

$b \rightarrow s\ell\ell$ A SEARCH FOR NEW PHYSICS IN
TRANSITIONS IN MESONIC AND
BARYONIC SYSTEMS AT THE LHCb
EXPERIMENT.

L. Pescatore

*Thesis submitted for the degree of
Doctor of Philosophy*



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

June 19, 2015

ABSTRACT

X was measured, we showed that $Y \neq Z$ and that $M_H = 126 \text{ GeV}/c^2$.

DECLARATION OF AUTHORS CONTRIBUTION

I did this, and that, and some of the other.

ACKNOWLEDGEMENTS

I would like to thank bla, and bla ...

v

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

Motto latino

Contents

1	INTRODUCTION	1
1.1	Electromagnetic and weak interactions	3
1.1.1	Flavour and the CKM matrix	4
1.2	The puzzles in the SM	7
1.2.1	The flavour problem	9
1.3	Beyond the Standard Model	10
1.4	Flavour and BSM theories	10
1.5	Rare decays: a tool to search for new physics	12
1.5.1	Theoretical framework: the effective Hamiltonian	12
1.5.2	Perturbative corrections	13
1.5.3	Phenomenology of $b \rightarrow s\ell\ell$ decays	15
1.5.4	Observables in $b \rightarrow s\ell\ell$ decays	16
1.6	Experimental status	16
1.6.1	Dimuon decays of b hadrons	16
1.6.2	Semileptonic $b \rightarrow s\ell\ell$ decays of b hadrons	17
1.6.3	Lepton Flavour Violation searches	18
2	The LHCb detector at the Large Hadron Collider	20
2.1	The Large Hadron Collider	20
2.2	The LHCb detector	22
2.2.1	Tracking system	24
2.2.2	Particle identification	26
2.2.2.1	Calorimeters	26
2.2.2.2	RICH	27
2.2.3	The muon system	29
2.2.4	Trigger and software	30
I	Branching fraction and angular analysis of the rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay	32
3	Introduction	33
3.1	Analysis strategy and q^2 regions	35
3.2	Candidate types	36
4	Selection	38
4.1	Neural Networks	40

4.2	MVA optimization	46
4.3	Trigger	47
4.4	Background from specific decays	49
5	Yield extraction	51
5.1	Fit description	51
5.2	Fit results	56
5.3	Efficiency	63
5.3.1	Decay Model	63
5.3.1.1	Effect of new physics on the decay model	64
5.3.2	Kinematic re-weighting	64
5.3.3	Event type	67
5.3.4	Geometric acceptance	67
5.3.5	Reconstruction and stripping efficiency	67
5.3.6	Neural Networks efficiency	68
5.3.7	Trigger efficiency	70
5.3.8	PID efficiency	71
5.3.9	Relative efficiency	72
6	Systematic uncertainties	76
6.1	Yields	76
6.2	Efficiencies	77
6.2.1	Simulation statistics	78
6.2.2	Production polarisation and decay structure	81
6.2.3	Λ_b^0 lifetime	81
6.2.4	Λ detection efficiency	82
6.2.5	Data-MC discrepancies	83
7	Summary and results	85
8	Angular analysis	90
8.1	Extraction of forward-backward asymmetry	90
8.1.1	Angular acceptance	91
8.1.2	Background parameterisation	93
8.1.3	Toy studies on a 3-dimensional fit	94
9	Systematics uncertainties on angular observables	99
9.0.4	Non-flat angular efficiency	99
9.0.5	Resolution	100
9.0.6	Angular efficiency	101
9.0.7	Background parameterisation	102
9.0.8	Polarisation	103
10	Angular analysis summary and results	105
10.1	Feldman-cousins plug-in method	105
10.2	J/ψ cross-check	107
10.3	Results	108

II The R_{K^*} analysis	112
11 Towards new physics: R_H	113
11.0.1 Experimental status	116
11.1 The R_{K^*} analysis	116
12 Selection	118
12.1 Choice of q^2 intervals and signal windows	119
12.2 Trigger and Stripping	119
12.3 PID	123
12.4 Bremsstrahlung recovery for electrons	124
12.5 Peaking backgrounds	125
12.5.1 Charmonium vetoes	126
12.5.2 Cascade $b \rightarrow c\ell$, $c \rightarrow \ell$ decays	126
12.5.3 ϕ veto	127
12.5.4 $B^+ \rightarrow K^+\ell^+\ell^-$ plus a random pion	127
12.5.5 Λ_b decays	128
12.5.5.1 Other peaking backgrounds	129
12.6 Multivariate analysis	129
12.7 MVA optimization	133
13 Mass fits	136
13.1 Mass fits: muonic channels	137
13.2 Mass fits: electronic channels	139
13.2.1 Signal PDFs for the electronic channels in the central q^2 interval	141
13.2.2 Background PDFs for the electronic channels in the central q^2 interval	143
13.2.3 Summary of the fit to electronic channels in the central q^2 interval	145
13.2.4 Electron channels fits in the high q^2 interval	145
13.3 Fit summary	149
14 Efficiency	152
14.1 Data-simulation discrepancies	155
14.2 Geometric efficiency	157
14.3 Reconstruction efficiency and bin migration	157
14.4 PID efficiency	158
14.5 Trigger efficiency	159
14.5.1 Muonic channels	159
14.5.2 Electronic channels	160
14.6 Neural Networks efficiency	162
15 Systematic uncertainties	164
16 Result extraction	165
16.1 $R_{J/\psi}$ sanity check	165
16.2 R_{K^*} result summary	166
16.3 Branching ratios and expectations	167

List of Tables

1.1	Fundamental forces of nature together with their gauge bosons, relative strengths and range. Gravity is not included in the SM and the graviton is hypothetical at the current time.	2
4.1	Summary of stripping requirements. Where two values are given, the main one applies to long candidates and the one in parenthesis to downstream candidates.	39
4.2	Summary of inputs to the neural network in order of importance. Under “Id” the indices used for the correlation matrix (see Fig. 4.2) are reported. Column “adds” gives correlation significance added by given input when adding it to list of those ranked above, “only this” provides power of given input alone and “loss” shows how much information is lost when removing only given input.	43
4.3	Summary of trigger lines which candidates have to pass at various trigger levels. Trigger is always required to be due to tracks of the candidate itself.	47
5.1	Predicted numbers of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in each considered q^2 interval.	55
5.2	Raw yield in q^2 bins for DD and LL events. Uncertainties shown are statistical only.	58
5.3	Raw yield in q^2 bins for DD and LL events. Uncertainties shown are statistical only.	61
5.4	Exponential slopes, b , found from the fits to rare date and resoution scale factors, c , fixed in the fits using MC for DD and LL events. . . .	62
5.5	Absolute geometrical acceptance in bins of q^2 from MC simulation. Errors shown are statistical only.	68
5.6	Absolute Detection and reconstruction plus stripping efficiencies. Reconstruction efficiency is given separately for DD and LL events. Errors shown are statistical only.	69
5.7	Neural network selection efficiency. Errors shown are statistical only.	69
5.8	Absolute trigger efficiencies for selected events as determined from the simulation separately for LL and DD events. Efficiency is derived as ratio of events selected and accepted by trigger and events selected. Errors shown are statistical only.	70
5.9	Absolute PID efficiencies in q^2 bins	71
5.10	Absolute efficiency values for $A_b^0 \rightarrow J/\psi \Lambda$. Errors shown are statistical only.	72

5.11	Relative geometric efficiency and Detection efficiency between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays. Uncertainty reflects statistics of both samples. Errors shown are statistical only.	72
5.12	Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for long events. Uncertainty reflects statistics of both samples. Errors shown are statistical only.	73
5.13	Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for downstream events. Uncertainty reflects statistics of both samples. Errors shown are statistical only.	74
5.14	Relative PID efficiencies in q^2 bins	74
6.1	Values of systematics due to the choice of signal and background shapes in bins of q^2	77
6.2	Absolute values of systematic uncertainties on relative geometric efficiency.	79
6.3	Absolute values of systematic uncertainties on relative detection efficiency.	79
6.4	Absolute values of systematic uncertainties on relative reconstruction efficiency for long-long and down-down events.	79
6.5	Absolute values of systematic uncertainties on relative trigger efficiency for long-long and down-down events.	80
6.6	Absolute values of systematic uncertainties on relative MVA efficiency for long-long and down-down events.	80
6.7	Values of DD vertexing systematic uncertainties on relative reconstruction, trigger and MVA efficiencies for down-down events.	80
7.1	Absolute values of the total relative efficiency and the absolute value of the uncorrelated error, together with relative values of correlated error.	86
7.2	Values of corrected relative branching fraction for DD and LL events with statistical, correlated and uncorrelated error shown separately.	88
7.3	Combined values of corrected relative branching fraction with statistical error on rare (stat) and normalisation (norm) channels and systematic error shown separately.	89
8.1	RMS values for toy experiments on the extraction of the three parameters of interest with the 1D or 3D fitting methods.	96
8.2	Absolute values of systematic uncertainties due to fit bias in bins of q^2	97
9.1	Values of simulated $\cos\theta_\ell$ and $\cos\theta_\Lambda$ resolutions and systematics on angular observables in bins of q^2	101
9.2	Values systematic errors due to limited knowledge of the efficiency function on the three angular observables in bins of q^2	101
9.3	Values systematic errors due to MC statistics on the three angular observables in bins of q^2	102
9.4	Values of systematics due to the choice of background parameterisation in bins of q^2	103

10.1	Measured values of lepton side angular observables with systematic errors. The statistical errors are reported in Fig.10.8 evaluated as 2D 68% confidence regions. Errors reported on this table are estimations obtained using the Feldman-Cousins method where one only of the two observables is treated as parameter of interest at a time.	110
10.2	Measured values of A_{FB}^h observable with statistical errors, showing 68% CL interval, and systematic errors.	111
11.1	Previous R_X measurement by the BaBar [1] and Belle [2] experiment.	116
12.1	Summary of the trigger lines used for the $\mu\mu$ and the ee analysis.	121
12.2	Summary of cuts made for pre-selection of good candidates.	122
12.3	Variables used as inputs for the NN training. Next to each variable the ID number in brackets provides the index reported in the correlation matrices shown in Fig. 12.6.	130
12.4	Summary of inputs to the neural network in order of importance. The 10 most discriminating variables are shown. Column “adds” gives correlation significance added by given input when adding it to list of those ranked above, “only this” provides power of given input alone and “loss” shows how much information is lost when removing only given input. Decay Tree Fit is performed using DecayTreeFitter tool on whole decay chain with constraining tracks to appropriate vertex topology and the $m(p\pi)$ invariant mass to the PDG value.	132
13.1	Percentages of events with 0, 1 and 2 emitted photons in the three trigger categories, extracted from simulated events.	142
13.2	Percentages of events with 0, 1 and 2 emitted photons in the three trigger categories, extracted from simulated events.	148
13.3	Raw yields of events found fitting invariant mass distributions of the rare and resonant events.	151
14.1	Absolute efficiencies for ee and $\mu^+\mu^-$ channels in the J/ψ q^2 interval.	153
14.2	Double ratios of efficiencies $(\varepsilon^{ee}/\varepsilon^{J/\psi \rightarrow ee})/(\varepsilon^{\mu^+\mu^-}/\varepsilon^{J/\psi \rightarrow \mu^+\mu^-})$ in the $1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.	153
14.3	Absolute efficiencies for ee and $\mu^+\mu^-$ channels in the $1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.	154
14.4	Relative efficiencies rare over resonant ($\varepsilon^{rel} = \varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$) for ee and $\mu^+\mu^-$ channels in the $1 < q^2 < 6$ and $q^2 > 15$ GeV^2/c^4 intervals.	154
14.5	Net bin migration amounts in the considered q^2 intervals. Positive values mean “net in”, negative values “net out”.	158
14.6	Ratio $\varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$ where the efficiency for the rare channel its calculated in the same q^2 rage used to select the resonant channel.	163
16.1	Fully corrected measured values of the ratio $\mathcal{B}(B^0 \rightarrow K^*J/\psi \rightarrow \mu\mu)/\mathcal{B}(B^0 \rightarrow K^*J/\psi \rightarrow ee)$ in the three electron trigger categories.	166
16.2	Measured values of R_{ee} , $R_{\mu\mu}$ and R_{K^*} ratios.	167

16.3	Measured absolute branching ratio of the rare $\mu\mu$ and ee channels in the central and high q^2 regions. Errors shown are statistical only. . .	167
16.4	Expected and observed ratios of raw event yields, $N_{\ell\ell}/N_{J/\psi}$	168

List of Figures

1.1	Diagram of SM particles with their properties [?].	3
1.2	(left) A representation of the unitarity triangle and its parameters. (right) A summary of the most up to date measurements of the unitarity triangle parameters [].	8
1.3	Loop Feynmann diagrams for the rare $b \rightarrow s$ decay.	12
1.4	A typical q^2 spectrum of $b \rightarrow s\ell\ell$ process characterised by the photon pole at very low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2	15
1.5	Summary of limits set in lepton flavour violation searches [].	19
2.1	Scheme of CERN accelerators.	21
2.2	A side view of the LHCb detector [3].	23
2.3	On the left VeLo sensors mounted in line and on the right a schematic view of one sensor [3].	25
2.4	On the left a sketch of Cherenkov light emission [?] and on the right the Cherenkov angle versus momentum for the two radiators of RICH1 and for different particles. One can see that they allow to separate particles in different momentum ranges.	28
2.5	The LHCb muon system [3].	29
2.6	Scheme of the LHCb trigger system [3].	31
3.1	Representation of the two Λ candidate types built from “long” and “downstream” tracks.	37
4.1	Graphical representation of the DIRA (left) and χ_{IP}^2 (right) variables.	41
4.2	Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of table ???. Correlation is calculated using all events without distinguishing signal and background.	44
4.3	(right) NN output distribution for training (solid) and test (stripes) samples, for MC signal and sideband data events. (left) Purity as a function of neural network output.	44
4.4	Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and MC signal (right) events.	45

4.5	Dependence of figure-of-merit on the requirement on neural network output in the low q^2 region (left) and high q^2 (right) regions. The vertical line corresponds to the chosen cut.	45
4.6	Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection.	45
4.7	Single trigger efficiency for high q^2 events (left) and low q^2 (right).	48
4.8	Invariant mass distribution of fully selected $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ MC events (left) and of generator level events from J/ψ radiative tail in the region $q^2 < 8 \text{ GeV}^2/c^4$ (right).	50
5.1	Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows simulated data and the blue line is the signal fit function.	53
5.2	Invariant mass distribution of simulated misreconstructed $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted with Double Crystal Ball functions.	54
5.3	Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and residuals of the fit for long-long (left) and down-down (right) events. Lower plots are the same as the upper ones but shown in linear scale. The histogram shows data. On the plot are shown the total fit (blue line), the signal (black dashed), the combinatorial background (red dashed) and the $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ background (green dashed).	57
5.4	Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows real data selected with low q^2 selection. The blue line is the signal fit function.	58
5.5	Invariant mass distribution of rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ with fit and residuals for DD events (left) and LL events (right) in the integrated 15-20 GeV^2/c^2 q^2 bin. The histogram shows data, the blue line the total fit function and the dashed red line represents combinatorial background.	58
5.6	Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ candidates in the integrated 0.1-6 GeV^2/c^2 q^2 bin, still blinded, for DD events (left) and LL events (right).	59
5.7	Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ candidates in q^2 bins [0.1,2], [2,4], [4,6], [6,8] for DD events.	59
5.8	Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ candidates in q^2 bins [11,12.5], [15,16], [16,18], [18,20] for DD events.	60
5.9	Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ candidates in q^2 bins [0.1,2], [2,4], [4,6], [6,8] for LL events.	60
5.10	Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ candidates in q^2 bins [11,12.5], [15,16], [16,18], [18,20] for LL events.	61
5.11	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.	64
5.12	Relative effect of different Wilson Coefficients sets on the total efficiency in bins of q^2	65

5.13	Histogram of weights used for the kinematical reweighting as a function of the momentum and transverse momentum of Λ_b^0	66
5.14	Relative efficiencies as a function of q^2 . geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) events, trigger efficiency for DD (e) and LL (f) and MVA efficiency for DD (g) and LL (h)	75
6.1	Ratio of reconstruction efficiency in Data and MC found using K_S events[4]	82
6.2	Difference between the efficiency calculated with and without the weight to match the PIDmu distributions in data and MC as a function of q^2	84
7.1	Corrected relative $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ yield as a function of q^2 bins for down-down and long-long events. Errors shown represent statistical and systematic.	87
7.2	Combined relative $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ yield as a function of q^2 bins . Errors shown represent statistical and systematic.	87
8.1	Efficiency as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events in the $15-20 GeV^2/c^2$ q^2 bin.	92
8.2	Efficiency as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events in the $15-20 GeV^2/c^2$ q^2 bin.	93
8.3	Efficiency as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events in the $1.1-6.0 GeV^2/c^2$ q^2 bin.	93
8.4	Efficiency as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events in the $1.1-6.0 GeV^2/c^2$ q^2 bin.	94
8.5	Efficiency as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events for J/ψ events.	94
8.6	Efficiency as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events for J/ψ events.	95
8.7	Background distribution as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events in the $15-20 GeV^2/c^2$ (top) and $1.1-6 GeV^2/c^2$ (bottom) q^2 bins.	96
8.8	Background distribution as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events in the $15-20 GeV^2/c^2$ (top) and $1.1-6 GeV^2/c^2$ (bottom) q^2 bins.	97
8.9	Distribution of observed parameters of interest over 500 toy experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the $15 - 20 q^2$ bin in real data.	98
8.10	Distribution of observed parameters of interest over 500 toy experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the $11 - 12.5 q^2$ bin in real data.	98

9.1	Plots show the absolute difference of the observables' values obtained fitting toy MC generated with a 2D distributions multiplied by a 2D efficiency and fitting 1D projections. From left to right for f_L , A_{FB}^ℓ and A_{FB}^h	100
9.2	2-dimensional efficiencies obtained from weighted MC for $\cos \theta_\ell$ and $\cos \theta_\Lambda$ versus $\cos \theta$	104
9.3	Plots show the absolute difference of the observables' values obtained fitting toy MC generated with polarisation set to -0.03 and $+0.15$. From left to right for f_L , A_{FB}^ℓ and A_{FB}^h	104
10.1	The plot shows the physical A_{FB}^ℓ vs f_L parameter space. The dark region corresponds to points where the pdf is positive in the whole $[-1, 1]$ interval.	107
10.2	p-values obtained with the plug-in method around the measured values in the high q^2 integrated bin for f_L (left) and A_{FB}^h (right). The red line marks the points at p-value 32% corresponding to a 68% CL.	107
10.3	Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and residuals of the fit for long-long (left) and down-down (right) events. The histogram shows data as in Fig.5.3 with an extra proton PID cut to remove K_s^0 background.	108
10.4	Fitted angular distribution as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events for J/ψ events.	109
10.5	Fitted angular distribution as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events for J/ψ events.	109
10.6	Fitted angular distribution as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events in the $15\text{-}20 GeV^2/c^2 q^2$ bin.	110
10.7	Fitted angular distribution as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events in the $15\text{-}20 GeV^2/c^2 q^2$ bin.	110
10.8	In black 2-dimensional 68% CL confidence regions as a function of A_{FB}^ℓ and f_L	111
11.1	Loop diagrams of the $B_d \rightarrow K/K^{*0} l^+ l^-$ process.	114
11.2	Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.	115
12.1	Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electronic (left) and muonic (right) channels in 2012 data.	120
12.2	Correct ID probability distributions for muons (left) and electrons (right) in 2012 data.	123
12.3	On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis shows the mis-ID probability.	124
12.4	Schematic view of the Bremsstrahlung recovery.	125

12.5 On the left the distribution of 2011 data events on the variables ($m_{K(\pi \rightarrow K)}$) and ($m_{K(\pi \rightarrow K)\mu\mu}$), where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. On the right the mass of the three-body system ($m_{K\mu\mu}$) where the peak due to the $B^+ \rightarrow K^+\mu\mu$ decay is visible.	128
12.6 Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of table ???. Correlation is calculated using all events without distinguishing signal and background.	131
12.7 NN output distributions for training (solid) and test (stripes) samples, for MC signal and data sideband events. For the muon training (left) and the electron one (right).	132
12.8 Average value of NN output as a function of B^0 mass for data sideband (left) and MC signal (right) events for the electron (top) and muon (bottom) training.	133
12.9 Dependence of figure-of-merit on the requirement on neural network output for electrons (left) and muons (right). The vertical lines corresponds to the chosen cuts.	134
12.10 Receiver operating characteristic (ROC) curves for electrons (left) and muons (right). They show the signal efficiency versus the background rejection.	134
13.1 Fitted $m(K\pi\mu\mu)$ mass spectrum for K^*J/ψ simulated events.	139
13.2 Fitted $m(K\pi\mu\mu)$ invariant mass spectrum for K^*J/ψ (left) and $K^*(J/\psi \rightarrow \mu^+\mu^-)$ (right). Dashed lines represent background components.	140
13.3 Simulated distributions of misreconstructed background events falling into the $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ sample coming from the hadronic (left) and leptonic (right) systems.	144
13.4 (left) Simulated 4-body invariant mass distributions for events involving higher K^* states and passing out full selection. (right) Simulated invariant mass distribution of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ events leaking into the central q^2 interval.	145
13.5 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.	146
13.6 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.	146
13.7 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.	147
13.8 Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0}J/\psi(J/\psi \rightarrow ee)$ real data events in the three trigger categories.	147

13.9 Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} e^+ e^-$ real data events in the three trigger categories and no photon emitted.	148
13.10 Simulated invariant mass of the $K\pi ee$ system in the $1 < q^2 < 6$ and $15 < q^2 < 20$ GeV^2/c^4 intervals.	150
13.11 Fitted $m(K\pi ee)$ invariant mass distribution in the $\psi(2S)$ interval, $12 < q^2 < 25$ GeV^2/c^4 and in the high q^2 interval, $15 < q^2 < 20$ GeV^2/c^4	150
14.1 Fitted 4-body invariant mass distributions after pre-selection for the electron (left) and muon (right) channels.	156
14.2 Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.	156
14.3 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).	156
14.4 Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^* ee$ events. . .	158
14.5 Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).	160
14.6 Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.	161

1

CHAPTER 1

2

3

INTRODUCTION

4

5 The Standard Model of Particle Physics (SM) is a quantum field theory (QFT) de-
6 scribing strong and electroweak (EW) interactions. It was formulated in its current
7 form in the mid-70s and has been an extremely successful and predictive theory
8 since then. Almost all known phenomena from 1 eV up to almost 200 GeV are well
9 described by the SM and experiments at the Large Hadron Collider (LHC) are now
10 probing the SM up to the TeV scale. Finally, in 2013 we were able to observe the
11 Higgs boson, one of the fundamental building blocks of the theory, giving a solid
12 theoretical basis to the theory. However, experimentally well established effects, like
13 neutrino oscillations and the presence of dark matter, are outside the reach of the
14 SM. Furthermore, the model does not include the description of gravity. Therefore
15 this motivates the search for New Physics (NP).

16 The SM is based on the symmetry groups of strong ($SU(3)_C$) and electroweak in-
17 teractions ($SU(2)_L \times U(1)_Y$). The subscripts C, L, and Y stand for colour, charge,
18 left-handed fields, and hyper-charge. The Lagrangian describing the SM has been

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

Table 1.1: Fundamental forces of nature together with their gauge bosons, relative strengths and range. Gravity is not included in the SM and the graviton is hypothetical at the current time.

¹⁹ then developed on invariance under the unitary product group $SU(3) \times SU(2) \times U(1)$,
²⁰ which reflects conservation laws such as the conservation of electric and strong
²¹ charge. The parameters of the model were then experimentally measured.

²² Particles included in the SM can be grouped under a few categories depending
²³ on their properties and ability to interact with each other. First of all we can
²⁴ distinguish between fermions (half-integer spin particles) and bosons (integer spin
²⁵ particles). Fermions constitute the basic building blocks of matter, while bosons are
²⁶ the mediators of the interaction between them. Since in the SM the concept
²⁷ of bosonic mediators of interactions arises because of gauge symmetry[?], they are
²⁸ called “gauge bosons”. The list of the known interactions with their force carrier
²⁹ and properties is reported in Table 1. The matter of which we are made is mainly
³⁰ composed of electrons and protons, which have spin 1/2; protons are then composed
³¹ of u and d quark, which again have spin 1/2. Among fermions one can then consider
³² two smaller groups: quarks and leptons. Quarks carry colour charge and therefore
³³ can interact through the, so called, strong interaction, while leptons, which do not
³⁴ carry colour charge, are insensitive to it. For each particle exists a corresponding
³⁵ anti-particle with opposite quantum numbers. Finally fermions seem to be divided
³⁶ in three families having similar properties but different mass. This last structure
³⁷ embedded in the SM is also calle flavour structure and it will be the main tool used
³⁸ in this thesis, a more detailed description of it is given in the next sections. A
³⁹ schematic view of the fundamental particles in the SM is shown in Fig. 1.1.

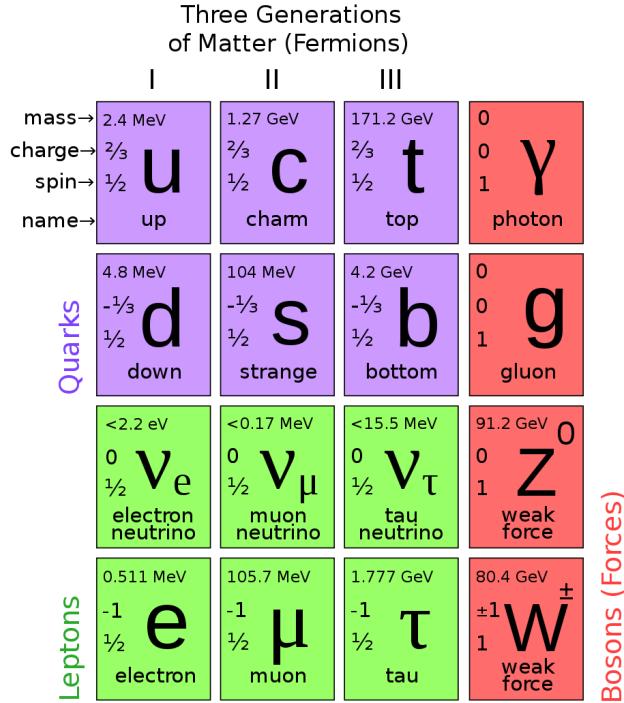


Figure 1.1: Diagram of SM particles with their properties [?].

40 1.1 Electromagnetic and weak interactions

41 The Electromagnetic (EM) force is responsible for binding electrons and nuclei to-
 42 gether in atoms. Its force carrier, the photon, is the gauge boson of the EM force.
 43 In the SM the photon must be massless, which also sets the range of the EM force to
 44 infinity, since it is proportional to the inverse of the mediator mass. In fact Heisen-
 45 berg's Uncertainty Principle tells us that $\Delta E \Delta t > \hbar$, namely virtual particles of
 46 energy ΔE are allowed to exist for time intervals inferior to Δt . Then, since they
 47 can move at maximum at the velocity of light this also sets a relation between the
 48 length of time and space in which a virtual photons can exist. The EM force has an
 49 infinite range as virtual photons can be very close to the mass shell, which results
 50 in a very long lifetime.

51 The weak interaction is responsible for the β decay of nuclei and all known fermions
 52 interact through the weak interaction. In the Standard Model of particle physics
 53 this interaction is caused by the emission or absorption of W^\pm and Z bosons. These

54 are much heavier than protons or neutrons (see Table 1) and this yields that the
55 weak force has a very short range. Using Heisenberg’s Principle together with Ein-
56 stein’s formula $\Delta E = mc^2$, which relates mass and energy, and knowing that the
57 maximum space that a particle can cover in a time Δt is $r = c\Delta t$, qualitatively
58 $r \sim \hbar/mc$. In this picture the carriers of the weak force can travel $r \sim 2 \cdot 10^{-3}$
59 fm. The weak interaction is also the only one that violates parity-symmetry, which
60 states that interactions are invariant under a reflection of all coordinates. This sym-
61 metry breaking arises from the fact that only left-handed fermions interact through
62 the weak interaction. The first experiment showing this was made by Wu in 1957
63 [?]. Similarly, the weak interaction is the only one that also breaks the CP sym-
64 metry, which combines parity transformations and “charge conjugation”. This is
65 interesting also because all interactions are invariant under CPT transformations,
66 which combines CP transformations and time reversal, hence, breaking CP the weak
67 interaction is also not invariant under time reversal.

68 In 1968 Salam, Glasow and Weinberg unified the weak and electromagnetic force
69 in a single theory called electroweak (EW), having a single coupling constant[5].
70 The EW interactions are divided into charged currents (CC) and neutral currents
71 (NC). In the first group, quarks and leptons interact with the W^\pm bosons, as in
72 the decays $\mu^+(\mu^-) \rightarrow e^+\nu_e\bar{\nu}_\mu(e^-\bar{\nu}_e\nu_\mu)$ and $n \rightarrow pe^-\bar{\nu}_e(\bar{p}e^+\nu_e)$. The study of these
73 processes confirmed that only the left-handed (right-handed) component of fermions
74 (anti-fermions) takes part in weak processes. The CC interaction have a peculiarity:
75 they are the only interactions in the SM that violate flavour conservation at tree
76 level (see next section), while any other interaction not conserving flavour has to
77 happen through loops. The second group of EW interactions, NC, corresponds to
78 interactions of the photon and Z boson with a fermion and its anti-fermion.

79 1.1.1 Flavour and the CKM matrix

80 “Flavour” in particle physics refers to the quark-lepton composition of a particle.
81 The introduction of flavour quantum numbers was motivated in order to explain

82 why some decays, although kinematically allowed, have never been observed. All
83 leptons have a quantum number $L_l = 1$ (where $l = e, \mu, \tau$), which is conserved by
84 all interactions. For example decays like $\mu^- \rightarrow e^-\gamma$, which is kinematically possible,
85 have never been observed, since the lepton number in the initial and final state are
86 different, while decays conserving the lepton number as $\mu^- \rightarrow e^-\nu_\mu\bar{\nu}_e$ have been
87 observed.

88 In the non leptonic sector particles carry flavour numbers described as follow:

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

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

96 Measuring branching fractions of weak decays like $\pi \rightarrow \mu\nu_\mu$ and $K \rightarrow \mu\nu_\mu$, suggested
97 the existence of more than one coupling constant. Cabibbo[5], in order to preserve
98 the universality of weak interactions, suggested that the difference in branching frac-
99 tion could arise from the fact that the doublets participating in the weak interactions
100 are a mixture of the flavour eigenstates. He therefore introduced the Cabibbo angle
101 θ_c considering that eigenstates participating to the weak interaction are rotated with
102 respect of the flavour eigenstates.

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

103 Considering a 6 quark system one angle is not enough to describe a rotation but the
104 mixing system can be generalised using a 3×3 unitary matrix, this is called CKM

matrix, from the names of Cabibbo, Kobayashi and Maskawa. The unitarity of the matrix is required in order to conserve the total probability. Theoretically, a $N \times N$ complex matrix is dependent on $2 \cdot N^2$ real parameters. Then, requiring unitarity ($AA^\dagger = A(A^*)^T = I$), the number of independent parameters left is $(N-1)^2$. A 3×3 depends then on 4 real parameters, which can be divided in 3 real constants and one imaginary phase. The imaginary phase generates the CP-violation which was observed in weak interactions. In Eq. 1.2 is reported a parametrisation of the CKM matrix together with the most recent measured values[5]. In this parametrisation ρ , A , and λ are the real constants and η the imaginary phase; in Eq. 1.3 are reported their relations with the 3 mixing angles.

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

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

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

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

It is interesting to note that the CKM matrix seems to be hierarchical, namely elements on the diagonal are approximately 1 and then get smaller and smaller going farther from the diagonal. An other feature to note is that, due to the unitarity of the matrix, the transformation have no effect on neutral interaction. As a result flavour-changing neutral currents are forbidden at tree level (in absence of closed

120 loops) in the SM.

121 The CKM matrix to preserve probability has to be unitary and this imposes con-
122 straints to its terms for the form:

$$\sum_i |V_{ik}|^2 = 1 \text{ and } \sum_k V_{ik} V_{jk}^* = 0. \quad (1.6)$$

123 This is a constraint of 3 complex numbers that can be viewed as the sides of a triangle
124 called the “unitarity triangle”. The precise measurement of the parameters of the
125 unitarity triangle is a powerful stability test of the standard model and sets a solid
126 base for new physics searches in the favour sector.

127 In Fig. 1.1.1 is shown a representation of the unitarity triangle together with a plot
128 summarising the most up to date constraints the the angles from measurements. One
129 of the main goals of the LHCb experiment is to precisely measure the angle γ , which
130 is currently the least constrained from measurements.

131 1.2 The puzzles in the SM

132 Despite the confirmation of many predictions of the SM, this theory has several
133 limitations and is unable to account for some observational facts.

- 134 • *Dark matter*: From experimental evidence the content of visible matter in the
135 universe is not enough to account the observed rotation of galaxies [?] in the
136 context of general relativity. Furthermore, studies of the fluctuations of the
137 cosmic microwave background indicate the existence of cold dark matter[?],
138 formed of particles which do not interact through the SM forces and for which
139 there is no SM candidate.

- 140 • *Matter-antimatter asymmetry*: We observe a large asymmetry between the
141 quantity of matter and antimatter in the Universe. Assuming that both were
142 equally created in the initial state of the Universe, a condition such as the

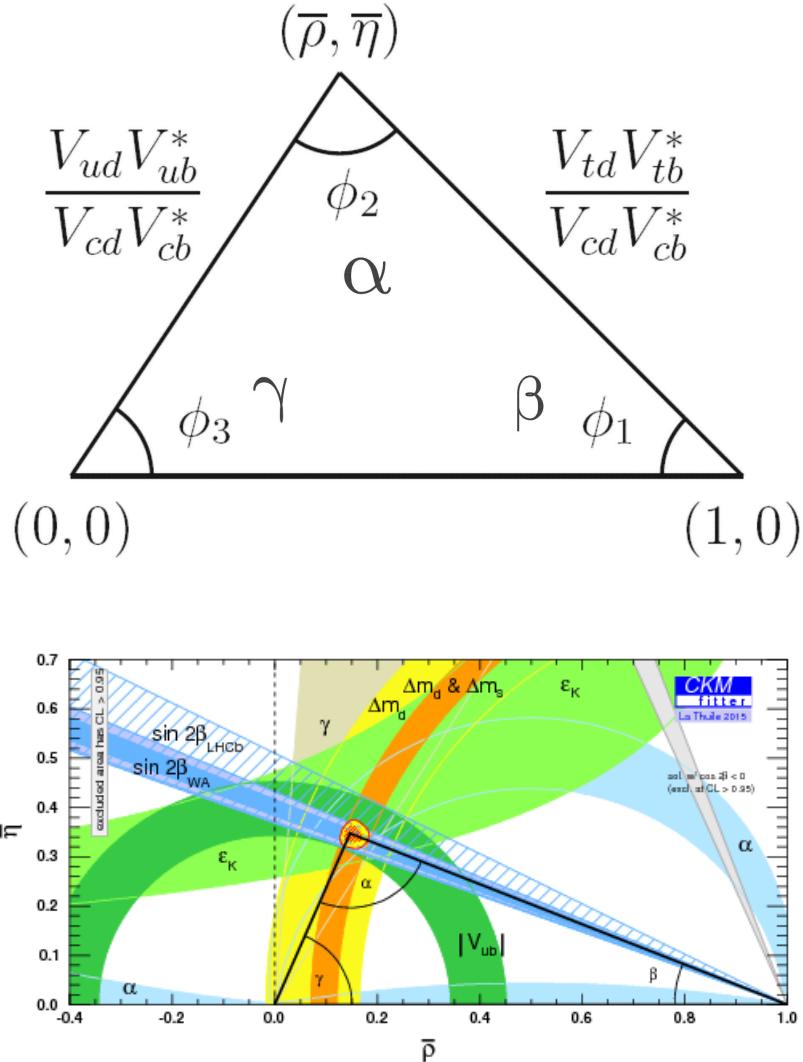


Figure 1.2: (left) A representation of the unitarity triangle and its parameters. (right) A summary of the most up to date measurements of the unitarity triangle parameters [1].

violation of the CP symmetry is necessary to account for such observed differences. However the magnitude of CP violation predicted by the SM is not enough to explain them [?].

- *Gravity:* There is not yet a consistent procedure to introduce gravity in the SM.
- *Neutrino oscillation:* By now, many measurements regarding solar and atmospheric neutrinos as well as neutrinos from nuclear reactor established that

neutrinos can change flavour while propagating in space. This is not predicted in the SM, in fact in the SM neutrinos are massless while an oscillation requires a non zero mass [?].

- *The mass hierarchy problem:* The mass of a scalar (spin 0 particle), such as the Higgs boson, suffers from quantum corrections of its mass due to the physics above a certain scale Λ , $m_{HSM}^2(\text{phys}) \sim m_{HSM}^2 + \frac{c}{16\pi}\Lambda^2$. Using the recently measured value for the Higgs mass $\sim 126 \text{ GeV}/c^2$ [?], Λ should be $\sim TeV$ in order to avoid a fine tuning of the bare mass term.

1.2.1 The flavour problem

The SM has been very successful in describing the observed particles and their interactions so far. However, because of its many puzzles, described in 1.2, it is believed only to be part of a more general theory or only to be valid up to a certain energy scale. Many theoretical models expect New Physics (NP) to enter at the TeV scale. For example, flavour conservation does not have a strong theoretical basis in the Standard Model and is mainly motivated by experimental evidence.

So far we have talked about Flavour Changing Charged Currents (FCCC) that are mediated by the W^\pm bosons. In the SM, they are the only sources of flavour changing interaction and, in particular, of generation changing interactions, where a quark or a lepton of a family transforms into one of another family. There is no fundamental reason why there cannot be Flavour Changing Neutral Currents (FCNCs). Yet, experimentally we see that FCNC processes are highly suppressed. The way in which the SM explains this is to forbid FCNC at tree level: since Z and γ interaction conserve flavour the only other way to have FCNC is through particle loops. This makes these interactions very sensitive to new physics, since its effects, which are expected to be small, are not disguised by dominating SM processes.

By now one of the possible explanations why we do not observe FCNC at tree level is the Minimal Flavour Violation (MFV) hypothesis[?][?], where FCNC are 'protected'

¹⁷⁷ by symmetry principles.

¹⁷⁸ 1.3 Beyond the Standard Model

¹⁷⁹ From the last two sections it is evident that, despite the great success of the SM,
¹⁸⁰ there is a need to explore new theories. Among the most promising approaches are
¹⁸¹ those invoking Super-Symmetry and extradimensions

¹⁸² In Super-Symmetry new degrees of freedom are introduced to suppress the diverging
¹⁸³ term of the scalar mass. This represents the main reasoning of Super-Symmetry,
¹⁸⁴ which assumes that for each fermion there is a corresponding boson. Since boson
¹⁸⁵ and fermions contribute with opposite sign to the mass term they would cancel out
¹⁸⁶ if for each fermion we can add a boson term [?].

¹⁸⁷ The idea to introduce extra dimensions was triggered by the fact that normally
¹⁸⁸ gravity is not relevant in particle physics at today's energy scales, which is at the
¹⁸⁹ EW scale (~ 100 GeV), and this is why it is neglected in the SM. However, adding
¹⁹⁰ extra dimensions to the normal 3 spatial dimensions, one can restore some of the
¹⁹¹ strength of gravity, yielding effects at the EW scale [?].

¹⁹² In all these approaches severe constraints on masses and couplings must be imposed
¹⁹³ to maintain compatibility with the SM at the electroweak scale.

¹⁹⁴ 1.4 Flavour and BSM theories

¹⁹⁵ Since most of the BSM theories predict processes violating flavour, the observation
¹⁹⁶ or non-observation of these processes can give important information about New
¹⁹⁷ Physics.

¹⁹⁸ BSM theories can be classified according to the amount of flavour violation they
¹⁹⁹ introduce. The first class of models to consider is the Minimal Flavour Violation

200 (MFV). These are models where the only sources of flavour violation in the SM and
201 in BSM are the different Yukawa couplings. The MFV paradigm provides a way
202 to resolve the tension between expectation, driven by naturalness arguments, that
203 NP should be at the TeVscales and limits on FCNC processes that point to much
204 higher scales. As hamiltonians for $b \rightarrow d$ and $b \rightarrow s$ share the same structure,
205 ratios between these transitions provide powerful tests of MFV. One particularly
206 important example is the ratio of B^0 and B_s^0 dimuon decay rates ??.

207 In the quest for New Physics an important role is also played by simplified models
208 as an intermediate model building step. Instead of building models valid up to the
209 GUT scale one could consider simplified models including the SM and a new sector
210 with a limited number of parameters. Such models are easier to constrain but can
211 nevertheless point in the right direction to build a complete theory. The choice of
212 the new sector to add can be driven by the need to explain existing discrepancies
213 between data and SM predictions or by theoretical prejudice.

214 Two models especially relevant for the discussion in this thesis are Z'-penguins and
215 leptoquarks.

216 A Z'-penguin is a FCNC process involving a neutral field and as for the SM pen-
217 guins it arises in loops and modifications of the effective couplings arise in most SM
218 extensions. A survey of Z' models can be found in Ref. [].

219 Leptoquarks are bosonic particles that carry one quark and one lepton flavour quan-
220 tum number. They can be spin 1 but they are more commonly assumed to be
221 scalar particles. A three level exchange of these particles induces processes such as
222 $b \rightarrow (s, d)\ell\ell$ and therefore we could observe an enhancement of their decay rates with
223 respect to the SM. Leptoquarks also provide a natural explanation for non-universal
224 couplings to leptons, introducing lepton flavour violation.

²²⁵ **1.5 Rare decays: a tool to search for new physics**

²²⁶ In the SM Flavour Changing Neutral Current processes, e.g. transitions from a
²²⁷ b quark with charge of $1/3$ to a s or d with a charge of $+2/3$, are forbidden at
²²⁸ tree level but can occur through loops, box or penguin decays, see Fig. 1.3. The
²²⁹ branching fractions of this kind of decays is $\sim 10^{-6}$ or lower and therefore they
²³⁰ are called "rare decays". Additional NP contributions to the virtual loops are not
²³¹ necessarily suppressed with respect to the SM component and this makes these
²³² decays very sensitive to New Physics. Furthermore, this approach to New Physics
²³³ searches is interesting as New Physics could be at a high mass scale not accessible
²³⁴ at colliders but its effect could be observed in loop effects. Radiative and penguin
²³⁵ decays are particularly interesting because they are theoretically well understood
²³⁶ which allows precise comparisons with measurements. Finally they provide a great
²³⁷ quantity of observables, not only decay rates, but also CP asymmetries and angular
²³⁸ observables can be affected by New Physics.

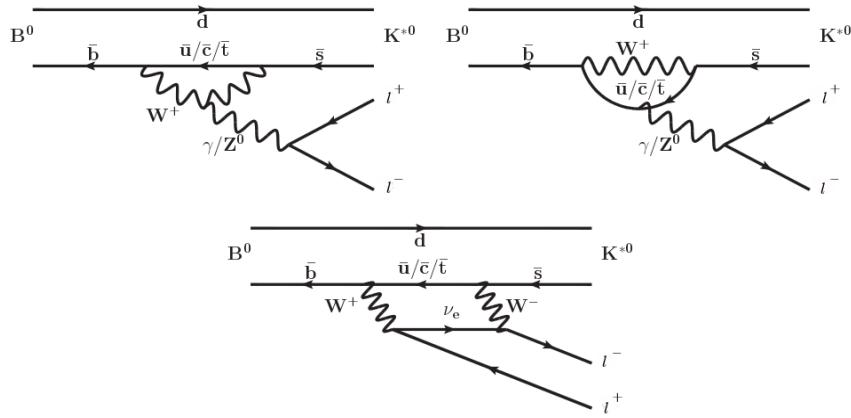


Figure 1.3: Loop Feynmann diagrams for the rare $b \rightarrow s$ decay.

²³⁹ **1.5.1 Theoretical framework: the effective Hamiltonian**

²⁴⁰ Rare B decays are governed by an interplay between weak and strong interactions.
²⁴¹ The QCD corrections that arise from hard gluon exchange bring large logarithms

242 of the form $\alpha_s^n(m_b) \log^m(m_b/M)$, where $M = m_t$ or $M = m_W$. The large masses of
243 W, Z and top quark compared to that of the b quark allow the construction of an
244 effective low-energy theory where the effective hamiltonian can be written as:

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} \sum C_i(\mu, M) \mathcal{O}_i(\mu) \quad (1.7)$$

245 The method of the Operator Product Expansion (OPE) allows the separation of the
246 decay amplitudes into two parts: the long-distance contributions, contained in the
247 operator matrix elements, \mathcal{O}_i , and the short-distance physics described by the so
248 called Wilson Coefficients, C_i . G_F denotes the Fermi coupling constant.

249 The weak coefficients at the weak scale can be obtained from matching amplitudes
250 of the full electroweak theory into \mathcal{H}_{eff} . Then one can derive a renormalization
251 group equation for the Wilson Coefficients

$$\mu \frac{d}{d\mu} C_i(\mu) = \gamma_{ij} C_j(\mu) \quad (1.8)$$

252 where the matrix γ is the anomalous dimensions matrix of the operators \mathcal{O}_i . At
253 leading order the solution is given by:

$$C_i(\mu) = \left[\frac{\alpha_s(\mu_W)}{\alpha_s(\mu)} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W) = \left[\frac{1}{1 + \beta_0 \frac{\alpha_s(\mu)}{4\pi} \ln \frac{\mu_W^2}{\mu^2}} \right]^{\frac{\gamma_{ii}^0}{2\beta_0}} C_i(\mu_W) \quad (1.9)$$

254 1.5.2 Perturbative corrections

255 The relevant efective Hamiltonian for $b \rightarrow s\ell^+\ell^-$ transitions is

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

where the V_{ub} and V_{bs} are the factors of the CKM matrix. The following local operators are particularly important for leptonic decays:

$$\mathcal{O}_7 = \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_R b F_{\mu\nu} \quad \mathcal{O}'_7 = \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_L b F_{\mu\nu} \quad (1.11)$$

$$\mathcal{O}_8 = g_s \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_R T^a b G_{\mu\nu}^a \quad \mathcal{O}'_8 = g_s \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_L T^a b G_{\mu\nu}^a \quad (1.12)$$

$$\mathcal{O}_9 = \bar{s} \gamma_\mu P_L b \bar{\ell} \gamma^\mu \ell \quad \bar{s} \mathcal{O}'_9 = \gamma_\mu P_R b \bar{\ell} \gamma^\mu \ell \quad (1.13)$$

$$\mathcal{O}_{10} = \bar{s} \gamma_\mu P_L b \bar{\ell} \gamma^\mu \gamma_5 \ell \quad \mathcal{O}'_{10} = \bar{s} \gamma_\mu P_R b \bar{\ell} \gamma^\mu \gamma_5 \ell \quad (1.14)$$

(1.15)

where $P_{L/R} = (1 \mp \gamma_5)/2$ denotes the left/right handed chiral projection, T^a are the QCD generators and $F_{\mu\nu}$ is the electromagnetic field tensor. The \mathcal{O}' operators correspond to right-handed coupling obtained by swapping P_R and P_L in the equations. The left-handedness of the weak interaction means that the C' Wilson Coefficients are suppressed by $\mathcal{O}(m_s/m_b)$ in the SM.

