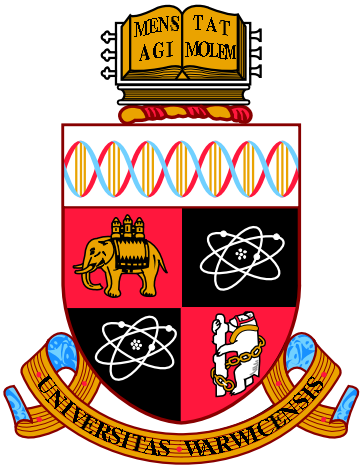

Branching fraction measurements of $B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ decays

by

Matthew M. Reid

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Physics

May 2014

THE UNIVERSITY OF WARWICK

Contents

List of Tables **vi**

List of Figures **xiv**

Acknowledgments **xxiii**

Declarations **xxv**

Abstract **xxvi**

Chapter 1 Introduction **1**

Chapter 2 Theory **4**

2.1 Observational motivation 4

2.2 The standard model 5

2.2.1 The SM Lagrangian - Quantum Chromodynamics 6

2.2.2 The SM Lagrangian - electroweak origins of CP violation . . 7

2.2.3 Higgs field, spontaneous symmetry breaking and gauge boson masses 8

2.2.4 Yukawa and fermion masses 11

2.3 The CKM matrix 12

2.4 Unitarity triangles 13

2.5 Why is *CP* violation important? 14

2.5.1 *CP* violation in decay 15

2.5.2 *CP* violation - mixing 16

2.5.3 *CP* violation - between mixing and decay 18

2.5.4 Measurement of β_s 19

Chapter 3 Experimental overview **21**

3.1 The LHC 21

3.2 LHCb experiment . . . 23
3.3 Tracking . . . 25
3.3.1 Tracking - **VE**rtex **LO**cator (**VELO**) . . . 26
3.3.2 Tracking - tracking stations . . . 30
3.3.3 Magnet . . . 33
3.4 Particle identification . . . 33
3.4.1 Particle identification - RICH . . . 34
3.4.2 Particle identification - calorimeters . . . 38
3.4.3 Particle identification - muon system . . . 41
3.5 Trigger system . . . 43
3.5.1 Trigger - level 0 . . . 43
3.5.2 Trigger - high level trigger . . . 44
3.6 Stripping and trigger . . . 44
3.6.1 K^0_S reconstruction . . . 45
3.6.2 J/ψ reconstruction . . . 45
3.6.3 Bachelor pions and kaons . . . 46
3.6.4 B^0 reconstruction . . . 46
3.6.5 Trigger lines . . . 46
3.7 Dataset . . . 47

Chapter 4 Analysis overview **49**

I $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ analysis **53**

Chapter 5 Backgrounds **54**
5.1 $B^0 \to J/\psi\, K^0_S$ backgrounds . . . 54
5.2 $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ backgrounds . . . 55
5.3 Backgrounds from mis-reconstructed photons - $B^0 \to J/\psi\, K^0_S \eta'$ . . . 57

Chapter 6 Final selection requirements **60**
6.1 Vetoes . . . 60
6.2 Data – Monte Carlo comparisons . . . 62
6.3 CROP optimisation . . . 63

Chapter 7 Fit model and yields **70**
7.1 Signal model . . . 70
7.2 Combinatorial background model . . . 72
7.3 Fit results . . . 72

7.4 Correcting yields for non-resonant $\psi(2S)$ decays . . . 74
7.5 Invariant mass combinations . . . 75

Chapter 8 Efficiencies and systematic uncertainties **81**
8.1 Harmonic weighting of efficiencies . . . 81
8.1.1 Geometric efficiency . . . 82
8.1.2 Reconstruction, selection and trigger efficiency . . . 83
8.1.3 Track reconstruction efficiency . . . 85
8.1.4 Particle identification efficiency . . . 85
8.2 Systematic uncertainties . . . 87
8.2.1 Track reconstruction . . . 87
8.2.2 Trigger efficiency . . . 87
8.2.3 Particle identification efficiency . . . 89
8.2.4 Fit model – $B^0 \to J/\psi K^0_S \pi^+\pi^-$ . . . 90
8.2.5 Simulated sample size . . . 91
8.2.6 Efficiency variation across the phase space . . . 91
8.3 Total systematic uncertainty . . . 92

Chapter 9 Measurement of $\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)/\mathcal{B}(B^0 \to J/\psi K^0_S)$ **93**
9.1 Cross check measurement of $\mathcal{B}(B^0 \to \psi(2S)K^0_S)/\mathcal{B}(B^0 \to J/\psi K^0_S)$ . . . 94

II $B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}$ analysis **96**

Chapter 10 Event Selection **97**
10.1 Vetoes . . . 98
10.2 Initial cuts . . . 99
10.3 Optimisation of $B^0 \to J/\psi K^0_S \pi^+\pi^-$ . . . 100
10.4 Tuning the selection for other modes . . . 101
10.4.1 Calculation of signal scale factor $f_{\rm sig}$ . . . 104
10.4.2 Calculation of background scale factor $f_{\rm bkg}$ . . . 105
10.5 Final selection . . . 105

Chapter 11 Fit model **109**
11.1 Signal model . . . 109
11.2 Signal cross-feed model . . . 111
11.3 Combinatorial background model . . . 112
11.4 Toy studies . . . 113
11.5 Unblinded fit results . . . 114

11.5.1 Bias correction 114
11.6 Invariant mass distributions 115

Chapter 12 Efficiencies and systematic uncertainties 132
12.1 Discussion of efficiencies 132
12.1.1 Geometric efficiency 133
12.1.2 Reconstruction, selection and trigger efficiency 133
12.1.3 Particle identification efficiency 134
12.2 Systematic uncertainties 135
12.2.1 Choice of PDF shapes (fit model) 135
12.2.2 Fit bias 137
12.2.3 Efficiencies 137
12.2.4 Particle identification 138
12.2.5 Efficiency variation over the phase space 138
12.2.6 Simulation sample size 140
12.2.7 Fragmentation fraction 140
12.2.8 Statistical uncertainty on $B^0 \to J/\psi K^0_S \pi^+\pi^-$ yields 140
12.3 Summary of systematic uncertainties 140

Chapter 13 Relative branching fraction measurements 146
13.1 Results 146
13.2 $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ branching fraction 149
13.2.1 $B^0 \to J/\psi K^0_S K^+ K^-$ branching fraction 149
13.2.2 $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ branching fraction 150
13.2.3 $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ branching fraction 151
13.2.4 $B^0_s \to J/\psi K^0_S K^+ K^-$ branching fraction 152

III Summary 156

Chapter 14 Summary and absolute branching fractions 157

Appendix A Variable alias dictionary 160

Appendix B Generator level efficiencies 161
B.1 Systematic due to generator level variation 161
B.2 Generator level efficiencies for Part II modes 163
B.2.1 Observed modes and their efficiency corrections 164
B.2.2 Unobserved mode and associated uncertainties 165

B.3 Observed invariant mass combinations 166

Appendix C Efficiencies for $B^0 \to \psi(2S)K^0_S$ decays **180**

C.0.1 Geometric efficiency 180

C.0.2 Reconstruction, selection and trigger efficiency 180

C.0.3 Track reconstruction efficiency $\psi(2S)$ 181

C.0.4 Particle identification efficiency 181

C.1 Systematic uncertainties for $\mathcal{B}(B^0 \to \psi(2S)K^0_S)/\mathcal{B}(B^0 \to J/\psi\, K^0_S)$ . . 182

C.1.1 Fit model – $B^0 \to \psi(2S)K^0_S$ 183

C.1.2 Total systematic uncertainty 183

List of Tables

1.1 States below 1500 MeV/c^2 listed in the PDG [15] that decay to $K^0_S K^{\pm}\pi^{\mp}$. The $h_1(1380)$, listed as needing confirmation, but seen in $K\bar{K}\pi$, is not included. 2

1.2 Previous measurements of $B^0_{(s)} \to J/\psi K^0_S h^+ h'^-$ decays [15]. 3

2.1 Summary of the fermions, organised by generation. Masses are taken from Ref [15]. The neutrino masses have been starred (*) since there have been significant observations that imply the masses are not in fact zero [34], however, as yet this is not included in the Standard Model. 5

2.2 Summary of the standard model boson particles. Masses are taken from Ref [15]. 6

2.3 Summary of the standard model quantum numbers for quarks and leptons. The quantum numbers describe how each of the fermion types transforms under the different symmetry groups of the SM, $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$. They can be found directly in Ref [15]. . 8

3.1 Numbers of events from 2011 dataset passing the stripping selection. Inclusive samples of signal and background. 47

3.2 Simulated datasets for all B decay modes including event type and both magnet polarisations. The $B^0 \to J/\psi K^0_S \pi^+\pi^-$ and $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ samples are composed of 50 % `PHSP`, with the remainder produced through resonances decaying to $K^0_S\pi^+\pi^-$: $K^0_1(1270)$ (18 %), ρK^0_S (17.5 %), ωK^0_S (14.5 %). The rest of the decays proceed through the `PHSP` model. 48

5.1 List of simulation decay files used for background characterisation of $B^0 \to J/\psi K^0_S$. The total number events consists of both magnet polarities. . . . 54

5.2 List of simulation decay files used for background characterisation of $B^0 \to J/\psi K^0_S \pi^+\pi^-$. The total number events consists of both magnet polarities. 56

5.3 Estimates for potential peaking background coming from the mode $B_s^0 \to J/\psi K_S^0 \eta$. 58

6.1 Veto requirements. The known values of all particles are taken from the PDG [15]. All cuts are the same for LL and DD samples except where indicated otherwise. 62
6.2 Cuts for the optimised selection on simulation for both DD and LL K_S^0 types. The list is in order of significance for the LL mode, for the DD mode the numbering is labelled in open brackets next to the cut value. 63

7.1 Fitted parameters and goodness of fit obtained from fits to MC for both $B^0 \to J/\psi K_S^0$ and $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ modes. The yield which was also floated is not included in this table. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. . . . 72
7.2 Results of the fits to the $B^0 \to J/\psi K_S^0$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. 73
7.3 Results of the fit to the $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. 73
7.4 Results of the fit to the $B^0 \to \psi(2S)K_S^0$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. The measures of the goodness of fit do not make much sense with low statistics and arbitrary binning scheme. 76
7.5 Results of the $J/\psi\pi\pi$ invariant mass spectrum which has been weighted using the ${}_s\mathcal{W}eight$-ed fit from the $B^0 \to \psi(2S)K_S^0$ mode. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. The measures of the goodness of fit do not make much sense with low statistics and arbitrary binning scheme. (Left) side show decays from K_S^0(DD) and (right) side for K_S^0(LL) type decays. 76

8.1 Number of events used in weighting procedure for $B^0 \to J/\psi K_S^0$. 82
8.2 Number of events used in weighting proceedure for $B^0 \to J/\psi K_S^0 \pi^+\pi^-$. . 82

8.3 Generator level cut efficiencies for $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. The Monte Carlo productions for $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ were produced with the LHCb generator level cuts `DaughtersInLHCb` `DaughtersInLHCbWithMinP` respectively, they are outlined in more detail in Sec. 3.7. 83

8.4 The integrated efficiency, $\epsilon^{\text{Selection}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S$. 85

8.5 The integrated efficiency, $\epsilon^{\text{Selection}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$. 85

8.6 Data/MC tracking corrections for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ as calculated using the tracking efficiency tables. Separated into LL/DD and magnet polarity categories. Note that the uncertainty comes from the size of the signal MC sample. 86

8.7 PID efficiency as calculated via the `PIDCalib` tool, for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ split by LL and DD and Up and Down magnet polarities, using the method described in the text. The uncertainty is from the statistics of the control sample only. 87

8.8 Comparison of the relative trigger efficiencies for MC truth, reconstructed MC and reconstructed 2011 data. This table highlights the efficiency ratios for the K^0_S(LL) for both the $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. Errors are statistical only. 88

8.9 Comparison of the relative trigger efficiencies for MC truth, reconstructed MC and reconstructed 2011 data. This table highlights the efficiency ratios for the K^0_S(DD) for both the $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. Errors are statistical only. 88

8.10 Difference in the final ratios of $\epsilon^{\text{trg/tot}}_{B^0 \to J/\psi K^0_S} / \epsilon^{\text{trg/tot}}_{B^0 \to J/\psi K^0_S \pi^+\pi^-}$ between MC truth and 2011 Data for the full trigger selection, L0 × HLT1 × HLT2. The K^0_S type and magnet polarity are secified for each mode. 89

8.11 Particle identification systematic error for combined π^+ and π^- tracks. . 90

8.12 Variation of the signal and background PDFs for $B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$ and $B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$ modes. 90

8.13 Number of events used in weighting procedure for $B^0 \to J/\psi K^0_S \pi^+\pi^-$. . . 91

8.14 Summary of the weighted-average values of all considered sources of systematic error for the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. The final value is the sum of the contributions in quadrature. 92

9.1 Relative trigger, selection and reconstruction efficiencies from simulation. Based on simulation optimised selection for $B^0 \to J/\psi K^0_S \pi^+\pi^-$. 94
9.2 Relative efficiencies of $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$. 95
9.3 Branching fraction measurements found in Ref. [15] and used in the calculation of the expected branching fraction for $B^0 \to \psi(2S)K^0_S$. 95

10.1 Efficiency when applying the Decay Tree Fit (DTF) constraint such that the K^0_S and J/ψ are mass constrained as well as the PV. Taken from signal simulated events where the errors are statistical only. 97
10.2 Veto requirements. The masses values of all particles are taken from the PDG [15]. All cuts are the same for LL and DD samples except where indicated otherwise. Those cuts marked with a (*) are only applied to the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ mode. 100
10.3 Set of loose cuts applied to the data passing the stripping selection. These cuts clean up the signal enough so that we can perform a rough fit to the data. An asterisk (*) indicates a cut only applied to K^0_S(LL). 100
10.4 List of cut variables used by CROP. Candidates' vertex reconstruction, B pointing direction and lifetime-related variables are considered. 101
10.5 Calculated values for the signal scale factors, based on the fraction of the signal yields. 105
10.6 Calculated values for the background scale factor, based on the fraction of the background yields in the upper side band $5425 < M_{J/\psi K^0_S h'h} < 5550\,\mathrm{MeV}/c^2$. 105
10.7 Cuts for the data optimised selection both K^0_S(LL) and K^0_S(DD) types. The list is in order of significance for the LL mode for the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ selection, for the DD mode the numbering is labelled in open brackets next to the cut value and the other modes are also indexed appropriately when required. 107
10.8 Number of multiple candidates in each of the data samples. 108

11.1 Fitted parameters from fits to MC for $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ modes. 111
11.2 Efficiencies of B candidate reconstructed under a different final state to that which was generated. The full selection and trigger have been applied here but not particle identification. The values are obtained using the total sample combining both magnet up and down polarities, the errors are statistical only. 113

11.3 Gaussian constraints that are applied to the cross-feed yields for all modes. If a mode is not present then its yield has been deemed negligible and thus omitted. The numerator and denominator are taken in an analagous fashion; $\epsilon^{\text{Generator}} \times \epsilon_X^{\text{Selection}} \times \epsilon_X^{\text{PIDCalib}}$ where X corresponds to whether it is a mis-identified (numerator) or truth (denominator) efficiency. 118

11.4 Pull mean and width values for parameters of the fit to data. The yields for the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ modes include goodness of fit information and correspond to figures 11.3 and 11.4 for B^0 and B^0_s modes respectively. 121

11.5 Fitted parameters from unblinded simultaneous fit to 2011 dataset. The significance of the signal yields are statistical only and do not include systematic corrections. 125

11.6 Combined statistical significance of the DD and LL yields from the simultaneous fit to 2011 dataset. Note that no systematic correction is applied. 126

