

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.

September 19, 2015

ABSTRACT

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

DECLARATION OF AUTHORS CONTRIBUTION

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

ACKNOWLEDGEMENTS

I thank everybody, evvvvvvvveverybody!

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

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

Contents

1	Introduction	1
1.1	The electroweak interaction	4
1.2	Flavour and the CKM matrix	5
1.3	The puzzles of the SM	9
1.3.1	The flavour problem	10
1.4	Beyond the Standard Model	11
1.4.1	Flavour and BSM theories	11
1.5	Rare decays: a tool to search for new physics	13
1.5.1	Theoretical framework: the effective Hamiltonian	14
1.5.2	Operators	16
1.5.3	Phenomenology of $b \rightarrow s\ell^+\ell^-$ decays	18
1.5.4	Observables in $b \rightarrow s\ell^+\ell^-$ decays	19
1.6	Experimental status	20
1.6.1	Dimuon decays of b hadrons	20
1.6.2	Semileptonic $b \rightarrow s\ell^+\ell^-$ decays of b hadrons	22
1.6.3	Lepton Flavour Violation searches	23
2	The LHCb detector at the Large Hadron Collider	25
2.1	The Large Hadron Collider	25
2.2	The LHCb detector	27
2.3	The magnet	29
2.4	Tracking system	29
2.5	Calorimeters	32
2.5.1	Bremsstrahlung recovery for electrons	34
2.6	RICH	35
2.7	The muon system	36
2.8	Particle identification	37
2.8.1	PID calibration	39
2.9	Trigger and software	39
2.10	Kinematical fits	41
2.11	Validation of hadronic processes in the simulation	42
2.11.1	Geometry and interaction probability	44
2.11.2	PDG prediction	45
2.11.3	Validation results	46
2.12	Material budget studies	49
3	Differential branching fraction of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$	51

3.1	Analysis strategy and q^2 regions	53
3.2	Candidate types	54
3.3	Simulation	55
3.3.1	Decay Model	56
3.3.2	Kinematic re-weighting	56
3.3.3	Event type	58
3.4	Selection	59
3.4.1	Pre-selection	59
3.4.2	Neural Networks	60
3.4.3	MVA optimisation	65
3.4.4	Trigger	67
3.4.5	Background from specific decays	68
3.5	Yield extraction	69
3.5.1	Fit description	70
3.5.2	Fit results	76
3.6	Efficiency	81
3.6.1	Geometric acceptance	81
3.6.2	Reconstruction and neural network efficiencies	82
3.6.3	Trigger efficiency	83
3.6.4	PID efficiency	84
3.6.5	Relative efficiencies	85
3.7	Systematic uncertainties	89
3.7.1	Systematic uncertainty on the yields	89
3.7.2	Systematic uncertainties on the efficiency determination	91
3.7.2.1	Effect of new physics on the decay model	91
3.7.2.2	Simulation statistics	91
3.7.2.3	Production polarisation and decay structure	92
3.7.2.4	Λ_b^0 lifetime	93
3.7.2.5	Downstream candidates reconstruction efficiency	93
3.7.2.6	Data-simulation discrepancies	94
3.8	Differential branching ratio extraction	94
4	Angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays	99
4.1	One-dimensional angular distributions	100
4.2	Multi-dimensional angular distributions	102
4.3	Angular resolution	104
4.4	Fit strategy	107
4.4.1	Feldman-cousins plug-in method	108
4.4.2	Modelling the angular distributions	109
4.4.3	Angular acceptance	110
4.4.4	Studies on a three-dimensional fit	112
4.5	Systematics uncertainties on angular observables	113
4.5.1	Angular correlations	114
4.5.2	Resolution	114
4.5.3	Efficiency description	115
4.5.4	Background parameterisation	116

4.5.5	Polarisation	117
4.6	J/ψ cross-check	118
4.7	Results	120
5	Testing lepton flavour universality with $R_{K^{*0}}$	124
5.1	Combining ratios	127
5.2	Experimental status	128
5.3	Analysis strategy	129
5.4	Choice of q^2 intervals	130
5.5	Data samples and simulation	131
5.5.1	Data-simulation corrections	131
5.6	Selection	132
5.6.1	Trigger and Stripping	134
5.6.2	PID	138
5.6.3	Peaking backgrounds	139
5.6.3.1	Charmonium vetoes	140
5.6.3.2	ϕ veto	140
5.6.3.3	$B^+ \rightarrow K^+ \ell^+ \ell^-$ plus a random pion	141
5.6.3.4	Λ_b decays	141
5.6.3.5	Other peaking backgrounds	142
5.6.4	Mis-reconstructed background	142
5.6.5	Multivariate analysis	143
5.6.6	MVA optimisation	147
5.7	Mass fits	147
5.7.1	Muon channels	149
5.7.2	Electron channels	151
5.7.2.1	Signal PDFs for the electron channels in the central- q^2 interval .	153
5.7.2.2	Background PDFs for the electron channels in the central- q^2 interval .	155
5.7.2.3	Summary: electron channels fit in the central- q^2 interval	156
5.7.2.4	Electron channels fits in the high- q^2 interval	157
5.7.3	Fit summary	162
5.8	Efficiency	164
5.8.1	Geometric efficiency	167
5.8.2	Reconstruction efficiency and bin migration	167
5.8.3	PID efficiency	169
5.8.4	Trigger efficiency	169
5.8.4.1	Electron channels	170
5.8.4.2	TISTOS cross-check	172
5.8.5	Neural Networks efficiency	173
5.9	Systematic uncertainties	175
5.9.1	Choice of PDF	175
5.9.2	Statistical error on the efficiency determination	176
5.9.3	TISTOS	176
5.9.4	Bin migration	176
5.10	Result extraction	178

5.10.1 $R_{J/\psi}$ sanity check	178
5.10.2 R_{K^*} result summary	179
5.10.3 Branching ratios and expectations	179
6 Conclusions	181
A Decay models	193
A.1 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution	193
A.2 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution	197
B Data-simulation comparison	198
C Systematic uncertainties on the efficiency calculation for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction analysis.	201
D Invariant mass fits to $B^0 \rightarrow K^{*0} e^+ e^-$ simulated candidates	204
E Extra	208
E.0.1 Quantum Electrodynamics: an example of gauge field theory	208
E.1 Anomalies	209

1

CHAPTER 1

2

3

Introduction

4

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

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

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

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

Particles included in the SM can be grouped under a few categories depending on their properties and ability to interact with each other. The first distinction is between fermions, half-integer spin particles, and bosons, integer spin particles. Fermions constitute the basic building blocks of matter, while bosons are the mediators of the interactions. Since the concept of bosonic mediators of interactions arises because of gauge symmetry [4], they are called “gauge bosons”. The list of the

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

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

known interactions with their force carrier and properties is reported in Tab. 1.2. The matter of which we are made of is mainly composed of electrons and protons, which have spin 1/2; protons are in turn composed of u and d quarks, which again have spin 1/2. Among fermions one can then consider two smaller groups: quarks and leptons. Quarks carry colour charge and therefore can interact through the,

³⁷ so called, strong interaction, while leptons, which do not carry colour charge, are
³⁸ insensitive to it. For each particle a corresponding anti-particle exists with opposite
³⁹ quantum numbers. Finally, fermions are divided into three families having similar
⁴⁰ properties but different masses. This last structure embedded in the SM is also
⁴¹ called “flavour structure” and it will be the main tool used in this thesis; a more
⁴² detailed description of it is given in the next sections. A schematic view of the fundamental particles in the SM is shown in Fig. 1.1. Due to the asymptotic freedom

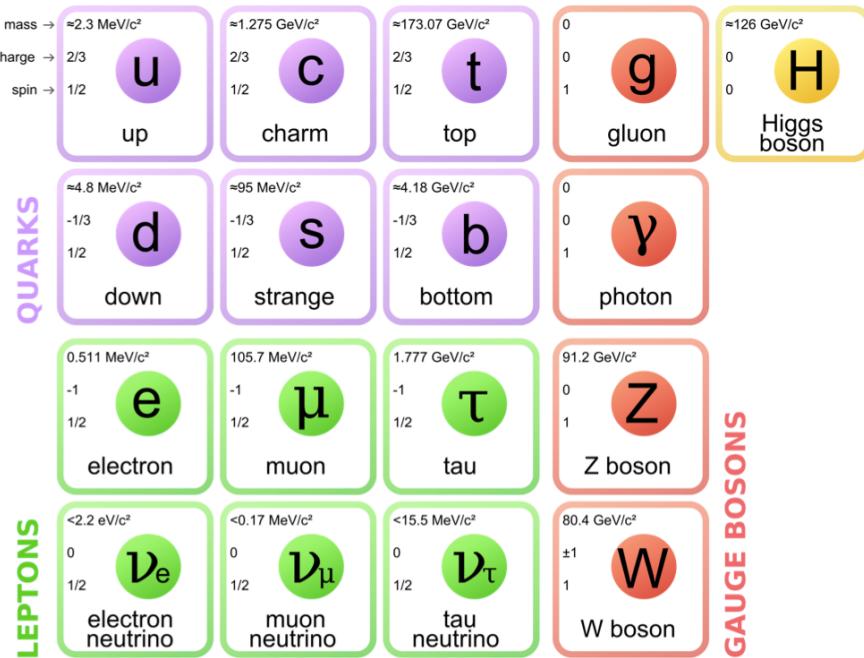


Figure 1.1: A scheme of the fundamental particles in the SM with their properties.

⁴³

⁴⁴ of the strong interaction quarks cannot be observed alone but are always combined
⁴⁵ with other quarks to form color singlets. Non-fundamental particles composed by
⁴⁶ quarks are called hadrons and can be divided in mesons, where the color singlet is
⁴⁷ achieved by the combination of a quark and its antiquark ($q \bar{q}$), and baryons formed
⁴⁸ by three quarks ($q q q$) of different colours. Recently, in 2014 and 2015 evidence for
⁴⁹ new states, formed by four and five quarks, was found [5, 6].

50 1.1 The electroweak interaction

51 The electromagnetic interaction is responsible for binding electrons and nuclei to-
52 gether in atoms and its mediator is the photon. The weak interaction is responsible
53 for the β decay of nuclei and is mediated by the emission or absorption of W^\pm and
54 Z bosons. Unlike the electromagnetic force, that affects only charged particles, all
55 known fermions interact through the weak interaction. The weak interaction is also
56 the only one that violates the parity symmetry, which states that interactions are
57 invariant under a reflection of all coordinates. This symmetry breaking arises from
58 the fact that only left-handed fermions interact through the weak interaction as dis-
59 covered by Wu in 1957 [7]. Similarly, the weak interaction is the only one that also
60 breaks the CP symmetry, which combines parity transformations and charge conju-
61 gation. This is particularly interesting because all interactions are invariant under
62 the CPT transformation, which combines the CP transformation and time reversal,
63 hence, breaking CP the weak interaction must also be not invariant under time re-
64 versal. In 1968 Salam, Glashow and Weinberg unified the weak and electromagnetic
65 forces in a single theory, where the coupling constants of the electromagnetic, e ,
66 and weak, g , interactions are linked by the weak mixing angle, θ_W by the relation
67 $g \sin \theta_W = e$ [2]. The electroweak symmetry is spontaneously broken by the Higgs
68 field [8] and this causes the W^\pm and Z bosons to become massive (see Tab. 1.2)
69 and consequently the weak force has a very short range. In fact using Heisenberg's
70 Principle ($\Delta E \Delta t > \hbar$) together with Einstein's formula $\Delta E = mc^2$, which relates
71 mass and energy, and knowing that the maximum space that a particle can cover
72 in a time Δt is $r = c\Delta t$, qualitatively $r \sim \hbar/mc$. In this picture the carriers of the
73 weak force can travel $r \sim 2 \cdot 10^{-3}$ fm. The photon must instead be massless in the
74 theory, which accounts for the long range of the electromagnetic force. The EW
75 interactions are divided into charged currents (CC) and neutral currents (NC). In
76 the first group, quarks and leptons interact with the W^\pm bosons, producing decays
77 such as $\mu^+(\mu^-) \rightarrow e^+ \nu_e \bar{\nu}_\mu (e^- \bar{\nu}_e \nu_\mu)$ and $n \rightarrow p e^- \bar{\nu}_e (\bar{p} e^+ \nu_e)$. The study of these pro-
78 cesses confirmed that only the left-handed (right-handed) component of fermions
79 (anti-fermions) takes part in weak processes. The CC interactions have a peculiar-

80 ity: they are the only interactions in the SM that violate flavour conservation at
81 tree level (see next section), while any other interaction not conserving flavour has
82 to happen through loops. The second group of EW interactions, NC, corresponds
83 to diagrams mediated by a photon or a Z boson interacting with a fermion and its
84 anti-fermion.

85 1.2 Flavour and the CKM matrix

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

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

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

102 Measuring branching fractions of weak decays like $\pi \rightarrow \mu\nu_\mu$ and $K \rightarrow \mu\nu_\mu$, corre-
103 sponding respectively to $ud \rightarrow \mu\nu_\mu$ and $us \rightarrow \mu\nu_\mu$ processes, suggested the existence
104 of more than one coupling constant for different quarks. Nicola Cabibbo [2], in order

to preserve the universality of weak interactions, suggested that the branching fraction differences could arise from the fact that the doublets participating in the weak interactions are an admixture of the flavour eigenstates. He therefore introduced the Cabibbo angle, θ_c , considering that eigenstates participating to the weak interaction are rotated with respect of the flavour eigenstates.

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

Considering a 6 quark system one angle is not enough to describe a rotation but the mixing can be generalised using a 3×3 unitary matrix, which is called CKM matrix, from the names of Cabibbo, Kobayashi and Maskawa. The unitarity of the matrix is required to preserve the universality of the weak interaction. Theoretically, a $N \times N$ complex matrix depends on $2 \cdot N^2$ real parameters. Requiring unitarity ($AA^\dagger = A(A^*)^T = I$), the number of independent parameters left is $(N - 1)^2$. Therefore a 3×3 matrix depends then on 4 real parameters, which can be divided in 3 real constants and one imaginary phase. The imaginary phase generates the CP-violation which was observed in weak interactions. Equation 1.2 reports a parametrisation of

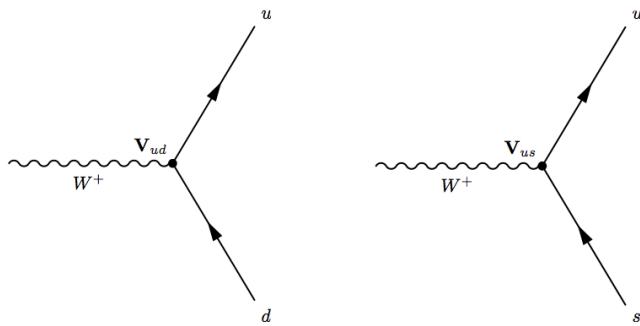


Figure 1.2: Feynman diagrams with CKM weights on weak interaction vertices

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

and Eq. 1.3 shows their relations with the 3 mixing angles.

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

110

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

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

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

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

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

118 These correspond to constraints to three complex numbers, which can be viewed
 119 as the sides of triangles in the (ρ, η) plane; these are called “unitarity triangles”.
 120 The most commonly used unitarity triangle arises from $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$.
 121 Figure 1.3 shows a representation of such triangle together with a plot summarising
 122 the most up to date experimental constraints to its parameters [9]. The precise
 123 measurement of the parameters of the CKM matrix is a powerful stability test of

¹²⁴ the SM and sets a solid base for new physics searches in the flavour sector. One of
¹²⁵ the main goals of the LHCb experiment is to precisely measure the angle γ , which
is currently the least constrained by measurements.

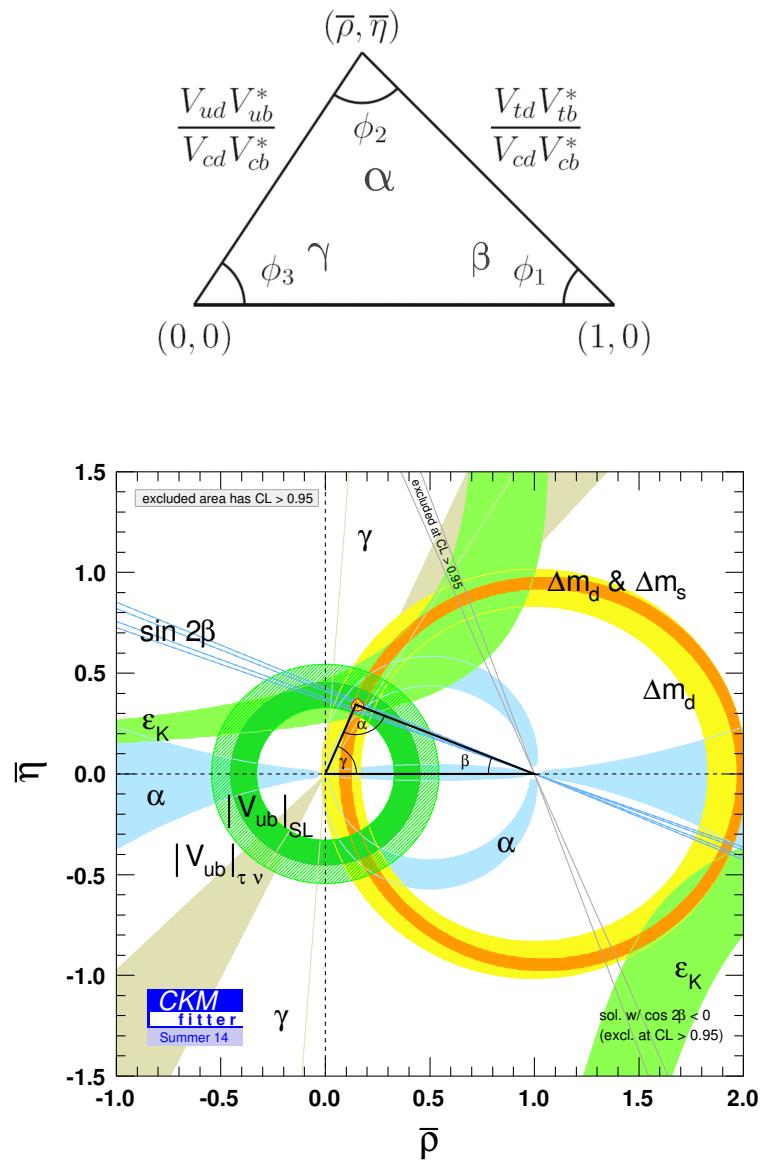


Figure 1.3: (top) A representation of the unitarity triangle and its parameters. (bottom) A summary of the most up to date measurements of the unitarity triangle parameters [9].

127 1.3 The puzzles of the SM

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

130 • *Dark matter*: experimental evidence tells us that the content of visible matter
131 in the universe is not enough to account for the observed rotation of galaxies [10]. The most natural way to solve the problem is the hypothesis of a form
132 of matter that interacts with the gravitational field but not with the other SM
133 interactions.

135 • *Matter-antimatter asymmetry*: a large asymmetry is observed between the
136 quantity of matter and antimatter in the universe, $O(10^{-9})$. Assuming that
137 both were equally created in the initial state of the universe, a condition such
138 as the violation of the CP symmetry is necessary to account for the observed
139 imbalance. However, the magnitude of CP violation predicted by the SM,
140 $O(10^{-20})$, is not enough to explain the observed imbalance [11].

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

147 • *Neutrino oscillation*: measurements regarding solar and atmospheric neutrinos
148 as wells as neutrinos from nuclear reactors established that neutrinos can
149 change flavour while propagating in space. This is not predicted in the SM, in
150 fact in the SM neutrinos are massless, while an oscillation requires a non zero
151 mass [12, 13, 14, 15].

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

¹⁵⁴ energy scales. As new physics can appear anywhere up to the Planck scale,
¹⁵⁵ $\sim 10^{19}$ GeV, at which gravity cannot be neglected any more, these corrections
¹⁵⁶ can be very large and it would require a high level of fine-tuning for them to
¹⁵⁷ cancel out and give such a small value as the one measured for the Higgs Mass,
¹⁵⁸ ~ 126 GeV/ c^2 [16, 17].

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

¹⁶³ 1.3.1 The flavour problem

¹⁶⁴ Flavour Changing Charged Currents (FCCC) that are mediated by the W^\pm bosons
¹⁶⁵ are the only sources of flavour changing interaction in the SM and, in particular, of
¹⁶⁶ generation changing interactions, where a quark or a lepton of a family transforms
¹⁶⁷ into one of another family. Another class of processes is the Flavour Changing
¹⁶⁸ Neutral Currents (FCNCs), e.g. transitions from a b quark with a charge of -1/3 to
¹⁶⁹ a s or d quark with the same charge. Examples of FCNC transitions in the quark
 and lepton sector are shown in Fig. 1.4. In the SM there is no fundamental reason

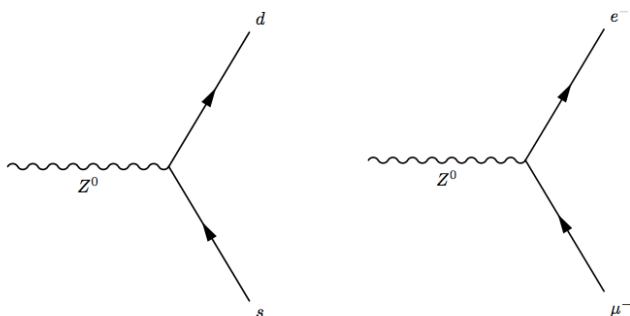


Figure 1.4: Feynman diagrams of FCNCs processes forbidden in the SM.

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

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

178 1.4 Beyond the Standard Model

179 From the last sections it is evident that, despite the great success of the SM, there
180 is a need to explore theories Beyond the SM (BSM). Among the most promising
181 approaches there are those involving Super-Symmetry and extra-dimensions. In
182 Super-Symmetry new degrees of freedom are introduced to suppress the diverging
183 terms of the scalar mass. This theory assumes that for each fermion there is a
184 corresponding boson and, since bosons and fermions contribute with opposite sign to
185 the mass term, these would naturally cancel out [18]. Supersymmetry also provides
186 a candidate for dark matter, the neutralino, which is a weakly interacting stable
187 particle. The idea to introduce extra-dimensions was triggered by the fact that,
188 normally, gravity is not relevant in particle physics but it would be natural if all
189 forces had similar strength. By adding extra dimensions to the normal 3 spatial
190 dimensions, one can restore the strength of gravity, as this could be dispersed by the
191 wider space available [19]. In all these approaches severe constraints to masses and
192 couplings must be imposed to maintain compatibility with the SM at the electroweak
193 scale.

194 1.4.1 Flavour and BSM theories

195 Most BSM theories predict processes violating flavour conservation. Therefore, the
196 observation or non-observation of these processes can give important information
197 about new physics. BSM theories can be classified according to the amount of flavour

violation they introduce. The first class of models to consider is those with Minimal Flavour Violation (MFV). These are models in which the only sources of flavour changing transitions are governed by the CKM matrix and the CKM phase is the only source of CP violation. These features can be assured by symmetry principles, which makes these types of models naturally compatible with the SM. Examples of such models include the MSSM with minimal flavour violation and the SM with one extra-dimension. Reviews of MFV models are presented in Refs. [20, 21]. The MFV paradigm provides a way to resolve the tension between expectation, driven by naturalness arguments, that NP should be at the TeV scale and limits on FCNC processes that point to much higher scales. A powerful test of MFV is provided by the study of ratios between $b \rightarrow d$ and $b \rightarrow s$ transitions, because their hamiltonians share the same structure. One particularly important example is the ratio of B^0 and B_s^0 dimuon decay rates [22], as this is a purely leptonic decay free from hadronic uncertainties. In the SM such ratios are approximately equal to $|V_{td}/V_{ts}| \sim 1/25$, only modified by phase space and hadronic matrix elements, while they can take very different values in non-MFV models.

In the quest for new physics an important role is also played by simplified models as an intermediate model building step. Instead of constructing models valid up to the GUT scale one can consider simplified models, where the SM is extended incorporating a new sector with a limited number of parameters. Such models are easier to constrain but can nevertheless point in the right direction to build more complete theories. The choice of the new sector to add can be driven by the need to explain existing tensions between data and SM predictions or by theoretical prejudice. Two models especially relevant when studying rare decays, which are the main topic of this thesis, are Z'-penguins and leptoquarks. A Z'-penguin is a FCNC process involving a neutral field arising from an extra U(1) gauge symmetry. As for the SM penguins, this field contributes in loops causing modifications of the effective couplings with respect to the SM. A survey of Z' models can be found in Ref. [23]. Leptoquarks are bosonic particles that carry both quark and lepton flavour quantum numbers. For simplicity they are commonly assumed to be scalar particles.

228 A tree level exchange of a leptoquark induces processes such as $b \rightarrow (s, d)\ell^+\ell^-$, and
229 therefore can result in an enhancement of their decay rates with respect to the
230 SM [24]. Leptoquarks would also provide a natural explanation for non-universal
231 couplings to leptons.

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

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

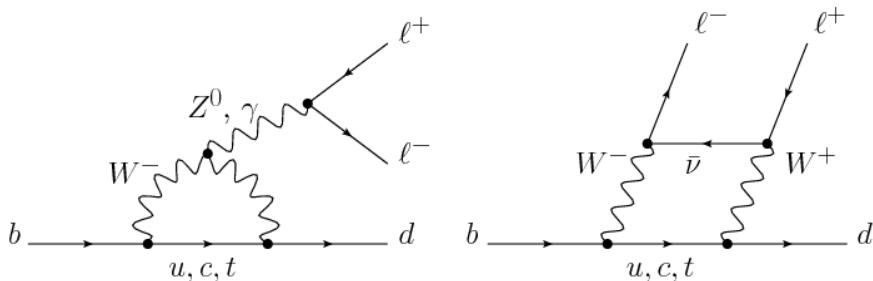


Figure 1.5: Loop Feynmann diagrams allowing $b \rightarrow d$ FCNC processes: penguin diagram (left) and W box (right).

²⁴⁷ 1.5.1 Theoretical framework: the effective Hamiltonian

²⁴⁸ Rare decays of b hadrons are governed by an interplay between weak and strong
²⁴⁹ interactions. The large masses of W , Z and top quark compared to that of the b quark
²⁵⁰ allow the construction of an effective theory that divides the problem of calculating
²⁵¹ weak decay amplitudes into two parts: “short-distance” and “long-distance” effects
²⁵² separated at an energy scale μ . The first part, dealing with short distance physics,
²⁵³ handles perturbative contributions due to energy scales above the b mass. The
²⁵⁴ second part typically deals with non-perturbative contributions. A classic example
²⁵⁵ of an effective theory is the Fermi theory of weak interactions which describes the
²⁵⁶ β decay in terms of a four-fermion interaction, where the short distance physics is
hidden into a point like vertex as illustrated in Fig. 1.6.

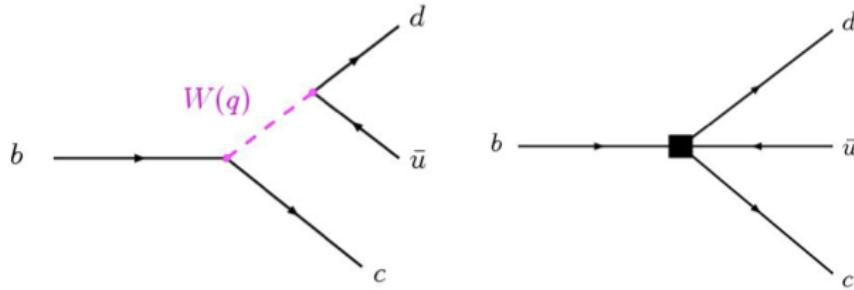


Figure 1.6: Example of a Fermi theory in which the full theory is divided between a short distance contribution, hidden in the vertex, and a long distance contribution.

²⁵⁷

²⁵⁸ The effective hamiltonian [25] relevant to $b \rightarrow s/d\gamma$ and $b \rightarrow s/d\ell^+\ell^-$ transitions
²⁵⁹ can be written as:

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

²⁶⁰ where G_F denotes the Fermi coupling constant and the λ constants are the CKM
²⁶¹ factors, $\lambda_q^t = V_{tb}V_{tq}^*$ and $\lambda_q^u = V_{ub}V_{uq}^*$. In $b \rightarrow s$ quark transitions, which are the
²⁶² main topic of this thesis, the doubly Cabibbo-suppressed contributions proportional
²⁶³ to λ_s^u can be neglected. To obtain this formula the Operator Product Expansion
²⁶⁴ (OPE) [26] method is used, which implements a summation over all contributing
²⁶⁵ operators weighted by corresponding constants called Wilson coefficients. In this

266 Hamiltonian the long-distance contributions are described by the operators, \mathcal{O}_i ,
267 while the short-distance physics is encoded in the Wilson Coefficients, C_i . Operators
268 and coefficients are evaluated at the renormalization scale μ . Any particle that
269 contributes to the decay and has a mass greater than the scale μ will affect the
270 value of at least one of the Wilson coefficients, including SM particles as the top
271 quark.

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

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

277 where the matrix γ is the anomalous dimensions matrix of the operators \mathcal{O}_i . At
278 leading order the solution is given by [27]:

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

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

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

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

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

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

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

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

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

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

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

²⁸⁹ source of uncertainty in theoretical calculations.

²⁹⁰ 1.5.2 Operators

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

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

²⁹² A complete basis is given by a set of 10 operators, where $\mathcal{O}_{1,2}$ are the tree level

²⁹³ W operators; $\mathcal{O}_{3-6,8}$ are penguin diagrams mediated by gluons; and $\mathcal{O}_{7,9,10}$, which

²⁹⁴ are the operators that are relevant for radiative and leptonic penguin processes are

²⁹⁵ defined as [22]:

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

²⁹⁶ where $P_{L/R} = (1 \mp \gamma_5)/2$ denote the left/right handed chiral projection, T^a are the

²⁹⁷ QCD generators and $F_{\mu\nu}$ is the electromagnetic field tensor. The \mathcal{O}' operators cor-

²⁹⁸ respond to right-handed coupling obtained by swapping P_R and P_L in the equations.

²⁹⁹ In the SM, as well as in MFV models where the flavour violation is entirely ruled by

³⁰⁰ the CKM matrix, the C' Wilson Coefficients are suppressed by the strange coupling,

³⁰¹ $C'_i \sim (m_s/m_b)C_i$. The operator \mathcal{O}_7 relates to penguin diagrams that are mediated

³⁰² via a photon. It represents the dominant contribution to the radiative $b \rightarrow s\gamma$

303 transition and contributes to $b \rightarrow s\ell^+\ell^-$ processes when the virtual photon decays
304 into a dilepton pair. The semileptonic \mathcal{O}_9 and \mathcal{O}_{10} correspond to penguin diagrams
305 mediated by a Z boson and W mediated box diagrams. These are the dominant
306 contributions in semileptonic $b \rightarrow s\ell^+\ell^-$ decays. The vertices corresponding to the
radiative and semileptonic operators are illustrated in Fig. 1.7

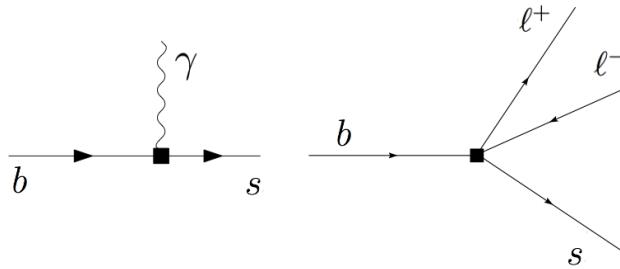


Figure 1.7: Interaction vertices corresponding to the radiative (left) and semileptonic (right) operators.

307

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

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

310 where the Wilson Coefficients are redefined as

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

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

317 1.5.3 Phenomenology of $b \rightarrow s\ell^+\ell^-$ decays

318 Semileptonic b hadron decays are characterised by two kinematic regimes which are
 319 treated theoretically in different ways; Table 1.3 shows a scheme of the q^2 spectrum.
 320 The ‘high q^2 ’ is the region of low hadron recoil, $q^2 > 15 \text{ GeV}^2/c^4$, and is charac-
 321 terised by the energy of the hadron being less than the energy scale of QCD interac-
 322 tions within the meson, $\Lambda_{QCD} \sim 1 \text{ GeV}$. In this region theoretical calculations of B
 323 meson decays can be simplified by working in the heavy quark limit, $m_b \rightarrow \infty$. In
 324 this limit a Heavy Quark Effective Theory (HQET) [28] can be constructed in which
 325 the heavy quark interacts only via ‘soft’ hadronic processes and an OPE in $1/m_b$ is
 326 valid. The ‘low q^2 ’ is instead the region where the light spectator quark is energetic
 327 and cannot be neglected. Furthermore, the light quark interacts not only via ‘soft’
 328 hadronic processes, as in HQET, but also via the so-called ‘collinear’ hadronic pro-
 329 cesses. The boundary of this region can be set at $\sim 7 \text{ GeV}^2/c^4$ which corresponds
 330 to the threshold for $c\bar{c}$ production, $(2m_c)^2$. In this region the hadronic interactions
 331 are handled by expanding in terms of the energy of the emitted energetic hadron,
 332 $1/E_h$, forming the so-called Soft-Collinear Effective Theory (SCET) [29]. In both
 333 regions decay rates can be predicted using the different methods and the biggest un-
 334 certainties come from the limited knowledge of hadronic transition matrix elements.
 335 The intermediate region is characterised by the presence of charmonium resonances,
 336 produced through tree level $b \rightarrow c\bar{c}s$ transitions and no precise theoretical calculation
 337 is available [30].

338 As can be seen in Fig. 1.8 the very low q^2 is characterised by a peak due to the
 339 virtual photon contribution, associated with C_7 . In the region $1 - 6 \text{ GeV}^2/c^4$ the
 340 interference between C_7 and C_9 becomes large, yielding sensitivity to NP in C_9 .
 341 The $7 - 15 \text{ GeV}^2/c^4$ interval is dominated by the charmonium resonances, J/ψ and
 342 $\psi(2S)$. Although these decays can be experimentally vetoed in principle charmonia

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

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

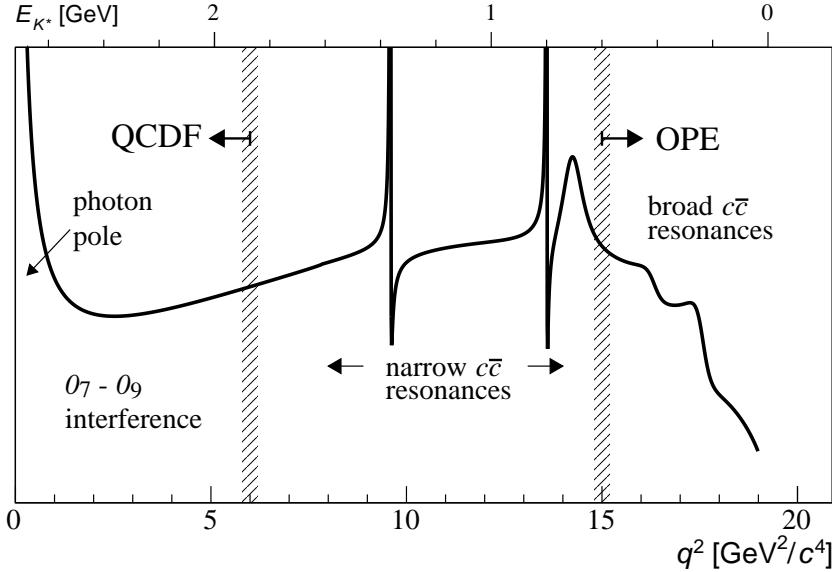


Figure 1.8: A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ process characterised by the photon pole at very low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2 .

³⁴³ affect the entire q^2 space. Finally, at high q^2 broad charmonium resonances can
³⁴⁴ contribute, like those observed by LHCb in $B^+ \rightarrow K^+\mu^+\mu^-$ decays [31].

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

³⁴⁶ Rare decays and especially semileptonic $b \rightarrow s\ell^+\ell^-$ processes offer a number of ob-
³⁴⁷ servables which can be used to benchmark BSM models. The most direct effects
³⁴⁸ appear in decay rates that can be enhanced by new physics but the precision on
³⁴⁹ these measurements is often limited by the uncertainty on form factor calculations.
³⁵⁰ Therefore, it is important to also look for different observables. One important
³⁵¹ class of observables are angular quantities that can often carry complementary in-
³⁵² formation with respect to branching ratio measurements. The most basic of these
³⁵³ observable are forward-backward asymmetries that characterise the angular distri-
³⁵⁴ bution of final particles. For the $B^0 \rightarrow K^*\mu^+\mu^-$ decay combinations of observables
³⁵⁵ have been proposed that are independent of form factor uncertainties at leading
³⁵⁶ order order [22].

357 An other way to build safe observables is to construct ratios between similar decays,
 358 in which uncertainties due to the hadronization process cancel out. These observ-
 359 ables include the R_H ratios, between B^0 decays into electrons and muons, that are
 360 described in detail in Ch. 5. It is also interesting to compare decays which go though
 361 the same fundamental process but where the spectator quark has a different flavour.
 362 This is the case of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays, which are both
 363 $b \rightarrow s$ transitions where the spectator quark is an u quark in the first case and a d
 364 quark in the second. The ratio of the branching fractions of these decays is called
 365 isospin asymmetry.

366 1.6 Experimental status

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

372 1.6.1 Dimuon decays of b hadrons

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

these decay rates are [32]:

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

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

The uncertainties on these values mainly come from the knowledge of the decay constants and CKM-elements. BSM models can produce significant enhancement to these decay rates. Furthermore, the measurement of their ratio is a stringent test of the MFV hypothesis. A combination of the LHCb and CMS results measured the values [33]:

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

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

373 Both decays were previously unobserved, while now the B_s^0 decay is observed with
374 a significance of 6σ and evidence for the B^0 decay is found at 3σ significance level.
375 The measured branching fractions are compatible with SM predictions within 2σ and
376 put strong constraints to the available parameter-space for BSM theories. Figure 1.9
377 shows the fit the dimuon invariant mass of B meson candidates where the peaks of
378 the two decays are visible.

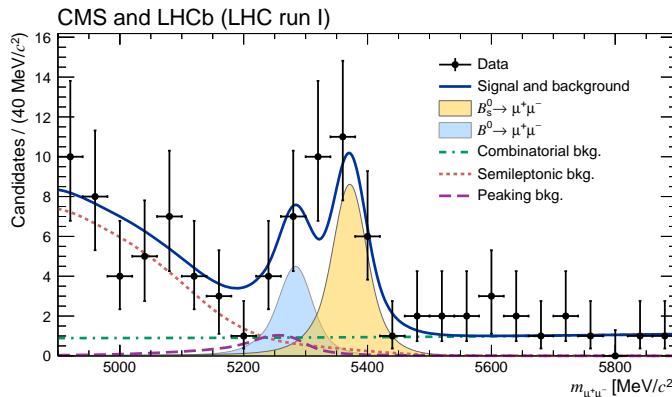


Figure 1.9: Dimuon invariant mass of B candidates showing peaks corresponding $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays [33].

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

³⁸⁰ At the LHC energies is now possible to collect large data samples of semileptonic
³⁸¹ decays, especially those with muons in the final state. Many branching fractions
³⁸² of semileptonic B meson decays were recently measured at the LHCb experiment,
³⁸³ including $B \rightarrow K\mu^+\mu^-$, $B \rightarrow K^{*0}\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ [34, 35, 36]. Baryon
³⁸⁴ decays were also studied at LHCb: including the rare $\Lambda_b \rightarrow \Lambda\mu^+\mu^-$ decay [1], whose
³⁸⁵ analysis is described in this thesis. Unlike for pure leptonic decays, SM predictions
³⁸⁶ for semileptonic decays are affected by the knowledge of hadronic form factors, which
³⁸⁷ results in relatively large uncertainties, $\mathcal{O}(30\%)$. As a result measurements are now
³⁸⁸ typically more precise than predictions.

³⁸⁹ Among the measurements of angular observables that can be affected by new physics,
³⁹⁰ particular interest was risen by the measurement of a set of observables in $B \rightarrow$
³⁹¹ $K^{*0}\mu^+\mu^-$ decays, free from form factors uncertainties at leading order [37]. Most of
³⁹² the measurements are found to be in agreement with SM predictions with the excep-
³⁹³ tion of the P'_5 observable, shown in Fig. 1.10, which presents a local 3.7σ deviation.
³⁹⁴ Attempts to build a consistent picture point to a new physics contribution to the
³⁹⁵ Wilson Coefficient C_9 [38]. An angular analysis of $B^+ \rightarrow K^+\mu^+\mu^-$ decays was also
³⁹⁶ performed, where observables are found to be compatible with SM predictions [39].

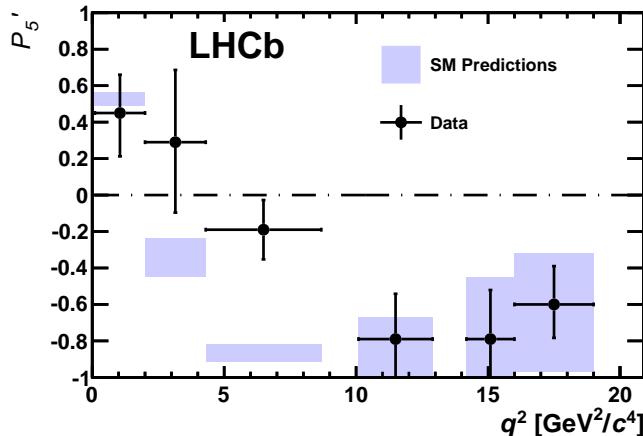


Figure 1.10: Measurement of the observable as a function of q^2 , showing a tension with SM predictions in the 2–6 GeV^2/c^4 region.

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

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

398 Other observables for which the sensitivity to form factors effects is reduced are the
 399 CP asymmetry between B and \bar{B} decays, \mathcal{A}_{CP} , and the isospin asymmetry between
 400 B^0 ad B^+ decays, \mathcal{A}_{CP} . Due to the small size of the corresponding CKM elements
 401 CP asymmetries of $B^0 \rightarrow K^{(*)} \mu^+ \mu^-$ decays are tiny in the SM, $O(10^{-3})$. In BSM
 402 models new sources of CP violation can arise and therefore \mathcal{A}_{CP} measurements are
 403 a powerful null test of the SM. The isospin asymmetry is not zero in the SM due
 404 to isospin breaking effects in the form factors. This is expected to be $\sim 1\%$ at low
 405 q^2 and increase to $\sim 10\%$ as q^2 tends to zero. The LHCb experiment, using the
 406 full dataset collected in Run I, corresponding to an integrated luminosity of 3 fb^{-1} ,
 407 measured both these asymmetries to be consistent with zero [34, 40], as reported in
 408 Tab. 1.4.

409 Recently, progress was made measuring also electron channels. The branching frac-
 410 tion of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay was measured to be $(3.1 \pm 1.3) \times 10^{-7}$ in the
 411 dilepton mass interval 30–1000 MeV/ c^2 [41]. Furthermore, for the first time angu-
 412 lar observables were measured for this decay and found to be consistent with SM
 413 predictions [42].

414 1.6.3 Lepton Flavour Violation searches

415 Several Lepton Flavour Violation (LFV) searches are linked to rare decays as they
 416 involve small branching ratios in the SM that can be enhanced by new physics. They
 417 are therefore a natural place to look for new physics. Lepton flavour conservation is
 418 well experimentally established measuring the branching ratios of decays of muons
 419 into electrons and no neutrinos but has no strong theoretical explanation in the

⁴²⁰ context of the SM. In fact it is already observed that flavour is not conserved in
⁴²¹ neutrino oscillations.

⁴²² The best-studied decays violating lepton flavour are rare muon decays including
⁴²³ $\mu^+ \rightarrow e^+\gamma$ and $\mu^+ \rightarrow e^+e^-e^+$. Since muons can be abundantly produced and
⁴²⁴ the final states are simple, these decays provide the best constraints to LFV. The
⁴²⁵ present best upper limits are 1.2×10^{-11} for the radiative decay and 1.0×10^{-12}
⁴²⁶ for $\mu^+ \rightarrow e^+e^-e^+$ obtained respectively by the MEGA [43] and SINDRUM [44] ex-
⁴²⁷ periments. Several LFV searches in the B sector have recently been performed at
⁴²⁸ the LHCb experiment including decays such as $B^0 \rightarrow e\mu$ [45] and τ decays such as
⁴²⁹ $\tau \rightarrow \mu^+\mu^-\mu^-$ [46]. None of these searches has found evidence of new physics so far
⁴³⁰ and therefore they set limits, constraining the parameter space available for BSM
⁴³¹ models. Figure 1.11 shows a summary of the best limits set at different times on
⁴³² LFV searches [47].

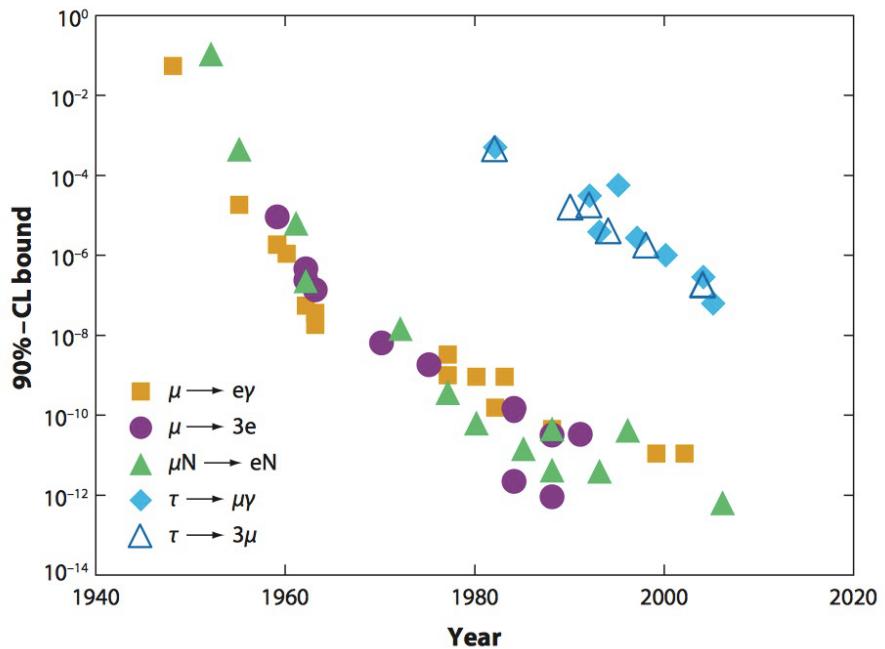


Figure 1.11: Summary of limits set in LFV searches as a function of time [47].

433

CHAPTER 2

434

435

The LHCb detector at the Large Hadron Collider

436

437

2.1 The Large Hadron Collider

438 The Large Hadron Collider (LHC) [48] is a circular particle accelerator with a cir-
439 cumference of 27 km located about 100 m underground at CERN in the surroundings
440 of Geneva, Switzerland. Two proton beams circulate in opposite directions around
441 the ring and cross each other in four points, in which particle detectors are placed.
442 These include two general-purpose detectors, ATLAS and CMS, sitting on opposites
443 sides of the ring and two smaller detectors, ALICE and LHCb that are designed to
444 study specific topics (see Fig. 2.1).

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

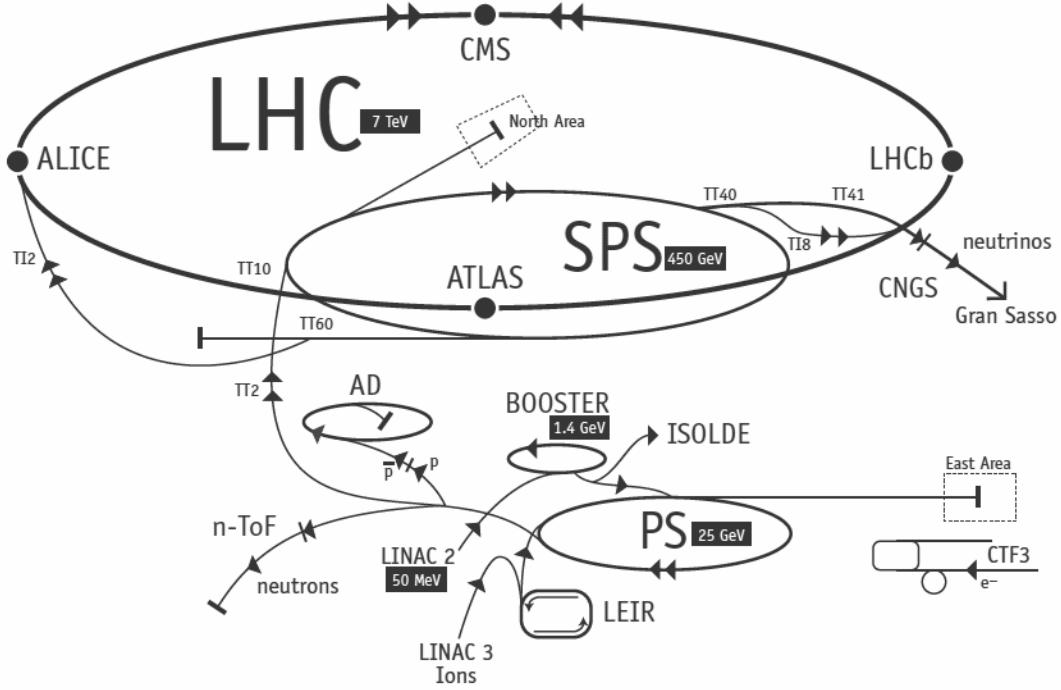


Figure 2.1: Scheme of CERN accelerators.

449 Protons are produced from hydrogen gas and are initially accelerated to the energy
 450 of 50 MeV in a linear accelerator (LINAC). Then they are injected into the Proton
 451 Synchrotron Booster (PSB), where they are boosted to an energy of 1.4 GeV, into
 452 the Proton Synchrotron (PS) to 25 GeV and into the Super Proton Synchrotron
 453 (SPS) to 450 GeV. Finally, protons enter into the LHC storage ring, where they are
 454 accelerated from injection energy to the final one by radio frequency (RF) cavities.
 455 The beams are steered around the ring by 8 T magnetic fields produced by 15 m
 456 long superconducting niobium-titanium dipole magnets and focused by quadrupole
 457 magnets. The LHC magnets use a design in which both proton beam pipes are
 458 contained in the same housing, allowing the same liquid helium to cool the system
 459 down for both. The LHC began colliding proton beams in physics mode in 2009 at
 460 a center of mass energy of $\sqrt{s} = 900$ GeV and from April 2010 to November 2011
 461 accelerated beams at $\sqrt{s} = 7$ TeV (3.5 TeV per proton beam) with a maximum
 462 instantaneous luminosity of $3 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, while in 2012 the energy was increased
 463 to 8 TeV. The LHC maximum design energy is 14 TeV, and its design luminosity is
 464 $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. After a long shut down to upgrade and maintain the machine, a new

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

468 **2.2 The LHCb detector**

469 The LHCb detector [49] was built to study decays of B and D mesons, mainly
470 looking for CP-violating processes. In 2011, running at a centre of mass energy of
471 7 TeV, the cross section for $b\bar{b}$ production was measured to be $284 \pm 53 \mu b$ [50],
472 while it will be $\sim 500 \mu b$ at the current LHC energy, 13 TeV. At these high energies,
473 proton-proton interactions produce highly boosted virtual gluons which produce $b\bar{b}$
474 pairs at small angles, close to the beam pipe. For this reason the LHCb detector is
475 designed to have a very forward angular coverage. The detector is fully instrumented
476 from 10 mrad to 300 mrad, corresponding to an interval $2 < \eta < 5$, where η is the
477 “pseudorapidity”, a quantity used in particle physics defined as:

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

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

479 At LHCb’s collision point the luminosity can be adjusted by displacing the beams
480 from head on collisions while keeping the same crossing angle allowing the experi-
481 ment to keep an approximately constant instantaneous luminosity. This means that
482 the average number of interactions per bunch crossing can be limited, which is im-
483 prtant because the detector efficiency, especially in detecting secondary vertices,
484 decreases for events with an high number of primary vertices (PV). Reducing the
485 particle occupancy through the detector also keeps radiation damage to a minimum.
486 Since the LHC started colliding protons in November 2009 until the end of 2011,

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. The origin, $(x, y, z) = (0, 0, 0)$, corresponds to the center of the interaction area.