In the SM at $\mu_s = m_b$ the Wilson Coefficients have values:

$$C_7^{SM} = -0.3, \quad C_9^{SM} = 4.2, C_{10}^{SM} = -4.2. \quad (1.16)$$

(1.17)

New Physics contributions appear in the Wilson Coefficients as additive factors

$$C_i = C_i^{NP} + C_i^{SM}$$

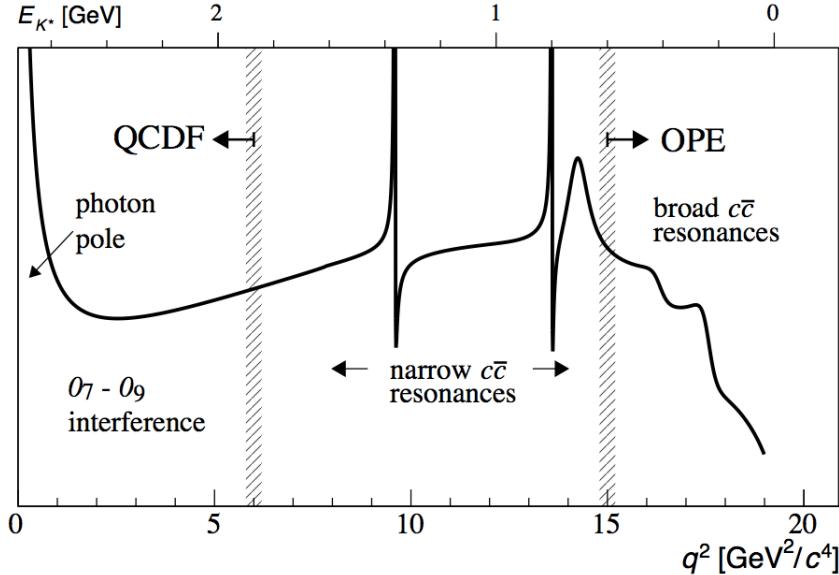


Figure 1.4: A typical q^2 spectrum of $b \rightarrow s\ell\ell$ process characterised by the photon pole at very low q^2 , charmonium resonances at central q^2 and broad resonances at high q^2 .

²⁶⁶ 1.5.3 Phenomenology of $b \rightarrow s\ell\ell$ decays

²⁶⁷ Semileptonic b hadron decays are characterised by two kinematic regimes: at low
²⁶⁸ q^2 , where the emitted hadron is energetic ($E > \Lambda_{QCD}$ in the b hadron rest frame),
²⁶⁹ QCD factorisation applied; at high q^2 , the region of low hadron recoil ($q^2 = O(m_b)$),
²⁷⁰ an Operator Product Expansion (OPE) in $1/m_b$ applies. In both regions decays can
²⁷¹ be predicted using the different methods and the biggest uncertainties come from
²⁷² the limited knowledge of hadronic transition form factors.

²⁷³ As can be seen in Fig. 1.4 at very low q^2 the virtual photon contribution, associated
²⁷⁴ with C_7 , dominates. In the region $1 - 6$ GeV^2/c^4 the interference between C_7 and
²⁷⁵ C_9 becomes large, yielding sensitivity to NP in C_9 . The $6 - 15$ GeV^2/c^4 interval is
²⁷⁶ dominated by charmonium resonances, J/ψ and $\psi(2S)$, though the tree level $b \rightarrow \bar{c}cs$
²⁷⁷ transition. Although the decays can be experimentally vetoed in principle charmonia
²⁷⁸ affect the entire q^2 space. Finally, at high q^2 broad charmonium resonances can
²⁷⁹ contribute, like those observed by LHCb in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays [].

²⁸⁰ 1.5.4 Observables in $b \rightarrow s\ell\ell$ decays

²⁸¹ Rare decays and especially semileptonic $b \rightarrow s\ell\ell$ processes offer a plethora of observ-
²⁸² ables which can be used to search for New Physics. The most direct effects appear
²⁸³ in decay rates that can be enhanced by NP but the precision on these measurements
²⁸⁴ is often by uncertainty of form factor calculations or charm loops. Therefore it is
²⁸⁵ important also to look for different observable. One important class of observables
²⁸⁶ are angular quantities that can carry information about NP, often complementary
²⁸⁷ to branching ratio measurements. The most basic of these observable are forward-
²⁸⁸ backward asymmetries that characterise the angular distribution of final particles.
²⁸⁹ For the $B^0 \rightarrow K^* \mu^+ \mu^-$ decay combination of observables have been proposed that
²⁹⁰ are independent of form factor uncertainties in the fits order \mathbb{O} . One more way to
²⁹¹ build stable observable is to construct ratios between similar decays in which, for
²⁹² example, uncertainties due to the hadronisation process cancel out. These observ-
²⁹³ ables include the R_H ratios, between B^0 decay into electrons and muons, that are
²⁹⁴ described in detail in Sec. 11.

²⁹⁵ 1.6 Experimental status

²⁹⁶ In order to set the background for the searches included in this thesis, in this section
²⁹⁷ a review of recent or important results of NP searches involving rare decays or lepton
²⁹⁸ flavour violation. Among these, results recently obtained by the LHCb experiment
²⁹⁹ show a series of anomalies with respect to the SM that have the potential to yield
³⁰⁰ to NP scenarios.

³⁰¹ 1.6.1 Dimuon decays of b hadrons

Decays of B mesons into two muons have been recently studies at the LHCb and CMS experiments. These are two-body decays where the two muons are back to back

in the hadron rest frame. The simple signatures of these decays 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 for the SM. The $B^0 \rightarrow \mu^+ \mu^-$ and $B^0_s \rightarrow \mu^+ \mu^-$ decays are exceedingly rare in the SM. First of all they can only happen in loops and furthermore they are CKM-suppressed. In addition to that the decay of a pseudo-scalar B meson into two muons has a significant helicity suppression. The latest SM predictions for these decay rates are []:

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

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

The uncertainties on these values mainly comes from the knowledge of the decay constants and CKM-elements. BSM models, for example models with extended Higgs sectors can produce significant enhancement to these decays. Furthermore the measurement of their ratio is a stringent test of the MFV hypothesis. A combination of the LHCb and CMS results resulted in the measured values []:

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

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

302 Both decays were unobserved and now the B^0_s decay was observed with a signifi-
303 cance of 6σ and evidence for the B^0 decay was found with a 3σ significance. These
304 are compatible with SM predictions within 2σ and put strong constraints to the
305 available parameter-space for BSM theories.

306 1.6.2 Semileptonic $b \rightarrow s\ell\ell$ decays of b hadrons

307 Many branching ratios of semileptonic B meson decays were recently measured at
308 the LHCb experiment, including $B \rightarrow K\mu^+\mu^-$, $B \rightarrow K^*\mu^+\mu^-$ and $B^0_s \rightarrow \phi\mu^+\mu^-$
309 []. Baryon decays were also studied at LHCb: including the branching ratio of the

³¹⁰ rare $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$ decay [], which is described in this thesis. For semileptonic decays,
³¹¹ unlike for dilepton decays, SM predictions are affected by the knowledge of hadronic
³¹² form factors, *describing phsyics bla bla*. This typically yields to uncertainties of
³¹³ $\mathcal{O}(30\%)$. As described in 1.5.4 many observables can be affected by

³¹⁴ 1.6.3 Lepton Flavour Violation searches

³¹⁵ Several LFV searches are linked to rare decays as they involve small branching ratios
³¹⁶ in the SM that can be enhanced by NP. They are therefore a natural place to look for
³¹⁷ NP. Lepton flavour conservation is well experimentally established but has no strong
³¹⁸ theoretical explanation and in fact we already know that flavour is not conserved in
³¹⁹ neutrino oscillations. In this section is reported a short review of Lepton Flavour
³²⁰ Violation searches. The best-studied decays violating lepton flavour are rare muon
³²¹ decays including $\mu^+ \rightarrow e^+ \gamma$ and $\mu^+ \rightarrow e^+ e^- e^+$. Since muons can be abundantly
³²² produced and the final states are simple, these decays provide the best constraints
³²³ to LFV. The present best-upper limits are 1.2×10^{-11} for the radiative decay and
³²⁴ 1.0×10^{-12} for $\mu^+ \rightarrow e^+ e^- e^+$ obtained respectively by the MEGA [] and SINDRUM
³²⁵ [] experiments.

³²⁶ Several LFV searches have been recently been performed at the LHCb experiment
³²⁷ including B meson decays such as $B^0 \rightarrow e \mu$ [] and τ decays such as $\tau \rightarrow \mu^+ \mu^- \mu$
³²⁸ []. None of these searches has found evidence of NP so far and therefore they set
³²⁹ limits, constraining the parameter space available for NP models. Fig. 1.6.3 reports
³³⁰ a summary of the best limits on LFV searches.

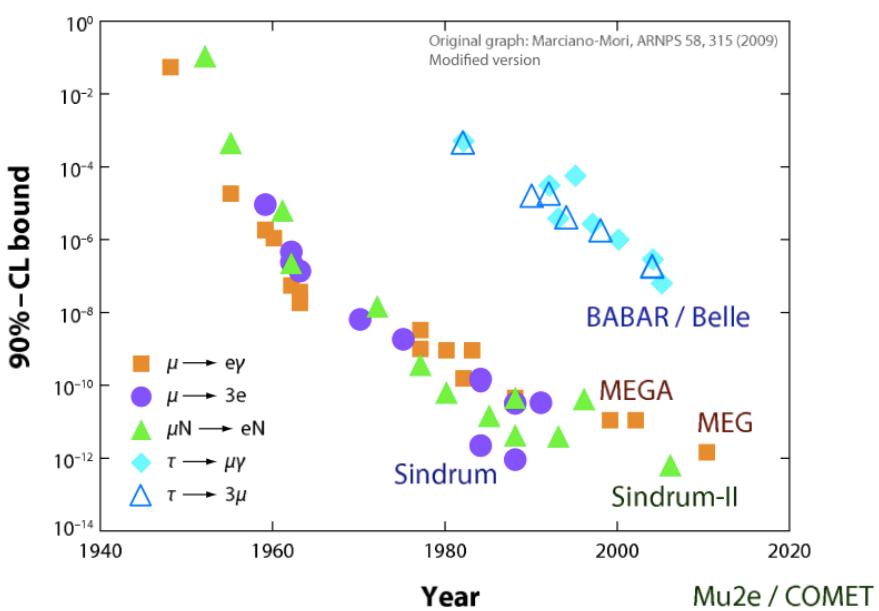


Figure 1.5: Summary of limits set in lepton flavour violation searches [].

CHAPTER 2

331

332

333

The LHCb detector at the Large Hadron Collider

334

335

2.1 The Large Hadron Collider

336 The Large Hadron Collider (LHC) is a circular particle accelerator with a circum-
337 ference of 27 km about 100 m underground. The two general-purpose detectors,
338 ATLAS and CMS, sit on opposite sides of the ring, while the two smaller specialty
339 detectors, ALICE and LHCb, are at the interaction points to either side of ATLAS
340 (see fig. 2.1).

341 Two proton beams circulate in opposite directions around the ring and cross each
342 other at several points, in which are placed huge particle detectors. Each beam
343 consists of a series of proton bunches, up to a maximum of 2835 in the beam. Each
344 bunch consists of about 10^{11} protons and the bunch spacing is such that the nominal
345 bunch crossing rate is 40 MHz. The beams are injected into pre-accelerators and then
346 led into LHC through the CERN acceleration system shown in figure 2.1. Protons

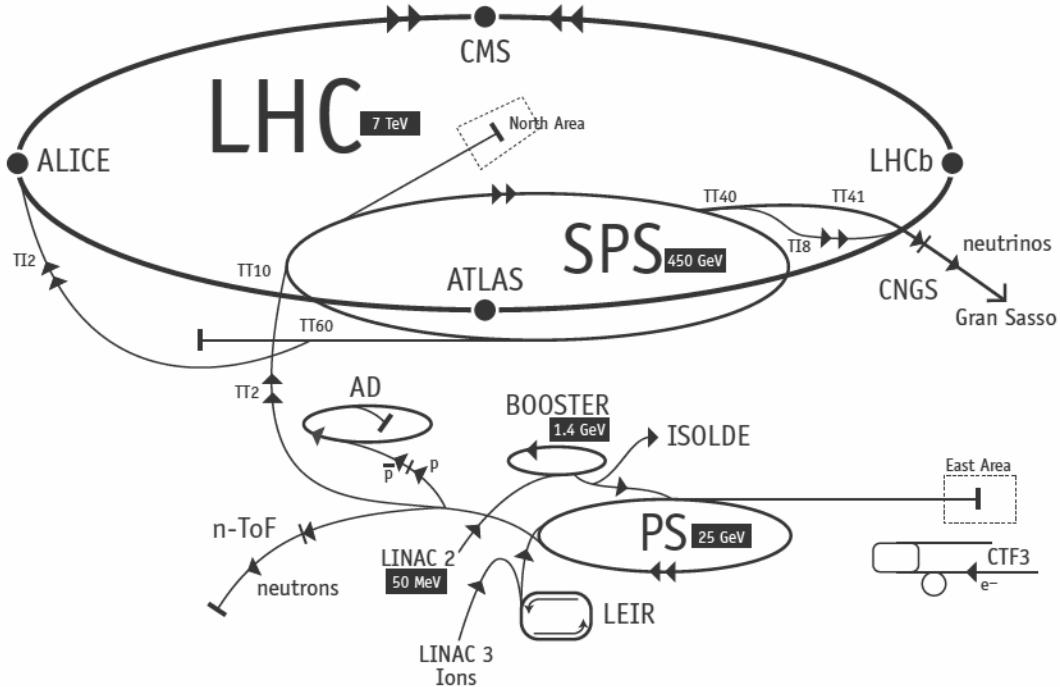


Figure 2.1: Scheme of CERN accelerators.

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

³⁶² maximum design energy is 14 TeV, and its design luminosity is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. After
³⁶³ a long shut down to upgrade and maintain the machine, a new run started in June
³⁶⁴ 2015 where protons are collided at an energy of $\sqrt{s} = 13 \text{ TeV}$ at this energy the total
³⁶⁵ proton-proton cross section is expected to be roughly 100 mb.

³⁶⁶ 2.2 The LHCb detector

³⁶⁷ The LHCb detector was built with the main purpose of studying the decays of B
³⁶⁸ and D mesons, looking in particular for CP-violating processes. In 2011, running at
³⁶⁹ a centre of mass energy of 7 TeV, the cross section of $b\bar{b}$ production was measured
³⁷⁰ to be $284 \pm 53 \mu b$ [?], while it will be $\sim 500 \mu b$ at the nominal LHC energy, 14 TeV.
³⁷¹ At these high energies, proton-proton interactions produce highly boosted virtual
³⁷² gluons which interact to produce $b\bar{b}$ pairs at small angles, close to the beam pipe. For
³⁷³ this reason the LHCb detector is designed to have a very forward angular coverage:
³⁷⁴ it is fully instrumented from approximately 10 mrad to 300 mrad, corresponding to
³⁷⁵ $2 < \eta < 5$, where η is a quantity used in particle physics and called “pseudorapidity”
³⁷⁶ and defined as:

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

³⁷⁷ In Eq. 2.1, θ is the angle between a particle’s momentum and the beam direction¹.

³⁷⁸ At the collision point of LHCb the luminosity can be adjusted by displacing the
³⁷⁹ beams from head on collisions while keeping the same crossing angle. This allows
³⁸⁰ the experiment to keep an approximately constant instantaneous luminosity. This
³⁸¹ also means that the average number of interactions per bunch crossing can be limited
³⁸² as LHCb efficiency, especially of detecting secondary vertices, decreases for events
³⁸³ with an high number of primary vertices (PV). Reducing the particle occupancy
³⁸⁴ through the detector also keeps radiation damage to a minimum. Since the LHC
³⁸⁵ started colliding protons in November 2009 until the end of 2011, the instantaneous

¹LHCb’s reference system has the z axis in the direction of the beam, the x axis directed to the centre of the accelerator and y is directed upward. Then we define θ as the angle with the beam direction and ϕ as the position around the beam in the xy plane, taking $\phi = 0$ on the x axis.

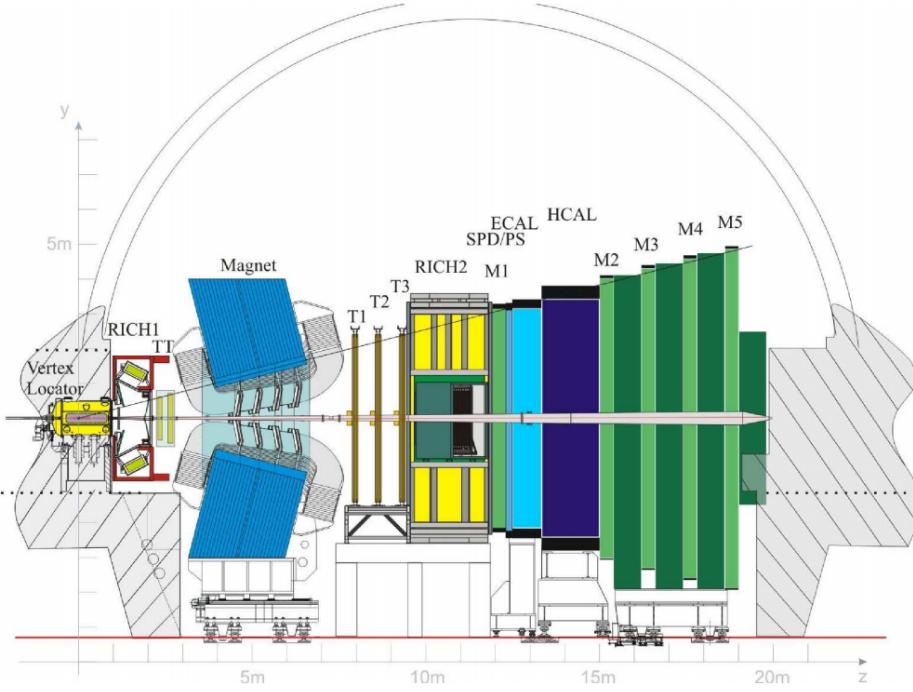


Figure 2.2: A side view of the LHCb detector [3].

386 luminosity was at an average of $3 \cdot 10^{32} \text{cm}^{-2}\text{s}^{-1}$, with an average number of 1.5
 387 vertices per bunch crossing in LHCb. At the end of 2011 LHCb had collected an
 388 integrated luminosity of 1 fb^{-1} ; in 2012 the luminosity was increased and 2 fb^{-1}
 389 more were collected.

390 Other B physics experiments, like BaBar at the Stanford Linear Accelerator (SLAC),
 391 Belle at KEK at J-PARC (Japan) and the Tevatron experiments at Fermilab have
 392 made accurate measurements in heavy flavour physics. All of these results have
 393 so far been consistent with the Standard Model predictions. However, some of the
 394 deviations from the Standard Model are expected to be very small, therefore LHCb
 395 has begun to make the most precise measurements in heavy flavour physics.

396 The LHCb detector[3] includes a high-precision tracking system consisting of a
 397 silicon-strip vertex detector surrounding the pp interaction region, a large-area silicon-
 398 strip detector located upstream of a dipole magnet with a bending power of about 4
 399 Tm, and three stations of silicon-strip detectors and straw drift tubes placed down-
 400 stream. The combined tracking system has momentum resolution $\Delta p/p$, that varies

401 from 0.4% at 5 GeV/c² to 0.6% at 100 GeV/c². Charged hadrons are identified us-
 402 ing two Ring-Imaging Cherenkov detectors (RICH)[6]. Photon, electron and hadron
 403 candidates are identified by a calorimeter system consisting of scintillating-pad and
 404 pre-shower detectors, an electromagnetic calorimeter and a hadronic calorimeter.
 405 Muons are identified by a system composed of alternating layers of iron and mul-
 406 tiwire proportional chambers[7]. A schematic view of the detector is shown in Fig.
 407 2.2.

408 2.2.1 Tracking system

409 The tracking system is made up of the Vertex Locator (VeLo), and 4 tracking sta-
 410 tions: the Tracker Turicensis (TT) which are located before the magnet and T1,
 411 T2 and T3 which are downstream of the magnet. Charged particle tracks are bent
 412 horizontally in the magnetic field so that their momentum can be measured from
 413 the curvature radius.

414 B mesons have lifetimes of approximately 1.5 ps. At the LHC energies, this means
 415 they travel about 1 cm before decaying and they form a displaced vertex. It is
 416 therefore important to be able to separate the particles produced at the primary
 417 $p - p$ vertex and the B decay vertex.

418 The VeLo accurately measures positions of tracks close to the interaction point so
 419 that production and decay vertices of bottom and charm hadrons can be recon-
 420 structed. The VeLo is made up of 21 staggered silicon modules which surround the
 421 beam axis and are positioned from $z = -18$ cm to +80 cm. It is able to detect
 422 particles within a pseudorapidity range $1.6 < \eta < 4.9$. The sensitive region of the
 423 VeLo starts at an inner diameter of 8mm from the beam axis. The VeLo is housed
 424 in its own vacuum vessel of thin aluminium foil which protects the vacuum of the
 425 beam pipe from any outgassing of the VeLo. VeLo stations consist of two modules,
 426 and each has two types of sensors: the ϕ -sensor which measures the azimuthal po-
 427 sition around the beam, and the R-sensor which measures the radial distance from

428 the beam axis. A sketch of the VeLo sensor is shown in Fig. 2.2.1. The sensors are
429 300 μm thick, approximately semicircular and are positioned on either side of the
430 beam axis. To ensure that they cover the full azimuthal angle the right-side module
431 is placed 1.5 cm behind the left-side module on the z-axis and they overlap. There
432 are two modules which cover the backward direction and were used as a veto for
433 multiple interactions in 2011, called the pileup veto.

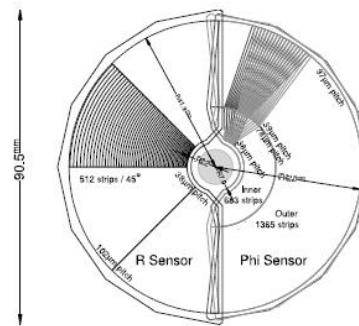
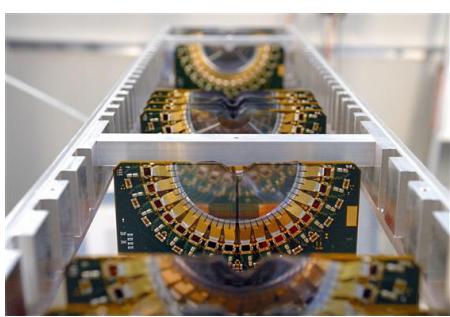


Figure 2.3: On the left Velo sensors mounted in line and on the right a schematic view of one sensor [3].

434

435 The LHCb dipole magnet is comprised of two coils supported on an iron yoke and
436 is wedge-shaped to fit the LHCb angular acceptance. It is a warm magnet so can be
437 ramped easily and the field can be reversed periodically. This is used to limit some
438 systematics that can arise from imperfections performance in different areas of the
439 detector.

440 The IT and TT both use silicon microstrip and together constitute the Silicon
441 Tracker (ST). Straw tubes are used in the outer regions of the tracking stations
442 which together are called the Outer Tracker (OT). The IT has a higher inner gran-
443 ularity because of the higher flux of particles nearer the beam pipe. Each ST station
444 has four detection layers, the first and last being vertical, measuring the track po-
445 sition in x. The second and third layer are rotated by a stereo angle of +5 and
446 -5 degrees, which allows the y-coordinate to be measured. The TT is placed up-
447 stream of the magnet which allows reconstruction of the tracks from low-momentum

⁴⁴⁸ particles which are swept out of the downstream acceptance.

⁴⁴⁹ 2.2.2 Particle identification

⁴⁵⁰ Particle identification in LHCb is performed in various ways. The calorimeter detects
⁴⁵¹ particles with high transverse momentum, the muon chambers identify muons and
⁴⁵² the Ring Imaging Cherenkov (RICH) detectors identify heavier charged particles.

⁴⁵³ 2.2.2.1 Calorimeters

⁴⁵⁴ The main purpose of the calorimeter system is to determine the energy of par-
⁴⁵⁵ ticles traversing the detector. The material in the calorimeter system is layered
⁴⁵⁶ with absorber and active material. The absorber makes particles interact and pro-
⁴⁵⁷ duces a cascade of secondaries, which multiply quickly and are detected by the
⁴⁵⁸ active part. The sensitive material consists of scintillating layers, where the light
⁴⁵⁹ detected is approximately proportional to the number of deposited particles. Cal-
⁴⁶⁰ibration is then used to calculate the deposited energy. The calorimeter system
⁴⁶¹ is essential for flavour tagging because it identifies electrons. In addition, it is re-
⁴⁶²quired for accurately reconstructing π^0 particles and prompt photons, which are
⁴⁶³ both needed for the study of B-meson decays. The LHCb calorimeter system con-
⁴⁶⁴sists of the Scintillator Pad Detector (SPD), the Pre-Shower Detector (PS) as well as
⁴⁶⁵ the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL).
⁴⁶⁶ All four detectors transmit scintillation light via wavelength-shifting fibres to photo-
⁴⁶⁷ multiplier tubes (PMTs). The SPD/PS cells are read out with MAPMTs (Multi-
⁴⁶⁸anode PMTs) located outside the LHCb acceptance. The ECAL and HCAL have
⁴⁶⁹ individual MAPMTs located on the modules. All four detectors vary the segmen-
⁴⁷⁰tation of their cells according to the distance from the beam pipe. The purpose of
⁴⁷¹ the SPD and PS is to separate the electrons from a high background of neutral and
⁴⁷² charged pions produced in the collisions. In order to obtain the highest energy reso-
⁴⁷³lution the showers from high energy photons must be fully absorbed. For this reason

⁴⁷⁴ the ECAL has a thickness of 25 radiation lengths and its resolution is measured to
⁴⁷⁵ be [3]

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

⁴⁷⁶ The trigger requirements on the HCAL resolution do not depend on the containment
⁴⁷⁷ of the hadron showers as much as for the ECAL, so due to a limited space, its
⁴⁷⁸ thickness is only 5.6 interaction lengths and its resolution

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

⁴⁷⁹ 2.2.2.2 RICH

⁴⁸⁰ The two RICH detectors are a special feature of LHCb, as it is the only experiment
⁴⁸¹ at LHC including them. These detectors take advantage of the Cherenkov light
⁴⁸² produced by particles passing in a medium with velocity higher than the velocity of
⁴⁸³ light in the medium. The Cherenkov light, as shown in Fig. 2.2.2.2, is produced in
⁴⁸⁴ cones with a specific angle depending on the velocity of the particle

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

⁴⁸⁵ where β is the velocity of the particle over c and n is the refraction index of the
⁴⁸⁶ medium.

⁴⁸⁷

⁴⁸⁸ RICH 1 is situated before the magnet in order to cover a large angular acceptance.
⁴⁸⁹ Its purpose is to ensure particle identification over the momentum range $1 < p < 70$
⁴⁹⁰ GeV. It uses two radiators, C_4F_{10} covers the momentum range $5 - 70$ GeV/c, and

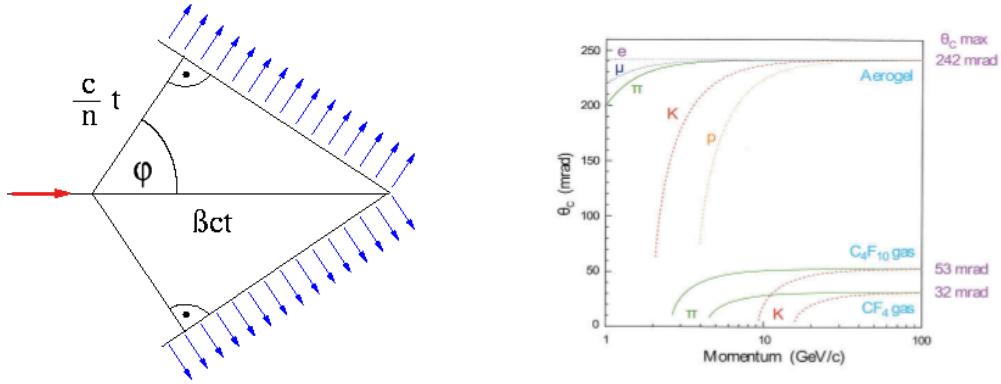


Figure 2.4: On the left a sketch of Cherenkov light emission [?] and on the right the Cherenkov angle versus momentum for the two radiators of RICH1 and for different particles. One can see that they allow to separate particles in different momentum ranges.

silica aerogel which covers $1 - 10 \text{ GeV}/c$. RICH 2 is situated after the magnet and tracking stations. It identifies higher momentum particles from approximately 20 GeV up to beyond 100 GeV using CF_4 as a radiator. The Cherenkov light produced when charged particles travel through the radiators, is reflected and focused using flat and spherical mirrors which are tilted so that the ring image is reflected onto arrays of photo-detectors. The radius of the ring becomes equivalent to 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.

For particle identification a particle type hypothesis is assigned to each charged track found in the tracking stations. Initially the hypothesis is for a pion, which is the most common particle type. The corresponding expected number and Cherenkov radii of the resulting photons are calculated and the likelihood is calculated. The hypothesis is then changed and the likelihood is recalculated. The case with the largest increase in likelihood is kept.

⁵⁰⁷ 2.2.3 The muon system

⁵⁰⁸ It is essential for many of the key physics analyses to be able to identify muons in
⁵⁰⁹ the final state. Muons are the most penetrating particles that can be detected at
⁵¹⁰ LHC experiments, so the muon chambers are the final subdetectors. There are five
⁵¹¹ stations (M1 - M5), the first one being located before the calorimeter in order to
⁵¹² improve the p_T measurements. A scheme of the muon system is shown in Fig. 2.2.3.

⁵¹³ The remaining four lay behind the HCAL and are separated by 1.2 m from each
⁵¹⁴ other, interleaved with iron block filters 80cm thick, which absorb hadrons, elec-
⁵¹⁵ trons and photons to ensure that only muons reach the final muon station. Only
⁵¹⁶ muons with a minimum momentum of 10 GeV/c traverse all of the five stations and
⁵¹⁷ for positive identification of a muon the trigger requires a signal in each of them.
⁵¹⁸ Each station has a detection efficiency of at least 95% and the detectors provide
⁵¹⁹ position measurements. Since there is a larger particle flux towards the beam pipe,
⁵²⁰ the stations are divided into four concentric rectangular regions (R1-R4), their size
⁵²¹ increasing according to the ratio 1 : 2 : 4 : 8. This means that there is a similar
⁵²² channel occupancy over the four regions. All of the muon stations use Multi Wire
⁵²³ Proportional Chambers (MWPC) except for the inner region of M1, where the par-

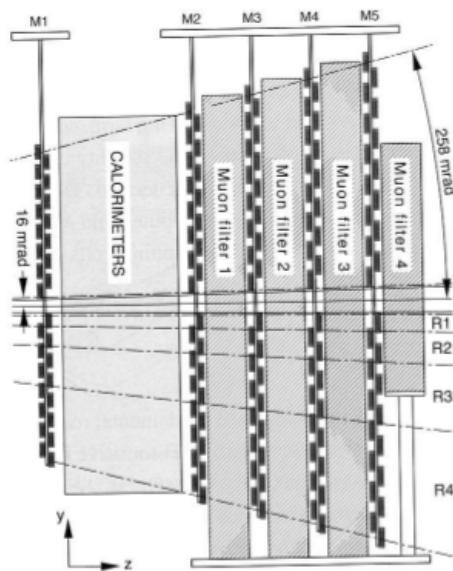


Figure 2.5: The LHCb muon system [3].

524 ticle flux is too high. In this region triple-GEM (Gas Electron Multiplier) detectors
525 are used instead because they have better ageing properties.

526 The Gas Electron Multiplier (GEM) detectors in the inner region of M1 have to
527 withstand a rate up to 500 kHz cm^{-2} of charged particles. Particles traversing
528 through the drift gap between the cathode and the first GEM foil produce ionisation
529 electrons which are then attracted by electric fields though all of the GEM foils and
530 they multiply. They then drift into the anode inducing a signal on the pads. A gas
531 mixture of Argon, CO_2 and CF_4 , is used to give a time resolution better than 3 ns.

532 2.2.4 Trigger and software

533 The LHCb trigger system[8] consists of a hardware stage (L0), based on information
534 from the calorimeter and muon systems, followed by a software stage (HLT), which
535 applies a full event reconstruction. To increase performances the HLT is split again
536 into stages (HLT1 and HLT2). The bunch crossing frequency is 40 MHz, which
537 corresponds to an instantaneous luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ for LHCb, and about
538 15% of the total number of $b\bar{b}$ pairs produced will have at least one B meson with
539 all of its decay products within the detector acceptance. This needs to be reduced
540 down to about 2 kHz so that the events can be written to disk for analysis. Fig.
541 2.2.4 shows a scheme of the trigger system.

542 The L0 reduces the rate of visible interactions from 10 MHz to a rate of 1 MHz
543 and uses mainly the information from the calorimeter dividing the events in the 5
544 categories: L0Photon, L0Electron, L0LocalPion, L0GlobalPion, L0Hadron. “local”
545 pions refer to π^0 reconstructed though decay in $\gamma\gamma$, where the two photons fall in the
546 same ECAL board, they are labelled “global” otherwise. The HLT1 uses information
547 from the VELO and trackers performing a partial reconstruction of the event and
548 reduces the rate to 2 kHz. Finally the HLT2 involves a full reconstruction of the
549 event and includes many “lines” designed to trigger specific decays.

550 LHCb also developed an extended simulation software in order to reconstruct ef-

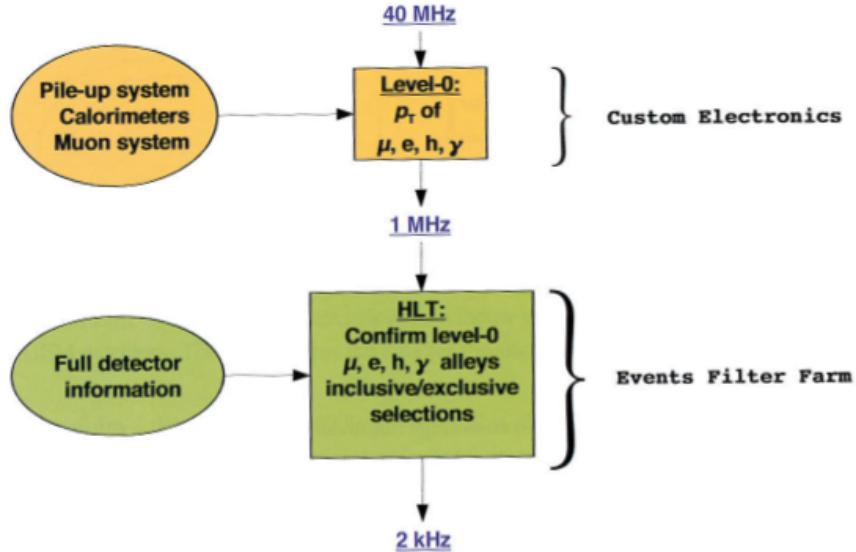


Figure 2.6: Scheme of the LHCb trigger system [3].

551 efficiencies and signal shapes. In the simulation, pp collisions are generated using
 552 PYTHIA 6.4[9] with a specific LHCb configuration[10]. Decays of hadronic parti-
 553 cles are described by EVTGEN[11], in which final state radiation is generated using
 554 PHOTOS[12]. The interaction of the generated particles with the detector and its
 555 response are implemented using the GEANT4 toolkit[13, ?] as described in [14].

556 For the analysis in this document, I used the ROOT framework[?] to analyse data,
 557 the RooFit package for fitting. The multivariate analysis is based on the NeuroBayes
 558 package [] which provides a framework for neural network training.

559

Part I

560

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

561

562

CHAPTER 3

563

564

Introduction

565

The rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decay is a FCNC decay governed by the $b \rightarrow s\mu^+\mu^-$ quark level transition. In the SM this decay proceeds through electroweak penguin and W box diagrams (see Fig. 1.3). Since this process happens only through loop diagrams, it is highly sensitive to new particles entering the loops. Moreover, as final state contains only a single long-lived hadron, the hadronic physics is easier to handle than in fully hadronic decays.

Interest in Λ_b^0 baryon decays arises from two important facts. First of all, as Λ_b^0 has non-zero initial spin, there is a potential to learn information about the helicity structure of the underlying Hamiltonian, which cannot be extracted from the meson decays [15, 16]. Second, as Λ_b^0 baryon is in first approximation composed of heavy quark and diquark formed of light quarks the hadronic physics significantly differs from that of the mesons. This itself provides possibility to better understand and test the hadronic physics in the theory, which could yield improved understanding and confidence also for mesons.

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

Theoretical aspects of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays were considered by a number of authors both in the SM and in different new physics scenarios [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]. All authors start from the same effective Hamiltonian already described in Sec. 1.5.1. However, form factors, describing hadronic physics, are not developed as well as in the meson case. Since there are not as many experimental constraints and form factors are still not well understood this leads to a relatively large spread in predicted branching fractions. An interesting quantity to study is the differential branching fraction as function of q^2 . This still suffers from knowledge of form factors, but as different approaches to form factors are applicable in different q^2 regions, this allows a more meaningful comparison to theory.

Experimentally, the decay $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ was observed for the first time in 2011 by the CDF collaboration [28], with signal yield of 24 ± 5 signal events. Later this was updated using the full CDF statistics [29]. Their preliminary result on full statistics yields $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-) = [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}$. CDF observed the signal only in the q^2 region above the square of the $\psi(2S)$ mass. Recently, the decay was observed also at LHCb [30] with a yield of 78 ± 12 signal events using 1 fb^{-1} of integrated luminosity collected in 2011. The signal was again found only in the high q^2 region. The LHCb result for the branching fraction relative to the $J/\psi \Lambda$ decay, used as normalisation channel, is

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

and for absolute branching fraction

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

597 This parts of the thesis describes the measurement of the differential branching
598 fraction of the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ normalised by $J/\psi \Lambda$ using 3 fb^{-1} of pp collisions
599 collected in 2011 and 2012. Furthermore an angular analysis of these decays is
600 performed, measuring observables including the forward-backward asymmetries in
601 the leptonic and hadronic systems.

602 3.1 Analysis strategy and q^2 regions

603 A typical q^2 spectrum of $b \rightarrow s\ell\ell$ decays was shown in Fig. ???. This is characterised
604 by the presence of the narrow peaks of the J/ψ and $\psi(2S)$ resonances. For this
605 analysis two regions are defined: the “low q^2 ” region, below the J/ψ resonance
606 ($q^2 < 8 \text{ GeV}^2/c^4$), where the signal is unobserved, and the “high q^2 ” region, above
607 the J/ψ resonance ($q^2 > 11 \text{ GeV}^2/c^4$). The decay $\Lambda_b^0 \rightarrow J/\psi \Lambda$, where J/ψ decays
608 into two muons, which had same final states as the signal, is used as a normalisation
609 channel and the branching fraction measurement is given in relative form fo limit
610 systematic uncertainties. In both cases the Λ decay mode into a pion and a proton,
611 $\Lambda \rightarrow p\pi$, is used to reconstruct the decays. The rare and normalisation channels
612 are naturally distinguished by the q^2 interval the fall into. The regions in which the
613 rare channel is studied include:

- 614 • $0.1 < q^2 < 8 \text{ GeV}^2/c^4$, where the selection is optimised to observed the signal
615 as explained in Sec. 4.2. The upper bound of this interval was chosen to be suf-
616 ficiently far from the J/ψ radiative tail at low masses, that could contaminate
617 the rare sample;
- 618 • $11 < q^2 < 12.5 \text{ GeV}^2/c^4$ in between two charmonium resonances and

619 $q^2 > 15 \text{ GeV}^2/c^4$, above $\psi(2S)$. In these two intervals the selection is optimised
620 to maximise the yield which is particularly important for a stable angular
621 analysis.

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

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

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

- 626 • 1.1-6.0 GeV^2/c^4 : this interval is theoretically clean since is far from the photon pole, which dominated at low q^2 , washing out the sensitivity to NP contributions. The lower bound of this interval is chosen excludes the possible contribution from the ϕ resonance, which appears at 1 GeV^2/c^4 . The upper bound of the interval is chosen to totally exclude a small contribution from the J/ψ resonance that leaks below 8 GeV^2/c^4 .
- 632 • 15.0-20.0 GeV^2/c^4 : this interval is the one that contains most of the statistics and it is used as a natural cross check that the analysis in smaller bins is stable.

635 3.2 Candidate types

636 This analysis deals with Λ baryons, which have a lifetime of $(2.632 \pm 0.020) \times 10^{-10}$
637 s [31]. These are considered long-lived particles in particle physics terms and can
638 travel into the detector for several meters generating well distinguished secondary
639 vertices. In LHCb Λ baryons can be reconstructed from tracks with out without
640 hits in the VELO and therefore with define two candidates types as follows:

- 641 • **Long candidates:** built from tracks which have hits in the VELO, “long
642 tracks”. These candidates, also denoted as “LL”, are characterised by a bet-
643 ter momentum resolution thanks to the longer leverage arm available to long
644 tracks.
- 645 • **Downstream candidates:** built from tracks without hits in the VELO,
646 “downstream tracks”, also denoted as “DD”.

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

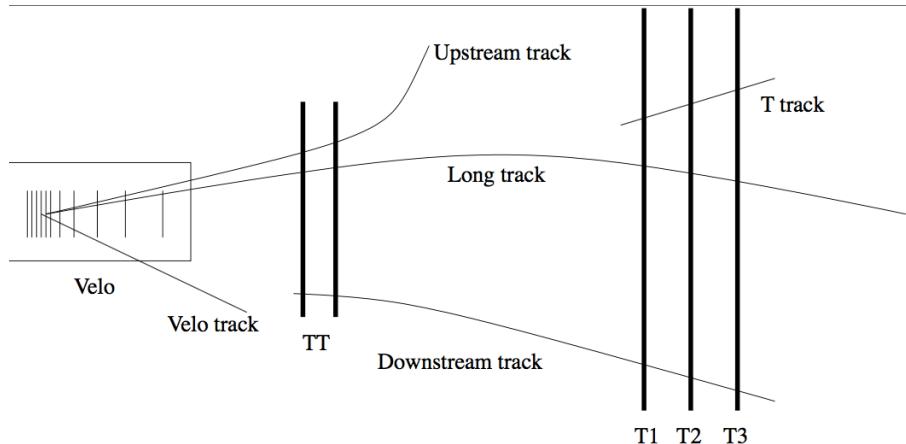


Figure 3.1: Representation of the two Λ candidate types built from “long” and “downstream” tracks.

CHAPTER 4

652

653

654

Selection

655

656 The reconstruction of $\Lambda_b^0 \rightarrow \Lambda \text{ mm}$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates begins with the
657 application of requirements on basic kinematic properties as p_T of the final particles
658 and quality requirements for the track and vertices from which the particles are
659 derived. This selection aims to first form a dimuon candidate from two oppositely
660 charged muons. Then, in events containing a dimuon candidate, two oppositely
661 charged tracks are combined together and retained as Λ candidate, if they form a
662 good vertex which is well separated from all primary vertices. In the final step,
663 dimuon candidates are combined with Λ candidates to form Λ_b^0 baryons and require-
664 ments are set on the properties of this combination. The full list of cuts is reported
665 in Tab. 4.1.

666 In the table χ_{IP}^2 is defined as the projected distance from the vertex divided by its
667 uncertainty, for example $B^0 \chi_{\text{IP}}^2(\text{primary}) > 4$ means that the B^0 vertex is at least
668 2 standard deviations away from the primary vertex. Another useful quantity to
669 remove combinatorial background is a pointing variable defined as the angle between

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

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

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

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

quantifies the probability of a particle being a kaon rather than a pion. For details about the definition of the variables used see Ref. [32]. A large mass window is kept around the Λ_b^0 peak in order to be able to fit the sideband, to train the multivariate analysis and better constrain backgrounds. Rare candidates are simply selected by the q^2 region requirements as described in Sec. 3.1 while resonant candidates are further constrained to have dimuon invariant mass in a $100 \text{ MeV}/c^2$ interval around the known J/ψ mass.

4.1 Neural Networks

The final selection is performed using neural network (NN) based on NeuroBayes package [33, 34]. The input to the neural network consists of 14 variables carrying

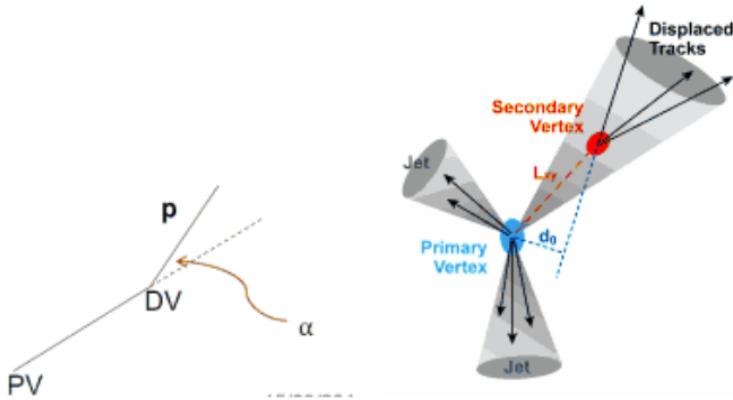


Figure 4.1: Graphical representation of the DIRA (left) and χ^2_{IP} (right) variables.

695 information about the kinematics of the decay, the quality of tracks and vertices
 696 and the PID of the muons. The list of inputs to the neural network is in table
 697 4.2. Together with it we give ranking and information on the importance of inputs.
 698 Variables related to Λ and its daughters are considered as different inputs depend-
 699 ing on the candidate type (long or downstream). This effectively corresponds to
 700 making a separate training for the two categories. Further details on the definition
 701 and calculation of the variables importance is available in Ref. [35]. The graphical
 702 representation of the correlation matrix is shown in Fig. 4.2, where the variable with
 703 $ID = 1$ is the NN output and the IDs of the other variables can be found in Tab.4.2.

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

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

715 only combinatorial background, while the lower sideband may contain partially re-
 716 constructed and misreconstructed events. In the background samples are still present
 717 J/ψ and $\psi(2S)$ peaks indicating that charmonium resonances can be combined with
 718 other random tracks. These candidates do not give a good description of purely
 719 combinatorial background and, in order to avoid biases, they are removed from the
 720 background training sample by rejecting events in a 100 MeV interval around the
 721 nominal J/ψ and $\psi(2S)$ masses ???. For the signal, the training is done combining
 722 simulated $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ events corresponding to beam conditions in the two years
 723 respectively. A total of 30000 total events was used for the taining. This corresponds
 724 $\sim 50\%$ of the available sideband data sample and $\sim 20\%$ of our full MC sample.
 725 The full simulated sample is not used as it will also be used to study efficiencies.
 726 Events are uniformly samples over full simulated sample.

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

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

Table 4.2: Summary of inputs to the neural network in order of importance. Under “Id” the indices used for the correlation matrix (see Fig. 4.2) are reported. Column “adds” gives correlation significance added by given input when adding it to list of those ranked above, “only this” provides power of given input alone and “loss” shows how much information is lost when removing only given input.

