

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.

March 15, 2016

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 (SM) 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 $b \rightarrow s\ell^+\ell^-$ decays are presented. First, $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays are analysed to measure their branching fraction as a function of the square 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 $R_{K^{*0}}$ 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. This chapter also includes studies performed to validate the hadronic physics in LHCb simulation software. Chapter 3 deals with the measurement of the differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay, while Chapter 4 describes its angular analysis. Finally, Chapter 5 reports the measurement of the $R_{K^{*0}}$ ratio.

DECLARATION OF AUTHORS CONTRIBUTION

I am one of the main authors of the two analysis reported in Chapters 3, 4 and 5. For the analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays I collaborated with Michal Kreps, who took care of implementing the decay model to re-weight the simulation and provided simulated samples. Furthermore, I want to thank him for the advice given throughout. The work on this analysis was published and can be found at Ref. [1]. For the $R_{K^{*0}}$ analysis, described in Ch. 5, I actively participated in most stages of the analysis collaborating with Simone Bifani. In particular I took care of the production of various simulated samples, participated in the definition of the selection, in the yields extraction and I provided a fit and data-reduction framework. Finally, as a service work for the LHCb experiment, I developed the tools used to perform the validation studies described in Sections 2.11–2.13.

ACKNOWLEDGEMENTS

I thank everybody, evvvvvvvveverybody!

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

*Nec per se quemquam tempus sentire fatendumst
semotum ab rerum motu placidaque quiete.
(Lucretius, De rerum natura, vv. 462-463)*

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	12
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	24
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	Constrained kinematic 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
2.13	Validation and material budget studies conclusions	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	57
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	63
3.4.4 Trigger	65
3.4.5 Background from specific decays	66
3.5 Yield extraction	68
3.5.1 Fit description	69
3.5.2 Fit results	73
3.6 Efficiency	79
3.6.1 Geometric acceptance	79
3.6.2 Reconstruction and neural network efficiencies	80
3.6.3 Trigger efficiency	80
3.6.4 PID efficiency	81
3.6.5 Relative efficiencies	82
3.7 Systematic uncertainties	85
3.7.1 Systematic uncertainty on the yields	85
3.7.2 Systematic uncertainties on the efficiency determination	87
3.7.2.1 Effect of new physics on the decay model	87
3.7.2.2 Simulation statistics	87
3.7.2.3 Production polarisation and decay structure	88
3.7.2.4 Λ_b^0 lifetime	89
3.7.2.5 Downstream candidates reconstruction efficiency	89
3.7.2.6 Data-simulation discrepancies	90
3.8 Differential branching ratio extraction	90
4 Angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays	95
4.1 One-dimensional angular distributions	96
4.2 Multi-dimensional angular distributions	98
4.3 Angular resolution	100
4.4 Fit strategy	102
4.4.1 Feldman-cousins plug-in method	103
4.4.2 Modelling the angular distributions	104
4.4.3 Angular acceptance	106
4.4.4 Studies on a three-dimensional fit	107
4.5 Systematics uncertainties on angular observables	108
4.5.1 Angular correlations	109
4.5.2 Resolution	110
4.5.3 Efficiency description	110

4.5.4	Background parameterisation	111
4.5.5	Polarisation	112
4.6	J/ψ cross-check	112
4.7	Results	114
5	Testing lepton flavour universality with $R_{K^{*0}}$	118
5.1	Combining ratios	121
5.2	Experimental status	122
5.3	Analysis strategy	123
5.4	Dilepton invariant mass intervals	124
5.4.1	Control channels	125
5.5	Data samples and simulation	125
5.5.1	Data-simulation corrections	125
5.6	Selection	128
5.6.1	Trigger and Stripping	129
5.6.2	PID	132
5.6.3	Peaking backgrounds	134
5.6.3.1	Charmonium vetoes	134
5.6.3.2	ϕ veto	135
5.6.3.3	$B^+ \rightarrow K^+ \ell^+ \ell^-$ plus a random pion	136
5.6.3.4	Λ_b decays	136
5.6.3.5	$B^0 \rightarrow (D^- \rightarrow K e^- \bar{\nu}) e^+ \nu$	136
5.6.3.6	$B^0 \rightarrow K^{*0} (\gamma \rightarrow e^+ e^-)$	137
5.6.3.7	Other peaking backgrounds	138
5.6.4	Partially-reconstructed background	139
5.6.5	Bremsstrahlung corrected mass	139
5.6.6	Multivariate analysis	141
5.6.7	Optimisation	144
5.7	Selection summary	147
5.8	Mass fits	148
5.8.1	Muon channels	148
5.8.1.1	$B^0 \rightarrow K^{*0} (J/\psi \rightarrow \mu^+ \mu^-)$ PDF	149
5.8.1.2	$B^0 \rightarrow K^{*0} \mu^+ \mu^-$ PDF	150
5.8.1.3	Summary	151
5.8.2	Electron channels	153
5.8.2.1	Signal PDFs for the electron channels	154
5.8.2.2	Background PDFs for the electron channels	157
5.8.2.3	Summary of the fit to the electron samples	160
5.8.3	Event yields	162
5.9	Efficiency	166
5.9.1	Geometric efficiency	168
5.9.2	Reconstruction efficiency and bin migration	168
5.9.2.1	Bin migration	168
5.9.3	PID efficiency	170
5.9.4	Trigger efficiency	170
5.9.4.1	Electron triggers	171

5.9.4.2	TISTOS cross-check	173
5.9.5	Neural networks and BCM efficiencies	175
5.10	Systematic uncertainties	176
5.10.1	Choice of signal and background PDFs	176
5.10.2	Efficiency determinations	178
5.10.3	Bin migration	178
5.11	Result extraction	179
5.11.1	$R_{J/\psi}$ sanity check	180
5.11.2	$R_{K^{*0}}$	180
6	Conclusions	181
A	Decay models	193
A.1	$\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ distribution	193
A.2	Bi-dimensional distribution parameters	196
A.3	$\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution	196
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	Improved predictions for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ observables.	204
E	Invariant mass fits to $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ simulated candidates	207
F	Invariant mass fits to $B^0 \rightarrow K^{*0} e^+ e^-$ candidates divided in trigger categories	212

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 its cur-
7 rent form in the mid-70s and has been an extremely successful predictive theory
8 since then. Almost all known phenomena from 1 eV up to several hundred GeV are
9 described well 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 2012 the Higgs boson, which is one
13 of the fundamental building blocks of the theory, was observed [3, 4]. This is a
14 critical ingredient of the SM as it introduces a mechanism that produces particles'
15 masses [5]. Despite the success of the SM, experimentally well-established effects,
16 like neutrino oscillations and the presence of dark matter, remain outside the reach
17 of this theory. Furthermore, the model does not include the description of gravity,
18 which can be neglected at the EW energy scale. This motivates the search for New

Physics (NP).

Table 1.1: Predicted and measured values of the decay widths of the Z^0 and W bosons [2].

Quantity	Predicted	Measured
Γ_{Z^0}	2.4960 ± 0.0002 GeV	2.4952 ± 0.0023 GeV
Γ_W	2.0915 ± 0.0005 GeV	2.085 ± 0.042 GeV

¹⁹

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 of the wave function under the unitary group transformations given by the product $SU(3)_C \otimes SU(2)_W \otimes U(1)_Y$, and leads to conservation laws such as the conservation of electric and strong charge. The model has then 26 free parameters, which have to be experimentally measured.

Particles included in the SM can be grouped into 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 local gauge symmetry [6], they are called “gauge bosons”. The list of the 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

Table 1.2: Fundamental forces of nature together with their gauge bosons, ranges and relative strengths, as they act on a pair of protons in an atomic nucleus [2]. Gravity is not included in the SM and the graviton is hypothetical at the current time.

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

³⁶ have spin 1/2. Among fermions one can then consider two smaller groups: quarks
³⁷ and leptons. Quarks carry colour charge and therefore can interact through the
³⁸ so-called strong interaction, while leptons, which do not carry colour charge, are
³⁹ insensitive to it. For each particle a corresponding anti-particle exists with opposite
⁴⁰ quantum numbers. Finally, fermions are divided into three families having similar
⁴¹ properties but different masses. This last classification embedded in the SM is also
⁴² called “flavour structure” and it will be the main tool used in this thesis; a more
⁴³ detailed description of it is given in the following sections. A schematic view of the
fundamental particles in the SM is shown in Fig. 1.1.

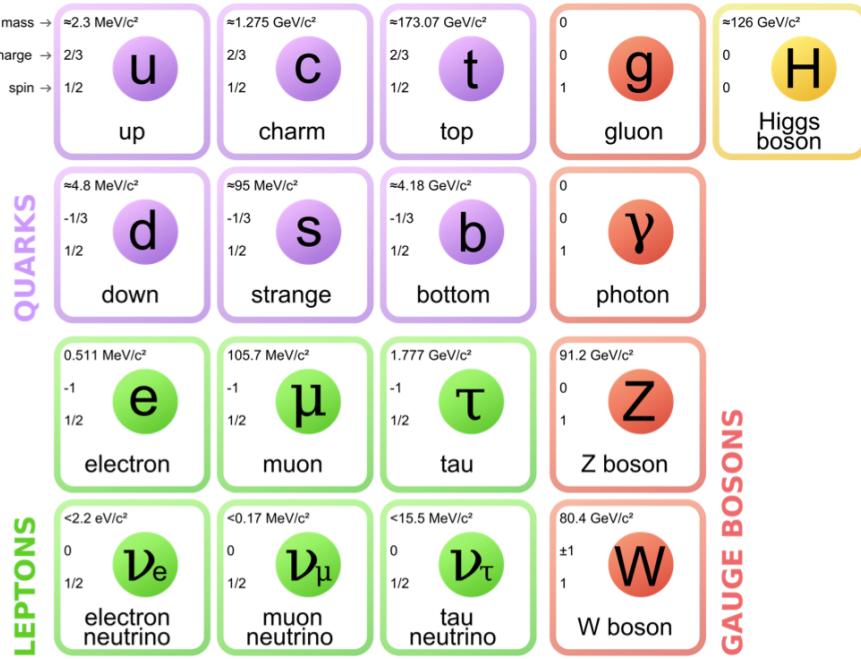


Figure 1.1: A scheme of the fundamental particles in the SM with their properties [7].

44

⁴⁵ Due to the asymptotic freedom of the strong interaction quarks cannot be observed
⁴⁶ alone but are always combined with other quarks to form color singlets [8]. Non-
⁴⁷ fundamental particles composed of quarks are called hadrons and are classified into
⁴⁸ two groups: mesons, where the color singlet is achieved by the combination of a
⁴⁹ quark and an antiquark ($q \bar{q}$), and baryons formed from three quarks ($q q q$) of
⁵⁰ different colours. Recently, in 2014 and 2015 evidence for new states, formed by
⁵¹ four and five quarks, was found [9, 10].

52 1.1 The electroweak interaction

53 The electromagnetic interaction is responsible for binding electrons and nuclei to
54 gether to form atoms and its mediator is the photon. The weak interaction is
55 responsible for the β decay of nuclei and is mediated by the exchange of W^\pm and
56 Z^0 bosons. Unlike the electromagnetic force, that affects only charged particles, all
57 known fermions interact through the weak interaction. The weak interaction is also
58 the only one that violates the parity symmetry, which states that interactions are
59 invariant under an inversion of spatial coordinates. This symmetry breaking arises
60 from the fact that only left-handed fermions interact through the weak interaction
61 as discovered by Wu in 1957 [11]. Similarly, the weak interaction is the only one that
62 also breaks the CP symmetry, which combines parity transformations and charge
63 conjugation. This is particularly interesting because all interactions are believed to
64 be invariant under the CPT transformation, which combines the CP transformation
65 and time reversal. Hence, breaking CP the weak interaction implies that the process
66 is also not invariant under time reversal transformations.

67 In 1968 Salam, Glashow and Weinberg unified the weak and electromagnetic forces
68 into a single theory, where the coupling constants of the electromagnetic, e , and
69 weak, g , interactions are related through the weak mixing angle, θ_W by the rela-
70 tion $g \sin \theta_W = e$ [2]. The electroweak symmetry is spontaneously broken by the
71 Higgs mechanism [12] and this causes the W^\pm and Z bosons to become massive (see
72 Tab. 1.2) and consequently the weak force has a very short range. In fact, using
73 Heisenberg's Principle, $\Delta E \Delta t > \hbar$, together with Einstein's formula $\Delta E = mc^2$,
74 which relates mass and energy, and knowing that the maximum space that a parti-
75 cle can cover in a time Δt is $r \sim c\Delta t$, qualitatively $r \sim \hbar/mc$. In this picture the
76 carriers of the weak force can travel $r \sim 2 \cdot 10^{-3}$ fm. In contrast, the photon must
77 be massless in the theory, which accounts for the long range of the electromagnetic
78 force. The EW interactions are divided into Charged Currents (CC) and Neutral
79 Currents (NC). In the first group, quarks and leptons interact with the W^\pm bosons,
80 producing decays such as $\mu^+(\mu^-) \rightarrow e^+ \nu_e \bar{\nu}_\mu (e^- \bar{\nu}_e \nu_\mu)$ and $n(\bar{n}) \rightarrow p e^- \bar{\nu}_e (\bar{p} e^+ \nu_e)$. The

study of these processes confirmed that only the left-handed (right-handed) component of fermions (anti-fermions) takes part in weak processes. The CC interactions have a peculiarity: they are the only interactions in the SM that violate flavour conservation at tree level, while any other interaction not conserving flavour has to proceed through higher order processes. The second group of EW interactions, NC, corresponds to diagrams mediated by a photon or a Z boson interacting with a fermion and its anti-fermion.

1.2 Flavour and the CKM matrix

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

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

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

Measuring branching fractions of weak decays like $\pi \rightarrow \mu\nu_\mu$ and $K \rightarrow \mu\nu_\mu$, corresponding respectively to $ud \rightarrow \mu\nu_\mu$ and $us \rightarrow \mu\nu_\mu$ processes, suggested the existence of more than one coupling constant for different quarks. Nicola Cabibbo, in order to

106 preserve the universality of weak interactions, suggested that the differences could
 107 arise from the fact that the doublets participating in the weak interactions are an
 108 admixture of the mass eigenstates [2, 13]. He therefore introduced the Cabibbo an-
 109 gle, θ_c , proposing that eigenstates participating to the weak interaction are rotated
 110 with respect to the flavour eigenstates.

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

111 In a six quark system one angle is not sufficient to describe a rotation but the mixing
 112 can be generalised using a 3×3 unitary matrix, called the CKM matrix, from the
 113 names of Cabibbo, Kobayashi and Maskawa [13, 14]. The unitarity of the matrix
 114 is required to preserve the universality of the weak interaction. Theoretically, a
 115 $N \times N$ complex matrix depends on $2 \cdot N^2$ real parameters. Requiring unitarity
 116 ($AA^\dagger = A(A^*)^T = I$), the number of independent parameters left is

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

117 Therefore a 3×3 matrix depends then on 4 real parameters: three real constants and
 118 one imaginary phase. The imaginary phase generates the CP-violation which was
 119 observed in weak interactions. Figure 1.2 displays examples of CC processes together
 with the CKM elements associated with their vertices. Equation 1.3 reports the most

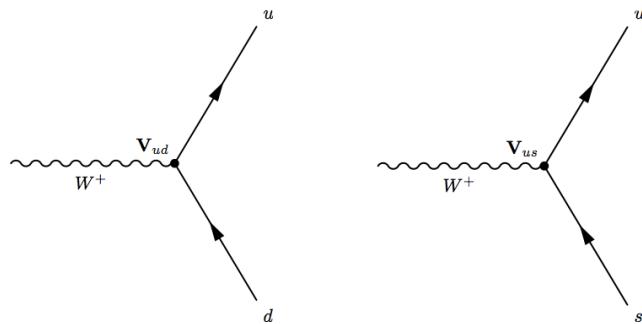


Figure 1.2: Feynman diagrams with CKM weights on weak interaction vertices as defined in Eq. 1.3.

recent measured values of its elements [2] together with the widely used Wolfenstein parametrisation which highlights the hierarchical structure of the matrix. In fact, elements on the diagonal, corresponding to transitions between quarks of the same generation, are approximately 1 and become smaller and smaller going farther from the diagonal. In the formula ρ , A , and λ are the real constants and η the imaginary phase and Eq. 1.4 shows how they are related to the three mixing angles; terms further from the diagonal are proportional to higher powers of λ .

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.9743 \pm 0.0002 & 0.2253 \pm 0.0007 & 0.0035^{+0.0002}_{-0.001} \\ 0.2252 \pm 0.0007 & 0.9734 \pm 0.0002 & 0.00412^{+0.0011}_{-0.0005} \\ 0.0087 \pm 0.0003 & 0.0404^{+0.0011}_{-0.0005} & 0.99915^{+0.00002}_{-0.00004} \end{pmatrix}$$

$$= \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \quad (1.3)$$

128

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

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

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

130 These correspond to constraints on three complex numbers, which can be viewed
131 as the sides of triangles in the (ρ, η) plane; these are called “unitarity triangles”.
132 The most commonly used unitarity triangle arises from $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$.
133 Figure 1.3 shows a representation of such triangle together with a plot summarising
134 the most up-to-date experimental constraints to its parameters [15]. Due to these
135 unitarity constraints flavour-changing neutral currents are forbidden at tree level in
136 the SM.

137 The precise measurement of the parameters of the CKM matrix is a powerful sta-

¹³⁸ bility test of the SM and sets a solid basis for new physics searches in the flavour
¹³⁹ sector. One of the main goals of the LHCb experiment is to measure precisely the angle γ , which is currently the least constrained by measurements.

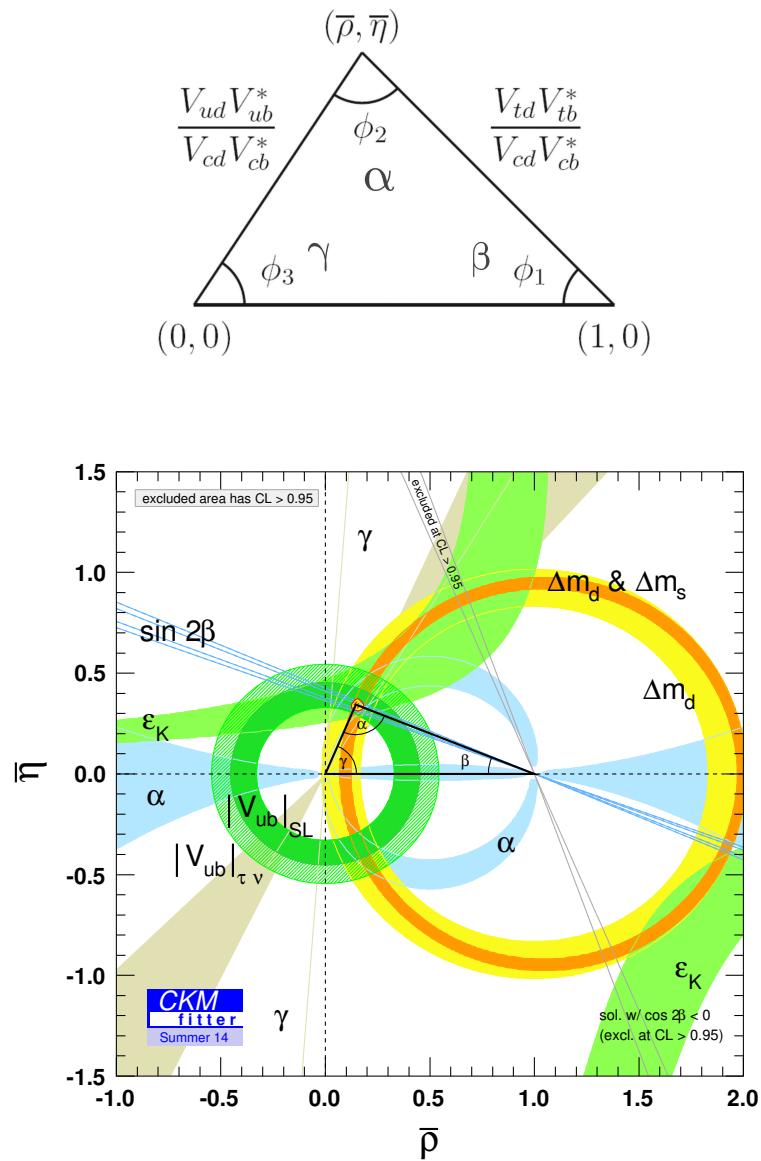


Figure 1.3: (top) A representation of the unitarity triangle and its parameters. (bottom) A summary of the most up-to-date measurements of the unitarity triangle parameters [15].

¹⁴¹ 1.3 The puzzles of the SM

¹⁴² Despite the experimental confirmation of many predictions of the SM, the theory has
¹⁴³ several limitations and is unable to account for some well-established experimental
¹⁴⁴ facts:

- ¹⁴⁵ • *Dark matter*: experimental evidence tells us that the content of visible matter
¹⁴⁶ in the universe is not sufficient to account for the observed rotation of galaxies [16]. The most natural way to solve the problem is the hypothesis of a form
¹⁴⁷ of matter that interacts with the gravitational field but not with the other SM
¹⁴⁸ interactions.
- ¹⁵⁰ • *Matter-antimatter asymmetry*: a large asymmetry is observed between the
¹⁵¹ quantity of matter and antimatter in the universe, $O(10^{-9})$. Assuming that
¹⁵² both were equally created in the initial state of the universe, a condition such
¹⁵³ as the violation of the CP symmetry is necessary to account for the observed
¹⁵⁴ imbalance. However, the magnitude of CP violation predicted by the SM,
¹⁵⁵ $O(10^{-20})$, is not sufficient to account for the observed asymmetry [17].
- ¹⁵⁶ • *Gravity*: even though the gravitational force was the first to be discovered
¹⁵⁷ this is not included in the SM. When introducing gravity in the framework of
¹⁵⁸ QFT the theory diverges. On the other hand gravity becomes irrelevant for
¹⁵⁹ the small masses of particles and can be neglected to a good approximation at
¹⁶⁰ the EW energy scale. Many attempts have been made but there is not yet a
¹⁶¹ consistent theoretical framework through which gravity can be introduced in
¹⁶² the SM [18].
- ¹⁶³ • *Neutrino oscillation*: measurements of solar and atmospheric neutrinos, as
¹⁶⁴ well as neutrinos from nuclear reactors, have established that neutrinos can
¹⁶⁵ change flavour while propagating in space. This is not predicted in the SM, in
¹⁶⁶ fact in the SM neutrinos are massless, while an oscillation requires a non-zero
¹⁶⁷ mass [19, 20, 21, 22].

- 168 • *The hierarchy problem*: the mass of a scalar (spin 0) particle, such as the
169 Higgs boson, suffers from quantum corrections due to the physics at high
170 energy scales. As new physics can appear anywhere up to the Planck scale,
171 $\sim 10^{19}$ GeV, at which gravity cannot be neglected any more, these corrections
172 can be very large and it would require a high level of fine-tuning for them to
173 cancel out and give such a small value as the one measured for the Higgs Mass,
174 ~ 126 GeV/ c^2 [23, 3].

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

179 1.3.1 The flavour problem

180 Flavour Changing Charged Currents (FCCC) that are mediated by the W^\pm bosons
181 are the only sources of flavour changing interaction in the SM and, in particular, of
182 generation changing interactions, where a quark or a lepton of a family transforms
183 into one of another family. Another class of processes is the Flavour Changing
184 Neutral Currents (FCNCs), e.g. transitions from a b quark with a charge of -1/3 to
185 a s or d quark with the same charge. Examples of FCNC transitions in the quark
186 and lepton sector are shown in Fig. 1.4. FCNCs are experimentally observed to be
187 highly suppressed which derives from the unitarity of the CKM matrix, however
188 there is no fundamental reason why there cannot be FCNCs at tree level. In fact
189 the CKM matrix could be part of a larger matrix involving for example quark-
190 lepton terms. This would introduce new sources of FCNCs but could also allow for
191 natural explanations of the equality of the proton and electron charges. Furthermore,
192 the observation of neutrino oscillation proves that flavour is not always conserved
193 suggesting flavour structures beyond the SM. Finally, the values of the terms of the
194 CKM matrix and the PMNS matrix [24, 25], which is the mixing-matrix equivalent to
195 the CKM in the lepton sector, are not explained in the SM but have to be measured

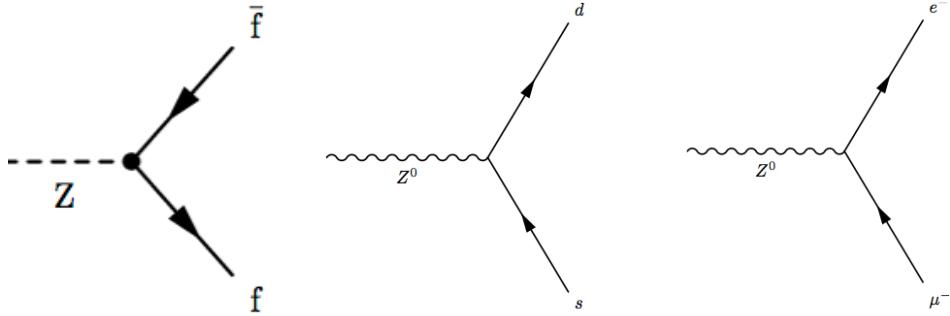


Figure 1.4: Feynman diagrams of a neutral current allowed in the SM (left), where f represents any fermion, and FCNCs processes forbidden in the SM (center-right).

¹⁹⁶ experimentally. These open problems motivate searches for flavour symmetries and
¹⁹⁷ deeper motivations for flavour conservation.

¹⁹⁸ 1.4 Beyond the Standard Model

¹⁹⁹ From the previous sections it is evident that, despite the great success of the SM,
²⁰⁰ there is a need to explore theories Beyond the SM (BSM). Among the most promis-
²⁰¹ ing approaches there are those involving Super-Symmetry (SUSY) [26] and extra-
²⁰² dimensions [27]. In SUSY new degrees of freedom are introduced to suppress the
²⁰³ diverging terms of the Higgs mass. This theory assumes that for each fermion there
²⁰⁴ is a corresponding boson and, since bosons and fermions contribute with opposite
²⁰⁵ sign to the mass term, these would naturally cancel out. Supersymmetry also pro-
²⁰⁶ vides a candidate for dark matter. In fact the lightest Super-Symmetric particle, the
²⁰⁷ neutralino, which in R-parity [28] conserving variants of the theory must be stable,
²⁰⁸ is a weakly interacting potentially heavy particle. The idea to introduce extra
²⁰⁹ dimensions was triggered by the fact that gravity is not relevant in particle physics
²¹⁰ but it would be natural if all forces had similar strength. By adding extra dimen-
²¹¹ sions to the normal three spatial dimensions, one can restore the strength of gravity,
²¹² as this could be dispersed by the wider space available. In all these approaches con-
²¹³ straints to masses and couplings must be imposed to maintain compatibility with
²¹⁴ the SM at the electroweak scale and the existing experimental observations.

215 1.4.1 Flavour and BSM theories

216 Most BSM theories predict processes violating flavour conservation. Therefore, the
 217 observation or non-observation of these processes can give important information
 218 about new physics. BSM theories can be classified according to the amount of
 219 flavour violation they introduce. The first class of models to consider is that with
 220 Minimal Flavour Violation (MFV). These are models in which the only sources of
 221 flavour changing transitions are governed by the CKM matrix and the CKM phase
 222 is the only source of CP violation. This definition is driven by the fact that usually
 223 a solution of the hierarchy problem is expected at the TeV scale, while the very
 224 small amount of flavour violation observed in measurements seems to indicate that
 225 the SM would remain valid up to much higher energy scales. It is therefore assumed
 226 that new physics must respect flavour symmetry principles, which also makes these
 227 types of models naturally compatible with the SM. Examples of such models include
 228 the MSSM with minimal flavour violation and the SM with one extra-dimension.
 229 Reviews of MFV models are presented in Refs. [29, 30]. A powerful test of MFV
 230 is provided by the study of ratios between $b \rightarrow d$ and $b \rightarrow s$ transitions, because
 231 their hamiltonians share the same structure. One particularly important example
 232 is the ratio between the decay rates of B^0 and B_s^0 into dimuons [31], as this is a
 233 purely leptonic decay free from hadronic uncertainties. In the SM such ratios are
 234 approximately equal to $|V_{td}/V_{ts}| \sim 1/25$, only modified by phase space and hadronic
 235 matrix elements, while they can take very different values in non-MFV models.

236 In the quest for new physics an important role is also played by simplified models
 237 as an intermediate model building step. Instead of constructing theories valid up
 238 to the GUT scale one can consider simplified models, where the SM is extended by
 239 the addition of new degrees of freedom with a limited number of parameters. Such
 240 models are easier to constrain but can nevertheless point in the right direction to
 241 build more complete theories. The choice of the new sector to add can be driven
 242 by the need to explain existing tensions between measurements and SM predictions
 243 or by theoretical prejudice. Two models especially relevant when studying rare

244 decays, which are the main topic of this thesis, are Z' -penguins and leptoquarks.
245 A Z' -penguin is a FCNC process involving a neutral field arising from an extra
246 $U(1)$ gauge symmetry, for example $U(1)_{B-L}$, where B and L are the baryon and
247 lepton numbers. As for the SM penguins, the Z' field contributes in loops causing
248 modifications of the effective couplings with respect to the SM. A survey of Z' models
249 can be found in Ref. [32]. Leptoquarks are bosonic particles that carry both quark
250 and lepton flavour quantum numbers, which for simplicity are assumed to be scalar.
251 A tree level exchange of a leptoquark induces processes such as $b \rightarrow (s, d)\ell^+\ell^-$,
252 and therefore can result in an enhancement of their decay rates with respect to the
253 SM [33]. Leptoquarks would also provide a natural explanation for non-universal
254 couplings to leptons.

255 1.5 Rare decays: a tool to search for new physics

256 In the Standard Model FCNC processes are forbidden at tree level but can occur
257 through loop diagrams such as penguin or W box diagrams (see Fig. 1.5). The
258 branching fractions of decays going through these processes are small, typically
259 $\sim 10^{-6}$ or lower, and therefore they are called “rare decays”. Additional contribu-
260 tions to the virtual loops are not necessarily suppressed with respect to the SM
261 component which makes these decays very sensitive to new physics. This approach
262 to new physics searches is interesting as new particles could be at high mass scales
263 that are not accessible via direct production at colliders but their effect could be
264 observed in loops. Radiative and penguin decays are particularly interesting because
265 they are theoretically well understood, which allows precise comparisons with mea-
266 surements. Furthermore, they provide a large quantity of observables that can be
267 affected by new physics, not only decay rates, but also CP asymmetries and angular
268 observables such as forward-backward asymmetries. The joint analysis of different
269 observables can help to build a consistent picture and rule out specific models.

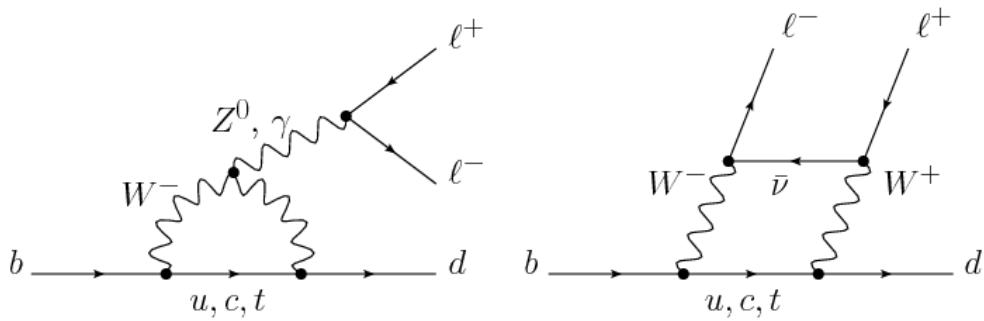


Figure 1.5: Loop Feynmann diagrams allowing $b \rightarrow d$ FCNC processes: penguin diagram (left) and W box (right).

²⁷⁰ 1.5.1 Theoretical framework: the effective Hamiltonian

²⁷¹ Rare decays of b hadrons are governed by an interplay between weak and strong
²⁷² interactions. The large masses of the W^\pm and Z^0 bosons and top quark compared
²⁷³ to that of the b quark allow the construction of an effective theory that divides the
²⁷⁴ problem of calculating weak decay amplitudes into two parts: “short-distance” and
²⁷⁵ “long-distance” effects separated at an energy scale μ . The first part, dealing with
²⁷⁶ short distance physics, handles perturbative contributions due to energy scales above
²⁷⁷ the b mass. The second part typically deals with non-perturbative contributions.
²⁷⁸ A classic example of an effective theory is the Fermi theory of weak interactions
²⁷⁹ which describes the β decay in terms of a four-fermion interaction, where the short
 distance physics is hidden into a point-like vertex as illustrated in Fig. 1.6.

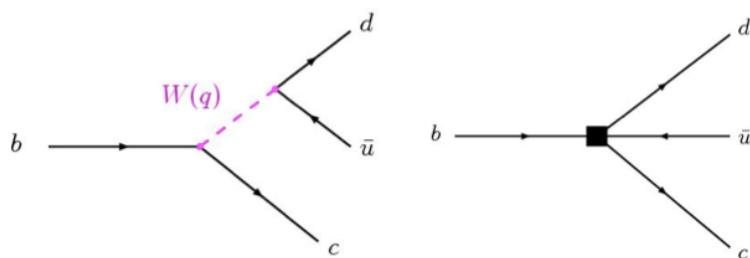


Figure 1.6: Example of a Fermi theory in which the full theory (left) is divided into (right) a short distance contribution, hidden in the vertex, and a long distance contribution.

281 The effective hamiltonian [34] relevant to $b \rightarrow s/d\gamma$ and $b \rightarrow s/d\ell^+\ell^-$ transitions
282 can be written as:

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

283 where G_F denotes the Fermi coupling constant and the λ constants are the CKM
284 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 main
285 topic of this thesis, the doubly Cabibbo-suppressed contributions can be neglected
286 as $\lambda_s^u \ll \lambda_s^t$. To obtain this formula the Operator Product Expansion (OPE) [35]
287 method is used, which implements a summation over all contributing operators
288 weighted by corresponding constants called Wilson coefficients. In this Hamiltonian
289 the long-distance contributions are described by the operators, \mathcal{O}_i , while the short-
290 distance physics is encoded in the Wilson Coefficients, C_i . Operators and coefficients
291 are evaluated at the renormalisation scale μ . Any particle that contributes to the
292 decay and has a mass greater than the scale μ will affect the value of at least one of
293 the Wilson coefficients, including SM particles as the top quark.

294 In order to describe SM processes the effective theory must be matched with the
295 SM by requiring the equality between each term in effective theory and the full the-
296 oretical calculation at a matching scale, typically the EW scale (μ_W). Then, using
297 the scale independence of the effective Hamiltonian, one can derive a renormalisa-
298 tion group equation for the Wilson Coefficients [36]. Taking into account only SM
299 contributions and using $\mu_W = m_b$, the Wilson Coefficients have values:

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

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

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

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

³⁰⁴ (e.g. $I = B^0$ and $F = K^{*0}\mu^+\mu^-$) the decay amplitude can be calculated as

$$A(I \rightarrow F) = \langle I | \mathcal{H}_{eff} | F \rangle = \\ = \frac{G_F}{\sqrt{2}} \sum V_{CKM}^i \underbrace{C_i(\mu)}_{\substack{\text{Perturbative} \\ \text{Includes new physics}}} \cdot \underbrace{\langle I | \mathcal{O}_i(\mu) | F \rangle}_{\substack{\text{Non-perturbative} \\ \text{Known physics}}}, \quad (1.9)$$

³⁰⁵ where $\langle I | \mathcal{O}_i(\mu) | F \rangle$ are the hadronic matrix elements also called “form factors”.
³⁰⁶ These can be evaluated using non perturbative methods such as lattice calculations.
³⁰⁷ However, due to the limitations of these methods, they represent the dominant
³⁰⁸ source of uncertainty in theoretical calculations.

³⁰⁹ 1.5.2 Operators

³¹⁰ Separating the left- and right-handed components the effective Hamiltonian is

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

³¹¹ A complete basis is given by a set of 10 operators, where $\mathcal{O}_{1,2}$ are the tree level
³¹² W operators; $\mathcal{O}_{3-6,8}$ are penguin diagrams mediated by gluons; and $\mathcal{O}_{7,9,10}$, which
³¹³ are the operators that are relevant for radiative and leptonic penguin processes are
³¹⁴ defined as [31]:

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

³¹⁵ where $P_{L/R} = (1 \mp \gamma_5)/2$ denote the left- and right-handed chiral projections and
³¹⁶ $F_{\mu\nu}$ is the electromagnetic field tensor. The \mathcal{O}' operators correspond to right-handed
³¹⁷ coupling obtained by swapping P_R and P_L in the equations. In the SM, as well as
³¹⁸ in MFV models where the flavour violation is entirely ruled by the CKM matrix,
³¹⁹ the C' Wilson Coefficients are suppressed by the strange coupling, $C'_i \sim (m_s/m_b)C_i$.

320 The operator \mathcal{O}_7 relates to penguin diagrams that are mediated via a photon. It
321 represents the dominant contribution to the radiative $b \rightarrow s\gamma$ transition and con-
322 tributes to $b \rightarrow s\ell^+\ell^-$ processes when the virtual photon decays into a dilepton
323 pair. The semileptonic \mathcal{O}_9 and \mathcal{O}_{10} correspond to penguin diagrams mediated by
324 a Z^0 boson and W mediated box diagrams. These are the dominant contributions
325 in semileptonic $b \rightarrow s\ell^+\ell^-$ decays. The vertices corresponding to the radiative and
semileptonic operators are illustrated in Fig. 1.7

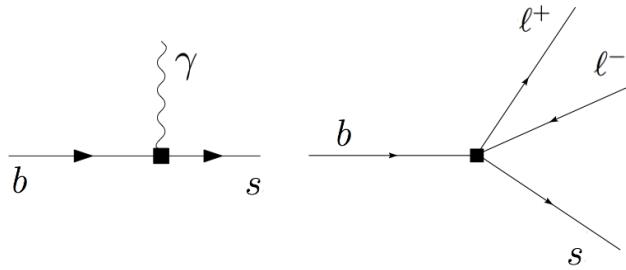


Figure 1.7: Interaction vertices corresponding to the radiative (left) and semileptonic (right) operators.

326

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

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

329 where the Wilson Coefficients are redefined as

$$\begin{aligned} C_{LL} &= C_9 - C_{10}, & C_{LR} &= C_9 + C_{10}, \\ C_{RR} &= C'_9 - C'_{10}, & C'_{RL} &= C'_9 + C_{10}. \end{aligned} \quad (1.13)$$

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

336 1.5.3 Phenomenology of $b \rightarrow s\ell^+\ell^-$ decays

337 Semileptonic b hadron decays are characterised by two kinematic regimes which are
 338 treated theoretically in different ways; Table 1.3 shows a scheme of the q^2 spec-
 339 trum. The ‘high q^2 ’ is the region of low hadron recoil, $q^2 > 15 \text{ GeV}^2/c^4$, and is
 340 characterised by the energy of the hadron being less than the energy scale of QCD in-
 341 teractions within the meson, $\Lambda_{QCD} \sim 1 \text{ GeV}$. In this region theoretical calculations
 342 of B meson decays can be simplified by working in the heavy quark limit, $m_b \rightarrow \infty$.
 343 In this limit a Heavy Quark Effective Theory (HQET) [37] can be constructed in
 344 which the heavy quark interacts only via ‘soft’ hadronic processes and an OPE in
 345 $1/m_b$ is valid. The ‘low q^2 ’ region is where the light spectator quark is energetic
 346 and cannot be neglected. Furthermore, the light quark interacts not only via ‘soft’
 347 hadronic processes, as in HQET, but also via the so-called ‘collinear’ hadronic pro-
 348 cesses. The boundary of this region can be set at $\sim 7 \text{ GeV}^2/c^4$ which corresponds
 349 to the threshold for $c\bar{c}$ production, $(2m_c)^2$. In this region the hadronic interactions
 350 are handled by expanding in terms of the energy of the emitted energetic hadron,
 351 $1/E_h$, forming the so-called Soft-Collinear Effective Theory (SCET) [38]. In both
 352 regions decay rates can be predicted using the different methods and the biggest un-
 353 certainties come from the limited knowledge of hadronic transition matrix elements.
 354 The intermediate region is characterised by the presence of charmonium resonances,
 355 produced though tree level $b \rightarrow \bar{c}cs$ transitions and no precise theoretical calculation
 356 is available [39].

357 As can be seen in Fig. 1.8 the very low q^2 region is characterised by a peak due to
 358 the virtual photon contribution, associated with C_7 . In the region $1 - 6 \text{ GeV}^2/c^4$ the

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

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

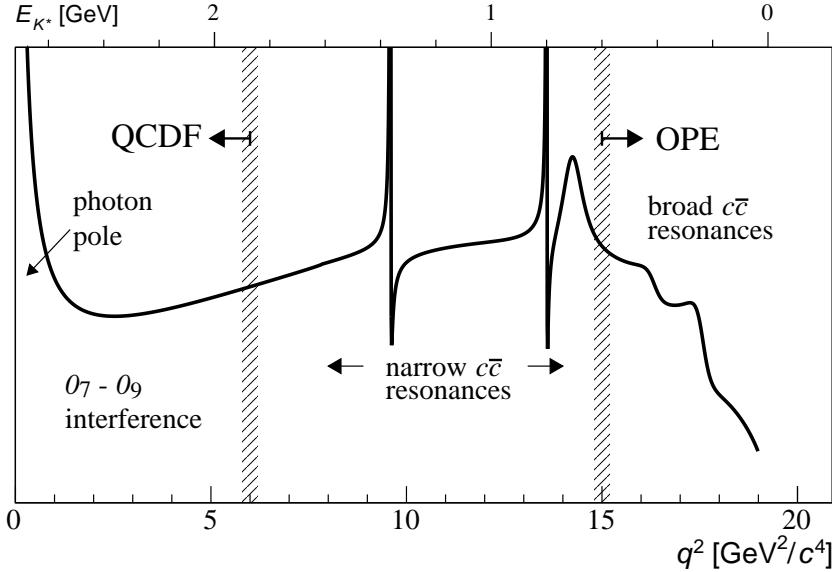


Figure 1.8: A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ processes characterised by the photon pole at low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2 [31].

359 interference between C_7 and C_9 becomes large, yielding sensitivity to new physics in
 360 C_9 . The $7 - 15$ GeV^2/c^4 interval is dominated by the charmonium resonances, J/ψ
 361 and $\psi(2S)$. Although these decays can be experimentally vetoed in principle char-
 362 monia affect the entire q^2 space. Finally, at high q^2 broad charmonium resonances
 363 can contribute, like those observed by LHCb in $B^+ \rightarrow K^+\mu^+\mu^-$ decays [40].

364 1.5.4 Observables in $b \rightarrow s\ell^+\ell^-$ decays

365 Rare decays and especially semileptonic $b \rightarrow s\ell^+\ell^-$ processes offer a number of ob-
 366 servables which can be used to study BSM models. The most direct effects appear
 367 in decay rates that can be enhanced by new physics but the precision on these
 368 measurements is often limited by uncertainties on the non-perturbative part of the
 369 calculations. Therefore, it is important to also look for different observables. One
 370 important class of observables are angular quantities that can often carry comple-
 371 mentary information with respect to branching ratio measurements. The most basic
 372 of these observable are forward-backward asymmetries that characterise the angular

³⁷³ distribution of final particles. For the $B^0 \rightarrow K^* \mu^+ \mu^-$ decay combinations of ob-
³⁷⁴ servables have been proposed that are independent of form factor uncertainties at
³⁷⁵ leading order order [31].

³⁷⁶ Another way to build safe observables is to construct ratios between similar decays,
³⁷⁷ in which uncertainties due to the hadronisation process cancel out. These observ-
³⁷⁸ ables include the R_H ratios, between B^0 decays into electrons and muons, that are
³⁷⁹ described in detail in Ch. 5. It is also interesting to compare decays which proceed
³⁸⁰ via the same fundamental process but where the spectator quark has a different
³⁸¹ flavour. This is the case of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays, which are
³⁸² both $b \rightarrow s$ transitions where the spectator quark is an u quark in the first case
³⁸³ and a d quark in the second. The normalised difference of the branching fractions
³⁸⁴ of these decays is called isospin asymmetry.

³⁸⁵ 1.6 Experimental status

³⁸⁶ To set the background for the analyses described in this thesis, this section reports a
³⁸⁷ brief review of recent results of new physics searches involving rare decays or lepton
³⁸⁸ flavour violation. Among these, results recently obtained by the LHCb experiment
³⁸⁹ show a series of anomalies with respect to the SM that have the potential to yield
³⁹⁰ to BSM scenarios.

³⁹¹ 1.6.1 Dimuon decays of b hadrons

Decays of B mesons into a pair of muons are two-body decays where the two muons are back to back in the hadron rest frame. The simple signatures of these decays makes them easy to study and the fact that they are unaffected by hadronic physics in the final state makes predictions very clean and precise. Therefore these are essential tests of the SM. The $B^0 \rightarrow \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^-$ decays are FCNCs that can only happen via loops and furthermore they are CKM-suppressed, which makes

them particularly rare. In addition to that the decay of a pseudo-scalar B meson into two muons has a significant helicity suppression. The latest SM predictions for these decay rates are [41]:

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

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

The uncertainties on these values are dominated by the knowledge of the decay constants and CKM-elements. BSM models can produce significant enhancement to these decay rates and the measurement of their ratio is a stringent test of the MFV hypothesis. A combination of the LHCb and CMS results measured the values [42]:

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

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

Neither decay had been previously observed, while now the B_s^0 decay is observed with a significance of 6σ and evidence for the B^0 decay is found at 3σ significance level. The measured branching fractions are compatible with SM predictions within 2σ and put strong constraints on the available parameter-space for BSM theories. Figure 1.9 shows the fit the dimuon invariant mass of B meson candidates where the peaks of the two decays are visible.

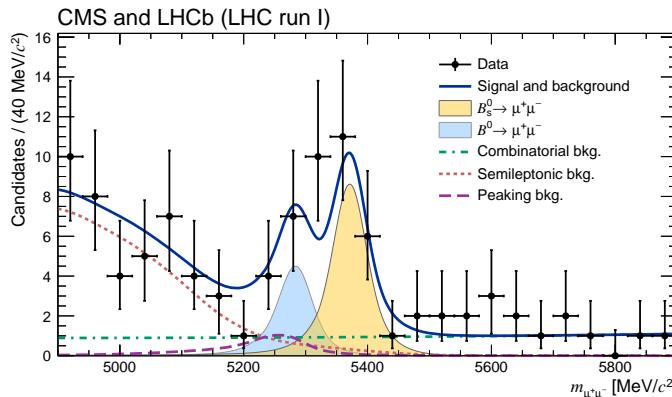


Figure 1.9: Dimuon invariant mass of B candidates showing peaks corresponding $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays [42].

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

³⁹⁹ At the LHC it is possible to collect large data samples of semileptonic decays, es-
⁴⁰⁰ pecially those with muons in the final state. Many branching fractions of semilep-
⁴⁰¹ tonic B meson decays were recently measured at the LHCb experiment, including
⁴⁰² $B \rightarrow K\mu^+\mu^-$, $B \rightarrow K^{*0}\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ [43, 44, 45]. Baryon decays were
⁴⁰³ also studied at LHCb: including the rare $\Lambda_b \rightarrow \Lambda\mu^+\mu^-$ decay [1], whose analysis is
⁴⁰⁴ described in this thesis. In contrast to purely leptonic decays, SM predictions for
⁴⁰⁵ semileptonic decays are affected by the knowledge of hadronic form factors, which
⁴⁰⁶ results in relatively large uncertainties, $\mathcal{O}(30\%)$. As a result measurements are now
⁴⁰⁷ typically more precise than predictions.

⁴⁰⁸ Among the measurements of angular observables that can be affected by new physics,
⁴⁰⁹ particular interest was raised by the measurement of a set of observables in $B^0 \rightarrow$
⁴¹⁰ $K^{*0}\mu^+\mu^-$ decays, free from form factors uncertainties at leading order [46]. Most
⁴¹¹ of the measurements are found to be in agreement with SM predictions with the
⁴¹² exception of the P'_5 observable, shown in Fig. 1.10, which presents a local 3.7σ de-
⁴¹³ viation confirmed by a recent analysis with higher statistics [47]. Attempts to build
⁴¹⁴ a consistent picture point to a new physics contribution to the Wilson Coefficient
⁴¹⁵ C_9 [48]. An angular analysis of $B^+ \rightarrow K^+\mu^+\mu^-$ decays was also performed, where
observables are found to be compatible with SM predictions [49]. Other observables

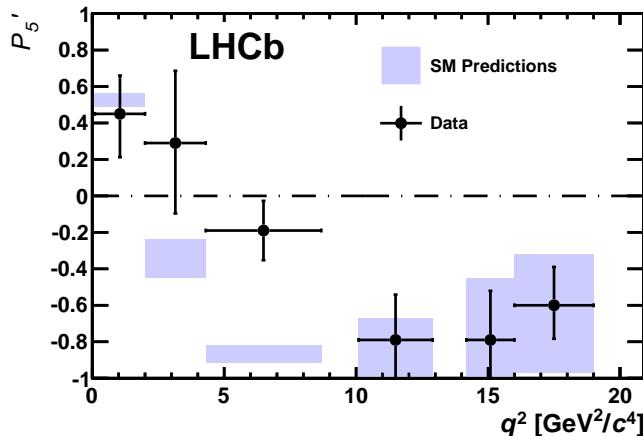


Figure 1.10: Measurement of the P'_5 observable as a function of q^2 , showing a tension with SM predictions in the 2–6 GeV^2/c^4 region [46].

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

Table 1.4: Measurement of CP and isospin asymmetry in $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays from the LHCb experiment [31].

for which the sensitivity to form factors effects is reduced are the CP asymmetry between B and \bar{B} decays, \mathcal{A}_{CP} , and the isospin asymmetry between B^0 and B^+ decays, \mathcal{A}_I . Due to the small size of the corresponding CKM elements, CP asymmetries of $B^0 \rightarrow K^{(*)} \mu^+ \mu^-$ decays are tiny in the SM, $O(10^{-3})$. In BSM models new sources of CP violation can arise and therefore \mathcal{A}_{CP} measurements are a powerful test of the SM. The isospin asymmetry is not zero in the SM due to isospin breaking effects in the form factors. This is expected to be $\sim 1\%$ at low q^2 and increase to $\sim 10\%$ as q^2 tends to zero. The LHCb experiment, using the full dataset collected in Run I, corresponding to an integrated luminosity of 3 fb^{-1} and $\sim 10^9$ B decays, measured both of these asymmetries to be consistent with zero [43, 50], as reported in Tab. 1.4. Recently, progress was made measuring also electron channels. The branching fraction of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay was measured to be $(3.1 \pm 1.3) \times 10^{-7}$ in the dilepton mass interval 30–1000 MeV/ c^2 [51]. Furthermore, for the first time angular observables were measured for this decay and found to be consistent with SM predictions [52].

Given the wide set of available measurements, theorists have implemented global fits including results from rare decays analyses, as well as inputs from B_s^0 mixing and Higgs measurements, in order to understand if the existing anomalies could be caused by a common factor. The results of such global fits agree that there is a tension with respect to the SM at the level of 1–4.2 standard deviations, depending on the set of assumptions made. In particular they favour a shift $C^{NP} \sim -1$ to the C_9 Wilson Coefficient, related with the penguin diagram mediated by a Z^0 boson [53, 48, 54].

439 1.6.3 Lepton Flavour Violation searches

440 Several Lepton Flavour Violation (LFV) searches are linked to rare decays as they
 441 involve small branching ratios in the SM that can be enhanced by BSM physics. Lep-
 442 ton flavour conservation is experimentally well-established measuring the branching
 443 ratios of decays of muons into electrons and no neutrinos, but has no strong the-
 444 oretical explanation in the context of the SM. In fact it is already observed that
 445 flavour is not conserved in neutrino oscillations. The best-studied decays violating
 446 lepton flavour are rare muon decays including $\mu^+ \rightarrow e^+ \gamma$ and $\mu^+ \rightarrow e^+ e^- e^+$. Since
 447 muons can be abundantly produced and the final states are simple, these decays
 448 provide the best constraints to LFV. The present best upper limits are 1.2×10^{-11}
 449 for the radiative decay and 1.0×10^{-12} for $\mu^+ \rightarrow e^+ e^- e^+$ obtained respectively by
 450 the MEGA [55] and SINDRUM [56] experiments. Several LFV searches in the B
 451 sector have recently been performed at the LHCb experiment including decays such
 452 as $B^0 \rightarrow e\mu$ [57] and τ decays such as $\tau \rightarrow \mu^+ \mu^- \mu$ [58]. None of these searches has
 453 found evidence of new physics so far and therefore they set limits, constraining the
 454 parameter space available for BSM models. Figure 1.11 shows a summary of the
 455 best limits set at different times on LFV searches [59].

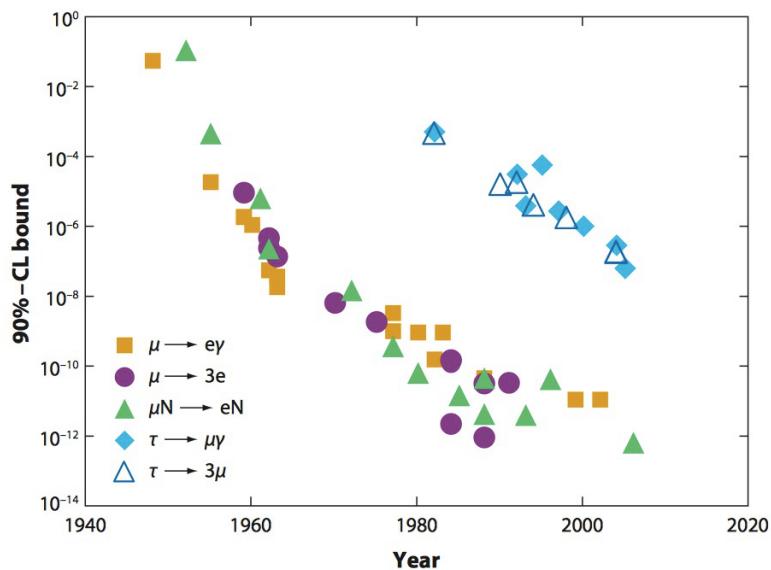


Figure 1.11: Summary of limits set in LFV searches as a function of time [59].