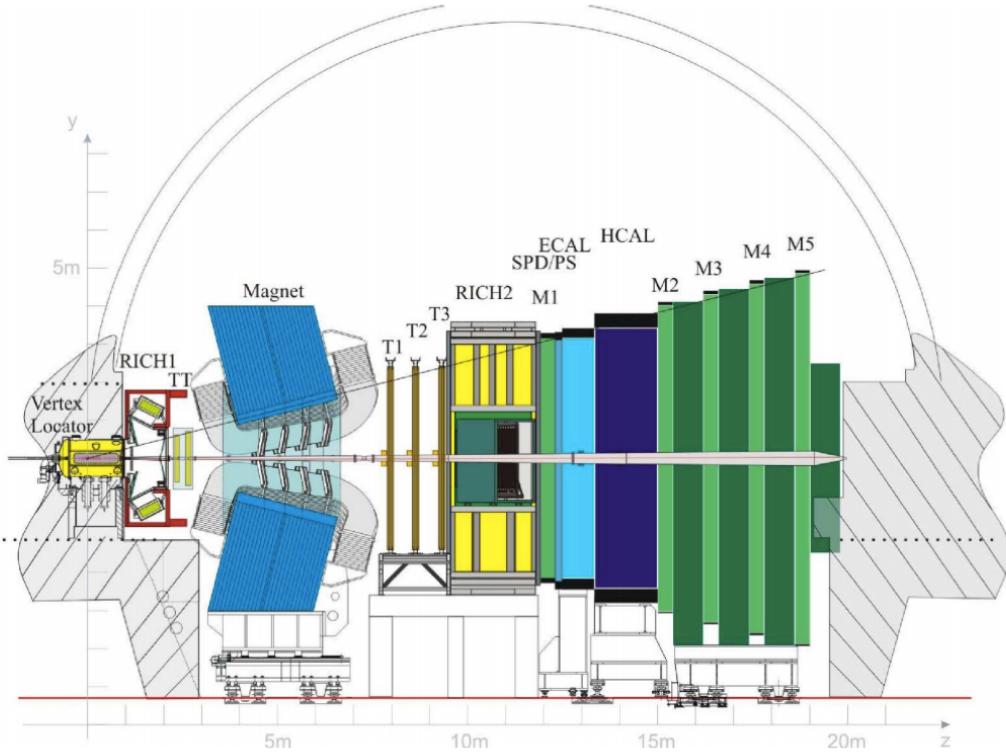


Figure 2.2: A side view of the LHCb detector [49].

487 the instantaneous luminosity was at an average of $3 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, corresponding
 488 to an average number of 1.5 PVs per bunch crossing. At the end of 2011 LHCb had
 489 collected an integrated luminosity of 1 fb^{-1} ; in 2012 the luminosity was increased
 490 and 2 fb^{-1} more were collected.

491 Other B physics experiments, like BaBar at the Stanford Linear Accelerator (SLAC),
 492 Belle at KEK at J-PARC (Japan) and the Tevatron experiments at Fermilab have
 493 made accurate measurements in heavy flavour physics. All of these results have
 494 so far been consistent with the Standard Model predictions. However, some of the
 495 deviations from the Standard Model are expected to be very small. Therefore LHCb
 496 was designed to make the most precise measurements in heavy flavour physics to
 497 test the consistency of the Standard Model and look for new physics.

498 The LHCb detector includes a high-precision tracking system consisting of a silicon-
 499 strip vertex detector surrounding the pp interaction region, and a larger silicon-strip
 500 and drift tubes detectors located on both sides of a dipole magnet with a bend-

ing power of about 4 Tm. Charged hadrons are identified using two Ring-Imaging Cherenkov detectors (RICH) [51]. Photon, electron and hadron candidates are identified by a calorimeter system and muons by a system composed of alternating layers of iron and multi-wire proportional chambers [52]. A schematic view of the detector is shown in Fig. 2.2 and more details on each sub-detector are given in the following sections.

2.3 The magnet

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

2.4 Tracking system

B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means they travel about 1 cm before decaying forming a displaced vertex. To study specific decays, it is therefore important to be able to separate the particles produced at the primary pp vertex and at the B decay secondary vertex (SV). The tracking

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

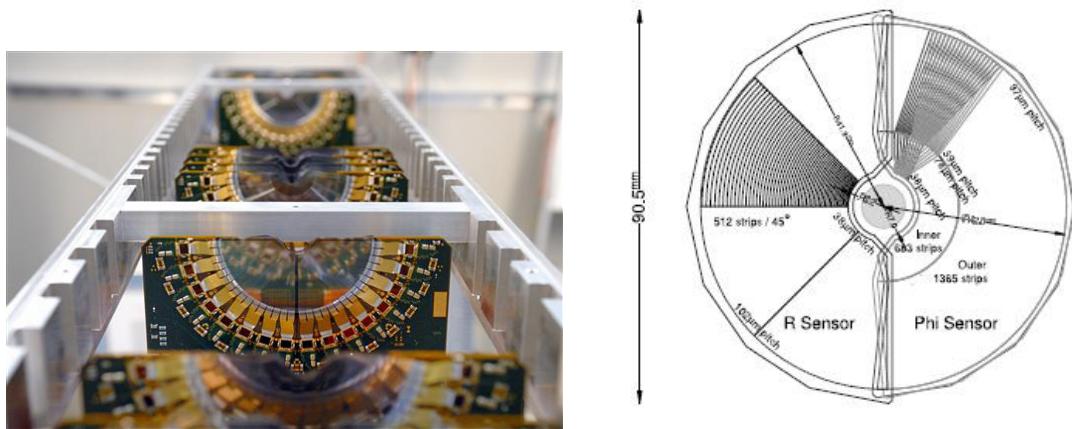


Figure 2.3: On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [49].

531

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

545 they overlap. There are two modules which cover the backward direction and are
546 used as a veto for multiple interactions, this is called the pileup system.

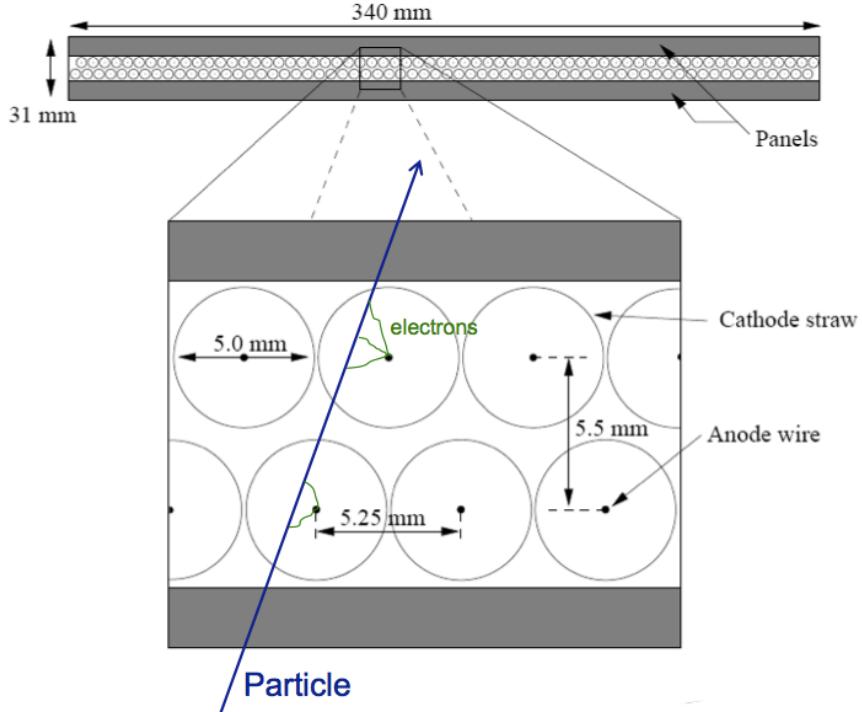


Figure 2.4: Sketch of the straw tubes which constitute the Outer Tracker layers [49].

547

548 The IT and TT both use silicon strips and together constitute the Silicon Tracker
549 (ST). Straw tubes are instead used in the OT, of which a sketch is shown in Fig. 2.4.
550 The IT requires an higher inner granularity because of the higher flux of particles
551 close to the beam pipe. In fact it covers only 1.3% of the total area of IT plus
552 OT but it contains about 20% of the tracks. Each ST station has four detection
553 layers: the first and last are vertical, measuring the track position in x , while the
554 second and third layers are rotated by an angle of +5 and -5 degrees, which allows
555 to measure the y coordinate. The TT is placed upstream of the magnet to allow
556 the reconstruction of tracks from low-momentum particles, which are bent out of
557 the downstream acceptance. Overall the tracking system provides a measurement of
558 momentum, p , with a relative uncertainty that varies from 0.4% at 5 GeV/ c to 1.0%
559 at 200 GeV/ c . The impact parameter (IP), namely the minimum distance of a track

to a primary vertex, is measured with a resolution of $(15 + 29/p_T) \mu m$, where p_T is the component of the momentum transverse to the beam, in GeV/c . The z -axis position of a PV reconstructed with 35–40 tracks can be measured with a precision of roughly $50\text{--}60 \mu m$. Accurate IP and vertex displacement measurements allow LHCb to effectively distinguish between B meson decays and background processes. In fact B mesons typically travel $\sim 1 \text{ cm}$ in the detector before decaying into lighter particles, which tend to have high IP as the B decay imparts transverse momentum to them.

2.5 Calorimeters

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

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

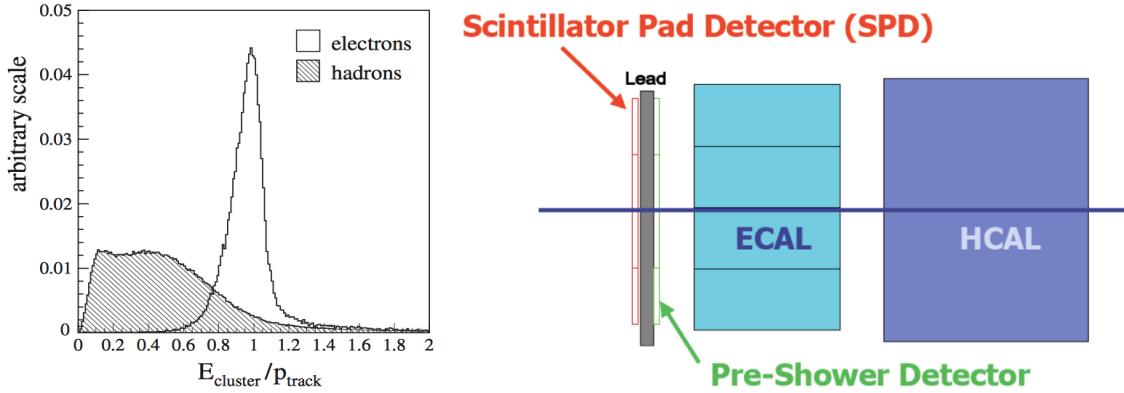


Figure 2.5: The ratio of the energy deposited in the ECAL and the particle momentum, which allows the separation between electrons and hadrons.

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

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 [49]

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

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

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

600 2.5.1 Bremsstrahlung recovery for electrons

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

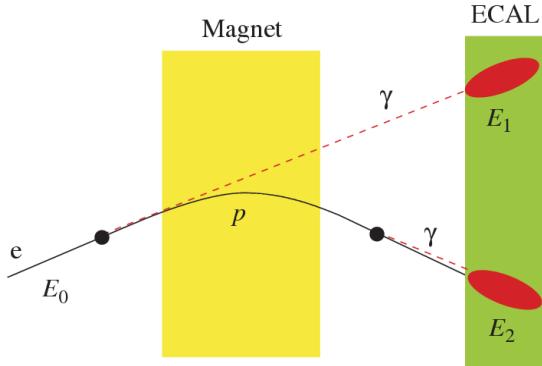


Figure 2.6: Schematic view of the bremsstrahlung recovery.

605
 606 photon will hit in the same calorimeter cells as the electron and the energy will be
 607 automatically recovered. However, if the photon is emitted before the magnet, the
 608 electron will be deflected by the magnetic field whereas the photon will continue
 609 on its initial trajectory, with its energy being deposited in a different part of the
 610 calorimeter. Missing this energy results in a poorer reconstructed invariant mass
 611 resolution, so it is desirable to recover these bremsstrahlung photons, when possible.
 612 A tool for bremsstrahlung recovery is available in the LHCb analysis software. This
 613 tool looks for other clusters in the calorimeter and, reconstructing the trajectory of
 614 the electron, checks if they may be associate with photons emitted by that. The
 615 photon energy is then added to the electron and its momentum is recalculated.
 616 Figure 2.6 displays a schematic view of the process. For more information see
 617 Ref. [53].

618 2.6 RICH

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

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

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

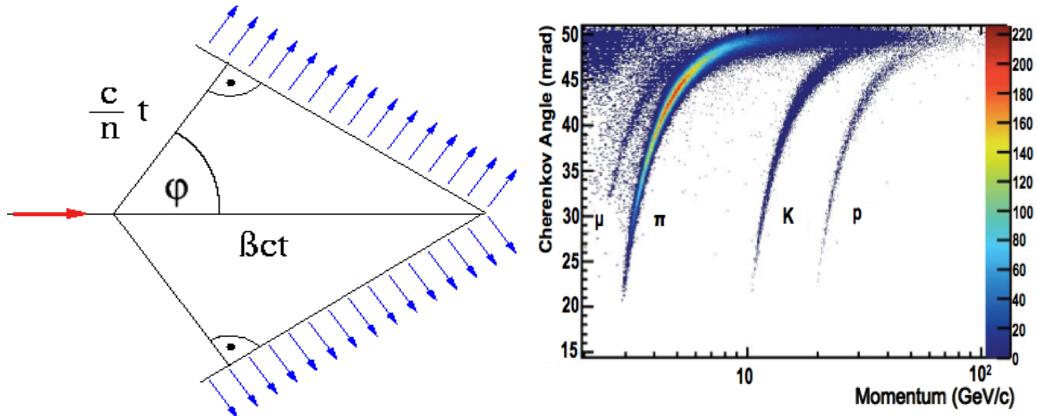


Figure 2.7: On the left a sketch of Cherenkov light emission and on the right the Cherenkov angle versus the particle momentum, where one can see that the study of the Cherenkov angle allows distinguish particles identities.

625

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

634 reflected onto arrays of PMTs. The radius of the ring can be used to measure the
 635 opening angle of the Cherenkov cone because of the known geometry. The photo-
 636 detectors are located outside of the LHCb acceptance in order to reduce the amount
 637 of material that the particles have to traverse. Pattern recognition algorithms are
 638 then used to reconstruct the Cherenkov rings.

639 2.7 The muon system

640 It is essential for many of the key physics analyses in LHCb to be able to identify
 641 muons in decay final states. Muons are the most penetrating particles that can be
 642 detected at LHC experiments, so the muon chambers are the farthest sub-detectors
 643 from the interaction point. The muon system is formed by five stations (M1 -
 644 M5), the first one being located before the calorimeters in order to improve p_T
 645 measurements. The remaining four stations are behind the HCAL and are separated
 646 from each other and interleaved with 80 cm thick iron blocks, which absorb hadrons,
 647 electrons and photons to ensure that only muons reach the final muon station. A
 648 scheme of the muon system is shown in Fig. 2.8. Only muons with a minimum
 649 momentum of $10 \text{ GeV}/c$ traverse all of the five stations and, for positive identification
 650 of a muon, the trigger requires a signal in each of them. Each station has a detection
 651 efficiency of at least 95% and the detectors also provide position measurements.
 652 Since there is a larger particle flux towards the beam pipe, the stations are divided
 653 into four concentric rectangular regions (R1-R4), their size increasing according to
 654 the ratio $1 : 2 : 4 : 8$. This results in a similar channel occupancy over the four
 655 regions. All of the muon stations use Multi Wire Proportional Chambers (MWPC)
 656 except for the inner region of M1, where the particle flux is too high. In this region
 657 triple-GEM (Gas Electron Multiplier) detectors are used because of their better
 658 ageing properties as they have to withstand a rate up to 500 kHz cm^{-2} of charged
 659 particles. In these detectors particles traversing through the drift gap between the
 660 cathode and the first GEM foil produce ionisation electrons, which are then attracted
 661 by electric fields through all of the GEM foils and multiply. They then drift into the

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

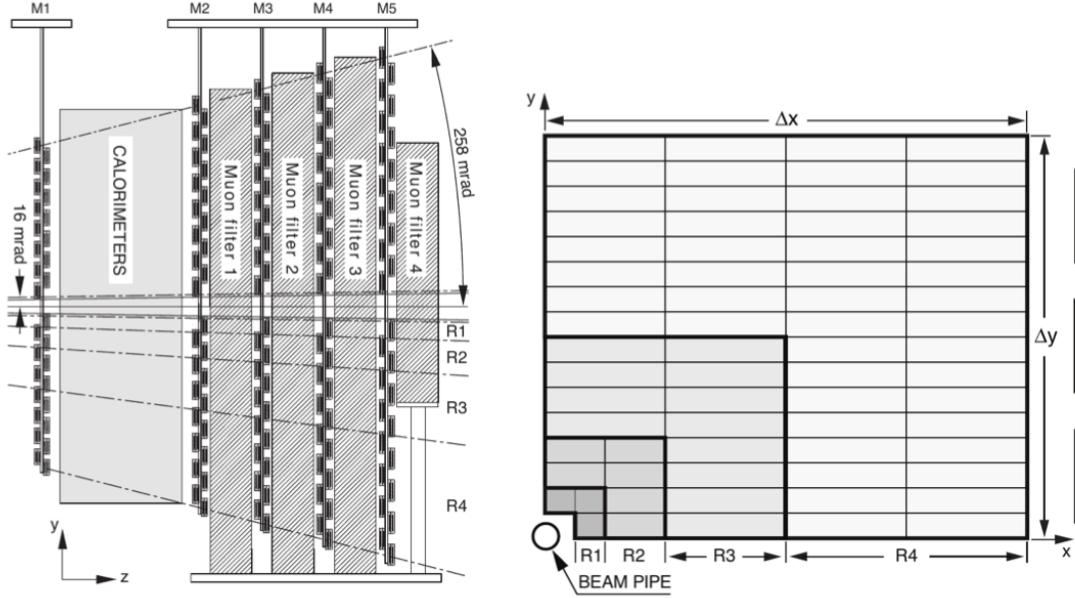


Figure 2.8: The LHCb muon system [49].

663

664 2.8 Particle identification

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

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

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

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

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

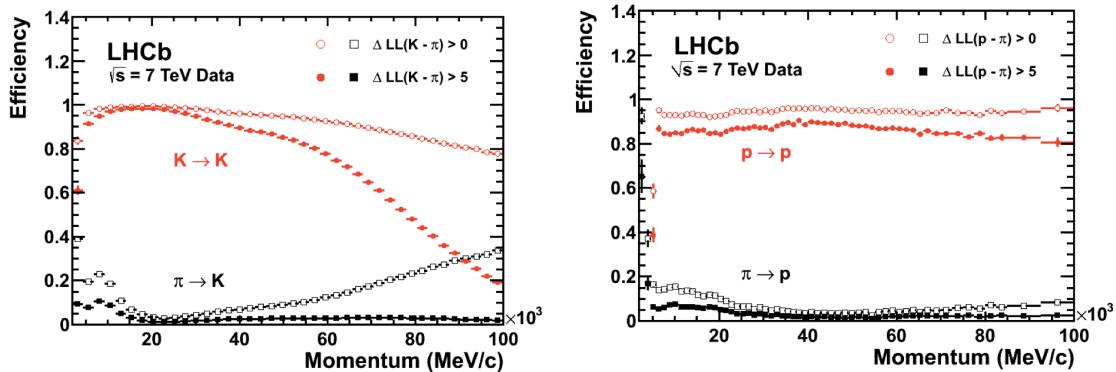


Figure 2.9: Particle Identification performances for kaons (left) and protons (right) as a function of the measured momentum of the particles.

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

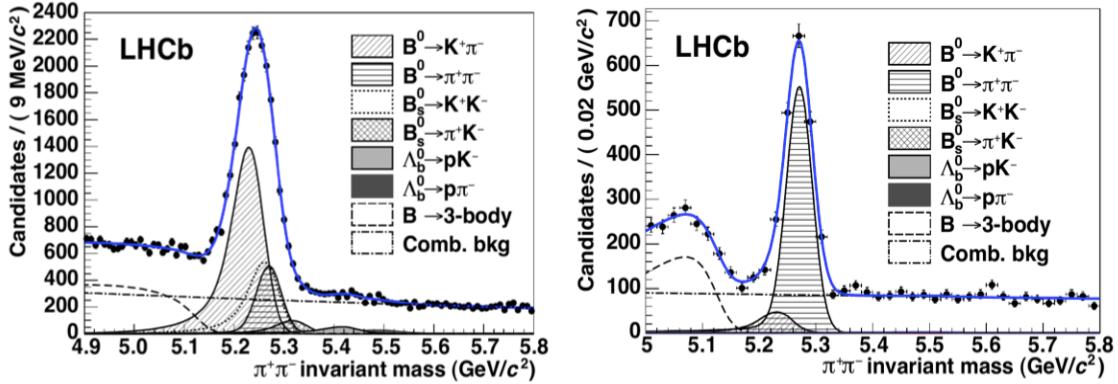


Figure 2.10: Invariant mass peak of the $B^0 \rightarrow \pi^+\pi^-$ decay before (left) and after (right) the application of PID requirements.

694 2.8.1 PID calibration

695 In order to be able to calculate detection efficiencies a data-driven method was
 696 developed. The calibration software is referred to as `PIDCalib` package [54]. This
 697 tool uses decays where final particles can be identified thanks to their kinematic
 698 properties. For example the $K_s^0 \rightarrow \pi^+\pi^-$ decay has a clear signature with a displaced
 699 vertex and can be easily singled out from other decays and used to test pion ID
 700 efficiency. The narrow peaks of the $J/\psi \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow e^+e^-$ decays allow to
 701 calibrate muon and electron efficiencies. A “tag-and-probe” method is used in this
 702 case, where only one of the two leptonic tracks is reconstructed requiring the correct
 703 identity and the other one is used to probe the PID efficiency. Finally, $\phi \rightarrow KK$
 704 samples and $D^{*+} \rightarrow D(\rightarrow K^-\pi^+)\pi^+$ decays, where the D^{*+} is used to tag the decay,
 705 are used to test the kaon efficiency. In all cases the residual background is subtracted
 706 using the $s\mathcal{P}$ lot technique [55].

707 2.9 Trigger and software

708 The LHCb trigger system [56] consists of a hardware stage, L0, based on information
 709 from the calorimeters and muon system, followed by a software stage, the High-
 710 Level Trigger (HLT), which applies a full reconstruction of the events. To increase

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

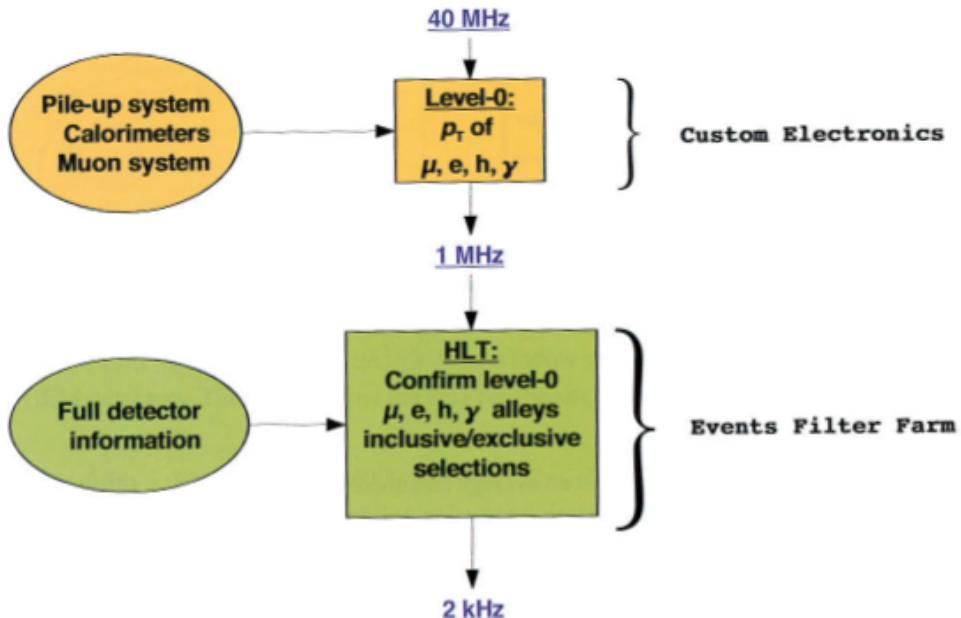


Figure 2.11: Scheme of the LHCb trigger system [49].

719

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

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

736 LHCb also developed an extended simulation software in order to reconstruct ef-
737 ficiencies and signal shapes. In the simulation, pp collisions are generated using
738 PYTHIA8 [57, 58] with a specific LHCb configuration [59]. Decays of hadronic par-
739 ticles are described by EVTGEN [60], and final state radiation is generated using
740 PHOTOS [61]. Finally, the interaction of the generated particles with the detec-
741 tor and its response are implemented using the GEANT4 toolkit [62] as described
742 in Ref. [63]. For this analysis in this thesis, the ROOT framework [64] is used to
743 analyse data and the RooFit package to perform maximum likelihood fits. A multi-
744 variate analysis is also performed based on the NeuroBayes package [65, 66], which
745 provides a framework for neural network training.

746 2.10 Kinematical fits

747 The resolution of key variables, such as the measured invariant mass of decaying
748 particles, can be improved by imposing constraints to the measured quantities to
749 remove redundant degrees of freedom. The four-momentum conservation can be
750 ensured at each vertex and the origin and decay vertices of a particle are related via
751 the momentum of the particle. Furthermore, additional constraints can be imposed
752 due to a particular decay hypothesis such as the known invariant masses of final and
753 intermediate particles. In order to do this the `DecayTreeFitter` tool was developed
754 by the BaBar experiment and later used by LHCb [67]. The algorithm takes a

complete decay chain and parameterises it in terms of vertex positions, decay lengths and momentum parameters. Then these parameters are fit simultaneously, taking into account the relevant constraints, including the information from photons. To perform the fit efficiently a Kalman filter is used. Figure 2.12 illustrates the effect of the application of the kinematical fit on the 4-body invariant mass of the final daughters of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay. The resolution in this case improves by over a factor of 2. Furthermore, the χ^2 from the kinematic fit can be used to quantify the likelihood of a specific decay structure, which helps to separate candidates where random particles from the event have been added to the decay tree or where one or more particles is not reconstructed or mis-identified.

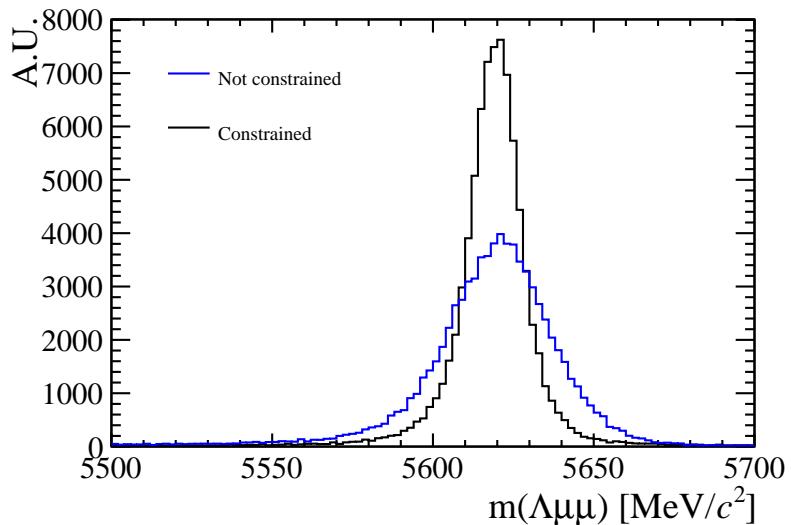


Figure 2.12: Invariant mass of the final daughters of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays calculated with and without constraints using the `DecayTreeFitter` tool.

764

2.11 Validation of hadronic processes in the simulation

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

770 The LHCb simulation software propagates particles through the detector using the
771 GEANT4 toolkit [49]. This offers a variety of models for physics processes over a
772 wide range of energies for both electromagnetic and strong interactions. Given a
773 combination of projectile, target and energy there can be several models applicable
774 with different reliability and computational costs. GEANT4 provides a number of
775 pre-packaged physics lists each representing complete and consistent sets of models
776 chosen to be appropriate for a given use case. In LHCb mainly two hadronic physics
777 lists are considered:

- 778 • **LHEP** (Low and High Energy Parametrisation): based on a parametrised
779 modelling of all hadronic interactions for all particles. This list combines
780 the High Energy Parametrised model (HEP) and the low energy one (LEP).
781 There is a sharp switch from the low to the high energy model at 25 GeV.
782 The modelling of elastic scattering off a nucleus and of nuclear capture also
783 proceeds via parametrised models.
- 784 • **FTFP_BERT**: includes the following models:
- 785 – Bertini cascade model (BERT) [68], which simulates the intra-nuclear cas-
786 cade, followed by pre-equilibrium and evaporation phases of the residual
787 nucleus, for protons, neutrons, pions and kaons interaction with nuclei
788 at kinetic energies below 9.9 GeV. The Bertini model produces more
789 secondary neutrons and protons than the LEP model, yielding a better
790 agreement with experiment data.
- 791 – FTFP model, which implements high energy inelastic scattering of hadrons
792 by nuclei using the FRITIOF model [69]. The change between the two
793 models happens with a linear shift from BERT to FTF that starts at 4
794 GeV and ends at 5 GeV.

795 Figure 2.13 summarises the composition of the different models.

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

2.11.1 Geometry and interaction probability

The results presented in the following sections are produced using the version v45r0 of the full LHCb framework for simulation, Gauss [63], interfaced to GEANT4 v95r2p1. A simple geometry setup is used in order to be able to calculate in a clean way the interaction cross sections in a specific material. This is constituted by a series of rectangular boxes filled with the most relevant materials for LHCb: Aluminium, Silicon and Beryllium. For each material three boxes are defined with different thicknesses (1mm, 10mm, 50mm). These values are chosen to represent a ballpark estimation of the LHCb interaction length.

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

$$P_{int} = \frac{N_{int}}{N_{tot}}, \quad (2.6)$$

where N_{int} is the number of particles which interacted in the material and N_{tot} is the number of generated particles. As GEANT4 provides an ID for the end process of a

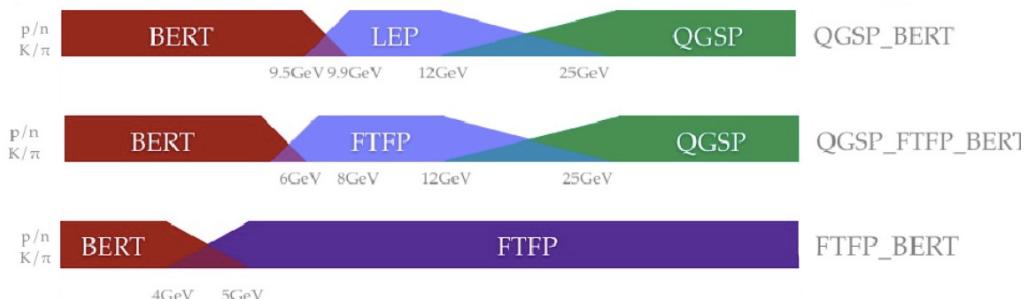


Figure 2.13: Diagram of LHEP, FTFP_BERT and QGSP_BERT models composition as a function of energy.

815 particle (e.g. 121 for inelastic interaction, 111 for elastic, 201 for decay) it is possible
816 to distinguish the inelastic and elastic probabilities of interaction and therefore cross
817 sections.

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

$$\sigma_{int} = \frac{A}{\rho N_A \Delta x} \cdot P_{int}, \quad (2.7)$$

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

822 2.11.2 PDG prediction

In the PDG book [2] cross sections of protons and neutrons are parametrised as:

$$\sigma_{tot}^{ab} = Z^{ab} + B^{ab} \log^2(s/s_M) + Y_1^{ab}(s_M/s)^{\eta_1} - Y_2^{ab}(s_M/s)^{\eta_2}, \quad (2.8)$$

$$\sigma_{tot}^{\bar{a}\bar{b}} = Z^{ab} + B^{ab} \log^2(s/s_M) + Y_1^{ab}(s_M/s)^{\eta_1} + Y_2^{ab}(s_M/s)^{\eta_2}, \quad (2.9)$$

823 where $s_M = (m_a + m_b + M)^2$ and $B^{ab} = \lambda \pi (\frac{\hbar c}{M})^2$. Some of the constants in these
824 equations are universal and valid for any kind of collision: $M = 2.15$, $\eta_1 = 0.462$, η_2
825 = 0.551, $\lambda = 1$ (for p, n and γ) and 1.63 (for d). The other ones are characteristic
826 of each type of collision and are listed in Tab. 2.11.2. In these formulae the particle-
827 antiparticle asymmetry arises from the last term which has opposite sign in the
828 two equations. This term becomes less and less important with increasing energies.
829 Therefore a net asymmetry is found at low energies, while the cross sections tend
830 to a common point at high energy and continue increasing logarithmically.

2.11.3 Validation results

This section reports particle and antiparticle cross sections and their ratios compared where available with PDG predictions and with data from the COMPASS experiment [70]. Figure 2.14 shows the probability of interaction for protons and anti-protons in 1mm of Aluminium using the FTFP_BERT and LHEP models compared with COMPASS data and Fig. 2.15 shows the ratios of $\sigma_{\bar{p}}^{tot}/\sigma_p^{tot}$ together with the PDG prediction. A difference of 40% is found between the two considered models for 1 GeV incoming anti-protons. This difference becomes negligible at higher energies. The discrepancies between the two physics lists for kaons and pions are of a few percents (2-3%) and usually constant with the energy. From the comparison with data and PDG predictions it can be qualitatively concluded that the FTFP_BERT model gives a better description of hadronic interactions at low energies, while both models give good results at high energy, above ~ 10 GeV.

The tool developed for this studies is not limited to cross-sections but can also give information on other simulated quantities: as an example Fig. 2.11.3 shows a comparison between the types of particles generated in inelastic collisions of protons and anti-protons into Aluminium using different models. Different physics lists can give very different results. For example the LHEP model does not produce photons in inelastic collisions. However, it is difficult to use these quantities for validation as there is no data available for comparison.

852

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

Table 2.1: Values for the constants Z^{ab} , Y_1^{ab} and Y_2^{ab} [2], which parametrise hadronic cross sections.

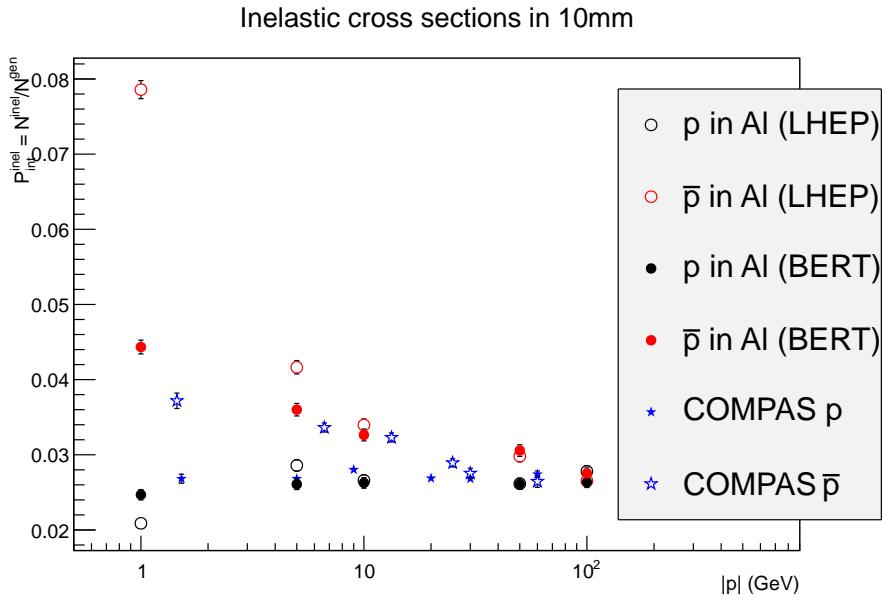


Figure 2.14: Probability of interaction for protons and anti-protons in Aluminium as a function of the projectile momentum. Two physics lists are used to generate events that can be compared with data from the COMPASS experiment.

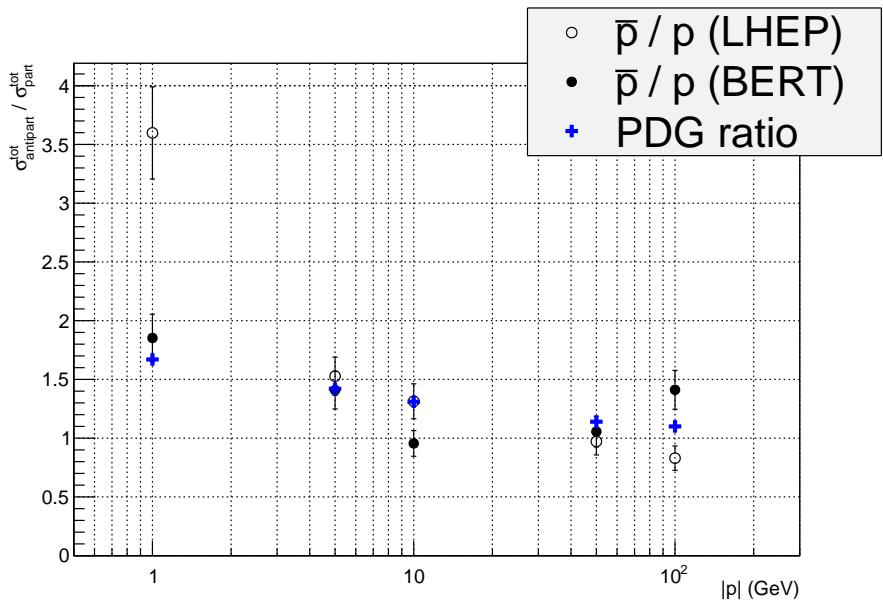


Figure 2.15: Ratio of antiproton over proton total interaction cross section as a function of energy compared with PDG predictions.

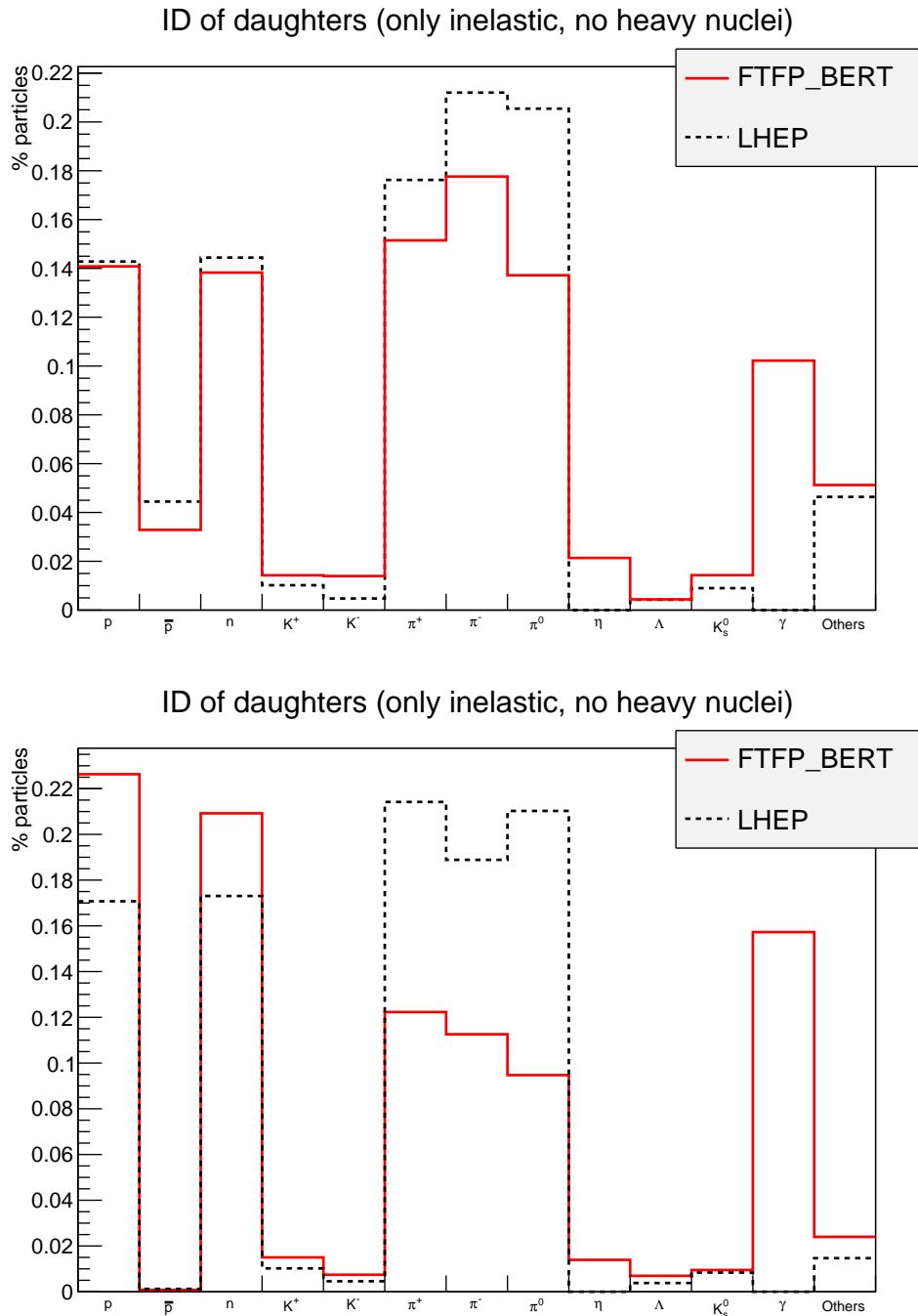


Figure 2.16: Composition of secondary particles produced in 100 GeV protons (top) and anti-protons (bottom) collisions in 1 mm Aluminium.

853 2.12 Material budget studies

854 It is important for many analysis to be aware of the material budget of the detector,
 855 for example to estimate the amount of multiple scattering. In GEANT4 particles
 856 are propagated in steps through the detector and for each step the framework anal-
 857 yses the geometry to understand in what material the particle is and modifies its
 858 trajectory accordingly. A tool was developed where neutrinos are used as probes to
 859 scan the detector summing the radiation length seen at each step up to a certain
 860 points. Neutrinos are used as they do not bend in magnetic field and do not interact
 861 with the detector. Thin air planes are inserted after each sub-detector. When these
 862 are traversed by the neutrinos the information about the cumulated radiation and
 863 interaction length is saved. In this way it is possible to obtain maps of the detector
 864 as the one shown in Fig. 2.17. Finally, Fig. 2.18 shows the average radiation length
 865 as a function of the distance from the interaction point.

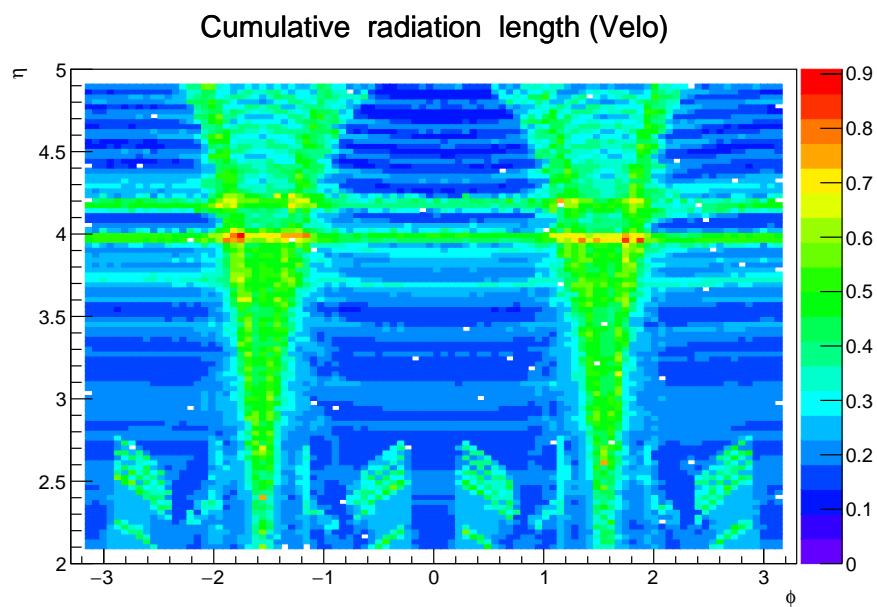


Figure 2.17: Map of cumulative radiation length traversed by a particle starting from the interaction point up to the end of the VeLo.

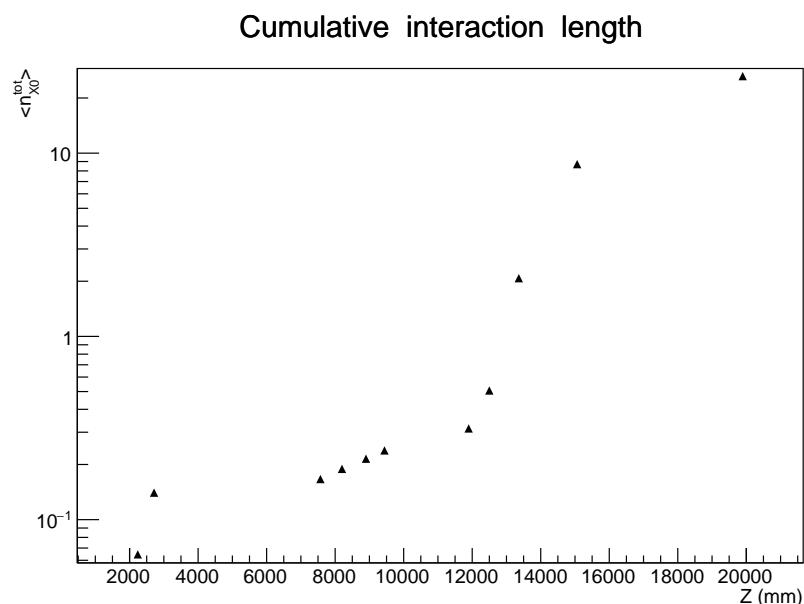


Figure 2.18: Average cumulative radiation length as a function of the horizontal distance from the interaction point. Each considered point corresponds to the end of a sub-detector: VeLo, RICH1, RICH2, tracking stations, ECAL and HCAL and muon detector.

CHAPTER 3

867

868

869

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

870

871 The rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay is a FCNC process governed by the $b \rightarrow s\mu^+\mu^-$ quark
872 level transition. In the SM this decay proceeds only through loop diagrams, elec-
873 troweak penguin and W box (see Fig. 1.5), and therefore it is highly sensitive to new
874 particles entering the loops. Interest in Λ_b^0 baryon decays arises from two important
875 facts. First of all, Λ_b^0 has non-zero initial spin, which allows to learn information
876 about the helicity structure of the underlying Hamiltonian, that cannot be extracted
877 from the meson decays [71, 72]. Secondly, the Λ_b^0 baryon can be considered in first
878 approximation as composed of an heavy quark and a light di-quark, therefore the
879 hadronic physics significantly differs from similar meson decays. This provides the
880 possibility to better understand and test the hadronic physics in the theory, which
881 could yield improved understanding and confidence also for the meson case.

882 With respect to B^0 decays going though the same transitions, such as $B^0 \rightarrow K^{*0}\mu^+\mu^-$,
883 Λ_b^0 decays can provide independent confirmations of the results as they involve the
884 same operators but different hadronic matrix elements. Furthermore, Λ decays

weakly, which results in complementary constraints with respect to B^0 decays. Finally, the narrow width approximation, used in theoretical calculations, is fully applicable in the Λ_b^0 case, which has $\Gamma_{\Lambda_b^0} \sim 2.5 \cdot 10^{-6}$ eV. This is not assured for $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays because the contribution from the non resonant channel $B^0 \rightarrow K\pi\mu^+\mu^-$ is unconstrained.

The theory of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays was considered by a number of authors both in the SM and in different new physics scenarios [73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83]. All authors start from the same effective Hamiltonian already described in Sec. 1.5.1. However, form factors, describing hadronic physics, are not developed as well as for the meson case because there are not as many experimental constraints. This leads to a relatively large spread in predicted branching fractions. For these reasons an interesting quantity to study is the differential branching fraction as function of q^2 . This still suffers from the knowledge of form factors, but, as different approaches to form factors calculations are applicable in different q^2 regions, it allows a more meaningful comparison with theory.

Experimentally, the decay $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ was observed for the first time in 2011 by the CDF collaboration [84], with a signal yield of 24 ± 5 events and was later updated using their full statistics [85]. CDF observed the signal only in the q^2 region above the square of the $\psi(2S)$ mass. Their result on full statistics yields $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-) = [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}$. Recently, the decay was also observed at LHCb [86] with a yield of 78 ± 12 signal events using 1 fb^{-1} of integrated luminosity collected in 2011. The signal was again found only in the high q^2 region, above $m_{\psi(2S)}^2$. The LHCb result for the branching fraction relative to the $J/\psi\Lambda$ decay, used as normalisation channel, is

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

and for absolute branching fraction

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

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

3.1 Analysis strategy and q^2 regions

A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ decays was shown in Fig. 1.8. This is characterised by the presence of the photon pole at low q^2 and the narrow peaks of the J/ψ and $\psi(2S)$ resonances at mid q^2 . For this analysis two regions are defined: the “low q^2 ” region, below the J/ψ resonance ($q^2 < 8 \text{ GeV}^2/c^4$), where the signal is unobserved, and the “high q^2 ” region, above the J/ψ resonance ($q^2 > 11 \text{ GeV}^2/c^4$). The decay $\Lambda_b^0 \rightarrow J/\psi \Lambda$, where J/ψ decays into two muons has the same final state as the signal and is used as a normalisation channel. 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. 3.4.3. The upper bound of this interval is chosen to be sufficiently far from the J/ψ radiative tail at low masses, that could contaminate the rare sample;
- $11 < q^2 < 12.5 \text{ GeV}^2/c^4$ in between two charmonium resonances and
- $q^2 > 15 \text{ GeV}^2/c^4$, above $\psi(2S)$.

In the latter two intervals the selection is optimised to maximise the yield which is particularly important for a stable angular analysis. The above regions are then divided in smaller intervals, as much as the available statistics allows, which results

925 in $\sim 2 \text{ GeV}^2/c^4$ wide bins. The binning used is the following:

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

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

- 927 • 1.1-6.0 GeV^2/c^4 : this interval is theoretically clean since it is far from the photon pole, which dominates at low q^2 washing out the sensitivity to new physics contributions. The lower bound of this interval is chosen to exclude the possible contribution from the ϕ resonance, which appears at 1 GeV^2/c^4 . The upper bound of the interval is chosen to totally exclude a small contribution from the J/ψ resonance that leaks below 8 GeV^2/c^4 .
- 933 • 15.0-20.0 GeV^2/c^4 : this interval is the one that contains most of the statistics and it is used as a natural cross check that the analysis in smaller bins is stable.

936 3.2 Candidate types

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

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

- 946 • **Downstream candidates:** built from tracks without hits in the VeLo, “down-
 947 stream tracks”, also denoted as “DD”.

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

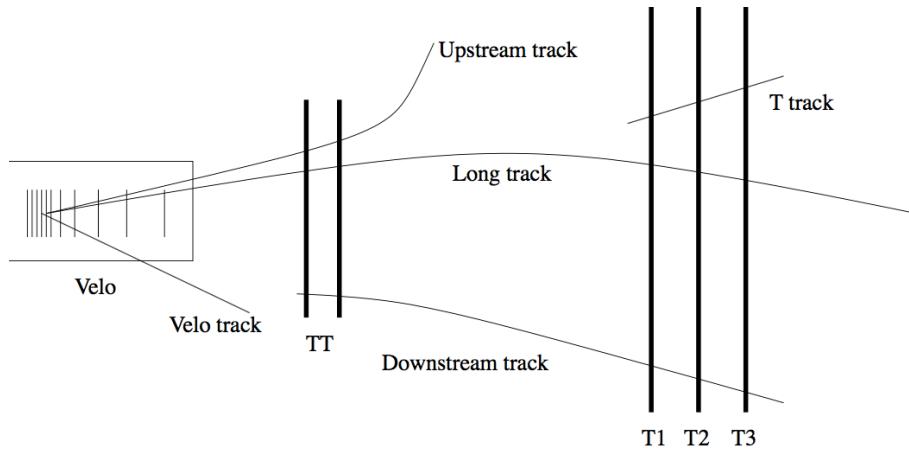


Figure 3.1: Representation of possible track types in LHCb. Candidates built from “long” and “downstream” tracks are used in this analysis.

953 3.3 Simulation

954 Samples of simulated events are needed in order to train the multivariate classifier
 955 (see Sec. 3.4.2), calculate the selection efficiency and study possible backgrounds;
 956 in particular for this analysis samples of ~ 2 millions $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and ~ 5 millions
 957 $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulated events are used. Samples of simulated $B^0 \rightarrow J/\psi K_s^0$,
 958 $B^0 \rightarrow K_s^0\mu^+\mu^-$ and $B^+ \rightarrow \mu^+\mu^-K^{*+}$ events are also used to study backgrounds
 959 from these decays. The events are generated using PYTHIA8, hadronic particle are
 960 decayed using EVTGEN and GEANT4 is used to simulate the interaction of final
 961 state particles with the detector. Simulated events are then reconstructed using
 962 the same reconstruction software used for real data. The L0 hardware trigger is

emulated in the simulation, while for the software stage, HLT (see Sec. 2.9), the same code can be used as for data. Events are simulated using both 2011 and 2012 beam and detector conditions in the same amounts in which data is available. While the simulation gives a generally good description of data some discrepancies remain. However it is important that the simulation gives an accurate description of the data, especially for quantitative estimations, e.g the extraction of efficiencies. The next sections describe corrections applied to the simulation in order to have a better description of data. In Appendix B data distributions are compared with simulated ones for variables relevant to this analysis.

3.3.1 Decay Model

Little is known about Λ_b^0 decays structure and therefore the simulation software generates events according to the phase space given by the available kinematic. To include a reasonably realistic q^2 dependence, the simulation is weighted using decay amplitudes based on the predictions in Ref. [87]. Equations in this paper are for the case of unpolarised Λ_b^0 production and for this analysis those are extended to include polarisation. Details about the models used are in Appendix A.1. The value of the Λ_b^0 production polarisation, P_b , used in the calculations is $P_b = 0.06$ as measured by LHCb [88]. Figure 3.2 shows the phase space q^2 distribution and the one obtained re-weighting the events. The latter can be qualitatively compared to the q^2 spectrum of a generic $b \rightarrow s\ell^+\ell^-$ decay shown in Fig. 1.8. For the normalisation mode, the decay model used is described in Appendix A.2, with amplitude magnitudes and production polarisation taken from the measurements in Ref. [88]. Phases are not yet measured and are all set to zero.