Input	Id	adds	only this	loss
$\Lambda_{DD} p_T$	15	143.11	143.11	29.20
χ^2 DecayTreeFitter	2	77.81	134.00	51.10
$\min(\chi^2_{IP} \mu)$	7	61.31	113.62	29.76
$\chi^2_{IP} \Lambda_b^0$	5	52.94	113.23	40.98
$\chi^2_{IP} \pi_{LL}$	16	20.29	60.72	12.82
$\min(\text{PID } \mu)$	8	17.91	59.11	13.44
$\tau_{\Lambda_b^0}$	3	16.24	35.36	11.24
Λ_b^0 DIRA	4	12.28	73.96	9.98
Λ_{DD} flight distance	14	9.47	86.75	11.24
$\chi^2_{IP} \Lambda_{DD}$	13	10.58	59.84	8.88
$\max(\chi^2_{IP} \mu)$	6	9.51	97.24	8.15
$\chi^2_{IP} \Lambda_{LL}$	10	7.31	54.27	10.32
$\max(\text{PID } \mu)$	9	6.99	69.33	6.87
$\pi_{LL} p_T$	18	6.13	47.03	7.12
$\Lambda_{LL} p_T$	12	5.58	49.64	5.86
$\chi^2_{IP} p_{LL}$	17	4.48	53.01	4.18
$\chi^2_{IP} p_{DD}$	20	3.43	55.09	3.31
Λ_{LL} flight distance	11	0.87	52.52	0.86
$p_{DD} p_T$	21	0.74	129.58	0.75
$\chi^2_{IP} \pi_{DD}$	19	0.24	70.43	0.24

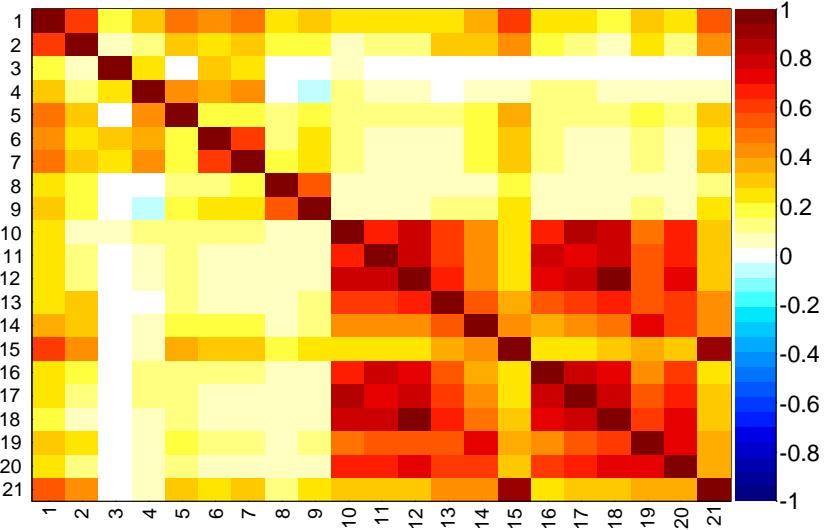


Figure 4.2: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of table ???.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

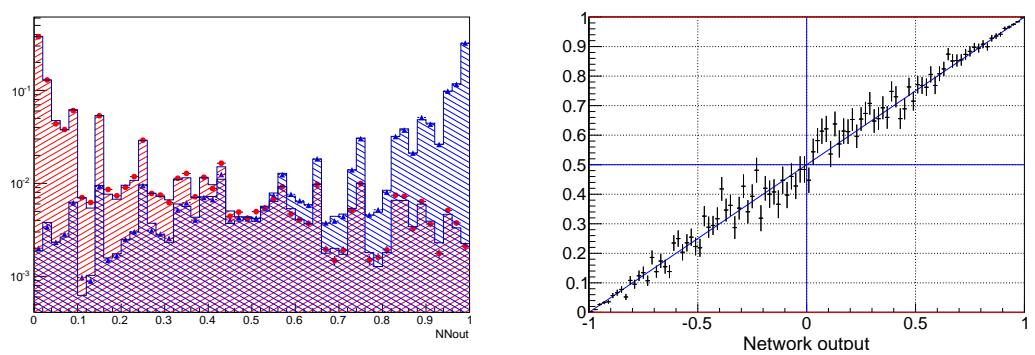


Figure 4.3: (right) NN output distribution for training (solid) and test (stripes) samples, for MC signal and sideband data events. (left) Purity as a function of neural network output.

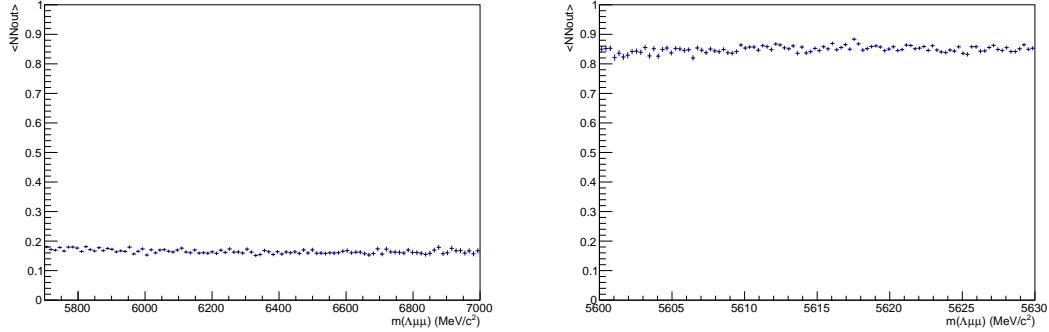


Figure 4.4: Average value of NN output as a function of Λ_b^0 mass for data sideband (left) and MC signal (right) events.

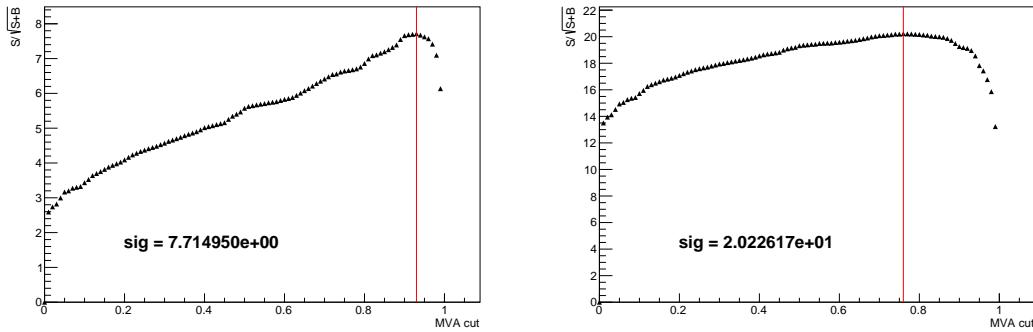


Figure 4.5: Dependence of figure-of-merit on the requirement on neural network output in the low q^2 region (left) and high q^2 (right) regions. The vertical line corresponds to the chosen cut.

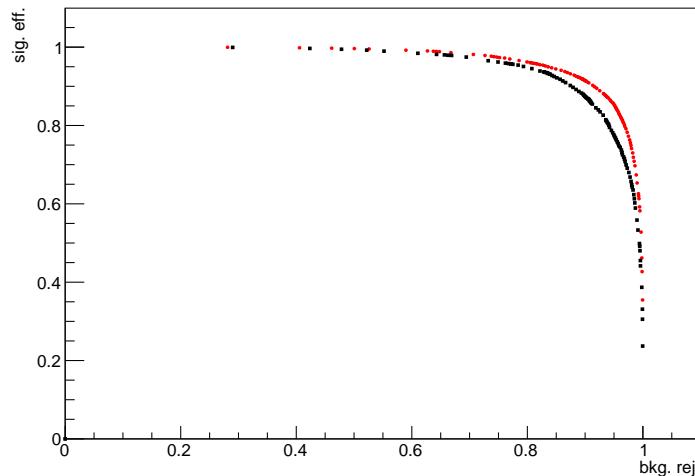


Figure 4.6: Receiver operating characteristic (ROC) curves for low q^2 (black) and high q^2 (red). They show the signal efficiency versus the background rejection.

741 4.2 MVA optimization

742 In the high q^2 region, where the signal is already observed, the final requirement
 743 on the neural network output is chosen in order to maximise the significance,
 744 $N_S/\sqrt{N_S + N_B}$, where N_S is number of expected signal candidates and N_B the num-
 745 ber of expected background candidates. N_S is derived from simulation but, as an
 746 arbitrary number of events can be generated, it needs to be normalised. To do this,
 747 the invariant mass distribution of real $\Lambda_b^0 \rightarrow J/\psi \Lambda$ candidates is fit after preselection
 748 (including all requirements but MVA). This is possible as the peak of the resonant
 749 channel is already well visible after preselection. Then the resonant yield is scaled
 750 by the ratio of between the $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ branching fractions as
 751 measured by LHCb on 2011 data

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

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

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

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

755 In the low q^2 region, where the signal is unobserved, the so called ‘‘Punzi figure
 756 of merit’’, $N_S/(n_\sigma/2 + \sqrt{N_B})$, is maximised [36]. This figure-of-merit is considered
 757 to be optimal for discovery and the parameter with n_σ corresponds to the number
 758 of expected standard deviations of significance, in this analysis $n_\sigma = 3$ is used.
 759 Moreover the Punzi shape does not depend on the relative normalisation between
 760 signal and background, which is important since the signal is still unobserved at
 761 low q^2 and existing predictions vary significantly for this region. The dependence
 762 of the figure-of-merit for both q^2 regions are shown in Fig. 4.5, and curves of signal
 763 efficiency versus background rejection are shown in Fig. 4.6.

Table 4.3: Summary of trigger lines which candidates have to pass at various trigger levels. Trigger is always required to be due to tracks of the candidate itself.

Trigger Level	Lines
L0	L0Muon
	L0DiMuon
Hlt1	Hlt1TrackAllL0
	Hlt1DiMuonHighMass
	Hlt1TrackMuon
Hlt2	Hlt2Topo2BodyBBDT
	Hlt2Topo3BodyBBDT
	Hlt2Topo4BodyBBDT
	Hlt2TopoMu2BodyBBDT
	Hlt2TopoMu3BodyBBDT
	Hlt2TopoMu4BodyBBDT
	Hlt2SingleMuon
	Hlt2DiMuonDetached

764 For final selection the neural network output is required to be larger than 0.81 for
 765 high q^2 region and 0.96 for the low q^2 one. Using these requirements the neural
 766 network retains approximately 96% (66 %) of downstream candidates and 97 %
 767 (82 %) of long candidates for the selection at high (low) q^2 , with respect to the
 768 preselected event sample. After the full selection $\sim 0.5\%$ of the events contain
 769 multiple candidates which are randomly rejected to keep only one candidate per
 770 event. To normalise the branching ratio measurement J/ψ events are selected using
 771 the low and high q^2 selection to normalise respectively low and high q^2 intervals.

772 4.3 Trigger

773 In addition specific trigger lines are selected, corresponding to events triggered by the
 774 muons of the reconstructed candidate. This is denoted as Trigger On Signal (TOS).
 775 The trigger lines used in the analysis are shown in Tab. 4.3. The logical *or* of lines
 776 on the same lever is required and the logical *and* and lined in different levels. The
 777 L0Muon trigger required hits in the muon detector. The Hlt1TrackAllL0 performs
 778 a partial reconstruction of the events end triggers if the L0 decision is confirmed.

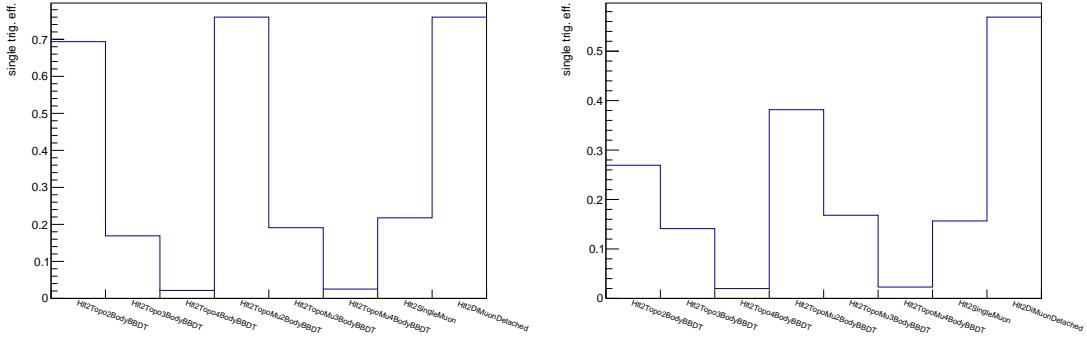


Figure 4.7: Single trigger efficiency for high q^2 events (left) and low q^2 (right).

Finally, at the Hlt2 level, a complete reconstruction is done and a multivariate analysis is used to identify decay structures. In Fig. 4.7 is shown the single trigger efficiency, defined as if each line was alone. More information on trigger simulation and efficiencies are contained in Sec. 5.3.7.

783 4.4 Background from specific decays

784 A survey of possible peaking backgrounds concluded that the only physics back-
785 ground to take into account is coming from misreconstructed decays of B^0 to K_s^0
786 with two muons, whether via J/ψ or not. The lack of background from other decays
787 is mainly due to the particular topology of the Λ decay which has a displaced vertex.
788 In order to study the effect of misreconstructed $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$
789 decays simulated samples are used, where the K_s^0 is reconstructed as a Λ with a
790 $p \rightarrow \pi$ identity swap and $m(p\pi)$ in the Λ mass window. On data the $B^0 \rightarrow J/\psi K_s^0$
791 contribution is clearly visible in the resonant channel mass distribution. This back-
792 ground is not suppressed with specific cuts in this analysis as its mass shape is
793 sufficiently distinct the from Λ_b^0 signal, which allows to reliably model its contribu-
794 tion in the mass fits (see Sec. 5.1). For rare case a rough estimate of the size is made
795 using the yield in the resonant channel rescaled the measured ratios between the rare
796 and resonant branching ratios. Details are given in Sec. 5.1 and numbers of events
797 predicted are reported in Tab. 5.1. This contribution, although close to negligible
798 is again considered in the fit. A possible pollution due to $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ decays,
799 where the K^{*+} further decays into $K_s^0 \pi$ is also investigated using a dedicated Monte
800 Carlo sample and found to be negligible. Finally, $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events radiating pho-
801 tons from the final state, can escape the J/ψ veto and be reconstructed in the rare
802 channel. The only contribution observed using simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events ends
803 up in the interval just below the J/ψ exclusion region, between 7 and 8 GeV/c^2 in
804 q^2 . Given that this contribution does not contribute in the region where we expect
805 Λ_b^0 peak, we do not attempt to exclude this but it is again modelled in the fit. In
806 Fig. 4.8 is reported the invariant mass distribution of simulated $\Lambda_b^0 \rightarrow J/\psi \Lambda$ events
807 falling into the rare q^2 region.

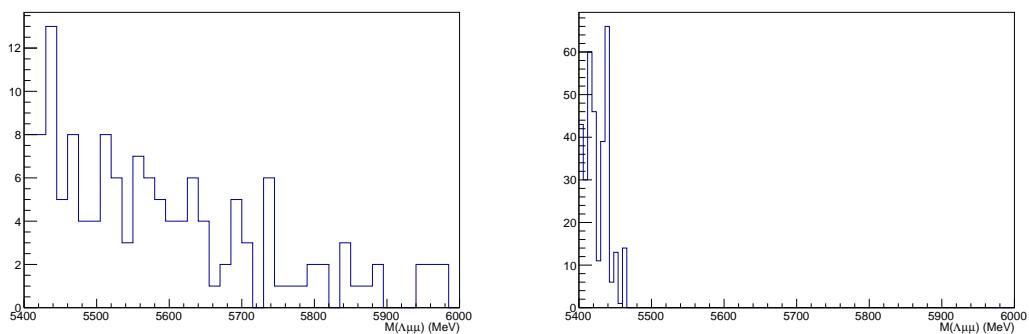


Figure 4.8: Inariant mass distribution of fully slected $B^+ \rightarrow \mu^+ \mu^- K^{*+}$ MC events (left) and of generator level events from J/ψ radiative tail in the region $q^2 < 8$ GeV^2/c^4 (right).

808

CHAPTER 5

809

810

Yield extraction

811

5.1 Fit description

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

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

814 where N_S , N_C and N_B are number of signal, combinatorial and K_s^0 background events
815 and $P_i(m_i)$ is the corresponding probability density function (PDF). From now on
816 when we refer to the invariant mass of the Λ_b^0 system we use the value obtained
817 from a kinematical fit of the full decay chain in which each particle is constrained
818 to point to its assigned origin vertex and the invariant mass of the $p\pi$ system is
819

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

824 For the resonant channel the signal is described as a sum of two Crystal Ball functions
 825 (CB) with common mean (m_0) and tail slope (n). A Crystal Ball function [37] is
 826 a probability density function commonly used to model various processes involving
 827 energy loss. In particular it is used to model the radiative tail which can be seen in
 828 many resonances' peaks. This function consists of a Gaussian core and a power-law
 829 tail, below a certain threshold. The function itself and its first derivative are both
 830 continuous and has form

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

831 where for normalisation and continuity

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

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

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

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

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

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

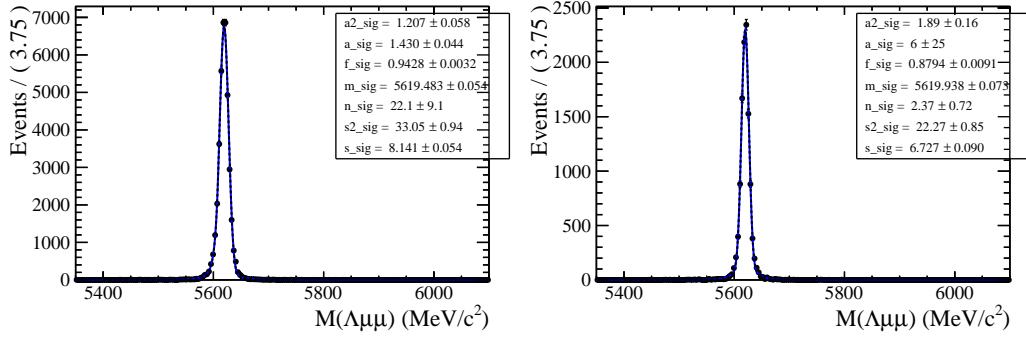


Figure 5.1: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows simulated data and the blue line is the signal fit function.

are modelled: the combinatorial background, parameterized by an exponential and the background from $B^0 \rightarrow J/\psi K_s^0$ decays. The K_s^0 background is described using the shape obtained using a $B^0 \rightarrow J/\psi K_s^0$ simulated sample and applying to it the full selection. The invariant distribution of these events is fit with a Double Crystal Ball function, which is then used to model the K_s^0 background in the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ fit. The fit to the simulated misreconstructed $B^0 \rightarrow J/\psi K_s^0$ events is reported in Fig. 5.2. When the K_s^0 shape is introduced in the final fit all parameters are fixed. This is particularly important when fitting long-long events, where the K_s^0 peak is less evident, which does not allow to constrain many parameters. On the other hand, in order to take in account possible data-simulation differences, an horizontal shift is added and left floating (by adding a constant to the central value, m_0 of the DCB). In summary, the free parameters in the fit to the resonant $\Lambda_b^0 \rightarrow J/\psi \Lambda$ sample are the yields of the signal and the combinatorial and K_s^0 backgrounds, the slope of the exponential and the horizontal shift of the K_s^0 shape. Notice that all parameters of the fit to the long and downstream samples are independent.

Finally, the rare $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ sample is fit. In this case a simultaneous fit to the long and downstream samples is performed to obtain a more stable convergence. In this fit the signal is modelled with the same shape used in the resonant case as there is no physical reason why they should be different. This method is also useful to limit systematic uncertainties. In fact the result will be given as a ratio between rare and resonant quantities. However, the low statistics for the rare sample does not allow

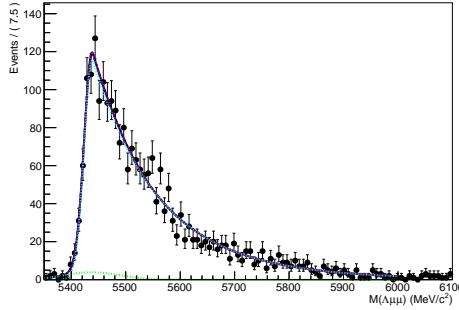


Figure 5.2: Invariant mass distribution of simulated misreconstructed $B^0 \rightarrow J/\psi K_s^0$ events after full selection fitted with Double Crystal Ball functions.

to constrain many parameters, especially when dividing data in q^2 bins. Therefore, all parameters of the signal shape are fixed to the ones derived from the fit on the normalisation channel. To account for possible differences, arising for example from a different resolution in different q^2 regions, a scale factor is multiplied to the width of the two gaussians cores at the signal DCB: $\sigma_1 \rightarrow c \cdot \sigma_1$ and $\sigma_2 \rightarrow c \cdot \sigma_2$, where the two scale factors are the same. This factors are fixed in the fit on data by fitting a $\Lambda \mu^+ \mu^-$ simulated sample in each q^2 bin and comparing its widths with the ones found on the fit to the resonant simulated sample, namely

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

Used values, reported in Tab. 5.4, are all around 2 corresponding to the fact that in the resonant case a further constrain on the dimuon mass is used. Instead the dependence on q^2 is found to be little. For fits on the DD and LL samples the parameters are always fixed to the corresponding J/ψ fit; in this analysis parameters are never shared between DD and LL fits.

The background components modelled in the resonant case are also in this case the combinatorial background, described with an exponential function. The slope of the background is visibly different depending on the q^2 interval. This is partly due to the fact that, at high q^2 , the combinatorial changes slope due to the kinematical limit at low masses. The exponential slopes are therefore left floating independently in each q^2 bins and also independently of the resonant channel and for the in DD and

Table 5.1: Predicted numbers of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in each considered q^2 interval.

q^2 bin	DD	LL
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

880 LL samples. The background component from $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays is modelled
 881 using the same shapes used for the resonant channel. However, in this case the
 882 horizontal shift is fixed to what found for the resonant channel. The expected
 883 amount of misreconstructed $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events is small and does not allow to
 884 determine reliably the yield. Therefore, in the default fit, this is fixed to the the yield
 885 of $B^0 \rightarrow J/\psi K_s^0$ decays, rescaling it by the expected ratio of branching fractions
 886 between the resonant and rare channels. The q^2 distribution of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$
 887 simulated events is then used to predict the yield as a function of q^2 . In Tab. 5.1
 888 is reported the number of predicted $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ events in each q^2 bin obtained
 889 with the following formula:

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

890 where $N(q^2)_{MC}$ is the number of simulated events in a q^2 bin after full selection
 891 and N_{MC}^{tot} is the total number of simulated events. The K_s^0 mm contribution is
 892 then completely taken out to study systematic uncertainties as described in Sec. 6
 893 Only for the 6-8 GeV $^2/c^4$ bin, a background component coming from the residual
 894 of the J/ψ radiative tail is added, modelled using a the shape obtained studying
 895 simulated events and smoothed using the `RooKeysPdf` method of `RooFit`. This is
 896 then removed from the final fit because it returns zero yield.

897 The fit on the rare sample is performed simultaneously on the LL and DD candidate

898 categories. Therefore the two separate yields are not separately floating but are but
899 are parameterised ad a function of the branching ratio with the following formula:

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

900 where $k = \text{LL,DD}$, Δq^2 is width of the q^2 bin and the only free paramater is the
901 branching fraction ratio rare over J/ψ . For the $J/\psi \rightarrow \mu^+\mu^-$ the value reported in
902 the PDG book [31], $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$. In this formula the
903 efficiencies and the normalisation yield appear as constants. These constants are
904 then varied in order to obtain systemaitcs on the final result as described in Sec. 6.

905 5.2 Fit results

906 The invariant mass distribution is shown in Fig. 5.3 for the normalisation channel
907 selected with high q^2 selection and in Fig. 5.4 for low q^2 selection. Fits performed
908 on MC samples are also reported in Fig. ??, fitted with the same fit function for the
909 signal, but with all parameters floating, and no function for the background.

910 The χ^2 value of the fit is 126 for LL and 112 for DD both with 140 degrees of freedom,
911 which corresponds to probability of 80% and 95%. In table 5.2 are reported raw
912 yields from the fits for the normalisation channel selected using the low and high q^2
913 selections, values for the other parameters are shown on Fig. 5.3 together with their
914 errors. Results of the fits on $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ sample, for the high q^2 region, are shown
915 in Fig. 5.5. For this with the χ^2 is 41 with 68 dof, corresponding to a probability
916 of 99.8%. In Fig. 5.6 the invariant mass is shown for the low q^2 region, where the
917 signal is still blinded. Corresponding plots in bins of q^2 are in Fig. 5.8 for DD events
918 and Fig. 5.8 for LL. The fitted yields are reported in table 5.3, notice that, since the
919 fit is simultanous on DD and LL events and the floating parameter is the branching

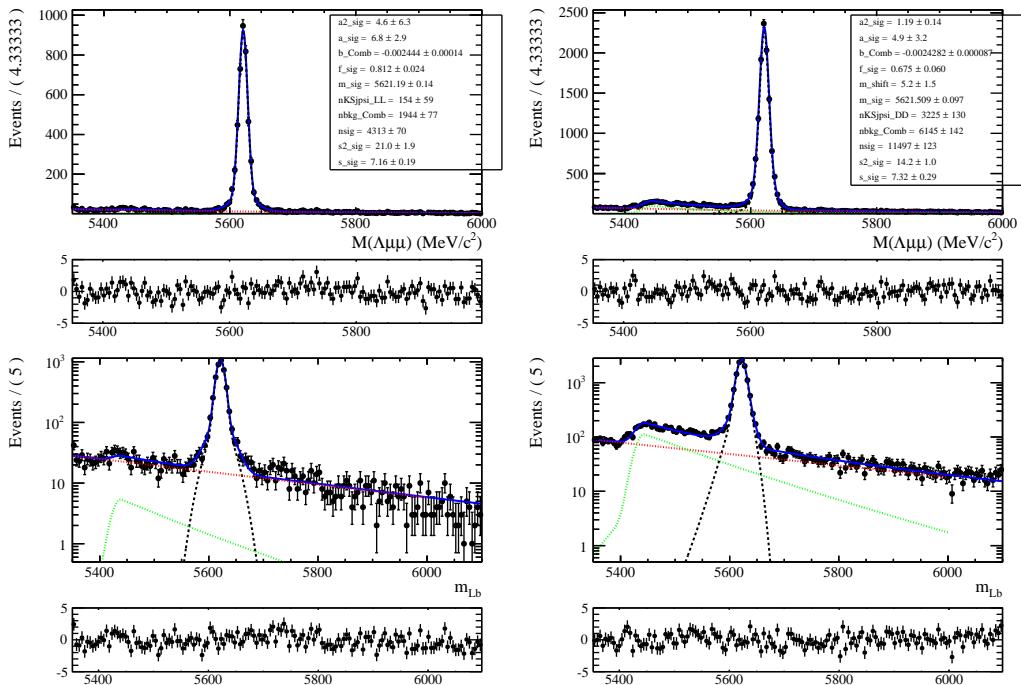


Figure 5.3: Invariant mass distribution of $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and residuals of the fit for long-long (left) and down-down (right) events. Lower plots are the same as the upper ones but shown in linear scale. The histogram shows data. On the plot are shown the total fit (blue line),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

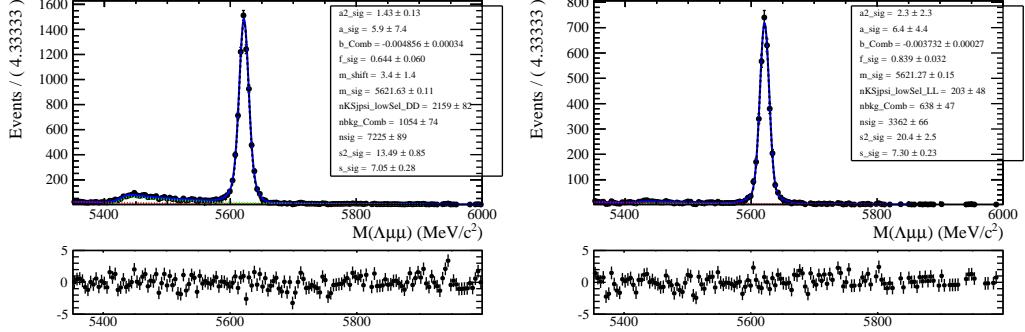


Figure 5.4: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda J/\psi$ with fit and residuals for DD events (left) and LL events (right). The histogram shows real data selected with low q^2 selection. The blue line is the signal fit function.

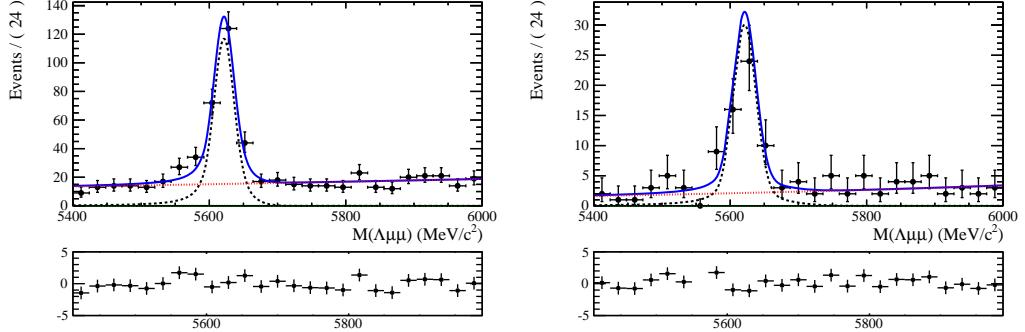


Figure 5.5: Invariant mass distribution of rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ with fit and residuals for DD events (left) and LL events (right) in the integrated $15-20 \text{ GeV}^2/c^2 q^2$ bin. The histogram shows data, the blue line the total fit function and the dashed red line represents combinatorial background.

ratio, these yields are not free to float independently. The exponential slopes found from the fits and the scale factors multiplied to the widths are reported in table 5.4.

Selection	N_{evts} (LL)	N_{evts} (DD)
high q^2	4312.7 ± 69.8	11497.2 ± 122.9
low q^2	3362.6 ± 58.9	7224.9 ± 89.3

Table 5.2: Raw yield in q^2 bins for DD and LL events. Uncertainties shown are statistical only.

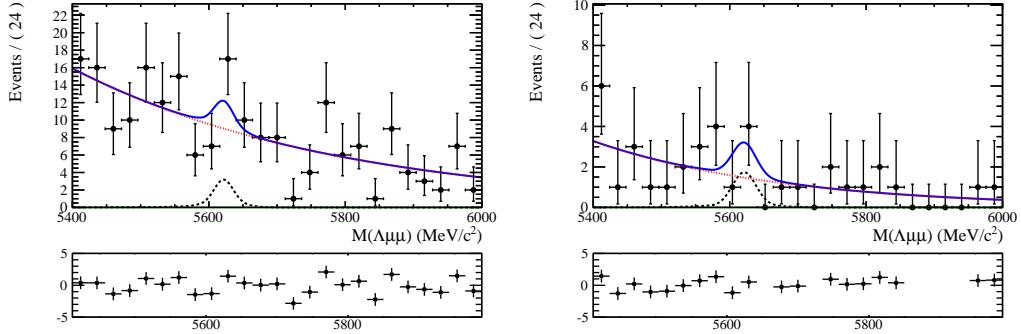


Figure 5.6: Invariant mass distribution of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in the integrated $0.1\text{-}6 \text{ GeV}^2/\text{c}^2 q^2$ bin, still blinded, for DD events (left) and LL events (right).

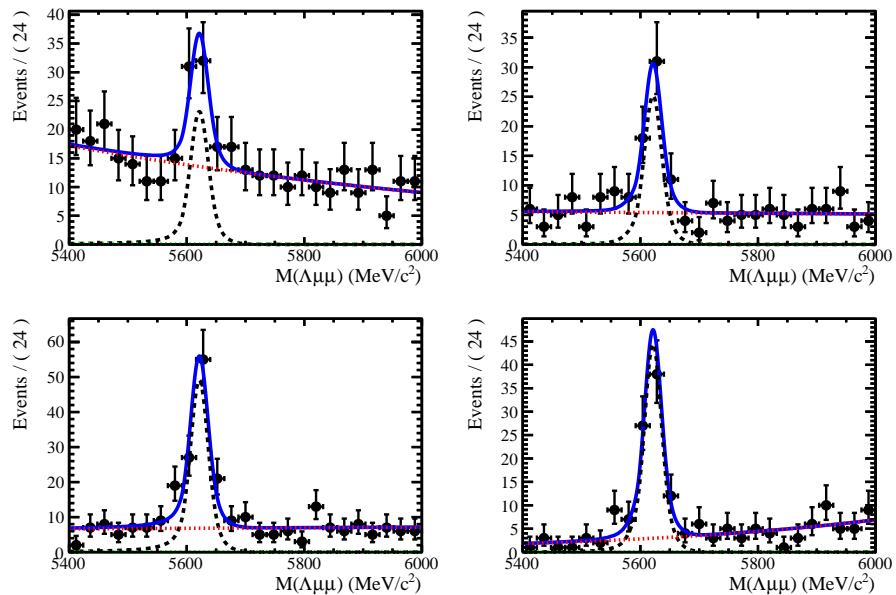


Figure 5.7: Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in q^2 bins $[0.1, 2]$, $[2, 4]$, $[4, 6]$, $[6, 8]$ for DD events.

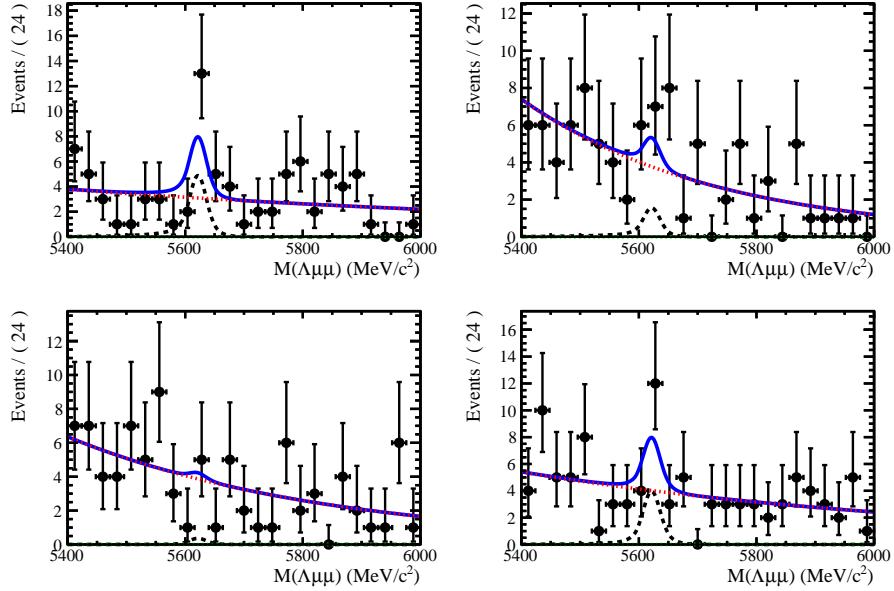


Figure 5.8: Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in q^2 bins [11,12.5], [15,16], [16,18], [18,20] for DD events.

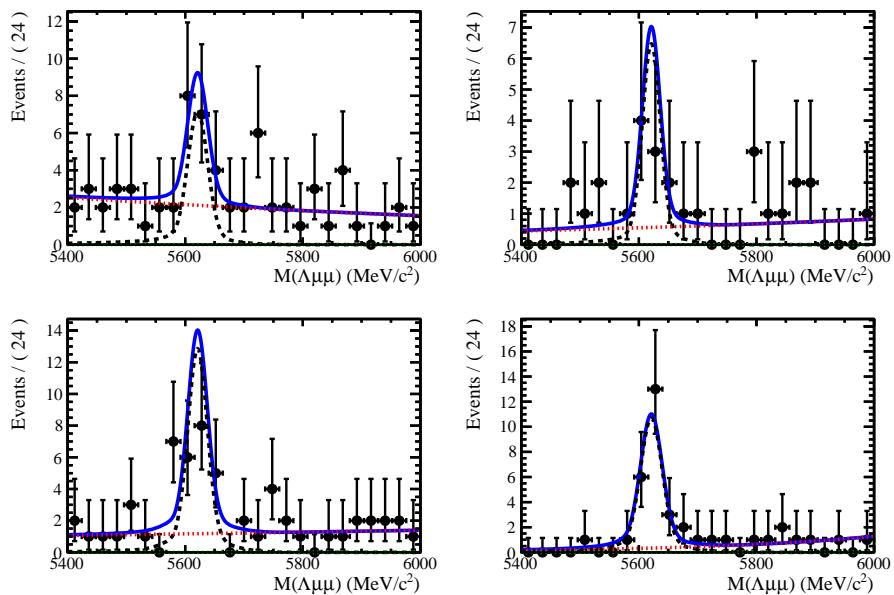


Figure 5.9: Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in q^2 bins [0.1,2], [2,4], [4,6], [6,8] for LL events.

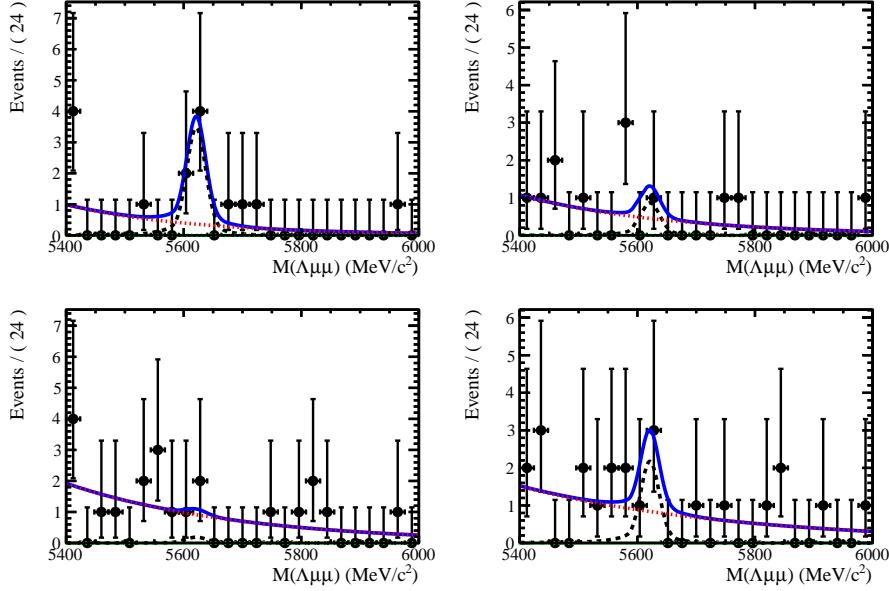


Figure 5.10: Invariant mass distributions with fit projection for rare $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidates in q^2 bins [11,12.5], [15,16], [16,18], [18,20] for LL events.

q^2 bin	N_{evts} (LL)	N_{evts} (DD)
0.1-2.0	6.9 ± 2.2	9.1 ± 3.0
2.0-4.0	1.8 ± 1.7	3.0 ± 2.8
4.0-6.0	0.4 ± 0.9	0.6 ± 1.4
6.0-8.0	4.3 ± 2.0	7.2 ± 3.3
11.0-12.5	14.6 ± 2.9	42.8 ± 8.5
15.0-16.0	13.5 ± 2.2	43.5 ± 7.2
16.0-18.0	28.6 ± 3.3	88.8 ± 10.1
18.0-20.0	22.4 ± 2.6	78.0 ± 8.9
1.1-6.0	3.6 ± 2.4	5.7 ± 3.8
15.0-20.0	64.6 ± 4.7	209.6 ± 15.3

Table 5.3: Raw yield in q^2 bins for DD and LL events. Uncertainties shown are statistical only.

Parameter	DD	LL
15.00-20.00		
b	0.0006 ± 0.0003	0.0012 ± 0.0008
c	1.9027 ± 0.0000	2.2910 ± 0.0000
N_{exp}	$393.2287^{22.7312}_{-21.8980}$	$64.7093^{9.2617}_{-8.5147}$
11.00-12.50		
b	-0.0011 ± 0.0003	-0.0008 ± 0.0008
c	1.9419 ± 0.0000	2.3497 ± 0.0000
N_{exp}	$317.8518^{19.7083}_{-18.8526}$	$50.0563^{8.0412}_{-7.2450}$
15.00-16.00		
b	-0.0001 ± 0.0005	0.0010 ± 0.0015
c	1.8794 ± 0.0000	2.0837 ± 0.0000
N_{exp}	$130.4247^{12.9083}_{-12.1713}$	$17.7568^{4.9209}_{-4.2204}$
16.00-18.00		
b	0.0002 ± 0.0005	0.0005 ± 0.0011
c	1.9104 ± 0.0000	2.2704 ± 0.0000
N_{exp}	$166.9400^{15.0846}_{-14.3628}$	$33.4597^{6.7251}_{-6.0169}$
18.00-20.00		
b	0.0022 ± 0.0007	$0.0035^{0.0022}_{-0.0019}$
c	1.8462 ± 0.0000	2.5404 ± 0.0000
N_{exp}	$96.1396^{11.3356}_{-10.7064}$	$12.7939^{4.4421}_{-3.6569}$
0.10-2.00		
b	-0.0009 ± 0.0007	$-0.0045^{0.0022}_{-0.0025}$
c	1.9421 ± 0.0000	2.1567 ± 0.0000
N_{exp}	$72.2655^{9.3030}_{-8.6567}$	$8.8273^{3.5813}_{-2.8547}$
2.00-4.00		
b	-0.0030 ± 0.0007	$-0.0039^{0.0019}_{-0.0021}$
c	1.9289 ± 0.0000	2.2536 ± 0.0000
N_{exp}	$84.3589^{10.0741}_{-9.4369}$	$10.8812^{3.8051}_{-3.1277}$
4.00-6.00		
b	-0.0022 ± 0.0006	-0.0034 ± 0.0014
c	1.8399 ± 0.0000	2.4371 ± 0.0000
N_{exp}	$87.4554^{9.8967}_{-9.2197}$	$20.5297^{5.1759}_{-4.4864}$
6.00-8.00		
b	-0.0013 ± 0.0006	-0.0027 ± 0.0014
c	1.9261 ± 0.0000	2.1144 ± 0.0000
N_{exp}	$92.0141^{10.4971}_{-9.7643}$	$19.5430^{5.0085}_{-4.3603}$
1.10-6.00		
b	-0.0026 ± 0.0004	-0.0036 ± 0.0011
c	1.9208 ± 0.0000	2.3504 ± 0.0000
N_{exp}	$203.1425^{14.9452}_{-14.6724}$	$33.7378^{6.5084}_{-5.7915}$

Table 5.4: Exponential slopes, b , found from the fits to rare date and resoution scale factors, c , fixed in the fits using MC for DD and LL events.

922 **5.3 Efficiency**

923 Efficiency for each of the decays is calculated according

$$\epsilon = \epsilon(\text{Geom})\epsilon(\text{Det}|\text{Geom})\epsilon(\text{Reco}|\text{Det})\epsilon(\text{MVA}|\text{Reco})\epsilon(\text{Trig}|\text{MVA}). \quad (5.8)$$

924 In this expression the first term gives efficiency to have final state particles in the
925 LHCb acceptance. The second term handles the possibility of Λ escaping the de-
926 tector or interacting with it and therefore never decaying in $p\pi$. We call this term
927 "Detection" efficiency. The third term carries information on reconstruction and
928 stripping efficiency (we keep these together given that boundaries between them
929 are completely artificial). Fourth part deals with the efficiency of the Neural Net-
930 work selection. And finally the last term handles trigger efficiency for those events
931 which are selected by the preselection process. All of those are evaluated on simu-
932 lated data with trigger efficiency for $\Lambda_b^0 \rightarrow J/\psi \Lambda$ being cross-checked using TISTOS
933 method on data. While we list all efficiencies for two decays separately including
934 many systematic uncertainties, for the analysis itself we need only relative efficiency
935 $\epsilon(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)/\epsilon(\Lambda_b^0 \rightarrow J/\psi \Lambda)$. Systematic uncertainties for relative efficiencies
936 take into account that some effects are correlated between two decays. All efficien-
937 cies are calculated using simulated samples generated with Pythia8.

938 **5.3.1 Decay Model**

939 To include reasonably realistic q^2 and angular dependence into simulation, we start
940 with fully simulated events generated according to phase space decay and weigh
941 those with amplitude based on prediction in Ref. [38]. Equations in this paper
942 are for case of unpolarised production and we extend those to include polarisation.
943 Details of this are in Appendix ???. In the calculation we use production polarisation
944 of $P_b = 0.06$ as measured by LHCb [39]. In the weight calculation we always
945 use generator level momenta to obtain q^2 and corresponding angles. Resulting q^2
946 distribution is shown in Fig. 5.11. For the normalisation mode, we use decay model

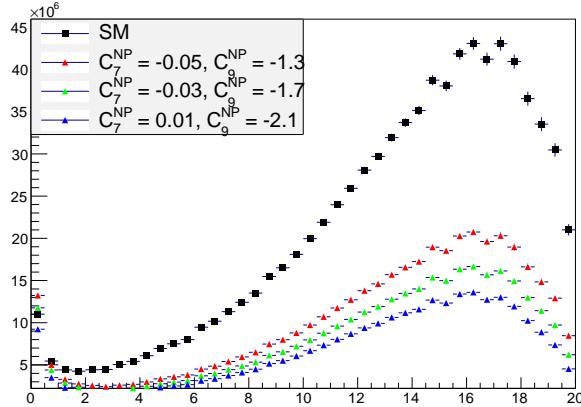


Figure 5.11: The q^2 spectrum of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ events weighted with models embedding different sets of Wilson Coefficients. The black distribution corresponds to the weighting used to calculate efficiencies.

described in Appendix ??, with amplitude magnitudes and production polarisation taken from measurement in Ref. [39]. Phases are not yet measured and we set all of them to zero.

5.3.1.1 Effect of new physics on the decay model

In order to check how new physics would affect the decay model we modified the Wilson Coefficients in the model adding new physics contribution to the C_7 and C_9 coefficients. We modified these coefficients by adding a NP component ($C_i \rightarrow C_i + C_i^{NP}$). The values used are reported on top of each plot. In Fig. 5.11 is reported the q^2 spectrum weighted for a model embedding the default and 3 modified sets of wilson coefficients. As can be seen the biggest effect is in the very low q^2 . In Fig. 5.12 are reported the relative effects of these different models on the total efficiency.

5.3.2 Kinematic re-weighting

In order to take into account data-MC differences, especially in the initial kinematic of the mother particle, Λ_b^0 , we re-weight our Monte Carlo. We take as variables the transverse momentum of Λ_b^0 and Λ and we compare data and MC performing a 2D

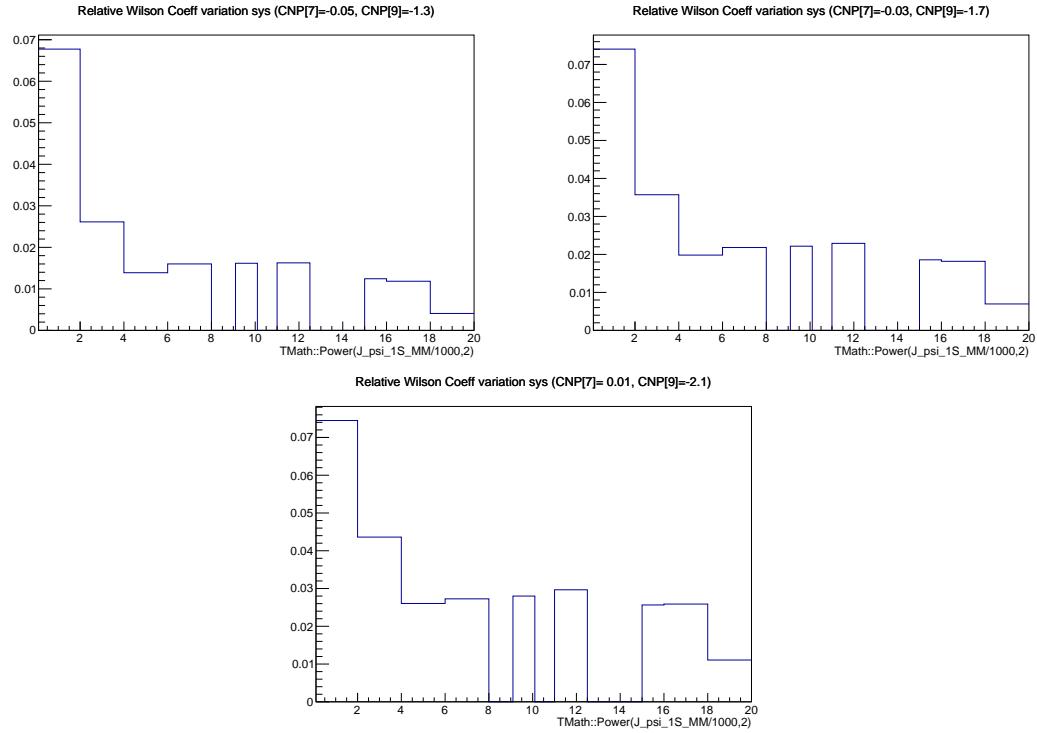


Figure 5.12: Relative effect of different Wilson Coefficients sets on the total efficiency in bins of q^2 .

