $b \rightarrow s$ transitions, because their hamiltonians
 211 share the same structure. One particularly important example is the ratio of B^0
 212 and B_s^0 dimuon decay rates [14], as this is a purely leptonic decay free from hadronic
 213 uncertainties. In the SM such ratios are approximately equal to $|V_{td}/V_{ts}| \sim 1/25$,
 214 modified by phase space and hadronic matrix elements, while they can take very
 215 different values in non-MFV models.

216 In the quest for New Physics an important role is also played by simplified models
 217 as an intermediate model building step. Instead of constructing models valid up
 218 to the GUT scale one can consider simplified models, which typically start from
 219 the SM and incorporate a new sector with a limited number of parameters. Such
 220 models are easier to constrain but can nevertheless point in the right direction to
 221 build more complete theories. The choice of the new sector to add can be driven
 222 by the need to explain existing tensions between data and SM predictions or by
 223 theoretical prejudice. Two models especially relevant when studying rare decays are
 224 Z'-penguins and leptoquarks. A Z'-penguin is a FCNC process involving a neutral
 225 field arising from an extra U(1) gauge symmetry. As for the SM penguins, this field
 226 contributes in loops causing modifications of the effective couplings with respect to
 227 the SM. A survey of Z' models can be found in Ref. [15]. Leptoquarks are bosonic
 228 particles that carry one quark and one lepton flavour quantum number. They can be
 229 spin 1 but they are commonly assumed to be scalar particles. A tree level exchange
 230 of a leptoquark induces processes such as $b \rightarrow (s, d)\ell^+\ell^-$, and therefore can result
 231 in an enhancement of their decay rates with respect to the SM [16]. Leptoquarks

²³² would also provide a natural explanation for non-universal couplings to leptons,
²³³ introducing lepton flavour violation.

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

²³⁵ In the Standard Model FCNC processes are forbidden at tree level but can oc-
²³⁶ cur trough loops diagrams such as W box or penguin diagrams (see Fig. 1.5). The
²³⁷ branching fractions decays going through these processes are small, typically $\sim 10^{-6}$
²³⁸ or lower, and therefore they are called “rare decays”. Additional NP contributions
²³⁹ to the virtual loops are not necessarily suppressed with respect to the SM compo-
²⁴⁰ nent and this makes these decays very sensitive to new physics. This approach to
²⁴¹ new physics searches is interesting as new particles could be at a high mass scale not
²⁴² accessible at colliders but its effect could be observed in loop effects. Radiative and
²⁴³ penguin decays are particularly interesting because they are theoretically well under-
²⁴⁴ stood, which allows precise comparisons with measurements. Finally, they provide
²⁴⁵ a great quantity of observables that can be affected by NP, not only decay rates,
²⁴⁶ but also CP asymmetries and angular observables such as forward-backward asym-
²⁴⁷ metries. The joint analysis of different observables can help building a consistent
 picture and rule out specific models.

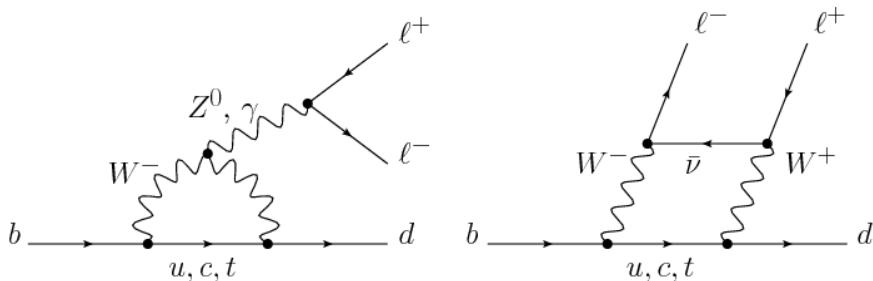


Figure 1.5: Loop Feynmann diagrams allowing $b \rightarrow d$ FCNC processes: penguin diagram (left) and W box (right).

²⁴⁹ 1.5.1 Theoretical framework: the effective Hamiltonian

²⁵⁰ Rare decays of b hadrons are governed by an interplay between weak and strong
²⁵¹ interactions. The large masses of W , Z and top quark compared to that of the
²⁵² b quark allow the construction of an effective theory divides the the problem of
²⁵³ calculating weak decay amplitudes into two parts, The first part deals with “short
²⁵⁴ distance” physics, namely perturbative contributions due to energy scales above
²⁵⁵ the b mass. The second parts handles “long distance”, typically non-perturbative,
²⁵⁶ contributions. Figure 1.6 illustrates an example of a Fermi theory where the short
 distance physics is hidden into a point like vertex. The effective hamiltonian [17] of

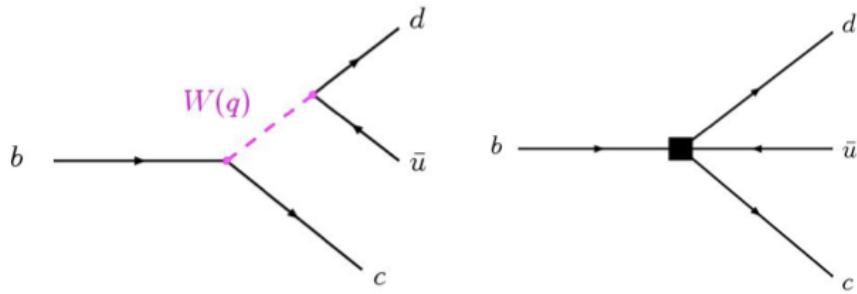


Figure 1.6: Example of a Fermi theory in which the full theory is divided between the short distance contribution, hidden in the vertex, and the long distance contribution.

²⁵⁷

²⁵⁸ such a theory relevant to $b \rightarrow s/d\gamma$ and $b \rightarrow s/d\ell^+\ell^-$ transitions can be written as:

$$\mathcal{H}_{eff} = \frac{-4G_F}{\sqrt{2}} \left[\lambda_q^t \sum C_i(\mu, M) \mathcal{O}_i(\mu) + \lambda_q^u \sum C_i(\mu, M) (\mathcal{O}_i(\mu) - \mathcal{O}_i^u(\mu)) \right] \quad (1.9)$$

²⁵⁹ where G_F denotes the Fermi coupling constan and the λ constants are the CKM
²⁶⁰ factors, $\lambda_q^t = V_{tb}V_{tq}^*$ and $\lambda_q^u = V_{ub}V_{uq}^*$. To obtain this formula the method of the
²⁶¹ Operator Product Expansion (OPE) [18] was used to allow the separation into the
²⁶² long-distance contributions, contained in the operator matrix elements, \mathcal{O}_i , and the
²⁶³ short-distance physics described by the so called Wilson Coefficients, C_i . Opera-
²⁶⁴ tors and coefficient are evaluated at the renormalization scale μ . In $b \rightarrow s$ quark
²⁶⁵ transitions, which are the main topic of this thesis, the doubly Cabibbo-suppressed
²⁶⁶ contributions proportional to λ_s^u can be neglected.

²⁶⁷ In order to describe SM processes the effective theory must be matched with the SM

268 at the EW scale, μ_W . Then, using the scale independence of the effective hamiltonian, one can derive a renormalization group equation for the Wilson Coefficients
269

$$\mu \frac{d}{d\mu} C_i(\mu) = \gamma_{ij} C_j(\mu) \quad (1.10)$$

270 where the matrix γ is the anomalous dimensions matrix of the operators \mathcal{O}_i . At
271 leading order the solution is given by [19]:

$$C_i(\mu) = \left[\frac{\alpha_s(\mu_W)}{\alpha_s(\mu)} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W) = \left[\frac{1}{1 + \beta_0 \frac{\alpha_s(\mu)}{4\pi} \ln \frac{\mu_W^2}{\mu^2}} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W) \quad (1.11)$$

272 where α_s is the strong coupling constant. In the SM at $\mu_s = m_b$ the Wilson Coeffi-
273 cients have values:

$$C_7^{SM} = -0.3, \quad C_9^{SM} = 4.2, \quad C_{10}^{SM} = -4.2. \quad (1.12)$$

274 New physics contributions appears in the Wilson Coefficients in the form of additive
275 factors: $C_i = C_i^{NP} + C_i^{SM}$.

276 Finally, the amplitudes of exclusive hadronic decays can be calculated as the expec-
277 tation values of the effective hamiltonian. Given an initial state I and a final state
278 F (e.g. $I = B$ and $F = K^{*0}\mu^+\mu^-$) they can be calculated as

$$A(M \rightarrow F) = \langle M | \mathcal{H}_{eff} | F \rangle = \mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} \sum V_{CKM}^i C_i(\mu) \langle M | \mathcal{O}_i(\mu) | F \rangle \quad (1.13)$$

279 where $\langle M | \mathcal{O}_i(\mu) | F \rangle$ are the hadronic matrix elements also called “form factors”.
280 These can be evaluated using non perturbative methods such as lattice calcula-
281 tions. However, due to the limitations of these methods, the dominant theoretical
282 uncertainties reside in the calculation of the matrix elements.

²⁸³ **1.5.2 Operators**

²⁸⁴ Separating the left- right-handed components the relevant effective Hamiltonian for
²⁸⁵ $b \rightarrow s\ell^+\ell^-$ transitions is

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{\alpha_e}{4\pi} \sum_{i=1}^{10} [C_i \mathcal{O}_i + C'_i \mathcal{O}'_i]. \quad (1.14)$$

²⁸⁶ The operators are that are important for electroweak penguin processes are the
²⁸⁷ following [14]:

$$\begin{aligned} \mathcal{O}_7 &= \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_R b) F_{\mu\nu} & \mathcal{O}'_7 &= \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_L b) F_{\mu\nu} \\ \mathcal{O}_8 &= g_s \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_R T^a b) G_{\mu\nu}^a & \mathcal{O}'_8 &= g_s \frac{m_b}{e} (\bar{s}\sigma^{\mu\nu} P_L T^a b) G_{\mu\nu}^a \\ \mathcal{O}_9 &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell) & \mathcal{O}'_9 &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu \ell) \\ \mathcal{O}_{10} &= (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell) & \mathcal{O}'_{10} &= (\bar{s}\gamma_\mu P_R b)(\bar{\ell}\gamma^\mu \gamma_5 \ell) \end{aligned} \quad (1.15)$$

²⁸⁸ where $P_{L/R} = (1 \mp \gamma_5)/2$ denotes the left/right handed chiral projection, T^a are
²⁸⁹ the QCD generators and $F_{\mu\nu}$ ($G_{\mu\nu}$) is the electromagnetic (chromo-magnetic) field
²⁹⁰ tensor. The \mathcal{O}' operators correspond to right-handed coupling obtained by swap-
²⁹¹ ping P_R and P_L in the equations. In the SM, as well as in MFV models where
²⁹² the flavour violation is entirely ruled by the CKM matrix, the results C' Wilson
²⁹³ Coefficients are suppressed by the strange coupling $C'_i \sim (m_s/m_b)C_i$. The operator
²⁹⁴ \mathcal{O}_7 is the dominant contribution to the radiative $b \rightarrow s\gamma$ transitions, while \mathcal{O}_9 and
²⁹⁵ \mathcal{O}_{10} are the dominant contributions in semileptonic $b \rightarrow s\ell^+\ell^-$ decays. The vertices
corresponding to these operators are shown in Fig. 1.7

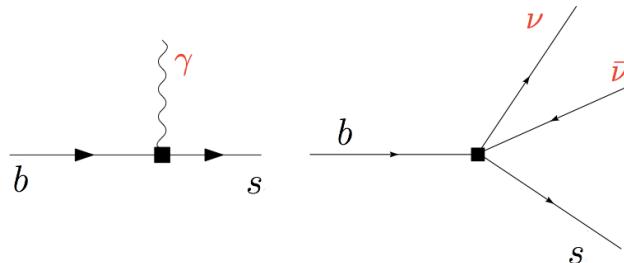


Figure 1.7: Interaction vertices corresponding to the radiative (left) and semileptonic (right) operators.

²⁹⁷ It is also common to express the semileptonic operators in a basis with left and right
²⁹⁸ projected leptons

$$\begin{aligned}\mathcal{O}_{LL} &= (\mathcal{O}_9 - \mathcal{O}_{10})/2 & \mathcal{O}_{LR} &= (\mathcal{O}_9 + \mathcal{O}_{10})/2 \\ \mathcal{O}_{RR} &= (\mathcal{O}'_9 - \mathcal{O}'_{10})/2 & \mathcal{O}'_{RL} &= (\mathcal{O}'_9 + \mathcal{O}'_{10})/2\end{aligned}\tag{1.16}$$

²⁹⁹ where the Wilson Coefficients are also redefined as

$$\begin{aligned}C_{LL} &= C_9 - C_{10} & C_{LR} &= C_9 + C_{10} \\ C_{RR} &= C'_9 - C'_{10} & C'_{RL} &= C'_9 + C_{10}\end{aligned}\tag{1.17}$$

³⁰⁰ This basis is particularly useful in frameworks where BSM physics at a high mass
³⁰¹ scale respects the $SU(2)_L$ part of the SM gauge symmetry group. For instance,
³⁰² instead of fitting the two parameters C_9 and C_{10} , the LL-hypothesis gives the con-
³⁰³ straint $C_9 + C_{10} = 0$.

³⁰⁴ Finally, in the picture presented in this section all operators were considered as
³⁰⁵ universal with respect of the flavour of the involved leptons. However, BSM mod-
³⁰⁶ els often contain courses of lepton flavour violation leading to a split of the same
³⁰⁷ operators into two groups depending on the lepton considered.

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

³⁰⁹ Semileptonic b hadron decays are characterised by two kinematic regimes which
³¹⁰ are treated theoretically in different ways: at low q^2 , where the emitted hadron is
³¹¹ energetic ($E > \Lambda_{QCD}$ in the b hadron rest frame), the QCD factorisation applies;
³¹² at high q^2 , the region of low hadron recoil ($q^2 = O(m_b)$), an Operator Product
³¹³ Expansion in $1/m_b$ is valid. In both regions decay rates can be predicted using the
³¹⁴ different methods and the biggest uncertainties come from the limited knowledge
³¹⁵ of hadronic transition matrix elements. As can be seen in Fig. 1.8 the very low q^2
³¹⁶ is characterised by a peak due to the virtual photon contribution, associated with
³¹⁷ C_7 . In the region $1 - 6 \text{ GeV}^2/c^4$ the interference between C_7 and C_9 becomes large,

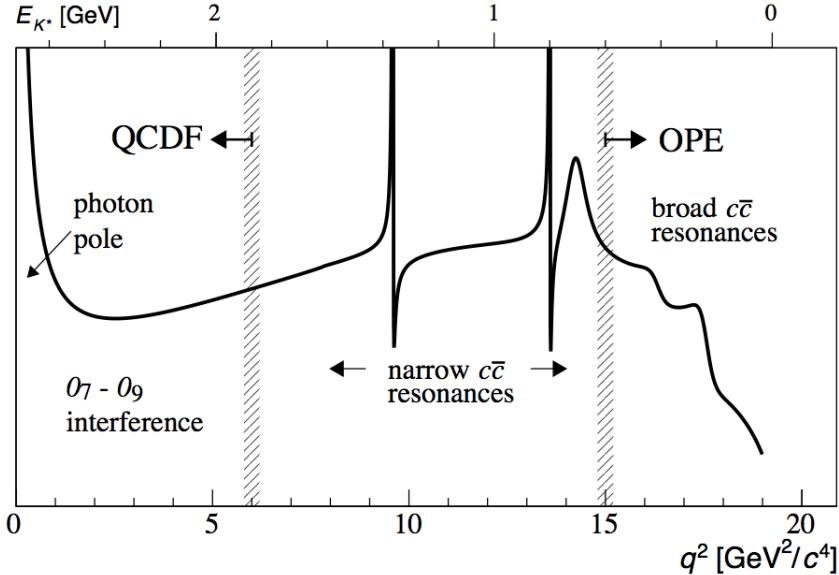


Figure 1.8: A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ process characterised by the photon pole at very low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2 .

³¹⁸ yielding sensitivity to NP in C_9 . The $6 - 15$ GeV^2/c^4 interval is dominated by the
³¹⁹ charmonium resonances, J/ψ and $\psi(2S)$, though the tree level $b \rightarrow \bar{c}cs$ transition.
³²⁰ Although the decays can be experimentally vetoed in principle charmonia affect the
³²¹ entire q^2 space. Finally, at high q^2 broad charmonium resonances can contribute,
³²² like those observed by LHCb in $B^+ \rightarrow K^+\mu^+\mu^-$ decays [20].

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

³²⁴ Rare decays and especially semileptonic $b \rightarrow s\ell^+\ell^-$ processes offer a number of ob-
³²⁵ servables which can be used to benchmark BSM models. The most direct effects
³²⁶ appear in decay rates that can be enhanced by NP but the precision on these mea-
³²⁷ surements is often limited by the uncertainty on form factor calculations. Therefore,
³²⁸ it is important to also look for different observables. One important class of observ-
³²⁹ ables are angular quantities that can often carry complementary information about
³³⁰ NP with respect to branching ratio measurements. The most basic of these observ-
³³¹ able are forward-backward asymmetries that characterise the angular distribution

332 of final particles. For the $B^0 \rightarrow K^*\mu^+\mu^-$ decay combinations of observables have
333 been proposed that are independent of form factor uncertainties at leading order
334 order [14].

335 One way to build safe observables is to construct ratios between similar decays, in
336 which uncertainties due to the hadronization process cancel out. These observables
337 include the R_H ratios, between B^0 decay into electrons and muons, that are described
338 in detail in Sec. 12. It is also interesting to compare decays which go through the
339 same fundamental process but where the spectator quark has a different flavour. This
340 is the case of $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K_s^0\mu^+\mu^-$ decays, which are both $b \rightarrow s$
341 transitions where the spectator quark is an u quark in the first case and a d quark
342 in the second. The ratio of the branching fractions of these decays is called isospin
343 asymmetry.

344

1.6 Experimental status

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

350

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 $B^0 \rightarrow \mu^+\mu^-$ 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 [21]:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9} \text{ and} \quad (1.18)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}. \quad (1.19)$$

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 [22]:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (2.8_{-0.6}^{+0.7}) \times 10^{-9} \text{ and} \quad (1.20)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (3.9_{-1.4}^{+1.6}) \times 10^{-10}. \quad (1.21)$$

- 351 Both decays were previously unobserved and now the B_s^0 decay was observed with
- 352 a significance of 6σ and evidence for the B^0 decay was found with a 3σ significance.
- 353 These are compatible with SM predictions within 2σ and put strong constraints to
- 354 the available parameter-space for BSM theories. Figure 1.9 shows the fit the dimuon
- 355 invariant mass of B meson candidates where the peaks of the two decays are visible.

356

357 1.6.2 Semileptonic $b \rightarrow s \ell^+ \ell^-$ decays of b hadrons

- 358 At the LHC energies is now possible to collect large data sample of semileptonic
- 359 decays, especially those with a dimuon pair in the final state. Many branching
- 360 fractions of semileptonic B meson decays were recently measured at the LHCb ex-
- 361 periment, including $B \rightarrow K \mu^+ \mu^-$, $B \rightarrow K^{*0} \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ [23, 24, 25].
- 362 Baryon decays where also studied at LHCb: including the branching fraction of the

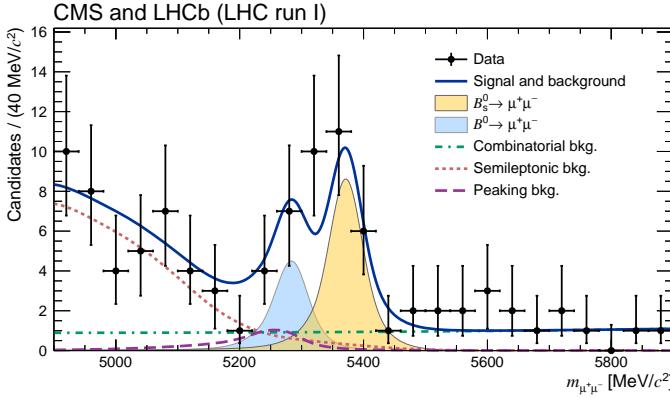


Figure 1.9: Dimuon invariant mass of B candidates showing peaks corresponding $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays [22].

363 rare $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$ decay [1], which is described in this thesis. Unlike for pure lep-
364 tonic decays, SM predictions for semileptonic decays are affected by the knowledge
365 of hadronic form factors, which yields in relatively large uncertainties, $\mathcal{O}(30\%)$. As
366 a result measurements are now typically more precise than predictions.

367 As described in Sec. 1.5.4 angular observables can be affected by new physics. Partic-
368 ular interest was risen by the measurement of a series of observables in $B \rightarrow K^{*0} \mu^+ \mu^-$ decays, free from form factors uncertainties at leading order [26]. Most of the mea-

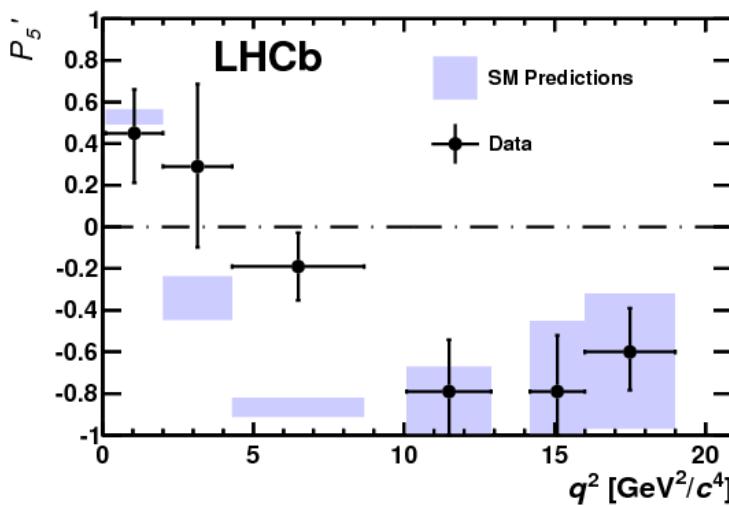


Figure 1.10: Measurement of the observable as a function of q^2 , showing a tension with SM predictions in the 2–6 GeV^2/c^4 region.

369 surements are found to be in agreement with SM predictions with the exception

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	$?0.094 \pm 0.047$	-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 1.2: Measurement of CP and isospin asymmetry in $B^0 \rightarrow K^{(*)} \mu^+ \mu^-$ decays from the LHCb experiment.

of the P'_5 observable, shown in Fig. 1.10, which presents a local 3.7σ deviation.
 Attempts to build a consistent picture point to a NP contribution to the Wilson Coefficient C_9 [27]. An angular analysis of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays was also performed, where observables are found to be compatible with SM predictions [28].

Other observables for which the sensitivity to form factors effects is reduced are the CP asymmetry between B and \bar{B} decays, \mathcal{A}_{CP} , and the isospin asymmetry between B^0 ad B^+ decays. Due to the small numerical size of the corresponding CKM elements CP asymmetries of $B^0 \rightarrow K^{(*)} \mu^+ \mu^-$ decays are tiny in the SM, $O(10^{-3})$. In BSM models new sources of CP violation can arise and therefore their measurement is a powerful null test of the SM. The isospin asymmetry, \mathcal{A}_I , between B^+ and B^0 is not zero in the SM due to isospin breaking effects in the form factors. This is expected to be $\sim 1\%$ at low q^2 and grow up to $\sim 1\%$ as q^2 tends to zero. The LHCb experiment, using the full dataset collected in Run I, corresponding to an integrated luminosity of 3 fb^{-1} measured both these asymmetries to be consistent with zero [29, 30], as reported in Tab. 1.2.

1.6.3 Lepton Flavour Violation searches

Several LFV searches are linked to rare decays as they involve small branching ratios in the SM that can be enhanced by new physics. They are therefore a natural place to look for NP. Lepton flavour conservation is well experimentally established measuring the branching ratios of decays of muons into electrons and no neutrinos but has no strong theoretical explanation in the context of the SM. In fact it is already observed know that flavour is not conserved in neutrino oscillations.

393 The best-studied decays violating lepton flavour are rare muon decays including
 394 $\mu^+ \rightarrow e^+ \gamma$ and $\mu^+ \rightarrow e^+ e^- e^+$. Since muons can be abundantly produced and the
 395 final states are simple, these decays provide the best constraints to LFV. The present
 396 best-upper limits are 1.2×10^{-11} for the radiative decay and 1.0×10^{-12} for $\mu^+ \rightarrow$
 397 $e^+ e^- e^+$ obtained respectively by the MEGA [31] and SINDRUM [32] experiments.
 398 Several LFV searches in the B sector have been recently been performed at the
 399 LHCb experiment including decays such as $B^0 \rightarrow e\mu$ [33] and τ decays such as
 400 $\tau \rightarrow \mu^+ \mu^- \mu$ [34]. None of these searches has found evidence of NP so far and
 401 therefore they set limits, constraining the parameter space available for NP models.
 402 Fig. ?? shows a summary of the best limits to date on LFV searches [35].

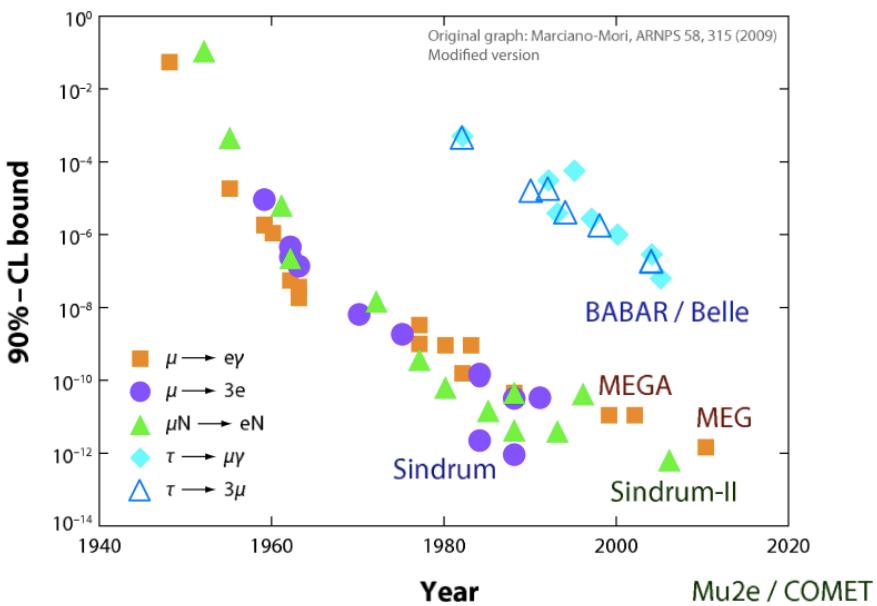


Figure 1.11: Summary of limits set in lepton flavour violation searches [35].

403

CHAPTER 2

404

405

The LHCb detector at the Large Hadron Collider

406

407

2.1 The Large Hadron Collider

408 The Large Hadron Collider (LHC) [36] is a circular particle accelerator with a
409 circumference of 27 km located about 100 m underground in the surroundings of
410 Geneva, Switzerland. Two proton beams circulate in opposite directions around the
411 ring and cross each other at several points, in which particle detectors are placed.
412 These include two general-purpose detectors, ATLAS and CMS, siting on opposites
413 sides of the ring and the two smaller specialty detectors, ALICE and LHCb, are at
414 the interaction points to either side of ATLAS (see fig. 2.1).

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

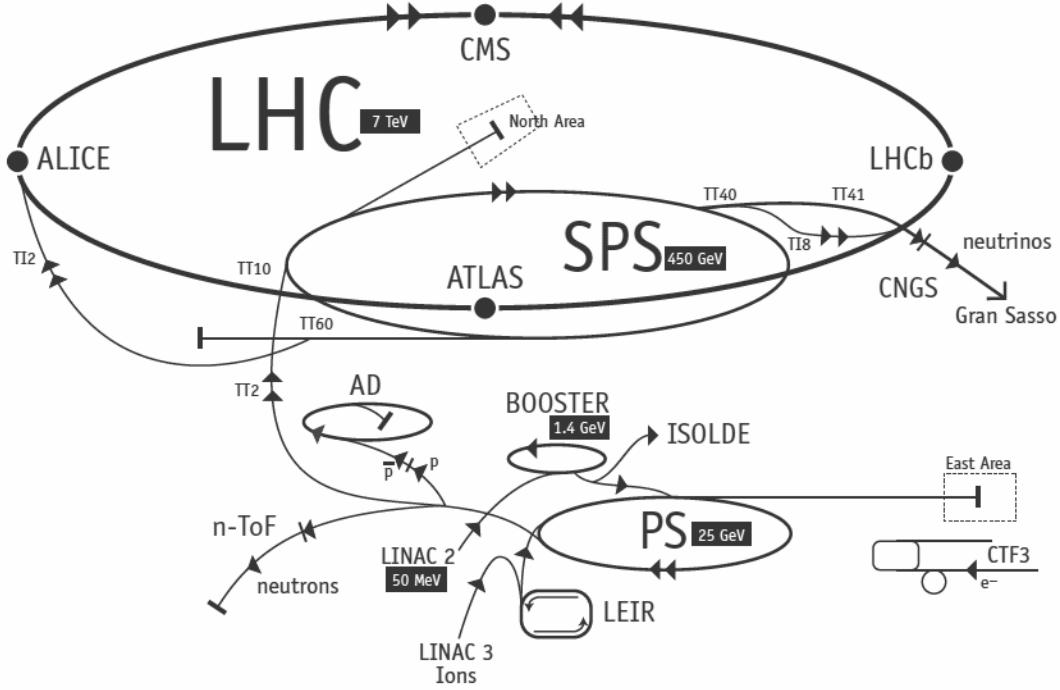


Figure 2.1: Scheme of CERN accelerators.

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

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

438 2.2 The LHCb detector

439 The LHCb detector [37] was built with the main purpose of studying the decays of
440 B and D mesons, looking in particular for CP-violating processes. In 2011, running
441 at a centre of mass energy of 7 TeV, the cross section of $b\bar{b}$ production was measured
442 to be $284 \pm 53 \mu b$ [?], while it will be $\sim 500 \mu b$ at the current LHC energy, 13 TeV.
443 At these high energies, proton-proton interactions produce highly boosted virtual
444 gluons which interact to produce $b\bar{b}$ pairs at small angles, close to the beam pipe. For
445 this reason the LHCb detector is designed to have a very forward angular coverage:
446 it is fully instrumented from approximately 10 mrad to 300 mrad, corresponding to
447 $2 < \eta < 5$, where η is the “pseudorapidity”, a quantity used in particle physics and
448 called and defined as:

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

449 In Eq. 2.1, θ is the angle between a particle’s momentum and the beam direction ¹.

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

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.

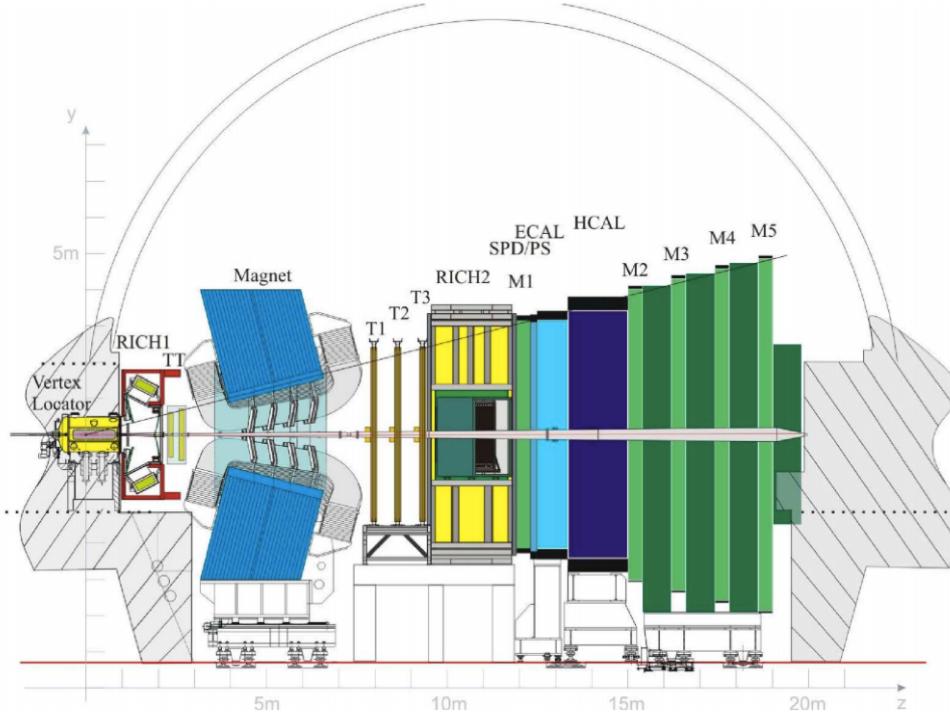


Figure 2.2: A side view of the LHCb detector [37].

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

462 Other B physics experiments, like BaBar at the Stanford Linear Accelerator (SLAC),
 463 Belle at KEK at J-PARC (Japan) and the Tevatron experiments at Fermilab have
 464 made accurate measurements in heavy flavour physics. All of these results have
 465 so far been consistent with the Standard Model predictions. However, some of the
 466 deviations from the Standard Model are expected to be very small, therefore LHCb
 467 has begun to make the most precise measurements in heavy flavour physics to test
 468 the consistency of the Standard Model and look for new physics.

469 The LHCb detector includes a high-precision tracking system consisting of a silicon-
 470 strip vertex detector surrounding the pp interaction region, a large-area silicon-strip
 471 detector located upstream of a dipole magnet with a bending power of about 4 Tm,
 472 and three stations of silicon-strip detectors and straw drift tubes placed downstream.

⁴⁷³ The combined tracking system has momentum resolution $\Delta p/p$, that varies from
⁴⁷⁴ 0.4% at 5 GeV/c² to 0.6% at 100 GeV/c². Charged hadrons are identified using
⁴⁷⁵ two Ring-Imaging Cherenkov detectors (RICH) [38]. Photon, electron and hadron
⁴⁷⁶ candidates are identified by a calorimeter system consisting of scintillating-pad and
⁴⁷⁷ pre-shower detectors, an electromagnetic calorimeter and a hadronic calorimeter.
⁴⁷⁸ Muons are identified by a system composed of alternating layers of iron and multiwire
⁴⁷⁹ proportional chambers [39]. A schematic view of the detector is shown in Fig. ??
⁴⁸⁰ and more details on each subdetector are given in the following sections.

⁴⁸¹ 2.3 The magnet

⁴⁸² Charged particle tracks are bent horizontally in the magnetic field so that their
⁴⁸³ momentum can be measured from the curvature radius. The LHCb dipole magnet
⁴⁸⁴ is comprised of two coils supported on an iron yoke and is shaped to fit the LHCb
⁴⁸⁵ angular acceptance. Unlike the other LHC experiments, LHCb uses a warm magnet,
⁴⁸⁶ so that it can be ramped easily and the field can be reversed periodically. When the
⁴⁸⁷ polarity is flipped and particles of a given sign are bent in the opposite direction.
⁴⁸⁸ This method is used to limit systematic uncertainties that can arise from
⁴⁸⁹ different performances in different areas of the detector and average out using data
⁴⁹⁰ taken in both polarities. In the magnet flow a 5.85 kA current that generates
⁴⁹¹ an integrated magnetic field of 4 Tm for 10 m long tracks. In order to achieve
⁴⁹² the required momentum precision the magnetic field must be mapped with a 10⁻⁴
⁴⁹³ precision. For this reason a grid of 60 sensors is positioned inside the magnet and
⁴⁹⁴ provides real time magnetic field maps.

⁴⁹⁵ 2.4 Tracking system

⁴⁹⁶ B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means
⁴⁹⁷ they travel about 1 cm before decaying and they form a displaced vertex. To study

498 specific decays, it is therefore important to be able to separate the particles produced
 499 at the primary pp vertex and the B decay vertex. The tracking system consists in the
 500 Vertex Locator (VeLo), and 4 tracking stations: the Tracker Turicensis (TT), which
 501 are located before the magnet and the T1, T2 and T3 stations, located downstream
 502 of the magnet. The latter three stations are in turn formed by two subsystems: the
 503 Inner Tracker (IT) close to the beam-line, where the particle density is greatest and
 504 the Outer Tracker (OT) covering the rest of the acceptance.

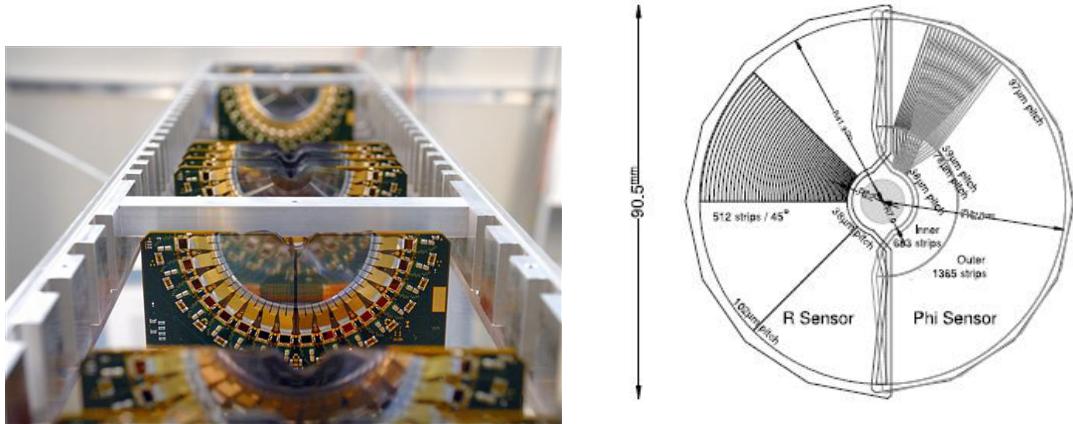


Figure 2.3: On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [37].

505

506 The VeLo accurately measures positions of tracks close to the interaction point so
 507 that production and decay vertices of bottom and charm hadrons can be recon-
 508 structed. The VeLo is composed by 21 staggered silicon modules which surround
 509 the beam axis and are positioned from $z = -18$ cm to $+80$ cm. It is able to detect
 510 particles within a pseudorapidity range $1.6 < \eta < 4.9$. The sensitive region of the
 511 VeLo starts at an inner diameter of only 8 mm from the beam axis. The VeLo is
 512 housed in its own vacuum vessel of thin aluminium foil which protects the vacuum
 513 of the beam pipe from any outgassing of the VeLo. The silicon layers composing
 514 the VeLo consist of two modules each including two types of sensors: the ϕ -sensor
 515 which measures the azimuthal position around the beam, and the R-sensor which
 516 measures the radial distance from the beam axis. A sketch of the VeLo sensor is

517 shown in Fig. 2.3. The sensors are $300 \mu\text{m}$ thick, approximately semicircular and
 518 are positioned on either side of the beam axis. To ensure that they cover the full
 519 azimuthal angle the right-side module is placed 1.5 cm behind the left-side module
 520 on the z-axis and they overlap. There are two modules which cover the backward
 521 direction and are used as a veto for multiple interactions, this is called the pileup
 522 system.

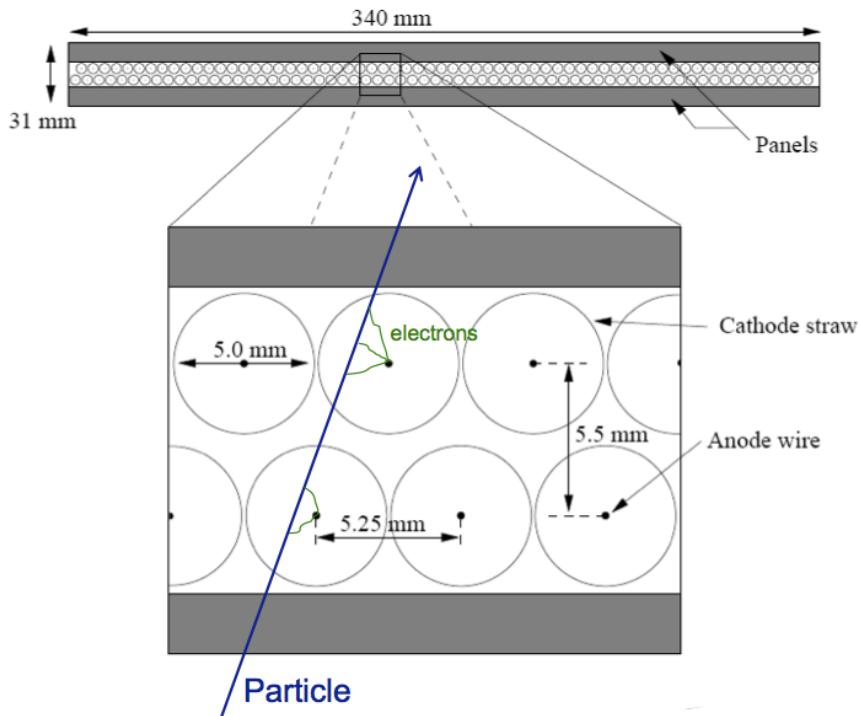


Figure 2.4: Sketch of the straw tubes which constitute the Outer Tracker layers [37].

523

524 The IT and TT both use silicon microstrips and together constitute the Silicon
 525 Tracker (ST). Straw tubes are used in the OT, of which a sketch is shown in Fig.2.4.
 526 The IT requires an higher inner granularity because of the higher flux of particles
 527 nearer the beam pipe, in fact it covers only 1.3% of the total area of IT plot OT
 528 but it contains about 20% of the tracks. Each ST station has four detection layers,
 529 the first and last being vertical, measuring the track position in x. The second and
 530 third layer are rotated by a stereo angle of +5 and -5 degrees, which allows the y-
 531 coordinate to be measured. The TT is placed upstream of the magnet, which allows

532 reconstruction of the tracks from low-momentum particles which are swept out of
533 the downstream acceptance. Overall the tracking system provides a measurement of
534 momentum, p , with a relative uncertainty that varies from 0.5% at low momentum
535 to 1.0% at 200 GeV/c. The impact parameter (IP), namely the minimum distance
536 of a track to a primary vertex, is measured with a resolution of $(15 + 29/p_T)$ μm ,
537 where p_T is the component of the momentum transverse to the beam, in GeV/c.

538

2.5 Particle identification

539 Particle identification is an important feature in LHCb and it is performed in various
540 ways. The electromagnetic calorimeters can distinguish between pions and electron,
541 the muon chambers identify muons and the Ring Imaging Cherenkov (RICH) de-
542 tectors can be used to identify heavier charged particles as protons and kaons.

543

2.5.1 Calorimeters

544 The main purpose of the calorimeter system is to determine the energy of parti-
545 cles traversing the detector. A calorimeter is composed by layers of absorber and
546 active material. The absorber makes particles interact and produces a cascade of
547 secondaries, which multiply quickly and are detected by the active part. In LHCb
548 the sensitive material consists of scintillating layers, where the light detected is ap-
549 proximately proportional to the number of deposited particles. Calibration is then
550 used to translate the signal into a measurement of deposited energy. The LHCb
551 calorimeter system consists of the Scintillator Pad Detector (SPD), the Pre-Shower
552 Detector (PS) as well as the Electromagnetic Calorimeter (ECAL) and the Hadronic
553 Calorimeter (HCAL). The most difficult identification is that of electrons. First of
554 all the rejection of a high background of charged pions requires a longitudinal seg-
555 mentation of the electromagnetic calorimeters. This is provided by the PS detector
556 added in front of the main electromagnetic calorimeter, ECAL. Electrons also have

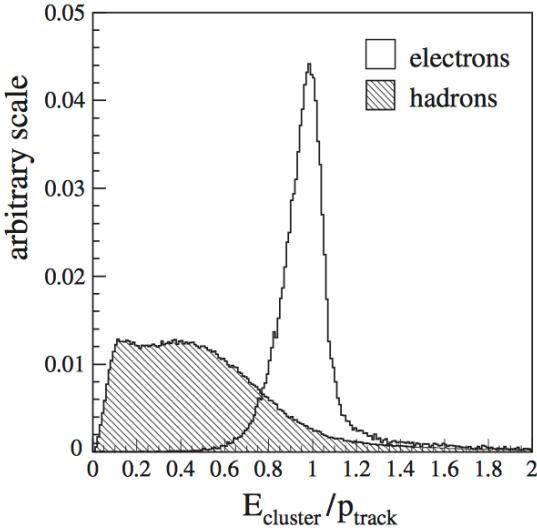


Figure 2.5: The ratio of the energy deposited in the ECAL and the particle momentum, which allows the separation between electrons and hadrons.

to be distinguished from high energy π^0 s. For this purpose the SPD calorimeter, detecting charged particles, is located in front of the PS and ECAL detectors. Fig. 2.5 shows how the ratio of the energy detected in the ECAL and the particle momentum allows the separation of electrons and 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 achieve 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 [37]

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

which results in a mass resolution of $\sim 70 \text{ MeV}/c^2$ or B mesons for $\sim 8 \text{ MeV}/c^2$ for π^0 . The trigger requirements on the HCAL resolution do not depend on the containment

572 of the hadron showers as much as for the ECAL, so due to a limited space, its
 573 thickness is only 5.6 interaction lengths and its resolution

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

574 2.5.1.1 Bremsstrahlung recovery for electrons

575 Bremsstrahlung is an electromagnetic radiation produced by particles, that deceler-
 576 ate or deviate because of the presence other charged particles. Typically electrons
 577 produce Bremsstrahlung when deflected by atomic nuclei. The probability of emit-
 578 ting bremsstrahlung radiation is proportional to the inverse of the squared mass of
 the particle ($1/m^2$) and therefore is relevant only for electrons. At these energies, if

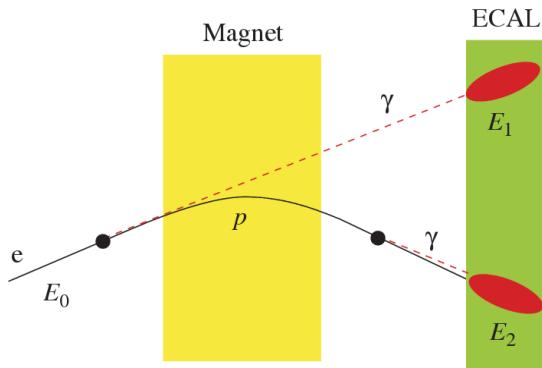


Figure 2.6: Schematic view of the Bremsstrahlung recovery.