456

CHAPTER 2

457

458

The LHCb detector at the Large Hadron Collider

459

460

2.1 The Large Hadron Collider

461 The Large Hadron Collider (LHC) [60] is a synchrotron particle accelerator with a
462 circumference of 27 km located about 100 m underground at CERN in the surround-
463 ings of Geneva, Switzerland. Two proton beams circulate in opposite directions
464 around the ring and cross each other in four points, in which particle detectors are
465 placed. These include two general-purpose detectors, ATLAS and CMS, sitting on
466 opposites sides of the ring and two smaller detectors, ALICE and LHCb that are
467 designed to study specific topics (see Fig. 2.1).

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

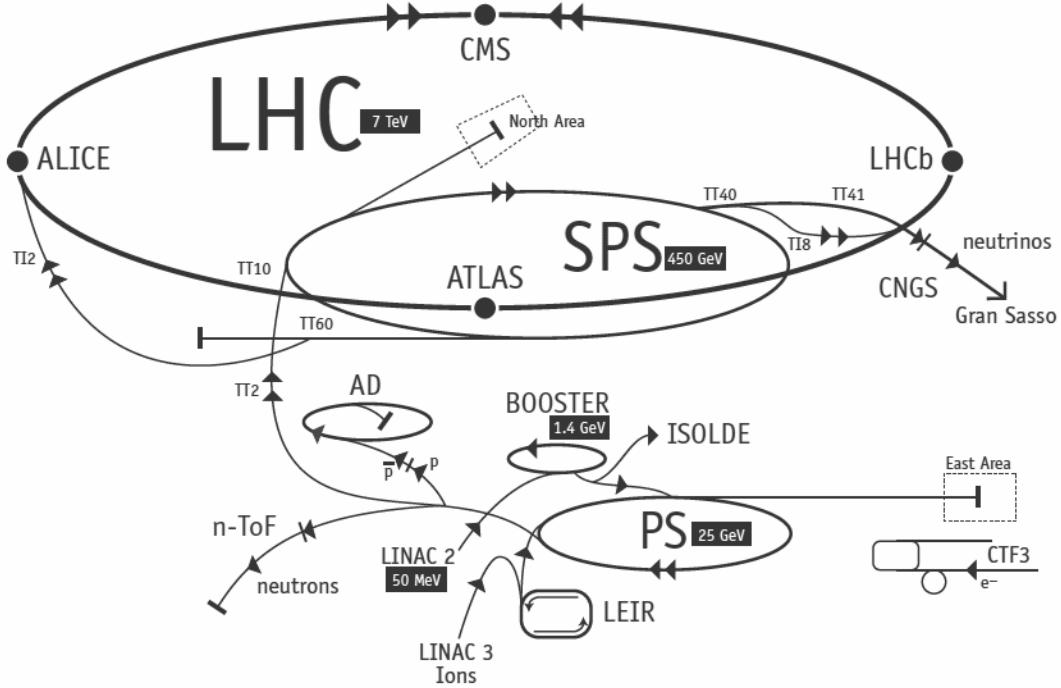


Figure 2.1: Schematic of CERN accelerators [61].

472 Protons are produced from hydrogen gas and are initially accelerated to an energy
 473 of 50 MeV in a linear accelerator (LINAC). Then they are injected into the Proton
 474 Synchrotron Booster (PSB), where they are boosted to an energy of 1.4 GeV, into
 475 the Proton Synchrotron (PS) to 25 GeV and into the Super Proton Synchrotron
 476 (SPS) to 450 GeV. Finally, protons enter into the LHC storage ring, where they are
 477 accelerated from injection energy to the final one by radio frequency (RF) cavities.
 478 The beams are steered around the ring by 8 T magnetic fields produced by 15 m
 479 long superconducting niobium-titanium dipole magnets and focused by quadrupole
 480 magnets. The LHC magnets use a design in which both proton beam pipes are
 481 contained in the same housing, allowing a common liquid helium cooling system to
 482 be used. The LHC began colliding proton beams in “physics mode” in 2009 at a
 483 centre of mass energy of $\sqrt{s} = 900$ GeV and from April 2010 to November 2011
 484 accelerated beams at $\sqrt{s} = 7$ TeV (3.5 TeV per proton beam) with a maximum
 485 instantaneous luminosity of $3 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, while in 2012 the energy was increased
 486 to 8 TeV. The LHC maximum design energy is 14 TeV, and its design luminosity is
 487 $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. After a long shut down to upgrade and maintain the machine, a new

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

491 **2.2 The LHCb detector**

492 The LHCb detector [62] was designed to study decays of B and D mesons, mainly
493 looking for CP-violating processes. In 2011, running at a centre of mass energy of 7
494 TeV, the cross-section for $b\bar{b}$ production was measured to be $284 \pm 53 \mu\text{b}$ [63], while
495 it will be $\sim 500 \mu\text{b}$ at the current LHC energy, 13 TeV. At these high energies,
496 proton-proton interactions produce highly boosted virtual gluons which produce $b\bar{b}$
497 pairs at small angles, close to the beam pipe. For this reason the LHCb detector is
498 designed to have a very forward angular coverage. The detector is fully instrumented
499 from 10 mrad to 300 mrad, corresponding to an interval $2 < \eta < 5$, where η is the
500 “pseudorapidity”, a quantity defined as:

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

501 where θ is the angle between a particle’s momentum and the beam direction¹.

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

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 centre of the interaction area.

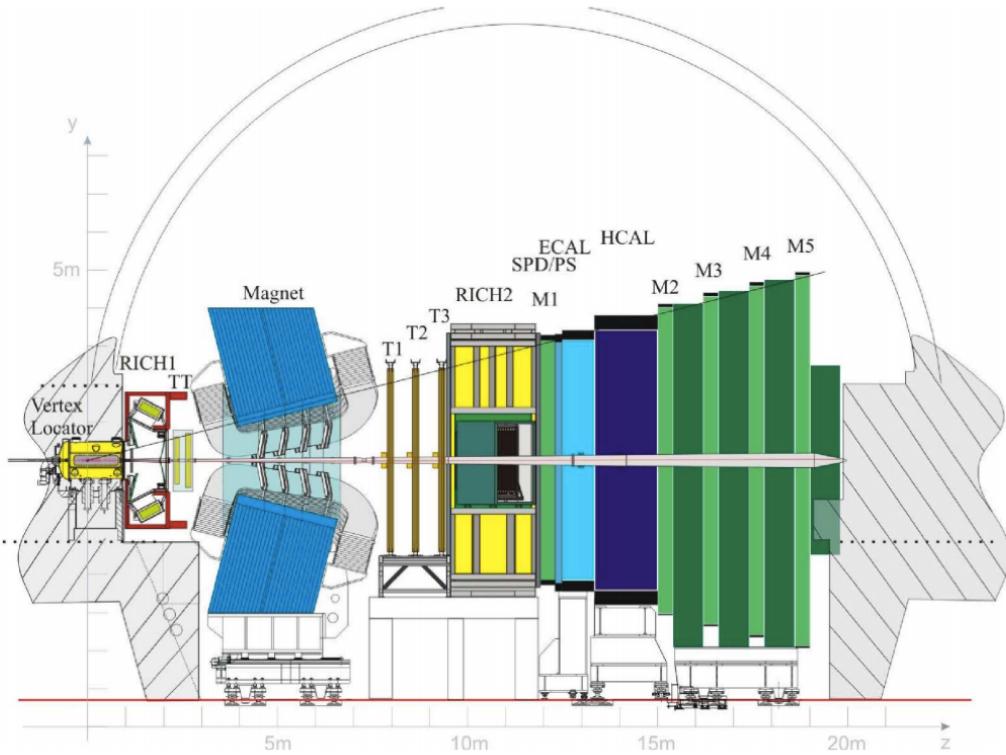


Figure 2.2: A side view of the LHCb detector [62].

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

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

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

524 Imaging Cherenkov detectors (RICH) [64]. Photon, electron and hadron candidates
525 are identified by a calorimeter system and muons by a system composed of alter-
526 nating layers of iron and multi-wire proportional chambers [65]. A schematic view
527 of the detector is shown in Fig. 2.2 and more details on each sub-detector are given
528 in the following sections.

529 2.3 The magnet

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

543 2.4 Tracking system

544 B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means
545 that they travel about 1 cm before decaying to form a displaced vertex. To study
546 specific decays, it is therefore important to be able to separate the particles pro-
547 duced at the primary pp vertex and at the B decay secondary vertex (SV). The
548 tracking system consists of the Vertex Locator (VeLo), and four tracking stations:

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

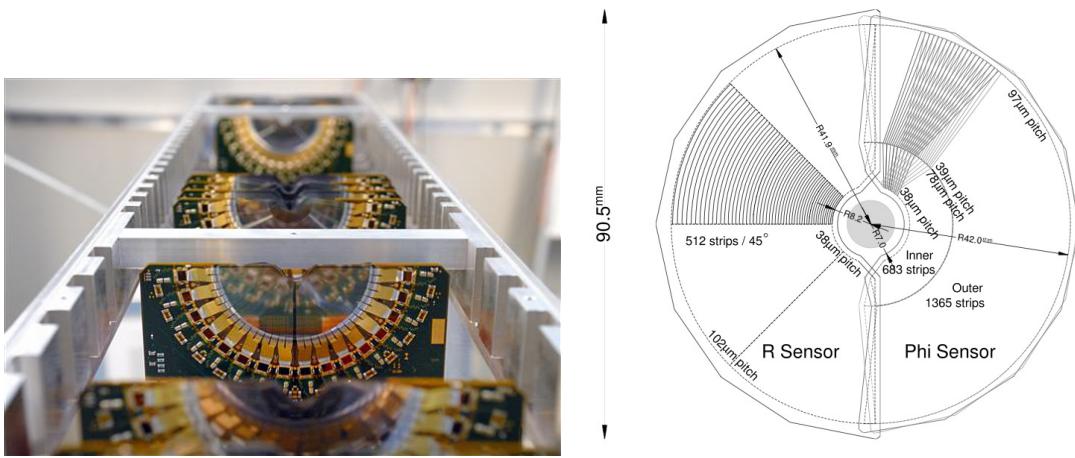


Figure 2.3: On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [62].

554

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

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

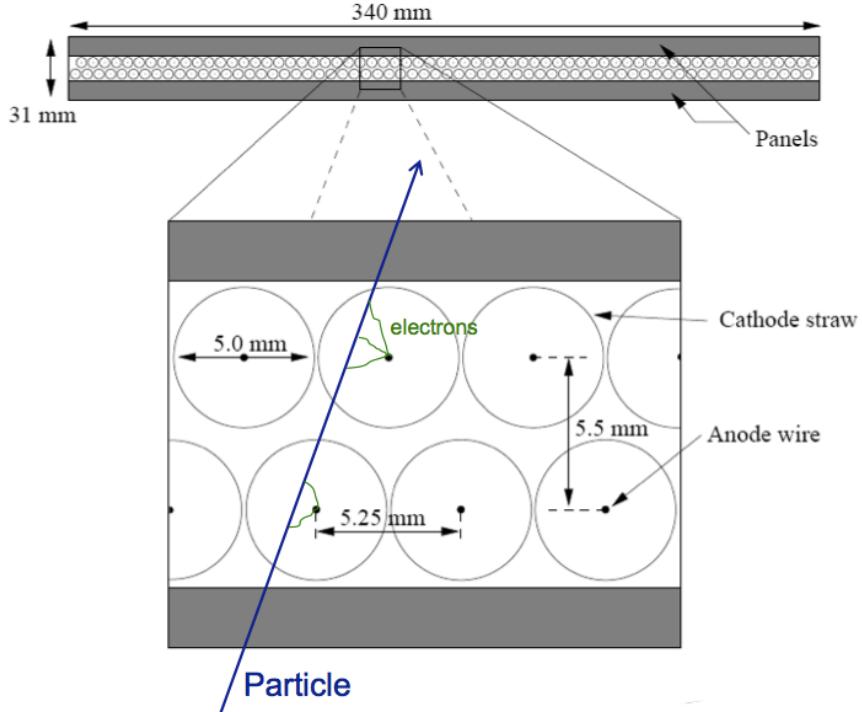


Figure 2.4: A sketch of the straw tubes which constitute the OT layers [62].

570

571 The IT and TT both use silicon strips and together constitute the Silicon Tracker
 572 (ST). Straw tubes are instead used in the OT, of which a sketch is shown in Fig. 2.4.
 573 The IT requires a higher inner granularity because of the greater flux of particles
 574 close to the beam pipe. In fact, it covers only 1.3% of the total area of IT plus
 575 OT but it contains about 20% of the tracks. Each ST station has four detection
 576 layers: the first and last are vertical, measuring the track position in x , while the
 577 second and third layers are rotated by an angle of +5 and -5 degrees, which allows
 578 the measurement of the y coordinate. The TT is placed upstream of the magnet to
 579 allow the reconstruction of tracks from low-momentum particles, which are bent out
 580 of the downstream acceptance. Overall the tracking system provides a measurement
 581 of momentum, p , with a relative uncertainty that varies from 0.4% at 5 GeV/ c to
 582 1.0% at 200 GeV/ c . The impact parameter (IP), namely the minimum distance of a

track to a primary vertex, is measured with a resolution of $(15 + 29/p_T) \mu\text{m}$, 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 50–60 μm . The decay products of B mesons tend to have high IP values because the B decay imparts transverse momentum to them. Therefore, accurate IP and vertex displacement measurements allow LHCb to distinguish effectively between B meson decays and background processes.

2.5 Calorimeters

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

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

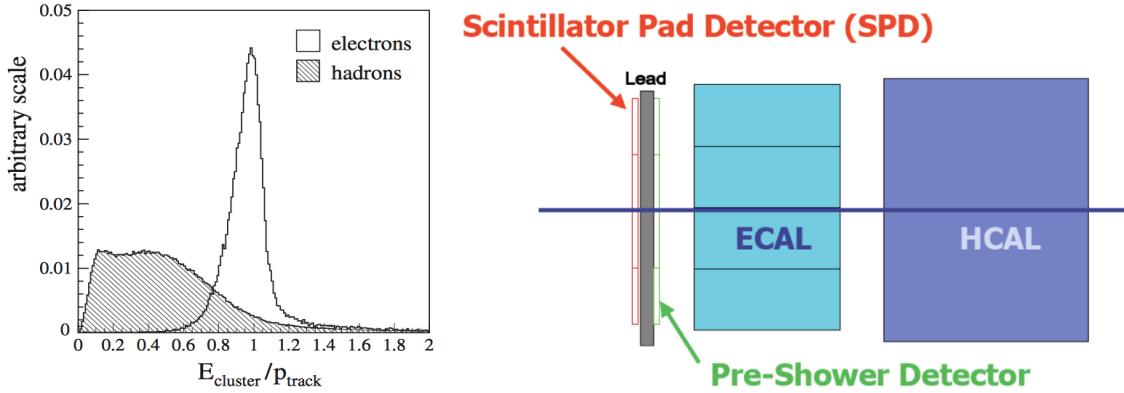


Figure 2.5: (left) The ratio of the energy deposited in the ECAL and the particle momentum, which allows the separation between electrons and hadrons [62]. (right) A schematic of the LHCb’s calorimeter system.

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

616 The ECAL is formed by 66 lead layers (2 mm thick) separated by 4 mm thick plastic
 617 scintillator layers. In order to obtain the highest energy resolution the showers
 618 from high energy photons must be fully absorbed. For this reason the ECAL has a
 619 thickness of 25 radiation lengths and its resolution is measured to be $\sigma_{\text{ECAL}}(E)/E =$
 620 $10\%/\sqrt{E(\text{GeV})} + 1\%$ [62], which results in a mass resolution of $\sim 70 \text{ MeV}/c^2$ for
 621 B mesons and $\sim 8 \text{ MeV}/c^2$ for π^0 . The HCAL is mainly used for triggering and
 622 it is similar to the ECAL but with 4 mm thick scintillator layers and 16 mm thick
 623 absorber layers. The trigger requirements on the HCAL resolution do not depend
 624 on the containment of the hadron showers as much as for the ECAL, therefore, due
 625 to space limits, its thickness is only 5.6 interaction lengths and its resolution is given
 626 by $\sigma_{\text{HCAL}}(E)/E = 69\%/\sqrt{E(\text{GeV})} + 9\%$.

⁶²⁷ 2.5.1 Bremsstrahlung recovery for electrons

628 Bremsstrahlung is an electromagnetic radiation produced by charged particles that
 629 undergo an acceleration. Typically electrons produce Bremsstrahlung when de-
 630 flected by atomic nuclei. The probability of emitting bremsstrahlung radiation is
 631 proportional to the inverse of the squared mass of the particle ($1/m^2$) and therefore
 it is most relevant for electrons. At LHC energies, if electrons radiate after the mag-

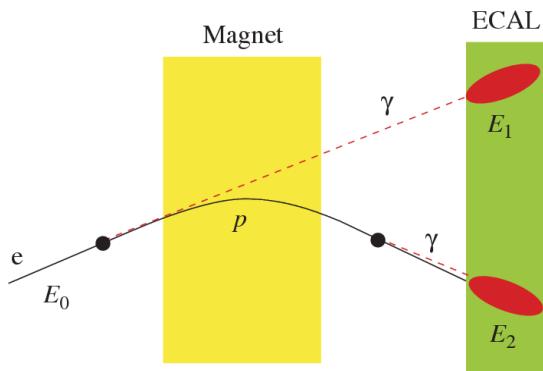


Figure 2.6: Schematic view of the bremsstrahlung recovery [62].

632 net, the photon will hit the same calorimeter cell as the electron and the energy will
 633 be automatically recovered, as illustrated in Fig. 2.6. However, if the photon is emit-
 634 ted before the magnet, the electron will be deflected by the magnetic field whereas
 635 the photon will continue on its initial trajectory, with its energy being deposited in
 636 a different part of the calorimeter. Missing this energy results in a poorer recon-
 637 structed invariant mass resolution, so it is desirable to recover these bremsstrahlung
 638 photons. A tool for bremsstrahlung recovery is available in the LHCb analysis soft-
 639 ware. This tool looks for other clusters in the calorimeter and, reconstructing the
 640 trajectory of the electron, checks if they may be associated with emitted photons.
 641 The photon energy is then added to the electron and its momentum is recalculated.
 642 For more information see Ref. [66].

644 2.6 RICH

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

$$\cos \theta = \frac{1}{\beta n}, \quad (2.2)$$

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

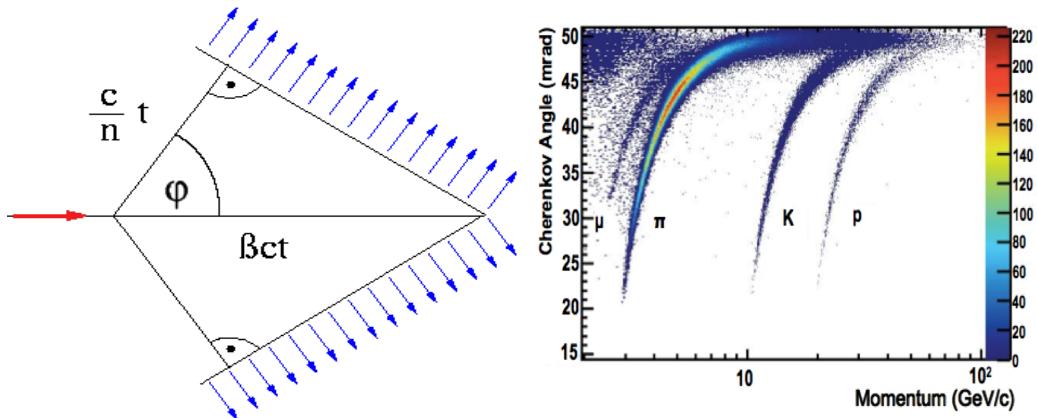


Figure 2.7: (left) A sketch of Cherenkov light emission [67]. (right) Measured Cherenkov angle as a function of particle momentum [62], where one can see that the study of the Cherenkov angle allows to distinguish particles' identities.

651

652 RICH 1 is located before the magnet in order to cover a larger angular accep-
 653 tance. Its purpose is to ensure particle identification over the momentum range
 654 $1 < p < 70 \text{ GeV}/c$. It uses two radiators: C_4F_{10} that covers the momentum range
 655 $5 - 70 \text{ GeV}/c$ and silica aerogel which covers $1 - 10 \text{ GeV}/c$. RICH 2 is positioned
 656 after the magnet and tracking stations and it identifies higher momentum particles
 657 from approximately $20 \text{ GeV}/c$ up to beyond $100 \text{ GeV}/c$ using CF_4 as a radiator.
 658 The Cherenkov light produced when charged particles travel through the radiators,
 659 is reflected and focussed using mirrors, which are tilted so that a ring image is re-

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

2.7 The muon system

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

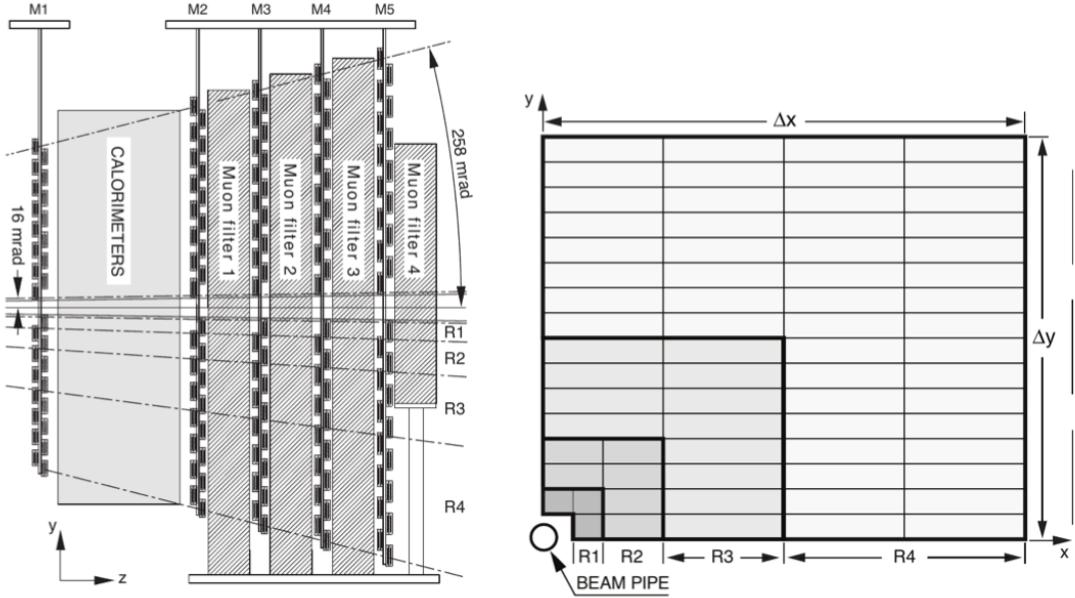


Figure 2.8: The LHCb muon system [62].

686 2.8 Particle identification

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

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

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

⁷⁰¹ Log-Likelihood (DLL) and denoted with ‘‘PID’’. For example,

$$\text{PID}_K = \text{DLL}_{K-\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi) \quad (2.3)$$

⁷⁰² quantifies the probability of a particle being a kaon rather than a pion. Figure 2.9
⁷⁰³ shows the efficiency for correctly identifying and mis-identifying kaons and protons as
⁷⁰⁴ a function of the measured momentum of the particle. For kaons the efficiency drops
⁷⁰⁵ at momenta below 10 GeV, where they fall below threshold for the gas radiators.
⁷⁰⁶ The DLL cuts enable LHCb physics analyses to distinguish between kinematically
⁷⁰⁷ similar decays with different final states. For example, Fig 2.10 illustrates the power
⁷⁰⁸ of particle identification, showing how the application of DLL cuts can be used to
isolate $B^0 \rightarrow \pi^+\pi^-$ decays from other two-body B decays. The identification of

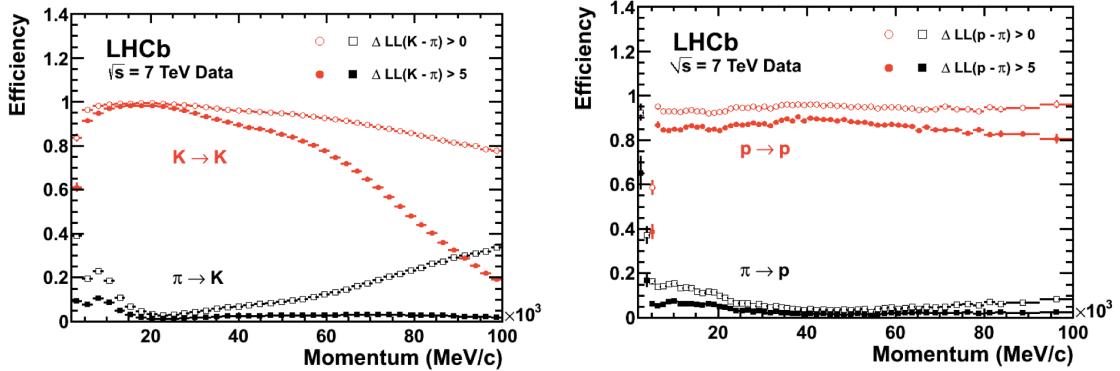


Figure 2.9: Particle identification performances for kaons (left) and protons (right) as a function of the measured momentum of the particles [62].

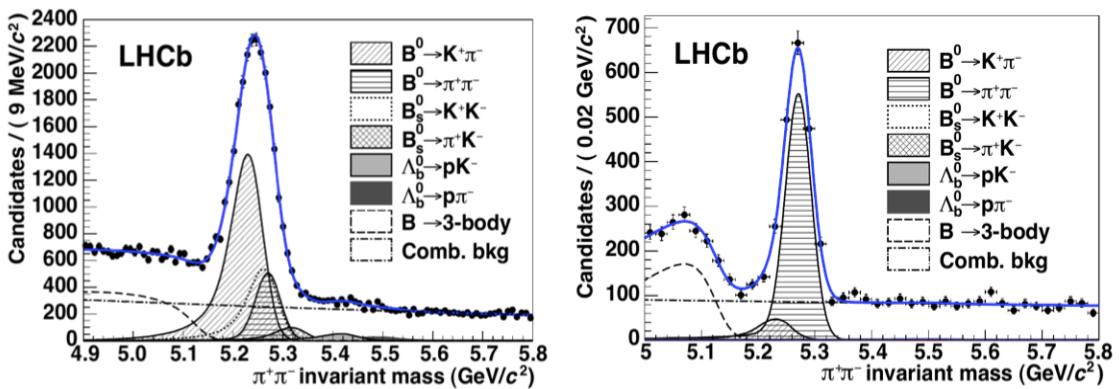


Figure 2.10: Invariant mass peak of the $B^0 \rightarrow \pi^+\pi^-$ decay before (left) and after (right) the application of PID requirements [68].

709

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

715 2.8.1 PID calibration

716 In order to be able to calculate detection efficiencies, a “data-driven” method was
717 developed. The calibration software is referred to as `PIDCalib` package [68]. This
718 tool uses decays where final particles can be identified thanks to their kinematic
719 properties. For example the $K_s^0 \rightarrow \pi^+\pi^-$ decay has a clear signature with a displaced
720 vertex and can be easily singled out from other decays and used to test pion ID
721 efficiency. The narrow peaks of the $J/\psi \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow e^+e^-$ decays allow
722 muon and electron efficiencies to be calibrated. A “tag-and-probe” method is used
723 in this case, where only one of the two leptonic tracks is reconstructed requiring
724 the correct identity and the other one is used to probe the PID efficiency. Finally,
725 $\phi \rightarrow KK$ samples and $D^{*+} \rightarrow D(\rightarrow K^-\pi^+)\pi^+$ decays, where the D^{*+} is used to tag
726 the decay, are used to test the kaon efficiency. In all cases the residual background
727 is subtracted using the `sPlot` technique [69].

728 2.9 Trigger and software

729 The LHCb trigger system [70] consists of a hardware stage, L0, based on information
730 from the calorimeters and muon system, followed by a software stage, the High-
731 Level Trigger (HLT), which applies a full reconstruction of the events. To increase
732 performance, the HLT is further split into two stages, HLT1 and HLT2. The HLT1
733 phase happens in real time and saves data to local disks while the HLT2 phase uses

the resources available during periods with no beam. The event selected by the HLT2 stage are then saved for offline analysis. Figure 2.11 shows a schematic of the trigger system. The bunch crossing frequency is 40 MHz, which corresponds to an instantaneous luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ for LHCb. About 15% of the total number of $b\bar{b}$ pairs produced will contain at least one B meson with all of its decay products within the detector acceptance. This rate needs to be reduced to about 2 kHz at which the events can be written to disk.

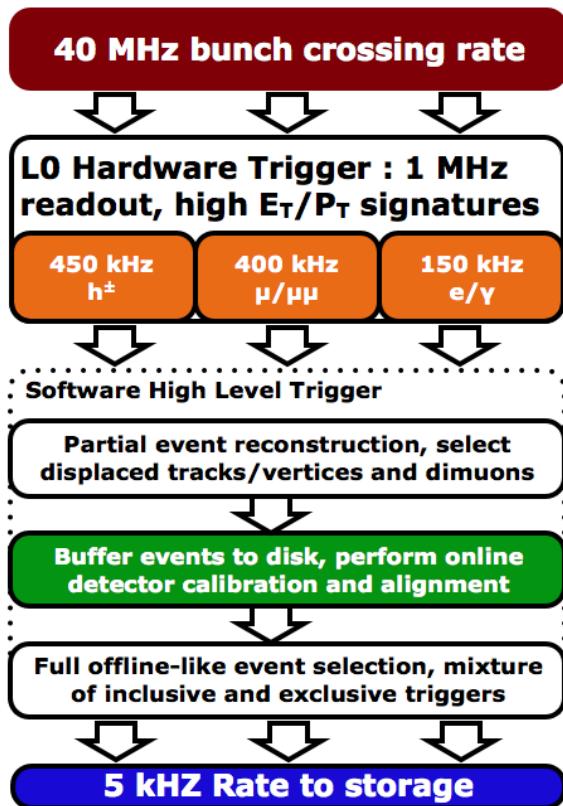


Figure 2.11: A schematic of the LHCb trigger system [62].

740

The L0 trigger reduces the rate of visible interactions from 10 MHz to 1 MHz. Due to the heavy mass of B mesons, they often produce particles with high energy and momentum. Therefore the trigger selects events with large deposits in the calorimeter or high p_T muons. The event is classified as L0Muon if it was triggered due to information from the muon detector, while the information from the calorimeters is used to divide the events into five categories: L0Photon, L0Electron, L0LocalPion, L0GlobalPion, L0Hadron. The PS detector information is converted to a photon

748 flag (`PS && !SPD`) or an electron flag (`PS && SPD`). The “local” label of the `L0Pion`
749 trigger refers to π^0 reconstructed through their $\gamma\gamma$ decay, where the two photons fall
750 in the same ECAL element, they are labelled “global” otherwise. The first four
751 calorimeter triggers require energy clusters in the ECAL, while `L0Hadron` requires
752 clusters also in the HCAL. The `HLT1` uses information from the VELO and trackers
753 performing a partial reconstruction of the event and reduces the rate to 2 kHz by
754 adding requirements on the IP and χ^2 of tracks. Finally, the `HLT2` involves a full
755 reconstruction of the event and includes many “lines” designed to select specific
756 decay structures.

757 LHCb also developed an extended simulation software in order to reconstruct ef-
758 ficiencies and signal shapes. In the simulation, pp collisions are generated using
759 `PYTHIA8` [71, 72] with a specific LHCb configuration [73]. Decays of hadronic par-
760 ticles are described by `EVTGEN` [74], and final state radiation is generated using
761 `PHOTOS` [75]. Finally, the interaction of the generated particles with the detec-
762 tor and its response are implemented using the `GEANT4` toolkit [76] as described
763 in Ref. [77]. For this analysis in this thesis, the `ROOT` framework [78] is used to
764 analyse data and the `RooFit` package to perform maximum likelihood fits. A multi-
765 variate analysis is also performed based on the `NeuroBayes` package [79, 80], which
766 provides a framework for neural network training.

767 2.10 Constrained kinematic fits

768 The resolution of key variables, such as the measured invariant mass of decaying
769 particles, can be improved by imposing constraints on the measured quantities to
770 remove redundant degrees of freedom. The four-momentum conservation can be
771 ensured at each vertex and the origin and decay vertices of a particle are related via
772 the momentum of the particle. Furthermore, additional constraints can be imposed
773 due to a particular decay hypothesis such as the known invariant masses of final and
774 intermediate particles. In order to do this the `DecayTreeFitter` tool was developed

775 by the BaBar experiment and later used by LHCb [81]. The algorithm takes a com-
 776 plete decay chain and parametrises it in terms of vertex positions, decay lengths
 777 and momentum parameters. These parameters are then fit simultaneously, taking
 778 into account the relevant constraints, including the information from photons. Fig-
 779 ure 2.12 illustrates the effect of the application of the kinematical fit on the 4-body
 780 invariant mass of the final daughters of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay. The resolution in this
 781 case improves by over a factor of 2. Furthermore, the χ^2 from the kinematic fit can
 782 be used to quantify the compatibility with a specific decay structure, which helps to
 783 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.

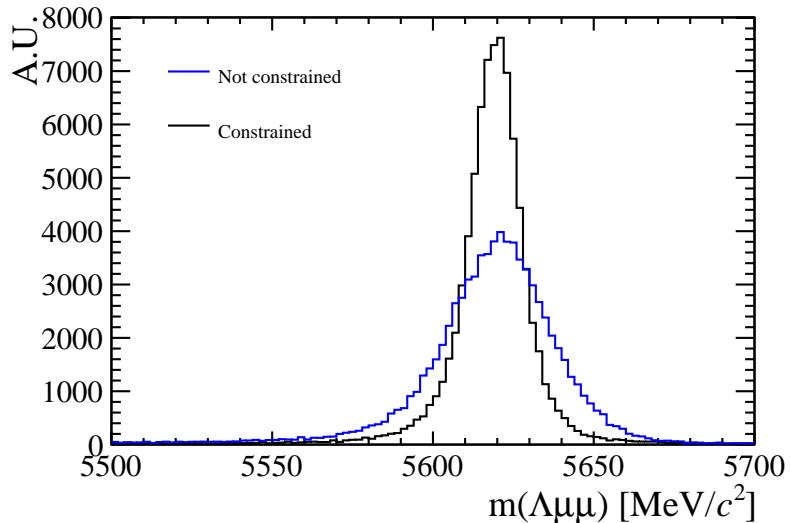


Figure 2.12: Invariant mass of the final daughters of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays calculated with and without constraints using the `DecayTreeFitter` tool.

784

785 2.11 Validation of hadronic processes in the simulation

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

790 The LHCb simulation software propagates particles through the detector using the
791 GEANT4 toolkit [62]. This offers a variety of models for physics processes over a
792 wide range of energies for both electromagnetic and strong interactions. Given a
793 combination of projectile, target and energy there can be several models applicable
794 with different reliability and computational costs. GEANT4 provides a number of
795 pre-packaged physics lists each representing complete and consistent sets of models
796 chosen to be appropriate for a given use case. In LHCb mainly two hadronic physics
797 lists are considered:

798 • **LHEP** (Low and High Energy Parametrisation): based on a parametrised
799 modelling of all hadronic interactions for all particles. This list combines
800 the High Energy Parametrised model (HEP) and the low energy one (LEP).

801 There is a sharp switch from the low to the high energy model at 25 GeV.
802 The modelling of elastic scattering off a nucleus and of nuclear capture also
803 proceeds via parametrised models.

804 • **FTFP_BERT**: includes the following models:

805 – Bertini cascade model (BERT) [82], which simulates the intra-nuclear cas-
806 cade, followed by pre-equilibrium and evaporation phases of the residual
807 nucleus, for protons, neutrons, pions and kaons interaction with nuclei
808 at kinetic energies below 9.9 GeV. The Bertini model produces more
809 secondary neutrons and protons than the LEP model, yielding a better
810 agreement with experiment data.

811 – FTFP model, which implements high energy inelastic scattering of hadrons
812 by nuclei using the FRITIOF model [83].

813 The change between the two models happens with a linear shift from BERT
814 to FTFP that starts at 4 GeV and ends at 5 GeV.

815 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** [77], which is interfaced to **GEANT4** v95r2p1. A simple geometry setup is used in order to be able to calculate in a clean way the interaction cross-sections in a specific material. This is constituted by a series of rectangular boxes filled with the most relevant materials for LHCb: Aluminium, Silicon and Beryllium. For each material three boxes are defined with different thicknesses (1mm, 10mm, 50mm). These values are chosen to be indicative of the amount of material present in the LHCb detector.

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

$$P_{int} = \frac{N_{int}}{N_{tot}}, \quad (2.4)$$

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

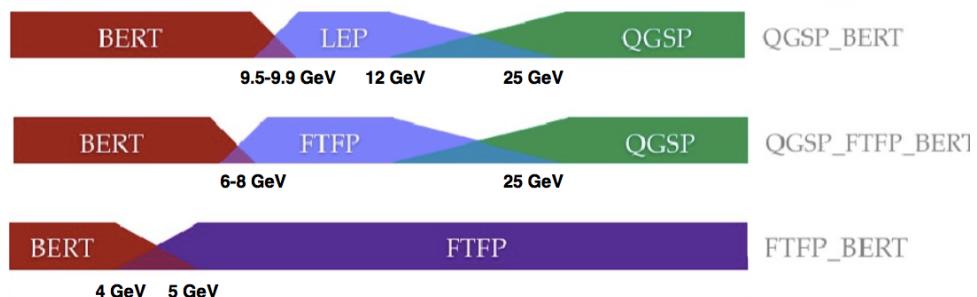


Figure 2.13: Diagram of LHEP, FTFP_BERT and QGSP_BERT models' composition as a function of energy.

835 of a particle (e.g. 121 for inelastic interaction, 111 for elastic, 201 for decay) it
836 is possible to distinguish the inelastic and elastic probabilities of interaction and
837 therefore cross-sections.

838 To compare simulation and data the cross-section and P_{int} are related through the
839 following formula valid for thin layers:

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

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

842 2.11.2 PDG prediction

In the Review of Particle Physics (PDG) [2] cross-sections of protons and neutrons
are parametrised as:

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

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

843 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
844 equations are universal and valid for any kind of collision: $M = 2.15$, $\eta_1 = 0.462$, η_2
845 = 0.551, $\lambda = 1$ (for p, n and γ) and 1.63 (for d). The other ones are characteristic
846 of each type of collision and are listed in Tab. 2.11.2. In these formulae the particle-
847 antiparticle asymmetry arises from the last term which has opposite sign in the
848 two equations. This term becomes less and less important with increasing energies.
849 Therefore a net asymmetry is found at low energies, while the cross-sections tend
850 to a common point at high energy and continue increasing logarithmically.

2.11.3 Validation results

This section reports particle and antiparticle cross-sections and their ratios compared, where available, with predictions and with data from the COMPASS experiment [84]. Figure 2.14 shows the probability of interaction for protons and anti-protons in 10 mm 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 these studies is not limited to cross-sections but can also give information on other simulated quantities. As an example, Fig. 2.11.3 shows a comparison between the types of particles generated in inelastic collisions of protons and anti-protons onto Aluminium using different models. Physics lists can give very different results, for example the LHEP model does not produce photons in inelastic collisions. However, it is difficult to use these quantities for validation as there is no data available for a comparison.

872

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.

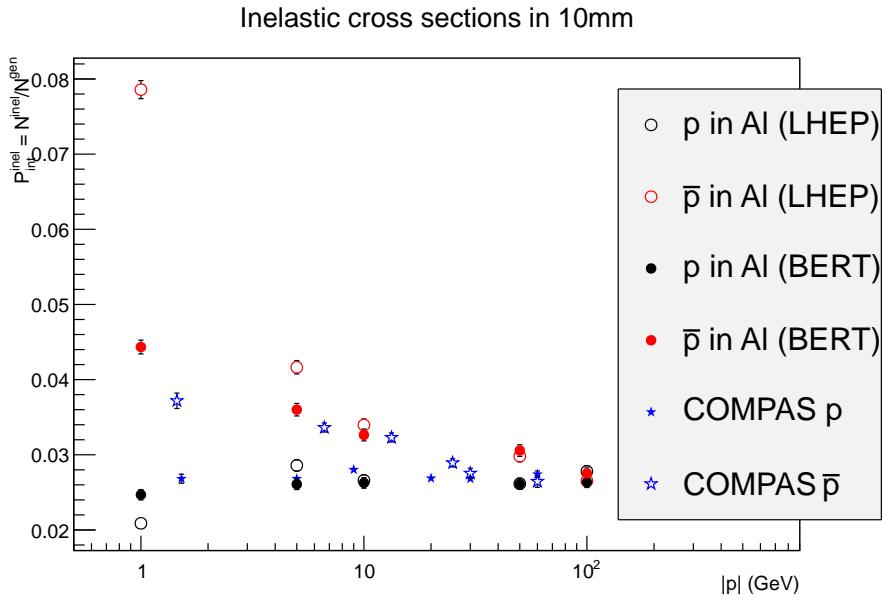


Figure 2.14: Probability of interaction for protons and anti-protons in Aluminium as a function of the projectile momentum. Two physics lists are used to generate events that can be compared with data from the COMPASS experiment.

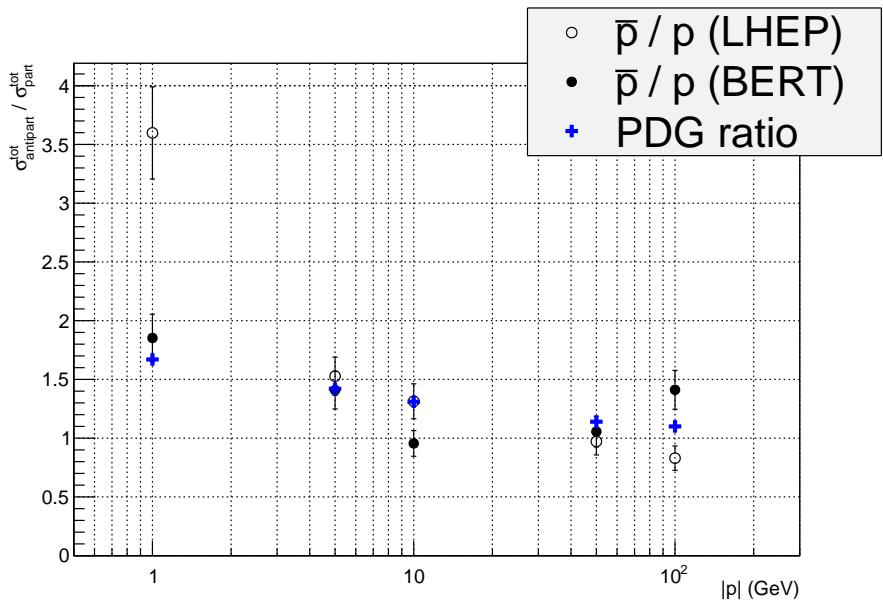


Figure 2.15: Ratio of antiproton over proton total interaction cross-section as a function of energy compared with PDG predictions.

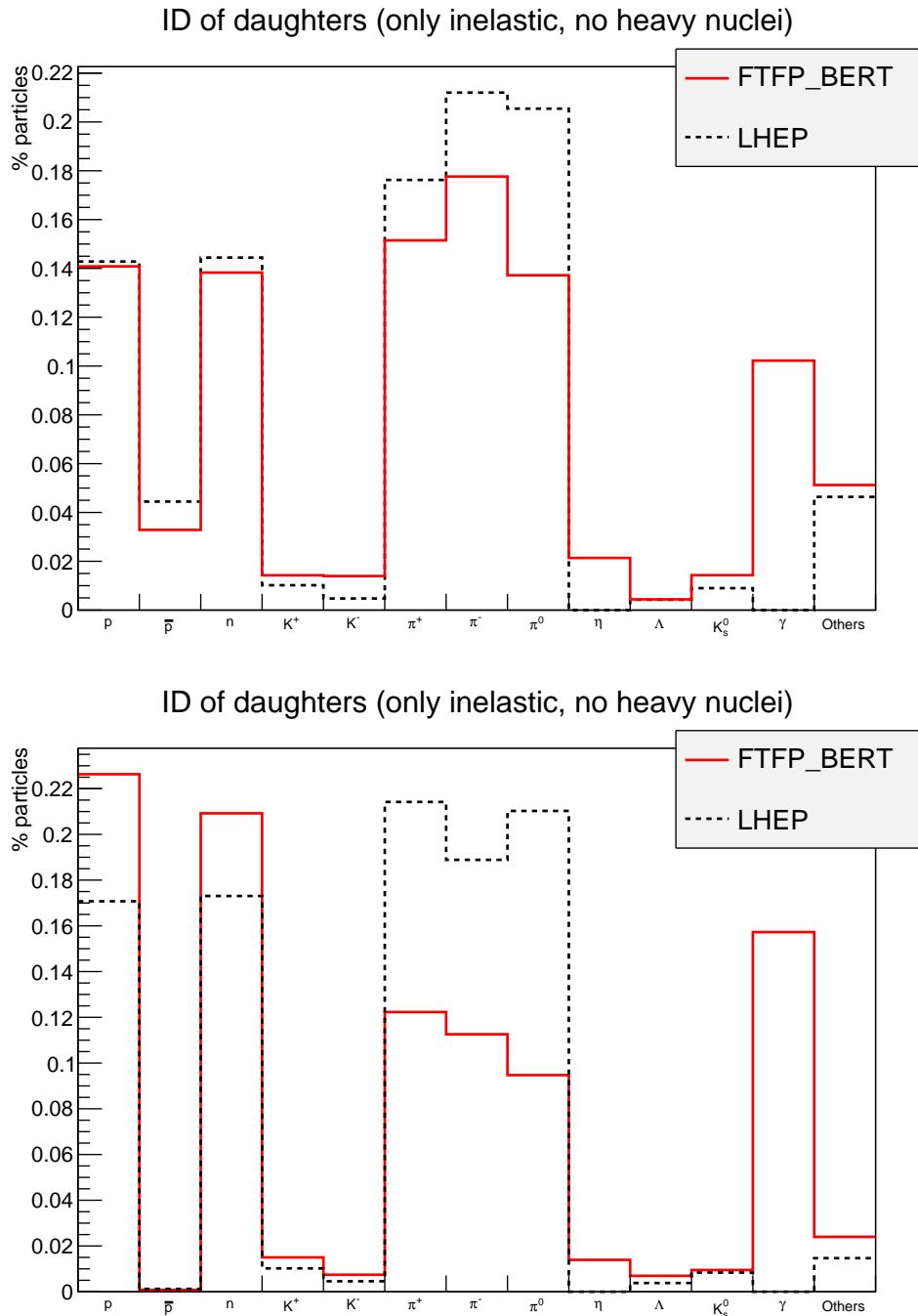


Figure 2.16: Composition of secondary particles produced in 100 GeV protons (top) and anti-protons (bottom) collisions in 1 mm Aluminium.

873 2.12 Material budget studies

874 It is important for many analysis to quantify the amount of material present in the
875 detector, for example to estimate the amount of multiple scattering. In GEANT4
876 particles are propagated in steps through the detector and for each step the frame-
877 work analyses the geometry to understand in what material the particle is and
878 modifies its trajectory accordingly. A tool was developed where neutrinos are used
879 as probes to scan the detector summing the radiation length seen at each step up
880 to a certain point. Neutrinos are used as they do not bend in magnetic field and
881 do not interact with the detector to any appreciable extent. Thin air planes are
882 inserted after each sub-detector. When these are traversed by the neutrinos, the
883 information about the accumulated radiation and interaction length is saved. In
884 this way it is possible to obtain maps of the detector, such as the one shown in
885 Fig. 2.17. Using the tool developed for this study it is also possible to obtain the
886 cumulative interaction length. As an example Fig. 2.18 shows the average radiation
887 length as a function of the distance from the interaction point. Furthermore, it is
888 possible to displace the primary vertex from its position, normally set at the origin,
889 in order to study how this translates into the amount of material traversed.

890 2.13 Validation and material budget studies conclusions

891 The studies outlined in the previous two sections are based on tools which are
892 now officially part of the LHCb simulation framework. These tools were used to
893 validate the framework when passing from GEANT4 version 9.5 to version 9.6. In
894 particular a patch was provided by the GEANT4 team including improved kaon cross
895 sections and it was verified these go into the right direction. The tool will continue
896 to be used in the future, in particular to validate the upgrade to GEANT4 10, in
897 2016. Furthermore, the tools can be used by analyses sensitive to the quality of the
898 simulation of particle and antiparticles cross-sections in order to study systematic
899 effects and uncertainties.

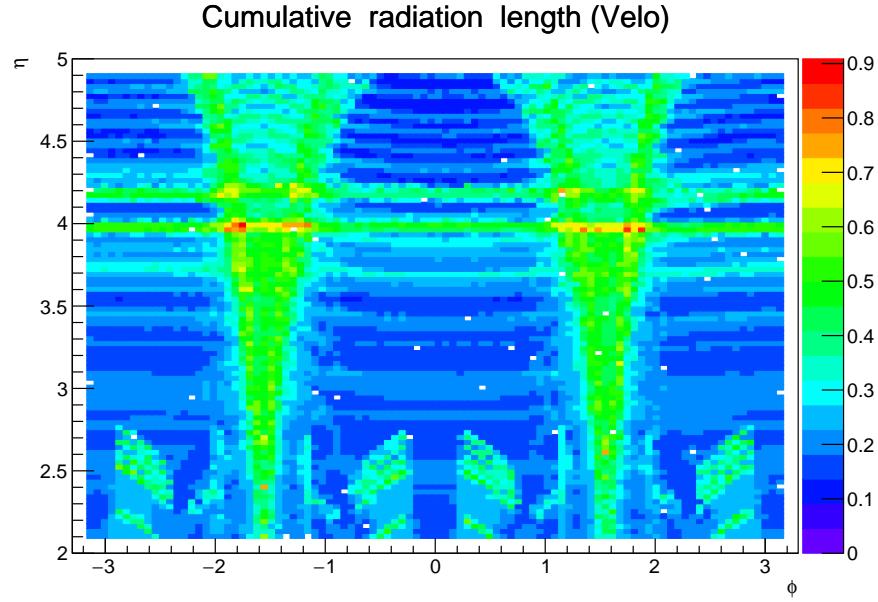


Figure 2.17: Map of cumulative radiation length seen by a particle starting from the interaction point up to the end of the VeLo.

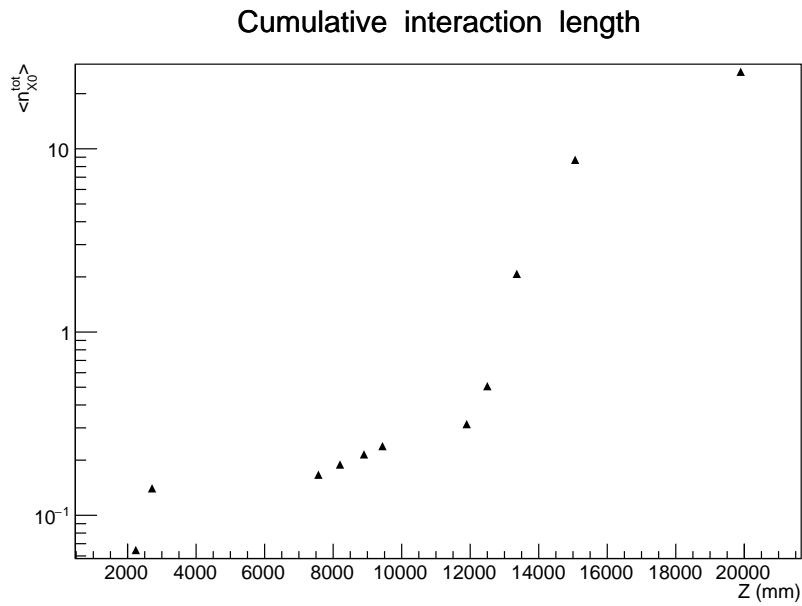


Figure 2.18: Average cumulative radiation length as a function of the horizontal distance from the interaction point. Each considered point corresponds to the end of a sub-detector: VeLo, RICH1, RICH2, tracking stations, ECAL and HCAL and muon detector.

900

CHAPTER 3

901

902

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

903

The $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay is a FCNC process governed by the $b \rightarrow s\mu^+\mu^-$ quark level transition which, in the SM, proceeds only through loop diagrams (electroweak penguin or W box) as discussed in Sec. 1.5, and therefore it is highly sensitive to new particles entering the loops. Interest in Λ_b^0 baryon decays arises, first of all, from the fact that the Λ_b^0 has non-zero initial spin, which allows to extract information about the helicity structure of the underlying Hamiltonian that cannot be obtained from meson decays [85, 86]. Secondly, the Λ_b^0 baryon can be considered to a first approximation as composed of a heavy quark and a light di-quark, therefore the hadronic physics differs significantly from B meson decays. This provides the possibility to better understand and test the hadronic physics in the theory, which could yield an improved understanding relevant also for the meson case.

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

918 decay weakly, which results in complementary constraints with respect to B^0 de-
 919 cays. Finally, the narrow width approximation, used in theoretical calculations, is
 920 fully applicable in the Λ_b^0 case, which has $\Gamma_{\Lambda_b^0} \sim 2.5 \cdot 10^{-6}$ eV. This is not the case
 921 for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays because the contribution from the non-resonant channel
 922 $B^0 \rightarrow K \pi \mu^+ \mu^-$ is unconstrained.

923 The theory of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays was widely investigated both in the context of
 924 the SM and in various BSM scenarios [87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97]. All
 925 authors start from the effective Hamiltonian outlined in Sec. 1.5.1. However, form
 926 factors, describing hadronic physics are not as well-developed as for the meson case
 927 because there are fewer experimental constraints, which leads to a relatively large
 928 spread in predicted branching fractions. For these reasons an interesting quantity
 929 to study is the differential branching fraction as a function of q^2 . This still suffers
 930 from the limited knowledge of form factors but, as different approaches to form
 931 factors calculations are applicable in different q^2 regions, it allows a more meaningful
 932 comparison with theory.

Experimentally, the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay was observed for the first time in 2011 by
 the CDF collaboration [98] and later updated in preliminary form using their full
 statistics [99]. The latter measurement yields $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-) = [1.95 \pm 0.34(\text{stat}) \pm$
 $0.61(\text{syst})] \times 10^{-6}$ and the signal was observed only in the q^2 region above the square
 of the $\psi(2S)$ mass. Recently, the decay was also observed at LHCb [100] with a yield
 of 78 ± 12 signal events using 1 fb^{-1} of integrated luminosity collected in 2011. The
 signal was also 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, which is used as a 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 the absolute branching fraction,

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

933 This chapter describes the measurement of the differential branching fraction of the
 934 $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay using 3 fb^{-1} of pp collisions collected by the LHCb experiment
 935 in 2011 and 2012.