962 re-weighting.

963 We extract real data distributions from real $J/\psi \Lambda$ events out of stripping. To work
 964 with a sample as clean as possible, we select a narrow interval around J/ψ and Λ_b^0
 965 peaks. Then we fit the Λ_b^0 invariant mass to extract the amount of background under
 966 the peak. Finally, we use the background fraction, f_b , extracted to statistically
 967 subtract the background from the kinematical distributions as described by the
 968 following equation

$$S(p_T) = T(p_T) - f_b \cdot B(p_t), \quad (5.9)$$

969 where $T(p_T)$ is the distribution of all events in the signal interval ($5605 < m(p\pi\mu^+\mu^-) <$
 970 $5635 \text{ MeV}/c^2$) and $B(p_T)$ is the distribution of background events extracted using
 971 events from the upper sideband ($m(p\pi\mu^+\mu^-) > 5800 \text{ MeV}/c^2$). Finally, $S(p_T)$ is the
 972 distribution of signal events.

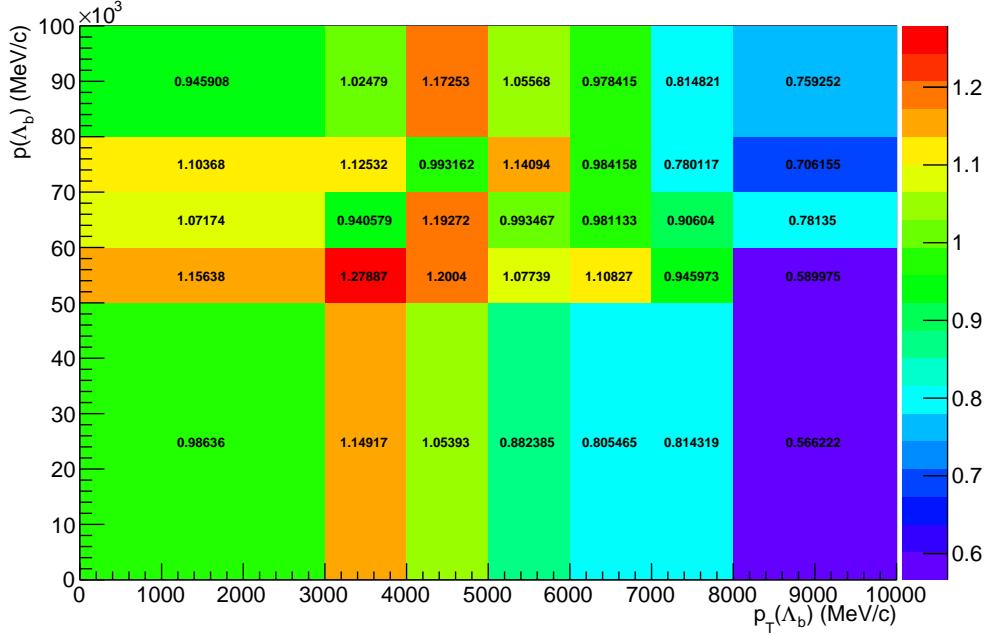


Figure 5.13: Histogram of weights used for the kinematical reweighting as a function of the momentum and transverse momentum of Λ_b^0 .

After obtaining the signal distributions from real data, we compare them with $\Lambda_b^0 \mu^+ \mu^-$ MC events. We take as a weight for the MC the ratio between data and MC distributions in 2 dimensions. The histogram used for the re-weighting is shown in Fig.5.13. In appendix ?? are shown distributions of sideband subtracted data in the signal and sideband regions and weighted and unweighted MC events for many variables. In these plots the Λ_b^0 distributions match by construction but the reweighting brings also the transverse momenta of all final particles to match between data and MC. Other variables, as the χ^2 of tracks, which have little dependence on the kinematics are unaffected by the weighting procedure. N.B.: Differences in the tails of the Λ_b^0 plots are due to the fact that the weight is binned. In particular in the tails we use wide bins because we need to have enough statistics in each bin to perform a reliable background subtraction on data.

985 5.3.3 Event type

986 In LHCb detector, Λ can be reconstructed using long tracks (LL), having also hits
987 in the VELO detector, or downstream tracks (DD). Distinction is mainly driven
988 by geometry of the detector. The fraction of Λ reconstructed from long tracks and
989 downstream tracks does not fully agree with simulation. For $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays
990 which pass all out selection, we determine on data that $(69.0 \pm 0.9)\%$ of Λ candidates
991 are reconstructed from downstream tracks. On contrary in simulation of the same
992 decay, $74.9 \pm 0.9\%$ of candidates are reconstructed from down tracks. Furthermore
993 the fraction of down and long tracks varies as a function of q^2 . The biggest differences
994 are found at low q^2 .

995 In order to deal with this difference we extract efficiencies separately for downstream
996 and long events and we do the analysis separately for the two categories, joining
997 results at the end.

998 5.3.4 Geometric acceptance

999 The simulated samples we use contain requirement that daughters are in LHCb
1000 acceptance. This corresponds to requirement for each of the muon to have polar
1001 angle θ between 10 and 400 mrad. In table 5.5 we list the efficiencies corresponding
1002 to this requirement in bins of q^2 .

1003 5.3.5 Reconstruction and stripping efficiency

1004 The efficiency to reconstruct our decays together with stripping selection is evaluated
1005 from simulated data. This component does not include the efficiency of the PID cut,
1006 because PID is known to be not well simulated and therefore we use a data driven
1007 method for this part (see 5.3.8). For the evaluation we use the most recent LHCb
1008 measurement of Λ_b^0 lifetime of 1.482 ± 0.03 ps[40] and Λ_b^0 polarisation 0.06 ± 0.09

q^2 bin	Geom. acc.
0.1-2.0	0.23591 ± 0.00084
2.0-4.0	0.20983 ± 0.00072
4.0-6.0	0.20076 ± 0.00074
6.0-8.0	0.19596 ± 0.00077
9.1-10.1	0.19265 ± 0.00115
11.0-12.5	0.18965 ± 0.00102
15.0-16.0	0.18955 ± 0.00151
16.0-18.0	0.18719 ± 0.00120
18.0-20.0	0.18704 ± 0.00156
1.1-6.0	0.20720 ± 0.00047
15.0-20.0	0.18763 ± 0.00082

Table 5.5: Absolute geometrical acceptance in bins of q^2 from MC simulation. Errors shown are statistical only.

[39]. Both $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ are reweighted with same physics rate as in previous subsection. In table 5.6 the reconstruction efficiency is reported in bins of q^2 and for low and high q^2 regions integrated. In the table the efficiency is subdivided in "Detection" and "Reconstruction and Stripping" efficiencies. In fact since Λ is a long lived particles there is a non-negligible probability that it interacts in the detector or escapes from it and therefore never decays in proton and pion. The efficiency for this to happen is what we call "Detection" efficiency. "Reconstruction and Stripping" efficiency include the efficiency of reconstructing tracks and the efficiency for events passing the stripping cuts. The errors shown are statistical.

5.3.6 Neural Networks efficiency

Finally, the MVA efficiency is again evaluated from weighted MC samples. Results are shown in table 5.7 in bins of q^2 . The error shown is due to to MC statistics. The sudden jump in efficiency before and after ~ 9 GeV/ c^2 is due to the fact that we use a different optimisation for the MVA cut in the low and high q^2 region.

q^2 bin	Detection	Reco and Strip (DD)	Reco and Strip (LL)
0.1-2.0	0.87928 ± 0.00046	0.05192 ± 0.00059	0.01939 ± 0.00039
2.0-4.0	0.88496 ± 0.00042	0.06635 ± 0.00062	0.01954 ± 0.00036
4.0-6.0	0.89020 ± 0.00044	0.07174 ± 0.00066	0.02092 ± 0.00038
6.0-8.0	0.89623 ± 0.00046	0.07562 ± 0.00069	0.02122 ± 0.00039
9.1-10.1	0.90218 ± 0.00069	0.07867 ± 0.00105	0.02198 ± 0.00058
11.0-12.5	0.90840 ± 0.00061	0.07989 ± 0.00091	0.02209 ± 0.00052
15.0-16.0	0.91868 ± 0.00089	0.07357 ± 0.00124	0.01788 ± 0.00065
16.0-18.0	0.92468 ± 0.00069	0.06959 ± 0.00095	0.01688 ± 0.00049
18.0-20.0	0.93177 ± 0.00085	0.06000 ± 0.00110	0.01356 ± 0.00055
1.1-6.0	0.88683 ± 0.00027	0.06843 ± 0.00041	0.02021 ± 0.00024
15.0-20.0	0.92601 ± 0.00046	0.06694 ± 0.00063	0.01589 ± 0.00032

Table 5.6: Absolute Detection and reconstruction plus stripping efficiencies. Reconstruction efficiency is given separately for DD and LL events. Errors shown are statistical only.

q^2 bin	MVA eff (DD)	MVA eff (LL)
0.1-2.0	0.62267 ± 0.00835	0.81337 ± 0.01096
2.0-4.0	0.58293 ± 0.00706	0.75671 ± 0.01144
4.0-6.0	0.58390 ± 0.00686	0.77630 ± 0.01085
6.0-8.0	0.58674 ± 0.00671	0.77757 ± 0.01053
9.1-10.1	0.90381 ± 0.00561	0.94811 ± 0.00781
11.0-12.5	0.88847 ± 0.00517	0.94381 ± 0.00732
15.0-16.0	0.88197 ± 0.00732	0.92942 ± 0.01233
16.0-18.0	0.84710 ± 0.00658	0.92803 ± 0.00946
18.0-20.0	0.83144 ± 0.00905	0.88888 ± 0.01621
1.1-6.0	0.58433 ± 0.00454	0.77248 ± 0.00721
15.0-20.0	0.84923 ± 0.00453	0.91650 ± 0.00722

Table 5.7: Neural network selection efficiency. Errors shown are statistical only.

q^2 bin	Trigger eff (DD)	Trigger eff (LL)
0.1-2.0	0.56038 ± 0.00752	0.57682 ± 0.01210
2.0-4.0	0.60563 ± 0.00644	0.65118 ± 0.01048
4.0-6.0	0.62319 ± 0.00625	0.67381 ± 0.01026
6.0-8.0	0.66879 ± 0.00614	0.70643 ± 0.01009
9.1-10.1	0.69987 ± 0.00701	0.72242 ± 0.01267
11.0-12.5	0.74441 ± 0.00574	0.73769 ± 0.01116
15.0-16.0	0.81816 ± 0.00752	0.82586 ± 0.01498
16.0-18.0	0.83580 ± 0.00590	0.85956 ± 0.01071
18.0-20.0	0.85743 ± 0.00750	0.86295 ± 0.01468
1.1-6.0	0.61038 ± 0.00412	0.65307 ± 0.00672
15.0-20.0	0.83867 ± 0.00398	0.85282 ± 0.00752

Table 5.8: Absolute trigger efficiencies for selected events as determined from the simulation separately for LL and DD events. Efficiency is derived as ratio of events selected and accepted by trigger and events selected. Errors shown are statistical only.

1024 5.3.7 Trigger efficiency

1025 Trigger efficiency is again calculated on simulated sample following same logic as
 1026 previous two efficiencies. The pre-scale emulator tool in DaVinci is also used. We
 1027 calculate this efficiency for events which are accepted by the full selection. Using the
 1028 resonant channel we crosschecked the efficiency obtained using the MC simulation
 1029 with the data driven TISTOS method. In LHCb triggered events can fall in two
 1030 categories: events triggered by a track which is part of the decay of interest, Trigger
 1031 On Signal (TOS), or by other tracks in the events, Trigger Independent of Signal
 1032 (TIS). All trigger lines used for this analysis are required to be TOS. The efficiency
 1033 for TOS trigger can be obtained by data by the following formula:

$$\epsilon_{TOS} = \frac{TOS \text{ and } TIS}{TIS} \quad (5.10)$$

1034 Using the data driven method we obtain $(70 \pm 5)\%$ efficiency, while from MC we
 1035 obtain $(73.33 \pm 0.02)\%$. Results are therefore compatible within 1σ . Efficiencies are
 1036 calculated using three samples where different TCKs were simulated. For 2011 we
 1037 use the official 0x40760037 TCK, corresponding to 0.944 fb^{-1} of data. For 2012
 1038 the official TCK 0x409f0045 describes only data after the June technical stop (1.442

q^2 bin	Eff.
0.1-2.0	97.3203 ± 0.01174
2.0-4.0	97.4231 ± 0.01168
4.0-6.0	97.5958 ± 0.01096
6.0-8.0	97.6993 ± 0.00982
11.0-12.5	98.0382 ± 0.00876
15.0-16.0	98.3119 ± 0.00562
16.0-18.0	98.1004 ± 0.00548
18.0-20.0	98.1121 ± 0.00071
1.1-6.0	97.4885 ± 0.00712
15.0-20.0	98.1656 ± 0.00306
J/ψ	97.8904 ± 0.00498

Table 5.9: Absolute PID efficiencies in q^2 bins

1039 fb^{-1}), when the topological HLT2 trigger was improved to have better efficiency on
1040 events including down tracks. Therefore this affected especially K_S and Λ decays.
1041 In order to describe pre-June data (0.517 fb^{-1}) we rerun trigger on the simulated
1042 sample for 2012 with 0x4097003d TCK. To evaluate trigger efficiency, samples with
1043 different TCKs are weighted to respect the percentage of luminosity taken with each
1044 configuration.

1045 5.3.8 PID efficiency

1046 For long long tracks we use a PID cut on protons ($PIDp > -5$). The MC is known
1047 not to describe particle ID well and therefore we use a data driven method to
1048 obtain this efficiency component. In order to do this we used the PIDCalib package.
1049 This package uses decays where particles can be identified due to their kinematic
1050 properties. In our case a sample of Λ particles was used where the proton can be
1051 identified because it always has the highest momentum. Then the package allows
1052 to divide the phase-space in bins (momentum and pseudorapidity) and reweight the
1053 efficiencies found to take in account that the decay channel under study could have
1054 different kinematical distributions than the calibration sample. Results are shown
1055 in table 5.9 and the relative efficiency over J/ψ is also shown in 5.14.

Efficiency	Value (DD)	Value (LL)
geom	0.18181 ± 0.00026	
det	0.90172 ± 0.00025	
reco	0.07240 ± 0.00036	0.02027 ± 0.00020
mva	0.88196 ± 0.00172	0.94220 ± 0.00240
trig	0.69680 ± 0.00259	0.73428 ± 0.00465
Full Selection	0.04449 ± 0.00029	0.01402 ± 0.00017
Total	0.00729 ± 0.00005	0.00230 ± 0.00003

Table 5.10: Absolute efficiency values for $\Lambda_b^0 \rightarrow J/\psi \Lambda$. Errors shown are statistical only.

q^2 bin	Geometric	Detection
0.1-2.0	1.29755 ± 0.00501	0.97512 ± 0.00058
2.0-4.0	1.15410 ± 0.00433	0.98141 ± 0.00054
4.0-6.0	1.10425 ± 0.00438	0.98722 ± 0.00056
6.0-8.0	1.07780 ± 0.00451	0.99391 ± 0.00058
9.1-10.1	1.05964 ± 0.00653	1.00051 ± 0.00082
11.0-12.5	1.04312 ± 0.00583	1.00741 ± 0.00073
15.0-16.0	1.04256 ± 0.00842	1.01881 ± 0.00103
16.0-18.0	1.02959 ± 0.00679	1.02547 ± 0.00081
18.0-20.0	1.02878 ± 0.00873	1.03332 ± 0.00098
1.1-6.0	1.13964 ± 0.00309	0.98349 ± 0.00041
15.0-20.0	1.03201 ± 0.00478	1.02694 ± 0.00059

Table 5.11: Relative geometric efficiency and Detection efficiency between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays. Uncertainty reflects statistics of both samples. Errors shown are statistical only.

5.3.9 Relative efficiency

In the previous tables efficiency values were given for the rare channel in different q^2 bins. In this section we list the corresponding relative efficiencies which will be used for the calculation of the final result. In order to link relative and absolute efficiencies we also report the absolute values of efficiency for the J/ψ channel in table 5.10. In table 5.11 we list relative geometric efficiency between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays. They are calculated using results on individual absolute efficiencies. Uncertainties reflect statistics of both samples. Systematic uncertainties are discussed in next section.

In tables 5.13 and 5.12 we show relative reconstruction, trigger and mva selection

q^2 bin	Reco and strip	MVA	Trigger	Full Selection
0.1-2.0	0.95977 ± 0.02168	0.86346 ± 0.01184	0.78591 ± 0.01722	0.65130 ± 0.02236
2.0-4.0	0.96880 ± 0.02006	0.80331 ± 0.01231	0.88722 ± 0.01535	0.69048 ± 0.02143
4.0-6.0	1.03869 ± 0.02145	0.82411 ± 0.01171	0.91805 ± 0.01515	0.78585 ± 0.02358
6.0-8.0	1.05419 ± 0.02194	0.82546 ± 0.01138	0.96251 ± 0.01505	0.83756 ± 0.02467
9.1-10.1	1.08433 ± 0.03068	1.00650 ± 0.00868	0.98428 ± 0.01836	1.07422 ± 0.03757
11.0-12.5	1.09602 ± 0.02798	1.00194 ± 0.00818	1.00509 ± 0.01649	1.10373 ± 0.03468
15.0-16.0	0.88647 ± 0.03336	0.98666 ± 0.01333	1.12522 ± 0.02162	0.98416 ± 0.04366
16.0-18.0	0.83711 ± 0.02567	0.98519 ± 0.01036	1.17113 ± 0.01638	0.96584 ± 0.03410
18.0-20.0	0.67408 ± 0.02788	0.94362 ± 0.01738	1.17576 ± 0.02135	0.74788 ± 0.03648
1.1-6.0	1.00225 ± 0.01543	0.82006 ± 0.00794	0.88979 ± 0.01076	0.73132 ± 0.01597
15.0-20.0	0.78841 ± 0.01767	0.97294 ± 0.00806	1.16195 ± 0.01263	0.89131 ± 0.02340

Table 5.12: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for long events. Uncertainty reflects statistics of both samples. Errors shown are statistical only.

efficiencies separately for DD and LL events. Since these three components are taken from the same MC sample their statistical errors are correlated. Therefore we also extract directly from MC the total of the three as a single efficiencies, which is labeled "Full Selection" efficiency in the table. Finally the PID efficiency is reported in table 5.14. As before, uncertainties reflect statistics of both samples. Results are obtained weighting both the resonant and rare MC for decay model and Λ_b^0 lifetime. In table 7.1 is reported the total of all relative efficiencies, which will be then used to correct the final result.

q^2 bin	Reco and strip	MVA	Trigger	Full Selection
0.1-2.0	0.72121 ± 0.00886	0.70604 ± 0.00956	0.80455 ± 0.01121	0.40968 ± 0.00942
2.0-4.0	0.91975 ± 0.00967	0.66098 ± 0.00811	0.86951 ± 0.00979	0.52861 ± 0.01041
4.0-6.0	0.99671 ± 0.01030	0.66208 ± 0.00788	0.89473 ± 0.00957	0.59043 ± 0.01125
6.0-8.0	1.04983 ± 0.01090	0.66530 ± 0.00772	0.96020 ± 0.00952	0.67066 ± 0.01238
9.1-10.1	1.09206 ± 0.01549	1.02482 ± 0.00667	1.00481 ± 0.01074	1.12455 ± 0.02127
11.0-12.5	1.11161 ± 0.01376	1.00743 ± 0.00619	1.06876 ± 0.00915	1.19688 ± 0.01946
15.0-16.0	1.01897 ± 0.01789	1.00006 ± 0.00853	1.17465 ± 0.01166	1.19700 ± 0.02621
16.0-18.0	0.96766 ± 0.01396	0.96052 ± 0.00770	1.19997 ± 0.00957	1.11532 ± 0.02044
18.0-20.0	0.83207 ± 0.01578	0.94276 ± 0.01042	1.23102 ± 0.01171	0.96567 ± 0.02310
1.1-6.0	0.94965 ± 0.00736	0.66257 ± 0.00531	0.87634 ± 0.00675	0.55140 ± 0.00747
15.0-20.0	0.92919 ± 0.00981	0.96294 ± 0.00547	1.20410 ± 0.00726	1.07738 ± 0.01446

Table 5.13: Relative efficiencies between $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow J/\psi\Lambda$ decays for downstream events. Uncertainty reflects statistics of both samples. Errors shown are statistical only.

q^2 bin	Rel. Eff.
0.1-2.0	0.99418 ± 0.00013
2.0-4.0	0.99523 ± 0.00013
4.0-6.0	0.99699 ± 0.00012
6.0-8.0	0.99805 ± 0.00011
11.0-12.5	1.00151 ± 0.00010
15.0-16.0	1.00431 ± 0.00008
16.0-18.0	1.00215 ± 0.00008
518.0-20.0	1.00226 ± 0.00005
1.1-6.0	0.99589 ± 0.00009
15.0-20.0	1.00281 ± 0.00006

Table 5.14: Relative PID efficiencies in q^2 bins

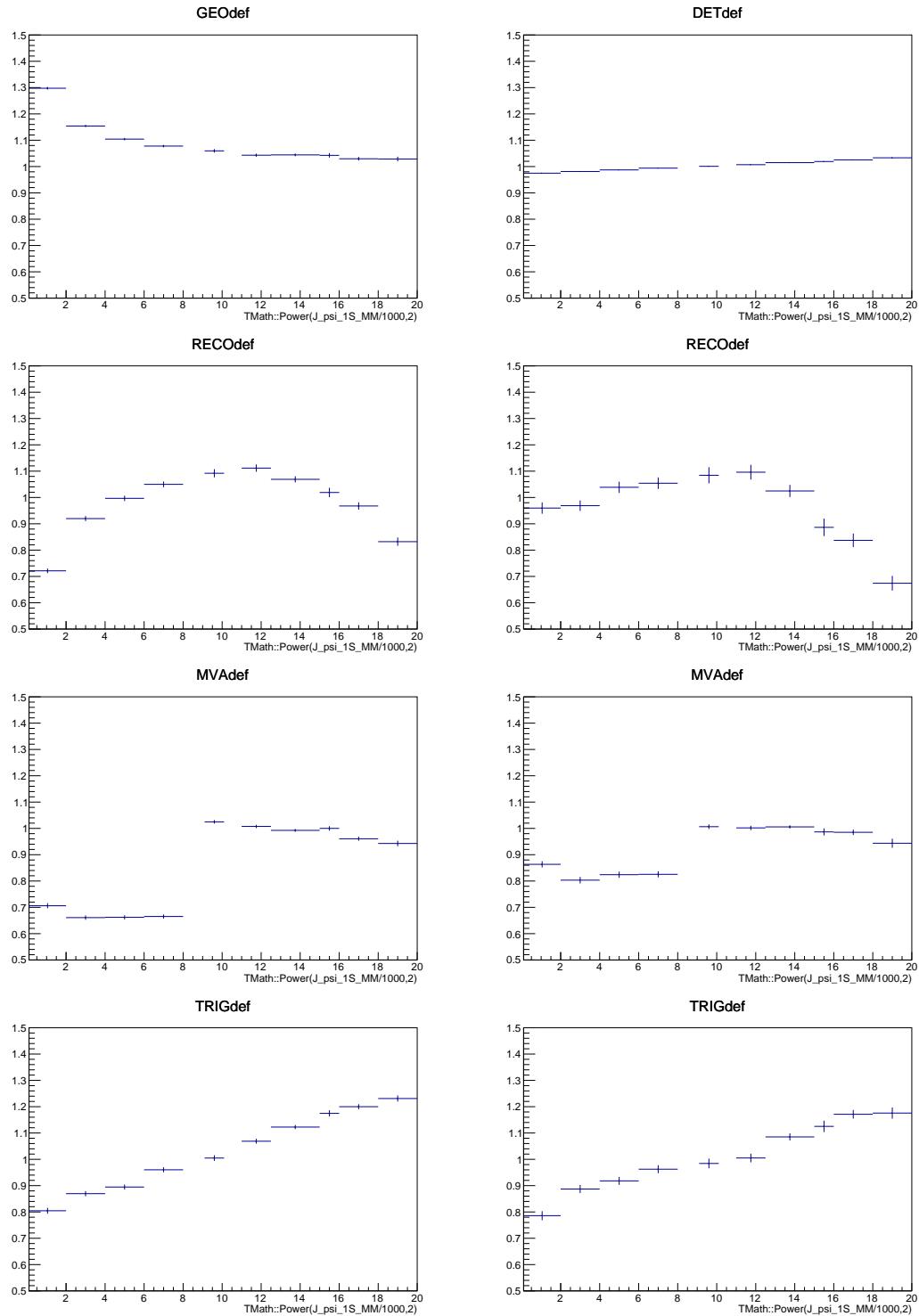


Figure 5.14: Relative efficiencies as a function of q^2 . geometric efficiency (a), detection efficiency (b), reconstruction efficiency for DD (c) and LL (d) events, trigger efficiency for DD (e) and LL (f) and MVA efficiency for DD (g) and LL (h).

1074

CHAPTER 6

1075

1076

Systematic uncertainties

1077

1078 **6.1 Yields**

1079 In order to evaluate the systematic error on the yield due to the choice of a PDF
1080 we generate toy MC experiments and we re-fit each toy data with different models.
1081 The toys are generated using a PDF, which is the default model fitted on data in
1082 each q^2 bin. In this way for each q^2 bins a specific shape is reproduced including the
1083 background level and slope. Furthermore for each q^2 bin we generate a number of
1084 events comparable to the one found in data. Finally we use the average bias (mean
1085 of the deviations) over 1000 toy experiments as systematic. In each case we fit rare
1086 and normalisation channels with the same signal model and, as for the default case,
1087 for the fits on the normalisation channel all parameters are left free, while we fix
1088 the rare parameters to what we find in the resonant case. All fits are performed
1089 minimising unbinned extended maximum likelihood.

1090 For the signal PDF systematic we refit each toy with the default model (Double
1091 Crystal Ball) and with a Double Gaussian, which is the only other signal PDF
1092 giving acceptable χ^2 . For the background PDF systematic we re-fit leaving the K_S
1093 component yield floating in the rare channel fit. This is fixed in the defalut rare fit
1094 and always floating in the resonant fit.

1095 Finally, we plot the deviations of the relative yield from the default fit $\Delta N/N =$
1096 $(N_{\text{default}}^{\mu\mu}/N_{\text{default}}^{J/\psi} - N_{\text{bkg/DGauss}}^{\mu\mu}/N_{\text{bkg/DGauss}}^{J/\psi})/(N_{\text{default}}^{\mu\mu}/N_{\text{default}}^{J/\psi})$. The distributions
1097 have approximately gaussian shape and the means are more than 2σ away from
1098 zero, therefore we consider these significant biases and we assign the means' values
1099 as systematic, values are reported in table 6.1. The most affected bin is the one in
1100 the middle of the charmonium resonances, where a combination of lower statistics
1101 and higher background leaves more freedom to the signal shape.

q^2 bin	Sig. PDF bias (%)	Bkg. PDF bias (%)	Tot. sys. (%)
Rare channel			
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 6.1: Values of systematics due to the choice of signal and background shapes in bins of q^2 .

1102

1103 6.2 Efficiencies

1104 In this subsection we discuss systematic uncertainties on absolute and relative ef-
1105 ficiencies. As we study q^2 dependence of the differential branching fraction, one

of the challenges is to properly account for correlations. There are three types of correlations we need to handle, one is correlation between normalisation and signal decay, second is correlation between different q^2 bins and finally as we split efficiency into three parts (geometry, selection and trigger), we should take also correlation amongst them into account. The last class of correlations is in principle rather easy to treat, but fully correct treatment of the first two classes would require to split the systematic uncertainties to correlated amongst q^2 bins and uncorrelated part, which given our statistics would complicate the presentation of the result. For simplicity we treat correlations where systematic uncertainty is significant and neglect correlations amongst q^2 bins when uncertainty for given source is small compared to dominant sources.

In the following sections we discuss all sources of systematic uncertainties we consider. Each subsection deals with a single source and both independent pieces of efficiency as well as total relative efficiency. It should be noted that not all sources contribute to each part of the efficiency. Numerical values of systematic uncertainties on relative efficiencies are in table 6.2 and for 6.3 for geometrical and detection efficiency; 6.4, 6.5 and 6.6 for reconstruction, trigger and neural networks for long-long and down-down events. Finally, we remind the reader that we always use fully simulated phase space sample we weight those, modelling polarisation and decay structure. The MC sample used is always the same thus within the single decay there is always correlation in our simulated events.

6.2.1 Simulation statistics

Simulated samples used to determine efficiencies have limited statistics. While it is not the dominant source of systematics, its size does not allow to completely neglect this uncertainty and thus it is propagated to the final systematic uncertainties. When reporting relative values we include both statistical uncertainty of signal as well as normalisation channel. While it would be useful to treat part from normalisation channel separately due to the correlation across q^2 bins, given its size we

q^2 bin	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 6.2: Absolute values of systematic uncertainties on relative geometric efficiency.

q^2 bin	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 6.3: Absolute values of systematic uncertainties on relative detection efficiency.

q^2 bin	DD			LL		
	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 6.4: Absolute values of systematic uncertainties on relative reconstruction efficiency for long-long and down-down events.

q^2 bin	DD			LL		
	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 6.5: Absolute values of systematic uncertainties on relative trigger efficiency for long-long and down-down events.

q^2 bin	DD			LL		
	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 6.6: Absolute values of systematic uncertainties on relative MVA efficiency for long-long and down-down events.

q^2 bin	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 6.7: Values of DD vertexing systematic uncertainties on relative reconstruction, trigger and MVA efficiencies for down-down events.

1134 decide to suppress it as in final presentation its effect is hard to see.

1135 6.2.2 Production polarisation and decay structure

1136 One of the main unknown which affects detection efficiencies is the angular struc-
1137 ture of the decays we study. And connected to this also the production polarisation,
1138 which is a parameter of the model. The normalisation decay $\Lambda_b^0 \rightarrow J/\psi \Lambda$ is also
1139 weighted and $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ is distributed according to the model described in ap-
1140 pendix ??.

1141 To assess the systematic uncertainty from production polarisation for signal $\Lambda_b^0 \rightarrow$
1142 $\Lambda \mu^+ \mu^-$ decay, we compare the efficiency obtained weighting the MC varying the
1143 polarisation in a range $\pm\sigma$ from the most recent LHCb measurement $P = 0.06 \pm$
1144 0.09[39]. We take as systematic the larger difference from two values. To assess
1145 systematic uncertainty due to decay structure we use alternative set of form factors
1146 based on lattice QCD calculation [41]. Details of this are explained in ???. We
1147 compare the two models and take the full difference as systematic uncertainty.

1148 6.2.3 Λ_b^0 lifetime

1149 The Λ_b^0 lifetime is known only with limited precision. For evaluation of the ef-
1150 ficiencies we use world average value of 1.482 ps^{-1} [42]. To evaluate systematic
1151 uncertainty, we vary this by $\pm 1\sigma$ and take larger difference as systematic uncer-
1152 tainty. We consider only cases where both signal and normalisation channel are
1153 varied in same direction. To evaluate uncertainty on total relative efficiency, we
1154 simple evaluate total relative efficiency with new Λ_b^0 lifetime and take difference to
1155 default as systematic uncertainty. This way, all correlations are taken into account.
1156 We do not attempt to separate out part correlated amongst q^2 bins. This source
1157 does not affect geometric acceptance, which is defined purely by requirements on
1158 angles of each muon with respect to beam axis.

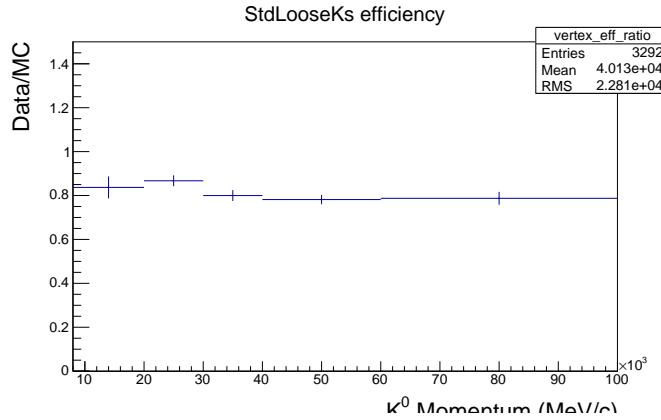


Figure 6.1: Ratio of reconstruction efficiency in Data and MC found using K_S events[4].

1159 6.2.4 Λ detection efficiency

- 1160 In the LHCb detector, Λ can be reconstructed using long or downstream tracks.
1161 The distinction is mainly driven by the geometry of the detector. A potential issue
1162 is that fraction of Λ reconstructed from long tracks and downstream tracks does
1163 not fully agree with simulation. For $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay we determine on data that
1164 $(26.44 \pm 0.70)\%$ of Λ candidates are reconstructed from long tracks. On contrary
1165 in simulation of the same decay, only $21.15 \pm 0.24\%$ of candidates are reconstructed
1166 from long tracks. The same ratio on phase space simulation of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ is
1167 $21.54 \pm 0.14\%$ (integrated over q^2). From this we conclude that this ratio is expected
1168 to be similar in the two samples and that simulation does not fully match data.
- 1169 Furthermore, an other effect, which may be related to this is affecting the simulation
1170 of the reconstruction efficiency for events with down-down tracks. Other analysis
1171 in LHCb, using particles reconstructed with downstream tracks, showed that the
1172 efficiency for these events is not well simulated in the Monte Carlo. For example in
1173 Fig. 6.1 we show the ratio between DD efficiency in data and MC which was found
1174 analysing K_S events[4]. This effect is not yet fully understood and currently under
1175 study. Main effects seem to be due to a poor simulation of the vertexing efficiency
1176 using downstream tracks.
- 1177 We deal with these effects in two steps. Firstly, we do the analysis completely

separately for down-down and long long events and we join the results at the end.
Since efficiencies are also calculated separately this takes in account for the poorly simulated ratio between the two. In a second step we assign a further systematic on the relative efficiency for down-down events only. In order to do this we weight our Monte Carlo, both for rare and resonant channel, by the efficiency ratio between data and MC found for K_S as a function of momentum. Then we compare corrected and uncorrected result taking the full difference as systematic. Dependencies due to the different momentum distributions of Λ and K_S are assumed to be negligible since the discrepancy shows little dependence on momentum. In table 6.7 are reported vertexing systematic values for reconstruction, trigger and MVA efficiencies.

6.2.5 Data-MC discrepancies

We reweighted the MC as described in sec.5.3.2. We checked how much this influences the efficiencies by comparing values obtained with and without reweighting. The effect is negligible with respect to other systematics considered. After the kinematical reweighting there could still be some data-MC discrepancies. In particular the PID variables, used in the Neural Networks, are not perfectly described in the MC. We checked if this could be a source of systematics by comparing sideband subtracted data Monte Carlo re-weighted for the A_b^0 kinematics and extracting a further weight to match the PID variable. Figure 6.2 shows the difference between the efficiency calculated with and without the further PID weight. In all q^2 bins the difference is always compatible with zero within one sigma and therefore we do not add any further systematic due to the PIDmu variable modelling.

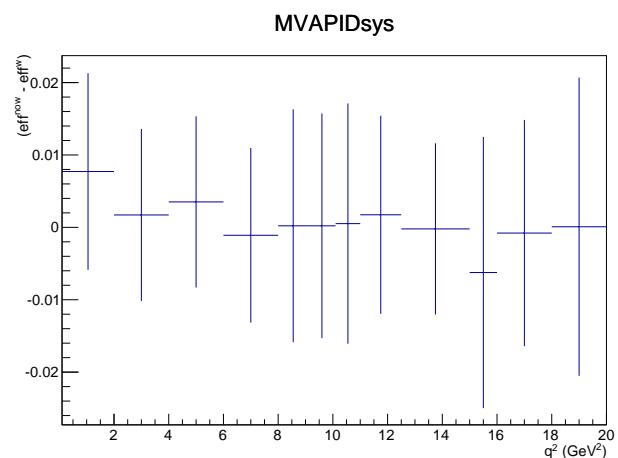


Figure 6.2: Difference between the efficiency calculated with and without the weight to match the PIDmu distributions in data and MC as a function of q^2 .

1200

CHAPTER 7

1201

1202

Summary and results

1203

1204 In this section we report values of corrected relative $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ branching ratio as
 1205 a function of q^2 . These values are directly obtained from the fit by parameterising
 1206 the DD and LL yields the following formula:

$$N(\Lambda\mu^+\mu^-)_k = \left[\frac{d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2}{\mathcal{B}(J/\psi\Lambda)} \right] \cdot N(J/\psi\Lambda)_k \cdot \epsilon_k^{\text{rel}} \cdot \frac{\Delta q^2}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}, \quad (7.1)$$

1207 where $k = \text{LL,DD}$, Δq^2 is width of the q^2 bin and the only free paramater is the
 1208 relative branching fraction ratio over the J/ψ channel. For the $J/\psi \rightarrow \mu^+\mu^-$ we use
 1209 the value reported in the PDG book, $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$.

1210 In table 7.1 are summarised relative efficiencies for DD and LL events together with
 1211 their correlated and uncorrelated errors, where the correlation is intended between
 1212 DD and LL results. In the table is reported the absolute value of the total relative

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

Table 7.1: Absolute values of the total relative efficiency and the absolute value of the uncorrelated error, together with relative values of correlated error.

efficiency and the absolute value of the uncorrelated error. This corresponds to the systematic error on the efficiency, including the systematic due to the variation of Wilson Coefficients in the decay model as described in sec. 5.3.1.1. The correlated error is instead given in a relative form since it can be applied to either DD, LL and their combination. This includes the pdf systematic and the systematic due to the uncertainty on $J/\psi \rightarrow \mu^+\mu^-$ branching ratio.

First of all we report in Fig. 7.1, the branching ratio obtained by fitting the DD and LL samples independently. Values are also reported in table 7.2, where we show separately the statistical error on the rare channel (stat), the statistical error on the normalisation channel (norm), which is correlated between q^2 bins, and the total systematic error.

The combined result, obtained fitting both samples simultaneously is shown in fig. 7.2 and values are reported in table 7.3.

The statistical error is calculated using the MINOS tool, returning an asymmetric interval. The normalisation and systematic errors are evaluated by pushing the efficiencies and normalisation yields up and down by their errors. DD and LL errors are different so each one is taken in account properly. Finally we look for the effect on the result and we take the full difference with the default fit as systematic in each direction.

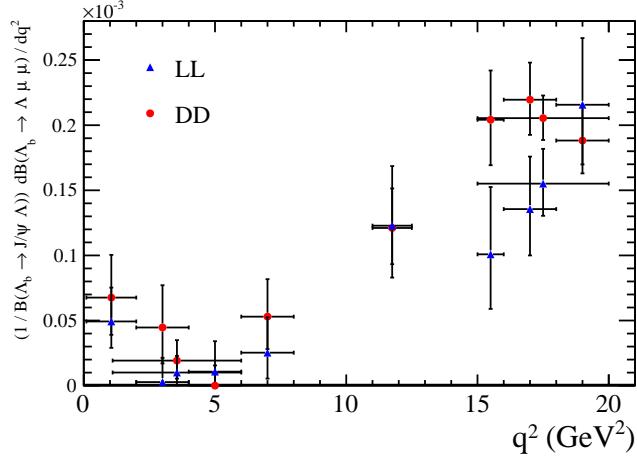


Figure 7.1: Corrected relative $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ yield as a function of q^2 bins for down-down and long-long events. Errors shown represent statistical and systematic.

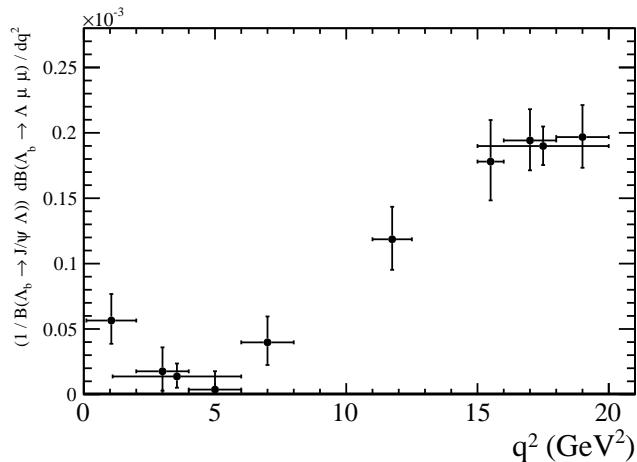


Figure 7.2: Combined relative $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ yield as a function of q^2 bins . Errors shown represent statistical and systematic.

q^2 bin	$d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2/\mathcal{B}(J/\psi\Lambda) (10^{-3})$
Down-down events	
0.1-2.0	$0.0676^{+0.0323}_{-0.0279}(\text{stat})^{+0.0055}_{-0.0064}(\text{sys})^{+0.0008}_{-0.0008}(\text{norm})$
2.0-4.0	$0.0447^{+0.0324}_{-0.0275}(\text{stat})^{+0.0021}_{-0.0022}(\text{sys})^{+0.0005}_{-0.0006}(\text{norm})$
4.0-6.0	$0.0000^{+0.0155}_{-0.0000}(\text{stat})^{+0.0000}_{-0.0000}(\text{sys})^{+0.0000}_{-0.0000}(\text{norm})$
6.0-8.0	$0.0530^{+0.0287}_{-0.0248}(\text{stat})^{+0.0020}_{-0.0021}(\text{sys})^{+0.0007}_{-0.0007}(\text{norm})$
11.0-12.5	$0.1211^{+0.0300}_{-0.0273}(\text{stat})^{+0.0049}_{-0.0053}(\text{sys})^{+0.0010}_{-0.0014}(\text{norm})$
15.0-16.0	$0.2042^{+0.0369}_{-0.0338}(\text{stat})^{+0.0082}_{-0.0084}(\text{sys})^{+0.0022}_{-0.0022}(\text{norm})$
16.0-18.0	$0.2196^{+0.0276}_{-0.0260}(\text{stat})^{+0.0066}_{-0.0068}(\text{sys})^{+0.0023}_{-0.0024}(\text{norm})$
18.0-20.0	$0.1882^{+0.0259}_{-0.0241}(\text{stat})^{+0.0070}_{-0.0071}(\text{sys})^{+0.0020}_{-0.0020}(\text{norm})$
1.1-6.0	$0.0192^{+0.0157}_{-0.0138}(\text{stat})^{+0.0010}_{-0.0011}(\text{sys})^{+0.0002}_{-0.0002}(\text{norm})$
15.0-20.0	$0.2054^{+0.0168}_{-0.0161}(\text{stat})^{+0.0039}_{-0.0040}(\text{sys})^{+0.0022}_{-0.0022}(\text{norm})$
Long-long events	
0.1-2.0	$0.0493^{+0.0258}_{-0.0200}(\text{stat})^{+0.0030}_{-0.0033}(\text{sys})^{+0.0009}_{-0.0010}(\text{norm})$
2.0-4.0	$0.0027^{+0.0186}_{-0.0000}(\text{stat})^{+0.0002}_{-0.0002}(\text{sys})^{+0.0001}_{-0.0001}(\text{norm})$
4.0-6.0	$0.0107^{+0.0234}_{-0.0000}(\text{stat})^{+0.0005}_{-0.0006}(\text{sys})^{+0.0002}_{-0.0002}(\text{norm})$
6.0-8.0	$0.0253^{+0.0271}_{-0.0198}(\text{stat})^{+0.0014}_{-0.0015}(\text{sys})^{+0.0005}_{-0.0005}(\text{norm})$
11.0-12.5	$0.1228^{+0.0451}_{-0.0391}(\text{stat})^{+0.0071}_{-0.0076}(\text{sys})^{+0.0020}_{-0.0020}(\text{norm})$
15.0-16.0	$0.1009^{+0.0513}_{-0.0413}(\text{stat})^{+0.0063}_{-0.0069}(\text{sys})^{+0.0016}_{-0.0017}(\text{norm})$
16.0-18.0	$0.1356^{+0.0397}_{-0.0348}(\text{stat})^{+0.0066}_{-0.0072}(\text{sys})^{+0.0022}_{-0.0022}(\text{norm})$
18.0-20.0	$0.2157^{+0.0497}_{-0.0438}(\text{stat})^{+0.0120}_{-0.0129}(\text{sys})^{+0.0035}_{-0.0035}(\text{norm})$
1.1-6.0	$0.0101^{+0.0127}_{-0.0099}(\text{stat})^{+0.0006}_{-0.0007}(\text{sys})^{+0.0002}_{-0.0002}(\text{norm})$
15.0-20.0	$0.1551^{+0.0260}_{-0.0239}(\text{stat})^{+0.0055}_{-0.0058}(\text{sys})^{+0.0025}_{-0.0026}(\text{norm})$

Table 7.2: Values of corrected relative branching fraction for DD and LL events with statistical, correlated and uncorrelated error shown separately.

q^2 bin	$d\mathcal{B}(\Lambda\mu^+\mu^-)/dq^2/\mathcal{B}(J/\psi\Lambda) (10^{-3})$
0.1-2.0	$0.0565^{+0.0198}_{-0.0172}(\text{stat})^{+0.0041}_{-0.0046}(\text{sys})^{+0.0009}_{-0.0009}(\text{norm})$
2.0-4.0	$0.0176^{+0.0183}_{-0.0148}(\text{stat})^{+0.0011}_{-0.0012}(\text{sys})^{+0.0003}_{-0.0004}(\text{norm})$
4.0-6.0	$0.0037^{+0.0141}_{-0.0037}(\text{stat})^{+0.0003}_{-0.0003}(\text{sys})^{+0.0000}_{-0.0000}(\text{norm})$
6.0-8.0	$0.0398^{+0.0197}_{-0.0172}(\text{stat})^{+0.0018}_{-0.0019}(\text{sys})^{+0.0006}_{-0.0007}(\text{norm})$
11.0-12.5	$0.1186^{+0.0242}_{-0.0226}(\text{stat})^{+0.0055}_{-0.0057}(\text{sys})^{+0.0015}_{-0.0015}(\text{norm})$
15.0-16.0	$0.1780^{+0.0306}_{-0.0283}(\text{stat})^{+0.0082}_{-0.0086}(\text{sys})^{+0.0022}_{-0.0022}(\text{norm})$
16.0-18.0	$0.1942^{+0.0228}_{-0.0216}(\text{stat})^{+0.0068}_{-0.0072}(\text{sys})^{+0.0024}_{-0.0024}(\text{norm})$
18.0-20.0	$0.1968^{+0.0230}_{-0.0218}(\text{stat})^{+0.0081}_{-0.0084}(\text{sys})^{+0.0023}_{-0.0024}(\text{norm})$
1.1-6.0	$0.0137^{+0.0099}_{-0.0085}(\text{stat})^{+0.0009}_{-0.0010}(\text{sys})^{+0.0002}_{-0.0002}(\text{norm})$
15.0-20.0	$0.1899^{+0.0141}_{-0.0136}(\text{stat})^{+0.0044}_{-0.0045}(\text{sys})^{+0.0023}_{-0.0023}(\text{norm})$

Table 7.3: Combined values of corrected relative branching fraction with statistical error on rare (stat) and normalisation (norm) channels and systematic error shown separately.