3.3.2 Kinematic re-weighting

Small data-simulation differences are found in the kinematic properties of the mother particle, Λ_b^0 , which also affect the final state particles. The simulation is re-weighted

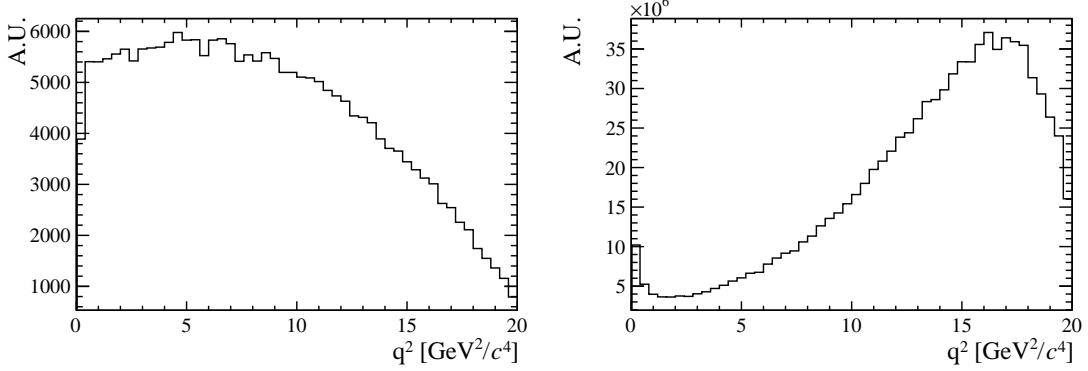


Figure 3.2: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulates events according to the phase space of the decay (left) and reweighted using the decay amplitudes (right).

989 by comparing the momentum and transverse momentum of Λ_b^0 between real and
 990 simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates which passed pre-selection (see Sec. 3.4). To do
 991 this a data sample as clean as possible is obtained selecting a narrow interval around
 992 J/ψ and Λ_b^0 peaks. Then the Λ_b^0 invariant mass is fitted to extract the amount of
 993 background under the peak. The background fraction, $f_b = B/(S + B)$, is then
 994 used to statistically subtract the background from the kinematical distributions as
 995 described by the equation:

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

996 where $S(p, p_T)$ is the distribution of pure signal events, which we want to obtain,
 997 $T(p, p_T)$ is the total distribution of signal plus background, namely the distribution
 998 of all events in the signal interval, $5605 < m(p\pi\mu^+\mu^-) < 5635 \text{ MeV}/c^2$, and $B(p, p_T)$
 999 is the pure background distribution obtained using events from the upper sideband,
 1000 $m(p\pi\mu^+\mu^-) > 5800 \text{ MeV}/c^2$.

1001 After obtaining the signal distributions from data these are compared with $\Lambda_b^0 \rightarrow J/\psi \Lambda$
 1002 simulated events and a weight, $w(p_{\Lambda_b^0}, p_{T\Lambda_b^0})$ is defined by taking the ratio of the two
 1003 dimensional (p, p_T) distributions. The result is shown in Fig. 3.3, while Appendix B
 1004 reports distributions of sideband subtracted data in the signal and sideband regions
 1005 together with weighted and unweighted simulated events. In these plots the $\Lambda_b^0 p$ and
 1006 p_T distributions match by construction but the re-weighting also improves the agree-

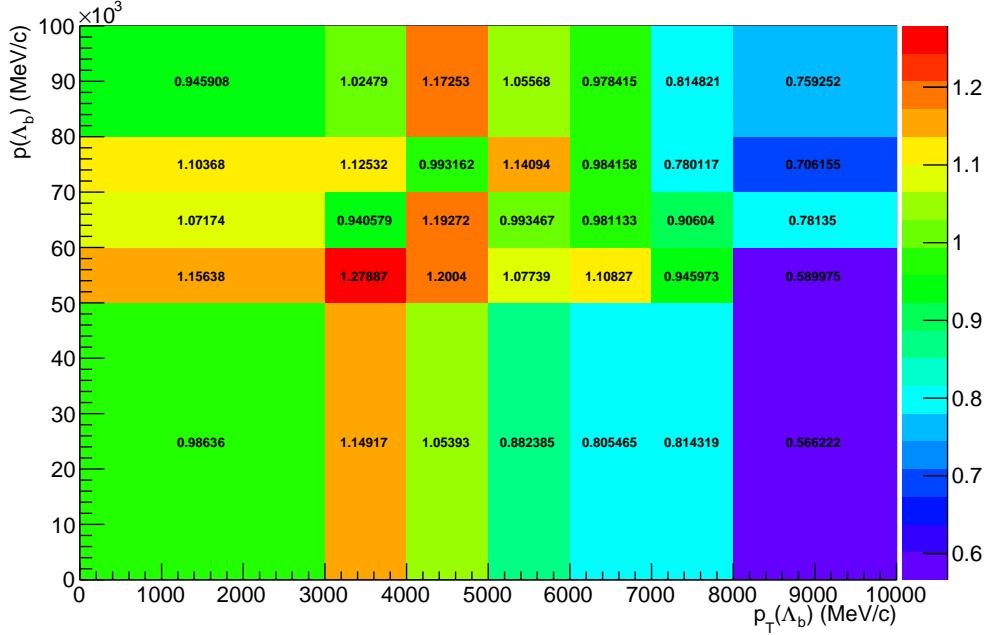


Figure 3.3: Weights used for the kinematical reweighting as a function of the momentum and transverse momentum of Λ_b^0 .

ment between the kinematical distributions of all final particles. Small differences remain due to the finite binning used for the weights calculation. Quality variables, such as the χ^2 of tracks and vertices, show little dependence on the kinematics and are relatively unaffected by the weighting procedure.

3.3.3 Event type

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

₁₀₂₁ **3.4 Selection**

₁₀₂₂ This section described the requirements applied to reconstruct $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and
₁₀₂₃ $\Lambda_b^0 \rightarrow J/\psi\Lambda$ candidates. The selection procedure is divided into two steps: a pre-
₁₀₂₄ selection, where cuts are applied in order to be able to work with manageable
₁₀₂₅ datasets and a multivariate analysis (MVA) which combines information from several
₁₀₂₆ variables. At first good tracks are selected using requirements on basic kinematic
₁₀₂₇ properties, as the p_T of the final particles, and quality requirements, as the track
₁₀₂₈ χ^2 . The selection then aims to first form a dimuon candidate from two oppositely
₁₀₂₉ charged muons. Then, in events containing a dimuon candidate, two oppositely
₁₀₃₀ charged tracks are combined and retained as Λ candidate if they form a good vertex
₁₀₃₁ which is well separated from all primary vertices. Finally the dimuon and Λ candi-
₁₀₃₂ dates are combined to form Λ_b^0 baryons and requirements are set on the properties
₁₀₃₃ of this combination.

₁₀₃₄ **3.4.1 Pre-selection**

₁₀₃₅ The full list of pre-selection cuts is reported in Tab. 3.1. In the table χ_{IP}^2 is defined
₁₀₃₆ as the projected distance from a vertex divided by its uncertainty, for example the
₁₀₃₇ $\chi_{IP}^2(primary) > n$ requirement on Λ_b^0 means that the Λ_b^0 vertex must be at least
₁₀₃₈ \sqrt{n} standard deviations away from the primary vertex. Another quantity, especially
₁₀₃₉ useful to remove combinatorial background, is a pointing variable called DIRA de-
₁₀₄₀ fined as the cosine of the angle between the direction of a particle's momentum
₁₀₄₁ and the flight direction from its mother vertex. Requiring a DIRA close to unity
₁₀₄₂ corresponds to the selection of particles with well-defined origin vertices. Graphical
₁₀₄₃ representation of the χ_{IP}^2 and DIRA variables are shown in Fig. 3.4. The variable
₁₀₄₄ χ_{FD}^2 represents the flight distance of a particle from its origin vertex divided by its
₁₀₄₅ uncertainty. The χ_{trk}^2/ndf and χ_{vtx}^2/ndf quantities are the χ^2 from the fit to tracks
₁₀₄₆ and vertices, which are used to quantify their quality. The `GhostProb` quantity
₁₀₄₇ describes the probability of a track being fake. By construction, cutting at k , re-

1048 moves $(1 - k) \cdot 100\%$ of fake tracks. The `hasRich`, `hasCalo` and `isMuon` variables are
1049 binary indicators that the information from the RICH/calorimeter/muon detector
1050 is available for the track. Loose PID requirements on the proton are also applied
1051 in pre-selection. Details about the quantification of the PID quality are given in
1052 Sec. 2.8. A large mass window is kept around the Λ_b^0 peak in order to be able to
1053 fit the sideband, to train the multivariate analysis and to better constrain back-
1054 grounds. Rare candidates are selected by the q^2 region requirements described in
1055 Sec. 3.1, while resonant candidates are further constrained to have dimuon invariant
1056 mass in a $100 \text{ MeV}/c^2$ interval around the known J/ψ mass [2].

1057 3.4.2 Neural Networks

1058 The final selection is performed using a neural network (NN) classifier based on the
1059 NeuroBayes package [65, 66]. The input to the neural network consists of 14 variables
1060 carrying information about the kinematics of the decay, the quality of tracks and
1061 vertices and the PID of the muons. The list of inputs is reported in Tab. 3.2, together
1062 with information about the importance of each input. Variables related to Λ and its
1063 daughters are considered as different inputs depending on the candidate type (long
1064 or downstream). This effectively corresponds to making a separate training for the
1065 two categories. The graphical representation of the correlation matrix is shown in
1066 Fig. 3.5, where the variable with ID= 1 is the neural network output and the IDs of

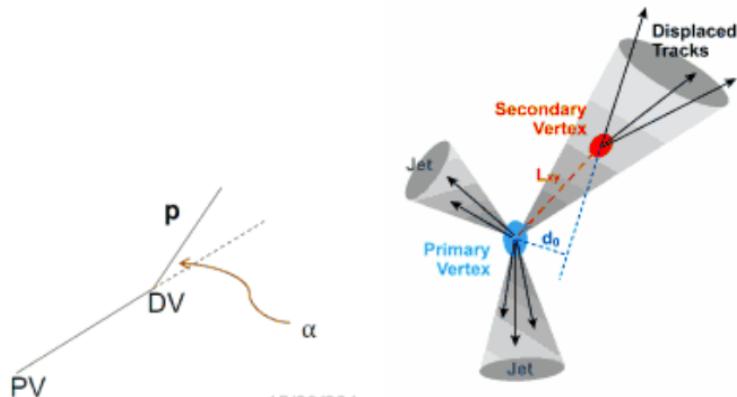


Figure 3.4: Graphical representation of the DIRA (left) and χ_{IP}^2 (right) variables.

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

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

1067 the other variables are listed in Tab. 3.2.

1068 The NN is trained using representative samples of signal and background. A sample
 1069 of simulated $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ events is used as a proxy for the signal, while for the
 1070 background a representative sample is given by candidates in the upper $m(p\pi\mu\mu)$
 1071 invariant mass sideband. Only the upper sideband, $m(p\pi\mu\mu) > 6 \text{ GeV}/c^2$, is used
 1072 since it contains only combinatorial background, while the lower sideband may con-
 1073 tain partially reconstructed and misreconstructed candidates. In the q^2 spectrum
 1074 of background samples the J/ψ and $\psi(2S)$ peaks are still present indicating that
 1075 charmonium resonances are often combined with other random tracks. These can-
 1076 didates do not give a good description of purely combinatorial background and, in
 1077 order to avoid biases, they are removed from the training sample by rejecting events
 1078 in a $100 \text{ MeV}/c^2$ interval around the nominal J/ψ and $\psi(2S)$ masses [2]. A total
 1079 of 30000 total events is used for the training from each sample. This corresponds

1080 $\sim 50\%$ of the available sideband data sample and $\sim 20\%$ of the simulated sample.
1081 The full simulated sample is not used as the same sample will also be used to study
1082 efficiencies. For reproducibility the events are sampled uniformly.

1083 The single most important variable used for downstream candidates is the transverse
1084 momentum of Λ , which allows to reject random combination of tracks as these have
1085 preferentially low p_T . For long candidates instead the best variable is the χ^2 from a
1086 kinematic fit that constrains the decay products of the Λ_b^0 , the Λ and the dimuon,
1087 to originate from their respective vertices performed using the `DecayTreeFitter`
1088 tool (see Sec. 2.10). Other variables that contribute significantly are the χ_{IP}^2 of Λ_b^0 ,
1089 Λ and muons, the separation between the Λ_b^0 and Λ vertices and, finally, the muon
1090 PID.

1091 Figure 3.6 shows distributions of neural network output for the signal and back-
1092 ground samples and purity, $P = N(\text{signal})/N(\text{background})$, as a function of the
1093 neural network output. The distributions from test samples are also overlaid in
1094 order to check for overtraining. The distributions follow the same shape but with
1095 different fluctuations indicating no significant overtraining. In general it can be con-
1096 cluded that the neural network is able to separate signal from background and the
1097 training converged properly.

1098 It can happen that too much information is given to the classifier, which becomes
1099 able to calculate the invariant mass of the candidates from the input variables.
1100 This can generate fake peaks and it is therefore important to check for correlations
1101 between the 4-body invariant mass and the NN output. Figure 3.7 reports the
1102 average neural network output as a function of the 4-body $m(p\pi\mu\mu)$ invariant mass
1103 for data and simulation. The distributions are flat indicating that no significant
1104 correlation is present.

Table 3.2: Summary of inputs to the neural network in order of importance. Column “ID” lists the indices used for the correlation matrix (see Fig. 3.5). Column “adds” gives the significance added by a given input when it is added to the list of those ranked above. Finally, column “only this” provides the power of a given input alone and “loss” shows how much information is lost when removing only a given input.

Input	ID	adds	only this	loss
$\Lambda_{DD} p_T$	15	143.11	143.11	29.20
χ_{DTF}^2	2	77.81	134.00	51.10
$\min(\chi_{IP}^2 \mu)$	7	61.31	113.62	29.76
$\chi_{IP}^2 \Lambda_b^0$	5	52.94	113.23	40.98
$\chi_{IP}^2 \pi_{LL}$	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
$\chi_{IP}^2 \Lambda_{DD}$	13	10.58	59.84	8.88
$\max(\chi_{IP}^2 \mu)$	6	9.51	97.24	8.15
$\chi_{IP}^2 \Lambda_{LL}$	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
$\chi_{IP}^2 p_{LL}$	17	4.48	53.01	4.18
$\chi_{IP}^2 p_{DD}$	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
$\chi_{IP}^2 \pi_{DD}$	19	0.24	70.43	0.24

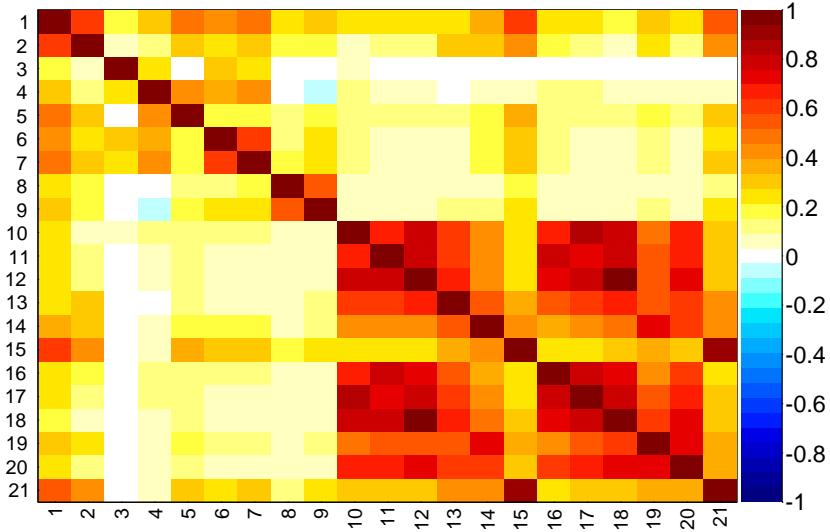


Figure 3.5: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of Tab. 3.2. Correlation is calculated using all events without distinguishing signal and background.

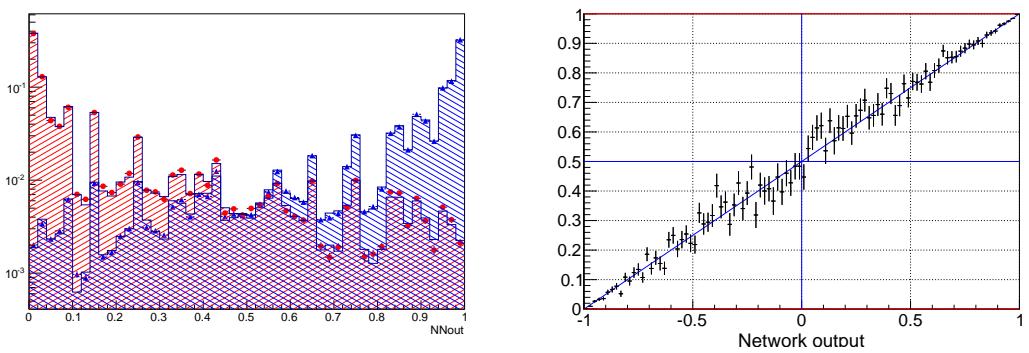


Figure 3.6: (left) NN output distribution for training (points) and test (stripes) samples, for signal and background events. (right) Purity as a function of neural network output.

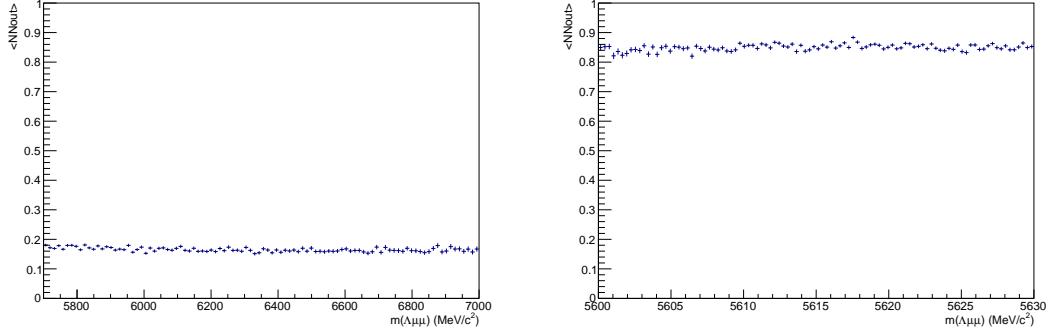


Figure 3.7: Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and simulated signal (right) events.

1105 3.4.3 MVA optimisation

1106 In the high q^2 region, where the signal is already observed, the requirement on
 1107 the neural network output is chosen maximising the significance, $N_S/\sqrt{N_S + N_B}$,
 1108 where N_S and N_B are the numbers of expected signal and background candidates
 1109 respectively. N_S is derived from simulation but, as an arbitrary number of events can
 1110 be generated, it needs to be normalised. To do this, the invariant mass distribution
 1111 of real $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates is fit after pre-selection (including all requirements
 1112 but MVA). This is possible as the peak of the resonant channel is already well visible
 1113 before the MVA cut. The resonant yield is then scaled by the ratio of between the
 1114 $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ branching fractions as measured by LHCb on 2011
 1115 data

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

1116 and by the $J/\psi \rightarrow \mu^+\mu^-$ branching fraction. In summary:

$$N_S = N_{J/\psi} \cdot \frac{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}. \quad (3.4)$$

1117 The number of expected background events instead is derived fitting the data side-
 1118 band with an exponential and extrapolating under the signal region.

1119 In the low q^2 region, where the signal is unobserved, the so called ‘‘Punzi figure-of-
 1120 merit’’, $N_S/(n_\sigma/2 + \sqrt{N_B})$, is maximised [89]. This figure-of-merit is considered to be

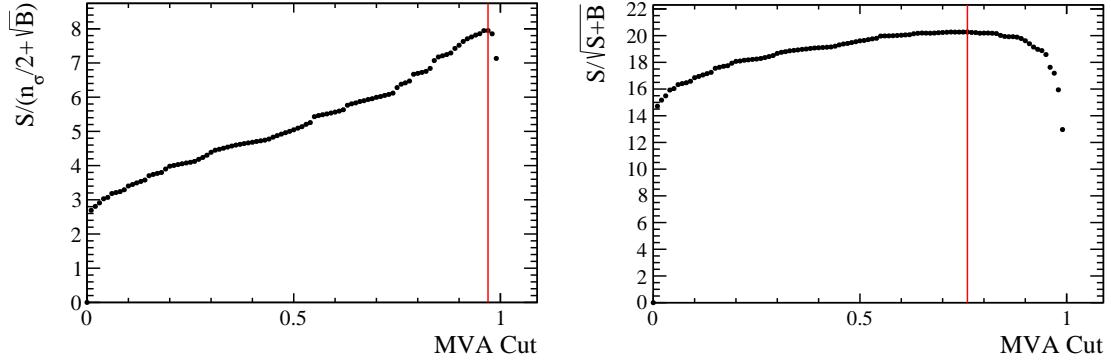


Figure 3.8: Dependence of the figure-of-merits on the neural network output requirement for the low q^2 (left) and high q^2 (right) regions. The vertical lines correspond to the chosen cuts.

optimal for discovery and the parameter n_σ corresponds to the number of expected standard deviations of significance, in this analysis $n_\sigma = 3$ is used. Moreover, the Punzi shape does not depend on the relative normalisation between signal and background, which is important since the signal is still unobserved at low q^2 and the existing predictions vary significantly for this region. The dependence of the figure-of-merit for both q^2 regions is shown in Fig. 3.8, and curves of signal efficiency versus background rejection are shown in Fig. 3.9.

For final selection the neural network output is required to be larger than 0.76 for candidates in the high q^2 region and 0.97 for the low q^2 ones. Using these requirements the neural network retains approximately 96% (66 %) of downstream candidates and 97 % (82 %) of long candidates for the high (low) q^2 selection, with respect to the pre-selected samples. After full selection $\sim 0.5\%$ of the events contain multiple candidates which are randomly rejected keeping only one candidate per event. To normalise the branching ratio measurement J/ψ events are selected using the low and high q^2 requirements to normalise respectively low and high q^2 intervals.

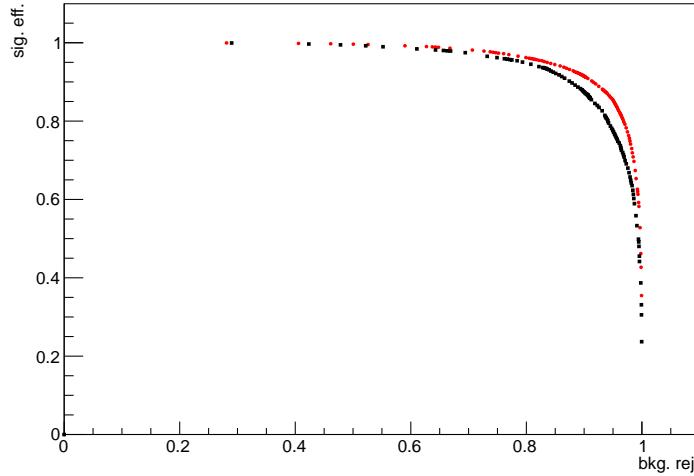


Figure 3.9: Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection. The optimal points on these curves are the closest ones to (1,1).

¹¹³⁷ 3.4.4 Trigger

¹¹³⁸ Finally, specific trigger lines are selected, corresponding to events triggered by muons
¹¹³⁹ which formed the reconstructed candidate. This is denoted as Trigger On Signal
¹¹⁴⁰ (TOS). The trigger lines used in the analysis are listed in Tab. 3.3. The logical *or*
¹¹⁴¹ of the lines on the same lever is required and the logical *and* of those on different
¹¹⁴² levels. The L0Muon trigger requires hits in the muon detector and triggers if a muon
¹¹⁴³ with $p_T > 1.5$ GeV/ c is identified. L0Dimuon imposes the same requirement on the
¹¹⁴⁴ sum of the transverse momenta of two tracks. The Hlt1TrackAllL0 performs a
¹¹⁴⁵ partial reconstruction of the events and applies basic requirements on the IP, χ^2
¹¹⁴⁶ and p_T of tracks; it triggers if the L0 decision is confirmed. Hlt1TrackMuon applies
¹¹⁴⁷ looser requirements but in addition requires the `isMuon` variable (see Sec. 2.8) to be
¹¹⁴⁸ true to limit the yield. Finally, at the Hlt2 level, a complete reconstruction is done
¹¹⁴⁹ and a multivariate analysis is used to identify decay structures. One of the main
¹¹⁵⁰ variables used at this stage is the Distance Of Closest Approach (DOCA), which is
¹¹⁵¹ required to be less than 0.2 mm to form a 2-body object.

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

Trigger Level	Lines
L0	LOMuon LODiMuon
Hlt1	Hlt1TrackAllL0 Hlt1TrackMuon
Hlt2	Hlt2Topo [2-4] BodyBBDT Hlt2TopoMu [2-4] BodyBBDT Hlt2SingleMuon Hlt2DiMuonDetached

1152 3.4.5 Background from specific decays

1153 Candidates from other decays can be reconstructed as the decays of interest if par-
 1154 ticles are not reconstructed or mis-identified. A survey of possible backgrounds
 1155 concluded that the only physics background to take into account comes from mis-
 1156 reconstructed decays of B^0 to K_s^0 with two muons in the final state, whether via
 1157 J/ψ or not, where the K_s^0 is reconstructed as a Λ with a $p \rightarrow \pi$ identity swap.
 1158 The lack of background from other decays is mainly due to the particular topol-
 1159 ogy of the Λ decay, which is long-lived and decays at a displaced vertex. To study
 1160 the effect of misreconstructed $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays simulated
 1161 samples are used. On data the $B^0 \rightarrow J/\psi K_s^0$ contribution is clearly visible in the
 1162 resonant channel mass distribution. This background is not suppressed with specific
 1163 cuts in this analysis as its mass shape is sufficiently distinct from the Λ_b^0 signal and
 1164 its contribution can be reliably modelled in the mass fits (see Sec. 3.5.1). For the
 1165 rare case a rough estimate of the K_s^0 background size is obtained using the yield in
 1166 the resonant channel rescaled by the measured ratio between the rare and resonant
 1167 branching fractions. Details are given in Sec. 3.5.1 and numbers of events predicted
 1168 are reported in Tab. 3.4. This contribution, although close to negligible is again
 1169 considered in the fit. A possible pollution due to $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ decays, where
 1170 the K^{*+} further decays into $K_s^0 \pi$ is also investigated using a dedicated simulated
 1171 sample and found to be negligible. Finally, $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events radiating photons
 1172 from the final state, can escape the J/ψ veto and be reconstructed in the rare chan-

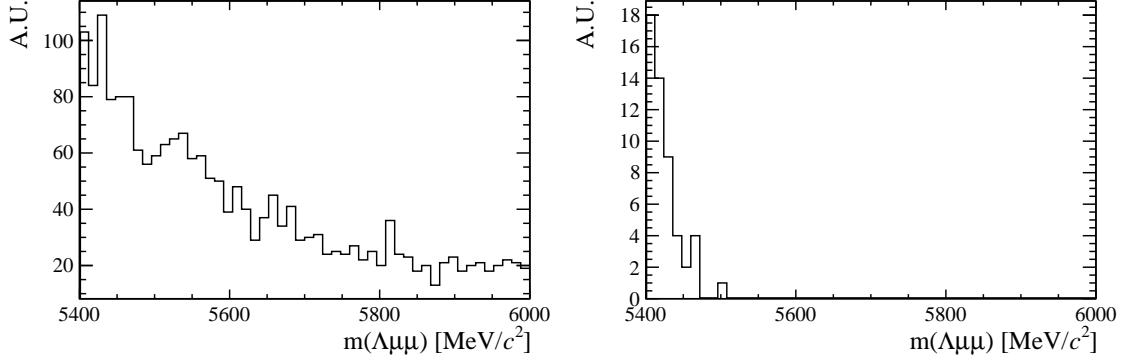


Figure 3.10: Invariant mass distributions of simulated $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ (left) and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ (right) candidates passing the full selection. Only $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed in $q^2 < 8 \text{ GeV}^2/c^4$ are selected. Distributions are shown in the invariant mass range relevant for the analysis (see Sec. 3.5.1).

nel sample. Analysing simulated events it was found that the only contribution is in the closest q^2 interval to the J/ψ tail, $6 < q^2 < 8 \text{ GeV}^2/c^4$. In this interval 1.3% of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates are reconstructed but only 0.06% fall into the 4-body invariant mass window used for the fits. This corresponds to ~ 6 events, 4 of which in the downstream category. Given the low yield and that these events do not peak under the signal but show a decaying distribution at the edge of the fit mass window, this background is considered as absorbed in the combinatorial background. Figure 3.10 shows the invariant mass distribution of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events falling into the rare q^2 region and the distribution of simulated $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ events mis-reconstructed as $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays.

1183 3.5 Yield extraction

1184 Extended unbinned maximum likelihood fits are used to extract the yields of the
1185 rare and resonant channels. The likelihood has the form:

$$\mathcal{L} = e^{-(N_S + N_C + N_B)} \times \prod_{i=1}^N [N_S P_S(m_i) + N_C P_C(m_i) + N_B P_B(m_i)] \quad (3.5)$$

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

3.5.1 Fit description

The fit is performed though the following steps:

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

In the first step simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distributions are fitted using the signal PDF alone. This is done separately for long and downstream candidates. Figure 3.11 shows distributions of candidates selected in the resonant sample with the fit function overlaid. The signal is described as the sum of two Crystal Ball functions (CB) with common mean (m_0) and tail slope (n). This is also known as Double Crystal Ball (DCB) function. A single Crystal Ball [90] is a probability density function commonly used to model processes involving energy loss. In particular it is used to describe resonances’ peaks with radiative tails. This function consists of a Gaussian

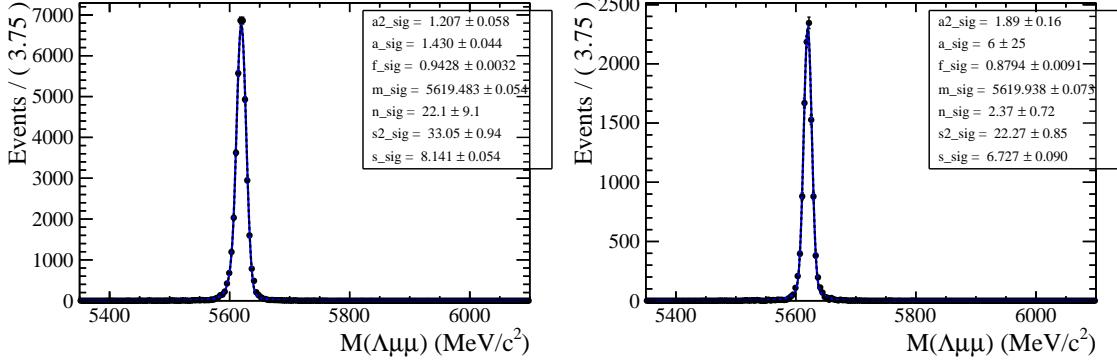


Figure 3.11: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ downstream (left) long (right) candidates. The points show simulated data and the blue line is the signal fit function.

₁₂₁₀ core and a power-law tail below a certain threshold and has form

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

₁₂₁₁ where for normalisation and continuity

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

₁₂₁₂ The full PDF for the resonant channel is therefore:

$$P_S(m; m_0, \alpha_1, \alpha_2, f, n) = f \text{CB}(m; m_0, \sigma_1, \alpha_1, n) + (1-f) \text{CB}(m; m_0, \sigma_2, \alpha_2, n), \quad (3.8)$$

₁₂₁₃ where f is the relative fraction of candidates falling into the first CB function.

₁₂₁₄ In a second step the fit to the resonant channel data sample is performed. For this fit
₁₂₁₅ the tail slope parameter, “ n ”, which is highly correlated with α_1 and α_2 , is fixed to
₁₂₁₆ the value found in the fit to simulated data. In this fit two background components
₁₂₁₇ are modelled: the combinatorial background, parameterized with an exponential
₁₂₁₈ and the background from $B^0 \rightarrow J/\psi K_s^0$ decays. The shape used to describe the
₁₂₁₉ K_s^0 background is obtained from a $B^0 \rightarrow J/\psi K_s^0$ simulated sample to which the
₁₂₂₀ full selection is applied. The invariant distribution of these events is fit with a DCB

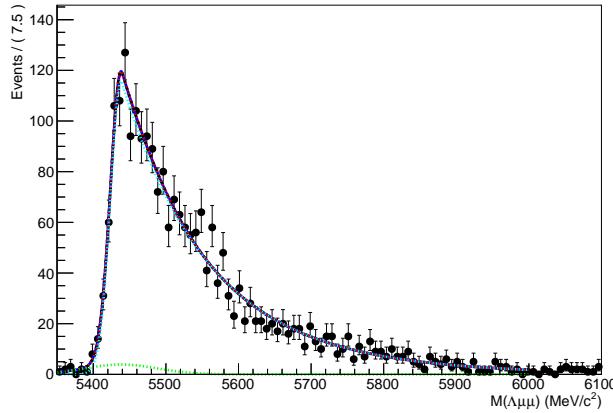


Figure 3.12: Invariant mass distribution of simulated $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted a Double Crystal Ball function.

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 $B^0 \rightarrow J/\psi K_s^0$ events is reported in Fig. 3.12. When the K_s^0 shape is introduced in the fit to the data all its parameters are fixed. This is particularly important when fitting long candidates, where the K_s^0 peak is less evident, which does not allow to constrain many parameters. On the other hand, in order to take into account possible data-simulation differences, an horizontal shift is added and left floating (by adding a constant to the central value of the DCB, $m_0 \rightarrow m_0 + m'$). 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. Note that all parameters of the fit to the long and downstream samples are independent.

Finally, the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ data sample is fit. In this case the fit to the long and downstream samples is performed simultaneously 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 as the result will be given as a ratio between rare and resonant quantities. However, the low statistics for the rare sample does not allow to constrain many parameters. Therefore, all parameters of the signal shape are fixed to the ones derived from the fit to the normalisation channel.

1240 However, to account for possible differences, arising from a different resolution in
 1241 different q^2 regions, a scale factor is multiplied to the widths of the two gaussian
 1242 cores of the signal DCB: $\sigma_1 \rightarrow c \cdot \sigma_1$ and $\sigma_2 \rightarrow c \cdot \sigma_2$, where the two scale factors
 1243 are the same. This factors are fixed in the fit to data by fitting rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$
 1244 simulated events in each q^2 bin and comparing the widths with the ones found on
 1245 the fit to the resonant simulated sample, namely

$$c = \sigma_{\mu^+\mu^-}^{MC} / \sigma_{J/\psi}^{MC}. \quad (3.9)$$

1246 Values obtained are ~ 1.9 for downstream candidates and ~ 2.3 for long candidates,
 1247 corresponding to the fact that in the resonant case a further constrain on the dimuon
 1248 mass is used, which improves the resolution by a factor of ~ 2 . The dependence of
 1249 the scaling factor on q^2 is found to be small. For the fits on the long and downstream
 1250 samples the parameters are always fixed to the corresponding J/ψ fit; in this analysis
 1251 shape parameters are never shared between the two candidate categories.

1252 Also in the rare case the modelled background components are the combinatorial
 1253 background, described with an exponential function and the K_s^0 background. The
 1254 slope of the background is visibly different depending on the q^2 interval. This is
 1255 partly due to the fact that at high q^2 the combinatorial changes slope because of
 1256 a kinematical limit at low 4-body masses imposed by the q^2 requirements. The
 1257 exponential slopes are therefore left as independent parameters in each q^2 interval.
 1258 The background component from $B^0 \rightarrow K_s^0\mu^+\mu^-$ decays is modelled using the same
 1259 shapes used for the resonant channel. However, in this case the horizontal shift is
 1260 fixed to what found for the resonant channel. The expected amount of misrecon-
 1261 structed $B^0 \rightarrow K_s^0\mu^+\mu^-$ events is small and does not allow to determine reliably the
 1262 yield. Therefore this is fixed to the yield of $B^0 \rightarrow J/\psi K_s^0$ decays rescaled by the
 1263 expected ratio of branching fractions between the resonant and rare channels. The
 1264 q^2 distribution of $B^0 \rightarrow K_s^0\mu^+\mu^-$ simulated events is used to predict the yield as a
 1265 function of q^2 . Table 3.4 reports the number of predicted $B^0 \rightarrow K_s^0\mu^+\mu^-$ events in

Table 3.4: 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

1266 each q^2 interval obtained with the following formula:

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

1267 where $N(q^2)_{MC}$ is the number of simulated rare candidates falling in a q^2 interval
1268 after full selection and N_{MC}^{tot} is the total number of simulated events.

1269 As the fit on the rare sample is performed simultaneously on long and downstream
1270 candidates, their two yields are not free to vary separately but are parameterised as
1271 a function of the common branching fraction using the following formula:

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

1272 where $k = (\text{LL}, \text{DD})$, Δq^2 is the width of the q^2 interval and the only free parameter
1273 is the relative branching fraction ratio of the rare over J/ψ channels. For the
1274 branching fraction of the $J/\psi \rightarrow \mu^+ \mu^-$ decay the value reported in the PDG book,
1275 $(5.93 \pm 0.06) \cdot 10^{-2}$ [2] is used and ε^{rel} corresponds to the relative efficiency between
1276 the rare and resonant channels obtained in Sec. 3.6. In this formula the efficiencies
1277 and the normalisation yield appear as constants, namely $N(\Lambda \mu^+ \mu^-)_k = C_k \cdot \mathcal{B}^{rel}$.

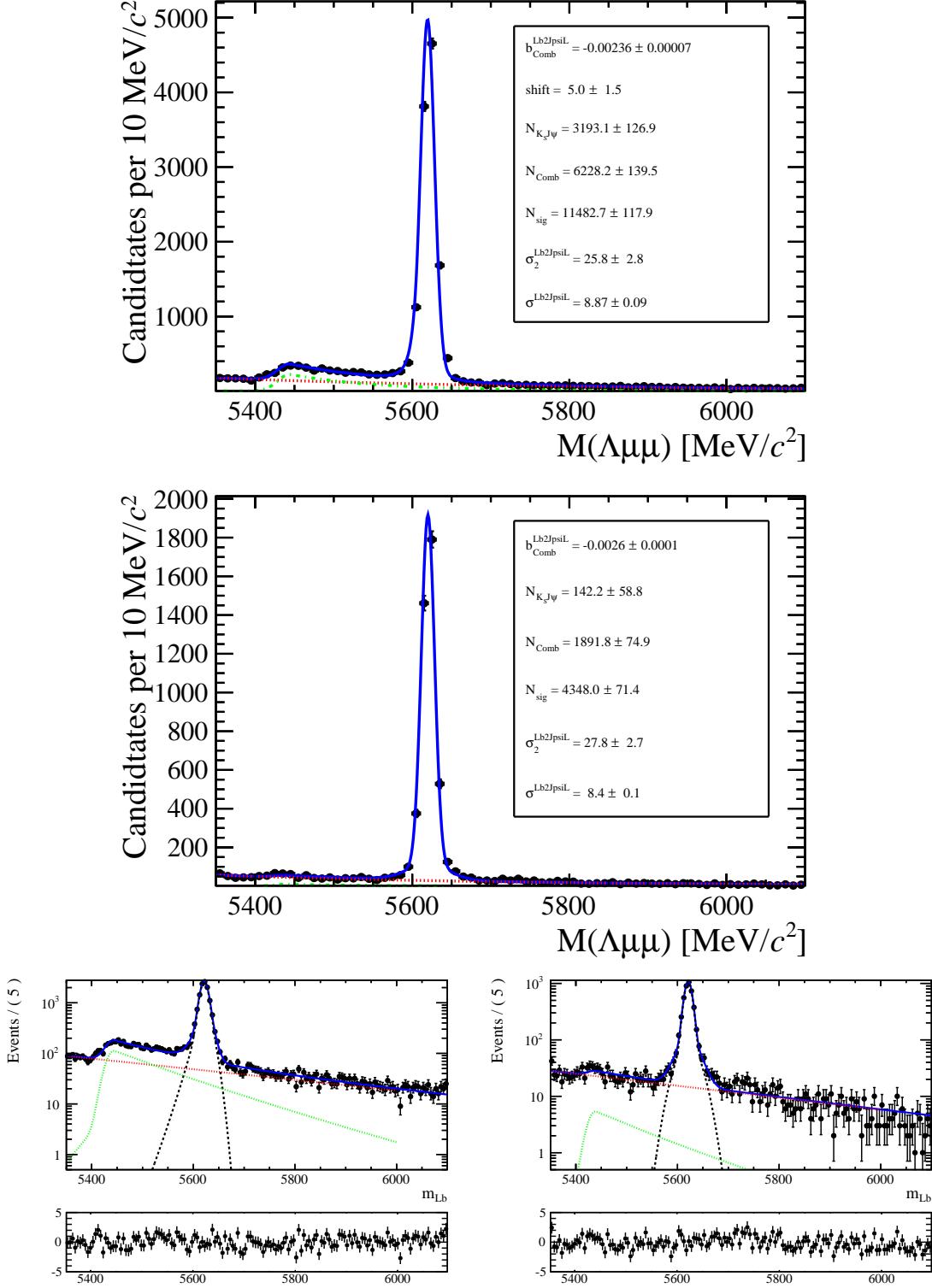


Figure 3.13: Invariant mass distributions of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ downstream (top) and long (middle) candidates selected with high q^2 requirements. Bottom plots are the same as the upper ones but shown in logarithmic scale. Black points show data. The blue solid line represents the total fit function, the black dashed line the signal, the red dashed line the combinatorial background and the green dashed line the $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ background.

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

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

1278 3.5.2 Fit results

1279 Figures 3.13 and 3.14 show fitted invariant mass distributions for the normalisation
 1280 channel, selected with the high q^2 and low q^2 requirements respectively. Table 3.5
 1281 reports the measured yields of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates found using the low and high
 1282 q^2 selections. Values for the signal shape parameters are shown on Fig. 3.13. Fits to
 1283 the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ samples are shown in Fig. 3.15 for the integrated $15 < q^2 < 20$
 1284 and $1.1 < q^2 < 6.0$ GeV $^2/c^4$ q^2 intervals, while fitted invariant mass distribution in
 1285 all other considered q^2 intervals are in Figs. 3.16 and 3.17 for downstream and long
 1286 candidates respectively. The yields of rare candidates obtained from the fit are
 1287 listed in Tab. 3.6 together with their significances. Most candidates are found in the
 1288 downstream sample, which comprises $\sim 80\%$ of the total yield. Note that, since the
 1289 fit is simultaneous to the two candidate categories, their yields are not parameters
 1290 free to float independently in the fit but are correlated via the branching ratio. The
 1291 statistical significance of the observed signal yields is evaluated as $\sqrt{2\Delta \ln \mathcal{L}}$, where
 1292 $\Delta \ln \mathcal{L}$ is the change in the logarithm of the likelihood function when the signal
 1293 component is excluded from the fit, relative to the nominal fit in which it is present.

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

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

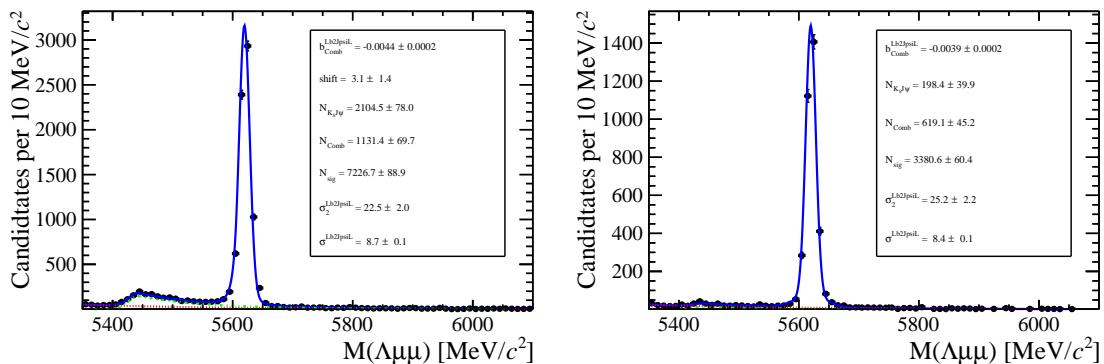


Figure 3.14: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ for downstream (left) and long (right) candidates selected with low q^2 requirements.

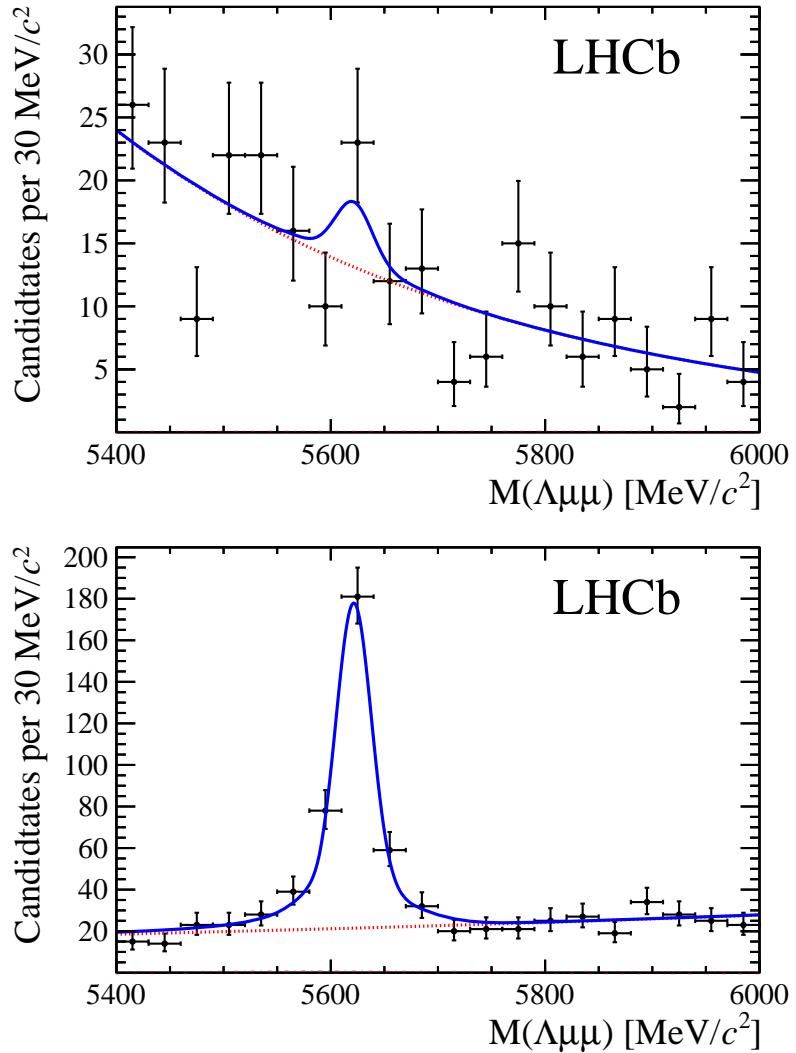


Figure 3.15: Invariant mass distributions of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the integrated 0.1 – $6.0 \text{ GeV}^2/c^4$ (top) and 15 – $20 \text{ GeV}^2/c^4$ (bottom) q^2 intervals. Points show data combining downstream and long candidates together. The blue solid line represents the total fit function and the dashed red line the combinatorial background.

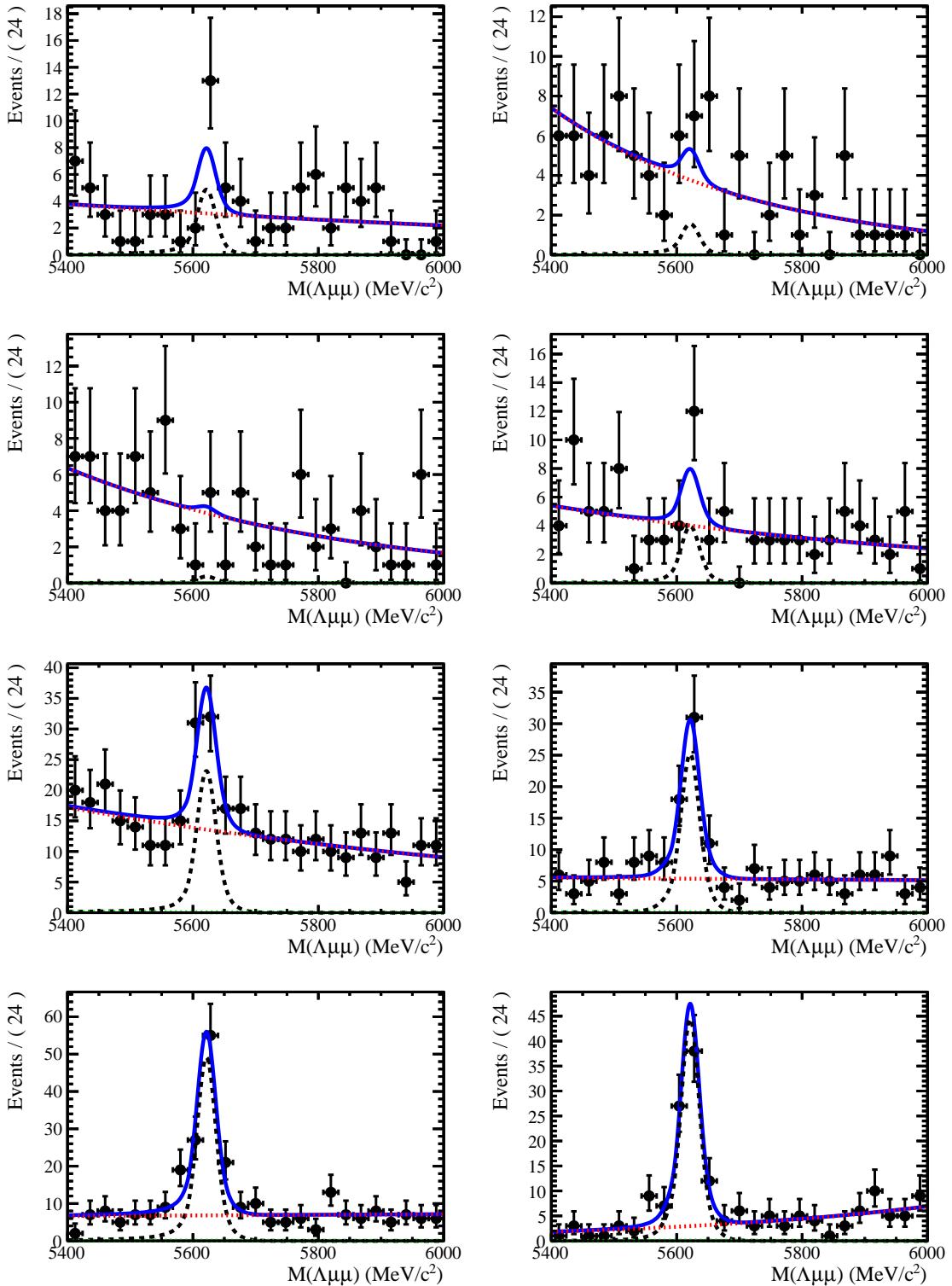


Figure 3.16: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for downstream candidates.

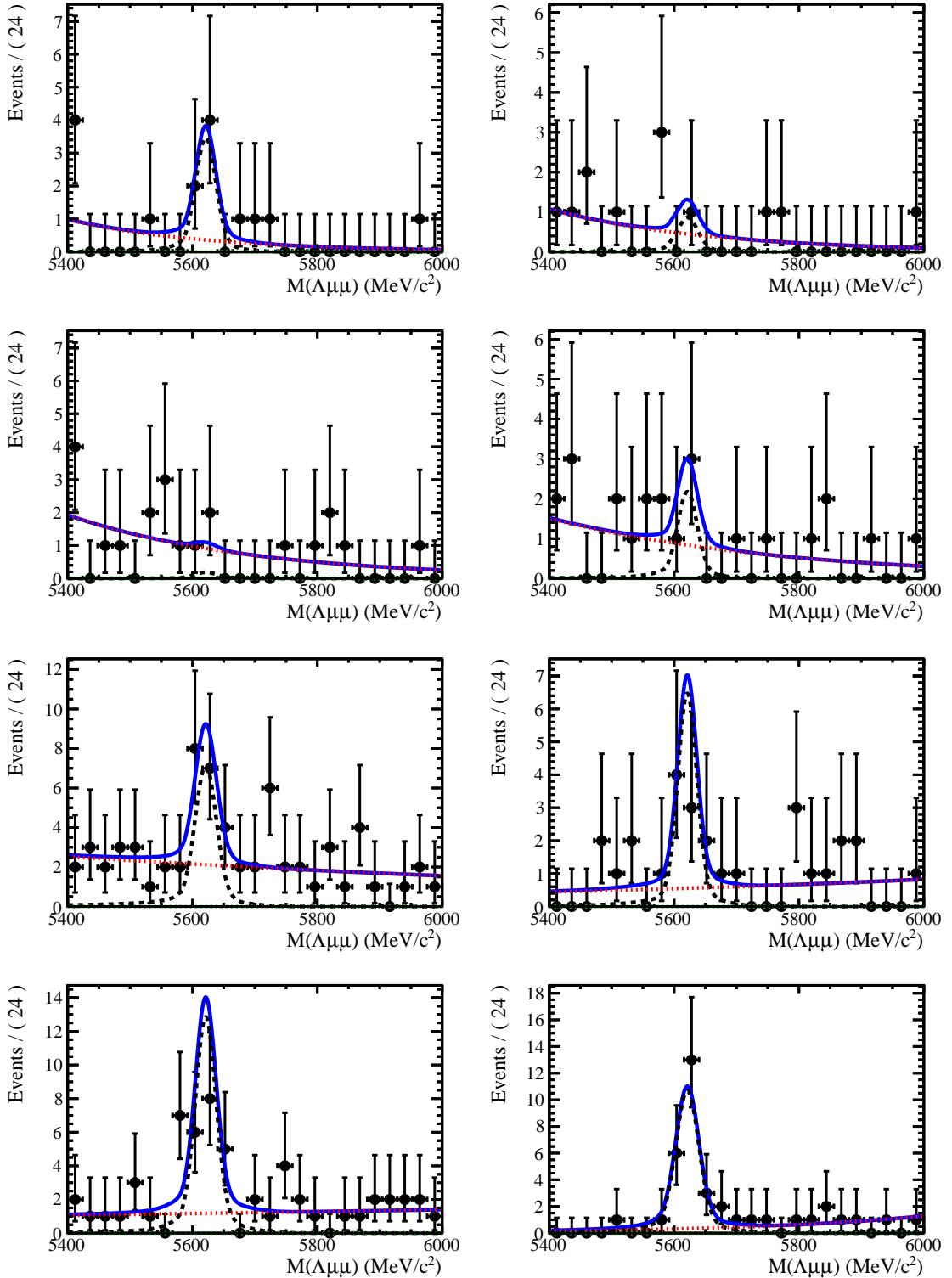


Figure 3.17: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for long candidates.