579
 580 electrons radiate after the magnet, the photon will hit in the same calorimeter cells
 581 as the electron and the energy will be automatically recovered. However, if the pho-
 582 ton is emitted before the magnet, the electron will be deflected by the magnetic field
 583 whereas the photon will continue on its initial trajectory, with its energy being de-
 584 posited in a different part of the calorimeter. Missing this energy results in a poorer
 585 reconstructed B^0 mass resolution, so it is desirable to recover these bremsstrahlung
 586 photons, when possible. A tool for bremsstrahlung recovery is available in the LHCb
 587 analysis software. This tool looks for other clusters in the calorimeter and recon-
 588 structing the trajectory of the electron checks if they may have been emitted by

that. Then the photon energy is added to the electron and its momentum recalculated. Figure ?? shows a schematic view of the process. For more information see Ref. [40].

2.5.2 RICH

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

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

where β is the particle velocity over c and n is the refraction index of the medium.

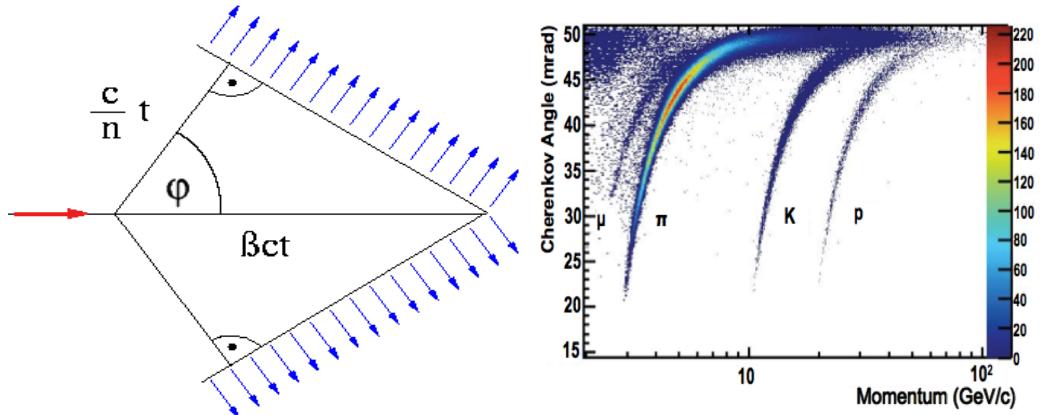


Figure 2.7: On the left a sketch of Cherenkov light emission and on the right the Cherenkov angle versus momentum for the two radiators of RICH1 and for different particles. One can see that they allow to separate particles in different momentum ranges.

599

RICH 1 is situated before the magnet in order to cover a large angular acceptance. Its purpose is to ensure particle identification over the momentum range $1 < p < 70$ GeV/c. It uses two radiators: C_4F_{10} that covers the momentum range 5 – 70

603 GeV/c and silica aerogel which covers $1 - 10$ GeV/ c . RICH 2 is positioned after
604 the magnet and tracking stations. It identifies higher momentum particles from
605 approximately 20 GeV/cup to beyond 100 GeV/cusing CF_4 as a radiator. The
606 Cherenkov light produced when charged particles travel through the radiators, is
607 reflected and focused using flat and spherical mirrors, which are tilted so that the ring
608 image is reflected onto arrays of photo-detectors. The radius of the ring can be used
609 to measure the opening angle of the Cherenkov cone because of the known geometry.
610 The photo-detectors are located outside of the LHCb acceptance in order to reduce
611 the amount of material that the particles have to traverse. Pattern recognition
612 algorithms are then used to reconstruct the Cherenkov rings.

613 2.6 The muon system

614 It is essential for many of the key physics analyses to be able to identify muons
615 in the final state. Muons are the most penetrating particles that can be detected
616 at LHC experiments, so the muon chambers are the farthest subdetectors from the
617 interaction point. The muon system is formed by five stations (M1 - M5), the first
618 one being located before the calorimeters in order to improve p_T measurements. A
619 scheme of the muon system is shown in Fig. ???. The remaining four stations lay
620 behind the HCAL and are separated by 1.2 m from each other, interleaved with iron
621 80 cm thick blocks, which absorb hadrons, electrons and photons to ensure that only
622 muons reach the final muon station. Only muons with a minimum momentum of 10
623 GeV/ctraverse all of the five stations and for positive identification of a muon the
624 trigger requires a signal in each of them. Each station has a detection efficiency of at
625 least 95% and the detectors provide position measurements. Since there is a larger
626 particle flux towards the beam pipe, the stations are divided into four concentric
627 rectangular regions (R1-R4), their size increasing according to the ratio 1 : 2 : 4 : 8.
628 This results in a similar channel occupancy over the four regions. All of the muon
629 stations use Multi Wire Proportional Chambers (MWPC) except for the inner region
630 of M1, where the particle flux is too high. In this region triple-GEM (Gas Electron

631 Multiplier) detectors are used because of their better ageing properties. The Gas
 632 Electron Multiplier (GEM) detectors in the inner region of M1 have to withstand a
 633 rate up to 500 kHz cm^{-2} of charged particles. In these detectors particles traversing
 634 through the drift gap between the cathode and the first GEM foil produce ionisation
 635 electrons, which are then attracted by electric fields though all of the GEM foils and
 636 multiply. They then drift into the anode inducing a signal on the pads. A gas
 637 mixture of Argon, CO_2 and CF_4 , is used to give a time resolution better than 3 ns.

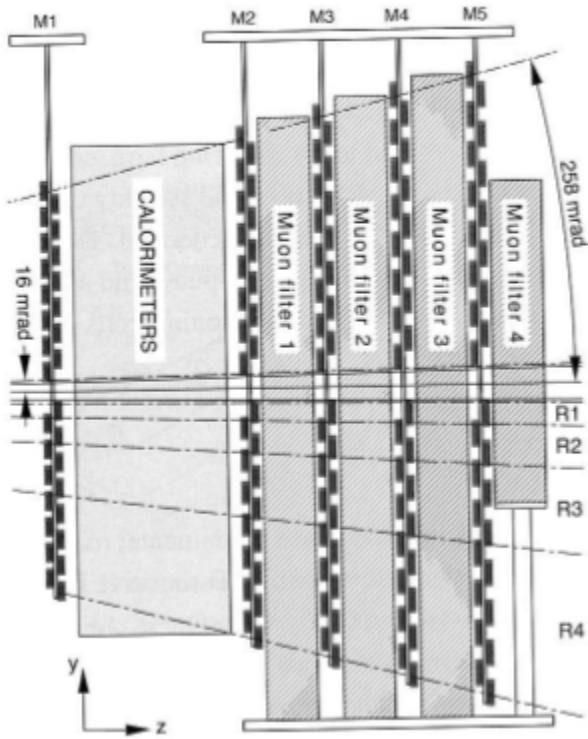


Figure 2.8: The LHCb muon system [37].

638

639 2.7 Trigger and software

640 The LHCb trigger system [41] consists of a hardware stage (L0), based on infor-
 641 mation from the calorimeters and muon system, followed by a software stage, the
 642 High-Level Trigger (HLT), which applies a full event reconstruction. To increase

643 performances the HLT is split again into stages (HLT1 and HLT2). The HLT1
 644 phase happens in real time and saves data in local disks while the HLT2 phases
 645 uses the resources available during periods with no beam. The event selected by
 646 the HLT2 stage are then saved for offline analysis. The bunch crossing frequency is
 647 40 MHz, which corresponds to an instantaneous luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ for
 648 LHCb. About 15% of the total number of $b\bar{b}$ pairs produced will contain at least
 649 one B meson with all of its decay products within the detector acceptance. This
 650 rate needs to be reduced down to about 2 kHz so that the events can be written to
 651 disk. Fig. 2.9 shows a scheme of the trigger system.

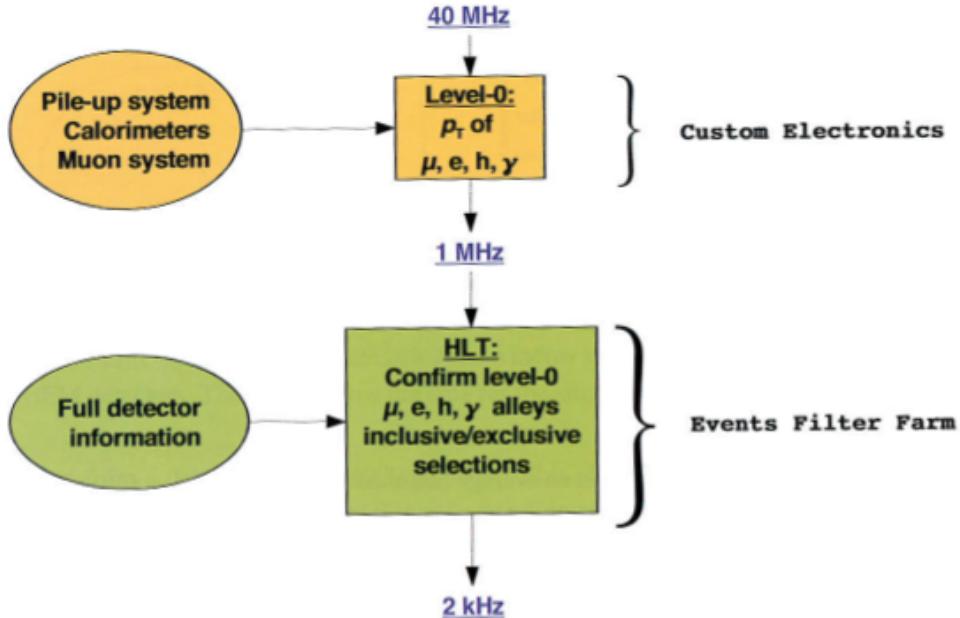


Figure 2.9: Scheme of the LHCb trigger system [37].

652 The L0 trigger reduces the rate of visible interactions from 10 MHz to 1 MHz. Due
 653 to the heavy mass of B mesons, they often produce particles with high p_T or E_T .
 654 Therefore the trigger selects events with large E_T deposits in the calorimeter of high
 655 p_T muons. The events is classified as L0Muon it was triggered due to information for
 656 the muon detector, while the information from the calorimeters is combined to divide
 657 the events in the 5 categories: L0Photon, L0Electron, L0LocalPion, L0GlobalPion,
 658 L0Hadron. The label “local” refers to π^0 reconstructed though their $\gamma\gamma$ decay,
 659 where the two photons fall in the same ECAL element, they are labelled “global”

660 otherwise. The HLT1 uses information from the Velo and trackers performing a
661 partial reconstruction of the event and reduces the rate to 2 kHz. Finally, the HLT2
662 involves a full reconstruction of the event and includes many “lines” designed to
663 trigger specific decays.

664 LHCb also developed an extended simulation software in order to reconstruct ef-
665 ficiencies and signal shapes. In the simulation, pp collisions are generated using
666 PYTHIA8 [42] with a specific LHCb configuration [43]. Decays of hadronic par-
667 ticles are described by EVTGEN [44], and final state radiation is generated using
668 PHOTOS [45]. Finally, the interaction of the generated particles with the detec-
669 tor and its response are implemented using the GEANT4 toolkit [46] as described
670 in Ref. [47]. For this analysis in this thesis, the ROOT framework [48] was used
671 to analyse data and the RooFit package to perform maximum likelihood fits. A
672 multivariate analysis is also used based on the NeuroBayes package [49, 50] which
673 provides a framework for neural network training.

674

Part I

675

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

676

CHAPTER 3

677

678

679

Introduction

680

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

687 Interest in Λ_b^0 baryon decays arises from two important facts. First of all, as Λ_b^0
688 has non-zero initial spin, there is a potential to learn information about the helicity
689 structure of the underlying Hamiltonian, which cannot be extracted from the meson
690 decays [51, 52]. Second, as Λ_b^0 baryon is in first approximation composed of heavy
691 quark and diquark formed of light quarks the hadronic physics significantly differs
692 from that of the mesons. This itself provides possibility to better understand and
693 test the hadronic physics in the theory, which could yield improved understanding
694 and confidence also for mesons.

With respect of B^0 decays going through the same transitions, such as $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, Λ can provide independent confirmations of the results as it involves the same operators but different hadronic matrix elements. Furthermore, Λ decays weakly and therefore complementary constraints with respect to B^0 decays can be extracted. Finally, the narrow width approximation, used in theoretical calculation is fully applicable in the Λ_b^0 case, which has $\Gamma_{\Lambda_b^0} \sim 2.5 \cdot 10^{-6}$ eV. This is not assured using $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ because the contribution from $B^0 \rightarrow K \pi \mu^+ \mu^-$ is unconstrained.

Theoretical aspects of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays were considered by a number of authors both in the SM and in different new physics scenarios [53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63]. All authors start from the same effective Hamiltonian already described in Sec. 1.5.1. However, form factors, describing hadronic physics, are not developed as well as in the meson case. Since there are not as many experimental constraints and form factors are still not well understood this leads to a relatively large spread in predicted branching fractions. An interesting quantity to study is the differential branching fraction as function of q^2 . This still suffers from knowledge of form factors, but as different approaches to form factors are applicable in different 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 [64], with signal yield of 24 ± 5 signal events. Later this was updated using the full CDF statistics [65]. Their preliminary 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 [66] 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. The LHCb result for the branching fraction relative to the $J/\psi \Lambda$ decay, used as normalisation channel, is

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

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 parts of the thesis describes the measurement of the differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ normalised by $J/\psi \Lambda$ using 3 fb^{-1} of pp collisions collected in 2011 and 2012. Furthermore an angular analysis of these decays is performed, 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. ???. This is characterised by the presence of the narrow peaks of the J/ψ and $\psi(2S)$ resonances. 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 had same final states 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. 4.2. 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

734 $q^2 > 15$ GeV $^2/c^4$, above $\psi(2S)$. In these two intervals the selection is optimised
735 to maximise the yield which is particularly important for a stable angular
736 analysis.

737 The above regions are then divided in smaller intervals, as much as the available
738 statistics allows, which results in bins ~ 2 GeV $^2/c^4$ wide. The binning used is the
739 following

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

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

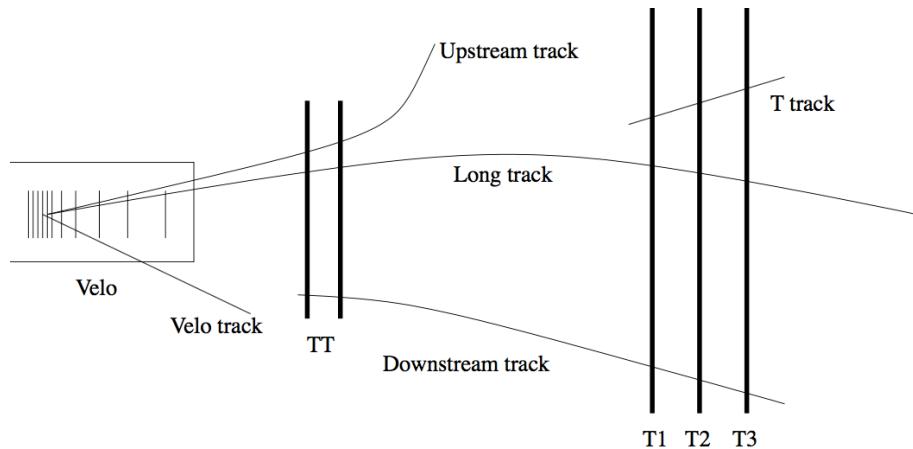
- 741 • 1.1-6.0 GeV $^2/c^4$: this interval is theoretically clean since is far from the pho-
742 ton pole, which dominated at low q^2 , washing out the sensitivity to NP con-
743 tributions. The lower bound of this interval it chosen excludes the possible
744 contribution from the ϕ resonance, which appears at 1 GeV $^2/c^4$. The
745 upper bound of the interval is chosen to totally exclude a small contribution
746 from the J/ψ resonance that leaks below 8 GeV $^2/c^4$.
- 747 • 15.0-20.0 GeV $^2/c^4$: this interval is the one that contains most of the statistics
748 and it is used as a natural cross check that the analysis in smaller bins is
749 stable.

750 3.2 Candidate types

751 This analysis deals with Λ baryons, which have a lifetime of $(2.632 \pm 0.020) \times 10^{-10}$
752 s [3]. These are considered long-lived particles in particle physics terms and can
753 travel into the detector for several meters generating well distinguished secondary
754 vertices. In LHCb Λ baryons can be reconstructed from tracks with or without
755 hits in the VELO and therefore define two candidate types as follows:

- 756 • **Long candidates:** built from tracks which have hits in the Velo, “long
 757 tracks”. These candidates, also denoted as “LL”, are characterised by a bet-
 758 ter momentum resolution thanks to the longer leverage arm available to long
 759 tracks.
- 760 • **Downstream candidates:** built from tracks without hits in the Velo,
 761 “downstream tracks”, also denoted as “DD”.

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



767 Figure 3.1: Representation of the two Λ candidate types built from “long” and
 “downstream” tracks.

767 3.3 Simulation

768 Samples of simulated events are needed in order to train the multivariate classifier
 769 (see Sec. 4.1), calculate the selection efficiency and study possible background. In
 770 particular for this analysis one sample of ~ 2 millions $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and ~ 5 millions
 771 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ simulated events are used. Samples of simulated $B^0 \rightarrow J/\psi K_s^0$,

772 $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ and $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ events are also used to study backgrounds
773 from these decays. The events are generated using Pythia8, hadronic particle are
774 decayed using EvtGen and Geant4 is used to simulate the interaction of final state
775 particles with the detector. Simulated events are then reconstructed using the same
776 reconstruction version used for real data. The L0 hardware trigger is emulated in the
777 simulation while for the software stage, Hlt1 and Hlt2 (see Sec. ??) the same code
778 can be used as done for real data. Events are simulated using both 2011 and
779 2012 conditions in the same amount in which data is available. It is important that
780 the simulation gives an accurate description of the data especially for quantitative
781 estimations, as the extraction of efficiencies. While the simulation gives a generally
782 good description of data some discrepancies remain. The next sections describe
783 corrections applied to the simulation in order to have a better description of data.
784 In App. B data distributions are compared with simulated ones.

785 3.3.1 Decay Model

786 As little is known about Λ_b^0 decays structure, the simulation software generates
787 events according to phase space decay. To include a reasonably realistic q^2 depen-
788 dence, the simulation is weighted using decays amplitudes based on the predictions
789 in Ref. [67]. Equations in this paper are for case of unpolarised production and
790 we extend those to include polarisation. Details of this are in Appendix A.2. The
791 value of the Λ_b^0 production polarisation used in the calculations is of $P_b = 0.06$ as
792 measured by LHCb [68]. Fig. 7.1 shows the phase space q^2 distribution and the
793 one obtained re-weighting the events. This can be qualitatively compared to the
794 q^2 spectrum of a generic $b \rightarrow s\ell\ell$ decay reported in Fig. 1.8 For the normalisation
795 mode, the decay model used is described in Appendix A.1, with amplitude magni-
796 tudes and production polarisation taken from the measurement in Ref. [68]. Phases
797 are not yet measured and are all set to zero.

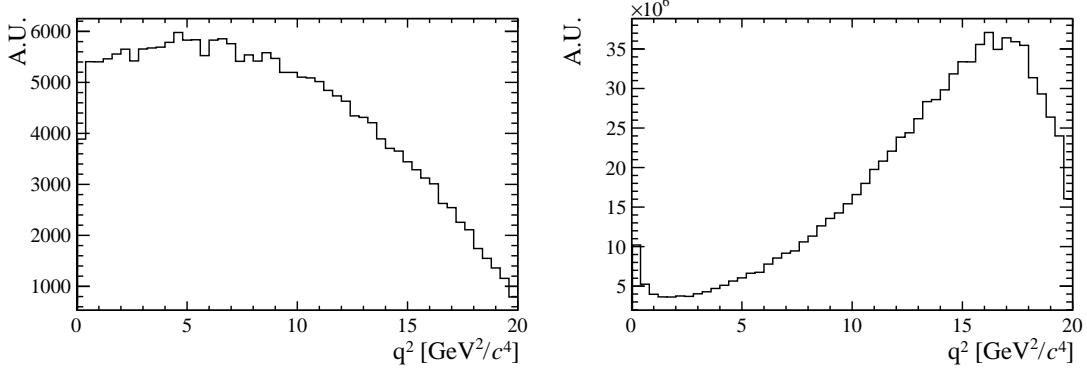


Figure 3.2: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulates events according to the phase space of the decay (left) and reweighted using the decay amplitudes (right).

3.3.2 Kinematic re-weighting

Small data-simulation differences are found in the kinematic properties of the mother particle, Λ_b^0 , which then affect also the final state particles. The Monte Carlo is re-weighted by comparing the two-dimensional momentum and transverse momentum of Λ_b^0 and Λ between real and simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ which passed pre-selection. To work with a data sample as clean as possible, a narrow interval around J/ψ and Λ_b^0 peaks is selected. Then the Λ_b^0 invariant mass is used to extract the amount of background under the peak. The background fraction with respect to the signal, f_b , is then used to statistically subtract the background from the kinematical distributions as described by the following 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 ($5605 < m(p\pi\mu^+\mu^-) < 5635 \text{ MeV}/c^2$) and $B(p, p_T)$ is the pure background distribution obtained using events from the upper sideband ($m(p\pi\mu^+\mu^-) > 5800 \text{ MeV}/c^2$).

After obtaining the signal distributions from data this is compared with $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulated events and a weight, $w(p_{\Lambda_b^0}, p_{T\Lambda_b^0})$ is defined by taking the ratio of the two

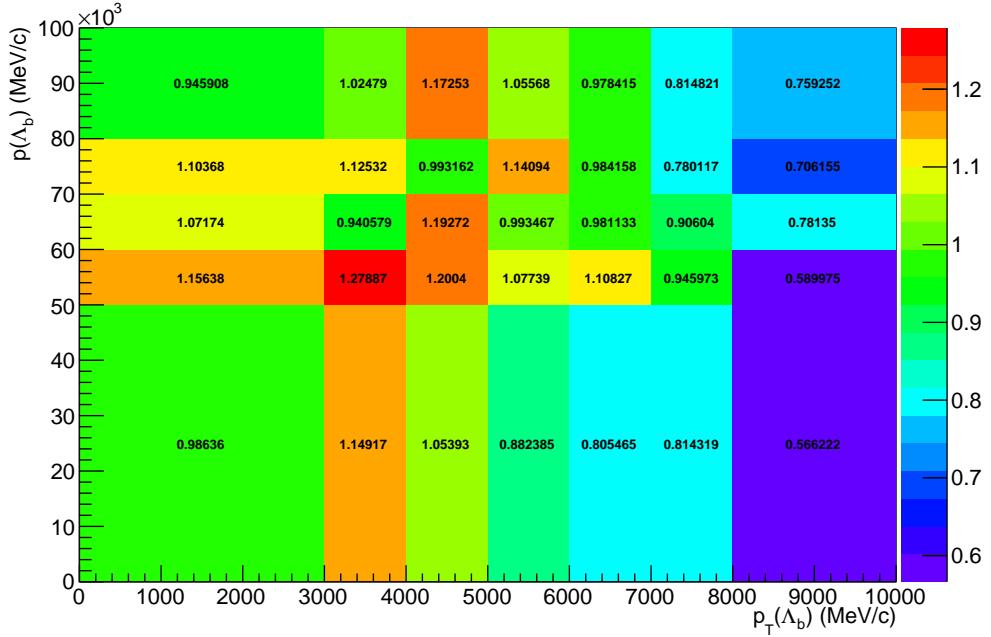


Figure 3.3: Weights used for the kinematical reweighting as a function of the momentum and transverse momentum of Λ_b^0 .

dimensional distributions. The result is shown in Fig. 3.3. In appendix B are shown distributions of sideband subtracted data in the signal and sideband regions and weighted and unweighted Monte Carlo events. In these plots the Λ_b^0 distributions match by construction but the re-weighting improves also the agreement between the transverse momentum 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 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 from downstream tracks. The fraction of downstream and long tracks

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

834

CHAPTER 4

835

836

Selection

837

838 The reconstruction of $\Lambda_b^0 \rightarrow \Lambda$ mm and $\Lambda_b^0 \rightarrow J/\psi$ Λ candidates begins with the
 839 application of requirements on basic kinematic properties as p_T of the final particles
 840 and quality requirements for the track and vertices from which the particles are
 841 derived. This selection aims to first form a dimuon candidate from two oppositely
 842 charged muons. Then, in events containing a dimuon candidate, two oppositely
 843 charged tracks are combined together and retained as Λ candidate, if they form a
 844 good vertex which is well separated from all primary vertices. In the final step,
 845 dimuon candidates are combined with Λ candidates to form Λ_b^0 baryons and require-
 846 ments are set on the properties of this combination. The full list of cuts is reported
 847 in Tab. 4.1.

848 In the table χ_{IP}^2 is defined as the projected distance from the vertex divided by its
 849 uncertainty, for example $B^0 \chi_{\text{IP}}^2(\text{primary}) > 4$ means that the B^0 vertex is at least
 850 2 standard deviations away from the primary vertex. Another useful quantity to
 851 remove combinatorial background is a pointing variable defined as the angle between

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

Table 4.1: Summary of stripping requirements. Where two values are given, the main one applies to long candidates and the one in parenthesis to downstream candidates.

the direction of the particle momentum and the flight direction from its mother vertex, called DIRA. Graphical representation of the χ^2_{IP} and DIRA variables are shown in Fig. 4.1. This allows the selection of particles with well-defined primary vertices. The $\chi^2_{\text{trk}}/\text{ndf}$ and $\chi^2_{\text{vtx}}/\text{ndf}$ quantities are the χ^2 from the fit to tracks and vertices, which are used to quantify their quality. The `GhostProb` quantity describes the probability of a track being fake. By construction, cutting at 0.4, removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. The `hasRich`, `hasCalo` and `isMuon` variables are binary indicators that the information from the RICH/calorimeter/muon detector is available for the track. Loose Particle Identification (PID) cuts are also applied in pre-selection to limit the size of the samples. To quantify the probability of particular particle identity a combined likelihood is calculated combining information from the calorimeters, the RICH and the Muon detectors. The pion hypothesis is used as a reference point and the probability of a specific ID is given in terms of the difference between the Log-Likelihood of the given hypothesis and the pion. This variable is called is called Delta Log-Likelihood (DLL) and denoted with `PID`. For example:

$$\text{PID}_K = \text{DLL}_{K-\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi) \quad (4.1)$$

quantifies the probability of a particle being a kaon rather than a pion. A large mass window is kept around the Λ_b^0 peak in order to be able to fit the sideband, to train the multivariate analysis and better constrain backgrounds. Rare candidates are simply selected by the q^2 region requirements as described in Sec. 3.1 while resonant candidates are further constrained to have dimuon invariant mass in a 100 MeV/ c^2 interval around the known J/ψ mass.

4.1 Neural Networks

The final selection is performed using neural network (NN) based on NeuroBayes package [49, 50]. The input to the neural network consists of 14 variables carrying information about the kinematics of the decay, the quality of tracks and vertices

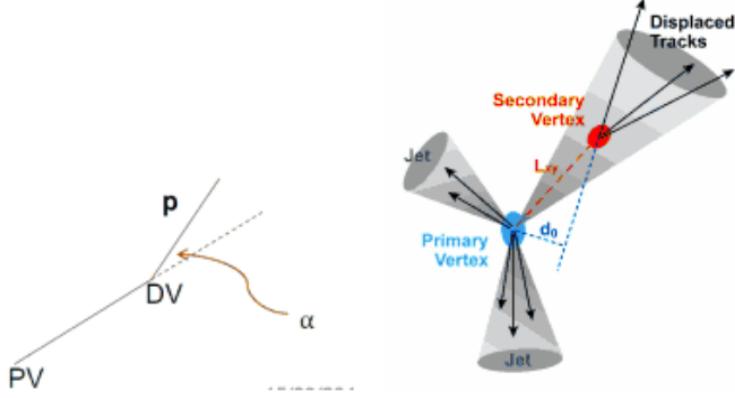


Figure 4.1: Graphical representation of the DIRA (left) and χ^2_{IP} (right) variables.

and the PID of the muons. The list of inputs to the neural network is in table 4.2. Together with it we give ranking and information on the importance of inputs. Variables related to Λ and its daughters are considered as different inputs depending on the candidate type (long or downstream). This effectively corresponds to making a separate training for the two categories. Further details on the definition and calculation of the variables importance is available in Ref. [69]. The graphical representation of the correlation matrix is shown in Fig. 4.2, where the variable with $ID = 1$ is the NN output and the IDs of the other variables can be found in Tab.4.2.

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

The NN is trained using representative samples for signal and background. For the signal a sample of simulated $\Lambda_b^0 \rightarrow \Lambda$ mm events is used. For the background a representative sample is given by candidates in the upper $m(K\pi\mu\mu)$ invariant mass sideband. Only the upper sideband, $m_{K\pi\mu\mu} > 6 \text{ GeV}/c^2$, is used since it contains only combinatorial background, while the lower sideband may contain partially re-

897 constructed and misreconstructed events. In the background samples are still present
898 J/ψ and $\psi(2S)$ peaks indicating that charmonium resonances can be combined with
899 other random tracks. These candidates do not give a good description of purely
900 combinatorial background and, in order to avoid biases, they are removed from the
901 background training sample by rejecting events in a $100 \text{ MeV}/c^2$ interval around the
902 nominal J/ψ and $\psi(2S)$ masses [3]. For the signal, the training is done combining
903 simulated $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ events corresponding to beam conditions in the two years
904 respectively. A total of 30000 total events was used for the taining. This corresponds
905 $\sim 50\%$ of the available sideband data sample and $\sim 20\%$ of our full MC sample.
906 The full simulated sample is not used as it will also be used to study efficiencies.
907 Events are uniformly samples over full simulated sample.

908 Figure 4.3 shows distributions of neural network output for the signal and back-
909 ground samples. and purity ($P = N(\text{signal})/N(\text{background})$) as a function of neural
910 network output. On this plot distributions from test samples are also overlaid in
911 order to check for overtraining. The distributions follow the same shape but with
912 different fluctuations so we conclude that we have no significant overtraining. In gen-
913 eral we conclude that the neural network is able to separate signal from background
914 and that the training converged properly.

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

Table 4.2: Summary of inputs to the neural network in order of importance. Under “Id” the indices used for the correlation matrix (see Fig. 4.2) are reported. 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.

Input	Id	adds	only this	loss
$\Lambda_{DD} p_T$	15	143.11	143.11	29.20
χ^2 DecayTreeFitter	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
Λ_b^0 DIRA	4	12.28	73.96	9.98
Λ_{DD} 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
Λ_{LL} 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

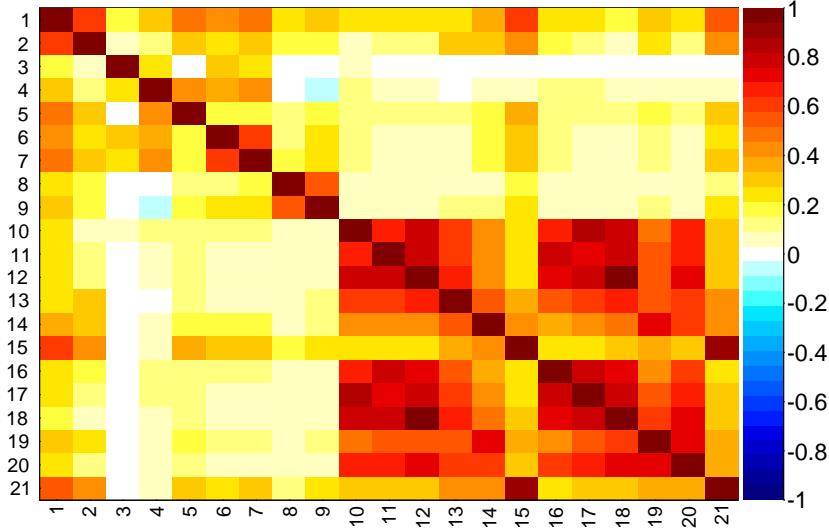


Figure 4.2: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of table ???.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

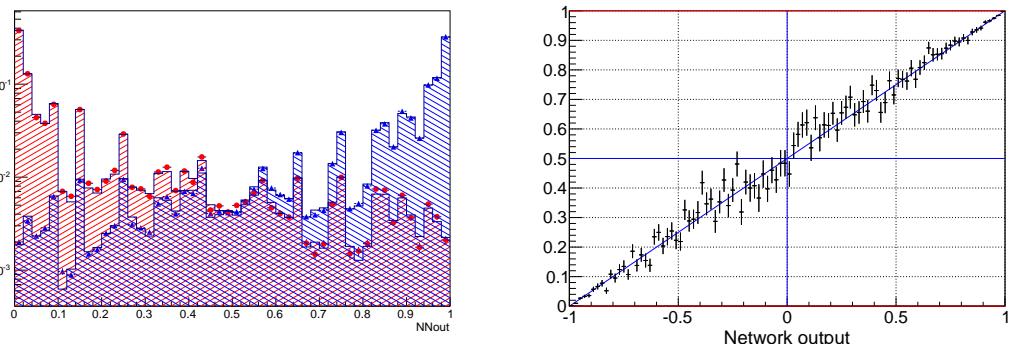


Figure 4.3: (right) NN output distribution for training (solid) and test (stripes) samples, for MC signal and sideband data events. (left) Purity as a function of neural network output.

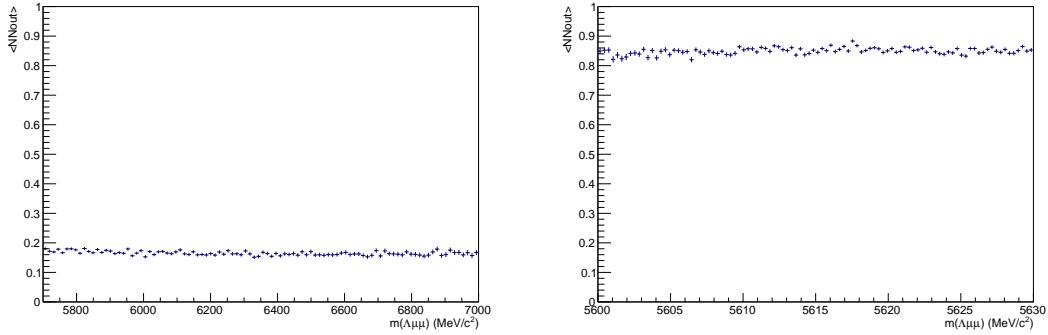


Figure 4.4: Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and MC signal (right) events.

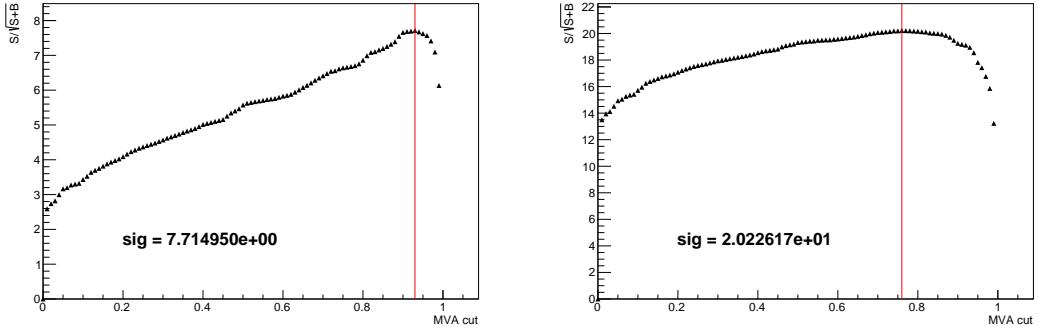


Figure 4.5: Dependence of figure-of-merit on the requirement on neural network output in the low q^2 region (left) and high q^2 (right) regions. The vertical line corresponds to the chosen cut.

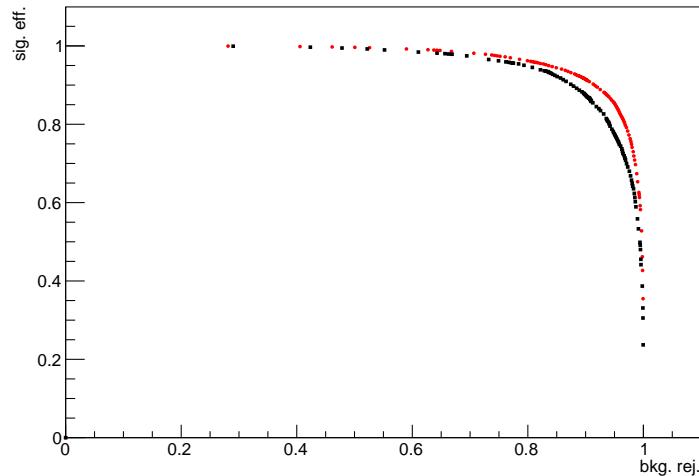


Figure 4.6: Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection.

922 4.2 MVA optimization

923 In the high q^2 region, where the signal is already observed, the final requirement
 924 on the neural network output is chosen in order to maximise the significance,
 925 $N_S/\sqrt{N_S + N_B}$, where N_S is number of expected signal candidates and N_B the num-
 926 ber of expected background candidates. N_S is derived from simulation but, as an
 927 arbitrary number of events can be generated, it needs to be normalised. To do this,
 928 the invariant mass distribution of real $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates is fit after preselection
 929 (including all requirements but MVA). This is possible as the peak of the resonant
 930 channel is already well visible after preselection. Then the resonant yield is scaled
 931 by the ratio of between the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ branching fractions as
 932 measured by LHCb on 2011 data

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) = 1.54 \times 10^{-3} \quad (4.2)$$

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

$$N_S = N_{J/\psi} \cdot \frac{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}. \quad (4.3)$$

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

936 In the low q^2 region, where the signal is unobserved, the so called “Punzi figure
 937 of merit”, $N_S/(n_\sigma/2 + \sqrt{N_B})$, is maximised [70]. This figure-of-merit is considered
 938 to be optimal for discovery and the parameter with n_σ corresponds to the number
 939 of expected standard deviations of significance, in this analysis $n_\sigma = 3$ is used.
 940 Moreover the Punzi shape does not depend on the relative normalisation between
 941 signal and background, which is important since the signal is still unobserved at
 942 low q^2 and existing predictions vary significantly for this region. The dependence
 943 of the figure-of-merit for both q^2 regions are shown in Fig. 4.5, and curves of signal
 944 efficiency versus background rejection are shown in Fig. 4.6.

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
	Hlt1DiMuonHighMass
	Hlt1TrackMuon
Hlt2	Hlt2Topo[2-4]BodyBBDT
	Hlt2TopoMu[2-4]BodyBBDT
	Hlt2SingleMuon
	Hlt2DiMuonDetached

945 For final selection the neural network output is required to be larger than 0.81 for
 946 high q^2 region and 0.96 for the low q^2 one. Using these requirements the neural
 947 network retains approximately 96% (66 %) of downstream candidates and 97 %
 948 (82 %) of long candidates for the selection at high (low) q^2 , with respect to the
 949 preselected event sample. After the full selection $\sim 0.5\%$ of the events contain
 950 multiple candidates which are randomly rejected to keep only one candidate per
 951 event. To normalise the branching ratio measurement J/ψ events are selected using
 952 the low and high q^2 selection to normalise respectively low and high q^2 intervals.

953 4.3 Trigger

954 In addition specific trigger lines are selected, corresponding to events triggered by the
 955 muons of the reconstructed candidate. This is denoted as Trigger On Signal (TOS).
 956 The trigger lines used in the analysis are shown in Tab. 4.3. The logical *or* of lines
 957 on the same lever is required and the logical *and* and lined in different levels. The
 958 L0Muon trigger required hits in the muon detector. The Hlt1TrackAllL0 performs
 959 a partial reconstruction of the events end triggers if the L0 decision is confirmed.
 960 Finally, at the Hlt2 level, a complete reconstruction is done ana a multivariate
 961 analysis is used to identify decay structures. In Fig. 4.7 is shown the single trigger
 962 efficiency, defined as if each line was alone. More information on trigger simulation

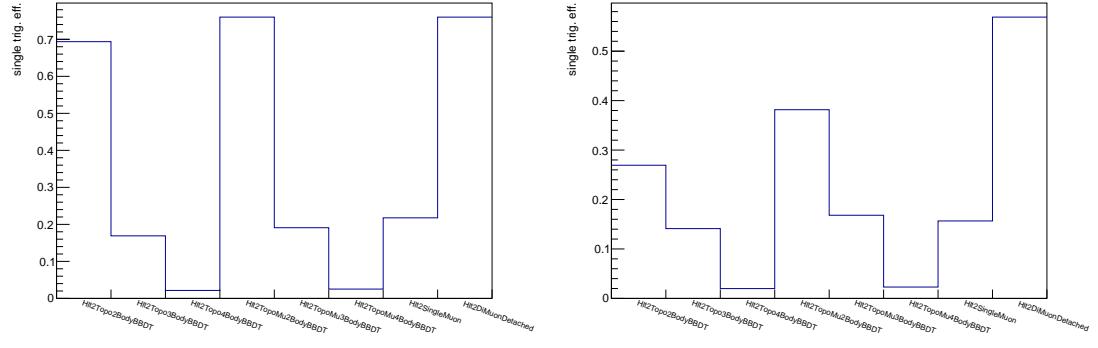


Figure 4.7: Single trigger efficiency for high q^2 events (left) and low q^2 (right).

and efficiencies are contained in Sec. 6.3.

4.4 Background from specific decays

A survey of possible peaking backgrounds concluded that the only physics background to take into account is coming from misreconstructed decays of B^0 to K_s^0 with two muons, whether via J/ψ or not. The lack of background from other decays is mainly due to the particular topology of the Λ decay which has a displaced vertex. In order to study the effect of misreconstructed $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays simulated samples are used, where the K_s^0 is reconstructed as a Λ with a $p \rightarrow \pi$ identity swap and $m(p\pi)$ in the Λ mass window. On data the $B^0 \rightarrow J/\psi K_s^0$ contribution is clearly visible in the resonant channel mass distribution. This background is not suppressed with specific cuts in this analysis as its mass shape is sufficiently distinct from Λ_b^0 signal, which allows to reliably model its contribution in the mass fits (see Sec. 5.1). For rare case a rough estimate of the size is made using the yield in the resonant channel rescaled the measured ratios between the rare and resonant branching ratios. Details are given in Sec. 5.1 and numbers of events predicted are reported in Tab. 5.1. This contribution, although close to negligible is again considered in the fit. A possible pollution due to $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ decays, where the K^{*+} further decays into $K_s^0 \pi$ is also investigated using a dedicated Monte Carlo sample and found to be negligible. Finally, $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events radiating

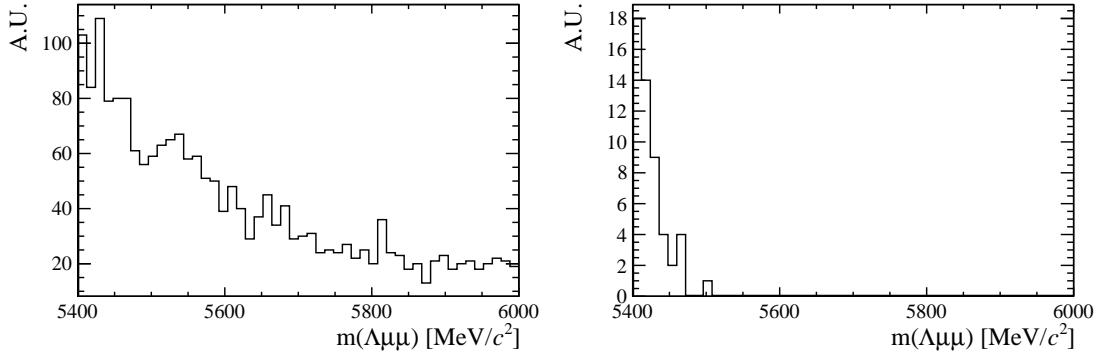


Figure 4.8: Invariant mass distributions of simulated $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ (left) and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ (right) candidates passing the full selection. Only $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed in $q^2 < 8 \text{ GeV}^2/c^4$ are selected. Distributions are shown in the invariant mass range relevant for the analysis (see Sec. 5.1).

photons from the final state, can escape the J/ψ veto and be reconstructed in the rare channel. Analysing simulated events it was found that the only contribution is in the closest q^2 interval to the J/ψ tail, $6 < q^2 < 8 \text{ GeV}^2/c^4$. In this interval 1.3% of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates are reconstructed but only 0.06% falls 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 in the fit mass window this background is considered as absorbed in the combinatorial background. In Fig. 4.8 is reported the invariant mass distribution of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events falling into the rare q^2 region.

992

CHAPTER 5

993

994

Yield extraction

995

5.1 Fit description

997 To extract the yields of the rare and resonant channels, an extended unbinned
998 maximum likelihood fits are used. The likelihood has the form:

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

999 where N_S , N_C and N_B are number of signal, combinatorial and K_s^0 background events
1000 and $P_i(m_i)$ is the corresponding probability density function (PDF). From now on
1001 when we refer to the invariant mass of the Λ_b^0 system we use the value obtained
1002 from a kinematical fit of the full decay chain in which each particle is constrained
1003 to point to its assigned origin vertex and the invariant mass of the $p\pi$ system is

1004 constrained to be equal to the world average Λ mass. In the resonant channel case
1005 a further constrain is used on the dimuon mass to be equal to the known J/ψ mass.
1006 This method allows to improve the mass resolution giving better defined peaks and
1007 therefore a more stable fit.