11.7 Fitted and corrected yields for $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ modes. 126

12.1 Generator level cut efficiencies for all six modes. Both magnet polarities are presented but they fall well within the statistical limitation of one another. The simulation productions listed here were all run with the `DaughtersInLHCbWithMinP` requirement. 133

12.2 Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (left) $B^0 \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$ and (right) $B^0 \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$. All uncertainties are statistical only. 134

12.3 Phase-space integrated relative trigger, selection and reconstruction efficiencies for $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ modes. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. All uncertainties are statistical only. (left) $B^0_s \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$ and (right) $B^0_s \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$. . 135

12.4 Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$. (left) $B^0 \to J/\psi\, K^0_S(\text{LL})K^\pm\pi^\mp$ and (right) $B^0 \to J/\psi\, K^0_S(\text{DD})K^\pm\pi^\mp$. All uncertainties are statistical only. 136

12.5 Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$. (left) $B^0_s \to J/\psi\, K^0_S(\text{LL})K^\pm\pi^\mp$ and (right) $B^0_s \to J/\psi\, K^0_S(\text{DD})K^\pm\pi^\mp$. All uncertainties are statistical only. 137

12.6 Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample,) following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$. (left) $B^0 \to J/\psi\, K^0_S(\text{LL})K^+K^-$ and (right) $B^0 \to J/\psi\, K^0_S(\text{DD})K^+K^-$. All uncertainties are statistical only. 138

12.7 Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$. (left) $B^0_s \to J/\psi\, K^0_S(\text{LL})K^+K^-$ and (right) $B^0_s \to J/\psi\, K^0_S(\text{DD})K^+K^-$. All uncertainties are statistical only. 139

12.8 PID efficiencies for all modes obtained using the PIDCalib procedure. The ±0.01 uncertainty due to MC statistics is neglected. 140
12.9 Mis-identification rates for all modes. The first number is the rate, second the statistical uncertainty and third is the systematic uncertainty due to the PIDCalib procedure. 141
12.10 Relative systematic uncertainties (in %) due to the change of functional form of the signal PDF, varying fixed parameters within their uncertainty, varying the B_s^0 width in accordance with the relative difference between the B^0-B_s^0 Q-values, and changing the functional form of the background PDF. The final column is the sum in quadrature of all these contributions. . . . 142
12.11 Summary of the systematic error induced by the fit model bias for all modes separated into K_S^0(LL) and K_S^0(DD) categories. 142
12.12 Summary of the PID systematics for all modes obtained using the PIDCalib procedure separated into K_S^0(LL) and K_S^0(DD) categories. 143
12.13 Systematic uncertainty due to efficiency variation over the phase space. The column ϵ^{gen} is due to the generator level efficiency, and the column $\epsilon^{\text{Selection}}$ is due to the selection efficiency. 143
12.14 Summary of the systematics induced by the limitation in the simulation sample size, separated into K_S^0(LL) and K_S^0(DD) categories. 144
12.15 Summary of the systematic uncertainties for each mode. Not included in the table are the uncertainties due to the fragmentation fraction.The total systematic uncertainty is the sum in quadrature of all contributions. . . . 144
12.16 Summary of the systematic uncertainties that effect the measurement of the branching fraction relative to that of $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ for each mode. Not included in the table are the uncertainties due to the fragmentation fraction. The total systematic uncertainty is the sum in quadrature of all contributions. 145

13.1 Relevant efficiencies, fragmentation fractions and yields entering in the relative branching fractions measurements for $B^0_{(s)} \to J/\psi K_S^0 h^+ h^{(\prime)-}$ decays for Down-Down and Long-Long K_S^0 categories. 147

A.1 Dictionary of variable names used in the analysis and their definition. . . . 160

B.1 Number of events used in weighting procedure for $B^0 \to J/\psi K_S^0 \pi^+\pi^-$. . . 164
B.2 Number of events used in weighting procedure for $B^0 \to J/\psi K_S^0 K^+K^-$. . 166
B.3 Number of events used in weighting procedure for $B_s^0 \to J/\psi K_S^0 K^\pm\pi^\mp$. . 166

B.4 $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ mode invariant mass standard deviation of efficiency variation. . . . 166
B.5 $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$ mode invariant mass standard deviation of efficiency variation. . . . 168
B.6 $B^0_s \to J/\psi K^0_S K^+ K^-$ mode invariant mass standard deviation of efficiency variation. . . . 169

C.1 Signal Monte Carlo datasets for $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)\ K^0_S$ decay modes including event type and both magnet polarisations. . . . 180
C.2 `DaughtersInLHCb` cut efficiency. . . . 181
C.3 Trigger, selection and reconstruction efficiencies for $B^0 \to \psi(2S)K^0_S$. The numbers of events surviving each cut are given, with the values in parentheses being the efficiencies calculated relative to the preceding requirement, following the order of the table. The overall efficiency and its uncertainty (statistical only) is also given. . . . 181
C.4 Trigger, selection and reconstruction efficiencies for $B^0 \to J/\psi K^0_S$. The numbers of events surviving each cut are given, with the values in parentheses being the efficiencies calculated relative to the preceding requirement, following the order of the table. The overall efficiency and its uncertainty (statistical only) is also given. . . . 182
C.5 Data/MC tracking efficiency corrections for $B^0 \to \psi(2S)\ K^0_S$ as calculated using the tracking efficiency tables. Separated by LL/DD and magnet polarity. Note that the uncertainty comes from the size of the signal MC samples. . . . 182
C.6 PID efficiency, for $B^0 \to \psi(2S)K^0_S$ split by K^0_S(LL), K^0_S(DD) and magnet polarities, using the method described in the text. The uncertainties are given in order as; statistical due from the control sample and induced systematic using the outlined method. . . . 183
C.7 Variation of the signal and background PDFs for $B^0 \to \psi(2S)K^0_S$(LL) and $B^0 \to \psi(2S)K^0_S$(DD) modes. . . . 183
C.8 Summary of the weighted-average values of all considered sources of systematic error for $B^0 \to \psi(2S)K^0_S\pi^+\pi^-$. The final value is the sum of the contributions in quadrature. . . . 184

List of Figures

1.1 Feynman diagrams showing possible tree-level decays. (top) $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ with $\eta(1405) \to KK\pi$. (bottom) $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ with $K_1(1270) \to K\rho$, $(K_0^*(1430) \to K\pi)\pi$. 3

2.1 Experimental constraints on the angles α, β, γ angles the latest combined results in the $(\bar{\rho}, \bar{\eta})$ plane of the unitary triangle. Figure taken from Ref. [45]. 20

3.1 The CERN accelerator complex (not to scale). Figure taken from [57]. . . 22
3.2 Schematic view of the LHCb detector in yz projection made at $x = 0$. Figure taken from [58]. 24
3.3 Results for fits to the 2011+2012 datasets for the rare $B \to \mu\mu$ decays by CMS [60] (left) and LHCb [61] (right). 25
3.4 A diagram of the VELO stations along the xz-plane, indicating the angular acceptance and the arrangement of the stations. The diagram shows a VELO station in both the open and closed position in the xy-plane. Taken from Ref. [58]. 26
3.5 Sketch of the $R\phi$ geometry of the VELO sensors, showing only a portion of the strips that are actually present on the real sensors. Taken from Ref. [58]. 27
3.6 Simulation of the VELO stations showing how the RF-foil encloses the system (bottom). Taken from Ref. [58]. 28
3.7 Hit resolution as a function of the inner-strip readout pitch as measured in 2011 (left). Primary vertex resolution as a function of event multiplicity (right). Impact parameter (IP_x) comparing 2011 simulation and data (bottom). Taken from Ref. [64]. 29
3.8 TT (left) and IT boxes (right). Both taken from Ref. [65]. 31
3.9 Cross-section of an OT module (left) and a full overview of the tracking systems excluding the VELO (right). Taken from Ref. [58]. 32

3.10 Schematic of the B_y-field component as a function of z-axis. The various types of reconstructible tracks are also displayed: long, upstream, downstream, VELO and T tracks. Figure is taken from Ref. [67]. 34
3.11 Invariant mass distribution for $B \to h^+h'^-$ decays [68] (where $h^{(\prime)} = \pi, K$) before the application of the RICH information (left) and after applying particle identification (right). The distributions seen here are the signal being investigated which is $B^0 \to \pi^+\pi^-$ (turquoise dotted line) then several background components: mis-identified $B^0 \to K^+\pi^-$ (red dashed-dotted line), partially reconstructed B three body modes (orange dashed line), double mis-identification from $B_s^0 \to K^+K^-$ (yellow line), single mis-identification from $B_s^0 \to \pi^+K^-$ (brown line), $\Lambda_b^0 \to pK$ (purple/magenta line) and $\Lambda_b^0 \to p\pi$ (green line). After applying particle identification only the signal and two small background contributions remain, all others are negligible. The solid grey line represents the combinatorial background in both figures. Taken from Ref. [69]. 35
3.12 (Left) RICH1 detector as viewed from the side and (right) RICH2 as viewed from above. Spherical and flat mirrors are indicated as well as the size of each detector. Taken from Ref. [58]. 37
3.13 Cherenkov angle data for the reconstructed Cherenkov angle as a function of track momentum for different radiators. Taken from Ref. [69] 38
3.14 The internals of the HCAL (left) and ECAL (right). The main difference being the HCAL uses scintillating tiles parallel to z whilst the ECAL has its placed perpendicular to the z-axis. This causes the readout electronics to be directed differently under each scenario. 40
3.15 Diagrams showing the segmentation of the calorimeter components. Segmentation of the SPD, PS and ECAL (left) and the segmentation of the HCAL sensors (right). Taken from [58]. 40
3.16 Side view of the muon system (left) and exploded view of a single muon chamber (right), notice the multiple gas gaps. Taken from [58]. 42

5.1 $\Lambda_b^0 \to J/\psi\,\Lambda$ simulated decays reconstructed in the $J/\psi\,K_S^0$ final state (left) before and (right) after a veto is applied. The top row shows the effect on the K_S^0(LL) reconstructed mode. The curves are only to guide the eye. . . 55
5.2 Simulated partially reconstructed background modes (left) $B^0 \to J/\psi\,K_S^0\pi^+\pi^-\pi^0$ and (right) $B^+ \to J/\psi\,K_S^0\pi^+\pi^-\pi^+$ when reconstructed as $B^0 \to J/\psi\,K_S^0\pi^+\pi^-$. Both samples above display the K_S^0(LL) category only. 57

5.3 Analysis of various cocktail MC samples to see if there are any peaking backgrounds we did not consider. Reconstructed under the $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ invariant mass hypothesis. (Top left) $B^0 \to J/\psi X$, (top right) $B^0_s \to J/\psi X$ and (bottom) $B^+ \to J/\psi X^+$. 58
5.4 Simulated partially reconstructed backgrounds from (left) $B^0 \to J/\psi K^0_S\eta$ and (right) $B^0 \to J/\psi K^0_S\eta'$ events that pass all selection requirements and are in the mass range used in the fit. Both samples above display the $K^0_S(\mathrm{LL})$ category only. 59
5.5 Simulated partially reconstructed backgrounds from (left) $B^0_s \to J/\psi K^0_S\eta$ and (right) $B^0_s \to J/\psi K^0_S\eta'$ events that pass all selection requirements and are in the mass range used in the fit. Both samples above display the $K^0_S(\mathrm{LL})$ category only. 59

6.1 Candidate (top) $B^0 \to J/\psi K^0_S\pi^+\pi^-$ and (bottom) $B^0 \to J/\psi K^0_S$ decays in the whole invariant mass range reconstructed with either pion track from the K^0_S candidate assigned the proton mass hypothesis. (Left) $K^0_S(\mathrm{LL})$ sample and (right) $K^0_S(\mathrm{DD})$ type. The blue lines show $|m(p\pi^-) - m_\Lambda|$ where the cut is at $\pm 10\,\mathrm{MeV}/c^2$ for $K^0_S(\mathrm{LL})$ and $\pm 25\,\mathrm{MeV}/c^2$ for $K^0_S(\mathrm{DD})$). 61
6.2 Candidate $B^0 \to J/\psi K^0_S\pi^+\pi^-$ decays in the whole invariant mass range reconstructed with either bachelor track assigned the proton mass hypothesis. (Left) LL sample, (right) DD sample. The blue lines show the vetoed region. 62
6.3 $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ normalised distributions comparing data and Monte Carlo, red line indicates MC and black is for signal ${}_s\mathcal{W}eighted$ data and blue is for background ${}_s\mathcal{W}eighted$ data. Units are omitted. There is a relatively good agreement between signal like variables. 65
6.4 Correlations between training variables for $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$; (top) signal MC, (bottom) data sideband. 66
6.5 Correlations between training variables for $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$; (top) signal MC, (bottom) data sideband. 67
6.6 Separation of variables for the $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ mode where green and blue is signal for different magnet polarities, red and black are background from the two magnet polarities. 68
6.7 Separation of variables for the $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ mode where green and blue is signal for different magnet polarities, red and black are background from the two magnet polarities. 69

7.1 Results of the fits to the MC samples used to obtain the signal PDF parameters. (Top) $B^0 \to J/\psi K^0_S$ and (bottom) $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (Left) LL and (right) DD samples. 71
7.2 Fit results for the $B^0 \to J/\psi K^0_S$(LL) sample on (left) linear and (right) logarithmic scales. 74
7.3 Fit results for the $B^0 \to J/\psi K^0_S$(DD) sample on (left) linear and (right) logarithmic scales. 74
7.4 Fit results for the $B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$ sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind. 75
7.5 Fit results for the $B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$ sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind. 75
7.6 Fit results for the $B^0 \to \psi(2S)K^0_S$(LL) sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind. 77
7.7 Fit results for the $B^0 \to \psi(2S)K^0_S$(DD) sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind. 77
7.8 Fit results of the $\psi(2S) \to J/\psi\pi^+\pi^-$ invariant mass spectrum having applied ${}_s\mathcal{W}eights$ calculated from Figs. 7.7 and 7.6 for K^0_S(DD) and K^0_S(LL) types respectively. The sample on (left) K^0_S(DD) and (right) K^0_S(LL). . . 78
7.9 $B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$ two body combinations of the daughters, Red indicates simulation and black is for ${}_s\mathcal{W}eighted$ data. There is an excess of $\rho(770) \to \pi^+\pi^-$ in the simulation due to the forced branching fraction set in the decay file. 79
7.10 $B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$ three body combinations of the daughters, Red indicates simulation and black is for ${}_s\mathcal{W}eighted$ data. There is an excess of $K_1(1270) \to K^0_S\pi^+\pi^-$ in the simulation due to the branching fraction set in the decay file. We see no obvious exotic charmonium states such as $X(3872) \to J/\psi\pi^+\pi^-$. 80

8.1 Generator level cut efficiencies of (left) `DaughtersInLHCb` requirement and (right) the additional requirement that pion and muon tracks have minimum momentum greater than 1600 MeV/c and 3000 MeV/c respectively (`DaughterInLHCbWithMinP` cut). The $\psi(2S)$ region is shown with vertical lines. 83
8.2 Reconstruction level (left) stripping line and (right) CROP selection efficiencies. These are for the DD K^0_S sample. The $\psi(2S)$ region is shown with vertical lines. 84

8.3 Reconstruction level (left) trigger and (right) total efficiency in bins of $m_{J/\psi\pi\pi}$. The total efficiency is the product of all the relative efficiency histograms that have been applied. These are for the DD K^0_S sample. The $\psi(2S)$ region is shown with vertical lines. 84

10.1 Candidate (top) $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and (bottom) $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ decays in the whole invariant mass range reconstructed with either daughter of the K^0_S candidate reconstructed under the proton mass hypothesis for 2011 stripped data. The blue lines show $|m(p\pi^-) - m_\Lambda| > 10\,\text{MeV}/c^2$ (the cut is at $\pm 10\,\text{MeV}/c^2$ for K^0_S(LL) and $\pm 25\,\text{MeV}/c^2$ for K^0_S(DD). 98