CHAPTER 8

Angular analysis

8.1 Extraction of forward-backward asymmetry

In this part we describe measurement of two forward-backward asymmetries A_{FB}^ℓ and A_{FB}^h defined as

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

where $\frac{d\Gamma}{dq^2 d\cos\theta_i}$ is two-dimensional differential rate and $\frac{d\Gamma}{dq^2}$ is rate integrated over angles. The observable A_{FB}^ℓ is the exact equivalent of the A_{FB} known from $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays. The A_{FB}^h is equivalent using baryon side rather than lepton side. The observable is usually not discussed in $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays where it is zero by definition, but as in $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ we deal with weakly decaying Λ , which has significant decay asymmetry, we are sensitive to it. The observables are extracted

from the fit of one-dimensional angular distributions. The distribution to extract A_{FB}^ℓ is (for details, see appendix ??)

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2 d \cos \theta_\ell} = \frac{d\Gamma}{dq^2} \left[\frac{3}{8} (1 + \cos^2 \theta_\ell) (1 - f_L) + A_{FB}^\ell \cos \theta_\ell + \frac{3}{4} f_L \sin^2 \theta_\ell \right], \quad (8.2)$$

where f_L is fraction of longitudinally polarised dimuons. The distribution for extraction of baryon side forward-backward asymmetry has form

$$\frac{d\Gamma(\Lambda_b \rightarrow \Lambda (\rightarrow p\pi^-) \ell^+ \ell^-)}{dq^2 d \cos \theta_\Lambda} = \text{Br}(\Lambda \rightarrow p\pi^-) \frac{d\Gamma(\Lambda_b \rightarrow \Lambda \ell^+ \ell^-)}{dq^2} \frac{1}{2} (1 + 2A_{FB}^h \cos \theta_\Lambda). \quad (8.3)$$

These expressions assume that Λ_b^0 is produced unpolarized. LHCb measurement of the production polarisation [43] suggests that is reasonable assumption. We will investigate effect of production polarisation in systematic uncertainties.

The extraction of forward-backward asymmetries is done by unbinned maximum likelihood fits. Fitting functions are composed by a theoretical shape given by eq. 8.3 and 8.2 multiplied by an acceptance function as described in sec. 8.1.1 and a background component described in sec 8.1.2. In order to limit systematic effects due to the background parameterisation, we perform the fit in a restricted mass region around the peak: $5580 < m(\Lambda \mu^+ \mu^-) < 5660$ MeV/ c^2 . In this way in most of the q^2 bins we have $\sim 20\%$ of background events. The signal over background ratio is fixed in the fit, this is done by fitting the mass distribution (in a wider mass interval) and using this fit to extract the fraction of background in the signal region.

8.1.1 Angular acceptance

The final effect we need to take into account is the detector efficiency, which can be non-uniform across angular variables. While in principle one could take it into account by an additional weight, in order to minimise the distortion of the uncertainties estimate, we incorporate the efficiency function $\epsilon(\cos \theta_{\ell,\Lambda})$ directly into the

1256 likelihood as

$$-\ln \mathcal{L} = -\sum_i \ln [f_s P(\cos \theta_{\ell,\Lambda}^i) \epsilon(\cos \theta_{\ell,\Lambda}) + (1 - f_s) P_B(\cos \theta_{\ell,\Lambda}^i)]. \quad (8.4)$$

1257 where $P(\cos \theta_{\ell,\Lambda}$ is the theoretical distribution and $P_B(\cos \theta_{\ell,\Lambda}$ the parameterised
 1258 background PDF. The efficiency is parametrised using a 2nd order polynomial. Us-
 1259 ing polynomial functions allows to calculate the PDF normalisation analytically.
 1260 The efficiency is extracted for downstream and long events separately, and a sepa-
 1261 rate model is built for each event type and each q^2 bin. The unbinned fit is then
 1262 performed simultaneously on all events and all efficiency parameters are fixed in the
 1263 final fit to the angular distributions.

1264 The efficiencies are obtained from a MC simulation using the same procedure used
 1265 for the branching ratio analysis and described in sec. ???. In Figs. 8.1, 8.2, 8.3, 8.4,
 1266 8.5 and 8.6 are reported efficiencies as a function of $\cos \theta_\Lambda$ and $\cos \theta_\ell$ using $J/\psi \Lambda$
 1267 MC and $\Lambda \mu^+ \mu^-$ MC in the 15-20 and 1.1-6 integrated q^2 bin.

1268 For the lepton side, even though the efficiency it's symmetric by construction, we
 1269 leave all parameters floating in the fit, namely we do not constrain it to be symmetric.

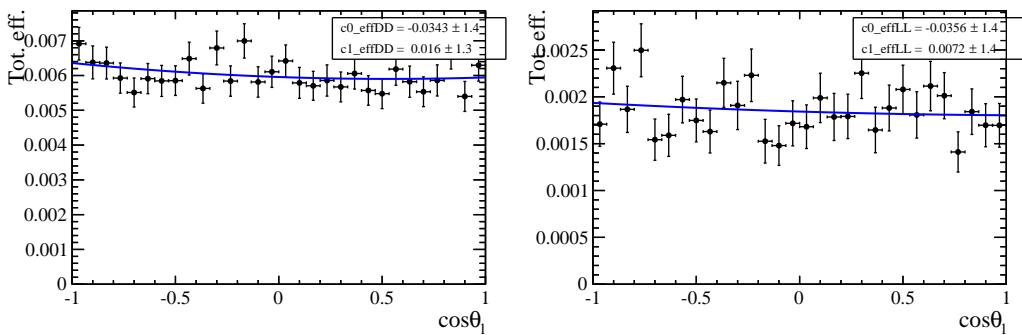


Figure 8.1: Efficiency as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events in the $15-20 GeV^2/c^2$ q^2 bin.

1270 Finally, since we are fitting 1-dimensional projections, this implies an integral over
 1271 the other angular variables. If the efficiency is not flat in these variables we could
 1272 have extra terms left in eq. 8.1 and 8.2. We take in account for this effect in the

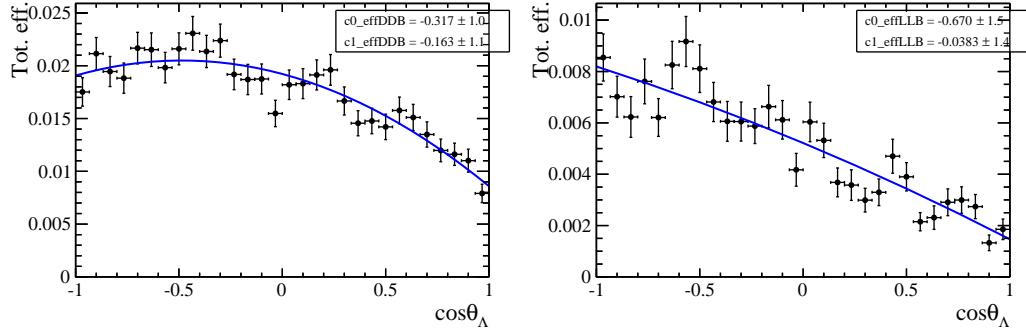


Figure 8.2: Efficiency as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events in the $15-20\text{GeV}^2/c^2 q^2$ bin.

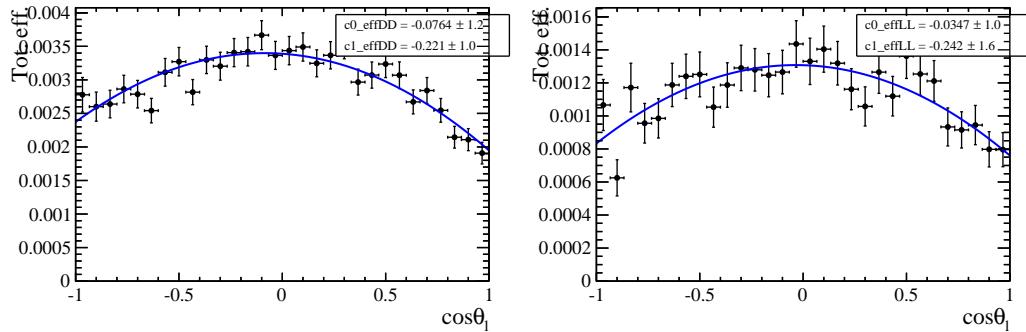


Figure 8.3: Efficiency as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events in the $1.1-6.0\text{GeV}^2/c^2 q^2$ bin.

1273 systematics as described in the following sections. In appendix ?? are reported
 1274 efficiencies as a function of $\cos\theta_\ell$ in bins of $\cos\theta_\Lambda$ in different q^2 bins and a 2D plot
 1275 of the efficiency as a function of $\cos\theta_\ell$ and $\cos\theta_\Lambda$.

1276 8.1.2 Background parameterisation

1277 We parameterise the background shape with a linear function times the efficiency.
 1278 A different efficiency shape is used for down-down and long-long events and for each
 1279 q^2 bin. The free parameter of this model is fitted on sideband events and fixed
 1280 for the final fit. All fitted sideband distributions are reported in appendix ???. A
 1281 systematic is then considered by using a different background shape (see sec.9.0.7).
 1282 In Figs. 8.7 and 8.8 are reported the background distributions in the sideband
 1283 ($m_{\Lambda\mu^+\mu^-} > 5700\text{MeV}$) for the high and low q^2 integrated bin with overlaid the

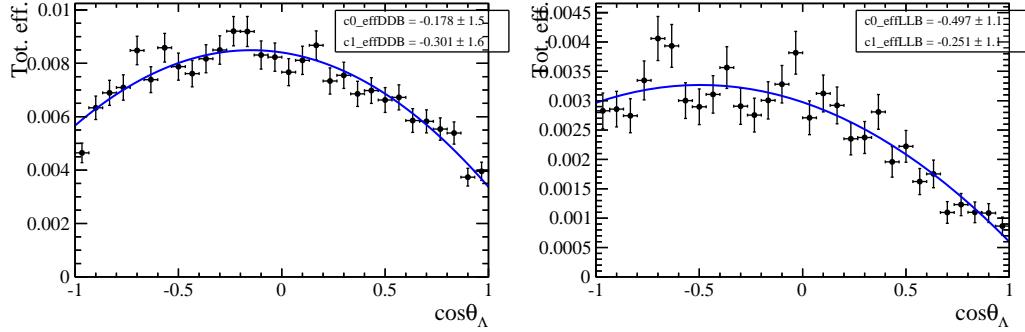


Figure 8.4: Efficiency as a function of $\cos\theta_\Lambda$ for down-down (left) and long-long (right) events in the $1.1\text{-}6.0\text{GeV}^2/c^2 q^2$ bin.

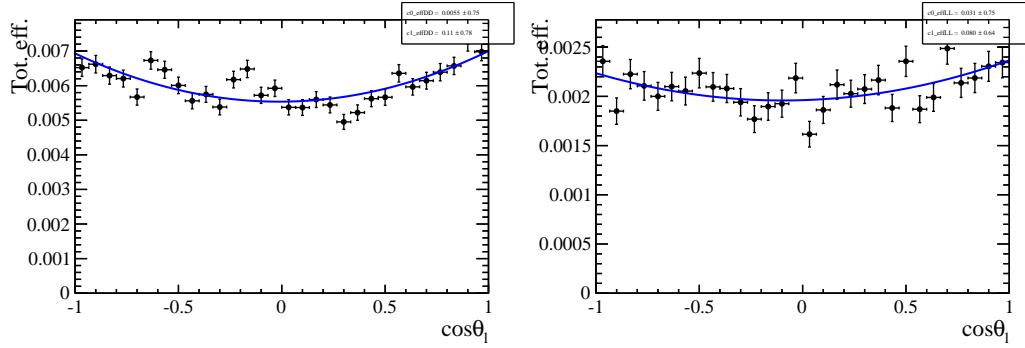


Figure 8.5: Efficiency as a function of $\cos\theta_\ell$ for down-down (left) and long-long (right) events for J/ψ events.

1284 efficiency functions alone. N.B.: In these plots the function is not fitted to the data
1285 but simply overlaid on them with a proper normalisation.

1286 In appendix ?? are reported the events in the sideband (for all q^2 bins) with overlaid
1287 the background function used for the angular fit.

1288 8.1.3 Toy studies on a 3-dimensional fit

1289 One other way of extracting the result would be to fit at the same time both angles
1290 and also mass in order to have a better handle on the level of background. In this
1291 case we can use more of the information available. On the other hand we have a
1292 larger mass window with more background in it and more parameters to fit. In the
1293 1D case we have 2 parameters of interest for the lepton case and 1 for the hadron

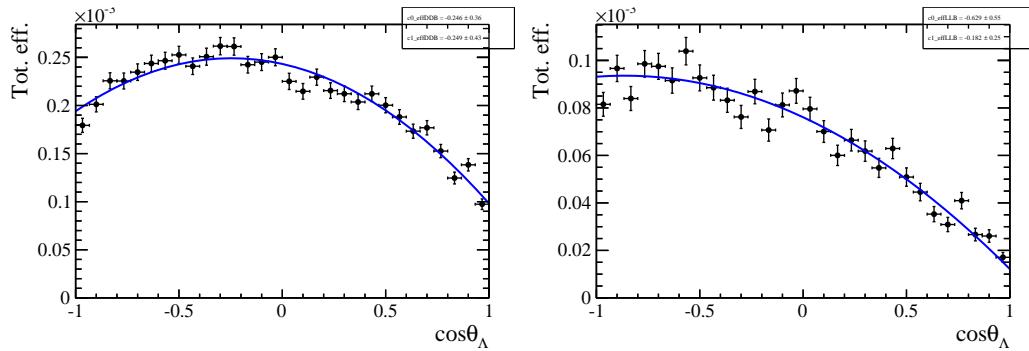


Figure 8.6: Efficiency as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events for J/ψ events.

1294 one plus 2 background fractions for DD and LL events. For the 3D case we have 3
 1295 parameters of interest plus 2 background fractions plus the 2 exponential slopes for
 1296 the mass background. Therefore we have a total of 4 free parameters for the lepton
 1297 case and 3 for hadron case fitting in 1D and 7 free parameters in the 3D fit. In both
 1298 cases we obtain the background fractons fitting mass alone and then we gaussian
 1299 constrain the fractions in the final fit. We remind that in these fits the parameters of
 1300 interest are treated differently from the others. A scan of the likelihood function is
 1301 performed on the parameters of interest, while for each point all the niusances are
 1302 fitted using a maximum likelihood fit.

1303 In order to check which one of the two methods gives the best sensitivity we generate
 1304 500 toy experiments. Events are generated in 3D using parameters taken from the
 1305 fit on real data. The generated values of the parameters of interest are $A_{FB}^\ell = 0$,
 1306 $f_L = 0.7$ and $A_{FB}^h = -0.37$. These are data-like values inspired to what we measure
 1307 in our highest statistics bin and in each toy the same values are used. The overall
 1308 statistics and the fraction of bacground events in the mass window it's obtained
 1309 from a fit on real data. This is done using the statistics of our highest statistics bin,
 1310 $15 - 20 \text{ GeV}^2/c^2$ in q^2 , and our lowest statistics unblinded bin, $11 - 12.5 \text{ GeV}^2/c^2$ in
 1311 q^2 . Each toy is fitted with both methods and we report in figure 8.9 the distibutions
 1312 of fitted parameters of interest in the 1D and 3D cases. We can use the RMS of
 1313 these distributions as a measure of how much sensitivity each method has. In table
 1314 8.1 RMSs from both methods can be compared. For all parameters of interest the

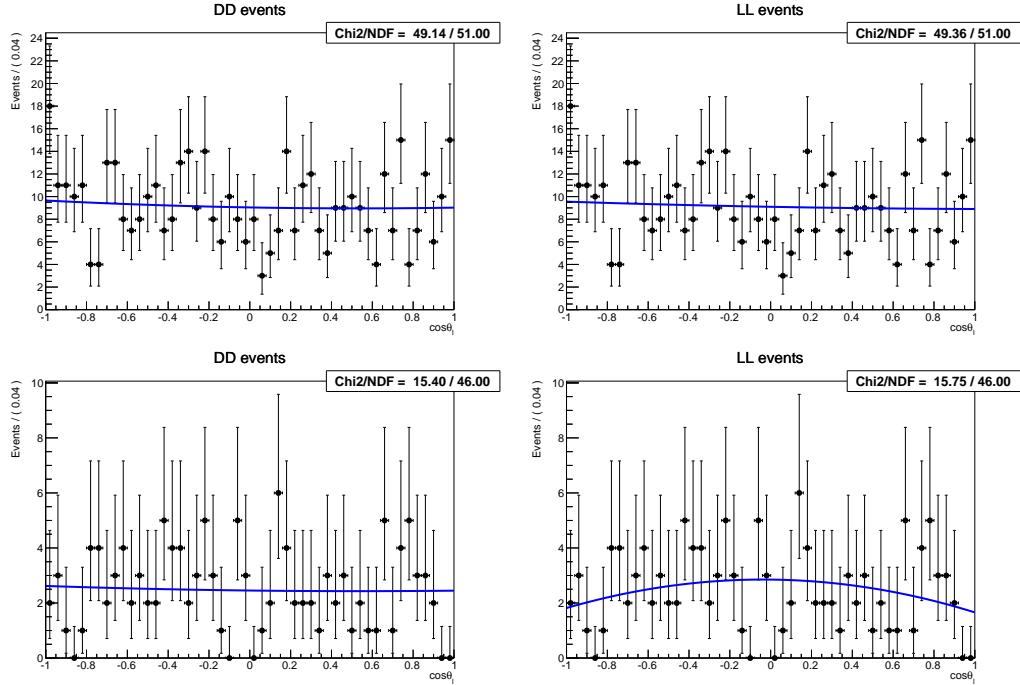


Figure 8.7: Background distribution as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events in the $15\text{-}20 \text{ } GeV^2/c^2$ (top) and $1.1\text{-}6 \text{ } GeV^2/c^2$ (bottom) q^2 bins.

q^2 bin	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

Table 8.1: RMS values for toy experiments on the extraction of the three parameters of interest with the 1D or 3D fitting methods.

1315 1D fit method gives a smaller RMS. N.B.: In these plots you can see a high peak
 1316 at -0.5 for the A_{FB}^h variable. This is the phase space boundary and the peak is due
 1317 to the fact that the fit is constrained inside the phase-space and therefore if the best
 1318 log-likelihood is found to be outside the fit chooses the point at the boundary as
 1319 best fit.

1320 In all cases we notice a slight bias in the fitting result. Therefore we assign as
 1321 a systematic the full difference we observe on toys. Values of the systematic are
 1322 reported in table 8.2 in bins of q^2 .

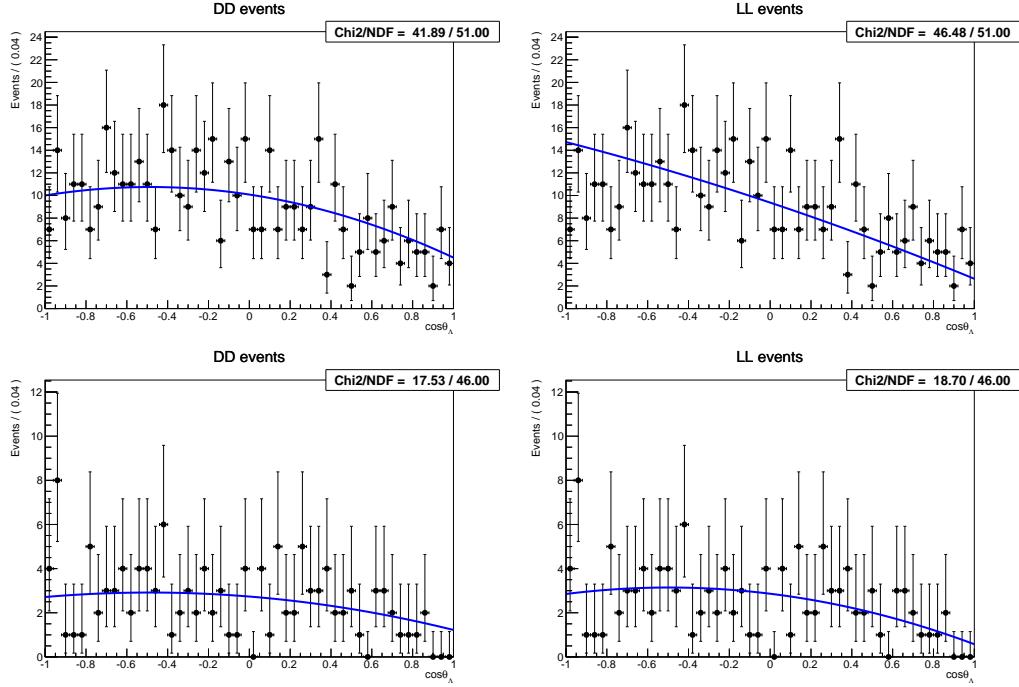


Figure 8.8: Background distribution as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events in the $15\text{-}20 \text{ GeV}^2/\text{c}^2$ (top) and $1.1\text{-}6 \text{ GeV}^2/\text{c}^2$ (bottom) q^2 bins.

q^2 bin	A_{FB}^ℓ	f_L	A_{FB}^h
0.1-2.0	-0.01126 ± 0.01272	-0.14299 ± 0.01551	0.04002 ± 0.00950
11.0-12.5	0.00123 ± 0.00574	-0.03601 ± 0.00885	0.02030 ± 0.00634
15.0-16.0	0.00592 ± 0.00555	-0.03112 ± 0.00880	0.00772 ± 0.00577
16.0-18.0	0.00973 ± 0.00389	0.00115 ± 0.00634	-0.00161 ± 0.00428
18.0-20.0	-0.00400 ± 0.00391	-0.00306 ± 0.00728	0.00258 ± 0.00458
15.0-20.0	-0.00052 ± 0.00247	-0.00424 ± 0.00441	-0.00100 ± 0.00316

Table 8.2: Absolute values of systematic uncertainties due to fit bias in bins of q^2 .

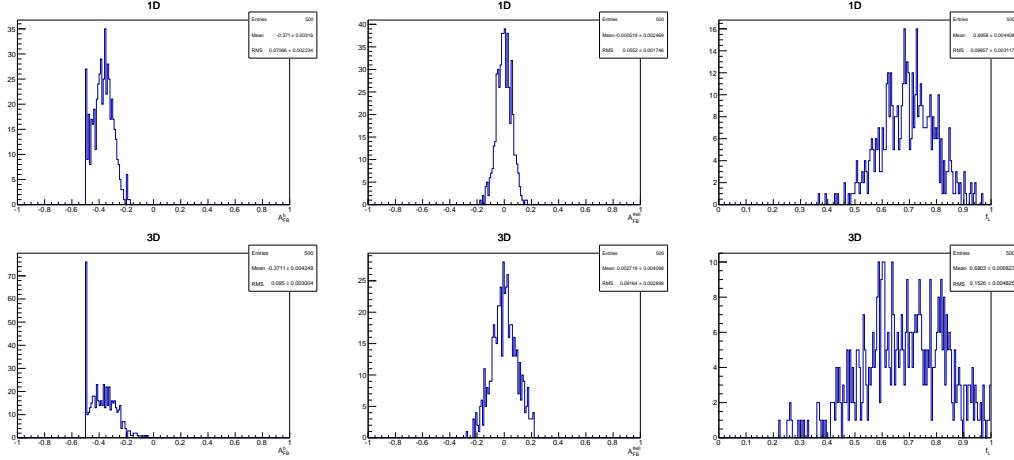


Figure 8.9: Distribution of observed parameters of interest over 500 toy experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the $15 - 20 q^2$ bin in real data.

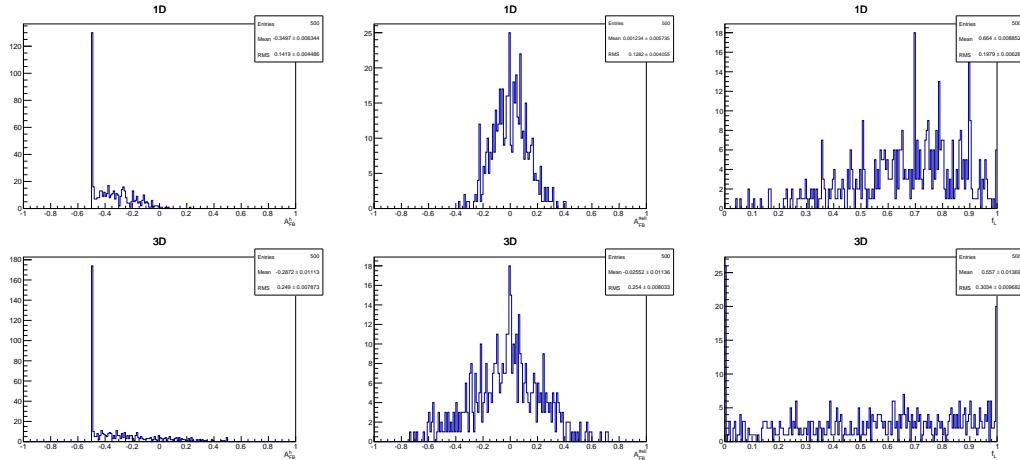


Figure 8.10: Distribution of observed parameters of interest over 500 toy experiments using the 1D fit method (top) and the 3D one (bottom). These toys correspond to events generated with parameters and statistics corresponding to what we observe in the $11 - 12.5 q^2$ bin in real data.

1323

CHAPTER 9

1324

1325

Systematics uncertainties on angular observables

1326

1327 In the following sections are explained the five main sources of systematic uncer-
1328 tainties that are considered for the A_{FB} measurement.

1329 9.0.4 Non-flat angular efficiency

1330 The angular efficiency is non-flat as a function of $\cos \theta_\ell$ and $\cos \theta_\Lambda$. Therefore, while
1331 integrating the full angular distribution, terms that cancel with perfect efficiency
1332 may remain and generate a bias in the final result. In order to deal with this effect
1333 we generate events in a 2D $(\cos \theta_\ell, \cos \theta_\Lambda)$ space using the theoretical distribution
1334 described by Eq.?? multiplied by the 2D efficiency histogram reported in Fig. ??.

1335 Then we take 1-dimensional projections and we fit them with the default 1D effi-
1336 ciency function. In Fig. 9.1 are shown the deviations from the true generated value
1337 $\Delta x = x_{true} - x_{measured}$. Since the mean of these distributions is non zero by more
1338 than 3σ , we take it as a systematic error.

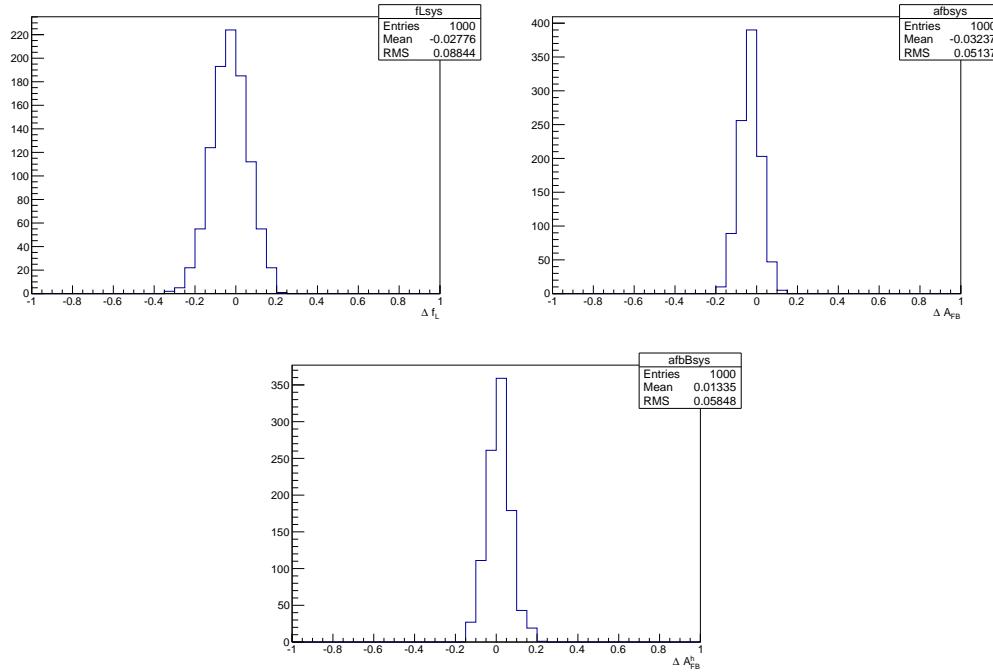


Figure 9.1: Plots show the absolute difference of the observables' values obtained fitting toy MC generated with a 2D distributions multiplied by a 2D efficiency and fitting 1D projections. From left to right for f_L , A_{FB}^l and A_{FB}^h .

1339 9.0.5 Resolution

1340 We investigated whether the angular resolution can bias the A_{FB} measurement.
 1341 If the angular distribution is asymmetric, a non-zero resolution may yield to an
 1342 asymmetric migration of events, which can then generate a bias in the value of
 1343 A_{FB} . This is especially important for $\cos \theta_\Lambda$, which has worse resolution and a
 1344 considerably asymmetric distribution. In order to study this systematic we used
 1345 toy experiments where events are generated following the measured distributions
 1346 (including efficiencies). Then these events are normally smeared by the angular
 1347 resolution. To be conservative we always use the case with biggest angular resolution,
 1348 down-down events, as reported in table ???. Finally, the smeared and not-smeared
 1349 distributions are fitted with the same angular pdf. The average deviations from the
 1350 default values are reported in table 9.1 as a function of q^2 . All are several σ
 1351 away from 0, therefore we assign these average deviations as systematics.

q^2 bin	σ_ℓ	σ_Λ	ΔA_{FB}^ℓ	Δf_L	ΔA_{FB}^h
0.1-2.0	0.0051	0.061	0.0011	-0.0022	-0.007
11.0-12.5	0.0055	0.067	0.0016	-0.0051	-0.013
15.0-16.0	0.0059	0.070	0.0006	-0.0054	-0.010
16.0-18.0	0.0064	0.070	0.0014	-0.0077	-0.010
18.0-20.0	0.0081	0.074	0.0014	-0.0062	-0.010
15.0-20.0	0.0066	0.072	0.0013	-0.0076	-0.011

Table 9.1: Values of simulated $\cos \theta_\ell$ and $\cos \theta_\Lambda$ resolutions and systematics on angular observables in bins of q^2 .

1352 9.0.6 Angular efficiency

1353 Selection requirements on the minimum momentum of the muons has the effect of
 1354 warping the $\cos \theta_\ell$ distribution removing candidates with extreme values of $\cos \theta_\ell$.
 1355 And similarly the impact parameter requirements affect $\cos \theta_\Lambda$ as very forward
 1356 hadrons tend to have lower IP. An imprecise determination of the reconstruction
 1357 and selection efficiency can introduce extra oddity and therefore bias the A_{FB} mea-
 1358 surement. In particular the baryon side distribution is asymmetric and the correct
 1359 evaluation of this asymmetry is important to avoid biases in A_{FB}^h .

1360 In order to asses this systematic we remove the kinematic re-weighting of the MC
 1361 described in sec.5.3.2 and we check what is the effect on the observables measure-
 1362 ment. We do this by fitting toy samples using the same theoretical pdf multiplied by
 1363 the efficiency function obtained with and without kinematical reweight. Finally, we
 1364 take the average bias on the final results over 1000 toy experiments as a systematic.
 1365 Results are shown in table 9.2.

q^2 bin	A_{FB}^h	A_{FB}^ℓ	f_L
0.1-2.0	0.0093	0.0020	0.0440
11.0-12.5	0.0069	0.0069	0.0027
15.0-16.0	0.0109	0.0018	0.0046
16.0-18.0	0.0159	0.0012	0.0043
18.0-20.0	0.0148	0.0030	0.0017
15.0-20.0	0.0138	0.0002	0.0046

Table 9.2: Values systematic errors due to limited knowledge of the efficiency function on the three angular observables in bins of q^2

1366 Furthermore, we take in account for the effect of the limited MC statistics. This
1367 is not already taken in account by the Feldman-Cousins plug-in method (see 10.1)
1368 since the efficiency parameters are always fixed in the fit. We generate toy MC events
1369 using the measured values of angular observables and the default efficiency. Then
1370 we vary the efficiency with its errors and we refit with the new efficiency model. In
1371 each toy we generate a number of events comparable to the one we have in data in
1372 bins of q^2 . The average deviation over 1000 toys is taken as systematic, values are
1373 reported in table 9.3. N.B.: Efficiency parameters are always fixed when fitting the
1374 angular distributions.

1375 Finally, we checked polarisation effects of the efficiency description by varying the
1376 polarisation parameter in the MC weighting from $-\sigma$ to $+\sigma$ for the central measured
1377 value, as done for the branching ratio analysis 6.2.2. Then again we generate toys and
1378 fit them using efficiency models obtained using with different values of polarisation.
1379 No significant effect is found on the final results so we don't assign extra systematics.

q^2 bin	A_{FB}^ℓ	f_L	A_{FB}^h
0.1-2.0	0.00151	0.00170	0.00213
11.0-12.5	0.00121	0.00154	0.00196
15.0-16.0	0.00004	0.00017	0.00103
16.0-18.0	0.00065	0.00246	0.00417
18.0-20.0	0.00023	0.00372	0.00162
15.0-20.0	0.00039	0.00091	0.00137

Table 9.3: Values systematic errors due to MC statistics on the three angular observables in bins of q^2

1380 9.0.7 Background parameterisation

1381 Since there is a certain degree of arbitrariness in the choice of a parameterisation for
1382 the background, especially in bins with low statistics, we study if this may be a
1383 source of systematic. We generate toy MC experiments using the shapes from real
1384 data fits and same statistics as in each q^2 bin. Then we fit each toy with 2 models:
1385 the default one, a "line times efficiency" function as described in 8.1.2, and a "flat
1386 times efficiency" function. We take the average bias over 100 toys as systematic.

¹³⁸⁷ Results are reported in table 9.4. The most affected observable is f_L , while the least
¹³⁸⁸ affected is A_{FB}^ℓ .

q^2 bin	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

Table 9.4: Values of systematics due to the choice of background parameterisation in bins of q^2 .

¹³⁸⁹

¹³⁹⁰ 9.0.8 Polarisation

¹³⁹¹ To study the effect of a non-zero Λ_b production polarisation we generate events using
¹³⁹² the distributions ?? and ?? as a function of our angular observable and $\cos \theta$, which is
¹³⁹³ sensitive to polarisation. Similarly to the procedure we used for the branching ratio
¹³⁹⁴ measurement, we generate events for three values of the polarisation corresponding
¹³⁹⁵ to $\pm \sigma$ from the recent LHCb measurement and its central value ($Pb = 0.06 \pm 0.09$
¹³⁹⁶ [39]). In the theoretical distributions $\cos \theta$ is always odd therefore with perfect
¹³⁹⁷ efficiency it always drops out by integrating over $\cos \theta$. Therefore the generated
¹³⁹⁸ distributions contain also the information of the 2-dimensional efficiency. In fig 9.2
¹³⁹⁹ are reported 2D efficiencies as a function of $\cos \theta_\ell$ and $\cos \theta_A$ versus $\cos \theta$. And in fig
¹⁴⁰⁰ 9.3 are reported distributions of the absolute difference between the observable value
¹⁴⁰¹ obtained fitting generated events with the two polarisation values. For the integrated
¹⁴⁰² high q^2 region we obtain the following average deviations: $\Delta f_L = 0.0016 \pm 0.0012$,
¹⁴⁰³ $\Delta A_{FB}^\ell = 0.00048 \pm 0.00074$ and $\Delta A_{FB}^h = 0.00014 \pm 0.00075$. Since all average
¹⁴⁰⁴ differences are consistent with zero within less than 1.5σ we do not assign extra
¹⁴⁰⁵ systematics for polarisation.

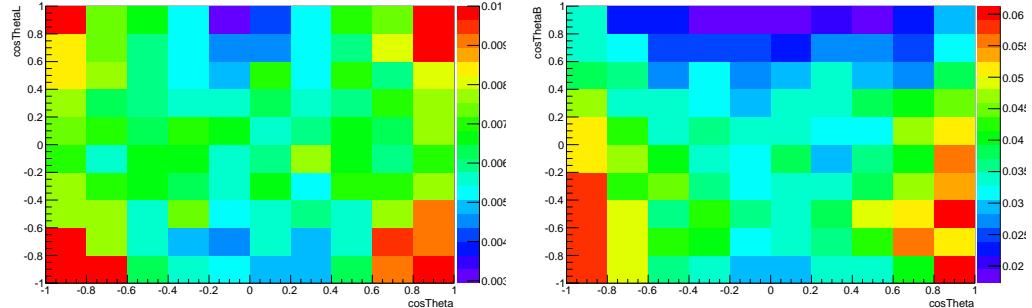


Figure 9.2: 2-dimensional efficiencies obtained from weighted MC for $\cos \theta_\ell$ and $\cos \theta_A$ versus $\cos \theta$.

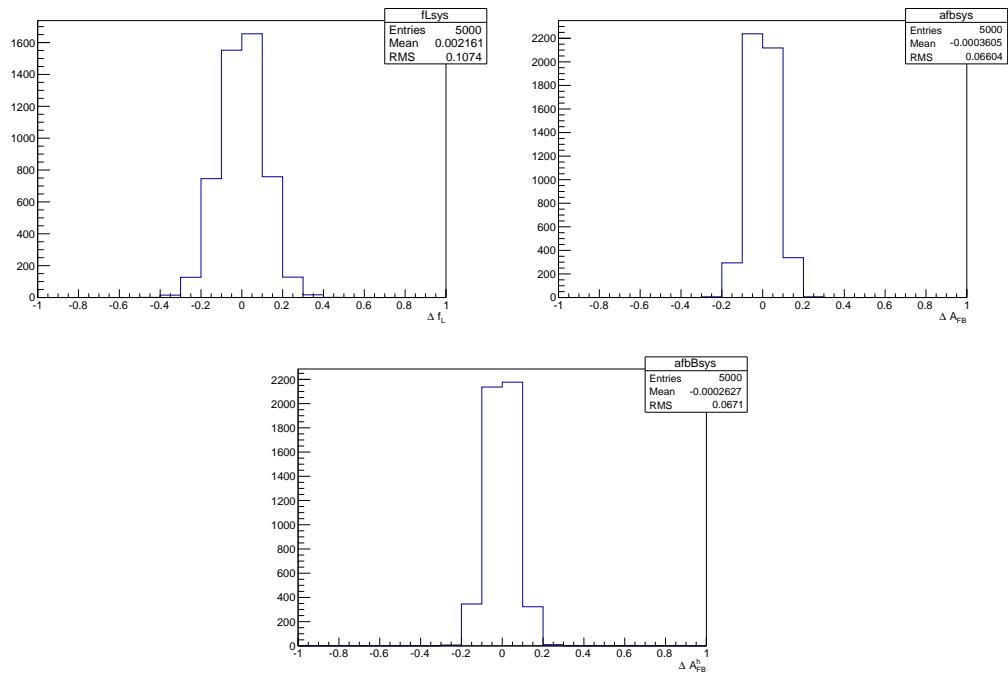


Figure 9.3: Plots show the absolute difference of the observables' values obtained fitting toy MC generated with polarisation set to -0.03 and $+0.15$. From left to right for f_L , A_{FB}^ℓ and A_{FB}^h .

1406

CHAPTER 10

1407

1408

Angular analysis summary and results

1409

1410 In this section is described the fitting technique applied on the angular distributions
1411 and results are reported.

1412 The lepton side pdf is positive only in the $|A_{FB}^\ell| < 3/4(f_L - 1)$ area. If the minimum
1413 is close to the border minuit doesn't always converge. For this reason we apply a
1414 "brute force" fitting technique. We scan the full allowed area manually looking for
1415 the minimum of the $\log\mathcal{L}$. Then we take the minimum found and we reiterate the
1416 procedure 2 more times in smaller intervals around it.

1417 10.1 Feldman-cousins plug-in method

1418 Feldman-Cousins plug-in method [44] is a unified method to calculate confidence
1419 intervals and upper/lower limits, based on toy Monte Carlo. We use this approach
1420 since the lepton side angular pdf is positive only for a restricted region of the pa-

¹⁴²¹ rament space, as shown in figure 10.1, and this fact can bias the error calculation.
¹⁴²² This method also has the advantage of having a well defined frequentist coverage.

¹⁴²³ One defines two classes of parameters: Parameters of Interest (PoI), which are those
¹⁴²⁴ for which one wants to construct confidence intervals, and nuisance parameters,
¹⁴²⁵ which are all the secondary parameters in a model. In our analysis the baryon side
¹⁴²⁶ has one parameter only, A_{FB}^h , which we consider PoI and the lepton side has two
¹⁴²⁷ which we alternatively consider PoI and nuisance.

¹⁴²⁸ The method is constituted by the following steps:

- ¹⁴²⁹ • fit the real data distribution with all parameters free;
- ¹⁴³⁰ • fit real data fixing the PoIs to a value of choice and nuisance parameters free;
- ¹⁴³¹ • generate Monte Carlo events following the distribution given by the fit model,
¹⁴³² where all nuisance parameters are taken from the fit in point 2 and PoIs are
¹⁴³³ fixed to the same value as in point 2;
- ¹⁴³⁴ • repeat the two fits made on data on each toy MC: fit with all parameters free
¹⁴³⁵ and fit with PoIs fixed;
- ¹⁴³⁶ • extract the value of the log-likelihoods at the minimum for all cases;
- ¹⁴³⁷ • calculate the percentage of times where the ratio $\log\mathcal{L}_{fixed}/\log\mathcal{L}_{free}$ is smaller
¹⁴³⁸ in data than in the toy MC.
- ¹⁴³⁹ • repeat the procedure for may values of the PoIs scanning around the best fit
¹⁴⁴⁰ values.

¹⁴⁴¹ The 1σ confidence interval at k% is given by the points (1-k)% of times the free-
¹⁴⁴² to-fixed likelihood ratio is bigger in data than MC. As an example, in figure 10.2 is
¹⁴⁴³ reported the p-values plot obtained with the plug-in method around the measured
¹⁴⁴⁴ values for A_{FB}^h . For the leptons case we have instead a 2-dimensional grid of p-values.

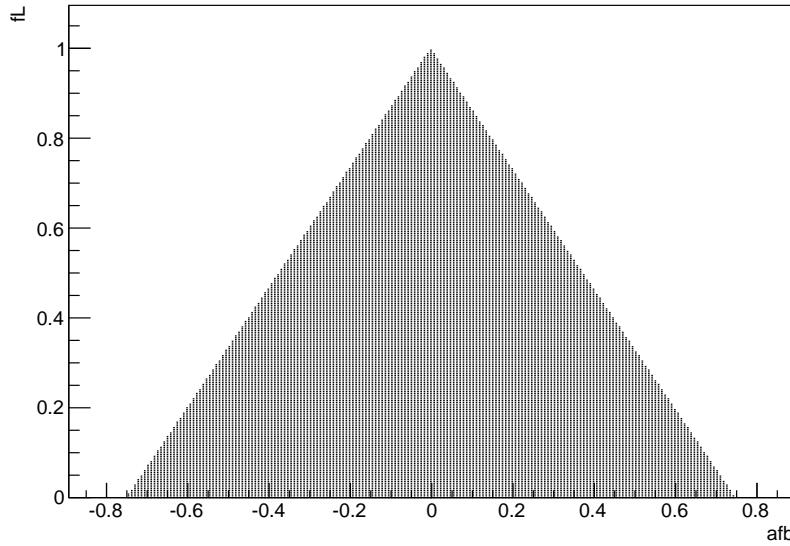


Figure 10.1: The plot shows the physical A_{FB}^ℓ vs f_L parameter space. The dark region corresponds to points where the pdf is positive in the whole $[-1, 1]$ interval.

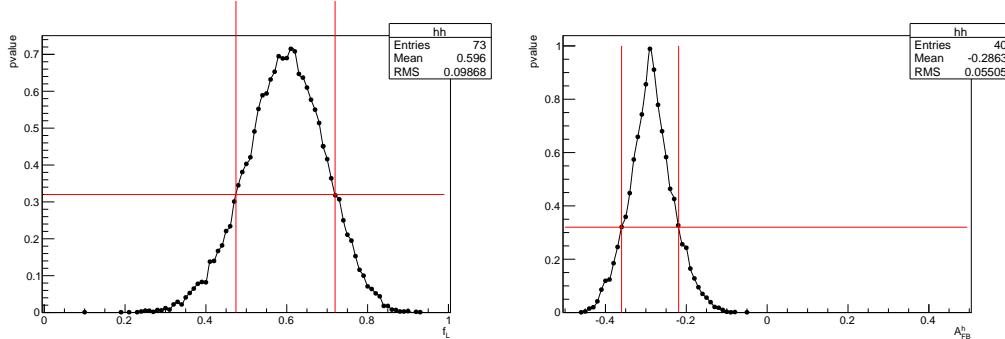
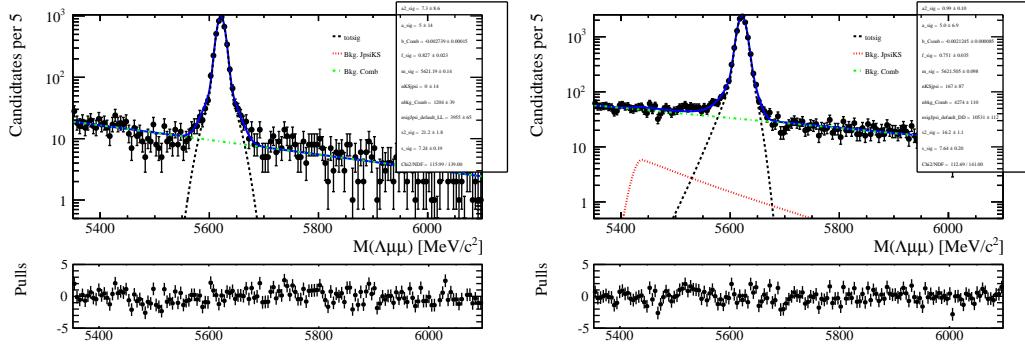


Figure 10.2: p-values obtained with the plug-in method around the measured values in the high q^2 integrated bin for f_L (left) and A_{FB}^h (right). The red line marks the points at p-value 32% corresponding to a 68% CL.