1008 For the resonant channel the signal is described as a sum of two Crystal Ball functions
1009 (CB) with common mean (m_0) and tail slope (n). A Crystal Ball function [71] is
1010 a probability density function commonly used to model various processes involving
1011 energy loss. In particular it is used to model the radiative tail which can be seen in
1012 many resonances' peaks. This function consists of a Gaussian core and a power-law
1013 tail, below a certain threshold. The function itself and its first derivative are both
1014 continuous and has form

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

1015 where for normalisation and continuity

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

1016 The full form of the PDF for the resonant channel is therefore:

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

1017 where f is the relative fraction of candidates falling into the first CB function.

1018 As a first step simulated $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distributions are fitted
1019 using the signal PDF separately for long and downstream candidates. Figure 5.1
1020 shows simulated distributions of resonant events with the fit function overlaid.

1021 In a second step the fit to the resonant channel data sample is performed. For the fit
1022 on data the tail slope parameter, “ n ”, which is highly correlated with the α s, is fixed it
1023 to the value found in the fit to simulated data. In this fit two background components

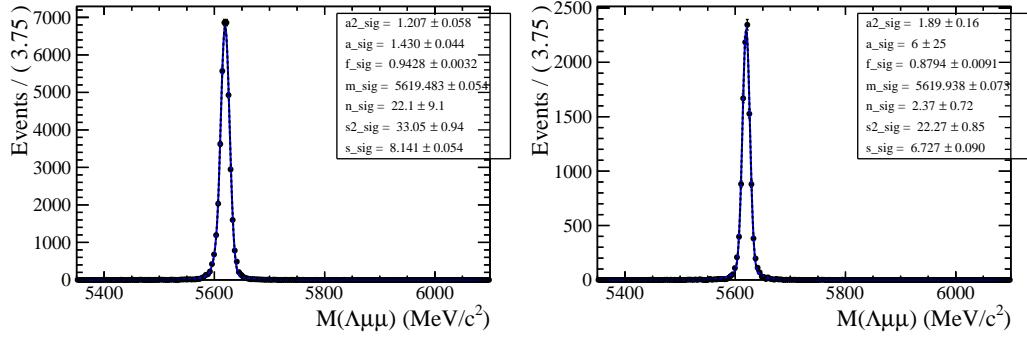


Figure 5.1: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows simulated data and the blue line is the signal fit function.

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 Double Crystal Ball 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 final fit all parameters are fixed. This is particularly important when fitting long-long events, 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 $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ 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 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

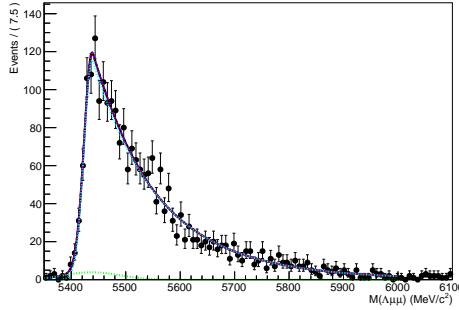


Figure 5.2: Invariant mass distribution of simulated misreconstructed $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted with Double Crystal Ball functions.

to constrain many parameters, especially when dividing data in q^2 bins. 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 for example 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 on data by fitting a $\Lambda \mu^+ \mu^-$ simulated sample in each q^2 bin and comparing its 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 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 parameters are never shared between DD and LL fits.

The background components modelled are also in this case the combinatorial background, described with an exponential function. 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 due to the kinematical limit at low masses. The exponential slopes are therefore left floating independently in each q^2 bins and

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

also independently of the resonant channel and for the in DD and LL 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, in the default fit, 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 . In Tab. 5.1 is reported the number of predicted $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in each q^2 bin 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 events in a q^2 bin 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

The fit on the rare sample is performed simultaneously on the LL and DD candidate categories. Therefore the two separate yields are not separately floating but are but are parameterised ad a function of the branching ratio with the following formula:

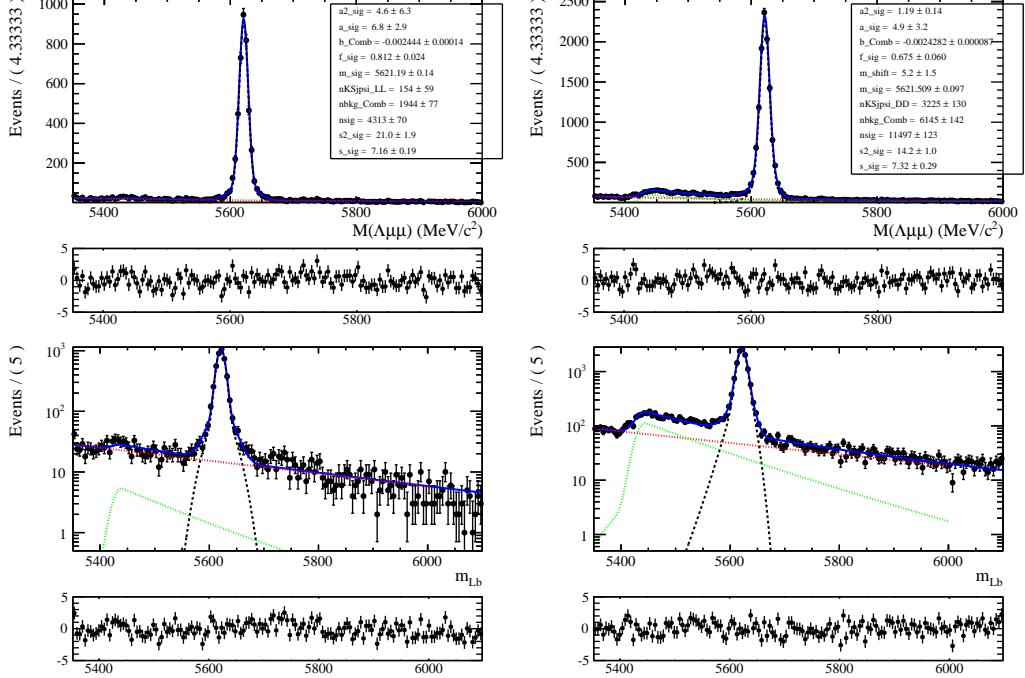


Figure 5.3: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and residuals of the fit for long-long (left) and down-down (right) events. Lower plots are the same as the upper ones but shown in linear scale. The histogram shows data. On the plot are shown the total fit (blue line), the signal (black dashed), the combinatorial background (red dashed) and the $B^0 \rightarrow K_s \mu^+ \mu^-$ background (green dashed).

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

where $k = \text{LL,DD}$, Δq^2 is width of the q^2 bin and the only free parameter is the branching fraction ratio rare over J/ψ . For the $J/\psi \rightarrow \mu^+\mu^-$ the value reported in the PDG book [3], $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$. In this formula the efficiencies and the normalisation yield appear as constants. These constants are then varied in order to obtain systematics on the final result as described in Sec. 7.

Table 5.2: Number of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays in the long and downstream categories found using the selection for low- and high- q^2 regions. Uncertainties shown are statistical only.

Selection	N_S (long)	N_S (downstream)
high- q^2	4313 ± 70	$11\,497 \pm 123$
low- q^2	3363 ± 59	7225 ± 89

1085 **5.2 Fit results**

1086 In Fig. 5.3 are shown fitted invariant mass distributions for the normalisation chan-
1087 nel, selected with the high q^2 requirements and in Fig. 5.4 for low q^2 requirements.
1088 Table 5.2 reports measured yields of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates found using the low and
1089 high q^2 selections. Values for the signal shape parameters are shown on Fig. 5.3.
1090 Results of the fit to the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ sample are shown in Fig. ?? for the
1091 integrated $15 < q^2 < 20$ GeV^2/c^4 interval and in Fig. 5.6 for the $1.1 < q^2 < 6.0$
1092 GeV^2/c^4 one. The exponential slopes, the scale factors multiplied to the widths and
1093 the number of combinatorial events found from these fits are reported in Tab. ??.
1094 Fitted invariant mass distribution in all other considered q^2 intervals are in Fig. ??
1095 for downstream candidates and Fig. ?? for long candidates together with their sig-
1096 nificances. The yields of rare events obtained from the fit are reported in Tab. 5.3.
1097 Most candidates are found in the downstream sample comprising $\sim 80\%$ of the
1098 total yield. Notice that, since the fit is simultanous on DD and LL candidates, the
1099 yields in the two categories yields are not parameters free to float independently in
1100 the fit but are correlated via the branching ratio. The statistical significance of the
1101 observed signal yields is evaluated as $\sqrt{2\Delta \ln \mathcal{L}}$, where $\Delta \ln \mathcal{L}$ is the change in the
1102 logarithm of the likelihood function when the signal component is excluded from
1103 the fit, relative to the nominal fit in which it is present.

Table 5.3: Signal decay 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

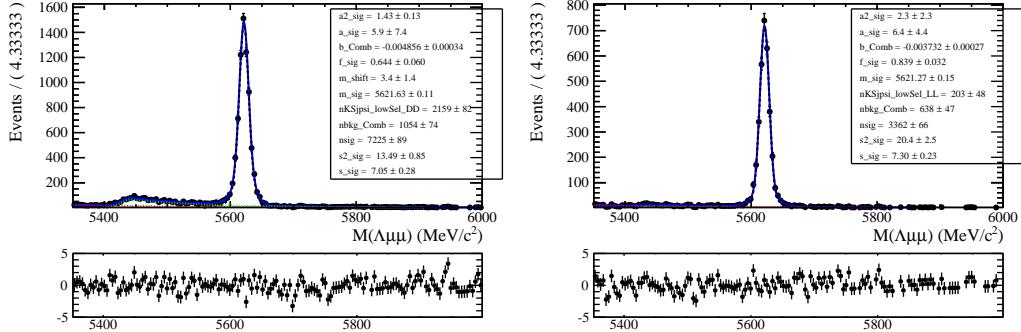


Figure 5.4: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows real data selected with low q^2 selection. The blue line is the total fit function.

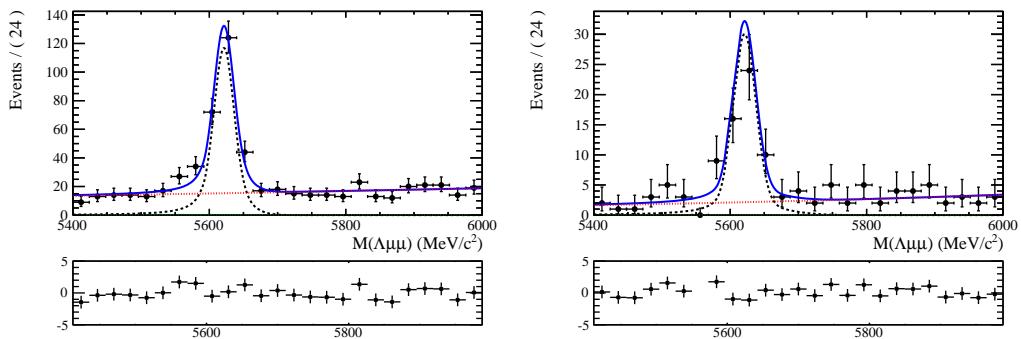


Figure 5.5: Invariant mass distribution of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ for downstream (left) and long (right) candidates in the integrated $15\text{--}20$ GeV^2/c^4 q^2 interval. The points show data, the blue line the total fit function and the dashed red line represents combinatorial background.

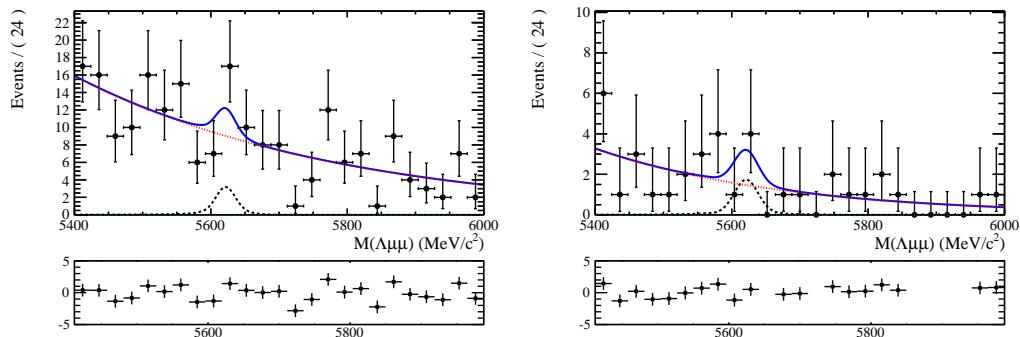


Figure 5.6: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the integrated 0.1 – 6.0 $\text{GeV}^2/c^4 q^2$ interval for downstream (left) and long (right) candidates. The points show data, the blue line the total fit function and the dashed red line represents combinatorial background.

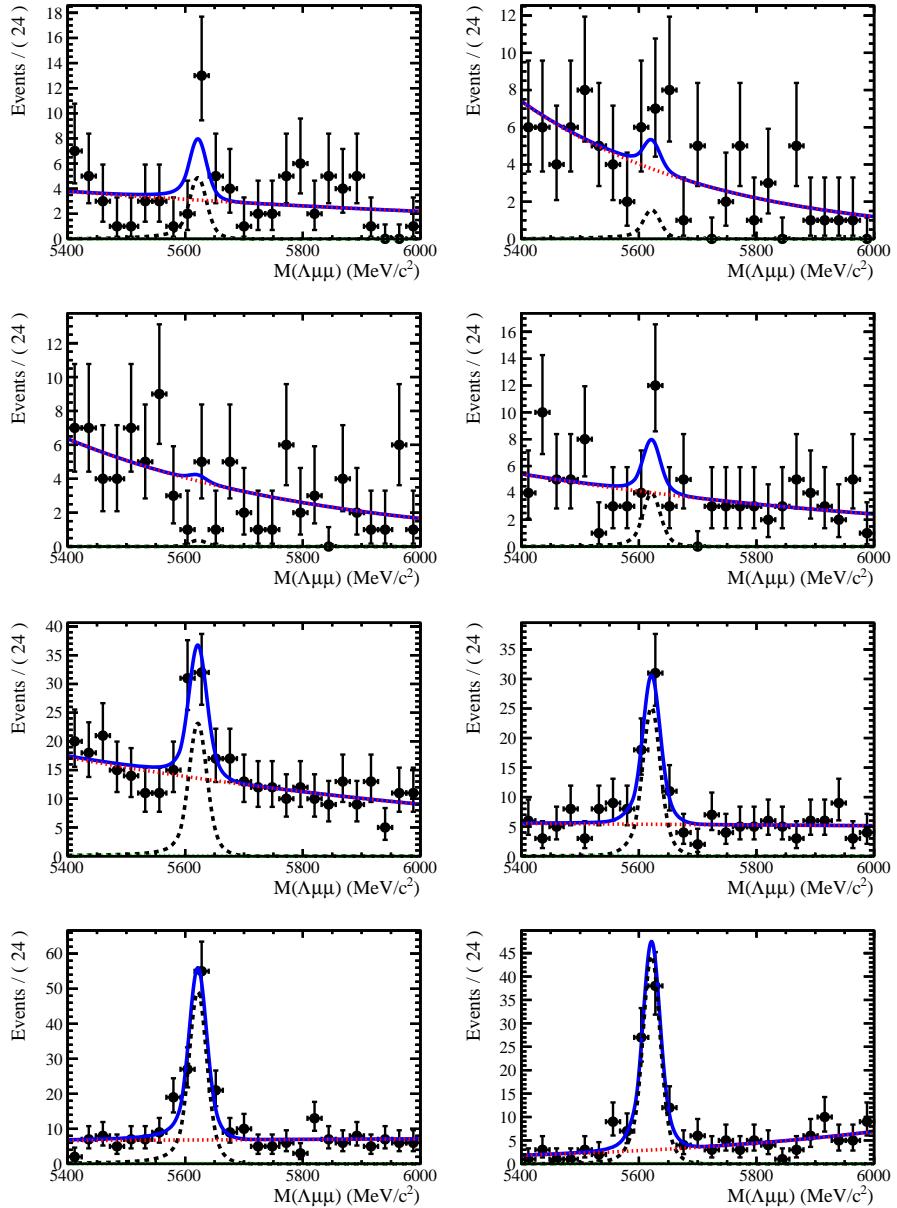


Figure 5.7: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for downstream candidates.

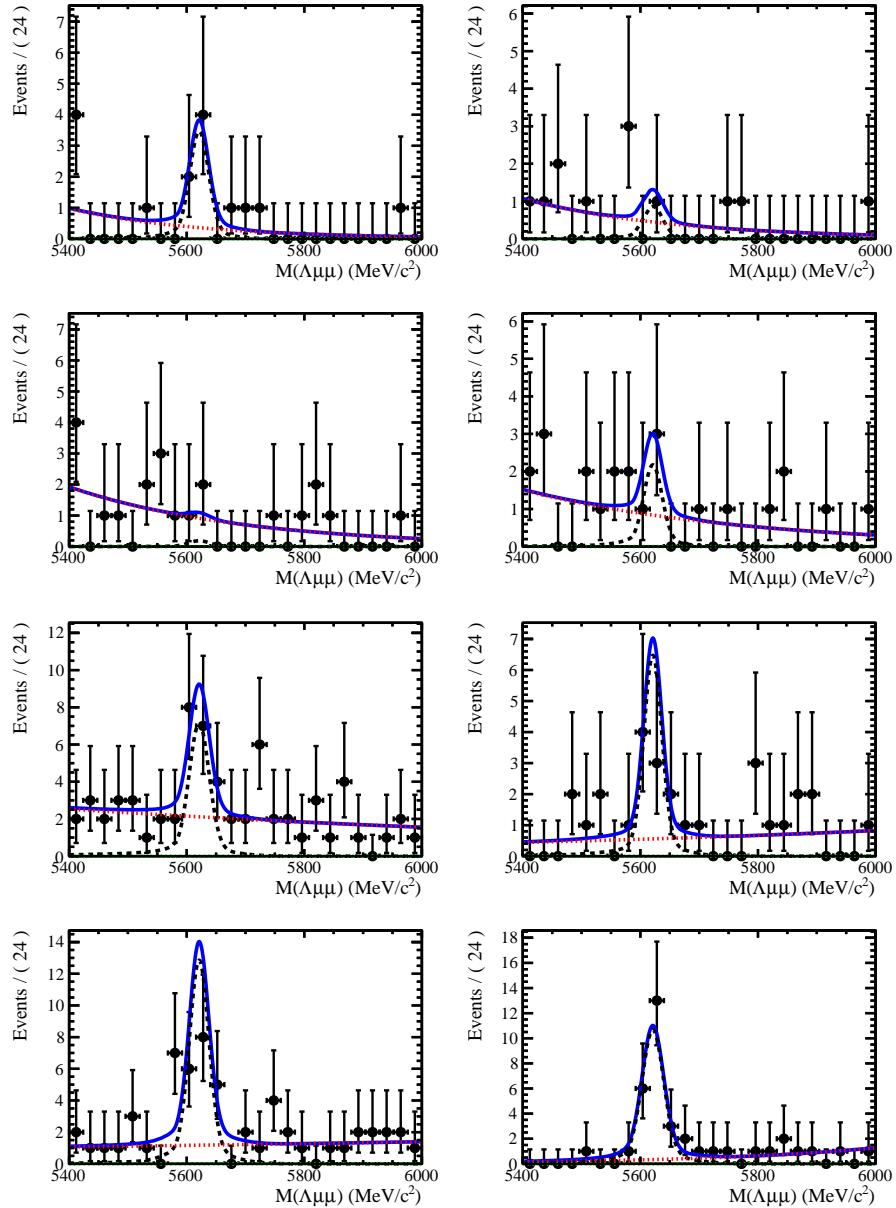


Figure 5.8: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for long candidates.

1104

CHAPTER 6

1105

1106

Efficiency

1107

1108 The efficiency for each of the decays is calculated according to the formula

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

1109 In this expression the first term gives the efficiency to have final state particles in the
1110 LHCb acceptance. The second term handles the possibility of Λ escaping the detec-
1111 tor or interacting with it and therefore never decaying into $p\pi$. This term is referred
1112 to as “Detection” efficiency. The third term carries information about the recon-
1113 struction and stripping efficiency which keep these together given that boundaries
1114 between them are completely artificial. The fourth part deals with the efficiency
1115 of the Neural Network for those events which passed the pre-selection . Finally,
1116 the last term handles the trigger efficiency. Most of the efficiency components are
1117 evaluated using the simulated events described in Sec. 3.3. Only the efficiency of
1118 cut on the PID of the proton, present in the stripping, is separately derived with a

Table 6.1: Absolute geometrical acceptance in bins of q^2 from MC simulation. Errors shown 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

1119 data–driven method because the simulation does not provide a good description of
1120 PID variables. For complete information, all absolute efficiencies for the two decays
1121 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ are separately listed in the next sections. However,
1122 for the analysis itself only relative efficiency, $\epsilon(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-) / \epsilon(\Lambda_b^0 \rightarrow J/\psi \Lambda)$, is
1123 used.

1124 6.1 Geometric acceptance

1125 In order to save disk space and time, simulated samples contain only events in which
1126 the final muons are in LHCb acceptance, and therefore can be reconstructed. This
1127 corresponds to requirement for each of the muons to be in an interval $10 < \theta <$
1128 400 mrad, where θ is the angle between the muon momentum and the beam line.
1129 The efficiency of this requirement is obtained by using a separate simulated sample
1130 where events are generated in the full space. In Tab. 6.1 the efficiencies due to the
1131 geometrical acceptance are listed in bins of q^2 for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays.

Table 6.2: Absolute Detection and reconstruction plus stripping efficiencies. Reconstruction efficiency is given separately for DD and LL events. Errors shown are statistical only.

q^2 [GeV $^2/c^4$]	Detection	Reco and Strip (DD)	Reco and Strip (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

1132 6.2 Reconstruction and neural network efficiencies

1133 The efficiency to reconstruct the decays together with the stripping selection is eval-
 1134 uated from simulated data. This component does not include the efficiency of the
 1135 PID cut that appears in Tab. 4.1, which is kept separate because PID variables are
 1136 not well described by the simulation and therefore a data-driven method is used
 1137 instead (see Sec. 6.4). In Tab. 6.2 the reconstruction efficiency is reported in bins of
 1138 q^2 and for long and downstream candidates. In the table the efficiency is subdivided
 1139 in “Detection” and “Reconstruction and Stripping” efficiencies. In fact since Λ is
 1140 a long lived particle there is a non-negligible probability that it interacts in the
 1141 detector or escapes from it and therefore never decays in proton and pion. The
 1142 efficiency for this to happen is what is called “Detection” efficiency. Reconstruc-
 1143 tion and Stripping” efficiency include the efficiency of reconstructing tracks and the
 1144 efficiency for events passing the stripping cuts.

1145 The MVA efficiency is again evaluated from simulated samples. Results are shown in
 1146 Tab. 6.3 in bins of q^2 . The sudden jump in efficiency before and after ~ 9 GeV/ c^2 is
 1147 due to the fact that a different optimisation is performed for the MVA cut in the
 1148 low and high q^2 regions.

Table 6.3: Neural network selection efficiency. Errors shown are statistical only.

q^2 [GeV $^2/c^4$]	MVA eff. (DD)	MVA 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

¹¹⁴⁹ 6.3 Trigger efficiency

¹¹⁵⁰ The trigger efficiency is again calculated on a simulated sample for events which are
¹¹⁵¹ accepted by the full selection. Using the resonant channel it is possible to crosscheck
¹¹⁵² on data the efficiency obtained using the simulation with the data driven TISTOS
¹¹⁵³ method. In LHCb triggered events can fall in two categories: events triggered by
¹¹⁵⁴ a track which is part of a signal candidate, Trigger On Signal (TOS), or by other
¹¹⁵⁵ tracks in the event, Trigger Independent of Signal (TIS). All trigger lines used for this
¹¹⁵⁶ analysis are required to be TOS. As the TIS and TOS categories are not exclusive
¹¹⁵⁷ the TIS sample provides a control sample which can be used to obtain the efficiency
¹¹⁵⁸ for TOS trigger. This is calculated with the formula:

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

¹¹⁵⁹ Using the data–driven method an efficiency of $(70 \pm 5)\%$ is obtained, while this
¹¹⁶⁰ is calculated to be $(73.33 \pm 0.02)\%$ using the simulation. Results are therefore
¹¹⁶¹ 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

1162 6.4 PID efficiency

1163 For long tracks a PID cut on protons ($\text{PIDp} > -5$) is used. The simulation is known
 1164 not to describe particle ID well and therefore data-driven method is used to obtain
 1165 this efficiency component. This is done using the `PIDCalib` package, which uses
 1166 decays where particles can be identified due to their kinematic properties. In the
 1167 case of protons a sample of Λ particles was used where the proton can be identified
 1168 because it always has the highest momentum. The package allows to divide the phase
 1169 space in bins of variables relevant for PID performances, in this analysis momentum
 1170 and pseudorapidity are used. Using the calibration sample the efficiency is derived
 1171 in each two-dimensional bin. To take in account that the decay channel under
 1172 study could have different kinematical distributions than the calibration sample
 1173 these efficiency tables are used to re-weight the simulation. Absolute PID efficiencies
 1174 are shown in Tab. 6.5 and the relative efficiency over the resonant channel are
 1175 reported in Tab. 6.10.

1176 6.5 Relative efficiencies

1177 In the previous sections absolute efficiencies values were given for the rare channel in
 1178 different q^2 bins. In this section the corresponding relative efficiencies with respecte

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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$. Errors shown 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

to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ channel are reported, which will be used for the differential branching ratio calculation. In Tab. 6.6 the absolute efficiency values for the J/ψ channel are also reported. Relative geometric efficiencies are listed in Tab. 6.7, In Tabs. 6.9 and 6.8 relative reconstruction, trigger and mva efficiencies are shown separately for downstream and long candidates. Since these three components are obtained from the same simulated sample their statistical errors are correlated. Therefore the total of the three is also reported as a single efficiency and labeled “Full Selection” in the table. Finally, the relative PID efficiency is reported in Tab. 6.10. Figure 6.1 shows the values in these tables in graphical form. Finally, in Tab. 8.1 is reported the total of all relative efficiencies, which will be then used to correct the raw yields and calculate the differential branching fraction. Uncertainties reflect statistics of both rare and resonant samples, while systematic uncertainties are discussed in next chapter.

Table 6.7: Relative geometric efficiency and Detection efficiency between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays. Uncertainty reflects statistics of both samples.

q^2 [GeV $^2/c^4$]	Geometric	Detection
0.1–2.0	1.2976 ± 0.0050	0.9751 ± 0.0006
2.0–4.0	1.1541 ± 0.0043	0.9814 ± 0.0005
4.0–6.0	1.1043 ± 0.0044	0.9872 ± 0.0006
6.0–8.0	1.0778 ± 0.0045	0.9939 ± 0.0006
11.0–12.5	1.0431 ± 0.0058	1.0074 ± 0.0007
15.0–16.0	1.0426 ± 0.0084	1.0188 ± 0.0010
16.0–18.0	1.0296 ± 0.0068	1.0255 ± 0.0008
18.0–20.0	1.0288 ± 0.0087	1.0333 ± 0.0010
1.1–6.0	1.1396 ± 0.0031	0.9835 ± 0.0004
15.0–20.0	1.0320 ± 0.0048	1.0269 ± 0.0006

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. Uncertainty reflects 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. Uncertainty reflects 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

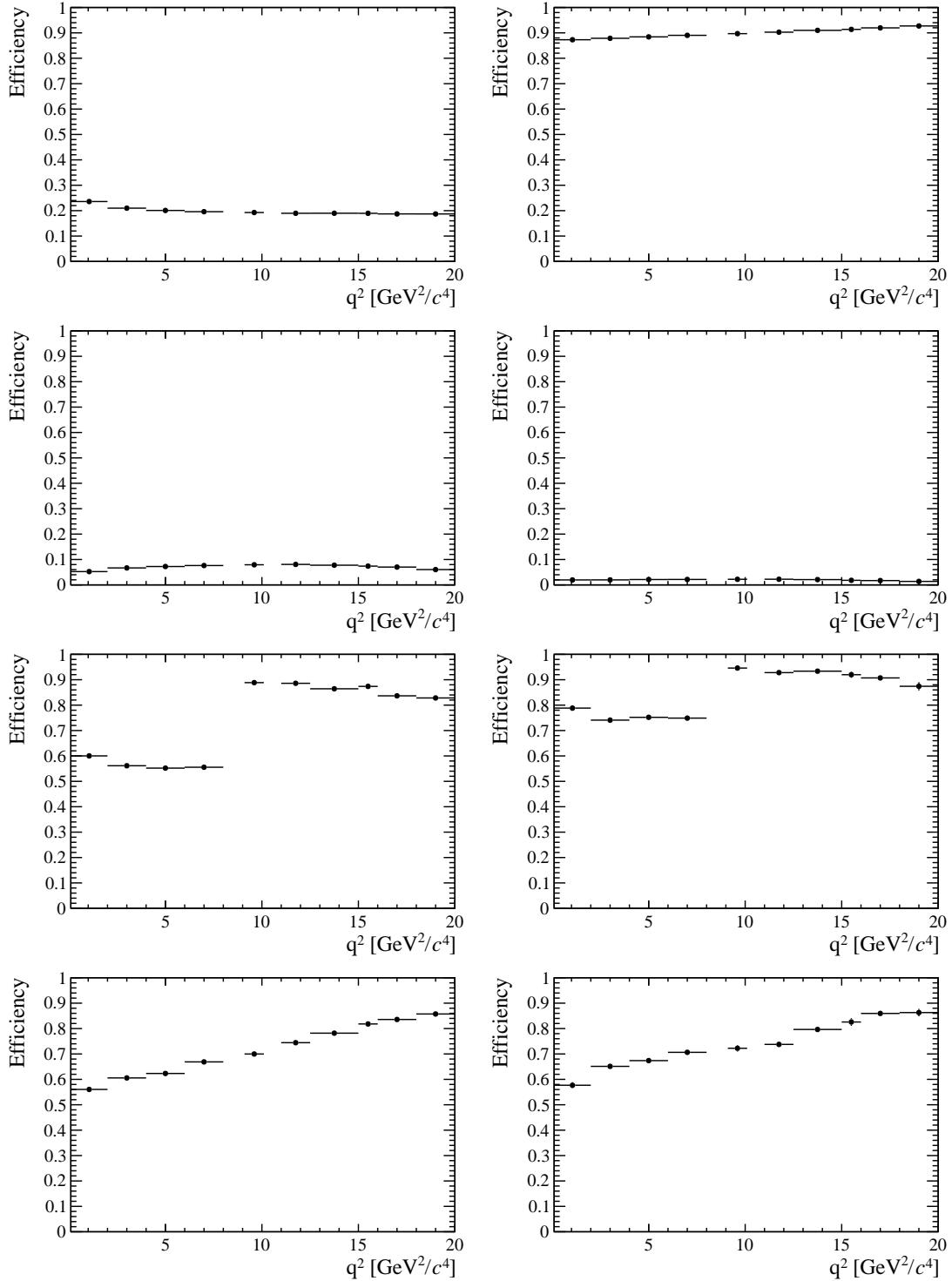


Figure 6.1: Relative efficiencies as a function of q^2 . geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) events, trigger efficiency for DD (e) and LL (f) and MVA efficiency for DD (g) and LL (h).

Table 6.10: Relative PID efficiencies in q^2 bins

q^2 [GeV $^2/c^4$]	Rel. PID Eff.
0.1–2.0	0.99418 ± 0.00013
2.0–4.0	0.99523 ± 0.00013
4.0–6.0	0.99699 ± 0.00012
6.0–8.0	0.99805 ± 0.00011
11.0–12.5	1.00151 ± 0.00010
15.0–16.0	1.00431 ± 0.00008
16.0–18.0	1.00215 ± 0.00008
518.0–20.0	1.00226 ± 0.00005
1.1–6.0	0.99589 ± 0.00009
15.0–20.0	1.00281 ± 0.00006

1192

CHAPTER 7

1193

1194

Systematic uncertainties

1195

1196 7.1 Yields

1197 The choice of a specific PDF to model the invariant mass distribution could result in
1198 a bias. In order to assess the effect of the signal PDF choice a number of models are
1199 tried on the $A_b^0 \rightarrow J/\psi A$ data sample in order to understand which ones are plausible.
1200 In Tab. 7.1 are reported the χ^2 relative probabilities obtained using different models
1201 including: the default model, a Double Crystal Ball function, a simple Gaussian
1202 function, a simple Crystal Ball function and the sum of two Gaussians. The only
1203 two models that give a reasonable p-value are the default DCB and the sum of two
1204 Gaussian functions (DG). As a second step simulated experiments are generated and
1205 fit the two chosen functions. Events are generated according to a density function
1206 given by the default model fitted on data separately for each q^2 interval. In this
1207 way, for each q^2 interval, a specific shape is reproduced including the background

1208 level and slope. Furthermore, a number of events comparable to the one found in
1209 data is generated. For each experiment a per cent bias is calculated as

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

1210 where $N_{\ell\ell}^{model}$ and $N_{J/\psi}^{model}$ are the numbers of rare and resonant events observed
1211 using a specific model. The distribution of biases have approximately gaussian
1212 shape. Finally, the average bias over 1000 pseudo-experiments is taken as systematic
1213 uncertainty. Notice that in each case the rare and normalisation channels are fit with
1214 the same signal model and, while for the default case the rare parameters are fixed
1215 to what found for the resonant channel, they are left free to float with the second
1216 model in order to asses at the same time the systematic due to the parameters
1217 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

1218

1219 For the background PDF systematic the rare channel is refit leaving the yield of
1220 K_s^0 component floating, which is fixed to the predicted value in the default fit. The
1221 same procedure as for the signal PDF is applied. Results are reported in Tab. 7.2.
1222 The most affected bin is the one in the middle of the charmonium resonances, where
1223 a combination of lower statistics and higher background leaves more freedom to the
1224 signal shape. Finally, a background component for $B^+ \rightarrow K^{*+}(K_s^0 \pi^+) \mu^+ \mu^-$ decays
1225 is added in the fit, modelled using the distribution of simulated events after full
1226 selection. No significant bias is found for this component.

1227

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 systematics due to the choice of signal and background shapes in bins of q^2 .

1228 7.2 Efficiencies

1229 Systematic uncertainties in the efficiency determination are due to limited knowl-
 1230 edge of the decay properties such as the Λ_b^0 lifetime and production polarisation.
 1231 The uncertainties are directly calculated on the relative efficiencies as these are the
 1232 ones that are actually used in the analysis. It should be noted that not all sources
 1233 contribute to each part of the efficiency. For brevity in this section are only re-
 1234 ported estimates of the systematic uncertainties obtained while the full information
 1235 is contained in App. C.

1236 7.2.1 Effect of new physics on the decay model

1237 New physics could affect the decay model modifying the Wilson Coefficients by
 1238 adding contributions to the C_7 and C_9 coefficients. This would result in a modifi-
 1239 cation of the q^2 spectrum and therefore of the efficiency. To asses this systematic
 1240 Wilson Coefficients are modified by adding a NP component ($C_i \rightarrow C_i + C_i^{\text{NP}}$).
 1241 In Fig. 7.1 are reported q^2 spectra obtained weighting the simulation for a model
 1242 embedding the default and 3 modified sets of wilson coefficients. Used values used
 1243 are reported on top of each plot and are inspired to maintain compatibility with
 1244 the recent LHCb result for the P'_5 observable [27]. The biggest effect is in the very

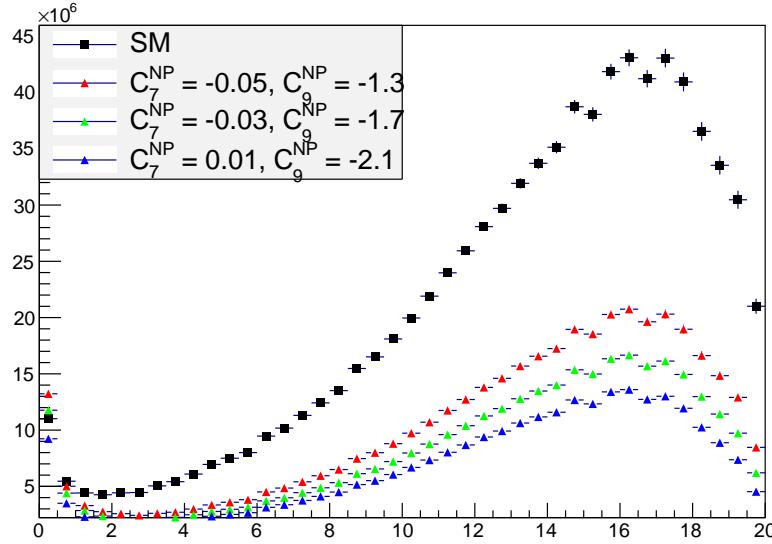


Figure 7.1: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

₁₂₄₅ low q^2 , belo 2 GeV^2/c^4 the efficiency can change up to 7%, 3-4 % between 3 and 4
₁₂₄₆ GeV^2/c^4 and 2-3 % in the resto of the spectrum. This values are given in order to
₁₂₄₇ provide the full information but are not added as systematic uncertainties. In fact
₁₂₄₈ the hypothesis of this analysis is that the decays are described by a the SM.

₁₂₄₉ 7.2.2 Simulation statistics

₁₂₅₀ The limited statistics of the simulated samples used to determine efficiencies is
₁₂₅₁ considered a source of systematic uncertainty. While it is not the dominant source
₁₂₅₂ of systematics, its size does not allow to completely neglect this uncertainty. When
₁₂₅₃ reporting relative efficiency values the statistical uncertainty due to the rare and
₁₂₅₄ resonant channels is always considered.

₁₂₅₅ 7.2.3 Production polarisation and decay structure

₁₂₅₆ One of the main unknown which affects the determination of the efficiencies is the
₁₂₅₇ angular structure of the decays. And connected to it also the production polarisa-
₁₂₅₈ tion, which is a parameter of the model. To assess the systematic uncertainty
₁₂₅₉ due to the knowledge of the production polarisation for $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays the
₁₂₆₀ polarisation parameter in the model is varied within one standard deviation of the
₁₂₆₁ most recent LHCb measurement $P = 0.06 \pm 0.09$ [68]. The full difference observed
₁₂₆₂ is taken as systematic uncertainty. To assess systematic uncertainty due to decay
₁₂₆₃ structure an alternative set of form factors is used based on lattice QCD calcula-
₁₂₆₄ tion [72]. Details of this are explained in A.2. The two models are compared and
₁₂₆₅ the full difference is taken as systematic uncertainty. In total this results in an
₁₂₆₆ uncertainty of $\sim 1.3\%$ for long candidates and $\sim 0.6\%$ for downstream candidates,
₁₂₆₇ mostly coming from the knowledge of the production polarisation.

₁₂₆₈ 7.2.4 Λ_b^0 lifetime

₁₂₆₉ The Λ_b^0 lifetime is known only with limited precision. For evaluation of the efficien-
₁₂₇₀ cies the world average value, 1.482 ps^{-1} [73] is used. To evaluate the systematic
₁₂₇₁ uncertainty, this values is varied within one standard deviation from the measured
₁₂₇₂ value. Only cases where both signal and normalisation channel are varied in same
₁₂₇₃ direction are considered. The larger difference with the default lifetime case is taken
₁₂₇₄ as systematic uncertainty, which is found to range from $\sim 0.4\%$ at low q^2 to $\sim 0.1\%$
₁₂₇₅ at high q^2 .

₁₂₇₆ 7.2.5 Downstream candidates reconstruction efficiency

₁₂₇₇ Other analysis in LHCb, using particles reconstructed with downstream tracks,
₁₂₇₈ showed that the efficiency for these events is not well simulated in the Monte Carlo.
₁₂₇₉ For example in Fig. 7.2 the ratio between the reconstruction efficiency for down-

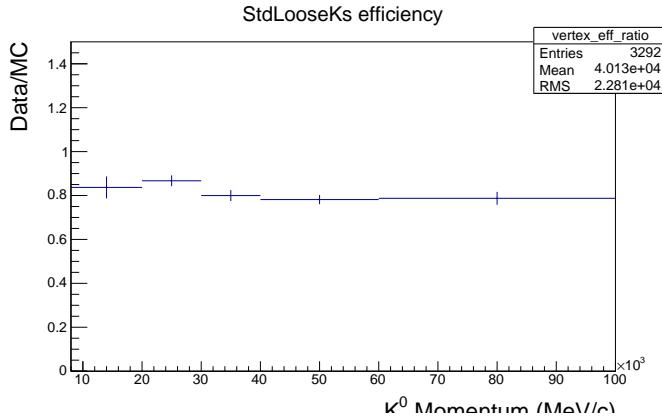


Figure 7.2: Ratio of reconstruction efficiency in Data and MC found using K_S events [74].

stream candidates in data and simulation found analysing K_S^0 events [74] is shown. This effect is not yet fully understood and is currently under study. The main effect seems to be due to a poor simulation of the vertexing efficiency for downstream tracks.

This effect is dealt with in two steps. Firstly, the analysis separately for downstream and long candidates. Since efficiencies are also calculated separately, the effect should mostly cancel in the ratio between the rare and resonant channels. In a second step a systematic uncertainty is assigned for down-down events only. To do this the simulation is re-weighted by the efficiency ratio between data and simulation found for K_S as a function of momentum and shown in Fig. 7.2. Then corrected and uncorrected efficiencies are compared and the full difference is taken as systematic uncertainty. Dependencies due to the different momentum distributions of Λ and K_S^0 are assumed to be negligible since the discrepancy shows little dependence on momentum. This results in an extra 0.4% systematic at low q^2 and 1.2 % at high q^2 , only for downstream candidates.

7.2.6 Data-simulation discrepancies

The simulation used to extract efficiency is re-weighted as described in sec.3.3.2. The influence on this procedure on the efficiencies was checked by comparing values

₁₂₉₈ obtained with and without re-weighting. The effect is negligible with respect to
₁₂₉₉ other systematics considered.

1300

CHAPTER 8

1301

1302

Differential branching ratio extraction

1303

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

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2}{\mathcal{B}(J/\psi\Lambda)} \right] \cdot N(J/\psi\Lambda)_k \cdot \epsilon_k^{\text{rel}} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}, \quad (8.1)$$

where $k = (\text{LL}, \text{DD})$, Δq^2 is width of the q^2 bin and the only free parameter is the relative branching fraction ratio. For the $J/\psi \rightarrow \mu^+\mu^-$ branching ratio the value reported in the PDG book, $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$ [3], is used. Tab. 8.1 summarises the total relative efficiencies for downstream and long candidates together with their correlated and uncorrelated errors, where the correlation is intended between the downstream and long samples. On the table the uncorrelated error corresponds to the total systematic error on the efficiency. The correlated

Table 8.1: Absolute values of the total relative efficiency and the absolute value of the uncorrelated error, together with relative values of correlated error.

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%

1315 error is given in per cent form since it can be applied to either downstream, long
1316 candidates or their combination. This includes the PDF systematic described in
1317 Sec. 7.1 and the systematic due to the uncertainty on $J/\psi \rightarrow \mu^+\mu^-$ branching ratio.

1318 In Fig. 8.1, the branching ratio obtained by fitting the downstream and long samples
1319 independently.

1320 The combined result, obtained fitting both samples simultaneously is shown in
1321 Fig. 8.2. Values are also reported are reported in Tab. 8.2, where the statistical
1322 error on the rare channel (stat) and the total systematic error (stat) are shown
1323 separately. The statistical error is calculated using the MINOS tool, which returns
1324 an asymmetric interval. The normalisation and systematic errors are evaluated by
1325 pushing the efficiencies and normalisation yields up and down by the value of their
1326 errors are re-performing the fit. The different efficiencies used have an effect on the
1327 branching ratio and the full difference with respect to the default fit is taken as
1328 systematic uncertainty in each direction.

1329 Finally, values for the absolute branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay are
1330 obtained by multiplying the relative branching fraction by the absolute branching
1331 fraction of the normalisation channel, $\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi\Lambda) = (6.3 \pm 1.3) \times 10^{-4}$ [?].
1332 Values are shown in Fig. 8.3 and summarised in Tab. 8.3, the uncertainty due to
1333 the knowledge of the normalisation channel (norm), which is correlated between q^2

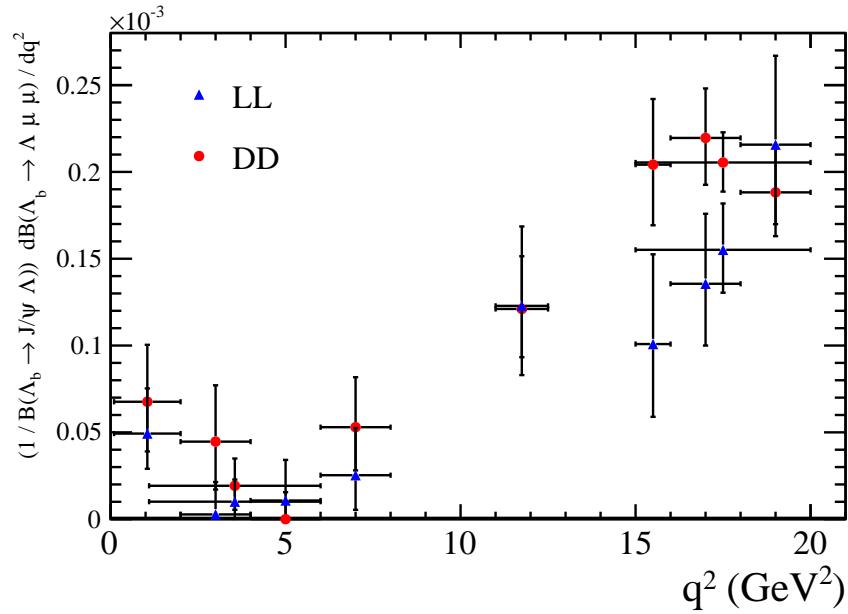


Figure 8.1: Measured values of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ relative to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay as a function of q^2 bins obtained fitting the downstream and long samples independently. Errors shown represent statistical and systematic uncertainties.