936 3.1 Analysis strategy and q^2 regions

937 A typical q^2 spectrum of $b \rightarrow s\ell^+\ell^-$ decays was shown in Fig. 1.8. This is charac-
 938 terised by the presence of the photon pole at low q^2 and the narrow peaks of the
 939 J/ψ and $\psi(2S)$ resonances at intermediate values of q^2 . In the analysis, $\Lambda_b^0 \rightarrow J/\psi\Lambda$
 940 decays, in which the J/ψ decays into two muons and therefore has the same final
 941 state as the signal, are used as the normalisation channel. The rare and normal-
 942 isation channels are naturally distinguished by the q^2 intervals in which they are
 943 reconstructed. The Λ decay mode into a pion and a proton, $\Lambda \rightarrow p\pi$, is always used
 944 to reconstruct the decays. The intervals in which the rare channel is studied are:

- 945 • $0.1 < q^2 < 8 \text{ GeV}^2/c^4$, where the signal is unobserved and the selection is opti-
 946 mised to observe it. The upper bound of this interval is chosen to be sufficiently
 947 far from the J/ψ radiative tail at low masses and reduce its contamination into
 948 the rare sample;
- 949 • $11 < q^2 < 12.5 \text{ GeV}^2/c^4$, between two charmonium resonances, and
- 950 • $q^2 > 15 \text{ GeV}^2/c^4$, above $\psi(2S)$.

951 The first interval, below the J/ψ resonance, is referred to as “low- q^2 ” region, while
 952 the two intervals above the J/ψ resonance ($q^2 > 11 \text{ GeV}^2/c^4$) are referred to as
 953 “high- q^2 ” regions. These regions are then sub-divided into smaller intervals, as the
 954 available statistics allows, which results in $\sim 2 \text{ GeV}^2/c^4$ wide bins. The binning used
 955 is the following:

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

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

- 957 • $1.1\text{--}6.0 \text{ GeV}^2/c^4$: this interval is theoretically favoured to observe new physics
958 effects since it is far from the photon pole, which dominates at low q^2 values,
959 reducing the sensitivity to new physics contributions. The lower bound of
960 this interval is chosen to exclude the possible contribution from the ϕ
961 resonance, which appears at $\sim 1 \text{ GeV}^2/c^4$. The upper bound of the interval
962 is chosen to exclude completely a small contribution from the J/ψ resonance
963 that leaks below $8 \text{ GeV}^2/c^4$.
- 964 • $15.0\text{--}20.0 \text{ GeV}^2/c^4$: this interval is the one that is expected to contain most
965 of the rare decays and it is used as a natural cross check that the analysis is
966 stable when performed in smaller bins.

967 3.2 Candidate types

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

- 973 • **Downstream candidates**: built from tracks without hits in the VeLo, “down-
974 stream tracks”, also denoted as “DD”.
- 975 • **Long candidates**: built from tracks which also have hits in the VeLo, “long
976 tracks”. These candidates, also denoted as “LL”, are characterised by a better
977 momentum resolution than the downstream ones thanks to the longer lever
978 arm available to their tracks.

979 Figure 3.1 shows the two types of candidates used in the analysis, together with
980 other possible track types in LHCb, which are not used in this analysis. As the long

and downstream candidate categories are characterised by different resolutions and kinematic properties, the analysis is performed separately on the two samples and the results are then combined.

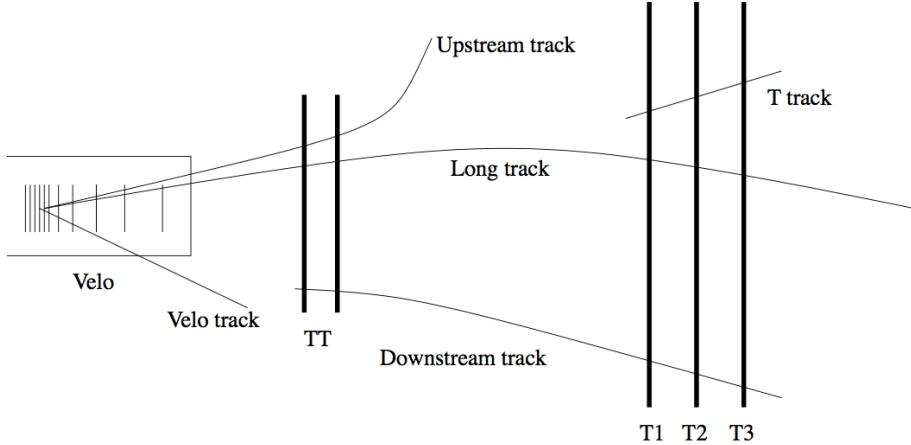


Figure 3.1: Representation of possible track types in LHCb. Candidates built from “long” and “downstream” tracks are used in this analysis [62].

3.3 Simulation

Samples of simulated events are needed in order to train a multivariate classifier, calculate the selection efficiency and study possible backgrounds; in particular for this analysis samples of $\sim 2 \cdot 10^6 \Lambda_b^0 \rightarrow J/\psi \Lambda$ and $\sim 5 \cdot 10^6 \Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ simulated decays are used. Samples of simulated $B^0 \rightarrow J/\psi K_s^0$, $B^0 \rightarrow K_s^0\mu^+\mu^-$ and $B^+ \rightarrow \mu^+\mu^-K^{*+}$ decays are also used to study their contribution to the background. The events are generated using PYTHIA8; hadronic particles are decayed using EVTGEN and GEANT4 is used to simulate the interaction of final state particles with the detector. Simulated events are then reconstructed by the same reconstruction software that is used for real data. The L0 hardware trigger is emulated in the simulation, while for the software stage, HLT, the same code can be used as for data. Events are simulated using both 2011 and 2012 beam and detector conditions, in the same proportion as in recorded data. While the simulation gives a generally good description of data, some discrepancies remain. It is important that the simulation gives an accurate description of the data, in particular for the extraction of efficiencies. The next

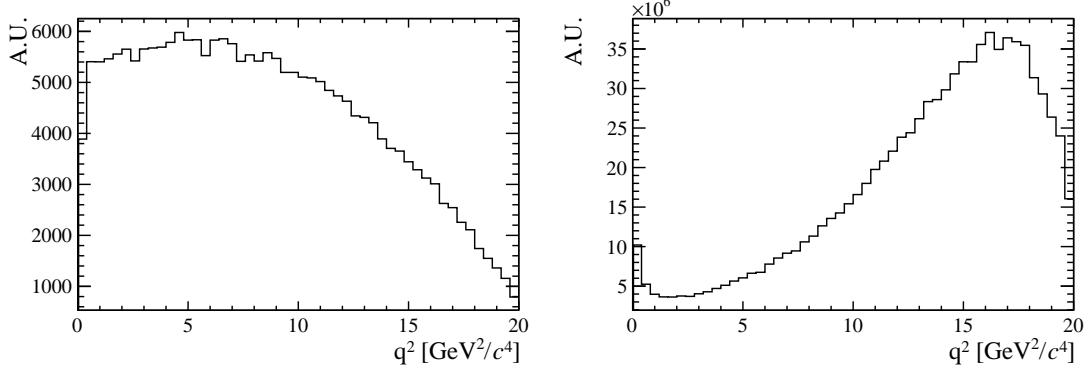


Figure 3.2: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ simulated events according to the phase space of the decay (left) and re-weighted using the decay amplitudes (right).

999 sections therefore describe corrections applied to the simulation in order to provide
1000 a better description of data. In Appendix B data distributions are compared with
1001 simulated ones for variables relevant to this analysis.

1002 3.3.1 Decay Model

1003 Little is known about the decay structure of Λ_b^0 decays and therefore the simulation
1004 software generates events according to the phase space given by the available kine-
1005 matics. To obtain a more realistic q^2 dependence, the simulation is weighted using
1006 decay amplitudes based on the predictions in Ref. [101]. Equations in this paper are
1007 for the case of unpolarised Λ_b^0 production and for this analysis they are extended to
1008 include polarisation. Details about the models used are given in Appendix A.1. The
1009 value of the Λ_b^0 production polarisation, P_b , used in the calculations is $P_b = 0.06$ as
1010 measured by LHCb [102]. Figure 3.2 shows the phase space q^2 distribution and the
1011 one obtained by re-weighting the events. The latter can be qualitatively compared
1012 to the q^2 spectrum of a generic $b \rightarrow s\ell^+\ell^-$ decay shown in Fig. 1.8. For the normal-
1013 isation mode, the decay model used is described in Appendix A.3, with amplitude
1014 magnitudes and production polarisation taken from the measurements in Ref. [102].
1015 Phases are not yet measured and are therefore set to zero.

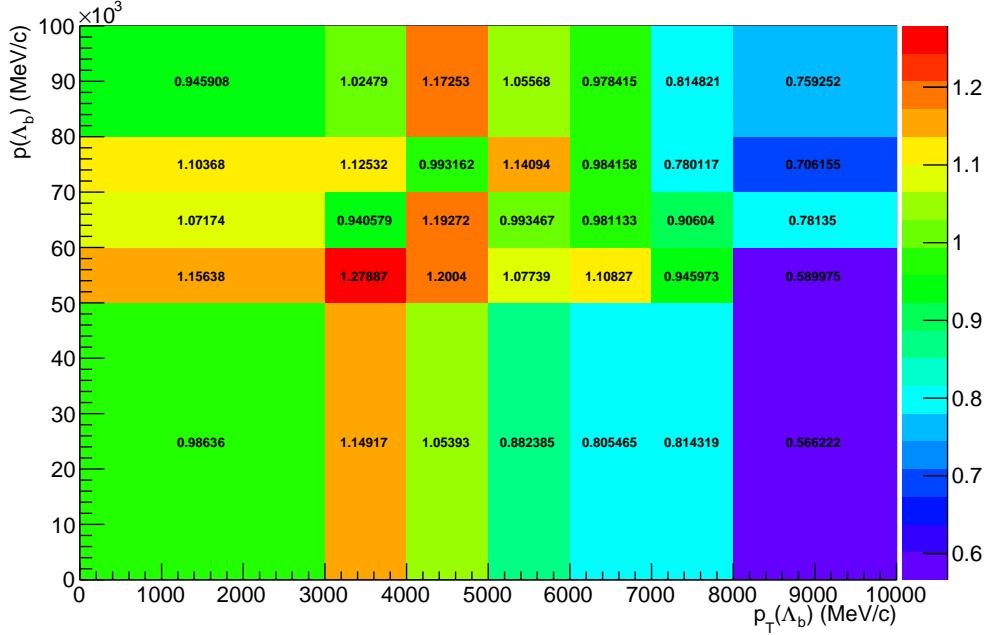
1016 3.3.2 Kinematic re-weighting

1017 Small data-simulation differences are found in the kinematic properties of the mother
1018 particle, Λ_b^0 , which also affect the final state particles. The simulation is re-weighted
1019 by comparing the momentum and transverse momentum of Λ_b^0 baryons in real and
1020 simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates that satisfy the pre-selection requirements (see
1021 Sec. 3.4). To do this a high purity data sample is obtained by selecting a nar-
1022 row invariant mass interval around the J/ψ and Λ_b^0 peaks; this contains about
1023 $4 \cdot 10^5$ candidates. The Λ_b^0 invariant mass distribution is then fitted to estimate
1024 the number of background decays under the peak. Finally, the background fraction,
1025 $f_b = B/(S + B)$, is used to statistically subtract the background from the kinemat-
1026 ical distributions as described by the equation:

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

1027 where $S(p, p_T)$ is the distribution of pure signal candidates, which we want to obtain,
1028 $T(p, p_T)$ is the total distribution of signal plus background, namely the distribution
1029 of all events in the signal interval, $5605 < m(p\pi\mu^+\mu^-) < 5635$ MeV/ c^2 , and $B(p, p_T)$
1030 is the pure background distribution obtained using candidates from the upper side-
1031 band, $m(p\pi\mu^+\mu^-) > 5800$ MeV/ c^2 .

1032 After the signal distributions have been obtained from data, they are compared
1033 with $\Lambda_b^0 \rightarrow J/\psi \Lambda$ simulated events and a weight, $w(p_{\Lambda_b^0}, p_{T\Lambda_b^0})$ is defined by taking
1034 the ratio of the two dimensional (p, p_T) distributions. The result is shown in Fig. 3.3,
1035 while Appendix B reports distributions of sideband subtracted data in the signal and
1036 sideband regions together with weighted and unweighted simulated events. In these
1037 plots the momentum and p_T distributions of Λ_b^0 baryons match by construction but
1038 the re-weighting also improves the agreement between the kinematical distributions
1039 of all final particles. Small differences remain due to the finite binning used for the
1040 weights calculation. Quality variables, such as the χ^2 of tracks and vertices, show
1041 little dependence on the kinematics and are relatively unaffected by the weighting
1042 procedure.



1054 3.4 Selection

1055 This section describes the requirements applied to reconstruct $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and
1056 $\Lambda_b^0 \rightarrow J/\psi\Lambda$ candidates. The selection procedure is divided into two steps: a
1057 pre-selection, where cuts are applied in order to be able to work with manage-
1058 able datasets and a multivariate analysis (MVA) which combines information from
1059 several variables. As a first step good quality tracks are selected by imposing re-
1060 quirements on their basic kinematic properties, such as the p_T of the final particles,
1061 and quality requirements, such as the track χ^2 . The selection then forms a dimuon
1062 candidate from two oppositely charged muons. In events containing a dimuon can-
1063 didate, two oppositely charged tracks are combined and retained as a Λ candidate
1064 if they form a good quality vertex which is well separated from all primary ver-
1065 tices. Finally, the dimuon and Λ candidates are combined to form Λ_b^0 baryons with
1066 requirements placed on the properties of this combination.

1067 3.4.1 Pre-selection

1068 The full list of pre-selection cuts is reported in Tab. 3.1. In the table χ_{IP}^2 is defined
1069 as the projected distance from a vertex divided by its uncertainty, for example the
1070 $\chi_{IP}^2(primary) > n$ requirement on Λ_b^0 means that the Λ_b^0 vertex must be at least
1071 \sqrt{n} standard deviations away from the primary vertex. Another quantity, found
1072 to be particularly powerful at removing combinatorial background, is a pointing
1073 variable called DIRA defined as the cosine of the angle between the direction of a
1074 particle's momentum and the flight direction from its mother vertex. Requiring a
1075 DIRA close to unity corresponds to the selection of particles with well-defined origin
1076 vertices. Figure 3.4 shows graphical representations of the χ_{IP}^2 and DIRA variables.
1077 The variable χ_{FD}^2 represents the flight distance of a particle from its origin vertex
1078 divided by the corresponding uncertainty. The χ_{trk}^2/ndf and χ_{vtx}^2/ndf quantities
1079 are the χ^2 from the fit to tracks and vertices, which are used to quantify their
1080 quality. The `GhostProb` quantity describes the probability of a track being fake. By

construction, cutting at a value of k , removes $(1 - k) \cdot 100\%$ of fake tracks. The `hasRich`, `hasCalo` and `isMuon` variables are binary indicators that the information from the RICH, calorimeter and muon detectors is available for the track. Loose PID requirements on the proton are also applied in the pre-selection. Details about PID quality estimators are given in Sec. 2.8. A large mass window around the Λ_b^0 peak is used to allow a fit to the sideband to be performed and to use sideband candidates to train a multivariate classifier. Rare candidates are selected by the q^2 region requirements described in Sec. 3.1, while resonant candidates are further constrained to have dimuon invariant masses in a $100 \text{ MeV}/c^2$ interval around the known J/ψ mass [2].

3.4.2 Neural Networks

The final selection is performed using a neural network classifier based on the NeuroBayes package. The input to the neural network consists of 14 variables carrying information about the kinematics of the decay, the quality of tracks and vertices and the PID of the muons. The list of the 10 most significant inputs is reported in Tab. 3.2, together with information about the importance of each input. Variables related to Λ and its daughters are considered as different inputs depending on the candidate type (long or downstream). This effectively corresponds to making a separate training for the two categories.

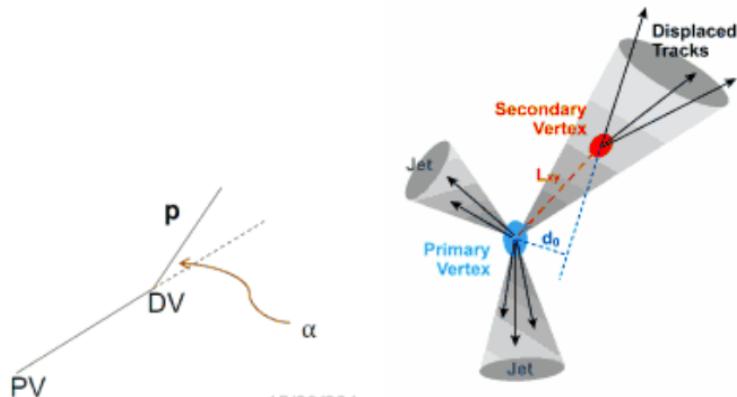


Figure 3.4: Graphical representation of the DIRA (left) and χ_{IP}^2 (right) variables.

Particle	Requirement
Λ_b^0	$4.6 < m(p\pi\mu\mu) < 7.0 \text{ GeV}/c^2$ $\text{DIRA} > 0.9999$ $\chi_{\text{IP}}^2 < 16.0$ $\chi_{\text{FD}}^2 > 121.0$ $\chi_{\text{vtx}}^2/ndf < 8.0$
Λ	$\chi_{\text{vtx}}^2/ndf < 30.0(25.0)$ $\text{Decay time} > 2 \text{ ps}$ $ m(p\pi) - m_A^{\text{PDG}} < 35(64) \text{ GeV}/c$
p/π	$p > 2 \text{ GeV}/c$ $p_T > 250 \text{ MeV}/c$ $\chi_{\text{IP}}^2 > 9(4)$
p (only long cand.)	hasRICH $\text{PID}p > -5$
μ	isMuon $\chi_{\text{trk}}^2/ndf < 5$ $\text{GhostProb} < 0.4$ $\text{PID}\mu > -3$ $\chi_{\text{IP}}^2 > 9.0$
Dimuon	$\chi_{\text{vtx}}^2/ndf < 12.0$ $ m(\mu\mu) - m_{J/\psi}^{\text{PDG}} < 100 \text{ MeV}/c^2 \quad (J/\psi \Lambda \text{ only})$

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

1100 The neural network is trained using representative samples of signal and background.
1101 A sample of simulated $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates is used as a proxy for the signal,
1102 while for the background a representative sample is given by candidates in the upper
1103 $m(p\pi\mu\mu)$ invariant mass sideband. Only the upper sideband, $m(p\pi\mu\mu) > 6 \text{ GeV}/c^2$,
1104 is used since it contains only combinatorial background, while the lower sideband
1105 may contain partially reconstructed and misreconstructed candidates. In the q^2
1106 spectrum of background samples the J/ψ and $\psi(2S)$ peaks are still present indicating
1107 that charmonium resonances are often combined with other random tracks. These
1108 candidates do not give a good description of purely combinatorial background and,
1109 in order to avoid biases, they are removed from the training sample by rejecting
1110 candidates in a $100 \text{ MeV}/c^2$ interval around the nominal J/ψ and $\psi(2S)$ masses [2].
1111 A total of $3 \cdot 10^4$ events is used for the training from each sample. This corresponds
1112 to approximately $\simeq 50\%$ of the available sideband data and $\simeq 20\%$ of the simulated

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

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

sample. The full simulated sample is not used in the training as the same sample will also be used to study efficiencies. For reproducibility the events are sampled uniformly.

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

Figure 3.5 shows distributions of neural network output for the signal and background samples and purity, $P = N_{\text{sig}}/N_{\text{bkg}}$, as a function of the neural network output. To check for potential overtraining, the distributions from test samples are also overlaid. These are found to follow the same shape but with different fluctuations giving no significant evidence of overtraining. In general it can be concluded that the neural network is able to separate signal from background and the training converged properly. It can happen that too much information is given

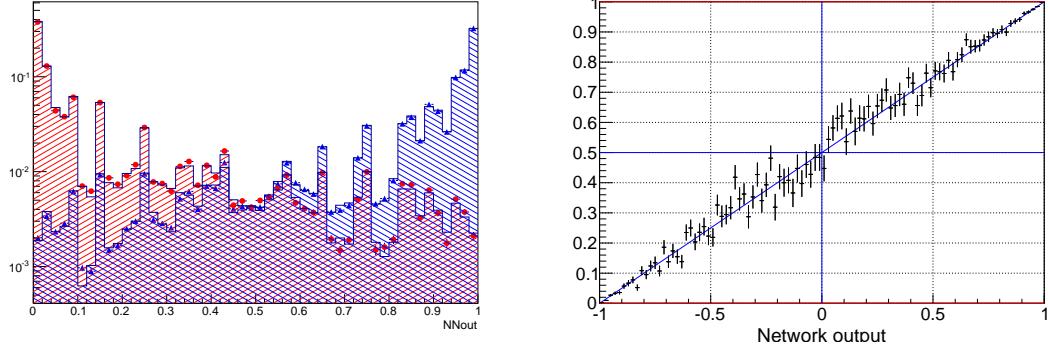


Figure 3.5: (left) Neural network output distribution for training (points) and test (stripes) samples, for signal (blue) and background (red) candidates. (right) Purity as a function of neural network output.

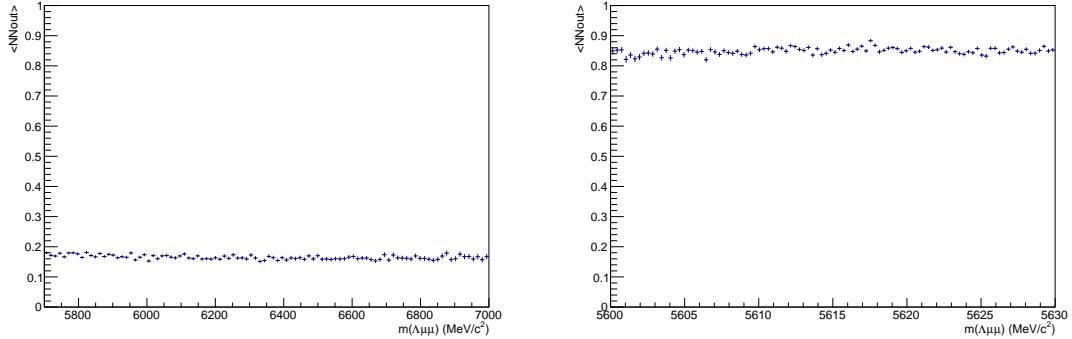


Figure 3.6: Average value of neural network output as a function of the 4-body invariant mass for data sideband (left) and simulated signal (right) candidates.

to the classifier, allowing the invariant mass of the candidates to be inferred from the input variables. This can generate fake peaks and it is therefore important to check for correlations between the 4-body invariant mass and the neural network output. Figure 3.6 reports the average neural network output as a function of the 4-body $m(p\pi\mu\mu)$ invariant mass for data and simulation. The distributions are flat indicating that no significant correlation is present.

3.4.3 MVA optimisation

In the high- q^2 region, where the signal is already observed, the requirement on the neural network output is chosen to maximise the significance, $N_S/\sqrt{N_S + N_B}$, where N_S and N_B are the numbers of expected signal and background candidates

respectively. N_S is derived from simulation but, as an arbitrary number of events can be generated, it needs to be normalised. To do this, the invariant mass distribution of real $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates is fit after pre-selection (including all requirements except the MVA selection). This is possible as the peak of the resonant channel is already clearly visible before the MVA requirement. The resonant yield is then scaled by the ratio of between the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ branching fractions as measured by LHCb on 2011 data,

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

and by the $J/\psi \rightarrow \mu^+\mu^-$ branching fraction, *i.e.*

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

The number of expected background candidates is derived by fitting the data side-band with an exponential and extrapolating into the signal region.

In the low- q^2 region, where the signal is unobserved, the so called ‘‘Punzi figure-of-merit’’, $N_S/(n_\sigma/2 + \sqrt{N_B})$, is maximised [103]. This figure-of-merit is considered to be optimal for discovery and the parameter n_σ corresponds to the number of expected standard deviations of significance, in this analysis $n_\sigma = 3$ is used. Moreover, the Punzi shape does not depend on the relative normalisation between signal and background, which is important since the signal is still unobserved at low- q^2 and the existing predictions vary significantly for this region. The dependence of the figure-of-merit for both q^2 regions is shown in Fig. 3.7, and curves of signal efficiency versus background rejection are shown in Fig. 3.8.

For the 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 97 % (82 %) of long candidates and 96% (66 %) of downstream candidates for the high- (low-) q^2 selection, with respect to the pre-selected samples. After full selection $\sim 0.5\%$ of the events con-

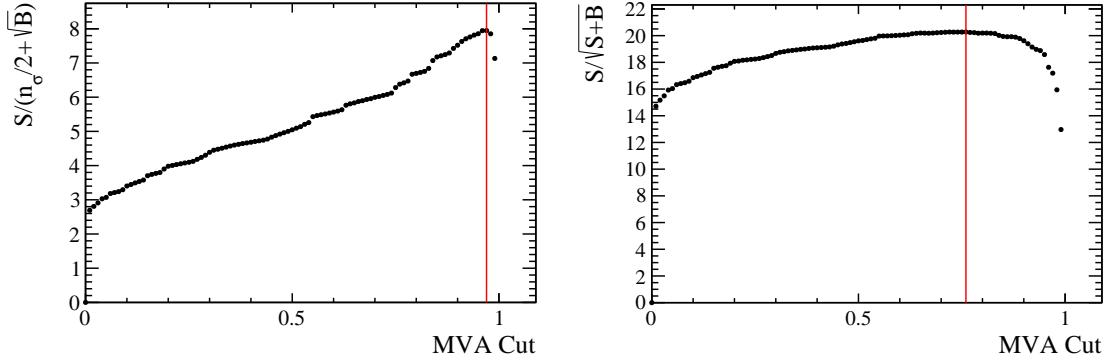


Figure 3.7: Dependence of the figure-of-merits on the neural network output requirement for the low- q^2 (left) and high- q^2 (right) regions. The vertical lines correspond to the chosen cuts.

tain multiple candidates which are randomly rejected keeping only one candidate per event. To normalise the branching ratio measurement $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates are selected using both, low- and high- q^2 , requirements to normalise respectively low and high- q^2 intervals.

3.4.4 Trigger

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

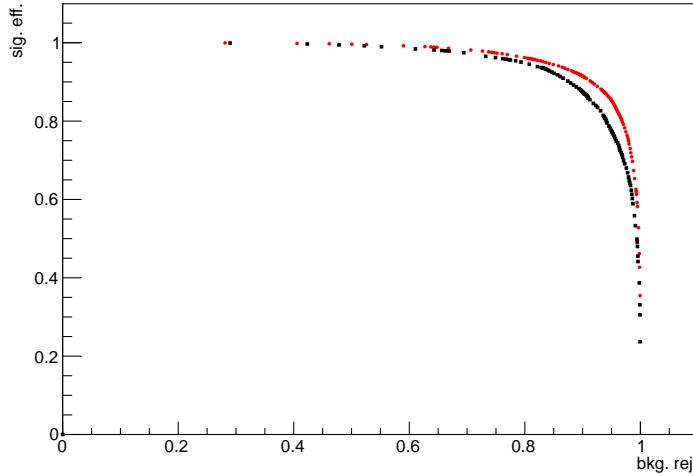


Figure 3.8: Receiver operating characteristic (ROC) curves for low- q^2 (black) and high- q^2 (red). They show the signal efficiency versus the background rejection. The optimal points on these curves are the closest ones to (1,1).

required to be less than 0.2 mm to form a 2-body object.

Table 3.3: Summary of the trigger lines used to select events at various levels. Trigger is always required to be due to the 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

1182

1183 3.4.5 Background from specific decays

1184 Candidates from other decays can be reconstructed as the decays of interest if par-
 1185 ticles are not reconstructed or mis-identified. A survey of possible backgrounds
 1186 concluded that the only physics background to take into account comes from mis-
 1187 reconstructed decays of B^0 to K_s^0 with two muons in the final state, whether via
 1188 J/ψ or not, where the K_s^0 is reconstructed as a Λ with a $p \rightarrow \pi$ identity swap.

1189 The lack of background from other decays is mainly due to the particular topol-
 1190 ogy of the Λ decay, which is long-lived and decays at a displaced vertex. To study
 1191 the effect of misreconstructed $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow K_s^0\mu^+\mu^-$ decays simulated
 1192 samples are used. On data the $B^0 \rightarrow J/\psi K_s^0$ contribution is clearly visible in the
 1193 resonant channel mass distribution. This background is not suppressed with specific
 1194 cuts in this analysis as its mass shape is sufficiently distinct from the Λ_b^0 signal and
 1195 its contribution can be reliably modelled in the mass fits (see Sec. 3.5.1). For the
 1196 rare case a rough estimate of the K_s^0 background size is obtained using the yield in
 1197 the resonant channel rescaled by the measured ratio between the rare and resonant
 1198 branching fractions. Details are given in Sec. 3.5.1 and numbers of events predicted
 1199 are reported in Tab. 3.4. This contribution, although close to negligible is again
 1200 considered in the fit. A possible pollution due to $B^+ \rightarrow \mu^+\mu^-K^{*+}$ decays, where
 1201 the K^{*+} further decays into $K_s^0\pi$ is also investigated using a dedicated simulated
 1202 sample and found to be negligible. Finally, $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events radiating photons
 1203 from the final state, can escape the J/ψ veto and be reconstructed in the rare chan-
 1204 nel sample. Analysing simulated events it was found that the only contribution is
 1205 in the closest q^2 interval to the J/ψ tail, $6 < q^2 < 8$ GeV $^2/c^4$. In this interval
 1206 1.3% of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates are reconstructed but only 0.06% fall into the
 1207 4-body invariant mass window used for the fits. This corresponds to ~ 6 events,
 1208 4 of which in the downstream category. Given the low yield and that these events
 1209 do not peak under the signal but show a decaying distribution at the edge of the
 1210 fit mass window, this background is considered as absorbed in the combinatorial
 1211 background. Figure 3.9 shows the 4-body invariant mass distribution of simulated
 1212 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events falling into the rare q^2 region and the distribution of simulated
 1213 $B^+ \rightarrow \mu^+\mu^-K^{*+}$ events mis-reconstructed as $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays.

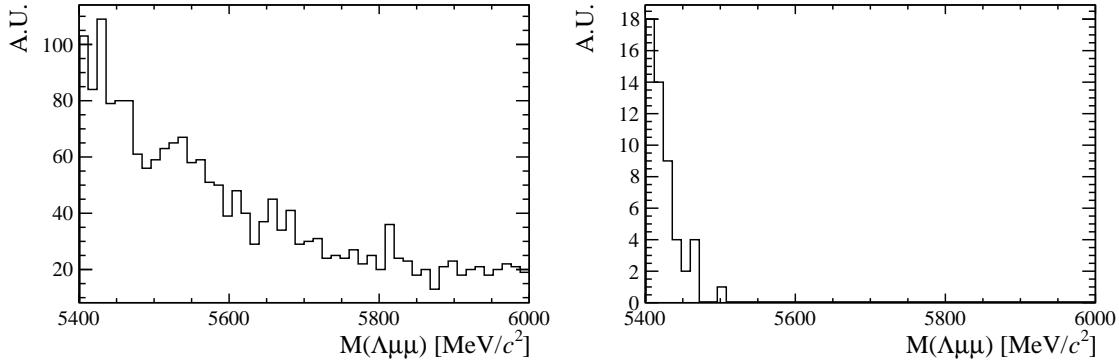


Figure 3.9: Invariant mass distributions of simulated $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ (left) and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ (right) candidates passing the full selection. Only $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed in $q^2 < 8 \text{ GeV}^2/c^4$ are selected.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

1214 3.5 Yield extraction

1215 Extended unbinned maximum likelihood fits are used to extract the yields of the
1216 rare and resonant channels. The likelihood has the form:

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

1217 where N_S , N_C and N_B are respectively the numbers of signal, combinatorial and
1218 K_s^0 background events and the $P_i(m_i)$ are the corresponding probability density
1219 functions (PDF). The fit variable is the 4-body $m(p\pi\mu\mu)$ invariant mass obtained
1220 from a kinematical fit of the full decay chain in which each particle is constrained
1221 to point to its assigned origin vertex and the invariant mass of the $p\pi$ system is
1222 constrained to be equal to the world average for the Λ baryon mass. In the resonant
1223 case a further constrain is used on the dimuon mass to be equal to the known J/ψ
1224 mass. This method allows to improve the mass resolution giving better defined
1225 peaks and therefore a more stable fit. For brevity, in the following these variables
1226 are simply referred to as ‘‘invariant mass’’.

₁₂₂₇ 3.5.1 Fit description

₁₂₂₈ The fit is performed through the following steps:

- ₁₂₂₉ • simulated distributions are fit to extract initial parameters;
- ₁₂₃₀ • the resonant data sample is fitted;
- ₁₂₃₁ • the rare sample is fitted fixing some parameters to those obtained in the pre-
- ₁₂₃₂ vious cases.

₁₂₃₃ In the first step simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distributions are fitted using the signal PDF
₁₂₃₄ alone. This is done separately for downstream and long candidates. Figure 3.10
₁₂₃₅ shows distributions of candidates selected in the resonant sample with the fit function
₁₂₃₆ overlaid. The signal is described as the sum of two Crystal Ball functions (CB)
₁₂₃₇ with common mean (m_0) and tail slope (n). This is also known as Double Crystal
₁₂₃₈ Ball (DCB) function. A single Crystal Ball [104] is a probability density function
₁₂₃₉ commonly used to model processes involving energy loss. In particular it is used to
₁₂₄₀ describe resonances' peaks with radiative tails. This function consists of a Gaussian
₁₂₄₁ core and a power-law tail below a certain threshold and has form

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

₁₂₄₂ where for normalisation and continuity

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

₁₂₄₃ The full PDF for the resonant channel is therefore:

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

₁₂₄₄ where f is the relative fraction of candidates falling into the first CB function.

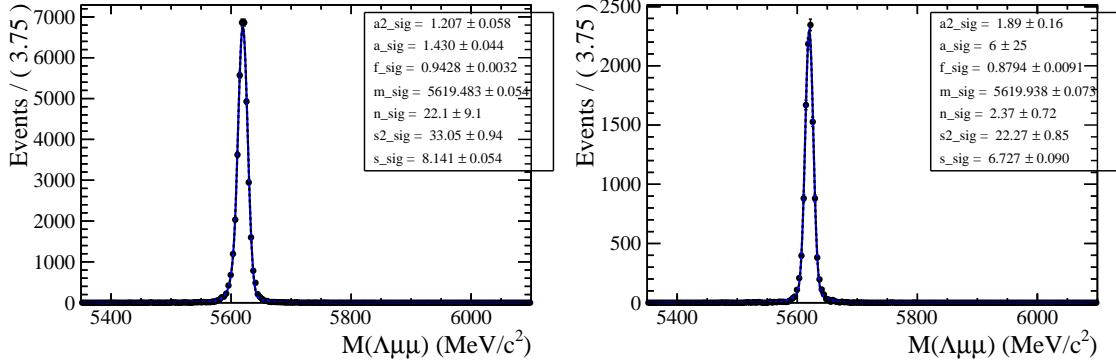


Figure 3.10: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ downstream (left) long (right) candidates. The points show simulated data and the blue line is the signal fit function.

1245 In a second step the fit to the resonant channel data sample is performed. For this fit
 1246 the tail slope parameter, “ n ”, which is highly correlated with α_1 and α_2 , is fixed to
 1247 the value found in the fit to simulated data. In this fit two background components
 1248 are modelled: the combinatorial background, parameterized with an exponential
 1249 and the background from $B^0 \rightarrow J/\psi K_s^0$ decays. The shape used to describe the
 1250 K_s^0 background is obtained from a $B^0 \rightarrow J/\psi K_s^0$ simulated sample to which the
 1251 full selection is applied. The invariant distribution of these events is fit with a DCB
 1252 function, which is then used to model the K_s^0 background in the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ fit. The
 1253 fit to the simulated $B^0 \rightarrow J/\psi K_s^0$ events is reported in Fig. 3.11. When the K_s^0 shape
 1254 is introduced in the fit to the data all its parameters are fixed. This is particularly
 1255 important when fitting long candidates, where the K_s^0 peak is less evident, which
 1256 does not allow to constrain many parameters. On the other hand, in order to take
 1257 into account possible data-simulation differences, an horizontal shift is added and
 1258 left floating (by adding a constant to the central value of the DCB, $m_0 \rightarrow m_0 + m'$).
 1259 In summary, the free parameters in the fit to the resonant $\Lambda_b^0 \rightarrow J/\psi \Lambda$ sample are
 1260 the yields of the signal and the combinatorial and K_s^0 backgrounds, the slope of the
 1261 exponential and the horizontal shift of the K_s^0 shape. Note that all parameters of
 1262 the fit to the long and downstream samples are independent.
 1263 Finally, the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ data sample is fit. In this case the fit to the long
 1264 and downstream samples is performed simultaneously to obtain a more stable con-

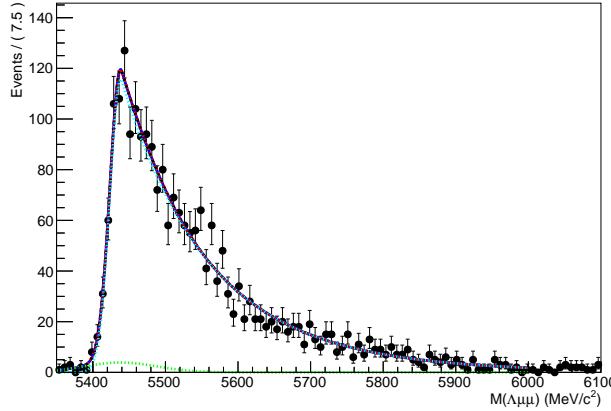


Figure 3.11: Invariant mass distribution of simulated $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted a Double Crystal Ball function.

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

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

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

Table 3.4: Predicted numbers of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ 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

1282 shape parameters are never shared between the two candidate categories.

1283 Also in the rare case the modelled background components are the combinatorial
1284 background, described with an exponential function and the K_s^0 background. The
1285 slope of the background is visibly different depending on the q^2 interval. This is
1286 partly due to the fact that at high q^2 the combinatorial changes slope because of
1287 a kinematical limit at low 4-body masses imposed by the q^2 requirements. The
1288 exponential slopes are therefore left as independent parameters in each q^2 interval.
1289 The background component from $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays is modelled using the same
1290 shapes used for the resonant channel. However, in this case the horizontal shift is
1291 fixed to what found for the resonant channel. The expected amount of misrecon-
1292 structed $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events is small and does not allow to determine reliably the
1293 yield. Therefore this is fixed to the yield of $B^0 \rightarrow J/\psi K_s^0$ decays rescaled by the
1294 expected ratio of branching fractions between the resonant and rare channels. The
1295 q^2 distribution of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ simulated events is used to predict the yield as a
1296 function of q^2 . Table 3.4 reports the number of predicted $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in
1297 each q^2 interval obtained with the following formula:

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

1298 where $N(q^2)_{MC}$ is the number of simulated rare candidates falling in a q^2 interval
1299 after full selection and N_{MC}^{tot} is the total number of simulated events.

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

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

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

3.5.2 Fit results

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

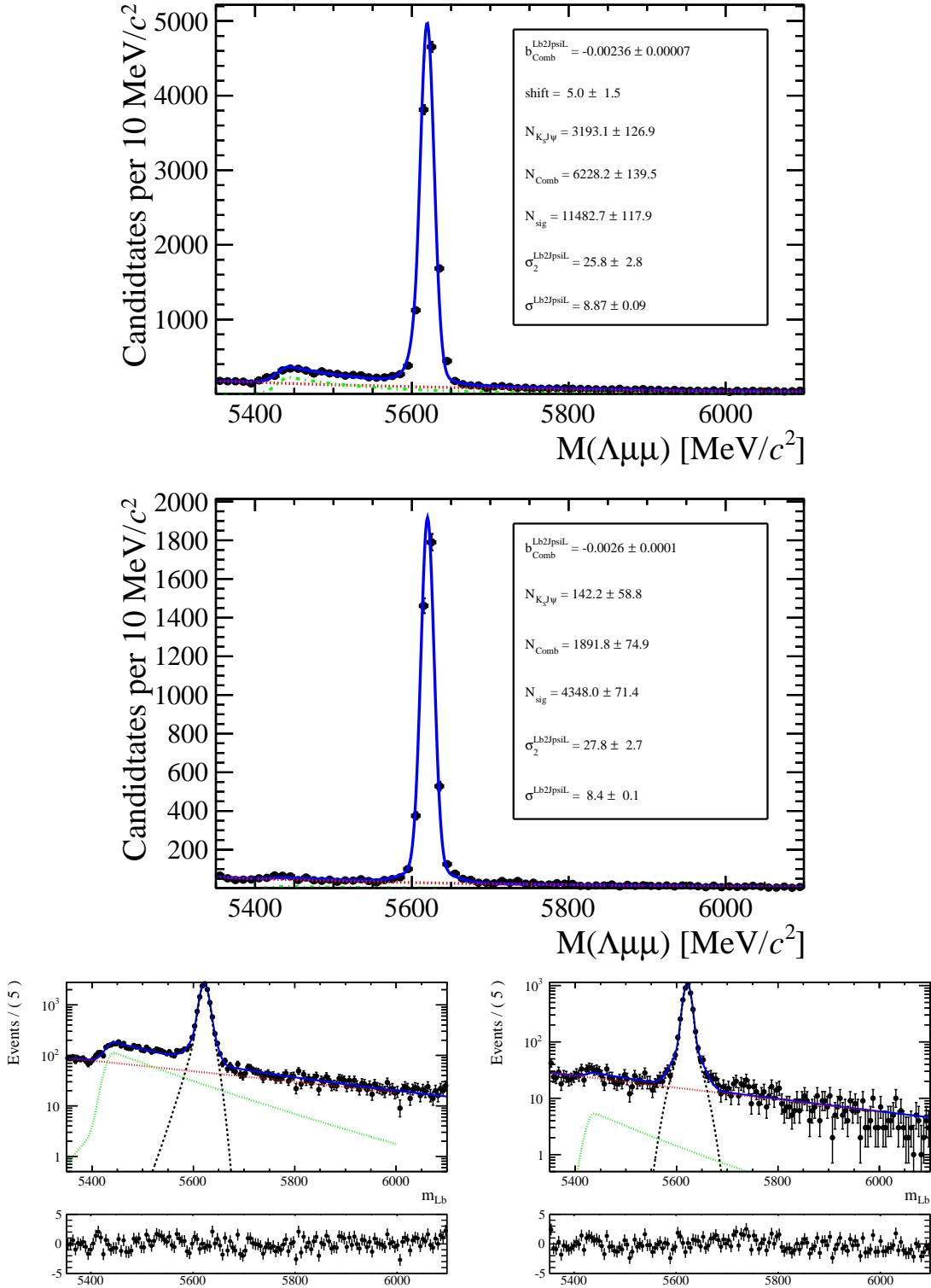


Figure 3.12: Invariant mass distributions of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ downstream (top) and long (middle) candidates selected with high q^2 requirements. Bottom plots are the same as the upper ones but shown in logarithmic scale. Black points show data. The blue solid line represents the 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 downstream and long categories found using the for low- and high- q^2 requirements; uncertainties are statistical only.

Selection	Long	Downstream
high- q^2	4313 ± 70	11497 ± 123
low- q^2	3363 ± 59	7225 ± 89

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

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

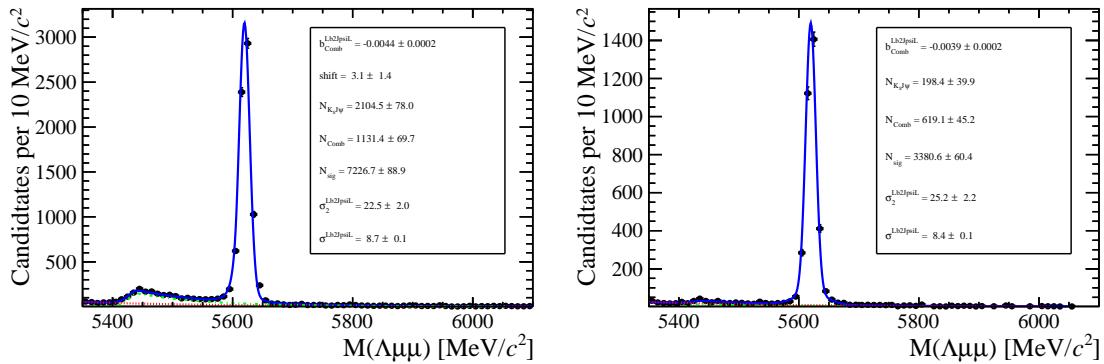


Figure 3.13: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ for downstream (left) and long (right) candidates selected with low q^2 requirements.

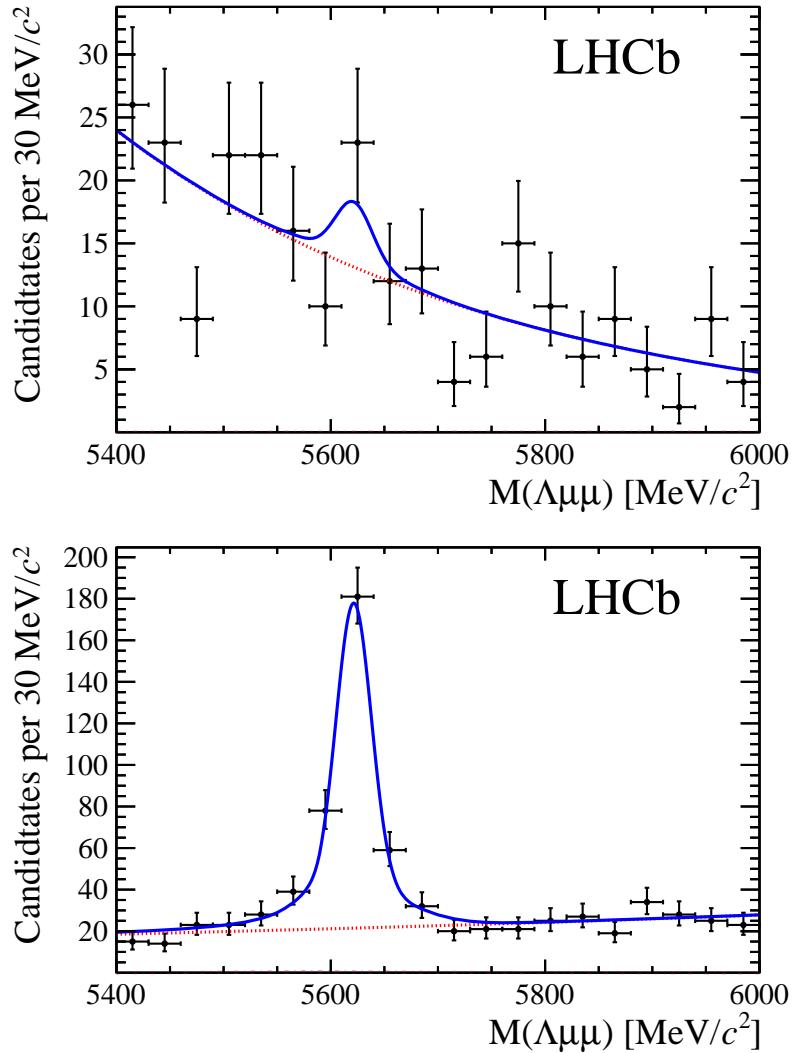


Figure 3.14: Invariant mass distributions of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the integrated $0.1\text{--}6.0 \text{ GeV}^2/c^4$ (top) and $15\text{--}20 \text{ GeV}^2/c^4$ (bottom) q^2 intervals. Points show data combining long and downstream candidates together. The blue solid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

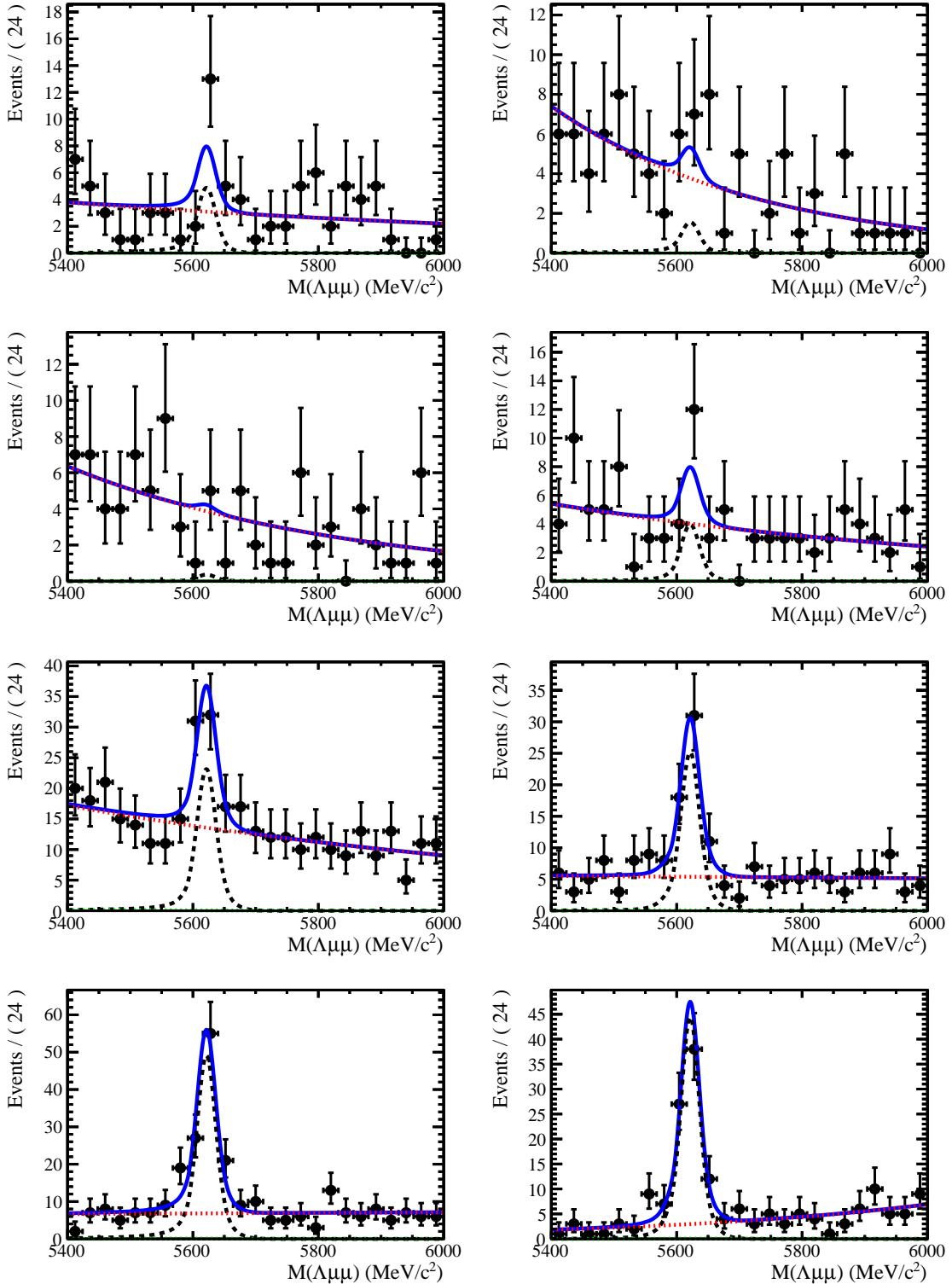


Figure 3.15: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for downstream candidates.

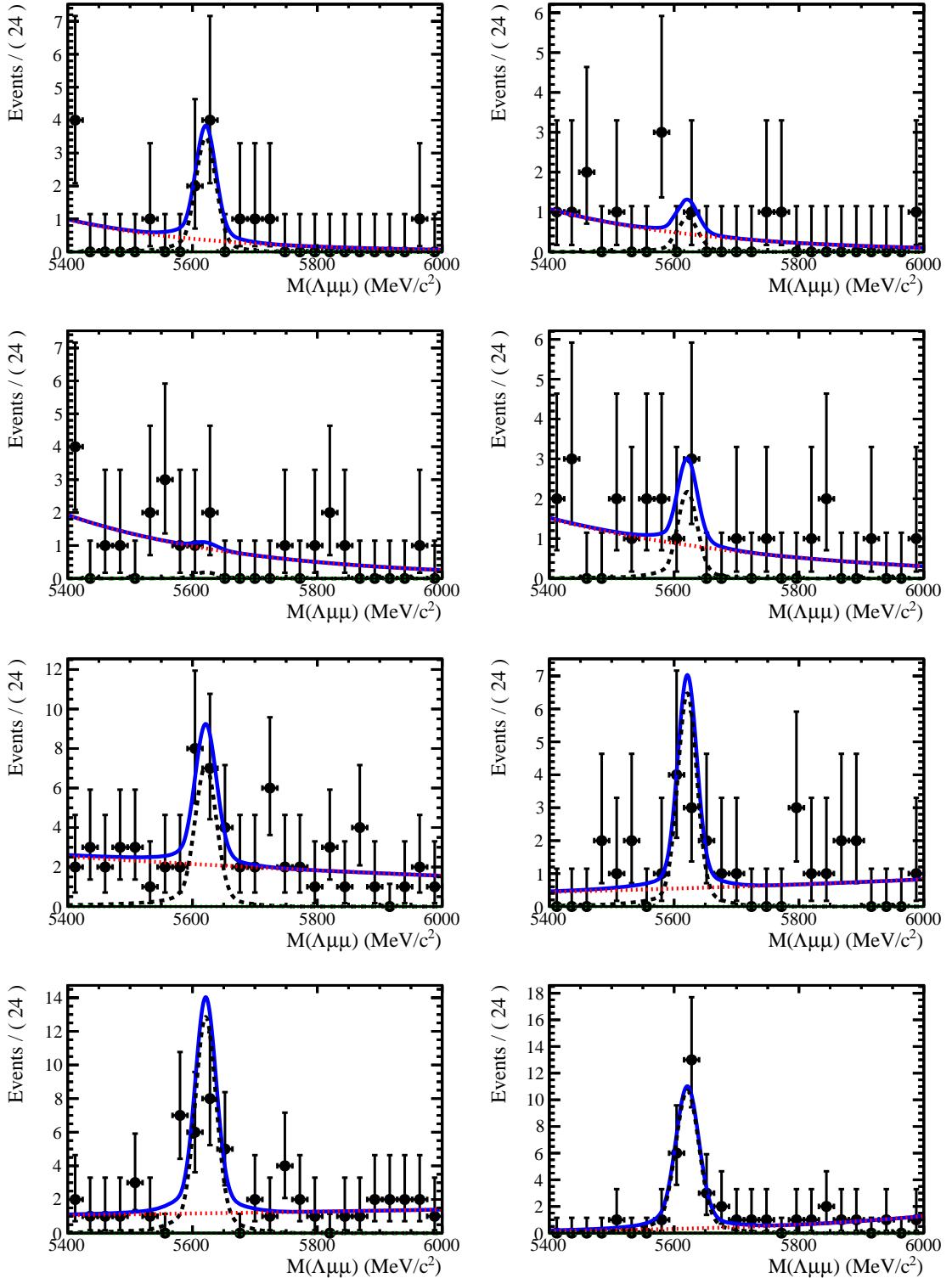


Figure 3.16: Invariant mass distributions of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the considered q^2 bins for long candidates.