¹⁴⁴⁵ 10.2 J/ψ cross-check

¹⁴⁴⁶ In order to cross-check our fitting procedure we apply it on J/ψ events, for which we
¹⁴⁴⁷ have higher statistics. To select these events we apply the same selection ad for the
¹⁴⁴⁸ branching fraction (see sec. 4) with the addition of a strong PID cut on the proton
¹⁴⁴⁹ ($PID_{p \rightarrow p} > 10$), needed to reduce the K_s^0 J/ψ background. This is particularly
¹⁴⁵⁰ important for the $\cos\Theta_\Lambda$ fit, since the K_s^0 events are not distributed in a flat
¹⁴⁵¹ way in this variable and therefore can bias the fit. In Fig. 10.3 are reported mass



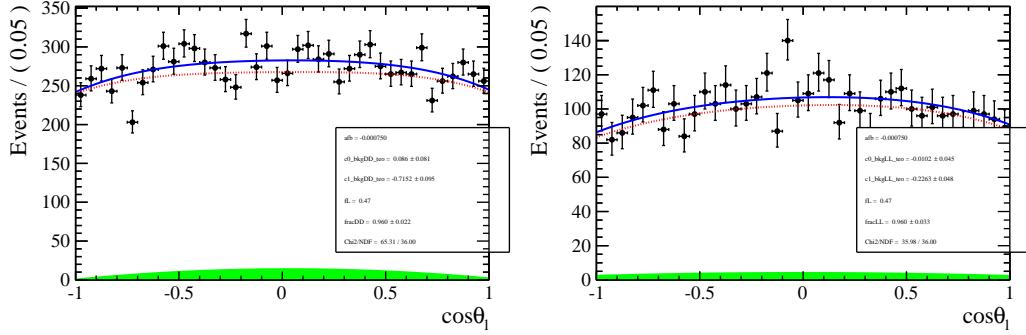


Figure 10.4: Fitted angular distribution as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events for J/ψ events.

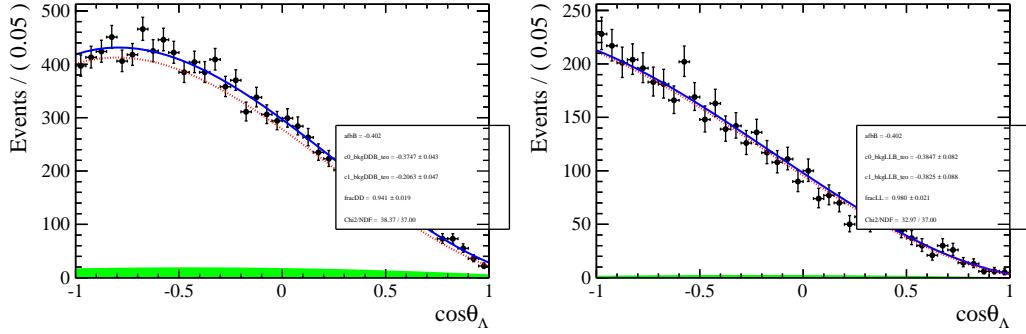


Figure 10.5: Fitted angular distribution as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events for J/ψ events.

In tables 10.1 and 10.2 are reported values of A_{FB}^l , A_{FB}^h and f_L in the same q^2 binning used for the branching fraction measurement. The statistical errors shown are asymmetric 68% CL intervals, obtained using the Feldman-Cousins method. For the $\cos \theta_{ell}$ case we report statistical errors as 2D confidence regions shown in Fig.10.8. Systematic errors are square root sum of the sources considered.

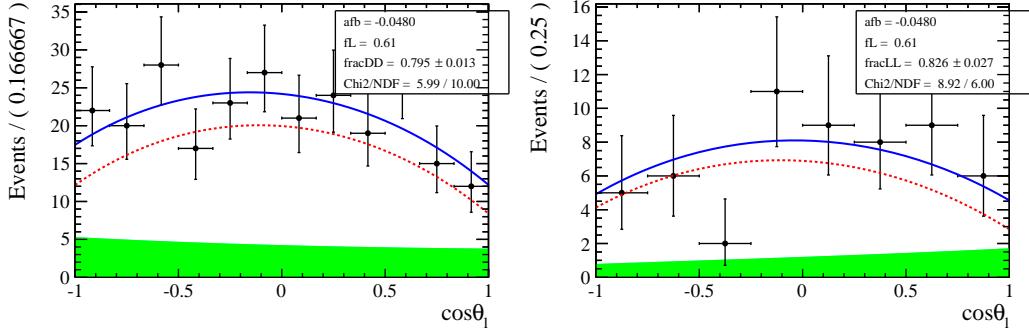


Figure 10.6: Fitted angular distribution as a function of $\cos \theta_\ell$ for down-down (left) and long-long (right) events in the $15\text{-}20\text{GeV}^2/c^2 q^2$ bin.

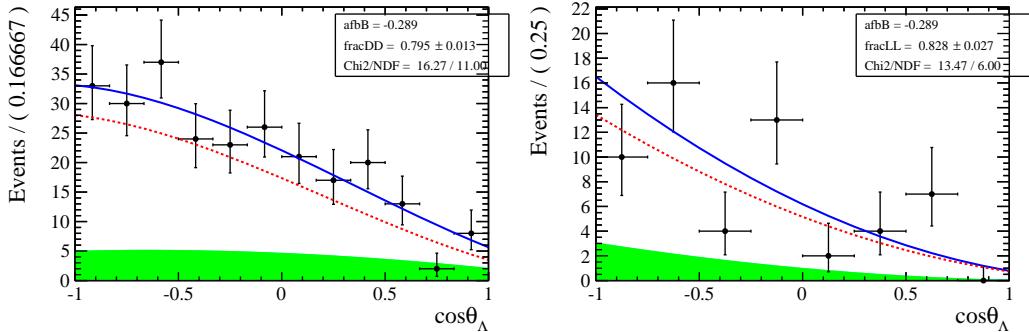


Figure 10.7: Fitted angular distribution as a function of $\cos \theta_\Lambda$ for down-down (left) and long-long (right) events in the $15\text{-}20\text{GeV}^2/c^2 q^2$ bin.

q^2 bin	A_{FB}^ℓ	f_L
0.1 - 2.0	$0.3710^{+0.370}_{-0.475}(\text{stat}) \pm 0.034(\text{sys})$	$0.559^{+0.231}_{-0.559}(\text{stat}) \pm 0.083(\text{sys})$
11.0 - 12.5	$0.0083^{+0.1944}_{-0.1807}(\text{stat}) \pm 0.056(\text{sys})$	$0.398^{+0.366}_{-0.359}(\text{stat}) \pm 0.058(\text{sys})$
15.0 - 16.0	$-0.1027^{+0.1821}_{-0.1649}(\text{stat}) \pm 0.034(\text{sys})$	$0.491^{+0.305}_{-0.301}(\text{stat}) \pm 0.048(\text{sys})$
16.0 - 18.0	$-0.0675^{+0.1312}_{-0.1178}(\text{stat}) \pm 0.043(\text{sys})$	$0.683^{+0.148}_{-0.210}(\text{stat}) \pm 0.046(\text{sys})$
18.0 - 20.0	$0.00675^{+0.1459}_{-0.143}(\text{stat}) \pm 0.035(\text{sys})$	$0.620^{+0.243}_{-0.271}(\text{stat}) \pm 0.042(\text{sys})$
15.0 - 20.0	$-0.048^{+0.086}_{-0.086}(\text{stat}) \pm 0.034(\text{sys})$	$0.612^{+0.108}_{-0.138}(\text{stat}) \pm 0.032(\text{sys})$

Table 10.1: Measured values of lepton side angular observables with systematic errors. The statistical errors are reported in Fig.10.8 evaluated as 2D 68% confidence regions. Errors reported on this table are estimations obtained using the Feldman-Cousins method where one only of the two observables is treated as parameter of interest at a time.

q^2 bin	A_{FB}^h
0.1 - 2.0	$-0.116^{+0.308}_{-0.284}$ (stat) ± 0.151 (sys)
11.0 - 12.5	$-0.500^{+0.104}_{-0.090}$ (stat) ± 0.043 (sys)
15.0 - 16.0	$-0.188^{+0.139}_{-0.161}$ (stat) ± 0.032 (sys)
16.0 - 18.0	$-0.437^{+0.096}_{-0.052}$ (stat) ± 0.031 (sys)
18.0 - 20.0	$-0.128^{+0.094}_{-0.124}$ (stat) ± 0.032 (sys)
15.0 - 20.0	$-0.289^{+0.070}_{-0.071}$ (stat) ± 0.026 (sys)

Table 10.2: Measured values of A_{FB}^h observable with statistical errors, showing 68% CL interval, and systematic errors.

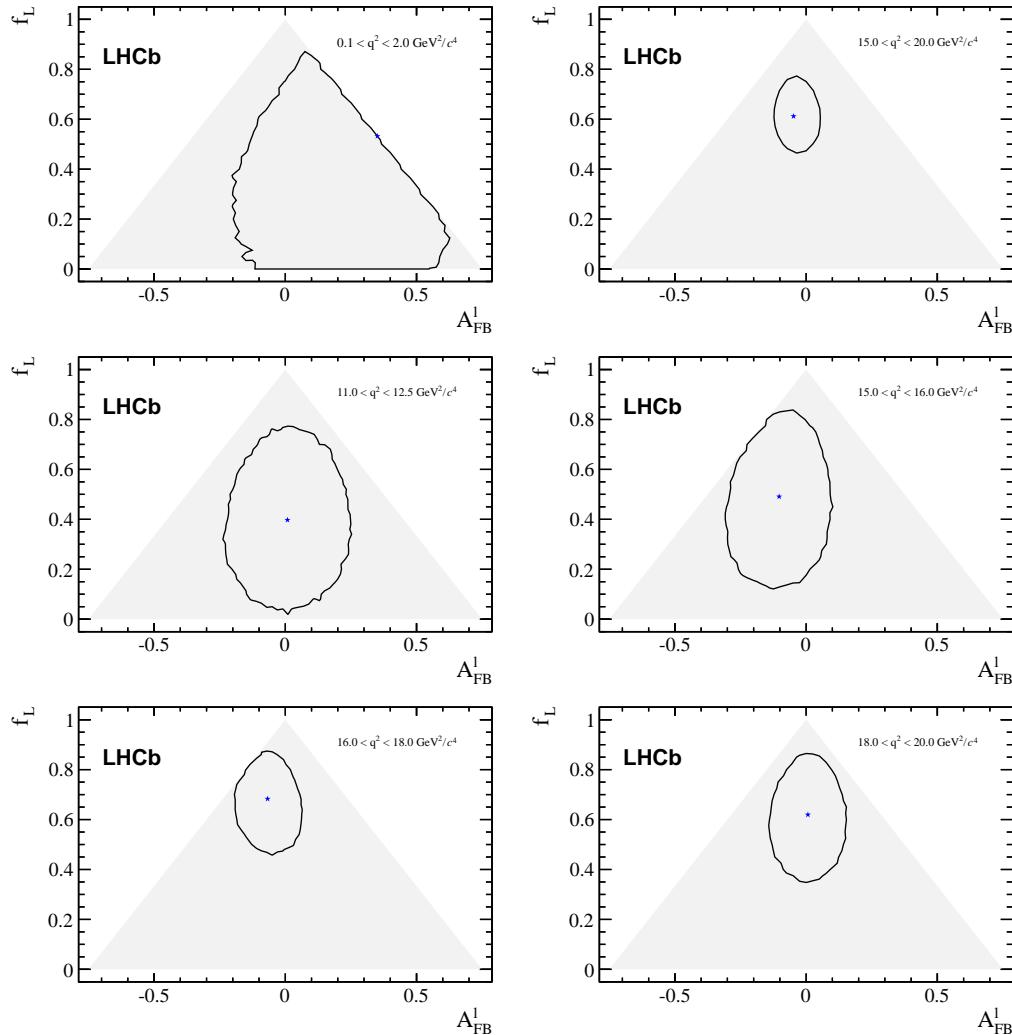


Figure 10.8: In black 2-dimensional 68% CL confidence regions as a function of A_{FB}^ℓ and f_L .

¹⁴⁷⁴

Part II

¹⁴⁷⁵

The R_{K^*} analysis

1476

CHAPTER 11

1477

1478

Towards new physics: R_H

1479

1480 Flavor-Changing Neutral Currents (FCNCs) processes, where a quark changes its
1481 flavour without altering its electric charge, are forbidden in the SM at tree level
1482 and arise only at one loop, typically by the exchange of a W boson. Hence, they
1483 are sensitive to quantum corrections by loops of heavy particles at and above the
1484 electroweak scale (~ 100 GeV). The rare decays $b \rightarrow s\gamma$, $b \rightarrow sg$, $b \rightarrow s\ell^+\ell^-$,
1485 where $\ell = e$ or μ , are good probes for these processes. In particular, in this work,
1486 decays of $b \rightarrow s\mu^+\mu^- (e^+e^-)$ type, are considered. These decays happen through
1487 loop diagrams, called “penguins diagrams”. Fig. 11.1 shows the possible Feynman
1488 diagrams producing semileptonic $B^0 \rightarrow K/K^*\ell^+\ell^-$ decays while Fig. 11.2 shows how
1489 the Feynman diagrams of these processes may include new particles and therefore
1490 be used to probe new physics. A series of recent LHCb measurements [45] shows
1491 tensions with SM predictions, which makes it interesting to further investigate these
1492 processes.

1493 In order to exploit the sensitivity of loop diagrams, in 2004 Hiller and Kruger pro-

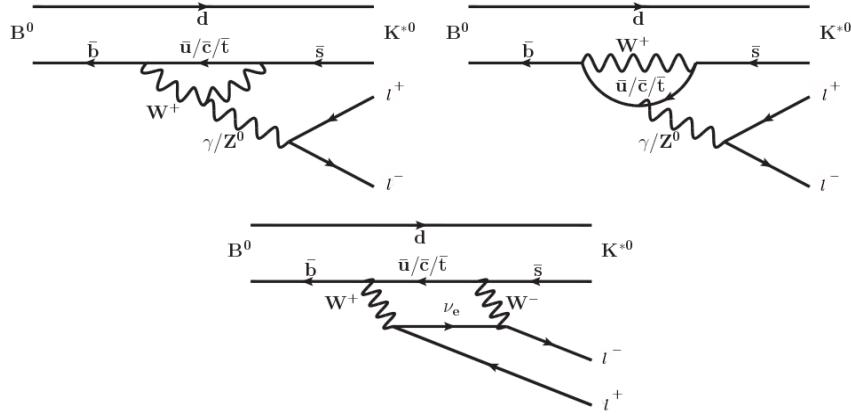


Figure 11.1: Loop diagrams of the $B_d \rightarrow K/K^{*0} l^+ l^-$ process.

1494 posed the measurement of the R_H ratio [46], defined in Eq. 11.1, where H can be
1495 an inclusive state containing an s quark (X_s) or an s -quark resonance like K or K^* .

$$R_{K^{*0}} = \frac{\int_{4m_\mu^2}^{s_{max}} \frac{d\mathcal{B}(B^0 \rightarrow K^* \mu^+ \mu^-)}{ds} ds}{\int_{4m_\mu^2}^{s_{max}} \frac{d\mathcal{B}(B^0 \rightarrow K^* e^+ e^-)}{ds} ds} \quad (11.1)$$

1496 In this quantity the decay width is integrated over the squared dilepton mass, $s =$
1497 q^2 , starting from $s_{min} = 4m_\mu$, which is the threshold for the $\mu\mu$ process, up to
1498 $s_{max} = m_b^2$. The notation $\mathcal{B}(X \rightarrow \text{final state})$ denotes the fraction of X particles
1499 which decays in the given final state, this is called “branching ratio”. For example
1500 $\mathcal{B}(B^0 \rightarrow K^* \mu^+ \mu^-)$ is the fraction of B^0 particles which decays into $K^* \mu^+ \mu^-$ with
1501 respect to all allowed B^0 decays.

1502 The advantage of using these observables is that, in the theoretical prediction,
1503 hadronic uncertainties cancel out. Furthermore, experimentally, some of the sys-
1504 tematics also cancel out in the ratio giving a better measurement. For example,
1505 what is measured is the number of $\mu\mu$ and ee decays which happen in a certain
1506 period of time and then the luminosity (\mathcal{L}) is used to obtain a cross section (σ),
1507 using $R = \mathcal{L}\sigma$, where R is the rate with which a decay happens. The luminosity
1508 measurement is usually a source of systematic uncertainty, however it appears on
1509 both sides of the ratio and therefore cancels out.

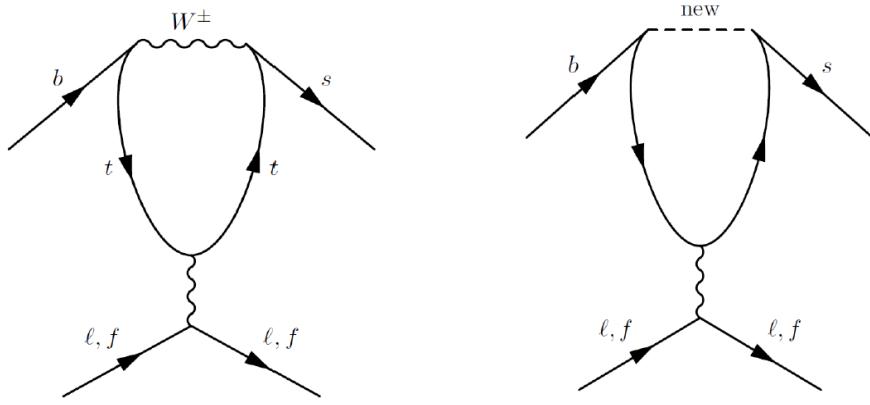


Figure 11.2: Example of penguin diagrams, on the left involving SM particles and on the right involving new possible particles.

1510 Since the Standard Model does not distinguish between leptons the predicted value
1511 for this ratio is $R_H = 1$ for massless leptons. Taking effects of order m_μ^2/m_b^2 into
1512 account Hiller and Kruger calculate that in the SM [46] and in the full q^2 range:

$$\begin{aligned} R_{X_s} &= 0.987 \pm 0.006 \\ R_K &= 1.0000 \pm 0.0001 \\ R_{K^*} &= 0.991 \pm 0.002 \end{aligned} \tag{11.2}$$

1513 Under the assumptions:

- 1514 • right-handed currents are negligible;
- 1515 • (pseudo-)scalar couplings are proportional to the lepton mass;
- 1516 • there are no CP-violating phases beyond the SM.

1517 The measurement of the R_H ratios is of particular interest after the recent measure-
1518 ment of the branching ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay [47] where no evidence of NP
1519 was found. In fact the quantities $R_H - 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ remain proportional
1520 with

$$\frac{R_H - 1}{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)} \sim 2 \cdot 10^{-5} \tag{11.3}$$

1521 A joint measurement of this two quantities can give much information and constrain
 1522 MFV models. If $R_X = 1$ and $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ is close to the SM prediction as it is
 1523 measured to be this will allow to put strong constraints on extensions of the SM. If
 1524 instead $R_H > 1$ and the equation above is not verified, this would mean that one of
 1525 the assumptions listed above are not verified, which can happen in some extensions
 1526 of the SM, for example Super-Symmetric models with broken R-parity.

1527 11.0.1 Experimental status

1528 The R_K and R_{K^*} have already been measured at the B-factories (hence in e^+e^-
 1529 collisions). And the R_K ratio has been also recently measured at LHCb [48] in the
 1530 $1 < q^2 < 6$ GeV $^2/c^4$, which provided the currently most precise measurement. This
 1531 measurement showed a 2.6σ deviation from the SM prediction. The experimental
 1532 status is summarised in Tab. 11.0.1. As for R_K LHCb is expected to reduce the
 1533 error on R_{K^*} by at least a factor of 2 with respect to the B-factories.

	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^*}	$0.93 \pm 0.46 \pm 0.12$	$0.98^{+0.30+0.08}_{-0.31-0.08}$	—

Table 11.1: Previous R_X measurement by the BaBar [1] and Belle [2] experiment.

1534

1535 11.1 The R_{K^*} analysis

1536 The aim of this analysis is to measure the R_{K^*} using $p - p$ collision data collected
 1537 by the LHCb detector in 2011 and 2012, corresponding to 3 fb $^{-1}$ of integrated
 1538 luminosity. In order to do this $B^0 \rightarrow K^*\mu^+\mu^-$ and $B^0 \rightarrow K^*e^+e^-$ (“rare channels”)
 1539 candidates are reconstructed. In both cases K^* is reconstructed through its decay
 1540 in a kaon and a pion of opposite signs.

1541 The analysis has to separate signal candidates from background candidates which
 1542 have similar observed properties. The selection presented in Sec. 12 aims to max-
 1543 imise the yield while minimising the background contamination. Two types of back-
 1544 grounds are identified: “peaking background” and “combinatorial background”. The
 1545 first comes from the mis-reconstruction of other decays or from partially recon-
 1546 structed events. This type of background, because its specific kinematic properties,
 1547 usually peaks in some variable, such as the invariant mass of all final particles,
 1548 therefore we can remove these events by removing the peak. The combinatorial
 1549 background instead comes from the random combination of particles and can be
 1550 lowered selecting events with good-quality tracks and vertices.

1551 Together with the rare channels the decays reaching the same final states via a J/ψ
 1552 resonance, $B^0 \rightarrow K^* J/\psi (\rightarrow \ell^+ \ell^-)$, are also reconstructed and referred as “charmo-
 1553 nium” or “resonant” channels. These decays have identical final states to the rare
 1554 channels, differing only in the invariant mass of the dilepton pair. As they have
 1555 much higher statistics they can be used as control samples.

1556 In Sec. 14 the efficiency of selecting and reconstructing each channel is extracted
 1557 and, finally, in Sec. 16 the R_{K^*} ratio defined is built as the double ratio of rare and
 1558 resonant channels:

$$R_{K^*} = \frac{N_{B^0 \rightarrow K^* ee}}{N_{B^0 \rightarrow K^* J/\psi \rightarrow ee}} \cdot \frac{N_{B^0 \rightarrow K^* J/\psi \mu^+ \mu^-}}{N_{B^0 \rightarrow K^* \mu^+ \mu^-}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^* J/\psi \rightarrow ee}}{\varepsilon_{B^0 \rightarrow K^* ee}} \cdot \frac{\varepsilon_{B^0 \rightarrow K^* \mu^+ \mu^-}}{\varepsilon_{B^0 \rightarrow K^* J/\psi \rightarrow \mu^+ \mu^-}} \quad (11.4)$$

1559 As no new physics is expected to affect charmonium resonances the ratio of the J/ψ
 1560 channels is 1 and therefore $R'_{K^*} = R_{K^*} \times R_{J/\psi} = R_{K^*}$. On the other hand using
 1561 the relative efficiencies between the rare and resonant channels allows to cancel out
 1562 many effects resulting in a better control of systematic uncertainties.

1563

CHAPTER 12

1564

1565

Selection

1566

1567 The selection process, described in the following sections, is divided into four steps.
1568 First of all events have to fall into the detector acceptance, produce hits and be
1569 selected on the basis of quality features, such as χ^2 of tracks and vertices. Secondly
1570 it is required that some specific trigger lines were switched on by the events. The
1571 third step consistsis in the application of particle identification (PID) conditions
1572 which remove a good part of misreconstructed background and clear the way for
1573 the last step where a neural network is used to remove combinatorial background.
1574 In order to minimise systematic uncertainties the same selection requirements are
1575 applied on the rare signal candidates and on their relative charmonium channel, a
1576 part from the q^2 cuts which serve to distinguish the two and are described in the
1577 next section.

1578 12.1 Choice of q^2 intervals and signal windows

1579 Two q^2 intervals are analysed in this work: the “central” q^2 region, [1.1,6.0] GeV^2/c^4 ,
1580 and the “high” q^2 region, above 15 GeV^2/c^4 . The central q^2 region is the most
1581 interesting place to look for new physics. In fact at low q^2 the photon pole dominates
1582 leaving little space for new physics to be found while the central region is the most
1583 theoretically clean (for more details see Sec. 1.5.3). The lower edge of this interval
1584 is chosen in order to cut out the contribution from $\phi \rightarrow \ell^+\ell^-$ decays that can dilute
1585 new physics effects. The upper edge of the interval is chosen to be sufficiently far
1586 away from the J/ψ radiative tail. In the high q^2 region the signals are still unobserved
1587 and it is therefore new territory to explore.

1588 In order to select the $B^0 \rightarrow K^*(J/\psi \rightarrow \mu^+\mu^-)$ channel a dilepton mass interval of
1589 100 MeV/c^2 around the nominal J/ψ peak [31] is selected. The electron resonant
1590 channel it is not possible to use a narrow cut at the q^2 and 4-body $m(K\pi e^+e^-)$
1591 invariant mass distributions are characterised by a long radiative tail at low masses
1592 due to bremsstrahlung radiation. Furthermore, a cut in q^2 also distorts the 4-body
1593 mass distribution at low masses and it is important to be able to fit a wide mass
1594 range to constrain backgrounds. For these reasons the interval chosen to select $B^0 \rightarrow$
1595 $K^*(J/\psi \rightarrow e^+e^-)$ candidates is chosen to go as low as possible without overlapping
1596 with the rare channel interval. The electronic resonant channel is therefore selected
1597 in the interval [6,11] GeV^2/c^4 . Fig. 12.1 shows two-dimensional distributions of q^2
1598 versus the 4-body $m(K\pi\ell^+\ell^-)$ invariant mass. On these plots horizontal lines can
1599 be seen at the q^2 corresponding to the J/ψ and $\psi(2S)$ resonances.

1600 12.2 Trigger and Stripping

1601 Events are triggered for the $\mu\mu$ and the ee channels by the trigger lines reported
1602 in Table 12.2, where the logical *and* of L0, Hlt1 and Hlt2 lines is required and the
1603 logical *or* of the lines on the same level. The candidates are required to be triggered-

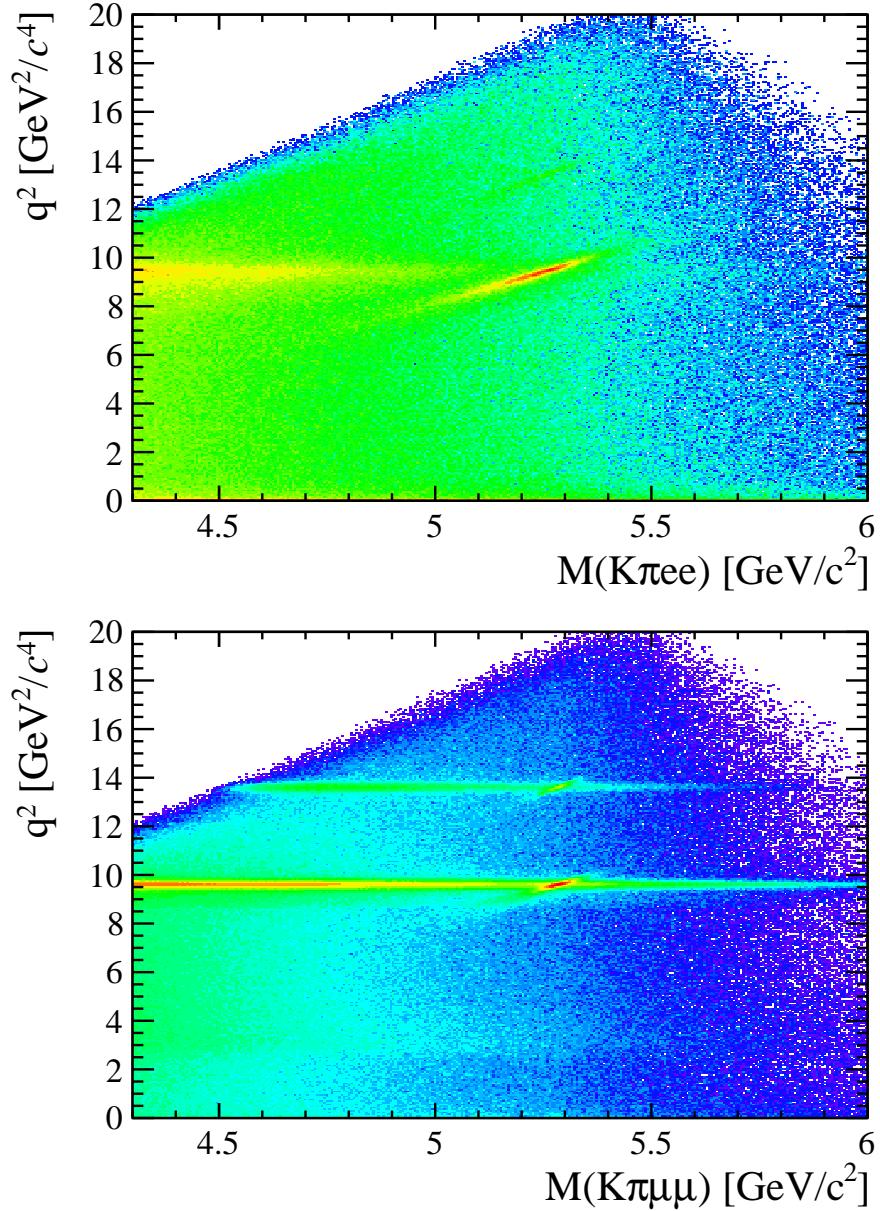


Figure 12.1: Two-dimensional distributions of q^2 versus 4-body $m(K\pi\ell\ell)$ invariant mass for the electronic (left) and muonic (right) channels in 2012 data.

1604 on-signal (TOS) for most of the stages, namely it is required for the particle which
 1605 triggered to be one of the particles used to build the signal candidates. Only for
 1606 L0Global, used in the electron case, we require a trigger-independent-of-signal (TIS),
 1607 this is aimed to collect all the possible statistics for the electron channels, which are
 1608 the most challenging. The L0Muon trigger requires hits in the muon detector, while
 1609 L0Electron and L0Hadron use information from the calorimeters; Hlt1TrackAllL0

1610 adds information from the trackers to the L0 candidates and triggers if the L0
1611 decision is confirmed; finally, Hlt2Topo[2,3]BodyBBDT uses a reconstruction of the
1612 event and a neural network trained on events with a specific topology in order to
1613 detect decays.

$\mu\mu$ candidates	ee candidates
L0Muon	L0Electron
L0DiMuon	L0Hadron
L0Hadron	L0Global (TIS)
Hlt1TrackAllL0	Hlt1TrackAllL0
Hlt1TrackMuon	
Hlt2Topo[2,4]BodyBBDT	Hlt2Topo[2,4]BodyBBDT
Hlt2TopoMu[2,4]BodyBBDT	Hlt2TopoMu[2,4]BodyBBDT
Hlt2DiMuonDetachedDecision	

Table 12.1: Summary of the trigger lines used for the $\mu\mu$ and the ee analysis.

1614 For the muonic channels the L0Muon trigger dominates instead for the electron
1615 channels all the L0 lines are important and, as they have different properties, the
1616 analysis is performed separately for three categories of events, depending on the L0
1617 trigger (for more details see Sec. 13.2).

1618 Candidates are then required to pass the kinematic and quality cuts summarised in
1619 Table 12.2. The meaning of variables in the table was already explained in Sec. ??.
1620 Loose PID cuts are applied in preselection to limit the size of the samples, while
1621 tighter cuts are applied in a second stage.

1622 Track-quality and vertex quality cuts are also applied using the χ^2_{track}/ndf , GhostProb,
1623 and χ^2_{vtx}/ndf variables. The GhostProb quantity describes the probability of a track
1624 being fake. By construction cutting at 0.4 removes $(1 - 0.4) \cdot 100 = 60\%$ of fake
1625 tracks. For details about the definition of the variables used see Ref. [32]. A large
1626 mass window is kept around the B^0 peak in order to be able to fit the sideband and
1627 to train the multivariate analysis and constrain backgrounds.

Particle	Cuts
All final	track $\chi_2/\text{ndf} < 3$ <code>GhostProb</code> < 0.4
π	$\chi_{\text{IP}}^2(\text{primary}) > 9$
K	$\text{PID}_K > -5$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>hasRICH</code>
K^*	$p_{\text{T}} > 500 \text{ MeV}/c$ $ m - m_{K^*}^{\text{PDG}} < 100 \text{ MeV}/c^2$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ Origin vertex $\chi_2/\text{ndf} < 25$
μ	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>isMuon</code>
e	$p_{\text{T}} > 300 \text{ MeV}/c$ $\chi_{\text{IP}}^2(\text{primary}) > 9$ <code>hasCalo</code> $PID_e > 0$
Dilepton	$m_{\ell\ell} < 5500 \text{ MeV}/c^2$ End vertex $\chi^2/\text{ndf} < 9$ Origin vertex χ^2 separation > 16
B^0	$\text{DIRA} > 0.9995$ End vertex $\chi^2/\text{ndf} < 9$ $\chi_{\text{IP}}^2(\text{primary}) < 25$ Primary vertex χ^2 separation > 100

Table 12.2: Summary of cuts made for pre-selection of good candidates.

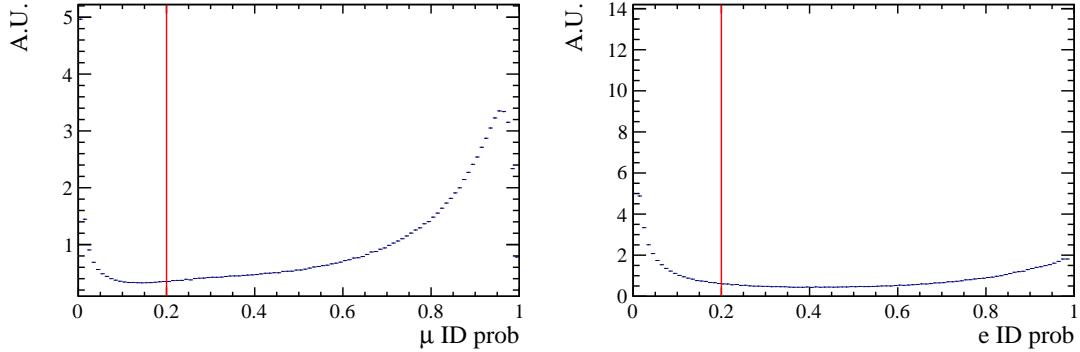


Figure 12.2: Correct ID probability distributions for muons (left) and electrons (right) in 2012 data.

1628 12.3 PID

1629 After preselection there still are high levels of misreconstructed background. In par-
1630 ticular, as the ID of kaons and pions are not constrained, the samples still contain
1631 both ID combinations for most candidates, therefore tighter PID cuts are applied.
1632 In the LHCb analysis framework the particle identification probability can be quan-
1633 tified using the “ProbNN” variables [49]. These variables are the output of a Neural
1634 Network which takes as input information from the calorimeters, the RICH detec-
1635 tors and the muon system. Unlike the DLL variables these are bounded from 0 to 1
1636 and can be therefore directly be interpreted as probabilities. For example `ProbNNk`
1637 is the probability for a reconstructed particle to be a kaon. Two tunes of the `ProbNN`
1638 variables, labelled V2 and V3, are available. Tune V3 was shown to be optimal for
1639 positive ID, while tune V3 was found to be optimal for background rejection and
1640 therefore it is used to quantify the mis-ID probability.

1641 Figure 12.3 shows distributions of the correct ID variables in the 2012 data sample
1642 while Fig. 12.3 shows in a two-dimensional plane the probability of correct identifi-
1643 cation and mis-identification of kaons and pions. These plots are characterised by
1644 clear peak at maximal ID probability and minimal mis-ID probability, corresponding
1645 to particles to which is possible to assign a well defined identification.

1646 In order to maximise the power of the PID cuts probabilities of correct ID and

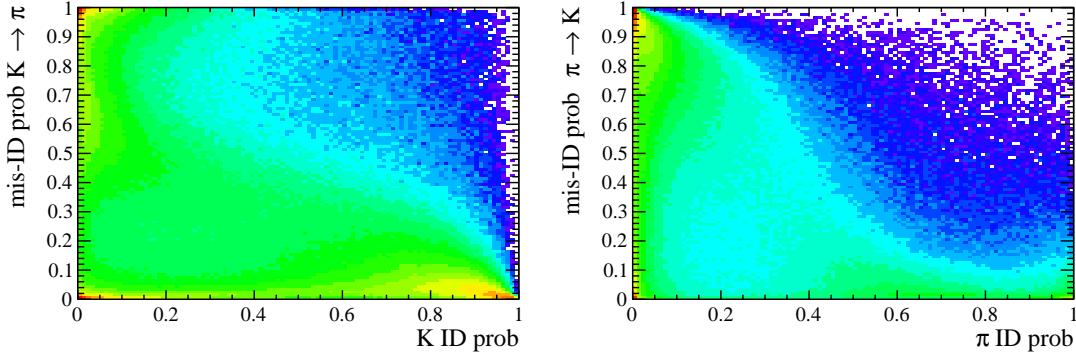


Figure 12.3: On the horizontal axis of these plots is shown the correct ID probabilities for kaons (left) and pions (right), while the vertical axis show the mis-ID probability.

¹⁶⁴⁷ mis-ID are combined using the following cuts:

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

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

¹⁶⁴⁸

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

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

¹⁶⁴⁹ In the first formula, for example, ProbNNpi is the probability of correctly identifying
¹⁶⁵⁰ the pion as a pion, while ProbNNk is the probability of mistaking it for a kaon. There-
¹⁶⁵¹ fore by maximising the quantity “ $\text{ProbNNpi} \times (1 - \text{ProbNNk})$ ”, one can maximise
¹⁶⁵² the correct ID probability and minimise at the same time the mis-ID probability. In
¹⁶⁵³ the kaon case we do not use requirements on the $K \rightarrow \pi$ mis-ID probability because
¹⁶⁵⁴ this cut was found to be unacceptably inefficient.

¹⁶⁵⁵ 12.4 Bremsstrahlung recovery for electrons

¹⁶⁵⁶ Bremsstrahlung is an electromagnetic radiation produced by particles, that deceler-
¹⁶⁵⁷ ate or deviate because of the presence other charged particles. Typically electrons
¹⁶⁵⁸ produce Bremsstrahlung when deflected by atomic nuclei. The probability of emit-
¹⁶⁵⁹ ting bremsstrahlung radiation is proportional to the inverse of the squared mass of
¹⁶⁶⁰ the particle ($1/m^2$) and therefore is relevant only for electrons. At these energies, if

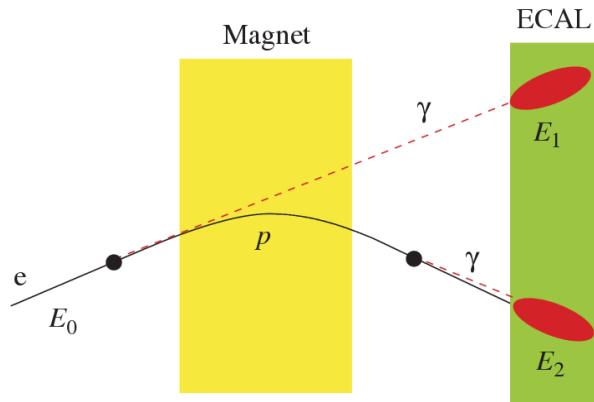


Figure 12.4: Schematic view of the Breamsstrahlung recovery.

1661 electrons radiate after the magnet, the photon will hit in the same calorimeter cells
1662 as the electron and the energy will be automatically recovered. However, if the pho-
1663 ton is emitted before the magnet, the electron will be deflected by the magnetic field
1664 whereas the photon will continue on its initial trajectory, with its energy being de-
1665 posited in a different part of the calorimeter. Missing this energy results in a poorer
1666 reconstructed B^0 mass resolution, so it is desirable to recover these bremsstrahlung
1667 photons, when possible. A tool for bremsstrahlung recovery is available in the LHCb
1668 analysis software. This tool looks for other clusters in the calorimeter and recon-
1669 structing the trajectory of the electron checks if they may have been emitted by that.
1670 Then the photon energy is added to the electron and its momentum recalculated.
1671 Figure 12.4 shows a schematic view of the process. For more information see Ref.
1672 [50].

1673 12.5 Peaking backgrounds

1674 Cuts are applied in order to remove background sources due to specific decays.
1675 These types of backgrounds usually peak in some variable because of their mass or
1676 distinctive kinematic properties and therefore they can be removed without signifi-
1677 cant signal efficiency loss. In the following sections are described the main sources
1678 of peaking background.

1679 12.5.1 Charmonium vetoes

1680 Charmonium resonances such as J/ψ and $\psi(2S)$ peak in q^2 . The choice of q^2 binning
 1681 described in Sec. 12.1 constitutes a natural veto for these decays. Simulated events
 1682 were used to check if resonant events leak inside the q^2 intervals chosen for the
 1683 rare channel analysis. For the muonic channels the leakage is negligible as the
 1684 peaks are sharper due to a better resolution and muons emit fewer bremsstrahlung
 1685 photons, resulting in shorter radiative tails. The electronic channels are instead
 1686 characterised by a worse resolution and at the same time electrons can radiate
 1687 several bremsstrahlung photons, yielding long tails at low q^2 . Analysing Monte
 1688 Carlo events it was found that 1.3 - 2% (depending on the trigger category) of
 1689 $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candidates leak into the $1.1 < q^2 < 6$ GeV^2/c^4 interval
 1690 and 1.8% of $\psi(2S)$ events leak above 15 GeV^2/c^4 . The contribution from these
 1691 candidates is modelled in the fit.

1692 12.5.2 Cascade $b \rightarrow c\ell, c \rightarrow \ell$ decays

1693 A source of mis-reconstructed background is due to cascade decays with a B^0 de-
 1694 caying semileptonically into a D meson which also decays semileptonically, e.g.
 1695 $B^0 \rightarrow D^-\ell^+\bar{\nu}_\ell$ followed by $D^- \rightarrow K^*\ell^-\nu_\ell$. The candidates built from these de-
 1696 cays tend to have a low 4-body invariant mass as two or more particles are not
 1697 reconstructed. This is in general true for any partially reconstructed background
 1698 from B decays.

1699 In order to remove this background in the muonic channels, the 4-body $m(K\pi\mu^+\mu^-)$
 1700 invariant mass is recalculated with a kinematical fit (in the resonant case this in-
 1701 cludes a constraint of the dilepton mass to be the J/ψ nominal mass). This con-
 1702 straint has the effect of pushing the misreconstructed events far from the B^0 peak.
 1703 Therefore, to avoid this background, it is sufficient to limit the analysis to 4-body
 1704 invariant masses above 5150 MeV/c^2 .

1705 In the electronic case it is instead important to fit a wider mass window to correctly
1706 constrain the background therefore one cannot eliminate this mis-reconstructed
1707 background which is then modelled in the fit (for details see Sec. 13.2.2).

1708 **12.5.3 ϕ veto**

1709 It can happen that a kaon from the decay $B_s \rightarrow \phi\ell^+\ell^-$, where the ϕ decays in two
1710 kaons, is mis-identified as a pion and therefore causes the ϕ to be reconstructed as
1711 a K^* . This results in a candidate with a value of $m_{K\pi}$ that is less than m_{K^*} but
1712 still high enough to enter into the mass selection. In Fig. 12.5.3 is reported the
1713 plot of $m_{K\pi}$ versus $m_{K\pi\ell\ell}$, where the mass of the kaon has been assigned to the
1714 pion. A peak can clearly be seen around the ϕ mass ($1020 \text{ MeV}/c^2$). To remove this
1715 background only candidates with $m_{K(\pi \rightarrow K)} > 1040$) are selected. This results in a
1716 98% signal rejection while keeping a 99% signal efficiency. This cut could be further
1717 optimised using PID information. On the other hand LHCb simulation struggles
1718 modelling the PID variables correctly. Therefore using PID in these cuts would add
1719 systematic uncertainties without significantly improving the signal efficiency which
1720 is already 99%. The ϕ could also constitute a background when it decays into two
1721 leptons but the branching ratio of this decay is small compared to the one into kaons
1722 and this contribution is taken into account by the choice of the q^2 intervals.

1723

1724 **12.5.4 $B^+ \rightarrow K^+\ell^+\ell^-$ plus a random pion**

1725 Some $B^+ \rightarrow K^+\ell^+\ell^-$ decays can contaminate the upper B^0 mass sideband if they are
1726 reconstructed with a soft pion from somewhere else in the event. It can also happen
1727 the same but with a kaon misidentified as a pion. In Fig. 12.5.3 the three-body
1728 invariant mass, $m_{K\mu\mu}$, distribution is shown. This is characterised by a narrow
1729 peak at the B^+ mass. Since these events have $m_{K\pi\ell\ell} > 5380 \text{ MeV}/c^2$ there is no

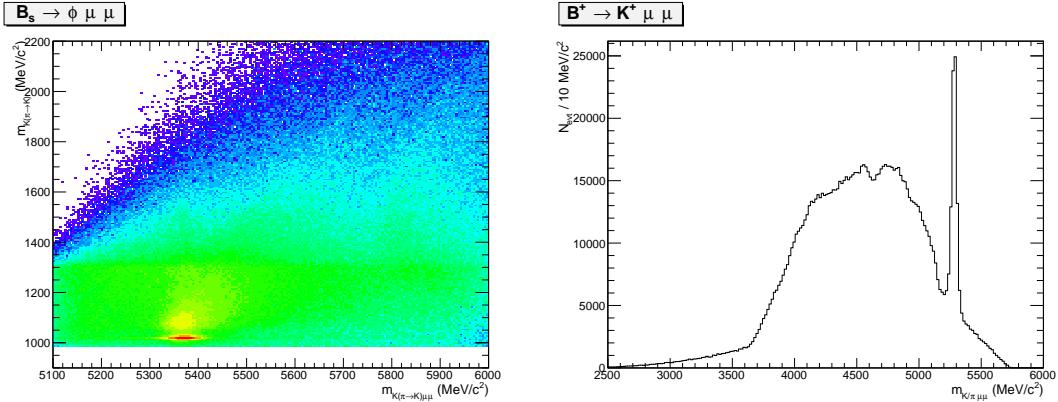


Figure 12.5: On the left the distribution of 2011 data events on the variables ($m_{K(\pi \rightarrow K)}$) and ($m_{K(\pi \rightarrow K)\mu\mu}$), where $\pi \rightarrow K$ means that the kaon mass is given to the pions too. On the right the mass of the three-body system ($m_{K\mu\mu}$) where the peak due to the $B^+ \rightarrow K^+ \mu\mu$ decay is visible.

1730 contribution under the B^0 peak, but they can cause problems when using sidebands
 1731 events to train the neural network. An effective veto for this decay was found to
 1732 be $\max(m_{K\ell\ell}, m_{K \rightarrow \pi\ell\ell}) < 5100 \text{ MeV}/c^2$, which results in 95% background rejection
 1733 while keeping 99% signal efficiency.

1734 12.5.5 Λ_b decays

1735 $\Lambda_b^0 \rightarrow \Lambda J/\psi$ decays are unlikely to be reconstructed as $B^0 \rightarrow K^* \ell^+ \ell^-$ because the
 1736 Λ is long-lived and decays further in the detector with a separate vertex. Simulated
 1737 events were used to check how many candidates fall into our B^0 samples and found
 1738 the amount negligible. The $\Lambda_b^0 \rightarrow J/\psi pK$ decay can instead contribute more easily
 1739 since the $m(pK)$ is above the Λ threshold and therefore they must come from Λ^*
 1740 resonances, which are not long-lived. This background is already reduced using
 1741 PID but a non-negligible contribution is still expected in the $\mu\mu$ sample, which is
 1742 modelled in the fit.

¹⁷⁴³ 12.5.5.1 Other peaking backgrounds

¹⁷⁴⁴ A possible background could come from $B^0 \rightarrow K^*\gamma$ decays where the photon con-
¹⁷⁴⁵ verts into two electrons while traversing the detector. In LHCb, around 40% of
¹⁷⁴⁶ photons convert before the calorimeter, but only a small fraction of these, $\sim 10\%$,
¹⁷⁴⁷ are reconstructed. Furthermore these events fall into a q^2 region well below the
¹⁷⁴⁸ intervals considered in these analysis and their contribution is therefore negligible.
¹⁷⁴⁹ Similar decays are also $B^0 \rightarrow K^*\eta$ and $B^0 \rightarrow K^*\pi^0$ where η and the pion decay
¹⁷⁵⁰ into two photons. Once again the contribution from these decays falls well below
¹⁷⁵¹ the considered q^2 intervals. Finally, a potentially dangerous background could come
¹⁷⁵² from events where the identity of the kaon and the pion are swapped as these candi-
¹⁷⁵³ dates peak under the signal. Their contribution is found to be small, 0.5%, however
¹⁷⁵⁴ the effect of their modelling into the fit is taken into account in the systematic
¹⁷⁵⁵ uncertainties.