¹²⁹⁴ **3.6 Efficiency**

¹²⁹⁵ The selection efficiency is calculated for each decay according to the formula

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

¹²⁹⁶ In this expression the first term gives the efficiency to have final state particles
¹²⁹⁷ in the LHCb acceptance. The second term handles the possibility of Λ escaping
¹²⁹⁸ the detector or interacting with it and therefore never decaying into $p\pi$; this term
¹²⁹⁹ is referred to as “detection” efficiency. The third term carries information about
¹³⁰⁰ the reconstruction and pre-selection efficiencies, which are kept together given that
¹³⁰¹ boundaries between them are completely artificial. The fourth part deals with the
¹³⁰² efficiency of the Neural Network for those events which passed the pre-selection.
¹³⁰³ Finally, the last term handles the trigger efficiency for events which are accepted
¹³⁰⁴ by the full selection. Most of the efficiency components are evaluated using the
¹³⁰⁵ simulated samples described in Sec. 3.3. Only the efficiency of the PID requirement
¹³⁰⁶ for the proton (see Tab. 3.1) is separately derived with a data–driven method because
¹³⁰⁷ the simulation does not provide a good description of PID variables. For complete
¹³⁰⁸ information, all absolute efficiencies for the two decays $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$
¹³⁰⁹ are separately listed in the next subsections. However, for the analysis itself only
¹³¹⁰ the relative efficiency, $\varepsilon(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)/\varepsilon(\Lambda_b^0 \rightarrow J/\psi\Lambda)$, is used.

¹³¹¹ **3.6.1 Geometric acceptance**

¹³¹² In order to save disk space and time only events are simulated in which the final
¹³¹³ muons are in the detector acceptance and therefore can be reconstructed. This corre-
¹³¹⁴ sponds to a requirement for each of the muons to be in an interval $10 < \theta < 400$ mrad,
¹³¹⁵ where θ is the angle between the muon momentum and the beam line. The efficiency
¹³¹⁶ of this requirement is obtained by using a separate simulated sample, where events
¹³¹⁷ are generated in the full space. In Tab. 3.7 the efficiencies due to the geometrical
¹³¹⁸ acceptance are listed in bins of q^2 for $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays.

Table 3.7: Absolute geometrical acceptance in bins of q^2 derived from generator level simulated samples. Uncertainties are statistical only.

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

1319 3.6.2 Reconstruction and neural network efficiencies

1320 The efficiency to reconstruct the decays together with the pre-selection requirements
 1321 is evaluated from simulated data. Table 3.8 reports values of reconstruction effi-
 1322 ciency in bins of q^2 for long and downstream candidates. In the table the efficiency
 1323 is subdivided in “Detection” and “Reconstruction and pre-selection” efficiencies. In
 1324 fact, since Λ is a long lived particle, there is a non-negligible probability that it
 1325 interacts in the detector or escapes from it and therefore never decays in proton
 1326 and pion. The reconstruction efficiency includes the efficiency of for the tracks to
 1327 produce observable signatures and the efficiency for candidates to pass the pre-
 1328 selection requirements. This component does not include the efficiency of the PID
 1329 cut that appears in Tab. 3.1, which is kept separate because PID variables are not
 1330 well described by the simulation. The NN selection efficiency is again evaluated
 1331 from simulated samples. Results are shown in Tab. 3.9 in bins of q^2 . The sudden
 1332 jump in efficiency at ~ 9 GeV/ c^2 is due to the fact that a different figure-of-merit is
 1333 used to optimise the NN cut in the low and high q^2 regions, which results in different
 1334 efficiencies.

Table 3.8: Absolute detection and reconstruction plus stripping efficiencies. Reconstruction efficiency is given separately for downstream and long candidates. Uncertainties are statistical only.

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

Table 3.9: Neural network selection efficiency. Uncertainties are statistical only.

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

3.6.3 Trigger efficiency

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

$$\varepsilon_{TOS} = \frac{TIS \& TOS}{TIS}. \quad (3.13)$$

1343 As data contains background the numbers of signal candidates in the “TIS” and
1344 “TIS $\&\&$ TOS” categories are not just determined by counting events but from a fit
1345 to the 4-body invariant mass, $m(p\pi\mu\mu)$. This procedure takes the name of TISTOS
1346 method. Using the data-driven method an efficiency of $(70 \pm 5)\%$ is obtained, while
1347 this is calculated to be $(73.33 \pm 0.02)\%$ using the simulation. Results are therefore
 compatible within 1σ .

Table 3.10: Absolute trigger efficiencies for selected events as determined from the simulation separately for downstream and long 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

1348

1349 3.6.4 PID efficiency

1350 For long tracks a PID requirement on protons ($\text{PID}_p > -5$) is applied. The simula-
1351 tion is known not to describe particle ID variables well and therefore a data-driven
1352 method is used to obtain this efficiency component. This is done using the **PIDCalib**
1353 package (see Sec. 2.8.1), which uses as calibrations samples decays where particles
1354 can be identified due to their kinematic properties. In the case of protons a sample
1355 of Λ particles is used, where the proton can be identified because it always has the
1356 highest momentum. The package allows to divide the phase space in bins of variables
1357 relevant for PID performances; in this analysis momentum and pseudorapidity are
1358 used. Using the calibration sample the efficiency is derived in each two-dimensional
1359 bin. Finally, to take into account that the decay channel under study could have
1360 different kinematical distributions than the calibration sample these efficiency tables

¹³⁶¹ are used to re-weight the simulation. Absolute PID efficiencies are listed in Tab. 3.11
in bins of q^2 .

Table 3.11: 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

¹³⁶²

¹³⁶³ 3.6.5 Relative efficiencies

¹³⁶⁴ In the previous sections absolute efficiencies values were given for the rare channel
¹³⁶⁵ in different q^2 intervals. Figure 3.18 contains a summary of those values in these
¹³⁶⁶ tables in graphical form. This section reports the corresponding relative efficiencies
¹³⁶⁷ with respect to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ channel, which will be used to correct the yields
¹³⁶⁸ and obtain the differential branching fraction. Table 3.12 reports the absolute effi-
¹³⁶⁹ ciency values for the J/ψ channel used to derive the relative efficiencies. Relative
¹³⁷⁰ geometric, detection and PID efficiencies are listed in Tab. 3.13, while Tabs. 3.15
¹³⁷¹ and 3.14 report relative reconstruction, trigger and NN efficiencies separately for
¹³⁷² downstream and long candidates. Since the latter three components are obtained
¹³⁷³ from the same simulated sample their statistical errors are correlated. Therefore the
¹³⁷⁴ total of the three is also reported as a single efficiency and labeled “Full Selection”.
¹³⁷⁵ Finally, Tab. 3.18 reports the total of all relative efficiencies, which will be then
¹³⁷⁶ used to correct the raw yields and calculate the differential branching fraction. Un-
¹³⁷⁷ certainties reflect the statistics of both rare and resonant samples, while systematic
¹³⁷⁸ uncertainties are discussed in next sections.

Table 3.12: Absolute efficiency values for $\Lambda_b^0 \rightarrow J/\psi \Lambda$. Uncertainties are statistical only.

Efficiency	Downstream	Long
Geometric	0.1818 ± 0.0003	
Detection	0.9017 ± 0.0003	
Reconstruction	0.0724 ± 0.0004	0.0203 ± 0.0002
MVA	0.882 ± 0.002	0.942 ± 0.002
Trigger	0.697 ± 0.003	0.734 ± 0.005
Full Selection	0.0445 ± 0.0003	0.0140 ± 0.0002
Total	0.00729 ± 0.00005	0.00230 ± 0.00003

Table 3.13: Relative geometric, detection and PID relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays. Uncertainties reflect the statistics of both samples.

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

Table 3.14: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays for long events. Uncertainties reflect the statistics of both samples.

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

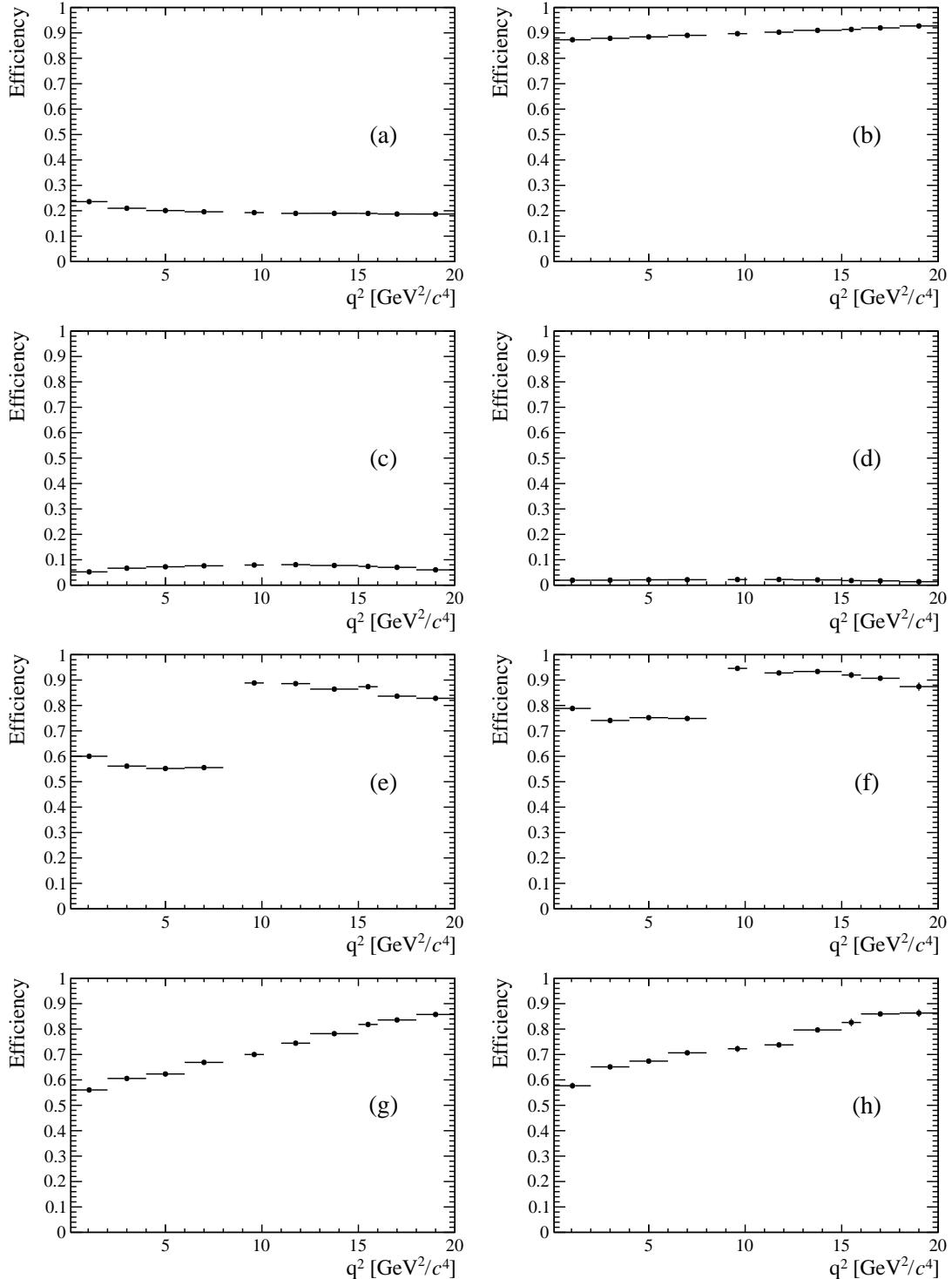


Figure 3.18: Absolute efficiencies as a function of q^2 : geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) candidates, NN efficiency for DD (e) and LL (f) and trigger efficiency for DD (g) and LL (h).

Table 3.15: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for downstream events. Uncertainties reflect the statistics of both samples.

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

¹³⁷⁹ **3.7 Systematic uncertainties**

¹³⁸⁰ This section describes the main considered sources of systematic uncertainty.

¹³⁸¹ **3.7.1 Systematic uncertainty on the yields**

¹³⁸² The choice of specific PDFs to model the invariant mass distribution could result in
¹³⁸³ a bias. To asses the effect of the signal PDF choice as a first step a number of models
¹³⁸⁴ are tried on the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ data sample to understand which ones are plausible.
¹³⁸⁵ Table 3.16 reports the χ^2 and relative probabilities obtained using different models
¹³⁸⁶ including: the default model (a DCB function), a simple Gaussian function, a single
¹³⁸⁷ Crystal Ball function and the sum of two Gaussians. The only two models that give
¹³⁸⁸ a reasonable p-value are the default DCB and the sum of two Gaussian functions
¹³⁸⁹ (DG). In a second step simulated experiments are generated and fit with the two
¹³⁹⁰ chosen models. Events are generated according to a density function given by the
¹³⁹¹ default model fitted on data separately for each q^2 interval. In this way, for each
¹³⁹² q^2 interval, a specific shape is reproduced including the background level and slope.
¹³⁹³ Furthermore, a number of events comparable to the one found in data is generated.
¹³⁹⁴ For each experiment a normalised bias is calculated as

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

¹³⁹⁵ where $N_{\ell\ell}^{model}$ and $N_{J/\psi}^{model}$ are the numbers of rare and resonant candidates observed
¹³⁹⁶ using a specific model. The average bias over 1000 pseudo-experiments is taken as
¹³⁹⁷ systematic uncertainty. Note that in each case the rare and normalisation channels
¹³⁹⁸ are fit with the same signal model and, while for the default case the rare parameters
¹³⁹⁹ are fixed to what found for the resonant channel, they are left free to vary in the
¹⁴⁰⁰ second model in order to asses at the same time the systematic due to the parameters
¹⁴⁰¹ constraints.

Table 3.16: 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

1402

1403 For the background PDF systematic the rare channel is re-fit leaving the yield of
 1404 the K_s^0 component free to vary; this is instead fixed to the predicted value in the
 1405 default fit. The same procedure as for the signal PDF is applied. Results are re-
 1406 ported in Tab. 3.17. The most affected q^2 interval is the one in the middle of the
 1407 charmonium resonances, where a combination of lower statistics and higher back-
 1408 ground leaves more freedom to the signal shape. Finally, a background component
 1409 for $B^+ \rightarrow K^{*+}(K_s^0\pi^+)\mu^+\mu^-$ decays is added to the fit, modelled using the distri-
 1410 bution of simulated events after full selection. No significant bias is found for this
 1411 component.

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

Table 3.17: Values of systematic uncertainties due to the choice of signal and background shapes in bins of q^2 .

1412

¹⁴¹³ 3.7.2 Systematic uncertainties on the efficiency determination

¹⁴¹⁴ Systematic uncertainties in the efficiency determination are due to the limited knowl-
¹⁴¹⁵ edge of the decay properties such as the Λ_b^0 lifetime and production polarisation. The
¹⁴¹⁶ uncertainties are directly calculated on the relative efficiencies as these are the ones
¹⁴¹⁷ that are actually used in the analysis. It should be noted that not all sources con-
¹⁴¹⁸ tribute to each part of the efficiency. For brevity, this section only reports estimates
¹⁴¹⁹ of the systematic uncertainties obtained while the full information is contained in
¹⁴²⁰ Appendix C.

¹⁴²¹ 3.7.2.1 Effect of new physics on the decay model

¹⁴²² New physics could affect the decay model by adding contributions to the C_7 and
¹⁴²³ C_9 Wilson Coefficients. This would result in a modification of the q^2 spectrum
¹⁴²⁴ and therefore of the efficiency. To asses this systematic the Wilson Coefficients are
¹⁴²⁵ modified by adding a new physics component ($C_i \rightarrow C_i + C_i^{\text{NP}}$). Figure 3.19 shows q^2
¹⁴²⁶ spectra obtained weighting the simulation for a model embedding the default and 3
¹⁴²⁷ modified sets of Wilson Coefficients. The used values, reported on top of each plot,
¹⁴²⁸ are inspired to maintain compatibility with the recent LHCb result about the P'_5
¹⁴²⁹ observable [38]. The biggest effect is observed in the very low q^2 , below 2 GeV^2/c^4 ,
¹⁴³⁰ where the efficiency can change up to 7%, while it changes 3-4 % between 3 and
¹⁴³¹ 4 GeV^2/c^4 and 2-3 % in the rest of the spectrum. As this analysis is performed under
¹⁴³² the hypothesis that the decays are described by a the SM, these values are given in
¹⁴³³ order to provide the full information but are not added as systematic uncertainties.

¹⁴³⁴ 3.7.2.2 Simulation statistics

¹⁴³⁵ The limited statistics of the simulated samples used to determine efficiencies is
¹⁴³⁶ considered as a source of systematic uncertainty. While it is not the dominant
¹⁴³⁷ source, its size does not allow to completely neglect it. When reporting relative

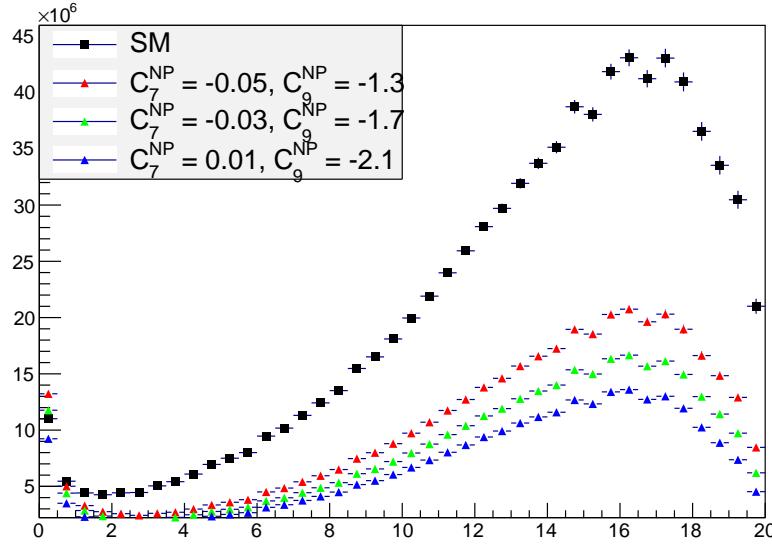


Figure 3.19: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

¹⁴³⁸ efficiency values the statistical uncertainty due to the rare and resonant channels is
¹⁴³⁹ always considered.

¹⁴⁴⁰ 3.7.2.3 Production polarisation and decay structure

¹⁴⁴¹ One of the main unknown, which affects the determination of the efficiencies, is
¹⁴⁴² the angular structure of the decays. And, connected to it, also the production
¹⁴⁴³ polarisation, which is a parameter of the model. To assess the systematic uncertainty
¹⁴⁴⁴ due to the knowledge of the production polarisation for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays the
¹⁴⁴⁵ polarisation parameter in the model is varied within one standard deviation from
¹⁴⁴⁶ the central value of the most recent LHCb measurement, $P_b = 0.06 \pm 0.09$ [88]. The
¹⁴⁴⁷ full observed difference is taken as systematic uncertainty. To assess the systematic
¹⁴⁴⁸ uncertainty due to the decay structure an alternative set of form factors is used based
¹⁴⁴⁹ on lattice QCD calculation [91]. Details of this are explained in Appendix A.1. The
¹⁴⁵⁰ two models are compared and the full difference is taken as systematic uncertainty.
¹⁴⁵¹ In total this results in an uncertainty of $\sim 1.3\%$ for long candidates and $\sim 0.6\%$

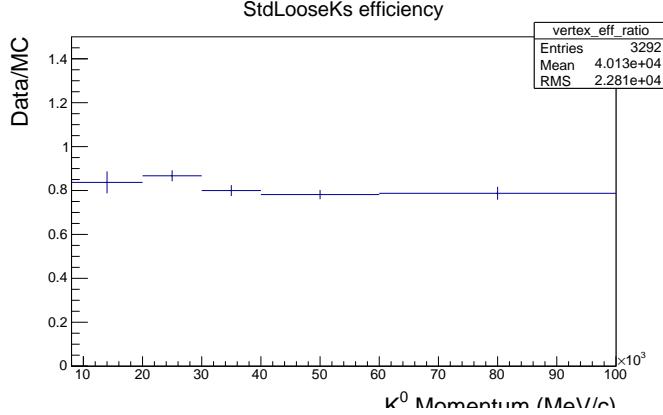


Figure 3.20: Ratio of reconstruction efficiency in Data and MC found using K_S events [93].

1452 for downstream candidates, mostly coming from the knowledge of the production
 1453 polarisation.

1454 3.7.2.4 Λ_b^0 lifetime

1455 The Λ_b^0 lifetime is known with limited precision. For evaluation of the efficiencies the
 1456 world average value, 1.482 ps^{-1} [92] is used. To evaluate the systematic uncertainty,
 1457 this is varied within one standard deviation from the measured value. Only the
 1458 case where both signal and normalisation channel are varied in same direction are
 1459 considered. The larger difference with the default lifetime case is taken as systematic
 1460 uncertainty, which is found to range from $\sim 0.4\%$ at low q^2 to $\sim 0.1\%$ at high q^2 .

1461 3.7.2.5 Downstream candidates reconstruction efficiency

1462 Other analysis in LHCb using particles reconstructed with downstream tracks showed
 1463 that the efficiency for these candidates is not well simulated. For example, Fig. 3.20
 1464 shows the ratio between the reconstruction efficiency for downstream candidates in
 1465 data and simulation found analysing K_S^0 events [93]. This effect is not yet fully
 1466 understood and is currently under study. It seems to be mainly due to a poor sim-
 1467 ulation of the vertexing efficiency for downstream tracks. This effect is dealt with

in two steps. Firstly, the analysis is performed separately for downstream and long candidates. Since efficiencies are also calculated separately, the effect mostly cancels in the ratio between the rare and resonant channels. In a second step a systematic uncertainty is assigned for downstream candidates only re-weighting the simulation by the efficiency ratio between data and simulation found for K_S as a function of momentum (see Fig. 3.20). The efficiencies obtained using the weighted and unweighted simulation are compared and the full difference is taken as systematic uncertainty. As the discrepancy shows little dependence on momentum, dependencies due to the different momentum distributions of Λ and K_S^0 are assumed to be negligible. This results in an extra 0.4% systematic uncertainty at low q^2 and 1.2% at high q^2 , only for downstream candidates.

3.7.2.6 Data-simulation discrepancies

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

3.8 Differential branching ratio extraction

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

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

where $k = (\text{LL}, \text{DD})$, Δq^2 is the width of the q^2 interval, $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$ [2] and the only free parameter is the relative branching fraction ratio. Table 3.18 sum-

Table 3.18: Absolute values of the total relative efficiency and the absolute value of the uncorrelated uncertainty, together with relative values of the correlated uncertainty.

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

1491 marises the total relative efficiencies, ε^{rel} , for downstream and long candidates to
1492 gether with their correlated and uncorrelated uncertainties, where the correlation is
1493 intended between the downstream and long samples. On the table the uncorrelated
1494 uncertainty corresponds to the total systematic uncertainty on the efficiency deter-
1495 mination. The correlated uncertainty is given in percent form since it can be applied
1496 to either downstream, long candidates or their combination. This includes the PDF
1497 systematic described in Sec. 3.7.1 and the systematic due to the uncertainty on the
1498 $J/\psi \rightarrow \mu^+\mu^-$ branching fraction.

1499 Figure 3.21 shows the branching fraction obtained by fitting the downstream and
1500 long samples independently, while the combined result, obtained fitting both samples
1501 simultaneously, is shown in Fig. 3.22. Values are also listed in Tab. 3.19, where
1502 the statistical uncertainty on the rare channel and the total systematic uncertainty
1503 are shown separately. The statistical uncertainty is calculated using the MINOS
1504 application of the MINUIT package [94], which provides an asymmetric interval. The
1505 normalisation and systematic uncertainties are evaluated by pushing the efficiencies
1506 and normalisation yields up and down by one standard deviation and re-performing
1507 the fit. The different efficiencies used translate into a different branching fraction and
1508 the full difference with respect to the default fit is taken as systematic uncertainty
1509 in each direction.

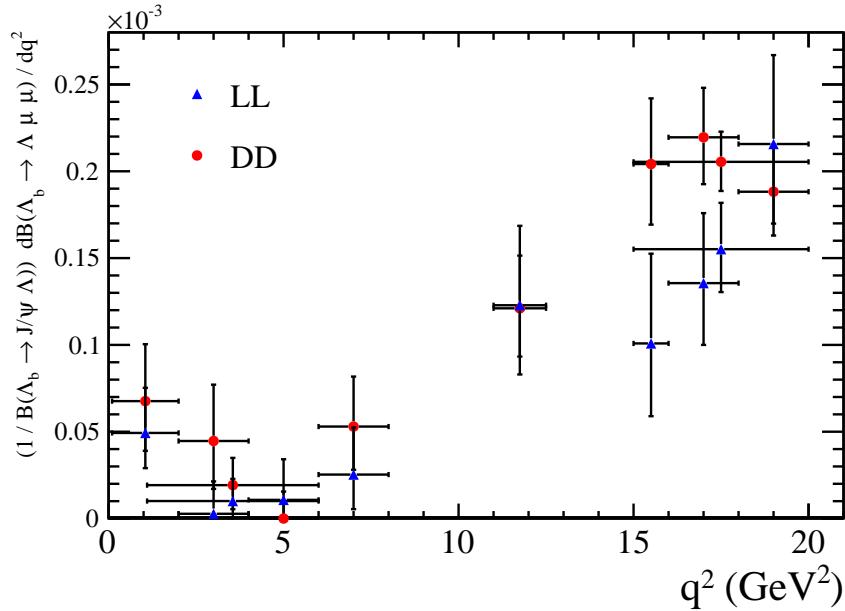


Figure 3.21: Measured values of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction relative to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay as a function of q^2 obtained fitting the downstream and long samples independently. Error bars represent the total statistical and systematic uncertainty.

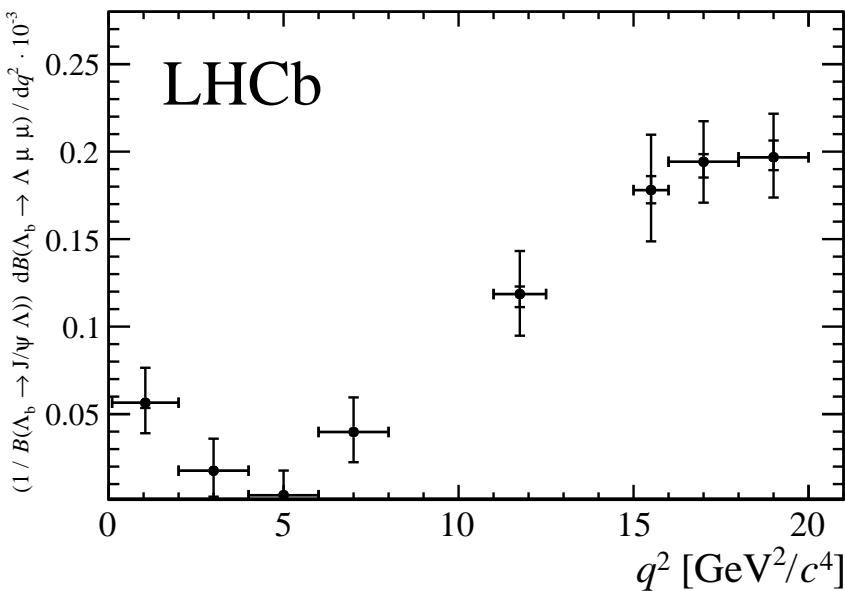


Figure 3.22: Branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay normalised to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ mode. The inner error bar represents the systematic uncertainty and the outer error bar includes the statistical uncertainty.

Table 3.19: 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

Finally, values for the absolute branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay are obtained by multiplying the relative branching fraction by the absolute branching fraction of the normalisation channel, $\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi\Lambda) = (6.3 \pm 1.3) \times 10^{-4}$ [2]. Values are shown in Fig. 3.23 and summarised in Tab. 3.20, where the uncertainty due to the knowledge of the normalisation channel (norm), which is correlated across q^2 , is shown separately. The SM predictions on the plot are obtained from Ref. [91].

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

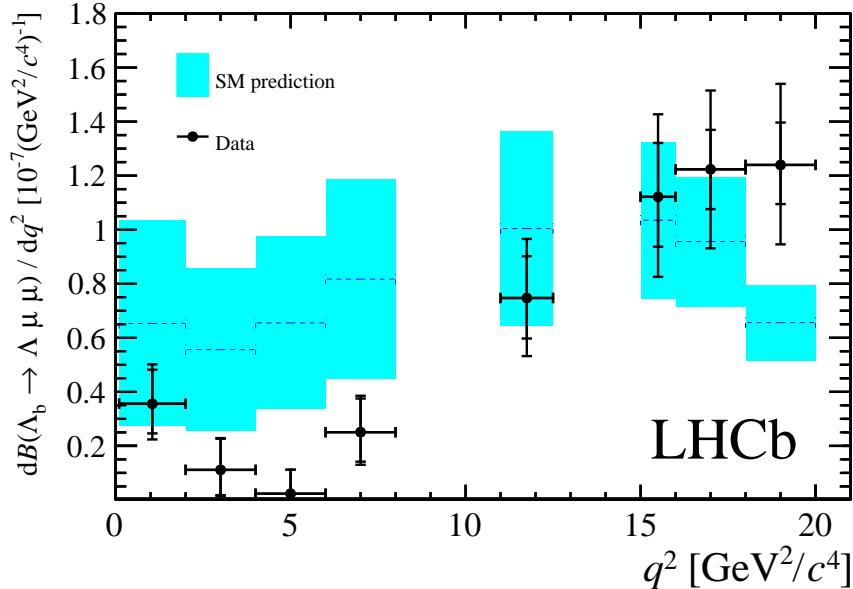


Figure 3.23: Measured $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction as a function of q^2 with the SM predictions [91] superimposed. The inner error bars on data points represent the total uncertainty on the relative branching fraction (statistical and systematic); the outer error bar also includes the uncertainties from the branching fraction of the normalisation mode.

Table 3.20: Measured differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay, where the uncertainties are statistical, systematic and due to the uncertainty on the normalisation mode, $\Lambda_b^0 \rightarrow J/\psi \Lambda$, respectively.

q^2 interval [GeV^2/c^4]	$d\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-)/dq^2 \cdot 10^{-7}[(\text{GeV}^2/c^4)^{-1}]$			
0.1 – 2.0	0.36	$+0.12$	$+0.02$	± 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

CHAPTER 4

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

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

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

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

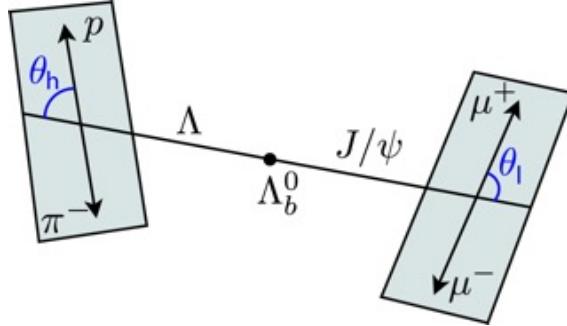


Figure 4.1: Graphical representation of the angles for the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay.

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

4.1 One-dimensional angular distributions

This section describes the derivation of the functional form of the differential distributions as a function of $\cos\theta_\ell$ and $\cos\theta_h$, which are used to measure the observables. The content of this section is based on the calculations in Ref. [87].

For unpolarised Λ_b^0 production, integrating over the three angles the differential branching fraction is given in Eq. 11 of Ref. [87] as

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

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

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

In these formulas m_ℓ is the mass of the lepton and $v = \sqrt{1 - 4m_\ell^2/q^2}$; U denotes the unpolarised-transverse contributions, L the longitudinal contributions and S the scalar contribution. The apices V and A represent respectively vector and axial-vector currents, with $X^{V+A} = X^V + X^A$. The authors of Ref. [87] define then the lepton-side forward-backward asymmetry as

$$A_{\text{FB}}^\ell(q^2) = -\frac{3}{2} \frac{v \cdot P^{VA}}{v^2 \cdot (U^{V+A} + L^{V+A}) + \frac{2m_\ell^2}{q^2} \cdot 3 \cdot (U^V + L^V + S^A)}. \quad (4.4)$$

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

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

and

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

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

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

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

$$f_L = \frac{L^{V+A}}{U^{V+A} + L^{V+A}}, \quad (4.8)$$

which leads to the functional form used in the analysis

$$\frac{d\Gamma}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) (1 - f_L) + A_{\text{FB}}^\ell \cos \theta_\ell + \frac{3}{4} \sin^2 \theta_\ell f_L \right]. \quad (4.9)$$

¹⁵⁵² Using the same steps the proton helicity distribution is given in Ref. [87] as

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda(\rightarrow p\pi^-)\ell^+\ell^-)}{dq^2 d \cos \theta_h} = \text{Br}(\Lambda \rightarrow p\pi^-) \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+\ell^-)}{dq^2} \left(\frac{1}{2} + A_{\text{FB}}^h \cos \theta_h \right), \quad (4.10)$$

¹⁵⁵³ and A_{FB}^h is defined as

$$A_{\text{FB}}^h = \frac{1}{2} \alpha_\Lambda P_z^\Lambda(q^2), \quad (4.11)$$

¹⁵⁵⁴ where $P_z^\Lambda(q^2)$ is the polarisation of the daughter baryon, Λ , and $\alpha_\Lambda = 0.642 \pm 0.013$ [2]

¹⁵⁵⁵ is the Λ decay asymmetry parameter.

¹⁵⁵⁶ These expressions assume that Λ_b^0 is produced unpolarised, which is in agreement
¹⁵⁵⁷ with the recent LHCb measurement [95]. Possible effects due to a non zero produc-
¹⁵⁵⁸ tion polarisation are investigated as systematic uncertainties (see Sec. 4.5.5).

¹⁵⁵⁹ 4.2 Multi-dimensional angular distributions

To incorporate effects of production polarisation this was introduced in the equations. In the modified version an angle θ is defined as the angle between the Λ direction in the Λ_b^0 rest frame with respect to $\hat{n} = \hat{p}_{\text{inc}} \times \hat{p}_{\Lambda_b^0}$, where \hat{p}_{inc} represents the direction of the incoming proton. This angle is sensitive to the production polarisation through the spin-density matrix in Eq. A.3. Integrating over all the angles but θ_ℓ results in the same distribution as in the unpolarised case (Eq. 4.3). Therefore, in the case of uniform efficiency, the lepton side forward-backward asymmetry, A_{FB}^ℓ , is unaffected by the production polarisation. To be able to estimate the effect of the production polarisation in the case of non-uniform efficiency, the differential distribution in θ and θ_ℓ is also derived, which in the massless leptons limit becomes

(up to a constant multiplicative factor)

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d\cos\theta d\cos\theta_\ell} = \frac{d\Gamma}{dq^2} \left\{ \frac{3}{8} (1 + \cos^2\theta_\ell) (1 - f_L) + A_{FB}^\ell \cos\theta_\ell + \frac{3}{4} \sin^2\theta_\ell f_L + P_b \cos\theta \left[-\frac{3}{4} \sin\theta_\ell^2 O_{Lp} + \frac{3}{8} (1 + \cos\theta_\ell^2) O_P \right. \right. \\ \left. \left. - \frac{3}{8} \cos\theta_\ell O_{UVA} \right] \right\}, \quad (4.12)$$

where three more observables are defined

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

- 1560 In the massless leptons approximation two of these quantities are related to the
1561 hadron side forward-backward asymmetry as

$$\frac{1}{2} \alpha_\Lambda (O_P + O_{Lp}) = A_{FB}^h. \quad (4.13)$$

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

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d(\cos\theta) d(\cos\theta_h)} = \frac{d\Gamma}{dq^2} [1 + 2A_{FB}^h \cos\theta_h + P_b (O_P - O_{Lp}) \cos\theta \\ + \alpha_A P_b (1 - 2f_L) \cos\theta \cos\theta_h]. \quad (4.14)$$

- 1562 In order to use these distributions, expectations for the three additional observables,
1563 which do not enter one-dimensional distributions, are needed. Expectations are
1564 calculated using form factors and numerical inputs from Ref. [87] and are shown in
1565 Tab. 4.1.

$q^2 [GeV^2/c^2]$	A_{FB}^ℓ	P_z^A	f_L	O_P	O_{Lp}	O_{UVA}
0.1 – 2.0	0.082	-0.9998	0.537	-0.463	-0.537	0.055
2.0 – 4.0	-0.032	-0.9996	0.858	-0.142	-0.857	-0.021
4.0 – 6.0	-0.153	-0.9991	0.752	-0.247	-0.752	-0.102
V.0 – VA.5	-0.348	-0.9834	0.508	-0.478	-0.505	-0.239
15.0 – 16.0	-0.384	-0.9374	0.428	-0.524	-0.413	-0.280
16.0 – 18.0	-0.377	-0.8807	0.399	-0.513	-0.368	-0.294
18.0 – 20.0	-0.297	-0.6640	0.361	-0.404	-0.260	-0.314
1.0 – 6.0	-0.040	-0.9994	0.830	-0.170	-0.830	-0.027
15.0 – 20.0	-0.339	-0.7830	0.385	-0.461	-0.3A	-0.302

Table 4.1: Prediction for angular observables entering two-dimensional angular distributions. Prediction is based on covariant quark model form factors from Ref. [87].

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

$$\begin{aligned} \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_h d \cos \theta_\ell} = & \frac{3}{8} + \frac{6}{16} \cos^2 \theta_\ell (1 - f_L) - \frac{3}{16} \cos^2 \theta_\ell f_L + A_{FB}^\ell \cos \theta_\ell + \\ & \left(\frac{3}{2} A_{FB}^h - \frac{3}{8} \alpha_A O_P \right) \cos \theta_h - \frac{3}{2} A_{FB}^h \cos^2 \theta_\ell \cos \theta_h - \frac{3}{16} f_L + \\ & \frac{9}{16} f_L \sin^2 \theta_\ell + \frac{9}{8} \alpha_A \cos^2 \theta_\ell \cos \theta_h O_P - \\ & \frac{3}{2} \alpha_A \cos \theta_\ell \cos \theta_h O_{UVA}. \end{aligned} \quad (4.15)$$

1566 4.3 Angular resolution

1567 This section describes a study of the angular resolution done in order to achieve
 1568 a better understanding of detector and reconstruction effects. This will be then
 1569 used to study systematic uncertainties (see Sec. 4.5.5). The study is performed by
 1570 analysing simulated events and comparing generated and reconstructed quantities.
 1571 Figure 4.2 shows plots of the difference between true and measured angular observ-
 1572 ables ($\cos \theta_\ell$ and $\cos \theta_h$) as a function of the observable itself. These are centred at
 1573 zero indicating no bias in the measurement. Figure 4.3 shows the angular resolution
 1574 in two-dimensional bins of q^2 and angular observables. In Fig. 4.2 the same differ-
 1575 ence is shown also as a function of q^2 revealing again no bias. The spread of these
 1576 distributions around the central value can be taken as an estimate of the angular

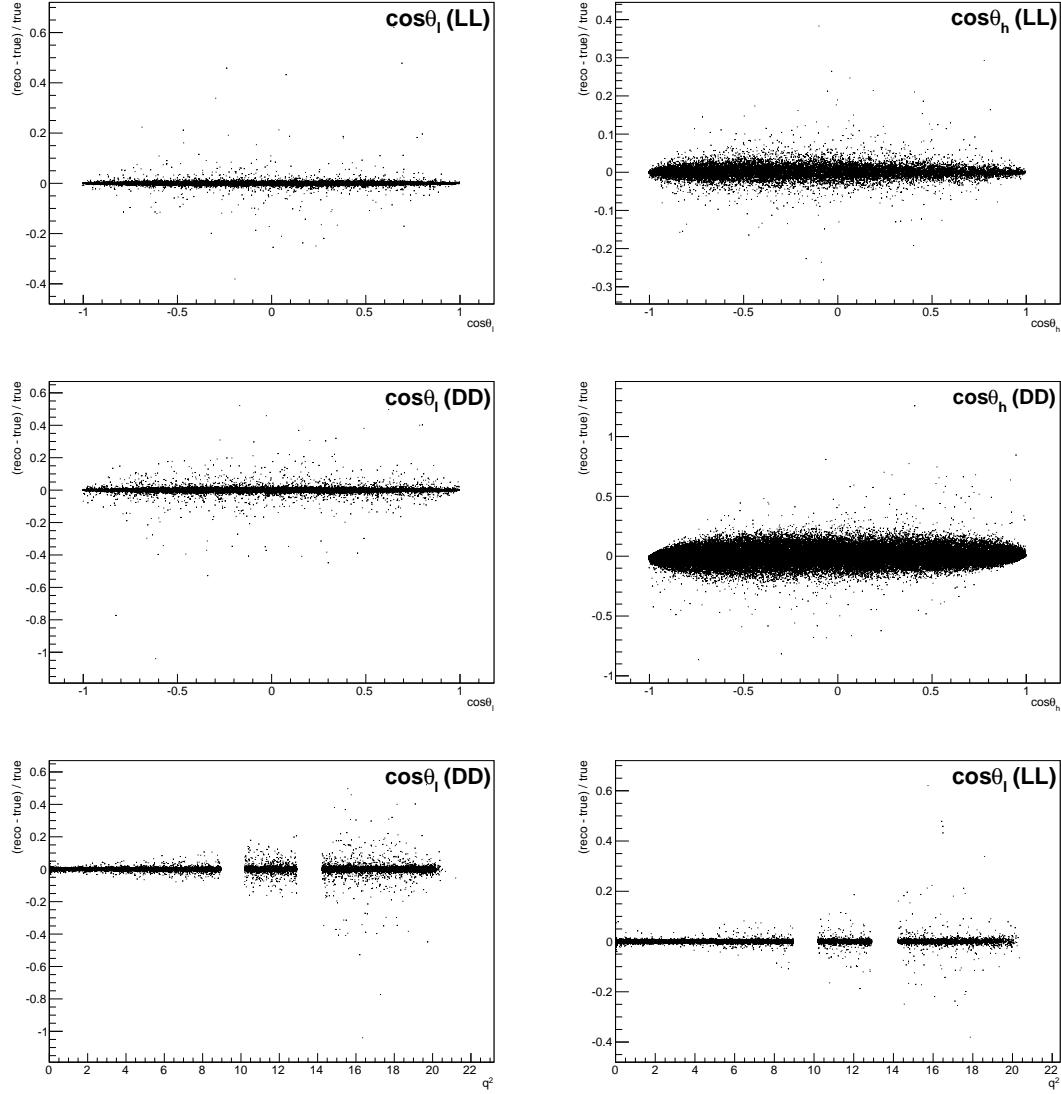


Figure 4.2: Difference of between generated and reconstructed angular observables as a function of the observables themselves for long (top) and downstream (bottom) candidates and as a function of q^2 for long (bottom left) and downstream (bottom right) candidates. As the plots are made using fully selected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

Table 4.2: Average angular resolutions integrated over the full interval and the full available q^2 .

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

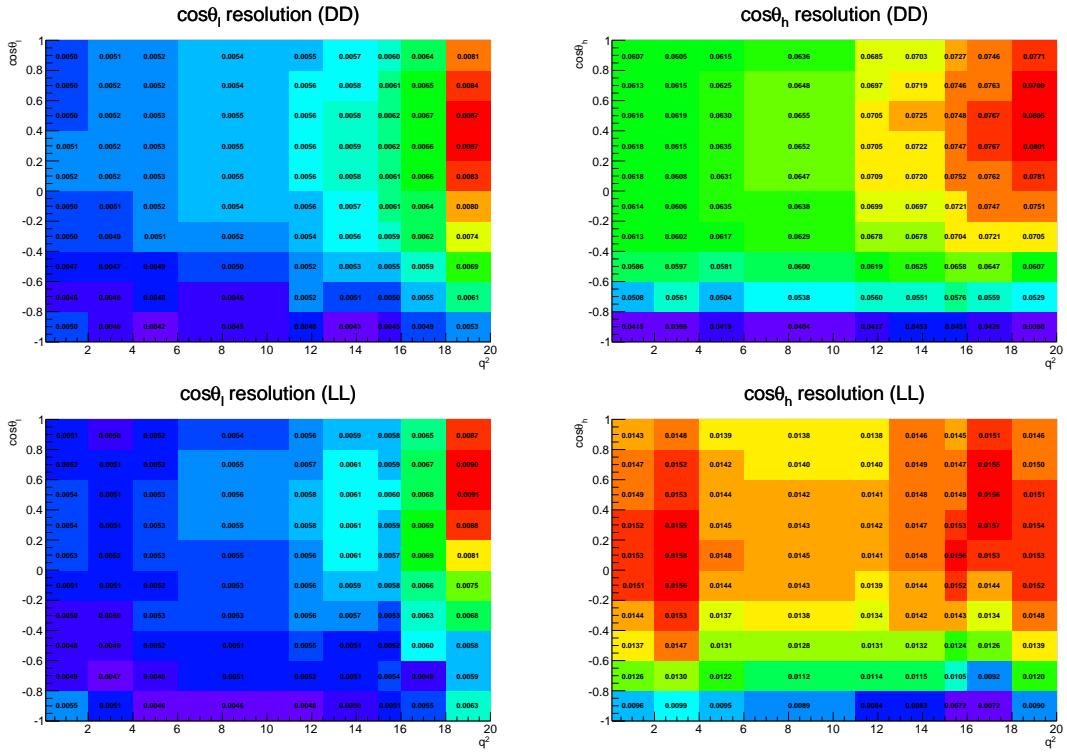


Figure 4.3: Angular resolution for $\cos \theta_\ell$ (left plots) and $\cos \theta_h$ (right plots) as a function of the angular observables and q^2 for downstream (upper plots) and long (lower plots) candidates. White bands correspond to the J/ψ and $\psi(2S)$ resonances which are excluded from the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

resolution. Taking vertical slices of the plots in Fig. 4.2 one obtains approximately gaussian distributions centred at zero. These are fit with a single gaussian and its width is interpreted as the angular resolution. Table 4.2 reports the average resolutions for the two angular observables separately for long and downstream candidates. As expected candidates built from long tracks are characterised by a better angular resolution due to a better momentum and vertex resolution.

1583 **4.4 Fit strategy**

1584 There are physical limits to the values of the parameters of interests: A_{FB}^h is limited
1585 in the $[-0.5, 0.5]$ interval and for the f_L and A_{FB}^ℓ parameters the physical region, given
1586 by $|A_{\text{FB}}^\ell| < 3/4(f_L - 1)$, is the triangle shown in Fig. 4.4. If the measured value is
1587 close to the border the fit does not always converge. Therefore a “brute force” fitting
1588 technique is applied. For this purpose fit parameters are divided into two categories:
1589 parameters of interest (PoIs), A_{FB}^ℓ , A_{FB}^h and f_L and all other parameters, which are
1590 referred to as “nuisances”. The value of the Log-Likelihood ($\log \mathcal{L}$) of the fit model
1591 with respect to data is evaluated in a grid of points in the PoIs allowed area to find
1592 the function minimum. A first coarse scan finds a candidate minimum and then the
1593 procedure is reiterated two more times in finer intervals around it. For each point
1594 all the nuisances are fitted using a maximum likelihood fit. Using this method the
1595 best fit point is therefore constrained inside the physical region. If the minimum
1596 of the log-likelihood is found to be outside it, the closest point on the boundary is
1597 chosen as the best fit.

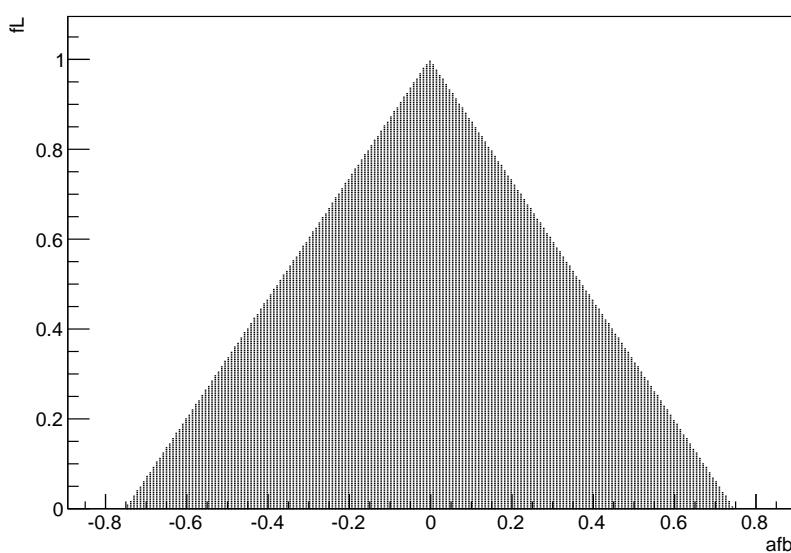


Figure 4.4: The physical $(A_{\text{FB}}^\ell, f_L)$ parameter space. The dark region corresponds to points where the PDF is positive in the whole $[-1, 1]$ interval.

1598 4.4.1 Feldman-cousins plug-in method

1599 Physical boundaries of the parameter space could result in a wrong estimation of the
 1600 uncertainties, especially if the measured value is close to the border. To deal with
 1601 this effect in this analysis the likelihood-ordering method [96] is used to estimate un-
 1602 certainties and nuisance parameters are accounted for using the plug-in method [97].
 1603 This is a unified method to calculate confidence intervals and upper/lower limits,
 1604 based on simulated experiments and has the advantage of having a well defined
 1605 frequentist coverage.

1606 The method is constituted by the following steps:

- 1607 1. fit real data distributions with all parameters free;
- 1608 2. fit real data fixing the PoIs to a value of choice while keeping nuisance param-
eters free;
- 1609 3. generate simulated samples following the distribution given by the fit model,
where all nuisance parameters are taken from the fit in point 2 and PoIs are
fixed to the same value used in point 2;
- 1610 4. repeat the two fits made on data (points 1 and 2) on each simulated sample:
fit with all parameters free and with fixed PoIs;
- 1611 5. extract the value of the Log-Likelihoods at the minimum for all cases;
- 1612 6. calculate the percentage of simulated experiments in which the free-to-fixed
likelihood ratio is bigger than in data: $\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free} > (\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free})_{data}$;
- 1613 7. repeat the procedure for many values of the PoIs scanning around the best fit
point.

1620 The confidence interval at $k\%$ is given by the points where the free-to-fixed likelihood
 1621 ratio is bigger in data than simulation for $(1 - k)\%$ of times. As an example, Fig. 4.5

1622 shows the p-values obtained with the plug-in method for A_{FB}^h and f_L . A two-
 1623 dimensional region can also be scanned giving a grid of p-values, which translates
 into two-dimensional confidence regions.

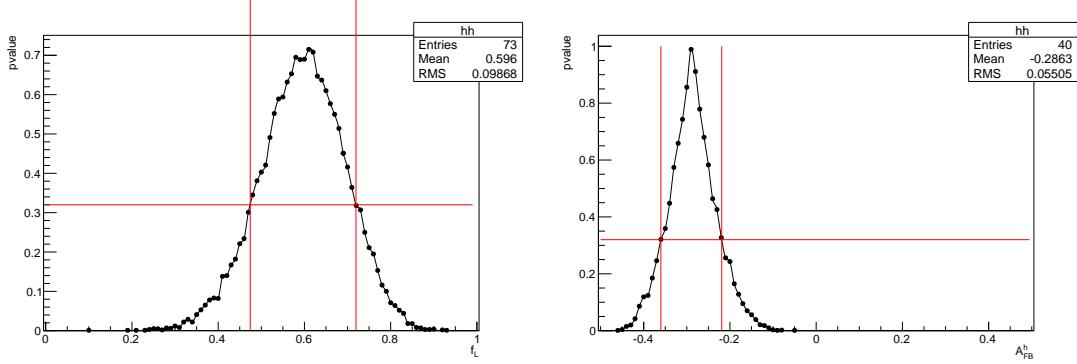


Figure 4.5: Dependence of the p-value from the values of the angular observables f_L (left) and A_{FB}^h (right) in simulated experiments. The red lines mark the points at p-value 32% corresponding to a 68% CL.