₁₃₂₅ **3.6 Efficiency**

₁₃₂₆ The selection efficiency is calculated for each decay according to the formula

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

₁₃₂₇ In this expression the first term gives the efficiency to have final state particles
₁₃₂₈ in the LHCb acceptance. The second term handles the possibility of Λ escaping
₁₃₂₉ the detector or interacting with it and therefore never decaying into $p\pi$; this term
₁₃₃₀ is referred to as “detection” efficiency. The third term carries information about
₁₃₃₁ the reconstruction and pre-selection efficiencies, which are kept together given that
₁₃₃₂ boundaries between them are completely artificial. The fourth part deals with the
₁₃₃₃ efficiency of the Neural Network for those events which passed the pre-selection.
₁₃₃₄ Finally, the last term handles the trigger efficiency for events which are accepted
₁₃₃₅ by the full selection. Most of the efficiency components are evaluated using the
₁₃₃₆ simulated samples described in Sec. 3.3. Only the efficiency of the PID requirement
₁₃₃₇ for the proton (see Tab. 3.1) is separately derived with a data–driven method because
₁₃₃₈ the simulation does not provide a good description of PID variables. For complete
₁₃₃₉ information, all absolute efficiencies for the two decays $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$
₁₃₄₀ are separately listed in the next subsections. However, for the analysis itself only
₁₃₄₁ the relative efficiency, $\varepsilon(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)/\varepsilon(\Lambda_b^0 \rightarrow J/\psi\Lambda)$, is used.

₁₃₄₂ **3.6.1 Geometric acceptance**

₁₃₄₃ In order to save disk space and time only events are simulated in which the final
₁₃₄₄ muons are in the detector acceptance and therefore can be reconstructed. This corre-
₁₃₄₅ sponds to a requirement for each of the muons to be in an interval $10 < \theta < 400$ mrad,
₁₃₄₆ where θ is the angle between the muon momentum and the beam line. The efficiency
₁₃₄₇ of this requirement is obtained by using a separate simulated sample, where events
₁₃₄₈ are generated in the full space. The geometric efficiency varies between 18% at
₁₃₄₉ high- q^2 and 20% at low- q^2 ; Fig. 3.17 shows the dependence of this efficiency as a

1350 function of q^2 .

1351 3.6.2 Reconstruction and neural network efficiencies

1352 The efficiency to reconstruct the decays together with the pre-selection requirements
1353 is evaluated from simulated data. The reconstruction efficiency is subdivided in “De-
1354 tection” and “Reconstruction and pre-selection” efficiencies. In fact, since Λ is a long
1355 lived particle, there is a non-negligible probability that it interacts in the detector or
1356 escapes from it and therefore never decays in proton and pion. The reconstruction
1357 efficiency includes the efficiency of for the tracks to produce observable signatures
1358 and the efficiency for candidates to pass the pre-selection requirements. This compo-
1359 nent does not include the efficiency of the PID cut that appears in Tab. 3.1, which
1360 is kept separate because PID variables are not well described by the simulation.
1361 The detection efficiency varies between 88% at high- q^2 and 20% at low- q^2 while the
1362 reconstruction efficiency for downstream candidates is almost flat at 6.6% and for
1363 long candidates it varies from 1.6% at high- q^2 to 2.0% at low- q^2 . Fig. 3.17 shows the
1364 dependence of these efficiencies as a function of q^2 . The NN selection efficiency is
1365 again evaluated from simulated samples and it is observed to vary from 58% to 84%
1366 for downstream candidates and from 77% to 92% for log candidates. Fig. 3.17 shows
1367 the dependence of this efficiency as a function of q^2 . The sudden jump in efficiency
1368 at $\sim 9 \text{ GeV}/c^2$ is due to the fact that a different figure-of-merit is used to optimise
1369 the NN cut in the low and high q^2 regions, which results in different efficiencies.

1370 3.6.3 Trigger efficiency

1371 The trigger efficiency is again calculated using a simulated sample and it varies
1372 between 61% and 84% for downstream candidates and from 65% to 85% for long
1373 candidates. Fig. 3.17 shows the dependence of this efficiency as a function of q^2 .
1374 Using the resonant channel it is possible to crosscheck on data the efficiency obtained
1375 using the simulation. In LHCb triggered events can fall in two categories: events

1376 triggered by a track which is part of a signal candidate, Trigger On Signal (TOS),
1377 or by other tracks in the event, Trigger Independent of Signal (TIS). As the TIS and
1378 TOS categories are not exclusive the TIS sample provides a control sample which
1379 can be used to obtain the efficiency for TOS trigger. This is calculated with the
1380 formula:

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

1381 As data contains background the numbers of signal candidates in the “TIS” and
1382 “TIS && TOS” categories are not just determined by counting events but from a fit
1383 to the 4-body invariant mass, $m(p\pi\mu\mu)$. This procedure takes the name of TISTOS
1384 method. Using the data–driven method an efficiency of $(70 \pm 5)\%$ is obtained, while
1385 this is calculated to be $(73.33 \pm 0.02)\%$ using the simulation. Results are therefore
1386 compatible within 1σ .

1387 3.6.4 PID efficiency

1388 For long tracks a PID requirement on protons ($\text{PID}_p > -5$) is applied. The simula-
1389 tion is known not to describe particle ID variables well and therefore a data-driven
1390 method is used to obtain this efficiency component. This is done using the `PIDCalib`
1391 package (see Sec. 2.8.1), which uses as calibrations samples decays where particles
1392 can be identified due to their kinematic properties. In the case of protons a sample
1393 of Λ particles is used, where the proton can be identified because it always has the
1394 highest momentum. The package allows to divide the phase space in bins of variables
1395 relevant for PID performances; in this analysis momentum and pseudorapidity are
1396 used. Using the calibration sample the efficiency is derived in each two-dimensional
1397 bin. Finally, to take into account that the decay channel under study could have
1398 different kinematical distributions than the calibration sample these efficiency tables
1399 are used to re-weight the simulation. The PID efficiency varies from 97.3% at low- q^2
1400 to 98.2% at high- q^2 .

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

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

1401 3.6.5 Relative efficiencies

1402 In the previous sections absolute efficiencies values were given for the rare channel
 1403 in different q^2 intervals. Figure 3.17 contains a summary of those values in these
 1404 tables in graphical form. This section reports the corresponding relative efficiencies
 1405 with respect to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ channel, which will be used to correct the yields and
 1406 obtain the differential branching fraction. Table 3.7 reports the absolute efficiency
 1407 values for the J/ψ channel used to derive the relative efficiencies. Relative geometric,
 1408 detection and PID efficiencies are listed in Tab. 3.8, while Tabs. 3.10 and 3.9 report
 1409 relative reconstruction, trigger and NN efficiencies separately for downstream and
 1410 long candidates. Since the latter three components are obtained from the same sim-
 1411 ulated sample their statistical errors are correlated. Therefore the total of the three
 1412 is also reported as a single efficiency and labeled “Full Selection”. Finally, Tab. 3.13
 1413 reports the total of all relative efficiencies, which will be then used to correct the
 1414 raw yields and calculate the differential branching fraction. Uncertainties reflect
 1415 the statistics of both rare and resonant samples, while systematic uncertainties are
 1416 discussed in next sections.

Table 3.8: 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 3.9: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for long candidates; uncertainties reflect the statistics of both samples.

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

Table 3.10: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for downstream candidates; uncertainties reflect the statistics of both samples.

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

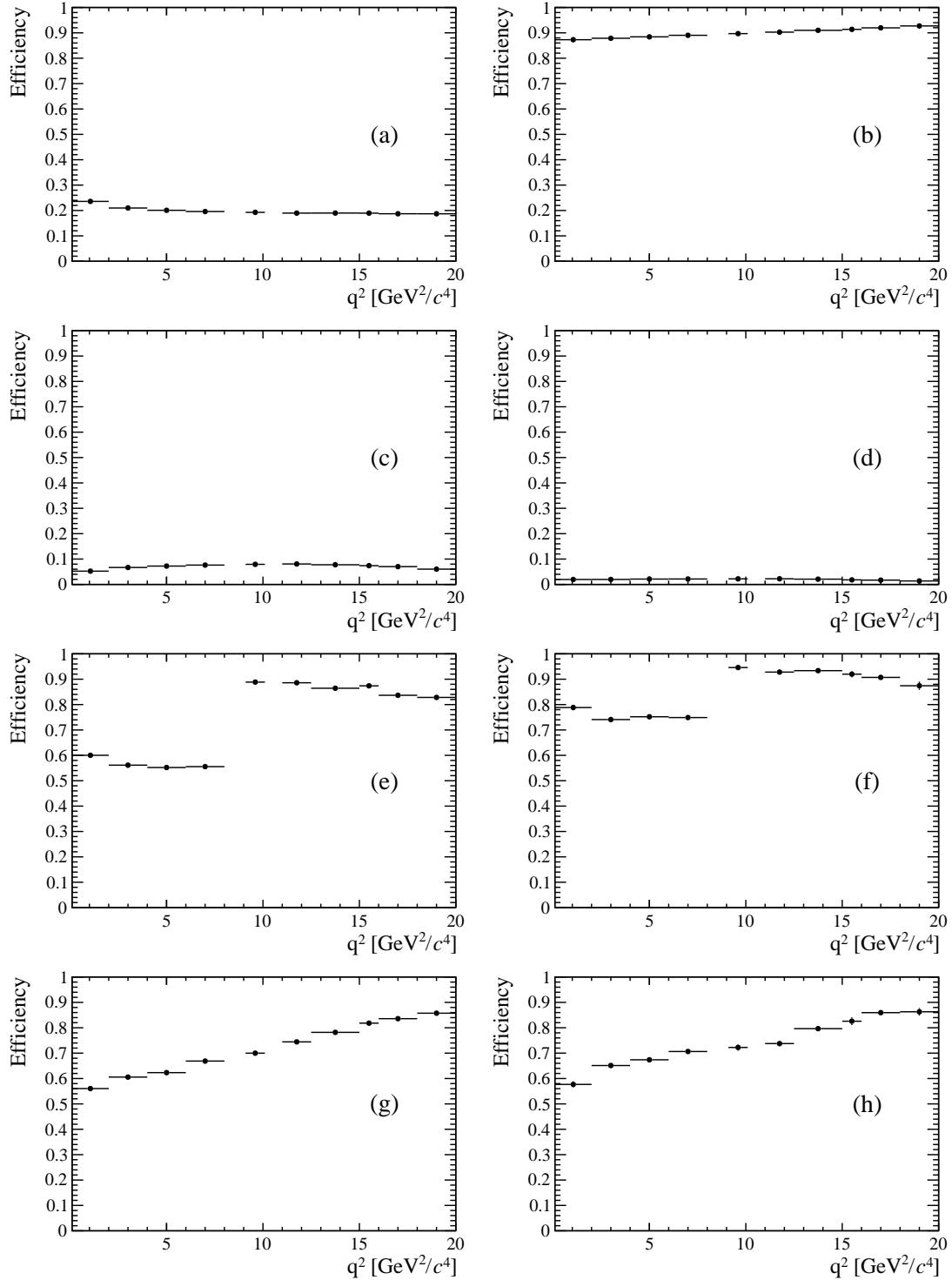


Figure 3.17: Absolute efficiencies as a function of q^2 : geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) candidates, NN efficiency for DD (e) and LL (f) and trigger efficiency for DD (g) and LL (h).

¹⁴¹⁷ **3.7 Systematic uncertainties**

¹⁴¹⁸ This section describes the main considered sources of systematic uncertainty.

¹⁴¹⁹ **3.7.1 Systematic uncertainty on the yields**

¹⁴²⁰ The choice of specific PDFs to model the invariant mass distribution could result in
¹⁴²¹ a bias. To asses the effect of the signal PDF choice as a first step a number of models
¹⁴²² are tried on the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ data sample to understand which ones are plausible.
¹⁴²³ Table 3.11 reports the χ^2 and relative probabilities obtained using different models
¹⁴²⁴ including: the default model (a DCB function), a simple Gaussian function, a single
¹⁴²⁵ Crystal Ball function and the sum of two Gaussians. The only two models that give
¹⁴²⁶ a reasonable p-value are the default DCB and the sum of two Gaussian functions
¹⁴²⁷ (DG). In a second step simulated experiments are generated and fit with the two
¹⁴²⁸ chosen models. Events are generated according to a density function given by the
¹⁴²⁹ default model fitted on data separately for each q^2 interval. In this way, for each
¹⁴³⁰ q^2 interval, a specific shape is reproduced including the background level and slope.
¹⁴³¹ Furthermore, a number of events comparable to the one found in data is generated.
¹⁴³² For each experiment a normalised bias is calculated as

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

¹⁴³³ where $N_{\ell\ell}^{model}$ and $N_{J/\psi}^{model}$ are the numbers of rare and resonant candidates observed
¹⁴³⁴ using a specific model. The average bias over 1000 pseudo-experiments is taken as
¹⁴³⁵ systematic uncertainty. Note that in each case the rare and normalisation channels
¹⁴³⁶ are fit with the same signal model and, while for the default case the rare parameters
¹⁴³⁷ are fixed to what found for the resonant channel, they are left free to vary in the
¹⁴³⁸ second model in order to asses at the same time the systematic due to the parameters
¹⁴³⁹ constraints.

Table 3.11: 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

1440

1441 For the background PDF systematic the rare channel is re-fit leaving the yield of
 1442 the K_s^0 component free to vary; this is instead fixed to the predicted value in the
 1443 default fit. The same procedure as for the signal PDF is applied. Results are re-
 1444 ported in Tab. 3.12. The most affected q^2 interval is the one in the middle of the
 1445 charmonium resonances, where a combination of lower statistics and higher back-
 1446 ground leaves more freedom to the signal shape. Finally, a background component
 1447 for $B^+ \rightarrow K^{*+}(K_s^0\pi^+)\mu^+\mu^-$ decays is added to the fit, modelled using the distri-
 1448 bution of simulated events after full selection. No significant bias is found for this
 1449 component.

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

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

1450

¹⁴⁵¹ 3.7.2 Systematic uncertainties on the efficiency determination

¹⁴⁵² Systematic uncertainties in the efficiency determination are due to the limited knowl-
¹⁴⁵³ edge of the decay properties such as the Λ_b^0 lifetime and production polarisation. The
¹⁴⁵⁴ uncertainties are directly calculated on the relative efficiencies as these are the ones
¹⁴⁵⁵ that are actually used in the analysis. It should be noted that not all sources con-
¹⁴⁵⁶ tribute to each part of the efficiency. For brevity, this section only reports estimates
¹⁴⁵⁷ of the systematic uncertainties obtained while the full information is contained in
¹⁴⁵⁸ Appendix C.

¹⁴⁵⁹ 3.7.2.1 Effect of new physics on the decay model

¹⁴⁶⁰ New physics could affect the decay model by adding contributions to the C_7 and
¹⁴⁶¹ C_9 Wilson Coefficients. This would result in a modification of the q^2 spectrum
¹⁴⁶² and therefore of the efficiency. To asses this systematic the Wilson Coefficients are
¹⁴⁶³ modified by adding a new physics component ($C_i \rightarrow C_i + C_i^{\text{NP}}$). Figure 3.18 shows q^2
¹⁴⁶⁴ spectra obtained weighting the simulation for a model embedding the default and 3
¹⁴⁶⁵ modified sets of Wilson Coefficients. The used values, reported on top of each plot,
¹⁴⁶⁶ are inspired to maintain compatibility with the recent LHCb result about the P'_5
¹⁴⁶⁷ observable [48]. The biggest effect is observed in the very low q^2 , below 2 GeV^2/c^4 ,
¹⁴⁶⁸ where the efficiency can change up to 7%, while it changes 3-4 % between 3 and
¹⁴⁶⁹ 4 GeV^2/c^4 and 2-3 % in the rest of the spectrum. As this analysis is performed under
¹⁴⁷⁰ the hypothesis that the decays are described by a the SM, these values are given in
¹⁴⁷¹ order to provide the full information but are not added as systematic uncertainties.

¹⁴⁷² 3.7.2.2 Simulation statistics

¹⁴⁷³ The limited statistics of the simulated samples used to determine efficiencies is
¹⁴⁷⁴ considered as a source of systematic uncertainty. While it is not the dominant
¹⁴⁷⁵ source, its size does not allow to completely neglect it. When reporting relative

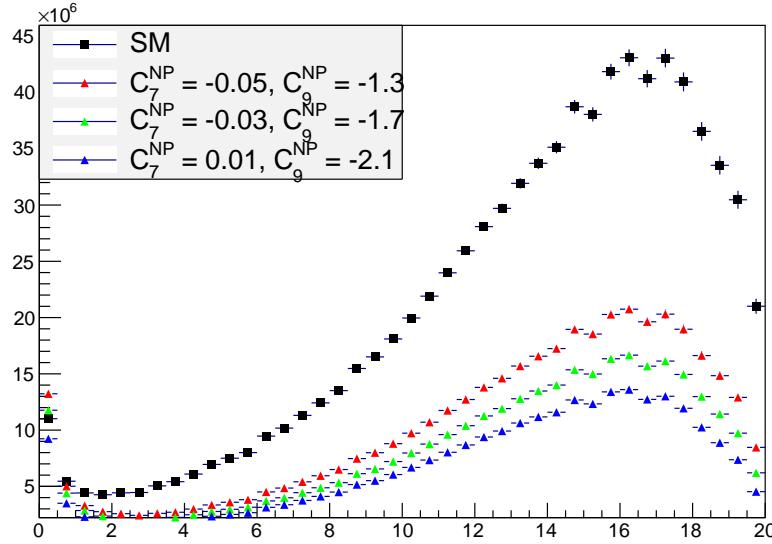


Figure 3.18: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

¹⁴⁷⁶ efficiency values the statistical uncertainty due to the rare and resonant channels is
¹⁴⁷⁷ always considered.

¹⁴⁷⁸ 3.7.2.3 Production polarisation and decay structure

¹⁴⁷⁹ One of the main unknown, which affects the determination of the efficiencies, is
¹⁴⁸⁰ the angular structure of the decays. And, connected to it, also the production
¹⁴⁸¹ polarisation, which is a parameter of the model. To assess the systematic uncertainty
¹⁴⁸² due to the knowledge of the production polarisation for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays the
¹⁴⁸³ polarisation parameter in the model is varied within one standard deviation from
¹⁴⁸⁴ the central value of the most recent LHCb measurement, $P_b = 0.06 \pm 0.09$ [102]. The
¹⁴⁸⁵ full observed difference is taken as systematic uncertainty. To assess the systematic
¹⁴⁸⁶ uncertainty due to the decay structure an alternative set of form factors is used based
¹⁴⁸⁷ on lattice QCD calculation [105]. Details of this are explained in Appendix A.1. The
¹⁴⁸⁸ two models are compared and the full difference is taken as systematic uncertainty.
¹⁴⁸⁹ In total this results in an uncertainty of $\sim 1.3\%$ for long candidates and $\sim 0.6\%$

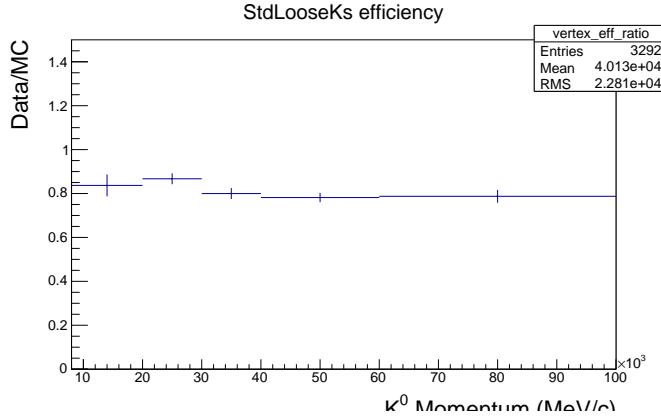


Figure 3.19: Ratio of reconstruction efficiency in Data and MC found using K_S events [107].

1490 for downstream candidates, mostly coming from the knowledge of the production
 1491 polarisation.

1492 3.7.2.4 Λ_b^0 lifetime

1493 The Λ_b^0 lifetime is known with limited precision. For evaluation of the efficiencies the
 1494 world average value, 1.482 ps^{-1} [106] is used. To evaluate the systematic uncertainty,
 1495 this is varied within one standard deviation from the measured value. Only the
 1496 case where both signal and normalisation channel are varied in same direction are
 1497 considered. The larger difference with the default lifetime case is taken as systematic
 1498 uncertainty, which is found to range from $\sim 0.4\%$ at low q^2 to $\sim 0.1\%$ at high q^2 .

1499 3.7.2.5 Downstream candidates reconstruction efficiency

1500 Other analysis in LHCb using particles reconstructed with downstream tracks showed
 1501 that the efficiency for these candidates is not well simulated. For example, Fig. 3.19
 1502 shows the ratio between the reconstruction efficiency for downstream candidates in
 1503 data and simulation found analysing K_S^0 events [107]. This effect is not yet fully
 1504 understood and is currently under study. It seems to be mainly due to a poor sim-
 1505 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. 3.19). The efficiencies obtained using the weighted and unweighted simulation are compared and the full difference is taken as systematic uncertainty. As the discrepancy shows little dependence on momentum, dependencies due to the different momentum distributions of Λ and K_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 3.13: Absolute values of the total relative efficiency and the absolute value of the uncorrelated uncertainty (σ_{uncorr}^k), together with relative values of the correlated uncertainty (σ_{corr}).

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

1529 Table 3.13 summarises the total relative efficiencies, ε^{rel} , for downstream and long
 1530 candidates together with their correlated and uncorrelated uncertainties, where the
 1531 correlation is intended between the downstream and long samples. On the table
 1532 the uncorrelated uncertainty corresponds to the total systematic uncertainty on the
 1533 efficiency determination. The correlated uncertainty is given in percent form since
 1534 it can be applied to either downstream, long candidates or their combination. This
 1535 includes the PDF systematic described in Sec. 3.7.1 and the systematic due to the
 1536 uncertainty on the $J/\psi \rightarrow \mu^+\mu^-$ branching fraction.

1537 Figure 3.20 shows the branching fraction obtained by fitting the downstream and
 1538 long samples independently, while the combined result, obtained fitting both sam-
 1539 ples simultaneously, is shown in Fig. 3.21. Values are also listed in Tab. 3.14, where
 1540 the statistical uncertainty on the rare channel and the total systematic uncertainty
 1541 are shown separately. The statistical uncertainty is calculated using the MINOS ap-
 1542 plication of the MINUIT package [108], which provides an asymmetric interval. The
 1543 normalisation and systematic uncertainties are evaluated by pushing the efficiencies
 1544 and normalisation yields up and down by one standard deviation and re-performing
 1545 the fit. The different efficiencies used translate into a different branching fraction and
 1546 the full difference with respect to the default fit is taken as systematic uncertainty
 1547 in each direction.

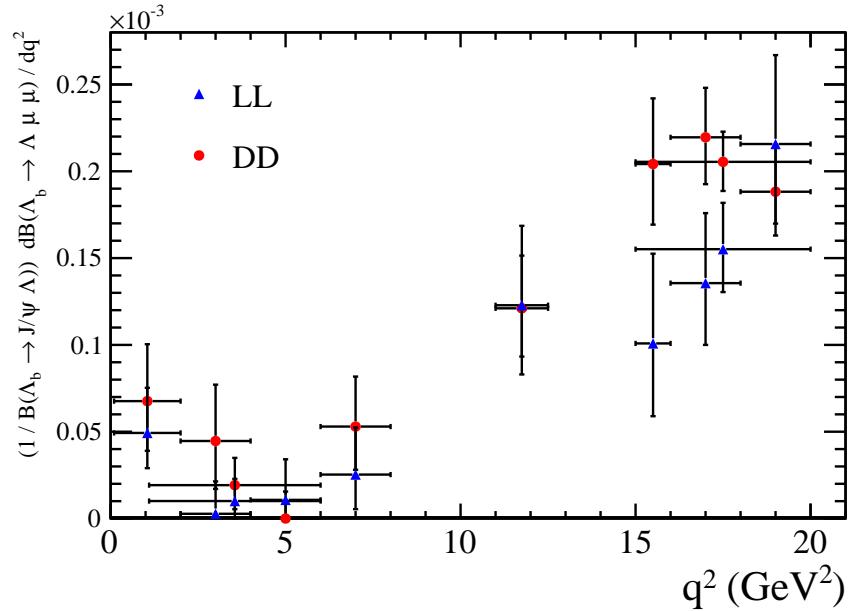


Figure 3.20: Measured values of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction relative to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay as a function of q^2 obtained fitting the downstream and long samples independently. Error bars represent the total statistical and systematic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

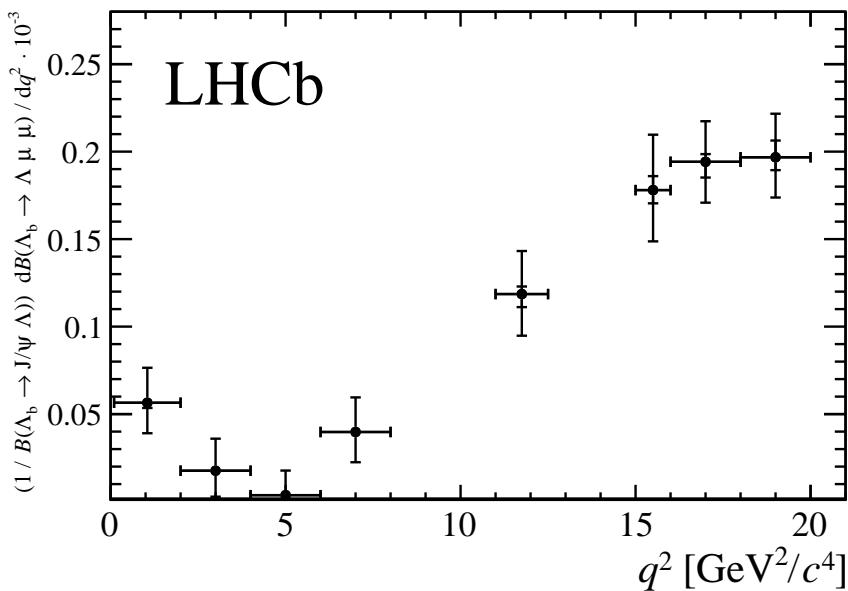


Figure 3.21: Branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay normalised to the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ mode. The inner error bar represents the systematic uncertainty and the outer error bar includes the statistical uncertainty.

Table 3.14: Measured differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay relative to $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays; 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 values listed in Tab. 3.14 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. 3.22 and summarised in Tab. 3.15, where the uncertainty due to the knowledge of the normalisation channel (norm), which is correlated across q^2 , is shown separately. The SM predictions on the plot are obtained from Ref. [105].

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

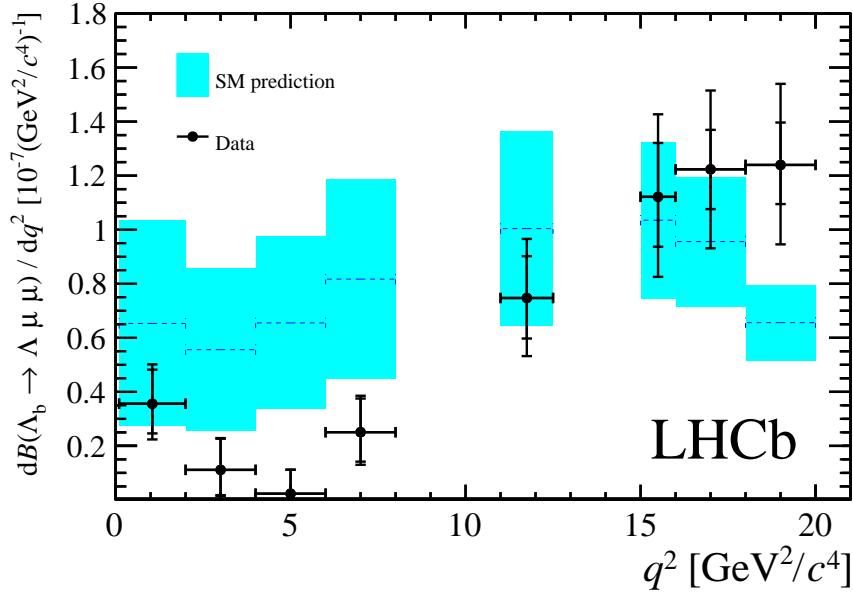


Figure 3.22: Measured $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction as a function of q^2 with SM predictions [105] superimposed. The inner error bars represent the total uncertainty on the relative branching fraction (statistical and systematic), while the outer error bar also includes the uncertainties due to the knowledge of the branching fraction of the normalisation mode.

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

q^2 interval [GeV^2/c^4]	$d\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-) / dq^2 \cdot 10^{-7} [(\text{GeV}^2/c^4)^{-1}]$	+/-	+/-	\pm
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 in the dimuon rest frame and the dimuon system direction in the Λ_b^0 ($\bar{\Lambda}_b^0$) rest frame, and, similarly, θ_h is defined as the angle between the proton and the Λ baryon directions, in the Λ and Λ_b^0 rest frames. 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

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

1566 where $d^2\Gamma/dq^2 d\cos\theta_i$ is the two-dimensional differential rate and $d\Gamma/dq^2$ is rate
1567 integrated over the angles.

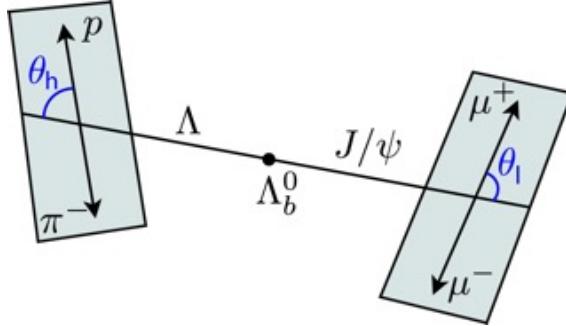


Figure 4.1: Graphical representation of the angles for the $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay.

1568 The A_{FB}^ℓ observable was also measured by LHCb in $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays which
1569 are going through the same quark level transition as $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays. Instead
1570 the hadronic asymmetry, A_{FB}^h , is interesting only in the Λ_b^0 case as it is zero by
1571 definition in B^0 decays, where K^{*0} decays strongly.

1572 4.1 One-dimensional angular distributions

1573 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.
1574 The content of this section is based on the calculations in Ref. [101].

1576 For unpolarised Λ_b^0 production, integrating over the three angles the differential
1577 branching fraction is given in Eq. 11 of Ref. [101] as

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

1578 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. [101] 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

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

1590 Using the same steps the proton helicity distribution is given in Ref. [101] as

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

1591 and A_{FB}^h is defined as

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

1592 where $P_z^\Lambda(q^2)$ is the polarisation of the daughter baryon, Λ , and $\alpha_\Lambda = 0.642 \pm 0.013$ [2]

1593 is the Λ decay asymmetry parameter.

1594 The above expressions assume that Λ_b^0 is produced unpolarised, which is in agree-
1595 ment with the recent LHCb measurement [109]. Possible effects due to a non zero
1596 production polarisation are investigated as systematic uncertainties (see Sec. 4.5.5).

1597 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 $\hat{n} = \hat{p}_{inc} \times \hat{p}_{\Lambda_b^0}$, where \hat{p}_{inc} represents the direction of the incoming proton. This angle is sensitive to the production polarisation through the spin-density matrix. 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

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

- 1598 In the massless leptons approximation two of these quantities are related to the
1599 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

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

- 1600 In order to use these distributions, expectations for the three additional observables,
1601 which do not enter one-dimensional distributions, are needed. Expectations are
1602 calculated using form factors and numerical inputs from Ref. [101] and are listed in
1603 Appendix A.1 in Tab. A.1.

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

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

1604 4.3 Angular resolution

1605 This section describes a study of the angular resolution done in order to achieve
 1606 a better understanding of detector and reconstruction effects. This will be then
 1607 used to study systematic uncertainties (see Sec. 4.5.5). The study is performed by
 1608 analysing simulated events and comparing generated and reconstructed quantities.
 1609 Figure 4.2 shows plots of the difference between true and measured angular observ-
 1610 ables ($\cos \theta_\ell$ and $\cos \theta_h$) as a function of the observable itself. These are centred at
 1611 zero indicating no bias in the measurement. Figure 4.3 shows the angular resolution
 in two-dimensional bins of q^2 and angular observables. In Fig. 4.2 the same differ-

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

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

1612
 1613 ence is shown also as a function of q^2 revealing again no bias. The spread of these
 1614 distributions around the central value can be takes as an estimate of the angular
 1615 resolution. Taking vertical slices of the plots in Fig. 4.2 one obtains approximately
 1616 gaussian distributions centred at zero. These are fit with a single gaussian and its
 1617 width is interpreted as the angular resolution. Table 4.1 reports the average resolu-
 1618 tions for the two angular observables separately for long and downstream candidates.
 1619 As expected candidates built from long tracks are characterised by a better angular

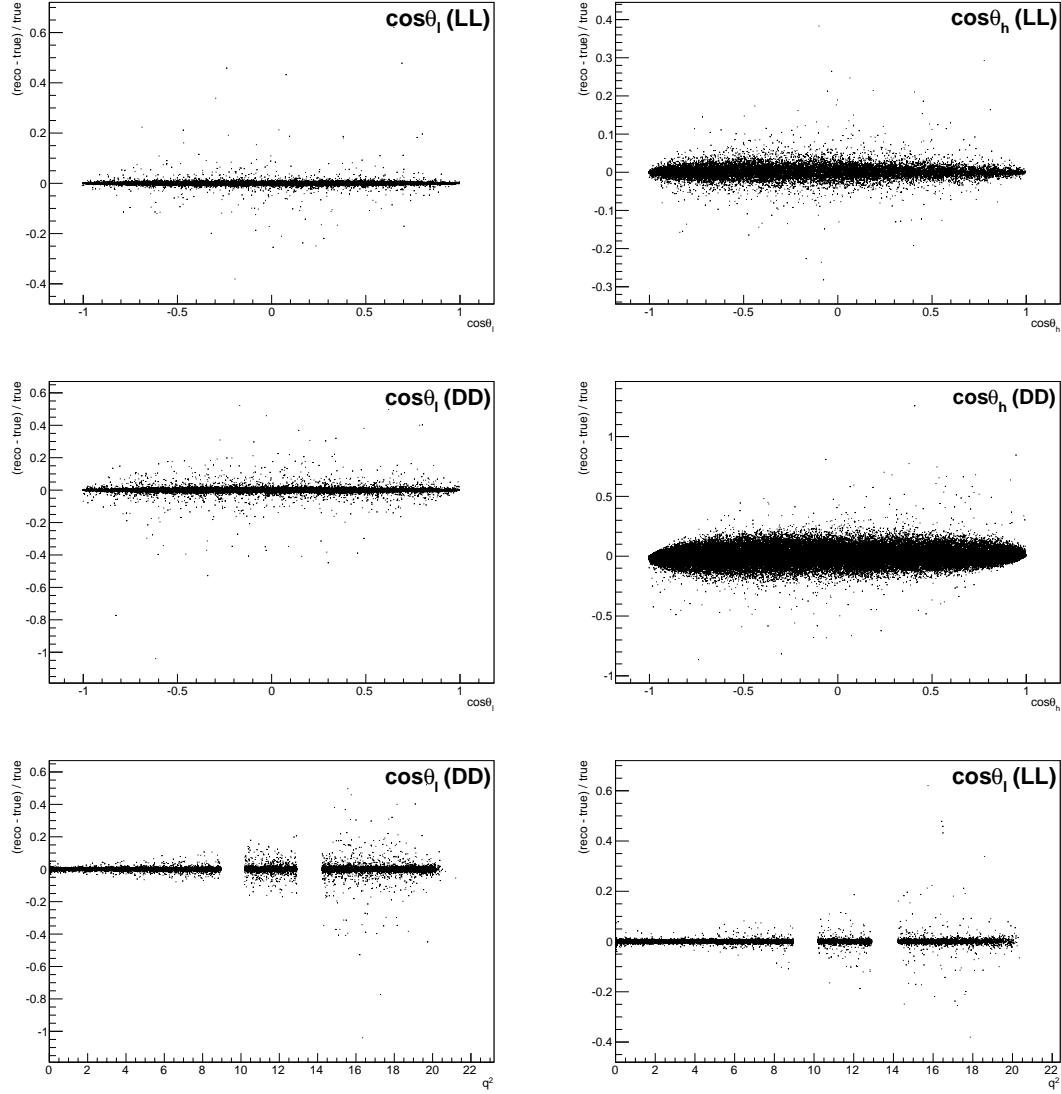


Figure 4.2: Difference of between generated and reconstructed angular observables as a function of the observables themselves for long (top) and downstream (bottom) candidates and as a function of q^2 for long (bottom left) and downstream (bottom right) candidates. As the plots are made using fully selected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

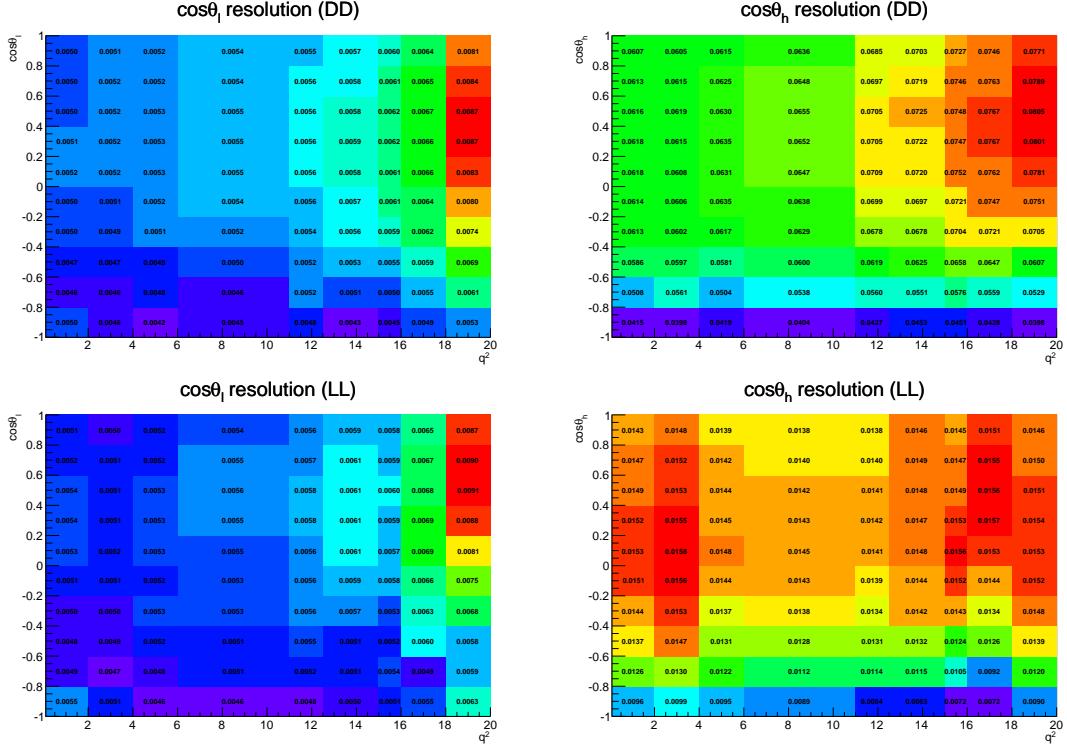


Figure 4.3: Angular resolution for $\cos \theta_\ell$ (left plots) and $\cos \theta_h$ (right plots) as a function of the angular observables and q^2 for downstream (upper plots) and long (lower plots) candidates. White bands correspond to the J/ψ and $\psi(2S)$ resonances which are excluded from the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

resolution due to a better momentum and vertex resolution.

4.4 Fit strategy

There are physical limits to the values of the parameters of interests: A_{FB}^h is limited in the $[-0.5, 0.5]$ interval and for the f_L and A_{FB}^ℓ parameters the physical region, given by $|A_{\text{FB}}^\ell| < 3/4(f_L - 1)$, is the triangle shown in Fig. 4.4. If the measured value is close to the border the fit does not always converge. Therefore a “brute force” fitting technique is applied. For this purpose fit parameters are divided into two categories: parameters of interest (PoIs), A_{FB}^ℓ , A_{FB}^h and f_L and all other parameters, which are referred to as “nuisances”. The value of the Log-Likelihood ($\log \mathcal{L}$) of the fit model with respect to data is evaluated in a grid of points in the PoIs allowed area to find the function minimum. A first coarse scan finds a candidate minimum and then the

procedure is reiterated two more times in finer intervals around it. For each point all the nuisances are fitted using a maximum likelihood fit. Using this method the best fit point is therefore constrained inside the physical region. If the minimum of the log-likelihood is found to be outside it, the closest point on the boundary is chosen as the best fit.

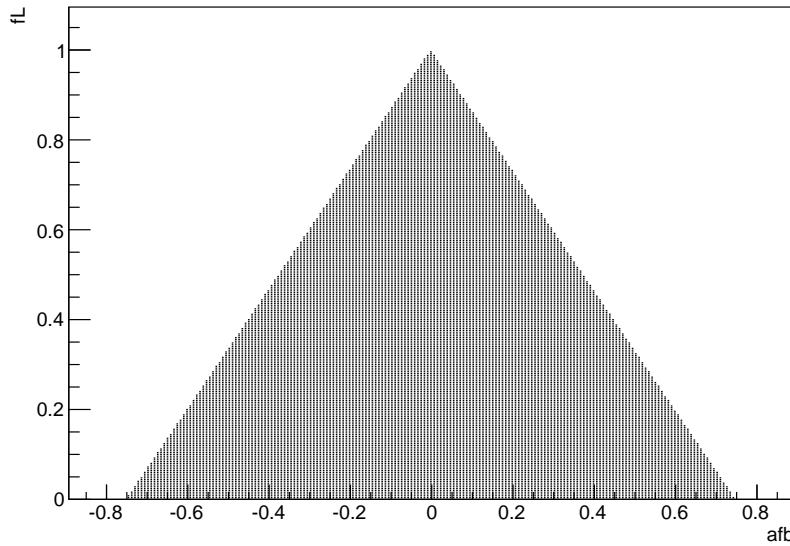


Figure 4.4: The physical (A_{FB}^{ℓ}, f_L) parameter space. The dark region corresponds to points where the PDF is positive in the whole $[-1, 1]$ interval.

1636 4.4.1 Feldman-cousins plug-in method

Physical boundaries of the parameter space could result in a wrong estimation of the uncertainties, especially if the measured value is close to the border. To deal with this effect in this analysis the likelihood-ordering method [110] is used to estimate uncertainties and nuisance parameters are accounted for using the plug-in method [111]. This is a unified method to calculate confidence intervals and upper/lower limits, based on simulated experiments and has the advantage of having a well defined frequentist coverage.

The method is constituted by the following steps:

-
- 1645 1. fit real data distributions with all parameters free;
- 1646 2. fit real data fixing the PoIs to a value of choice while keeping nuisance param-
- 1647 eters free;
- 1648 3. generate simulated samples following the distribution given by the fit model,
- 1649 where all nuisance parameters are taken from the fit in point 2 and PoIs are
- 1650 fixed to the same value used in point 2;
- 1651 4. repeat the two fits made on data (points 1 and 2) on each simulated sample:
- 1652 fit with all parameters free and with fixed PoIs;
- 1653 5. extract the value of the Log-Likelihoods at the minimum for all cases;
- 1654 6. calculate the percentage of simulated experiments in which the free-to-fixed
- 1655 likelihood ratio is bigger than in data: $\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free} > (\log \mathcal{L}_{fixed} / \log \mathcal{L}_{free})_{data}$;
- 1656 7. repeat the procedure for many values of the PoIs scanning around the best fit
- 1657 point.

1658 The confidence interval at $k\%$ is given by the points where the free-to-fixed likelihood

1659 ratio is bigger in data than simulation for $(1 - k)\%$ of times. As an example, Fig. 4.5

1660 shows the p-values obtained with the plug-in method for A_{FB}^h and f_L . A two-

1661 dimensional region can also be scanned giving a grid of p-values, which translates

1662 into two-dimensional confidence regions.

1663 4.4.2 Modelling the angular distributions

- 1664 The observables are extracted from fits to one-dimensional angular distributions.
- 1665 The PDFs used to model the data are defined as

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

1666 where $k = (\text{LL}, \text{DD})$. The signal function is composed by a theoretical shape (P_S)

1667 given by Eq. 4.10 and 4.9, which is multiplied by an acceptance function ε described

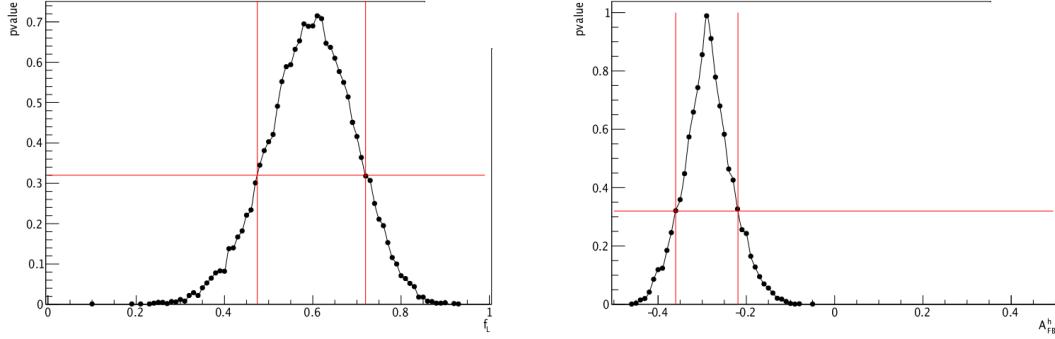


Figure 4.5: Dependence of the p-value from the values of the angular observables f_L (left) and A_{FB}^h (right) in simulated experiments. The red lines mark the points at p-value 32% corresponding to a 68% CL.

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

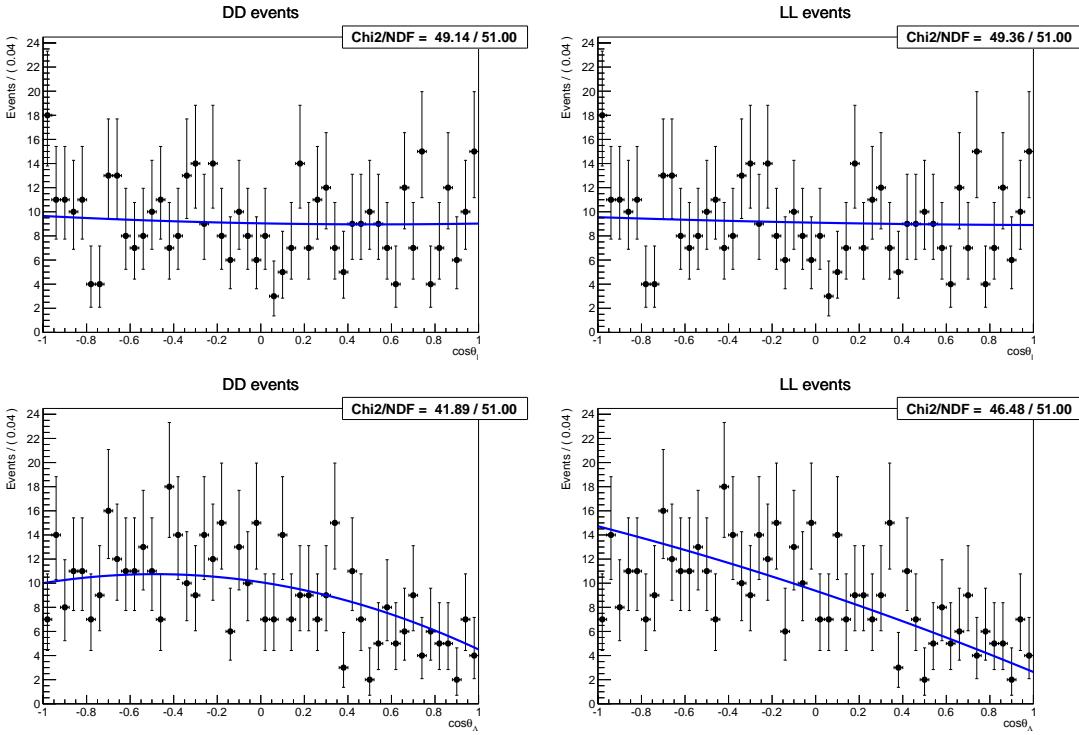


Figure 4.6: Background distribution as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

1684 4.4.3 Angular acceptance

1685 Selection requirements on the minimum momentum of the muons can distort the
 1686 $\cos \theta_\ell$ distribution by removing candidates with extreme values of the angle. Simi-
 1687 larly, the impact parameter requirements affect $\cos \theta_h$ because very forward hadrons
 1688 tend to have smaller impact parameter values. While in principle one could take this
 1689 into account by an additional weight, to minimise the distortion of the uncertain-
 1690 ties estimate, the efficiency function is incorporated in the fit model. The angular
 1691 efficiency is parametrised using a second-order polynomial, whose parameters are de-
 1692 termined separately for long and downstream candidates by fitting simulated events.
 1693 An independent set of parameters is obtained also for each q^2 interval. These pa-
 1694 rameters are then fixed for the fits to data. Using polynomial functions allows to
 1695 calculate the PDF normalisation analytically. Figure 4.7 shows the acceptance as a
 1696 function of $\cos \theta_h$ and $\cos \theta_\ell$ for the $15.0\text{--}20.0$ integrated q^2 interval obtained using
 1697 a $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ simulated sample. For the lepton side, even though the efficiency

1698 is symmetric by construction, all parameters are left free to float, namely it is not
1699 constrained to be symmetric.

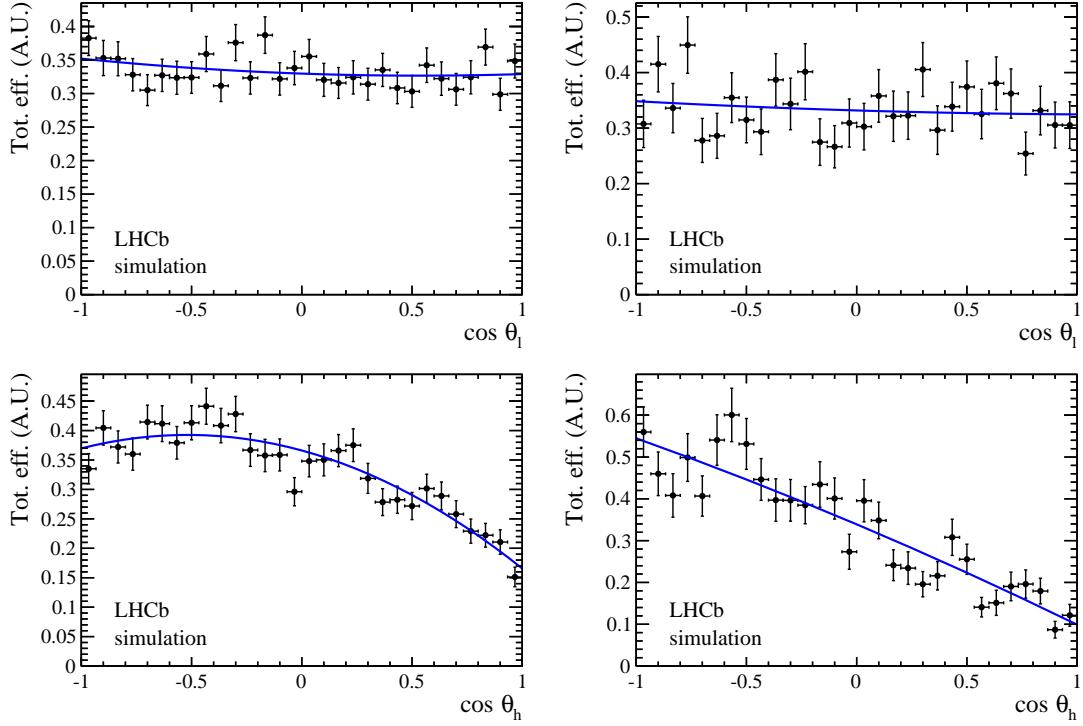


Figure 4.7: Efficiency as a function of $\cos \theta_\ell$ (top) and $\cos \theta_h$ (bottom) for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

1700 4.4.4 Studies on a three-dimensional fit

1701 One other way of extracting the angular observables would be to fit at the same
1702 time both angles and also the invariant mass distribution in order to have a better
1703 handle on the level of background. In this case one can use more of the information
1704 available. On the other hand it is necessary to use a larger mass window including
1705 more background and this method involves more parameters to fit. In the 1D case
1706 the free parameters are the two parameters of interest (A_{FB}^ℓ and f_L) for the lepton
1707 case and one (A_{FB}^h) for the hadron one. For the 3D case in addition to the three
1708 PoIs there are two background fractions and the two exponential slopes for the
1709 invariant mass background. Furthermore, to take correctly into account correlations
1710 three more observables enter the fit (see Eq. 4.12). As an high number of free

1711 parameters is difficult to constrain with the very limited statistics available, pseudo-
1712 experiments are used to check which method gives the best sensitivity. Events are
1713 generated in a 3D $(\cos \theta_\ell, \cos \theta_h, m_{p\pi\mu\mu})$ space. The generated values of the PoIs
1714 are $A_{\text{FB}}^\ell = 0$, $f_L = 0.7$ and $A_{\text{FB}}^h = -0.37$, which are data-like values inspired to
1715 a preliminary measurement in the highest statistics interval. The overall statistics
1716 and the fraction of background events in the mass window are generated to be data-
1717 like using information from the preliminary fit to data. Each pseudo-experiment is
1718 fitted with both methods and Fig. 4.8 reports distributions of parameters of interest
1719 obtained from the fit in the 1D and 3D cases. The RMS of these distributions can
1720 be taken as a measure of the sensitivity of each method. Table 4.2 lists the RMSs
1721 obtained from both methods; for all parameters of interest the 1D fit method gives
1722 a smaller RMS, hence a better sensitivity.