10.2 Candidate (top) $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and (bottom) $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ decays in the whole invariant mass range reconstructed with either bachelor track assigned the proton mass hypothesis for 2011 stripped data. The blue lines show the vetoed region. (Left) LL sample, (right) DD sample. . . . 99

10.3 Gaussian with linear background mass fit after applying loose cuts outlined in Table 10.3. (Left) $B^0 \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$ data, (right) $B^0_s \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$ data. 101

10.4 Correlations between variables using *sWeights* technique applied to the $B^0 \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$ training data; (top) signal weighted, (bottom) background weighted. 102

10.5 Correlations between variables using *sWeights* technique applied to the $B^0 \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$ training data; (top) signal weighted, (bottom) background weighted. 103

11.1 Various mis-identification invariant mass distributions for $B^0 \to J/\psi K^0_S h^+ h^{(\prime)-}$ where a $\pi \Leftrightarrow K$ is interchanged under the wrong mass hypothesis. The left side column are K^0_S(LL) and right column K^0_S(DD) KS types. From top to bottom we have: $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi K^0_S \pi^+\pi^-$, $B^0 \to J/\psi K^0_S \pi^+\pi^-$ reconstructed as $J/\psi K^0_S K^+\pi^-$, $B^0 \to J/\psi K^0_S K^+ K^-$ reconstructed as $J/\psi K^0_S K^+\pi^-$ and finally $B^0 \to J/\psi K^0_S K^\pm\pi^\mp$ reconstructed as $J/\psi K^0_S K^+ K^-$. The red lines are for visual purposes and indicate our nominal fitting range. 116

11.2 Various mis-identification invariant mass distributions for $B^0_s \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ where a $\pi \Leftrightarrow K$ is interchanged under the wrong mass hypothesis. The left side column are K^0_S(LL) and right column K^0_S(DD) KS types. From top to bottom we have: $B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi\, K^0_S \pi^+\pi^-$, $B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$ reconstructed as $J/\psi\, K^0_S K^+\pi^-$, $B^0_s \to J/\psi\, K^0_S K^+ K^-$ reconstructed as $J/\psi\, K^0_S K^+\pi^-$ and finally $B^0_s \to J/\psi\, K^0_S K^\pm\pi^\mp$ reconstructed as $J/\psi\, K^0_S K^+K^-$. The red lines are for visual purposes and indicate our nominal fitting range. 117
11.3 Unblinded pull distribution plots for B^0 signal yields. (Top) $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$, (middle) $B^0 \to J/\psi\, K^0_S K^\pm\pi^\mp$ and (bottom) $B^0 \to J/\psi\, K^0_S K^+K^-$. (Left) LL and (right) DD samples respectively. 119
11.4 Pull distribution plots for B^0_s signal yields. (Top) $B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$, (middle) $B^0_s \to J/\psi\, K^0_S K^\pm\pi^\mp$ and (bottom) $B^0_s \to J/\psi\, K^0_S K^+K^-$. (Left) LL and (right) DD samples respectively. 120
11.5 Invariant mass fits to the $B^0_{(s)} \to J/\psi\, K^0_S \pi^+\pi^-$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively. 122
11.6 Unblinded invariant mass fits to the $B^0_{(s)} \to J/\psi\, K^0_S K^\pm\pi^\mp$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively. 123
11.7 Unblinded invariant mass fits to the $B^0_{(s)} \to J/\psi\, K^0_S K^+K^-$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively. 124
11.8 Background-subtracted distributions of the possible two-body invariant mass combinations in $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ decays. Contributions from the $\rho(770)^0$ and $K^*(892)^\pm$ mesons can be seen in the $m(\pi^+\pi^-)$ and $m(K^0_S\pi^\pm)$ distributions respectively. 127
11.9 Background-subtracted distributions of the possible three-body invariant mass combinations in $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ decays. A possible enhancement from the $K_1(1400)$ state can be seen in the $m(K^0_S\pi^+\pi^-)$ distribution. . . 128
11.10 Background-subtracted distributions of the possible two-body invariant mass combinations in $B^0_s \to J/\psi\, K^0_S K^\pm\pi^\mp$ decays. Contributions from the $K^*(892)^0 + \overline{K}^*(892)^0$ and $K^*(892)^\pm$ mesons can be seen in the $m(K^\pm\pi^\mp)$ and $m(K^0_S\pi^\pm)$ distributions respectively. 129
11.11 Background-subtracted distributions of the possible three-body invariant mass combinations in $B^0_s \to J/\psi\, K^0_S K^\pm\pi^\mp$ decays. No clear signatures of narrow resonances are observed. 130

11.12A zoom in of the lower invariant mass range of the background-subtracted $m(K^0_S K^+\pi^-)$ invariant mass combination in $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ decays. There are hints of the $f_1(1420)$ resonance but no significant observation is observed. 130
11.13Background-subtracted distributions of the possible two-body invariant mass combinations in $B^0 \to J/\psi K^0_S K^+ K^-$ decays. The $\phi(1020)$ resonance is clearly seen in the $m(K^+K^-)$ distribution. 131
11.14Background-subtracted distributions of the possible three-body invariant mass combinations in $B^0 \to J/\psi K^0_S K^+ K^-$ decays. No clear signatures of narrow resonances are observed. 131

13.1 Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0 \to J/\psi K^0_S(\mathrm{LL}) K^\pm \pi^\mp$ and (top right) $B^0 \to J/\psi K^0_S(\mathrm{DD}) K^\pm \pi^\mp$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included. 150
13.2 Posterior likelihood for $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %. 151
13.3 Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0 \to J/\psi K^0_S(\mathrm{LL}) K^+ K^-$ and (top right) $B^0 \to J/\psi K^0_S(\mathrm{DD}) K^+ K^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included. 152
13.4 Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0_s \to J/\psi K^0_S(\mathrm{LL}) \pi^+ \pi^-$ and (top right) $B^0_s \to J/\psi K^0_S(\mathrm{DD}) \pi^+ \pi^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included. 153
13.5 Profile likelihood for $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %. 153
13.6 Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0_s \to J/\psi K^0_S(\mathrm{LL}) K^\pm \pi^\mp$ and (top right) $B^0_s \to J/\psi K^0_S(\mathrm{DD}) K^\pm \pi^\mp$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included. 154
13.7 Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0_s \to J/\psi K^0_S(\mathrm{LL}) K^+ K^-$ and (top right) $B^0_s \to J/\psi K^0_S(\mathrm{DD}) K^+ K^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included. 154

13.8 Posterior likelihood for $B_s^0 \to J/\psi K_S^0 K^+ K^-$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %. 155

B.1 Efficiency of `DaughtersInLHCb` cut across various invariant mass combinations. From left to right and top to bottom: $m(K_S^0\pi^+)$, $m(K_S^0\pi^-)$, $m(J/\psi\pi^+\pi^-)$, $m(K_S^0\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2). 162

B.2 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations. From left to right and top to bottom: $m(K_S^0\pi^+)$, $m(K_S^0\pi^-)$, $m(J/\psi\pi^+\pi^-)$, $m(K_S^0\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2). 163

B.3 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0 \to J/\psi K_S^0 K^+ K^-$ mode. From left to right and top to bottom: $m(K_S^0 K^+)$, $m(K_S^0 K^-)$, $m(J/\psi K^+ K^-)$, $m(K_S^0 K^+ K^-)$, $m(K^+ K^-)$ (in MeV/c^2). 165

B.4 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ mode. From left to right and top to bottom: $m(K_S^0 K^+)$, $m(K_S^0 \pi^-)$, $m(J/\psi K^+ \pi^-)$, $m(K_S^0 K^+ \pi^-)$, $m(K^+ \pi^-)$ (in MeV/c^2). 167

B.5 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ mode. From left to right and top to bottom: $m(K_S^0 K^+)$, $m(K_S^0 \pi^-)$, $m(J/\psi K^+ \pi^-)$, $m(K_S^0 K^+ \pi^-)$, $m(K^+ \pi^-)$ (in MeV/c^2). 168

B.6 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B_s^0 \to J/\psi K_S^0 \pi^+ \pi^-$ mode. From left to right and top to bottom: $m(K_S^0\pi^+)$, $m(K_S^0\pi^-)$, $m(J/\psi\pi^+\pi^-)$, $m(K_S^0\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2). 169

B.7 Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B_s^0 \to J/\psi K_S^0 K^+ K^-$ mode. From left to right and top to bottom: $m(K_S^0 K^+)$, $m(K_S^0 K^-)$, $m(J/\psi K^+ K^-)$, $m(K_S^0 K^+ K^-)$, $m(K^+ K^-)$ (in MeV/c^2). 170

B.8 $B^0 \to J/\psi K_S^0(\mathrm{DD})\pi^+\pi^-$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 171

B.9 $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 172
B.10 $B^0_s \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 173
B.11 $B^0_s \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 174
B.12 $B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 175
B.13 $B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units. 176
B.14 $B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ example showing invariant mass combinations as a function of $\epsilon^{\mathrm{Selection}}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements. 177
B.15 $B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp$ example showing invariant mass combinations as a function of $\epsilon^{\mathrm{Selection}}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements. 178
B.16 $B^0_s \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ example showing invariant mass combinations as a function of $\epsilon^{\mathrm{Selection}}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements. 179

Acknowledgments

Firstly I should thank Warwick Experimental Physics department, in particular, my supervisors Tim Gershon and Michal Kreps for giving me this opportunity. Their endless support and wisdom has made my PhD one of my most enjoyable endeavours to date. When I finally win the lottery I will be sure get you a new sports car.

I am indebted to Thomas Latham, who made me feel welcome when I started at Warwick and has put up with all my questions over the years and helped me grow as a programmer and scientist.

To the friends I have made both at my institute; Mark, Rafael, Tom, Eugenia, Tomas, multiple Daniel's, Charlotte, Callum, Andy and Nicola, you're all a great eccentric bunch and I wouldn't of had it any other way. To the friends I made whilst on long term attachment at CERN, in particular; Daniel, Heinrich, Tomas (Pilar), Jack and other LTA crew, I owe you all for keeping me sane and providing hours of entertainment. I also thank my school friends who have encouraged and supported me and kept in touch.

I would like to thank the great people I have had the privilege of working with on the VELO project including Paula Collins, Lars Eklund, Martin van Beuzekom, Karol Hennessy, Eduardo Rodrigues, Silvia Borghi and Chris Parkes. You made my time on the collaboration exciting and interesting.

A massive thank you must go to my family to which I dedicate this thesis. Both my immediate family; my parents, sister, my cool little dude nephew Torin, my cousins, aunts and uncles; as well as my extended family; Glenys, Vaughan and Alex. Your love and kindness is what has got me where I am today and I thank you all. Finally, Lyndsey, I suppose you should get some special mention. Given

that I left you for two years in the UK whilst I went to do all of this at CERN, I should apologise. We are stronger than ever and I look forward to going on a new journey with you in hand. Bring on the next adventure although this time around we should stick together.

Declarations

The work presented in this thesis is all my own work, unless it is specifically referenced to the contrary. This thesis has not been submitted, in any form, to this or any other university for another qualification.

Abstract

This thesis presents the branching fractions measurements of $B^0_{(s)} \to J/\psi\, K^0_{\rm S} h^+ h^{(\prime)-}$ ($h^{(\prime)} = K, \pi$) decays, corresponding to an integrated luminosity of $1.0\,{\rm fb}^{-1}$ of data recorded at a centre-of-mass energy of $\sqrt{s} = 7\,{\rm TeV}$ collected with the LHCb detector at CERN Large Hadron Collider (LHC) during 2011. The first observation of the $B^0_s \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp$ decay is reported, with a significance greater than 10 standard deviations. The $B^0 \to J/\psi\, K^0_{\rm S} K^+ K^-$ decay is also observed for the first time. The branching fraction of $B^0 \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-$ is determined, to significantly better precision than previous measurements, using $B^0 \to J/\psi\, K^0_{\rm S}$ as a normalisation channel. Branching fractions and upper limits of the other $B^0_{(s)} \to J/\psi\, K^0_{\rm S} h^+ h^{(\prime)-}$ modes are determined relative to that of the $B^0 \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-$ decay. The findings of this thesis have been submitted to JHEP as detailed in Ref. [1].

1

Introduction

"In the beginning there was nothing, which exploded."

– Terry Pratchett, *Lords and Ladies*.

All the current averaged experimental measurements of *CP* violation in the quark sector are well described by the Cabibbo-Kobayashi-Maskawa mechanism [2, 3] which is embedded in the framework of the standard model (SM). However, it is believed that the size of *CP* violation in the SM is not sufficient to account for the asymmetry between matter and antimatter in the Universe [4]; hence, additional sources of *CP* violation are being searched for as manifestations of physics beyond the SM.

The measurement of the phase ϕ_s associated with B_s^0 oscillations is of fundamental interest (see, *e.g.*, Ref. [5] and references therein) due to the constraints it can provide on new physics. To date only the decays $B_s^0 \to J/\psi\,\phi$ [6, 7, 8, 9, 10], $B_s^0 \to J/\psi\,\pi^+\pi^-$ [11, 12] and $B_s^0 \to \phi\phi$ [13] have been used to measure ϕ_s. To maximise the sensitivity to effects of physics beyond the SM, which might couple preferentially to states with certain quantum numbers, it would be useful to study more decay processes. Decay channels involving J/ψ mesons are well-suited for such studies since the $J/\psi \to \mu^+\mu^-$ decay provides a clean trigger signature and allows good measurement of the vertex position. Observation of the decay $B_s^0 \to J/\psi\,\pi^+\pi^-\pi^+\pi^-$, with a significant contribution from $J/\psi\,f_1(1285)$ has recently been reported by LHCb [14]. Searches for other channels are well motivated. There are several unflavoured mesons, detailed in Table 1.1, that are known to decay to $K^0_{\rm S}K^\pm\pi^\mp$ [15], and that could in principle be produced in B_s^0 decays together with a J/ψ meson. If such decays are discovered, they could be used in future analyses to search for *CP* violation.

No measurements exist of the branching fractions of $B^0_{(s)} \to J/\psi\,K^0_{\rm S}K^\pm\pi^\mp$

Table 1.1: States below 1500 MeV/c^2 listed in the PDG [15] that decay to $K^0_S K^\pm \pi^\mp$. The $h_1(1380)$, listed as needing confirmation, but seen in $K\bar{K}\pi$, is not included.

State	$I^G\, J^{PC}$	Mass (MeV/c^2)	Width (MeV/c^2)
$a_1(1260)$	$1^-\, 1^{++}$	1230 ± 40	250 to 600
$f_1(1285)$	$0^+\, 1^{++}$	1281.8 ± 0.6	24.3 ± 1.1
$\eta(1405)$	$0^+\, 0^{-+}$	1409.8 ± 2.5	51.1 ± 3.4
$f_1(1420)$	$0^+\, 1^{++}$	1426.4 ± 0.9	54.9 ± 2.6
$\eta(1475)$	$0^+\, 0^{-+}$	1476 ± 4	85 ± 9

decays. The decays $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ [16, 17, 18] and $B^0 \to J/\psi\, K^0_S K^+ K^-$ [19, 20] have been previously studied, though the measurements of their branching fractions have large uncertainties. In addition to being potential sources of particle misidentification background to $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$, these decays allow studies of potential exotic charmonia states. For example, the decay chain $B^+ \to X(3872) K^+$ with $X(3872) \to J/\psi\, \pi^+\pi^-$ has been seen by several experiments [21, 22, 23], and it is of interest to see if production of the $X(3872)$ state in B^0 decays follows the expectation from isospin symmetry. Another claimed state, dubbed the $X(4140)$, has been seen in the decay chain $B^+ \to X(4140) K^+$, $X(4140) \to J/\psi\, \phi$ by some experiments [24, 25, 26] but not others [27], and further experimental studies are needed to understand if the structures in the $J/\psi\, \phi$ system in $B^+ \to J/\psi\, \phi K^+$ decays are of resonant nature or otherwise. In addition, it has been noted that the relative production of isoscalar mesons in association with a J/ψ particle in B^0 and B^0_s decays can provide a measurement of the mixing angle between $\frac{1}{\sqrt{2}} \left| u\bar{u} + d\bar{d} \right\rangle$ and $|s\bar{s}\rangle$ components of the meson's wavefunction [28, 29, 30], and therefore studies of $B^0_{(s)} \to J/\psi\, K^0_S K^\pm \pi^\mp$ may provide further insights into light meson spectroscopy.