1624

1625 4.4.2 Modelling the angular distributions

1626 The observables are extracted from fits to one-dimensional angular distributions.

1627 The PDFs used to model the data are defined as

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

1628 where $k = (\text{LL}, \text{DD})$. The signal function is composed by a theoretical shape (P_S)
 1629 given by Eq. 4.10 and 4.9, which is multiplied by an acceptance function ε described
 1630 in Sec. 4.4.3. The background function, P_B , is parameterised with a linear func-
 1631 tion times the efficiency shape: $P_B^k(\cos \theta_{\ell/h}) = (cx + q) \times \varepsilon^k(\cos \theta_{\ell/h})$. The free
 1632 parameter of this model is fixed by fitting candidates in the sideband which con-
 1633 tains only background. Finally, f_b is the background fraction: $f_b = B/(S + B)$.

1634 To limit systematic effects due to the background parameterisation the fit is per-
 1635 formed in a restricted mass region around the Λ_b^0 mass peak dominated by the
 1636 signal: $5580 < m(\Lambda\mu^+\mu^-) < 5660$ MeV/ c^2 (“signal region”). The background frac-
 1637 tion, f_b , is obtained by looking at the 4-body $m(p\pi\mu\mu)$ invariant mass distribution

in a wider interval and fitting it to extract the fraction of background in the signal region. In the fit to the angular distributions this is then gaussian constrained to the obtained value. Figure 4.6 shows the background distributions in the sideband, $m(p\pi\mu^+\mu^-) > 5700 \text{ MeV}/c^2$, for the high q^2 integrated interval with overlaid the background function. Note that a different acceptance shape is used for downstream and long events and for each q^2 interval. In summary the only free fit parameter in each of the final fits to data is the forward-backward asymmetry (and f_L in the leptonic case).

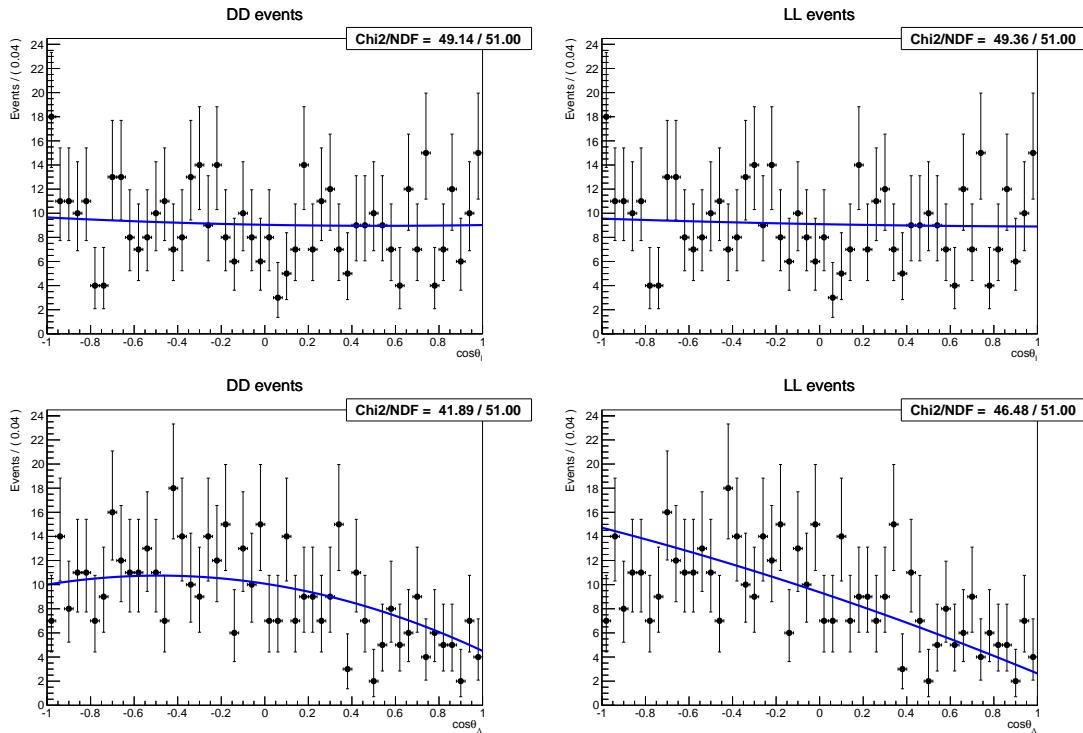


Figure 4.6: Background distribution as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15-20 \text{ GeV}^2/c^4 q^2$ interval.

1645

1646 4.4.3 Angular acceptance

1647 Selection requirements on the minimum momentum of the muons may distort the
 1648 $\cos \theta_\ell$ distribution by removing candidates with extreme values of $\cos \theta_\ell$. Similarly,
 1649 the impact parameter requirements affect $\cos \theta_h$ as very forward hadrons tend to
 1650 have smaller impact parameter values. While in principle one could take this into

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

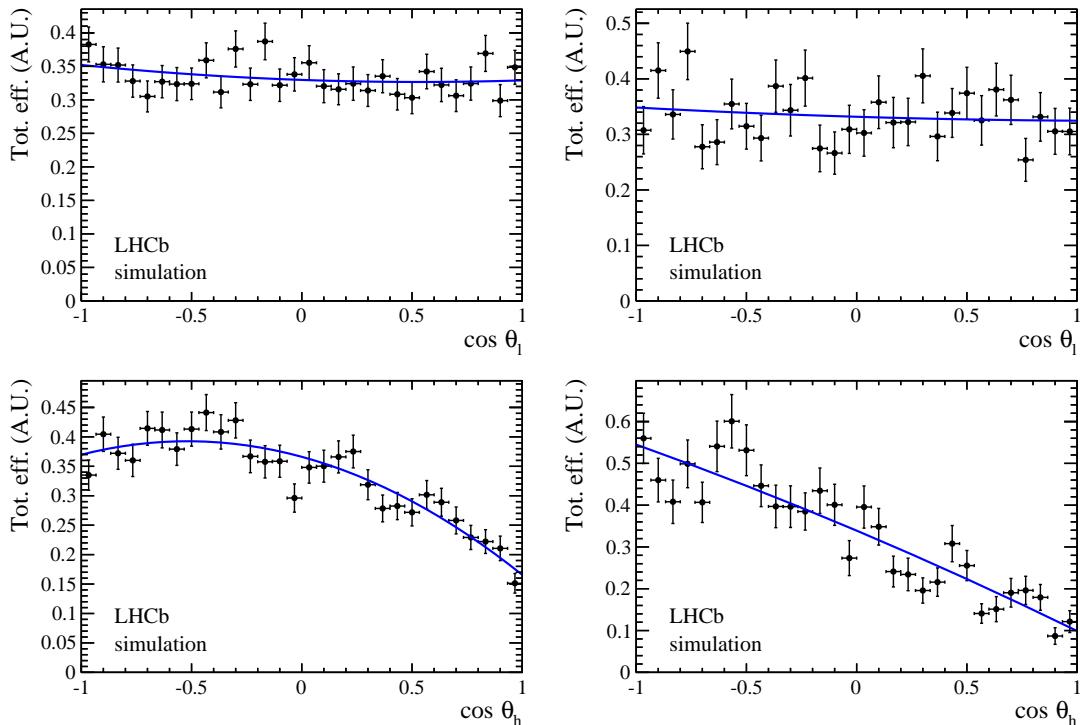


Figure 4.7: Efficiency as a function of $\cos\theta_\ell$ (top) and $\cos\theta_h$ (bottom) for downstream (left) and long (right) candidates in the 15–20 GeV^2/c^4 q^2 interval.

1662 4.4.4 Studies on a three-dimensional fit

1663 One other way of extracting the angular observables would be to fit at the same
 1664 time both angles and also the invariant mass distribution in order to have a better
 1665 handle on the level of background. In this case one can use more of the information
 1666 available. On the other hand it is necessary to use a larger mass window including
 1667 more background and this method involves more parameters to fit. In the 1D case
 1668 the free parameters are the two parameters of interest (A_{FB}^{ℓ} and f_L) for the lepton
 1669 case and one (A_{FB}^h) for the hadron one. For the 3D case in addition to the three
 1670 PoIs there are two background fractions and the two exponential slopes for the
 1671 invariant mass background. Furthermore, to take correctly into account correlations
 1672 three more observables enter the fit (see Eq. 4.12). As an high number of free
 1673 parameters is difficult to constrain with the very limited statistics available, pseudo-
 1674 experiments are used to check which method gives the best sensitivity. Events are
 1675 generated in a 3D $(\cos \theta_\ell, \cos \theta_h, m_{p\pi\mu\mu})$ space. The generated values of the PoIs
 1676 are $A_{\text{FB}}^{\ell} = 0$, $f_L = 0.7$ and $A_{\text{FB}}^h = -0.37$, which are data-like values inspired to
 1677 a preliminary measurement in the highest statistics interval. The overall statistics
 1678 and the fraction of background events in the mass window are generated to be data-
 1679 like using information from the preliminary fit to data. Each pseudo-experiment is
 1680 fitted with both methods and Fig. 4.8 reports distributions of parameters of interest
 1681 obtained from the fit in the 1D and 3D cases. The RMS of these distributions can
 1682 be taken as a measure of the sensitivity of each method. Table 4.3 lists the RMSs
 1683 obtained from both methods; for all parameters of interest the 1D fit method gives
 a smaller RMS, hence a better sensitivity.

Table 4.3: RMS values for toy experiments on the extraction of the three parameters of interest with the 1D or 3D fitting methods.

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

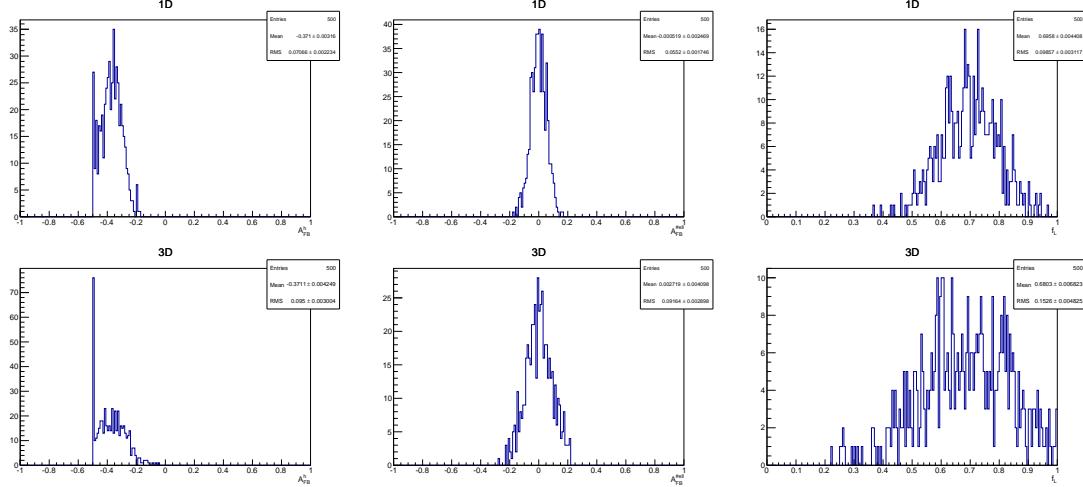


Figure 4.8: Distribution of observed parameters of interest over 500 pseudo-experiments obtained using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what is observed in the 15–20 q^2 ninterval.

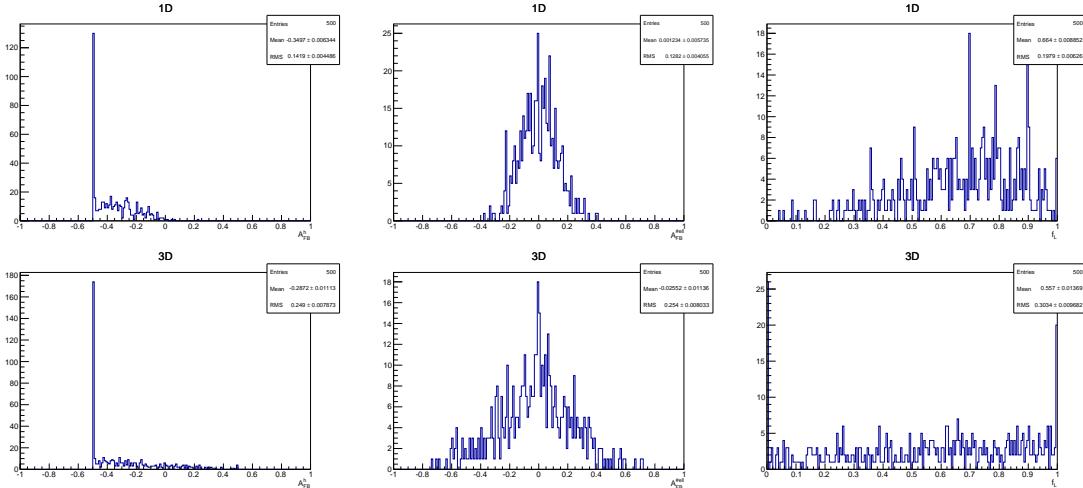


Figure 4.9: Distribution of observed parameters of interest over 500 pseudo-experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the 11–12.5 q^2 interval.

1685 4.5 Systematics uncertainties on angular observables

1686 The following section describes the five main sources of systematic uncertainties
 1687 that are considered for the angular observables measurement and, finally, results
 1688 are reported in Sec. 4.7. Results are derived only for q^2 intervals where the signal
 1689 significance, shown in Tab. 3.6, is above 3 standard deviations. This includes all

₁₆₉₀ q^2 intervals above the J/ψ resonance and the lowest q^2 interval, where an increased
₁₆₉₁ yield is due to the presence of the photon pole.

₁₆₉₂ 4.5.1 Angular correlations

₁₆₉₃ The angular acceptance is non-flat as a function of $\cos \theta_\ell$ and $\cos \theta_h$. Therefore, while
₁₆₉₄ integrating the full angular distribution, terms that cancel with perfect efficiency
₁₆₉₅ may remain and generate a bias in the final result. In order to deal with this effect
₁₆₉₆ simulated events are generated in a two-dimensional $(\cos \theta_\ell, \cos \theta_h)$ space according to
₁₆₉₇ the theoretical distribution described by Eq. 4.15 multiplied by the two-dimensional
₁₆₉₈ efficiency function obtained from simulation and reported in Fig. 4.10. Then, one-
₁₆₉₉ dimensional projections are taken and fit using the default one-dimensional efficiency
₁₇₀₀ functions. Figure 4.11 shows the distribution of observed deviations from the gener-
₁₇₀₁ ated value, $\Delta x = x_{true} - x_{measured}$. Since the mean of these distributions is non-zero
₁₇₀₂ by more than 3σ , they are taken as a systematic uncertainties.

₁₇₀₃ 4.5.2 Resolution

₁₇₀₄ The angular resolution could bias the observables measurement generating an asym-
₁₇₀₅ metric migration of events. This is especially important in the $\cos \theta_h$ case, because
₁₇₀₆ it has worse resolution and a considerably asymmetric distribution. Simulated ex-
₁₇₀₇ periments are used to asses this systematic. Events are generated according to
₁₇₀₈ the measured distributions including efficiencies. The generated events are then
₁₇₀₉ smeared by the angular resolution (gaussian smearing). To be conservative the case
₁₇₁₀ with biggest angular resolution, downstream candidates, is always used. Finally,
₁₇₁₁ the smeared and not-smeared distributions are fit with the same PDF. The average
₁₇₁₂ deviation from the default values are reported in Tab. 4.4 as a function of q^2 and
₁₇₁₃ assigned as systematic uncertainties.

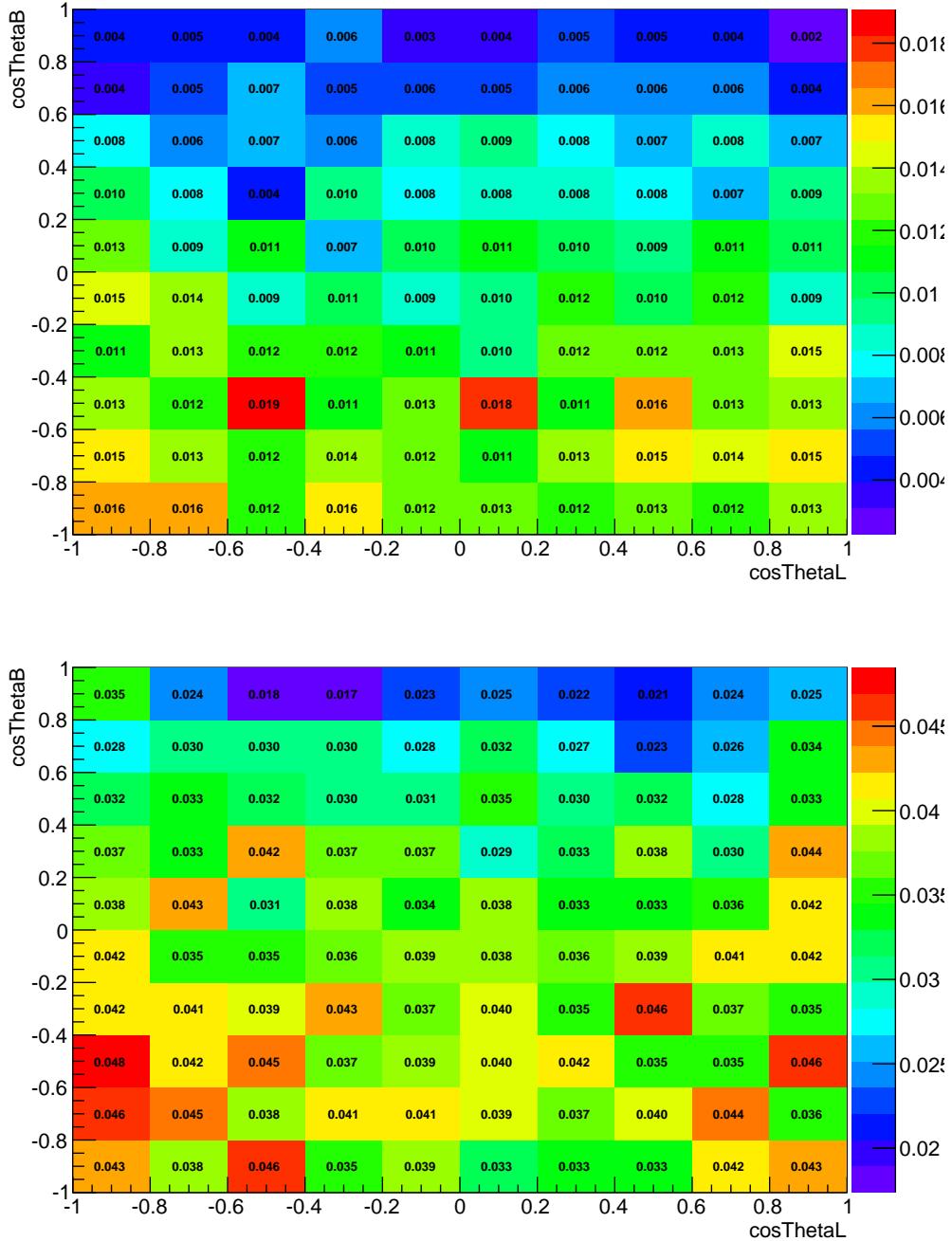


Figure 4.10: Angular acceptance as a function of $\cos \theta_\ell$ and $\cos \theta_h$ for long (top) and downstream (bottom) candidates, integrated over the full available q^2 range.

4.5.3 Efficiency description

¹⁷¹⁴ An imprecise determination of the reconstruction and selection efficiency can in-

¹⁷¹⁵ troduce an extra oddity and therefore bias the measurement. To asses this effect

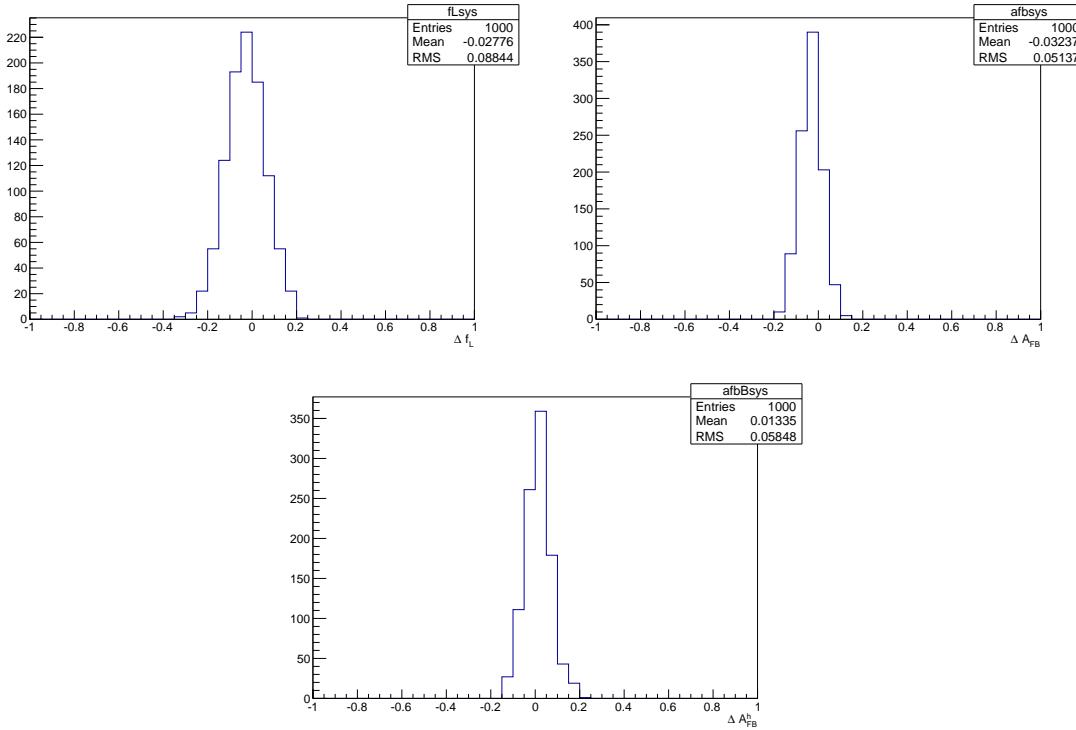


Figure 4.11: Deviations of the observables' values obtained fitting simulated events generated with a 2D distribution multiplied by a 2D efficiency and fitting 1D projections with respect to generated values. For f_L (top left), A_{FB}^ℓ (top right) and A_{FB}^h (bottom).

the kinematic re-weighting described in Sec. 3.3.2 is removed from the simulation and the efficiency is determined again. Simulated events are then fit using the same theoretical PDF but multiplied by the efficiency function obtained with and without kinematical weights. As in the previous cases the average bias is taken as systematic uncertainty; results are shown in Tab. 4.5. Furthermore, the effect of the limited simulated statistics is taken into account and added to the systematic uncertainty.

4.5.4 Background parameterisation

There is a certain degree of arbitrariness in the choice of a parameterisation for the background, especially for q^2 intervals with low statistics. To assess possible biases due to the PDF choice, simulated experiments are generated using the shapes from data fits and the same statistics as observed in data for each q^2 interval. Each

Table 4.4: Values of simulated $\cos \theta_\ell$ and $\cos \theta_\Lambda$ resolutions and systematic uncertainties on angular observables due to the finite resolution in bins of q^2 .

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

Table 4.5: Values systematic uncertainties due to limited knowledge of the efficiency function on the three angular observables in bins of q^2

q^2 [GeV $^2/c^4$]	A_{FB}^h	A_{FB}^ℓ	f_L
0.1–2.0	0.0093	0.0020	0.0440
11.0–12.5	0.0069	0.0069	0.0027
15.0–16.0	0.0109	0.0018	0.0046
16.0–18.0	0.0159	0.0012	0.0043
18.0–20.0	0.0148	0.0030	0.0017
15.0–20.0	0.0138	0.0002	0.0046

¹⁷²⁸ pseudo-experiment is fit with two models: the default one, a “line times efficiency”
¹⁷²⁹ function and the efficiency function alone, corresponding to the assumption that
¹⁷³⁰ background distributions are originally flat and only modified by the interaction
¹⁷³¹ with the detector. The average bias with respect to the default model is taken as
¹⁷³² systematic uncertainty; results are reported in Tab. 4.7.

¹⁷³³

¹⁷³⁴ 4.5.5 Polarisation

¹⁷³⁵ To study the effect of a non-zero Λ_b^0 production polarisation simulated events are
¹⁷³⁶ generated using the distributions given by Eqs. 4.12 and 4.14 as a function of the
¹⁷³⁷ angle under study ($\cos \theta_\ell$ or $\cos \theta_h$) and $\cos \theta$, defined in Sec. 4.2, which is sensitive to
¹⁷³⁸ polarisation. Similarly to the procedure used for the branching ratio measurement,
¹⁷³⁹ events are generated using values of the polarisation corresponding to $\pm \sigma$ from the
¹⁷⁴⁰ LHCb measurement [88]. In the theoretical functions $\cos \theta$ is always odd therefore

Table 4.6: 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

Table 4.7: Values of systematic uncertainties due to the choice of background parameterisation in bins of q^2 .

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

with perfect efficiency it always drops out by integrating over $\cos\theta$. Therefore the generated distributions are also multiplied by the two-dimensional efficiency function. No significant bias is found.

4.6 J/ψ cross-check

To cross-check the fitting procedure this is applied on the high statistics $\Lambda_b^0 \rightarrow J/\psi \Lambda$ sample. For this purpose events are selected with an additional requirement on the proton PID, $\text{PID}_p > 10$. This is needed to reduce the $B^0 \rightarrow K_s^0 J/\psi$ background, which is particularly important for the hadronic side fit, since the K_s^0 events are not distributed in a flat way in the $\cos\theta_h$ variable and would therefore bias the fit. Figure 4.12 shows the invariant mass distributions after this requirement is applied, which can be compared with the ones in Fig. 3.13. After the PID cut there are 0.2% of K_s^0 events left in the downstream sample and a fraction compatible with zero in the long sample. The signal model used for this fit is the same used for the rare case

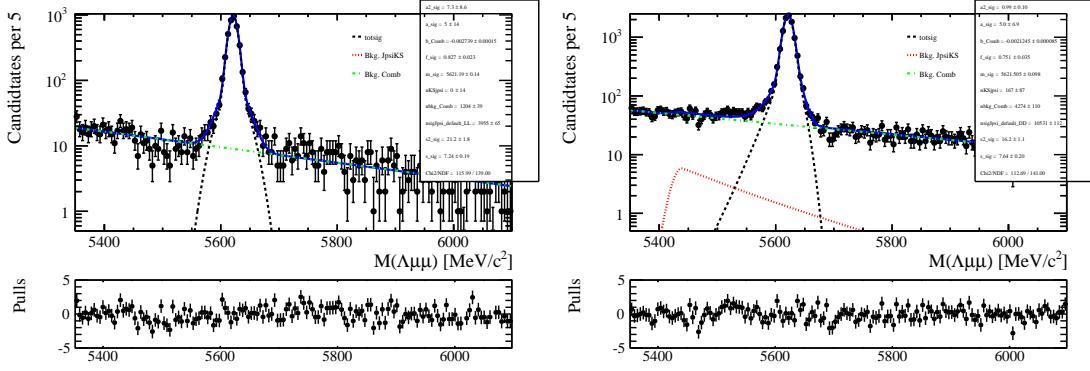


Figure 4.12: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ long (left) and downstream (right) candidates with an extra proton PID cut to remove K_s^0 background.

and described in Sec. 4.4.2. For the background instead the higher statistics allows to leave more freedom to the fit. Therefore a second-order Chebyschev polynomial is used, where the two parameters are free to vary. As for the rare case the background fractions are gaussian-constrained to what found from the invariant mass fit. Figures 4.13 and 4.14 show fitted angular distributions for the J/ψ channel. The measured values of the observables are $A_{FB}^\ell = -0.002^{+0.011}_{-0.011}$, $A_{FB}^h = -0.402^{+0.010}_{-0.009}$ and $f_L = 0.485^{+0.019}_{-0.020}$, where the uncertainties are 68% Feldman Cousins confidence intervals. The lepton side asymmetry as expected is measured to be zero.

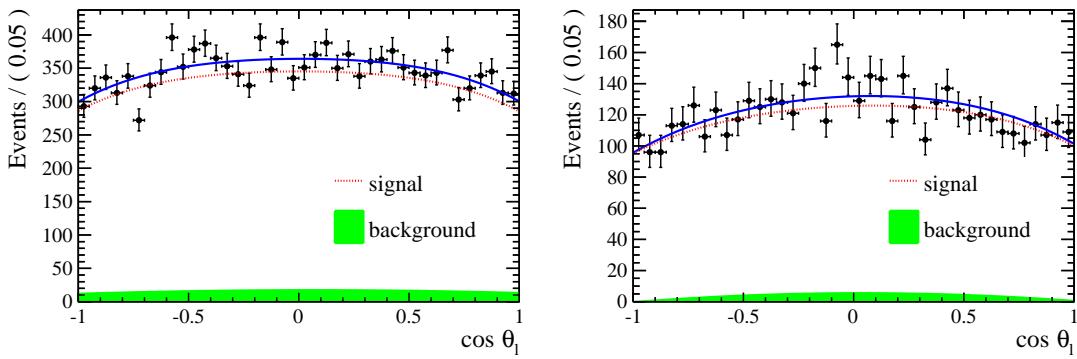


Figure 4.13: Fitted angular distribution as a function of $\cos \theta_\ell$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

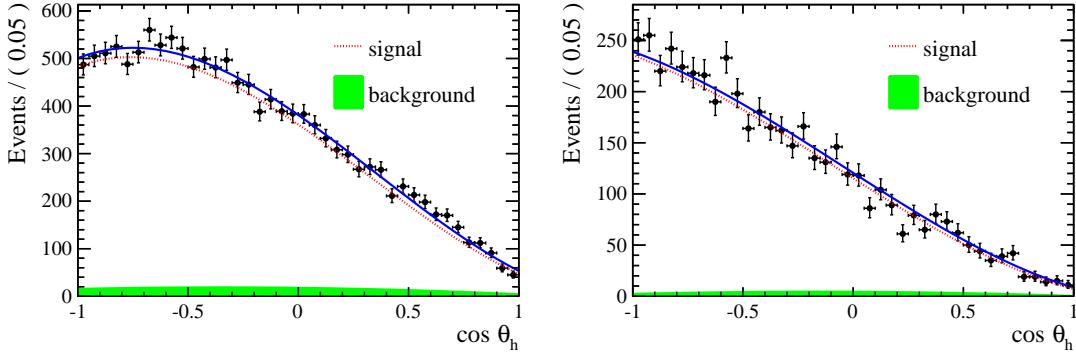


Figure 4.14: Fitted angular distribution as a function of $\cos \theta_h$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

4.7 Results

Figures 4.15 and 4.16 show fits to the angular distributions for the 15-20 $\text{GeV}^2/c^4 q^2$ interval and Tab. 4.8 reports measured values of A_{FB}^ℓ , A_{FB}^h and f_L . The asymmetries are also shown in Fig. 4.17 together with SM predictions obtained from Ref. [91]. The statistical uncertainties on these tables are obtained using the likelihood-ratio ordering method described in Sec. 4.4.1, where only one of the two observables is treated as the PoI at a time. The statistical uncertainties on A_{FB}^ℓ and f_L are also reported in Fig. 4.18 as two-dimensional 68 % confidence level (CL) regions, where the likelihood-ratio ordering method is applied by varying both observables and therefore taking correlations into account. Total systematic uncertainties correspond to the square root sum of the single considered sources.

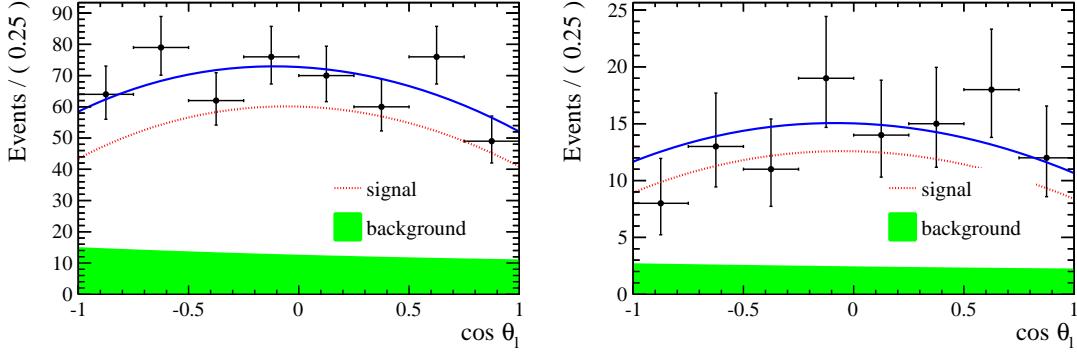


Figure 4.15: Fitted angular distributions as a function of $\cos \theta_\ell$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

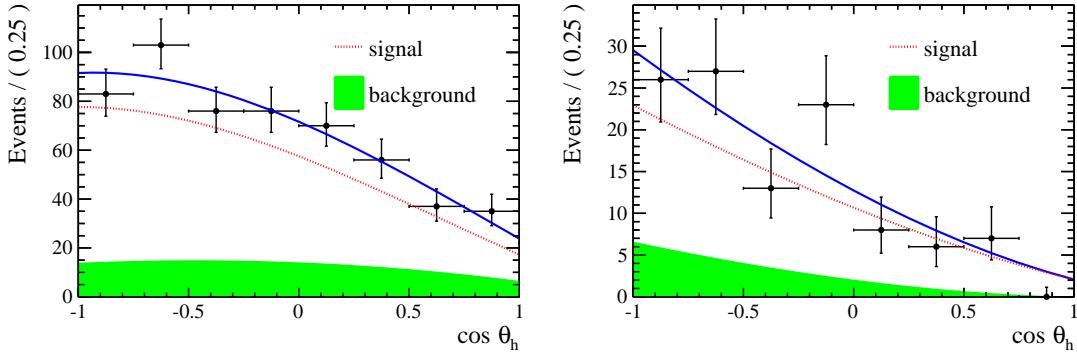


Figure 4.16: Fitted angular distributions as a function of $\cos \theta_h$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

Table 4.8: 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	$-0.05^{+0.09}_{-0.09} \pm 0.03$	$0.61^{+0.11}_{-0.14} \pm 0.03$	$-0.29^{+0.07}_{-0.07} \pm 0.03$

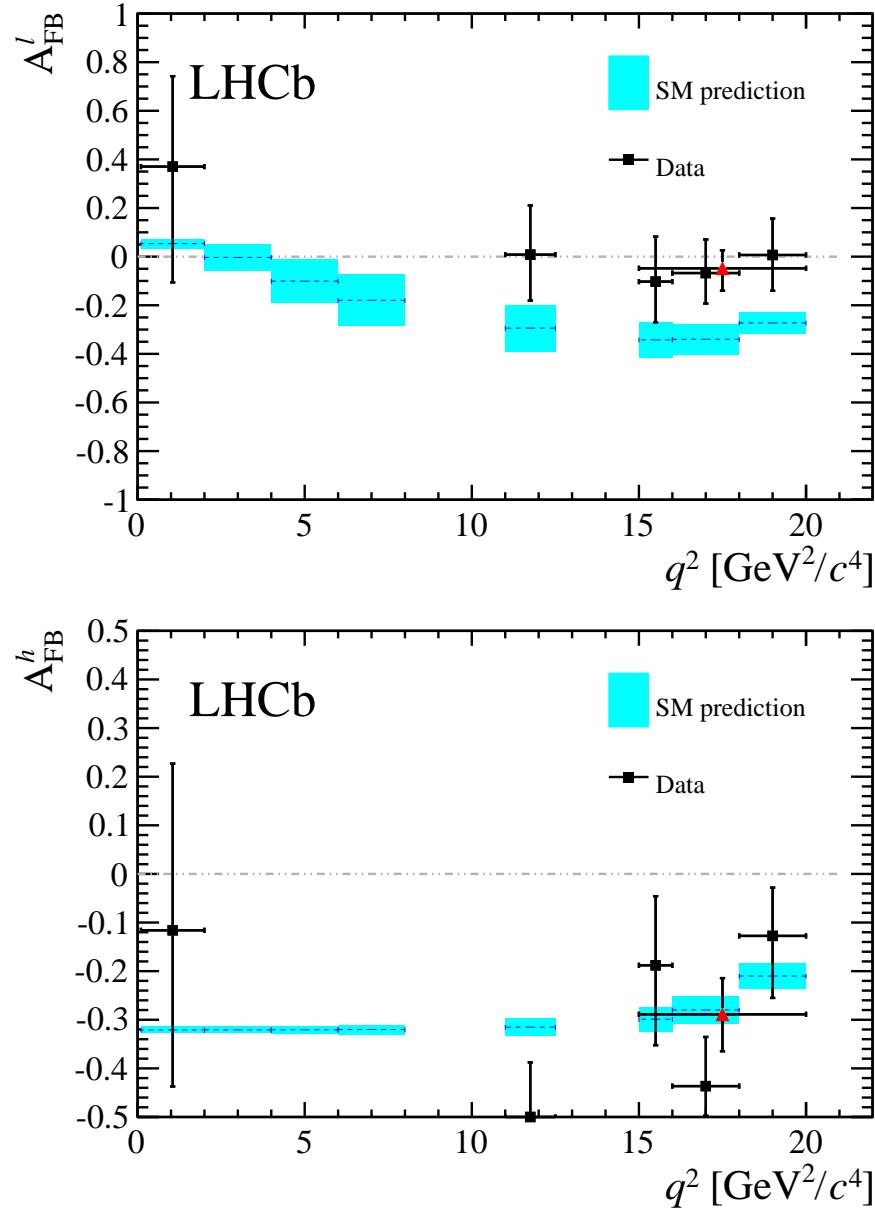


Figure 4.17: Measured values of the leptonic (top) and the hadronic (bottom) forward-backward asymmetries in bins of q^2 . Data points are only shown for q^2 intervals where a statistically significant signal yield is found, see text for details. The (red) triangle represents the values for the $15 < q^2 < 20 \text{ GeV}^2/c^4$ interval. Standard Model predictions are obtained from Ref. [98].

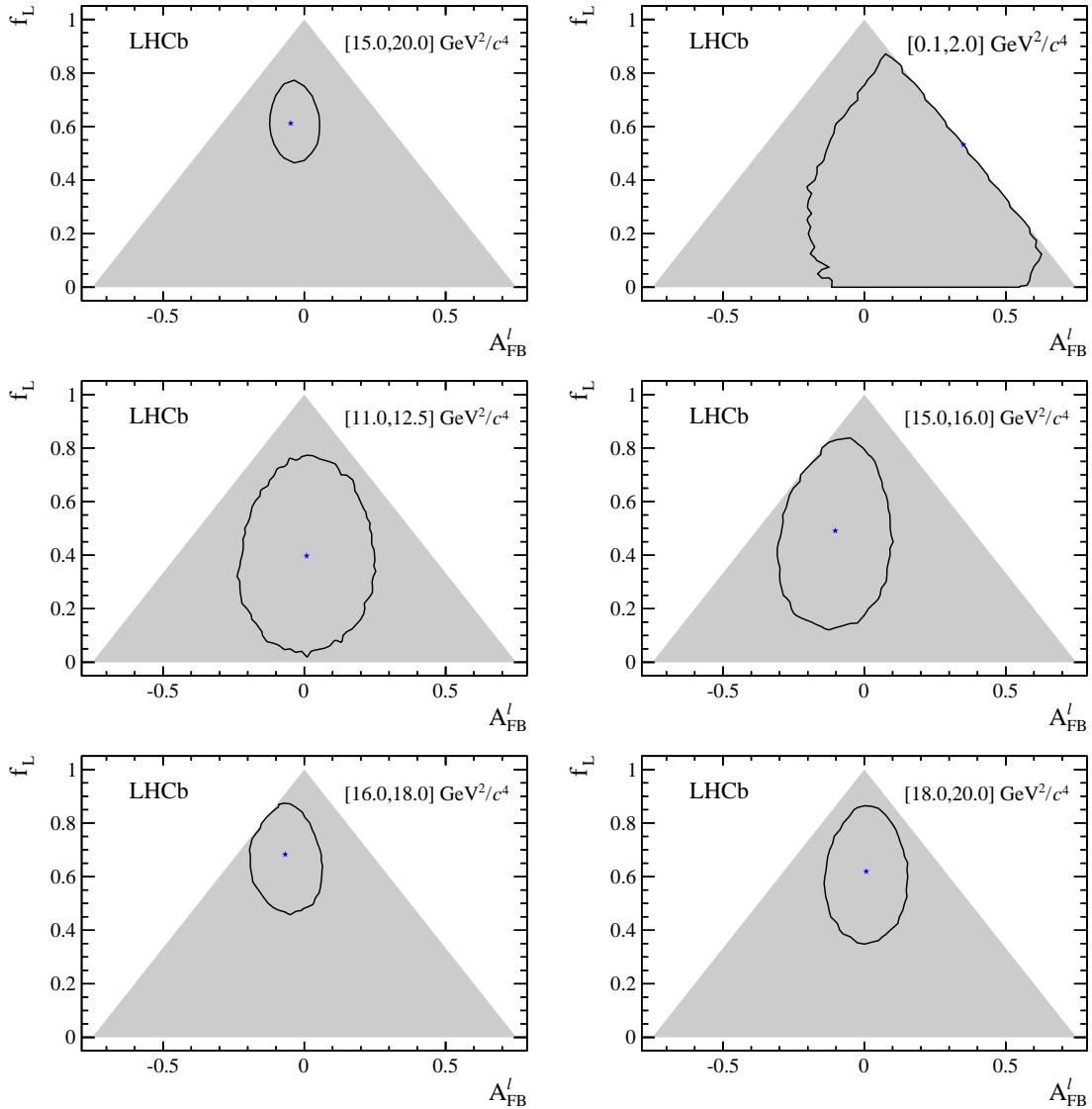


Figure 4.18: Two-dimensional 68 % CL regions (black) as a function of A_{FB}^ℓ and f_L . The shaded areas represent the regions in which the PDF is positive over the complete $\cos \theta_\ell$ range. The best fit points are indicated by the (blue) stars.

1773

CHAPTER 5

1774

1775

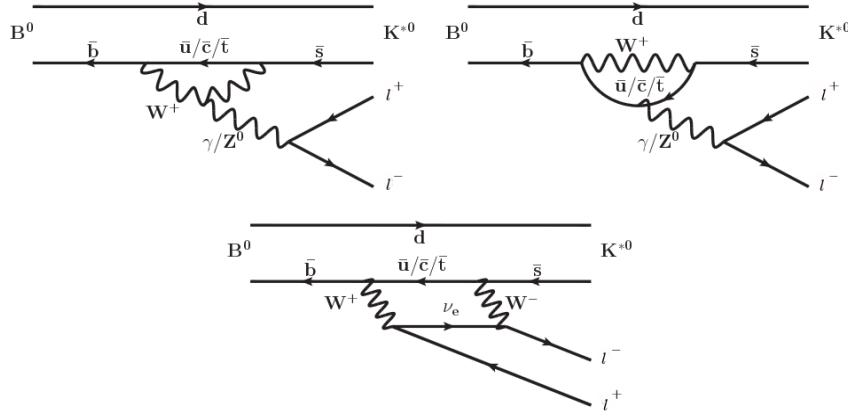
Testing lepton flavour universality with $R_{K^{*0}}$

1776

1777 Lepton Favour Universality (LFU) is the equality of the weak coupling constants for
1778 all leptons. FCNCs processes, forbidden in the SM at tree level and happening only
1779 at loop level, are an ideal laboratory for studying LFU as NP in the loops could
1780 break the flavour symmetry.

1781 In this work, decays of $b \rightarrow s\mu^+\mu^- (e^+e^-)$ type, are considered to test LFU between
1782 electrons and muons using penguin decays. In particular, B^0 meson semileptonic
1783 decays to $B^0 \rightarrow K^{*0}\ell^+\ell^-$ are considered. Figure 5.1 shows the possible Feynman di-
1784 agrams producing such decays while Fig. 5.2 illustrates how the Feynman diagrams
1785 of these processes may include new particles. A series of recent LHCb measure-
1786 ments [22] points to a tension with SM predictions, which make these processes very
1787 interesting to better understand the nature of the discrepancy.

1788 In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger pro-
1789 posed the measurement of the R_H ratio [99], defined in Eq. 5.1, where H can be an

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

1790 inclusive state containing an s quark (X_s) or an s -quark resonance like K or K^{*0} .

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

1791 In this quantity the decay width is integrated over the squared dilepton invari-
 1792 ant mass, q^2 , from $q_{min}^2 = 4m_\mu^2$, which is the threshold for the $\mu\mu$ process, up to
 1793 $q_{max}^2 = m_b^2$.

1794 The advantage of using ratios of branching fractions as observables is that, in the
 1795 theoretical prediction, hadronic uncertainties cancel out. Furthermore, experimen-
 1796 tally, some of the systematic uncertainties on the ratios are reduced giving a better
 1797 measurement. For example, what is measured is the number of $\mu\mu$ and ee decays
 1798 which happen in a certain period of time. Then, the luminosity, \mathcal{L} , is used to obtain
 1799 a cross section, σ , using $R = \mathcal{L}\sigma$, where R is the rate at which the decays happen.
 1800 The luminosity measurement is usually a source of systematic uncertainty, however
 1801 it appears on both sides of the ratio and therefore cancels out.

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

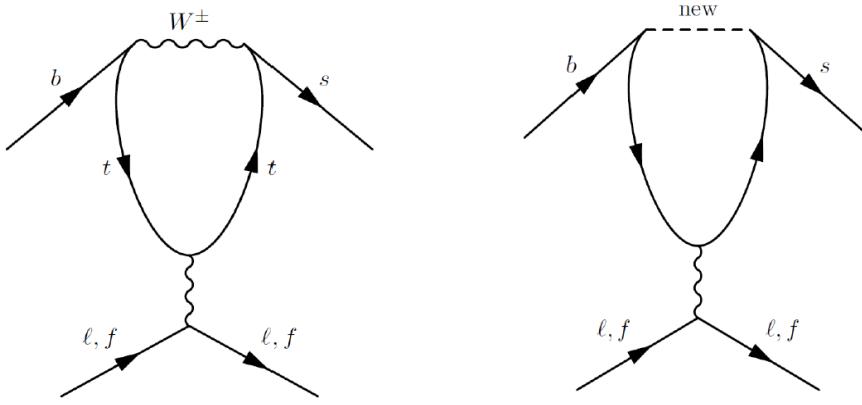


Figure 5.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

range [99]:

$$R_{X_s} = 0.987 \pm 0.006 \quad (5.2)$$

$$R_K = 1.0000 \pm 0.0001 \quad (5.3)$$

$$R_{K^{*0}} = 0.991 \pm 0.002 \quad (5.4)$$

$$(5.5)$$

1802 under the assumptions that:

- 1803 • right-handed currents are negligible;
- 1804 • (pseudo-)scalar couplings are proportional to the lepton mass;
- 1805 • there are no CP-violating phases beyond the SM.

1806 The measurement of the R_H ratios is of particular interest after the recent measure-
1807 ment of the branching ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay [33], where no evidence of NP
1808 was found. In fact the quantities $(R_H - 1)$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ remain proportional
1809 with

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

1810 A joint measurement of this two quantities can give much information and constrain
1811 MFV models. If $R_H = 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ is close to the SM prediction as it is

measured to be this will allow to put strong constraints on extensions of the SM. If instead $R_H > 1$ and the equation above is not verified, this would mean that one of the assumptions listed above are not verified, which can happen in some extensions of the SM as Super-Symmetric models with broken R-parity. A series of recent LHCb measurements [22] shows tensions with SM predictions, which makes it interesting to further investigate these processes.

5.1 Combining ratios

The full power of the R_H ratios in understanding new physics scenarios comes from their combinations. In Ref. [100] Hiller and Schmaltz propose the measurement of the double ratios, $X_H = R_H/R_K$, which not only can test LFU but also allow to disentangle the kind of new physics that lies behind. These ratios are in fact sensitive to FCNCs of right-handed currents. Furthermore, in Ref. [100] the study is extended to B_s^0 decays such as $B_s^0 \rightarrow \phi\ell^+\ell^-$ or $B_s^0 \rightarrow \eta\ell^+\ell^-$.

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

$$\begin{aligned} C + C' : & K, K_{\perp}^*, \dots \\ C - C' : & K_0(1430), K_{0,\parallel}^*, \dots \end{aligned} \tag{5.7}$$

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

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

1832 where the Δ_{\pm} quantities are combinations of Wilson coefficients described in Eq. 10
1833 of Ref. [100] and the parameter p is the polarisation of K^* that in Ref. [100] is
1834 determined to be close to 1 simplifying the formula to $R_{K^*} \simeq 1 + \Delta_-$. In particular
1835 one can observe the following correlations:

- 1836 • $R_K < 1$, as it is measured to be, and $X_{K^*} > 1$ points to dominant BSM
1837 contributions into C_{LR} (see definition in Sec. 1.5.2);
- 1838 • a SM like $R_K \sim 1$ together with $X_{K^*} \neq 1$ requires BSM with $C_{LL} + C_{RL} \simeq 0$;
- 1839 • $R_K \neq 1$ and $X_{K^*} \simeq 1$ corresponds to new physics in C_{LL} .

1840 5.2 Experimental status

1841 The R_K and $R_{K^{*0}}$ ratios have already been measured at the B-factories [101, 102],
1842 and the R_K ratio has been also recently measured at LHCb [103] in the $1 < q^2 <$
1843 $6 \text{ GeV}^2/c^4$ q^2 interval, which represents the most precise measurement to date. This
1844 measurement manifests a 2.6σ deviation from the SM prediction. The current exper-
1845 imental status is summarised in Tab. 5.1. By profiting of the large dataset collected
1846 during Run-I, the LHCb experiment is expected to reduce the uncertainty on $R_{K^{*0}}$
1847 by at least a factor of 2 with respect to the B-factories.

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

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

¹⁸⁴⁹ **5.3 Analysis strategy**

¹⁸⁵⁰ The aim of this analysis is to measure the $R_{K^{*0}}$ ratio using pp collision data collected
¹⁸⁵¹ by the LHCb detector in 2011 and 2012, corresponding to a total of 3 fb^{-1} of
¹⁸⁵² integrated luminosity. The $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$, “rare channels”,
¹⁸⁵³ are reconstructed with the K^{*0} decaying into a kaon and a pion with opposite
¹⁸⁵⁴ charges.

¹⁸⁵⁵ The analysis has to separate signal candidates from background candidates which
¹⁸⁵⁶ have similar observed properties. The selection presented in Sec. 5.6 aims to max-
¹⁸⁵⁷ imise the yield while minimising the background contamination. Two types of back-
¹⁸⁵⁸ grounds are identified: “peaking background” and “combinatorial background”. The
¹⁸⁵⁹ first comes from the mis-reconstruction of other decays or from partially recon-
¹⁸⁶⁰ structed events. This type of background, because its specific kinematic properties,
¹⁸⁶¹ usually peaks in some variable, such as the invariant mass of all final particles,
¹⁸⁶² therefore we can remove these events by removing the peak. The combinatorial
¹⁸⁶³ background instead comes from the random combination of particles and can be
¹⁸⁶⁴ lowered selecting events with good-quality tracks and vertices.

¹⁸⁶⁵ To further reduce systematic uncertainties the measurement is performed as a dou-
¹⁸⁶⁶ ble ratio as shown in Eq. 5.9 where decays reaching the same final states as the
¹⁸⁶⁷ rare channels via a J/ψ resonance, $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$, also referred as “char-
¹⁸⁶⁸ monium” or “resonant” channels, are used as control samples. These decays are
¹⁸⁶⁹ distinguished from the rare channels using the invariant mass of the dilepton pair.

¹⁸⁷⁰ In Sec. 5.8 the efficiency of selecting and reconstructing each channel is extracted
¹⁸⁷¹ and, finally, in Sec. 5.10 the R_{K^*} ratio defined is built as the double ratio of rare
¹⁸⁷² and resonant channels:

$$R_{K^{*0}} = \frac{N_{B^0 \rightarrow K^{*0}\mu^+\mu^-}}{N_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}} \cdot \frac{N_{B^0 \rightarrow K^{*0}J/\psi e^+e^-}}{N_{B^0 \rightarrow K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow \mu^+\mu^-}}{\varepsilon_{B^0 \rightarrow K^{*0}\mu^+\mu^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^{*0}e^+e^-}}{\varepsilon_{B^0 \rightarrow K^{*0}J/\psi \rightarrow e^+e^-}} \quad (5.9)$$

1873 As NP is expected not to affect charmonium resonances the ratio of the J/ψ channels
1874 is 1 and therefore $R'_{K^{*0}} = R_{K^{*0}} \times R_{J/\psi} = R_{K^{*0}}$. On the other hand using the relative
1875 efficiencies between the rare and resonant channels allows to cancel out many effects
1876 resulting in a better control of systematic uncertainties.

1877 For brevity, the rare channels will also be denoted as “ $\ell\ell$ ”, or specifically “ ee ” and
1878 “ $\mu\mu$ ”, and the resonant channels as “ $J/\psi(\ell\ell)$ ”, or “ $J/\psi(ee)$ ” and “ $J/\psi(\mu\mu)$ ”.