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

q^2 [GeV $^2/c^4$]	Fit type	A_{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

1722

1723 4.5 Systematics uncertainties on angular observables

1724 The following section describes the five main sources of systematic uncertainties
1725 that are considered for the angular observables measurement and, finally, results
1726 are reported in Sec. 4.7. Results are derived only for q^2 intervals where the signal
1727 significance, shown in Tab. 3.6, is above 3 standard deviations. This includes all
1728 q^2 intervals above the J/ψ resonance and the lowest q^2 interval, where an increased
1729 yield is due to the presence of the photon pole.

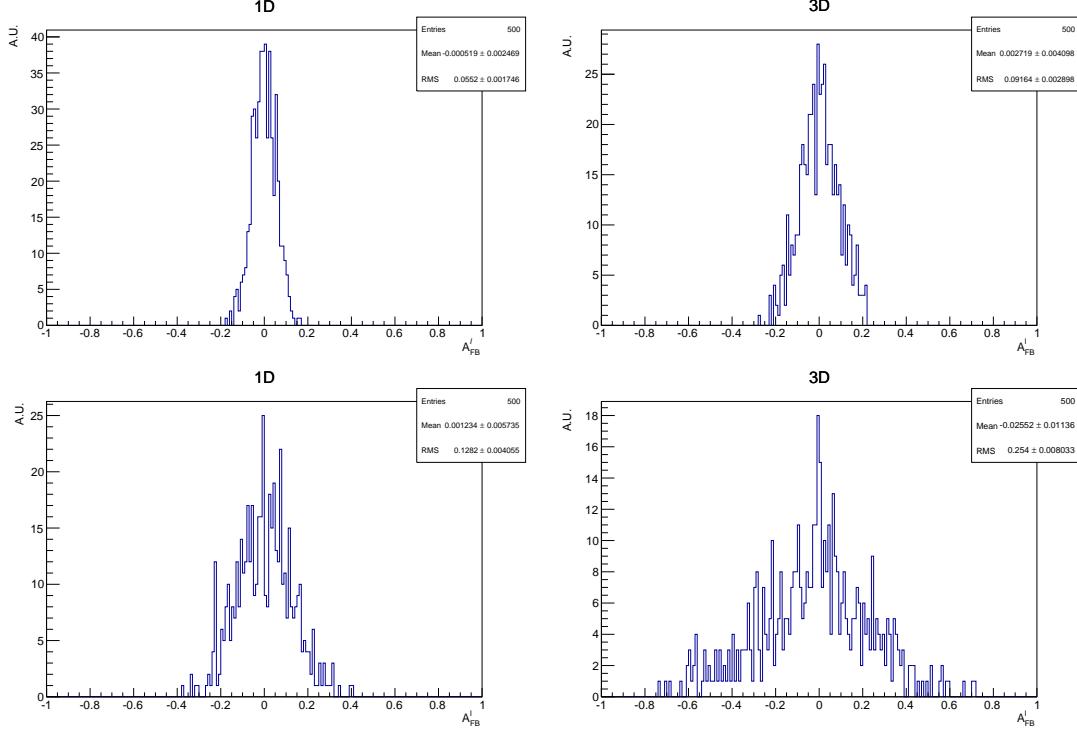


Figure 4.8: Values of the A_{FB}^l parameter observed over pseudo-experiments with input $A_{FB}^l = 0$ using the 1D fit method (left) and the 3D one (right). Events are generated with parameters and statistics corresponding to what is observed in the highest statistics interval (top), 15–20 GeV²/c⁴, and in the lowest statistics one, 11–12.5 GeV²/c⁴.

1730 4.5.1 Angular correlations

1731 The angular acceptance is non-flat as a function of $\cos\theta_\ell$ and $\cos\theta_h$. Therefore, while
 1732 integrating the full angular distribution, terms that cancel with perfect efficiency
 1733 may remain and generate a bias in the final result. In order to deal with this effect
 1734 simulated events are generated in a two-dimensional ($\cos\theta_\ell, \cos\theta_h$) space according to
 1735 the theoretical distribution described by Eq. 4.15 multiplied by the two-dimensional
 1736 efficiency function obtained from simulation. Then, one-dimensional projections are
 1737 taken and fit using the default one-dimensional efficiency functions. The distribu-
 1738 tions of observed deviations from the generated value, $\Delta x = x_{true} - x_{measured}$, are
 1739 approximately gaussian and their mean is non-zero by more than 3σ . Therefore the
 1740 mean biases are taken as systematic uncertainties.

1741 4.5.2 Resolution

1742 The angular resolution could bias the observables measurement generating an asym-
 1743 metric migration of events. This is especially important in the $\cos \theta_h$ case, because
 1744 it has worse resolution and a considerably asymmetric distribution. Simulated ex-
 1745 periments are used to asses this systematic. Events are generated according to
 1746 the measured distributions including efficiencies. The generated events are then
 1747 smeared by the angular resolution (gaussian smearing). To be conservative the case
 1748 with biggest angular resolution, downstream candidates, is always used. Finally,
 1749 the smeared and not-smeared distributions are fit with the same PDF. The average
 1750 deviation from the default values are reported in Tab. 4.3 as a function of q^2 and
 assigned as systematic uncertainties.

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

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

1751

1752 4.5.3 Efficiency description

1753 An imprecise determination of the reconstruction and selection efficiency can intro-
 1754 duce an extra oddity and therefore bias the measurement. To asses this effect the
 1755 kinematic re-weighting described in Sec. 3.3.2 is removed from the simulation and
 1756 the efficiency is determined again. Simulated events are then fit using the same the-
 1757 oretical PDF but multiplied by the efficiency functions obtained with and without
 1758 kinematical weights. As in the previous cases the average bias is taken as systematic
 1759 uncertainty; results are shown in Tab. 4.4. Furthermore, the effect of the limited
 1760 simulated statistics is taken into account and added to the systematic uncertainty.

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

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

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

q^2 [GeV $^2/c^4$]	A_{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

¹⁷⁶¹ 4.5.4 Background parameterisation

¹⁷⁶² There is a certain degree of arbitrariness in the choice of a parameterisation for the
¹⁷⁶³ background, especially for q^2 intervals with low statistics. To assess possible biases
¹⁷⁶⁴ due to the PDF choice, simulated experiments are generated using the shapes from
¹⁷⁶⁵ data fits and the same statistics as observed in data for each q^2 interval. Each
¹⁷⁶⁶ pseudo-experiment is fit with two models: the default one, a “line times efficiency”
¹⁷⁶⁷ function and the efficiency function alone, corresponding to the assumption that
¹⁷⁶⁸ background distributions are originally flat and only modified by the interaction
¹⁷⁶⁹ with the detector. The average bias with respect to the default model is taken as
¹⁷⁷⁰ systematic uncertainty; results are reported in Tab. 4.6.

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

q^2 [GeV $^2/c^4$]	A_{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

1772 4.5.5 Polarisation

1773 To study the effect of a non-zero Λ_b^0 production polarisation simulated events are
 1774 generated using the distributions given by Eqs. 4.12 and 4.14 as a function of the
 1775 angle under study ($\cos \theta_\ell$ or $\cos \theta_h$) and $\cos \theta$, defined in Sec. 4.2, which is sensitive to
 1776 polarisation. Similarly to the procedure used for the branching ratio measurement,
 1777 events are generated using values of the polarisation corresponding to $\pm \sigma$ from the
 1778 LHCb measurement [102]. In the theoretical functions $\cos \theta$ is always odd therefore
 1779 with perfect efficiency it always drops out by integrating over $\cos \theta$. Therefore
 1780 the generated distributions are also multiplied by the two-dimensional efficiency
 1781 function. No significant bias is found.

1782 4.6 J/ψ cross-check

1783 To cross-check the fitting procedure this is applied on the high statistics $\Lambda_b^0 \rightarrow J/\psi \Lambda$
 1784 sample. For this purpose events are selected with an additional requirement on the
 1785 proton PID, $\text{PID}_p > 10$. This is needed to reduce the $B^0 \rightarrow K_s^0 J/\psi$ background,
 1786 which is particularly important for the hadronic side fit, since the K_s^0 events are
 1787 not distributed in a flat way in the $\cos \theta_h$ variable and would therefore bias the fit.
 1788 Figure 4.9 shows the invariant mass distributions after this requirement is applies,
 1789 which can be compared with the ones in Fig. 3.12. After the PID cut there are
 1790 0.2% of K_s^0 events left in the downstream sample and a fraction compatible with

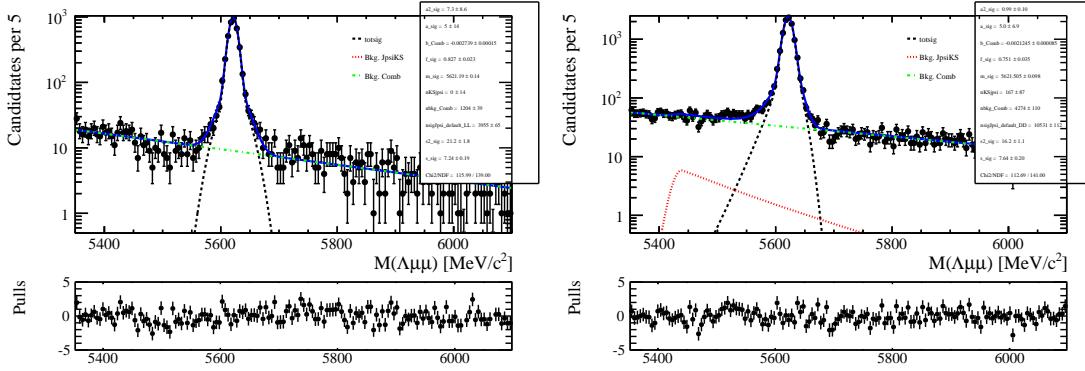


Figure 4.9: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ long (left) and downstream (right) candidates with an extra proton PID cut to remove K_s^0 background.

zero in the long sample. The signal model used for this fit is the same used for the

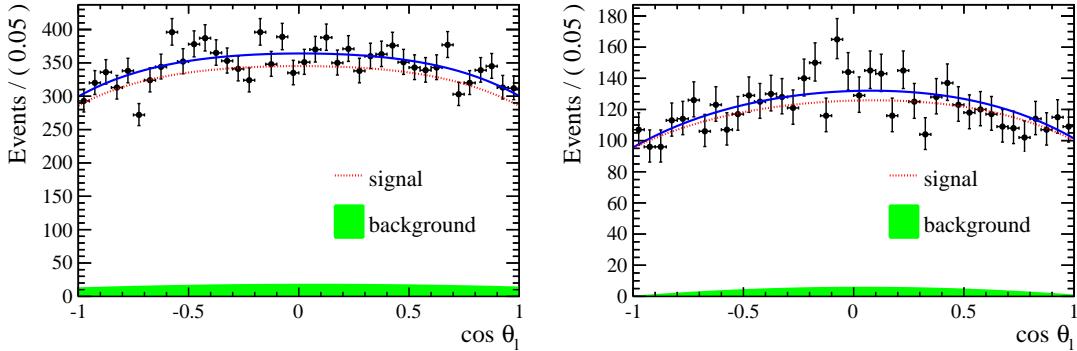


Figure 4.10: Fitted angular distribution as a function of $\cos \theta_\ell$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

1791

1792 rare case and described in Sec. 4.4.2. For the background instead the higher statis-
 1793 tics allows to leave more freedom to the fit. Therefore a second-order Chebyshev
 1794 polynomial is used, where the two parameters are free to vary. As for the rare case
 1795 the background fractions are gaussian-constrained to what found from the invariant
 1796 mass fit. Figures 4.10 and 4.11 show fitted angular distributions for the J/ψ channel.
 1797 The measured values of the observables are $A_{\text{FB}}^\ell = -0.002^{+0.011}_{-0.011}$, $A_{\text{FB}}^h = -0.402^{+0.010}_{-0.009}$
 1798 and $f_L = 0.485^{+0.019}_{-0.020}$, where the uncertainties are 68% Feldman Cousins confidence
 1799 intervals. The lepton side asymmetry as expected is measured to be zero.

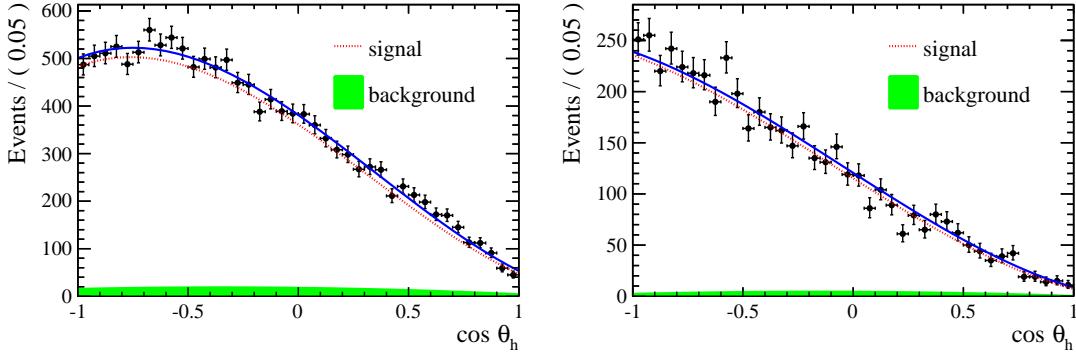


Figure 4.11: Fitted angular distribution as a function of $\cos \theta_h$ for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates reconstructed using downstream (left) and long (right) tracks.

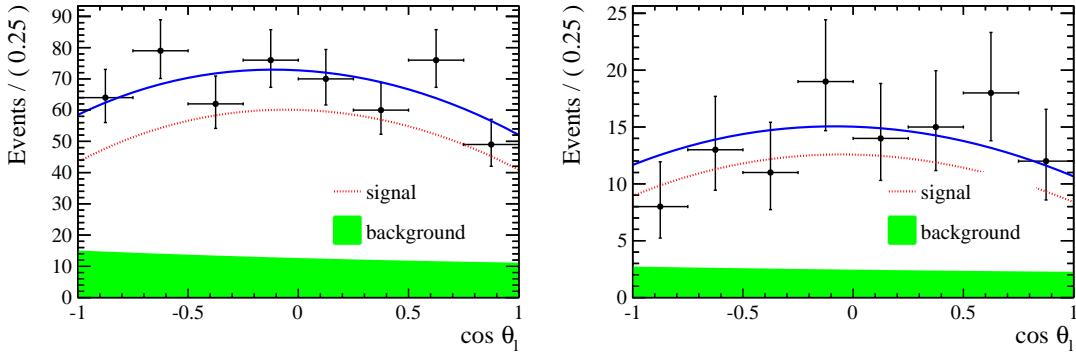


Figure 4.12: Fitted angular distributions as a function of $\cos \theta_\ell$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

4.7 Results

Figures 4.12 and 4.13 show fits to the angular distributions for the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval and Tab. 4.7 reports measured values of A_{FB}^ℓ , A_{FB}^h and f_L . The asymmetries are also shown in Fig. 4.14 together with SM predictions obtained from Ref. [105]. 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. 4.15 as two-dimensional 68 % confidence level (CL) regions, where the likelihood-ratio ordering method is applied by varying both observables at the same time and therefore taking correlations into account. Total systematic uncertainties correspond to the square root sum of the single considered sources.

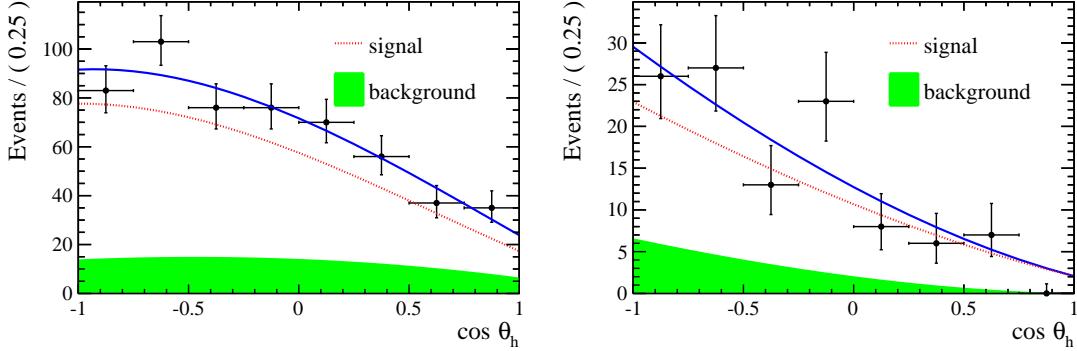


Figure 4.13: Fitted angular distributions as a function of $\cos \theta_h$ for downstream (left) and long (right) candidates in the $15\text{--}20 \text{ GeV}^2/c^4 q^2$ interval.

Table 4.7: Measured values of leptonic and hadronic angular observables; uncertainties are statistical and 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	$-0.05^{+0.09}_{-0.09} \pm 0.03$	$0.61^{+0.11}_{-0.14} \pm 0.03$	$-0.29^{+0.07}_{-0.07} \pm 0.03$

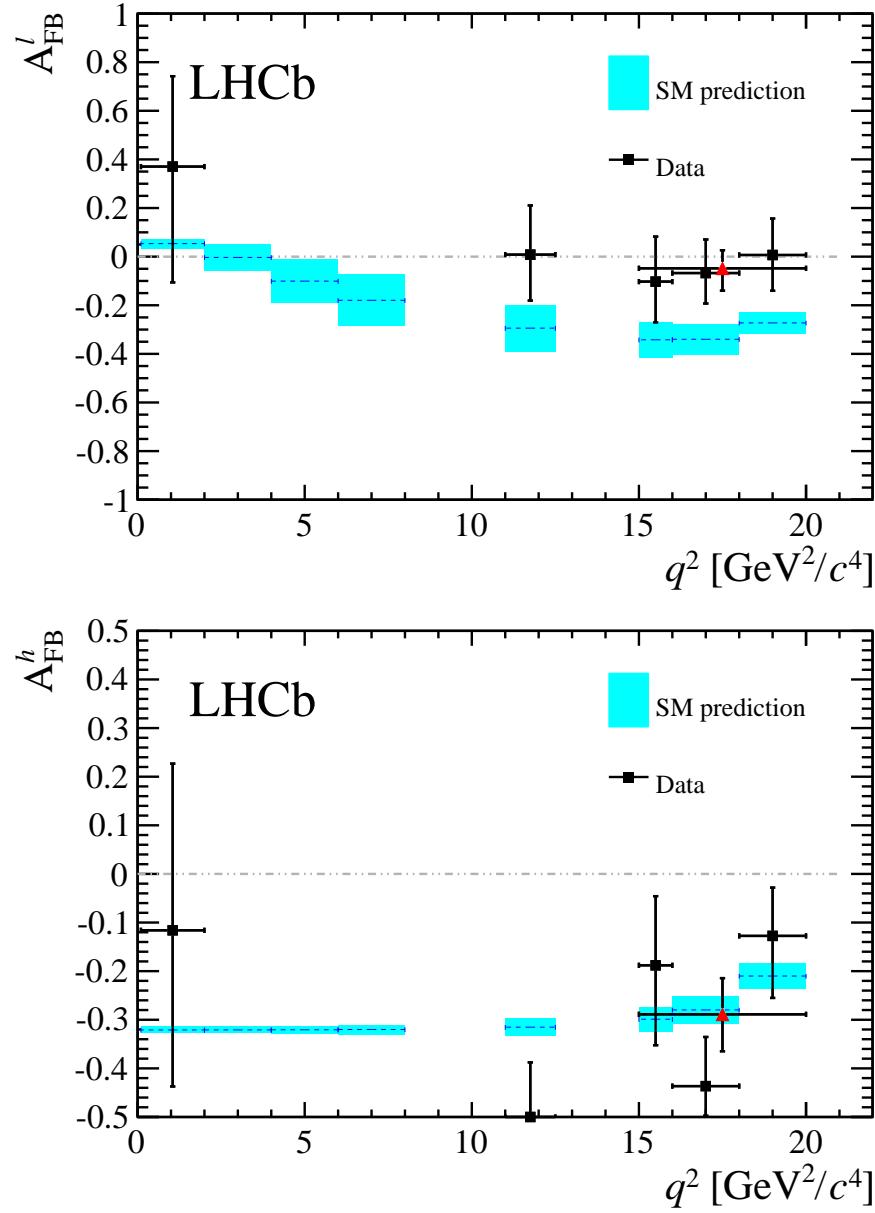


Figure 4.14: Measured values of the leptonic (top) and the hadronic (bottom) forward-backward asymmetries. Data points are only shown for q^2 intervals where the signal yield is found to be statistically significant, see text for details. The (red) triangle represents the values for the $15 < q^2 < 20$ GeV $^2/c^4$ interval. Standard Model predictions are obtained from Ref. [112].

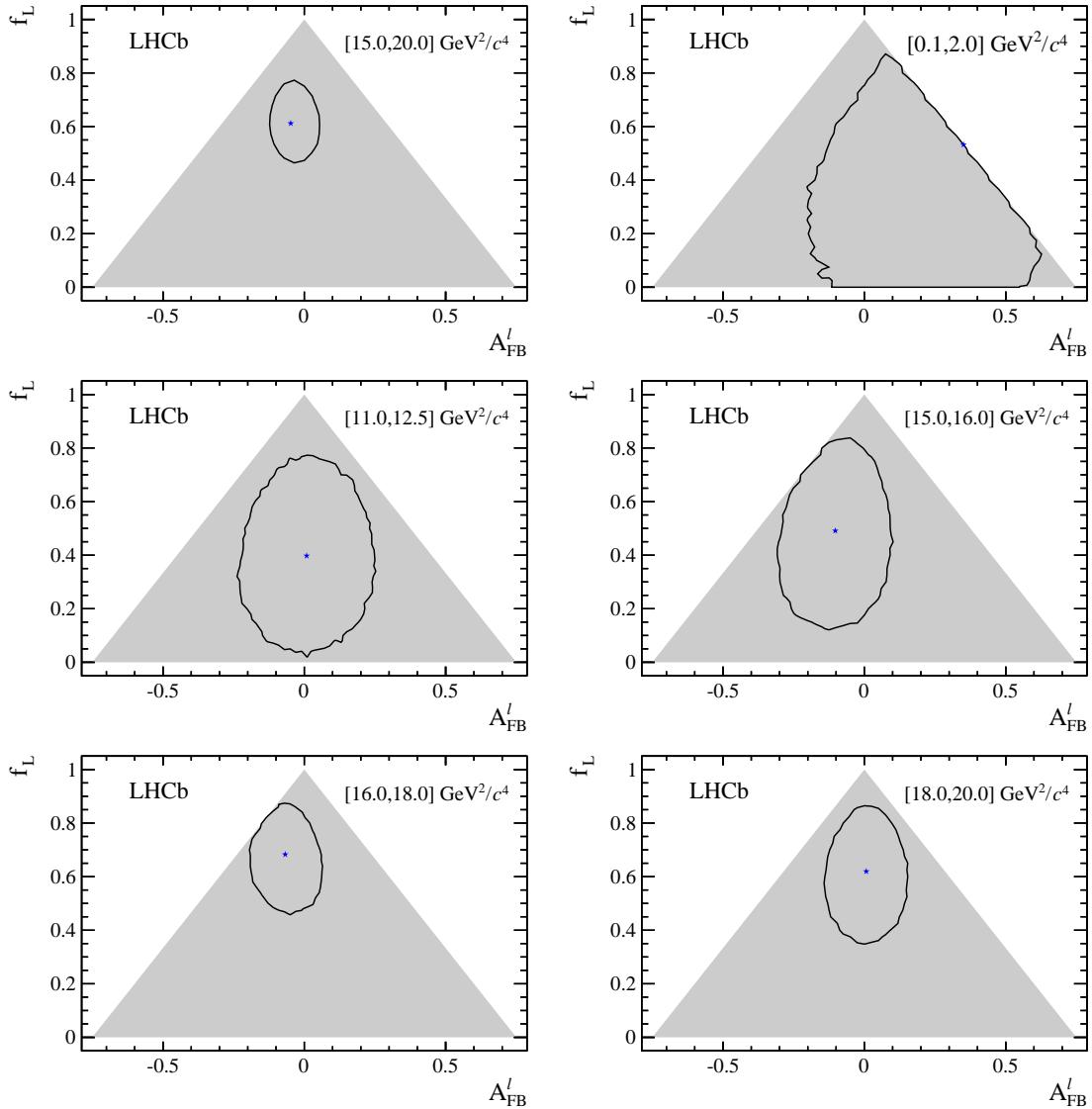


Figure 4.15: Two-dimensional 68 % CL regions (black) as a function of A_{FB}^ℓ and f_L . The shaded areas highlight the region in which the PDF is positive over the whole $\cos \theta_\ell$ range. The best fit points are indicated by the (blue) stars.

CHAPTER 5

1811

1812

1813

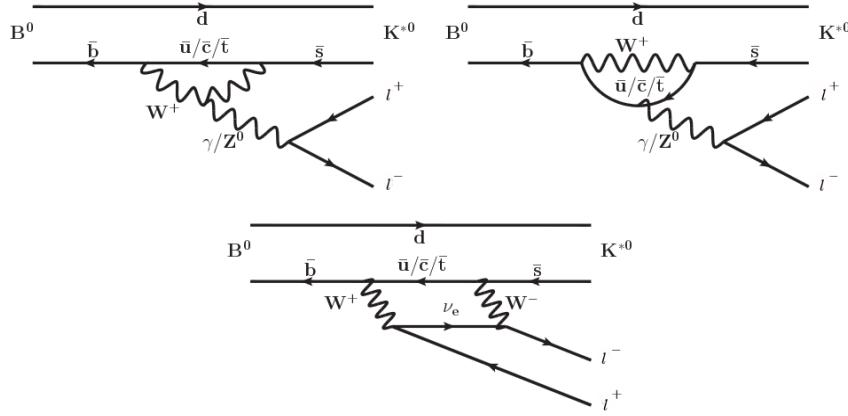
Testing lepton flavour universality with $R_{K^{*0}}$

1814

1815 Lepton Flavour Universality (LFU) is the equality of the weak coupling constants
1816 for all leptons. FCNC processes, which are forbidden in the SM at tree level and
1817 happen only at loop level, are an ideal laboratory to study LFU as new physics in
1818 the loops could break the flavour symmetry.

1819 In this work $b \rightarrow s\mu^+\mu^- (e^+e^-)$ decays are studied to test LFU between electrons and
1820 muons using penguin decays. In particular, the B^0 meson semileptonic decays $B^0 \rightarrow$
1821 $K^{*0}\ell^+\ell^-$ are considered. Figure 5.1 shows the possible Feynman diagrams producing
1822 such decays while Fig. 5.2 illustrates how these Feynman diagrams may include new
1823 particles. A series of recent LHCb measurements [31] points to a tension with SM
1824 predictions, which make these processes very interesting to better understand the
1825 nature of the discrepancy.

1826 In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger pro-
1827 posed the measurement of the R_H ratio [113], 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.

1828 inclusive state containing an s quark (X_s) or an s -quark resonance like K or K^{*0} .

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

1829 In this quantity the differential branching ratio is integrated over the squared dilepton invariant mass, q^2 , from $q_{min}^2 = 4m_\mu^2$, which is the threshold for the $\mu\mu$ process,
1830 up to $q_{max}^2 = m_b^2$.

1832 The advantage of using ratios of branching fractions as observables is that, in the
1833 theoretical prediction, hadronic uncertainties cancel out. Furthermore, experimentally,
1834 some of the systematic uncertainties on the ratios are reduced giving a better
1835 measurement. For example, what is measured is the number of $\mu\mu$ and ee decays
1836 happening in a certain period of time. Then, the luminosity, \mathcal{L} , is used to obtain
1837 a cross section, σ , using $R = \mathcal{L}\sigma$, where R is the rate at which the decays occur.
1838 The luminosity measurement is usually a source of systematic uncertainty, but it
1839 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

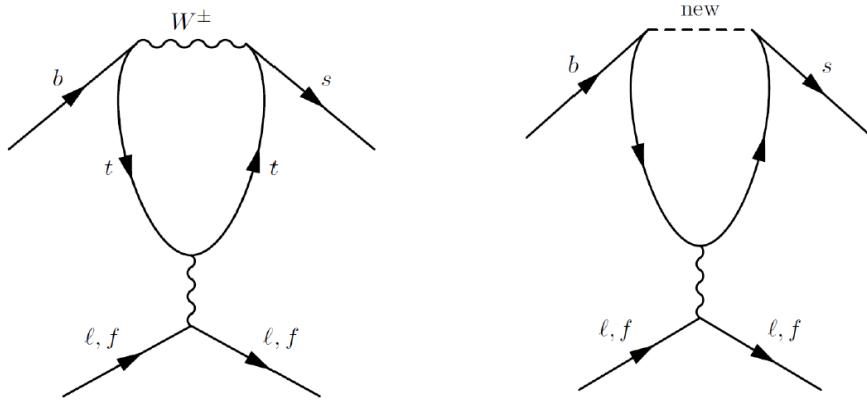


Figure 5.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

range [113]:

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

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

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

$$(5.5)$$

₁₈₄₀ under the assumptions that:

- ₁₈₄₁ • right-handed currents are negligible;
- ₁₈₄₂ • (pseudo-)scalar couplings are proportional to the lepton mass;
- ₁₈₄₃ • there are no CP-violating phases beyond the SM.

₁₈₄₄ The measurement of the R_H ratios is of particular interest after the recent measurement of the branching ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay [42], where no evidence of ₁₈₄₅ new physics was found. In fact the quantities $(R_H - 1)$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ remain ₁₈₄₆ proportional with ₁₈₄₇

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

₁₈₄₈ A joint measurement of these two quantities can give much information and constrain ₁₈₄₉ 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 [31] 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. [114] 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. [114] 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

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

where the Δ_{\pm} quantities are combinations of Wilson coefficients described in Eq. 10 of Ref. [114] and the parameter p is the polarisation of K^* that in Ref. [114] is determined to be close to 1 simplifying the formula to $R_{K^*} \simeq 1 + \Delta_-$. In particular one can observe the following correlations:

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

5.2 Experimental status

The R_K and $R_{K^{*0}}$ ratios have been measured at the B-factories [115, 116], and the R_K ratio has been recently measured also at LHCb [117] in the $1 < q^2 < 6$ GeV $^2/c^4$ q^2 interval, which represents the most precise measurement to date. This measurement manifests a 2.6σ deviation from the SM prediction. The current experimental status is summarised in Tab. 5.1. By profiting of the large dataset collected during Run-I, the LHCb experiment is expected to reduce the uncertainty on $R_{K^{*0}}$ by at least a factor of 2 with respect to the B-factories.

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

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

1886

1887 It is also worth mentioning the measurement of the $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$
 1888 ratio which was measured to be 2.1 standard deviations larger than the value ex-
 1889 pected from the assumption of lepton universality in the SM [118].

1890 5.3 Analysis strategy

1891 The aim of the analysis in this chapter is to measure the $R_{K^{*0}}$ ratio using pp collision
 1892 data collected by the LHCb detector in 2011 and 2012, corresponding to 3 fb^{-1} of
 1893 integrated luminosity. The $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$, “rare channels”,
 1894 are reconstructed via the K^{*0} decay into a kaon and a pion with opposite charges.

1895 The analysis has to separate signal candidates from background candidates which
 1896 have similar observed properties. The selection presented in Sec. 5.6 aims to max-
 1897 imise the yield while minimising the background contamination. Two types of back-
 1898 grounds are identified: “peaking background” and “combinatorial background”. The
 1899 first comes from the mis-reconstruction of other decays or from partially recon-
 1900 structed events. This type of background, because its specific kinematic properties,
 1901 usually peaks in some variable, such as the invariant mass of all final particles.
 1902 Therefore these candidates can be removed using specific cuts. The combinatorial
 1903 background instead comes from the random combination of particles and can be
 1904 lowered selecting events with good-quality tracks and vertices.

1905 To further reduce the systematic uncertainties the measurement is performed as the
 1906 double ratio

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

1907 where decays reaching the same final states as the rare channels via a J/ψ resonance,
 1908 $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$, also referred to as “charmonium” or “resonant” channels,
 1909 are used as control samples. These decays are distinguished from the rare channels
 1910 using the invariant mass of the dilepton pair.

1911 As new physics is expected not to affect charmonium resonances the ratio of the J/ψ
 1912 channels is 1 and therefore $R'_{K^{*0}} = R_{K^{*0}} \times R_{J/\psi} = R_{K^{*0}}$. On the other hand using
 1913 the relative efficiencies between the rare and resonant channels allows to cancel out
 1914 many effects resulting in a better control of systematic uncertainties. For brevity,

1915 the rare channels will also be denoted as “ $\ell\ell$ ”, or specifically “ ee ” and “ $\mu\mu$ ”, and
1916 the resonant channels as “ $J/\psi(\ell\ell)$ ”, or “ $J/\psi(ee)$ ” and “ $J/\psi(\mu\mu)$ ”.

1917 5.4 Dilepton invariant mass intervals

1918 Three q^2 intervals are considered in this work:

- 1919 • the “low- q^2 ” region between 0.0004 and 1.1 GeV^2/c^4 , where the $b \rightarrow s\ell^+\ell^-$
1920 process is dominated by the photon pole;
- 1921 • the “central- q^2 ” region, [1.1,6.0] GeV^2/c^4 ;
- 1922 • the “high- q^2 ” region, above 15 GeV^2/c^4 .

1923 The central- q^2 region is the most interesting place to look for new physics. In fact,
1924 at low q^2 , below 1 GeV^2/c^4 the photon pole dominates leaving little space for new
1925 physics to be found 1.5.3. The choice of the lower limit of the low- q^2 bin is driven by
1926 the need to reject the background due to the $B^0 \rightarrow K^{*0}\gamma$ decay where the photon
1927 converts into electrons in the material. The lower bound of the central interval is
1928 set at 1.1 GeV^2/c^4 , in order to exclude the contribution from $\phi \rightarrow \ell^+\ell^-$ decays, that
1929 can dilute new physics effects. The upper bound of the central interval is chosen
1930 to be sufficiently far away from the J/ψ radiative tail, where predictions cannot be
1931 cleanly obtained. The 6–15 GeV^2/c^4 region is characterised by the presence of the
1932 narrow peaks of the J/ψ and $\psi(2S)$ resonances. The lower bound of the high- q^2
1933 region, where the signal in the electron channel is still unobserved, is chosen to
1934 be sufficiently far from the $\psi(2S)$ resonance. Rare and normalisation channels are
1935 selected depending on the q^2 interval they fall into (for details see Sec. 5.6).

¹⁹³⁶ 5.4.1 Control channels

¹⁹³⁷ Beyond the normalisation channels, $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-/\mu^+\mu^-)$, extra-control
¹⁹³⁸ channels are used to perform cross-checks and better constrain some of the back-
¹⁹³⁹ ground components in the electron fit. In particular, $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$, also
¹⁹⁴⁰ denoted as “ $\gamma(ee)$ ”, and $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$, also denoted as “ $\psi(2S)(ee)$ ”.
¹⁹⁴¹ All the normalisation and control channels are distinguished depending on the q^2
¹⁹⁴² interval they fall into (for details see Sec. 5.6).

¹⁹⁴³ 5.5 Data samples and simulation

¹⁹⁴⁴ Simulated samples are used to study the background properties, determine efficien-
¹⁹⁴⁵ cies and to train the multivariate analysis. The hard interactions are generated with
¹⁹⁴⁶ Pythia8 hadronic particles are decayed using EvtGen and, finally, propagated into
¹⁹⁴⁷ the detector using Geant4 and reconstructed with the same software used for data.
¹⁹⁴⁸ Samples are generated with both 2011 and 2012, magnet up and down conditions
¹⁹⁴⁹ and are combined in the right proportions, according to the luminosity registered on
¹⁹⁵⁰ data. The next section describes the corrections applied to the simulation to obtain
¹⁹⁵¹ a better description of data.

¹⁹⁵² 5.5.1 Data-simulation corrections

¹⁹⁵³ Since the multivariate classifier training (see Sec. 5.6.6) and the calculation of most
¹⁹⁵⁴ of the efficiency components (see Sec. 5.9) are obtained from the study of simulated
¹⁹⁵⁵ events it is important to verify that the simulation provides a reliable reproduction
¹⁹⁵⁶ the data. In particular it is important to match data and Monte Carlo in the
¹⁹⁵⁷ kinematics of the final particles and the occupancy of the detector. The kinematics
¹⁹⁵⁸ of the decays is characterised by the transverse momentum spectrum of the B^0 .
¹⁹⁵⁹ Discrepancies in this distribution cause also the spectra of the final particles to

1960 differ from data and affect the efficiency determination as its value often depends
 1961 on the momentum of the final particles. The occupancy of the detector is relevant
 1962 as it is correlated to the invariant mass shape of the signal because of the addition
 1963 of energy clusters in the electromagnetic calorimeter, which affects the electron's
 1964 momenta especially when bremsstrahlung photons emitted before the magnet. The
 1965 hit multiplicity in the SPD detector is a proxy for the detector occupancy.

1966 Since it is important that these quantities are well modelled, the simulation is
 1967 reweighted so that the distributions in data and simulation match for these vari-
 1968 ables. The weight is calculated using resonant $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ candidates,
 1969 for which the signal peak is already visible in data after pre-selection (see Sec. 5.6).
 1970 However, the data still includes a high level of background and distributions cannot
 1971 be directly compared. The $s\mathcal{P}$ lot technique [69] is used to statistically subtract the
 1972 background from data and obtain pure signal distributions using the invariant mass
 1973 as control variable. Figure 5.3 shows fits to the 4-body invariant mass of candidates
 1974 after pre-selection. Data and simulation are then compared and the ratio between
 1975 the two distributions is used to re-weight the simulation. The discrepancy in the
 1976 SPD hits multiplicity is solved as a first step and then the B^0 transverse momentum
 1977 distributions are compared between data and simulation reweighted for the SPD
 1978 multiplicity only. Distributions of B^0 transverse momentum and SPD multiplicity
 1979 are reported in Fig. 5.4 and ratios of these distribution, which are used to re-weight
 1980 the simulation, are reported in Fig. 5.5. The weights for the SPD multiplicity are
 1981 calculated separately for 2011 and 2012 events, because distributions are signifi-
 1982 cantly different in the two years. The binnings for these distributions are chosen
 1983 to have approximately the same number of events in each bin to limit fluctuations.
 1984 Further corrections are made by re-weighting the simulation for PID efficiency using
 1985 the PIDCalib package as described in Sec. 5.9.3 and, finally, ee samples are also
 1986 reweighted for L0 trigger efficiency as described in Sec. 5.9.4. Weights are always
 1987 applied throughout unless specified.

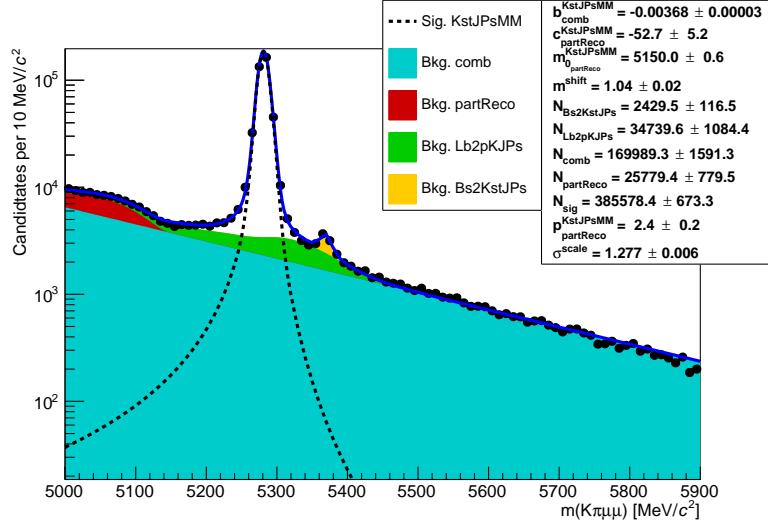


Figure 5.3: Fitted 4-body invariant mass distributions of muonic resonant candidates.

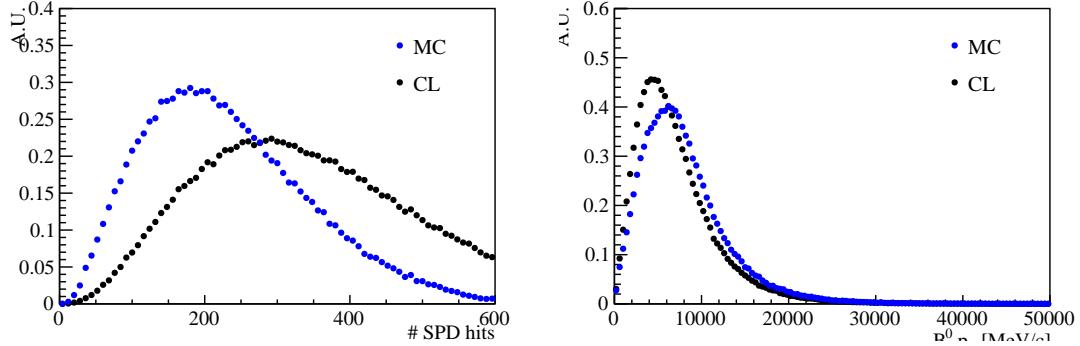


Figure 5.4: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

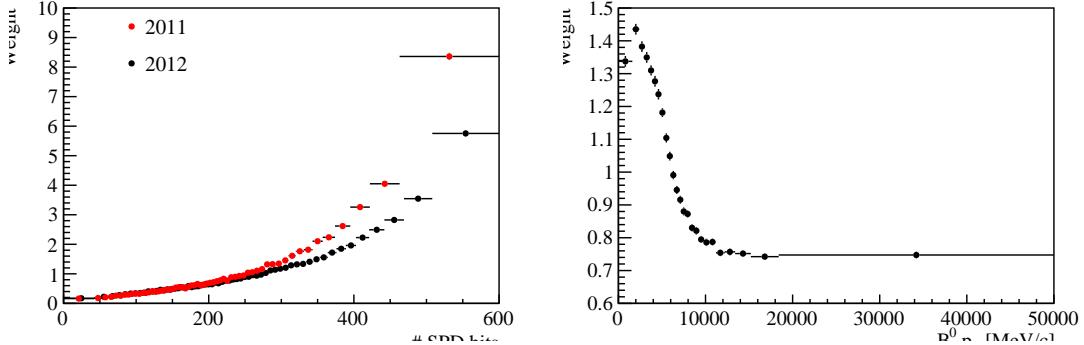


Figure 5.5: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

1988 5.6 Selection

1989 The selection process, described in this section, is divided into several steps:

- 1990 • first of all candidates have to fall into the detector acceptance, produce hits and
1991 be selected on the basis of quality features, such as χ^2 of tracks and vertices
1992 and basic kinematic cuts. This stage is called “stripping”. Furthermore, it
1993 is required that the events are triggered by specific trigger lines and cuts are
1994 applied to remove backgrounds from specific decays. All these first three steps
1995 are referred to as “pre-selection”;
- 1996 • secondly, particle identification requirements are applied to remove part of
1997 misreconstructed background and clear the way for the last step;
- 1998 • in the final step a neural network is used to remove combinatorial background.
1999 Furthermore, for the electron channels, which are more challenging, the kine-
2000 matic structure of the decays is also used to improve the samples purity.

2001 To identify the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ candidates a dilepton mass interval of
2002 $100 \text{ MeV}/c^2$ around the nominal J/ψ peak [2] is selected. On the other hand it
2003 is not possible to use a narrow interval around $J/\psi(ee)$ mass peak as the invari-
2004 ant mass distribution is characterised by a long radiative tail at low masses due
2005 to bremsstrahlung radiation. Furthermore, a requirement in q^2 distorts the 4-body
2006 $m(K\pi ee)$ mass distribution which is not advisable as is important to be able to fit
2007 a wide mass range to constrain the backgrounds. For these reasons the interval to
2008 select $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates is chosen to go as low as possible without
2009 overlapping with the rare channel interval. Candidates are therefore identified as
2010 $J/\psi(ee)$ if they fall in the q^2 interval $6 < q^2 < 11 \text{ GeV}^2/c^4$. Similarly, candidates
2011 are identified as $\psi(2S)(ee)$ if they fall into $11 < q^2 < 15 \text{ GeV}^2/c^4$ and $\gamma(ee)$ if they
2012 fall into $q^2 < 0.004 \text{ GeV}^2/c^4$. Table 5.2 summarises the requirements to distinguish
2013 sample from different channels. Figure 5.6 shows two-dimensional distributions of q^2
2014 versus the 4-body invariant mass for candidates passing the full selection. Horizontal

Table 5.2: Summary of the channel categories.

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

2015 bands can be clearly seen at q^2 values corresponding to the J/ψ and $\psi(2S)$ 2016 resonances. On the plot for muons it is also evident a vertical band which corresponds 2017 to rare decay of interest.

2018 5.6.1 Trigger and Stripping

2019 Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported 2020 in Tab. 5.3, where the logical *and* of L0, HLT1 and HLT2 lines is required and the 2021 logical *or* of the lines on the same level. The candidates are required to be triggered-2022 on-signal (TOS) for most of the stages, namely it is required for the particle which 2023 triggered to be one of the particles used to build the signal candidates. Only for 2024 **L0Global**, used in the electron case, we require a trigger-independent-of-signal (TIS), 2025 this is aimed to collect all the possible statistics for the electron channels, which are 2026 the most challenging. The **L0Muon** trigger requires hits in the muon detector, while 2027 **L0Electron** and **L0Hadron** use information from the calorimeters; **HLT1TrackAllL0** 2028 adds information from the trackers and triggers if the L0 decision is confirmed; 2029 finally, **HLT2Topo[2,3]BodyBBBDT** uses a full reconstruction of the event and a neural 2030 network trained on events with a specific topology in order to detect specific decay

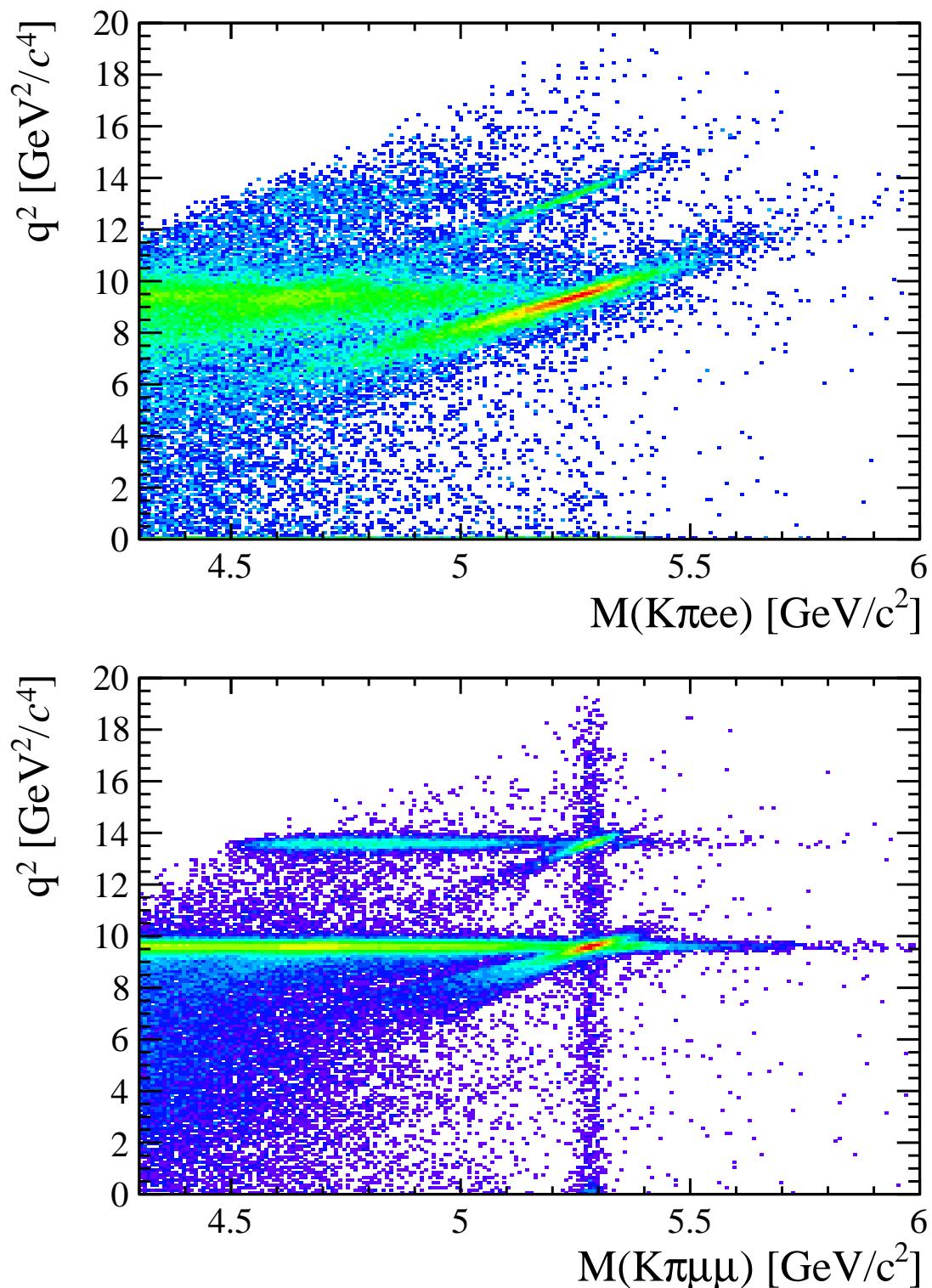


Figure 5.6: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electron (top) and muonic (bottom) channels in 2012 data.

2031 structures.

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

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

2032 For the electron channels the L0 lines have different properties, therefore the analysis
 2033 is performed separately for three categories of events, depending on the L0 trigger
 2034 that fired them. These categories are defined to be exclusive in the following way:

- 2035 • **L0E**: events triggered by at least one of the electrons in the signal candidate
 2036 (`L0Electron_TOS`);
- 2037 • **L0H**: events triggered by at least one of the hadrons in the signal candidate
 2038 and not by `L0Electron` (`L0Hadron_TOS && !L0Electron_TOS`);
- 2039 • **L0I**: events triggered by particles not in the signal candidate and not by the
 2040 previous cases (`L0_TIS && !(L0Electron_TOS || L0Hadron_TOS)`).

2041 The majority of the selected events falls in the L0E category, while the L0H category
 2042 is more efficient at low q^2 were the K^{*0} has more momentum. Because L0I is defined
 2043 to be independent of the signal candidate, the corresponding signal efficiency is the
 2044 same in the rare and resonant cases and cancels out in their ratio.

2045 Candidates are then required to pass the kinematic and quality cuts summarised
 2046 in Tab. 5.4. The meaning of the variables in the table was already explained in
 2047 Sec. 3.4. Loose PID cuts are applied in preselection to limit the size of the samples,
 2048 while tighter cuts are applied in a second stage. A large mass window is kept

Table 5.4: Summary of stripping requirements.

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

around the B^0 peak in order to be able to use the sideband to train the multivariate analysis and to constrain the backgrounds. Track and vertex quality cuts are also applied using the $\chi^2_{\text{track}}/\text{ndf}$, `GhostProb`, and $\chi^2_{\text{vtx}}/\text{ndf}$ variables. The `GhostProb` quantity describes the probability of a track being fake. By construction cutting at 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake tracks. For details about the definition of the variables used see Ref. [119].

5.6.2 PID

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

In the LHCb analysis framework the particle identification probability can be quantified using the “`ProbNN`” variables [120]. These variables are the output of a neural network which takes as input information from the calorimeters, the RICH detectors the muon system and the tracking system. Unlike the DLL variables (see Sec. 2.8) the `ProbNN` are bound from 0 to 1 and can be therefore directly be interpreted as probabilities. For example `ProbNNk` is the probability for a reconstructed particle to be a kaon. Figure 5.7 shows distributions of the correct ID variables in the

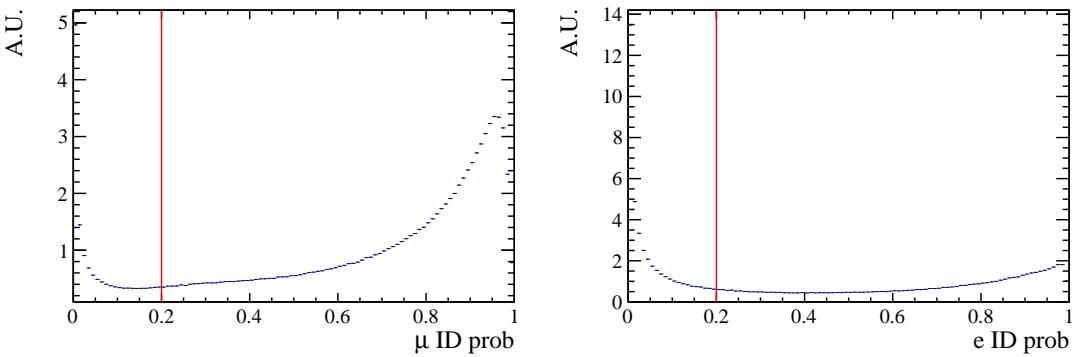


Figure 5.7: Correct ID probability distributions for muons (left) and electron (right) in 2012 data.

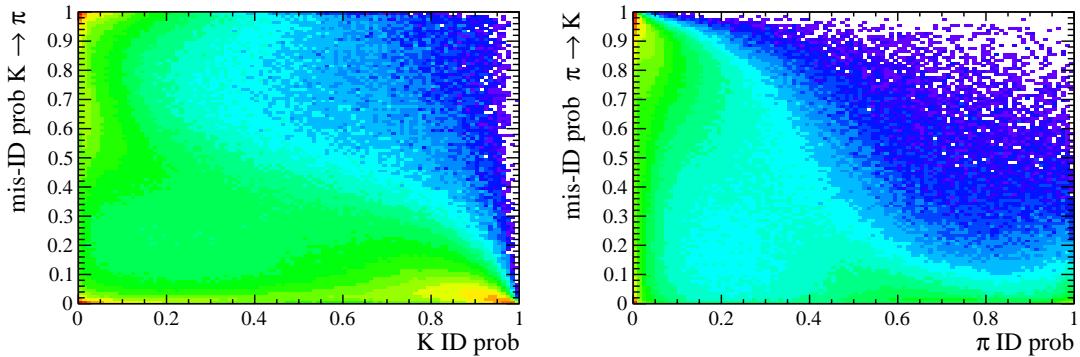


Figure 5.8: On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis show the mis-ID probability.