In this thesis, the first search for B meson decays to $J/\psi\, K^0_S K^\pm \pi^\mp$ is reported. A summary of all previous attempts of branching fraction measurements related to this analysis is provided in Table 1.2. Possible decay modes are shown in the Feynman diagram in Figs. 1.1 for $B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$ and $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ respectively. All $J/\psi\, K^0_S h^+ h^{(\prime)-}$ final states are included in the analysis, where $h^{(\prime)} = K, \pi$. The analysis strategy is to reconstruct the decays with minimal bias on the phase-space of the decay, in order to be able to retain all possible resonant contributions in the relevant invariant mass distributions. In case contributions from broad resonances overlap an amplitude analysis will be necessary to resolve them. Such a study would require a dedicated analysis and is beyond the scope of this thesis.

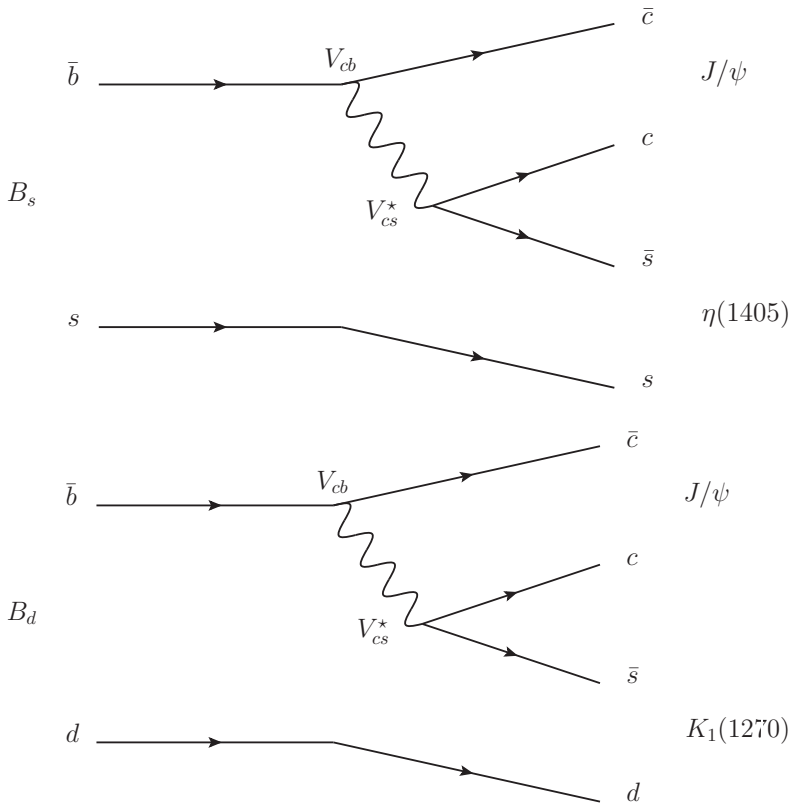

Figure 1.1: Feynman diagrams showing possible tree-level decays. (top) $B_s^0 \to J/\psi K_S^0 K^{\pm}\pi^{\mp}$ with $\eta(1405) \to KK\pi$. (bottom) $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ with $K_1(1270) \to K\rho$, $(K_0^*(1430) \to K\pi)\pi$.

Table 1.2: Previous measurements of $B^0_{(s)} \to J/\psi K^0_S h^+ h'^-$ decays [15].

Decay	Branching fraction (10^{-5})	Reference
	$B^0 \to J/\psi K^0_S \pi^+\pi^-$	
$B^0 \to J/\psi K^0_S \pi^+\pi^-$	$103 \pm 33 \pm 15$	CDF [16]
$B^0 \to J/\psi K_1(1270)^0$	$130 \pm 34 \pm 32$	Belle [17]
$B^0 \to J/\psi K^{*+}\pi^-$	$77 \pm 41 \pm 13$	CDF [16]
$B^0 \to X(3872) K^0_S$	$0.43 \pm 0.12 \pm 0.04$	Belle [18]
	$B^0 \to J/\psi K^0_S K^+ K^-$	
$B^0 \to J/\psi K^0_S \phi$	9.4 ± 2.6	BaBar [19], Belle [20]

2

Theory

2.1 Observational motivation

We begin with a basic assumption that at the instance the universe came into existence, affectionately known as the Big Bang, there was no asymmetry present, favouring matter over anti-matter. Yet here you are, reading this using your matter-assembled eyes. The modern view is that the observed excess of baryonic matter relative to antimatter is likely to be due to the properties of particles and their interactions during the expansion and cooling of the early universe. This mechanism is called *baryogenesis* since the matter excess is generated some time after the Big Bang. In 1967, Sakharov [31] described three necessary conditions:

1. The existence of baryon number violation.
2. *CP* symmetry violation.
3. A departure from thermal equilibrium.

If all three conditions are satisfied then this could lead to baryogenesis after the Big Bang, *i.e.*, the dynamic generation of a baryon asymmetry from an initially symmetric universe. The first of these is self explanatory which allows for an excess of baryons over antibaryons, the second is required otherwise the baryon asymmetry is washed out by charge and parity conjugate processes, and finally the departure from thermal equilibrium is needed else any baryon asymmetry can become overwhelmed by the reverse process.

Today one of the big unanswered questions is: why did the universe pick out matter over antimatter? Our current understanding of the ways in which particles interact is governed by the standard model (**SM**) of particle physics. It has held

Generation	Fermion name	Mass (MeV/c^2)	Electric charge
I	Electron	0.510998910(13)	-1
	Electron neutrino	0*	0
	Up quark	$2.5^{+0.6}_{-0.8}$	$2/3$
	Down quark	$5.0^{+0.7}_{-0.9}$	$-1/3$
II	Muon	105.658367(4)	-1
	Muon neutrino	0*	0
	Charm quark	1290^{+50}_{-110}	$2/3$
	Strange quark	100^{+30}_{-20}	$-1/3$
III	Tau	1776^{+600}_{-900}	-1
	Tau neutrino	0*	0
	Top quark	172900^{+600}_{-900}	$2/3$
	Bottom quark	4190^{+180}_{-60}	$-1/3$

Table 2.1: Summary of the fermions, organised by generation. Masses are taken from Ref [15]. The neutrino masses have been starred (*) since there have been significant observations that imply the masses are not in fact zero [34], however, as yet this is not included in the Standard Model.

up to years of experimental scrutiny, however, we know that it is not the full story. This chapter aims to highlight certain theoretical aspects of the SM that lead to naturally occurring *CP* violation and then discuss the tools available with which we can examine *CP* violation through dedicated experiments.

2.2 The standard model

The SM is a quantum field theory that unites three of the fundamental forces of nature under the symmetry group $SU(3)_C \otimes SU(2)_I \otimes U(1)_Y$ where $SU(3)_C$ is the colour symmetry of strong interactions, $SU(2)_I$ describes the weak isospin I for the unified electroweak interactions and $U(1)_Y$ the invariance under hypercharge Y transformations. Table 2.1 presents the group of particles called *fermions*: quarks and leptons with fractional spin. Table 2.2 summarises the gauge boson of three fundamental symmetries and the mediators of interactions in the standard model: strong, weak, electromagnetic as well as the more recent discoveries [32, 33] of a bosonic particle which, fits well the profile of the SM Higgs boson. Note the first and most prominent flaw in the SM is the lack of any description of gravity. We begin by expanding the SM Lagrangian to see how *CP* violation originates.

Boson name	**Associated force**	**Mass (GeV/c^2)**
Photon	Electromagnetic	0
Gluon	Strong nuclear	0
$W^{\pm}$	Weak nuclear	80.399 ± 0.023
Z^0	Weak nuclear	91.1876 ± 0.0021
H	Higgs	125.9 ± 0.4

Table 2.2: Summary of the standard model boson particles. Masses are taken from Ref [15].

2.2.1 The SM Lagrangian - Quantum Chromodynamics

Since we are interested in hadronic final states in this thesis, it is natural to discuss the forces that govern quark interactions. One additional characteristic of quark interactions is colour which is carried by gluons, and is refered to as quantum chromodynamics. The Lagrangian of the $\mathrm{SU}(3)_C$ component of the SM is given as

$$\mathcal{L}_{\mathrm{SU}(3)_C} = -\frac{1}{4} G^i_{\mu\nu} G^{i\mu\nu} + \bar{q}_{ra} i\gamma^\mu D^a_{\mu b} q^b_r \tag{2.1}$$

The first term leads to gluon self interactions with $G^i_{\mu\nu} = \partial_\mu G^i_\nu - \partial_\nu G^i_\mu - g_s f_{ijk} G^j_\mu G^k_\nu$ where g_s is the QCD gauge coupling constant and f_{ijk} is the structure constant determining the commutation relations between the generators of the group. The second term is the gauge covariant derivative for quarks (recalling that QCD acts only on the quark sector), where q_r is the r^{th} quark flavour and $a, b = 1, 2, 3$ provide the colour indices. The covariant derivative is of the form

$$D^a_{\mu b} = (D_\mu)_{ab} = \partial_\mu \delta_{ab} + i g_s G^i_\mu \frac{\lambda^i_{ab}}{2} \tag{2.2}$$

where the λ^i corresponds to the 8 Gell-Mann [35] matrices forming the generators of $\mathrm{SU}(3)_C$. The colour interactions can be seen above as diagonal along the flavour indices but can change the colour. From this point on in this thesis, we will omit the QCD components of the SM, albeit an important factor in theoretical predictions relating to flavour physics. In fact, there is a direct component in the SM relating to *CP* violation occurring naturally in the QCD sector, which is known as the *strong CP* problem [1] [36].

[1] This is a fine-tuning problem and stems from the QCD vacuum expectation whereby tunnelling from one state to another can give rise to *CP* violating terms that can only be removed by setting coefficients to zero, a process deemed unnatural by some in the theoretical community.

2.2.2 The SM Lagrangian - electroweak origins of CP violation

In the standard model it is the electroweak sector that gives rise to *CP* violation. These interactions are governed by a complete $\mathrm{SU}(2) \otimes \mathrm{U}(1)$ theory developed by Glashow-Weinberg-Salam [37, 38, 39] who were awarded the Nobel Prize in Physics 1979. The theorem states

$$\mathcal{L}_{\mathrm{SU(2)\otimes U(1)}} = \mathcal{L}_{\mathrm{Gauge}} + \mathcal{L}_{\mathrm{Matter}} + \mathcal{L}_{\mathrm{Yukawa}} + \mathcal{L}_{\mathrm{Higgs}} \tag{2.3}$$

The pure gauge interactions contain the kinetic energies of the gauge bosons and their self-interactions; the matter terms form the basis of the quark and lepton Dirac like terms; the Yukawa interactions give rise to the fermion masses; and finally the Higgs scalar term provides Higgs interactions and the mechanism under which masses are acquired. Expanding the terms of Eq. 2.3 in turn, we will discuss their importance. Starting with gauge bosons, the corresponding lagrangian is

$$\mathcal{L}_{\mathrm{Gauge}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.4}$$

here, the gauge fields are expanded to $W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu$ and $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ where g is the $\mathrm{SU}(2)_I$ coupling constant and ϵ_{ijk} is the anti-symmetric Levi-Civita matrix.

It is the gauging of the theory that leads to the quantised particles that we can observe. However, in order to think about interactions between particles we examine the matter term in Eq. 2.3

$$\begin{aligned}
\mathcal{L}_{\mathrm{Matter}} = {} & i\bar{Q^i}_{\mathrm{L}} \gamma^\mu \left(\partial_\mu + igW^i_\mu \frac{\sigma^i}{2} + iYg'B_\mu \right) Q^i_L \\
& + i\bar{q^i}_{\mathrm{R}} \gamma^\mu \left(\partial_\mu + iYg'B_\mu \right) q^i_R \\
& + i\bar{L^i}_{\mathrm{L}} \gamma^\mu \left(\partial_\mu + igW^i_\mu \frac{\sigma^i}{2} + iYg'B_\mu \right) L^i_L \\
& + i\bar{l^i}_{\mathrm{R}} \gamma^\mu \left(\partial_\mu + iYg'B_\mu \right) l^i_R
\end{aligned} \tag{2.5}$$

where again g is the $\mathrm{SU}(2)_I$ coupling constant and notice the parameter g' which is the $U(1)_Y$ coupling constant. The fermions are grouped into either left-handed doublets Q^i_L, L^i_L or right-handed singlets q^i_R, l^i_R for quarks and leptons respectively. The index i refers to the different flavour generations [2]. Massive vector gauge bosons

[2] For simplicity the QCD colour indices a, b have been dropped.

				$\mathrm{SU}(3)_C$	$\mathrm{SU}(2)_I$	$\mathrm{U}(1)_Y$
$Q_L^i =$	$\begin{pmatrix} u \\ d \end{pmatrix}_L$,	$\begin{pmatrix} c \\ s \end{pmatrix}_L$,	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	3	2	$\frac{1}{6}$
$q_R^i =$	u_R,	c_R,	t_R	$\bar{3}$	1	$\frac{2}{3}$
$q_R^i =$	d_R,	s_R,	b_R	$\bar{3}$	1	$-\frac{1}{3}$
$L_L^i =$	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$,	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$,	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	1	2	$-\frac{1}{2}$
$l_R^i =$	e_R,	μ_R,	τ_R	1	1	-1

Table 2.3: Summary of the standard model quantum numbers for quarks and leptons. The quantum numbers describe how each of the fermion types transforms under the different symmetry groups of the SM, $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$. They can be found directly in Ref [15].

are introduced which govern interactions between these matter particles and can be read off once fully expanding the covariant derivatives and transforming to a more convenient basis (which turn out to be the familiar W, Z and a massless A). We discuss these bosons in Section 2.2.3. To understand the quantum numbers assigned to these fermions we list them in Table 2.3. The left-handed and right-handed fermions have different charges under $\mathrm{U}(1)_Y$ which provide a consistent picture for their electric charges $Q = Y + T^3$, where $T^3 = \frac{\sigma^3}{2}$ is the property of weak isospin and acts on SU(2) doublets.

2.2.3 Higgs field, spontaneous symmetry breaking and gauge boson masses

Higgs [40] and Englert [41] were awarded the Nobel Prize in Physics 2013 for their work on "... the origin of mass of subatomic particles ...". Not only is it appropriately timed to mention this wonderful theory, but vital in our quest to generate *CP* violation in the SM. The Higgs field remains invariant under $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ transformations and is a weak isospin doublet made up of complex scalar fields, written in the following form

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4) \\ \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \end{pmatrix} \tag{2.6}$$

where ϕ^+ and ϕ^0 are complex scalar fields, which can be written in terms of four real scalar components, ϕ_i. The field itself carries quantum numbers with hypercharge

$Y = \frac{1}{2}$ thus the covariant gauge derivative is given by

$$D_\mu = \left(\partial_\mu - \frac{i}{2} g \sigma_i W_\mu^i - \frac{i}{2} g' B_\mu\right) \tag{2.7}$$

with the σ_i matrices being the usual Pauli $SU(2)$ generators and W_μ^i, B_μ the massless $SU(2)$ and $U(1)$ gauge bosons. The Higgs field ϕ also interacts with fermions with strength G_f. This is the Yukawa term in the Lagrangian, so the overall Lagrangian containing the Higgs field and a Yukawa coupling is

$$\mathcal{L} = \underbrace{-\frac{\mu^2}{2}\phi^\dagger\phi - \frac{\lambda}{4}\left(\phi^\dagger\phi\right)^2}_{\text{Higgs potential}} \underbrace{- G_f\left(L\phi R + R\phi^\dagger L\right)}_{\text{Yukawa interaction}} \tag{2.8}$$

where the μ^2 has units of mass squared, λ is a dimensionless parameter responsible for Higgs self interaction, G_f is the Fermi constant, L is a left-handed doublet and R a right-handed singlet in the Yukawa term. The potential term gives rise to the classic *Mexican hat* scenario when $\mu^2 < 0$ since we find degenerate minima about the trough of the hat and an unstable point in the middle at the peak. The potential is rotationally invariant in ϕ space. In terms of a classical analogy we can imagine placing a ball at the top of a pointed Mexican hat, and describe it by a classical Lagrangian. We can rotate the hat itself in ϕ-space and the ball remains invariant under $\mathrm{SU}(2)_I$ transformations, similarly we can apply a $\mathrm{U}(1)_Y$ transformation, which is essentially a rotation of the ball about its own axis by some local phase. The top of the hat is a metastable point and any movement of the ball would render it unstable causing the ball to roll down to the true vacuum, where it would oscillate about the trough of the hat.