1879 5.4 Choice of q^2 intervals

1880 Two q^2 intervals are considered in this work:

- 1881 • the “central- q^2 ” region, $[1.1, 6.0]$ GeV^2/c^4 ;
- 1882 • the “high- q^2 ” region, above 15 GeV^2/c^4 .

1883 The central- q^2 region is the most interesting place to look for new physics. In fact,
1884 at low q^2 , below 1 GeV^2/c^4 the photon pole dominates leaving little space for NP
1885 to be found [22]. The upper bound of this interval is set at 1.1 GeV^2/c^4 , in order
1886 entirely include the contribution from $\phi \rightarrow \ell^+ \ell^-$ decays, that can dilute new physics
1887 effects, into the low q^2 interval. The upper bound of the central interval is chosen
1888 to be sufficiently far away from the J/ψ radiative tail, where predictions cannot be
1889 cleanly extracted. The 6–15 GeV^2/c^4 region is characterised by the presence of the
1890 narrow peaks of the J/ψ and $\psi(2S)$ resonances. The lower bound of the high- q^2
1891 region, where the signal in the electron channel is still unobserved, is chosen to be
1892 sufficiently far from the $\psi(2S)$ resonance. Rare and resonant channels are selected
1893 depending on which q^2 interval they fall in (for details see Sec. 5.6).

1894 **5.5 Data samples and simulation**

1895 This analysis is based on a data set corresponding to 3 fb^{-1} of integrated luminosity
1896 collected by the LHCb detector in 2011 and 2012. In order to study the background
1897 properties, determine efficiencies and to train the multivariate analysis simulated
1898 events are used. After the hard interactions are generated with Pythia8 hadronic
1899 particles are decayed using EvtGen and, finally, propagated into the detector us-
1900 ing Geant4 and reconstructed with the same software used for data. Samples are
1901 generated with both 2011 and 2012, magnet up and down conditions and are com-
1902 bined in the right proportions, according to the luminosity registered on data. The
1903 next section describes the corrections applied to the simulation to obtain a better
1904 description of data.

1905 **5.5.1 Data-simulation corrections**

1906 Since the multivariate classifier training (see Sec. 5.6.5) and the calculation of most
1907 of the efficiency components (see Sec. 5.8) are obtained from the study of simulated
1908 events it is important to verify that the simulation is a reliable reproduction the
1909 data. In particular it is important to match data and Monte Carlo in the kinemat-
1910 ics of the final particles and the occupancy of the detector. The kinematics of the
1911 decays is characterised by the transverse momentum spectrum of the B^0 . Discrep-
1912 ancies in this distribution cause also the spectra of the final particles to differ from
1913 data and affect the efficiency determination as its value often depends on the mo-
1914 mentum distribution of final particles. The occupancy of the detector is correlated
1915 to the invariant mass shape of the signal because the addition of energy clusters in
1916 the electromagnetic calorimeter, affects the electron momenta for bremsstrahlung
1917 photons emitted before the magnet. The hit multiplicity in the SPD detector is a
1918 proxy for the detector occupancy.

1919 Since it is important that these quantities are well modelled, the simulation is
1920 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. 5.6). However, the data still includes a high level of background and distributions cannot be directly compared. The $s\mathcal{P}$ lot technique [55] is used to statistically subtract the background from data and obtain pure signal distributions using the invariant mass as control variable. Fig. 5.3 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. 5.4 and ratios of these distribution, which are used to re-weight the simulation, are reported in Fig. 5.5. The weights for the SPD multiplicity are calculated separately for 2011 and 2012 events, because distributions are significantly different in the two years. 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. 5.8.3 and, finally, ee samples are also reweighted for L0 trigger efficiency as described in Sec. 5.8.4. Weights are always applied throughout unless specified.

5.6 Selection

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

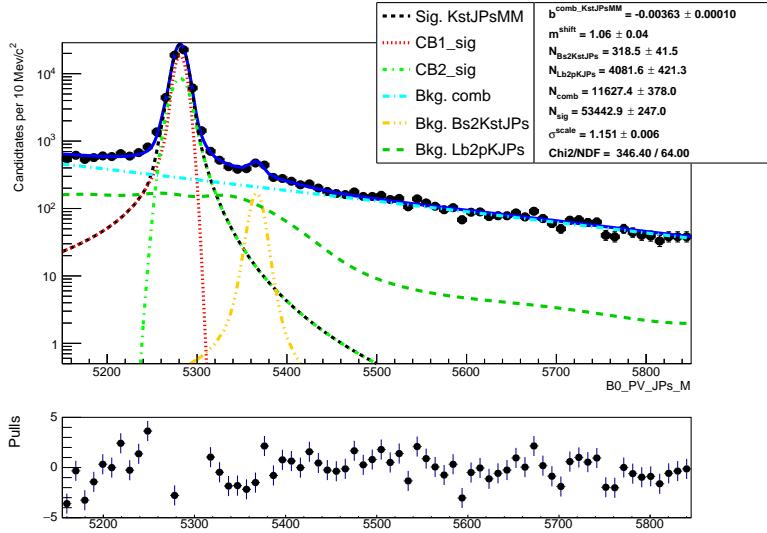


Figure 5.3: Fitted 4-body invariant mass distributions of muonic resonant candidates.

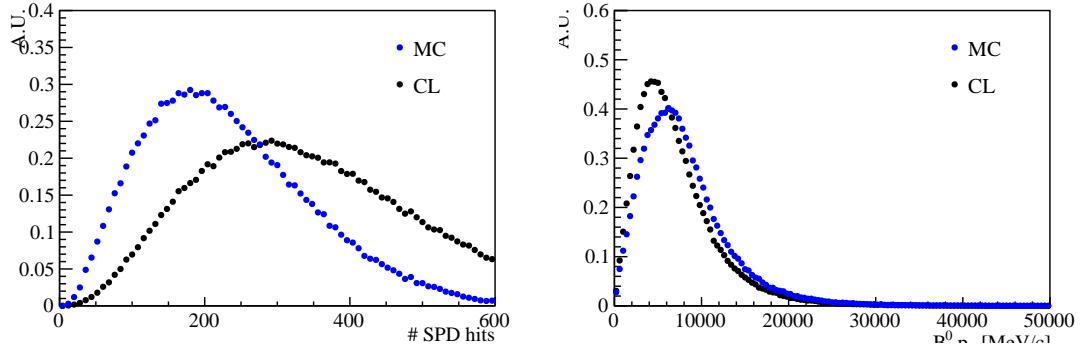


Figure 5.4: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

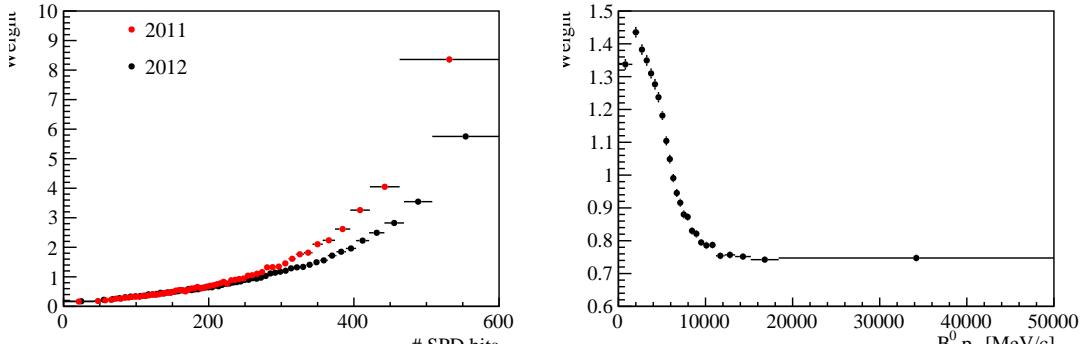


Figure 5.5: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

1949 The next step consists in the application of particle identification (PID) conditions
 1950 which remove a good part of misreconstructed background and clear the way for
 1951 the last step where a neural network is used to remove combinatorial background.
 1952 In order to minimise systematic uncertainties the same selection requirements are
 1953 applied to the rare signal candidates and on their relative charmonium channel, a
 1954 part from the q^2 cuts which serve to distinguish them. In order to identify the
 1955 $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ channel a dilepton mass interval of $100 \text{ MeV}/c^2$ around the
 1956 nominal J/ψ peak [2] is selected. For the electron resonant channel it is not possible
 1957 to use a narrow cut on the q^2 and 4-body $m(K\pi e^+e^-)$ invariant mass distributions
 1958 are characterised by a long radiative tail at low masses due to bremsstrahlung ra-
 1959 diation. Furthermore, a cut in q^2 also distorts the 4-body mass distribution at low
 1960 masses and it is important to be able to fit a wide mass range to constrain back-
 1961 grounds. For these reasons the interval to select $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates
 1962 is chosen to go as low as possible without overlapping with the rare channel interval.
 1963 The electron resonant channel is therefore selected in the interval [6,11] GeV^2/c^4 .
 1964 Figure 5.6 shows two-dimensional distributions of q^2 versus the 4-body $m(K\pi\ell^+\ell^-)$
 1965 invariant mass for events which pass the full selection. On these plots horizontal
 1966 bands can be seen at the q^2 corresponding to the J/ψ and $\psi(2S)$ resonances. On
 1967 the plot for muons it is also evident a vertical band which corresponds to rare decay
 1968 of interest.

1969 5.6.1 Trigger and Stripping

1970 Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported in
 1971 Tab. 5.2, where the logical *and* of L0, Hlt1 and Hlt2 lines is required and the logical
 1972 *or* of the lines on the same level. The candidates are required to be triggered-on-
 1973 signal (TOS) for most of the stages, namely it is required for the particle which
 1974 triggered to be one of the particles used to build the signal candidates. Only for
 1975 L0Global, used in the electron case, we require a trigger-independent-of-signal (TIS),
 1976 this is aimed to collect all the possible statistics for the electron channels, which are

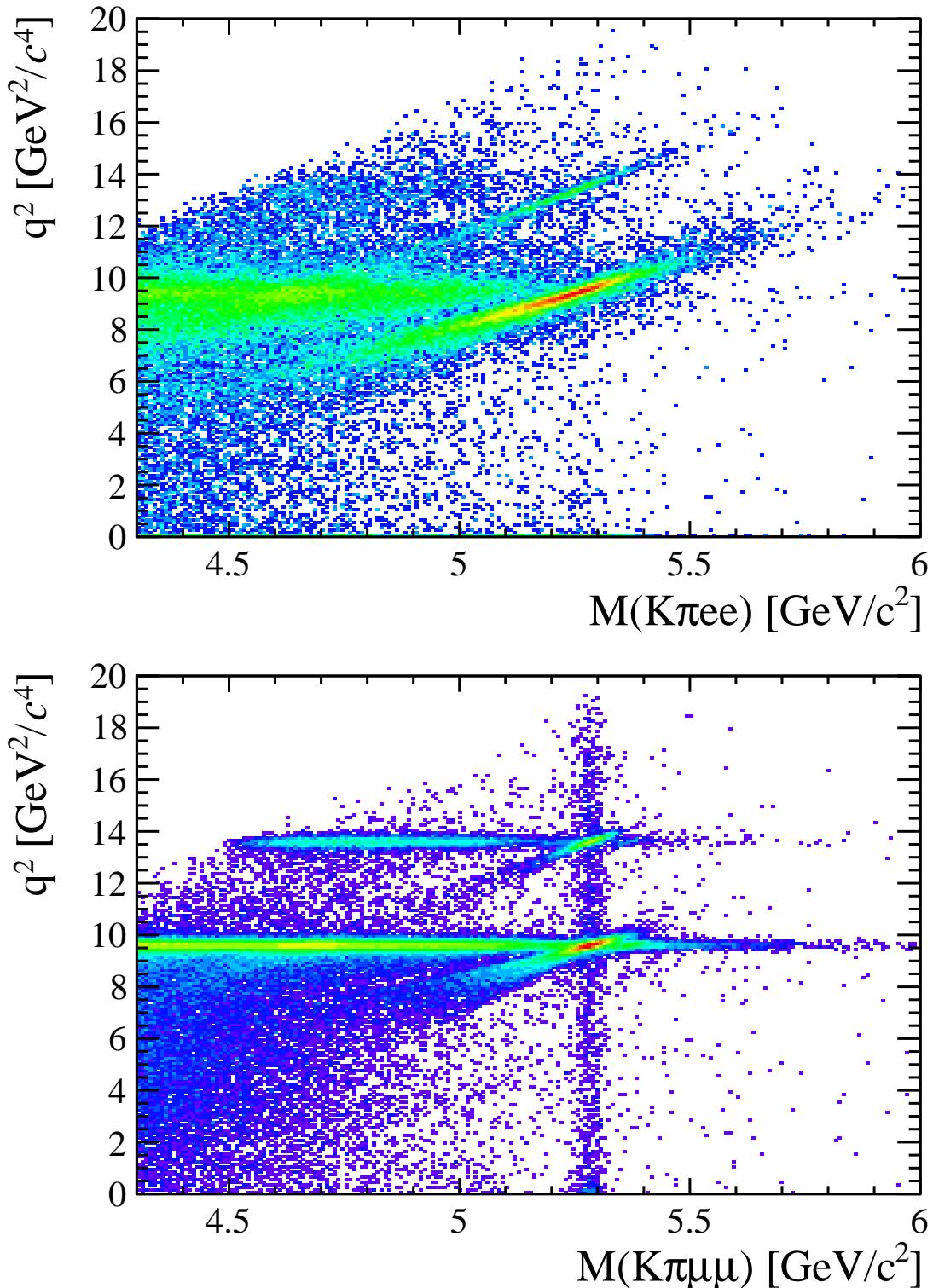


Figure 5.6: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electron (top) and muonic (bottom) channels in 2012 data.

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.

Table 5.2: Summary of the trigger lines used to select the $\mu\mu$ and the ee channels. Where not explicitly indicated, the lines are required to be TOS.

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

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

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

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

Table 5.3: Summary of stripping requirements used for the central and high q^2 regions.

Particle	Requirements
π	$\chi_{\text{IP}}^2(\text{primary}) > 9$
K	$\text{PID}_K > -5$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ hasRICH
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	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ isMuon
e	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ hasCalo $PID_e > 0$
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

1996 Candidates are then required to pass the kinematic and quality cuts summarised in
 1997 Tab. 5.3. The meaning of variables in the table was already explained in Sec. 3.4.
 1998 Loose PID cuts are applied in preselection to limit the size of the samples, while
 1999 tighter cuts are applied in a second stage. A large mass window is kept around
 2000 the B^0 peak in order to be able to fit the sideband and to train the multivariate
 2001 analysis and constrain backgrounds. Track-quality and vertex quality cuts are also
 2002 applied using the $\chi_{\text{track}}^2/\text{ndf}$, **GhostProb**, and $\chi_{\text{vtx}}^2/\text{ndf}$ variables. The **GhostProb**
 2003 quantity describes the probability of a track being fake. By construction cutting at
 2004 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. For details about the definition of
 2005 the variables used see Ref. [104].

2006 5.6.2 PID

2007 After preselection there still are high levels of misreconstructed background. In
 2008 particular, as the ID of kaons and pions are not constrained, the samples still contain
 2009 both ID combinations for most candidates, therefore tighter PID cuts are applied. In
 2010 the LHCb analysis framework the particle identification probability can be quantified
 2011 using the “ProbNN” variables [105]. These variables are the output of a Neural
 2012 Network which takes as input information from the calorimeters, the RICH detectors
 2013 and the muon system. Unlike the DLL variables these are bounded from 0 to 1 and
 2014 can be therefore directly be interpreted as probabilities. For example `ProbNNk` is
 2015 the probability for a reconstructed particle to be a kaon. Two tunes of the ProbNN
 2016 variables, labelled V2 and V3, are available. Tune V3 was shown to be optimal for
 2017 positive ID, while tune V3 was found to be optimal for background rejection and
 therefore it is used to quantify the mis-ID probability.

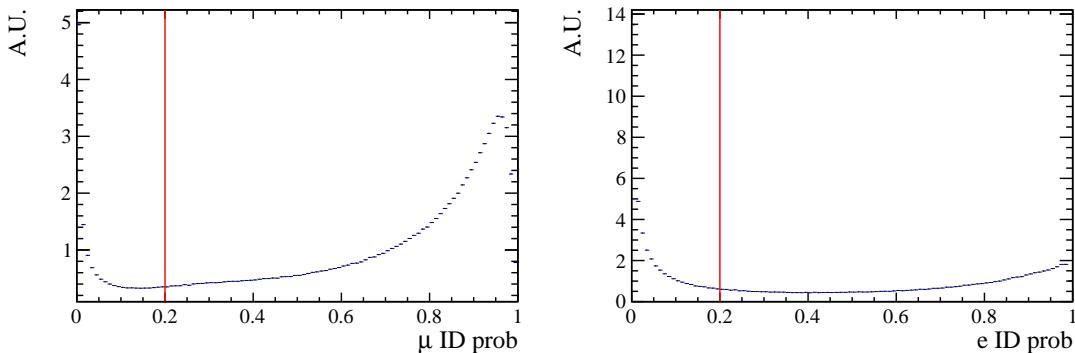


Figure 5.7: Correct ID probability distributions for muons (left) and electron (right) in 2012 data.

2018

2019 Figure 5.7 shows distributions of the correct ID variables in the 2012 data sample
 2020 while Fig. 5.8 shows in a two-dimensional plane the probability of correct identifica-
 2021 tion and mis-identification of kaons and pions. These plots are characterised by clear
 2022 peak at maximal ID probability and minimal mis-ID probability, corresponding to
 2023 particles to which is possible to assign a well defined identification.

2024 In order to maximise the power of the PID cuts probabilities of correct ID and
 2025 mis-ID are combined using the following cuts:

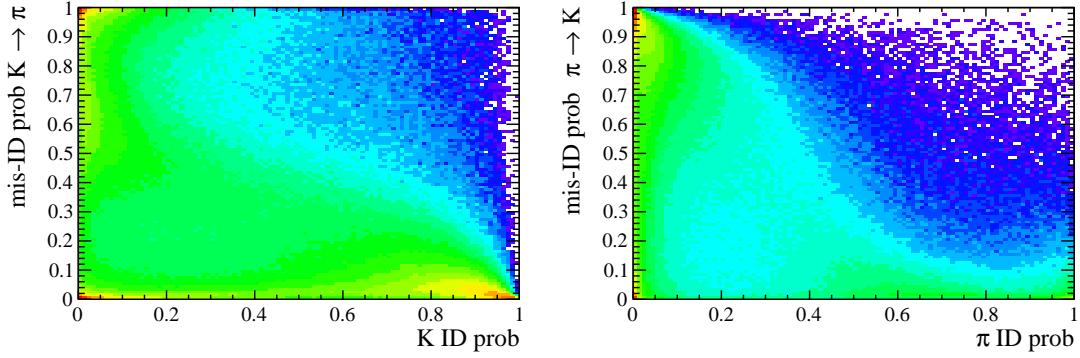


Figure 5.8: On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis show the mis-ID probability.

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

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

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

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

2026

2027 In the first formula, for example, ProbNNpi is the probability of correctly identifying
 2028 the pion as a pion, while ProbNNk is the probability of mistaking it for a kaon. There-
 2029 fore by maximising the quantity “ $\text{ProbNNpi} \times (1 - \text{ProbNNk})$ ”, one can maximise
 2030 the correct ID probability and minimise at the same time the mis-ID probability.

2031 5.6.3 Peaking backgrounds

2032 Cuts are applied in order to remove background sources due to specific decays. These
 2033 types of backgrounds usually peak in some variable because of their mass or dis-
 2034 tinctive kinematic properties and therefore they can be removed without significant
 2035 signal efficiency loss. In the following subsections are described the main sources of
 2036 peaking background.

²⁰³⁷ 5.6.3.1 Charmonium vetoes

²⁰³⁸ Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2 . The choice of q^2 binning
²⁰³⁹ described in Sec. 5.4 constitutes a natural veto for these decays. Simulated events
²⁰⁴⁰ were used to check if resonant events leak inside the q^2 intervals chosen for the rare
²⁰⁴¹ channel analysis. For the muonic channels the leakage is negligible as the peaks are
²⁰⁴² sharper due to a better resolution and muons emit fewer bremsstrahlung photons, re-
²⁰⁴³ sulting in shorter radiative tails. The electronic channels are instead characterised by
²⁰⁴⁴ a worse resolution and at the same time electrons can radiate several bremsstrahlung
²⁰⁴⁵ photons, yielding long tails at low q^2 . Analysing Monte Carlo events it was found
²⁰⁴⁶ that 1.3–2% (depending on the trigger category) of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candi-
²⁰⁴⁷ dates leak into the $1.1 < q^2 < 6$ GeV^2/c^4 interval and 1.8% of $\psi(2S)$ events leak
²⁰⁴⁸ above 15 GeV^2/c^4 . The contribution from these candidates is modelled in the fit.

²⁰⁴⁹ 5.6.3.2 ϕ veto

²⁰⁵⁰ It can happen that a kaon from the decay $B_s \rightarrow \phi\ell^+\ell^-$, where the ϕ decays in two
²⁰⁵¹ kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as a
²⁰⁵² K^{*0} . This results in a candidate with a value of $m(K\pi)$ that is less than the nominal
²⁰⁵³ K^{*0} mass but still high enough to pass the selection requirements. In Fig. 5.9 is
²⁰⁵⁴ reported the plot of $m(K\pi)$ versus $m(K\pi\ell\ell)$, where kaon mass hypothesis is assigned
²⁰⁵⁵ to the pion. A peak can clearly be seen around the ϕ mass ($1020 \text{ MeV}/c^2$). To remove
²⁰⁵⁶ this background only candidates with $m_{K(\pi \rightarrow K)} > 1040 \text{ MeV}/c^2$ are selected. This
²⁰⁵⁷ results in a 98% background rejection while keeping a 99% signal efficiency. The ϕ
²⁰⁵⁸ could also constitute a background when it decays into two leptons but the branching
²⁰⁵⁹ ratio of this decay is small compared to the one into kaons and this contribution is
²⁰⁶⁰ taken into account by the choice of the q^2 intervals.

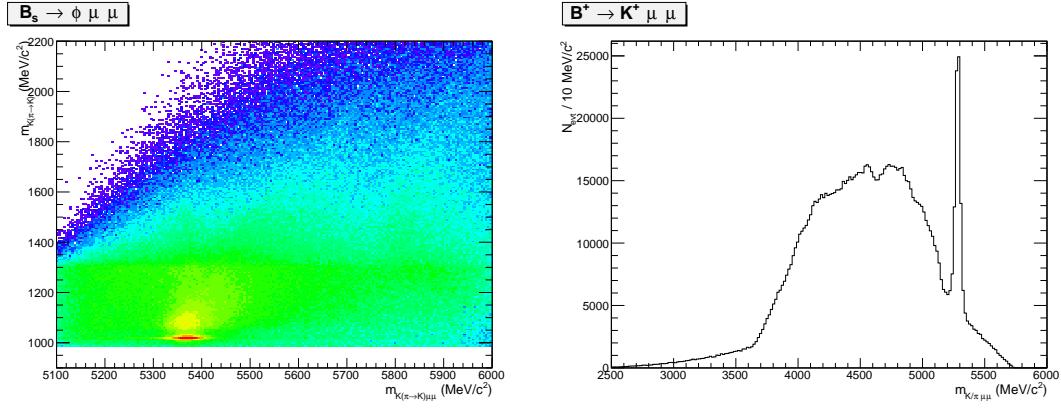


Figure 5.9: (left) Distribution of 2011 data events as a function of the variables ($m_{K(\pi \rightarrow K)}$) and ($m_{K(\pi \rightarrow K)\mu\mu}$), where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. (right) The invariant mass distribution of the three-body system ($K\mu\mu$),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

2062 5.6.3.3 $B^+ \rightarrow K^+\ell^+\ell^-$ plus a random pion

2063 Some $B^+ \rightarrow K^+\ell^+\ell^-$ decays can contaminate the upper B^0 mass sideband if they
 2064 are combined with a soft pion from somewhere else in the event and therefore re-
 2065 constructed as a B^0 decay. The same can also happen with a kaon misidentified
 2066 as a pion combined with another kaon in the event. In Fig. 5.9 the invariant
 2067 mass distribution of the three-body $K\mu^+\mu^-$ system, $m(K\mu\mu)$, is shown. This is
 2068 characterised by a narrow peak at the B^+ mass. Since these candidates have
 2069 $m(K\pi\ell\ell) > 5380$ MeV/ c^2 there is no contribution under the B^0 peak, but they
 2070 can cause problems when using sidebands events to train the neural network. An
 2071 effective veto for this decay was found to be $\max(m_{K\ell\ell}, m_{(K \rightarrow \pi)\ell\ell}) < 5100$ MeV/ c^2 ,
 2072 which results in 95% background rejection while keeping 99% signal efficiency.

2073 5.6.3.4 Λ_b decays

2074 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays are unlikely to be reconstructed as $B^0 \rightarrow K^{*0}\ell^+\ell^-$ because the
 2075 Λ is long-lived and decays further in the detector with a separate vertex. Simulated
 2076 events were used to check how many candidates fall into the B^0 samples, which
 2077 results to be negligible. The $\Lambda_b^0 \rightarrow J/\psi pK$ decay can instead contribute more easily

2078 since the $m(pK)$ is above the Λ threshold and therefore they must come from Λ^*
 2079 resonances, which are not long-lived. This background is already reduced using
 2080 PID but a non-negligible contribution is still expected in the $\mu\mu$ sample, which is
 2081 modelled in the fit.

2082 5.6.3.5 Other peaking backgrounds

2083 A possible background could come from $B^0 \rightarrow K^{*0}\gamma$ decays where the photon con-
 2084 verts into two electrons while traversing the detector. In LHCb, around 40% of
 2085 photons convert before the calorimeter, but only a small fraction of these, $\sim 10\%$,
 2086 are reconstructed. Furthermore these events fall into a q^2 region well below the
 2087 intervals considered in these analysis and their contribution is therefore negligible.
 2088 Similar decays are also $B^0 \rightarrow K^{*0}\eta$ and $B^0 \rightarrow K^{*0}\pi^0$ where η and the pion decay
 2089 into two photons. Once again the contribution from these decays falls well below
 2090 the considered q^2 intervals. Finally, a potentially dangerous background could come
 2091 from events where the identity of the kaon and the pion are swapped as these candi-
 2092 dates peak under the signal. Their contribution is found to be small, 0.5%, however
 2093 the effect of their modelling into the fit is taken into account in the systematic
 2094 uncertainties.

2095 5.6.4 Mis-reconstructed background

2096 A source of mis-reconstructed background is due to cascade decays with a B^0 de-
 2097 caying semileptonically into a D meson which also decays semileptonically, e.g.
 2098 $B^0 \rightarrow D^-\ell^+\bar{\nu}_\ell$ followed by $D^- \rightarrow K^{*0}\ell^-\nu_\ell$. The candidates built from these de-
 2099 cays tend to have a low 4-body invariant mass as two or more particles are not
 2100 reconstructed.

2101 In order to remove this background in the muonic channels, the 4-body $m(K\pi\mu^+\mu^-)$
 2102 invariant mass is recalculated with a kinematical fit using the `DecayTreeFitter`
 2103 package. In the resonant case this includes a constraint of the dilepton mass to

2104 be the J/ψ nominal mass and in both rare and resonant cases each particles is
2105 constrained to point to its origin vertex. This constraint has the effect of pushing the
2106 misreconstructed events far from the B^0 peak. Therefore, to avoid this background,
2107 it is sufficient to limit the analysis to 4-body invariant masses above $5150 \text{ MeV}/c^2$.

2108 In the electron case it is instead important to fit a wider mass window to correctly
2109 constrain the background therefore one cannot eliminate this mis-reconstructed
2110 background which is then modelled in the fit (for details see Sec. 5.7.2.2).

2111 5.6.5 Multivariate analysis

2112 The final selection is performed using a Neural Network classifier (NN) based on the
2113 NeuroBayes package [65, 66]. The multivariate analysis is intended to remove some
2114 combinatorial background and obtain a clearer signal peak.

2115 For the final selection in the central and high q^2 intervals a Neural Network clas-
2116 sifier (NN) is used based on the NeuroBayes package [65, 66]. Representative
2117 samples of the signal and background are needed to train the classifier. For the
2118 signal, fully reconstructed $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$ simulated events
2119 can be used. A sample representative of the background can be obtained taking
2120 real data events in the upper B^0 sideband: $m(K\pi\mu^+\mu^-) > 5400 \text{ MeV}/c^2$ and
2121 $m(K\pi e^+e^-) > 5600 \text{ MeV}/c^2$. The lower sideband is not used in the training as it
2122 contains a significant fraction of mis-reconstructed background. All pre-selection
2123 cuts are applied to the background samples used for the training. As L0 and PID
2124 variables are not well described these cuts are not applied in the Monte Carlo sam-
2125 ples but their effect is taken into account by the event weight. To train the classifier
2126 50% of the sideband events was used, keeping the other 50% for testing. For the sig-
2127 nal sample a number of Monte Carlo events was used equal to the number available
2128 for the background sample.

2129 The input to the NN consists of 22 variables containing information about the
2130 kinematic of the decays and the quality of tracks and vertices. All the variables

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 5.4: 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. 5.10.

used are listed in Tab. 5.4. The graphical representation of the correlation matrices are shown in Fig. 5.10, in these figures the variable with ID = 1 is the NN output and the other IDs are reported in Tab. 5.4. The single most discriminating variable used is the χ^2 of a kinematic fit that constrains the decay product of the B^0 , the K^{*0} and the dimuon, to originate from their respective vertices. Other variables that contribute significantly are the χ_{IP}^2 of J/ψ and K^{*0} , the transverse momentum of the B^0 and the pointing direction (DIRA) of the reconstructed B^0 to the primary vertex. The list the 10 most important variables is reported in Tab. 5.5, together with information on the relative importance of each input. The meaning of the column headings in this table was already explained in Sec. 3.4.2.

Figure 5.11 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 5.12 reports plots of the average NN output as a function of the B^0 mass on sideband data and simulated signal events. The

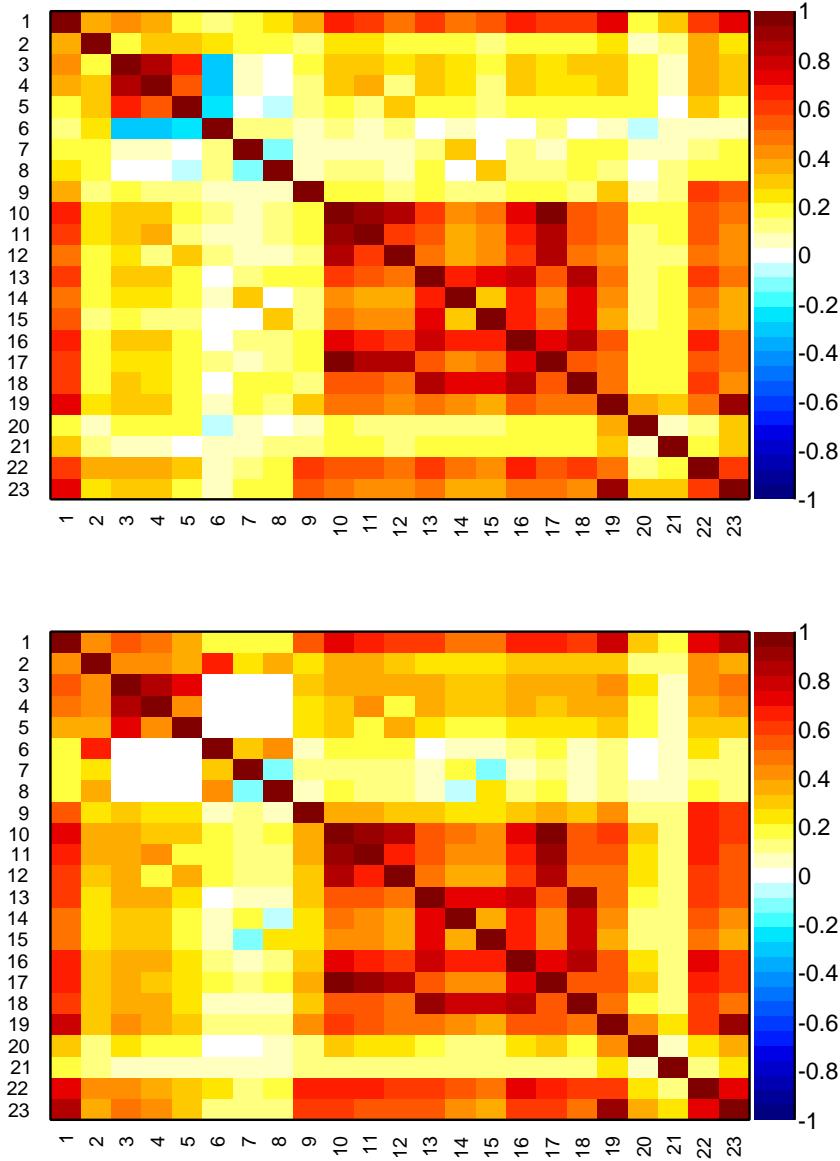


Figure 5.10: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of Tab. 5.5. Correlation is calculated using all events without distinguishing signal and background.

Table 5.5: 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

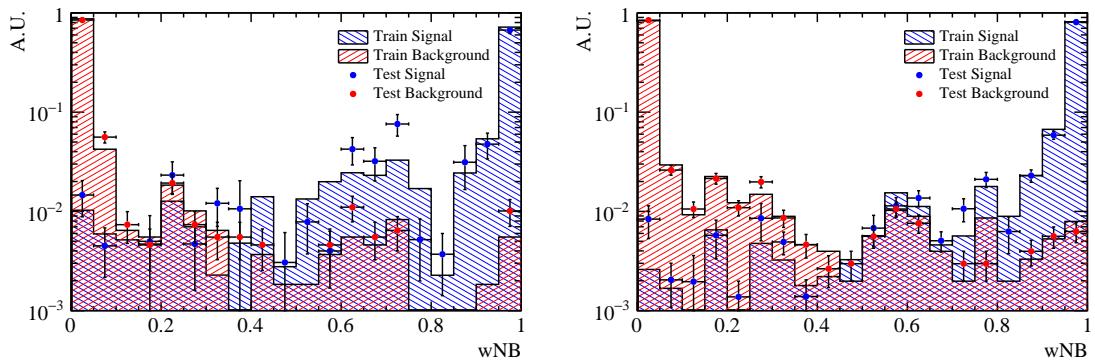


Figure 5.11: NN output distributions for training (solid) and test (stripes) samples, for simulated signal and data sideband events. For the electron (left) and muon (right) training.

2152 distributions are flat showing that no significant correlation is present.

2153 5.6.6 MVA optimisation

2154 In order to optimise the cut on our neural network output the expected signal
2155 significance, $N_S/\sqrt{N_S + N_B}$, was maximised. In this formula N_S is number of rare
2156 signal events and N_B the number of background events.

2157 The number of signal events accepted for a given NN output cut is determined
2158 exploiting the resonant channel and simulation. First, as an arbitrary number of
2159 events can be simulated, this has to be rescaled to the expected yield. This is done
2160 by fitting $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ events after pre-selection, including all selection
2161 cuts except MVA. The resonant yield is then scaled down by the expected ratio
2162 between the rare and the resonant channels. The number of background events is
2163 instead derived by fitting the combinatorial background in the sideband with an
2164 exponential function and extrapolating the fit function below the signal peak.

2165 The dependence of the figure-of-merit for both the electron and muon trainings are
2166 shown in Fig.5.13, where the red line indicate the chosen cut: 0.75 for both samples.
2167 Curves of signal efficiency versus background rejection are shown in Fig. 5.13. Using
2168 the described MVA cuts the signal efficiency is $\sim 91\%$ for the muon channels and
2169 $\sim 84\%$ for the electron channels (for more details see Sec. 5.8), while the background
2170 rejections is $\sim 98\%$ on both samples.

2171 After full selection about $\sim 3\%$ of events still contain multiple candidates which are
2172 removed at random keeping only a single candidate per event.

2173 5.7 Mass fits

2174 In order to extract the signal yields a simultanous unbinned maximum likelihood
2175 fit to the 4-body invariant masses, $m(K\pi\ell\ell)$, is performed on the rare and resonant

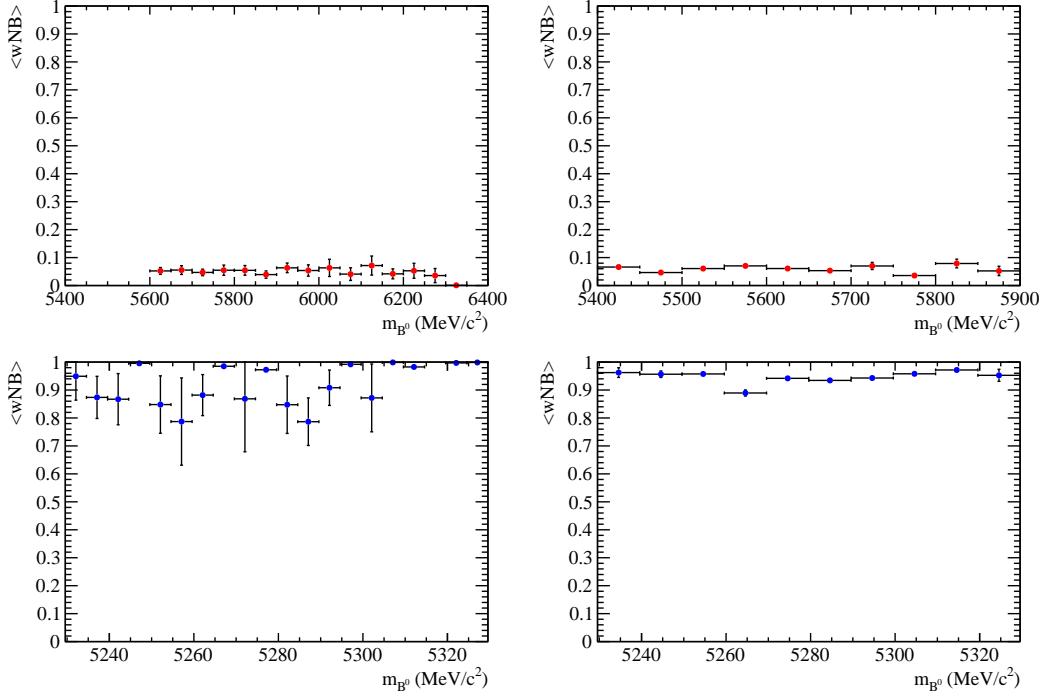


Figure 5.12: Average value of NN output as a function of B^0 mass for data sideband (top) and simulated signal (bottom) events for the electron (left) and muon (right) training.

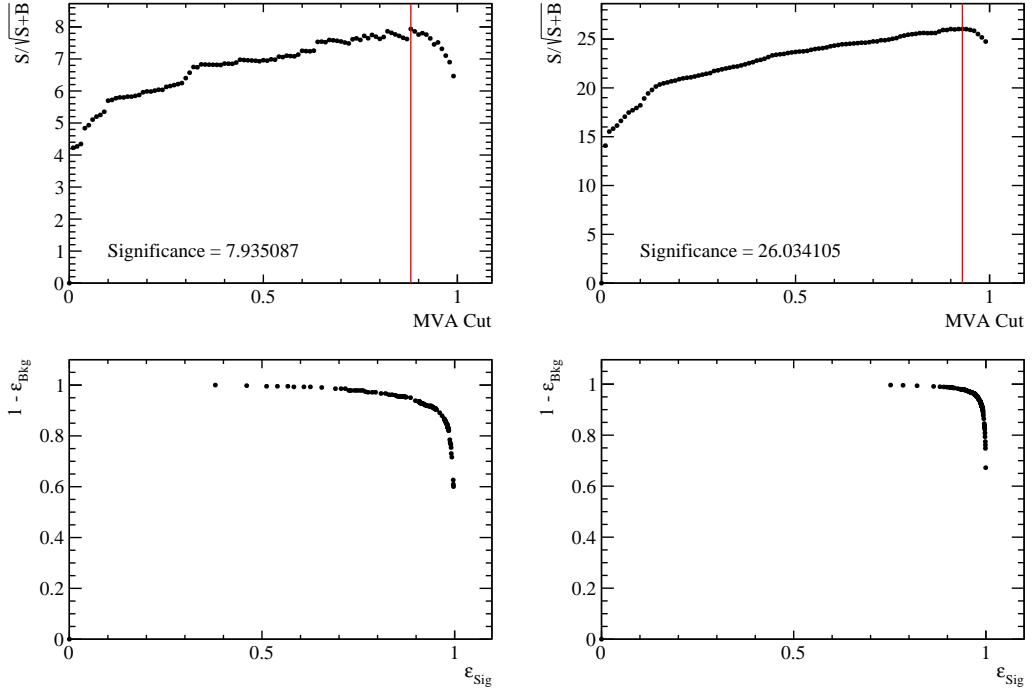


Figure 5.13: (top) Dependence of figure-of-merit on the requirement on neural network output. Vertical lines corresponds to the chosen cuts. (bottom) Signal efficiency versus the background rejection. Plots correspond to the electron (left) and muons (right) samples.

samples. The following subsubsections 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

$$N_{\ell\ell} = N_{J/\psi} \cdot \varepsilon^{\text{rel}} \cdot R_{\ell\ell}, \quad (5.10)$$

where ε^{rel} is the relative efficiency between the rare and resonant channels given in Tab. 5.12 and $R_{\ell\ell}$ corresponds to the efficiency corrected ratio of the raw rare and resonant yields:

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

The two ratios, R_{ee} and $R_{\mu\mu}$, are then used to determine the $R_{K^{*0}}$ quantity, as described in Sec. 5.10.

5.7.1 Muon channels

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

2198 The PDF chosen to describe the signal in both the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and its relative
2199 J/ψ channel is a Double Crystal Ball function already described in Sec. 3.5.1 and
2200 also in this case the mean value (m_0) is kept in common:

2201 As a first step simulated distributions are fit using the signal model to extract
2202 parameters to be constrained in the fit to data. The fitted MC distribution for the
2203 resonant channel is reported in Fig. 5.14.

2204 For the fit to real data the signal parameters are fixed to the ones found for the
2205 simulated samples. In order to account for possible data-simulation discrepancies
2206 a scale factor is multiplied to the widths and a shift is added to the masses. In
2207 summary the PDFs used for the signal fits on data are defined as

$$\begin{aligned} P(m; c, m') = & f^* \cdot C(m; \alpha_1^*, n_1^*, c \cdot \sigma_1^*, m_0^* + m') \\ & + (1 - f^*) \cdot C(m; \alpha_2^*, n_2^*, c \cdot \sigma_2^*, m_0^* + m') \end{aligned} \quad (5.12)$$

2208 where the free parameters are the width scale factor, c , and the mass shift, m' ,
2209 which are common between the rare and resonant samples. The other parameters,
2210 denoted with $*$, are taken from the fit to simulated events, separately for the rare and
2211 resonant samples and are fixed in the fit on data. The parameter f^* in the formula
2212 is the relative fraction of candidates falling in the first Crystal Ball function.

2213 To model the combinatorial background an exponential function was used. This is
2214 the only background component for the rare channel. In the normalisation channel
2215 fit the $B_s^0 \rightarrow K^{*0}J/\psi$ background is described using the same PDF used for the sig-
2216 nal but a different central value, m , which is set at the B_s^0 nominal mass [2]. Finally,
2217 a $\Lambda_b^0 \rightarrow J/\psi pK$ background component is modelled using simulated $\Lambda_b^0 \rightarrow J/\psi pK$
2218 events to which the full $B^0 \rightarrow K^{*0}J/\psi$ selection is applied. The invariant mass dis-
2219 tribution of these candidates is a broad shape under the signal peak. The simulated
2220 distribution is smoothed using a kernel estimation method (using the `RooKeysPdf`
2221 class of the `RooFit` package [106]).

2222 In summary the free parameters in the simultaneous fit to rare and resonant $\mu^+\mu^-$

2223 data samples are: the signal and background yields, the combinatorial background
 2224 slopes, the widths scale, c , and the the mass shift, m' .

2225 Fig. 5.15 reports fits to real data distributions for the rare and resonant $\mu^+\mu^-$
 channels. Values of fitted parameters are reported on the plots.

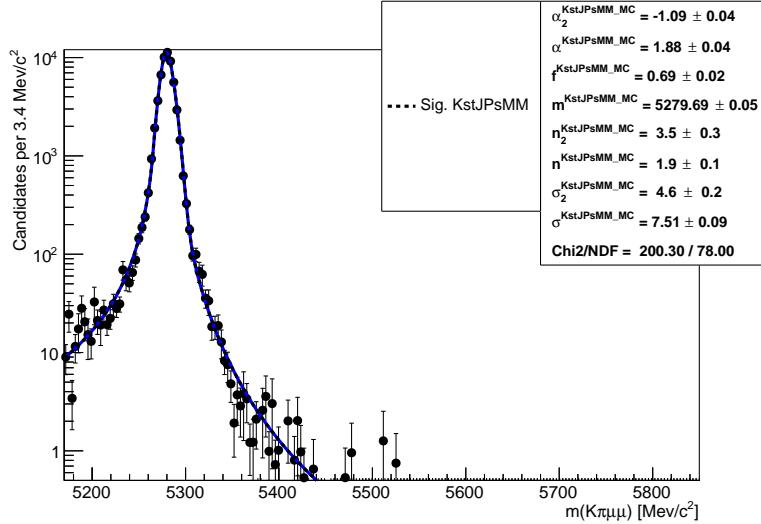


Figure 5.14: Fitted $m(K\pi\mu\mu)$ mass spectrum for $K^{*0}J/\psi$ simulated events.

2226

2227 5.7.2 Electron channels

2228 In the electron case the variable we fit is the $m(K\pi e^+e^-)$ invariant mass coming
 2229 from the kinematic fit where all vertices are required to point to their mother par-
 2230 ticle. In contrast to the muon channel, the constraint to the dilepton mass to the
 2231 nominal J/ψ nominal value is not applied. In fact, due to the longer bremsstrahlung
 2232 tail, the J/ψ mass constraint distorts the invariant mass distribution and makes it
 2233 is hard to model it. Furthermore, mis-reconstructed background enters in the rare
 2234 channel sample and its amount can be constrained by exploiting the higher statistics
 2235 resonant channel, but this implies the usage of the same variable in both fits. In
 2236 order to better constrain the parameters modelling the radiative tail and the mis-
 2237 reconstructed backgrounds a wide mass window is used [4500,5800] MeV/ c^2 . The
 2238 lower limit is given by the point in which the q^2 cut (at 6 GeV $^2/c^4$) to separate the

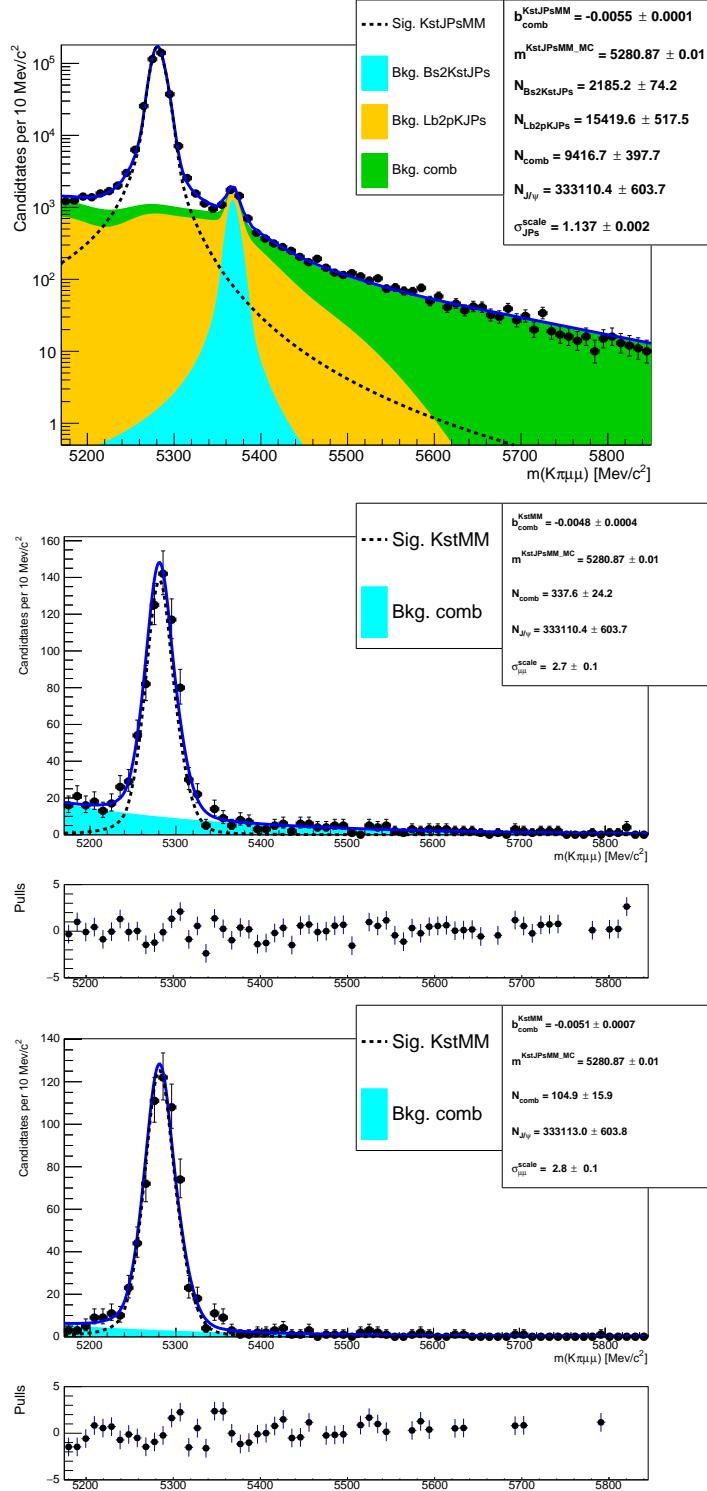


Figure 5.15: Fitted $m(K\pi\mu\mu)$ invariant mass distribution for $K^{*0}J/\psi$ candidates (top) and for rare candidates in the central (middle left) and high (middle right) q^2 intervals. Dashed black lines represent the signal PDFs and filled shapes the background components.

2239 rare and resonant channels) starts to affect the 4-body invariant mass distribution.

2240 In the electron case the invariant mass distributions are different depending on which
 2241 hardware trigger was used and especially how many bremsstrahlung photons were
 2242 recovered. Therefore our sample is divided in 3 trigger categories, as described in
 2243 Sec. 5.6.1, and 3 bremsstrahlung categories defined as:

- 2244 • 0 γ : events with no photon emitted
- 2245 • 1 γ : events with one photon by either of the electrons
- 2246 • 2 γ : events with one photon emitted by each electron

2247 The three samples, divided by trigger, are fitted simultaneously. This allows a better
 2248 use of statistics as the simultaneous fit gathers information from the three categories
 2249 at the same time and is more stable. Furthermore using this method the results for
 2250 the three categories are naturally combined in a single R_{ee} ratio.

2251 The PDFs used to fit the invariant mass distributions are described in the next
 2252 subsection.

2253 5.7.2.1 Signal PDFs for the electron channels in the central- q^2 interval

2254 As for the muonic channel simulated events are fitted first to constrain the shapes
 2255 for the subsequent fit on data. The signal PDFs are built using the following method:

- 2256 • Simulated $B^0 \rightarrow K^{*0}J/\psi(ee)$ and $B^0 \rightarrow K^{*0}ee$ events are divided in each
 2257 trigger and bremsstrahlung category and an independent fit is performed to
 2258 each sample.
- 2259 • For each trigger category a PDF is built as the sum of the three PDFs for each
 2260 bremsstrahlung category.

$$P(m)^{\text{trg}} = f_{0\gamma}^{\text{trg}} P_{0\gamma}^{\text{trg}}(m) + f_{1\gamma}^{\text{trg}} P_{1\gamma}^{\text{trg}}(m) + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) P_{2\gamma}^{\text{trg}}(m). \quad (5.13)$$

Table 5.6: 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 %

2261 where the $P(x)_{n\gamma}^{trg}$ functions are the chosen PDFs for each trigger and bremsstrahlung
 2262 category and the $f_{n\gamma}^{trg}$ parameters are the relative fractions of events falling in
 2263 each category.

- 2264 • Most parameters are fixed (details later) and the combined PDF, $P(m)$, is
 2265 used to fit real data divided only in trigger categories.

2266 The 0γ category is characterised by a better resolution and a sharp tail on the right-
 2267 hand side and it is fitted with a simple Crystal Ball function (CB). Instead the 1γ
 2268 and 2γ samples are modelled using the sum of a Crystal Ball and a Gaussian func-
 2269 tions (CBG) with all parameters independent. When the combined PDF, $P(m)^{trg}$,
 2270 is built all parameters are fixed leaving one global mass shift and one scale factor
 2271 for the widths free to vary, as done for the muonic samples.

2272 Finally, combining the three bremsstrahlung PDFs one needs to specify in which
 2273 fractions they contribute to the total. These fractions have been shown to be in
 2274 good agreement between resonant data and Monte Carlo and therefore they are
 2275 fixed to the values found on simulation, separately for the normalisation channel
 2276 and each q^2 bin. In Tab. 5.6 are reported percentages of events with 0, 1 and 2
 2277 recovered photons in the three trigger categories.