2065
2066 2012 data sample while Fig. 5.8 shows in a two-dimensional plane the probability
2067 of correct identification and mis-identification of kaons and pions. These plots are
2068 characterised by clear peak at maximal ID probability and minimal mis-ID probabil-
2069 ity, corresponding to particles to which a well defined identification can be assigned.
2070 In order to maximise the power of the PID requirements probabilities of correct ID
2071 and mis-ID are combined using the following cuts:

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

2073 In the first formula, for example, `ProbNNpi` is the probability of correctly identifying
2074 the pion as a pion, while `ProbNNk` is the probability of mistaking it for a kaon. There-
2075 fore by maximising the quantity “`ProbNNpi` \times (1 - `ProbNNk`)”, one can maximise
2076 the correct ID probability and minimise at the same time the mis-ID probability.

2077 5.6.3 Peaking backgrounds

2078 Backgrounds due to specific decays usually peak in some variable because of their
2079 distinctive kinematic properties and therefore they can be removed without sig-
2080 nificant signal efficiency loss. The following sections describe the main sources of
2081 peaking background. The same cuts are applied to the muon and electron channels,
2082 unless specified.

2083 5.6.3.1 Charmonium vetoes

2084 Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2 . The choice of q^2 bin-
2085 ning described in Sec. 5.4 constitutes a natural veto for these decays. Simulated
2086 events were used to check if resonant candidates leak inside the q^2 intervals cho-
2087 sen for the rare channel analysis. For the muonic channels the leakage is negli-
2088 gible as the peaks are sharper due to a better resolution and muons emit fewer
2089 bremsstrahlung photons, resulting in shorter radiative tails. The electronic chan-
2090 nels are instead characterised by a worse resolution and at the same time electrons
2091 can radiate several bremsstrahlung photons, yielding long tails at low q^2 . Analysing
2092 Monte Carlo events it was found that 1.3–2% (depending on the trigger category)
2093 of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candidates leak into the $1.1 < q^2 < 6$ GeV $^2/c^4$ interval

2094 and 1.8% of $\psi(2S)$ events leak above $15 \text{ GeV}^2/c^4$. The contribution from these
 2095 candidates is modelled in the fit.

2096 5.6.3.2 ϕ veto

2097 It can happen that a kaon from the decay $B_s \rightarrow \phi\ell^+\ell^-$, where the ϕ decays in two
 2098 kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as a
 2099 K^{*0} . This results in a candidate with a value of $m(K\pi)$ that is less than the nominal
 2100 K^{*0} mass but still high enough to pass the selection requirements. Figure 5.9 shows
 2101 the plot of $m(K\pi)$ versus $m(K\pi\ell\ell)$, where the kaon mass hypothesis is assigned to
 2102 the pion. A peak can clearly be seen around the ϕ mass ($1020 \text{ MeV}/c^2$). To remove
 2103 this background only candidates with $m_{K(\pi \rightarrow K)} > 1040 \text{ MeV}/c^2$) are selected. This
 2104 results in a 98% background rejection while keeping a 99% signal efficiency. The ϕ
 2105 could also constitute a background when it decays into two leptons but the branching
 2106 ratio of this decay is small compared to the one into kaons and this contribution is
 2107 taken into account by the choice of the q^2 intervals (see Sec. 5.4).

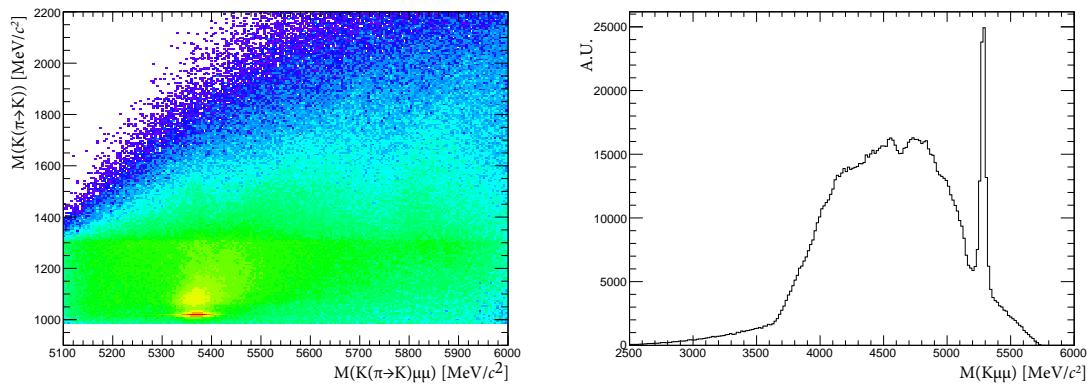


Figure 5.9: (left) Distribution of 2011 data events as a function of the variables $(m_{K(\pi \rightarrow K)})$ and $(m_{K(\pi \rightarrow K)\mu\mu})$, where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. (right) The invariant mass distribution of the three-body system $(K\mu\mu)$,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

2109 5.6.3.3 $B^+ \rightarrow K^+ \ell^+ \ell^-$ plus a random pion

2110 $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays can contaminate the upper B^0 mass sideband if they are com-
2111 bined with a soft pion from somewhere else in the event and therefore reconstructed
2112 as a B^0 decay. Similarly a kaon can be mis-identified as a pion and combined with
2113 an other kaon in the event. Figure 5.9 shows the invariant mass distribution of the
2114 three-body $K\mu^+\mu^-$ system, $m(K\mu\mu)$. This is characterised by a narrow peak at
2115 the B^+ mass. Since these candidates have $m(K\pi\ell\ell) > 5380$ MeV/ c^2 there is no
2116 contribution under the B^0 peak, but they can cause problems when using sidebands
2117 events to train the neural network. An effective veto for this decay was found to
2118 be $\max(m_{K\ell\ell}, m_{(K \rightarrow \pi)\ell\ell}) < 5100$ MeV/ c^2 , which results in 95% background rejection
2119 while keeping 99% signal efficiency.

2120 5.6.3.4 Λ_b decays

2121 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays are unlikely to be reconstructed as $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ because the
2122 Λ is long-lived and decays further in the detector with a separate vertex. How-
2123 ever, simulated events were used to check how many candidates fall into the B^0
2124 samples, which results to be negligible. The $\Lambda_b^0 \rightarrow J/\psi pK$ decay, when the pro-
2125 ton is mis-identified, can instead contribute more easily since the $m(pK)$ is above
2126 the Λ threshold and therefore they must come from Λ^* resonances, which are not
2127 long-lived. This background is already reduced by the PID requirements but a
2128 non-negligible contribution is still expected, which is modelled in the fit.

2129 5.6.3.5 $B^0 \rightarrow (D^- \rightarrow Ke^-\bar{\nu})e^+\nu$

2130 The $B^0 \rightarrow D^- e^+ \nu$ decay, where the D^- in turn decays semileptonically to $K^{*0} e^- \nu$
2131 has the same final particles as the $B^0 \rightarrow K^{*0} e^+ e^-$ decay plus two neutrinos which
2132 are not reconstructed. This decay has a branching ratio four orders of magnitude
2133 larger than $B^0 \rightarrow K^{*0} e^+ e^-$ in the low- q^2 region and it may pass the selection

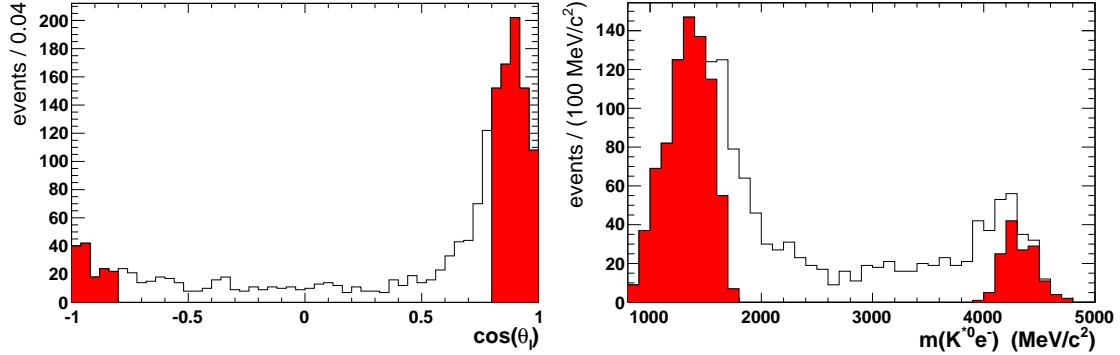


Figure 5.10: Distribution of (left) $\cos(\theta_\ell)$ and of (right) the $m(K^{*0}e^-)$ invariant mass, where the $B^0 \rightarrow (D^- \rightarrow K e^- \bar{\nu}) e^+ \nu$ background is selected by requiring $m(K^{*0}e^+e^-) < 4800 \text{ MeV}/c^2$. The red distribution corresponds to events with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

2134 requirements when the two neutrinos carry a low momentum. To lower the level
 2135 of this background the angle θ_ℓ is used, which is defined as the angle between the
 2136 direction of the e^+ (e^-) in the di-electron rest frame and the direction of the di-
 2137 electron in the B^0 (\bar{B}^0) rest frame. Low momentum neutrinos demand the D^- and
 2138 the e^+ to be almost back-to-back in the B^0 rest frame giving the e^+ a relatively
 2139 large energy compared to the e^- . As a consequence, the direction of the e^+ is close
 2140 to the direction of the di-electron pair, thus the θ_ℓ angle is close to 0. This explains
 2141 why the distribution of background selected in data with an invariant mass cut of
 2142 $m(K^{*0}ee) < 4800 \text{ MeV}/c^2$ is asymmetric towards higher $\cos(\theta_\ell)$ values as it can be
 2143 seen in Fig. 5.10(left). The cut is chosen to be $|\cos(\theta_\ell)| < 0.8$, and is not applied in
 2144 the high- q^2 bin as the variable loses its discriminating power.

2145 In the muon channels the background from $B^0 \rightarrow (D^- \rightarrow K \mu^- \bar{\nu}) \mu^+ \nu$ decays is
 2146 suppressed by the choice of the fitting range.

2147 5.6.3.6 $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$

2148 For the low- q^2 region, a potentially dangerous peaking background is due to the
 2149 $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ decay followed by a conversion of the photon in the detector.
 2150 The branching fraction of $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ has been measured to be $\mathcal{B} = (4.33 \pm$

2151 $0.15) \times 10^{-5}$ and when the photon converts to an electron and a positron has similar
2152 characteristics to $B^0 \rightarrow K^{*0}e^+e^-$. In LHCb around 40% of the photons convert
2153 before the calorimeter. Although only a small fraction of these, $\sim 10\%$, converts in
2154 the VELO and are reconstructed as long tracks, the resulting B^0 mass should peak
2155 under that of the signal, making it a dangerous background. To veto this signal-
2156 like background an effective veto is in the reconstructed invariant mass window for
2157 the e^+e^- -pair that was chosen above $20 \text{ MeV}/c^2$. Furthermore, the e^+e^- -pair from
2158 $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ has a vertex at the point of conversion of the photon, but it
2159 may still be reconstructed as originating from the B^0 decay when the e^+e^- -vertex is
2160 determined with a large error. Therefore a requirement is applied on the uncertainty
2161 of the reconstructed z -coordinate of the e^+e^- -pair: $\sigma_z(e^+e^-) < 30\text{mm}$. iSimulation
2162 is used to predict the contamination from $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ events in the signal
2163 region which is found to be $(3.2 \pm 1.6)\%$.

2164 5.6.3.7 Other peaking backgrounds

2165 Contamination from $B^0 \rightarrow K^{*0}\eta$ and $B^0 \rightarrow K^{*0}\pi^0$ where η and the pion decay
2166 into two photons was considered and found to be small. Furthermore, a potentially
2167 dangerous background could come from events where the identity of the kaon and
2168 the pion are swapped as these candidates peak under the signal. Their contribution
2169 is found to be small, 0.5%, however the effect of their modelling into the fit is
2170 taken into account in the systematic uncertainties. Finally, charmonium decays
2171 where the identity of the kaon, or the pion, and one of the muons are swapped is
2172 a potentially dangerous background. These decays are rejected by requiring that
2173 the hadron- μ invariant mass $m((h \rightarrow \mu)\mu)$, where the muon mass hypothesis is
2174 assigned to the hadron, is not compatible with a J/ψ ($\psi(2S)$) resonance: $|m((h \rightarrow$
2175 $\mu)\mu) - m_{J/\psi,(\psi(2S))}| > 60 \text{ MeV}/c^2$.

2176 5.6.4 Partially-reconstructed background

2177 Partially-reconstructed candidates are defined as decays where one or more particles
 2178 in the final state are not reconstructed, resulting in $m(K\pi\ell\ell)$ values smaller than
 2179 m_{B^0} , but with tails that can still contaminate the signal peak. Sources of mis-
 2180 reconstructed background are decays involving higher hadronic states such as $B^0 \rightarrow$
 2181 $(Y \rightarrow K\pi X)(J/\psi \rightarrow e^+e^-)$, where X represents one or more not reconstructed
 2182 particles. The Y state can be a K^* resonance as well as D mesons that decay
 2183 semileptonically (*e.g.* $B^0 \rightarrow D^-\ell^+\bar{\nu}_\ell$ followed by $D^- \rightarrow K^{*0}\ell^-\nu_\ell$). In case of the
 2184 $J/\psi(ee)$ channel, an additional source of mis-reconstructed background are decays
 2185 of higher $c\bar{c}$ resonances, $B^0 \rightarrow (K^{*0} \rightarrow K\pi)(Y \rightarrow (J/\psi \rightarrow e^+e^-)X)$. To reject
 2186 this backgrounds in the $\mu^+\mu^-$ channels the 4-body invariant mass $m(K\pi\mu\mu)$ is
 2187 recalculated using `DecayTreeFitter` with a vertex constraint. For the resonant
 2188 case this also includes a J/ψ mass constraint to the dilepton pair; in this case the
 2189 4-body mass is denoted as $m(K\pi\mu\mu)_{J/\psi}$. Partially-reconstructed candidates are
 2190 pushed towards low constrained-mass values, resulting in no contamination above
 2191 5150 MeV/ c^2 . This requirement is implicitly applied for the muon channels by the
 2192 definition of the invariant mass fit-windows. The same requirement is also applied
 2193 to select the $J/\psi(ee)$ and $\psi(2S)(ee)$ samples. However, the vertex constrain alone
 2194 is not sufficient to cleanup the electron rare channels. Furthermore, to correctly
 2195 model the long radiative tail of the mass shapes, a fit region that extends down to
 2196 4500 MeV/ c^2 is needed. As a consequence the partially-reconstructed background is
 2197 still relevant for the electron rare channels and extends below the signal peak. For
 2198 this reason this background is modelled in the fit for the rare channels (for details
 2199 see Sec. 5.8.2.2).

2200 5.6.5 Bremsstrahlung corrected mass

2201 An additional handle against backgrounds that contaminate the ee channels is pro-
 2202 vided the analysis of the kinematics of the decay. In fact for the B^0 daughters the

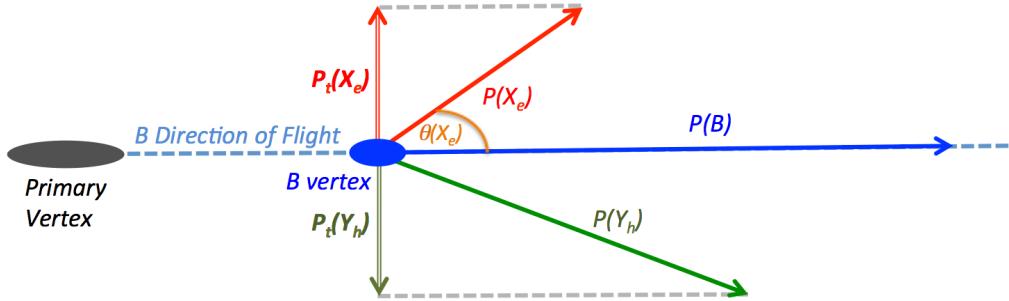


Figure 5.11: Schema of the kinematic of a $B \rightarrow Y_h X_e$ decay, highlighting the quantities relevant for the definition of the bremsstrahlung correction factor, α .

momentum component orthogonal to the flight direction of the B^0 meson should cancel out. The flight direction is defined using the primary and the decay vertices and sketch is shown in Fig. 5.11.

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

$$\alpha = \frac{p_T(K^*)}{p_T(e^+e^-)}.$$

When α deviates from one, some energy is missing in the final state. For signal events, the missing energy is most likely carried away by bremsstrahlung photons emitted by the electrons. Therefore we can use α to correct the electron momentum as

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

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

The corrected momenta can be used to re-calculate the invariant mass of the B^0 candidate, which in the following will be called Bremsstrahlung Corrected Mass (m_{BCM}). The resolution of m_{BCM} depends on the quality of the vertex reconstruction

and on the B^0 lifetime, and degrades as a function of q^2 . Figure 5.12 shows the dependence of the $B^0 \chi_{\text{FD}}^2$ (flight distance χ^2) as a function of m_{BCM} in the considered q^2 regions.

As the correction does not work properly for backgrounds this leads the candidates to spread out making m_{BCM} a discriminating variable between signal and background shapes. A two-dimensional cut is adopted

$$m_{\text{BCM}} > a_{\text{BCM}} + b_{\text{BCM}} \cdot \log(\chi_{\text{FD}}^2)$$

where the a_{BCM} and b_{BCM} coefficients are optimised as described in Sec. 5.6.7.
No cut is applied at high- q^2 nor on the muon channels for which the bremsstrahlung radiation is negligible.

5.6.6 Multivariate analysis

The final selection is performed using a Neural Network classifier (NN) based on the NEUROBAYES package [79, 80]. The multivariate analysis is intended to remove some combinatorial background and obtain a clearer signal peak. In order to avoid biases, a k -fold approach is adopted to train and optimise the classifier, using $k = 10$. This method consists in dividing the samples in k equally sized subsamples; k classifiers are then trained and optimised each on $(k - 1)$ samples and applied to the k th one. This approach ensures that a classifier is never applied to the events used for its training. Each classifier is trained on half of the events included in the $(k - 1)$ samples and optimised using the other half, which ensures that events used for training are not used for optimisation.

Samples:

Representative samples of the signal and background are needed to train the classifier. For the signal, fully reconstructed $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$ sim-

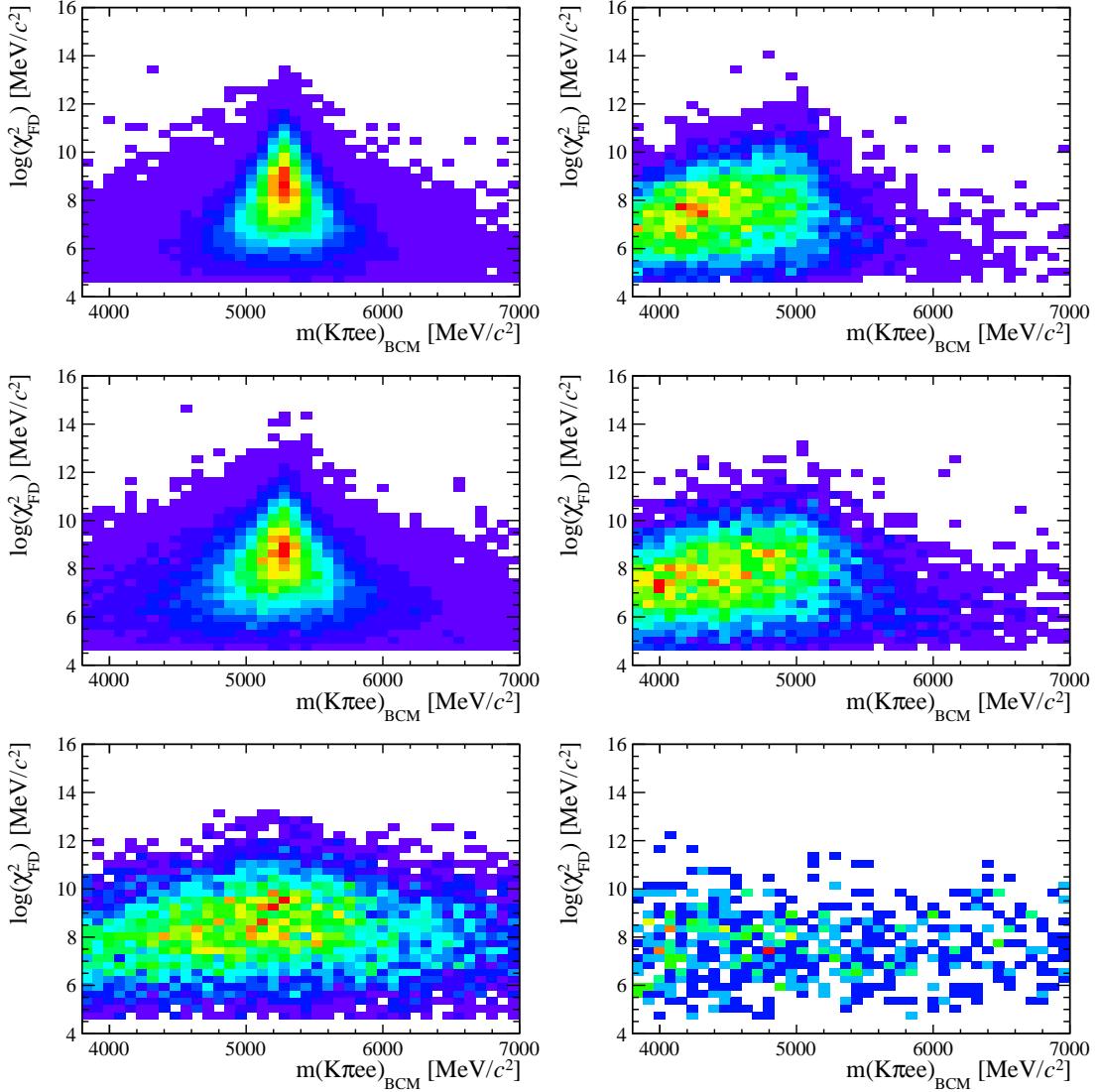


Figure 5.12: Two-dimensional distribution of χ^2_{FD} vs. m_{BCM} for (left) $B^0 \rightarrow K^{*0} e^+ e^-$ signal and (right) partially-reconstructed background. From top to bottom the low-, central- and high- q^2 intervals.

ulated events can be used. Instead a sample representative of the background can be obtained using real data candidates in the upper B^0 sideband: $m(K\pi\mu\mu) > 5400$ MeV/c² and $m(K\pi ee) > 5600$ MeV/c². The lower sideband is not used in the training as it contains a significant fraction of mis-reconstructed background. All pre-selection cuts are applied to the background samples used for the training. As L0 and PID variables are not well described in simulation these cuts are not applied to the simulation but their effect is taken into account by the event weight. An approximately equal number of signal and background events is used for the

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

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

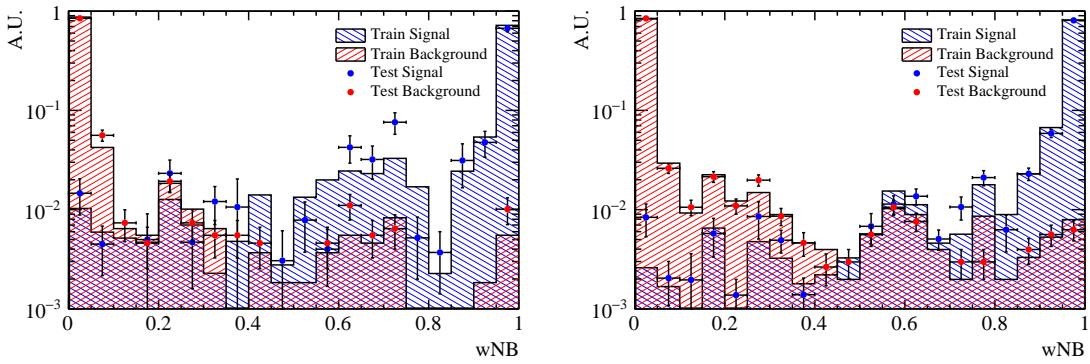


Figure 5.13: NN output distributions for training (solid) and test (stripes) samples, for simulated signal and data sideband events. For the electron (left) and muon (right) training.

2243 training which corresponds to about 1000 events for the electron case and 10,000 for
 2244 the muon one.

2245 Training:

2246 The the neural-network input consists of 24 variables containing information about
 2247 the kinematic of the decays and the quality of tracks and vertices. All the vari-
 2248 ables used are listed in Tab. 5.5. In these figures the variable with ID = 1 is the
 2249 neural-network output and the other IDs are reported in Tab. 5.5. The single most
 2250 discriminating variable used is the χ^2 of a kinematic fit that constrains the decay
 2251 product of the B^0 , the K^{*0} and the dimuon, to originate from their respective ver-
 2252 tices. Other variables that contribute significantly are the χ_{IP}^2 of J/ψ and K^{*0} , the
 2253 transverse momentum of the B^0 and the pointing direction (DIRA) of the recon-
 2254 structed B^0 to the primary vertex.

2255 Figure 5.13 shows neural network output distributions for signal and background.

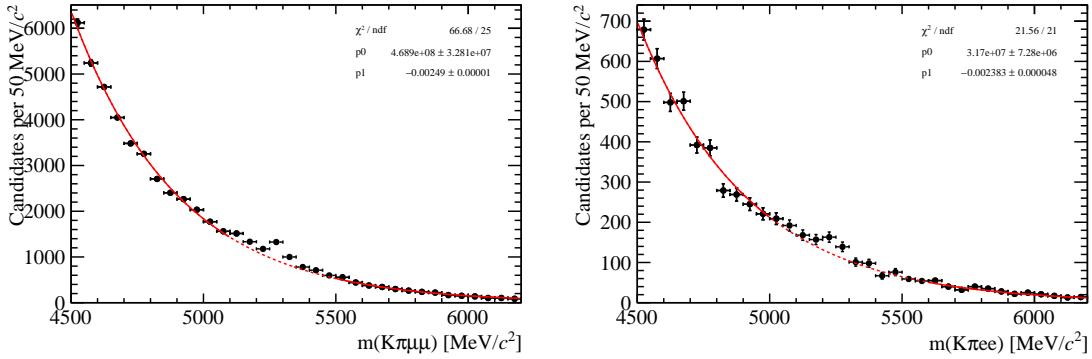


Figure 5.14: Fit to the data sidebands performed to estimate the amount of residual background in the signal mass window for (left) muons and (right) electrons. The region corresponding to the dashed line is excluded from the fit.

2256 On this plot the distributions from the test samples are also overlaid in order to
 2257 check for overtraining. The distributions follow the same shape but with different
 2258 fluctuations indicating no significant overtraining. In general it can be concluded
 2259 that the neural network is able to separate signal from background and that the
 2260 training converged properly.

2261 It can happen that too much information is given to the classifier, which becomes
 2262 able to calculate the invariant mass of the candidates from its inputs. This could
 2263 generate fake peaks and it is therefore important to check for correlations between
 2264 the B^0 mass and the neural-network output. Figure 5.15 reports plots of the average
 2265 neural-network output as a function of the B^0 mass on sideband data and simulated
 2266 signal events. The distributions are flat showing that no significant correlation is
 2267 present.

2268 5.6.7 Optimisation

2269 In order to optimise the requirements on the m_{BCM} and the neural network output
 2270 the expected signal significance, $N_S/\sqrt{N_S + N_B}$, is maximised, where N_S (N_B) is
 2271 number of rare signal (background) candidates. When the BCM requirement is
 2272 applied, the optimisation is performed in a three-dimensional space (t_{MVA} , a_{BCM} ,
 2273 b_{BCM}) where t_{MVA} is the neural-network output threshold below which a candidate

²²⁷⁴ is considered background, and a_{BCM} and b_{BCM} are the parameters of the BCM cut
²²⁷⁵ described in Sec. 5.6.5. Otherwise, only the MVA cut is optimised (for all muons
²²⁷⁶ samples and the high- q^2 electron sample).

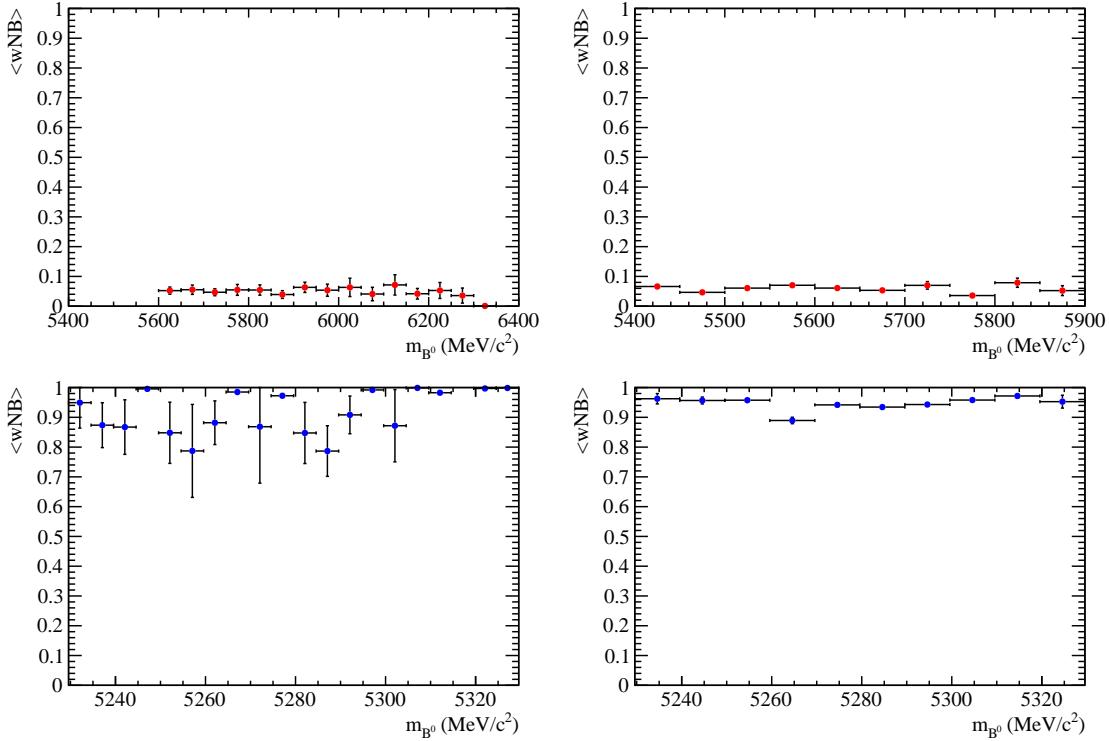


Figure 5.15: Average value of neural-network output as a function of B^0 mass for data sideband (top) and simulated signal (bottom) events for the electron (left) and muon (right) training.

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

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

²²⁷⁷ The number of background events is also derived from data by fitting the back-
²²⁷⁸ ground in the lower- and upper-mass sidebands with an exponential function, and
²²⁷⁹ extrapolating the residual yield in the signal region (Fig. 5.14). Because the back-

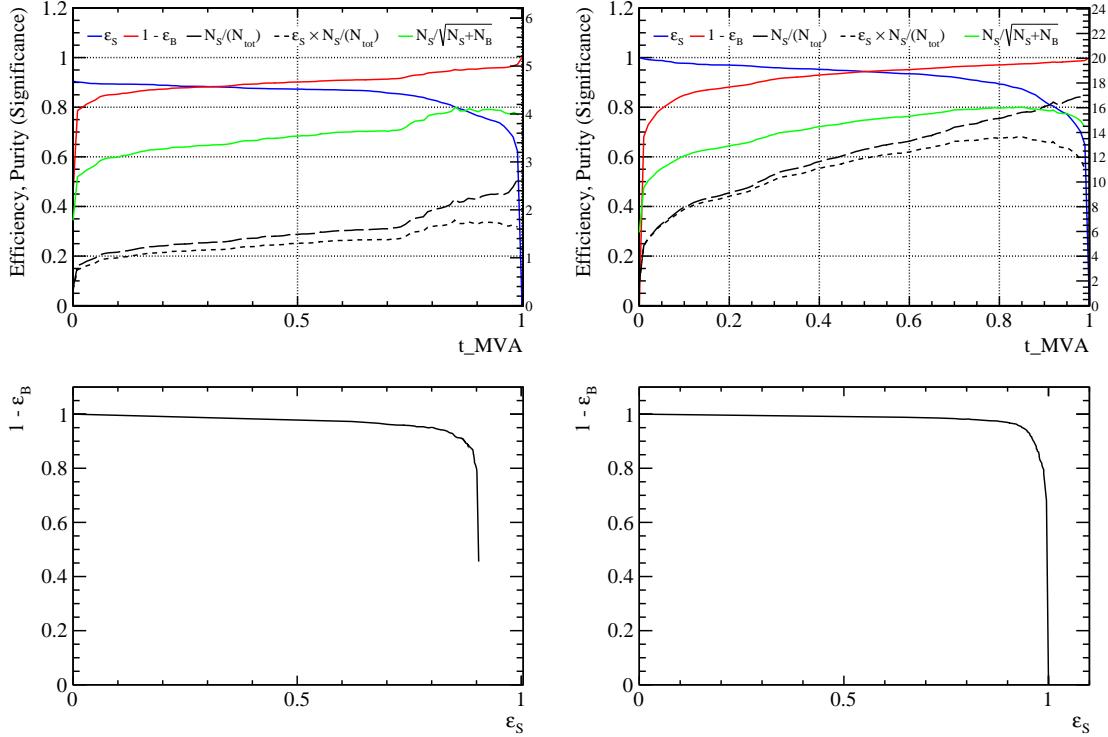


Figure 5.16: (top) Dependence of figure-of-merit on the requirement on neural network output. (bottom) Signal efficiency versus background rejection. Plots correspond to the electron (left) and muons (right) samples.

ground shape changes as a function of the requirement that is being optimised, the sidebands are refitted for each considered cut value.

The cut optimisation is performed in a signal mass window of ± 100 MeV/ c^2 around the nominal B^0 mass for muons, and between 5000 and 5400 MeV/ c^2 for electrons. The average result of the k-fold optimisations is taken as the nominal requirement. The variation of the signal and background efficiency, signal purity and figure-of-merit as a function of the neural-network output requirement for the central- q^2 is shown in Fig. 5.16 together with curves of the background rejection as a function of the signal efficiency. After full selection about $\sim 3\%$ of events still contain multiple candidates which are removed at random keeping only a single candidate per event.

2290 **5.7 Selection summary**

2291 Table 5.6 summarises the requirements applied for each cut after stripping.

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

Type		Requirement	q^2
Quality	All tracks	$\chi^2/\text{ndf} < 3$	all
		$\text{GhostProb} < 0.4$	all
ID	K^{*0}	$ m(K\pi) - m_{K^{*0}}^{\text{PDG}} < 100 \text{ MeV}/c^2$	all
PID	K	$\text{ProbNNk} \cdot (1 - \text{ProbNNp}) > 0.05$	all
	π	$\text{ProbNNpi} \cdot (1 - \text{ProbNNk}) \cdot (1 - \text{ProbNNp}) > 0.1$	all
	μ	$\text{ProbNNmu} > 0.2$	all mm
	e	$\text{ProbNNe} > 0.2$	all e^+e^-
BKG	Swap	$ m((h \rightarrow \mu)\mu) - m_{J/\psi(2S)}^{\text{PDG}} > 60 \text{ MeV}/c^2$	all
	$B^+ \rightarrow K^+\ell^+\ell^-$	$\max(m(K\ell\ell), m((\pi \rightarrow K)\ell\ell)) < 5.1 \text{ GeV}/c^2$	all
	$B_s^0 \rightarrow \phi\ell^+\ell^-$	$m(K(\pi \rightarrow K)) > 1040 \text{ MeV}/c^2$	all
	$B^0 \rightarrow D^-e^+\nu$	$ \cos \theta_\ell < 0.8$	except ee high-
	$B^0 \rightarrow K^{*0}\gamma$	$\sigma_z(e^+e^-) < 30 \text{ mm}$	except $\gamma(ee)$
	Comb	$\text{NNout} > 0.68$	$\mu\mu$ low-
		$\text{NNout} > 0.64$	ee low-
		$\text{NNout} > 0.85$	$\mu\mu$ central-
		$\text{NNout} > 0.97$	ee central-
		$\text{NNout} > 0.40$	$\mu\mu$ high-
		$\text{NNout} > 0.93$	ee high-
		$\text{NNout} > 0.06$	$J/\psi(\mu\mu)$
		$\text{NNout} > 0.20$	$J/\psi(ee)$
	Part-reco	$\text{NNout} > 0.16$	$\gamma(ee)$
		$\text{NNout} > 0.68$	$\psi(2S)(ee)$
Comb, part-reco		$m(K\pi\ell\ell)_{J/\psi} > 5150 \text{ MeV}/c^2$	$J/\psi(ee)$
		$m_{\text{BCM}} > 4680 + 31 \cdot \log(\chi_{\text{FD}}^2)$	ee low-
		$m_{\text{BCM}} > 4437 + 64 \cdot \log(\chi_{\text{FD}}^2)$	ee central-
		$m_{\text{BCM}} > 3380 + 140 \cdot \log(\chi_{\text{FD}}^2)$	$\gamma(ee)$

2292 5.8 Mass fits

2293 The signal yields are extracted using a simultaneous unbinned maximum likelihood
 2294 fit to the 4-body invariant mass, $m(K\pi\ell\ell)$, of the rare and normalisation
 2295 samples. The simultaneous fit allows to share parameters e.g. those describing
 2296 data-simulation differences. The yields of the rare channels are parameterised as a
 2297 function of the corresponding J/ψ yields as

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

2298 where ε^{rel} is the relative efficiency between the rare and resonant channels (given
 2299 in Tab. 5.10). Consequently, $r_{\ell\ell}$ corresponds to the efficiency corrected ratio of the
 2300 raw rare and resonant yields:

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

2301 The two ratios, R_{ee} and $R_{\mu\mu}$, are then used to determine the $R_{K^{*0}}$ quantity, as
 2302 described in Sec. 5.11. The following subsections contain a description of the line
 2303 shapes used to model the signal and background components in each sample.

2304 5.8.1 Muon channels

2305 For the rare and resonant $\mu\mu$ channels the fitted variable is the $m(K\pi\mu\mu)$ invariant
 2306 mass coming from a kinematic fit where all vertices are required to point to
 2307 their mother particle. In the resonant case it is beneficial to also constrain the the
 2308 dimuon mass to the known J/ψ mass; in this case the invariant mass is referred to
 2309 as $m(K\pi\mu\mu)_{J/\psi}$. The effect of the kinematical constraint is to improve the mass
 2310 resolution by roughly a factor of 2, which results in a more stable fit. Furthermore,
 2311 mis-reconstructed background candidates are pushed away from the B^0 peak, which
 2312 allows to use a wider mass window to better constrain the combinatorial background
 2313 slope. The mass spectrum is fitted in the range 5150–5800 MeV/ c^2 with the lower

2314 limit chosen to totally exclude partially reconstructed background. As it is not
2315 needed to model partially reconstructed backgrounds in the fit this also eliminates
2316 the systematic uncertainties associated with the knowledge of their shape.

2317 5.8.1.1 $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ PDF

The signal PDF adopted to describe the reconstructed $m(K\pi\mu\mu)$ invariant mass of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ candidates is the sum of a Double Crystal Ball [104] (DCB) function with opposite-side tails and a Gaussian function with a common mean, μ :

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

2318 where $f_{\text{CB}i}$ is the relative fraction of candidates falling in the i^{th} Crystal Ball func-
2319 tion, σ_i is the width, α_i and n_i are the parameters controlling the power law tail of
2320 each CB, and σ_3 is the width of the Gaussian function.

2321 As a first step, the parameters of the signal PDF are extracted by fitting the
2322 $m(K\pi\mu\mu)$ distribution on $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ simulation and fixed for the
2323 fit to the data. Figure E.1 shows the fitted simulated distribution for the normal-
2324 isation channel, while fits or the rare channel in the three q^2 bins are reported in
2325 Appendix E. In order to account for possible discrepancies in the invariant mass
2326 distribution between data and simulation, the mass is allowed to shift, $\mu \rightarrow \mu + m'$,
2327 and the widths are allowed to scale, $\sigma_i \rightarrow c \cdot \sigma_i$, where the scale factor c is common
2328 between the three σ s.

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

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

2329 where the only free parameters are the mass shift, m' and the width scale factor, c .

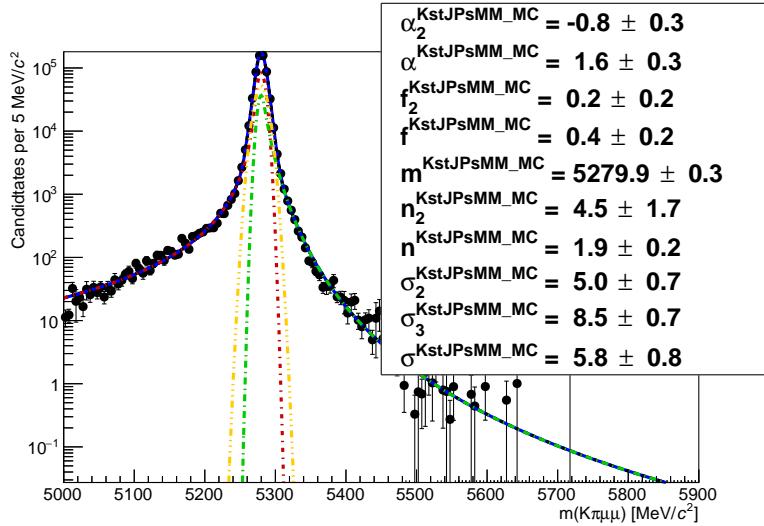


Figure 5.17: Fitted $m(K\pi\mu\mu)$ mass spectrum for $K^{*0}J/\psi$ simulated events.

2330 The following backgrounds are considered:

- 2331 • *Combinatorial*: modelled with an exponential function;
- 2332 • $\Lambda_b^0 \rightarrow pK(J/\psi \rightarrow \mu^+\mu^-)$: described using simulated events to which the
2333 full selection selection and weights for the pK Dalitz plot are applied; this
2334 distribution has a broad shape under the signal peak and is smoothed using
2335 the RooKeysPdf class of the ROOFIT [121] package;
- 2336 • $B_s^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$: described using the same PDF adopted for the
2337 signal, but a different central value, μ , which is set at the B_s^0 nominal mass.
2338 The same shift m' is used as for the signal.

2339 5.8.1.2 $B^0 \rightarrow K^{*0}\mu^+\mu^-$ PDF

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

common factor:

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

where f_{core,q^2} is the relative fraction of candidates falling in the first Crystal Ball function, m'_{q^2} is the mass shift and c_{q^2} is the width scale. The subscript “ q^2 ” indicates that independent parameters are used for each q^2 interval. The background is described by an exponential function in all the three q^2 bins.

5.8.1.3 Summary

In summary, the free parameters of the simultaneous fit to the $J/\psi(\mu\mu)$ and mm candidates are the signal and background yields, the combinatorial background slopes, the mass shifts and the width scales. Figure 5.19 shows the results of the fit to the rare and resonant $\mu\mu$ candidates. Values of the fitted parameters are reported on the plots.

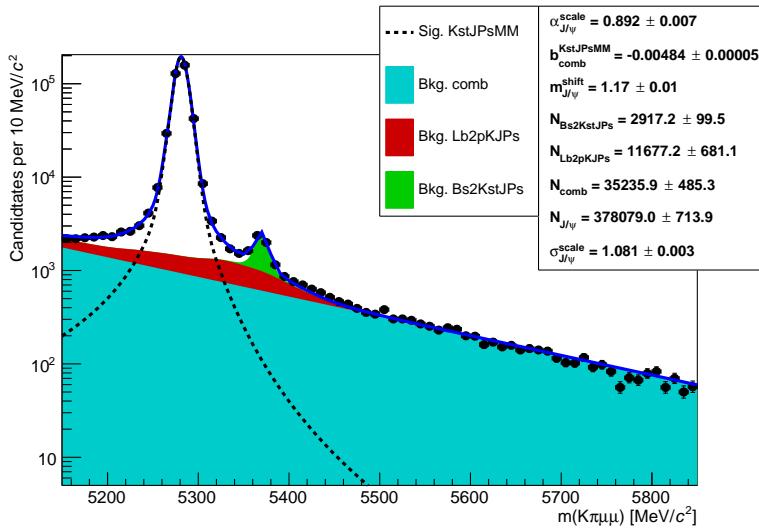


Figure 5.18: Fitted $m(K\pi\mu\mu)_{J/\psi}$ invariant mass distribution for $J/\psi(\mu\mu)$ candidates in logarithmic scale. Dashed black lines represent the signal PDFs and filled shapes the background components.

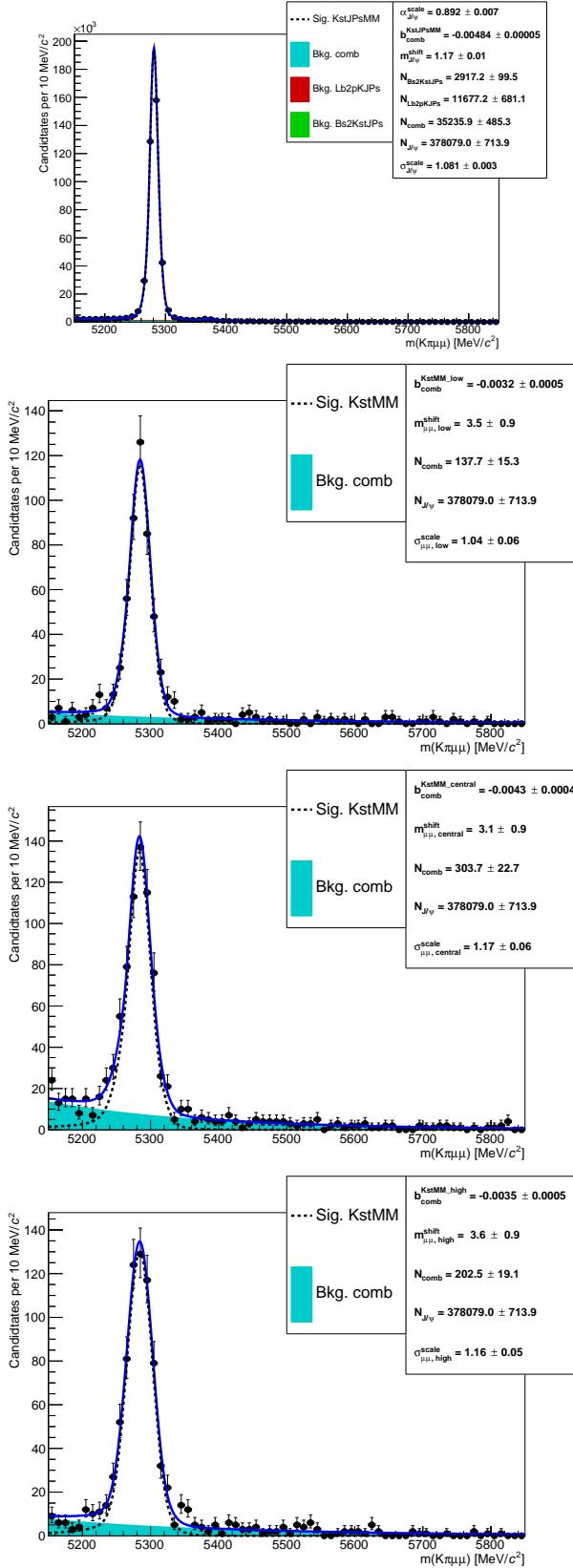


Figure 5.19: From top to bottom fitted $m(K\pi\mu\mu)_{J/\psi}$ invariant mass distributions for $J/\psi(\mu\mu)$ candidates and fitted $m(K\pi\mu\mu)$ distributions for rare candidates in the low-, central- and high- q^2 intervals. Dashed black lines represent the signal PDFs and filled shapes the background components.

²³⁵⁰ 5.8.2 Electron channels

²³⁵¹ Also in the electron case the fitted variable is the 4-body invariant mass coming
²³⁵² from a kinematic fit. In general, this does not include constraints to intermediate
²³⁵³ resonances, unless specified. When constraints to intermediate resonances are ap-
²³⁵⁴ plied the invariant mass is referred to as $m(K\pi ee)_R$, where $R = J/\psi$ or $\psi(2S)$.
²³⁵⁵ The reconstructed invariant mass of the B^0 depends on which L0 line triggered
²³⁵⁶ the event. For this reason, a simultaneous fit to the 4-body invariant mass of the
²³⁵⁷ $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}e^+e^-$ channels in the three trigger categories
²³⁵⁸ is performed. In each trigger category, the $J/\psi(ee)$ and ee yields are extracted from
²³⁵⁹ the following signal channel categories:

- ²³⁶⁰ • $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$, with a J/ψ mass constraint, $m(K\pi ee)_{J/\psi}$;
- ²³⁶¹ • $B^0 \rightarrow K^{*0}e^+e^-$ in the three q^2 intervals.

²³⁶² Extra control channels are fit simultaneously:

- ²³⁶³ • $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ to constrain the yield of partially-reconstructed back-
²³⁶⁴ ground in the low- q^2 and the leakage of $B^0 \rightarrow K^{*0}\gamma$ into the low- q^2 ;
- ²³⁶⁵ • $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$, without the J/ψ mass constraint, to constrain
²³⁶⁶ the leakage into the central- q^2 and the parameters that model residual data-
²³⁶⁷ simulation discrepancies;
- ²³⁶⁸ • $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$, with a $\psi(2S)$ mass constraint, $m(K\pi ee)_{\psi}$, to
²³⁶⁹ constrain the leakage to lower and higher q^2 values.

²³⁷⁰ When fitting the variable without a J/ψ mass constraint it is important to fit a
²³⁷¹ wider mass range to better constrain the parameters modelling the radiative tail
²³⁷² and the backgrounds; a mass window [4500,6200] MeV/ c^2 is used. The lower limit
²³⁷³ is given by the point in which the q^2 cut (at 6 GeV $^2/c^4$ to separate the rare and
²³⁷⁴ resonant channels) starts to affect the 4-body invariant mass distribution.

2375 The invariant mass distributions are different depending on the trigger category
2376 and also on the number of bremsstrahlung photons recovered. Therefore, our sam-
2377 ples are divided in three trigger categories, as described in Sec. 5.6.1, and three
2378 bremsstrahlung categories defined as:

- 2379 • 0γ : candidates with no photon emitted
- 2380 • 1γ : candidates with one photon by either of the electrons
- 2381 • 2γ : candidates with one photon emitted by each electron

2382 All samples are fitted simultaneously, which allows a better use of the available
2383 statistics as the simultaneous fit gathers information from the three categories at
2384 the same time. Furthermore, using this method the results for the three categories
2385 are naturally combined in a single r_{ee} ratio. The PDFs used to fit the invariant mass
2386 distributions are described in the next subsections.

2387 5.8.2.1 Signal PDFs for the electron channels

2388 As for the muon channels, simulated candidates are fitted first to constrain the shape
2389 parameters for the subsequent fit to data. The signal PDFs are built using the
2390 following method:

- 2391 • Simulated $B^0 \rightarrow K^{*0} J/\psi (ee)$ and $B^0 \rightarrow K^{*0} ee$ events are divided in each
2392 trigger and bremsstrahlung category and an independent fit is performed to
2393 each sample. A different fit is also performed for the central, J/ψ and high q^2
2394 intervals. In the case of the high- q^2 interval it is particularly important to keep
2395 signal tail parameters independent from J/ψ channel ones because, as can be
2396 seen in Fig. 5.20, the invariant mass distributions are significantly different for
2397 the two intervals.
- 2398 • For each trigger category a PDF is built as the sum of the three PDFs of the

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

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

2399 bremsstrahlung categories:

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

2400 where the $\mathcal{P}(m)_{n\gamma}^{L0}$ functions are the chosen PDFs for the trigger and bremsstrahlung
 2401 categories and the $f_{n\gamma}^{L0}$ parameters are the relative fractions of events falling
 2402 in each category.

- 2403 • Most parameters are fixed (details later) and the combined PDF, $P(m)$, is
 2404 used to fit real data divided only in trigger categories.

2405 The distribution of the $m(K\pi ee)$ mass in the 0γ category is characterised by a
 2406 sharp tail on the righthand side and is described with a Crystal Ball function (CB),

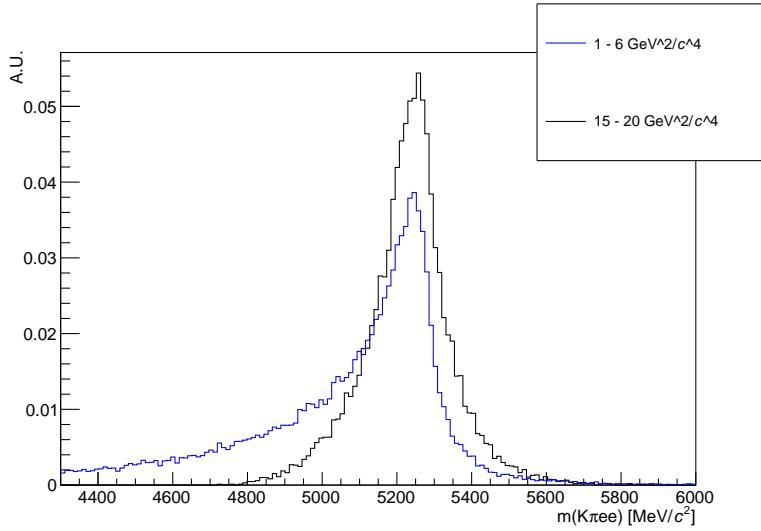


Figure 5.20: Simulated invariant mass of the $K\pi ee$ system in the $1.1 < q^2 < 6$ and $q^2 > 15 \text{ GeV}^2/c^4$ intervals.

while the 1γ and 2γ categories are modelled using the sum of a Crystal Ball and a Gaussian function (CBG) with independent parameters. In all the bremsstrahlung categories the distribution of the 4-body invariant mass with a mass with the J/ψ mass constraint is modelled using the sum of a DCB and a Gaussian functions as done in the muon fit. To account for possible data-simulation discrepancies, the mass (widths) of each trigger PDF is allowed to shift (scale), similarly to the muon channels. However, due to the larger background contamination these parameters are shared between the rare and the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ control sample (no J/ψ mass constraint). The tail parameters are similar between the $J/\psi(ee)$ and the central- q^2 but this is not the case at high- q^2 , as can be seen in Fig. 5.20, due to the migration of candidates in the tail to lower reconstructed q^2 . For this reason the initial parameters for each candidate type are obtained fitting a simulated sample of the same candidate type.