¹⁷⁵⁶ 12.6 Multivariate analysis

¹⁷⁵⁷ The final selection is performed using a Neural Network classifier (NN) based on the
¹⁷⁵⁸ NeuroBayes package [33, 34]. The multivariate analysis is intended to remove some
¹⁷⁵⁹ combinatorial background and obtain a clearer signal peak.

¹⁷⁶⁰ Representative samples of the signal and background are needed to train the classi-
¹⁷⁶¹ fier. For the signal, fully reconstructed $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$ Monte
¹⁷⁶² Carlo events can be used. To the classifier is also passed the event weight calculated
¹⁷⁶³ to improve the data-simulation agreement (see Sec. 14.1). A sample representative
¹⁷⁶⁴ of the background can be obtained taking real data events in the upper B^0 side-
¹⁷⁶⁵ band ($m_{B^0} > 5400$ MeV/ c^2). The lower sideband it is not used in the training as
¹⁷⁶⁶ it contains a significant fraction of mis-reconstructed background. All pre-selection
¹⁷⁶⁷ cuts are applied to the background samples used for the training. As L0 and PID
¹⁷⁶⁸ variables are not well described these cuts are not applied in the Monte Carlo sam-

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

Table 12.3: Variables used as inputs for the NN training. Next to each variable the ID number in brackets provides the index reported in the correlation matrices shown in Fig. 12.6.

1769 ples but their effect is taken into account by the event weight. For the training 50%
 1770 of the sideband events was used, keeping the other 50% for testing. For the signal
 1771 sample a number of Monte Carlo events was used equal to the number available for
 1772 the background sample. This corresponds to ~ 1600 events for the $\mu^+\mu^-$ training
 1773 and ~ 500 for e^+e^- .

1774 The input to the NN consists of 22 variables containing information about the
 1775 kinematic of the decays and the quality of tracks and vertices. All the variables used
 1776 are listed in Tab. 12.3. The graphical representation of the correlation matrices are
 1777 shown in Fig. 12.6, in these figures the variable with ID = 1 is the NN output and
 1778 the other IDs are reported in Tab. 12.3. The single most discriminating variable
 1779 used is the χ^2 of a kinematical fit that constrains the decay product of the B^0 , the
 1780 K^* and the dimuon, to originate from their respective vertices. Other variables that
 1781 contribute significantly are the χ_{IP}^2 of J/ψ and K^* , the transverse momentum of
 1782 the B^0 and the pointing direction (DDIRA) of the reconstructed B^0 to the primary
 1783 vertex. The list the 10 most important variables is reported in Tab. 12.4, together
 1784 with information on the relative importance of each input. The meaning of the
 1785 column headings in this table was already explained in Sec. 4.1.

1786 Figure 12.7 shows distributions of neural network output for signal and background.
 1787 On this plot distributions from test samples are also overlaid in order to check for
 1788 overtraining. The distributions follow the same slope but with different fluctuations
 1789 so we conclude that we have no significant overtraining. In general we conclude that

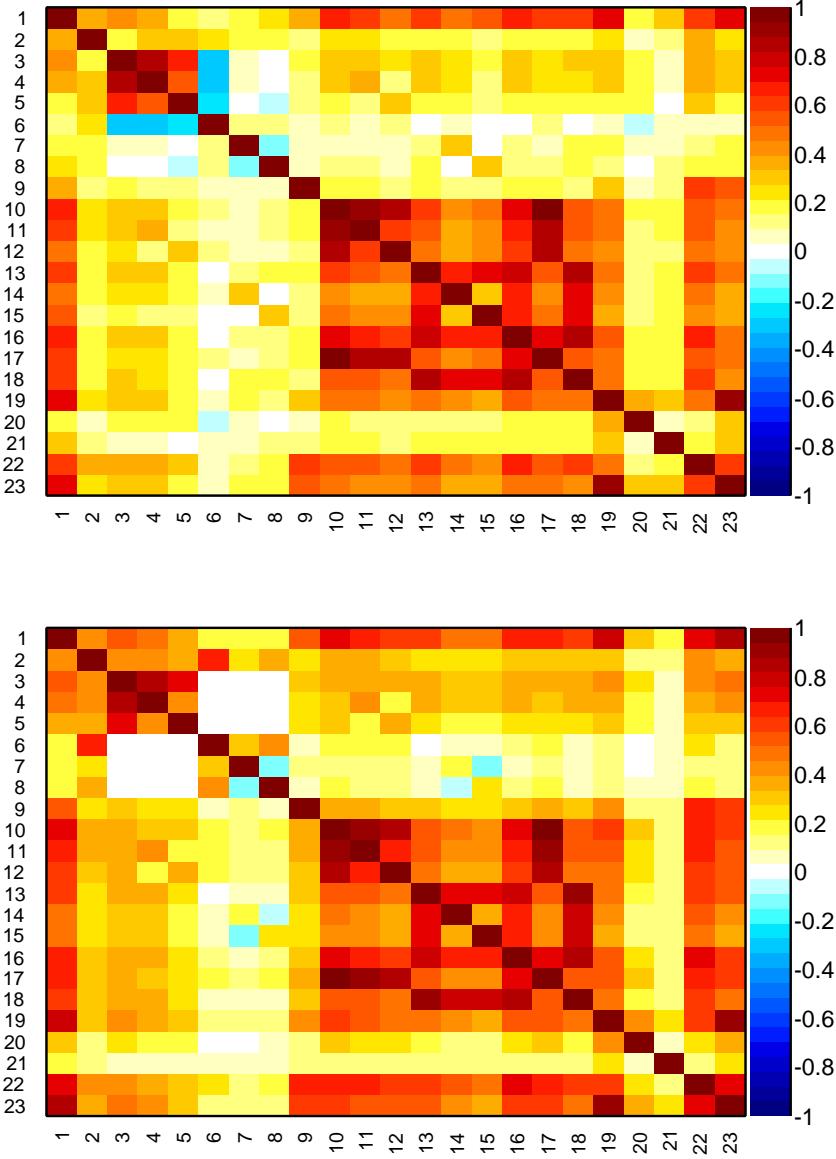


Figure 12.6: Graphical representation of correlation matrix between truth and neural network inputs. Column/row number 1 is correlation to the truth (whether candidate is signal or background). All others give correlation between inputs with numbering scheme corresponding to the id column of table ???. Correlation is calculated using all events without distinguishing signal and background.

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

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

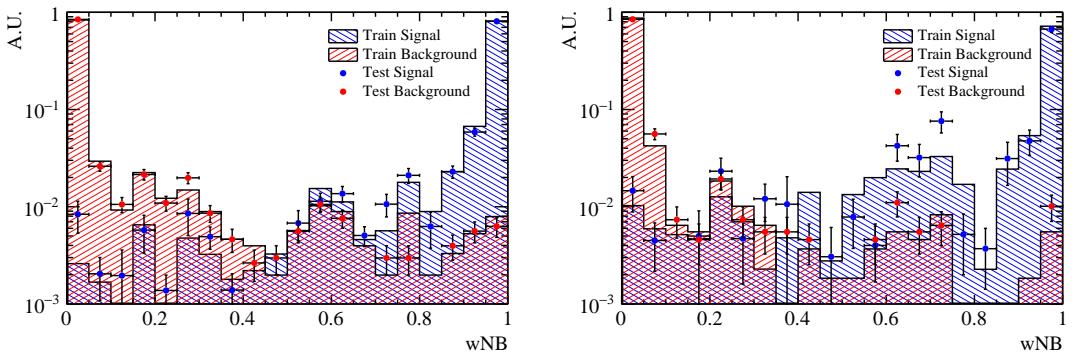


Figure 12.7: NN output distributions for training (solid) and test (stripes) samples, for MC signal and data sideband events. For the muon training (left) and the electron one (right).

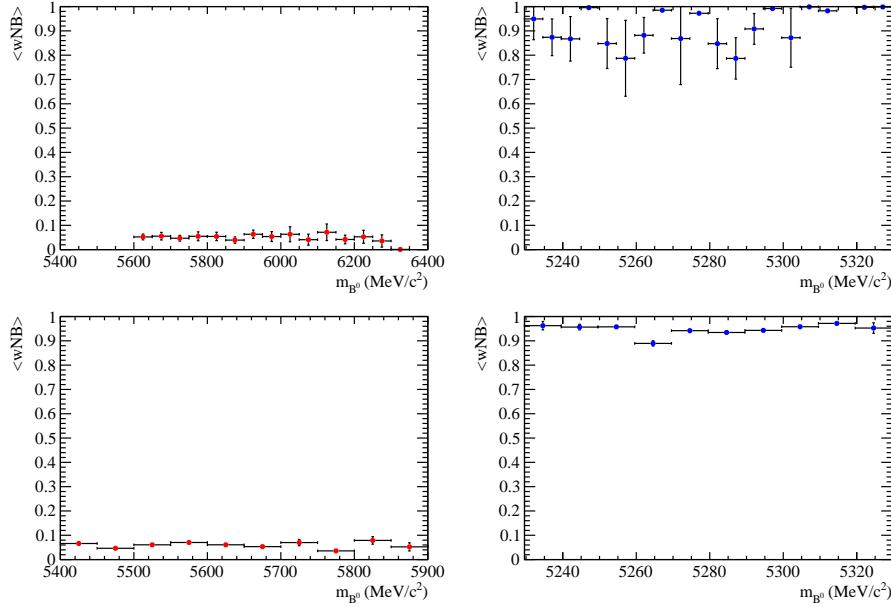


Figure 12.8: Average value of NN output as a function of B^0 mass for data sideband (left) and MC signal (right) events for the electron (top) and muon (bottom) training.

1790 the neural network is able to separate signal from background and that the training
1791 converged properly.

1792 It can happen that too much information is given to the classifier which becomes
1793 able to calculate the invariant mass of the candidates generating fake peaks. It
1794 is therefore important to check for correlations between the B^0 mass and the NN
1795 output. Fig 12.8 reports plots of the average NN output as a function of the B^0
1796 mass on sideband data and simulated signal events. The distributions do not show
1797 potential for peaks, so we conclude that no significant correlation is present.

1798 12.7 MVA optimization

1799 In order to optimise the cut on our neural network output the expected signal
1800 significance, $N_S/\sqrt{N_S + N_B}$, was maximised. In this formula N_S is number of rare
1801 signal events and N_B the number of background events.

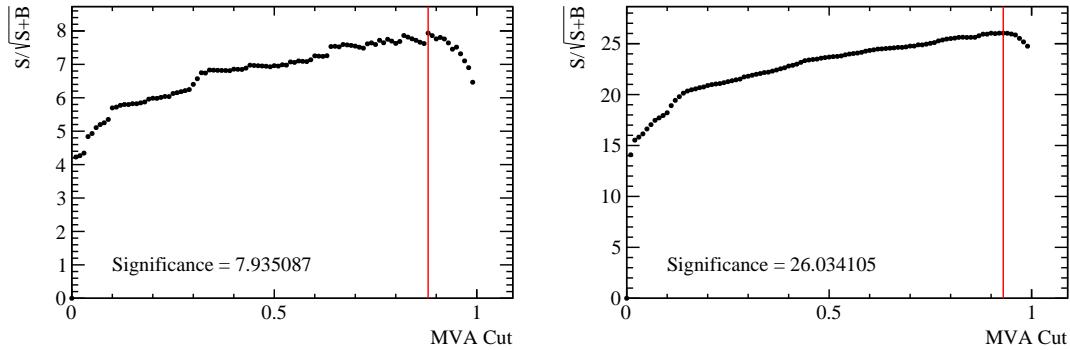


Figure 12.9: Dependence of figure-of-merit on the requirement on neural network output for electrons (left) and muons (right). The vertical lines corresponds to the chosen cuts.

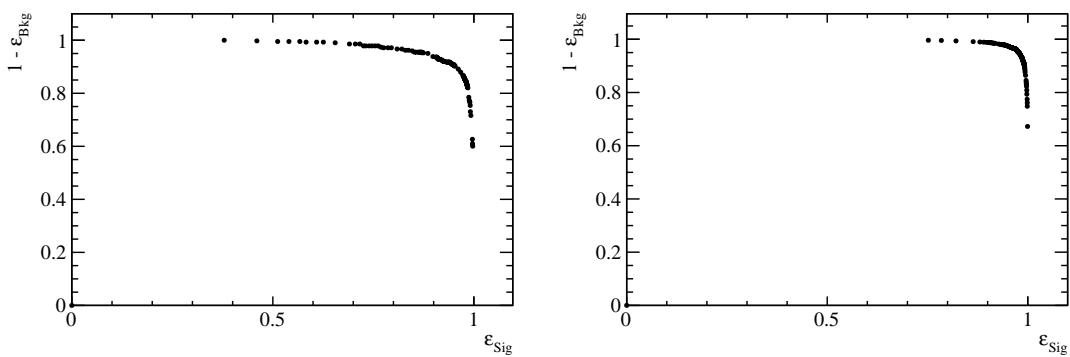


Figure 12.10: Receiver operating characteristic (ROC) curves for electrons (left) and muons (right). They show the signal efficiency versus the background rejection.

1802 The number of signal events accepted after a certain MVA output cut is derived from
1803 simulation. However, as an arbitrary number of events can be simulated, this has
1804 to be rescaled to the expected yield. This is done by fitting $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+ \ell^-)$
1805 events straight after pre-selection and scaling it down by the expected ratio between
1806 the rare and the resonant channels. The number of background events is instead
1807 derived by fitting the combinatorial background in the sideband with an exponential
1808 function and extrapolating the fit function below the signal peak.

1809 The dependence of the figure-of-merit for both the electron and muon trainings are
1810 shown in Fig.12.9, where the red line indicate the chosen cut: 0.75 for both samples.
1811 Curves of signal efficiency versus background rejection are shown in Fig. 12.10. Using
1812 the described MVA cuts the signal efficiency is $\sim 91\%$ for the muon channels and
1813 $\sim 84\%$ for the electron channels (for more details see Sec. 14), while the background
1814 rejections is $\sim 98\%$ on both samples.

1815 After full selection about $\sim 3\%$ of events still contain multiple candidates which are
1816 removed at random keeping only a single candidate per event.

CHAPTER 13

1817

1818

1819

Mass fits

1820

1821 In order to extract the yields of the rare and normalisation channels unbinned max-
1822 imum likelihood fits to the 4-body invariant masses $m(K\pi\ell\ell)$ are performed. The
1823 following sections contain a description of the line shapes used to model the signal
1824 and background components in each sample. These fits are performed simultane-
1825 ously on the resonant and rare channels. This method allows to share parameters
1826 between the two e.g. those describing data-simulation differences. The yields of the
1827 rare channels are parameterised as a function of the corresponding J/ψ yields as

$$N_{\ell\ell} = N_{J/\psi} \cdot \varepsilon^{ref} \cdot R_{\ell\ell}. \quad (13.1)$$

1828 In this formula ε^{ref} is the relative efficiency given in Tab. 14.4 and $R_{\ell\ell}$ corresponds
1829 to the efficiency corrected ratio of the raw rare and resonant yields:

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

1830 The two ratios R_{ee} and $R_{\mu\mu}$ are then used to build the R_{K^*} quantity, as described
1831 in Sec. 16.

1832 13.1 Mass fits: muonic channels

1833 For the rare and resonant $\mu^+\mu^-$ channels the fitted variable is the $m(K\pi\mu\mu)$ in-
1834 variant mass coming from a kinematic fit where all vertices are required to point
1835 to their mother particle. In the resonant case it is beneficial to also constrain the
1836 dimuon mass to the known J/ψ mass. The effect of the kinematical fit is to
1837 improve the mass resolution by roughly a factor of 2, which results in a more stable
1838 fit. Furthermore, misreconstructed events are pushed away from the B^0 peak, which
1839 allows to use a wider mass window to better constrain the combinatorial background
1840 slope. The mass spectrum is fitted in the range $5150 - 5800$ MeV/c² with the lower
1841 limit of the mass range chosen to exclude partially reconstructed background. As it
1842 is not needed to model misreconstructed backgrounds in the fit this also eliminates
1843 systematic uncertainties associated with the knowledge of its shape.

1844 The PDF chosen to describe the signal in both the $B^0 \rightarrow K^*\mu^+\mu^-$ and its relative
1845 J/ψ channel is a Double Crystal Ball function already described in 5.1.

1846 As a first step simulated distributions are fit using the signal model. The fitted MC
1847 distribution for the resonant channel is reported in Fig. 13.1.

1848 For the fit to real data the signal parameters are fixed to the ones found for the
1849 simulated samples. In order to account for possible data-simulation discrepancies
1850 a scale factor is multiplied to the widths and a shift is added to the masses. In
1851 summary the PDFs used for the signal fits on data are defined as

$$P(x; c, m_0) = f^* \cdot C(x; \alpha_1^*, n_1^*, c \cdot \sigma_1^*, m^* + m_0) \quad (13.3)$$

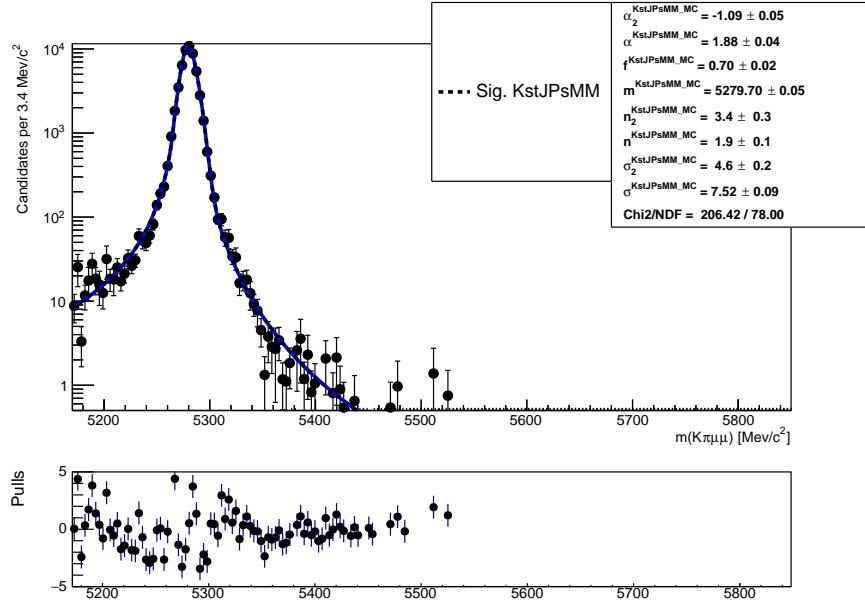
$$+ (1 - f^*) \cdot C(x; \alpha_2^*, n_2^*, c \cdot \sigma_2^*, m^* + m_0) \quad (13.4)$$

where the free parameters are the width scale factor, c , and the mass shift, m_0 , which are common between the rare and resonant samples. The other parameters, denoted with $*$, are taken from the fit to simulated events, separately for the rare and resonant samples and are fixed in the fit on data. The parameter f^* in the formula is the relative fraction of candidates falling in the first Crystal Ball function.

To model the combinatorial background an exponential function was used. This is the only background component for the rare channel. In the normalisation channel fit the $B_s^0 \rightarrow K^* J/\psi$ background is described using the same PDF used for the signal but a different central value, m , which is set at the B_s^0 nominal mass [31]. Finally, a $\Lambda_b^0 \rightarrow J/\psi pK$ background component is modelled using simulated $\Lambda_b^0 \rightarrow J/\psi pK$ events to which the full $B^0 \rightarrow K^* J/\psi$ selection is applied. The invariant mass distribution of these candidates is a broad flat shape under the signal peak. The simulated distribution is smoothed using a kernel estimation method (using the RooKeysPdf class of the RooFit package [51]).

In summary the floating variables in the simultaneous fit to rare and resonant $\mu^+ \mu^-$ samples are: the signal and background yields, the combinatorial background slopes, the widths scale c and the the mass shift m_0 .

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

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

1871 13.2 Mass fits: electronic channels

1872 In the electronic case the variable we fit is the $m(K\pi e^+e^-)$ invariant mass com-
 1873 ing from the kinematic fit where all vertices are required to point to their mother
 1874 particle. While in the muon case a further constraint was used for the resonant
 1875 fit, constraining the dilepton mass to the nominal J/ψ nominal mass, this is not
 1876 applied in the electronic case. In fact, due to the longer bremsstrahlung tail, the
 1877 J/ψ mass constraint distorts the invariant mass distribution and makes it is hard
 1878 to model it. Furthermore mis-reconstructed background enters in the rare channel
 1879 fit and can be constrained by looking at the higher statistics resonant channel, but
 1880 this implies the usage of the same variable in both fits. In order to better constrain
 1881 the parameters modelling the radiative tail and the misreconstructed backgrounds
 1882 a wide mass window is used $[4500, 5800]$ MeV/ c^2 . The lower limit is given by the
 1883 point in which the q^2 cut (at 6 GeV $^2/c^4$ to separate the rare and resonant channels)
 1884 starts to affect the 4-body invariant mass distribution.

1885 In the electronic case the invariant mass distributions are different depending on
 1886 which hardware trigger was used and especially how many bremsstrahlung photons

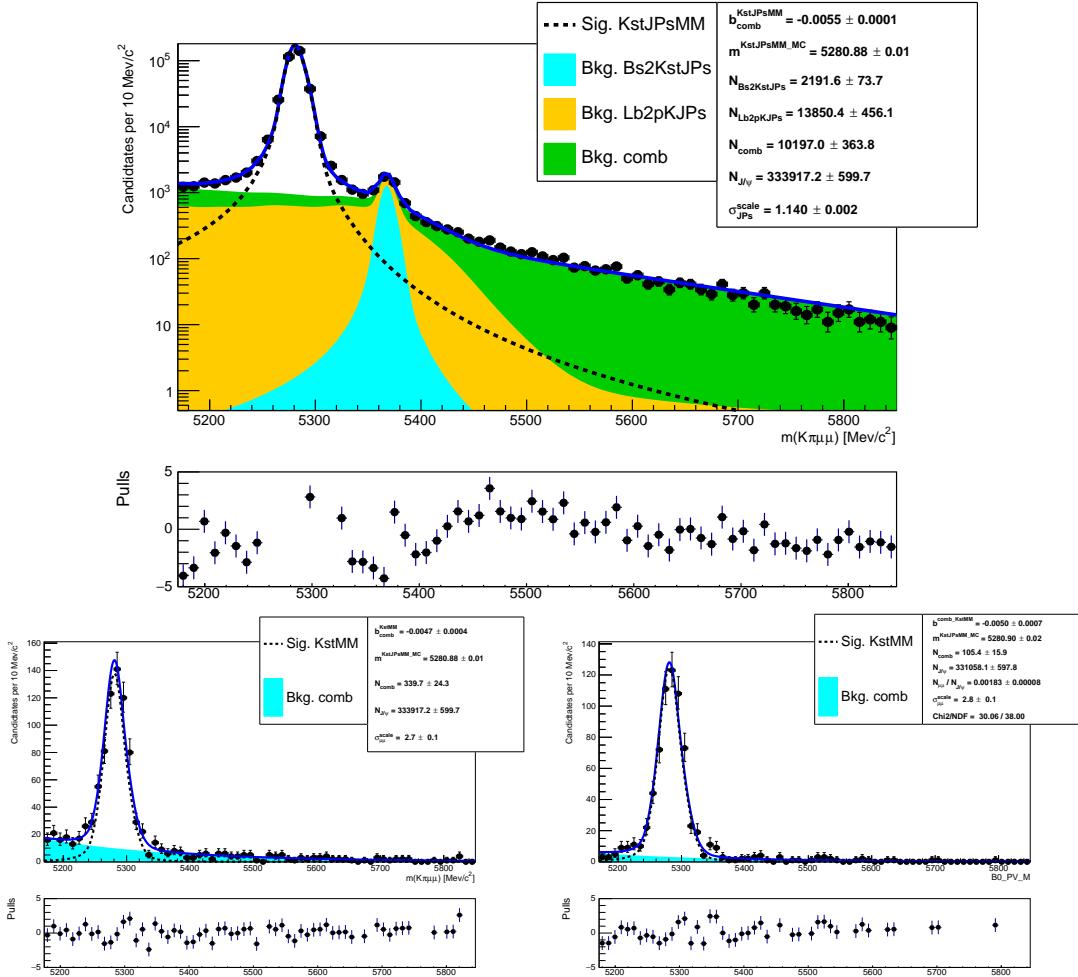


Figure 13.2: Fitted $m(K\pi\mu\mu)$ invariant mass spectrum for K^*J/ψ (left) and $K^*(J/\psi \rightarrow \mu^+\mu^-)$ (right). Dashed lines represent background components.

were reconstructed. Therefore our sample is divided in: 3 trigger categories and 3 bremsstrahlung categories. The three trigger categories are defined, to be exclusive, in the following way:

- Events triggered by an electron in the signal candidate:
1890 LOElectron_TOS
- Events triggered by L0Hadron in the signal candidate and not L0Electron:
1891 LOHadron_TOS and not LOElectron_TOS
- Events triggered by particles not in the signal candidate (Trigger Independent
1892 of Signal, TIS) and not by the previous cases:
1893

1896 L0_TIS and not (L0Electron_TOS or L0Hadron_TOS)

1897 The majority of the selected events falls in the L0Electron category, triggered by
1898 the electron. The L0Hadron category is more efficient at low q^2 were the K^* has
1899 more momentum.

1900 The three breamstrhalung categories are:

- 1901 • 0γ : events with no photon emitted
- 1902 • 1γ : events with one photon by either of the electrons
- 1903 • 2γ : events with one photon emitted by each electron

1904 The three samples, divided by trigger, are fitted simultaneously. This allows a better
1905 use of statistics as the simultaneous fit gathers information from the three categories
1906 at the same time and is more stable. Furthermore using this method the results for
1907 the three categories are naturally combined in a single R_{ee} ratio.

1908 In the next sections the PDFs used to fit the invariant mass distributions in the
1909 central and high q^2 intervals are described.

1910 13.2.1 Signal PDFs for the electronic channels in the central q^2 interval

1911 As for the muonic channel simulated events are fitted first to constrain the shapes
1912 for the subsequent fit on data. The signal PDFs are built using the following method:

- 1913 • Simulated $B^0 \rightarrow K^* J/\psi (ee)$ and $B^0 \rightarrow K^* ee$ events divided in each trigger and
1914 bremsstrahlung category and an independent fit is performed to each sample.
- 1915 • For each trigger category a PDF is built as the sum of the three PDFs for each
1916 bremsstrahlung category.

$$P(x)^{\text{trg}} = f_{0\gamma}^{\text{trg}} P(x)_{0\gamma}^{\text{trg}} + f_{1\gamma}^{\text{trg}} P(x)_{1\gamma}^{\text{trg}} + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) P(x)_{2\gamma}^{\text{trg}}. \quad (13.5)$$

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

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

1917 where the $P(x)_{n\gamma}^{trg}$ functions are the chosen PDFs for each trigger and bremsstrahlung
 1918 category and the $f_{n\gamma}^{trg}$ parameters are the relative fractions of events falling in
 1919 each category.

- 1920 • Most parameters are fixed (details later) and this joint PDFs are used to fit
 1921 real data divided only in trigger categories.

1922 The 0γ caterogy is catacterised by a better resoluton and a sharp tail on the right-
 1923 hand side and it is fitted with a simple Crystal Ball function (CB). While the 1γ and
 1924 2γ samples are modelled using the sum of a Crystal Ball and a Gaussian functions
 1925 (CBG) with all parameters independent. When the joint PDF, $P(x)^{trg}$, is build we
 1926 all parameters are fixed leaving one global mass shift and one scale factor for the
 1927 widths to float, as done for the muonic samples.

1928 Finally, when constructing the sum of the three breamsstrahlung components one
 1929 needs to specify in which fractions they contribute to the total. These frانctions have
 1930 been shown to be in good agreement between data and Monte Carlo and therefore
 1931 they are fixed to the values found on simulation, separately for the normalisation
 1932 channel and each q^2 bin. In Tab. 13.1 are reported percentages of events with 0, 1
 1933 and 2 emitted photons in the three trigger categories.

1934 In summary the signal PDF for the fit on data is defined as:

$$P(x; c, m_0)^{\text{trg}} = f_{0\gamma}^{\text{trg}} \text{CB}(x; c, m_0)_{0\gamma}^{\text{trg}} + f_{1\gamma}^{\text{trg}} \text{CBG}(x; c, m_0)_{1\gamma}^{\text{trg}} + (1 - f_{0\gamma}^{\text{trg}} - f_{1\gamma}^{\text{trg}}) \text{CBG}(x; c, m_0)_{2\gamma}^{\text{trg}} \quad (13.6)$$

1935 where the free parameters are: c , the scaling factor for the widths, and m_0 , the mass
1936 shift.

1937 **13.2.2 Background PDFs for the electronic channels in the central q^2**
1938 **interval**

1939 In the fit to the resonant sample three background components are modelled: com-
1940 binatorial background, and misreconstructed background coming from the hadronic
1941 and the leptonic systems. The combinatorial is described with an exponential func-
1942 tion.

1943 The misreconstructed background is split in two categories, that involving higher
1944 hadronic resonances, $B^0 \rightarrow (Y \rightarrow K\pi X)(J/\psi \rightarrow e^+e^-)$, and that coming from
1945 higher $c - \bar{c}$ resonances, $B^0 \rightarrow (K^* \rightarrow K\pi)(Y \rightarrow (J/\psi \rightarrow e^+e^-)X)$ where X is not
1946 reconstructed. The first component also includes decays from D chains described in
1947 Sec. 12.5.2. These backgrounds are modelled using inclusive $B^0 \rightarrow J/\psi X$ simulated
1948 samples to which the full selection is applied. The distributions for the hadronic
1949 (leptonic) background are defined selecting candidates where the K^* (dimuon) is
1950 not a direct daughter of the B^0 . The invariant mass distributions of these events,
1951 shown in Fig. 13.3, are smoothed using a kernel estimation method and their yields
1952 are left floating in the fit. Given the low statistics available, the same shape was is
1953 used for the three trigger categories.

1954 In the fit for the rare sample in the central q^2 interval the modelled backgrounds are:
1955 combinatorial background, again modelled with an exponential; misreconstructed
1956 background coming from the hadronic system and the leakage of the J/ψ radiative
1957 tail into the lower q^2 interval. The shape for the misreconstructed component is

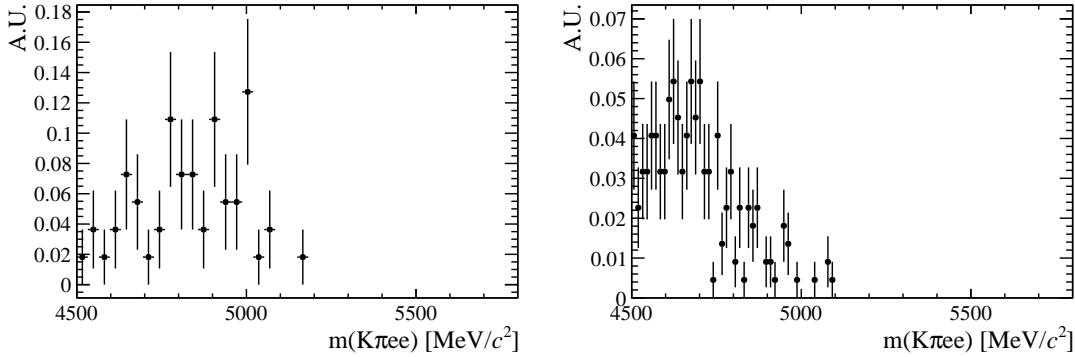


Figure 13.3: Simulated distributions of misreconstructed background events falling into the $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ sample coming from the hadronic (left) and leptonic (right) systems.

1958 obtained from simulated distributions similarly to what described for the resonant
 1959 channel. However, as there are no inclusive samples for the rare case, a sample
 1960 including higher K^* resonances is generated, including $K_1^+(1400)$ and $K_2^+(1460)$.
 1961 The yield of this component is not floating independently but its relative proportion
 1962 with respect to the signal yield is constrained to be the same as in the resonant
 1963 sample, namely:

$$N_{\ell\ell}^{mis-reco} = N_{ee} \cdot k = N_{ee} \cdot \frac{N_{J/\psi}^{mis-reco}}{N_{J/\psi}}. \quad (13.7)$$

1964 Notice that, as the fit is simultaneous for the rare and resonant samples, this fraction
 1965 is not fixed in the fit but floats using information from both samples.

1966 The shape to describe the J/ψ tail leakage is obtained using simulated $B^0 \rightarrow J/\psi K^*$
 1967 candidates and selecting those falling in q^2 below $6 \text{ GeV}^2/c^4$. The 4-body invariant
 1968 mass distribution of these events is reported in Fig. 13.4. The yield of this component
 1969 again is not floating independently but it is liked to the yield found in the resonant
 1970 fit as follows

$$N_{\ell\ell}^{leak} = N_{J/\psi} \cdot k^{MC} = N_{J/\psi} \cdot \frac{N_{leak}^{MC}}{N_{J/\psi}^{MC}} \quad (13.8)$$

1971 where k is the ratio between $N_{J/\psi}^{MC}$, the number of J/ψ events that fall into the
 1972 J/ψ q^2 window ($6-11 \text{ GeV}^2/c^4$) in the simulation and N_{leak}^{MC} , the number of J/ψ events
 1973 leaking below $6 \text{ GeV}^2/c^4$ in the simulation. In this case k is previously extracted
 1974 from simulated events and fixed in the fit on data.

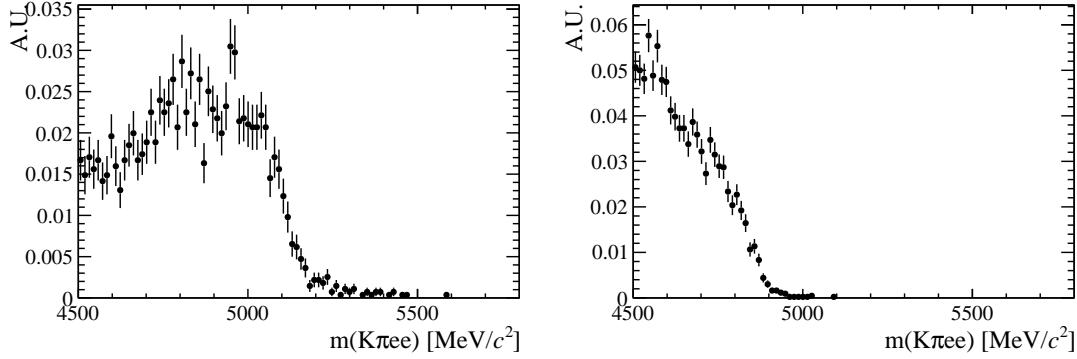


Figure 13.4: (left) Simulated 4-body invariant mass distributions for events involving higher K^* states and passing out full selection. (right) Simulated invariant mass distribution of $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ events leaking into the central q^2 interval.

¹⁹⁷⁵ 13.2.3 Summary of the fit to electronic channels in the central q^2 interval

¹⁹⁷⁶ In summary in the resonant fit on data the floating parameters are the yields of
¹⁹⁷⁷ all the components in the resonant channel, a common R_{ee} ratio, the combinatorial
¹⁹⁷⁸ background yield in the rare sample, one scale factor c , one mass shift m_0 and the
¹⁹⁷⁹ combinatorial background slopes.

¹⁹⁸⁰ In Fig. 13.5 are reported fits on simulated $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ candidates for
¹⁹⁸¹ all trigger categories and no photons emitted, in Fig. 13.6 for one photon emitted
¹⁹⁸² and in Fig. 13.7 for two photons emitted. Finally, in Fig. 13.8 and 13.9 are reported
¹⁹⁸³ fits on real $B^0 \rightarrow K^*(J/\psi \rightarrow e^+e^-)$ and $B^0 \rightarrow K^*e^+e^-$ candidates ($1.1\text{--}6$ GeV^2/c^4
¹⁹⁸⁴ interval) in the three trigger categories. Values of fitted parameters are reported on
¹⁹⁸⁵ the plots.

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

¹⁹⁸⁷ In the high q^2 interval, above 15 GeV^2/c^4 , the efficiency for the L0Hadron trigger
¹⁹⁸⁸ becomes very low as the K^* has very low momentum. In this region only 9 candidates
¹⁹⁸⁹ are found spread in the interval $4500 < m(K\pi ee) < 6000$ MeV/c^2 . In the L0TIS
¹⁹⁹⁰ category, even if the yield is bigger a clear peak cannot be seen, therefore only
¹⁹⁹¹ L0Electron triggered events are used in this region.

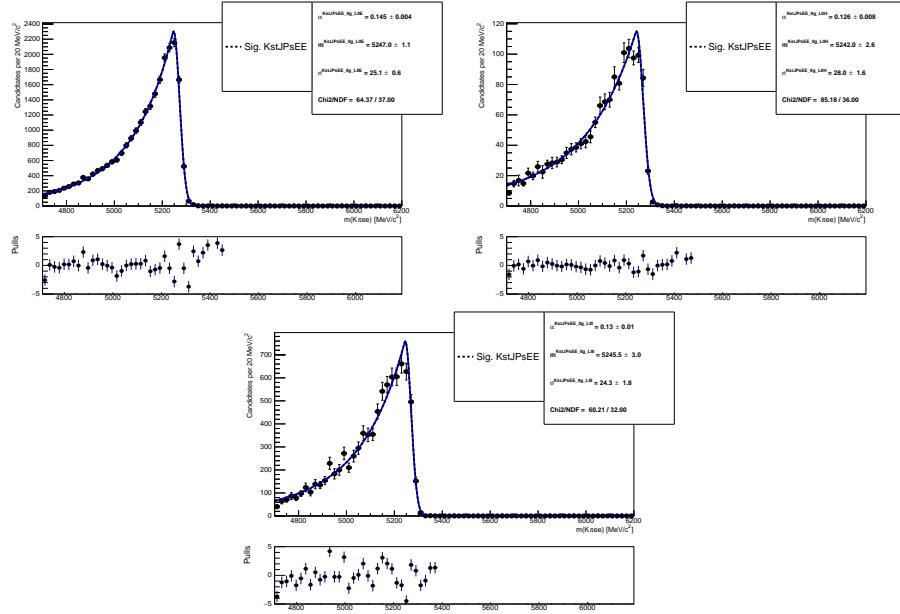


Figure 13.5: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and no photon emitted.

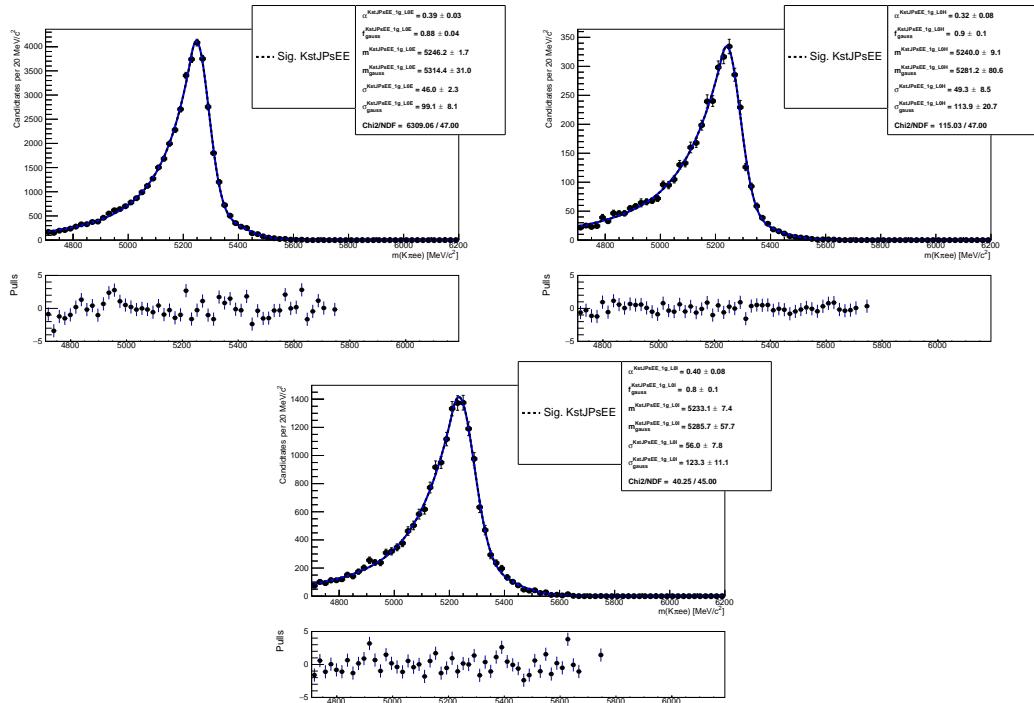


Figure 13.6: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and one photon emitted.

1992 The signal PDF is described in the same way as for the central bin. Simulated
1993 events are divided in three bremsstrahlung categories and fitted using the same

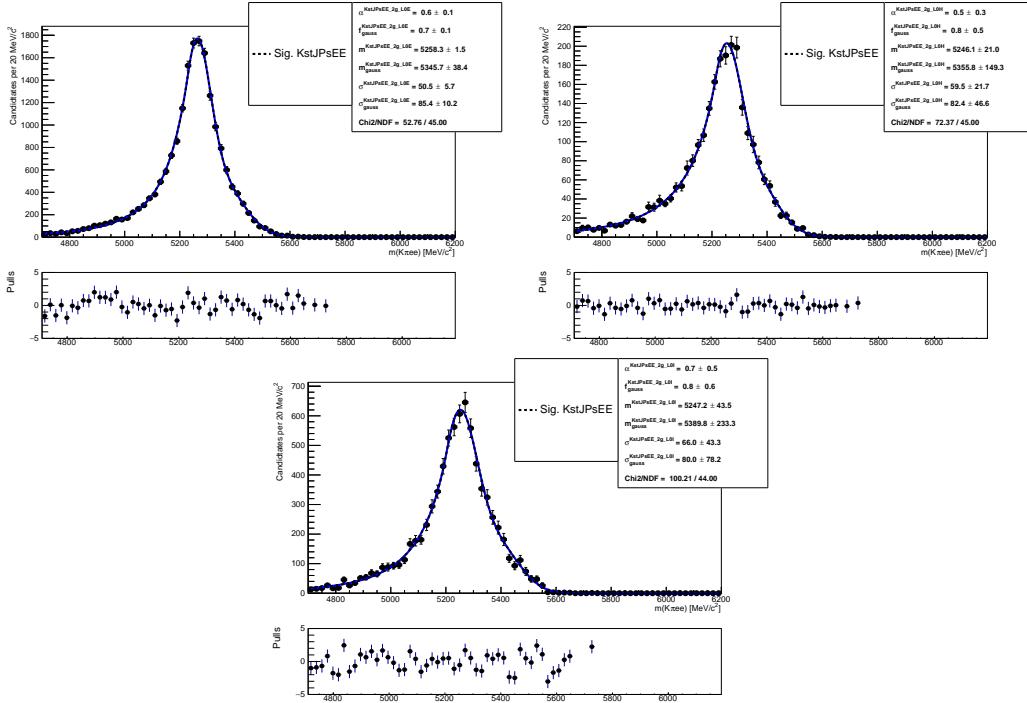


Figure 13.7: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ simulated events in the three trigger categories and two photons emitted.

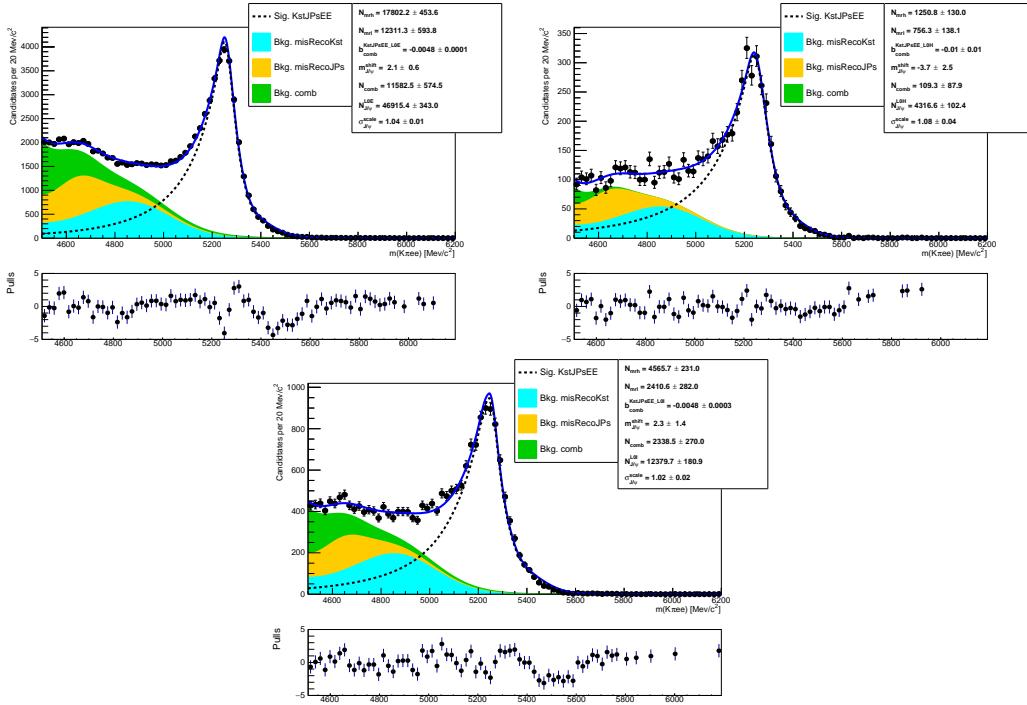


Figure 13.8: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} J/\psi (J/\psi \rightarrow ee)$ real data events in the three trigger categories.

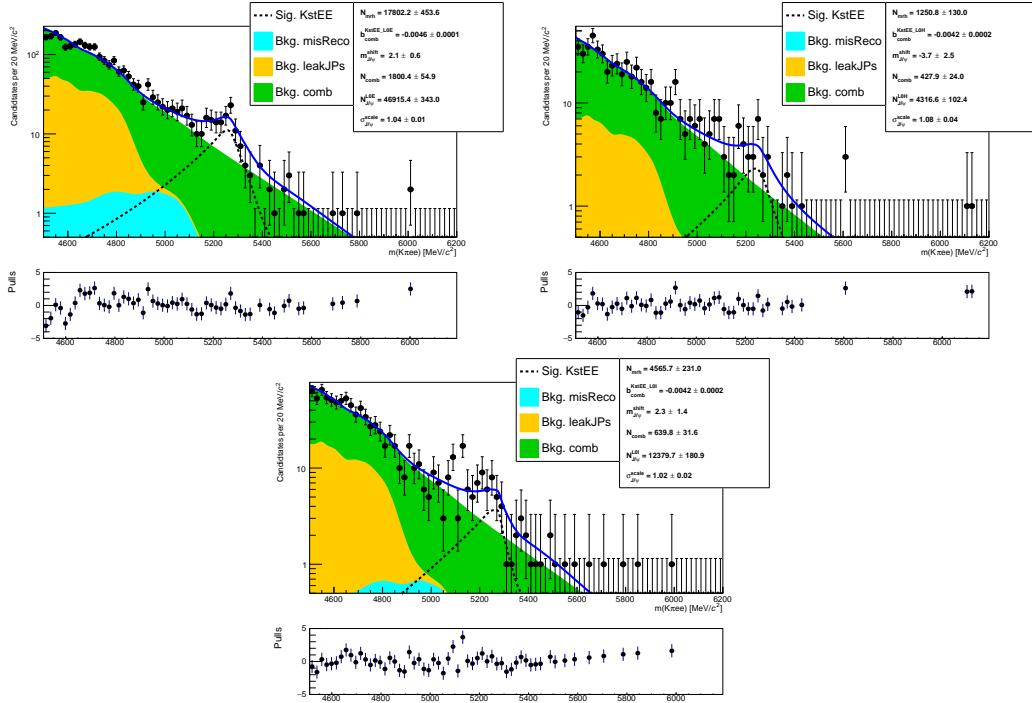


Figure 13.9: Fitted $m(K\pi ee)$ mass spectrum of $B^0 \rightarrow K^{*0} e^+ e^-$ real data events in the three trigger categories and no photon emitted.