2278 In summary the signal PDF for the fit on data is defined as:

$$P(m; c, m')^{\text{trg}} = f_{0\gamma}^{\text{trg}} \text{CB}_{0\gamma}^{\text{trg}}(m; c, m') + f_{1\gamma}^{\text{trg}} \text{CBG}_{1\gamma}^{\text{trg}}(m; c, m') + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) \text{CBG}_{2\gamma}^{\text{trg}}(m; c, m') \quad (5.14)$$

2279 where the free parameters are: c , the scaling factor for the widths, and m' , the mass
2280 shift.

2281 5.7.2.2 Background PDFs for the electron channels in the central- q^2 interval

2282 In the fit to the resonant sample three background components are modelled: combinatorial background, and misreconstructed background coming from the hadronic
2283 and the leptonic systems. The combinatorial is described with an exponential function.
2284

2285 The misreconstructed background is split in two categories, that involving higher
2286 hadronic resonances and that coming from higher $c\bar{c}$ resonances. These backgrounds
2287 are modelled using inclusive $B^0 \rightarrow J/\psi X$ simulated samples to which the full selec-
2288 tion is applied. The distributions for the hadronic (leptonic) background are defined
2289 selecting candidates where the K^{*0} (dimuon) is not a direct daughter of the B^0 . The
2290 invariant mass distributions of these events, shown in Fig. 5.16, are smoothed using
2291 a kernel estimation method and their yields are left floating in the fit. Given the
2292 low statistics available, the same shape is used for the three trigger categories.
2293

2294 In the fit for the rare sample in the central- q^2 interval the modelled backgrounds are:
2295 combinatorial background, again modelled with an exponential; misreconstructed
2296 background coming from the hadronic system and the leakage of the J/ψ radiative
2297 tail into the lower q^2 interval. The shape for the misreconstructed component is
2298 obtained from simulated distributions similarly to what described for the resonant
2299 channel. However, as there are no inclusive samples for the rare case, a sample
2300 including higher K^* resonances is generated, including $K_1^+(1400)$ and $K_2^+(1460)$.
2301 The yield of this component is not floating independently but its relative proportion
2302 with respect to the signal yield is constrained to be the same as in the resonant

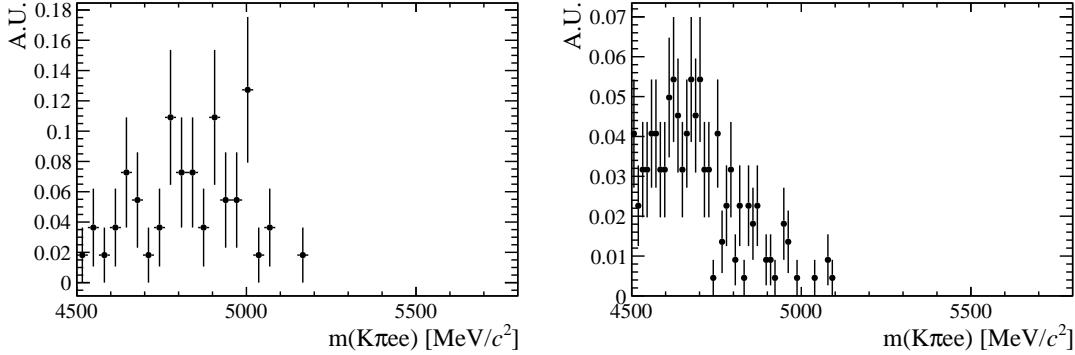


Figure 5.16: Simulated distributions of misreconstructed background events falling into the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ sample coming from the hadronic (left) and leptonic (right) systems.

sample, namely:

$$N_{\ell\ell}^{mis-reco} = N_{ee} \cdot k = N_{ee} \cdot \frac{N_{J/\psi}^{mis-reco}}{N_{J/\psi}}. \quad (5.15)$$

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

The shape to describe the J/ψ tail leakage is obtained using simulated $B^0 \rightarrow J/\psi K^{*0}$ candidates and selecting those falling in q^2 below $6 \text{ GeV}^2/c^4$. The 4-body invariant mass distribution of these events is reported in Fig. 5.17. The yield of this component again is not floating independently but it is linked to the yield found in the resonant fit as follows

$$N_{\ell\ell}^{leak} = N_{J/\psi} \cdot k^{MC} = N_{J/\psi} \cdot \frac{N_{leak}^{MC}}{N_{J/\psi}^{MC}} \quad (5.16)$$

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

5.7.2.3 Summary: electron channels fit in the central- q^2 interval

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

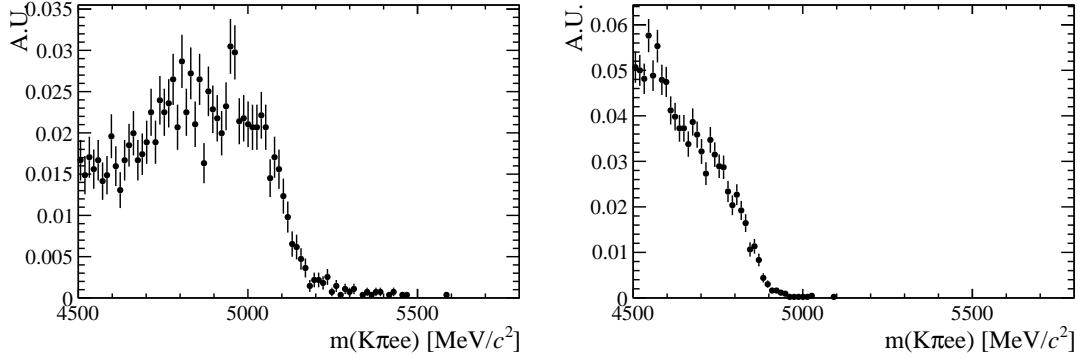


Figure 5.17: (left) Simulated 4-body invariant mass distributions for events involving higher K^* states and passing out full selection. (right) Simulated invariant mass distribution of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ events leaking into the central- q^2 interval.

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

Fits on simulated $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates are shown in Appendix D. Figure 5.18 and 5.19 present fits on real $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}e^+e^-$ candidates (central- q^2 interval) in the three trigger categories. Values of fitted parameters are reported on the plots.

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

In the high- q^2 interval, above $15 \text{ GeV}^2/c^4$, the efficiency for the L0Hadron trigger becomes very low as the K^* has very low momentum. In this region only 9 candidates are found spread in the interval $4500 < m(K\pi ee) < 6000 \text{ MeV}/c^2$. In the L0TIS category, even if the yield is bigger a clear peak cannot be seen, therefore only L0Electron triggered events are used in this region.

The signal PDF is described in the same way as for the central bin. Simulated events are divided in three bremsstrahlung categories and fitted using the same PDFs described in Sec. 5.7.2.1. While the signal tail parameters are similar for the J/ψ and central- q^2 samples in the case of the high- q^2 interval it is particularly important to keep them independent. In fact, as can be seen in Fig. 5.20, the invariant mass distributions are significantly different for the two intervals. The

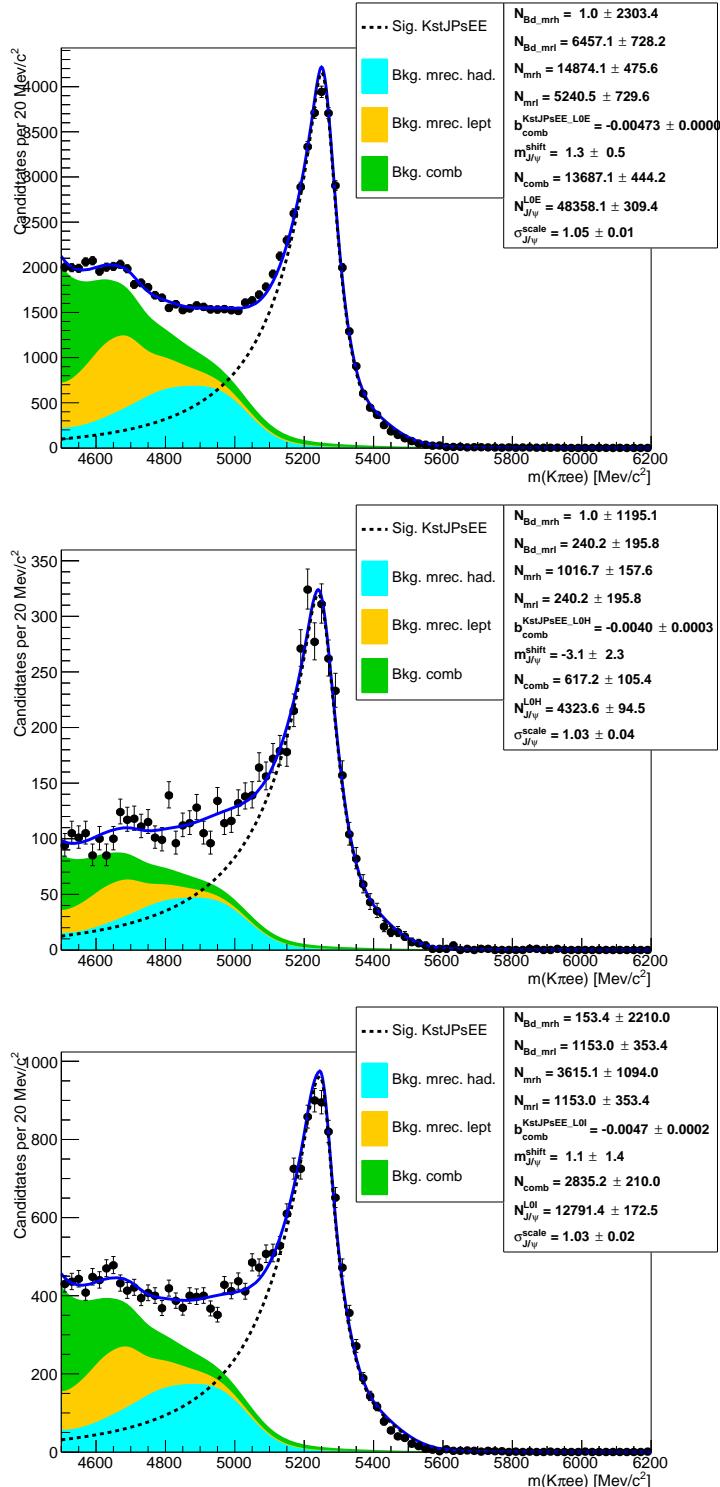


Figure 5.18: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

2336 fractions of 0, 1 and 2 γ components used to build the total PDF are also in this
 2337 case taken from simulated events are reported in Tab. 5.7.

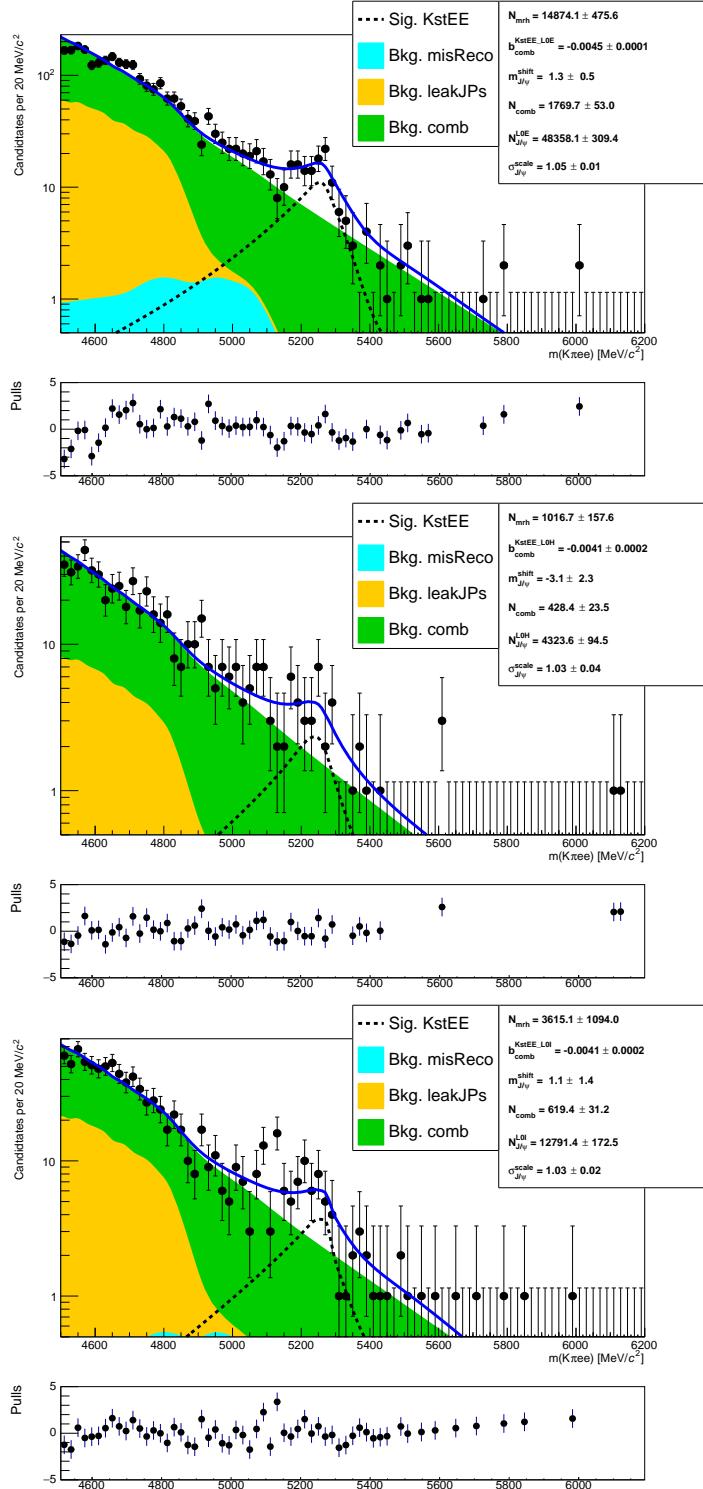


Figure 5.19: Fit to the $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} e^+ e^-$ real data events in the three trigger categories. From top to bottom L0E, L0H and L0I.

2338 The background components, as for the central- q^2 interval, include a combinatorial
 2339 background and a misreconstructed background coming from the hadronic system.

Table 5.7: 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 %
$15-20 \text{ GeV}^2/c^4$ (L0E)	20.7 %	51.7 %	27.6 %

Furthermore there is a leakage due to the $\psi(2S)$ resonance, that is wide enough to contribute in q^2 above $15 \text{ GeV}^2/c^4$. The combinatorial background is modelled using the shape obtained by reversing the NN output cut. In Fig. 5.21 are shown invariant mass distributions for different NN output cuts using the electron and muon high- q^2 samples. The shapes are normalised to the same integral. This plot shows that the shape is similar for the two samples and as a function of the NN output cut, which supports the hypothesis that these distributions contain mostly combinatorial background and that its shape has no strong dependency on the NN output cut. The shape is therefore obtained from the muon sample, which has higher statistics, selecting events with NN output < 0.1 . The distribution is then smoothed using the `RooKeysPdf` method.

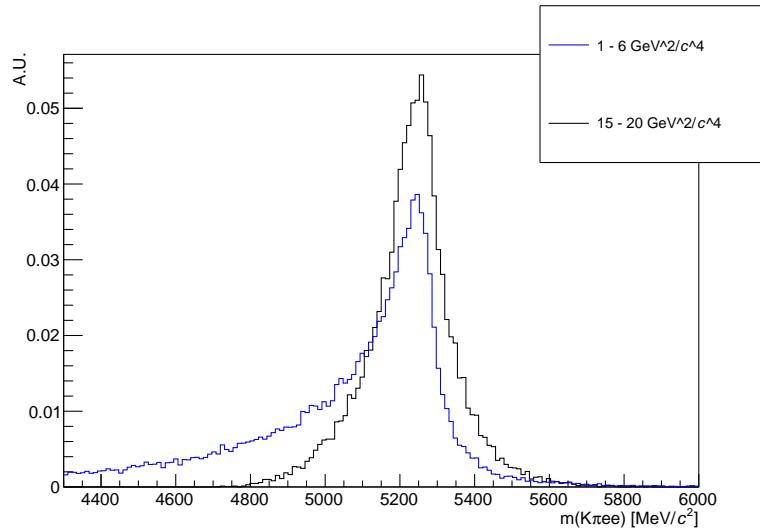


Figure 5.20: Simulated invariant mass of the $K\pi ee$ system in the $1.1 < q^2 < 6$ and $q^2 > 15 \text{ GeV}^2/c^4$ intervals.

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

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

$$N_{\ell\ell}^{leak} = N_{\psi(2S)} \cdot k^{MC} = N_{\psi(2S)} \cdot \frac{N_{\ell\ell}^{MC}}{N_{\psi(2S)}^{MC}}. \quad (5.17)$$

In order to do this the $\psi(2S)$ yield, $N_{\psi(2S)}$, is obtained from a fit to the $\psi(2S)$ invariant mass peak. Since we are only interested in the $\psi(2S)$ yield we fit the $m(K\pi ee)$ obtained from a kinematic fit where the dimuon mass is constrained to the known $\psi(2S)$ mass. This allows to eliminate the misreconstructed background form the fit mass window and use a simple model composed by a signal component and a combinatorial background component. The signal is described with a Double Crystal Ball function with opposite tails already described the A_b^0 fits (see Sec. 3.5.1), and the combinatorial background is described with an exponential. The fit to the $\psi(2S)$ peak is reported in Fig. 5.22 together with the fit to the $B^0 \rightarrow K^*e^+e^-$ candidates in the high- q^2 interval.

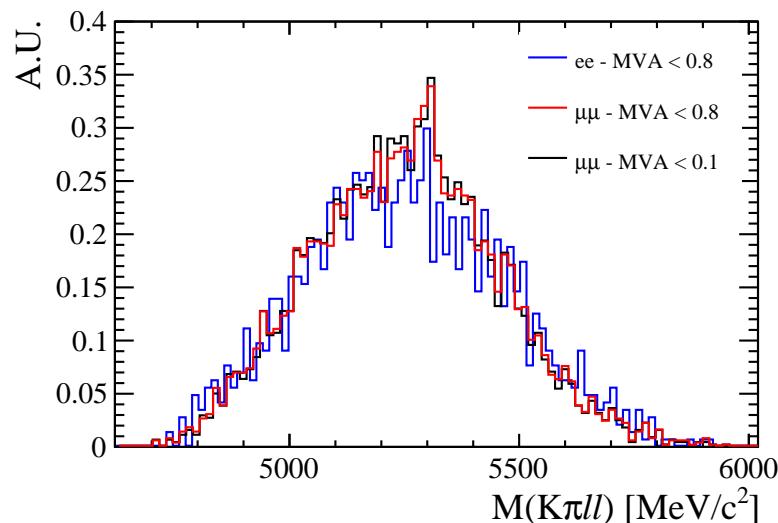


Figure 5.21: Invariant mass distribution of candidates selected by reversed NN output cuts.

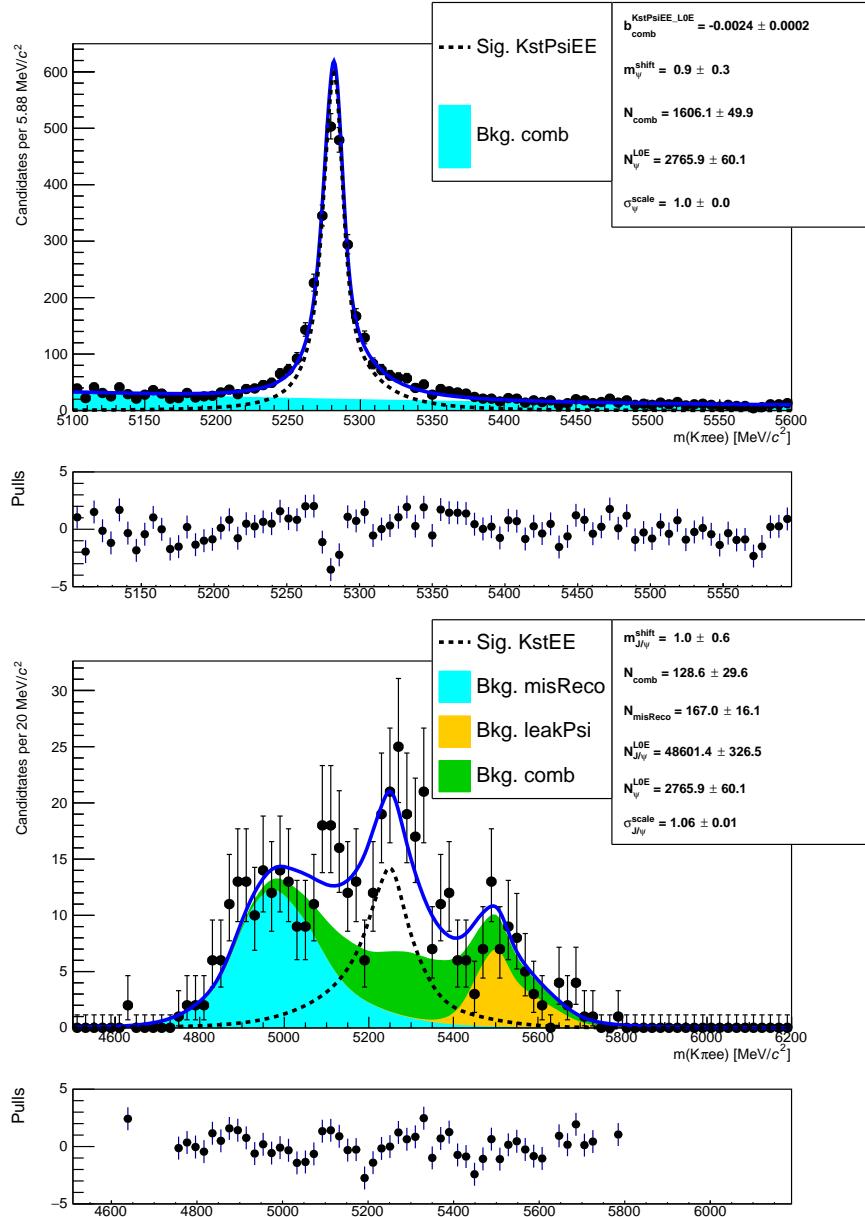


Figure 5.22: Fitted $m(K\pi ee)$ invariant mass distribution in the $\psi(2S)$ interval, $11 < q^2 < 15 \text{ GeV}^2/c^4$ and in the high- q^2 interval.

2367 5.7.3 Fit summary

2368 In Tab. 5.8 are reported raw yields obtained from the fits described in the previous
 2369 subsections. The values for the rare channels are not directly floating in the fits
 2370 but as described in Sec. 5.7 they are parameterised as a function of the number of
 2371 resonant events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare
 2372 branching fractions. Measured values of these ratios are reported in Tab. 5.18.

Sample	1–6 GeV^2/c^4	15–20 GeV^2/c^4	J/ψ
$\mu\mu$	626.47 ± 29.60	605.09 ± 27.44	333112.99 ± 603.77
ee L0E	131.62 ± 17.11	136.69 ± 27.34	48601.38 ± 326.48
ee L0H	31.65 ± 4.16	—	4323.62 ± 94.49
ee L0I	49.59 ± 6.48	—	12791.37 ± 172.47

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

2373 5.8 Efficiency

2374 The efficiency for each of the decay channels is calculated according to the formula

$$\varepsilon^{tot} = \varepsilon(geom)\varepsilon(reco|geom)\varepsilon(PID|reco)\varepsilon(trig|PID)\varepsilon(MVA|trig). \quad (5.18)$$

2375 In this expression the first term is the efficiency to have final state particles in the
 2376 LHCb detector acceptance. The second term carries information on reconstruction
 2377 and stripping efficiency (we keep these together given that boundaries between them
 2378 are completely artificial). The third part corresponds to the efficiency of the PID
 2379 requirements. The fourth term handles the trigger efficiency for those events which
 2380 are selected by the preselection process. Finally, the latter term deals with the
 2381 efficiency of the NN classifier. Reconstruction, trigger and MVA efficiencies are
 2382 evaluated on simulated data with the trigger efficiency for $B^0 \rightarrow K^* J/\psi$ being
 2383 cross-checked using the data-driven TISTOS method as described in Sec. 3.6.3. The
 2384 PID efficiency is calculated with a data-driven method as described in Sec. 5.8.3.

2385 All absolute efficiencies for the muon and electron rare channels are separately listed
 2386 in Tab. 5.11 for the central and high q^2 intervals and in Tab. 5.9 for the resonant
 2387 channels. However for the analysis itself only efficiencies relative to the resonant
 2388 channels are used in order to limit systematic uncertainties.

2389 Tab. 5.11 reports relative efficiencies between the rare and resonant channels, $\varepsilon(B^0 \rightarrow$
 2390 $K^{*0}\ell^+\ell^-)/\varepsilon(B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-))$. Finally, in Tab. 5.10 are listed ratios of rel-
 2391 ative efficiencies for the ee and $\mu\mu$ channels, $[ee/(J/\psi \rightarrow ee)]/[\mu\mu/(J/\psi \rightarrow \mu\mu)]$.

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

ε	$\mu\mu$	ee		
		L0E	L0H	L0I
Geom	0.1598 ± 0.0005		0.1589 ± 0.0005	
Reco	0.0947 ± 0.0001		0.0603 ± 0.0001	
PID	0.8148 ± 0.0000		0.8222 ± 0.0000	
Trig	0.7511 ± 0.0005	0.1939 ± 0.0005	0.0163 ± 0.0002	0.0707 ± 0.0003
MVA	0.8944 ± 0.0004	0.8597 ± 0.0007	0.8983 ± 0.0006	0.8276 ± 0.0017
Tot	0.0083 ± 0.0000	0.0013 ± 0.0000	0.0001 ± 0.0000	0.0005 ± 0.0000

Table 5.10: Double ratios of efficiencies $(\varepsilon^{ee}/\varepsilon^{J/\psi \rightarrow ee})/(\varepsilon^{\mu^+\mu^-}/\varepsilon^{J/\psi \rightarrow \mu^+\mu^-})$ in the central and high q^2 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
	1.022 ± 0.012			1.022 ± 0.012
	0.931 ± 0.006			0.443 ± 0.443
	0.982 ± 0.000			0.970 ± 0.000
Geom	0.882 ± 0.008	2.459 ± 0.050	1.257 ± 0.016	1.416 ± 0.014
Reco	0.970 ± 0.003	0.936 ± 0.003	0.970 ± 0.006	1.064 ± 0.006
PID	1.146 ± 0.023	3.084 ± 0.084	1.633 ± 0.037	0.858 ± 0.020
Tot				

Table 5.11: Absolute efficiencies for the rare ee and $\mu\mu$ channels in the central and high q^2 intervals.

ε	$\mu\mu$		ee		15–20 GeV^2/c^4 L0E
	1–6 GeV^2/c^4	15–20 GeV^2/c^4	1–6 GeV^2/c^4 L0H	15–20 GeV^2/c^4 L0I	
q^2	0.2142 ± 0.0015	0.1552 ± 0.0013	0.1493 ± 0.0012	0.1657 ± 0.0012	0.1196 ± 0.0011
Geom	0.1630 ± 0.0014	0.1630 ± 0.0014	0.0105 ± 0.0000	0.7750 ± 0.0001	0.1657 ± 0.0012
Reco	0.0177 ± 0.0001	0.0110 ± 0.0001	0.8420 ± 0.0001		0.0031 ± 0.0000
PID	0.7824 ± 0.0002				0.8239 ± 0.0001
Trig	0.6720 ± 0.0029	0.8348 ± 0.0029	0.1531 ± 0.0012	0.0358 ± 0.0006	0.0795 ± 0.0009
MVA	0.9072 ± 0.0022	0.8249 ± 0.0032	0.8460 ± 0.0021	0.8530 ± 0.0020	0.8141 ± 0.0047
Tot	0.0064 ± 0.0001	0.0067 ± 0.0001	0.0012 ± 0.0000	0.0003 ± 0.0000	0.0006 ± 0.0000

Table 5.12: Relative efficiencies, $\varepsilon^{rel} = \varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$, for the ee and $\mu\mu$ channels in the central and high q^2 intervals.

ε	$\mu\mu$		ee		15–20 GeV^2/c^4 L0E
	1–6 GeV^2/c^4 L0E	1–6 GeV^2/c^4 L0H	1–6 GeV^2/c^4 L0I	$\mu\mu$	
Geom	1.0200 ± 0.0091		1.0429 ± 0.0084		1.0200 ± 0.0091
Reco	0.1873 ± 0.0011		0.1743 ± 0.0006		0.1159 ± 0.0009
PID	0.9602 ± 0.0002		0.9425 ± 0.0001		1.0334 ± 0.0001
Trig	0.8947 ± 0.0039	0.7893 ± 0.0065	2.2005 ± 0.0438	1.1247 ± 0.0138	1.1115 ± 0.0039
MVA	1.0143 ± 0.0025	0.9841 ± 0.0025	0.9495 ± 0.0023	0.9838 ± 0.0060	0.9223 ± 0.0036
Tot	0.7773 ± 0.0108	0.8912 ± 0.0131	2.3973 ± 0.0558	1.2694 ± 0.0230	0.8070 ± 0.0126

2392 5.8.1 Geometric efficiency

2393 The simulated samples used contain the requirement that daughters are in the LHCb
2394 detector acceptance. This corresponds to the requirement for each of the final par-
2395 ticles to have polar angle θ between 10 and 400 mrad. The efficiency of this cuts is
2396 obtained using a generator level simulated sample.

2397 5.8.2 Reconstruction efficiency and bin migration

2398 The reconstruction efficiency is here defined as the efficiency to reconstruct each
2399 decay channel given that its daughters are into the geometrical acceptance of the
2400 detector. This includes both the probability that a particle generates observable sig-
2401 natures and the efficiency of all the preselection cuts described in Sec. 5.6, including
2402 those done to remove peaking backgrounds. The efficiency of the PID cuts is kept
2403 separate as it is known to be not well simulated and there are reliable data-driven
2404 methods which can be used to extract it (see Sec. 5.8.3).

2405 It can happen that events generated in a q^2 interval are reconstructed in a different
2406 one, this is referred to as “bin migration”. Two different effects can cause bin
2407 migration. First of all, as the resolution of real detectors is not perfect, events
2408 close to the edges of the considered intervals can fall on the wrong side of the
2409 edge. This effect is only important in case of non-flat true distributions, as the
2410 amount of bin migration in the two directions is different. The second possible
2411 source of bin migration are systematic effects due, for example, to the presence
2412 of bremsstrahlung photons that cannot be recovered. It is particularly important
2413 to take into account the bin migration in the electron channels case because more
2414 photons are radiated from the final state. Figure 5.23 reports the response matrix for
2415 simulated $B^0 \rightarrow K^{*0} e^+ e^-$ events, which shows the correlation between reconstructed
2416 and generated q^2 . In the ideal case of perfect resolution this plot would look like a
2417 diagonal line and in case no bias is present its slope would be 1. Table 5.13 reports

²⁴¹⁸ net amounts of bin migration, M_{net} , in the considered q^2 intervals defined as:

$$M_{net} = N(\text{in} \rightarrow \text{in}) + N(\text{out} \rightarrow \text{in}) - N(\text{in} \rightarrow \text{out}) \quad (5.19)$$

²⁴¹⁹ where $N(\text{in} \rightarrow \text{in})$ is the number of candidates that are generated and reconstructed
²⁴²⁰ inside the considered interval, $N(\text{out} \rightarrow \text{in})$ the number of candidates that are
²⁴²¹ generated outside the interval but reconstructed inside and $N(\text{in} \rightarrow \text{out})$ the number
²⁴²² of candidates generated inside that fall outside.

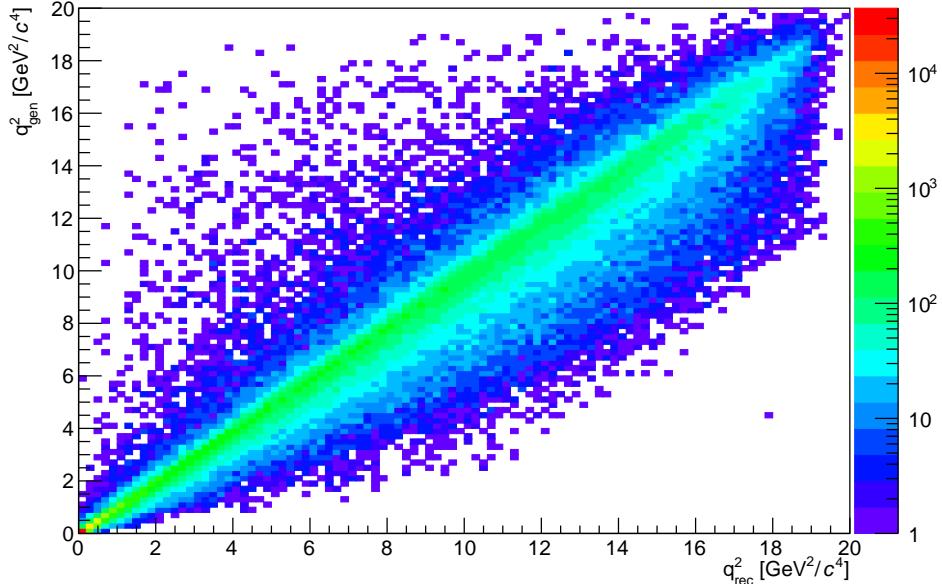


Figure 5.23: Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^* e^+ e^-$ events.

Table 5.13: Net bin migration amounts (M_{net}) in the considered q^2 intervals. Positive values indicate “net in”, negative values “net out”.

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

²⁴²³ The reconstruction efficiency is calculated comparing generated to reconstructed
²⁴²⁴ samples and therefore already includes bin migration effects. Nevertheless, it is
²⁴²⁵ useful to single out this component to better asses the corresponding systematic
²⁴²⁶ uncertainties.

²⁴²⁷ 5.8.3 PID efficiency

²⁴²⁸ The Monte Carlo is known not to reliably describe particle ID variables and therefore
²⁴²⁹ a data-driven method is used to obtain this efficiency component. This is done
²⁴³⁰ using the `PIDCalib` described in Sec. 2.8.1. Furthermore, the same method is used
²⁴³¹ to weight the MC in order to extract MVA and trigger efficiencies. The package
²⁴³² `PIDCalib` allows to divide the phase-space in bins and obtain a data-driven efficiency
²⁴³³ for each bin. For this analysis the phase-space is divided in equi-populated bins
²⁴³⁴ of momentum and pseudorapidity of the particle under study. Figure 5.24 shows
²⁴³⁵ performance tables for pions, kaons, muons and electrons.

²⁴³⁶ The decay channel under study generally has different kinematical distributions than
²⁴³⁷ the calibration sample. Therefore, once the efficiency table is obtained for each
²⁴³⁸ particle, the total efficiency for each candidate is calculated as the product of the
²⁴³⁹ four final particles efficiencies. $\varepsilon^{ev} = \varepsilon_K \cdot \varepsilon_\pi \cdot \varepsilon_{\ell_1} \cdot \varepsilon_{\ell_2}$. Finally, the total efficiency is
²⁴⁴⁰ found by averaging over all simulated events.

$$\varepsilon_{PID} = \frac{1}{N} \sum_i^N \varepsilon_K(p_K^i, \eta_K^i) \cdot \varepsilon_\pi(p_\pi^i, \eta_\pi^i) \cdot \varepsilon_\ell(p_{\ell_1}^i, \eta_{\ell_1}^i) \cdot \varepsilon_K(p_{\ell_2}^i, \eta_{\ell_2}^i) \quad (5.20)$$

²⁴⁴¹ 5.8.4 Trigger efficiency

²⁴⁴² While the muon channels the trigger efficiency is calculated using simulated events,
²⁴⁴³ for the electron channels a combination of simulation and data-driven methods is
²⁴⁴⁴ used. The efficiency of the software stage, HLT, is always obtained from simulation,
²⁴⁴⁵ while the efficiency of the hardware stage, L0, is obtained using a data-driven method
²⁴⁴⁶ as described in the next subsection. For both muon and electron channels it is
²⁴⁴⁷ possible to use the resonant sample to crosscheck the efficiency obtained using the
²⁴⁴⁸ simulation as explained in Sec. 5.8.4.2.

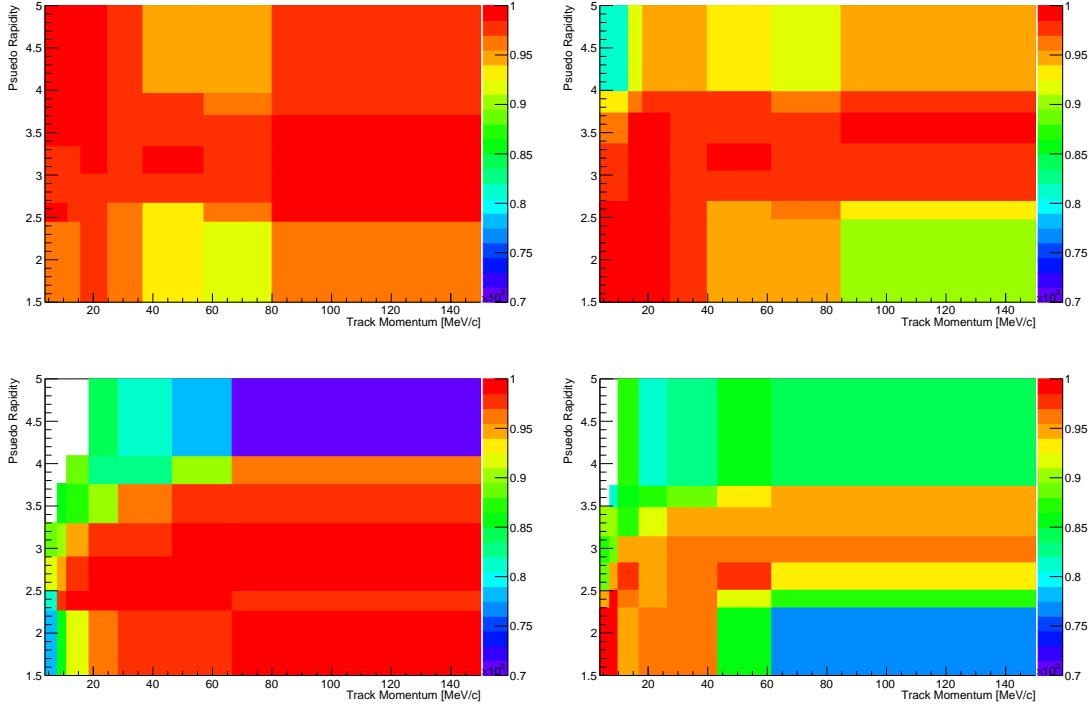


Figure 5.24: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

2449 5.8.4.1 Electron channels

2450 For the electron channels data is fitted separately in three trigger categories: L0Electron,
 2451 L0Hadon and L0TIS. Therefore we need to extract the efficiency separately for each
 2452 category.

2453 While the Hlt (1 and 2) efficiency is still computed using simulated events, the
 2454 L0Electron and L0Hadron efficiencies cannot be modelled with the Monte Carlo.
 2455 The discrepancy between data and simulation is mainly due to the ageing of the
 2456 calorimeters, on which the decision of these triggers relies. The ageing is not simu-
 2457 lated in the Monte Carlo and affects the L0 trigger efficiency which, therefore, must
 2458 be calibrated using data driven-methods. Tables of efficiencies are obtained applying
 2459 the TIS-TOS method to a calibration sample.

2460 For each trigger category these tables contain efficiency as a function of p_T of the
 2461 considered particle and are given for different calorimeter regions as these have

2462 different properties (e.g. cell size) due to the different position with respect to the
 2463 beam line. Regions considered are inner and outer HCAL, and inner, middle and
 2464 outer ECAL. Figure 5.25 shows data-driven efficiencies for the L0Electron trigger in
 the three ECAL regions.

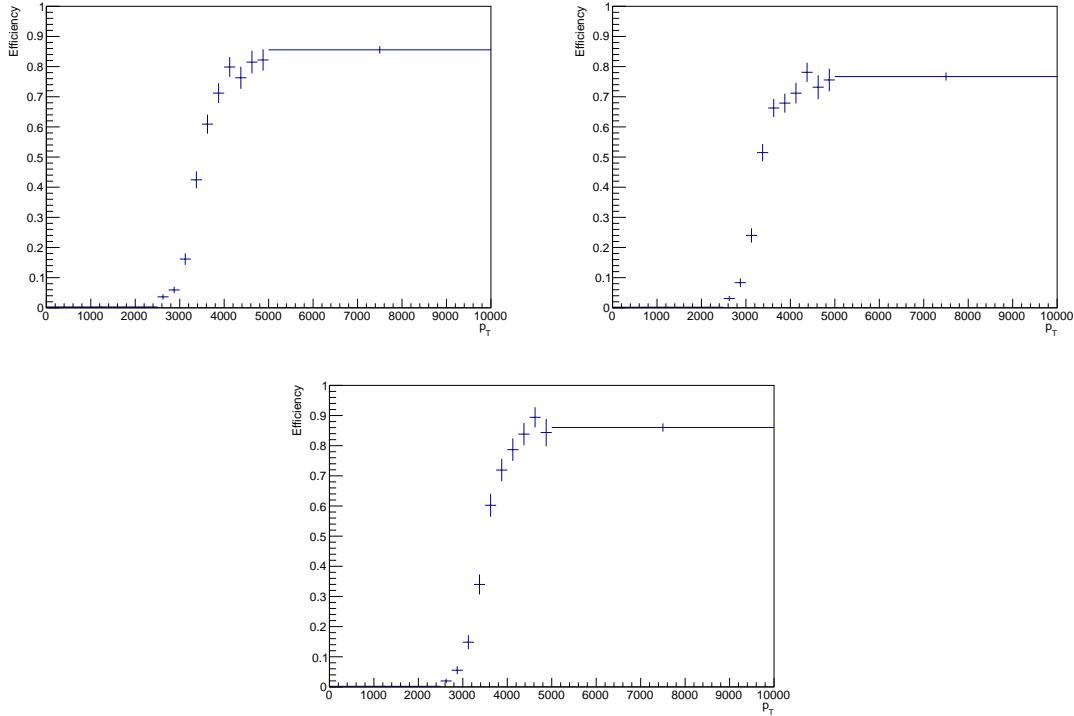


Figure 5.25: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

2465

2466 The probability of L0Electron trigger is calculated for each event as $P_{L0E} = \varepsilon(e^+) +$
 2467 $\varepsilon(e^-) - \varepsilon(e^+)\varepsilon(e^-)$. Similarly, the L0Hadron trigger probability is $P_{L0H} = \varepsilon(\pi) +$
 2468 $\varepsilon(K) - \varepsilon(\pi)\varepsilon(K)$. For L0TIS a probability $P_{L0I} = (1 - P_{L0E}) \cdot (1 - P_{L0H})$ is calculated.
 2469 Note that the probability of TIS trigger is defined to be independent of the signal
 2470 and therefore must be the same in the rare and resonant channels and cancel in
 2471 their ratio.

2472 Then event by event efficiencies for the three trigger categories are defined to be
 2473 exclusive in the following way:

- 2474 • L0E: $\varepsilon^{L0E} = P_{L0E}$, namely the probability that at least one electron triggered;

- 2475 • L0H: $\varepsilon^{L0H} = P_{L0H} \cdot (1 - P_{L0E})$, namely the probability that at least one hadron
2476 triggered but none of the electrons;
- 2477 • L0I: $\varepsilon^{L0I} = (1 - P_{L0H}) \cdot (1 - P_{L0E})$, namely the probability that neither the
2478 hadrons or the electrons in the event triggered. Note that in this case ε^{L0I}
2479 does not correspond to the efficiency of L0TIS trigger.

2480 As in the PID case, the total efficiency is found averaging over all events of a
2481 simulated sample:

$$\varepsilon^{\text{trg}} = \frac{1}{N} \sum_i^N \varepsilon^{\text{trg}}(p_T^i) \quad (5.21)$$

2482 where “trg” is a label indicating the trigger category under consideration.

2483 5.8.4.2 TISTOS cross-check

2484 The efficiency obtained using the simulation is crosschecked on resonant data us-
2485 ing the TISTOS method already described in Sec. 3.6.3. For this purpose the
2486 logical *and* of L0Global_TIS, Hlt1Phys_TIS and Hlt2Phys_TIS are used as con-
2487 trol sample (TIS). As data also contains non negligible amounts of background the
2488 DecayTreeFitter tool is used to constrain the J/ψ mass to its known value and a
2489 narrow interval around the peak, dominated by the signal, is selected. Finally, the
2490 $s\mathcal{P}\text{lot}$ method is used to remove residual background in the data sample. Results
2491 are shown in Tab. 5.14, where the efficiency obtained using the TISTOS method is
2492 compared between data and simulation. These are found to be in agreement for the
2493 muon channel, while they show deviations in the electron channels. In particular a
2494 significant discrepancy is found, for the L0I category, for which the procedure ex-
2495 plained in Sec. 5.8.4.1 does not ensure a correct calibration. The table also reports
2496 a correction factor obtained according to the formula

$$f = 1 + \frac{\varepsilon_{\text{date}}^{\text{TISTOS}} - \varepsilon_{\text{MC}}^{\text{TISTOS}}}{\varepsilon_{\text{MC}}^{\text{TISTOS}}}, \quad (5.22)$$

which is used to correct the resonant yields described in Sec. 5.10.1. To check the effect of this correction on the ratios between rare and resonant channels, the efficiency obtained using the TISTOS method is plotted as a function of the kinematic quantity relevant for each trigger category: the maximum p_T of the electrons for L0E, the maximum p_T of p and π for L0H, and the maximum p_T of all the final particles for L0I. Results are shown in Fig. 5.26. The p_T distributions of the rare and resonant simulated samples are then used to find the average efficiency in the two cases. This corresponds to making an average of the efficiency weighted by the kinematical distributions in the two samples. The ratios $\varepsilon_{\ell\ell}^{\text{tistos}}/\varepsilon_{J/\psi}^{\text{tistos}}$ obtained on data and simulation are then compared and found to be fully compatible. This means that even though the TISTOS correction has an effect on the absolute efficiency, this becomes negligible on the ratio between the rare and resonant channels. Therefore, no correction is applied for the extraction of the $R_{K^{*0}}$ ratio.

Table 5.14: Trigger efficiencies obtained using the TISTOS method on simulated and real $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \ell^+ \ell^-)$ decays.

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

2509

2510 5.8.5 Neural Networks efficiency

2511 The NN efficiency is again evaluated from fully weighted Monte Carlo samples. For
2512 the electron channels it is obtained separately for each trigger category.

2513 In order to cross check that this efficiency component is extracted correctly one
2514 can compare the efficiency obtained using $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+ \ell^-)$ events and rare
2515 $B^0 \rightarrow K^{*0}\ell^+\ell^-$ events in the same q^2 region selected for the resonant case. The ratio
2516 between the two should be close to 1 with small deviations due to the fact that the bin
2517 width is finite and the events are distributed differently inside the bin. This ratio is

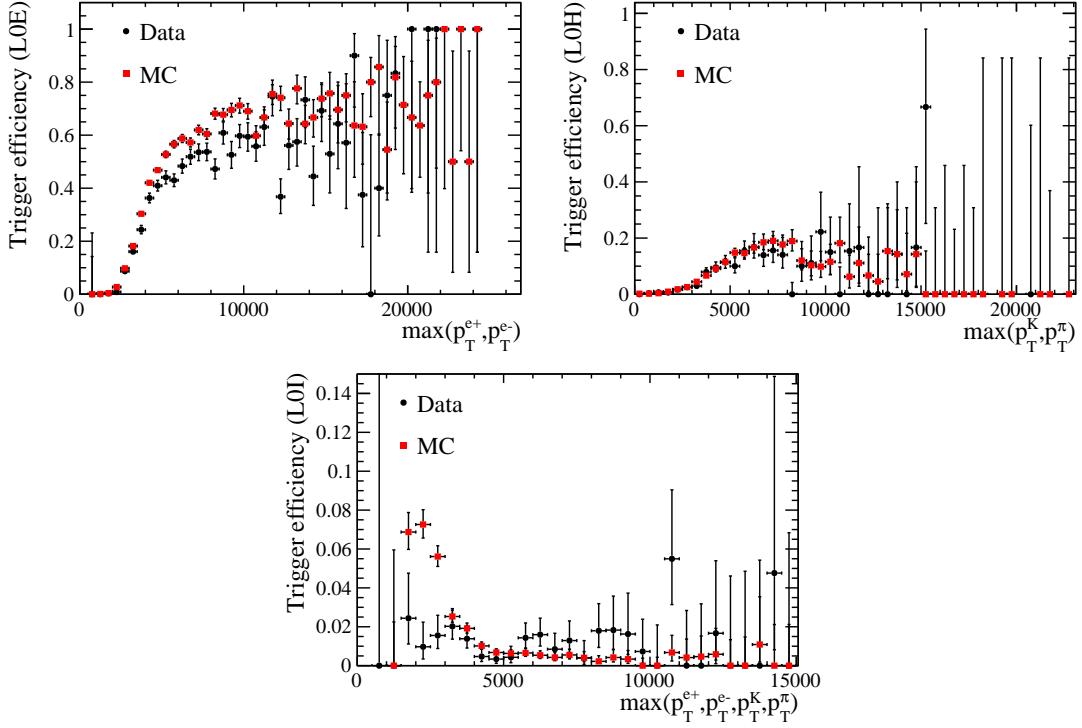


Figure 5.26: Trigger efficiency obtained applying the TISTOS method on $B^0 \rightarrow J/\psi K^*$ candidates as a function of the maximum p_T of the electrons for the L0E category (top left), the maximum p_T of p and π for L0H (top right) and the maximum p_T of all the final particles for L0I (bottom).

2518 found to be 0.997 ± 0.004 for the $\mu\mu$ channels and on average 0.981 ± 0.005 for the
 2519 ee channels. Values for the electron channels show a small deviation from one due
 2520 to the very large q^2 interval used to select the resonant channel ($[6,11]$ GeV^2/c^4).

2521 **5.9 Systematic uncertainties**

2522 The next subsections describe the four main sources of systematic uncertainties con-
2523 sidered. Other sources, which would matter in absolute quantities measurements,
2524 cancel in the ratio between the rare and resonant channels. Table 5.15 summarises
2525 the various sources of systematics and their effect on the R_{K^*} ratios. The total sys-
2526 tematic uncertainty is calculated as the square root sum of the single components.

Table 5.15: Summary of percent systematic uncertainties.

Source	1–6 GeV^2/c^4	15–20 GeV^2/c^4
Add swap	0.0	0.1
Free misreco	0.3	—
DCB	0.7	1.3
Eff.	2.1	2.4
Bin migration	5.5	6.9
Total	5.9	7.3

2527 **5.9.1 Choice of PDF**

2528 There is a certain arbitrariness in the choice of PDFs to model signal and background
2529 contributions in the invariant mass fits, which could translate in a bias on the final
2530 result. To assess this systematic the signal function for the electron channels is varied
2531 from the default CB plus Gaussian to a Double Crystal Ball function. This results
2532 in an uncertainty of $\sim 0.1\%$ for the central- q^2 region and $\sim 2.2\%$ for the high- q^2
2533 interval. Attempts to modify the PDF for the fit to muon channels result in a
2534 negligible systematic with respect to what found for the electron channels. The
2535 background PDF is changed in two ways. First of all a background component is
2536 added using the shape of simulated events, where the kaon and pion IDs have been
2537 swapped. Only the case in which this component is added both for the muon and
2538 electron channels is considered. This adds a $\sim 1\%$ systematic uncertainty. Secondly,
2539 the misreconstructed background yield in the rare electron channel is freed from its
2540 constraint to the background yield in the J/ψ sample. This is only relevant for the

2541 electron channels as there is no misreconstructed background in the muon ones.
2542 Furthermore, it is only relevant for the central- q^2 interval as in the fit to high- q^2
2543 candidates this yield is already free to float. This results in a 1.4% systematic
2544 uncertainty.

2545 5.9.2 Statistical error on the efficiency determination

2546 The statistical error on the determination of the efficiency is considered as a source
2547 of systematic uncertainty. This is $\sim 1.5\%$ on R_{ee} and $R_{\mu\mu}$ and, as it is due to fluctua-
2548 tions, it does not cancel in their ratio. This yields $\sim 2.5\%$ systematic uncertainty
2549 on the R_{K^*} measurement.

2550 5.9.3 TISTOS

2551 A further source of systematic, which is considered, is due to the discrepancies found
2552 in Sec. 5.8.4.2 on the determination of the trigger efficiency. As explained in that
2553 subsection the efficiency is derived as a function of the relevant kinematic quantities
2554 for each trigger category. The efficiency is then obtained for the rare and resonant
2555 sample by a weighted average. These are found to be compatible indicating that
2556 the effect on the ratios between rare and resonant channels is negligible. Therefore,
2557 no systematic uncertainty is added for this source.

2558 5.9.4 Bin migration

2559 The determination of the reconstruction efficiency is affected by the knowledge of
2560 the amount of bin migration as explained in Sec. 5.8.2. This amount depends on
2561 the shape of the q^2 distribution, which depends on the simulated model. In order
2562 to asses this systematic, simulated samples are generated using different models
2563 corresponding to different form factors [107, 108]. The generator level distributions

obtained using each model are compared with the ones obtained using the default one [109]. These ratios are shown in Fig. 5.27 as a function of q^2 and are used to re-weight the simulation. The amount of bin migration is recalculated using the simulation reweighted to reproduce each model. Table 5.16 reports the percent variation obtained with respect to the default model. The largest difference between two values is taken as systematic uncertainty. This results in a $\sim 5\%$ uncertainty for the central- q^2 interval and $\sim 11\%$ for the high- q^2 one.

Table 5.16: Percent variation on the bin migration amount obtained using different form factors models.