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

²⁴²⁴ In summary the signal PDF for the fit on data is defined as:

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

²⁴²⁵ where the free parameters are: c , the scaling factor for the widths, and m' , the mass
²⁴²⁶ shift.

²⁴²⁷ 5.8.2.2 Background PDFs for the electron channels

²⁴²⁸ This section reports the background components considered for each fitted sample.

²⁴²⁹ $B^0 \rightarrow K^{*0} e^+ e^-$ low- q^2

²⁴³⁰ • *Combinatorial*: described using an exponential function; the yield and slope
²⁴³¹ parameters are free to vary in the fit;

²⁴³¹ • *Partially-reconstructed* (hadronic): the shape is obtained from a $K_1^+(1270)$ simulated samples smoothed with a `RooKeysPdf`. Simulated distributions are shown in Fig. 5.21. The fraction of partially-reconstructed candidates with respect to signal ones is expected to be very similar to that in $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+ e^-)$ and therefore the normalisation is fixed as:

$$N_{e^+ e^-, \text{low}}^{\text{part-reco}} = N_{e^+ e^-} \cdot \frac{N_{\gamma(ee)}^{\text{part-reco}}}{N_{\gamma(ee)}},$$

²⁴³² where $N_{\gamma(ee)}^{\text{part-reco(hadronic)}} / N_{\gamma(ee)}$ is the fraction of the hadronic partially-reconstructed
²⁴³³ background relative to the signal yield in the $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+ e^-)$ channel;

²⁴³³ • $B^0 \rightarrow K^{*0} \gamma$ leakage: the leakage from the $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+ e^-)$ decay in the low- q^2 region is modelled using a simulated candidates that pass the low- q^2 requirements: the distribution is smoothed using a `RooKeysPdf`; the

normalisation is fixed to the $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ yield, $N_{\gamma(ee)}$ as:

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

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

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

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

$$N_{e^+e^-, \text{central}}^{\text{leak}} = N_{J/\psi ee} \cdot f_{J/\psi ee}^{\text{leak, MC}},$$

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

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

- *Combinatorial*: modelled using a shape obtained by reversing the NN output cut on data, which has the effect of selecting background candidates instead of signal ones. Figure 5.22 shows the invariant mass distributions for different anti-cuts on the electron and muon samples at high- q^2 . The shapes are very similar between the two samples and as a function of the cut value. In order

2449 to have a larger statistics, the shape is taken from the muon sample with a
 2450 tight NN output anti-cut at 0.1 and smoothed with a `RooKeysPdf`;

- 2451 • *Partially-reconstructed* (hadronic): modelled using simulation as described for
 2452 the previous intervals; the normalisation is left free to vary.
- $B^0 \rightarrow K^{*0}\psi(2S)$ leakage: the leakage from the $\psi(2S)$ radiative tail is modelled
 using simulated $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ events in the high- q^2 region. The
 normalisation is fixed to the $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ yield, $N_{\psi(2S)(ee)}$ as:

$$N_{e^+e^-, \text{high}}^{\text{leak}} = N_{\psi(2S)(ee)} \cdot f_{\psi(2S)(ee)}^{\text{leak, MC}},$$

2453 where $f_{\psi(2S)(ee)}^{\text{leak, MC}}$ is the fraction of $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ simulated candi-
 2454 dates leaking in the high- q^2 interval.

2455 $B^0 \rightarrow K^{*0}\gamma$

- 2456 • *Combinatorial*: described using an exponential function; the yield and slope
 2457 parameters are free to vary in the fit;
- *Partially-reconstructed* (hadronic): modelled using simulation as described for
 the previous intervals; the normalisation is left free to vary and used to con-
 strain the fraction of partially-reconstructed candidates in the low- q^2 interval.
- $B^0 \rightarrow K^{*0}e^+e^-$ leakage: as the $K^{*0}\gamma$ was added to the low- q^2 also the low- q^2
 leakage is added to $K^*\gamma$. The yield is constrained to the N_{ee}^{low} yield.

2463 $B^0 \rightarrow K^{*0}J/\psi$ and $B^0 \rightarrow K^{*0}\psi(2S)$

2464 The following backgrounds are considered for the fits to the invariant mass of $B^0 \rightarrow$
 2465 $K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates:

- 2466 • *Combinatorial*: described using an exponential function. The yield and slope
 2467 parameters are free to vary in the fit;

- 2468 • $\Lambda_b^0 \rightarrow pK(J/\psi \rightarrow e^+e^-)$: described using simulated events to which the full
 2469 selection is applied. This distribution has a broad shape under the signal peak
 2470 and is smoothed using a `RooKeysPdf`. The normalisation is constrained to
 2471 the $\Lambda_b^0 \rightarrow pK(J/\psi \rightarrow \mu^+\mu^-)$ yield returned by the $\mu\mu$ fit after correcting for
 2472 efficiency differences between final states with muons and electrons.
- 2473 • $B_s^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$: described using the same PDF adopted for the signal,
 2474 but a different central value, m_0 , which is set at the B_s^0 nominal mass. The
 2475 normalisation is constrained to the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \mu^+\mu^-)$ yield returned
 2476 by the $\mu\mu$ fit after correcting for efficiency differences between final states with
 2477 muons and electrons;

2478 The J/ψ mass constraint has the effect of pushing the partially-reconstructed back-
 2479 ground away from the peak outside the fit window. The J/ψ control sample is se-
 2480 lected using the requirement that the 4-body mass constrained using `DecaTreeFitter`
 2481 is above 5150 MeV/ c , which explicitly removes the partially-reconstructed back-
 2482 ground; this cut does not produce significant distortion to the unconstrained in-
 2483 variant mass distribution in the considered window. For these reasons this back-
 2484 ground does not need to be modelled in either of these cases. For the fit to
 2485 $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$, which includes a $\psi(2S)$ mass constraint, only the com-
 2486 binatorial background is considered and described using an exponential function.

2487

2488 5.8.2.3 Summary of the fit to the electron samples

2489 In summary, the free parameters in the fit to data are:

- 2490 • the $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$, $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}(\gamma \rightarrow$
 2491 $e^+e^-)$ yields in each trigger category;
- 2492 • the r_{ee} ratio common to all trigger categories; one for the low, one for the
 2493 central- and one for the high- q^2 region;

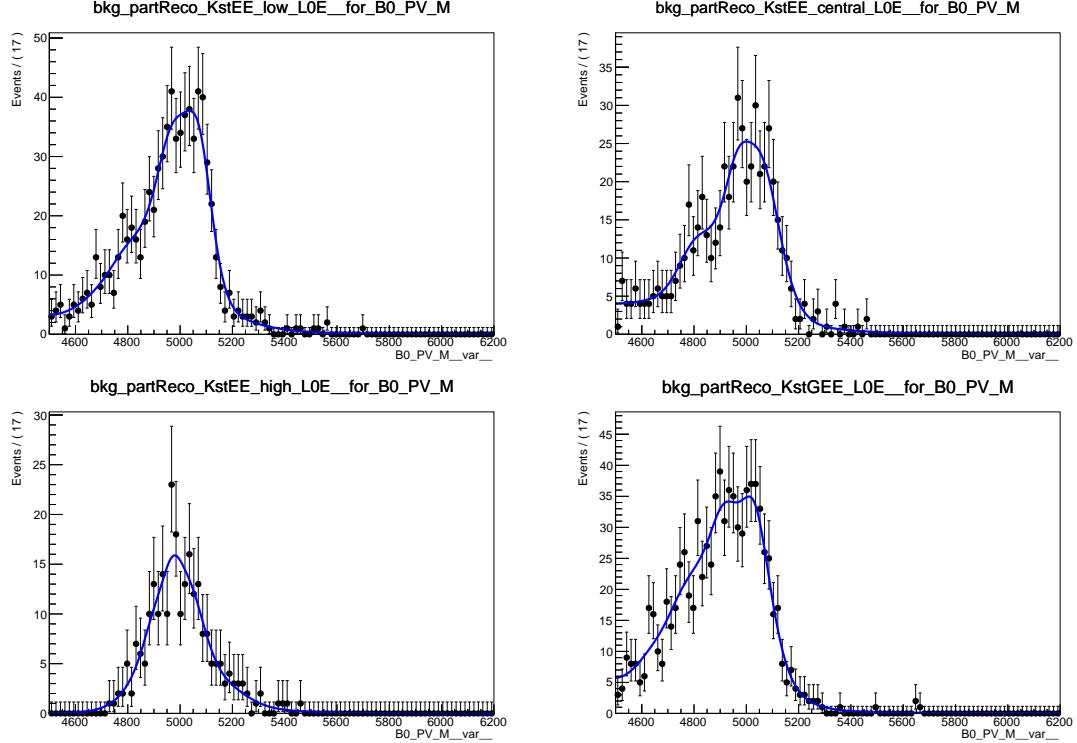


Figure 5.21: Distributions of the $m(K\pi ee)$ invariant mass of decays involving higher K^{*0} resonances for the (top left) low- (top right) central-, (bottom left) high- q^2 intervals and (bottom right) $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$.

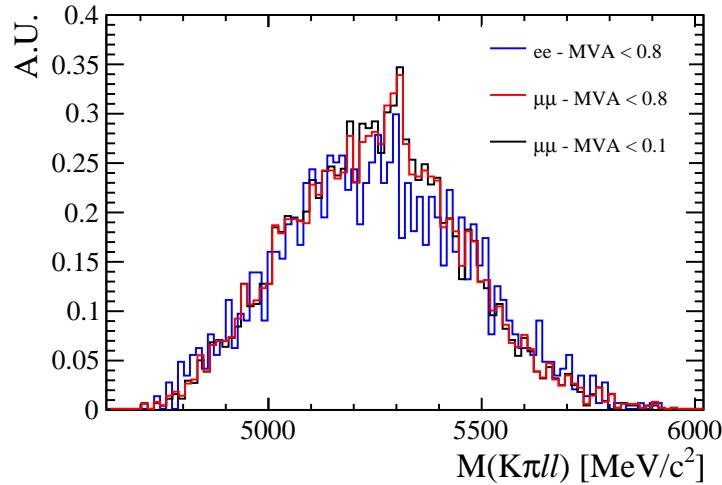


Figure 5.22: Distributions of the $m(K\pi\ell\ell)$ invariant mass for $B^0 \rightarrow K^{*0}\ell^+\ell^-$ candidates selected with a reversed cuts on the NN output.

- 2494 • one mass shift, m' , and one width scale factor, c , for the signal PDF common
 2495 between $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}e^+e^-$, but different for the

2496 three trigger categories and for $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}(\gamma \rightarrow$
2497 $e^+e^-)$;

- 2498 • the yield and slope, when applicable (e.g. no slope at high- q^2), of the combinatorial background in each trigger category and for each channel;
- 2500 • the yield of the backgrounds when not fixed as described in the previous section.

2502 Fits to simulated $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates are shown in Appendix E,
2503 while fits to real candidates are shown in Fig. 5.23 for the normalisation channel, in
2504 Fig. 5.24 for the rare channel and in Fig. 5.25 for the control channels. For simplicity
2505 the latter two figures show the sub of the three trigger categories, while the separate
2506 plots are reported in Appendix F, where fitted parameters are also reported on the
2507 plots. In the high- q^2 interval, above $15 \text{ GeV}^2/c^4$, the efficiency for the L0Hadron
2508 trigger becomes very low as the K^* has very low momentum. In this region only
2509 9 candidates are found spread in the interval $4500 < m(K\pi ee) < 6000 \text{ MeV}/c^2$.
2510 Therefore only L0E and L0I triggered events are fitted in this region.

2511 5.8.3 Event yields

2512 Table 5.8 reports raw yields obtained from the fits described in the previous subsec-
2513 tions. The values for the rare channels are not directly floating in the fits but, as
2514 described in Sec. 5.8, they are parameterised as a function of the number of resonant
2515 events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare branching
2516 fractions.

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

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

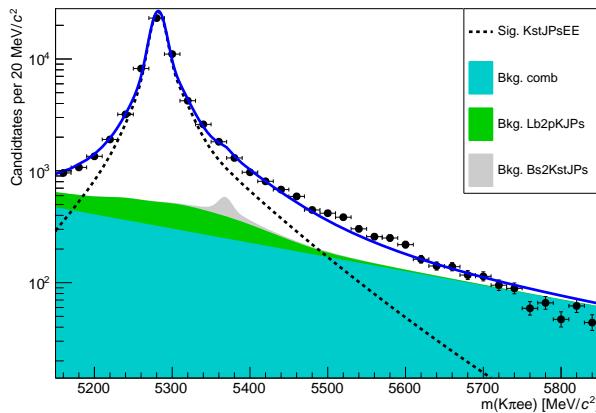


Figure 5.23: Fit to the mass constrained $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates. The dashed black line (shaded shapes) represents the signal (background) PDF.

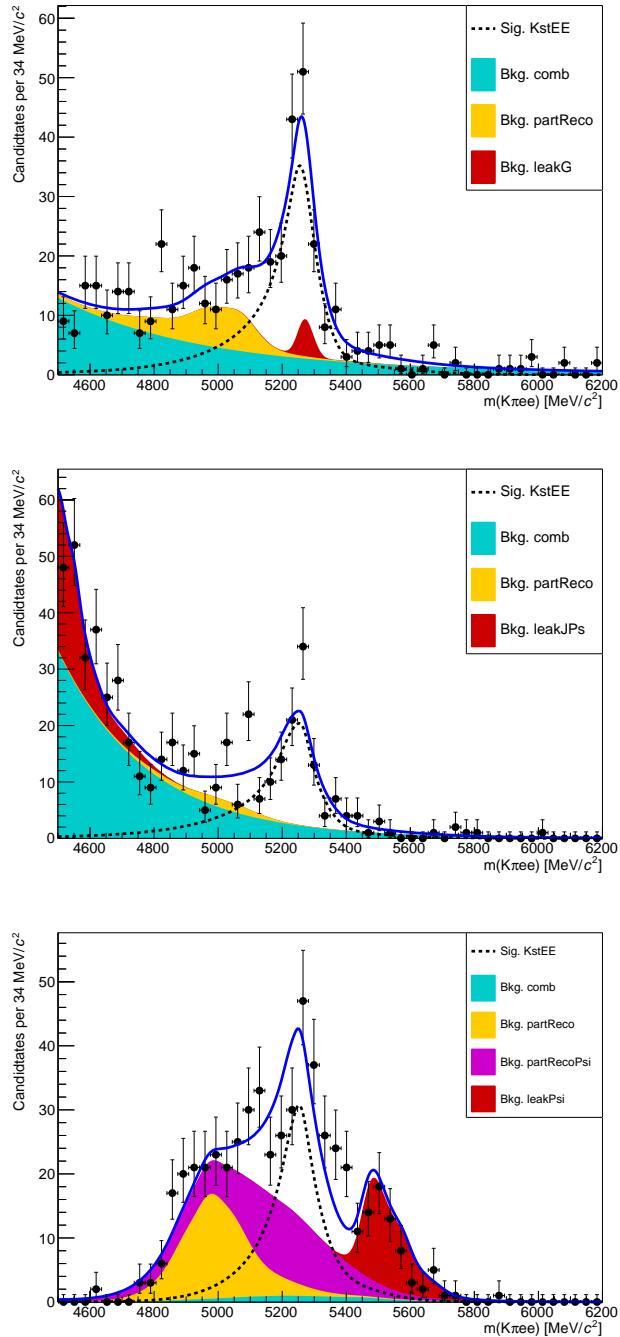


Figure 5.24: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0} e^+ e^-$ candidates. From top to bottom for the low-, central- and high- q^2 intervals. The dashed black line (shaded shapes) represents the signal (background) PDF.

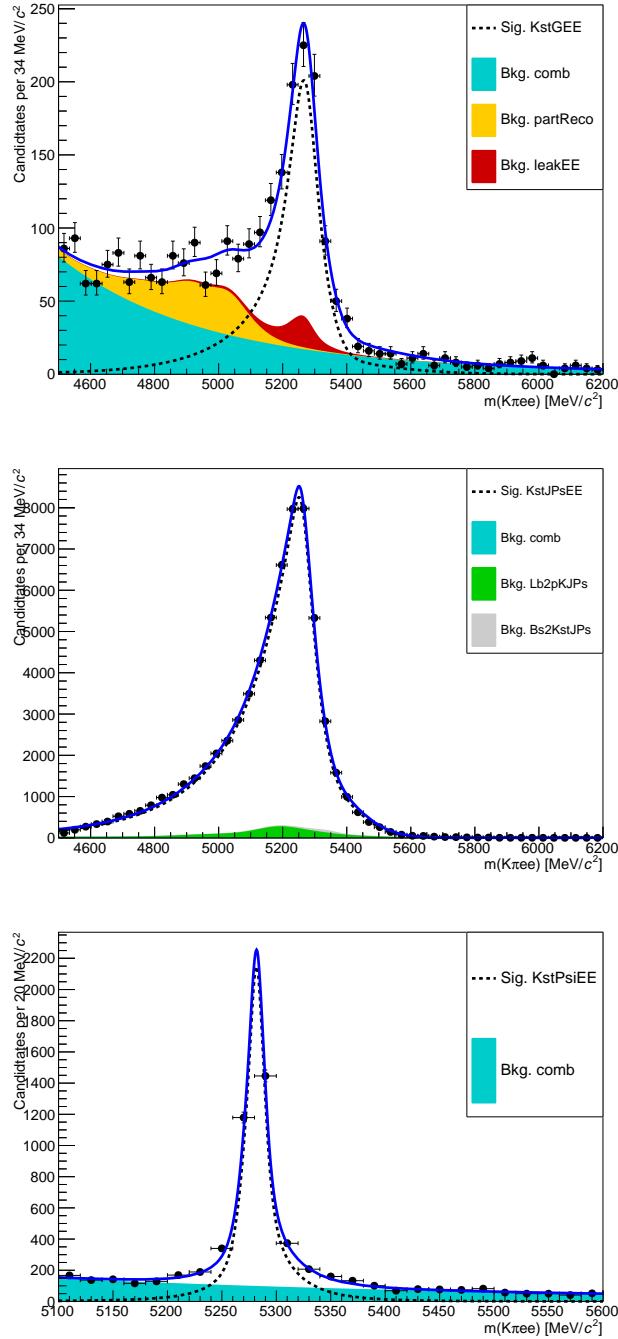


Figure 5.25: Fit to the $m(K\pi ee)$ invariant mass of control channel candidates. From top to bottom: invariant mass distribution without mass constraint of $B^0 \rightarrow K^{*0}(\gamma \rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates and mass constrained mass of $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ candidates. The dashed black line (shaded shapes) represents the signal (background) PDF.

2517 **5.9 Efficiency**

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

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

2518, where the first term is the efficiency to have final state particles in the LHCb
2519 detector acceptance; the second term ($\varepsilon^{\text{reco|geom}}$) carries information on reconstruc-
2520 tion and stripping efficiency; the third ($\varepsilon^{\text{PID|reco}}$) corresponds to the efficiency of
2521 the PID requirements; the fourth ($\varepsilon^{\text{trig|PID}}$) handles the trigger efficiency for those
2522 events which are selected by the pre-selection process; and, finally, the latter term
2523 deals with the efficiency of the neural network classifier. Reconstruction, trigger,
2524 MVA and BCM efficiencies are evaluated using simulated data samples with the
2525 trigger efficiency for $B^0 \rightarrow K^* J/\psi$ being cross-checked using the data-driven TIS-
2526 TOS method as described in Sec. 3.6.3. The PID efficiency is calculated with a
2527 data-driven method as described in Sec. 5.9.3.

2528 Absolute efficiencies for the muon and electron normalisation channel are reported in
2529 Tab. 5.9 and relative efficiencies between the rare and resonant channel, $\varepsilon(\ell\ell)/\varepsilon(J/\psi(\ell\ell))$,
2530 are listed in Tab. 5.10; these are the efficiencies which are used in the fit.

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

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

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

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

2531 5.9.1 Geometric efficiency

2532 In order to save disk space, simulated samples only contain decays with final daugh-
 2533 ters in the LHCb detector acceptance, which can therefore be reconstructed. This
 2534 corresponds to the requirement for each of the final particles to have polar angle θ
 2535 between 10 and 400 mrad. The efficiency of this cuts is obtained using a generator
 2536 level simulated sample.

2537 5.9.2 Reconstruction efficiency and bin migration

2538 The reconstruction efficiency is here defined as the efficiency to reconstruct each
 2539 decay channel given that its daughters are into the geometrical acceptance of the
 2540 detector. This includes both the probability that the final particles generate ob-
 2541 servable signatures and the efficiency of all the pre-selection requirements described
 2542 in Sec. 5.6, including those done to remove peaking backgrounds. The efficiency of
 2543 the PID cuts is kept separate as it is known to be not well simulated and there are
 2544 reliable data-driven methods which can be used to extract it (see Sec. 5.9.3).

2545 5.9.2.1 Bin migration

2546 It can happen that events generated in a q^2 interval are reconstructed in a different
 2547 one, this is referred to as “bin migration” and can be due to two different effects.
 2548 First of all, as the resolution of real detectors is not perfect, events close to the edges
 2549 of the considered intervals can fall on the wrong side of the edge. This effect is only
 2550 important in case of non-flat true distributions, as the amount of bin migration in the
 2551 two directions is different. The second possible source of bin migration are systematic
 2552 effects due, for example, to the presence of bremsstrahlung photons that cannot be
 2553 recovered. It is particularly important to take into account the bin migration in the
 2554 electron channels case because more photons are radiated from the final state and
 2555 the mass resolution is worse. Figure 5.26 shows the response matrix for simulated

2556 $B^0 \rightarrow K^{*0} e^+ e^-$ events, which represents the correlation between reconstructed and
2557 generated q^2 . In the ideal case of perfect resolution this plot would look like a
2558 diagonal line and in case no bias is present its slope would be 1. Table 5.11 lists the
2559 net amounts of bin migration, M_{net} , in the considered q^2 intervals defined as:

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

2560 where $N(\text{in} \rightarrow \text{in})$ is the number of candidates that are generated and reconstructed
2561 inside the considered interval, $N(\text{out} \rightarrow \text{in})$ the number of candidates that are gen-
2562 erated outside the interval but reconstructed inside and $N(\text{in} \rightarrow \text{out})$ the number of
2563 candidates generated inside that fall outside. The reconstruction efficiency is calcu-
2564 lated comparing generated to reconstructed samples and therefore already includes
2565 bin migration effects. Nevertheless, it is useful to single out this component to better
 asses the corresponding systematic uncertainty.

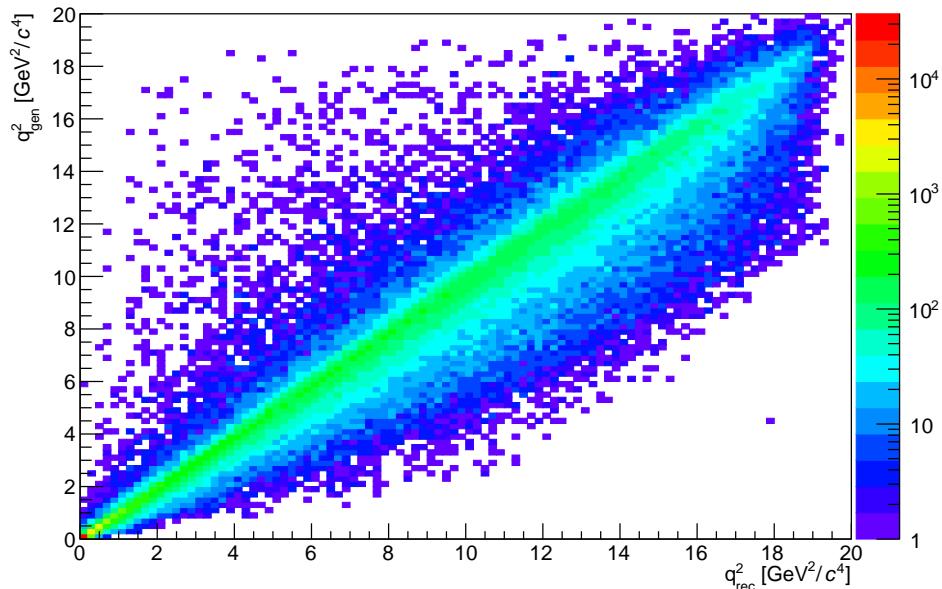


Figure 5.26: Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^{*0} e^+ e^-$ events.

2567 5.9.3 PID efficiency

2568 The simulation is known not to reliably describe particle ID variables and therefore
 2569 a data-driven method is used to obtain this efficiency component. This is done using
 2570 the **PIDCalib** package described in Sec. 2.8.1. Furthermore, the same method is used
 2571 to weight the simulation in order to calculate the MVA and trigger efficiencies. The
 2572 package **PIDCalib** allows to divide the phase-space in intervals of quantities relevant
 2573 for the determination of the PID efficiency and obtain a data-driven efficiency for
 2574 each interval. For this analysis the phase-space is divided in equi-populated bins
 2575 of momentum and pseudorapidity of the particle under study. Figure 5.27 shows
 2576 performance tables for pions, kaons, muons and electrons. Once the efficiency tables
 2577 are obtained for each particle, the total efficiency is calculated for each candidate
 2578 as the product of the four final particles efficiencies. $\varepsilon^{ev} = \varepsilon_K \cdot \varepsilon_\pi \cdot \varepsilon_{\ell_1} \cdot \varepsilon_{\ell_2}$. Finally,
 2579 as the decay channel under study generally has different kinematical distributions
 2580 than the calibration sample, the total efficiency is found by averaging over simulated
 2581 events.

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

2582

2583 5.9.4 Trigger efficiency

2584 While the trigger efficiency for the muon channels is calculated using simulated
 2585 events, for the electron channels a combination of simulation and data-driven meth-
 2586 ods is used. The efficiency of the software stage, HLT, is always obtained from

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

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

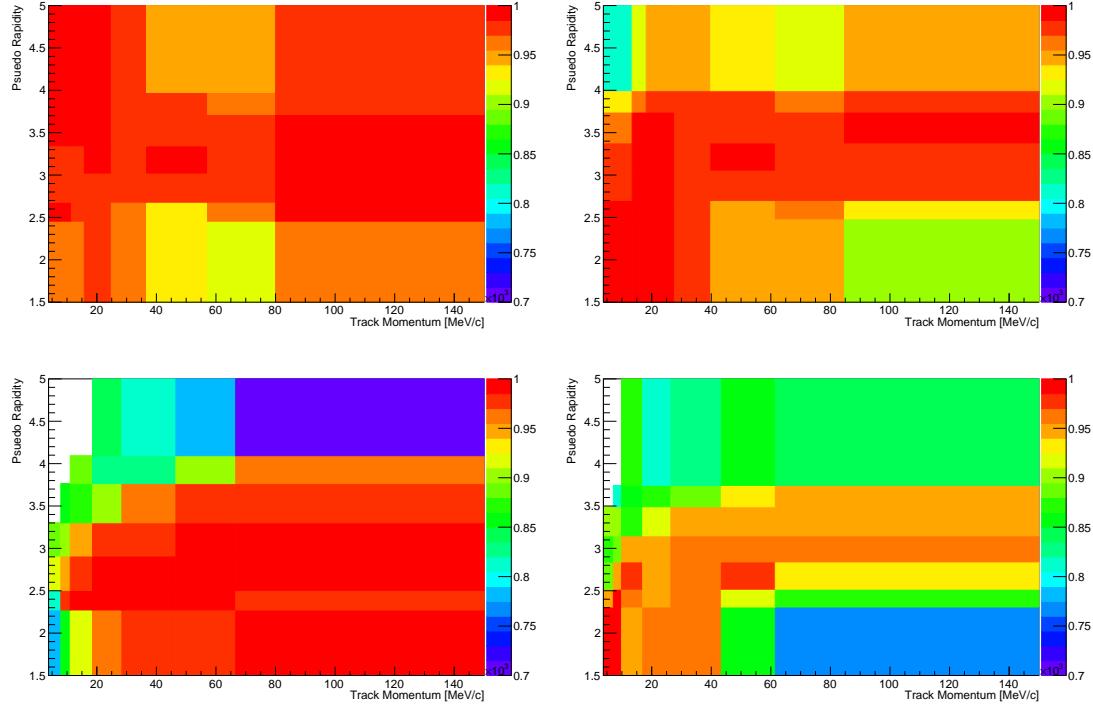


Figure 5.27: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

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

5.9.4.1 Electron triggers

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

Tables of efficiencies are obtained applying the TISTOS method to a calibration sample. For each trigger category these tables contain the efficiency as a function of p_T of the considered particle and are given for different calorimeter regions as these have different properties (e.g. cell size) due to the different position with respect to the beam line. The considered regions are the inner and outer HCAL, and the inner, middle and outer ECAL. Figure 5.28 shows data-driven efficiencies for the L0Electron trigger in the three ECAL regions.

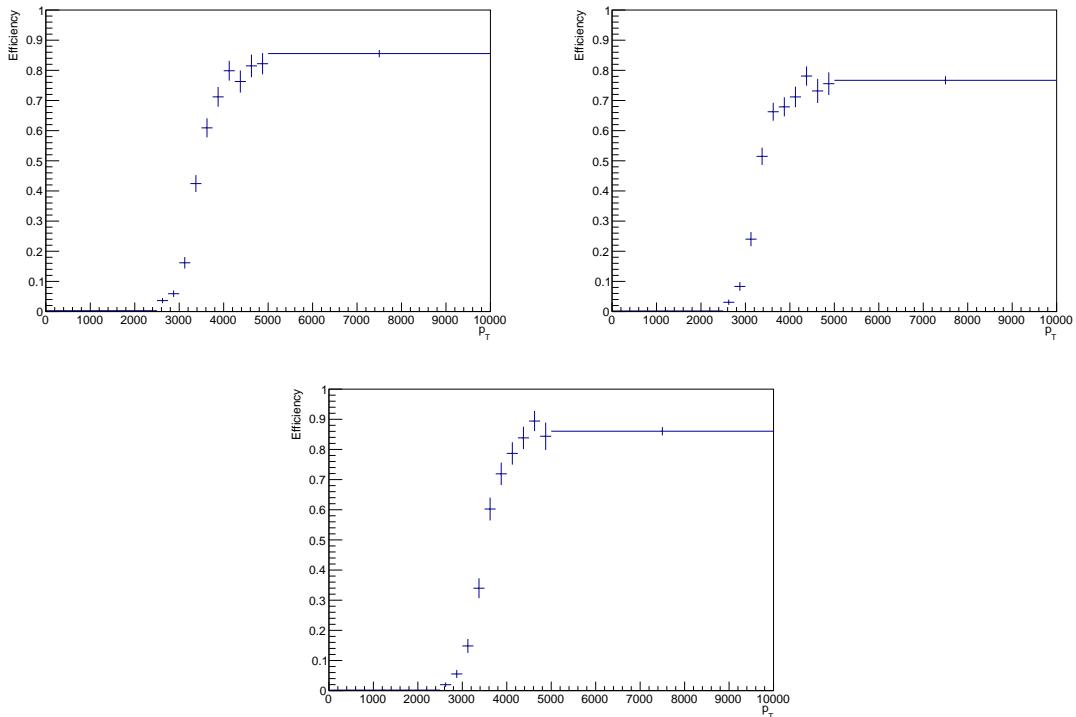


Figure 5.28: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

2606

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

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

2607 The probability of TIS trigger is defined to be independent of the signal and therefore
2608 must be the same in the rare and resonant channels and cancel in their ratio.

2609 Then event by event efficiencies for the three trigger categories are defined to be
2610 exclusive in the following way:

- 2611 • L0E: $\varepsilon^{L0E} = P_{L0E}$, namely the probability that at least one electron triggered;
- 2612 • L0H: $\varepsilon^{L0H} = P_{L0H} \cdot (1 - P_{L0E})$, namely the probability that at least one hadron
2613 triggered but none of the electrons;
- 2614 • L0I: $\varepsilon^{L0I} = (1 - P_{L0H}) \cdot (1 - P_{L0E})$, namely the probability that neither the
2615 hadrons or the electrons in the candidate triggered. Note that in this case ε^{L0I}
2616 does not correspond to the efficiency of TIS trigger but to the probability that
2617 the event does not fall into the L0E or L0H categories.

2618 Finally, as in the PID case, the total efficiency is found averaging over all events of
2619 a simulated sample:

$$\varepsilon^{\text{trg}} = \frac{1}{N} \sum_i^N \varepsilon^{\text{trg}}(p_T^i) \quad (5.16)$$

2620 where “trg” is a label indicating the trigger category under consideration.

2621 5.9.4.2 TISTOS cross-check

2622 The efficiency obtained using the simulation is cross-checked applying the TISTOS
2623 method, already described in Sec. 3.6.3, to resonant data. For this purpose a sample
2624 of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$ candidates triggered independent-of-signal (TIS) is used
2625 as control sample. As data also contains non negligible amounts of background
2626 a narrow interval around the peak, dominated by the signal, is selected and the
2627 *sPlot* method is used to remove residual background in the data sample. Results
2628 are shown in Tab. 5.12, where the efficiency obtained using the TISTOS method is
2629 compared between data and simulation. These are found to be in agreement for the
2630 muon channel, while they show deviations in the electron channels. In particular
2631 a significant discrepancy is found, for the L0I category, for which the procedure
2632 explained in Sec. 5.9.4.1 does not ensure a correct calibration. The table also reports

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

Sample	MC	Data	Correction factor
$J/\psi \rightarrow \mu\mu$	0.797 ± 0.002	0.803 ± 0.004	1.0073
$J/\psi \rightarrow ee$ LOE	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

2633 a correction factor obtained according to the formula

$$f = 1 + \frac{\varepsilon_{data}^{\text{TISTOS}} - \varepsilon_{MC}^{\text{TISTOS}}}{\varepsilon_{MC}^{\text{TISTOS}}}, \quad (5.17)$$

2634 which can be used to correct the absolute resonant yields. On the other hand, even
2635 though discrepancies are present, they should cancel out in the ratio between the
2636 rare and J/ψ channels; only the residual discrepancy on this ratio is relevant for
2637 the measurement of $R_{K^{*0}}$. In order to check if discrepancies cancel out we need to
2638 obtain a data-driven efficiency also for the rare channels. To do this the TISTOS
2639 efficiency obtained on $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+ \ell^-)$ candidates must be reweighted for
2640 the difference in the kinematics between the rare and resonant channels. This is
2641 done by determining the TISTOS efficiency as function of the maximum p_T of the
2642 particles that fired L0 (the leptons for `L0Electron` and `L0Muon`, the kaon and the
2643 pion for `L0Hadron`, and all final state particles for `L0Global`). Results are shown
2644 in Fig. 5.29 and used to re-weight the distribution of rare simulated candidates.
2645 The ratios $\varepsilon_{\ell\ell}^{\text{tistos}} / \varepsilon_{J/\psi}^{\text{tistos}}$ obtained using the data-driven method and simulation are
2646 compared and found to be fully compatible. This means that, even though the
2647 TISTOS correction has an effect on the absolute efficiency of each channel, this
2648 becomes negligible on their ratio. Therefore, no correction due to this effect is
2649 applied for the calculation of the $R_{K^{*0}}$ ratio.

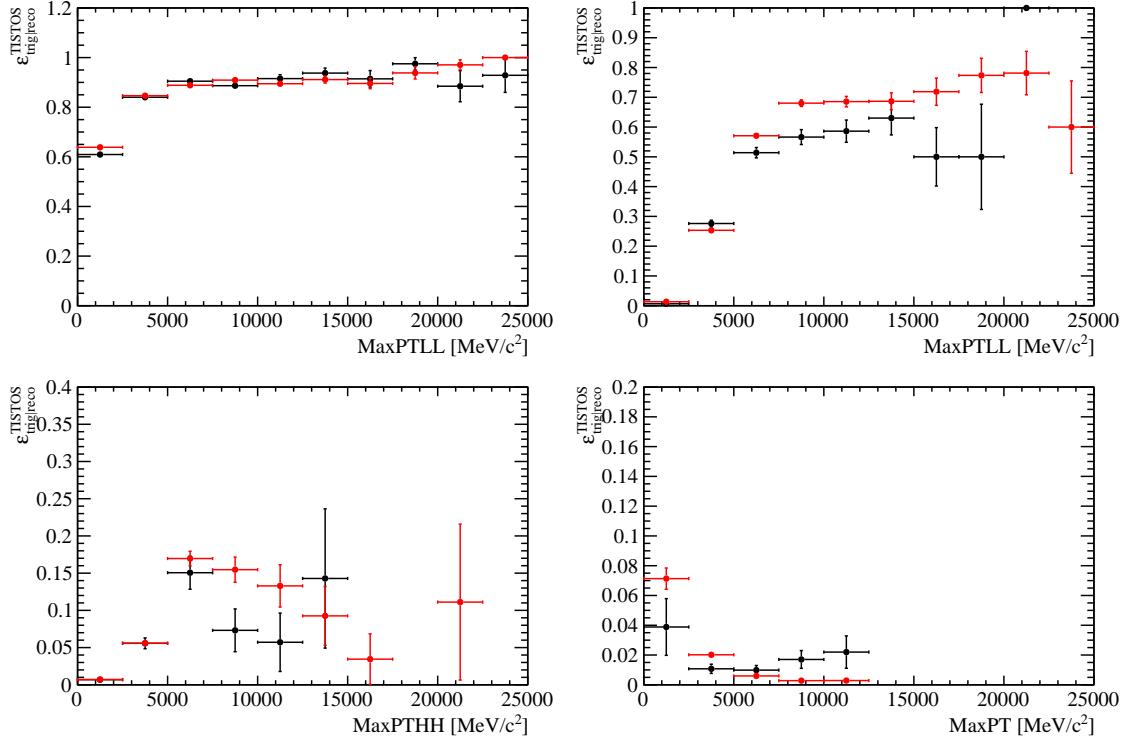


Figure 5.29: Trigger efficiency obtained applying the TISTOS method on $B^0 \rightarrow J/\psi K^*$ candidates as a function of the maximum p_T of the two muons (top left) or the two electrons for the L0E category (top right), the maximum p_T of p and π for L0H (bottom left) and the maximum p_T of all the final particles for L0I (bottom right).

2650 5.9.5 Neural networks and BCM efficiencies

2651 The neural network and BCM efficiencies are evaluated from fully weighted
 2652 simulated samples, and separately for each trigger category for the electron channels.
 2653 In order to check for biases one can compare the efficiency obtained using $B^0 \rightarrow$
 2654 $K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ events and rare $B^0 \rightarrow K^{*0}\ell^+\ell^-$ events in the same q^2 region
 2655 selected for the resonant case. The ratio between the two should be close to unity
 2656 with small deviations due the fact that the q^2 interval width is finite and the events
 2657 are distributed differently inside the interval. This ratio is found to be 0.997 ± 0.004
 2658 for the $\mu\mu$ channels and on average 0.981 ± 0.005 for the ee channels. Values for the
 2659 electron channels show a small deviation from one due to the very large q^2 interval
 2660 used to select the resonant channel ($6\text{--}11$ GeV $^2/c^4$).

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

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

2661 5.10 Systematic uncertainties

2662 This section describes the main sources of systematic uncertainties considered. Other
 2663 sources, which would matter in measurements of absolute quantities, cancel in the
 2664 ratio between the rare and resonant channels. A list of the systematic uncertainties
 2665 that are considered and their effect on the $R_{K^{*0}}$ ratio is summarised in Tab. 5.13.
 2666 The total uncertainty is evaluated by summing in quadrature the single components.

2667 5.10.1 Choice of signal and background PDFs

2668 There is a certain arbitrariness in the choice of PDFs to model signal and background
 2669 contributions in the invariant mass fits, which could translate in a bias on the final
 2670 result. The systematic uncertainty due to the parameterisation of line shapes is
 2671 studied in the following ways.

2673 For the signal PDF:

- 2674 • *Shape*: in the electron channels the PDF is changed from a Crystal Ball and
2675 Gaussian to a Double Crystal Ball. Modifying the PDF has a negligible effect
2676 in the muon modes, while it affects the electron ones. Furthermore the data-
2677 simulation discrepancy parameters (m' and c) are constrained using the $B^0 \rightarrow$
2678 $K^{*0}(\gamma \rightarrow e^+e^-)$ sample instead of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+e^-)$.
- 2679 • *Bremsstrahlung categories*: gaussian constraints are applied to the relative
2680 fractions of the bremsstrahlung categories, instead of fixing them to the values
2681 observed on simulation. This yields a $\sim \%$ systematic on $R_{K^{*0}}$ in the central-
2682 and high- q^2 region.

2683 For the background PDFs:

- 2684 • *Swaps*: a component that describes candidates where the particle identities are
2685 swapped is added both to the muon and electron resonant fits, and constrained
2686 to the number of candidates expected from simulation. This amounts to a $\sim \%$
2687 variation on $R_{K^{*0}}$ in the central- and high- q^2 region.
- 2688 • $\Lambda_b^0 \rightarrow pK J/\psi (\rightarrow e^+e^-)$: the normalisation is left free to vary. This results in
2689 a $\sim \%$ variation on $R_{K^{*0}}$ in the central- and high- q^2 region.
- 2690 • *Partially-reconstructed*: the yield of the mis-reconstructed background to $B^0 \rightarrow$
2691 $K^{*0}e^+e^-$ is left free to vary in the fit. This only applies to the central- q^2 in-
2692 terval as this contribution is already free to vary in the high- q^2 range. This
2693 yields a $\sim \%$ systematic on $R_{K^{*0}}$.
- 2694 • *Combinatorial*: the PDF at high- q^2 is changed from an exponential (anti-MVA
2695 cut) to an anti-MVA cut (exponential) for the $\mu\mu$ (ee) mode. This amounts
2696 to a $\sim \%$ variation on $R_{K^{*0}}$ in the central- and high- q^2 region.
- 2697 • $\Lambda_b^0 \rightarrow pK\ell^+\ell^-$: this background is added to the fit to the rare channel and
2698 returns zero yield for both the muon and the electron samples. Therefore this
2699 yields no systematic uncertainty.

- 2700 • *Leakage*: gaussian constraints are applied to the amounts of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow$
 2701 $e^+e^-)$ leakage in the central- q^2 region and to the $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$
 2702 leakage in the high- q^2 region, which are fixed in the default fit. This results in
 2703 a $\sim \%$ variation on $R_{K^{*0}}$ in the central- and high- q^2 region.

2704 Furthermore in all case where a simulated sample is used and smoothed to obtain a
 2705 PDF the kernel of the density estimation is varied by ± 0.1 from the value used in
 2706 the nominal fit.

2707 5.10.2 Efficiency determinations

2708 The statistical uncertainty on the efficiency determinations is taken as the corre-
 2709 sponding systematic uncertainty. The correlation among the electron trigger cate-
 2710 gories is taken into account (e.g. L0E and L0H are anti-correlated). A further source
 2711 of systematic uncertainty associated to the trigger efficiency is estimated using the
 2712 data-simulation differences observed in Sec. 5.9.4.2. Ratios of efficiencies for the
 2713 rare to resonant decays are found to be compatible between the electron and muon
 2714 modes, indicating that the effect on $R_{K^{*0}}$ is negligible, but the statistical precision
 2715 on the determinations is taken as an extra systematic uncertainty.

2716 5.10.3 Bin migration

2717 The determination of the reconstruction efficiency is affected by the knowledge of
 2718 the amount of bin migration as explained in Sec. 5.9.2. This amount depends on
 2719 the shape of the q^2 distribution, which in turn depends on the simulated $B^0 \rightarrow$
 2720 $K^{*0}e^+e^-$ decay model. In order to asses this systematic, simulated samples are
 2721 generated using different models corresponding to different form factors [122, 123].
 2722 The q^2 distributions obtained using each model are compared with the ones obtained
 2723 using the default one [124]. Figure 5.30 shows normalised ratios between these
 2724 q^2 distributions and the default one, which are used to re-weight the simulation.

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

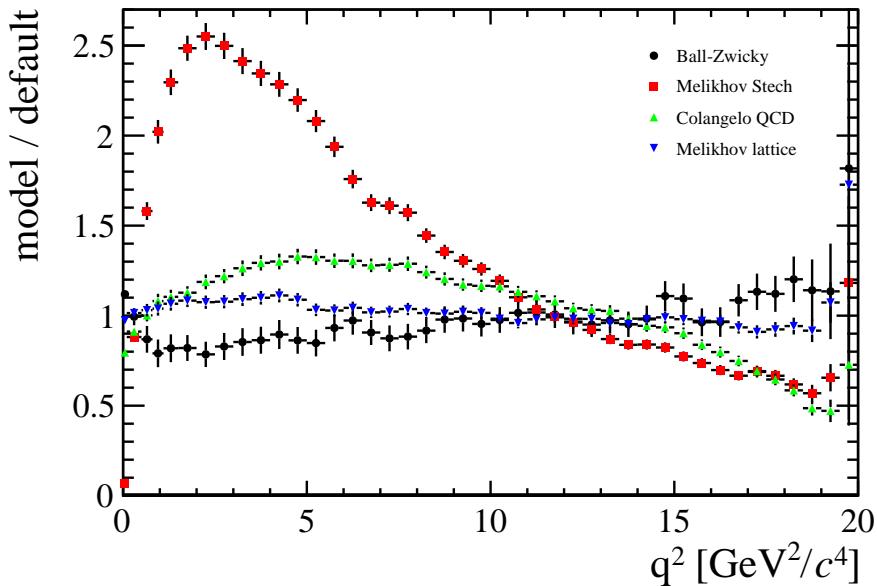


Figure 5.30: Ratios between the q^2 distributions obtained using different form factors models with respect to the default model.

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

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

2729

2730 5.11 Result extraction

2731 This section presents the procedure to obtain the $R_{K^{*0}}$ ratio together with methods
2732 to check robustness of the analysis.

²⁷³³ **5.11.1 $R_{J/\psi}$ sanity check**

²⁷³⁴ In order to cross-check the analysis procedure, the ratio between the measured
²⁷³⁵ branching ratio of the electron and muon resonant channels is calculated:

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

²⁷³⁶ Unlike absolute branching fractions calculations, the determination of $R_{J/\psi}$ repre-
²⁷³⁷ sents a better sanity test as it is not affected by uncertainties due to the knowledge
²⁷³⁸ of the amount of collected luminosity, \mathcal{L} , or the fragmentation fraction, f_d , the
²⁷³⁹ probability for a b quark to produce a B^0 meson. These quantities come with large
²⁷⁴⁰ uncertainties but they cancel in the $r_{J/\psi}$ ratio.

²⁷⁴¹ **5.11.2 $R_{K^{*0}}$**

²⁷⁴² The ratio $R_{K^{*0}}$ is extracted by dividing the r_{ee} and $r_{\mu\mu}$ parameters described in
²⁷⁴³ Sec. 5.8. These ratios are direct parameters of the fit but they can also be built from
²⁷⁴⁴ the yields in Tab. 5.8 and the efficiencies in Tab. 5.10. In summary the definition of
²⁷⁴⁵ the $R_{K^{*0}}$ ratio is the following:

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

²⁷⁴⁶ As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three
²⁷⁴⁷ electron categories its value is already a combination of the three samples.

2748

CHAPTER 6

2749

2750

Conclusions

2751

2752 In this work rare decays were analysed in order to look for hints of new physics using
2753 data collected by the LHCb detector at centre-of-mass energies of 7 and 8 TeV and
2754 corresponding to a total integrated luminosity of 3.0 fb^{-1} .

2755 First, a measurement of the differential branching fraction of the rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$
2756 decay was performed together with the first measurement of angular observables
2757 for this decay. Evidence for the signal was found for the first time in the q^2 region
2758 between the two charmonium resonances and below the square of the J/ψ mass,
2759 in particular in the $0.1 < q^2 < 2.0 \text{ GeV}^2/c^4$ interval, where an increased yield is
2760 expected due to the vicinity of the photon pole. Due to a larger data sample and
2761 a better control of systematic effects, the uncertainty on the measurement in the
2762 $15 < q^2 < 20 \text{ GeV}^2/c^4$ interval are reduced by approximately a factor of three with
2763 respect to the previous LHCb measurements. The branching fraction measurements
2764 are compatible with SM predictions in the high- q^2 region, above the square of the
2765 J/ψ mass, but lie below the predictions in the low- q^2 region. In the angular analysis

of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays two forward-backward asymmetries, in the $p\pi$ and dimuon systems, were measured. The measurements of the A_{FB}^h observable are in good agreement with the SM predictions, while for the A_{FB}^ℓ observable they are consistently above the predictions. Following the publication of these studies improved predictions were produced, which are reported in Appendix D. Theoretical values are now compatible with the branching fraction measurements at low- q^2 and overestimate the experimental values at high- q^2 . The situation regarding the angular observables is unchanged but the significance of the existing discrepancies is enhanced due to the improved uncertainties on the predicted values.

Secondly, an analysis is set up to test flavour universality between electrons and muons exploiting rare decays. A set of requirements is defined to select the rare and normalisation, $B^0 \rightarrow K^{*0}\ell^+\ell^-$ and $B^0 \rightarrow K^{*0}(J/\psi \rightarrow \ell^+\ell^-)$ modes in both electron and muons channels, which includes the definition of a multivariate classifier. Efficiencies are obtained for this selection for all channels. A study of the possible backgrounds to these channels is performed, which results in a set of requirements to lower their yield in the selected samples and a set of PDFs to model the remaining contributions in the invariant mass fits. Preliminary fits to the 4-body invariant mass distributions of all channels are performed. Finally, a preliminary study of the systematic uncertainties is presented and a set a procedure to obtain the result and check its robustness are defined. The results are currently blinded.

2786

2787

REFERENCES

2788

- [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](https://arxiv.org/abs/1503.07138)].
- [2] **Particle Data Group** Collaboration, K. Olive et al., *Review of Particle Physics*, *Chin.Phys.* **C38** (2014) 090001.
- [3] **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](https://arxiv.org/abs/1207.7214)].
- [4] **CMS** Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [[arXiv:1207.7235](https://arxiv.org/abs/1207.7235)].
- [5] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, *Phys.Rev.* **D20** (1979) 2619–2625.
- [6] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961) 579–588.
- [7] “Wikipedia.” https://en.wikipedia.org/wiki/Standard_Model. Accessed: 2016-02-26.
- [8] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.
- [9] **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](https://arxiv.org/abs/1404.1903)].

- 2811 [10] **LHCb** Collaboration, R. Aaij et al., *Observation of $J/\psi p$ resonances*
 2812 *consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi K^- p$ decays*,
 2813 [arXiv:1507.03414](https://arxiv.org/abs/1507.03414).
- 2814 [11] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental*
 2815 *Test of Parity Conservation in Beta Decay*, *Phys. Rev.* **105** (1957) 1413–1414.
- 2816 [12] F. Strocchi, *Spontaneous Symmetry Breaking in Local Gauge Quantum Field*
 2817 *Theory: The Higgs Mechanism*, *Commun. Math. Phys.* **56** (1977) 57.
- 2818 [13] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10**
 2819 (1963) 531–533. [,648(1963)].
- 2820 [14] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory*
 2821 *of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.
- 2822 [15] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, et al.,
 2823 *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$*
 2824 *New Physics*, *Phys. Rev.* **D91** (2015), no. 7 073007, [[arXiv:1501.05013](https://arxiv.org/abs/1501.05013)].
- 2825 [16] F. Zwicky, *Spectral displacement of extra galactic nebulae*, *Helv. Phys. Acta* **6**
 2826 (1933) 110–127.
- 2827 [17] M. Gavela and Hernandez, *Standard model CP violation and baryon*
 2828 *asymmetry*, *Mod. Phys. Lett.* **A9** (1994) 795–810, [[hep-ph/9312215](https://arxiv.org/abs/hep-ph/9312215)].
- 2829 [18] S. Carlip, *Quantum gravity: A Progress report*, *Rept. Prog. Phys.* **64** (2001)
 2830 885, [[gr-qc/0108040](https://arxiv.org/abs/gr-qc/0108040)].
- 2831 [19] M. Maltoni, *Status of three-neutrino oscillations*, *PoS EPS-HEP2011*
 2832 (2011) 090.
- 2833 [20] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee,
 2834 P. S. Wildenhain, and J. Ullman, *Measurement of the solar electron neutrino*
 2835 *flux with the Homestake chlorine detector*, *Astrophys. J.* **496** (1998) 505–526.
- 2836 [21] **Super-Kamiokande** Collaboration, Y. Fukuda et al., *Evidence for*
 2837 *oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567,
 2838 [[hep-ex/9807003](https://arxiv.org/abs/hep-ex/9807003)].
- 2839 [22] **KamLAND** Collaboration, K. Eguchi et al., *First results from KamLAND:*
 2840 *Evidence for reactor anti-neutrino disappearance*, *Phys. Rev. Lett.* **90** (2003)
 2841 021802, [[hep-ex/0212021](https://arxiv.org/abs/hep-ex/0212021)].
- 2842 [23] J. L. Feng, *Naturalness and the Status of Supersymmetry*,
 2843 *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 351–382, [[arXiv:1302.6587](https://arxiv.org/abs/1302.6587)].
- 2844 [24] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of*
 2845 *Leptonic Charge*, *Sov. Phys. JETP* **26** (1968) 984–988. [*Zh. Eksp. Teor.*
 2846 *Fiz.* **53**, 1717 (1967)].

- [25] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870–880.
- [26] P. Fayet and S. Ferrara, *Supersymmetry*, *Phys.Rept.* **32** (1977) 249–334.
- [27] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys.Rev.Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [28] J. R. Ellis, G. Gelmini, C. Jarlskog, G. G. Ross, and J. W. F. Valle, *Phenomenology of Supersymmetry with Broken R-Parity*, *Phys. Lett.* **B150** (1985) 142.
- [29] G. Isidori and D. M. Straub, *Minimal Flavour Violation and Beyond*, *Eur.Phys.J.* **C72** (2012) 2103, [[arXiv:1202.0464](#)].
- [30] A. J. Buras, *Minimal flavor violation*, *Acta Phys.Polon.* **B34** (2003) 5615–5668, [[hep-ph/0310208](#)].
- [31] 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](#)].
- [32] A. J. Buras, D. Buttazzo, J. Girkbach-Noe, and R. Knegjens, *Can we reach the Zeptouniverse with rare K and B_{s,d} decays?*, *JHEP* **1411** (2014) 121, [[arXiv:1408.0728](#)].
- [33] G. Hiller and M. Schmaltz, *R_K and future b → sℓℓ physics beyond the standard model opportunities*, *Phys.Rev.* **D90** (2014) 054014, [[arXiv:1408.1627](#)].
- [34] 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](#)].
- [35] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev.Mod.Phys.* **68** (1996) 1125–1144, [[hep-ph/9512380](#)].
- [36] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, [hep-ph/9806471](#).
- [37] 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](#)].
- [38] 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](#)].
- [39] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in B → K^(*)ℓ⁺ℓ⁻ and B → K^{*}γ*, *JHEP* **1009** (2010) 089, [[arXiv:1006.4945](#)].