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

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

1994 PDFs described in Sec. 13.2.1. While the signal tail parameters are similar for the
 1995 J/ψ and central q^2 samples in the case of the high q^2 interval it is particularly
 1996 important to keep them independent. In fact, as can be seen in Fig. 13.10, the
 1997 invariant mass distributions are significantly different for the two intervals. The
 1998 fractions of 0, 1 and 2 γ components used to build the total PDF are also in this
 1999 case taken from simulated events are reported in Tab. 13.2.

2000 The background components, as for the central q^2 interval, include a combinatorial
 2001 background and a misreconstructed background coming from the hadronic system.
 2002 Furthermore there is a leakage due to the $\psi(2S)$ resonance, that is wide enough to
 2003 contribute in q^2 above 15 GeV $^2/c^4$. The combinatorial background is modelled with
 2004 *comb model for high q2*.

2005 The misreconstructed component is modelled in the same way described for the
2006 central q^2 interval. However, in this case, its yield is not linked to the resonant fit as
2007 it is not guaranteed that the same fraction of misreconstructed background will be
2008 present in the J/ψ and high q^2 intervals. On the other hand the misreconstructed
2009 background shape is better defined at high q^2 and therefore its yield can be left
2010 floating in the fit.

2011 The $\psi(2S)$ leakage component is modelled from $B^0 \rightarrow K^*(\psi(2S) \rightarrow e^+e^-)$ simulated
2012 events with the same method used for the J/ψ leakage in the central q^2 interval. The
2013 yield of this component is fixed to the yield of $\psi(2S)$ as

$$N_{\ell\ell}^{leak} = N_{\psi(2S)} \cdot k^{MC} = N_{\psi(2S)} \cdot \frac{N_{leak}^{MC}}{N_{\psi(2S)}^{MC}}. \quad (13.9)$$

2014 In order to do this the $\psi(2S)$ yield, $N_{\psi(2S)}$, is obtained from a fit to the $\psi(2S)$
2015 invariant mass peak. Since we are only interested in the $\psi(2S)$ yield we fit the
2016 $m(K\pi ee)$ obtained from a kinematic fit where the dimuon mass is constrained to
2017 the known $\psi(2S)$ mass. This allows to eliminate the misreconstructed background
2018 from the fit mass window and use a simple model composed by a signal component
2019 and a combinatorial background component. The signal is described with a Double
2020 Crystal Ball function with opposite tails already described the Λ_b^0 fits (see Sec. 5.1),
2021 and the combinatorial background is described with an exponential. The fit to the
2022 $\psi(2S)$ peak is reported in Fig. 13.11 together with the fit to the $B^0 \rightarrow K^*e^+e^-$
2023 candidates in the high q^2 interval.

2024 13.3 Fit summary

2025 In Tab. 13.3 are reported raw yields obtained from the fits described in the previous
2026 sections. The values for the rare channels are not directly floating in the fits but as
2027 described in Sec. 13 they are parameterised as a function of the number of resonant
2028 events found and the ratios R_{ee} and $R_{\mu\mu}$ between the resonant and rare branching

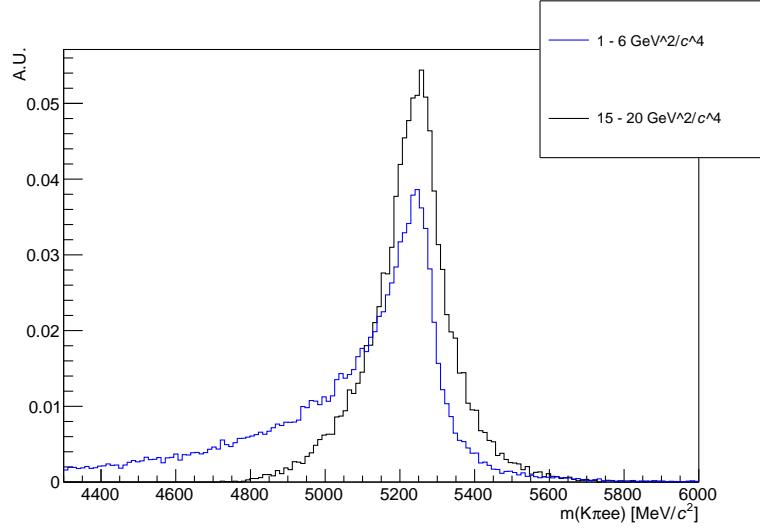


Figure 13.10: Simulated invariant mass of the $K\pi ee$ system in the $1 < q^2 < 6$ and $15 < q^2 < 20$ GeV^2/c^4 intervals.

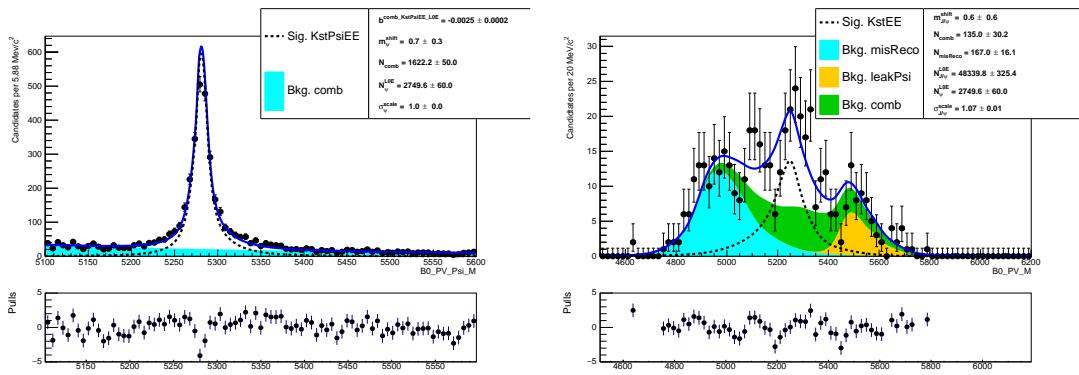


Figure 13.11: Fitted $m(K\pi ee)$ invariant mass distribution in the $\psi(2S)$ interval, $12 < q^2 < 25$ GeV^2/c^4 and in the high q^2 interval, $15 < q^2 < 20$ GeV^2/c^4

fractions. Measured values of these ratios are reported in Tab. 16.2.

2029

Sample	1–6 GeV^2/c^4	15–20 GeV^2/c^4	J/ψ
$\mu\mu$	625.38 ± 29.60	606.87 ± 27.56	333917.20 ± 599.73
ee L0Electron	131.77 ± 18.06	132.28 ± 27.92	48103.10 ± 329.77
ee L0Hadron	32.50 ± 4.50	—	4439.51 ± 98.38
ee L0TIS	48.53 ± 6.68	—	12683.18 ± 174.25

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

2030

CHAPTER 14

2031

2032

Efficiency

2033

2034 The efficiency for each of the decay channels is calculated according to the formula

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

2035 In this expression the first term is the efficiency to have final state particles in the
 2036 LHCb detector acceptance. The second term carries information on reconstruction
 2037 and stripping efficiency (we keep these together given that boundaries between them
 2038 are completely artificial). The third part corresponds to the efficiency of the PID
 2039 requirements. The fourth term handles the trigger efficiency for those events which
 2040 are selected by the preselection process. Finally, the latter term deals with the
 2041 efficiency of the NN classifier. Reconstruction, trigger and MVA efficiencies are
 2042 evaluated on simulated data with the trigger efficiency for $B^0 \rightarrow K^* J/\psi$ being
 2043 cross-checked using the data-driven TISTOS method as described in Sec. ???. The
 2044 PID efficiency is calculated with a data-driven method as described in Sec. 14.4.

Comp	$\mu\mu$	ee		
		L0Electron	L0Hadron	L0TIS
gen	0.1598 ± 0.0005		0.1589 ± 0.0005	
rec	0.0896 ± 0.0001		0.0583 ± 0.0001	
pid	0.8148 ± 0.0000		0.8222 ± 0.0000	
trg	0.7833 ± 0.0005	0.1831 ± 0.0004	0.0150 ± 0.0001	0.0565 ± 0.0002
mva	0.8958 ± 0.0004	0.8586 ± 0.0007	0.8974 ± 0.0006	0.8260 ± 0.0017
tot	0.0082 ± 0.0000	0.0012 ± 0.0000	0.0001 ± 0.0000	0.0004 ± 0.0000

Table 14.1: Absolute efficiencies for ee and $\mu^+\mu^-$ channels in the J/ψ q^2 interval.

Comp	L0Electron	L0Hadron	L0TIS	L0Electron	L0TIS
				$1-6 \text{ GeV}^2/c^4$	$15-20 \text{ GeV}^2/c^4$
q^2		0.70 ± 0.01		0.77 ± 0.01	
gen		1.02 ± 0.01		1.02 ± 0.01	
rec		0.91 ± 0.01		0.45 ± 0.00	
pid		0.98 ± 0.00		0.97 ± 0.00	
trg	0.89 ± 0.01	2.45 ± 0.05	1.24 ± 0.02	1.42 ± 0.01	0.71 ± 0.02
mva	0.97 ± 0.00	0.94 ± 0.00	0.97 ± 0.01	1.06 ± 0.01	0.96 ± 0.01
tot	1.12 ± 0.02	3.00 ± 0.08	1.57 ± 0.04	0.87 ± 0.02	0.39 ± 0.01

Table 14.2: Double ratios of efficiencies $(\varepsilon^{ee}/\varepsilon^{J/\psi \rightarrow ee})/(\varepsilon^{\mu^+\mu^-}/\varepsilon^{J/\psi \rightarrow \mu^+\mu^-})$ in the $1 < q^2 < 6$ and $q^2 > 15 \text{ GeV}^2/c^4$ intervals.

2045 All absolute efficiencies for the muonic and electronic rare channels are separately
 2046 listed in Tab. 14.3 for the central and high q^2 intervals and in Tab. 14.1 for the
 2047 resonant channels. However for the analysis itself only efficiencies relative to the
 2048 resonant channels are used in order to limit systematic uncertainties.

2049 In Tab. 14.3 are reported relative efficiencies between the rare and resonant chan-
 2050 nels, $\varepsilon(B^0 \rightarrow K^*\ell^+\ell^-)/\varepsilon(B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-))$. Finally, in Tab. 14.2 are
 2051 reported ratios of relative efficiencies for the ee and $\mu\mu$ channels, $(ee/(J/\psi \rightarrow$
 2052 $ee))/(\mu^+\mu^-/(J/\psi \rightarrow \mu^+\mu^-))$.

Comp	$\mu\mu$		ee			
	1–6 GeV/ c^4	15–20 GeV/ c^4	1–6 GeV/ c^4	L0Electron	L0Hadron	L0TIS
q^2	0.2142 ± 0.0015	0.1552 ± 0.0013	0.1493 ± 0.0012			0.1196 ± 0.0011
gen	0.1630 ± 0.0014	0.1630 ± 0.0014	0.1657 ± 0.0012			0.1657 ± 0.0012
rec	0.0170 ± 0.0001	0.0108 ± 0.0001	0.0100 ± 0.0000			0.0032 ± 0.0000
pid	0.7824 ± 0.0002	0.8420 ± 0.0001	0.7750 ± 0.0001			0.8239 ± 0.0001
trg	0.7044 ± 0.0029	0.8693 ± 0.0029	0.1465 ± 0.0011	0.0330 ± 0.0006	0.0629 ± 0.0007	0.2877 ± 0.0026
mva	0.9097 ± 0.0022	0.8298 ± 0.0032	0.8447 ± 0.0021	0.8571 ± 0.0020	0.8156 ± 0.0046	0.8436 ± 0.0033
tot	0.0065 ± 0.0001	0.0069 ± 0.0001	0.0011 ± 0.0000	0.0002 ± 0.0000	0.0004 ± 0.0000	0.0009 ± 0.0000

Table 14.3: Absolute efficiencies for ee and $\mu^+\mu^-$ channels in the $1 < q^2 < 6$ and $q^2 > 15$ GeV $^2/c^4$ intervals.

Comp	$\mu\mu$		ee		
	1–6 GeV/ c^4	L0Electron	1–6 GeV/ c^4	L0Hadron	L0TIS
gen	1.0200 ± 0.0091		1.0429 ± 0.0084		1.0200 ± 0.0091
rec	0.1896 ± 0.0012		0.1716 ± 0.0006		0.1201 ± 0.0009
pid	0.9602 ± 0.0002		0.9425 ± 0.0001		1.0334 ± 0.0001
trg	0.8993 ± 0.0038	0.8002 ± 0.0065	2.2025 ± 0.0434	1.1138 ± 0.0136	1.1098 ± 0.0037
mva	1.0154 ± 0.0025	0.9838 ± 0.0025	0.9551 ± 0.0023	0.9874 ± 0.0060	0.9262 ± 0.0036
tot	0.7918 ± 0.0110	0.8894 ± 0.0130	2.3764 ± 0.0550	1.2423 ± 0.0225	0.8382 ± 0.0131

Table 14.4: Relative efficiencies rare over resonant ($\varepsilon^{rel} = \varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$) for ee and $\mu^+\mu^-$ channels in the $1 < q^2 < 6$ and $q^2 > 15$ GeV $^2/c^4$ intervals.

2053 14.1 Data-simulation discrepancies

2054 Since most of the efficiency components are obtained from the study of simulated
2055 events it is important to verify that the simulation is a reliable reproduction of real-
2056 ity. In particular it is important to match data and Monte Carlo in the kinematics of
2057 the final particles and the occupancy of the detector. The kinematics of the decays
2058 is characterised by the transverse momentum spectrum of the B^0 . Discrepancies in
2059 this distribution cause also the spectra of the final particles to differ from reality
2060 and affect the efficiency estimation as its value often depends on the momentum
2061 distribution of final particles. The occupancy of the detector is correlated to the
2062 invariant mass shape of the signal because the addition of energy clusters in the
2063 electromagnetic calorimeter, affects the electron momenta for bremsstrahlung pho-
2064 tons emitted before the magnet. A way to quantify the detector occupancy is using
2065 the hits multiplicity in the SPD detector (see Sec. 2.2.2.1) distribution.

2066 Since it is important that these quantities are well simulated, the Monte Carlo is
2067 reweighted so that the distributions in data and simulation match for these variables.
2068 This can be done using resonant $B^0 \rightarrow K^* J/\psi$ events, for which the signal peak
2069 is already visible in data after pre-selection. However, the data includes also a
2070 high level of background and distributions cannot be directly compared. The sPlot
2071 technique [52] is used to statistically subtract the background from data and obtain
2072 pure signal distributions. This method is based on an estimation od the signal
2073 and background densities based on a fit to a control variable where the two are well
2074 distinct, usually the invariant mass. Fig. 14.1 shows fits to the 4-body invariant mass
2075 of candidates after preselection done in order to estimate the signal density. Data
2076 and simulation are then compared and the ratio between the distributions is used
2077 to reweight the Monte Carlo. The discrepancy in the SPD hits multiplicity is solved
2078 in the first place and then the B^0 transverse momentum are compared between
2079 data and simulation reweighted for the SPD multiplicities only. Distributions of B^0
2080 transverse momentum and SPD multiplicities are reported in Fig. 14.2 and ratios
2081 of these distribution, which are used to reweight the simulation, are reported in

2082 Fig. 14.3. Binnings for these distributions are chosen to have approximately the
 2083 same number of events in each bin to limit fluctuations.

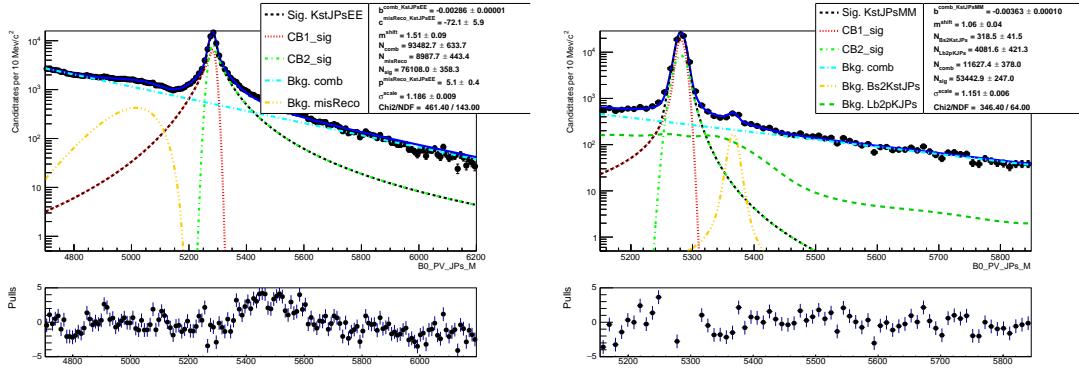


Figure 14.1: Fitted 4-body invariant mass distributions after pre-selection for the electron (left) and muon (right) channels.

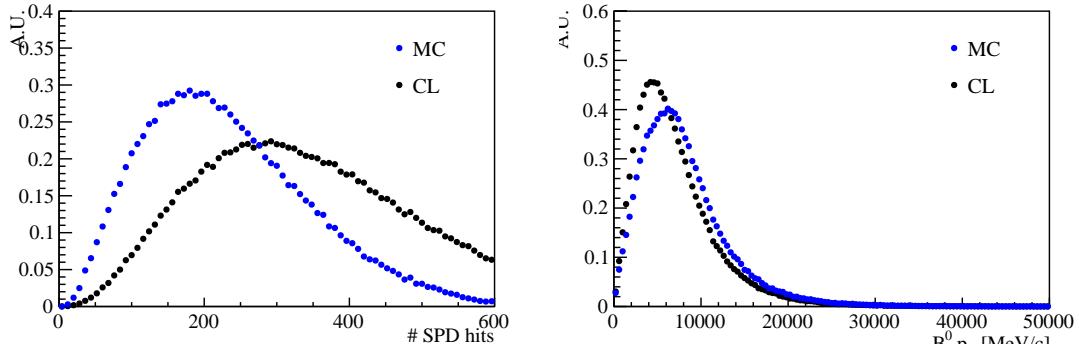


Figure 14.2: Distributions of number of SPD hits (left) and B^0 transverse momentum (right) in data and MC.

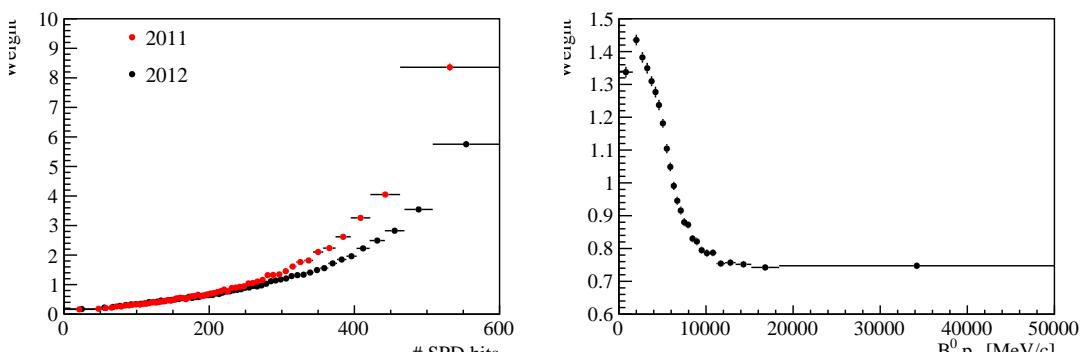


Figure 14.3: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

2084 14.2 Geometric efficiency

2085 The simulated samples used contain the requirement that daughters are in the LHCb
2086 detector acceptance. This corresponds to the requirement for each of the final par-
2087 ticles to have polar angle θ between 10 and 400 mrad. The efficiency of this cuts is
2088 obtained using a generator level Monte Carlo sample.

2089 14.3 Reconstruction efficiency and bin migration

2090 The reconstruction efficiency is here defined as the efficiency to reconstruct each
2091 decay channel given that its daughters are into the geometrical acceptance of the
2092 detector. This includes both the probability that a particle generates observable sig-
2093 natures and the efficiency of all the preselection cuts described in Sec. 12, including
2094 those done to remove peaking backgrounds. This efficiency is evaluated on simulated
2095 events to which the full list of weights described in Sec. ?? is applied. The efficiency
2096 of the PID cuts is kept separate as it is known to be not well simulated and there
2097 are reliable data-driven methods which can be used to extract it (see Sec. 14.4).

2098 One effect which must be considered is the “bin migration”, namely the possibility
2099 that events generated in a q^2 interval will be reconstructed in a different one. Two
2100 different effects can cause bin migration. First of all an imperfect resolution can
2101 cause events at the edges of the considered interval to fall on the wrong side of
2102 the edge. This effect is only important in case of non-flat true distributions, as
2103 the amount of bin migration in the two directions is different. The second possible
2104 source of bin migration are systematic effects due, for example, to the presence of
2105 bremsstrahlung photons that cannot be reconstructed. It is particularly important
2106 to take the bin migration into account in the electron channels case because the q^2
2107 resolution is worse and at the same time more photons are radiated from the final
2108 state. Figure 14.4 shows the correlation between reconstructed and generated q^2 in
2109 simulated $B^0 \rightarrow K^* ee$ events. In the ideal case of perfect resolution this plot would

look like a diagonal line and in the case of no bias its slope would be 1. Table 14.3 reports net bin migration amounts in the considered q^2 intervals. The reconstruction efficiency is calculated comparing generated to reconstructed samples and therefore already includes bin migration effects. Nevertheless, it is useful to single out this component in order to be able to assess systematic uncertainties.

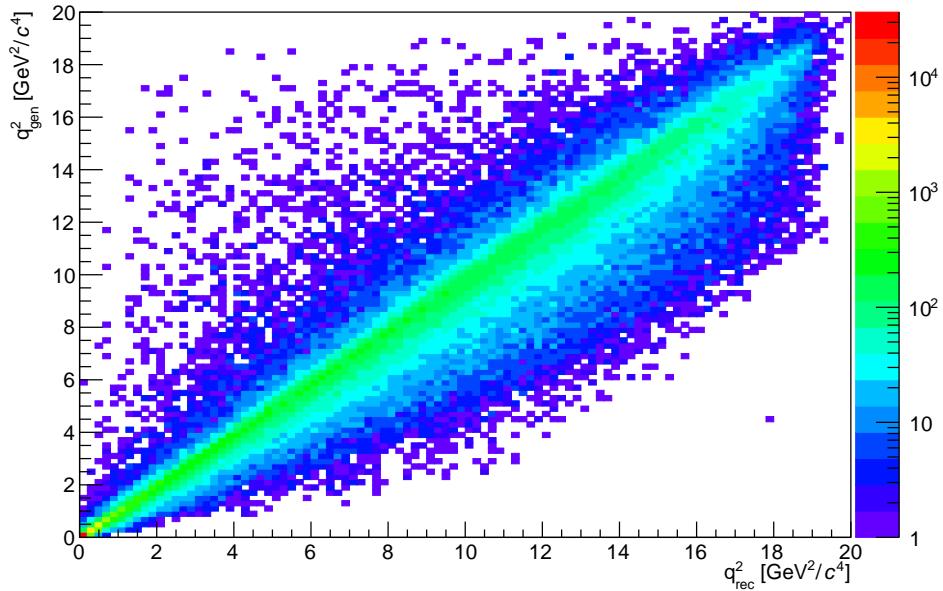


Figure 14.4: Generated versus reconstructed q^2 in simulated $B^0 \rightarrow K^* ee$ events.

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

Table 14.5: Net bin migration amounts in the considered q^2 intervals. Positive values mean “net in”, negative values “net out”.

14.4 PID efficiency

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

In order to do this the PIDCalib package [53] was used. This tool uses decays where final particles can be identified thanks to their kinematic properties. For example $K_s^0 \rightarrow \pi^+ \pi^-$ has a clear signature with a displaced vertex and can be easily singled out from other decays and used to test pion PID efficiency. The narrow peaks of the $J/\psi \rightarrow \mu^+ \mu^-$ and $J/\psi \rightarrow e^+ e^-$ decays allow to calibrate muon and electron efficiencies. Finally, $\phi \rightarrow KK$ and $D^* \rightarrow D(\rightarrow K\pi)\pi$ decays are used to test the kaon efficiency. Residual background in this decays is subtracted using the $s\mathcal{P}$ lot technique [52].

The package allows to divide the phase-space in bins and obtain a data-driven efficiency for each bin. For this analysis the phase-space was divided in equi-populated bins of momentum and pseudorapidity of the particle under study. Figure 14.5 shows performance tables for pions, kaons, muons and electrons.

The decay channel under study generally has different kinematical distributions than the calibration sample. Therefore, once the efficiency table is obtained for each particle, the total efficiency for each candidate is calculated 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, the total efficiency is found by averaging over all simulated events.

$$\varepsilon_{PID} = \frac{1}{N} \sum_i^N \varepsilon_K(p_K^i, \eta_K^i) \cdot \varepsilon_\pi(p_\pi^i, \eta_\pi^i) \cdot \varepsilon_\ell(p_{\ell_1}^i, \eta_{\ell_1}^i) \cdot \varepsilon_K(p_{\ell_2}^i, \eta_{\ell_2}^i) \quad (14.2)$$

14.5 Trigger efficiency

14.5.1 Muonic channels

For the muonic channels the trigger efficiency is calculated using simulated events. Using the resonant channel the efficiency obtained using the simulation was cross-checked with the data driven TISTOS method as already described in Sec. 5.3.7.

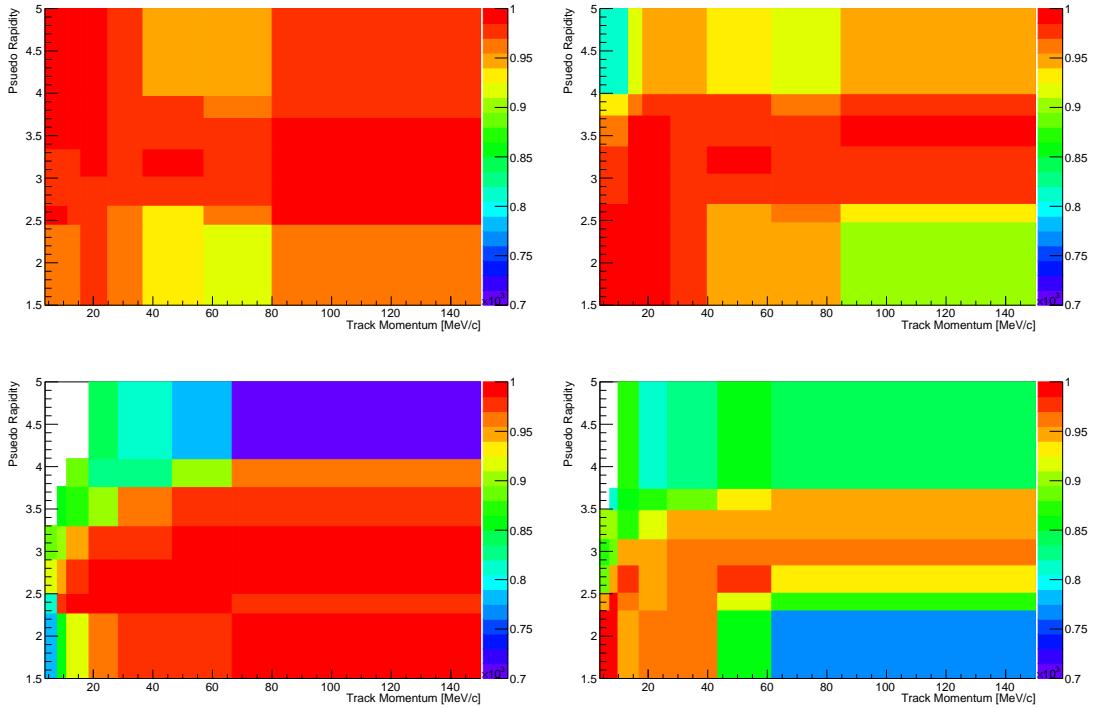


Figure 14.5: Performance tables obtained with data-driven methods for pions (top left), kaons (top right), muons (bottom left) and electrons (bottom right).

2142 *Results of TISTOS*

2143 14.5.2 Electronic channels

2144 For the electronic channels data is fitted separately in three trigger categories:
2145 L0Electron, L0Hadon and L0TIS. Therefore we need to extract the efficiency sepa-
2146 rately for each category.

2147 While the Hlt (1 and 2) efficiency is still computed using simulated events, the
2148 L0Electron and L0Hadron efficiencies cannot be modelled with the Monte Carlo.
2149 The discrepancy between data and simulation is mainly due to the ageing of the
2150 calorimeters, on which the decision of these triggers relies. The ageing is not simu-
2151 lated in the Monte Carlo and affects the L0 trigger efficiency which, therefore, must
2152 be calibrated using data driven-methods. Tables of efficiencies are obtained applying
2153 the TIS-TOS method to a calibration sample.

For each trigger category these tables contain 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. Regions considered are inner and outer HCAL, and inner, middle and outer ECAL. Figure 14.6 shows data-driven efficiencies for the L0Electron trigger in the three ECAL regions.

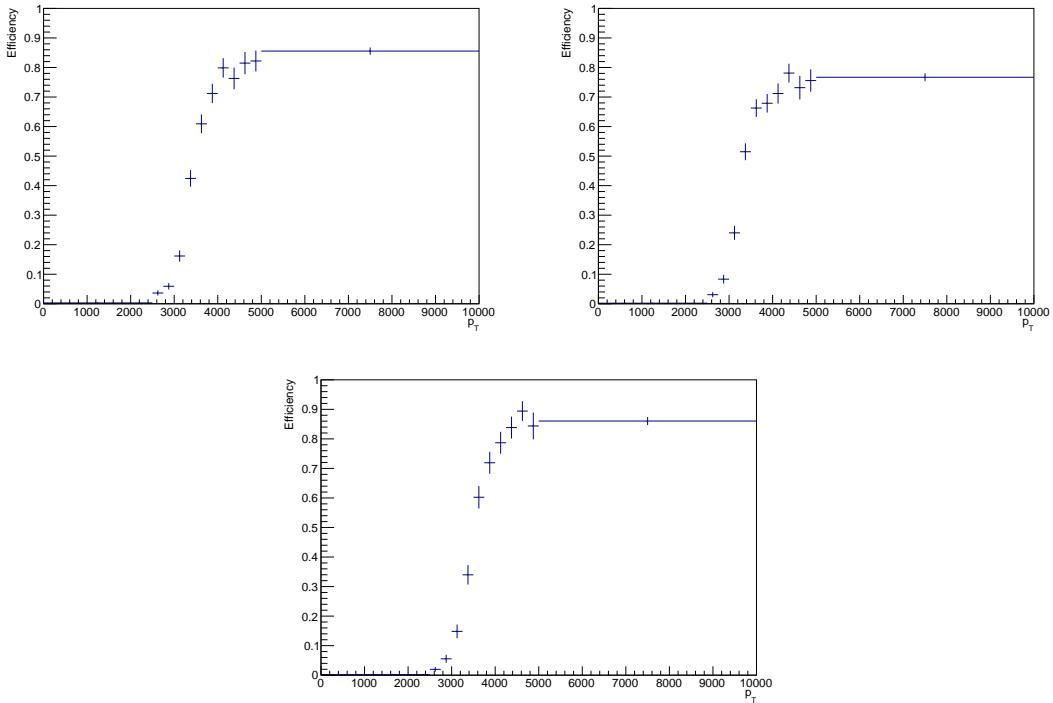


Figure 14.6: Data-driven L0Electron trigger efficiencies as a function of the transverse momentum of the electrons for the three ECAL regions.

The probability of L0Hadron trigger is calculated for each event as $P_{L0Had} = \varepsilon(\pi) + \varepsilon(K) - \varepsilon(\pi)\varepsilon(K)$. Similarly the L0Electron trigger probability is $P_{L0Ele} = \varepsilon(e^+) + \varepsilon(e^-) - \varepsilon(e^+)\varepsilon(e^-)$. Notice that no weight P_{L0TIS} can be defined using the TIS-TOS technique. On the other hand the probability of TIS trigger is defined to be independent of the signal and therefore must be the same in the rare and resonant channels and cancel in their ratio.

Then weights for the three trigger categories are then defined to be exclusive in the following way:

- 2168 • L0ElectronTOS: $\varepsilon^{L0E} = P_{L0Ele}$,
- 2169 • L0HadronTOS: $\varepsilon^{L0H} = P_{L0Had} \cdot (1 - P_{L0Ele})$, namely the probability that one
2170 hadron triggered but none of the electrons,
- 2171 • L0TIS: $\varepsilon^{L0I} = (1 - P_{L0Had}) \cdot (1 - P_{L0Ele})$, namely the probability that neither
2172 the hadrons or the electrons in the event triggered.

2173 As in the PID case, the total efficiency is found averaging over all events of a
2174 simulated sample:

$$\varepsilon^{cat} = \frac{1}{N} \sum_i^N \varepsilon^{cat}(p_T^i) \quad (14.3)$$

2175 where *cat* is a label indicating the trigger category under consideration.

2176 14.6 Neural Networks efficiency

2177 The NN efficiency is again evaluated from fully weighted Monte Carlo samples. For
2178 the electron channels it is obtained separately for each trigger category, because the
2179 yield is extracted independently for each of the three trigger categories and therefore
2180 these have to be independently corrected.

2181 In order to cross check that this efficiency component is extracted correctly one can
2182 compare the efficiency obtained using $B^0 \rightarrow J/\psi K^*$ events and rare $B^0 \rightarrow K^* \ell^+ \ell^-$
2183 events in the same q^2 region selected for the resonant case. The ratio between the
2184 two should be close to 1 with small deviations due the fact that the bin width is finite
2185 and the events are distributed differently inside the bin. These ratios are reported
2186 in Tab. 14.6; values are not exactly one for the electron channels due to the very
2187 large q^2 interval used to select the resonant channel ($[6.0, 11.0]$ GeV^2/c^4).

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

Table 14.6: Ratio $\varepsilon^{\ell\ell}/\varepsilon^{J/\psi}$ where the efficiency for the rare channel its calculated in the same q^2 rage used to select the resonant channel.

2188

CHAPTER 15

2189

2190

Systematic uncertainties

2191

CHAPTER 16

2192

2193

2194

Result extraction

2195

2196 In the following sections are presented the final results of this analysis together with
2197 the description of sanity checks performed to verify the stability of the methods
2198 used.

2199 16.1 $R_{J/\psi}$ sanity check

2200 In order to cross-check the analysis procedure, the ratio between the measured
2201 branching ratio of the electronic and muonic resonant channels is calculated:

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

2202 One could also look at the single absolute branching fractions but this ratio is a
2203 better defined sanity check for our analysis. In fact the absolute branching fractions

2204 can be calculated from the raw yields as

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

2205 where \mathcal{L} is the luminosity, $\sigma_{b\bar{b}}$ is the cross section for $b\bar{b}$ production and f_d is the
2206 fragmentation fraction, the probability for a b quark to produce a B^0 meson. All
2207 this quantities come with large uncertainties but they cancel in the $R_{J/\psi}$ ratio.

2208 Measured values of the $R_{J/\psi}$ ratio are reported in Tab. 16.1, where the error shown is
2209 statistical only. Notice that on these values systematic uncertainties, which cancel
2210 doing the ratio between the rare and resonant channels with same leptonic final
2211 state, do not cancel.

$R_{J/\psi}$ (L0E)	1.01618 ± 0.00896
$R_{J/\psi}$ (L0H)	0.94276 ± 0.02315
$R_{J/\psi}$ (L0I)	1.14471 ± 0.01758

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

2212 16.2 R_{K^*} result summary

2213 The ratio R_{K^*} is extracted by dividing the R_{ee} and $R_{\mu\mu}$ parameters described in
2214 Sec. 13. These ratios are floating in the fit but they can also be built from the yields
2215 in Tab. 13.3 and the efficiencies in Tab. 14.2. In summary the definition of the R_{K^*}
2216 ratio is the following :

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

2217 where $B^0 \rightarrow K^*$ has been neglected for brevity.

2218 As the electron ratio R_{ee} is a shared parameter in the simultaneous fit to the three

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

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

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

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

2219 electron categories its value is already a combination of the three samples. Results
 2220 are shown in Tab. 16.2.

2221 16.3 Branching ratios and expectations

2222 Multiplying the ratios R_{ee} and $R_{\mu\mu}$ by the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ [31]
 2223 branching ratios one can obtain absolute branching ratios for the rare channels:

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

2224 Table 16.3 reports absolute branching ratio values for the rare channels in the con-
 2225 sidered q^2 intervals, where the errors are statistical only.

2226 The results for the central q^2 bin can be compared also with SM predictions obtained
 2227 from Ref. [54]. This paper reports branching predicted ratios in the $1 < q^2 < 6$
 2228 GeV^2/c^4 interval for the rare electronic and muonic channels. These are rescaled
 2229 to the range $1.1 < q^2 < 6$ GeV^2/c^4 using the simulation. Finally, the measured
 2230 value of the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+\ell^-)$ decay is used. The predicted ratio
 2231 is found to be 0.75 ± 0.14 , which is in agreement with our measurement within one

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

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

standard deviation. In Tab. 16.4 are also reported observed and expected ratios of rare over resonant raw numbers of candidates ($N_{\ell\ell}/N_{J/\psi}$). In this table the observed ratios are simply obtained dividing the rare and resonant yields in Tab. 13.3 and the expected ones are obtained using the predicted rare channel branching ratios and the measured $B^0 \rightarrow K^*(J/\psi \rightarrow \ell^+ \ell^-)$ branching ratio, rescaled by the relative efficiencies in Tab. 14.4.

2238

2239

REFERENCES

2240

- 2241 [1] J. Lees *et al.*, “Measurement of Branching Fractions and Rate Asymmetries in
2242 the Rare Decays $B \rightarrow K^{(*)}l^+l^-$,” *Phys.Rev.*, vol. D86, p. 032012, 2012.
- 2243 [2] J.-T. Wei *et al.*, “Measurement of the Differential Branching Fraction and
2244 Forward-Backword Asymmetry for $B \rightarrow K^{(*)}l^+l^-$,” *Phys.Rev.Lett.*, vol. 103,
2245 p. 171801, 2009.
- 2246 [3] A. A. Alves Jr. *et al.*, “The LHCb detector at the LHC,” *JINST*, vol. 3,
2247 p. S08005, 2008.
- 2248 [4] T. Blake, S. Coquereau, M. Chrzaszcz, S. Cunliffe, C. Parkinson, K. Petridis,
2249 and M. Tresch, “The $B_0 \rightarrow K_0^*\mu\mu$ selection using 3fb^{-1} of LHCb data,” Tech.
2250 Rep. LHCb-INT-2013-058. CERN-LHCb-INT-2013-058, CERN, Geneva, Nov
2251 2013.
- 2252 [5] J. Beringer *et al.*, “Review of particle physics,” *Phys. Rev.*, vol. D86, p. 010001,
2253 2012. and 2013 partial update for the 2014 edition.
- 2254 [6] M. Adinolfi *et al.*, “Performance of the LHCb RICH detector at the LHC,” *Eur.
2255 Phys. J.*, vol. C73, p. 2431, 2013.
- 2256 [7] A. A. Alves Jr. *et al.*, “Performance of the LHCb muon system,” *JINST*, vol. 8,
2257 p. P02022, 2013.
- 2258 [8] R. Aaij *et al.*, “The LHCb trigger and its performance in 2011,” *JINST*, vol. 8,
2259 p. P04022, 2013.
- 2260 [9] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual,”
2261 *JHEP*, vol. 05, p. 026, 2006.

- [10] I. Belyaev *et al.*, “Handling of the generation of primary events in GAUSS, the LHCb simulation framework,” *Nuclear Science Symposium Conference Record (NSS/MIC)*, vol. IEEE, p. 1155, 2010.
- [11] D. J. Lange, “The EvtGen particle decay simulation package,” *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.
- [12] P. Golonka and Z. Was, “PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays,” *Eur.Phys.J.*, vol. C45, pp. 97–107, 2006.
- [13] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.*, “Geant4 developments and applications,” *IEEE Trans.Nucl.Sci.*, vol. 53, p. 270, 2006.
- [14] M. Clemencic *et al.*, “The LHCb simulation application, GAUSS: design, evolution and experience,” *J. Phys. Conf. Ser.*, vol. 331, p. 032023, 2011.
- [15] G. Hiller, M. Knecht, F. Legger, and T. Schietinger, “Photon polarization from helicity suppression in radiative decays of polarized Lambda(b) to spin-3/2 baryons,” *Phys.Lett.*, vol. B649, pp. 152–158, 2007.
- [16] T. Mannel and S. Recksiegel, “Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \rightarrow \Lambda \gamma$,” *J.Phys.*, vol. G24, pp. 979–990, 1998.
- [17] M. J. Aslam, Y.-M. Wang, and C.-D. Lu, “Exclusive semileptonic decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in supersymmetric theories,” *Phys.Rev.*, vol. D78, p. 114032, 2008.
- [18] Y.-m. Wang, Y. Li, and C.-D. Lu, “Rare Decays of $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ in the Light-cone Sum Rules,” *Eur.Phys.J.*, vol. C59, pp. 861–882, 2009.
- [19] C.-S. Huang and H.-G. Yan, “Exclusive rare decays of heavy baryons to light baryons: $\Lambda_b^0 \rightarrow \Lambda \gamma$ and $\Lambda_b^0 \rightarrow \Lambda l^+l^-$,” *Phys.Rev.*, vol. D59, p. 114022, 1999.
- [20] C.-H. Chen and C. Geng, “Rare $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decays with polarized lambda,” *Phys.Rev.*, vol. D63, p. 114024, 2001.
- [21] C.-H. Chen and C. Geng, “Baryonic rare decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$,” *Phys.Rev.*, vol. D64, p. 074001, 2001.
- [22] C.-H. Chen and C. Geng, “Lepton asymmetries in heavy baryon decays of $\Lambda_b^0 \rightarrow \Lambda l^+l^-$,” *Phys.Lett.*, vol. B516, pp. 327–336, 2001.
- [23] F. Zolfagharpour and V. Bashiry, “Double Lepton Polarization in $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ Decay in the Standard Model with Fourth Generations Scenario,” *Nucl.Phys.*, vol. B796, pp. 294–319, 2008.
- [24] L. Mott and W. Roberts, “Rare dileptonic decays of Λ_b^0 in a quark model,” *Int.J.Mod.Phys.*, vol. A27, p. 1250016, 2012.
- [25] T. Aliev, K. Azizi, and M. Savci, “Analysis of the $\Lambda_b^0 \rightarrow \Lambda l^+l^-$ decay in QCD,” *Phys.Rev.*, vol. D81, p. 056006, 2010.

- [2297] [26] R. Mohanta and A. Giri, “Fourth generation effect on Λ_b decays,” *Phys.Rev.*, vol. D82, p. 094022, 2010.
- [2299] [27] S. Sahoo, C. Das, and L. Maharana, “Effect of both Z and Z'-mediated flavor-changing neutral currents on the baryonic rare decay $\Lambda_b^0 \rightarrow \Lambda l^+ l^-$,” *Int.J.Mod.Phys.*, vol. A24, pp. 6223–6235, 2009.
- [2302] [28] T. Aaltonen *et al.*, “Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$,” *Phys.Rev.Lett.*, vol. 107, p. 201802, 2011.
- [2304] [29] T. Aaltonen *et al.*, “Precise measurements of exclusive $b \rightarrow s \mu^+ \mu^-$ decay amplitudes using the full cdf data set.”
- [2306] [30] R. Aaij *et al.*, “Measurement of the differential branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$,” *Phys. Lett.*, vol. B725, p. 25, 2013.
- [2308] [31] K. Olive *et al.*, “Review of Particle Physics,” *Chin.Phys.*, vol. C38, p. 090001, 2014.
- [2310] [32] “Lhcb loki twiki.” <https://twiki.cern.ch/twiki/bin/view/LHCb/LoKiHybridFilters>. Accessed: 2015-09-30.
- [2312] [33] M. Feindt and U. Kerzel, “The NeuroBayes neural network package,” *Nucl.Instrum.Meth.*, vol. A559, pp. 190–194, 2006.
- [2314] [34] M. Feindt, “A Neural Bayesian Estimator for Conditional Probability Densities,” 2004.
- [2316] [35] C. Lazzeroni, N. Watson, and M. Kreps, “Measurement of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ differential branching fraction.”
- [2318] [36] G. Punzi, “Sensitivity of searches for new signals and its optimization,” in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003.
- [2321] [37] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02.
- [2324] [38] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, “Rare baryon decays $\Lambda_b \rightarrow \Lambda l^+ l^- (l = e, \mu, \tau)$ and $\Lambda_b \rightarrow \Lambda \gamma$: differential and total rates, lepton- and hadron-side forward-backward asymmetries,” *Phys.Rev.*, vol. D87, p. 074031, 2013.
- [2328] [39] R. Aaij *et al.*, “Measurements of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV,” *Phys.Lett.*, vol. B724, p. 27, 2013.
- [2330] [40] R. Aaij *et al.*, “Precision measurement of the Λ_b^0 baryon lifetime,” *Phys. Rev. Lett.*, vol. 111, p. 102003, 2013.

- [41] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, “ $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ form factors and differential branching fraction from lattice QCD,” *Phys. Rev.*, vol. D87, no. 7, p. 074502, 2013.
- [42] R. Aaij *et al.*, “Precision measurement of the Λ_b^0 baryon lifetime,” *Phys. Rev. Lett.*, vol. 111, p. 102003, 2013.
- [43] R. Aaij *et al.*, “Measurements of the $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV,” *Phys. Lett.*, vol. B724, p. 27, 2013.
- [44] T. M. Karbach, “Feldman-Cousins Confidence Levels - Toy MC Method,” 2011.
- [45] T. Blake, T. Gershon, and G. Hiller, “Rare b hadron decays at the LHC,” 2015.
- [46] G. Hiller and F. Kruger, “More model independent analysis of $b \rightarrow s$ processes,” *Phys. Rev.*, vol. D69, p. 074020, 2004.
- [47] V. Khachatryan *et al.*, “Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data,” *Nature*, 2015.
- [48] R. Aaij *et al.*, “Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays,” 2014. to appear in *Phys. Rev. Lett.*
- [49] “Probnn presentation at ppt meeting.” <https://indico.cern.ch/event/226062/contribution/1/material/slides/0.pdf>. Accessed: 2015-09-30.
- [50] “LHCb technical design report: Reoptimized detector design and performance,” 2003.
- [51] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” *eConf*, vol. C0303241, p. MOLT007, 2003.
- [52] M. Pivk and F. R. Le Diberder, “SPlot: A Statistical tool to unfold data distributions,” *Nucl. Instrum. Meth.*, vol. A555, pp. 356–369, 2005.
- [53] R. e. a. Aaij, “LHCb Detector Performance,” *Int. J. Mod. Phys. A*, vol. 30, p. 1530022. 82 p, Dec 2014.
- [54] A. Ali, E. Lunghi, C. Greub, and G. Hiller, “Improved model independent analysis of semileptonic and radiative rare B decays,” *Phys. Rev.*, vol. D66, p. 034002, 2002.