The same principle can be applied to the Higgs such that quantum fluctuations cause the Higgs field to acquire a vacuum expectation. This "rolling" down from the peak of the hat picks out a direction in ϕ-space that breaks the $\mathrm{SU}(2)_I$ symmetry hence we can no longer rotate the hat and the "ball" remains where it is. It is down to convention that we can choose the isospin frame such that the lowest energy state occurs at $\frac{\partial V(\phi^\dagger\phi)}{\partial\phi^\dagger\phi}|_0 = -\frac{\mu^2}{\lambda}$. An arbitrary isospin frame is selected such that only one component of the complex field is real, $(\phi_1)|_0 = v = \pm\sqrt{\frac{2\mu^2}{\lambda}}$ and that the other components are $(\phi_2)|_0 = (\phi_3)|_0 = (\phi_4)|_0 = 0$. By expanding about the minima we find in general terms

$$\phi(x) = \exp\left\{i\frac{\sigma_i}{2}\theta^i(x)\right\}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.9}$$

where v is the vacuum expectation value and then four real scalar fields: $h(x)$ known as the Higgs field and $\theta^i(x)$ are Goldstone bosons. The crucial point is that the local $\mathrm{SU}(2)_I$ invariance of the Lagrangian allows us to rotate away any dependence on $\theta^i(x)$. Now we have a field which under spontaneous symmetry breaking changes from a $\mathrm{SU}(2)_I \otimes \mathrm{U}(1)_Y$ to a $\mathrm{U}(1)_Y$ symmetry: this is important for coupling the left-handed doublet and right-handed singlets as we shall see when examining the Yukawa Lagrangian. After spontaneous symmetry breaking, we recover the electroweak mass eigenstates. These emerge from the gauge term $(D_\mu\phi)^\dagger(D^\mu\phi)$ as shown below

$$D_\mu\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{i}{2}g\left(W_\mu^1 - iW_\mu^2\right)(v+h) \\ \partial_\mu h + \frac{i}{2}\left(g'B_\mu - gW_\mu^3\right) \end{pmatrix}, \tag{2.10}$$

hence

$$\begin{aligned}(D_\mu\phi)^\dagger (D^\mu\phi) = \frac{1}{2}(\partial_\mu h)^2 + \frac{(v+h)^2}{8}\left(g'B_\mu - gW_\mu^3\right)^2 \\ + \frac{g^2(v+h)^2}{8}\left[\left(W_\mu^1\right)^2 + \left(W_\mu^2\right)^2\right].\end{aligned} \tag{2.11}$$

Eq. 2.11 contains the couplings to the Higgs as well as the mass eigenstates of the vector bosons, terms involving v^2.

A suitable basis can be chosen such that we retrieve the Z^0, A_μ and $W^\pm$ bosons.

$$\begin{pmatrix} Z_\mu(x) \\ A_\mu(x) \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} W_\mu^2 \\ B_\mu \end{pmatrix} \quad \tan\theta_w \equiv \frac{g'}{g}$$
$$W^\pm(x) = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right)$$

where θ_w is the weak mixing angle. Physically, A_μ is the EM field, the massless quanta of which are photons, the $W^\pm$ are linear combinations of the $W^{1,2}$ fields that form the charged massive vector bosons with mass $M_W = \frac{gv}{2}$ and Z_μ is the neutral weak field whose quanta are Z^0 particles with mass $\sqrt{g^2+g'^2}\frac{v}{2} = \frac{M_W}{\cos\theta_w}$.

2.2.4 Yukawa and fermion masses

In order to generate fermion mass we need terms similar to those in the Dirac Lagrangian, i.e. we have

$$\mathcal{L}_{\text{Dirac}} = -m_f \left(\bar{f}_L f_R + \bar{f}_R f_L \right) \tag{2.12}$$

However, in the SM $\text{SU}(2)_I \otimes \text{U}(1)_Y$ representation, left-handed fermions are contained in doublets thus they transform differently from the right-handed singlets. It is clear that without some mechanism the SM Lagrangian cannot contain any mass terms. Mass terms can, however, be generated by couplings to the Higgs field. Since the Higgs acquires a vacuum expectation value (vev) and breaks $\text{SU}(2)_I$ symmetry, it couples together the left- and right-handed fermions to generate the much needed Dirac style mass terms in the Lagrangian

$$\mathcal{L}_{\text{Yukawa}} = -\underbrace{Y^d_{ij}\bar{Q}^i_L \phi d^j_R - Y^u_{ij}\epsilon^{ab}\bar{Q}^i_{La}\phi^*_b u^j_R}_{\text{Quark terms}} - Y^l_{ij}\bar{L}^i_L \phi e^j_R + h.c.\,, \tag{2.13}$$

where ϕ is the Higgs field, the quark couplings are given by $Y^{u,d}$ and the lepton sector via Y^l. Since we are not explicitly interested in leptons, we will ignore them but their treatment is almost identical. The Yukawa couplings provided here are complex and therefore not physical. In order to determine the physical quantities we must diagonalise the Yukawa coupling matrix $Y^{u,d}$. Writing the complex matrices in terms of unitary matrices and a diagonal matrix (of positive eigenvalues) and using a strategic change of basis for the left- and right-handed quark fields, we find

$$\left(Y^{u,d}\right)_{ij} = \left(V^{u,d}\right)_{ia}\left(D^{u,d}\right)_{ab}\left[\left(W^{u,d}\right)_{bj}\right]^\dagger, \tag{2.14}$$

$$u^i_L \to V^u_{ij}u^j_L \quad d^i_L \to V^d_{ij}d^j_L \quad u^i_R \to W^u_{ij}u^j_R \quad d^i_R \to W^d_{ij}d^j_R\,, \tag{2.15}$$

where both V, W are U(1) unitary matrices. We can utilise these two strategic changes of basis by applying Eq. 2.15 to the Yukawa Lagrangian in Eq. 2.13 to give

$$\mathcal{L}_{mass} = \underbrace{\frac{-v}{\sqrt{2}}\left[D^u_{ij}\bar{u}^i_L u^j_R + D^d_{ij}\bar{d}^i_L d^j_R\right]}_{\text{Fermion mass terms}} - \underbrace{\frac{h(x)}{\sqrt{2}}\left[D^u_{ij}\bar{u}^i_L u^j_R + D^d_{ij}\bar{d}^i_L d^j_R\right]}_{\text{Coupling to Higgs field}}. \tag{2.16}$$

This process generates physical mass terms in the SM that resemble the familiar Dirac masses. The mass eigenvalues that we try to deduce from measurements in our detectors are then $m^{u,d}_{ij} = \frac{v}{\sqrt{2}}D^{u,d}_{ij}$. We will find that upon closer inspection of

the charged currents through the weak interaction, that this choice of basis means that not all these matrices cancel out and, in fact, this is what gives rise to the CKM matrix.

2.3 The CKM matrix

In 1963, Nicola Cabibbo first came up with a 2×2 mixing matrix depending on only one real parameter (the Cabibbo angle), which therefore cannot account for *CP* violation [2]. It consisted of four quark flavours; up, down, strange and charm quarks. In 1973 the CKM matrix was presented by Kobayashi and Maskawa [3] who predicted a third generations of quarks; bottom and top. All of the quarks necessary for their model were not discovered until 1995 when the observation of the top quark [42, 43] was made. The charged current Lagrangian is found from Eq. 2.5 after insertion of the $W^{\pm}$ eigenstates

$$\mathcal{L}_{\mathrm{CC}} = \frac{g}{\sqrt{2}} \left(\bar{\nu}_L \gamma^\mu W^+_\mu e_L + \bar{u}_L \gamma^\mu W^+_\mu d_L + h.c. \right) , \tag{2.17}$$

where *h.c.* corresponds to the hermitian conjugate or W^- current. The choice of rotation W, V from Eq. 2.15, which diagonalises the Yukawa matrices, leaves additional terms in the charged currents that do not cancel as we have mixing of the u and d type unitary matrices $V^{u,d}$

$$\bar{u}^i_L \gamma^\mu d^i_L \rightarrow \bar{u}^a_L \gamma^\mu \underbrace{(V^u_{ia})^\dagger \left(V^d_{ib} \right)}_{\bar{V}^u_{ai} V^d_{ib} = V_{ab}} d^b_L \, , \tag{2.18}$$

where V_{ab} is known as the CKM matrix and is transforming the mass eigenstates of the down-type quarks, to the weak-interaction isospin partners of the up-type quarks (the weak eigenstates). Since the weak and mass eigenvalues are mixed, decays cross generations and we can use the CKM matrix to describe these mixing couplings as a direct test of the SM. For example, the weak eigenstate of a strange quark, s', is a mixture between down d, strange s and bottom b mass eigenstates such that $|s'\rangle = V_{cd}|d\rangle + V_{cs}|s\rangle + V_{cb}|b\rangle$. This is represented in matrix notation for all 3 generations of quarks in a 3×3 matrix as

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix} . \tag{2.19}$$

The elements of this matrix represent the transition of a quark of one type i to another type j with probability $|V_{ij}|^2$. The best known averages of the magnitudes of these elements are [15],

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000046}_{-0.000021} \end{pmatrix} . \tag{2.20}$$

These elements are not predicted by the SM therefore it is imperative that we obtain the most accurate result possible from our experiments. Transitions along the off-diagonals are known as "Cabbibo suppressed" as their values are much less than 1. Since the matrix is unitary it follows that

$$V_{ij}V_{ik}^* = \delta^{jk} \quad V_{ij}V_{kj}^* = \delta_{ik} \, . \tag{2.21}$$

Because the unitarity condition holds we end up with twelve relations; six that equal one and six degenerate relations that are equal to zero. One of the most famous of these relations that equal zero is the one we call "*the unitarity triangle*" for which

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \, . \tag{2.22}$$

This is simply a representation of a triangle in the complex plane and can be discussed in terms of two free parameters, $(\bar{\rho}, \bar{\eta})$, which come about by dividing the triangle by one of its sides such that one side is always of unit length which shifts the the focus to the other two intersections. The Jarlskog invariant [44], is a useful measure of CP violation and is characterised by half of the total area of any given triangle and is currently $J = \Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = (2.96^{+0.20}_{-0.16}) \times 10^{-5}$.

2.4 Unitarity triangles

In addition to the relation in Equation 2.22, there are five other triangles. We have a tendency to discuss two of them more than the others since the relative angles inside the triangles are large and of the same order making measurements easier. In total there are six triangles that are interrelated since each triangle shares its sides with one of the other five triangles. Three angles that have historically received the most attention and are related to the *unitarity triangle* from Eq. 2.22 are α, β, γ

defined as

$$\begin{aligned} \alpha &= \arg\left(-V_{td}V_{tb}^{*}/V_{ud}V_{ub}^{*}\right) , \\ \beta &= \arg\left(-V_{cd}V_{cb}^{*}/V_{td}V_{tb}^{*}\right) , \\ \gamma &= \arg\left(-V_{ud}V_{ub}^{*}/V_{cd}V_{cb}^{*}\right) . \end{aligned} \tag{2.23}$$

Note that the angles α, β, γ all sum to 180° by definition. Another related angle that could be targeted with this analysis is $\beta_s = \arg\left(-V_{ts}V_{us}^{*}/V_{td}V_{ud}^{*}\right)$.

The angles and lengths of the sides of the CKM matrix can all be measured directly or indirectly by various processes. Figure 2.1 shows the current constraints coming from experiments to date. An example of such a measurement is the angle β which is determined through the interference between mixing and decay of neutral B^0 mesons.

2.5 Why is CP violation important?

CP violation is a fascinating and elusive characteristic that we believe must be present since matter dominates over antimatter in the universe. The standard model allows and experiment has confirmed its existence, however, at a level that it is far, far too weak to explain baryogenesis. In fact, it is around seven orders of magnitude [4] from what is observed if the baryon asymmetry is considered relative to the number of photons in the universe.

There must be new physics out there that introduces additional *CP* violation. And there are some more fundamental philosophical questions to ask, such as why do we have three generations in the SM? Why are there not more? In fact, the theory of the SM in no way attempts to explain where *CP* violation comes from. Using the CKM matrix we can make remarkable predictions on the branching fractions of various decays and estimate relative amounts of *CP* violation, but it bestows nothing on the underlying mechanism causing it, leaving us to question why, in fact, do the generations mix at all?

We have shown that by some "cute" mathematical tricks , the diagonalisation of the Yukawa couplings gives rise to the fermion masses, which, in turn, leads to the CKM matrix in the interactions with the W boson. This does not really explain why *CP* violation occurs at all. Why do all the quarks have such different masses? The standard model may not be forthcoming in all its intricate detail, but we can still probe as far as possible to see when and where it breaks down, which will ultimately drive the discovery of new physics and the theories that come with them.

In general there are three methods for testing *CP* violation in the SM through experiment which we discuss next (see [46] for a comprehensive treatment).

2.5.1 *CP* violation in decay

CP violation in decay (historically known as direct *CP* violation) occurs when the decay amplitude of a particular interaction is different to that of the conjugate process (i.e. $\Gamma[X \to f] \neq \Gamma[\bar{X} \to \bar{f}]$). We can consider the decay amplitudes of such a process as follows

$$\begin{aligned} A_f &= \langle f|H|X\rangle = \sum_j A_j e^{i(\delta_j - \phi_j)}\,, \\ \bar{A}_{\bar{f}} &= \langle \bar{f}|H|\bar{X}\rangle = \sum_j A_j e^{i(\delta_j + \phi_j)}\,, \end{aligned} \tag{2.24}$$

where H denotes the Hamiltonian of the governing process, $A_j \leftrightarrow \bar{A}_{\bar{j}}$ are the *CP* conjugate decays of one another and δ_j, and ϕ_j are the strong (or *CP*-conserving) phase related to final state interactions and the weak (or *CP*-violating) phase relating to the phases of the weak interacting quark mixing matrix of a contributing process. Weak and strong phases do not necessarily refer to the weak and strong forces, instead a weak phase is defined as a phase that changes sign under *CP* and a strong phase does not. A more convenient variable is the ratio of the amplitudes

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{\sum_j A_j e^{i(\delta_j+\phi_j)}}{\sum_j A_j e^{i(\delta_j-\phi_j)}}\right| \tag{2.25}$$

CP violation in decay can be observed if the decay proceeds via at least two different decay mechanisms which have different weak and strong phases, resulting in interference which, in turn results in different decay rates. The condition for *CP* violation in decays is thus true if, and only if,

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1\,, \tag{2.26}$$

which happens when two or more amplitudes have different weak and strong phases. The difference between the decay of matter and anti-matter particles is called *direct CP violation* and is the only means for *CP* violation in charged baryon and meson systems. A prime example can be seen with two recent LHCb measurements $B^\pm \to K^\pm\pi^+\pi^-$ and $B^\pm \to K^\pm K^+ K^-$ decays [47] and $B^\pm \to K^+K^-\pi^\pm$ and $B^\pm \to \pi^+\pi^-\pi^\pm$ decays [48], where the significance of the $\mathcal{A}^{CP}(B^\pm \to K^\pm K^+ K^-)$ mode for

instance, exceeds three standard deviations and is the first evidence of an inclusive *CP* asymmetry in charmless three-body B decays.