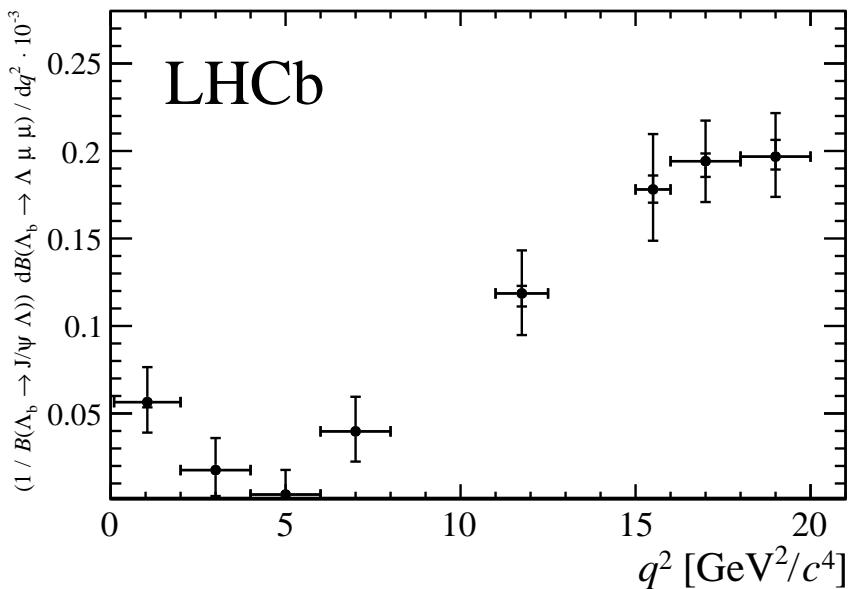


Figure 8.2: Branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay normalised to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ mode. The inner error represents the systematic error and the outer error the total error.

Table 8.2: 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

¹³³⁴ bins, is also shown. The SM predictions on the plot are obtained from Ref. [72].

¹³³⁵ Evidence for signal is found in the q^2 region between the charmonium resonances
¹³³⁶ and in the interval $0.1 < q^2 < 2.0 \text{ GeV}^2/c^4$, where an increased yield is expected due
¹³³⁷ to the proximity of the photon pole. The uncertainty on the branching fraction is
¹³³⁸ dominated by the precision of the branching fraction for the normalisation channel,
¹³³⁹ while the uncertainty on the relative branching fraction is dominated by the size of
¹³⁴⁰ the data sample available. The data are consistent with the theoretical predictions
¹³⁴¹ in the high- q^2 region but lie below the predictions in the low- q^2 region.

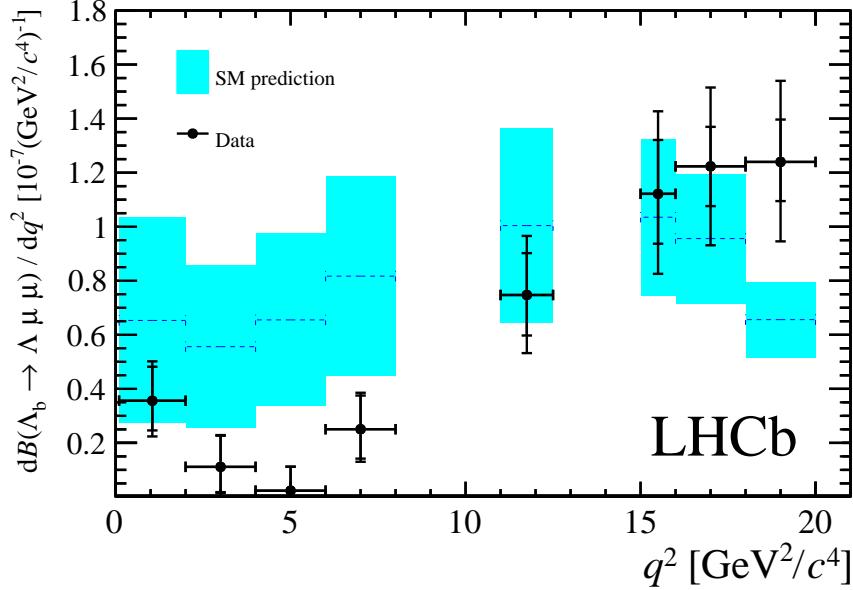


Figure 8.3: Measured $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction as a function of q^2 with the predictions of the SM [72] superimposed. The inner error bars on data points represent the total uncertainty on the relative branching fraction (statistical and systematic); the outer error bar also includes the uncertainties from the branching fraction of the normalisation mode.

Table 8.3: Measured differential branching fraction of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$, 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	$+0.12$	$+0.02$	± 0.07
2.0 – 4.0	0.11	$+0.12$	$+0.01$	± 0.02
4.0 – 6.0	0.02	$+0.09$	$+0.01$	± 0.01
6.0 – 8.0	0.25	$+0.12$	$+0.01$	± 0.05
11.0 – 12.5	0.75	$+0.15$	$+0.03$	± 0.15
15.0 – 16.0	1.12	$+0.19$	$+0.05$	± 0.23
16.0 – 18.0	1.22	$+0.14$	$+0.03$	± 0.25
18.0 – 20.0	1.24	$+0.14$	$+0.06$	± 0.26
1.1 – 6.0	0.09	$+0.06$	$+0.01$	± 0.02
15.0 – 20.0	1.20	$+0.09$	$+0.02$	± 0.25

1342

CHAPTER 9

1343

1344

Angular analysis

1345

The $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay angular distributions can be described as a function of three angles as defined in Fig. 9.1, 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. Finally, χ is the angle between dimuon and Λ decay planes, which is integrated over in this analysis. This part of the analysis performs 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

$$A_{\text{FB}}^i(q^2) = \frac{\int_0^1 \frac{d^2\Gamma}{dq^2 d\cos\theta_i} d\cos\theta_i - \int_{-1}^0 \frac{d^2\Gamma}{dq^2 d\cos\theta_i} d\cos\theta_i}{d\Gamma/dq^2}, \quad (9.1)$$

1346 where $d^2\Gamma/dq^2 d\cos\theta_i$ is the two-dimensional differential rate and $d\Gamma/dq^2$ is rate
1347 integrated over the angles.

1348 The A_{FB}^ℓ observable is also measured in $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays, going through the

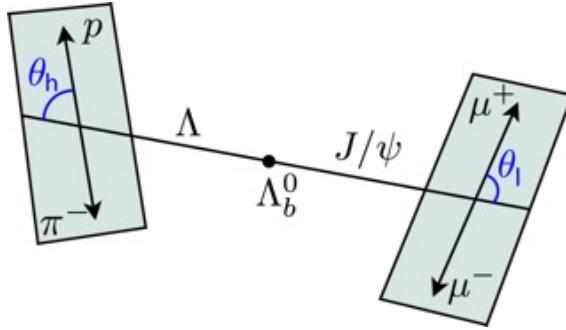


Figure 9.1: Graphical representation of the angles for the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay.

₁₃₄₉ same quark traditions as $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays. Instead the hadronic asymmetry,
₁₃₅₀ A_{FB}^h , is interesting only in the Λ_b^0 case as it is zero by definition in B^0 decays where
₁₃₅₁ K^{*0} decays strongly.

₁₃₅₂ 9.1 One-dimensional angular distributions

₁₃₅₃ In this section the derivation of the functional form of the angular distributions as
₁₃₅₄ a function of the $\cos\theta_\ell$ and $\cos\theta_h$ which are used to measure the observables. The
₁₃₅₅ content of this section is based on the calculations in Ref. [67]. For unpolarised Λ_b^0
₁₃₅₆ production, integrating over three angles the differential branching fraction is given
₁₃₅₇ in Eq. 11 of Ref. [67] as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2} = \frac{v^2}{2} \cdot \left(U^{11+22} + L^{11+22} \right) + \frac{2m_\ell^2}{q^2} \cdot \frac{3}{2} \cdot \left(U^{11} + L^{11} + S^{22} \right), \quad (9.2)$$

₁₃₅₈ and the lepton helicity angle θ_ℓ distribution is given in Eq. 15 as

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos\theta_\ell} &= v^2 \cdot \left[\frac{3}{8} (1 + \cos^2\theta_\ell) \cdot \frac{1}{2} U^{11+22} + \frac{3}{4} \sin^2\theta_\ell \cdot \frac{1}{2} L^{11+22} \right] \\ &- v \cdot \frac{3}{4} \cos\theta_\ell \cdot P^{12} + \frac{2m_\ell^2}{q^2} \cdot \frac{3}{4} \cdot \left[U^{11} + L^{11} + S^{22} \right]. \end{aligned} \quad (9.3)$$

₁₃₅₉ In these formulas m_ℓ is the mass of the lepton and $v = \sqrt{1 - 4m_\ell^2/q^2}$, U denotes
₁₃₆₀ the unpolarised-transverse contributions, L the longitudinal contributions and S
₁₃₆₁ the scalar contribution. The apices ¹¹ and ²² represent respectively vector and axial-

¹³⁶² vector currents, with $X^{11+22} = X^{11} + X^{22}$. The authors of Ref. [67] define then the
¹³⁶³ lepton-side forward-backward asymmetry as

$$A_{FB}^\ell(q^2) = -\frac{3}{2} \frac{v \cdot P^{12}}{v^2 \cdot (U^{11+22} + L^{11+22}) + \frac{2m_\ell^2}{q^2} \cdot 3 \cdot (U^{11} + L^{11} + S^{22})}. \quad (9.4)$$

Using these results as a starting point one can rewrite Eq. 9.3 as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_\ell} = \frac{3}{8} \frac{d\Gamma}{dq^2} (1 + \cos^2 \theta_\ell) U^{11+22} + \frac{d\Gamma}{dq^2} A_{FB}^\ell \cos \theta_\ell + \frac{3}{8} \sin^2 \theta_\ell v^2 (L^{11+22}) (U^{11} + L^{11} + S^{22}) \frac{3m_\ell^2}{q^2} \left(\frac{1}{8} - \frac{3}{8} \cos^2 \theta_\ell \right) \quad (9.5)$$

¹³⁶⁴ For this analysis the massless leptons limits is used, $m_\ell \rightarrow 0$. This is a good
¹³⁶⁵ approximation except at very low q^2 . In the massless limit the differential rates
¹³⁶⁶ simplify to

$$\frac{d\Gamma}{dq^2} = \frac{v^2}{2} \cdot \left(U^{11+22} + L^{11+22} \right) \quad (9.6)$$

and

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{3}{8} \frac{d\Gamma}{dq^2} (1 + \cos^2 \theta_\ell) U^{11+22} + \frac{d\Gamma}{dq^2} A_{FB}^\ell \cos \theta_\ell + \frac{3}{8} \sin^2 \theta_\ell v^2 (L^{11+22}). \quad (9.7)$$

Equations 9.6 and 9.7 can be then combined to achieve the form

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) \frac{U^{11+22}}{U^{11+22} + L^{11+22}} + A_{FB}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell \frac{L^{11+22}}{U^{11+22} + L^{11+22}} \right]. \quad (9.8)$$

¹³⁶⁷ The amplitude combination in the last term can be viewed as ratio between longitudi-
¹³⁶⁸ dinal and sum of longitudinal and unpolarized transverse contributions and therefore
¹³⁶⁹ one can define the longitudinal fraction

$$f_L = \frac{L^{11+22}}{U^{11+22} + L^{11+22}}, \quad (9.9)$$

which leads to the distribution used in the analysis

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) (1 - f_L) + A_{FB}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell f_L \right]. \quad (9.10)$$

¹³⁷⁰ Using the same steps the proton helicity distribution is given in Ref. [67] as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda(\rightarrow p\pi^-)\ell^+\ell^-)}{dq^2 d \cos \theta_h} = \text{Br}(\Lambda \rightarrow p\pi^-) \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+\ell^-)}{dq^2} \left(\frac{1}{2} + A_{FB}^h \cos \theta_h \right), \quad (9.11)$$

¹³⁷¹ where

$$A_{FB}^h = \frac{1}{2} \alpha_\Lambda P_z^\Lambda(q^2). \quad (9.12)$$

¹³⁷² where $P_z^\Lambda(q^2)$ is the polarisation of the daughter baryon, Λ , and $\alpha_\Lambda = 0.642 \pm 0.013$ [3]

¹³⁷³ is the Λ decay asymmetry parameter.

¹³⁷⁴ These expressions assume that Λ_b^0 is produced unpolarised, which is in agreement
¹³⁷⁵ with the recent LHCb measurement of the production polarisation [75]. Possible
¹³⁷⁶ effects due to a non zero production polarisation are investigated in systematic
¹³⁷⁷ uncertainties.

¹³⁷⁸ 9.2 Multi-dimensional angular distributions

To incorporate effects of production polarisation this was introduced in the equations. In the modified version, the angle θ is sensitive to the production polarisation through the spin-density matrix in Eq. A.4. Integrating over θ_h , ϕ_Λ , ϕ_ℓ and θ results in the same distribution as in the unpolarised case (Eq. 9.3). Therefore, in the case of uniform efficiency, the lepton side forward-backward asymmetry, A_{FB}^ℓ , is unaffected by production polarization. To estimate effect of the production polarization, the two-dimensional distribution in θ and θ_ℓ is also derived, which in the massless

leptons limit becomes (up to a constant multiplicative factor)

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d(\cos \theta) d(\cos \theta_\ell)} = \frac{d\Gamma}{dq^2} \left\{ \frac{3}{8} (1 + \cos^2 \theta_\ell) (1 - f_L) + A_{FB}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell f_L + P_b \cos(\theta) \left[-\frac{3}{4} \sin(\theta_\ell)^2 O_{Lp} + \frac{3}{8} (1 + \cos(\theta_\ell)^2) O_P \right. \right. \\ \left. \left. - \frac{3}{8} \cos(\theta_\ell) O_{U12} \right] \right\}, \quad (9.13)$$

where three more observables are defined

$$O_{Lp} = \frac{L_P^{11} + L_P^{22}}{U^{11+22} + L^{11+22}}, \\ O_P = \frac{P^{11} + P^{22}}{U^{11+22} + L^{11+22}}, \\ O_{U12} = \frac{U^{12}}{U^{11+22} + L^{11+22}}.$$

- ¹³⁷⁹ In the massless leptons approximation two of these quantities are related to hadron
¹³⁸⁰ side forward-backward asymmetry as

$$\frac{1}{2} \alpha_\Lambda (O_P + O_{Lp}) = A_{FB}^h. \quad (9.14)$$

Following the same steps as lepton side case, after integrating over four angles one finds that the hadron side, A_{FB}^h , is also unaffected by the production polarization in case of uniform efficiency. The two dimensional distribution in θ and θ_h has form

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d(\cos \theta) d(\cos \theta_B)} = \frac{d\Gamma}{dq^2} [1 + 2A_{FB}^h \cos \theta_B + P_b (O_P - O_{Lp}) \cos \theta \\ + \alpha_A P_b (1 - 2f_L) \cos \theta \cos \theta_B]. \quad (9.15)$$

- ¹³⁸¹ In order to use two-dimensional distributions, expectations for the three additional
¹³⁸² observables, which do not enter one-dimensional distributions are needed. Expecta-
¹³⁸³ tions are calculated using form factors and numerical inputs from Ref. [67] and are
¹³⁸⁴ shown in Tab. 9.1.

$q^2 [GeV^2/c^2]$	A_{FB}^ℓ	P_z^A	f_L	O_P	O_{Lp}	O_{U12}
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
11.0 – 12.5	-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	-0.322	-0.302

Table 9.1: Prediction for angular observables entering two-dimensional angular distributions. Prediction is based on covariant quark model form factors from Ref. [67].

For completeness, the two-dimensional distribution in $\cos \theta_L$ - $\cos \theta_B$ has the form

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d(\cos \theta_B) d(\cos \theta_L)} = & \frac{3}{8} + \frac{6}{16} \cos^2 \theta_L (1 - f_L) - \frac{3}{16} \cos^2 \theta_L f_L + A_{FB}^\ell \cos \theta_L + \\ & \left(\frac{3}{2} A_{FB}^h - \frac{3}{8} \alpha_A O_P \right) \cos \theta_B - \frac{3}{2} A_{FB}^h \cos^2 \theta_L \cos \theta_B - \frac{3}{16} f_L + \\ & \frac{9}{16} f_L \sin^2 \theta_L + \frac{9}{8} \alpha_A \cos^2 \theta_L \cos \theta_B O_P - \\ & \frac{3}{2} \alpha_A \cos \theta_L \cos \theta_B O_{U12}. \end{aligned} \quad (9.16)$$

1385 9.3 Angular resolution

1386 In this section is reported a study of the angular resolution done in order to achieve
 1387 a better understanding of detector and reconstruction effects. This will be then
 1388 used to study systematic uncertainties. The study is done by analysing simulated
 1389 events and comparing true and reconstructed quantities. In Fig. 9.2 plots of the
 1390 difference between true and measured angular observables ($\cos \theta_\ell$ and $\cos \theta_h$) are
 1391 reported as a function of the observable itself. These are centred at 0 indicating no
 1392 bias in the measurement. In Fig. 9.2 the same difference is shown also as a function
 1393 of q^2 showing again no bias. The spread the distributions on these plots around
 1394 the central value is an estimate of the angular resolution. Taking vertical slices of
 1395 the distributions in Fig. 9.2 approximately gaussian distributions centred at 0 are

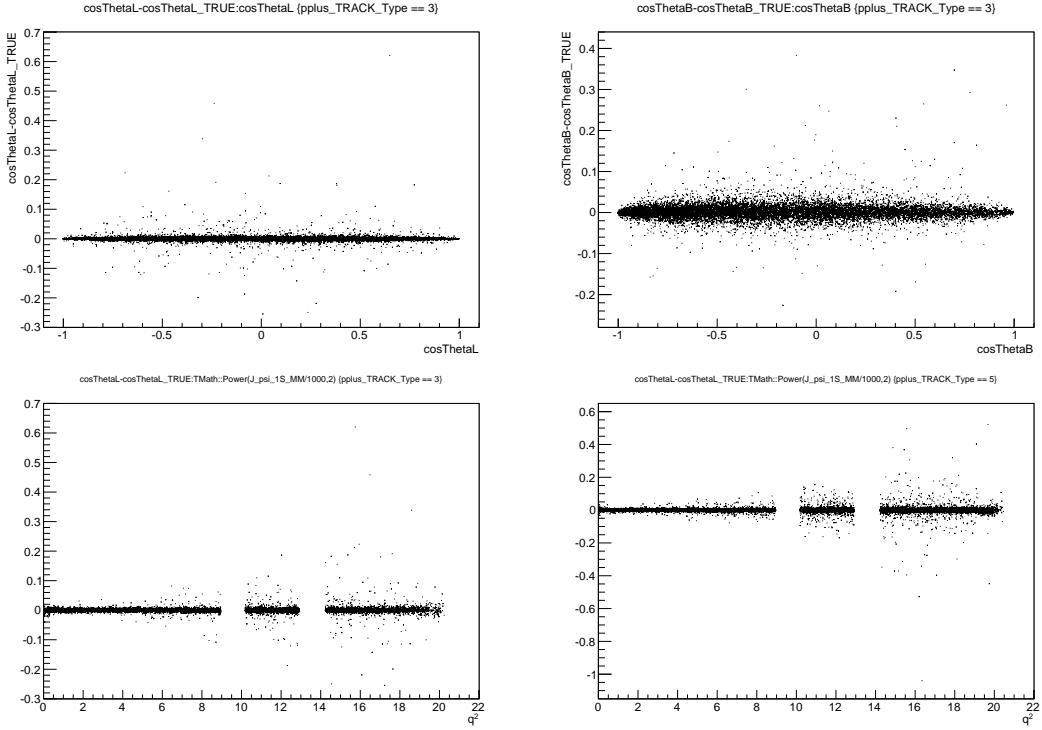


Figure 9.2: Difference of “true minus reconstructed” angular observables as a function of themselves long candidates (top) and as a function of q^2 for long (bottom left) and downstream (bottom right) candidates.

1396 obtained. These distributions are fit with a single gaussian and its σ is interpreted
 1397 as angular resolution. In Tab. 9.2 he average resolutions are reported for the two
 1398 angular variables separately for the long and downstream candidates. As expected
 1399 candidates built from long tracks are characterised by a better resolution due to a
 1400 better momentum and vertex resolution. In Fig. 9.3 responde matrices, showing the
 1401 correlation between reconstructed and generated angular observables, are shown.

Table 9.2: Average angular resolutions integrated over the full interval and the full available q^2 .

Observable	DD	LL
$\cos\theta_\ell$	0.015	0.01
$\cos\theta_h$	0.066	0.014

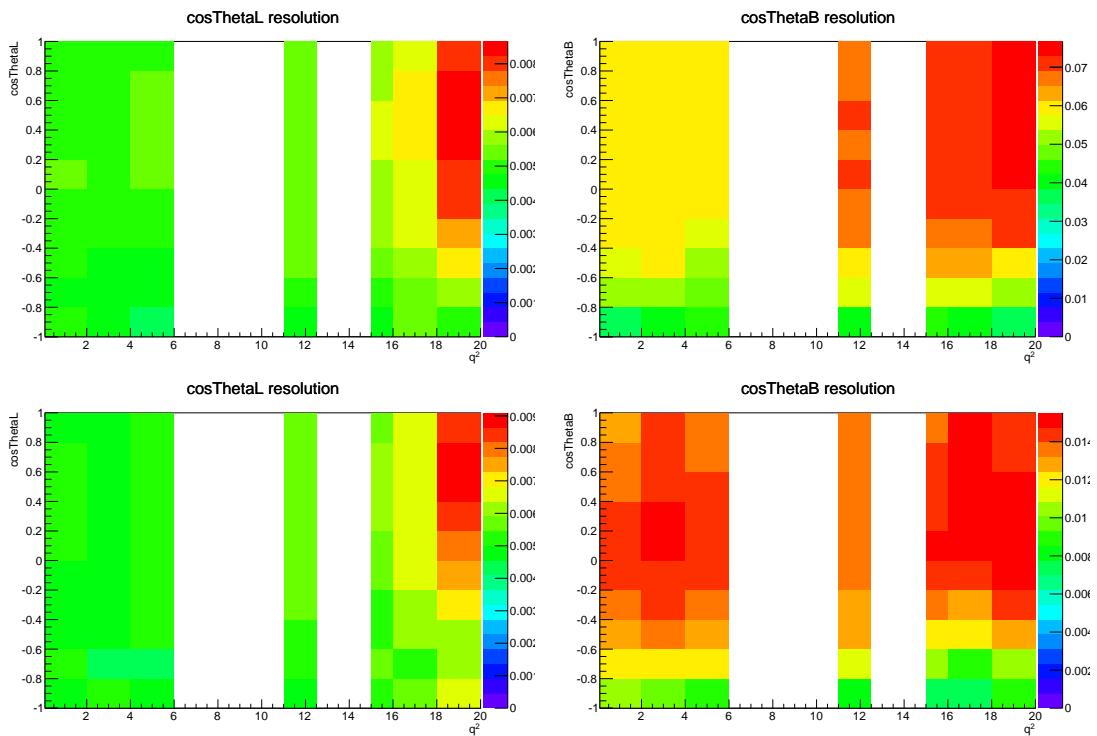


Figure 9.3: Angular resolution for $\cos \theta_\ell$ (left plots) and $\cos \theta_h$ (right plots) as a function of the angular variables and q^2 for down-down (upper plots) and long-long (lower plots) events.

1402

CHAPTER 10

1403

1404

Angular fit strategy

1405

1406 In this chapter is described the fitting technique applied to extract the angular
1407 observables. There are physical limits to the values of the parameters of interests:
1408 A_{FB}^h is limited in the $[-0.5, 0.5]$ interval and for the f_L and A_{FB}^ℓ parameters the
1409 physical region, given by $|A_{\text{FB}}^\ell| < 3/4(f_L - 1)$, is the triangle shown in Fig. 10.1. If
1410 the minimum is close to the border minuit does not always converge. For this reason
1411 a "brute force" fitting technique is applied. For this purpose fit parameters are
1412 divided into two categories: parameters of interest (PoIs), A_{FB}^ℓ , A_{FB}^h and f_L and all
1413 other parameters referred to as nuisances. The Log-Likelihood, $\log\mathcal{L}$, of the fit model
1414 with respect of data is calculated. The full PoIs allowed area is scanned looking for
1415 the minimum of the Log-Likelihood. A first coarse scan finds a candidate minimum
1416 and then the procedure is reiterated two more times in finer intervals around it. For
1417 each point all the niusances are fitted using a maximum likelihood fit. The fit is
1418 therefore constrained inside the physical region, if the best log-likelihood is found
1419 to be outside it, the point at the boundary is chosen as the best fit.

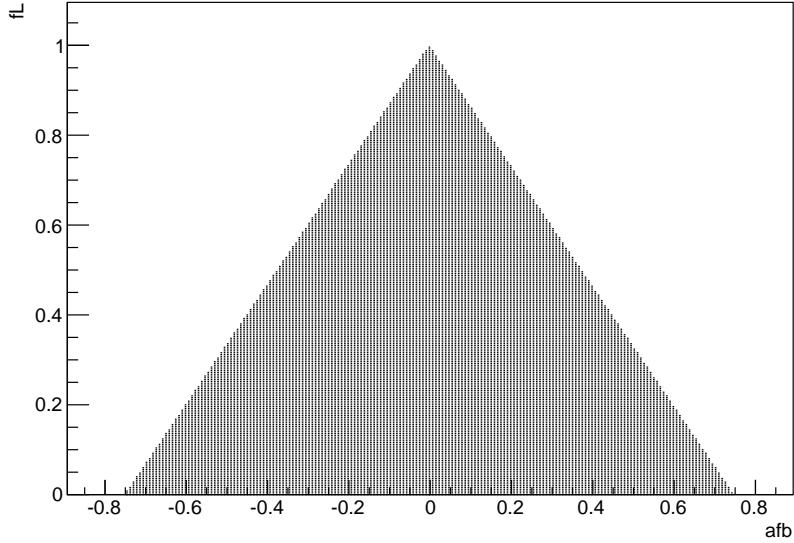


Figure 10.1: The plot shows the physical A_{FB}^ℓ vs f_L parameter space. The dark region corresponds to points where the pdf is positive in the whole $[-1, 1]$ interval.

¹⁴²⁰ 10.1 Feldman-cousins plug-in method

¹⁴²¹ Since physical boundaries of the parameter space could result in a wrong estimation
¹⁴²² of the uncertainties, especially if the measured value is close to the border, in this
¹⁴²³ analysis the likelihood-ordering method [?] is used. Nuisance parameters are ac-
¹⁴²⁴ counted for using the plug-in method [?, 76]. This is a unified method to calculate
¹⁴²⁵ confidence intervals and upper/lower limits, based on simulated experiments and
¹⁴²⁶ has the advantage of having a well defined frequentist coverage.

¹⁴²⁷ The method is constituted by the following steps:

- ¹⁴²⁸ 1. fit real data distributions with all parameters free;
- ¹⁴²⁹ 2. fit real data fixing the PoIs to a value of choice and keeping nuisance param-
¹⁴³⁰ eters free;
- ¹⁴³¹ 3. generate simulated events following the distribution given by the fit model,
¹⁴³² where all nuisance parameters are taken from the fit in point 2 and PoIs are
¹⁴³³ fixed to the same value used in point 2;

1434 4. repeat the two fits made on data on each simulated sample: fit with all pa-
1435 rameters free and fit with PoIs fixed;

1436 5. extract the value of the Log-Likelihoods at the minimum for all cases;

1437 6. calculate the percentage of times in which the ratio $\log\mathcal{L}_{fixed}/\log\mathcal{L}_{free}$ is smaller
1438 in data than in the simulated experiments.

1439 7. repeat the procedure for many values of the PoIs scanning around the best fit
1440 values.

1441 The confidence interval at k% is given by the points where the free-to-fixed likelihood
1442 ratio is bigger in data than simulation for (1-k)% of times. As an example, in
1443 Fig. 10.2 are reported the p-values obtained with the plug-in method for A_{FB}^h and
1444 f_L . For the leptons case we have instead a two-dimensional grid of p-values which
gives two-dimensional confidence regions.

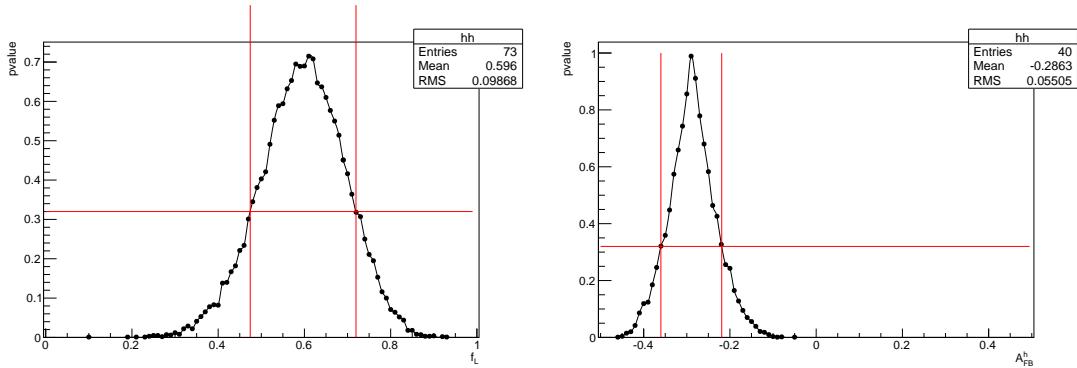


Figure 10.2: Distribution of measured values of angular observables f_L (left) and A_{FB}^h (right) in simulated experiments. The red lines mark the points at p-value 32% corresponding to a 68% CL.

1446 10.2 Modelling the angular distributions

- 1447 The observables are extracted from fits to one-dimensional angular distributions.
 1448 The PDFs used to model the data is defined as

$$P^k(\cos \theta_{\ell/h}) = f_b P_S(\cos \theta_{\ell/h}) \times \varepsilon^k(\cos \theta_{\ell/h}) + (1 - f_b) P_B^k(\cos \theta_{\ell/h}), \quad (10.1)$$

1449 where $k = (\text{LL}, \text{DD})$, P_S is the signal function composed by a theoretical shape given
 1450 by Eq. 9.11 and ?? multiplied by an acceptance function ε described in Sec. 10.3 and
 1451 P_B is a background component. To limit systematic effects due to the background
 1452 parameterisation, the fit is performed in a restricted mass region around the peak:
 1453 $5580 < m(\Lambda\mu^+\mu^-) < 5660 \text{ MeV}/c^2$ (“signal region”), which is dominated by the
 1454 signal. The background fraction, f_b , is obtained by looking at the invariant mass
 1455 distribution in a wider interval and fitting it to extract the fraction of background
 1456 in the signal region. In the fit to the angular distributions this is then gaussian
 1457 constrained to obtained value. The background shape is parameterised with a
 1458 linear function times the efficiency shape. A different efficiency shape is used for
 1459 downstream and long events and for each q^2 bin. The free parameter of this model is
 1460 fitted on sideband candidates which contain only background and fixed for the fit to
 1461 the signal region. In Figs. 10.3 and 10.4 are reported the background distributions
 1462 in the sideband ($m(\Lambda\mu^+\mu^-) > 5700 \text{ MeV}/c^2$) for the high and low q^2 integrated
 1463 intervals with overlaid the background functions alone. In summary the only fit
 1464 parameter in the total fit function is the forward-backward asymmetry.

1465 10.3 Angular acceptance

- 1466 Selection requirements on the minimum momentum of the muons may distort the
 1467 $\cos \theta_\ell$ distribution by removing candidates with extreme values of $\cos \theta_\ell$. Similarly,
 1468 the impact parameter requirements affect $\cos \theta_h$ as very forward hadrons tend to have
 1469 smaller impact parameter values. While in principle one could take it into account

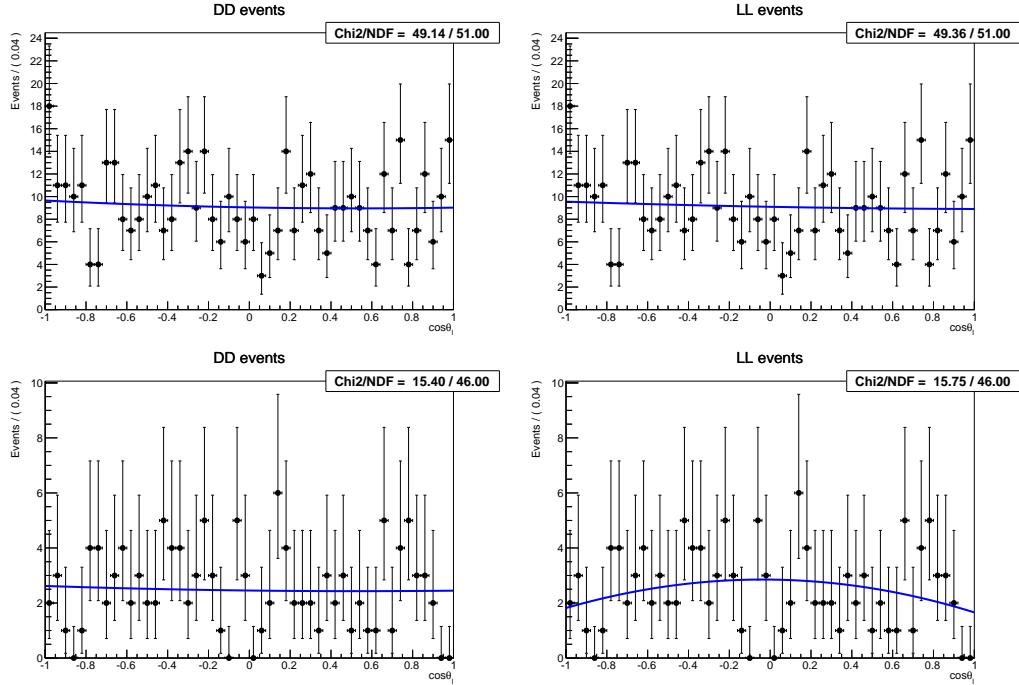


Figure 10.3: Background distribution as a function of $\cos \theta_\ell$ for downstream (left) and long (right) events in the $15\text{-}20 \text{ GeV}^2/c^4$ (top) and $1.1\text{-}6 \text{ GeV}^2/c^4$ (bottom) q^2 bins.

1470 by an additional weight, to minimise the distortion of the uncertainties estimate,
 1471 the efficiency function is incorporated in the fit model. The angular efficiency is
 1472 parametrised using a second-order polynomial and determined separately for down-
 1473 stream and long candidates by fitting simulated events, with an independent set of
 1474 parameters obtained for each q^2 interval. These parameters are fixed in the fits to
 1475 data. Using polynomial functions allows to calculate the PDF normalisation ana-
 1476 lytically. In Figs. ?? and 10.6 efficiencies are reported, as a function of $\cos \theta_h$ and
 1477 $\cos \theta_\ell$ using $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ simulated events in the $15.0\text{-}20.0$ and $1.1\text{-}6.0 \text{ GeV}^2/c^4$
 1478 integrated q^2 intervals. For the lepton side, even though the efficiency is symmetric
 1479 by construction, all parameters floating in the fit, namely it is not constrained to be
 1480 symmetric.

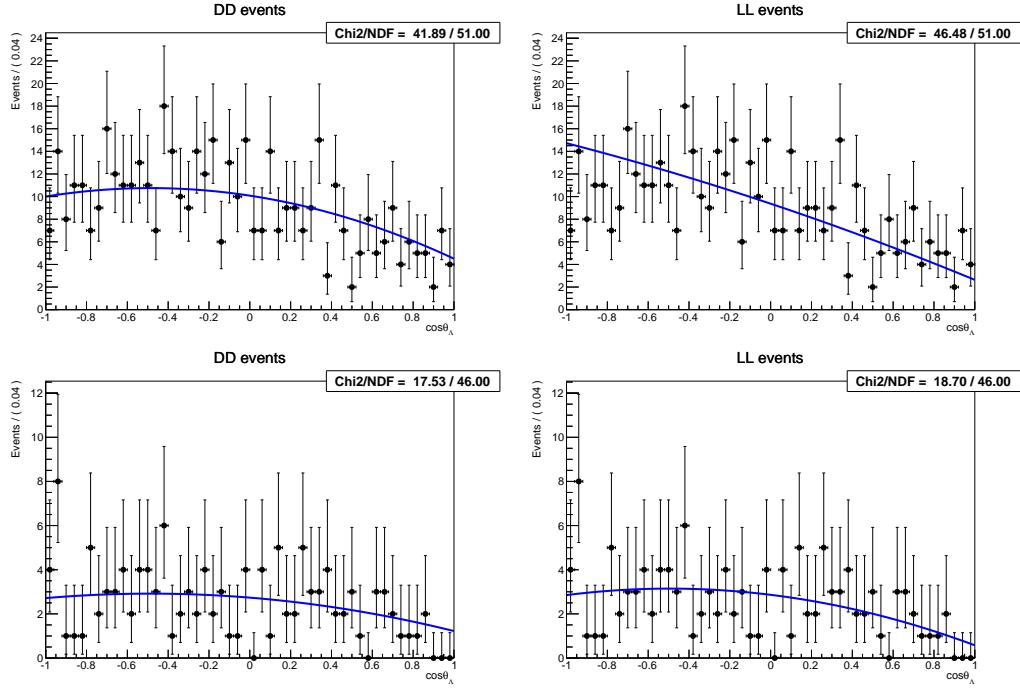


Figure 10.4: Background distribution as a function of $\cos \theta_\Lambda$ for downstream (left) and long (right) events in the 15-20 GeV^2/c^4 (top) and 1.1-6 GeV^2/c^4 (bottom) q^2 bins.

1481 10.4 Studies on a three-dimensional fit

1482 One other way of extracting the angular observables would be to fit at the same
 1483 time both angles and also the invariant mass distribution in order to have a better
 1484 handle on the level of background. In this case one can use more of the information
 1485 available. On the other hand it is necessary to use a larger mass window with more
 1486 background in it and more parameters to fit. In the 1D case the free parameters are
 1487 the two parameters of interest for the lepton case and one for the hadron one. For the
 1488 3D case the free parameters are the three parameters of interest plus two background
 1489 fractions and the two exponential slopes for the invariant mass background. An high
 1490 number of free parameters is difficult to constrain with the very limited statistics
 1491 available. To check which method gives the best sensitivity 500 pseudo-experiments
 1492 are generated. Events are generated in 3D using shapes taken from the fit on real
 1493 data. The generated values of the parameters of interest are $A_{FB}^\ell = 0$, $f_L = 0.7$ and
 1494 $A_{FB}^h = -0.37$. These are data-like values inspired to what is measured in the highest

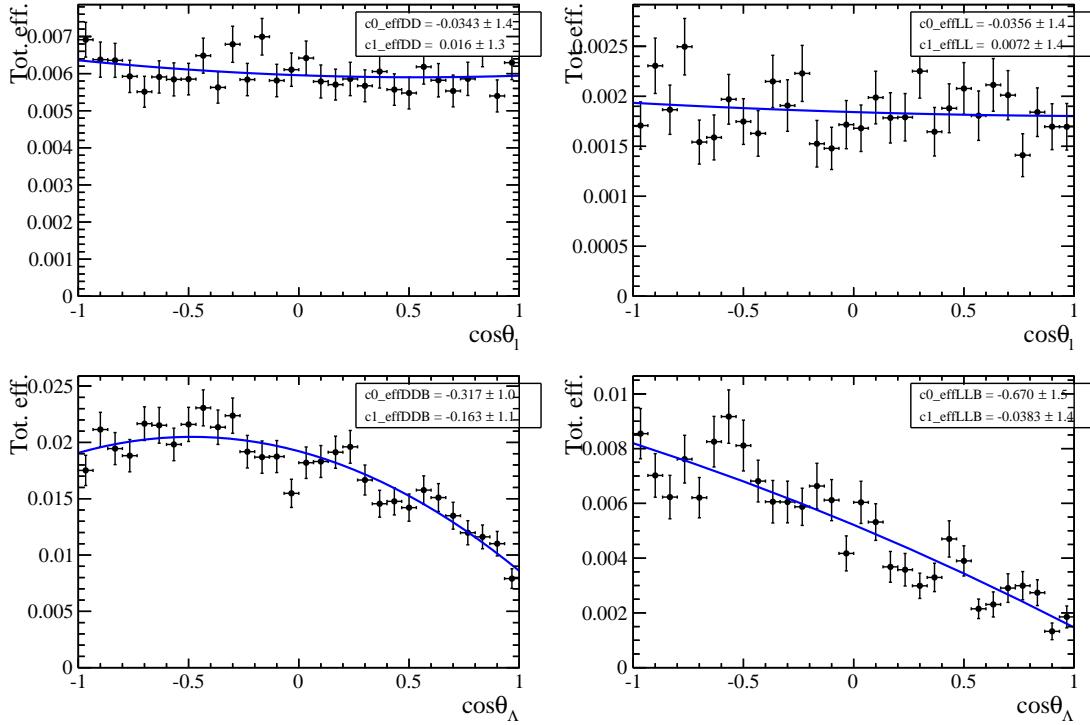


Figure 10.5: Efficiency as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

Table 10.1: RMS values for toy experiments on the extraction of the three parameters of interest with the 1D or 3D fitting methods.

$q^2 [\text{GeV}^2/c^4]$	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

statistics interval. The overall statistics and the fraction of background events in the mass window are constrained to what obtained from the fit on real data. Each pseudo-experiment is fitted with both methods and Fig. 10.7 reports distributions of parameters of interest obtained from the fit in the 1D and 3D cases. The RMS of these distributions can be taken as a measure of the sensitivity of each method. In Tab. 10.1 RMSs from both methods can be compared. For all parameters of interest the 1D fit method gives a smaller RMS, hence a better sensitivity.

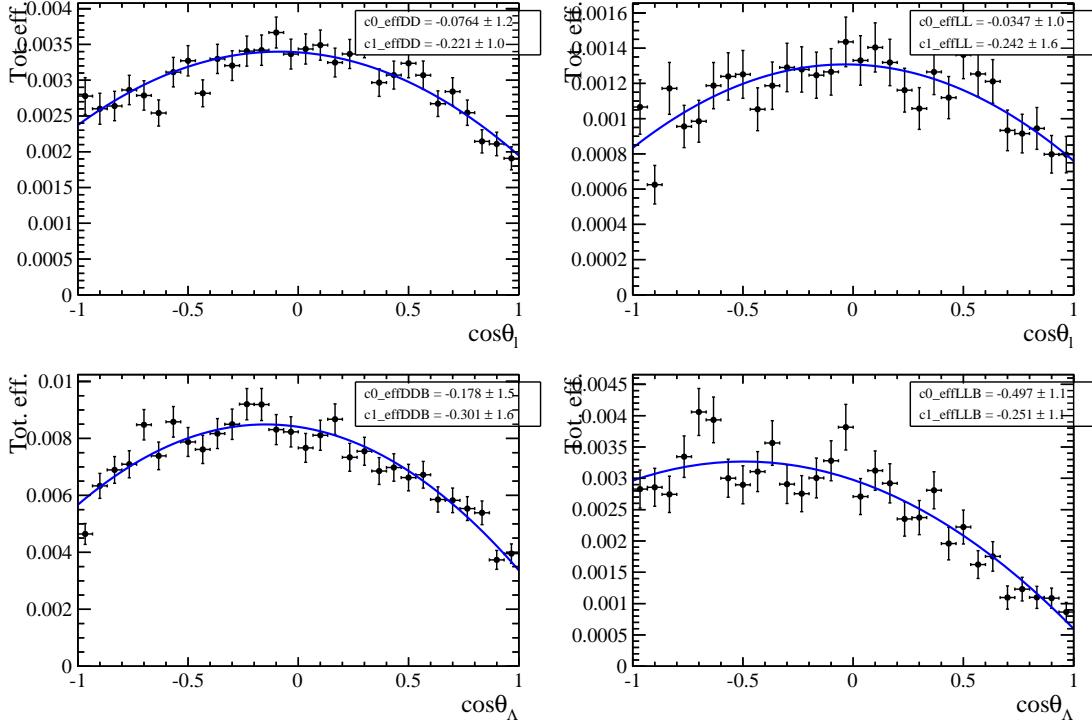


Figure 10.6: Efficiency as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $1.1\text{--}6.0$ $\text{GeV}^2/c^4 q^2$ interval.

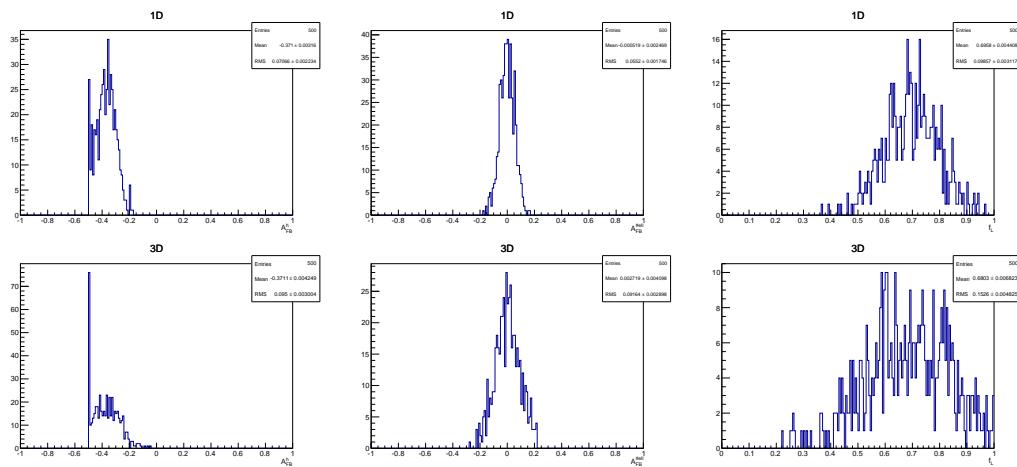


Figure 10.7: Distribution of observed parameters of interest over 500 pseudo-experiments obtained using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what is observed in the $15\text{--}20$ q^2 interval.