- [40] **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](#)].
- [41] 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](#)].
- [42] **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](#)].
- [43] **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](#)].
- [44] **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](#)].
- [45] **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](#)].
- [46] **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](#)].
- [47] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, *JHEP* **02** (2016) 104, [[arXiv:1512.04442](#)].
- [48] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \rightarrow K^* \mu^+ \mu^-$ Anomaly*, *Phys. Rev.* **D88** (2013) 074002, [[arXiv:1307.5683](#)].
- [49] **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](#)].
- [50] **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](#)].
- [51] **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](#)].
- [52] **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.
- [53] W. Altmannshofer and D. M. Straub, *New physics in $b \rightarrow s$ transitions after LHC run 1*, [arXiv:1411.3161](#).

- [54] T. Hurth, F. Mahmoudi, and S. Neshatpour, *On the anomalies in the latest LHCb data*, arXiv:1603.00865.
- [55] **MEGA** Collaboration, M. Ahmed et al., *Search for the lepton family number nonconserving decay $\mu^+ \rightarrow e^+ \gamma$* , Phys. Rev. **D65** (2002) 112002, [hep-ex/0111030].
- [56] **SINDRUM** Collaboration, U. Bellgardt et al., *Search for the Decay $\mu^+ \rightarrow e^+ e^+ e^-$* , Nucl. Phys. **B299** (1988) 1.
- [57] **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].
- [58] **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].
- [59] W. J. Marciano, T. Mori, and J. M. Roney, *Charged Lepton Flavor Violation Experiments*, Ann. Rev. Nucl. Part. Sci. **58** (2008) 315–341.
- [60] L. Evans, *The LHC machine*, PoS **EPS-HEP2009** (2009) 004.
- [61] C. Lefévre, “The CERN accelerator complex. Complexe des accélérateurs du CERN.” Dec, 2008.
- [62] **LHCb** Collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [63] **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].
- [64] M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759].
- [65] A. A. Alves Jr. et al., *Performance of the LHCb muon system*, JINST **8** (2013) P02022, [arXiv:1211.1346].
- [66] **LHCb** Collaboration, R. e. a. Aaij, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030.
- [67] “Wikipedia.” https://en.wikipedia.org/wiki/Cherenkov_radiation. Accessed: 2016-02-26.
- [68] **LHCb** Collaboration, R. e. a. Aaij, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (Dec, 2014) 1530022. 82 p.
- [69] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356–369, [physics/0402083].

- [70] R. Aaij et al., *The LHCb trigger and its performance in 2011*, *JINST* **8** (2013) P04022, [[arXiv:1211.3055](#)].
- [71] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].
- [72] 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](#)].
- [73] 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.
- [74] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.
- [75] 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](#)].
- [76] **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.
- [77] M. Clemencic et al., *The LHCb simulation application, GAUSS: design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023.
- [78] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, *Conf.Proc.* **C000917** (2000) 11–42.
- [79] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, *Nucl.Instrum.Meth.* **A559** (2006) 190–194.
- [80] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, [physics/0402093](#).
- [81] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, *Nucl.Instrum.Meth.* **A552** (2005) 566–575, [[physics/0503191](#)].
- [82] H. W. Bertini, *Low-Energy Intranuclear Cascade Calculation*, *Phys. Rev.* **131** (1963) 1801–1821.
- [83] B. Andersson, G. Gustafson, and H. Pi, *The FRITIOF model for very high-energy hadronic collisions*, *Z. Phys.* **C57** (1993) 485–494.
- [84] **COMPASS** Collaboration, P. Abbon et al., *The COMPASS experiment at CERN*, *Nucl. Instrum. Meth.* **A577** (2007) 455–518, [[hep-ex/0703049](#)].
- [85] 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](#)].

- [86] 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].
- [87] 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].
- [88] 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].
- [89] 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].
- [90] 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].
- [91] 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].
- [92] 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].
- [93] 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].
- [94] 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].
- [95] 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].
- [96] R. Mohanta and A. Giri, *Fourth generation effect on Λ_b decays*, *Phys.Rev.* **D82** (2010) 094022, [arXiv:1010.1152].
- [97] 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].
- [98] **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].
- [99] **CDF** Collaboration, S. Behari, *CDF results on $b \rightarrow s \mu \mu$ decays*, arXiv:1301.2244.

- 3027 [100] **LHCb** Collaboration, R. Aaij et al., *Measurement of the differential*
 3028 *branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$* , *Phys. Lett.* **B725** (2013) 25,
 3029 [[arXiv:1306.2577](#)].
- 3030 [101] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and
 3031 P. Santorelli, *Rare baryon decays $\Lambda_b \rightarrow \Lambda l^+l^- (l = e, \mu, \tau)$ and $\Lambda_b \rightarrow \Lambda\gamma$:
 3032 differential and total rates, lepton- and hadron-side forward-backward
 3033 asymmetries*, *Phys. Rev.* **D87** (2013) 074031, [[arXiv:1301.3737](#)].
- 3034 [102] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decay
 3035 amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*,
 3036 [*Phys. Lett.* **B724** (2013) 27, [[arXiv:1302.5578](#)]].
- 3037 [103] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in
 3038 *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*
 3039 (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003.
 3040 [[physics/0308063](#)].
- 3041 [104] T. Skwarnicki, *A study of the radiative cascade transitions between the
 3042 Upsilon-prime and Upsilon resonances*. PhD thesis, Institute of Nuclear
 3043 Physics, Krakow, 1986. DESY-F31-86-02.
- 3044 [105] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, *$\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ form
 3045 factors and differential branching fraction from lattice QCD*, *Phys. Rev.* **D87**
 3046 (2013), no. 7 074502, [[arXiv:1212.4827](#)].
- 3047 [106] **LHCb** Collaboration, R. Aaij et al., *Precision measurement of the Λ_b^0 baryon
 3048 lifetime*, *Phys. Rev. Lett.* **111** (2013) 102003, [[arXiv:1307.2476](#)].
- 3049 [107] T. Blake, S. Coquereau, M. Chrzaszcz, S. Cunliffe, C. Parkinson, K. Petridis,
 3050 and M. Tresch, *The $B_0 \rightarrow K_0^*\mu\mu$ selection using $3fb^{-1}$ of LHCb data*, Tech.
 3051 Rep. LHCb-INT-2013-058. CERN-LHCb-INT-2013-058, CERN, Geneva,
 3052 Nov, 2013.
- 3053 [108] F. James and M. Roos, *Minuit: A System for Function Minimization and
 3054 Analysis of the Parameter Errors and Correlations*, *Comput. Phys.
 3055 Commun.* **10** (1975) 343–367.
- 3056 [109] **LHCb** Collaboration, R. Aaij et al., *Measurements of the $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decay
 3057 amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys.
 3058 Lett.* **B724** (2013) 27, [[arXiv:1302.5578](#)].
- 3059 [110] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical
 3060 statistical analysis of small signals*, *Phys. Rev.* **D57** (1998) 3873–3889,
 3061 [[physics/9711021](#)].
- 3062 [111] T. M. Karbach, *Feldman-Cousins Confidence Levels - Toy MC Method*,
 3063 [[arXiv:1109.0714](#)].

- 3064 [112] S. Meinel, *Flavor physics with Λ_b baryons*, *PoS LATTICE2013* (2014) 024,
3065 [arXiv:1401.2685].
- 3066 [113] G. Hiller and F. Kruger, *More model independent analysis of $b \rightarrow s$*
3067 *processes*, *Phys.Rev.* **D69** (2004) 074020, [hep-ph/0310219].
- 3068 [114] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \rightarrow s\ell\ell$* ,
3069 *JHEP* **1502** (2015) 055, [arXiv:1411.4773].
- 3070 [115] **BaBar Collaboration** Collaboration, J. Lees et al., *Measurement of*
3071 *Branching Fractions and Rate Asymmetries in the Rare Decays*
3072 *$B \rightarrow K^{(*)}l^+l^-$* , *Phys.Rev.* **D86** (2012) 032012, [arXiv:1204.3933].
- 3073 [116] **BELLE Collaboration** Collaboration, J.-T. Wei et al., *Measurement of the*
3074 *Differential Branching Fraction and Forward-Backword Asymmetry for*
3075 *$B \rightarrow K^{(*)}l^+l^-$* , *Phys.Rev.Lett.* **103** (2009) 171801, [arXiv:0904.0770].
- 3076 [117] **LHCb Collaboration**, R. Aaij et al., *Test of lepton universality using*
3077 *$B^+ \rightarrow K^+\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601,
3078 [arXiv:1406.6482].
- 3079 [118] **LHCb Collaboration**, R. Aaij et al., *Measurement of the ratio of branching*
3080 *fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115**
3081 (2015), no. 11 111803, [arXiv:1506.08614]. [Addendum: *Phys. Rev.*
3082 *Lett.*115,no.15,159901(2015)].
- 3083 [119] “Lhc b loki twiki.”
3084 <https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters>.
3085 Accessed: 2015-09-30.
- 3086 [120] “Probnn presentation at ppts meeting.” <https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf>. Accessed: 2015-09-30.
- 3088 [121] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf C0303241* (2003) MOLT007, [physics/0306116].
- 3090 [122] P. Ball and R. Zwicky, *New results on $B \rightarrow \pi, K, \eta$ decay form factors from*
3091 *light-cone sum rules*, *Phys. Rev.* **D71** (2005) 014015, [hep-ph/0406232].
- 3092 [123] D. Melikhov and B. Stech, *Weak form-factors for heavy meson decays: An*
3093 *Update*, *Phys. Rev.* **D62** (2000) 014006, [hep-ph/0001113].
- 3094 [124] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *A Comparative study of the*
3095 *decays $B \rightarrow (K, K^*)\ell^+\ell^-$ in standard model and supersymmetric theories*,
3096 *Phys. Rev.* **D61** (2000) 074024, [hep-ph/9910221].
- 3097 [125] J. Hrivnac, R. Lednicky, and M. Smizanska, *Feasibility of beauty baryon*
3098 *polarization measurement in $\Lambda^0 J/\psi$ decay channel by ATLAS LHC*,
3099 *J.Phys.G* **G21** (1995) 629–638, [hep-ph/9405231].

- 3100 [126] W. Detmold and S. Meinel, $\Lambda_b \rightarrow \Lambda \ell^+ \ell^-$ form factors, differential branching
3101 fraction, and angular observables from lattice QCD with relativistic b quarks,
3102 arXiv:1602.01399.

3103

APPENDIX A

3104

3105

Decay models

3106

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

3107 The q^2 and angular dependancies of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays are modelled based on
 3108 Ref. [101], where the angular distribution for unpolarised Λ_b^0 production is defined
 3109 as

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

3110 In this formula θ_ℓ and θ_B correspond to the lepton and proton helicity angles, χ
 3111 is angle between dimuon and Λ decay planes (for unpolarised production we are
 3112 sensitive only to difference in azimuthal angles), $d_{i,j}^J$ are Wigner d-functions and h ,
 3113 h^B and H are helicity amplitudes for virtual dimuon, Λ and Λ_b^0 decays. The sum
 3114 runs over all possible helicities with the dimuon being allowed in spin 0 and 1 states
 3115 (J and J'). The m and m' indices run over the vector and axial-vector current
 3116 contributions.

3117 The production polarisation is introduced by removing $e^{i(\lambda_j - \lambda'_j)\chi}$ from the expression,
 3118 swapping small Wigner d-functions $d_{i,j}^J$ to the corresponding capital ones $D_{i,j}^J$ which
 3119 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. [101], 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. [101], 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

3129 terms of dimensionless quantities in eqs. C8 and C9 in Ref. [101]. In our actual
3130 implementation form factors calculated in the covariant quark model [101] are used
3131 and for the numerical values of the Wilson coefficients Ref. [101] 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. [105]. The form factors relations are found by comparing eqs. 66 and 68 in Ref. [101] to eq. 51 in Ref. [105]. 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

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

3132 In the calculations $\mu = m_b$ is used. For the strong coupling constant, we start
3133 from the world average value at the Z mass, $\alpha_s(m_Z^2) = 0.1185 \pm 0.0006$ [2], and we
3134 translate it to the scale m_b^2 by

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

3135 where $n_f = 5$. The LQCD form factors F_1^L and F_2^L can be then taken directly from
3136 Ref. [105] and plugged into the code implementing the calculation from Ref. [101].

₃₁₃₇ **A.2 Bi-dimensional distribution parameters**

₃₁₃₈ Expectations values for parameters in the bi-dimensional angular distribution for
₃₁₃₉ the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay calculated using form factors and numerical inputs from
₃₁₄₀ Ref. [101].

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

Table A.1: Prediction for angular observables entering two-dimensional angular distributions. Prediction is based on covariant quark model form factors from Ref. [101].

₃₁₄₁ **A.3 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ distribution**

₃₁₄₂ The angular distribution of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay is modelled using Ref. [125]. The
₃₁₄₃ differential rate is written as

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

₃₁₄₄ The expression uses four observables (angles) and depends on four complex amplitudes a_+, a_-, b_+, b_- and two real valued parameters for the production polarisation,
₃₁₄₅ P_b , and the Λ decay asymmetry, α_Λ . The angle θ is the angle of the Λ momentum
₃₁₄₆ in Λ_b^0 rest frame with respect to the vector $\vec{n} = \frac{\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}}{|\vec{p}_{inc} \times \vec{p}_{\Lambda_b^0}|}$, where \vec{p}_{inc} and $\vec{p}_{\Lambda_b^0}$ are the
₃₁₄₇ momenta of incident proton and Λ_b^0 in the center of mass system. The angles θ_1 and
₃₁₄₈ ϕ_1 are polar and azimuthal angle of the proton coming from the Λ decay in the Λ
₃₁₄₉ rest frame with axis defined as $z_1 \uparrow\uparrow \vec{p}_\Lambda$, $y_1 \uparrow\uparrow \vec{n} \times \vec{p}_\Lambda$. Finally, the angles θ_2 and ϕ_2
₃₁₅₀ are the angles of the momenta of the muons in J/ψ rest frame with axes defined as
₃₁₅₁ $z_2 \uparrow\uparrow \vec{p}_{J/\psi}$, $y_2 \uparrow\uparrow \vec{n} \times \vec{p}_{J/\psi}$.

₃₁₅₃ The distribution depends on the Λ decay asymmetry parameter, α_Λ , the production
₃₁₅₄ polarisation P_b and four complex amplitudes. The α_Λ is measured to be 0.642 ± 0.013
₃₁₅₅ for Λ . The production polarisation P_b and magnitudes of a_+, a_-, b_+ and b_- are

₃₁₅₆ measured in Ref. [109]. Phases are not measured therefore, as default all phases are
₃₁₅₇ set to zero and then they are randomly varied to calculate the systematic uncertainty.

APPENDIX B

Data-simulation comparison

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

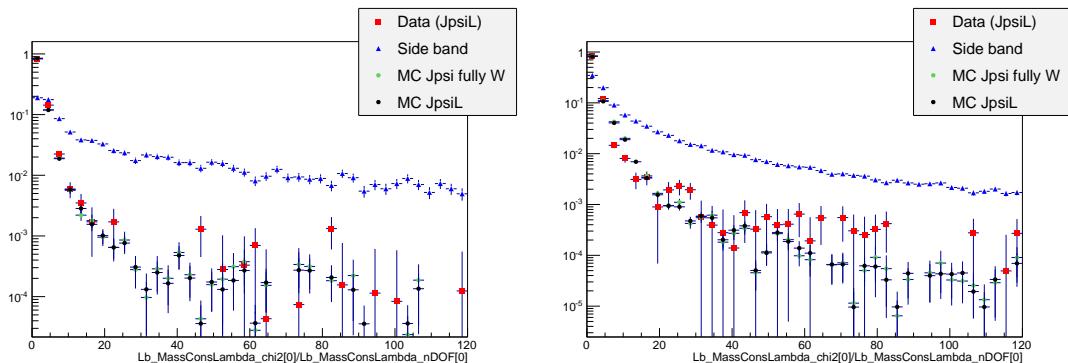


Figure B.1: Distributions of χ^2/NdF of the kinematic fit in data and simulation for LL (left) and DD (right) events.

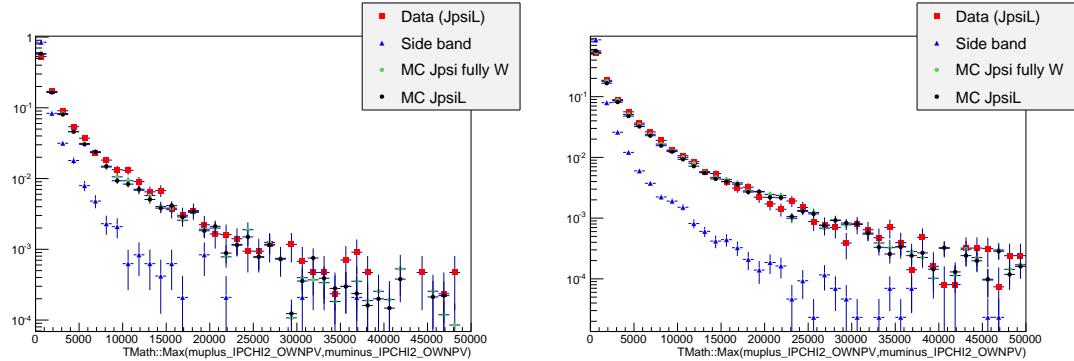


Figure B.2: Distributions of maximum muon $IP\chi^2$ variable in data and simulation for LL (left) and DD (right) events.

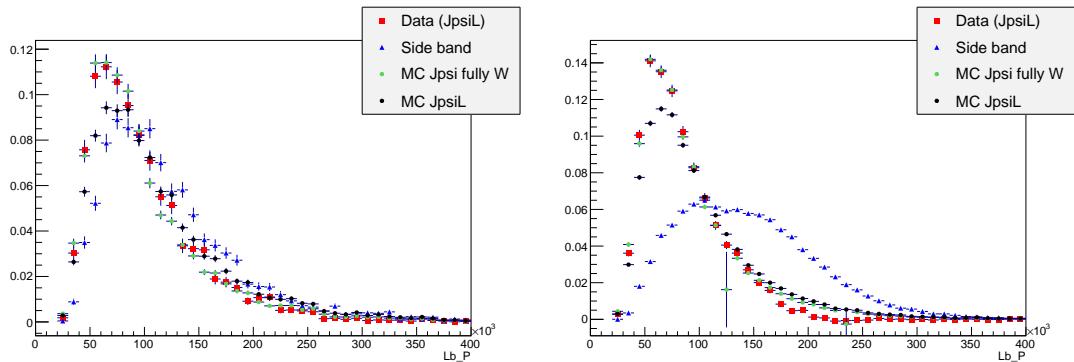


Figure B.3: Distributions of Λ_b^0 momentum variable in data and simulation for LL (left) and DD (right) events.

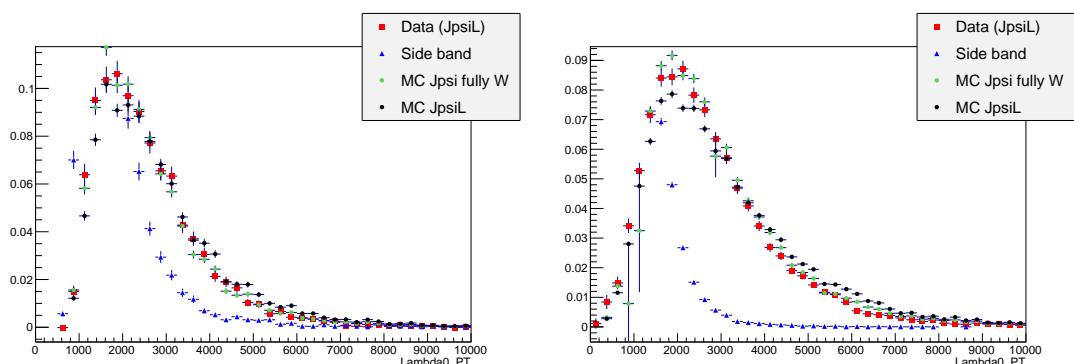


Figure B.4: Distributions of Λ transverse momentum variable in MC, data signal and data background for LL (left) and DD (right) events.

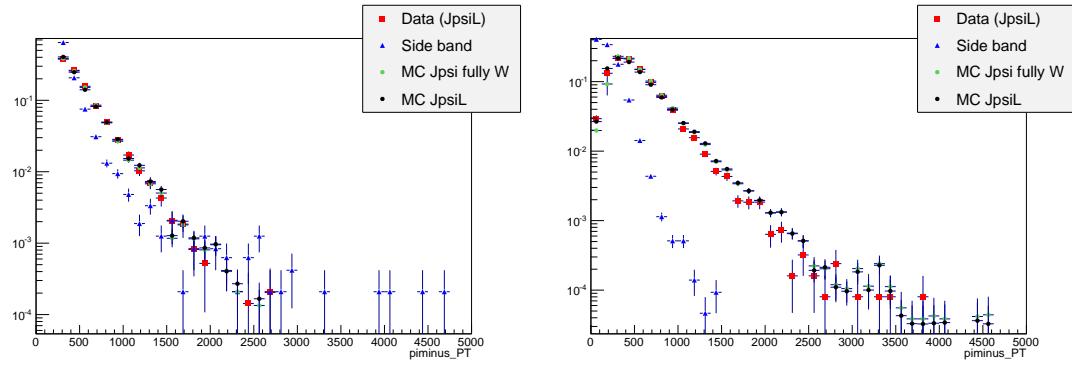


Figure B.5: Distributions of pion transverse momentum variable in data and simulation for LL (left) and DD (right) events.

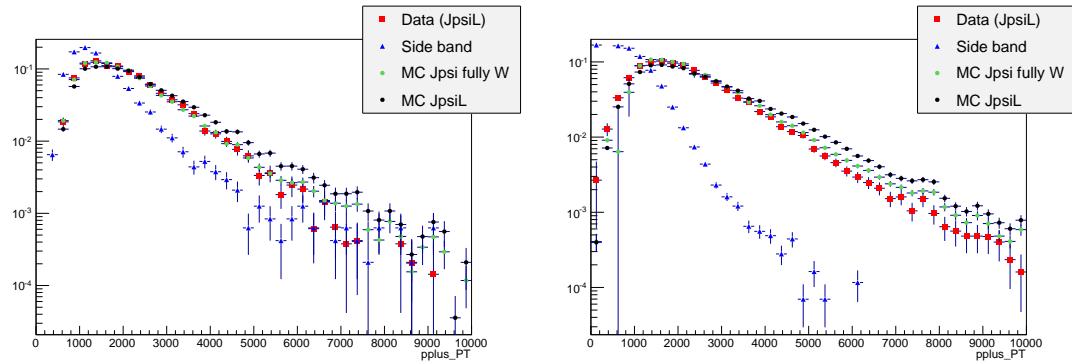


Figure B.6: Distributions of proton transverse momentum variable in data and simulation for LL (left) and DD (right) events.

3171

APPENDIX C

3172

**3173 Systematic uncertainties on the efficiency calculation for the
 3174 $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ branching fraction analysis.**

3175

3176 This appendix reports systematic uncertainties on absolute and relative efficiencies
3177 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.

APPENDIX D

Improved predictions for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ observables.

The publication of the results included in this thesis triggered interest in the theory community, which produced improved lattice calculations and predictions [126]. This section reports the measured quantities with the new predictions overlaid as reported in Ref. [126].

	Prediction	Measurement
$\langle d\mathcal{B}/dq^2 \rangle_{[15, 20]}$	0.756 ± 0.070	1.20 ± 0.27
$\langle F_L \rangle_{[15, 20]}$	0.409 ± 0.013	$0.61^{+0.11}_{-0.14}$
$\langle A_{\text{FB}}^\ell \rangle_{[15, 20]}$	-0.350 ± 0.013	-0.05 ± 0.09
$\langle A_{\text{FB}}^\Lambda \rangle_{[15, 20]}$	-0.2710 ± 0.0092	-0.29 ± 0.08

Table D.1: Comparison of predictions for the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ observables with the LHCb data presented in this thesis in the interval $[15, 20]$ GeV^2/c^4 , where the measurement is most precise.

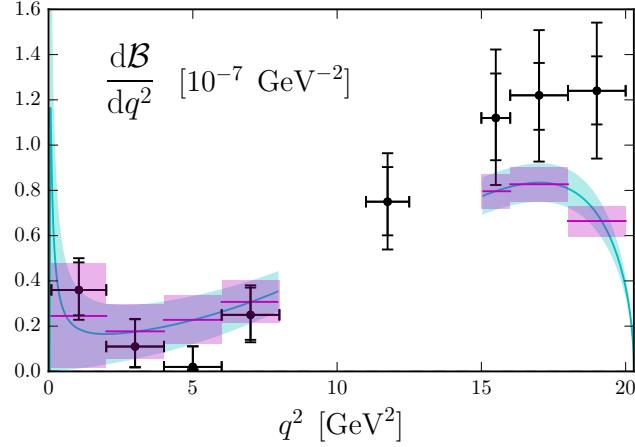


Figure D.1: Measurement of the differential branching fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decay as a function of q^2 already presented in Ch. 3 with improved Standard Model predictions from Ref. [126] overlaid.

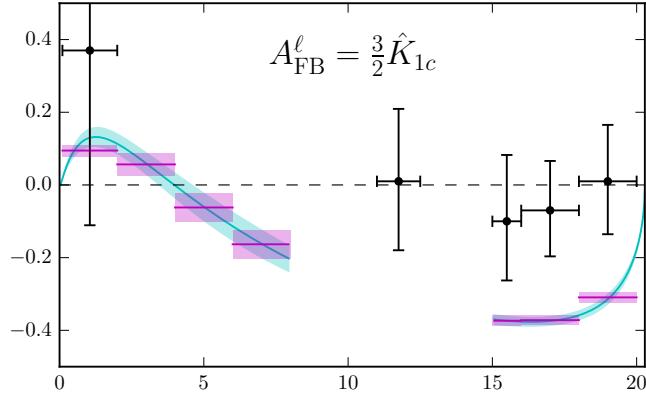


Figure D.2: Measurement of the lepton side forward-backward asymmetry, A_{FB}^ℓ , as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [126] overlaid.

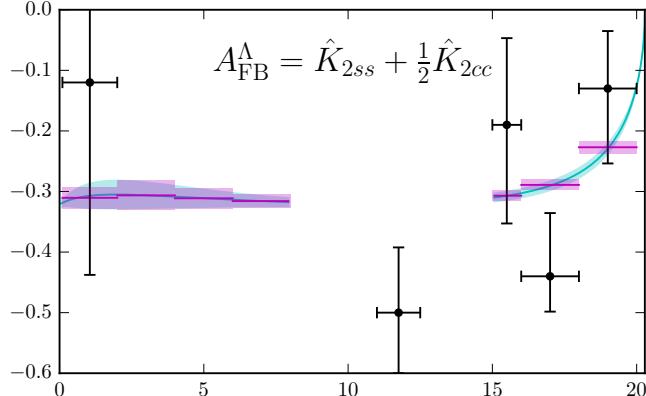


Figure D.3: Measurement of the hadron side forward-backward asymmetry, A_{FB}^h , as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [126] overlaid.

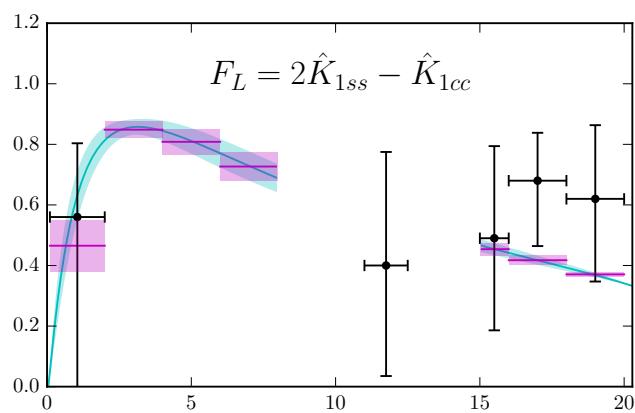


Figure D.4: Measurement of the fraction of longitudinally polarised dimuons, f_L , as a function of q^2 already presented in Ch. 4 with improved Standard Model predictions from Ref. [126] overlaid.

3186

APPENDIX E

3187

3188 Invariant mass fits to $B^0 \rightarrow K^{*0}\ell^+\ell^-$ simulated candidates

3189

3190 This appendix contains fits to the $m(K\pi\mu\mu)$ and $m(K\pi ee)$ invariant mass of $B^0 \rightarrow$
3191 $K^{*0}\ell^+\ell^-$ simulated candidates used to constrain parameters in the fit to data.

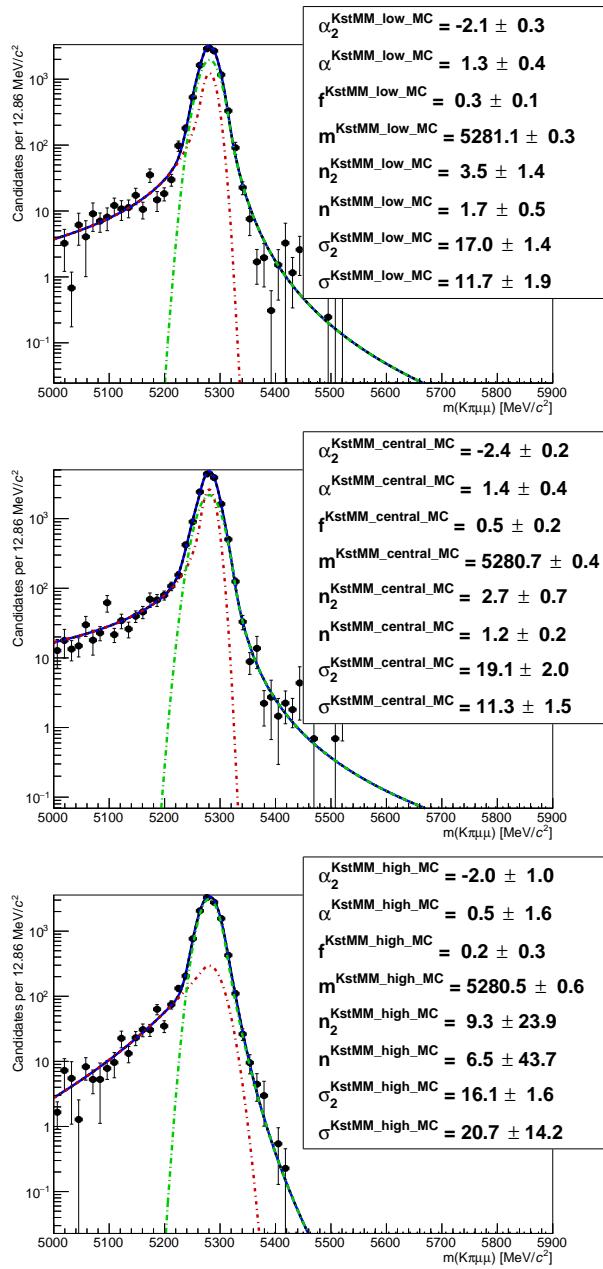


Figure E.1: Fitted $m(K\pi\mu\mu)$ mass spectrum for simulated events in the low (top), central (medium) and high (bottom) q^2 intervals.

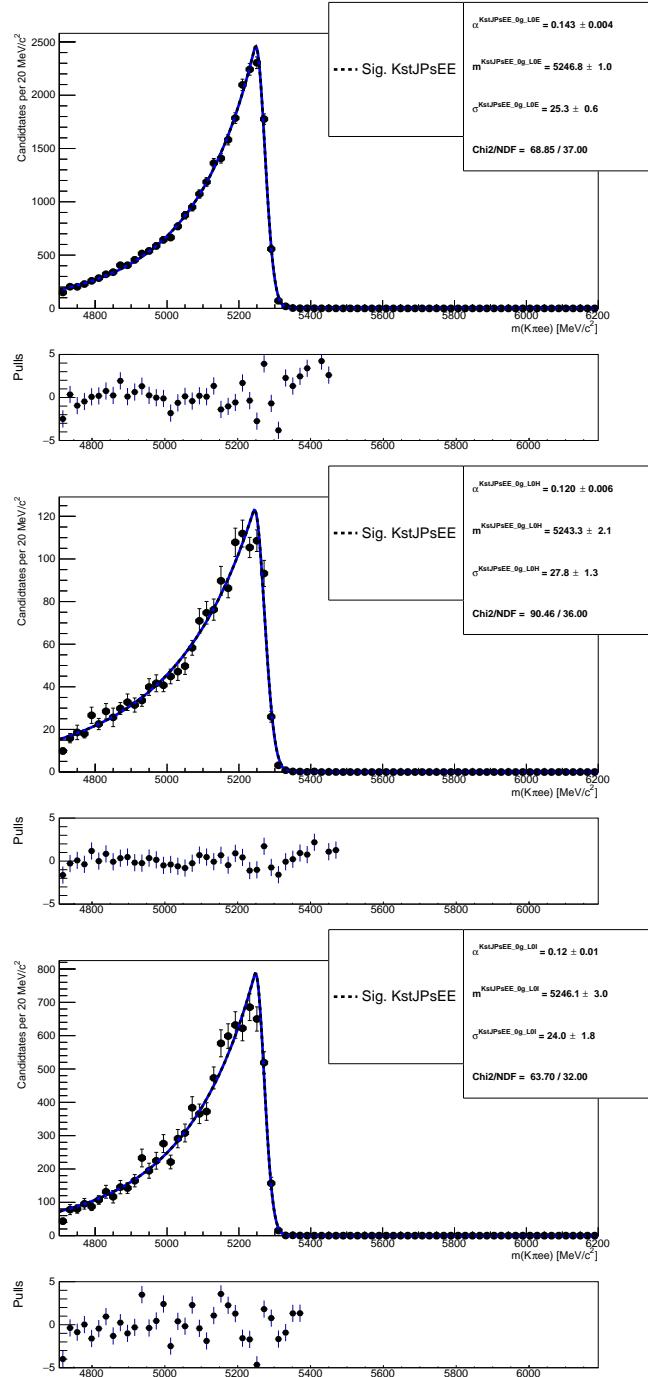


Figure E.2: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and no photon emitted.

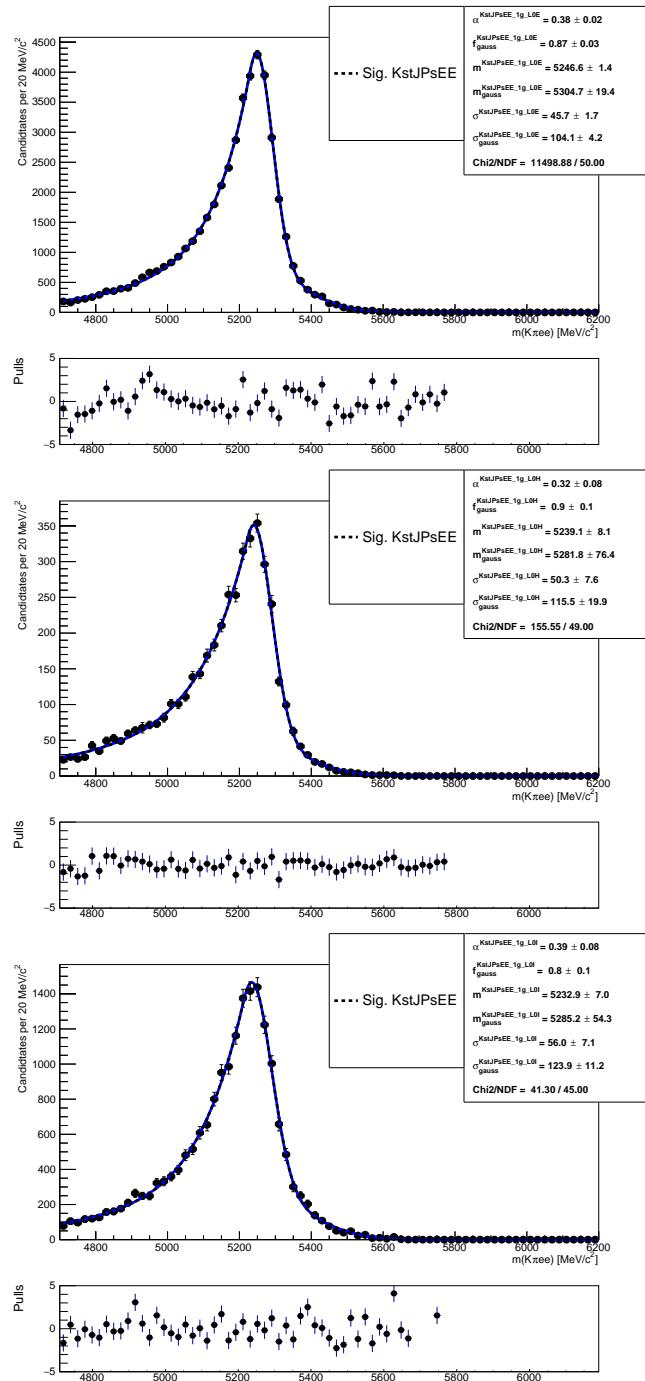


Figure E.3: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and one photon emitted.

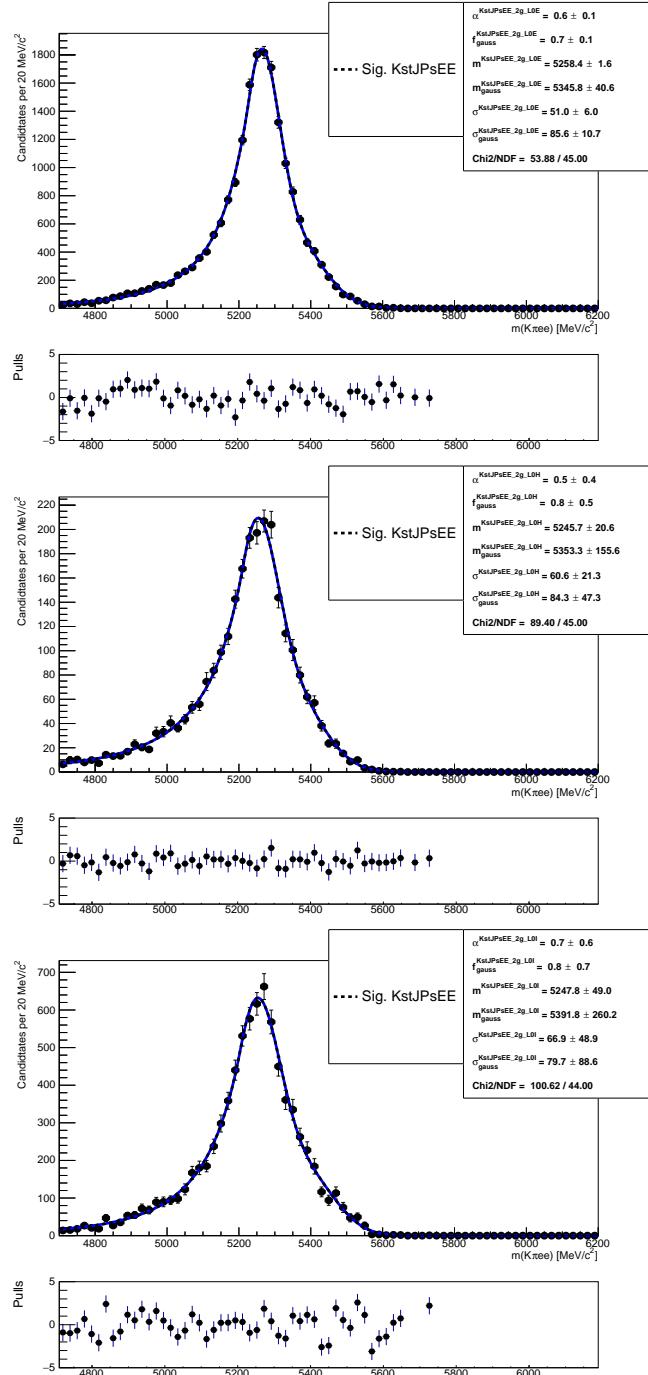


Figure E.4: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and two photons emitted.

APPENDIX F

3194 Invariant mass fits to $B^0 \rightarrow K^{*0} e^+ e^-$ candidates divided in 3195 trigger categories

3197 This appendix contains fits to the $m(K\pi ee)$ invariant mass of rare and control
3198 channel candidates separately in the tree trigger categories. Each trigger category
3199 is always fit with its own PDF but in the main text only their sum is shown for
3200 simplicity.

213 APPENDIX F. INVARIANT MASS FITS TO $B^0 \rightarrow K^{*0} E^+ E^-$ CANDIDATES DIVIDED IN TRIGGER CATEGORIES

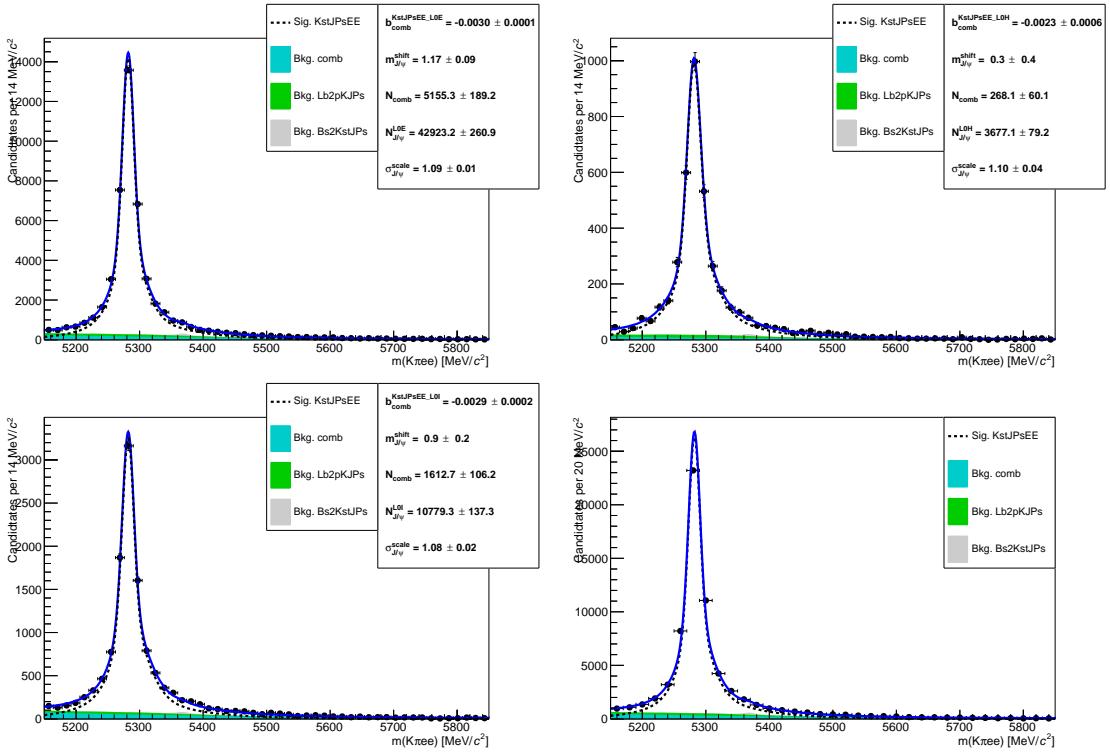


Figure F.1: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}(J/\psi \rightarrow e^+ e^-)$ candidates in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

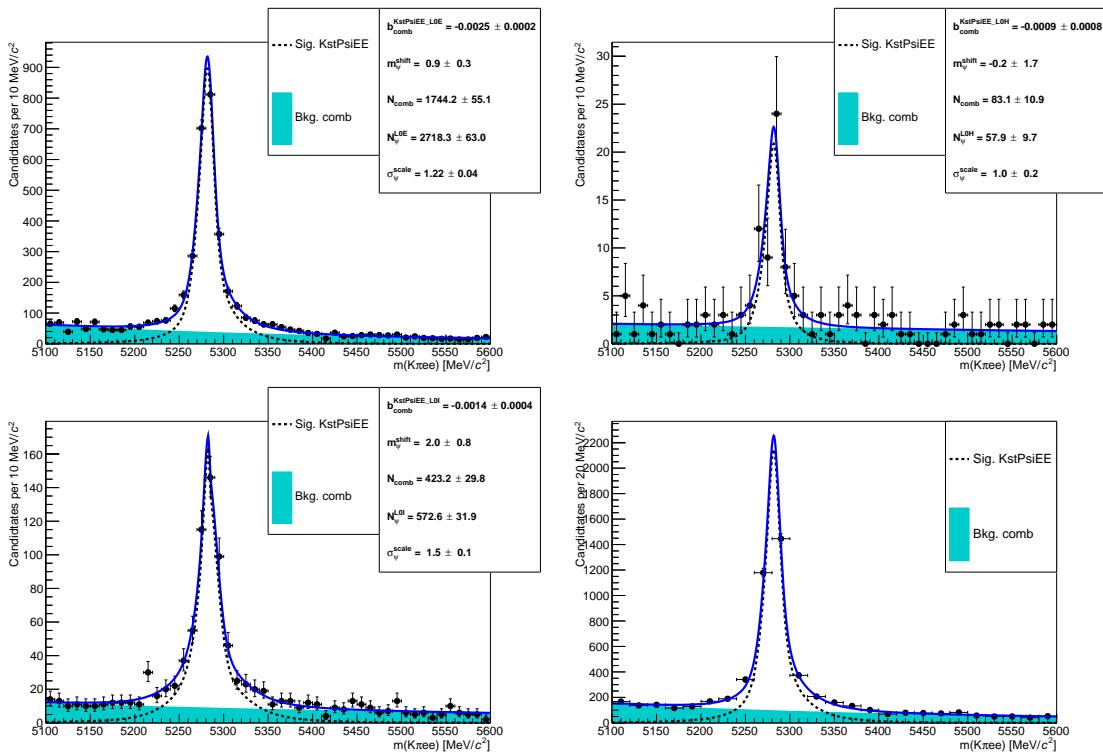


Figure F.2: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}(\psi(2S) \rightarrow e^+e^-)$ candidates in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

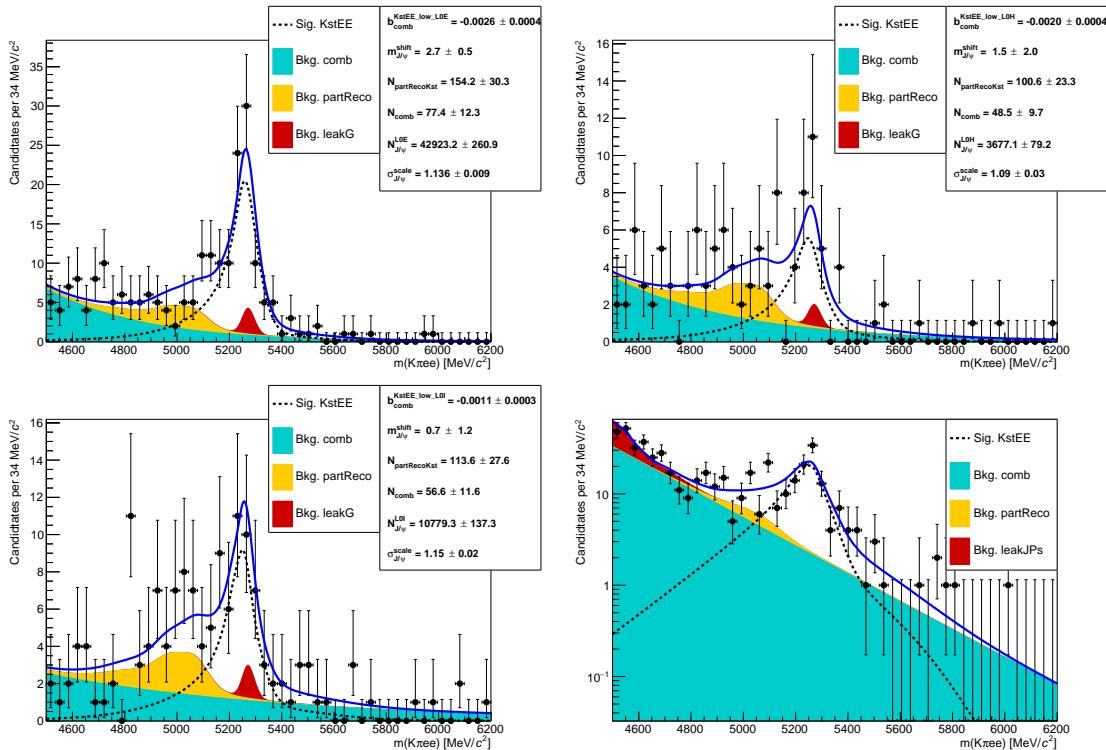


Figure F.3: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0} e^+ e^-$ candidates at low- q^2 in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

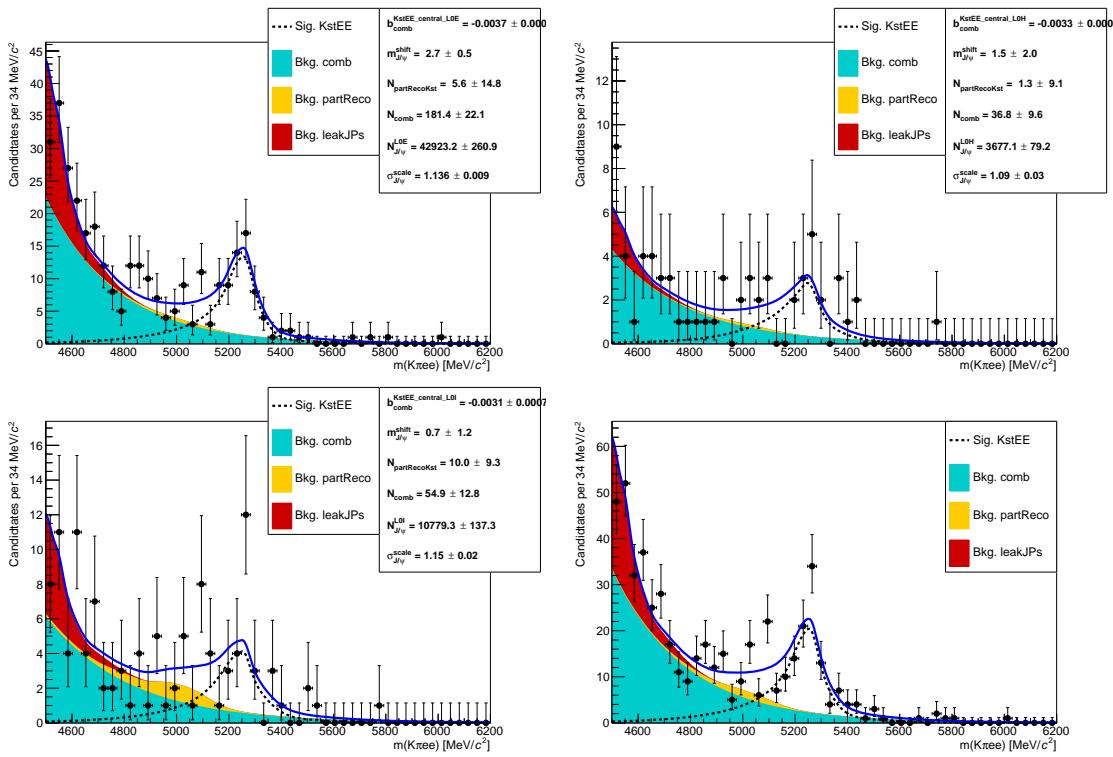


Figure F.4: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0}e^+e^-$ candidates at central- q^2 in the three trigger categories (L0E, L0H and L0I) separately, and (bottom right) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.

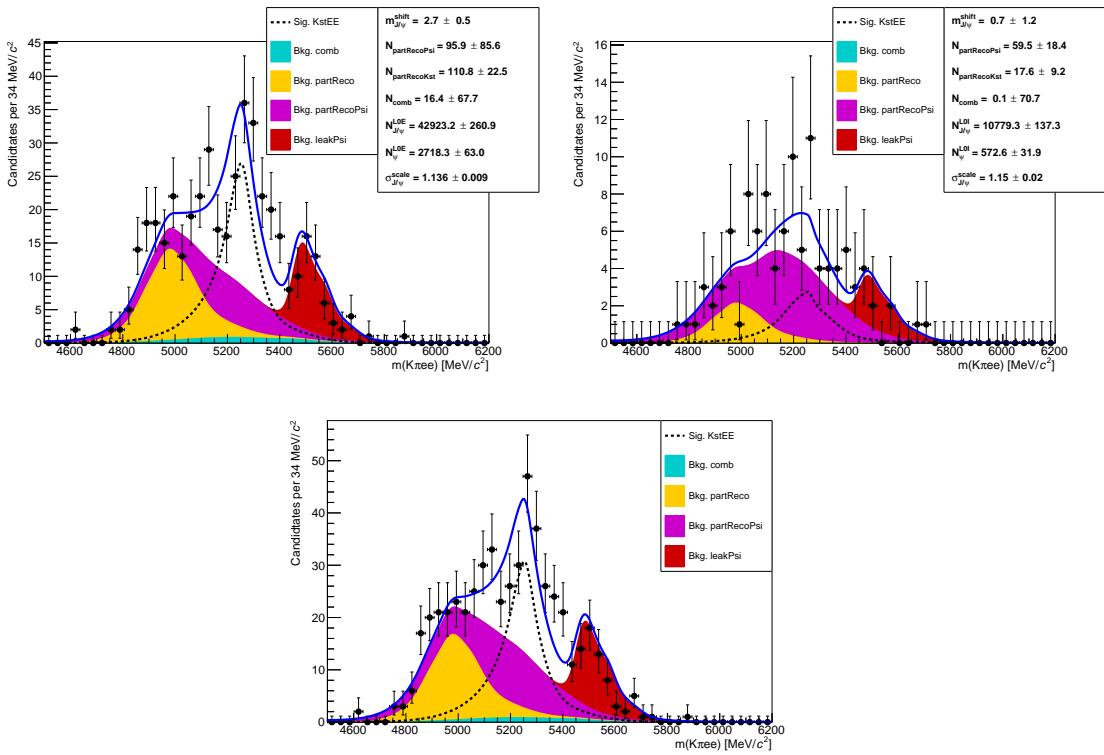


Figure F.5: Fit to the $m(K\pi ee)$ invariant mass of $B^0 \rightarrow K^{*0} e^+ e^-$ candidates at high- q^2 in the L0E and L0I trigger categories (top) separately, and (bottom) combined. The dashed black line (shaded shapes) represents the signal (background) PDF.