2.5.2 *CP* violation - mixing

The other type of *CP* violating decay is induced by interference between mixing and decay (historically called indirect *CP* violation). The first part occurs through mixing where a neutral particle oscillates into its conjugate state (i.e. the probability of $X^0 \to \bar{X}^0$ is different to that of $\bar{X}^0 \to X^0$) and the second when interference occurs between both mixing and the decay amplitude, again leading to *CP* violation. Experimental measurements of the parameters associated with oscillations are an important test of the SM, since energy scales several orders of magnitude greater than the collider energy can be probed by means of virtual particles entering in loop diagrams. This provides a sensitive probe of new physics. Over time the number of $X^0 \to \bar{X}^0$ transitions may be fewer than the converse process.

In the SM, neutral flavoured meson [3] pairs ($K^0/\bar{K}^0$, $D^0/\bar{D}^0$, $B^0/\bar{B}^0$ and $B_s^0/\bar{B}_s^0$), can be thought of as decaying two component quantum states. Since the formalism is the same for all pairs we will simply denote them by $X^0/\bar{X}^0$. *CP* violation in mixing occurs via the exchange of two charged weak currents through box diagrams. The time evolution of the X^0-meson system is governed by the Schrödinger equation defined by

$$i\frac{d}{dt}\begin{pmatrix} |X^0(t)\rangle \\ |\bar{X}^0(t)\rangle \end{pmatrix} = \left(\mathbf{M} + \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} |X^0(t)\rangle \\ |\bar{X}^0(t)\rangle \end{pmatrix}, \tag{2.27}$$

where $|X^0(t)\rangle$ and $|\bar{X}^0(t)\rangle$ are the flavour (*i.e.* strong) eigenstates and $\mathbf{M}$ and $\mathbf{\Gamma}$ are 2×2 hermitian mass and decay mixing matrices. It is the off-diagonal elements that give rise to mixing since the CPT theorem requires that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$, or more simply put, particle and antiparticle have identical masses and lifetimes. The light and heavy physical mass eigenstates X_L and X_H respectively can be written at time $t = 0$ as

$$\begin{aligned} |X_L\rangle &= p|X\rangle + q|\bar{X}\rangle\,, \\ |X_H\rangle &= p|X\rangle - q|\bar{X}\rangle\,, \end{aligned} \tag{2.28}$$

where p and q represent the amount of mixing and are complex numbers that are

[3] Charged mesons are forbidden to mix due to conservation of charge and hence exhibit only direct *CP* violation.

normalised to unity via the relation $|p|^2 + |q|^2 = 1$. By diagonalising the effective Hamiltonian $\mathbb{H} = \mathbf{M} + \frac{i}{2}\mathbf{\Gamma}$, we obtain eigenvalues

$$\lambda_{L,H} = M_{L,H} - \frac{i}{2}\Gamma_{L,H} = M_{11} - \frac{i}{2}\Gamma_{11} \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \quad (2.29)$$

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}, \quad (2.30)$$

Given an initially pure state of X^0 or $\bar{X}^0$ the system will evolve as

$$\begin{aligned} |X(t)\rangle &= g_+(t)|X\rangle + \frac{q}{p}g_-(t)|\bar{X}\rangle \\ |\bar{X}(t)\rangle &= \frac{p}{q}g_-(t)|X\rangle + g_+(t)|\bar{X}\rangle \end{aligned}$$

$$g_\pm(t) = \frac{1}{2}\left(e^{-i\left(M_L - \frac{i}{2}\Gamma_L\right)t} \pm e^{-i\left(M_H - \frac{i}{2}\Gamma_H\right)t}\right) = \frac{1}{2}e^{-Mt}\left(e^{\frac{i}{2}\Delta m t - \frac{1}{2}\Gamma_L} \pm e^{-\frac{i}{2}\Delta m t - \frac{1}{2}\Gamma_H}\right),$$

where we have expressed the final answer in terms of the effective mass $M = (M_H + M_L)/2$ and the mass width $\Delta m = M_H - M_L$. The mass difference is also a useful tool for comparing experiment with theory, as well as the lifetime difference $\Delta\Gamma$ defined below

$$\Delta m \equiv M_H - M_L = 2|M_{12}|\left(1 - \frac{1}{8}\frac{|\Gamma_{12}|^2}{|M_{12}|^2}\sin^2\phi_M\right) \quad (2.31)$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L = 2|\Gamma_{12}|\cos\phi\left(1 - \frac{1}{8}\frac{|\Gamma_{12}|^2}{|M_{12}|^2}\sin^2\phi_M\right) \quad (2.32)$$

where ϕ is the relative phase. As M_{12} corresponds to the virtual part of the box diagrams, measurement of Δm is a sensitive probe of new physics. The decay constants $\Gamma_{L,H}$, also provide the different lifetimes of the meson, where the effective lifetime is $\tau_s = 1/\Gamma_s = 2/(\Gamma_L + \Gamma_H)$. In the limit of small $\frac{|\Gamma_{12}|^2}{|M_{12}|^2}$ we have three identities

$$\Delta m = 2|M_{12}|, \quad \Delta\Gamma = 2|\Gamma_{12}|\cos\phi, \quad \phi = arg\left(-\frac{M_{12}}{\Gamma_{12}}\right). \quad (2.33)$$

Given that the time dependent decay rate is $\Gamma_{X^0\to f}(t) = \mathcal{N}|\langle f|H|X^0\rangle|^2$, where $\mathcal{N}$ is a time independent normalisation factor, we can use this information and find an extension to the mixing phenomenology defined above as we can have additional interference through decays to some final state f. In general one can consider four possible decay amplitudes of meson X and $\bar{X}$ to some final state f or $\bar{f}$. There are

a set of master equations that govern neutral meson decay rates

$$\Gamma_{X^0\to f}(t) = \mathcal{N}|A_f|^2\left(1+|\lambda_f|^2\right)\frac{e^{-\Gamma_S t}}{2} \left(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t + C_f\cos\Delta m t - S_f\sin\Delta m t\right), \quad (2.34)$$

$$\Gamma_{\bar{X}^0\to f}(t) = \mathcal{N}|A_f|^2\left|\frac{p}{q}\right|^2\left(1+|\lambda_f|^2\right)\frac{e^{-\Gamma_S t}}{2} \left(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t - C_f\cos\Delta m t - S_f\sin\Delta m t\right), \quad (2.35)$$

with

$$D_f = \frac{2\Re\lambda_f}{1+|\lambda_f|^2}, \quad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad S_f = \frac{2\Im\lambda_f}{1+|\lambda_f|^2}, \quad (2.36)$$

the parameter λ_f is used to parameterise this kind of interference. Once known we can fully describe the decay of the neutral particle and is defined by

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \quad (2.37)$$

The parameters outlined in Eq. 2.35 correspond to different types of *CP* violating effects in such neutral meson decays. The three letters represent

- D_f – mixing through oscillation.
- C_f – direct *CP* violating component. If the decay amplitudes and q/p are different this parameter effects the total decay rate.
- S_f – mixing without oscillation, generated by interference from cross terms.

2.5.3 *CP* violation - between mixing and decay

A particle X^0 and anti-particle $\bar{X}^0$ have the ability to decay to the same final state f. In this scenario we get a combination of mixing followed by decay interference. The governing parameter λ_f becomes

$$\lambda_{f_{CP}} = \frac{q}{p}\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}}, \quad (2.38)$$

where the decay amplitudes are as follows

$$A_{f_{CP}} = \langle f_{CP}|H|X^0\rangle,$$
$$\bar{A}_{f_{CP}} = \langle f_{CP}|H|\bar{X}^0\rangle,$$

with $A_{f_{CP}}$ and $\bar{A}_{f_{CP}}$ representing the amplitudes for $X^0 \to f_{CP}$ and $\bar{X}^0 \to f_{CP}$, p and q having the usual definition already presented. *CP* is violated when $\Gamma_{X^0 \to \bar{X}^0 \to f}(t) \neq \Gamma_{\bar{X}^0 \to X^0 \to f}(t)$. The asymmetry between the decay widths is given as

$$A_{f_{CP}}(t) = \frac{\Gamma_{X^0 \to f_{CP}} - \Gamma_{\bar{X}^0 \to f_{CP}}}{\Gamma_{X^0 \to f_{CP}} + \Gamma_{\bar{X}^0 \to f_{CP}}} = \frac{2C_{f_{CP}} \cos \Delta m t - 2S_{f_{CP}} \sin \Delta m t}{2 \cosh \frac{1}{2}\Delta\Gamma t + 2D_{f_{CP}} \sinh \frac{1}{2}\Delta\Gamma t} \tag{2.39}$$

Even if there is no *CP* present in either mixing or decay separately for the meson, $|\lambda_{f_{CP}}| = 1$, and that $\frac{q}{p} = 1$, a phase difference between the decay and mixing amplitudes can exist thus still permitting *CP* violation if $\Im(\lambda_{f_{CP}}) \neq 0$, since $D_f = \Re\lambda_{f_{CP}}$, $C_{f_{CP}} = 0$ and $S_{f_{CP}} = \Im\lambda_{f_{CP}}$. In this situation the time-dependent *CP* asymmetry can be defined as:

$$A_{f_{CP}} = \Im(\lambda_{f_{CP}}) \sin(\Delta m t)\,. \tag{2.40}$$

This kind of measurement has been made by BaBar, Belle and LHCb, see [49, 50, 51] respectively.

2.5.4 Measurement of β_s

The experimental values for β_s still have large uncertainties (by approximately a factor two [52]). Analysis of modes such as $B_s^0 \to J/\psi\, K_S^0 K^\pm \pi^\mp$ offer the chance to improve upon these measurements. A measurement of β_s could be made with the $B_s^0 \to J/\psi\, K_S^0 K^\pm \pi^\mp$ channel in this analysis. This, however, would be dependent on several things, such as the requirement for a large number of events since one would need to do a time-dependent (tagged) analysis and angular analysis to differentiate *CP*-even and *CP*-odd states. A recent LHCb measurement obtained from $1\,\text{fb}^{-1}$ data carried out an analysis on the $J/\psi\,\pi\pi$ decay and recorded a measurement of $\beta_s = 0.001 \pm 0.044$ [53]. This is consistent with the SM, yet within sizeable uncertainties. This uncertainty is still more than twice the SM central value and 40 times the standard model uncertainty; thus a lot will be learned from higher precision measurements in the future.

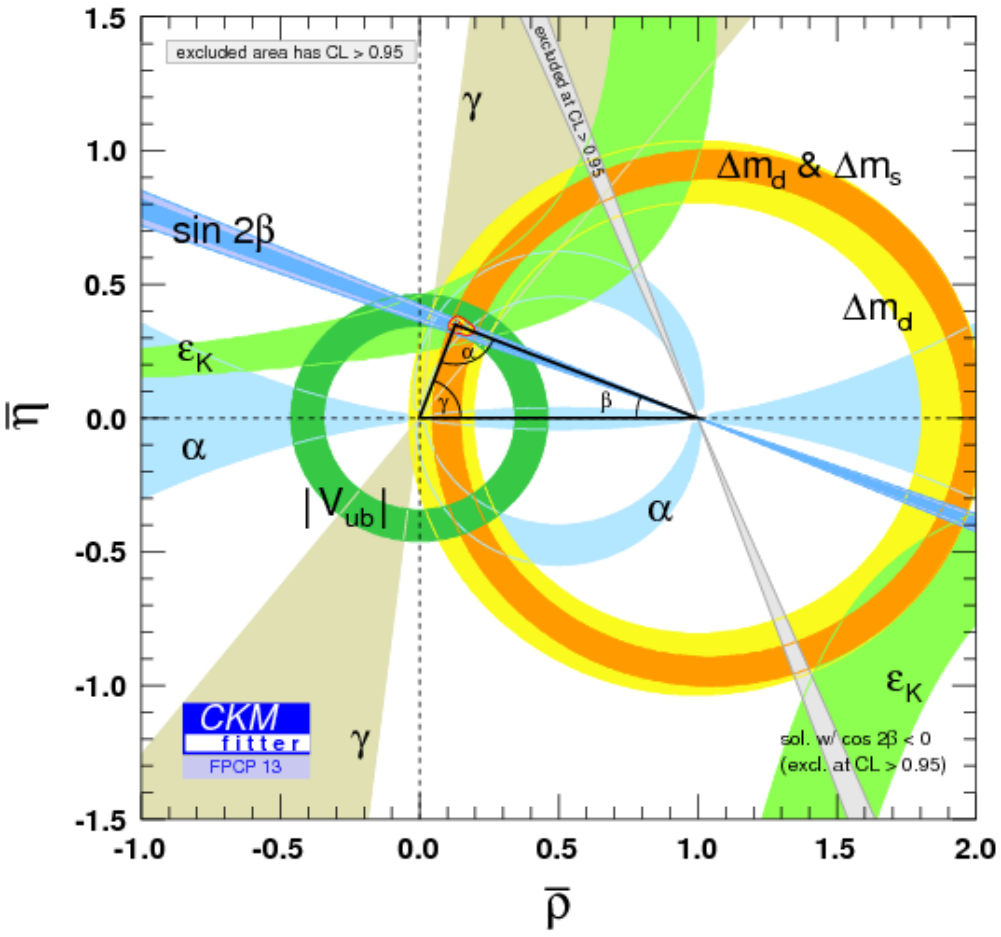

Figure 2.1: Experimental constraints on the angles α, β, γ angles the latest combined results in the $(\bar{\rho}, \bar{\eta})$ plane of the unitary triangle. Figure taken from Ref. [45].

3

Experimental overview

LHCb is the only experiment at the LHC that is dedicated to the study of heavy flavour physics. This chapter will provide a brief description of the LHC ring and an overview of the LHCb detector.

3.1 The LHC

The Large Hadron Collider (LHC) is a 27 km circumference synchrotron located beneath the Franco-Swiss border and is the most powerful particle accelerator built to date. It operates by colliding proton beams head on, where the design centre of mass energy is 14 TeV with each proton beam carrying 7 TeV. For safety reasons the LHC has not yet been run at its designed energy, instead the centre of mass energy has been carefully ramped up in stages. Thus there have been several datasets taken by the collaborations based at CERN. The first main sample was taken during 2011 where each beam had an energy of 3.5 TeV and the second during 2012 where the beam energy was raised to 4 TeV (small data samples were also taken in 2009, 2010 and 2013).