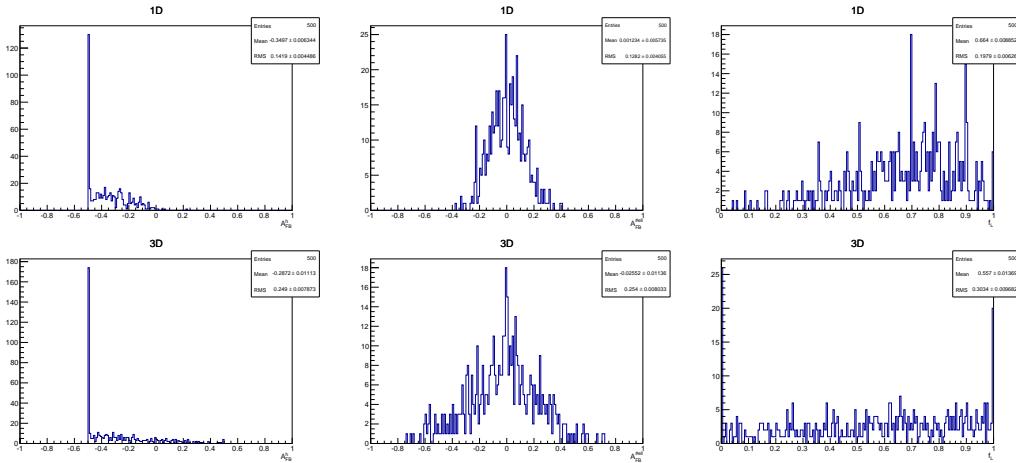


Figure 10.8: Distribution of observed parameters of interest over 500 pseudo-experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the 11–12.5 q^2 interval.

1502

CHAPTER 11

1503

1504

Systematics uncertainties on angular observables and results

1505

1506 In the following sections are described the five main sources of systematic uncertain-
1507 ties that are considered for the angular observables measurement and, finally, results
1508 are reported in Sec.11.3. Results are shown only for q^2 intervals where the signal
1509 significance, shown in Tab. 5.3 is more than 3 standard deviations. This includes
1510 all of the q^2 intervals above the J/ψ resonance and the lowest q^2 bin.

1511 11.1 Non-flat angular efficiency

1512 The angular efficiency is non-flat as a function of $\cos \theta_\ell$ and $\cos \theta_h$. Therefore,
1513 while integrating the full angular distribution, terms that cancel with perfect effi-
1514 ciency may remain and generate a bias in the final result. In order to deal with
1515 this effect simulated events are generated in a two-dimensional $(\cos \theta_\ell, \cos \theta_\Lambda)$ space
1516 according to the theoretical distribution described by Eq. 9.16 multiplied by the

1517 two-dimensional efficiency histograms obtained from simulation and reported in
 1518 Fig. 11.1. Then one-dimensional projections are taken and fit using the default
 1519 1D efficiency functions. In Fig. 11.2 deviations from the true generated value
 1520 $\Delta x = x_{true} - x_{measured}$ are shown. Since the mean of these distributions is non
 1521 zero by more than 3σ , they are taken as systematic uncertainties.

1522 **11.1.1 Resolution**

1523 The angular resolution could bias the observables measurement generating an asym-
 1524 metric migration of events. This is especially important in the $\cos \theta_h$ case, because
 1525 this has worse resolution and a considerably asymmetric distribution. To study this
 1526 systematic simulated experiments are used, where events are generated following
 1527 the measured distributions (including efficiencies). The generated events are then
 1528 smeared by the angular resolution (gaussian smearing). To be conservative the case
 1529 with biggest angular resolution, downstream events, is always used. Finally, the
 1530 smeared and not-smeared distributions are fitted with the same PDF. The average
 1531 deviation from the default values are reported in Tab. 11.1 as a function of q^2 and
 1532 assigned as systematic uncertainties.

Table 11.1: Values of simulated $\cos \theta_\ell$ and $\cos \theta_\Lambda$ resolutions and systematic uncer-
tainties on angular observables due to the finite resolution in bins of q^2 .

q^2 [GeV^2/c^4]	σ_ℓ	σ_Λ	$\Delta A_{\text{FB}}^\ell$	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

1533 **11.1.2 Angular efficiency**

1534 An imprecise determination of the reconstruction and selection efficiency can in-
 1535 troduce extra oddity and therefore bias the measurement. To asses this effect the

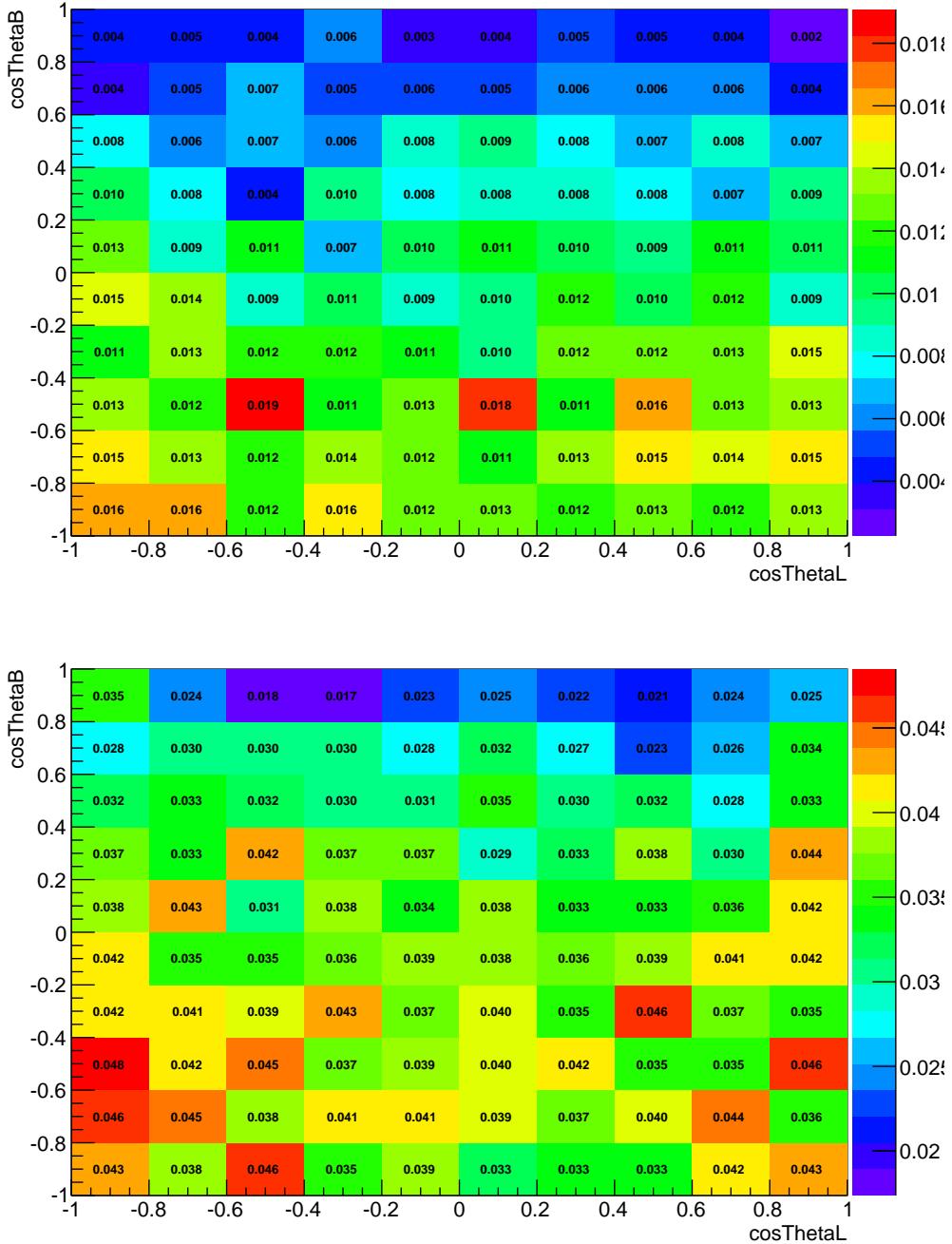


Figure 11.1: Angular acceptance as a function of $\cos \theta_\ell$ and $\cos \theta_h$ for long (left) and downstream (right) candidates, integrated over the full available q^2 range.

1536 kinematic re-weighting described in 3.3.2 is removed from the simulation and the
 1537 efficiency is determined again. Simulated events are then fit using the same the-
 1538 oretical PDF and multiplied by the efficiency function obtained with and without
 1539 kinematical weights. As in the previous cases the average bias is taken as systematic

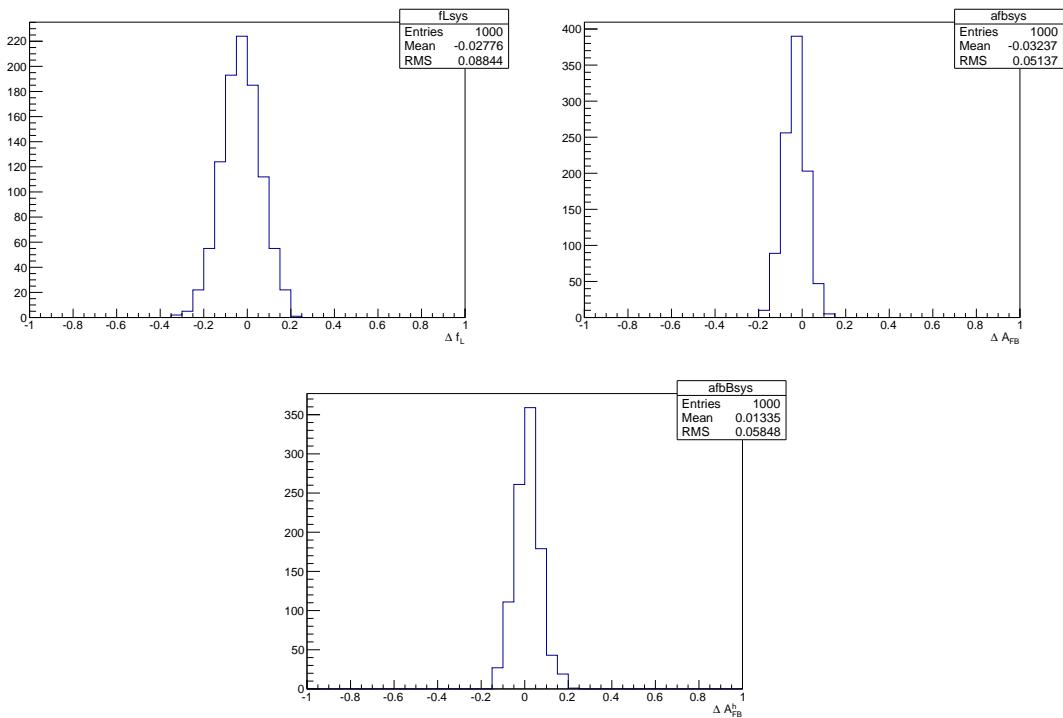


Figure 11.2: Deviations of the observables' values obtained fitting simulated events generated with a 2D distribution multiplied by a 2D efficiency and fitting 1D projections with respect to generated values. From left to for f_L (top left), A_{FB}^ℓ (top right) and A_{FB}^h (bottom).

1540 uncertainty. Results are shown in Tab. 11.2.

Table 11.2: Values systematic errors 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

1541 Furthermore, for the effect of the limited simulated statistics is taken into account
1542 and found to be negligible with respect to other sources.

Table 11.3: 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

1543 11.1.3 Background parameterisation

1544 There is a certain degree of arbitrariness in the choice of a parameterisation for the
1545 background, especially in bins with low statistics. To assess possible biases due
1546 to this simulated experiments are generated using the shapes from real data fits
1547 and the same statistics observed in data for each q^2 bin. Each toy is fit with two
1548 models: the default one, a “line times efficiency” function and the efficiency function
1549 alone, corresponding to the assumption that background distributions are originally
1550 flat and only modified by the interaction with the detector. The average bias the
1551 generated experiments is taken as systematic uncertainties. Results are reported in
1552 Tab. 11.4.

Table 11.4: Values of systematics 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

1553

1554 11.1.4 Polarisation

1555 To study the effect of a non-zero Λ_b^0 production polarisation simulated events are
 1556 generated using the distributions 9.13 and 9.15 as a function of our angular observ-
 1557 able and $\cos\theta$, which is sensitive to polarisation. Similarly to the procedure used
 1558 for the branching ratio measurement, events are generated using values of the polar-
 1559 isation corresponding to $\pm\sigma$ from the LHCb measurement [68]). In the theoretical
 1560 distributions $\cos\theta$ is always odd therefore with perfect efficiency it always drops out
 1561 by integrating over $\cos\theta$. Therefore the generated distributions must also contain
 1562 the information of the two-dimensional efficiency. No significant bias is found.

1563 11.2 J/ψ cross-check

1564 To cross-check the fitting procedure this is applied on the high statistics $\Lambda_b^0 \rightarrow J/\psi$
 1565 Λ sample. To select these events the same selection as for the branching fraction is
 1566 used (see Sec. 4) with the addition of a strong PID cut on the proton ($\text{PID}p > 10$),
 1567 needed to reduce the $K_s^0 J/\psi$ background. This is particularly important for the
 1568 $\cos\Theta_h$ fit, since the K_s^0 events are not distributed in a flat way in this variable
 1569 would therefore can bias the fit. In Fig. 11.3 invariant mass plots after this cut are
 1570 reported, which can be compared with Fig. 5.3. After the PID cut there are 0.2% of

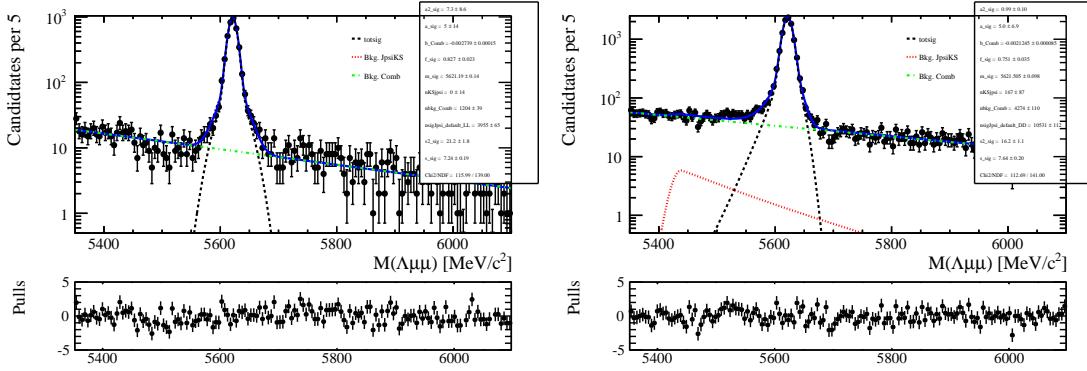


Figure 11.3: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ long (left) and downstream (right) candidates with an extra proton PID cut to remove K_s^0 background.

1571 K_s^0 events left in the downstream sample and a fraction compatible with zero in the
1572 long sample. The signal fit model is the same used for the rare case and described in
1573 10.2. For the background instead the higher statistics allows to leave more freedom
1574 to the fit. Therefore a second-order Chebyschev polynomial is used, where the two
1575 parameters are free to float. As for the rare case the background fractions are
1576 gaussian-constrained to what found in the invariant mass fit. In Figs. 11.4, 11.5
1577 fitted angular distributions are reported for the J/ψ channel. The measured values of
1578 the observables are $A_{FB}^\ell = -0.002^{+0.011}_{-0.011}$, $A_{FB}^h = -0.402^{+0.010}_{-0.009}$ and $f_L = 0.485^{+0.019}_{-0.020}$,
1579 where the errors are 60% Feldman Cousins confidence intervals. The lepton side
1580 asymmetry as expected is measured to be zero.

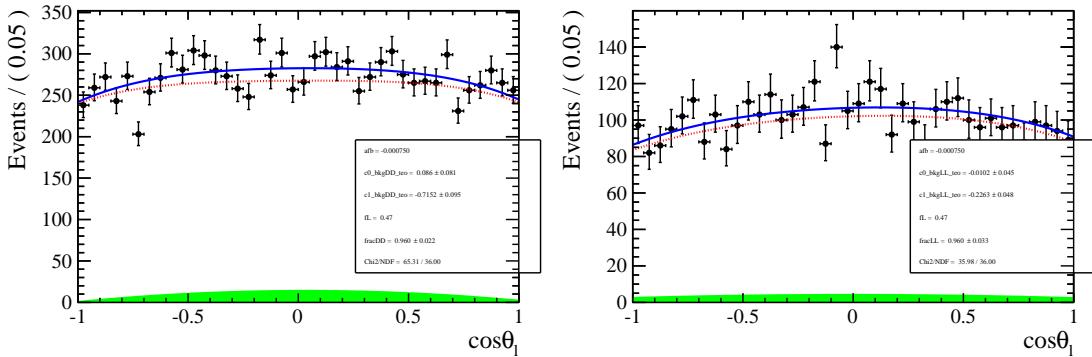


Figure 11.4: Fitted angular distribution as a function of $\cos\theta_\ell$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

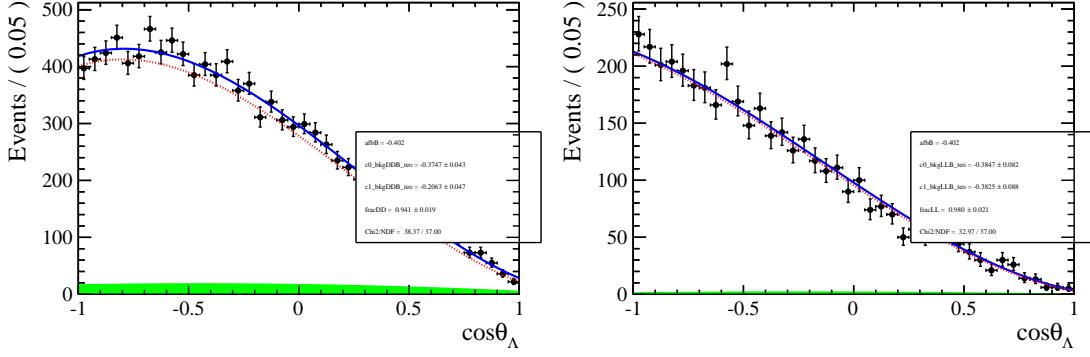


Figure 11.5: Fitted angular distribution as a function of $\cos \theta_h$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

1581 11.3 Results

1582 In Figs. 11.6 and ?? are reported fits to angular distributions for the 15–20 $\text{GeV}^2/c^4 q^2$
 1583 interval. Tab. 11.5 reports measured values of A_{FB}^ℓ , A_{FB}^h and f_{L} , with the asymme-
 1584 tries shown in Fig. 11.8. The statistical uncertainties on these tables are obtained
 1585 using the likelihood-ratio ordering method described in Sec. 10.1, where only one of
 1586 the two observables at a time is treated as the parameter of interest. In Fig. 11.9 the
 1587 statistical uncertainties on A_{FB}^ℓ and f_{L} are also reported as two-dimensional 68 %
 1588 confidence level (CL) regions, where the likelihood-ratio ordering method is applied
 1589 by varying both observables and therefore taking correlations into account. Total
 1590 systematic uncertainties correspond to the square root sum of the single considered
 1591 sources. The SM predictions on the plots are obtained from Ref. [72].

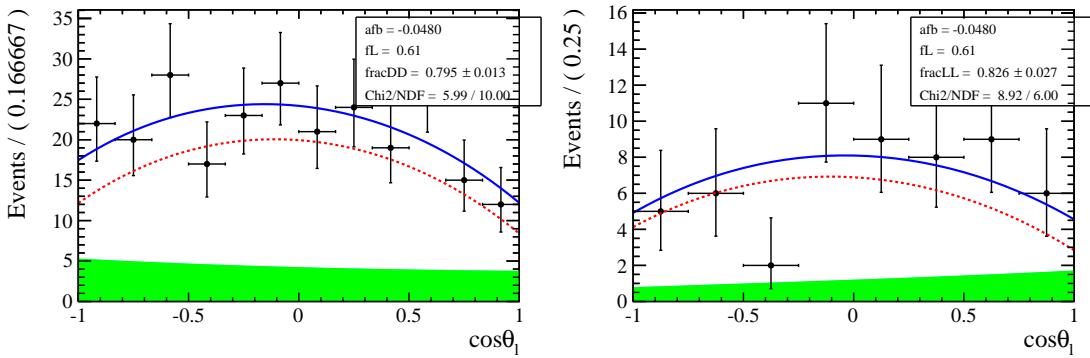


Figure 11.6: Fitted angular distributions as a function of $\cos \theta_\ell$ for downstream (left) and long (right) candidates in the 15–20 $\text{GeV}^2/c^4 q^2$ interval.

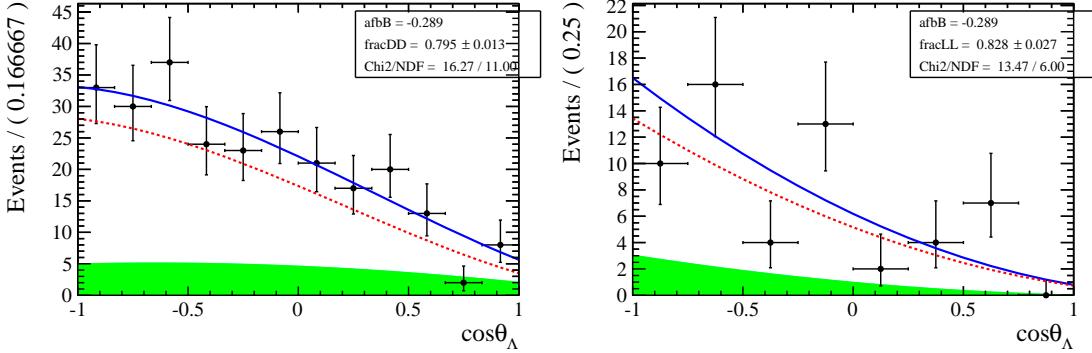


Figure 11.7: Fitted angular distributions as a function of $\cos \theta_h$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

Table 11.5: 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	$-0.05^{+0.09}_{-0.09} \pm 0.03$	$0.61^{+0.11}_{-0.14} \pm 0.03$	$-0.29^{+0.07}_{-0.07} \pm 0.03$

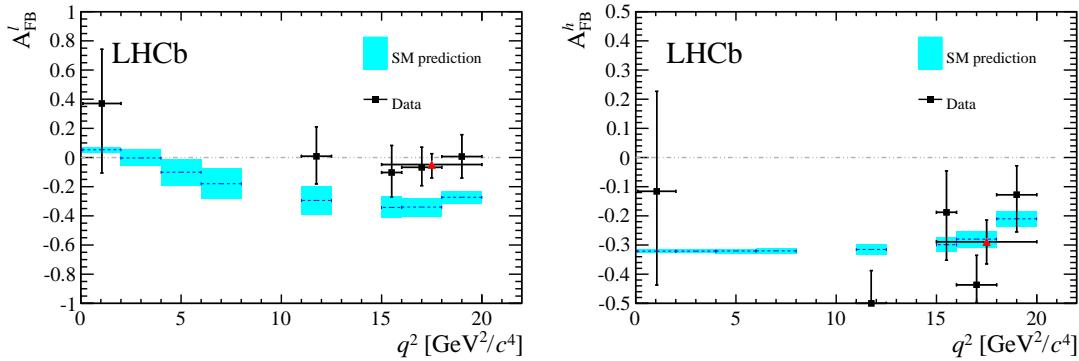


Figure 11.8: Measured values of (left) the leptonic and (right) the hadronic forward-backward asymmetries in bins of q^2 . Data points are only shown for q^2 intervals where a statistically significant signal yield is found, see text for details. The (red) triangle represents the values for the $15 < q^2 < 20 \text{ GeV}^2/c^4$ interval. Standard Model predictions are obtained from Ref. [?].

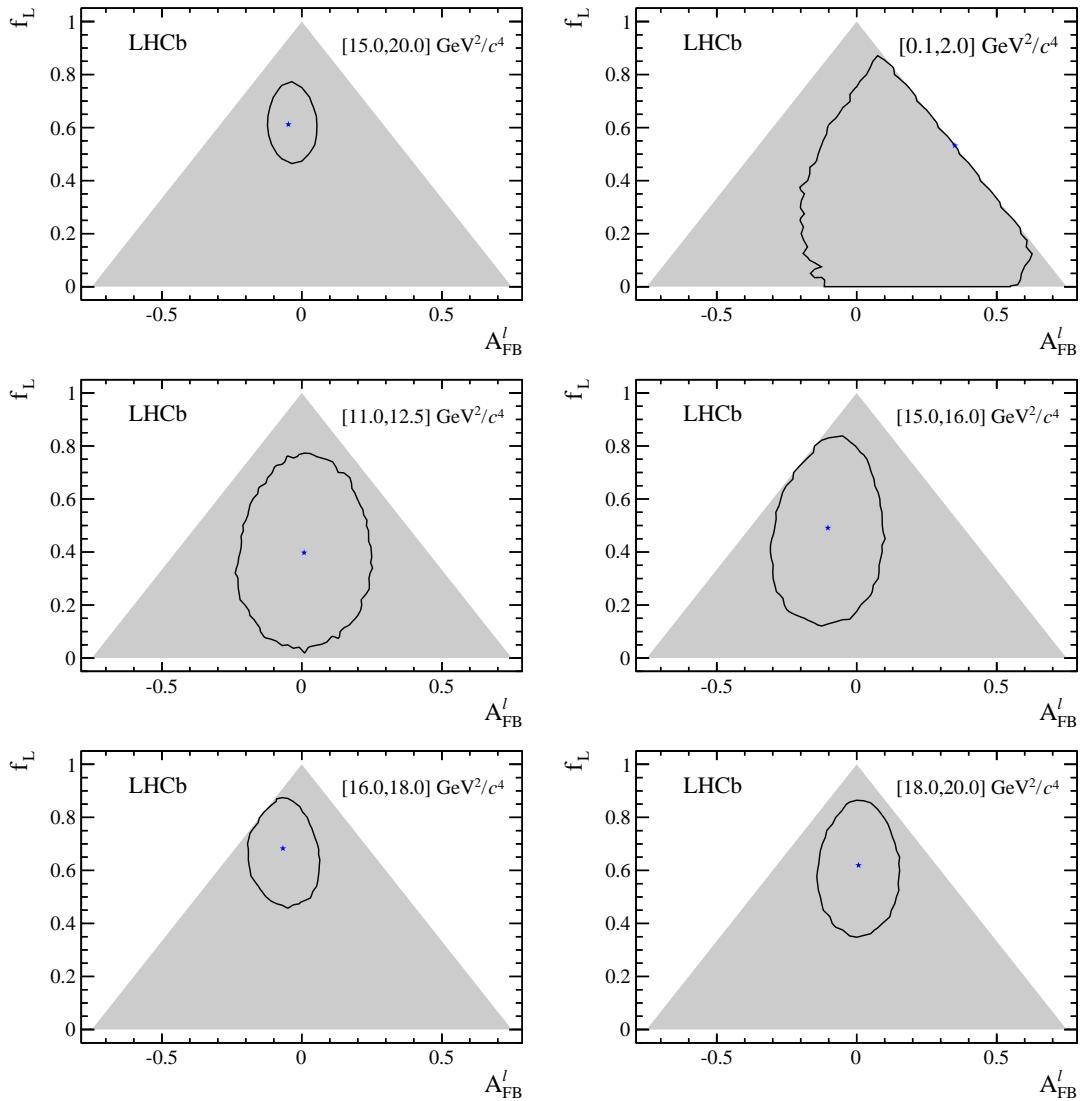


Figure 11.9: Two-dimensional 68 % CL regions (black) as a function of A_F^l and f_L . The shaded areas represent the regions in which the PDF is positive over the complete $\cos \theta_\ell$ range. The best fit points are indicated by the (blue) stars.

Part II

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

1594

CHAPTER 12

1595

1596

Testing lepton flavour universality with R_H

1597

1598 Lepton Favour Universality (LFU) is the equality of the weak coupling constants for
1599 all leptons and can be broken in NP scenarios. Flavor-Changing Neutral Currents
1600 (FCNCs) processes, forbidden in the SM at tree level and happening only at one
1601 loop are sensitive to quantum corrections are an idea laboratory for studying LFU.
1602 In fact NP in the loops could break the flavour symmetry.

1603 In this work, decays of $b \rightarrow s\mu^+\mu^- (e^+e^-)$ type, are considered to test LFU between
1604 electrons and muons using penguin decays. In particular, B^0 meson decays decaying
1605 semileptonically to $B^0 \rightarrow K/K^*\ell^+\ell^-$ are considered. Figure 12.1 shows the possible
1606 Feynman diagrams producing such decays while Fig. 12.2 illustrates how the Feyn-
1607 man diagrams of these processes may include new particles. Furthermore, a series of
1608 recent LHCb measurements [14] points a tension with SM predictions, which make
1609 these processes very interesting to better understand the nature of the discrepancy.

1610

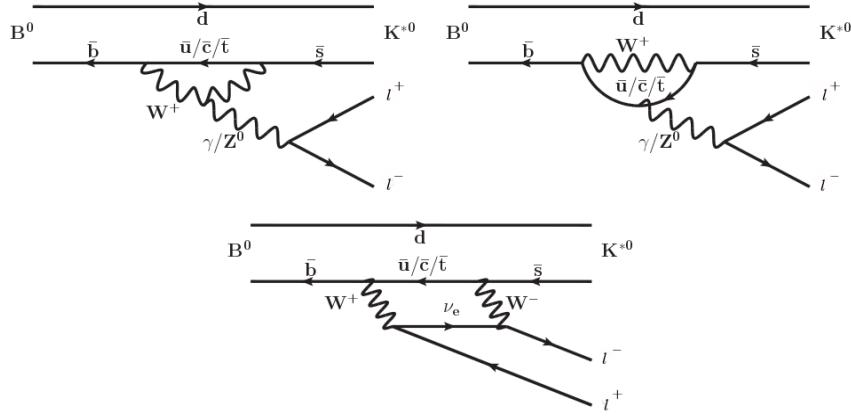


Figure 12.1: Loop diagrams of the $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ process.

1611 In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger pro-
1612 posed the measurement of the R_H ratio [77], defined in Eq. 12.1, where H can be an
1613 inclusive state containing an s quark (X_s) or an s -quark resonance like K or K^{*0} .

$$R_H = \frac{\int_{4m_\mu^2} \frac{d\mathcal{B}(B^0 \rightarrow H \mu^+ \mu^-)}{dq^2} dq^2}{\int_{4m_\mu^2} \frac{d\mathcal{B}(B^0 \rightarrow H e^+ e^-)}{dq^2} dq^2} \quad (12.1)$$

1614 In this quantity the decay width is integrated over the squared dilepton invariant
1615 mass, q^2 , from $q_{min}^2 = 4m_\mu^2$, which is the threshold for the $\mu\mu$ process, up to $q_{max}^2 =$
1616 m_b^2 . The advantage of using ratios of branching fractions is that, in the theoretical
1617 prediction, hadronic uncertainties cancel out.

1618 The advantage of using these observables is that, in the theoretical prediction,
1619 hadronic uncertainties cancel out. Furthermore, experimentally, some of the sys-
1620 tematic uncertainties on the ratios are reduced giving a better measurement. For
1621 example, what is measured is the number of $\mu\mu$ and ee decays which happen in a
1622 certain period of time and then the luminosity (\mathcal{L}) is used to obtain a cross section
1623 (σ), using $R = \mathcal{L}\sigma$, where R is the rate with which a decay happens. The luminosity
1624 measurement is usually a source of systematic uncertainty, however it appears on
1625 both sides of the ratio and therefore cancels out.

1626 Since the SM does not distinguish between lepton flavours the predicted value of

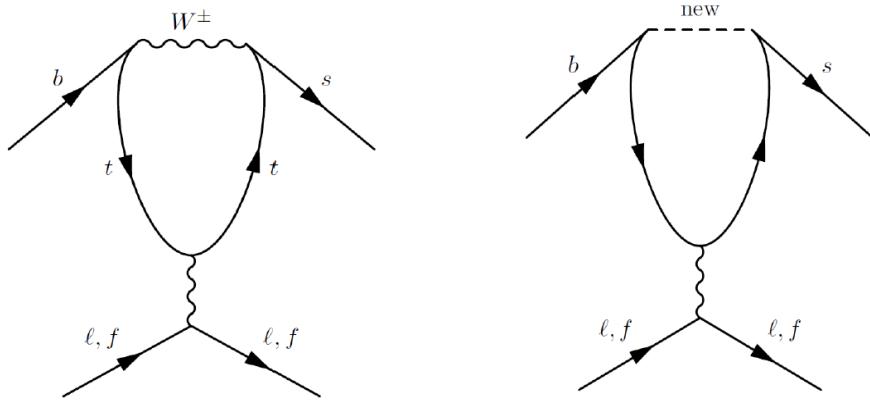


Figure 12.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

1627 this ratio is $R_H = 1$, under the assumption of massless leptons. Taking into account
1628 effects of order m_μ^2/m_b^2 Hiller and Kruger calculate that in the SM and in the full q^2
1629 range [77]:

$$R_{X_s} = 0.987 \pm 0.006 \quad (12.2)$$

$$R_K = 1.0000 \pm 0.0001 \quad (12.3)$$

$$R_{K^{*0}} = 0.991 \pm 0.002 \quad (12.4)$$

$$(12.5)$$

1630 under the assumptions that:

- 1631 • right-handed currents are negligible;
- 1632 • (pseudo-)scalar couplings are proportional to the lepton mass;
- 1633 • there are no CP-violating phases beyond the SM.

1634 The measurement of the R_H ratios is of particular interest after the recent measure-
1635 ment of the branching ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay [22] where no evidence of NP

1636 was found. In fact the quantities $R_H - 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ remain proportional
1637 with

$$\frac{R_H - 1}{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)} \sim 2 \cdot 10^{-5} \quad (12.6)$$

1638 A joint measurement of this two quantities can give much information and constrain
1639 MFV models. If $R_X = 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ is close to the SM prediction as it is
1640 measured to be this will allow to put strong constraints on extensions of the SM. If
1641 instead $R_H > 1$ and the equation above is not verified, this would mean that one of
1642 the assumptions listed above are not verified, which can happen in some extensions
1643 of the SM, for example Super-Symmetric models with broken R-parity. A series of
1644 recent LHCb measurements [14] shows tensions with SM predictions, which makes
1645 it interesting to further investigate these processes.

1646 12.0.1 Combining ratios

1647 The full power of the R_H ratios in understanding new physics scenarios comes from
1648 their combinations. In Ref. [78] Hiller and Schmaltz propose the measurement of
1649 the of double ratios, $X_H = R_H/R_K$, which not only can test LFU but also allow
1650 to disentangle the kind of new physics that lies behind. These ratios are in fact
1651 sensitive also to FCNCs of right-handed quarks, Furthermore in Ref. [78] the study
1652 is extended to B_s^0 decays such as $B_s^0 \rightarrow \phi \ell^+ \ell^-$ or $B_s^0 \rightarrow \eta \ell^+ \ell^-$.

1653 Parity and Lorentz invariance require that the Wilson Coefficients with left-handed
1654 chirality (C) and their right-handed counterparts (C') appear in the decay amplitude
1655 of exclusive decays in determined combinations, e.g.

$$\begin{aligned} C + C' : & K, K_{\perp}^*, \dots \\ C - C' : & K_0(1430), K_{0,\parallel}^*, \dots \end{aligned} \quad (12.7)$$

1656 where the labels for the K^* meson represent its longitudinal (0), parallel (\parallel) and
1657 perpendicular (\perp) transversity components. The C contributions are universal to
1658 all decays and therefore X_H double ratios are sensitive to right-handed currents. In

1659 fact the R_H ratios can be expressed in terms of their deviation from unity as

$$\begin{aligned} R_K &\simeq 1 + \Delta_+ \\ R_{K_0(1430)} &\simeq 1 + \Delta_- \\ R_K^* &\simeq 1 + p(\Delta_- - \Delta_+) + \Delta_+ \end{aligned} \tag{12.8}$$

1660 where the Δ_\pm quantities are combinations of Wilson coefficients reported in Eq.
1661 10 of Ref. [78] and the parameter p is the polarisation of K^* that in Ref. [78] is
1662 determined to be close to 1 simplifying the formula to $R_K^* \simeq 1 + \Delta_-$. In particular
1663 one can observe the following correlations: $R_K < 1$, as it is measured to be, and
1664 $X_{K^*} > 1$ points to dominant BSM contributions into C_{LR} (see definition in Sec. ??),
1665 a SM like $R_K \sim 1$ together with $X_{K^*} \neq 1$ requires BSM with $C_{LL} + C_{RL} \simeq 0$ and,
1666 finally, $R_K \neq 1$ and $X_{K^*} \simeq 1$ corresponds to new physics in C_{LL} .

1667 12.0.2 Experimental status

1668 The R_K and $R_{K^{*0}}$ ratios have already been measured at the B-factories [79, 80],
1669 and the R_K ratio has been also recently measured at LHCb [81] in the $1 < q^2 < 6$
1670 GeV^2/c^4 , which represents the most precise measurement to date. This measurement
1671 manifests a 2.6σ deviation from the SM prediction. The current experimental status
1672 is summarised in Tab. ???. By profiting of the large dataset collected during Run-I,
1673 the LHCb experiment is expected to reduce the uncertainty on $R_{K^{*0}}$ by at least a
1674 factor of 2 with respect to the B-factories.

Table 12.1: Experimental status of the $R_{K^{(*)0}}$ measurements.

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

1676 12.1 Analysis strategy

1677 The aim of this analysis is to measure the $R_{K^{*0}}$ ratio using pp collision data collected
 1678 by the LHCb detector in 2011 and 2012, corresponding to a total of 3 fb^{-1} of
 1679 integrated luminosity. The $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$, “rare channels”,
 1680 are reconstructed with the K^{*0} decaying into a kaon and a pion with opposite
 1681 charges.

1682 The analysis has to separate signal candidates from background candidates which
 1683 have similar observed properties. The selection presented in Sec. 13 aims to max-
 1684 imise the yield while minimising the background contamination. Two types of back-
 1685 grounds are identified: “peaking background” and “combinatorial background”. The
 1686 first comes from the mis-reconstruction of other decays or from partially recon-
 1687 structed events. This type of background, because its specific kinematic properties,
 1688 usually peaks in some variable, such as the invariant mass of all final particles,
 1689 therefore we can remove these events by removing the peak. The combinatorial
 1690 background instead comes from the random combination of particles and can be
 1691 lowered selecting events with good-quality tracks and vertices.

1692 To further reduce systematic uncertainties the measurement is performed as a dou-
 1693 ble ratio as shown in Eq. 12.9 where decays reaching the same final states as the
 1694 rare channels via a J/ψ resonance, $B^0 \rightarrow K^{*0}J/\psi(\rightarrow \ell^+\ell^-)$, also referred as “char-
 1695 monium” or “resonant” channels, are used as control samples. These decays are
 1696 distinguished from the rare channels using the invariant mass of the dilepton pair.

1697 In Sec. 15 the efficiency of selecting and reconstructing each channel is extracted
 1698 and, finally, in Sec. 17 the R_{K^*} ratio defined is built as the double ratio of rare and
 1699 resonant channels:

$$R_{K^{*0}} = \frac{N_{B^0 \rightarrow K^{*0}\mu^+\mu^-}}{N_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}} \cdot \frac{N_{B^0 \rightarrow K^{*0}J/\psi e^+e^-}}{N_{B^0 \rightarrow K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}}{\varepsilon_{B^0 \rightarrow K^{*0}\mu^+\mu^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}e^+e^-}}{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow e^+e^-}} \quad (12.9)$$

¹⁷⁰⁰ As NP is expected not to affect charmonium resonances the ratio of the J/ψ channels
¹⁷⁰¹ is 1 and therefore $R'_{K^{*0}} = R_{K^{*0}} \times R_{J/\psi} = R_{K^{*0}}$. On the other hand using the relative
¹⁷⁰² efficiencies between the rare and resonant channels allows to cancel out many effects
¹⁷⁰³ resulting in a better control of systematic uncertainties.

¹⁷⁰⁴ For brevity, the rare channels will also be denoted as “ $\ell\ell$ ”, or specifically “ ee ” and
¹⁷⁰⁵ “ $\mu\mu$ ”, and the resonant channels as “ $J/\psi(\ell\ell)$ ”, or “ $J/\psi(ee)$ ” and “ $J/\psi(\mu\mu)$ ”.

¹⁷⁰⁶ 12.2 Choice of q^2 intervals

¹⁷⁰⁷ Two q^2 intervals are considered in this work:

- ¹⁷⁰⁸ • the “central” q^2 region, $[1.1, 6.0]$ GeV^2/c^4 ;
- ¹⁷⁰⁹ • the “high” q^2 region, above 15 GeV^2/c^4 .

¹⁷¹⁰ The central q^2 region is the most interesting place to look for new physics. In fact,
¹⁷¹¹ at low q^2 , below 1 GeV^2/c^4 the photon pole dominates leaving little space for NP
¹⁷¹² to be found [14]. The upper bound of this interval is set at 1.1 GeV^2/c^4 , in order
¹⁷¹³ entirely include the contribution from $\phi \rightarrow \ell^+ \ell^-$ decays, that can dilute new physics
¹⁷¹⁴ effects, into the low q^2 interval. The upper bound of the central interval is chosen
¹⁷¹⁵ to be sufficiently far away from the J/ψ radiative tail, where predictions cannot be
¹⁷¹⁶ cleanly extracted. The 6 – 15 GeV^2/c^4 region is characterised by the presence of the
¹⁷¹⁷ narrow peaks of the J/ψ and $\psi(2S)$ resonances. The lower bound of the high q^2
¹⁷¹⁸ region, where the signal in the electron channel is still unobserved, is chosen to be
¹⁷¹⁹ sufficiently far from the $\psi(2S)$ resonance. Rare and resonant channels are selected
¹⁷²⁰ depending on which q^2 interval they fall in (for details see Sec. 13).

₁₇₂₁ **12.3 Data samples and simulation**

₁₇₂₂ This analysis is based on a data set corresponding to 3 fb^{-1} of integrated luminosity
₁₇₂₃ collected by the LHCb detector in 2011 and 2012. In order to study the background
₁₇₂₄ properties, determine efficiencies and to train the multivariate analysis simulated
₁₇₂₅ events are used. After the hard interactions are generated with Pythia8 hadronic
₁₇₂₆ particles are decayed using EvtGen and, finally, propagated into the detector us-
₁₇₂₇ ing Geant4 and reconstructed with the same software used for data. Samples are
₁₇₂₈ generated with both 2011 and 2012, magnet up and down conditions and are com-
₁₇₂₉ bined in the right proportions, according to the luminosity registered on data. The
₁₇₃₀ next section describes the corrections applied to the simulation to obtain a better
₁₇₃₁ description of data.

₁₇₃₂ **12.3.1 Data-simulation corrections**

₁₇₃₃ Since the multivariate classifier training (see Sec. 13.4) and the calculation of most
₁₇₃₄ of the efficiency components (see Sec. ??) are obtained from the study of simulated
₁₇₃₅ events it is important to verify that the simulation is a reliable reproduction the
₁₇₃₆ data. In particular it is important to match data and Monte Carlo in the kinemat-
₁₇₃₇ ics of the final particles and the occupancy of the detector. The kinematics of the
₁₇₃₈ decays is characterised by the transverse momentum spectrum of the B^0 . Discrep-
₁₇₃₉ ancies in this distribution cause also the spectra of the final particles to differ from
₁₇₄₀ data and affect the efficiency determination as its value often depends on the mo-
₁₇₄₁ mentum distribution of final particles. The occupancy of the detector is correlated
₁₇₄₂ to the invariant mass shape of the signal because the addition of energy clusters in
₁₇₄₃ the electromagnetic calorimeter, affects the electron momenta for bremsstrahlung
₁₇₄₄ photons emitted before the magnet. The hit multiplicity in the SPD detector is a
₁₇₄₅ proxy for the detector occupancy.

₁₇₄₆ Since it is important that these quantities are well modelled, the simulation is
₁₇₄₇ reweighted so that the distributions in data and simulation match for these vari-

ables. This can be done using resonant $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \ell^+ \ell^-)$ events, for which the signal peak is already visible in data after pre-selection (see Sec. 13). However, the data still includes a high level of background and distributions cannot be directly compared. The $s\mathcal{P}$ lot technique [82] is used to statistically subtract the background from data and obtain pure signal distributions using the invariant mass as control variable. Fig. 12.3 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. 12.4 and ratios of these distribution, which are used to re-weight the simulation, are reported in Fig. 12.5. 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. 15.3 and, finally, ee samples are also reweighted for L0 trigger efficiency as described in Sec. 15.4. Weights are always applied throughout unless specified.

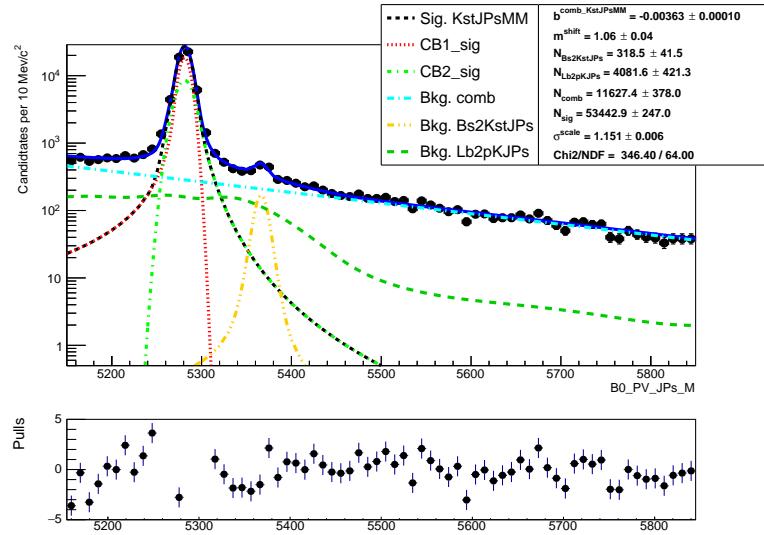


Figure 12.3: Fitted 4-body invariant mass distributions of muonic resonant candidates.