Model	1–6 GeV^2/c^4	15–20 GeV^2/c^4
Ball-Zwicky (6)	1.8	0.2
Melikhov-Stech	-3.7	6.6
Colangelo QCD (3)	0.3	0.8
Melikhov lattice (4)	-0.5	-0.4

2570

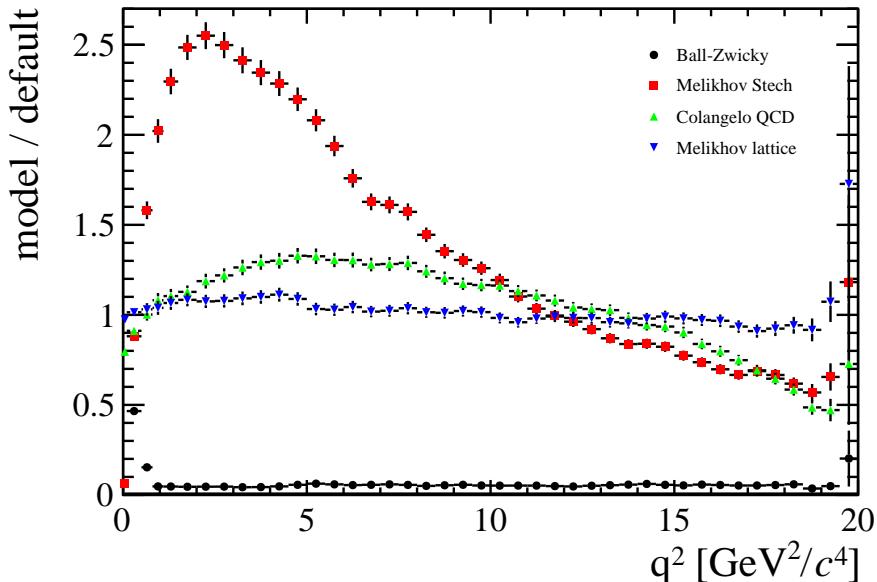


Figure 5.27: Ratios between the q^2 distributions obtained using different form factors models with respect to the default model.

2571 5.10 Result extraction

2572 In the following subsections are presented the final results of this analysis together
 2573 with the description of sanity checks performed to verify the stability of the methods
 2574 used.

2575 5.10.1 $R_{J/\psi}$ sanity check

2576 In order to cross-check the analysis procedure, the ratio between the measured
 2577 branching ratio of the electron and muonic resonant channels is calculated:

$$R_{J/\psi} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-))} = \frac{\varepsilon_{J/\psi(\mu\mu)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)}}{\varepsilon_{J/\psi(ee)} \cdot N_{B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-)}}. \quad (5.23)$$

2578 Unlike absolute branching fractions calculations, the determination of $R_{J/\psi}$ repre-
 2579 sents a better sanity test ad it is not affected by uncertainties due to the knowledge
 2580 of the amount of collected luminosity, \mathcal{L} , or the fragmentation fraction, f_d , the
 2581 probability for a b quark to produce a B^0 meson. These quantities come with large
 2582 uncertainties but they cancel in the $R_{J/\psi}$ ratio.

2583 Measured values of the $R_{J/\psi}$ ratio are reported in Tab. 5.17, where the error shown
 2584 is statistical only. For this purpose the trigger efficiencies are corrected using the
 2585 factors obtained in Sec. 5.8.4.2. Note that systematic uncertainties, which cancel
 2586 when doing the ratio between the rare and resonant channels with same leptonic
 2587 final state, do not cancel in this case. A reasonable agreement with unity is found.

Table 5.17: Fully corrected measured values of the ratio $R_{J/\psi}$ in the three electron trigger categories.

Trigger	$R_{J/\psi}$
L0E	1.028 ± 0.022
L0H	0.986 ± 0.072
L0I	0.973 ± 0.128

Table 5.18: 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}	0.00305 ± 0.00040	0.00406 ± 0.00081
$R_{\mu\mu}$	0.00242 ± 0.00011	0.00225 ± 0.00010
$R_{K^{*0}}$	blind	blind

2588 5.10.2 R_{K^*} result summary

2589 The ratio R_{K^*} is extracted by dividing the R_{ee} and $R_{\mu\mu}$ parameters described in
 2590 Sec. 5.7. These ratios are floating in the fit but they can also be built from the
 2591 yields in Tab. 5.8 and the efficiencies in Tab. 5.10. In summary the definition of the
 2592 R_{K^*} ratio is the following :

$$R_{K^*} = \frac{R_{ee}}{R_{\mu\mu}} = \frac{N_{ee}}{N_{J/\psi(ee)}} \cdot \frac{N_{J/\psi(\mu\mu)}}{N_{\mu\mu}} \cdot \frac{\varepsilon_{J/\psi(ee)}}{\varepsilon_{ee}} \cdot \frac{\varepsilon_{\mu\mu}}{\varepsilon_{J/\psi(\mu\mu)}}. \quad (5.24)$$

2593 As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three
 2594 electron categories its value is already a combination of the three samples. Results
 2595 are shown in Tab. 5.18.

2596 5.10.3 Branching ratios and expectations

2597 Multiplying the ratios R_{ee} and $R_{\mu\mu}$ by the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ [2]
 2598 branching ratios one can obtain absolute branching ratios for the rare channels:

$$\begin{aligned} \mathcal{B}(B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)) &= \mathcal{B}(B^0 \rightarrow K^*J/\psi) \times \mathcal{B}(B^0 \rightarrow \ell\ell) \\ &= (1.32 \pm 0.06)10^{-3} \times (5.96 \pm 0.03)10^{-2} = (7.87 \pm 0.36) \times 10^{-5} \end{aligned} \quad (5.25)$$

2599 Table 5.19 reports absolute branching ratio values for the rare channels in the con-
 2600 sidered q^2 intervals, where the errors are statistical only.

2601 The results for the central q^2 bin can be compared also with SM predictions obtained
 2602 from Ref. [110]. This paper reports branching predicted ratios in the $1 < q^2 < 6$ GeV^2/c^4

Table 5.19: 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	$(1.80 \pm 0.24) \times 10^{-7}$	$(3.19 \pm 0.64) \times 10^{-7}$
$\mu\mu$	$(2.07 \pm 0.10) \times 10^{-7}$	$(1.92 \pm 0.09) \times 10^{-7}$

2603 interval for the rare electronic and muonic channels. These are rescaled to the range
 2604 $1.1 < q^2 < 6 \text{ GeV}^2/c^4$ using the simulation. Finally, the measured value of the
 2605 measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ decay is used. The predicted ratio is found to
 2606 be 0.75 ± 0.14 , which is in agreement with our measurement within one standard
 2607 deviation. In Tab. 5.20 are also reported observed and expected ratios of rare over
 2608 resonant raw numbers of candidates ($N_{\ell\ell}/N_{J/\psi}$). In this table the observed ratios are
 2609 simply obtained dividing the rare and resonant yields in Tab. 5.8 and the expected
 2610 ones are obtained using the predicted rare channel branching ratios and the mea-
 2611 sured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ branching ratio, rescaled by the relative efficiencies in
 2612 Tab. 5.12.

Table 5.20: Expected and observed ratios of raw event yields, $N_{\ell\ell}/N_{J/\psi}$.

Sample	Expected	Observed	Obs / exp ratio
$\mu\mu$	0.00253 ± 0.00084	0.00188 ± 0.00009	0.74309 ± 0.24866
ee (L0E)	0.00269 ± 0.00084	0.00271 ± 0.00035	
ee (L0H)	0.00723 ± 0.00227	0.00732 ± 0.00098	1.00826 ± 0.34265
ee (L0I)	0.00383 ± 0.00120	0.00388 ± 0.00051	

2613

CHAPTER 6

2614

2615

Conclusions

2616

2617

2618

REFERENCES

2619

- [1] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays*, *JHEP* **1506** (2015) 115, [[arXiv:1503.07138](#)].
- [2] **Particle Data Group** Collaboration, K. Olive et al., *Review of Particle Physics*, *Chin.Phys. C* **38** (2014) 090001.
- [3] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, *Phys.Rev.* **D20** (1979) 2619–2625.
- [4] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961) 579–588.
- [5] **LHCb** Collaboration, R. Aaij et al., *Observation of the resonant character of the $Z(4430)^-$ state*, *Phys. Rev. Lett.* **112** (2014), no. 22 222002, [[arXiv:1404.1903](#)].
- [6] **LHCb** Collaboration, R. Aaij et al., *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi K^- p$ decays*, [arXiv:1507.03414](#).
- [7] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, *Phys.Rev.* **105** (1957) 1413–1414.
- [8] F. Strocchi, *Spontaneous Symmetry Breaking in Local Gauge Quantum Field Theory: The Higgs Mechanism*, *Commun.Math.Phys.* **56** (1977) 57.
- [9] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, et al., *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, *Phys.Rev.* **D91** (2015), no. 7 073007, [[arXiv:1501.05013](#)].

- [10] F. Zwicky, *Spectral displacement of extra galactic nebulae*, *Helv.Phys.Acta* **6** (1933) 110–127.
- [11] M. Gavela and Hernandez, *Standard model CP violation and baryon asymmetry*, *Mod.Phys.Lett.* **A9** (1994) 795–810, [[hep-ph/9312215](#)].
- [12] M. Maltoni, *Status of three-neutrino oscillations*, *PoS EPS-HEP2011* (2011) 090.
- [13] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, *Astrophys. J.* **496** (1998) 505–526.
- [14] **Super-Kamiokande** Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [[hep-ex/9807003](#)].
- [15] **KamLAND** Collaboration, K. Eguchi et al., *First results from KamLAND: Evidence for reactor anti-neutrino disappearance*, *Phys. Rev. Lett.* **90** (2003) 021802, [[hep-ex/0212021](#)].
- [16] J. L. Feng, *Naturalness and the Status of Supersymmetry*, *Ann.Rev.Nucl.Part.Sci.* **63** (2013) 351–382, [[arXiv:1302.6587](#)].
- [17] **ATLAS** Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [[arXiv:1207.7214](#)].
- [18] P. Fayet and S. Ferrara, *Supersymmetry*, *Phys.Rept.* **32** (1977) 249–334.
- [19] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys.Rev.Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [20] G. Isidori and D. M. Straub, *Minimal Flavour Violation and Beyond*, *Eur.Phys.J.* **C72** (2012) 2103, [[arXiv:1202.0464](#)].
- [21] A. J. Buras, *Minimal flavor violation*, *Acta Phys.Polon.* **B34** (2003) 5615–5668, [[hep-ph/0310208](#)].
- [22] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*, *Ann.Rev.Nucl.Part.Sci.* **65** (2015) 8007, [[arXiv:1501.03309](#)].
- [23] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, *Can we reach the Zeptouniverse with rare K and $B_{s,d}$ decays?*, *JHEP* **1411** (2014) 121, [[arXiv:1408.0728](#)].
- [24] G. Hiller and M. Schmaltz, *R_K and future $b \rightarrow s\ell\ell$ physics beyond the standard model opportunities*, *Phys.Rev.* **D90** (2014) 054014, [[arXiv:1408.1627](#)].

- [25] K. G. Chetyrkin, M. Misiak, and M. Munz, *Weak radiative B meson decay beyond leading logarithms*, *Phys.Lett.* **B400** (1997) 206–219, [[hep-ph/9612313](#)].
- [26] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev.Mod.Phys.* **68** (1996) 1125–1144, [[hep-ph/9512380](#)].
- [27] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, [hep-ph/9806471](#).
- [28] M. Della Morte, J. Heitger, H. Simma, and R. Sommer, *Non-perturbative Heavy Quark Effective Theory: An application to semi-leptonic B -decays*, *Nucl.Part.Phys.Proc.* **261-262** (2015) 368–377, [[arXiv:1501.03328](#)].
- [29] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, *Phys.Rev.* **D63** (2001) 114020, [[hep-ph/0011336](#)].
- [30] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \rightarrow K^{(*)}\ell^+\ell^-$ and $B \rightarrow K^*\gamma$* , *JHEP* **1009** (2010) 089, [[arXiv:1006.4945](#)].
- [31] **LHCb** Collaboration, R. Aaij et al., *Observation of a resonance in $B^+ \rightarrow K^+\mu^+\mu^-$ decays at low recoil*, *Phys. Rev. Lett.* **111** (2013) 112003, [[arXiv:1307.7595](#)].
- [32] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, et al., *$B_{s,d} \rightarrow l+l-$ in the Standard Model with Reduced Theoretical Uncertainty*, *Phys.Rev.Lett.* **112** (2014) 101801, [[arXiv:1311.0903](#)].
- [33] **CMS, LHCb** Collaboration, V. Khachatryan et al., *Observation of the rare $B_s^0 \rightarrow \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, *Nature* **522** (2015) 68–72, [[arXiv:1411.4413](#)].
- [34] **LHCb** Collaboration, R. Aaij et al., *Differential branching fractions and isospin asymmetry of $B \rightarrow K^{(*)}\mu^+\mu^-$ decays*, *JHEP* **06** (2014) 133, [[arXiv:1403.8044](#)].
- [35] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$* , *JHEP* **07** (2013) 084, [[arXiv:1305.2168](#)].
- [36] **LHCb** Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$* , *JHEP* **08** (2013) 131, [[arXiv:1304.6325](#)].
- [37] **LHCb** Collaboration, R. Aaij et al., *Measurement of form-factor-independent observables in the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$* , *Phys. Rev. Lett.* **111** (2013) 191801, [[arXiv:1308.1707](#)].

- ²⁷¹⁴ [38] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \rightarrow K^* \mu^+ \mu^-$ Anomaly*, *Phys. Rev.* **D88** (2013), no. 7 074002, [[arXiv:1307.5683](#)].
- ²⁷¹⁶ [39] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of charged and neutral $B \rightarrow K \mu^+ \mu^-$ decays*, *JHEP* **05** (2014) 082, [[arXiv:1403.8045](#)].
- ²⁷¹⁸ [40] **LHCb** Collaboration, R. Aaij et al., *Measurement of CP asymmetries in the decays $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^+ \rightarrow K^+ \mu^+ \mu^-$* , *JHEP* **1409** (2014) 177, [[arXiv:1408.0978](#)].
- ²⁷²¹ [41] **LHCb** Collaboration, R. Aaij et al., *Measurement of the $B^0 \rightarrow K^{*0} e^+ e^-$ branching fraction at low dilepton mass*, *JHEP* **05** (2013) 159, [[arXiv:1304.3035](#)].
- ²⁷²⁴ [42] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay in the low- q^2 region*, *JHEP* **04** (Jan, 2015) 064. 18 p.
- ²⁷²⁶ [43] **MEGA** Collaboration, M. Ahmed et al., *Search for the lepton family number nonconserving decay $\mu u^+ \rightarrow e^+ \gamma$* , *Phys. Rev.* **D65** (2002) 112002, [[hep-ex/0111030](#)].
- ²⁷²⁹ [44] **SINDRUM** Collaboration, U. Bellgardt et al., *Search for the Decay $\mu u^+ \rightarrow e^+ e^+ e^-$* , *Nucl. Phys.* **B299** (1988) 1.
- ²⁷³¹ [45] **LHCb** Collaboration, R. Aaij et al., *Search for the lepton-flavour-violating decays $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow e^\pm \mu^\mp$* , *Phys. Rev. Lett.* **111** (2013) 141801, [[arXiv:1307.4889](#)].
- ²⁷³⁴ [46] **LHCb** Collaboration, R. Aaij et al., *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, *Phys. Lett.* **B724** (2013) 36, [[arXiv:1304.4518](#)].
- ²⁷³⁷ [47] W. J. Marciano, T. Mori, and J. M. Roney, *Charged Lepton Flavor Violation Experiments*, *Ann. Rev. Nucl. Part. Sci.* **58** (2008) 315–341.
- ²⁷³⁹ [48] L. Evans, *The LHC machine*, *PoS EPS-HEP2009* (2009) 004.
- ²⁷⁴⁰ [49] **LHCb** Collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.
- ²⁷⁴² [50] **LHCb** Collaboration, R. Aaij et al., *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, *Phys. Lett.* **B694** (2010) 209–216, [[arXiv:1009.2731](#)].
- ²⁷⁴⁵ [51] M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J.* **C73** (2013) 2431, [[arXiv:1211.6759](#)].
- ²⁷⁴⁷ [52] A. A. Alves Jr. et al., *Performance of the LHCb muon system*, *JINST* **8** (2013) P02022, [[arXiv:1211.1346](#)].

- [53] **LHCb** Collaboration, R. e. a. Aaij, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030.
- [54] **LHCb** Collaboration, R. e. a. Aaij, *LHCb Detector Performance*, *Int. J. Mod. Phys. A* **30** (Dec, 2014) 1530022. 82 p.
- [55] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl.Instrum.Meth.* **A555** (2005) 356–369, [[physics/0402083](#)].
- [56] R. Aaij et al., *The LHCb trigger and its performance in 2011*, *JINST* **8** (2013) P04022, [[arXiv:1211.3055](#)].
- [57] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].
- [58] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [[arXiv:0710.3820](#)].
- [59] I. Belyaev et al., *Handling of the generation of primary events in GAUSS, the LHCb simulation framework*, *Nuclear Science Symposium Conference Record (NSS/MIC)* **IEEE** (2010) 1155.
- [60] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.
- [61] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur.Phys.J.* **C45** (2006) 97–107, [[hep-ph/0506026](#)].
- [62] **Geant4 collaboration** Collaboration, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, et al., *Geant4 developments and applications*, *IEEE Trans.Nucl.Sci.* **53** (2006) 270.
- [63] M. Clemencic et al., *The LHCb simulation application, GAUSS: design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023.
- [64] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, *Conf.Proc.* **C000917** (2000) 11–42.
- [65] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, *Nucl.Instrum.Meth.* **A559** (2006) 190–194.
- [66] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, [physics/0402093](#).
- [67] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, *Nucl.Instrum.Meth.* **A552** (2005) 566–575, [[physics/0503191](#)].
- [68] H. W. Bertini, *Low-Energy Intranuclear Cascade Calculation*, *Phys. Rev.* **131** (1963) 1801–1821.

- [69] B. Andersson, G. Gustafson, and H. Pi, *The FRITIOF model for very high-energy hadronic collisions*, *Z. Phys.* **C57** (1993) 485–494.
- [70] COMPASS Collaboration, P. Abbon et al., *The COMPASS experiment at CERN*, *Nucl. Instrum. Meth.* **A577** (2007) 455–518, [[hep-ex/0703049](#)].
- [71] G. Hiller, M. Knecht, F. Legger, and T. Schietinger, *Photon polarization from helicity suppression in radiative decays of polarized Lambda(b) to spin-3/2 baryons*, *Phys.Lett.* **B649** (2007) 152–158, [[hep-ph/0702191](#)].
- [72] T. Mannel and S. Recksiegel, *Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \rightarrow \Lambda \gamma$* , *J.Phys.* **G24** (1998) 979–990, [[hep-ph/9701399](#)].
- [73] M. J. Aslam, Y.-M. Wang, and C.-D. Lu, *Exclusive semileptonic decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in supersymmetric theories*, *Phys.Rev.* **D78** (2008) 114032, [[arXiv:0808.2113](#)].
- [74] Y.-m. Wang, Y. Li, and C.-D. Lu, *Rare Decays of $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in the Light-cone Sum Rules*, *Eur.Phys.J.* **C59** (2009) 861–882, [[arXiv:0804.0648](#)].
- [75] C.-S. Huang and H.-G. Yan, *Exclusive rare decays of heavy baryons to light baryons: $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Rev.* **D59** (1999) 114022, [[hep-ph/9811303](#)].
- [76] C.-H. Chen and C. Geng, *Rare $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decays with polarized lambda*, *Phys.Rev.* **D63** (2001) 114024, [[hep-ph/0101171](#)].
- [77] C.-H. Chen and C. Geng, *Baryonic rare decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Rev.* **D64** (2001) 074001, [[hep-ph/0106193](#)].
- [78] C.-H. Chen and C. Geng, *Lepton asymmetries in heavy baryon decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Phys.Lett.* **B516** (2001) 327–336, [[hep-ph/0101201](#)].
- [79] F. Zolfagharpour and V. Bashiry, *Double Lepton Polarization in $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ Decay in the Standard Model with Fourth Generations Scenario*, *Nucl.Phys.* **B796** (2008) 294–319, [[arXiv:0707.4337](#)].
- [80] L. Mott and W. Roberts, *Rare dileptonic decays of Λ_b^0 in a quark model*, *Int.J.Mod.Phys.* **A27** (2012) 1250016, [[arXiv:1108.6129](#)].
- [81] T. Aliev, K. Azizi, and M. Savci, *Analysis of the $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decay in QCD*, *Phys.Rev.* **D81** (2010) 056006, [[arXiv:1001.0227](#)].
- [82] R. Mohanta and A. Giri, *Fourth generation effect on Λ_b decays*, *Phys.Rev.* **D82** (2010) 094022, [[arXiv:1010.1152](#)].
- [83] S. Sahoo, C. Das, and L. Maharana, *Effect of both Z and Z'-mediated flavor-changing neutral currents on the baryonic rare decay $\Lambda_b^0 \rightarrow \Lambda l^+l^-$* , *Int.J.Mod.Phys.* **A24** (2009) 6223–6235, [[arXiv:1112.4563](#)].

- [84] **CDF Collaboration** Collaboration, T. Aaltonen et al., *Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \rightarrow \Lambda\mu^+\mu^-$* , *Phys.Rev.Lett.* **107** (2011) 201802, [[arXiv:1107.3753](#)].
- [85] **CDF** Collaboration, S. Behari, *CDF results on $b \rightarrow s\mu\mu$ decays*, [arXiv:1301.2244](#).
- [86] **LHCb** Collaboration, R. Aaij et al., *Measurement of the differential branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$* , *Phys. Lett.* **B725** (2013) 25, [[arXiv:1306.2577](#)].
- [87] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, *Rare baryon decays $\Lambda_b \rightarrow \Lambda l^+l^- (l = e, \mu, \tau)$ and $\Lambda_b \rightarrow \Lambda\gamma$: differential and total rates, lepton- and hadron-side forward-backward asymmetries*, *Phys.Rev.* **D87** (2013) 074031, [[arXiv:1301.3737](#)].
- [88] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys.Lett.* **B724** (2013) 27, [[arXiv:1302.5578](#)].
- [89] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. [physics/0308063](#).
- [90] 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.
- [91] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, *$\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ form factors and differential branching fraction from lattice QCD*, *Phys. Rev.* **D87** (2013), no. 7 074502, [[arXiv:1212.4827](#)].
- [92] **LHCb** Collaboration, R. Aaij et al., *Precision measurement of the Λ_b^0 baryon lifetime*, *Phys.Rev.Lett.* **111** (2013) 102003, [[arXiv:1307.2476](#)].
- [93] 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.
- [94] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343–367.
- [95] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Lett.* **B724** (2013) 27, [[arXiv:1302.5578](#)].

- 2858 [96] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical*
 2859 *statistical analysis of small signals*, *Phys.Rev.* **D57** (1998) 3873–3889,
 2860 [[physics/9711021](#)].
- 2861 [97] T. M. Karbach, *Feldman-Cousins Confidence Levels - Toy MC Method*,
 2862 [[arXiv:1109.0714](#)].
- 2863 [98] S. Meinel, *Flavor physics with Λ_b baryons*, *PoS LATTICE2013* (2014) 024,
 2864 [[arXiv:1401.2685](#)].
- 2865 [99] G. Hiller and F. Kruger, *More model independent analysis of $b \rightarrow s$*
 2866 *processes*, *Phys.Rev.* **D69** (2004) 074020, [[hep-ph/0310219](#)].
- 2867 [100] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \rightarrow s\ell\ell$* ,
 2868 *JHEP* **1502** (2015) 055, [[arXiv:1411.4773](#)].
- 2869 [101] **BaBar Collaboration** Collaboration, J. Lees et al., *Measurement of*
 2870 *Branching Fractions and Rate Asymmetries in the Rare Decays*
 2871 *$B \rightarrow K^{(*)}l^+l^-$* , *Phys.Rev.* **D86** (2012) 032012, [[arXiv:1204.3933](#)].
- 2872 [102] **BELLE Collaboration** Collaboration, J.-T. Wei et al., *Measurement of the*
 2873 *Differential Branching Fraction and Forward-Backward Asymmetry for*
 2874 *$B \rightarrow K^{(*)}l^+l^-$* , *Phys.Rev.Lett.* **103** (2009) 171801, [[arXiv:0904.0770](#)].
- 2875 [103] **LHCb** Collaboration, R. Aaij et al., *Test of lepton universality using*
 2876 *$B^+ \rightarrow K^+\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601,
 2877 [[arXiv:1406.6482](#)].
- 2878 [104] “Lhcb loki twiki.”
 2879 <https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters>.
 2880 Accessed: 2015-09-30.
- 2881 [105] “Probnn presentation at ppts meeting.” <https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf>. Accessed: 2015-09-30.
- 2883 [106] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf C0303241* (2003) MOLT007, [[physics/0306116](#)].
- 2885 [107] P. Ball and R. Zwicky, *New results on $B \rightarrow \pi, K, \eta$ decay form factors from*
 2886 *light-cone sum rules*, *Phys. Rev.* **D71** (2005) 014015, [[hep-ph/0406232](#)].
- 2887 [108] D. Melikhov and B. Stech, *Weak form-factors for heavy meson decays: An*
 2888 *Update*, *Phys. Rev.* **D62** (2000) 014006, [[hep-ph/0001113](#)].
- 2889 [109] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *A Comparative study of the*
 2890 *decays $B \rightarrow (K, K^*)\ell^+\ell^-$ in standard model and supersymmetric theories*,
 2891 *Phys. Rev.* **D61** (2000) 074024, [[hep-ph/9910221](#)].
- 2892 [110] A. Ali, E. Lunghi, C. Greub, and G. Hiller, *Improved model independent*
 2893 *analysis of semileptonic and radiative rare B decays*, *Phys.Rev.* **D66** (2002)
 2894 034002, [[hep-ph/0112300](#)].

- 2895 [111] J. Hrivnac, R. Lednický, and M. Smizanska, *Feasibility of beauty baryon*
2896 *polarization measurement in $\Lambda^0 J/\psi$ decay channel by ATLAS LHC,*
2897 *J.Phys.G* **G21** (1995) 629–638, [[hep-ph/9405231](#)].
- 2898 [112] **LHCb** Collaboration, R. Aaij et al., *Determination of the quark coupling*
2899 *strength $|V_{ub}|$ using baryonic decays*, [arXiv:1504.01568](#).
- 2900 [113] A. Crivellin and S. Pokorski, *Can the differences in the determinations of*
2901 *V_{ub} and V_{cb} be explained by New Physics?*, *Phys. Rev. Lett.* **114** (2015),
2902 no. 1 011802, [[arXiv:1407.1320](#)].
- 2903 [114] W. Altmannshofer and D. M. Straub, *New physics in $b \rightarrow s$ transitions after*
2904 *LHC run 1*, [arXiv:1411.3161](#).

2905

APPENDIX A

2906

2907

Decay models

2908

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

2910 The q^2 and angular dependancies of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays are modelled based on
 2911 Ref. [87], where the angular distribution for unpolarised Λ_b^0 production is defined as

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

2912 In this formula θ_ℓ and θ_B correspond to the lepton and proton helicity angles, χ
 2913 is angle between dimuon and Λ decay planes (for unpolarised production we are
 2914 sensitive only to difference in azimuthal angles), $d_{i,j}^J$ are Wigner d-functions and h ,
 2915 h^B and H are helicity amplitudes for virtual dimuon, Λ and Λ_b^0 decays. The sum
 2916 runs over all possible helicities with the dimuon being allowed in spin 0 and 1 states
 2917 (J and J'). The m and m' indices run over the vector and axial-vector current
 2918 contributions.

2919 The production polarisation is introduced by removing $e^{i(\lambda_j - \lambda'_j)\chi}$ from the expression,
 2920 swapping small Wigner d-functions $d_{i,j}^J$ to the corresponding capital ones $D_{i,j}^J$ which
 2921 are related as

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

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

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

Those changes lead to the formula

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

The lepton amplitudes come directly from Ref. [87], eq. 3. The Λ decay amplitudes are related to the Λ decay asymmetry parameter as

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

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

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

Finally, the remaining amplitudes are expressed in terms of form factors (Ref. [87], eq. C6) as

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

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

2930 terms of dimensionless quantities in eqs. C8 and C9 in Ref. [87]. In our actual
2931 implementation form factors calculated in the covariant quark model [87] are used
2932 and for the numerical values of the Wilson coefficients Ref. [87] 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. [91]. The form factors relations are found by comparing eqs. 66 and 68 in Ref. [87] to eq. 51 in Ref. [91]. Denoting LQCD form factors by F_i^L and dimensionless covariant quark model ones by f_i^{XX} we have

$$\begin{aligned} f_1^V &= c_\gamma(F_1^L + F_2^L), \\ f_2^V &= -2c_\gamma F_2^L, \\ f_3^V &= c_v(F_1^L + F_2^L), \\ f_1^A &= c_\gamma(F_1^L - F_2^L), \\ f_2^A &= -2c_\gamma F_2^L, \\ f_3^A &= -c_v(F_1^L - F_2^L), \\ f_1^{TV} &= c_\sigma F_2^L, \\ f_2^{TV} &= -c_\sigma F_1^L, \\ f_1^{TA} &= c_\sigma F_2^L, \\ f_2^{TA} &= -c_\sigma F_1^L, \end{aligned}$$

where

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

2933 In the calculations $\mu = m_b$ is used. For the strong coupling constant, we start
2934 from the world average value at the Z mass, $\alpha_s(m_Z^2) = 0.1185 \pm 0.0006$ [2], and we
2935 translate it to the scale m_b^2 by

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

2936 where $n_f = 5$. The LQCD form factors F_1^L and F_2^L can be then taken directly from
2937 Ref. [91] and plugged into the code implementing the calculation from Ref. [87].

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

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

2938 **A.2 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution**

2939 The angular distribution of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay is modelled using Ref. [111]. The
2940 differential rate is written as

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

2941 where f_{1i} , f_{2i} and F_i are listed in Tab. A.1. The expression uses four observables
2942 (angles) and depends on four complex amplitudes a_+ , a_- , b_+ , b_- and two real valued
2943 parameters for the production polarisation, P_b , and the Λ decay asymmetry, α_Λ . The
2944 angle θ is the angle of the Λ momentum in Λ_b^0 rest frame with respect to the vector
2945 $\vec{n} = \frac{\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}}{|\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}|}$, where \vec{p}_{inc} and $\vec{p}_{\Lambda_b^0}$ are the momenta of incident proton and Λ_b^0 in the
2946 center of mass system. The angles θ_1 and ϕ_1 are polar and azimuthal angle of the
2947 proton coming from the Λ decay in the Λ rest frame with axis defined as $z_1 \uparrow\uparrow \vec{p}_\Lambda$,
2948 $y_1 \uparrow\uparrow \vec{n} \times \vec{p}_\Lambda$. Finally, the angles θ_2 and ϕ_2 are the angles of the momenta of the
2949 muons in J/ψ rest frame with axes defined as $z_2 \uparrow\uparrow \vec{p}_{J/\psi}$, $y_2 \uparrow\uparrow \vec{n} \times \vec{p}_{J/\psi}$.

2950 The distribution depends on the Λ decay asymmetry parameter, α_Λ , the production
2951 polarisation P_b and four complex amplitudes. The α_Λ is measured to be 0.642 ± 0.013
2952 for Λ . The production polarisation P_b and magnitudes of a_+ , a_- , b_+ and b_- are
2953 measured in Ref. [95]. Phases are not measured therefore, as default all phases are
2954 set to zero and then they are randomly varied to calculate the systematic uncertainty.

2955

APPENDIX B

2956

2957

Data-simulation comparison

2958

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

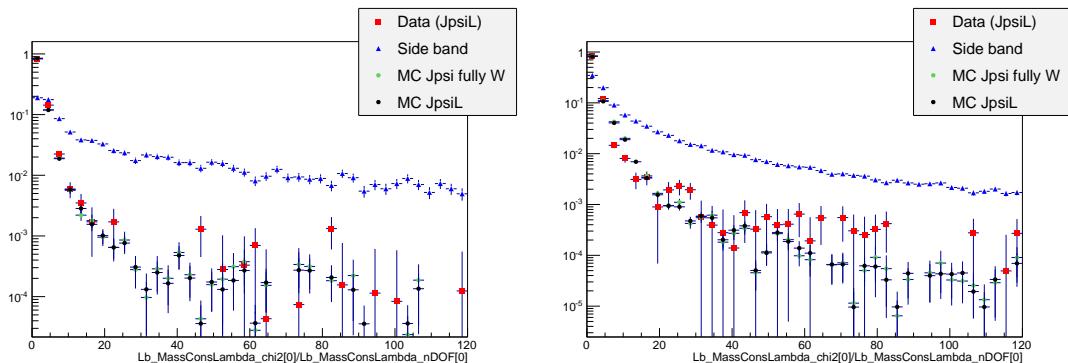


Figure B.1: Distributions of χ^2/NdF of the kinematic fit in data and simulation for LL (left) and DD (right) events.

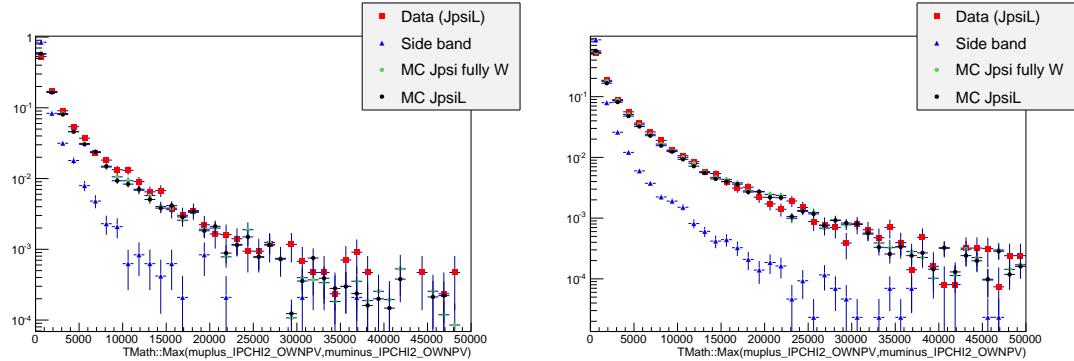


Figure B.2: Distributions of maximum muon $IP\chi^2$ variable in data and simulation for LL (left) and DD (right) events.

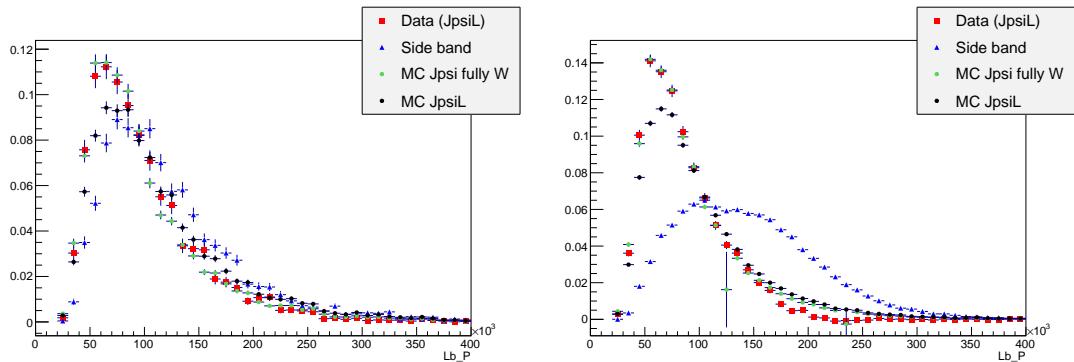


Figure B.3: Distributions of Λ_b^0 momentum variable in data and simulation for LL (left) and DD (right) events.

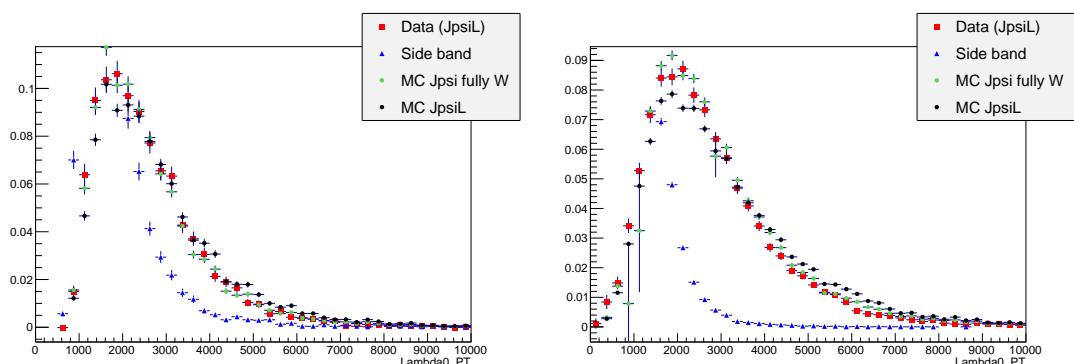


Figure B.4: Distributions of Λ transverse momentum variable in MC, data signal and data background for LL (left) and DD (right) events.

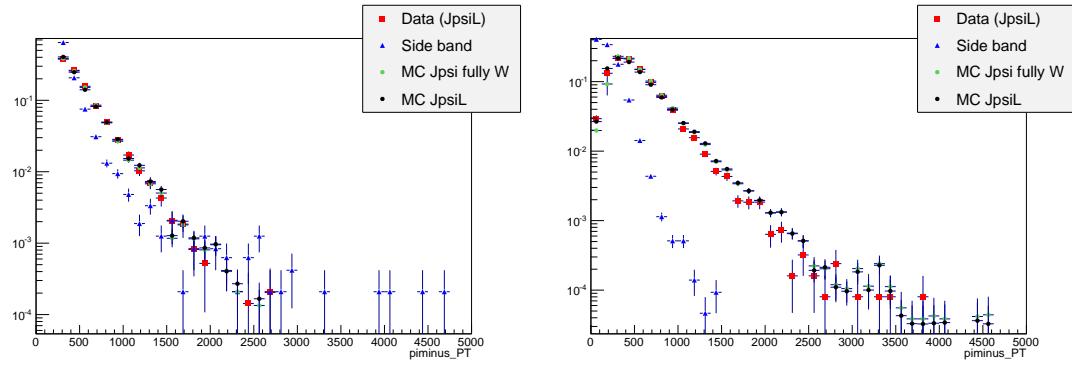


Figure B.5: Distributions of pion transverse momentum variable in data and simulation for LL (left) and DD (right) events.

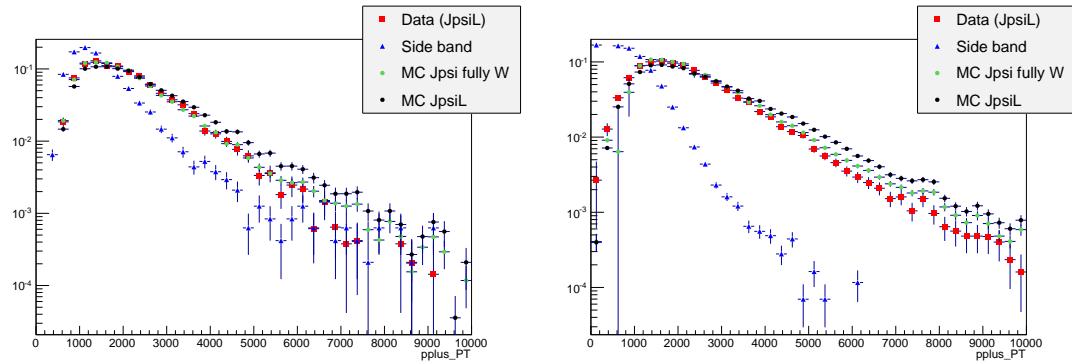


Figure B.6: Distributions of proton transverse momentum variable in data and simulation for LL (left) and DD (right) events.

2968

APPENDIX C

2969

2970 **Systematic uncertainties on the efficiency calculation for the**
 2971 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ **branching fraction analysis.**

2972

2973 This appendix reports systematic uncertainties on absolute and relative efficiencies
 2974 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.

2975

APPENDIX D

2976

2977 Invariant mass fits to $B^0 \rightarrow K^{*0}e^+e^-$ simulated candidates

2978

2979 This appendix contains fits to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}e^+e^-$ simulated candidates used to constrain parameters in the fit to data.

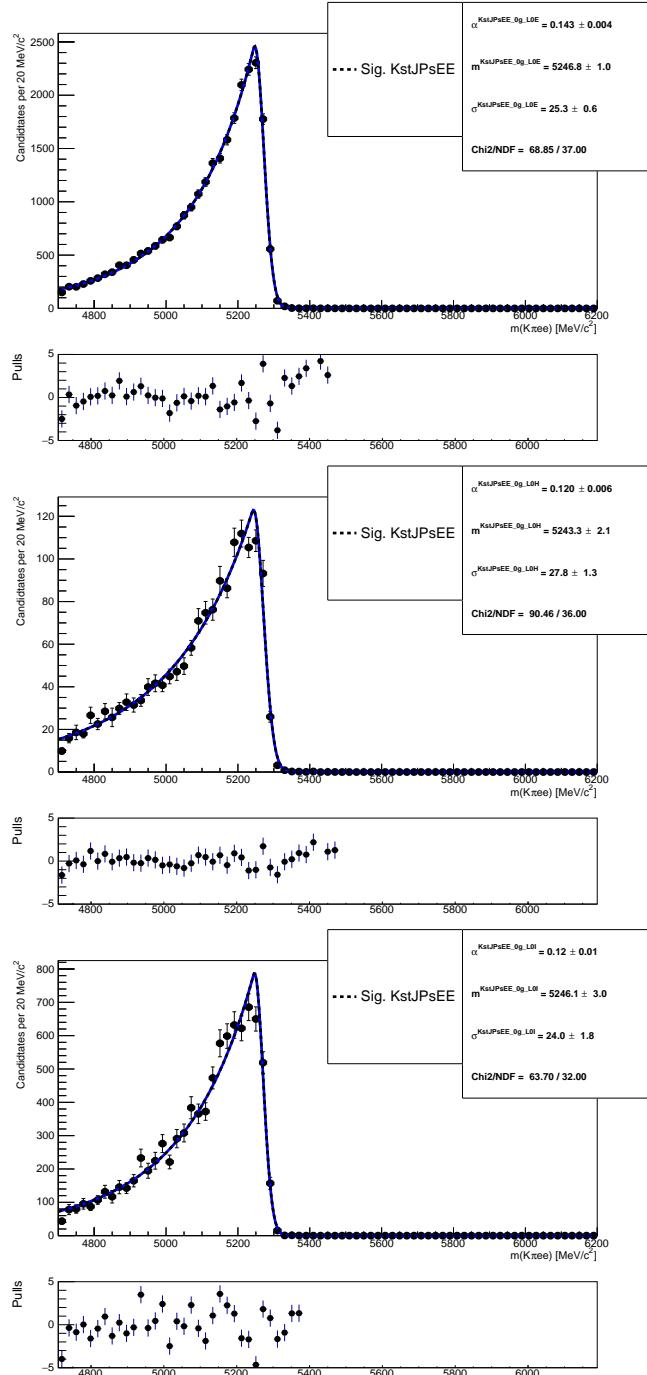


Figure D.1: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and no photon emitted.

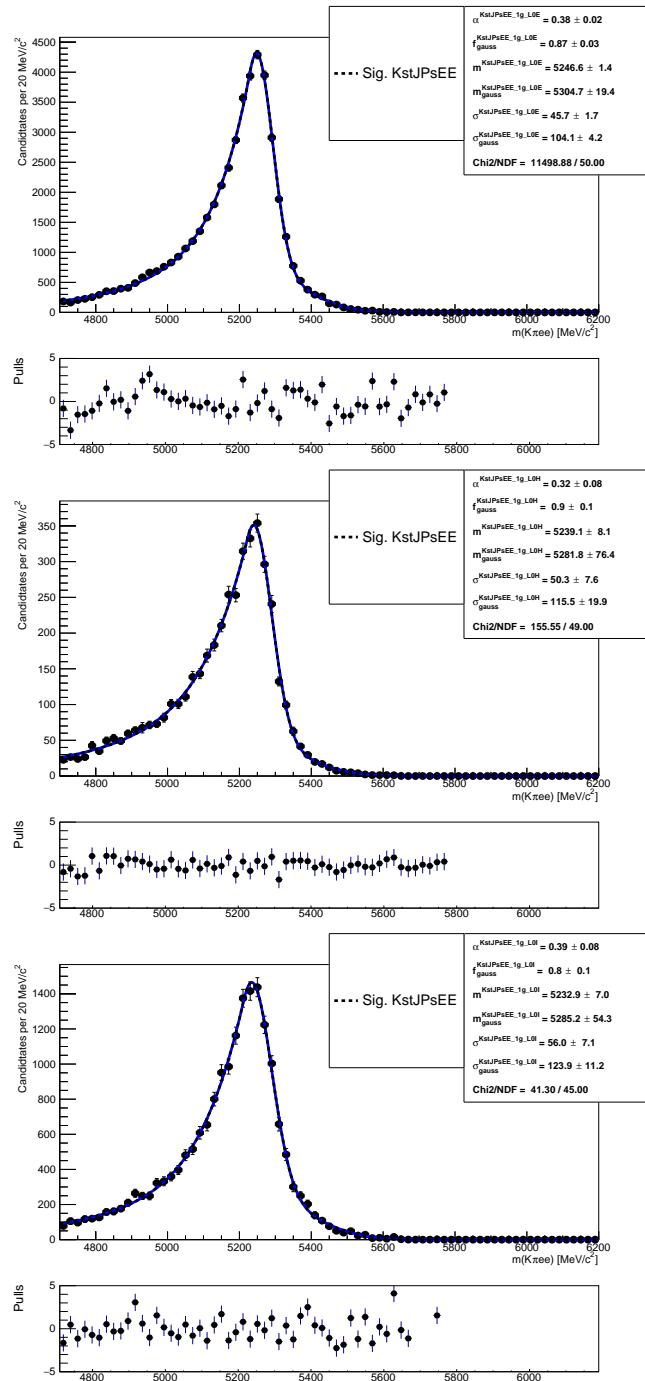


Figure D.2: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and one photon emitted.

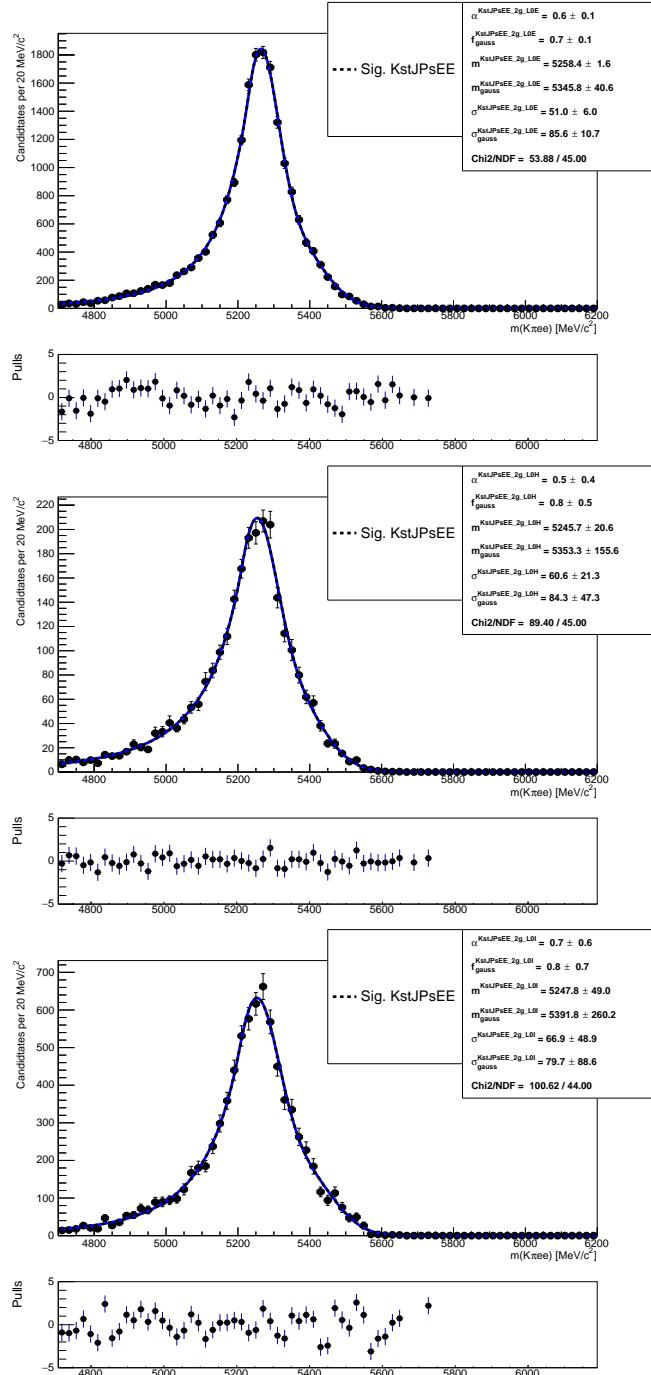


Figure D.3: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and two photons emitted.

APPENDIX E

2981

2982

2983

Extra

2984

2985 E.0.1 Quantum Electrodynamics: an example of gauge field theory

2986 *Possible addition around Sec. 1.1.*

2987 The EM theory results from requiring the fermion Lagrangian to be invariant under local gauge transformations. The electron and positron free fermionic fields
2988 are defined as $\psi(x)$ and $\bar{\psi}(x)$, where x is a relativistic four vector. A local gauge
2989 transformation can always be written as:
2990

$$\psi(x) \rightarrow \psi(x') = e^i \alpha(x) \psi(x) \quad (\text{E.1})$$

2991 where $\alpha(x)$ can be any function of space and/or time. The free fermion Lagrangian,
2992 given by

$$\mathcal{L} = -i\bar{\psi}(x)\gamma^\mu \partial_\mu \psi(x) - m\bar{\psi}(x)\psi \quad (\text{E.2})$$

2993 is not invariant under such transformations. Greek indices denote space-time directions
2994 and imply summation and γ^i are the Dirac matrices. If we apply the local
2995 gauge transformation and we subtract the initial Lagrangian we get a remaining
2996 term

$$\Delta\mathcal{L} = \mathcal{L}' - \mathcal{L} = -i\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha(x) \quad (\text{E.3})$$

2997 In order to make the Lagrangian invariant we can introduce a vector field A , which
2998 transforms as described by Eq. E.4, to

$$A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha(x) \quad (\text{E.4})$$

2999 Redefining then the field derivative $D_\mu = \partial_\mu - ieA_\mu$, we obtain the invariant La-
3000 grangian:

$$\mathcal{L} = -i\bar{\psi}(x)\gamma^\mu D_\mu\psi(x) - m\bar{\psi}(x)\psi = -i\bar{\psi}(x)(\gamma^\mu\partial_\mu - m)\psi(x) + e\bar{\psi}\gamma^\mu\psi A_\mu \quad (\text{E.5})$$

3001 In this Lagrangian the first term describes the behaviours of a free fermion and the
3002 second term, $\bar{\psi}\gamma^\mu\psi A_\mu$, the electromagnetic interaction through the photon repre-
3003 sented by A_μ .

3004 E.1 Anomalies

3005 *Possible addition around Sec. 1.6.*

3006 Various anomalies were observed in the past years with respect to SM predictions.
3007 This section reports a brief review of these anomalies, limiting to B physics.

3008 The measurement of the CKM matrix elements V_{ub} and V_{cb} is vital for analysis in the
3009 flavour sector. Both these quantities can be measured using tree level transitions,
3010 which are assumed to be free from NP. Decays such as $B \rightarrow D^*\ell\nu$ are used to
3011 measure V_{cb} and $B \rightarrow \pi\ell\nu$ for V_{ub} as well as inclusive decays. Several measurements,
3012 mainly from BaBar and LHCb [112, ?], observe a discrepancy at 2σ level between
3013 the values found using the exclusive and inclusive approaches. This has recently
3014 increased to 3σ level due to improvements in form factor calculations [113]. NP can
3015 modify the values of the CKM matrix elements as described in Ref. [].

3016 Secondly a series of anomalies was found in recent LHCb measurements of semilep-
3017 tonic B decays. The branching ratios of the $B \rightarrow K\mu^+\mu^-$, $B \rightarrow K^{*0}\mu^+\mu^-$ and
3018 $B_s^0 \rightarrow \phi\mu^+\mu^-$ [34, 35, 36] are all found to be slightly below the predicted values. Al-
3019 though taken by itself each measurement does not present relevant discrepancies,
3020 the systematic deviation seems to indicate a more general picture. Angular analy-
3021 sis were also performed for these decays and, while most observables are found to
3022 agree with SM predictions, the measurement of the P'_5 observable in $B \rightarrow K^{*0}\mu^+\mu^-$

3023 resulted in a local 3.7σ deviation with respect to predictions []. At the same time
3024 the measurement of the R_K ratio, between the branching fractions of the $B^0 \rightarrow K^{*0}$
3025 $\mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$ decays, showed a 2.6σ deviation from unity, indicating
3026 the possibility of a violation of lepton flavour universality. Authors of Ref. [114]
3027 performed a global fit taking into account of several measurements and found that
3028 a model with a NP component in C_9 is preferred with respect to the SM at 4.3σ
3029 level. Finally, one more discrepancy linked to this picture is the branching fraction
3030 of the $h \rightarrow \mu\tau$ decay, which is found to be different from zero at 2.4σ level.