Protons are ideal to accelerate to high energies because they have a large mass, and therefore emit almost negligible synchrotron radiation. However proton collisions are extremely messy due to QCD interactions [54]. This is in contrast to previous generation B-factories such as e^+-e^- colliders at BaBar [55] and Belle [56]. At the LHC the proton source is simply a bottle of hydrogen gas and an electric field is used to strip an electron from the hydrogen atom to yield protons. The first accelerator in the chain is the Linear Accelerator (LINAC 2), which takes the protons from rest in a straight line to an energy of 50 MeV. This beam is then injected into the Proton Synchrotron Booster (PSB), which accelerates the protons

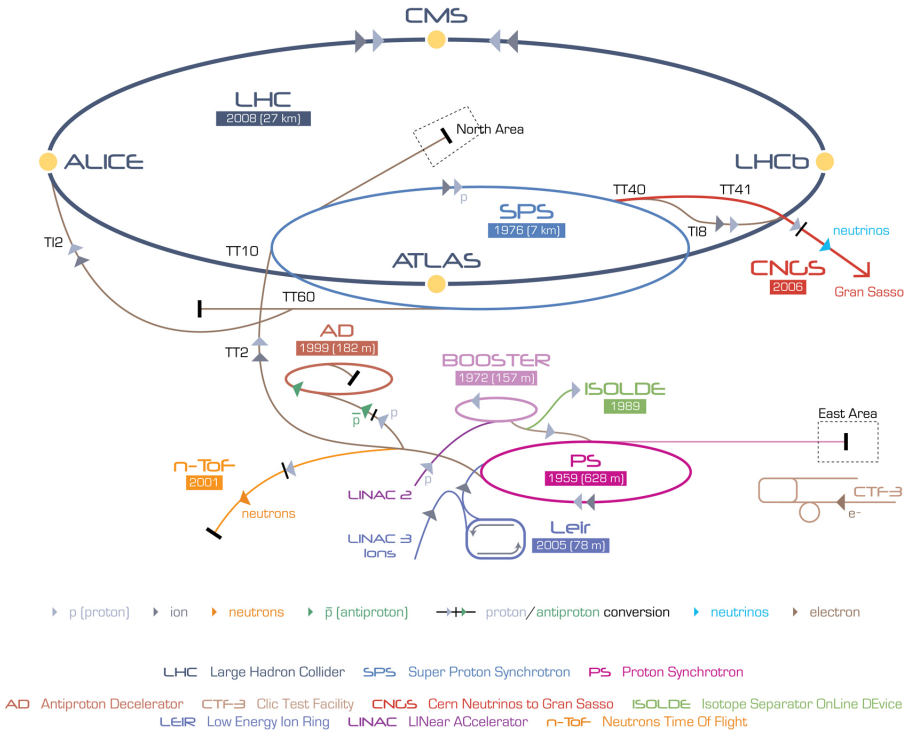

Figure 3.1: The CERN accelerator complex (not to scale). Figure taken from [57].

to 1.4 GeV, followed by the Proton Synchrotron (PS), which pushes the beam to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS) where they are accelerated further to 450 GeV. This is where the "friendly" handshake with the LHC ring is initiated and the protons can be injected safely into the LHC beam pipes which circulate two beams of protons (and occasionally heavy ions) in opposite directions (see figure 3.1 for a schematic of the CERN complex). Once inside the LHC storage ring, the beams are accelerated up to their maximum energy and then squeezed to form stable orbits that collide at various points where detectors are stationed. The two beams collide inside four detectors – ALICE, ATLAS, CMS and LHCb.

This analysis uses the data collected during 2011 running where the beam energy was 3.5 TeV. Each "fill" of the LHC ring can remain stable and collide on average for at least 8 hours under normal operating conditions, providing long periods of data taking. During 2011, ATLAS and CMS collected $5\,\text{fb}^{-1}$ of integrated luminosity whereas LHCb collected just over $1\,\text{fb}^{-1}$ for the same period. In contrast to the general purpose detectors ATLAS and CMS, LHCb is not designed to run at the maximum LHC luminosity of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. Instead the beams undergo a process called luminosity levelling whereby they are mis-aligned slightly in the y-

axis, to obtain the design luminosity of $2 - 5 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ remaining constant over the entire LHC fill. This has been chosen to limit the number of pp interactions occurring per bunch crossing, a term known as "pile-up", making it easier to trace primary and secondary vertices as well as limiting the radiation damage and detector occupancy. The LHC machine was briefly shutdown for upgrades in November 2011 and restarted in March 2012, with a beam energy of 4 TeV.

3.2 LHCb experiment

LHCb has been designed with several key aspects in mind:

1. Separation of primary and secondary vertices – this is paramount since the $B_s^0 - \overline{B}_s^0$ oscillation frequency is fast, necessitating a proper decay time resolution of the order $\mathcal{O}(50\,\mathrm{fs})$ for time-dependent CP studies. See section 3.3.1.

2. Excellent momentum resolution $\delta p/p \approx 0.4 - 0.6\,\%$ – this provides good invariant mass resolution enabling us to distinguish signal decays from background, which means the momentum of the tracks must be well measured. See section 3.3.2.

3. Particle identification capability (PID) – distinction between $e^{\pm}, \mu^{\pm}, K^{\pm}, \pi^{\pm}$ and $p(\overline{p})$ as many heavy flavour decays have similar topologies. See section 3.4

4. Trigger – a pp collision leads to high multiplicity events and thus LHCb must filter events efficiently without incurring deadtime for the next bunch crossing. Decay channels of interest must be kept with the highest efficiency possible. See section 3.5.

The LHCb detector [58] is located approximately 100 m underground. LHCb can be thought of as a general purpose detector in the forward direction.[1] It is the only single-arm spectrometer experiment at the LHC with an angular coverage from $\approx$ 15 mrad to 300(250) mrad in the bending (non-bending) plane. It has been designed specifically to take advantage of the fact that $b\bar{b}$ pairs are boosted in the direction of the higher energy parton in the laboratory rest frame, resulting in highly correlated b and $\bar{b}$ flight directions. To capitalise on this, the LHCb detector has been designed to capture one of these $b\bar{b}$ pairs in the forward direction only, hence its projective shape along the z-axis shown in figure 3.2. The $b\bar{b}$ cross-section for pp collisions is $\sigma(pp \to b\bar{b}X) = (75.3 \pm 5.4 \pm 13.0)\,\mu\mathrm{b}$ as measured by LHCb [59]

[1] In fact $b\bar{b}$ pairs are produced in the forward and backward direction, but due to cost, space and timing only pairs in one direction could be fully exploited.

for a centre of mass energy $\sqrt{s} = 7\,\text{TeV}$. Given a reasonable efficiency, the LHC experiments can quickly accumulate the world's largest dataset of B meson decays (after hadronising) making the LHC an ideal arena for their study. LHCb uses a

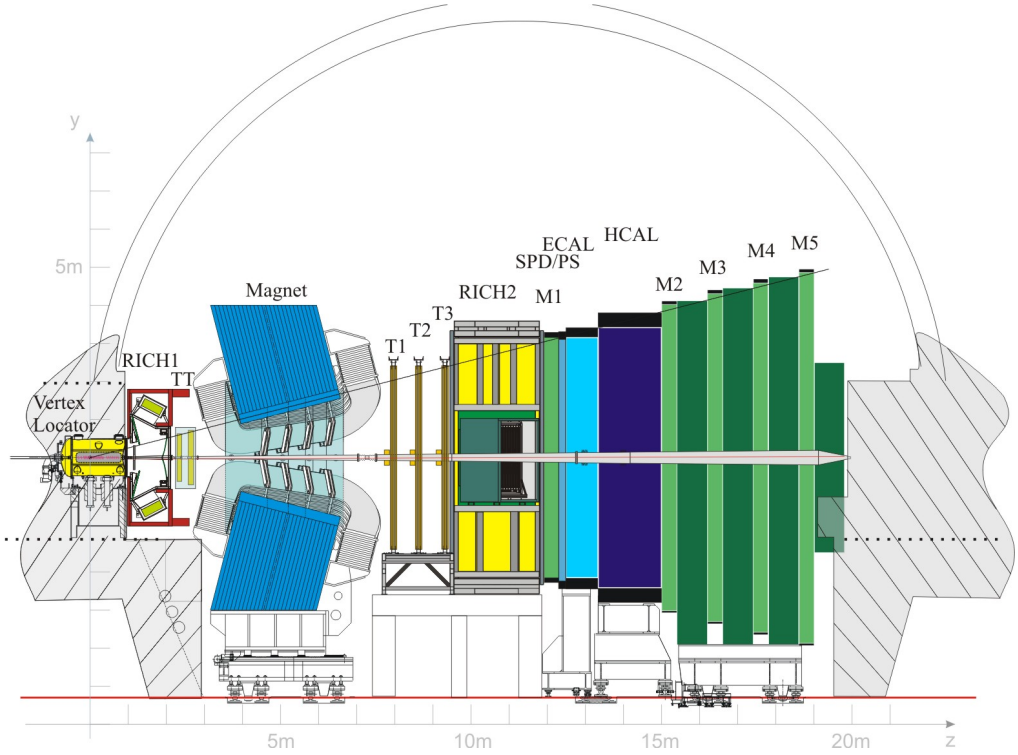

Figure 3.2: Schematic view of the LHCb detector in yz projection made at $x = 0$. Figure taken from [58].

right-handed Cartesian coordinate system. The z-axis runs through the beam pipe, pointing from the interaction region towards the muon detectors at the far end of the detector, with positive y defined in the vertical direction. The detector region before the magnet and near to the proton interaction region is referred to as *upstream*, and particles travelling after the magnet towards the other LHCb sub-detectors are said to travel *downstream*. Collisions occur at approximately $z = 0$, which is known as the primary vertex (PV). The PV is surrounded by the VErtex LOcator (or VELO) which measures the positions of charged particles.

Travelling along the detector in increasing $+z$ is the first of two RICH stations used to determine the velocity of particles and to distinguish between different particles (π,K, *etc.*). The RICH1 station is located immediately after the VELO and after RICH1 is the Tracker Turicensis (TT), which is positioned before the magnet. After the magnet are three tracking stations (T1-T3), which are positioned just before RICH2. Finally, after RICH2, are the calorimeter and muon systems, which

extend from $z = 12\,\mathrm{m}$ onwards.

The peak luminosity for 2012 was of the order $70 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at which point ATLAS and CMS recorded $23\,\mathrm{fb}^{-1}$ and LHCb recorded more than $2\,\mathrm{fb}^{-1}$ at a instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. LHCb was designed to read out events at a rate of 40 MHz, corresponding to the nominal bunch spacing inside the LHC.

3.3 Tracking

Tracking is of the utmost importance in precision heavy flavour physics. As mentioned we need to be able to distinguish decays from background, for which excellent vertex resolution is necessary and be able to measure their lifetime. The tracking

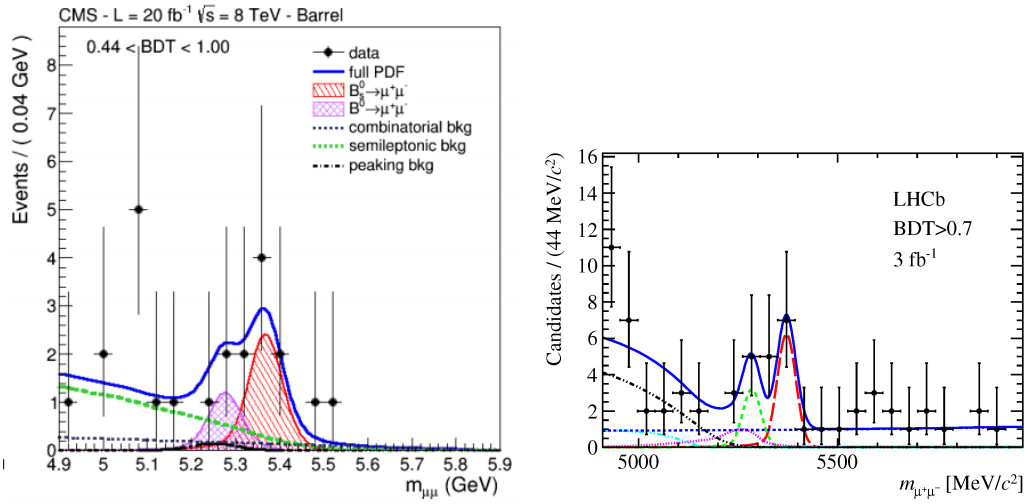

Figure 3.3: Results for fits to the 2011+2012 datasets for the rare $B \to \mu\mu$ decays by CMS [60] (left) and LHCb [61] (right).

detectors are required to have excellent spatial resolution in order to efficiently reconstruct particle trajectories (in particular by keeping the amount of material to a minimum) and keep the rate of fake track reconstruction as low as possible. The tracking system at LHCb consists of the VELO, the magnet and four main tracking stations; the Tracker Turicensis (TT) station placed before the magnet and three T stations (T1, T2 and T3) placed downstream of the magnet. A nice comparison between general purpose detectors such as CMS or ATLAS can be seen when comparing the recent (and exciting) measurements for $B_s^0 \to \mu\mu$ decays with those made at the LHCb experiment. From figure 3.3 we can see that the width of the signal peaks seen at LHCb are much narrower when compared to those made at CMS.

This highlights why LHCb is known as a precision flavour physics experiment, part of which can be attributed to the VELO, which is discussed next.

3.3.1 Tracking - VErtex LOcator (VELO)

The VELO [62] is required to have exceptional spatial resolution and a sufficient number of sensors to fully reconstruct a track trajectory whilst minimising the material budget. The VELO is rather different when compared with other vertex detectors at the LHC, and is built in two halves placed either side of the beam pipe. Most of the tracks reconstructed at LHCb belong to particles that originate from the primary vertex, *i.e.* they are prompt. Due to the high Lorentz boost, B mesons travel about 1 cm before decaying to form a secondary vertex. We refer to the displaced secondary vertices of long-lived particles, such as a B meson, using the nomenclature SV. The VELO is essential to the physics programme at LHCb since it is vital to the reconstruction of vertices lending itself to the identification of decay topologies and proper decay time information. One important feature of

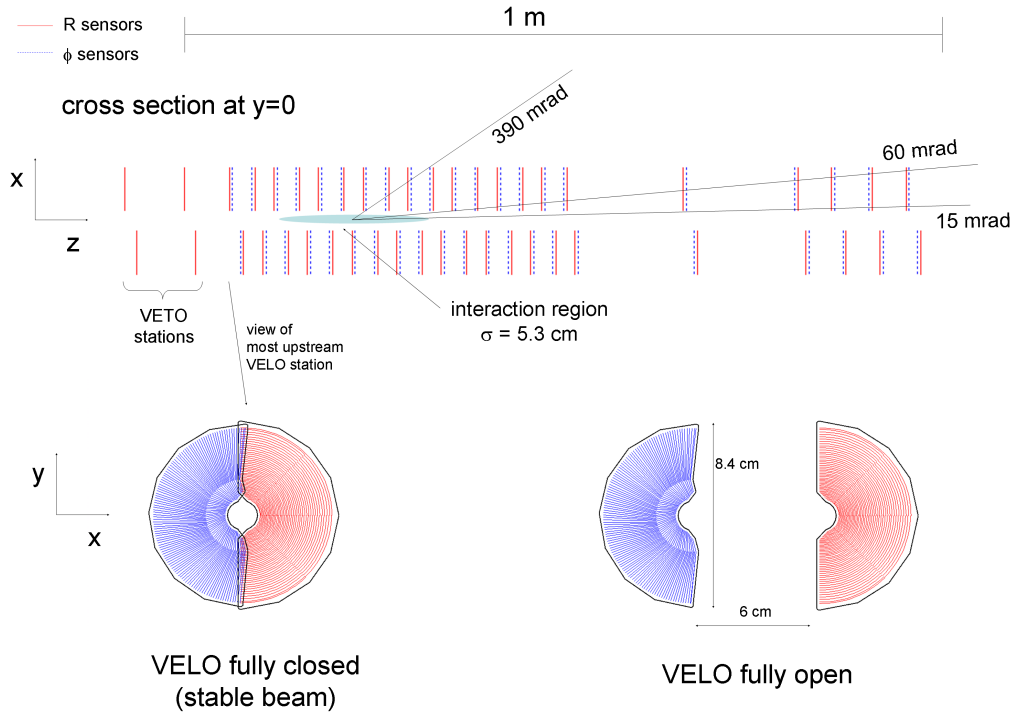

Figure 3.4: A diagram of the VELO stations along the xz-plane, indicating the angular acceptance and the arrangement of the stations. The diagram shows a VELO station in both the open and closed position in the xy-plane. Taken from Ref. [58].

the VELO is that the impact parameter (IP), defined as the distance of closest approach between the particle trajectory and the PV, has extremely good resolution thus allowing one to discriminate from tracks that do not originate from the PV.