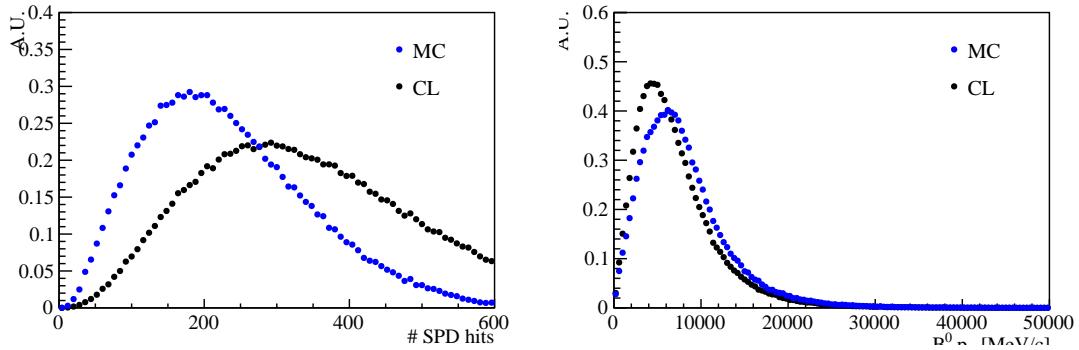


Figure 12.4: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

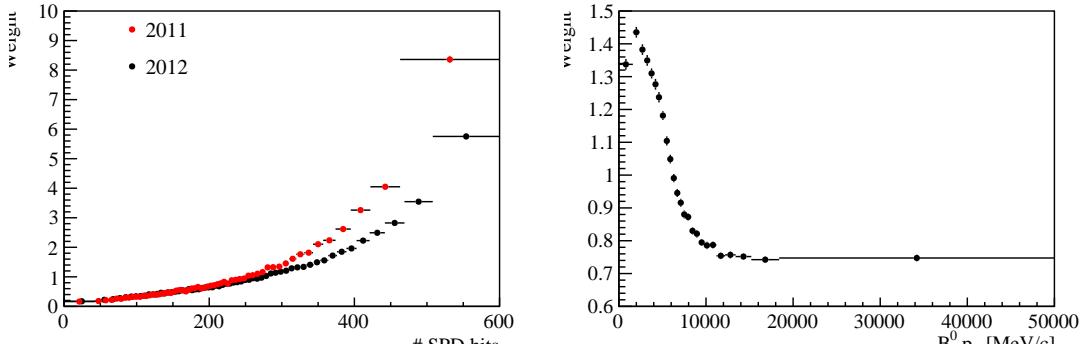


Figure 12.5: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

1768

CHAPTER 13

1769

1770

Selection

1771

1772 The selection process, described in the following subsections, is divided into several
1773 steps. First of all events have to fall into the detector acceptance, produce hits
1774 and be selected on the basis of quality features, such as χ^2 of tracks and vertices,
1775 this stage is called “stripping”. Secondly it is required that some specific trigger
1776 lines were fired by the events. After the trigger and stripping requirements, cuts
1777 are applied to remove backgrounds from specific decays. These first three steps are
1778 referred to as “pre-selection”.

1779 The next step consists in the application of particle identification (PID) conditions
1780 which remove a good part of misreconstructed background and clear the way for
1781 the last step where a neural network is used to remove combinatorial background.

1782 In order to minimise systematic uncertainties the same selection requirements are
1783 applied to the rare signal candidates and on their relative charmonium channel, a
1784 part from the q^2 cuts which serve to distinguish them. In order to identify the
1785 $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ channel a dilepton mass interval of 100 MeV/ c^2 around the

nominal J/ψ peak [3] 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 ?? 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.

13.1 Trigger and Stripping

Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported in Tab. ??, 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.

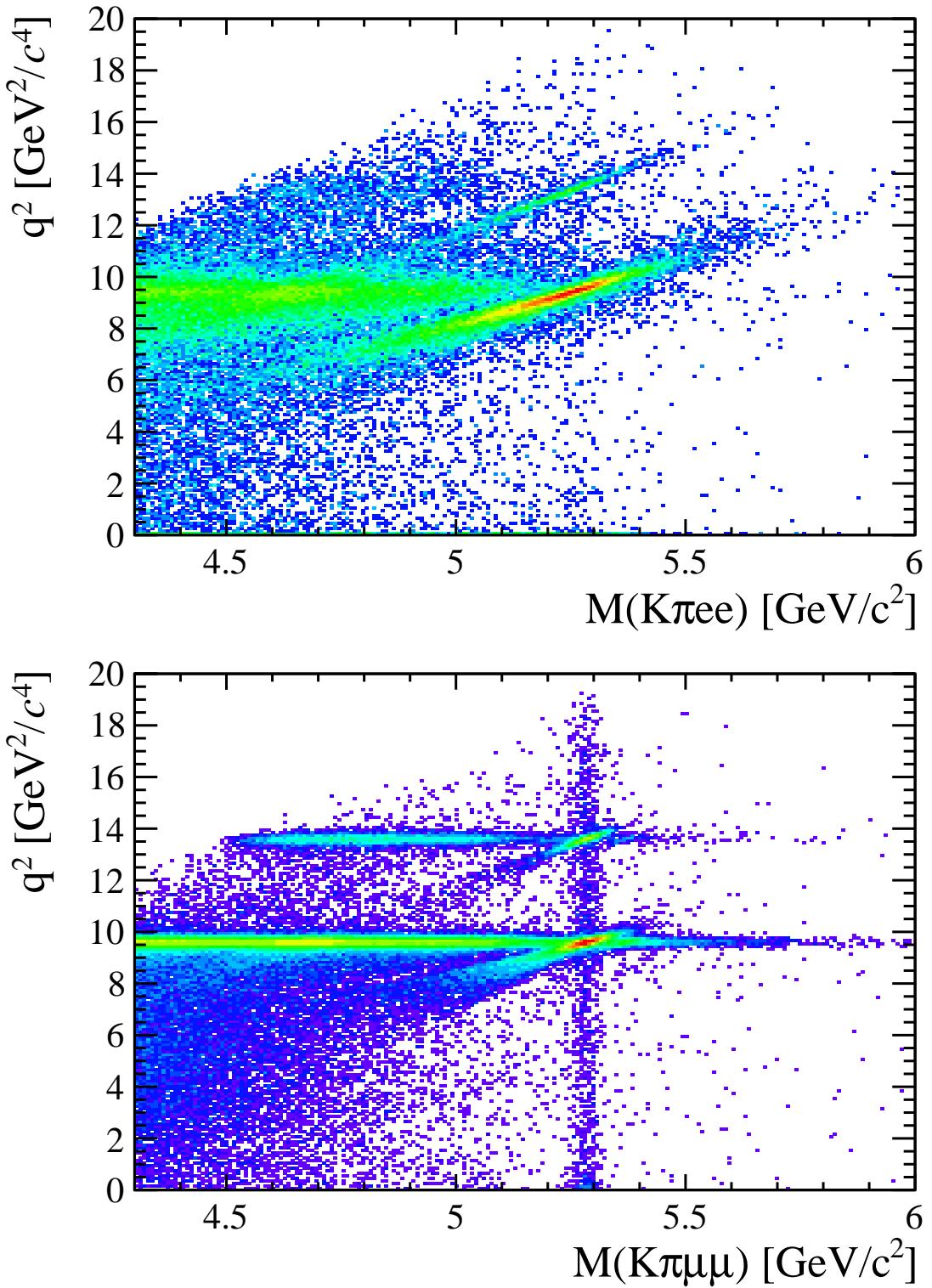


Figure 13.1: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electron (top) and muonic (bottom) channels in 2012 data.

Table 13.1: 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

1813 For the electron channels the L0 lines have different properties, therefore the analysis
1814 is performed separately for three categories of events, depending on the L0 trigger
1815 that fired them. These categories are defined, to be exclusive, in the following way:

- 1816 • Events triggered by at least one of the electrons in the signal candidate (L0E):
1817 `L0Electron_TOS`
- 1818 • Events triggered by at least one of the hadrons in the signal candidate and
1819 not by L0Electron (L0H):
1820 `L0Hadron_TOS && !L0Electron_TOS`
- 1821 • Events triggered by particles not in the signal candidate (Trigger Independent
1822 of Signal, TIS) and not by the previous cases (L0I):
1823 `L0_TIS && !(L0Electron_TOS || L0Hadron_TOS)`

1824 The majority of the selected events falls in the L0Electron category. The L0Hadron
1825 category is more efficient at low q^2 were the K^{*0} has more momentum.

1826 Candidates are then required to pass the kinematic and quality cuts summarised in
1827 Tab. ???. The meaning of variables in the table was already explained in Sec. ???.
1828 Loose PID cuts are applied in preselection to limit the size of the samples, while
1829 tighter cuts are applied in a second stage. A large mass window is kept around
1830 the B^0 peak in order to be able to fit the sideband and to train the multivariate

Table 13.2: Summary of stripping cuts used for the central and high q^2 regions.

Particle	Cuts
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> $PID_e > 0$
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

1831 analysis and constrain backgrounds. Track-quality and vertex quality cuts are also
1832 applied using the $\chi_{\text{track}}^2/\text{ndf}$, `GhostProb`, and $\chi_{\text{vtx}}^2/\text{ndf}$ variables. The `GhostProb`
1833 quantity describes the probability of a track being fake. By construction cutting at
1834 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. For details about the definition of
1835 the variables used see Ref. [83].

1836 13.2 PID

1837 After preselection there still are high levels of misreconstructed background. In par-
1838 ticular, as the ID of kaons and pions are not constrained, the samples still contain
1839 both ID combinations for most candidates, therefore tighter PID cuts are applied.
1840 In the LHCb analysis framework the particle identification probability can be quan-

1841 tified using the “ProbNN” variables [84]. These variables are the output of a Neural
1842 Network which takes as input information from the calorimeters, the RICH detec-
1843 tors and the muon system. Unlike the DLL variables these are bounded from 0 to 1
1844 and can be therefore directly be interpreted as probabilities. For example `ProbNNk`
1845 is the probability for a reconstructed particle to be a kaon. Two tunes of the ProbNN
1846 variables, labelled V2 and V3, are available. Tune V3 was shown to be optimal for
1847 positive ID, while tune V3 was found to be optimal for background rejection and
therefore it is used to quantify the mis-ID probability.

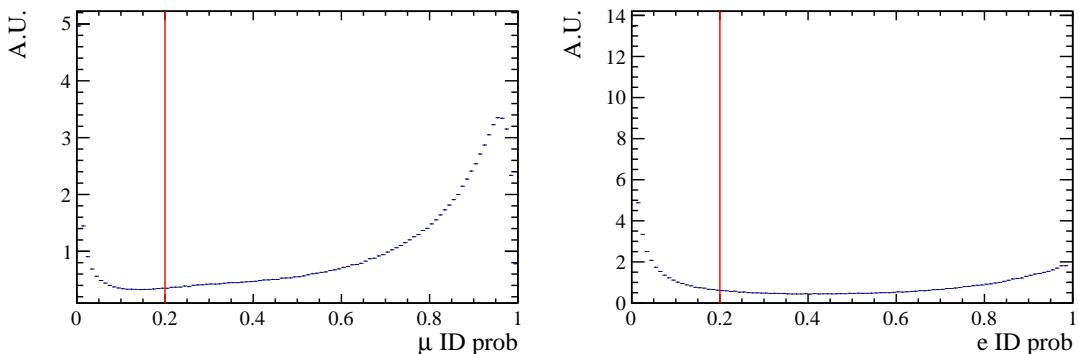


Figure 13.2: Correct ID probability distributions for muons (left) and electron (right) in 2012 data.

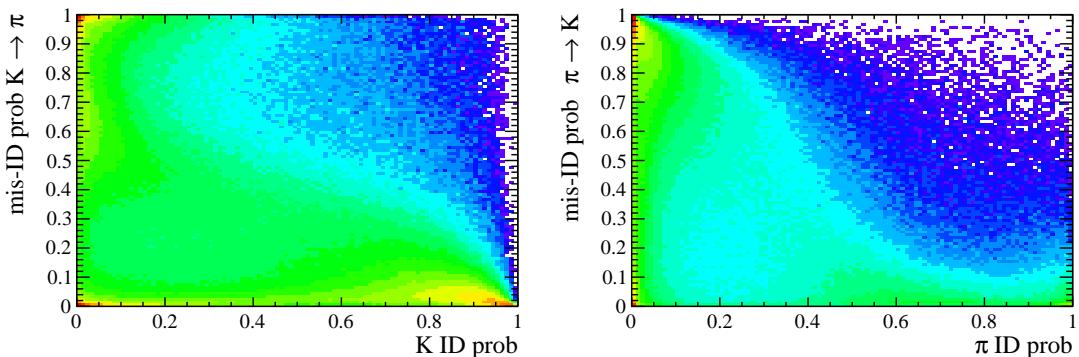


Figure 13.3: On the horizontal axis of these plots is shown the correct ID proba-
bilities for kaons (left) and pions (right), while the vertical axis show the mis-ID
probability.

1848

1849 Figure 13.2 shows distributions of the correct ID variables in the 2012 data sample
1850 while Fig. 13.3 shows in a two-dimensional plane the probability of correct identifi-
1851 cation and mis-identification of kaons and pions. These plots are characterised by

1852 clear peak at maximal ID probability and minimal mis-ID probability, corresponding
1853 to particles to which is possible to assign a well defined identification.

1854 In order to maximise the power of the PID cuts probabilities of correct ID and
1855 mis-ID are combined using the following cuts:

$$\pi \rightarrow \text{ProbNNpi-V3} \times (1 - \text{ProbNNk-V2}) \times (1 - \text{ProbNNp-V2}) > 0.1$$

$$K \rightarrow \text{ProbNNk-V3} \times (1 - \text{ProbNNp-V2}) > 0.05$$

$$\mu \rightarrow \min(\text{ProbNNmu-V3}, \text{ProbNNmu-V3}) > 0.2$$

$$e \rightarrow \min(\text{ProbNNe-V3}, \text{ProbNNe-V3}) > 0.2$$

1857 In the first formula, for example, ProbNNpi is the probability of correctly identifying
1858 the pion as a pion, while ProbNNk is the probability of mistaking it for a kaon. There-
1859 fore by maximising the quantity “ $\text{ProbNNpi} \times (1 - \text{ProbNNk})$ ”, one can maximise
1860 the correct ID probability and minimise at the same time the mis-ID probability.

1861 13.3 Peaking backgrounds

1862 Cuts are applied in order to remove background sources due to specific decays.
1863 These types of backgrounds usually peak in some variable because of their mass or
1864 distinctive kinematic properties and therefore they can be removed without signifi-
1865 cant signal efficiency loss. In the following sections are described the main sources
1866 of peaking background.

1867 13.3.1 Charmonium vetoes

1868 Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2 . The choice of q^2 binning
1869 described in Sec. 12.2 constitutes a natural veto for these decays. Simulated events
1870 were used to check if resonant events leak inside the q^2 intervals chosen for the
1871 rare channel analysis. For the muonic channels the leakage is negligible as the

peaks are sharper due to a better resolution and muons emit fewer bremsstrahlung photons, resulting in shorter radiative tails. The electronic channels are instead characterised by a worse resolution and at the same time electrons can radiate several bremsstrahlung photons, yielding long tails at low q^2 . Analysing Monte Carlo events it was found that 1.3 - 2% (depending on the trigger category) of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candidates leak into the $1.1 < q^2 < 6$ GeV^2/c^4 interval and 1.8% of $\psi(2S)$ events leak above 15 GeV^2/c^4 . The contribution from these candidates is modelled in the fit.

13.3.2 ϕ veto

It can happen that a kaon from the decay $B_s \rightarrow \phi\ell^+\ell^-$, where the ϕ decays in two kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as a K^{*0} . This results in a candidate with a value of $m_{K\pi}$ that is less than $m_{K^{*0}}$ but still high enough to pass the selection requirements. In Fig. 13.4 is reported the plot of $m_{K\pi}$ versus $m_{K\pi\ell\ell}$, where kaon mass hypothesis is assigned to the pion. A peak can clearly be seen around the ϕ mass ($1020 \text{ MeV}/c^2$). To remove this background only candidates with $m_{K(\pi \rightarrow K)} > 1040 \text{ MeV}/c^2$ are selected. This results in a 98% background rejection while keeping a 99% signal efficiency. The ϕ could also constitute a background when it decays into two leptons but the branching ratio of this decay is small compared to the one into kaons and this contribution is taken into account by the choice of the q^2 intervals.

1892

13.3.3 $B^+ \rightarrow K^+\ell^+\ell^-$ plus a random pion

Some $B^+ \rightarrow K^+\ell^+\ell^-$ decays can contaminate the upper B^0 mass sideband if they are combined with a soft pion from somewhere else in the event and therefore reconstructed as a B^0 decay. The same can also happen with a kaon misidentified as a

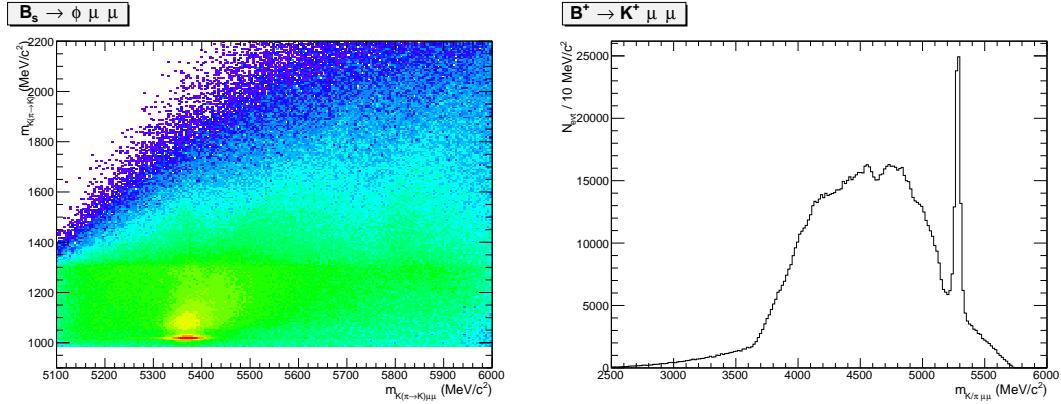


Figure 13.4: On the left the distribution of 2011 data events on the variables ($m_{K(\pi \rightarrow K)}$) and ($m_{K(\pi \rightarrow K)\mu\mu}$), where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. On the right the mass of the three-body system ($m_{K\mu\mu}$) where the peak due to the $B^+ \rightarrow K^+ \mu\mu$ decay is visible.

1897 pion combined with an other kaon in the event. In Fig. 13.4 the invariant mass dis-
1898 tribution of the three-body $K\mu^+\mu^-$ system, $m_{K\mu\mu}$, is shown. This is characterised
1899 by a narrow peak at the B^+ mass. Since these candidates have $m_{K\pi\ell\ell} > 5380$
1900 MeV/c² there is no contribution under the B^0 peak, but they can cause problems
1901 when using sidebands events to train the neural network. An effective veto for this
1902 decay was found to be $\max(m_{K\ell\ell}, m_{K \rightarrow \pi\ell\ell}) < 5100$ MeV/c², which results in 95%
1903 background rejection while keeping 99% signal efficiency.

1904 13.3.4 Λ_b decays

1905 $\Lambda_b^0 \rightarrow \Lambda J/\psi$ decays are unlikely to be reconstructed as $B^0 \rightarrow K^{*0}\ell^+\ell^-$ because the
1906 Λ is long-lived and decays further in the detector with a separate vertex. Simulated
1907 events were used to check how many candidates fall into the B^0 samples, which
1908 results to be negligible. The $\Lambda_b^0 \rightarrow J/\psi pK$ decay can instead contribute more easily
1909 since the $m(pK)$ is above the Λ threshold and therefore they must come from Λ^*
1910 resonances, which are not long-lived. This background is already reduced using
1911 PID but a non-negligible contribution is still expected in the $\mu\mu$ sample, which is
1912 modelled in the fit.

¹⁹¹³ 13.3.4.1 Other peaking backgrounds

¹⁹¹⁴ A possible background could come from $B^0 \rightarrow K^*\gamma$ decays where the photon con-
¹⁹¹⁵ verts into two electrons while traversing the detector. In LHCb, around 40% of
¹⁹¹⁶ photons convert before the calorimeter, but only a small fraction of these, $\sim 10\%$,
¹⁹¹⁷ are reconstructed. Furthermore these events fall into a q^2 region well below the
¹⁹¹⁸ intervals considered in these analysis and their contribution is therefore negligible.
¹⁹¹⁹ Similar decays are also $B^0 \rightarrow K^*\eta$ and $B^0 \rightarrow K^*\pi^0$ where η and the pion decay
¹⁹²⁰ into two photons. Once again the contribution from these decays falls well below
¹⁹²¹ the considered q^2 intervals. Finally, a potentially dangerous background could come
¹⁹²² from events where the identity of the kaon and the pion are swapped as these candi-
¹⁹²³ dates peak under the signal. Their contribution is found to be small, 0.5%, however
¹⁹²⁴ the effect of their modelling into the fit is taken into account in the systematic
¹⁹²⁵ uncertainties.

¹⁹²⁶ 13.3.5 Mis-reconstructed background

¹⁹²⁷ A source of mis-reconstructed background is due to cascade decays with a B^0 de-
¹⁹²⁸ caying semileptonically into a D meson which also decays semileptonically, e.g.
¹⁹²⁹ $B^0 \rightarrow D^-\ell^+\bar{\nu}_\ell$ followed by $D^- \rightarrow K^{*0}\ell^-\nu_\ell$. The candidates built from these de-
¹⁹³⁰ cays tend to have a low 4-body invariant mass as two or more particles are not
¹⁹³¹ reconstructed.

¹⁹³² In order to remove this background in the muonic channels, the 4-body $m(K\pi\mu^+\mu^-)$
¹⁹³³ invariant mass is recalculated with a kinematical fit using DecayTreeFitter [?]. In
¹⁹³⁴ the resonant case this includes a constraint of the dilepton mass to be the J/ψ nom-
¹⁹³⁵ inal mass and in both rare and resonant cases each particles is constrained to point
¹⁹³⁶ to its origin vertex. This constraint has the effect of pushing the misreconstructed
¹⁹³⁷ events far from the B^0 peak. Therefore, to avoid this background, it is sufficient to
¹⁹³⁸ limit the analysis to 4-body invariant masses above $5150 \text{ MeV}/c^2$.

1939 In the electron case it is instead important to fit a wider mass window to correctly
1940 constrain the background therefore one cannot eliminate this mis-reconstructed
1941 background which is then modelled in the fit (for details see Sec. 14.2.2).

1942 13.4 Multivariate analysis

1943 The final selection is performed using a Neural Network classifier (NN) based on the
1944 NeuroBayes package [49, 50]. The multivariate analysis is intended to remove some
1945 combinatorial background and obtain a clearer signal peak.

1946 For the final selection in the central and high q^2 intervals a Neural Network classifier
1947 (NN) is used based on the NeuroBayes package [49, 50]. Representative samples of
1948 the signal and background are needed to train the classifier. For the signal, fully
1949 reconstructed $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$ simulated events can be used. A
1950 sample representative of the background can be obtained taking real data events in
1951 the upper B^0 sideband: ($m_{K\pi\mu^+\mu^-} > 5400$ MeV/ c^2 and $m_{K\pi e^+e^-} > 5600$ MeV/ c^2).
1952 The lower sideband is not used in the training as it contains a significant fraction of
1953 mis-reconstructed background. All pre-selection cuts are applied to the background
1954 samples used for the training. As L0 and PID variables are not well described these
1955 cuts are not applied in the Monte Carlo samples but their effect is taken into account
1956 by the event weight. To train the classifier 50% of the sideband events was used,
1957 keeping the other 50% for testing. For the signal sample a number of Monte Carlo
1958 events was used equal to the number available for the background sample.

1959 The input to the NN consists of 22 variables containing information about the
1960 kinematic of the decays and the quality of tracks and vertices. All the variables used
1961 are listed in Tab. 13.3. The graphical representation of the correlation matrices are
1962 shown in Fig. 13.5, in these figures the variable with ID = 1 is the NN output and
1963 the other IDs are reported in Tab. 13.3. The single most discriminating variable
1964 used is the χ^2 of a kinematic fit that constrains the decay product of the B^0 , the
1965 K^* and the dimuon, to originate from their respective vertices. Other variables that

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 13.3: 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. 13.5.

contribute significantly are the χ_{IP}^2 of J/ψ and K^* , 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. 13.4, together with information on the relative importance of each input. The meaning of the column headings in this table was already explained in Sec. 4.1.

Figure 13.6 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 the neural network is able to separate signal from background and that the training converged properly.

It can happen that too much information is given to the classifier, which becomes able to calculate the invariant mass of the candidates from its inputs. This could generate fake peaks and it is therefore important to check for correlations between the B^0 mass and the NN output. Fig 13.7 reports plots of the average NN output as a function of the B^0 mass on sideband data and simulated signal events. The distributions are flat showing that no significant correlation is present.

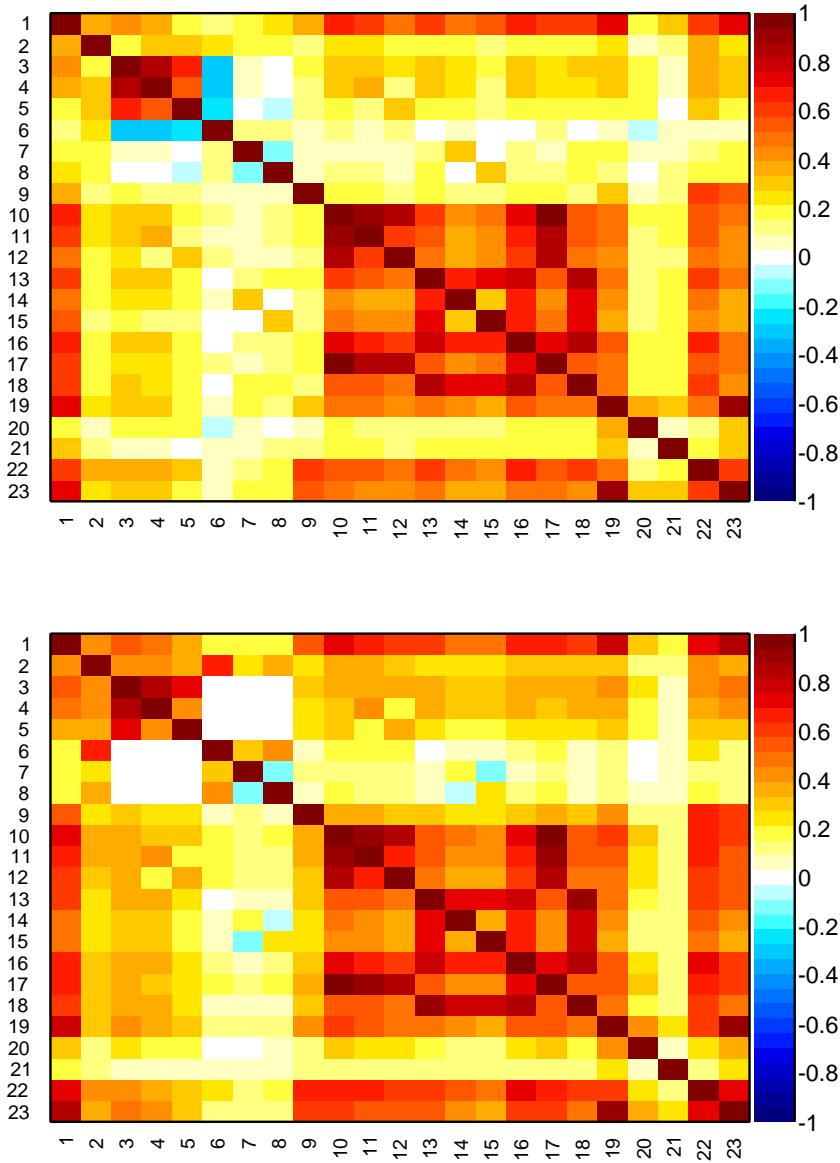


Figure 13.5: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of Tab. 13.4. Correlation is calculated using all events without distinguishing signal and background.

Table 13.4: 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

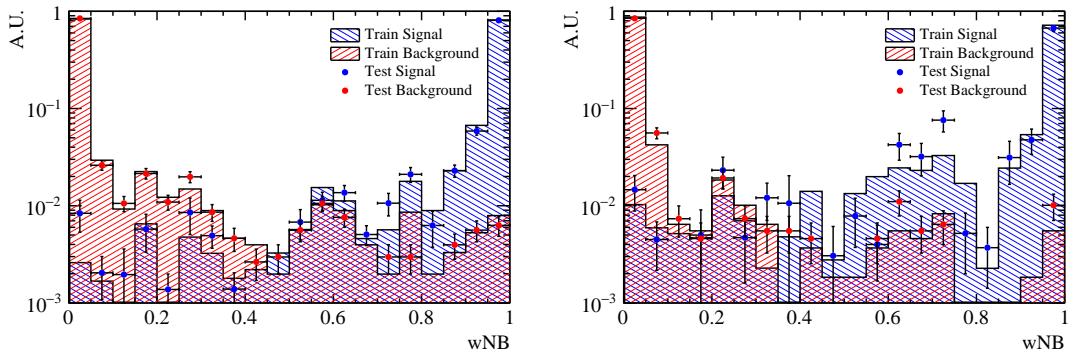


Figure 13.6: NN output distributions for training (solid) and test (stripes) samples, for MC signal and data sideband events. For the muon training (left) and the electron one (right).

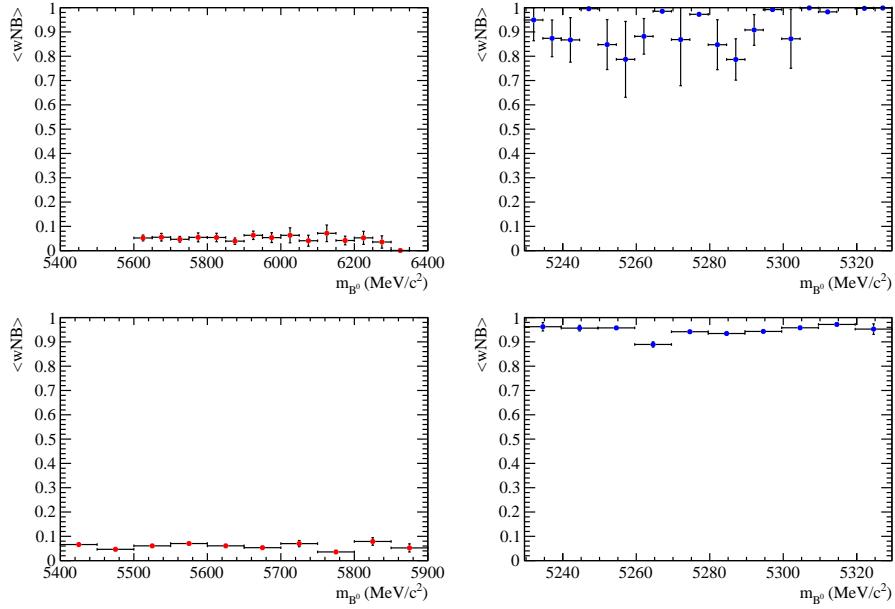


Figure 13.7: Average value of NN output as a function of B^0 mass for data side-band (left) and MC signal (right) events for the electron (top) and muon (bottom) training.

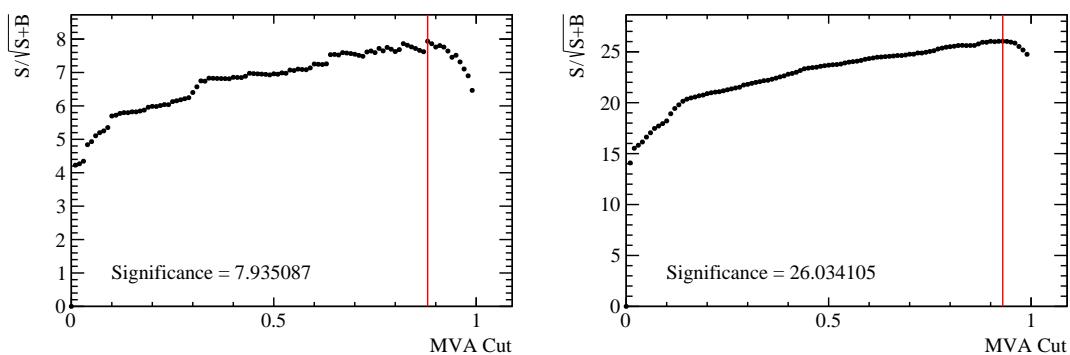


Figure 13.8: Dependence of figure-of-merit on the requirement on neural network output for electrons (left) and muons (right). The vertical lines corresponds to the chosen cuts.

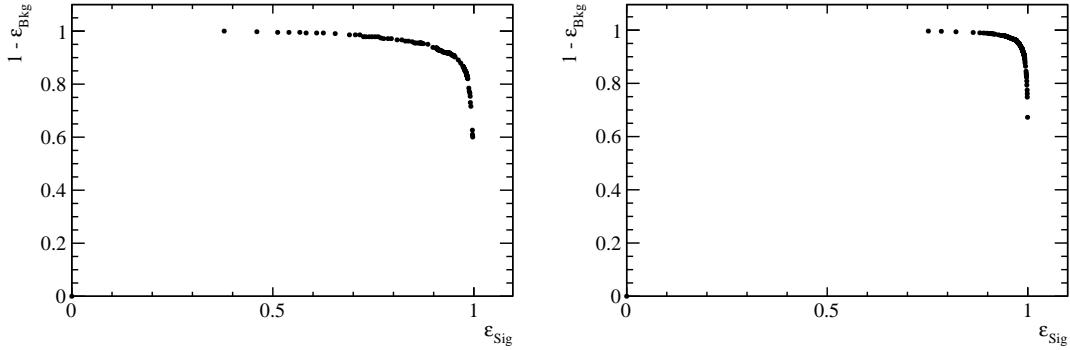


Figure 13.9: Receiver operating characteristic (ROC) curves for electrons (left) and muons (right). They show the signal efficiency versus the background rejection.

1983 13.5 MVA optimisation

1984 In order to optimise the cut on our neural network output the expected signal
 1985 significance, $N_S/\sqrt{N_S + N_B}$, was maximised. In this formula N_S is number of rare
 1986 signal events and N_B the number of background events.

1987 The number of signal events accepted for a given NN output cut is determined
 1988 exploiting the resonant channel and simulation. First, as an arbitrary number of
 1989 events can be simulated, this has to be rescaled to the expected yield. This is done
 1990 by fitting $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+ \ell^-)$ events after pre-selection, including all selection
 1991 cuts except MVA. The resonant yield is then scaled down by the expected ratio
 1992 between the rare and the resonant channels. The number of background events is
 1993 instead derived by fitting the combinatorial background in the sideband with an
 1994 exponential function and extrapolating the fit function below the signal peak.

1995 The dependence of the figure-of-merit for both the electron and muon trainings are
 1996 shown in Fig.13.8, where the red line indicate the chosen cut: 0.75 for both samples.
 1997 Curves of signal efficiency versus background rejection are shown in Fig. 13.9. Using
 1998 the described MVA cuts the signal efficiency is $\sim 91\%$ for the muon channels and
 1999 $\sim 84\%$ for the electron channels (for more details see Sec. 15), while the background
 2000 rejections is $\sim 98\%$ on both samples.

2001 After full selection about $\sim 3\%$ of events still contain multiple candidates which are

2002 removed at random keeping only a single candidate per event.

2003

CHAPTER 14

2004

2005

Mass fits

2006

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

$$N_{\ell\ell} = N_{J/\psi} \cdot \varepsilon^{\text{rel}} \cdot R_{\ell\ell}, \quad (14.1)$$

where ε^{rel} is the relative efficiency between the rare and resonant channels given in Tab. 15.4 and $R_{\ell\ell}$ corresponds to the efficiency corrected ratio of the raw rare and resonant yields:

$$R_{\ell\ell} = \frac{N_{\ell\ell}/\varepsilon^{\ell\ell}}{N_{J/\psi}/\varepsilon^{J/\psi}}. \quad (14.2)$$

2017 The two ratios, R_{ee} and $R_{\mu\mu}$, are then used to determine the $R_{K^{*0}}$ quantity, as
2018 described in Sec. 17.

2019 14.1 Muonic channels

2020 For the rare and resonant $\mu^+\mu^-$ channels the fitted variable is the $m(K\pi\mu\mu)$ in-
2021 variant mass coming from a kinematic fit where all vertices are required to point
2022 to their mother particle. In the resonant case it is beneficial to also constrain the
2023 the dimuon mass to the known J/ψ mass. The effect of the kinematical fit is to
2024 improve the mass resolution by roughly a factor of 2, which results a more stable
2025 fit. Furthermore, misreconstructed events are pushed away from the B^0 peak, which
2026 allows to use a wider mass window to better constrain the combinatorial background
2027 slope. The mass spectrum is fitted in the range $5150 - 5800$ MeV/c² with the lower
2028 limit of the mass range chosen to exclude partially reconstructed background. As it
2029 is not needed to model misreconstructed backgrounds in the fit this also eliminates
2030 systematic uncertainties associated with the knowledge of its shape.

2031 The PDF chosen to describe the signal in both the $B^0 \rightarrow K^*\mu^+\mu^-$ and its relative
2032 J/ψ channel is a Double Crystal Ball function already described in Sec. 5.1 and also
2033 in this case the mean value (m_0) is kept in common:

2034 As a first step simulated distributions are fit using the signal model to extract
2035 parameters to be constrained in the fit to data. The fitted MC distribution for the
2036 resonant channel is reported in Fig. 14.1.

2037 For the fit to real data the signal parameters are fixed to the ones found for the
2038 simulated samples. In order to account for possible data-simulation discrepancies
2039 a scale factor is multiplied to the widths and a shift is added to the masses. In

2040 summary the PDFs used for the signal fits on data are defined as

$$\begin{aligned} P(m; c, m') = & f^* \cdot C(m; \alpha_1^*, n_1^*, c \cdot \sigma_1^*, m_0^* + m') \\ & + (1 - f^*) \cdot C(m; \alpha_2^*, n_2^*, c \cdot \sigma_2^*, m_0^* + m') \end{aligned} \quad (14.3)$$

2041 where the free parameters are the width scale factor, c , and the mass shift, m' ,
 2042 which are common between the rare and resonant samples. The other parameters,
 2043 denoted with $*$, are taken from the fit to simulated events, separately for the rare and
 2044 resonant samples and are fixed in the fit on data. The parameter f^* in the formula
 2045 is the relative fraction of candidates falling in the first Crystal Ball function.

2046 To model the combinatorial background an exponential function was used. This is
 2047 the only background component for the rare channel. In the normalisation channel
 2048 fit the $B_s^0 \rightarrow K^* J/\psi$ background is described using the same PDF used for the
 2049 signal but a different central value, m , which is set at the B_s^0 nominal mass [3].
 2050 Finally, a $\Lambda_b^0 \rightarrow J/\psi pK$ background component is modelled using simulated $\Lambda_b^0 \rightarrow$
 2051 $J/\psi pK$ events to which the full $B^0 \rightarrow K^* J/\psi$ selection is applied. The invariant
 2052 mass distribution of these candidates is a broad shape under the signal peak. The
 2053 simulated distribution is smoothed using a kernel estimation method (using the
 2054 `RooKeysPdf` class of the `RooFit` package [85]).

2055 In summary the free parameters in the simultaneous fit to rare and resonant $\mu^+ \mu^-$
 2056 data samples are: the signal and background yields, the combinatorial background
 2057 slopes, the widths scale c and the the mass shift m' .

2058 Fig. 14.2 reports fits to real data distributions for the rare and resonant $\mu^+ \mu^-$
 2059 channels. Values of fitted parameters are reported on the plots.

2060 14.2 Electron channels

2061 In the electron case the variable we fit is the $m(K\pi e^+ e^-)$ invariant mass coming
 2062 from the kinematic fit where all vertices are required to point to their mother par-

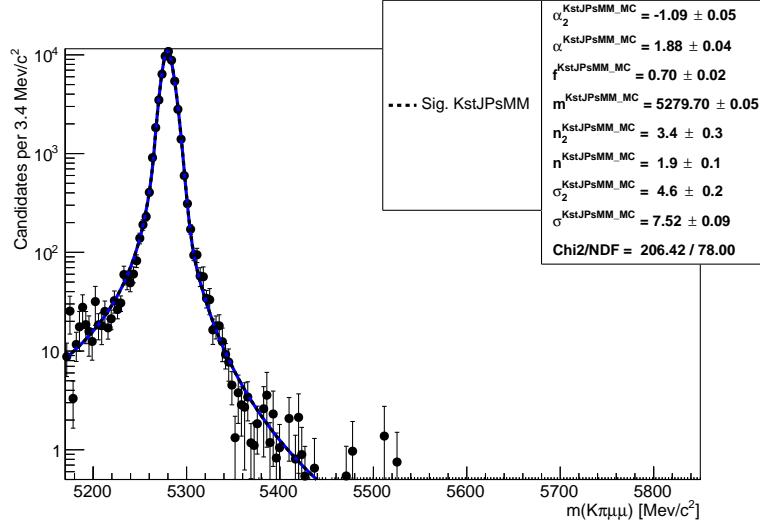


Figure 14.1: Fitted $m(K\pi\mu\mu)$ mass spectrum for $K^{*0}J/\psi$ simulated events.

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 $6 \text{ GeV}^2/c^4$ 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. ??, 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

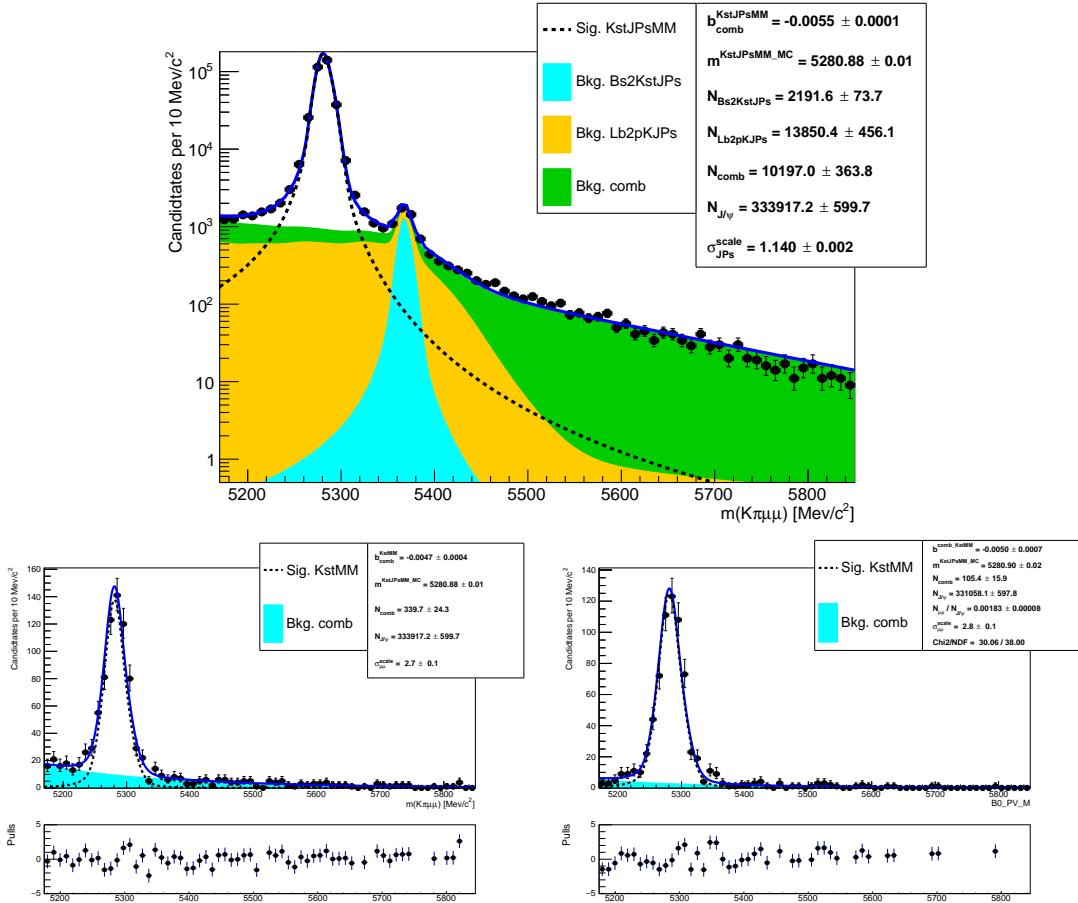


Figure 14.2: Fitted $m(K\pi\mu\mu)$ invariant mass distribution for $K^{*0}J/\psi$ candidates (top) and for rare candidates in the central (bottom left) and high (bottom right) q^2 intervals. Dashed black lines represent the signal PDFs and filled shapes the background components.

- 2080 The three samples, divided by trigger, are fitted simultaneously. This allows a better
 2081 use of statistics as the simultaneous fit gathers information from the three categories
 2082 at the same time and is more stable. Furthermore using this method the results for
 2083 the three categories are naturally combined in a single R_{ee} ratio.
 2084 The PDFs used to fit the invariant mass distributions are described in the next
 2085 section.

2086 14.2.1 Signal PDFs for the electron channels in the central q^2 interval

2087 As for the muonic channel simulated events are fitted at first to constrain the shapes
2088 for the subsequent fit on data. The signal PDFs are built using the following method:

- 2089 • Simulated $B^0 \rightarrow K^{*0} J/\psi(ee)$ and $B^0 \rightarrow K^{*0} ee$ events are divided in each
2090 trigger and bremsstrahlung category and an independent fit is performed to
2091 each sample.
- 2092 • For each trigger category a PDF is built as the sum of the three PDFs for each
2093 bremsstrahlung category.

$$P(m)^{\text{trg}} = f_{0\gamma}^{\text{trg}} P_{0\gamma}^{\text{trg}}(m) + f_{1\gamma}^{\text{trg}} P_{1\gamma}^{\text{trg}}(m) + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) P_{2\gamma}^{\text{trg}}(m). \quad (14.4)$$

2094 where the $P(x)_{n\gamma}^{\text{trg}}$ functions are the chosen PDFs for each trigger and bremsstrahlung
2095 category and the $f_{n\gamma}^{\text{trg}}$ parameters are the relative fractions of events falling in
2096 each category.

- 2097 • Most parameters are fixed (details later) and the combined PDF, $P(m)$, is
2098 used to fit real data divided only in trigger categories.

2099 The 0γ category is characterised by a better resolution and a sharp tail on the right-
2100 hand side and it is fitted with a simple Crystal Ball function (CB). Instead the 1γ
2101 and 2γ samples are modelled using the sum of a Crystal Ball and a Gaussian func-
2102 tions (CBG) with all parameters independent. When the combined PDF, $P(m)^{\text{trg}}$,
2103 is built all parameters are fixed leaving one global mass shift and one scale factor
2104 for the widths free to vary, as done for the muonic samples.

2105 Finally, combining the three bremsstrahlung PDFs one needs to specify in which
2106 fractions they contribute to the total. These fractions have been shown to be in
2107 good agreement between resonant data and Monte Carlo and therefore they are
2108 fixed to the values found on simulation, separately for the normalisation channel
2109 and each q^2 bin. In Tab. 14.1 are reported percentages of events with 0, 1 and 2

Table 14.1: 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 %

2110 recovered photons in the three trigger categories.

2111 In summary the signal PDF for the fit on data is defined as:

$$P(m; c, m')^{\text{trg}} = f_{0\gamma}^{\text{trg}} \text{CB}_{0\gamma}^{\text{trg}}(m; c, m') + f_{1\gamma}^{\text{trg}} \text{CBG}_{1\gamma}^{\text{trg}}(m; c, m') + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) \text{CBG}_{2\gamma}^{\text{trg}}(m; c, m') \quad (14.5)$$

2112 where the free parameters are: c , the scaling factor for the widths, and m' , the mass
2113 shift.

2114 14.2.2 Background PDFs for the electron channels in the central q^2 2115 interval

2116 In the fit to the resonant sample three background components are modelled: com-
2117 binatorial background, and misreconstructed background coming from the hadronic
2118 and the leptonic systems. The combinatorial is described with an exponential func-
2119 tion.

2120 The misreconstructed background is split in two categories, that involving higher
2121 hadronic resonances and that coming from higher $c\bar{c}$ resonances. These backgrounds
2122 are modelled using inclusive $B^0 \rightarrow J/\psi X$ simulated samples to which the full selec-
2123 tion is applied. The distributions for the hadronic (leptonic) background are defined
2124 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. 14.3, 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.

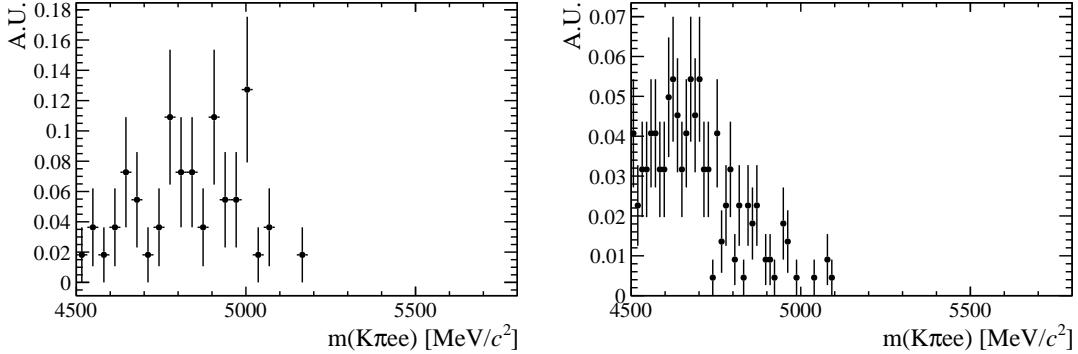


Figure 14.3: Simulated distributions of misreconstructed background events falling into the $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ sample coming from the hadronic (left) and leptonic (right) systems.

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:

$$N_{\ell\ell}^{mis-reco} = N_{ee} \cdot k = N_{ee} \cdot \frac{N_{J/\psi}^{mis-reco}}{N_{J/\psi}}. \quad (14.6)$$

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 $B^0 \rightarrow J/\psi K^*$ 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. 14.4. The yield of this component again is not floating independently but it is linked to the yield found in the resonant

²¹⁴⁴ fit as follows

$$N_{\ell\ell}^{leak} = N_{J/\psi} \cdot k^{MC} = N_{J/\psi} \cdot \frac{N_{leak}^{MC}}{N_{J/\psi}^{MC}} \quad (14.7)$$

²¹⁴⁵ where k is the ratio between $N_{J/\psi}^{MC}$, the number of J/ψ events that fall into the
²¹⁴⁶ J/ψ q^2 window ($6-11$ GeV^2/c^4) in the simulation and N_{leak}^{MC} , the number of J/ψ events
²¹⁴⁷ leaking below 6 GeV^2/c^4 in the simulation. In this case k is previously extracted
²¹⁴⁸ from simulated events and fixed in the fit on data.

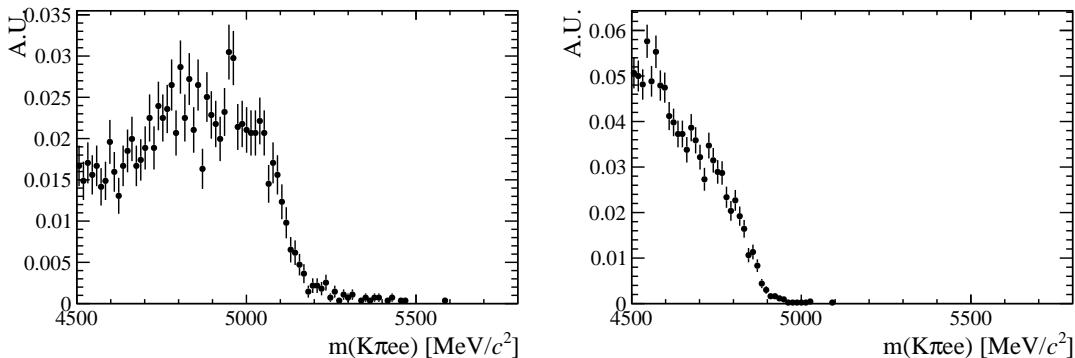


Figure 14.4: (left) Simulated 4-body invariant mass distributions for events involving higher K^* states and passing out full selection. (right) Simulated invariant mass distribution of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ events leaking into the central q^2 interval.

²¹⁴⁹ 14.2.3 Summary of the fit to electron channels in the central q^2 interval

²¹⁵⁰ In summary in the resonant fit on data the floating parameters are the yields of
²¹⁵¹ all the components in the resonant channel, a common R_{ee} ratio, the combinatorial
²¹⁵² background yield in the rare sample, one scale factor c , one mass shift m_0 and the
²¹⁵³ combinatorial background slopes.

²¹⁵⁴ Fits on simulated $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates are shown in Appendix D.
²¹⁵⁵ Figure 14.5 and 14.6 present fits on real $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow$
²¹⁵⁶ $K^{*0}e^+e^-$ candidates (central- q^2 interval) in the three trigger categories. Values of
²¹⁵⁷ fitted parameters are reported on the plots.

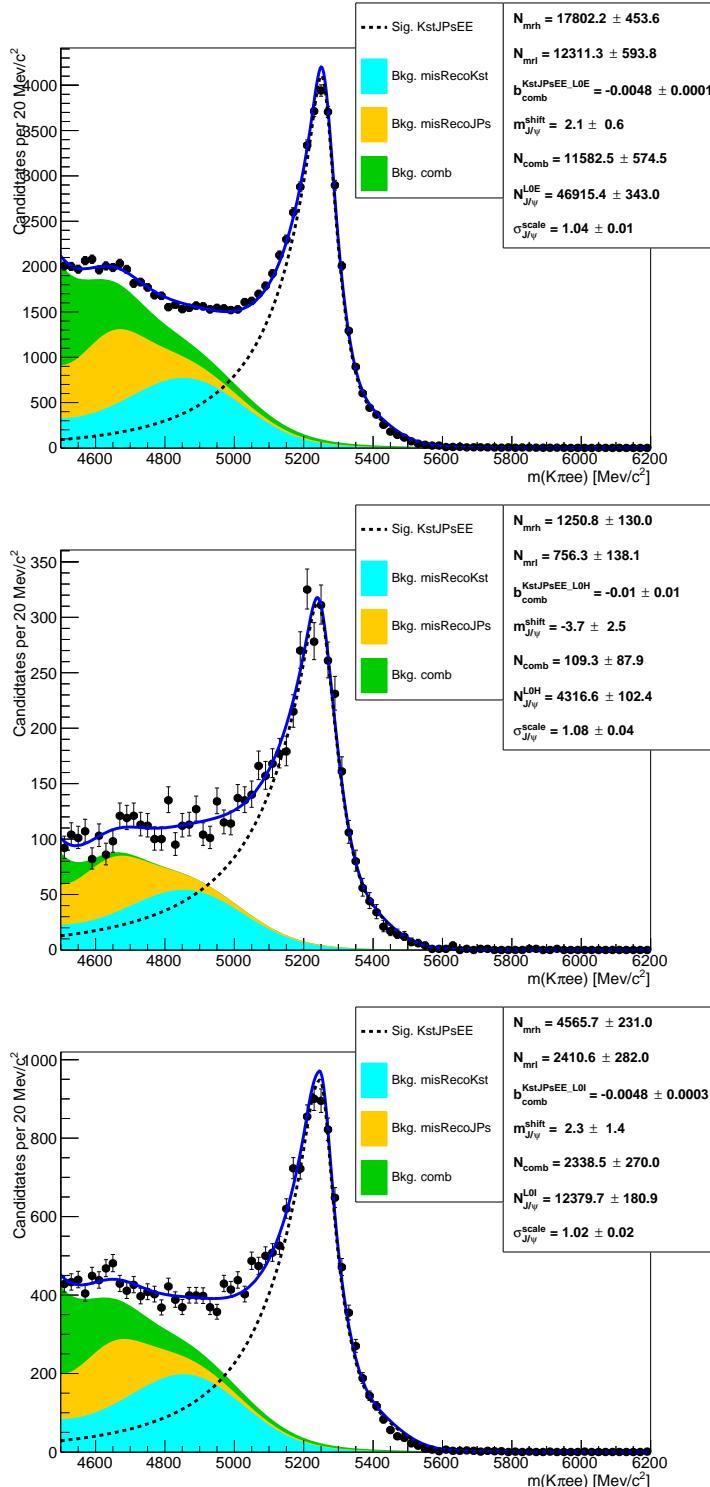


Figure 14.5: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+e^-)$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

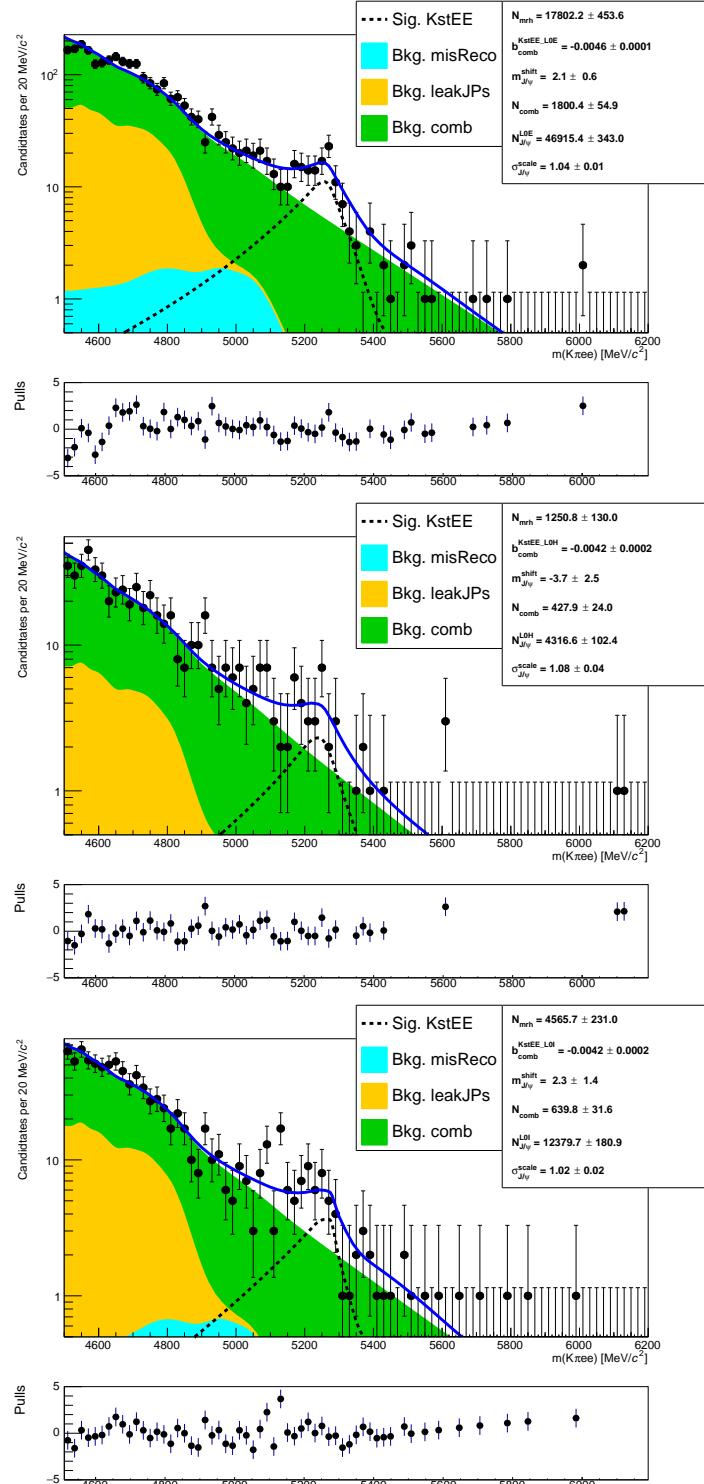


Figure 14.6: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} e^+ e^-$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

2158 14.2.4 Electron channels fits in the high q^2 interval

2159 In the high q^2 interval, above 15 GeV²/c⁴, the efficiency for the L0Hadron trigger
 2160 becomes very low as the K^* has very low momentum. In this region only 9 candidates

Table 14.2: 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 %

2161 are found spread in the interval $4500 < m(K\pi ee) < 6000 \text{ MeV}/c^2$. In the L0TIS
2162 category, even if the yield is bigger a clear peak cannot be seen, therefore only
2163 L0Electron triggered events are used in this region.

2164 The signal PDF is described in the same way as for the central bin. Simulated
2165 events are divided in three bremsstrahlung categories and fitted using the same
2166 PDFs described in Sec. 14.2.1. While the signal tail parameters are similar for the
2167 J/ψ and central q^2 samples in the case of the high q^2 interval it is particularly
2168 important to keep them independent. In fact, as can be seen in Fig. 14.7, the
2169 invariant mass distributions are significantly different for the two intervals. The
2170 fractions of 0, 1 and 2 γ components used to build the total PDF are also in this
2171 case taken from simulated events are reported in Tab. 14.2.

2172 The background components, as for the central q^2 interval, include a combinatorial
2173 background and a misreconstructed background coming from the hadronic system.
2174 Furthermore there is a leakage due to the $\psi(2S)$ resonance, that is wide enough to
2175 contribute in q^2 above 15 GeV^2/c^4 . The combinatorial background is modelled using
2176 the shape obtained by reversing the NN output cut. In Fig. 14.8 are shown invariant
2177 mass distributions for different NN output cuts using the electron and muon high- q^2
2178 samples. The shapes are normalised to the same integral. This plot shows that
2179 the shape is similar for the two samples and as a function of the NN output cut,
2180 which supports the hypothesis that these distributions contain mostly combinatorial
2181 background and that its shape has no strong dependency on the NN output cut.
2182 The shape is therefore obtained from the muon sample, which has higher statistics,
2183 selecting events with NN output < 0.1 . The distribution is then smoothed using the
2184 `RooKeysPdf` method.

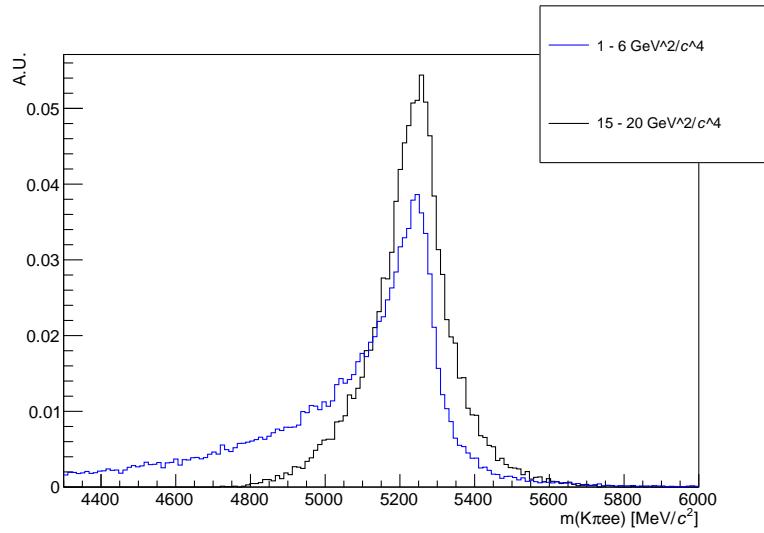


Figure 14.7: Simulated invariant mass of the $K\pi ee$ system in the $1.1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.

2185 The misreconstructed component is modelled in the same way described for the
 2186 central q^2 interval. However, in this case, its yield is not constrained to the resonant
 2187 channel.

2188 The $\psi(2S)$ leakage component is modelled from $B^0 \rightarrow K^*(\psi(2S) \rightarrow e^+e^-)$ simulated
 2189 events with the same method used for the J/ψ leakage in the central q^2 interval. The

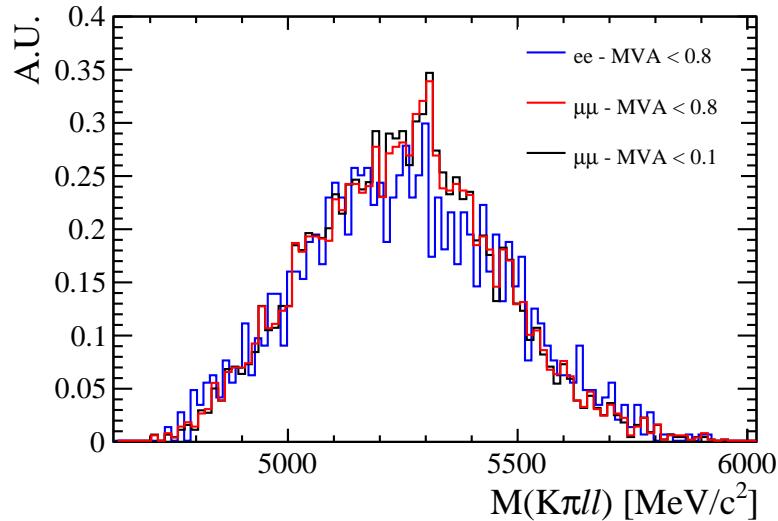


Figure 14.8: Invariant mass distribution of candidates selected by reversed NN output cuts.

Sample	$1\text{--}6 \text{ GeV}^2/c^4$	$15\text{--}20 \text{ GeV}^2/c^4$	J/ψ
$\mu\mu$	625.38 ± 29.60	606.87 ± 27.56	333917.20 ± 599.73
ee L0E	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 14.3: Raw yields of events found fitting invariant mass distributions of the rare and resonant events.

2190 yield of this component is fixed to the yield of $\psi(2S)$ as

$$N_{\ell\ell}^{leak} = N_{\psi(2S)} \cdot k^{MC} = N_{\psi(2S)} \cdot \frac{N_{\psi(2S)}^{MC}}{N_{\psi(2S)}^{MC}}. \quad (14.8)$$

2191 In order to do this the $\psi(2S)$ yield, $N_{\psi(2S)}$, is obtained from a fit to the $\psi(2S)$
2192 invariant mass peak. Since we are only interested in the $\psi(2S)$ yield we fit the
2193 $m(K\pi ee)$ obtained from a kinematic fit where the dimuon mass is constrained to
2194 the known $\psi(2S)$ mass. This allows to eliminate the misreconstructed background
2195 form the fit mass window and use a simple model composed by a signal component
2196 and a combinatorial background component. The signal is described with a Double
2197 Crystal Ball function with opposite tails already described the Λ_b^0 fits (see Sec. 5.1),
2198 and the combinatorial background is described with an exponential. The fit to the
2199 $\psi(2S)$ peak is reported in Fig. 14.9 together with the fit to the $B^0 \rightarrow K^* e^+ e^-$
2200 candidates in the high q^2 interval.

2201 14.3 Fit summary

2202 In Tab. 14.3 are reported raw yields obtained from the fits described in the previous
2203 sections. The values for the rare channels are not directly floating in the fits but as
2204 described in Sec. 14 they are parameterised as a function of the number of resonant
2205 events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare branching
2206 fractions. Measured values of these ratios are reported in Tab. 17.2.

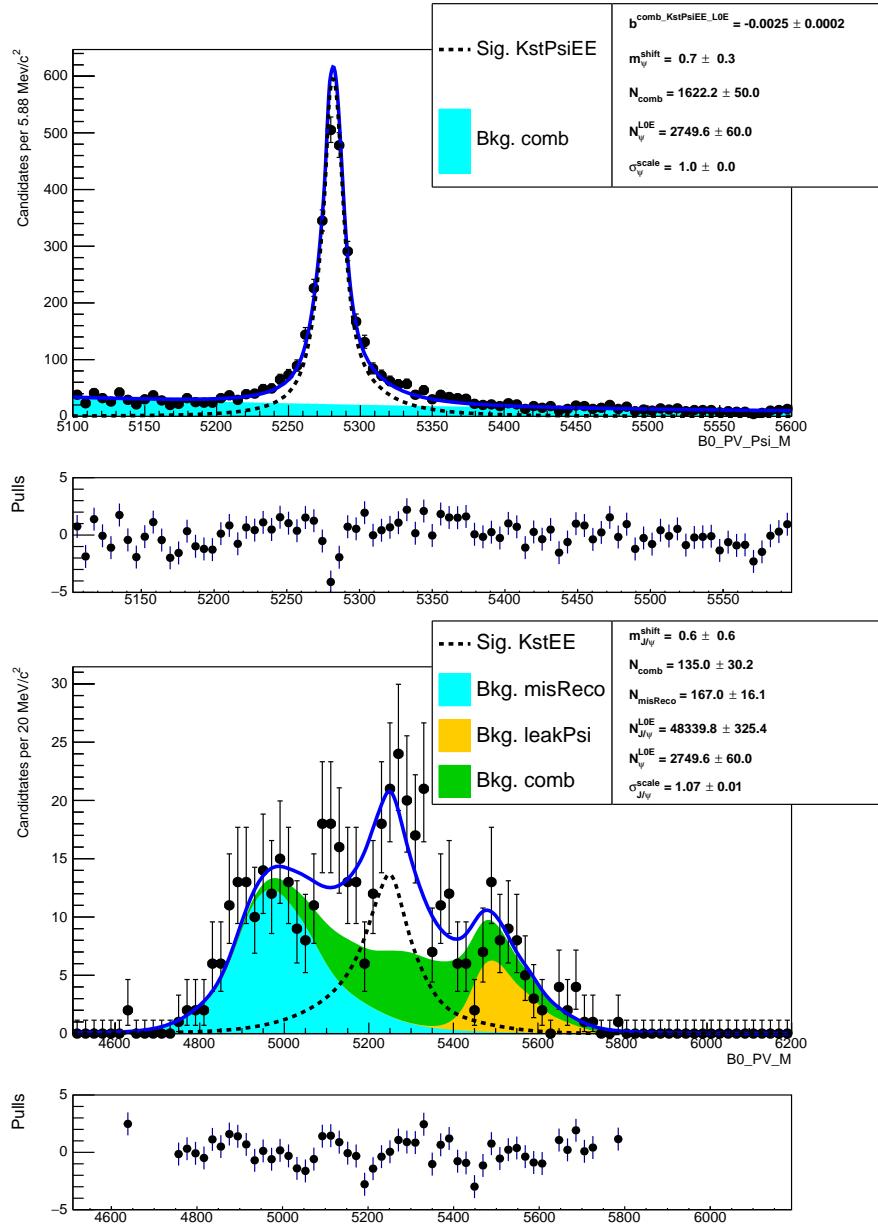


Figure 14.9: Fitted $m(K\pi ee)$ invariant mass distribution in the $\psi(2S)$ interval, $11 < q^2 < 15$ GeV^2/c^4 and in the high- q^2 interval.

CHAPTER 15

2207

2208

2209

Efficiency

2210

2211 The efficiency for each of the decay channels is calculated according to the formula

$$\varepsilon^{tot} = \varepsilon(geom)\varepsilon(reco|geom)\varepsilon(PID|reco)\varepsilon(trig|PID)\varepsilon(MVA|trig). \quad (15.1)$$

2212 In this expression the first term is the efficiency to have final state particles in the
2213 LHCb detector acceptance. The second term carries information on reconstruction
2214 and stripping efficiency (we keep these together given that boundaries between them
2215 are completely artificial). The third part corresponds to the efficiency of the PID
2216 requirements. The fourth term handles the trigger efficiency for those events which
2217 are selected by the preselection process. Finally, the latter term deals with the
2218 efficiency of the NN classifier. Reconstruction, trigger and MVA efficiencies are
2219 evaluated on simulated data with the trigger efficiency for $B^0 \rightarrow K^* J/\psi$ being
2220 cross-checked using the data-driven TISTOS method as described in Sec. ???. The
2221 PID efficiency is calculated with a data-driven method as described in Sec. 15.3.

2222 All absolute efficiencies for the muon and electron rare channels are separately listed
2223 in Tab. 15.3 for the central and high q^2 intervals and in Tab. 15.1 for the resonant
2224 channels. However for the analysis itself only efficiencies relative to the resonant
2225 channels are used in order to limit systematic uncertainties.

2226 Tab. 15.3 reports relative efficiencies between the rare and resonant channels, $\varepsilon(B^0 \rightarrow$
2227 $K^*\ell^+\ell^-)/\varepsilon(B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-))$. Finally, in Tab. 15.2 are listed ratios of relative
2228 efficiencies for the ee and $\mu\mu$ channels, $(ee/(J/\psi \rightarrow ee))/(\mu^+\mu^-/(J/\psi \rightarrow \mu^+\mu^-))$.

Table 15.1: Absolute efficiencies for the resonant for ee and $\mu\mu$ channels.

ε	$\mu\mu$	ee		
		L0E	L0H	L0I
Geom	0.1598 ± 0.0005		0.1589 ± 0.0005	
Reco	0.0896 ± 0.0001		0.0583 ± 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	0.0080 ± 0.0000	0.0013 ± 0.0000	0.0001 ± 0.0000	0.0005 ± 0.0000

Table 15.2: Double ratios of efficiencies $(\varepsilon^{ee}/\varepsilon^{J/\psi \rightarrow ee})/(\varepsilon^{\mu^+\mu^-}/\varepsilon^{J/\psi \rightarrow \mu^+\mu^-})$ in the $1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.

ε	1–6 GeV^2/c^4			15–20 GeV^2/c^4
	L0E	L0H	L0I	L0E
q^2		0.697 ± 0.007		0.770 ± 0.010
Geom		1.022 ± 0.012		1.022 ± 0.012
Reco		0.905 ± 0.006		0.451 ± 0.451
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	1.123 ± 0.023	3.001 ± 0.081	1.569 ± 0.036	0.871 ± 0.020

Table 15.3: Absolute efficiencies for the rare ee and $\mu\mu$ channels in the $1.1 < q^2 < 6$ and $q^2 > 15 \text{ GeV}^2/c^4$ intervals.

ε	$\mu\mu$		ee			$15-20 \text{ GeV}^2/c^4$
	$1-6 \text{ GeV}^2/c^4$	$15-20 \text{ GeV}^2/c^4$	L0E	L0H	L0I	
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	0.0170 ± 0.0001	0.0108 ± 0.0001		0.0100 ± 0.0000		0.0032 ± 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	0.3105 ± 0.0028
MVA	0.9097 ± 0.0022	0.8298 ± 0.0032	0.8447 ± 0.0021	0.8571 ± 0.0020	0.8156 ± 0.0046	0.8436 ± 0.0033
Tot	0.0063 ± 0.0001	0.0067 ± 0.0001	0.0012 ± 0.0000	0.0003 ± 0.0000	0.0006 ± 0.0000	0.0009 ± 0.0000

Table 15.4: Relative efficiencies rare over resonant ($\varepsilon^{rel} = \varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$) for ee and $\mu\mu$ channels in the $1.q < q^2 < 6$ and $q^2 > 15 \text{ GeV}^2/c^4$ intervals.

ε	$1-6 \text{ GeV}^2/c^4$			$15-20 \text{ GeV}^2/c^4$		
	$\mu\mu$	ee		$\mu\mu$	ee	
	L0E	L0H	L0I		L0E	
Geom	1.0200 ± 0.0091		1.0429 ± 0.0084		1.0200 ± 0.0091	1.0429 ± 0.0084
Reco	0.1896 ± 0.0012		0.1716 ± 0.0006		0.1201 ± 0.0009	0.0541 ± 0.0003
PID	0.9602 ± 0.0002		0.9425 ± 0.0001		1.0334 ± 0.0001	1.0021 ± 0.0001
Trig	0.8993 ± 0.0038	0.8002 ± 0.0065	2.2025 ± 0.0434	1.1138 ± 0.0136	1.1098 ± 0.0037	1.5715 ± 0.0145
MVA	1.0154 ± 0.0025	0.9838 ± 0.0025	0.9551 ± 0.0023	0.9874 ± 0.0060	0.9262 ± 0.0036	0.9825 ± 0.0039

2229 15.1 Geometric efficiency

2230 The simulated samples used contain the requirement that daughters are in the LHCb
 2231 detector acceptance. This corresponds to the requirement for each of the final par-
 2232 ticles to have polar angle θ between 10 and 400 mrad. The efficiency of this cuts is
 2233 obtained using a generator level Monte Carlo sample.

2234 15.2 Reconstruction efficiency and bin migration

2235 The reconstruction efficiency is here defined as the efficiency to reconstruct each
 2236 decay channel given that its daughters are into the geometrical acceptance of the
 2237 detector. This includes both the probability that a particle generates observable sig-
 2238 natures and the efficiency of all the preselection cuts described in Sec. 13, including
 2239 those done to remove peaking backgrounds. The efficiency of the PID cuts is kept
 2240 separate as it is known to be not well simulated and there are reliable data-driven
 2241 methods which can be used to extract it (see Sec. 15.3).

2242 It can happen that events generated in a q^2 interval are reconstructed in a different
 2243 one, this is referred to as “bin migration”. Two different effects can cause bin
 2244 migration. First of all, as the resolution of real detectors is not perfect, events
 2245 close to the edges of the considered intervals can fall on the wrong side of the
 2246 edge. This effect is only important in case of non-flat true distributions, as the
 2247 amount of bin migration in the two directions is different. The second possible
 2248 source of bin migration are systematic effects due, for example, to the presence
 2249 of bremsstrahlung photons that cannot be recovered. It is particularly important
 2250 to take into account the bin migration in the electron channels case because more
 2251 photons are radiated from the final state. Figure 15.1 reports the response matrix for
 2252 simulated $B^0 \rightarrow K^{*0} ee$ events, which shows the correlation between reconstructed
 2253 and generated q^2 . In the ideal case of perfect resolution and no bias this plot would
 2254 look like a diagonal line and in the case of no bias its slope would be 1. Table ??

2255 reports net bin migration amounts in the considered q^2 intervals. Table ?? reports
 2256 net amounts of bin migration, M_{net} , in the considered q^2 intervals defined as:

$$M_{net} = N(\text{in} \rightarrow \text{in}) + N(\text{out} \rightarrow \text{in}) - N(\text{in} \rightarrow \text{out}) \quad (15.2)$$

2257 where $N(\text{in} \rightarrow \text{in})$ is the number of candidates that are generated and reconstructed
 2258 inside the considered interval, $N(\text{out} \rightarrow \text{in})$ the number of candidates that are
 2259 generated outside the interval but reconstructed inside and $N(\text{in} \rightarrow \text{out})$ the number
 2260 of candidates generated inside that fall outside.

2261 The reconstruction efficiency is calculated comparing generated to reconstructed
 2262 samples and therefore already includes bin migration effects. Nevertheless, it is
 2263 useful to single out this component to better asses the corresponding systematic
 2264 uncertainties.

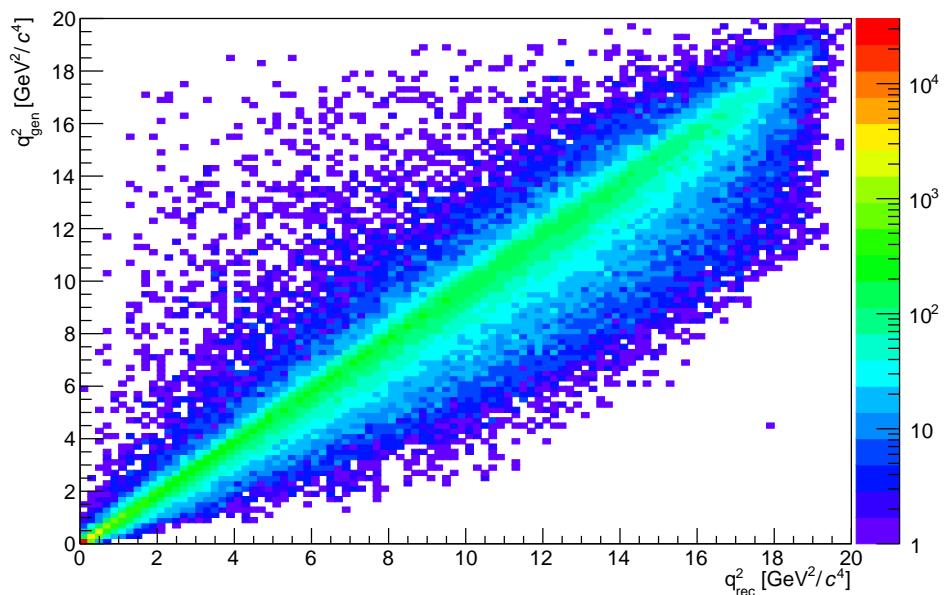


Figure 15.1: Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^*ee$ events.

2265 15.3 PID efficiency

2266 The Monte Carlo is known not to reliably describe particle ID variables and therefore
 2267 a data-driven method is used to obtain this efficiency component. Furthermore
 2268 the same method is used to weight the MC in order to extract MVA and trigger
 2269 efficiencies.

2270 In order to do this the `PIDCalib` package [86] was used. This tool uses decays where
 2271 final particles can be identified thanks to their kinematic properties. For example
 2272 the $K_s^0 \rightarrow \pi^+\pi^-$ decay has a clear signature with a displaced vertex and can be easily
 2273 singled out from other decays and used to test pion PID efficiency. The narrow peaks
 2274 of the $J/\psi \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow e^+e^-$ decays allow to calibrate muon and electron
 2275 efficiencies. A “tag-and-probe” method is used where only one of the two leptonic
 2276 tracks is reconstructed requiring the correct identity and the other one is used to
 2277 probe the PID efficiency. Finally, $\psi \rightarrow KK$ samples and $D^{*+} \rightarrow D(\rightarrow K^-\pi^+)\pi^+$
 2278 decays, where the D^* is used to tag the decay, are used to test the kaon efficiency.
 2279 Residual background in this decays is subtracted using the `sPlot` technique [82].

2280 The package allows to divide the phase-space in bins and obtain a data-driven effi-
 2281 ciency for each bin. For this analysis the phase-space was divided in equi-populated
 2282 bins of momentum and pseudorapidity of the particle under study. Figure 15.2
 2283 shows performance tables for pions, kaons, muons and electrons.

2284 The decay channel under study generally has different kinematical distributions than
 2285 the calibration sample. Therefore, once the efficiency table is obtained for each
 2286 particle, the total efficiency for each candidate is calculated as the product of the
 2287 four final particles efficiencies. $\varepsilon^{ev} = \varepsilon_K \cdot \varepsilon_\pi \cdot \varepsilon_{\ell_1} \cdot \varepsilon_{\ell_2}$. Finally, the total efficiency is

Table 15.5: Net bin migration amounts 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

2288 found by averaging over all simulated events.

$$\varepsilon_{PID} = \frac{1}{N} \sum_i^N \varepsilon_K(p_K^i, \eta_K^i) \cdot \varepsilon_\pi(p_\pi^i, \eta_\pi^i) \cdot \varepsilon_\ell(p_{\ell_1}^i, \eta_{\ell_1}^i) \cdot \varepsilon_K(p_{\ell_2}^i, \eta_{\ell_2}^i) \quad (15.3)$$

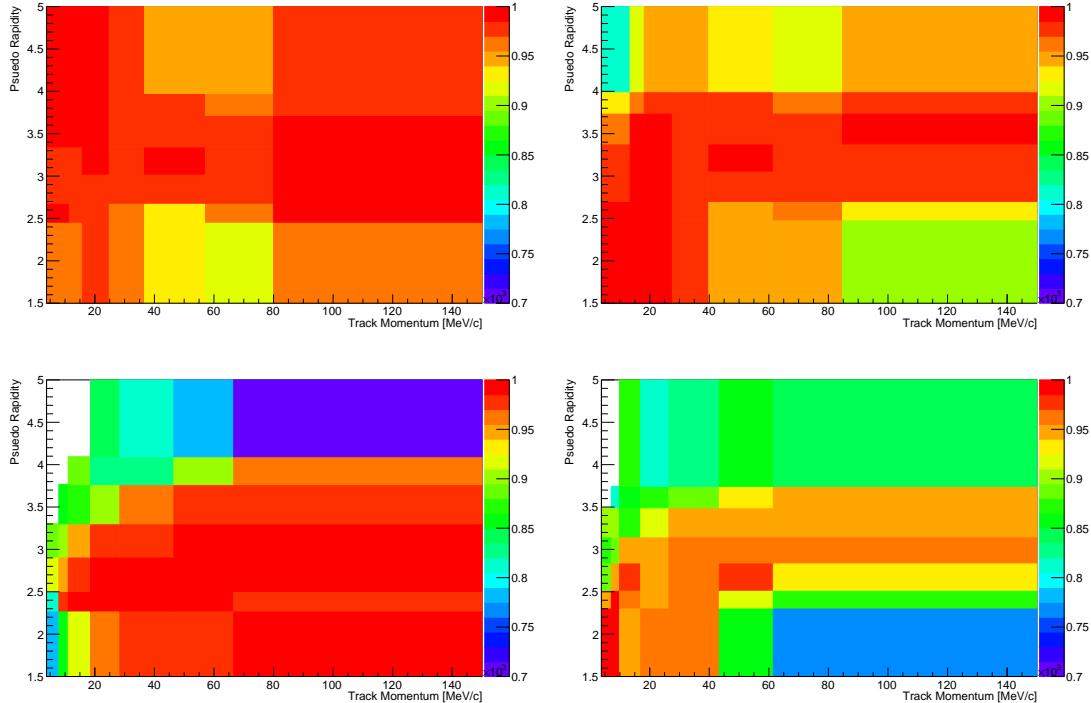


Figure 15.2: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

2289 15.4 Trigger efficiency

2290 15.4.1 muon channels

2291 For the muon channels the trigger efficiency is calculated using simulated events.

2292 Using the resonant channel the efficiency obtained using the simulation was cross-
2293 checked with the data driven TISTOS method as already described in Sec. 6.3.

2294 Results are shown in Tab. 15.6, where the efficiency obtained using the TISTOS

2295 method is compared between sWeighted data and simulation. As data also contains

2296 non negligible amounts of background the **DecayTreeFitter** tool is used to constrain
2297 the J/ψ mass to its known value and a narrow interval around the peak, dominated
2298 by the signal, is selected. Efficiencies obtained on simulation and data are in good
2299 agreement.

Table 15.6: 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
$J/\psi(\mu\mu)$	0.797 ± 0.002	0.803 ± 0.004
$J/\psi(ee)$ L0E	0.268 ± 0.002	0.255 ± 0.004
$J/\psi(ee)$ L0H	0.028 ± 0.001	0.026 ± 0.002
$J/\psi(ee)$ L0I	0.017 ± 0.001	0.011 ± 0.001

2300 15.4.2 Electron channels

2301 For the electron channels data is fitted separately in three trigger categories: L0Electron,
2302 L0Hadon and L0TIS. Therefore we need to extract the efficiency separately for each
2303 category.

2304 While the Hlt (1 and 2) efficiency is still computed using simulated events, the
2305 L0Electron and L0Hadron efficiencies cannot be modelled with the Monte Carlo.
2306 The discrepancy between data and simulation is mainly due to the ageing of the
2307 calorimeters, on which the decision of these triggers relies. The ageing is not simu-
2308 lated in the Monte Carlo and affects the L0 trigger efficiency which, therefore, must
2309 be calibrated using data driven-methods. Tables of efficiencies are obtained applying
2310 the TIS-TOS method to a calibration sample.

2311 For each trigger category these tables contain efficiency as a function of p_T of the
2312 considered particle and are given for different calorimeter regions as these have
2313 different properties (e.g. cell size) due to the different position with respect to the
2314 beam line. Regions considered are inner and outer HCAL, and inner, middle and
2315 outer ECAL. Figure 15.3 shows data-driven efficiencies for the L0Electron trigger in
2316 the three ECAL regions.

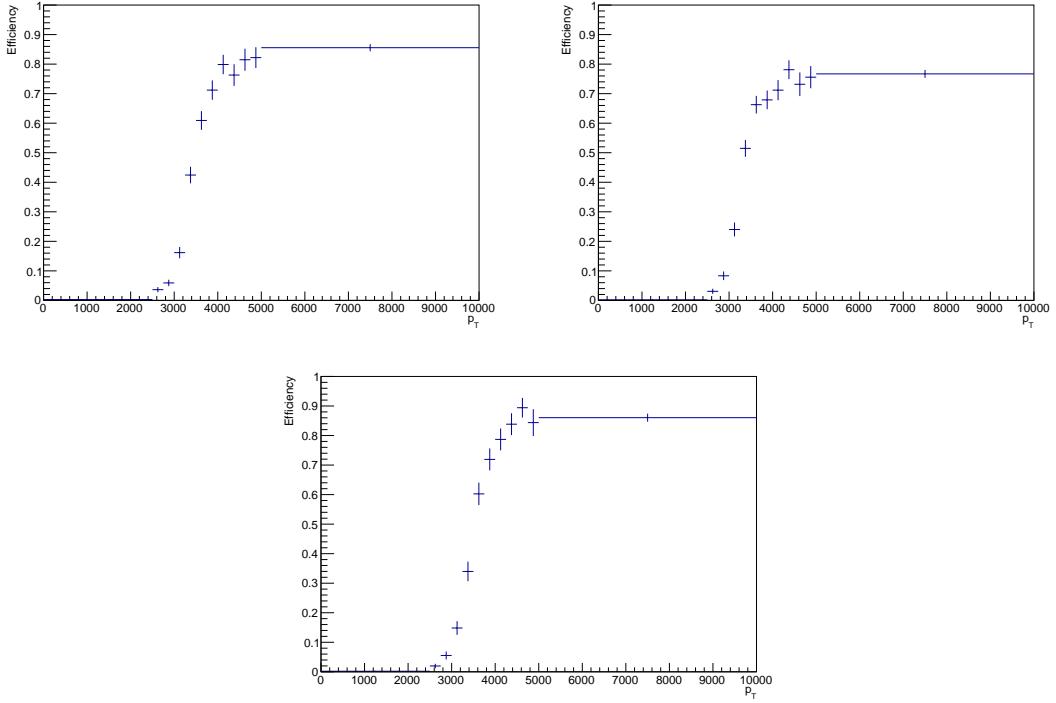


Figure 15.3: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

2317 The probability of L0Electron trigger is calculated for each event as $P_{L0E} = \varepsilon(e^+) +$
 2318 $\varepsilon(e^-) - \varepsilon(e^+)\varepsilon(e^-)$. Similarly, the L0Hadron trigger probability is $P_{L0H} = \varepsilon(\pi) +$
 2319 $\varepsilon(K) - \varepsilon(\pi)\varepsilon(K)$. For L0TIS a probability $P_{L0I} = (1 - P_{L0E}) \cdot (1 - P_{L0H})$ is calculated.
 2320 Note that the probability of TIS trigger is defined to be independent of the signal
 2321 and therefore must be the same in the rare and resonant channels and cancel in
 2322 their ratio.

2323 Then event by event efficiencies for the three trigger categories are defined to be
 2324 exclusive in the following way:

- 2325 • L0E: $\varepsilon^{L0E} = P_{L0E}$;
- 2326 • L0H: $\varepsilon^{L0H} = P_{L0H} \cdot (1 - P_{L0E})$, namely the probability that at least one hadron
 2327 triggered but none of the electrons;
- 2328 • L0I: $\varepsilon^{L0I} = (1 - P_{L0H}) \cdot (1 - P_{L0E})$, namely the probability that neither the
 2329 hadrons or the electrons in the event triggered.

2330 As in the PID case, the total efficiency is found averaging over all events of a
2331 simulated sample:

$$\varepsilon^{trg} = \frac{1}{N} \sum_i^N \varepsilon^{trg}(p_T^i) \quad (15.4)$$

2332 where trg is a label indicating the trigger category under consideration.

2333 15.5 Neural Networks efficiency

2334 The NN efficiency is again evaluated from fully weighted Monte Carlo samples. For
2335 the electron channels it is obtained separately for each trigger category, because the
2336 yield is extracted independently for each of the three trigger categories and therefore
2337 these have to be independently corrected.

2338 In order to cross check that this efficiency component is extracted correctly one can
2339 compare the efficiency obtained using $B^0 \rightarrow J/\psi K^*$ events and rare $B^0 \rightarrow K^* \ell^+ \ell^-$
2340 events in the same q^2 region selected for the resonant case. The ratio between the
2341 two should be close to 1 with small deviations due the fact that the bin width is finite
2342 and the events are distributed differently inside the bin. These ratios are reported
2343 in Tab. 15.7; values are not exactly one for the electron channels due to the very
2344 large q^2 interval used to select the resonant channel ($[6.0, 11.0] \text{ GeV}^2/c^4$).

Table 15.7: Ratio $\varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$ where the efficiency for the rare channel its calculated in the same q^2 rage used to select the resonant channel.

Comp	$\mu\mu$	ee		
		L0Electron	L0Hadron	L0TIS
mva	0.9969 ± 0.0039	0.9771 ± 0.0023	0.9794 ± 0.0019	0.9856 ± 0.0057

2345

CHAPTER 16

2346

2347

Systematic uncertainties

2348

CHAPTER 17

2349

2350

2351

Result extraction

2352

2353 In the following sections are presented the final results of this analysis together with
2354 the description of sanity checks performed to verify the stability of the methods
2355 used.

2356 17.1 $R_{J/\psi}$ sanity check

2357 In order to cross-check the analysis procedure, the ratio between the measured
2358 branching ratio of the electron and muonic resonant channels is calculated:

$$R_{J/\psi} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-))} = \frac{\varepsilon_{J/\psi(\mu\mu)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)}}{\varepsilon_{J/\psi(ee)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-)}}. \quad (17.1)$$

2359 One could also look at the single absolute branching fractions but this ratio is a
2360 better defined sanity check for our analysis. In fact the absolute branching fractions

2361 can be calculated from the raw yields as

$$\mathcal{B} = \mathcal{L} \cdot \sigma_{b\bar{b}} \cdot f_d \cdot \varepsilon \cdot N_{raw} \quad (17.2)$$

2362 where \mathcal{L} is the luminosity, $\sigma_{b\bar{b}}$ is the cross section for $b\bar{b}$ production and f_d is the
2363 fragmentation fraction, the probability for a b quark to produce a B^0 meson. All
2364 this quantities come with large uncertainties but they cancel in the $R_{J/\psi}$ ratio.

2365 Measured values of the $R_{J/\psi}$ ratio are reported in Tab. 17.1, where the error shown is
2366 statistical only. Notice that on these values systematic uncertainties, which cancel
2367 doing the ratio between the rare and resonant channels with same leptonic final
2368 state, do not cancel.

Table 17.1: 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.

$R_{J/\psi}$ (LOE)	1.01618 ± 0.00896
$R_{J/\psi}$ (LOH)	0.94276 ± 0.02315

2369 17.2 R_{K^*} result summary

2370 The ratio R_{K^*} is extracted by dividing the R_{ee} and $R_{\mu\mu}$ parameters described in
2371 Sec. 14. These ratios are floating in the fit but they can also be built from the yields
2372 in Tab. 14.3 and the efficiencies in Tab. 15.2. In summary the definition of the R_{K^*}
2373 ratio is the following :

$$R_{K^*} = \frac{R_{ee}}{R_{\mu\mu}} = \frac{N_{ee}}{N_{J/\psi(ee)}} \cdot \frac{N_{J/\psi(\mu\mu)}}{N_{\mu\mu}} \cdot \frac{\varepsilon_{J/\psi(ee)}}{\varepsilon_{ee}} \cdot \frac{\varepsilon_{\mu\mu}}{\varepsilon_{J/\psi(\mu\mu)}}. \quad (17.3)$$

2374 As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three
2375 electron categories its value is already a combination of the three samples. Results
2376 are shown in Tab. 17.2.

Table 17.2: 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

2377 17.3 Branching ratios and expectations

2378 Multiplying the ratios R_{ee} and $R_{\mu\mu}$ by the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ [3]
 2379 branching ratios one can obtain absolute branching ratios for the rare channels:

$$\begin{aligned} \mathcal{B}(B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)) &= \mathcal{B}(B^0 \rightarrow K^*J/\psi) \times \mathcal{B}(B^0 \rightarrow \ell\ell) \\ &= (1.32 \pm 0.06)10^{-3} \times (5.96 \pm 0.03)10^{-2} = (7.87 \pm 0.36) \times 10^{-5} \end{aligned} \quad (17.4)$$

2380 Table 17.3 reports absolute branching ratio values for the rare channels in the con-
 sidered q^2 intervals, where the errors are statistical only.

Table 17.3: 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	$(2.37 \pm 0.34) \times 10^{-7}$	$(3.09 \pm 0.67) \times 10^{-7}$
$\mu\mu$	$(1.47 \pm 0.10) \times 10^{-7}$	$(1.42 \pm 0.09) \times 10^{-7}$

2381

2382 The results for the central q^2 bin can be compared also with SM predictions obtained
 2383 from Ref. [87]. This paper reports branching predicted ratios in the $1 < q^2 < 6$
 2384 GeV^2/c^4 interval for the rare electronic and muonic channels. These are rescaled
 2385 to the range $1.1 < q^2 < 6 \text{ GeV}^2/c^4$ using the simulation. Finally, the measured
 2386 value of the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ decay is used. The predicted ratio
 2387 is found to be 0.75 ± 0.14 , which is in agreement with our measurement within one
 2388 standard deviation. In Tab. 17.4 are also reported observed and expected ratios of
 2389 rare over resonant raw numbers of candidates ($N_{\ell\ell}/N_{J/\psi}$). In this table the observed
 2390 ratios are simply obtained dividing the rare and resonant yields in Tab. 14.3 and
 2391 the expected ones are obtained using the predicted rare channel branching ratios
 2392 and the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ branching ratio, rescaled by the relative

²³⁹³ efficiencies in Tab. 15.4.

Table 17.4: 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

REFERENCES

- [1] R. Aaij *et al.*, “Differential branching fraction and angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays,” 2015.
- [2] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl.Phys.*, vol. 22, pp. 579–588, 1961.
- [3] K. Olive *et al.*, “Review of Particle Physics,” *Chin.Phys.*, vol. C38, p. 090001, 2014.
- [4] F. Strocchi, “Spontaneous Symmetry Breaking in Local Gauge Quantum Field Theory: The Higgs Mechanism,” *Commun.Math.Phys.*, vol. 56, p. 57, 1977.
- [5] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, “Experimental Test of Parity Conservation in Beta Decay,” *Phys.Rev.*, vol. 105, pp. 1413–1414, 1957.
- [6] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, *et al.*, “Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics,” *Phys.Rev.*, vol. D91, no. 7, p. 073007, 2015.
- [7] F. Zwicky, “Spectral displacement of extra galactic nebulae,” *Helv.Phys.Acta*, vol. 6, pp. 110–127, 1933.
- [8] M. Gavela and Hernandez, “Standard model CP violation and baryon asymmetry,” *Mod.Phys.Lett.*, vol. A9, pp. 795–810, 1994.
- [9] M. Maltoni, “Status of three-neutrino oscillations,” *PoS*, vol. EPS-HEP2011, p. 090, 2011.
- [10] J. L. Feng, “Naturalness and the Status of Supersymmetry,” *Ann.Rev.Nucl.Part.Sci.*, vol. 63, pp. 351–382, 2013.

- ²⁴¹⁹ [11] P. Fayet and S. Ferrara, “Supersymmetry,” *Phys.Rept.*, vol. 32, pp. 249–334, 1977.
- ²⁴²¹ [12] L. Randall and R. Sundrum, “A Large mass hierarchy from a small extra dimension,” *Phys.Rev.Lett.*, vol. 83, pp. 3370–3373, 1999.
- ²⁴²³ [13] G. Isidori and D. M. Straub, “Minimal Flavour Violation and Beyond,” *Eur.Phys.J.*, vol. C72, p. 2103, 2012.
- ²⁴²⁵ [14] T. Blake, T. Gershon, and G. Hiller, “Rare b hadron decays at the LHC,” 2015.
- ²⁴²⁶ [15] A. J. Buras, D. Buttazzo, J. Girkbach-Noe, and R. Knegjens, “Can we reach the Zeptouniverse with rare K and $B_{s,d}$ decays?,” *JHEP*, vol. 1411, p. 121, 2014.
- ²⁴²⁸ [16] G. Hiller and M. Schmaltz, “ R_K and future $b \rightarrow s\ell\ell$ physics beyond the standard model opportunities,” *Phys.Rev.*, vol. D90, p. 054014, 2014.
- ²⁴³⁰ [17] K. G. Chetyrkin, M. Misiak, and M. Munz, “Weak radiative B meson decay beyond leading logarithms,” *Phys.Lett.*, vol. B400, pp. 206–219, 1997.
- ²⁴³² [18] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, “Weak decays beyond leading logarithms,” *Rev.Mod.Phys.*, vol. 68, pp. 1125–1144, 1996.
- ²⁴³⁴ [19] A. J. Buras, “Weak Hamiltonian, CP violation and rare decays,” pp. 281–539, 1998.
- ²⁴³⁶ [20] R. Aaij *et al.*, “Observation of a resonance in $B^+ \rightarrow K^+\mu^+\mu^-$ decays at low recoil,” *Phys. Rev. Lett.*, vol. 111, p. 112003, 2013.
- ²⁴³⁸ [21] 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.*, vol. 112, p. 101801, 2014.
- ²⁴⁴¹ [22] 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*, 2015.
- ²⁴⁴³ [23] R. Aaij *et al.*, “Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$,” *JHEP*, vol. 07, p. 084, 2013.
- ²⁴⁴⁵ [24] R. Aaij *et al.*, “Differential branching fraction and angular analysis of the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$,” *JHEP*, vol. 08, p. 131, 2013.
- ²⁴⁴⁷ [25] R. Aaij *et al.*, “Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)}\mu^+\mu^-$ decays,” *JHEP*, vol. 1406, p. 133, 2014.
- ²⁴⁴⁹ [26] R. Aaij *et al.*, “Measurement of form-factor-independent observables in the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$,” *Phys. Rev. Lett.*, vol. 111, p. 191801, 2013.
- ²⁴⁵¹ [27] S. Descotes-Genon, J. Matias, and J. Virto, “Understanding the $B \rightarrow K^*\mu^+\mu^-$ Anomaly,” *Phys.Rev.*, vol. D88, no. 7, p. 074002, 2013.

- [2453] [28] R. Aaij *et al.*, “Angular analysis of charged and neutral $B \rightarrow K\mu^+\mu^-$ decays,” *JHEP*, vol. 05, p. 082, 2014.
- [2455] [29] R. Aaij *et al.*, “Differential branching fractions and isospin asymmetry of $B \rightarrow K^{(*)}\mu^+\mu^+$ decays,” *JHEP*, vol. 06, p. 133, 2014.
- [2457] [30] R. Aaij *et al.*, “Measurement of A_{CP} and ΔA_{FB} in $B^0 \rightarrow K^*\mu^+\mu^-$ and $B^+ \rightarrow K^+\mu^+\mu^-$ decays,” 2014. in preparation.
- [2459] [31] M. Ahmed *et al.*, “Search for the lepton family number nonconserving decay mu+ —> e+ gamma,” *Phys.Rev.*, vol. D65, p. 112002, 2002.
- [2461] [32] U. Bellgardt *et al.*, “Search for the Decay mu+ —> e+ e+ e-,” *Nucl.Phys.*, vol. B299, p. 1, 1988.
- [2463] [33] 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.*, vol. 111, p. 141801, 2013.
- [2465] [34] R. Aaij *et al.*, “Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb,” *Phys. Lett.*, vol. B724, p. 36, 2013.
- [2467] [35] W. J. Marciano, T. Mori, and J. M. Roney, “Charged Lepton Flavor Violation Experiments,” *Ann.Rev.Nucl.Part.Sci.*, vol. 58, pp. 315–341, 2008.
- [2469] [36] L. Evans, “The LHC machine,” *PoS*, vol. EPS-HEP2009, p. 004, 2009.
- [2470] [37] A. A. Alves Jr. *et al.*, “The LHCb detector at the LHC,” *JINST*, vol. 3, p. S08005, 2008.
- [2472] [38] M. Adinolfi *et al.*, “Performance of the LHCb RICH detector at the LHC,” *Eur. Phys. J.*, vol. C73, p. 2431, 2013.
- [2474] [39] A. A. Alves Jr. *et al.*, “Performance of the LHCb muon system,” *JINST*, vol. 8, p. P02022, 2013.
- [2476] [40] “LHCb technical design report: Reoptimized detector design and performance,” 2003.
- [2478] [41] R. Aaij *et al.*, “The LHCb trigger and its performance in 2011,” *JINST*, vol. 8, p. P04022, 2013.
- [2480] [42] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual,” *JHEP*, vol. 05, p. 026, 2006.
- [2482] [43] I. Belyaev *et al.*, “Handling of the generation of primary events in GAUSS, the LHCb simulation framework,” *Nuclear Science Symposium Conference Record (NSS/MIC)*, vol. IEEE, p. 1155, 2010.
- [2485] [44] D. J. Lange, “The EvtGen particle decay simulation package,” *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.

- [45] P. Golonka and Z. Was, “PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays,” *Eur.Phys.J.*, vol. C45, pp. 97–107, 2006.
- [46] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.*, “Geant4 developments and applications,” *IEEE Trans.Nucl.Sci.*, vol. 53, p. 270, 2006.
- [47] M. Clemencic *et al.*, “The LHCb simulation application, GAUSS: design, evolution and experience,” *J. Phys. Conf. Ser.*, vol. 331, p. 032023, 2011.
- [48] R. Brun, F. Rademakers, and S. Panacek, “ROOT, an object oriented data analysis framework,” *Conf.Proc.*, vol. C000917, pp. 11–42, 2000.
- [49] M. Feindt and U. Kerzel, “The NeuroBayes neural network package,” *Nucl.Instrum.Meth.*, vol. A559, pp. 190–194, 2006.
- [50] M. Feindt, “A Neural Bayesian Estimator for Conditional Probability Densities,” 2004.
- [51] 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.*, vol. B649, pp. 152–158, 2007.
- [52] T. Mannel and S. Recksiegel, “Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \rightarrow \Lambda \gamma$,” *J.Phys.*, vol. G24, pp. 979–990, 1998.
- [53] 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.*, vol. D78, p. 114032, 2008.
- [54] 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.*, vol. C59, pp. 861–882, 2009.
- [55] 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.*, vol. D59, p. 114022, 1999.
- [56] C.-H. Chen and C. Geng, “Rare $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decays with polarized lambda,” *Phys.Rev.*, vol. D63, p. 114024, 2001.
- [57] C.-H. Chen and C. Geng, “Baryonic rare decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$,” *Phys.Rev.*, vol. D64, p. 074001, 2001.
- [58] C.-H. Chen and C. Geng, “Lepton asymmetries in heavy baryon decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$,” *Phys.Lett.*, vol. B516, pp. 327–336, 2001.
- [59] 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.*, vol. B796, pp. 294–319, 2008.
- [60] L. Mott and W. Roberts, “Rare dileptonic decays of Λ_b^0 in a quark model,” *Int.J.Mod.Phys.*, vol. A27, p. 1250016, 2012.
- [61] T. Aliev, K. Azizi, and M. Savci, “Analysis of the $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decay in QCD,” *Phys.Rev.*, vol. D81, p. 056006, 2010.

- [62] R. Mohanta and A. Giri, “Fourth generation effect on Λ_b decays,” *Phys.Rev.*, vol. D82, p. 094022, 2010.
- [63] 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.*, vol. A24, pp. 6223–6235, 2009.
- [64] T. Aaltonen *et al.*, “Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$,” *Phys.Rev.Lett.*, vol. 107, p. 201802, 2011.
- [65] T. Aaltonen *et al.*, “Precise measurements of exclusive $b \rightarrow s \mu^+ \mu^-$ decay amplitudes using the full cdf data set.”
- [66] R. Aaij *et al.*, “Measurement of the differential branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$,” *Phys. Lett.*, vol. B725, p. 25, 2013.
- [67] 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.*, vol. D87, p. 074031, 2013.
- [68] 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.*, vol. B724, p. 27, 2013.
- [69] C. Lazzeroni, N. Watson, and M. Kreps, “Measurement of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ differential branching fraction.”
- [70] 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.
- [71] 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.
- [72] 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.*, vol. D87, no. 7, p. 074502, 2013.
- [73] R. Aaij *et al.*, “Precision measurement of the Λ_b^0 baryon lifetime,” *Phys.Rev.Lett.*, vol. 111, p. 102003, 2013.
- [74] 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.
- [75] 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.*, vol. B724, p. 27, 2013.

- 2560 [76] T. M. Karbach, “Feldman-Cousins Confidence Levels - Toy MC Method,” 2011.
- 2561 [77] G. Hiller and F. Kruger, “More model independent analysis of $b \rightarrow s$ processes,”
2562 *Phys.Rev.*, vol. D69, p. 074020, 2004.
- 2563 [78] G. Hiller and M. Schmaltz, “Diagnosing lepton-nonuniversality in $b \rightarrow s\ell\ell$,”
2564 *JHEP*, vol. 1502, p. 055, 2015.
- 2565 [79] J. Lees *et al.*, “Measurement of Branching Fractions and Rate Asymmetries in
2566 the Rare Decays $B \rightarrow K^{(*)}l^+l^-$,” *Phys.Rev.*, vol. D86, p. 032012, 2012.
- 2567 [80] J.-T. Wei *et al.*, “Measurement of the Differential Branching Fraction and
2568 Forward-Backword Asymmetry for $B \rightarrow K^{(*)}l^+l^-$,” *Phys.Rev.Lett.*, vol. 103,
2569 p. 171801, 2009.
- 2570 [81] R. Aaij *et al.*, “Test of lepton universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays,” 2014.
2571 to appear in Phys. Rev. Lett.
- 2572 [82] M. Pivk and F. R. Le Diberder, “SPlot: A Statistical tool to unfold data
2573 distributions,” *Nucl.Instrum.Meth.*, vol. A555, pp. 356–369, 2005.
- 2574 [83] “Lhcb loki twiki.” <https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters>. Accessed: 2015-09-30.
- 2576 [84] “Probnn presentation at ppts meeting.” <https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf>. Accessed: 2015-09-30.
- 2578 [85] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” *eConf*,
2579 vol. C0303241, p. MOLT007, 2003.
- 2580 [86] R. e. a. Aaij, “LHCb Detector Performance,” *Int. J. Mod. Phys. A*, vol. 30,
2581 p. 1530022. 82 p, Dec 2014.
- 2582 [87] A. Ali, E. Lunghi, C. Greub, and G. Hiller, “Improved model independent
2583 analysis of semileptonic and radiative rare B decays,” *Phys.Rev.*, vol. D66,
2584 p. 034002, 2002.
- 2585 [88] J. Hrivnac, R. Lednicky, and M. Smizanska, “Feasibility of beauty baryon po-
2586 larization measurement in $\Lambda^0 J/\psi$ decay channel by ATLAS LHC,” *J.Phys.G*,
2587 vol. G21, pp. 629–638, 1995.
- 2588 [89] J. Beringer *et al.*, “Review of particle physics,” *Phys. Rev.*, vol. D86, p. 010001,
2589 2012. and 2013 partial update for the 2014 edition.

2590

APPENDIX A

2591

2592

Decay models

2593

A.1 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution

2595 The angular distribution of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay is modelled using Ref. [88]. The
 2596 differential rate is written as

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

2597 where f_{1i} , f_{2i} and F_i are listed in Tab. A.1. The expression uses four observables
 2598 (angles) and depends on four complex amplitudes a_+ , a_- , b_+ , b_- and two real valued
 2599 parameters for the production polarisation, P_b , and the Λ decay asymmetry, α_Λ . The
 2600 angle θ is the angle of the Λ momentum in Λ_b^0 rest frame with respect to the vector
 2601 $\vec{n} = \frac{\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}}{|\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}|}$, where \vec{p}_{inc} and $\vec{p}_{\Lambda_b^0}$ are the momenta of incident proton and Λ_b^0 in the
 2602 center of mass system. The angles θ_1 and ϕ_1 are polar and azimuthal angle of the
 2603 proton coming from the Λ decay in the Λ rest frame with axis defined as $z_1 \uparrow\uparrow \vec{p}_\Lambda$,
 2604 $y_1 \uparrow\uparrow \vec{n} \times \vec{p}_\Lambda$. Finally, the angles θ_2 and ϕ_2 are the angles of the momenta of the
 2605 muons in J/ψ rest frame with axes defined as $z_2 \uparrow\uparrow \vec{p}_{J/\psi}$, $y_2 \uparrow\uparrow \vec{n} \times \vec{p}_{J/\psi}$.

2606 The distribution depends on the Λ decay asymmetry parameter, α_Λ , the production
 2607 polarization P_b and four complex amplitudes. The α_Λ is measured to be 0.642 ± 0.013
 2608 for Λ . The production polarization P_b and magnitudes of a_+ , a_- , b_+ and b_- are
 2609 measured in Ref. [75]. Phases are not measured therefore, as default all phases are
 2610 set to zero and then they are randomly varied to calculate the systematic uncertainty.

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.1.

₂₆₁₁ **A.2 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution**

₂₆₁₂ The q^2 and angular dependances of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays are modelled based on
₂₆₁₃ Ref. [67], where the angular distribution for unpolarized Λ_b^0 production is defined as

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

₂₆₁₄ In this formula θ_ℓ and θ_B correspond to the lepton and proton helicity angles, χ
₂₆₁₅ is angle between dimuon and Λ decay planes (for unpolarized production we are
₂₆₁₆ sensitive only to difference in azimuthal angles), $d_{i,j}^J$ are Wigner d-functions and h ,
₂₆₁₇ h^B and H are helicity amplitudes for virtual dimuon, Λ and Λ_b^0 decays. The sum
₂₆₁₈ runs over all possible helicities with the dimuon being allowed in spin 0 and 1 states
₂₆₁₉ (J and J'). The m and m' indices run over the vector and axial-vector current
₂₆₂₀ contributions.

₂₆₂₁ The production polarization is introduced by removing $e^{i(\lambda_j - \lambda'_j)\chi}$ from the expression,
₂₆₂₂ swapping small Wigner d-functions $d_{i,j}^J$ to the corresponding capital ones $D_{i,j}^J$ which
₂₆₂₃ are related as

$$D_{i,j}^J(\theta, \phi) = d_{i,j}^J(\theta) e^{i\phi(i-j)} \quad (\text{A.3})$$

₂₆₂₄ and substitute spin density matrix for $\delta_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} \delta_{JJ'}$. The spin density matrix
₂₆₂₅ itself is given by

$$\rho_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} = \frac{1}{2} \begin{pmatrix} 1 + P_b \cos \theta & P_b \sin \theta \\ P_b \sin \theta & 1 - P_b \cos \theta \end{pmatrix}. \quad (\text{A.4})$$

₂₆₂₆ Those changes lead to the formula

$$\begin{aligned} W(\theta\ell, \theta_B, \chi) &\propto \sum_{\lambda_1, \lambda_2, \lambda_j, \lambda'_j, J, J', m, m', \lambda_\Lambda, \lambda'_\Lambda, \lambda_p} h_{\lambda_1 \lambda_2}^m(J) h_{\lambda_1 \lambda_2}^{m'}(J') \\ &\times \rho_{\lambda_j - \lambda_\Lambda, \lambda'_j - \lambda'_\Lambda} D_{\lambda_j, \lambda_1 - \lambda_2}^J(\theta\ell, \phi_L) D_{\lambda'_j, \lambda_1 - \lambda_2}^{J'}(\theta\ell, \phi_L) H_{\lambda_\Lambda \lambda_j}^m(J) H_{\lambda'_\Lambda \lambda'_j}^{m'\dagger}(J') \\ &\times D_{\lambda_\Lambda \lambda_p}^{1/2}(\theta_B, \phi_B) D_{\lambda'_\Lambda \lambda_p}^{1/2}(\theta_B, \phi_B) h_{\lambda_p 0}^B h_{\lambda_p 0}^{B\dagger}. \end{aligned} \quad (\text{A.5})$$

₂₆₂₇ The lepton amplitudes come directly from Ref. [67], eq. 3. The Λ decay amplitudes
₂₆₂₈ are related to the Λ decay asymmetry parameter as

$$\alpha_\Lambda = \frac{|h_{\frac{1}{2}0}^B|^2 - |h_{-\frac{1}{2}0}^B|^2}{|h_{\frac{1}{2}0}^B|^2 + |h_{-\frac{1}{2}0}^B|^2}. \quad (\text{A.6})$$

₂₆₂₉ Finally, the Λ_b^0 decay amplitudes receive contributions from vector and axial-vector

²⁶³⁰ currents and can be written as

$$H_{\lambda_2, \lambda_j}^m = H_{\lambda_2, \lambda_j}^{Vm} - H_{\lambda_2, \lambda_j}^{Am}. \quad (\text{A.7})$$

Finally, the remaining amplitudes are expressed in terms of form factors (Ref. [67], eq. C6) as

$$\begin{aligned} H_{\frac{1}{2}t}^{Vm} &= \sqrt{\frac{Q_+}{q^2}} \left(M_- F_1^{Vm} + \frac{q^2}{M_1} F_3^{Vm} \right), \\ H_{\frac{1}{2}1}^{Vm} &= \sqrt{2Q_-} \left(F_1^{Vm} + \frac{M_+}{M_1} F_2^{Vm} \right), \\ H_{\frac{1}{2}0}^{Vm} &= \sqrt{\frac{Q_-}{q^2}} \left(M_+ F_1^{Vm} + \frac{q^2}{M_1} F_2^{Vm} \right), \\ H_{\frac{1}{2}t}^{Am} &= \sqrt{\frac{Q_-}{q^2}} \left(M_+ F_1^{Am} - \frac{q^2}{M_1} F_3^{Am} \right), \\ H_{\frac{1}{2}1}^{Am} &= \sqrt{2Q_+} \left(F_1^{Am} - \frac{M_-}{M_1} F_2^{Am} \right), \\ H_{\frac{1}{2}0}^{Am} &= \sqrt{\frac{Q_+}{q^2}} \left(M_- F_1^{Am} - \frac{q^2}{M_1} F_2^{Am} \right), \end{aligned} \quad (\text{A.8})$$

²⁶³¹ where $M_\pm = M_1 \pm M_2$, $Q_\pm = M_\pm^2 - q^2$. The form factors F are expressed in
²⁶³² terms of dimensionless quantities in eqs. C8 and C9 in Ref. [67]. In our actual
²⁶³³ implementation form factors calculated in the covariant quark model [67] are used
²⁶³⁴ and for the numerical values of the Wilson coefficients Ref. [67] 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. [72]. The form factors relations are found by comparing eqs. 66 and 68 in Ref. [67] to eq. 51 in Ref. [72]. 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

$$\begin{aligned} c_\gamma &= 1 - \frac{\alpha_s(\mu^2)}{\pi} \left[\frac{4}{3} + \ln \left(\frac{\mu}{m_b} \right) \right], \\ c_v &= \frac{2}{3} \frac{\alpha_s(\mu^2)}{\pi}, \\ c_\sigma &= 1 - \frac{\alpha_s(\mu^2)}{\pi} \left[\frac{4}{3} + \frac{5}{3} \ln \left(\frac{\mu}{m_b} \right) \right]. \end{aligned} \quad (\text{A.9})$$

2635 In the calculations $\mu = m_b$ is used. For the strong coupling constant, we start from
 2636 the world average value at the Z mass, $\alpha_s(m_Z^2) = 0.1185 \pm 0.0006$ [89], and we
 2637 translate it to the scale m_b^2 by

$$\alpha_s(\mu^2) = \frac{\alpha_s(m_Z^2)}{1 + \frac{\alpha_s(m_Z^2)}{12\pi} (33 - 2n_f) \ln \left(\frac{\mu^2}{m_Z^2} \right)}, \quad (\text{A.10})$$

2638 where $n_f = 5$. The LQCD form factors F_1^L and F_2^L can be then taken directly from
 2639 Ref. [72] and plugged into the code implementing the calculation from Ref. [67].

2640

APPENDIX B

2641

2642

Monte Carlo and data comparison

2643

2644 This appendix reports a comparison between distributions in data and simulated Λ_b^0
 2645 $\rightarrow J/\psi \Lambda$ events. In the plots what is labeled as "Data" is real data in a 20 MeV
 2646 interval around the Λ_b^0 mass, where a sideband subtraction technique to remove
 2647 background. "Side" is real data for masses above 6 GeV containing mostly combi-
 2648 natorial backgrounds. These can be compared to the previous sample to see which
 2649 variables differ the most. "MC" corresponds to Pythia8 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ simulated
 2650 events. Finally, the label "MC fully W" refers to the same simulated sample but
 2651 weighted for the Λ_b^0 and Λ kinematics (Sec. 3.3.2) and the decay model (Sec. 3.3.1).
 2652 Distributions are shown separately for long and downstream events.

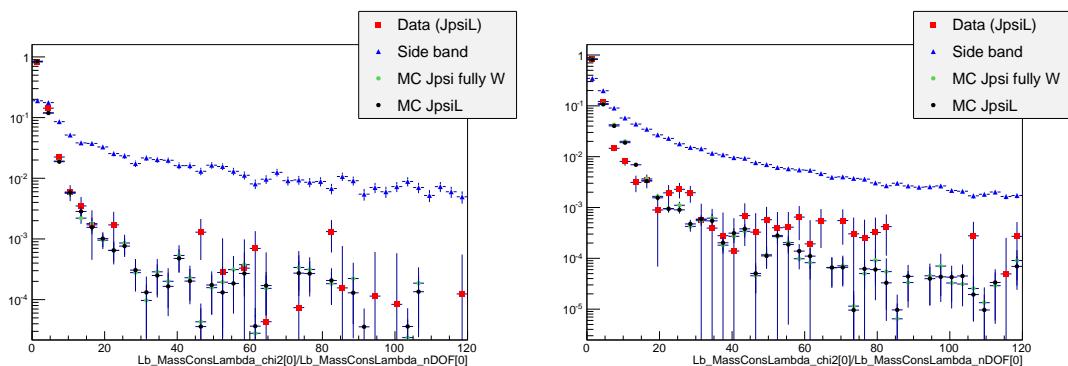


Figure B.1: Distributions of χ^2/NdF of the kinematic fit in data and simulation for LL (left) and DD (right) events.

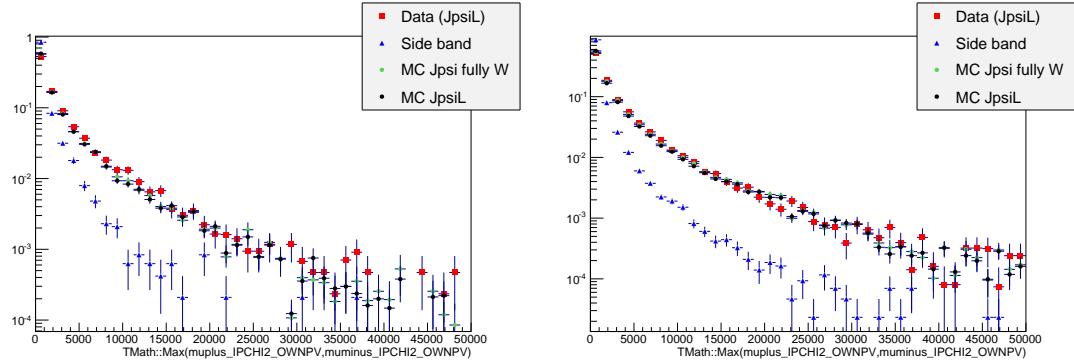


Figure B.2: Distributions of maximum muon $IP\chi^2$ variable in data and simulation for LL (left) and DD (right) events.

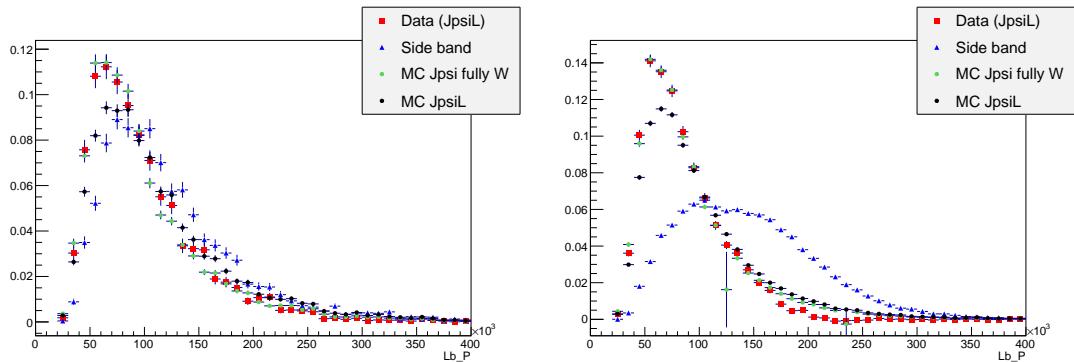


Figure B.3: Distributions of Λ_b^0 momentum variable in data and simulation for LL (left) and DD (right) events.

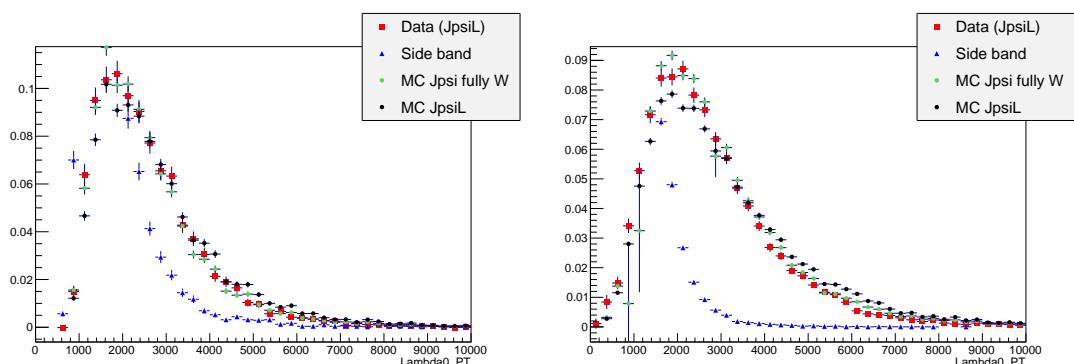


Figure B.4: Distributions of Λ transverse momentum variable in MC, data signal and data background for LL (left) and DD (right) events.

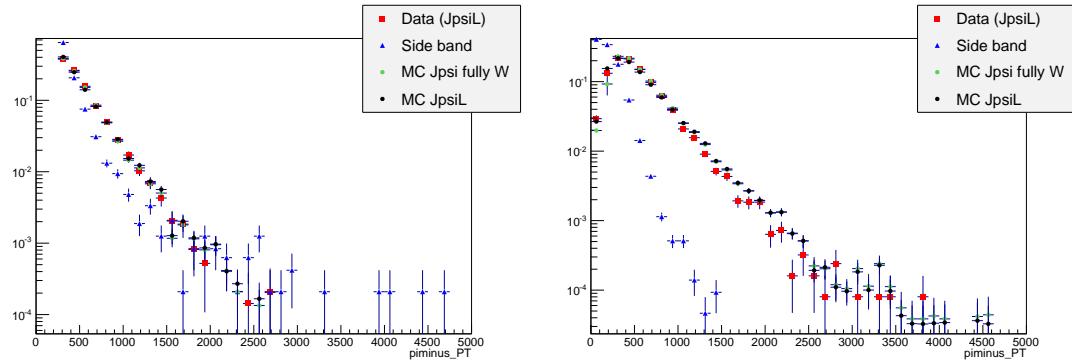


Figure B.5: Distributions of pion transverse momentum variable in data and simulation for LL (left) and DD (right) events.

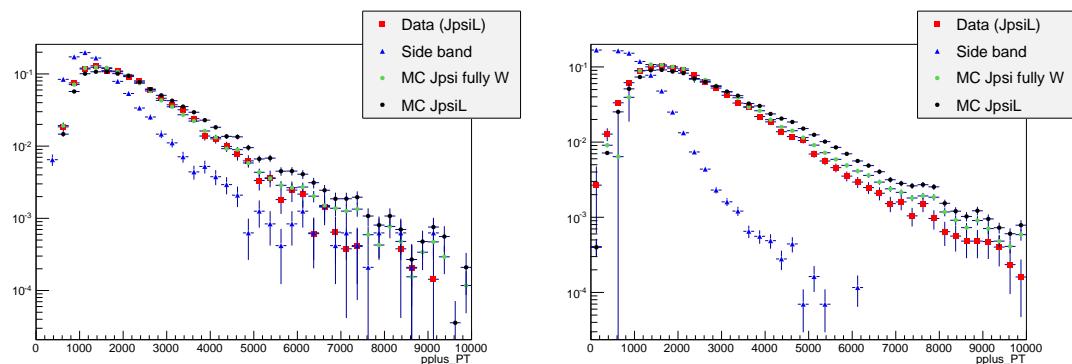


Figure B.6: Distributions of proton transverse momentum variable in data and simulation for LL (left) and DD (right) events.

2653

APPENDIX C

2654

2655 **Systematic uncertainties on the efficiency calculation for the**
 2656 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction analysis.

2657

2658 This appendix reports systematic uncertainties on absolute and relative efficiencies
 2659 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.

2660

APPENDIX D

2661

2662 **Invariant mass fits to $B^0 \rightarrow K^{*0}e^+e^-$ simulated candidates**

2663

2664 This appendix contains fits to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}e^+e^-$ sim-
2665 ulated candidates used to constrain parameters in the fit to data.

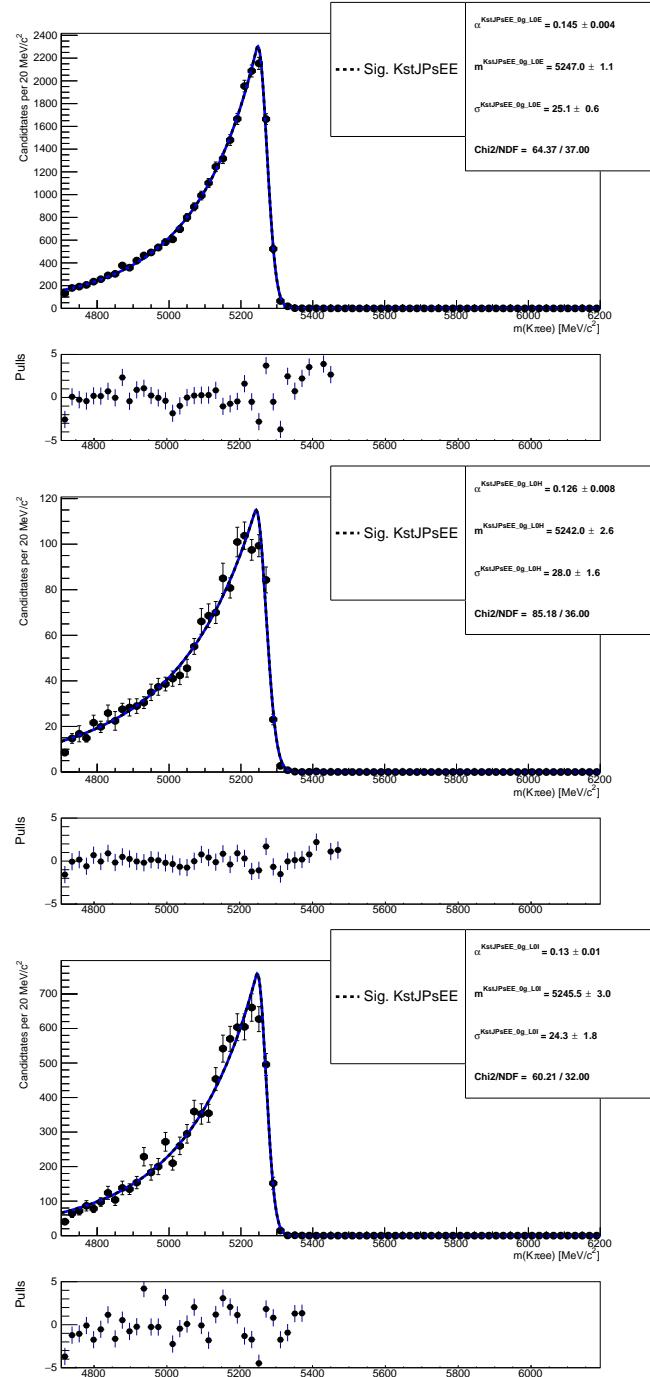


Figure D.1: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and no photon emitted.

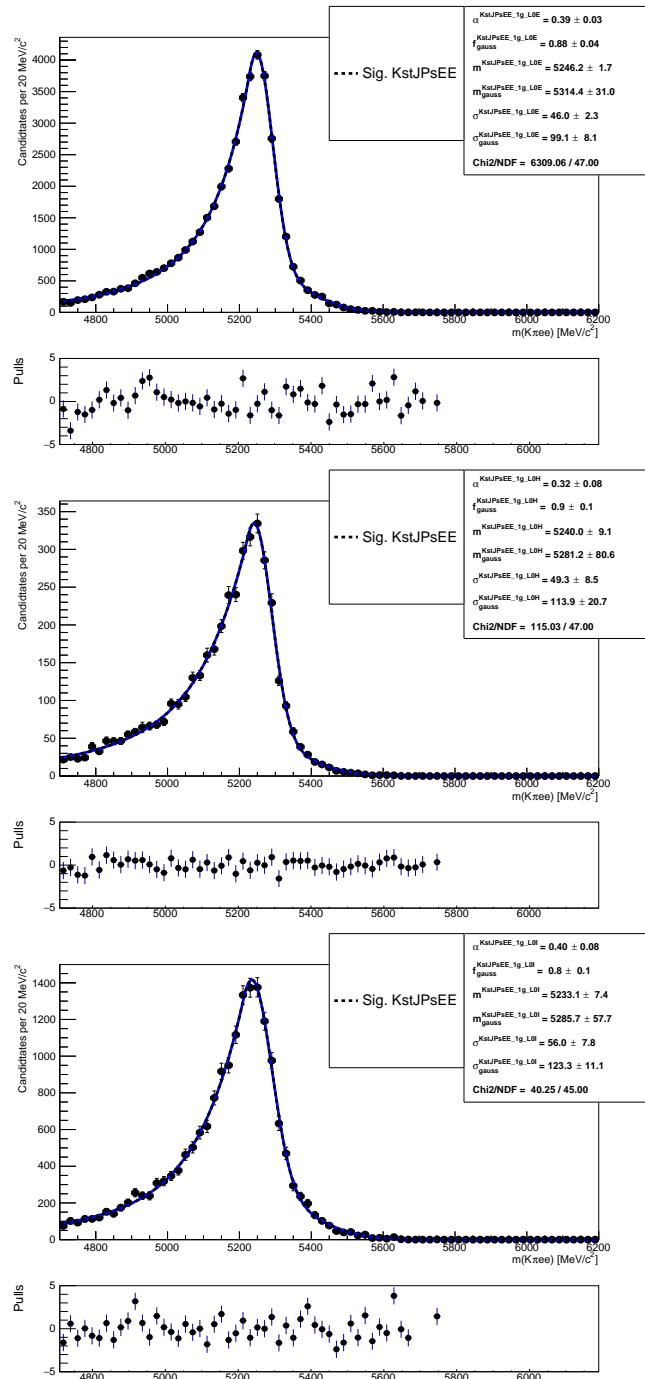


Figure D.2: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and one photon emitted.

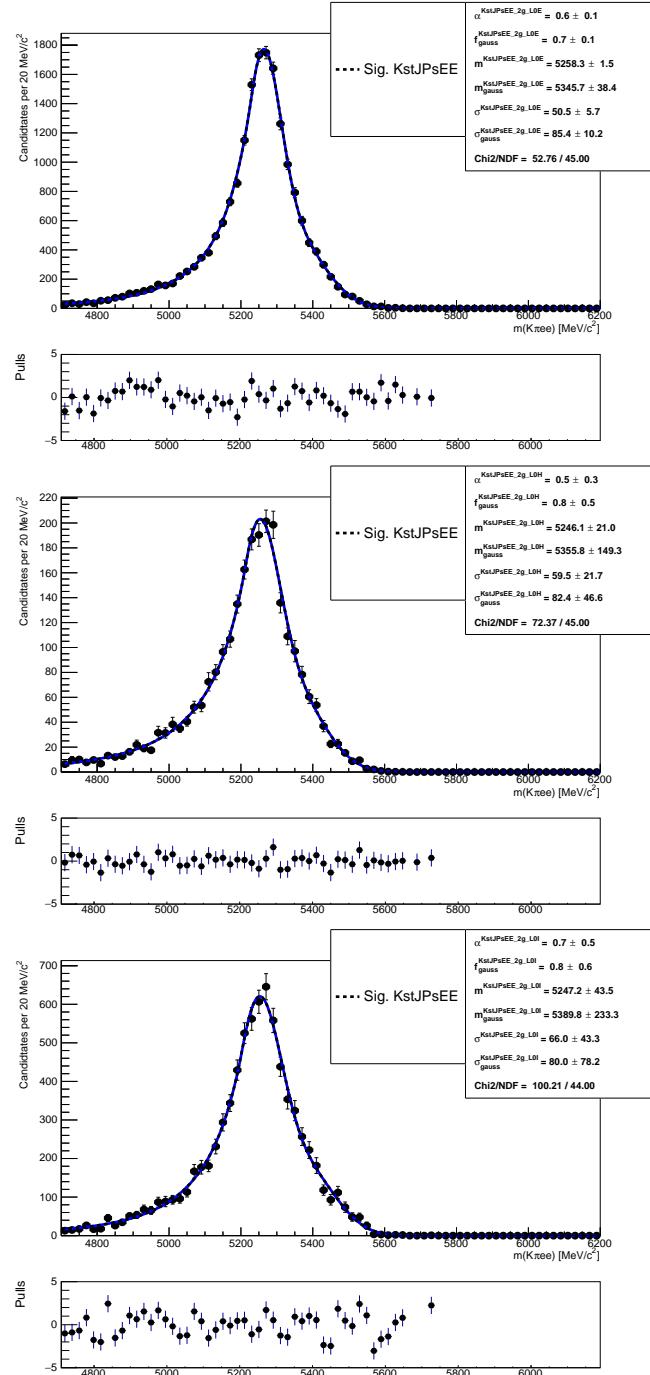


Figure D.3: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and two photons emitted.