The geometry of the VELO consists of two halves each containing 21 modules where each module is made up of two silicon half discs with strips in either radial, R, or polar, ϕ coordinates. When the two silcon sensors are combined they lead to a (x, y, z)-position in space. Each half is slightly offset from one another as shown in figure 3.4. Both the R and ϕ sensors are made from 300 µm thick silicon and span 7 mm < R < 44 mm with active area from 8.2 mm $\rightarrow$ 43 mm and each module providing 182° angular coverage. A sketch of the $R\phi$ geometry of the sensors can be seen in figure 3.5. The close proximity to the beam minimises the extrapolation distance from the PV to the first measurement allowing better resolution and IP discrimination. All but two of the sensors employ n^+-on-n silicon technology (meaning n-doped silicon implants in n-doped silicon bulk), and the remaining two use n^+-on-p technology, utilising p-type bulk. The VELO is approximately 1 m

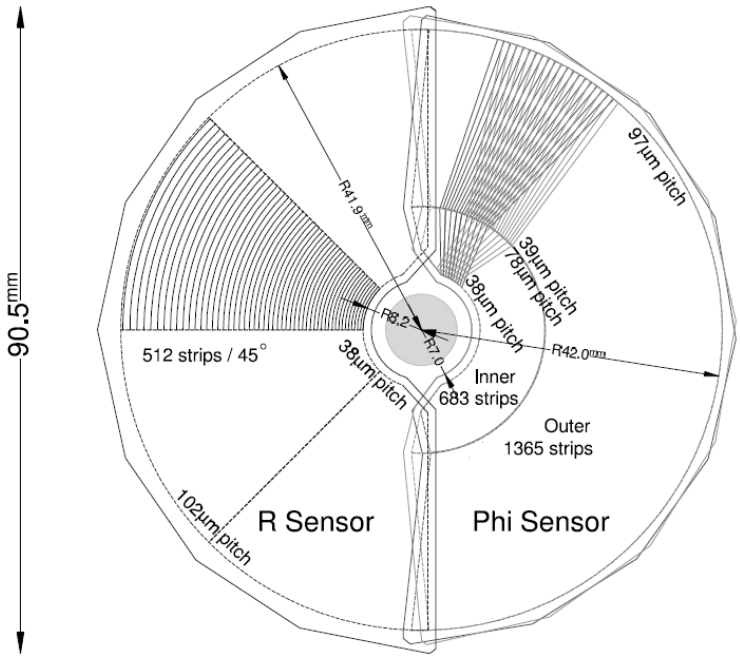

Figure 3.5: Sketch of the $R\phi$ geometry of the VELO sensors, showing only a portion of the strips that are actually present on the real sensors. Taken from Ref. [58].

long and covers a z range from $-0.2 \rightarrow 0.8$ m. There is a small amount of backward coverage by two R stations either side of the primary interaction which can be used to veto excessive pile-up events as LHCb typically runs with an average of $1.4 - 1.7$ interactions per bunch crossing. Unlike the other experiments at the LHC,

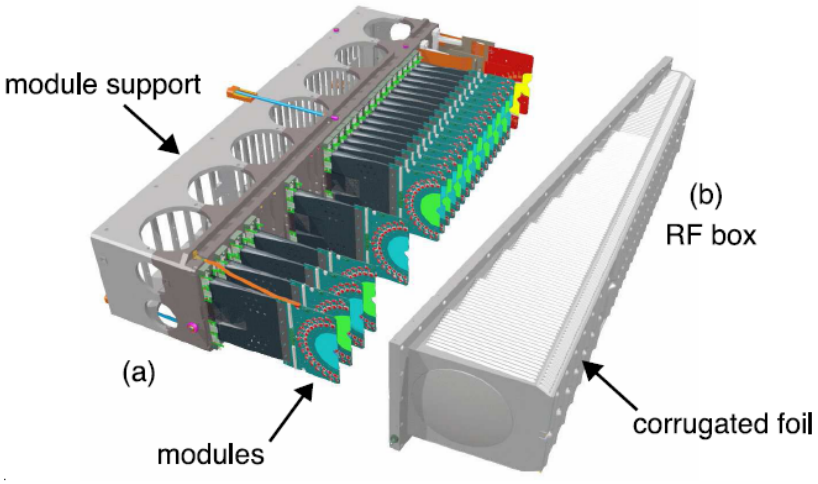

Figure 3.6: Simulation of the VELO stations showing how the RF-foil encloses the system (bottom). Taken from Ref. [58].

the VELO detector is located inside the LHC vacuum. Moreover the two halves of the VELO need to be mechanically moved closer to one another once stable beams have been declared. This is due to the relatively large dispersion of the beams upon entering the LHC which could cause damage to the sensors. Each half starts out retracted by 29 mm and then closes around the beam until the inner most strip approaches a distance of 8.2 mm from the beam. Since the VELO is so close to the PV it must be radiation hard [63] and able to operate for its lifetime, which is approximately 5 years of data taking, equivalent to about $8\,\text{fb}^{-1}$ of data. A replica VELO is on standby for installation should the current VELO become overly irradiated or undergo mechanical damage. As mentioned, the VELO lies before the magnet therefore tracks are considered to have straight trajectories since there is such a small bending influence. The sensors themselves are not in direct contact with the LHC vacuum but are instead separated by a radio-frequency (RF) foil which minimises RF pickup from the orbiting proton beams in the stations, seen in figure 3.6. When the VELO is closed about the beam axis, the inner radius of the RF-foil is just 5.5 mm from the collision region. The RF-foil has a complex geometry and undulates along the beam axis in multiple directions to avoid touching the sensors. Due to the close proximity to the beam, the silicon sensors must be kept cool, below $-10°C$. This is achieved using a bi-phase CO_2 Peltier system. Figure 3.6 shows one half of the VELO before it is placed about the beam axis.

Every sensor is made up of silicon strips that discretise space, thus there is an associated hit resolution for all measurements we make. The hit resolution varies

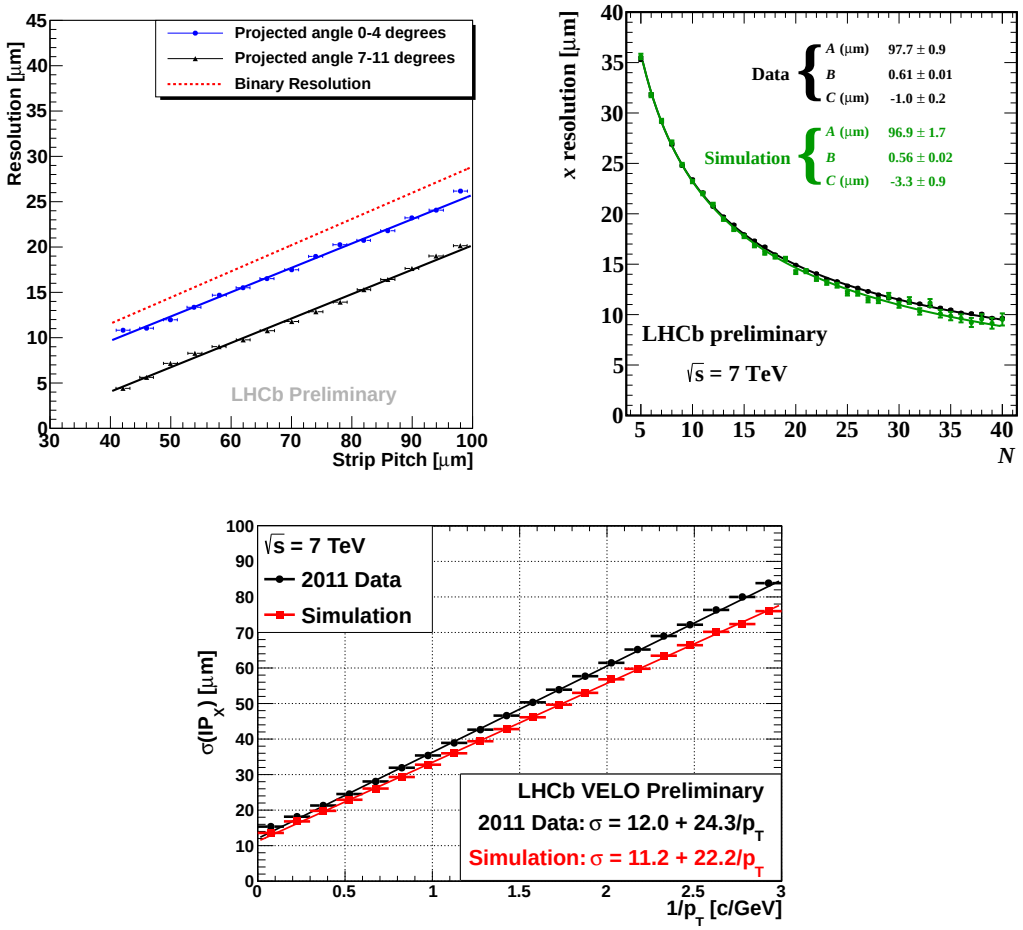

Figure 3.7: Hit resolution as a function of the inner-strip readout pitch as measured in 2011 (left). Primary vertex resolution as a function of event multiplicity (right). Impact parameter (IP_x) comparing 2011 simulation and data (bottom). Taken from Ref. [64].

as a function of the inner most strip pitch of the sensor. In figure 3.7 we see the hit resolution for various track angle intercepts with a sensor. We see that the hit resolution improves with increasing angle since more charge sharing between neighbouring strips will occur. However, after a point the angle increases such that the charge will be shared over multiple strips, thus the resolution degrades as it becomes difficult to accurately determine the position a particle traversed.

The inter-strip readout pitch for a VELO sensor varies between approximately 40 and 100 μm, and the hit resolution varies linearly with the inter-strip

pitch. In figure 3.7 (top left) the hit resolution is shown as a function of the inter-strip pitch, for two different ranges of the projected angle. The binary resolution is also shown for comparison. The best hit resolution is measured at a pitch of 40 µm and a projected angle of 8°, and is equal to 4 µm. It is the best hit resolution measured at the LHC. A typical IP resolution can be found in figure 3.7 (top right) As mentioned, the impact parameter is an extremely useful variable for rejecting prompt background from the PV; the smaller the resolution the closer you can make a vertex nearer to the PV whilst being assured it is not prompt. It has been recorded that the x (and y) component of the IP resolution is below 35 µm for particles with p_{T} of around 1 GeV/c.

Much work has gone into studying the difference of gradient between data and simulation and most outcomes suggest it is the material distribution that is key, see Appendix. The PV resolution (in the x-axis) as a function of track multiplicity can be found in figure 3.7 (bottom). For a typical primary vertex of 25 tracks, the PV resolution is about 13 µm in both the x and y-directions and about 71 µm in the z-direction.

3.3.2 Tracking - tracking stations

Further tracking stations are vital to be able to obtain momentum information. There is one positioned just before the magnet and the other three follow after the magnet, which are used to aid the determination of a particle's momentum. As previously mentioned, the four stations consist of: the TT, positioned before the magnet, and T1, T2 and T3 downstream of the magnet. The particle flux drops off with radius from the z-axis, so detectors closest to the beam will experience the largest amount of radiation damage and as such will need to have better positional resolution in order to maintain reliable pattern recognition. For this reason the TT and inner regions of T1-T3, which are known as the Inner Tracker IT, both use the same silicon sensor technology (p^{+}-on-n type) retaining a high granularity. These stations are commonly referred to as the Silicon Tracker (ST). This then leaves the outer regions of T1-T3 tracking stations, which are known as the Outer Tracker (OT), suffering significantly less irradiation, and so cheaper "straw tube" drift-time sensors are used.

Tracking - silicon tracker

The Tracker Turicensis (TT) and Inner Tracker (IT) illustrated in figure 3.8 (on the left and right respectively). The silicon detectors that make up the ST were designed

to achieve good spatial resolution of around 50 µm and high signal-to-noise ratios while minimising the occupancy, material budget, cost and signal shaping time in a moderately high radiation environment. The TT is a system of two modules,

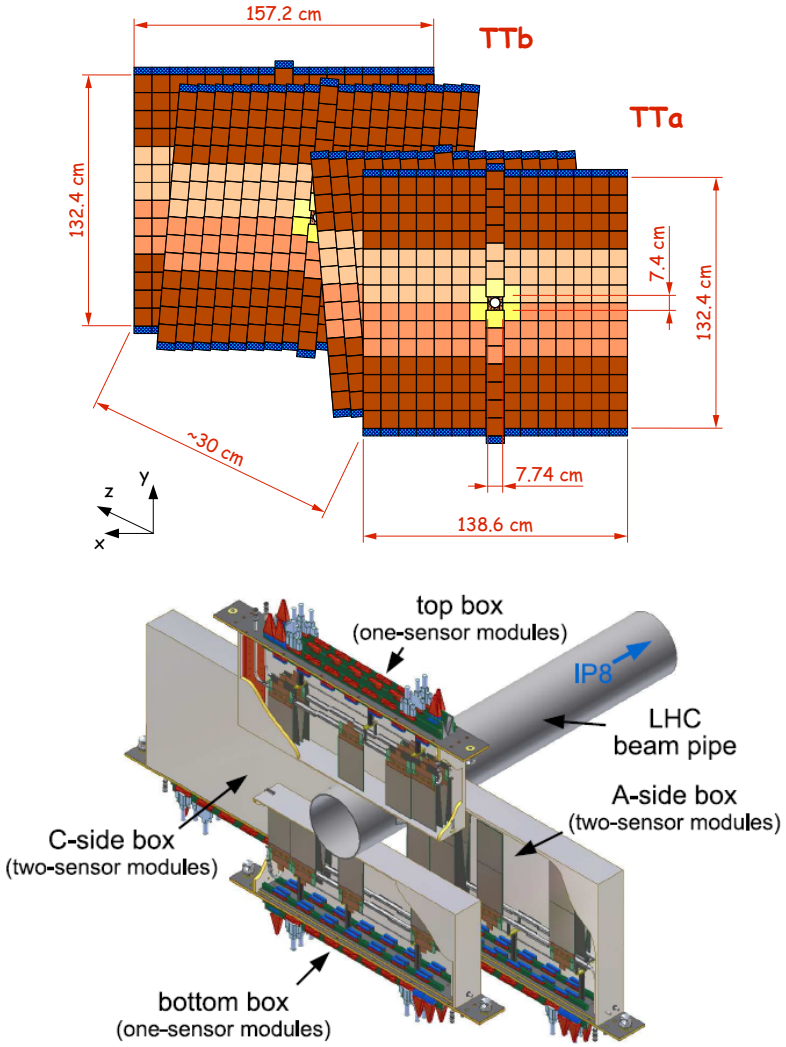

Figure 3.8: TT (left) and IT boxes (right). Both taken from Ref. [65].

each constructed from four layers of silicon strip sensors that cover the height of the LHCb acceptance. It is a planar tracking station, 150 cm wide and 130 cm high, containing 143360 readout strips covering an active area of 8.4 m^2. The TT is situated before the magnet, as shown in figure 3.2. There was a proactive decision to use a $(x-u-v-x)$ layout, shown in figure 3.8; x layers have their strips orientated vertically, whilst the u and v layers are positioned at stereo angles, rotated through $-5°$ and $+5°$ to the vertical respectively. The layout allows all infrastructure (front-

end readout hybrids, cooling and support structure) to be kept outside of the LHCb acceptance thus minimising material effects.

The Inner Tracker (IT) makes up the inner region of the three downstream tracking stations T1-T3. There are four boxes placed around the beam pipe in a cross shape, 120 cm wide and 40 cm high as shown in figure 3.8. Each box contains four silicon sensor layers orientated in an analogous manner to the TT. The unbiased resolution of the sensors has been measured to be 58 µm [66].

Tracking - outer tracker

The Outer Tracker (OT) is composed of four modules with each module containing two densely packed planes of straw tube drift chambers (128 tubes in total), staggered in order to ensure overlap. Figure 3.9 shows the cross section of one such

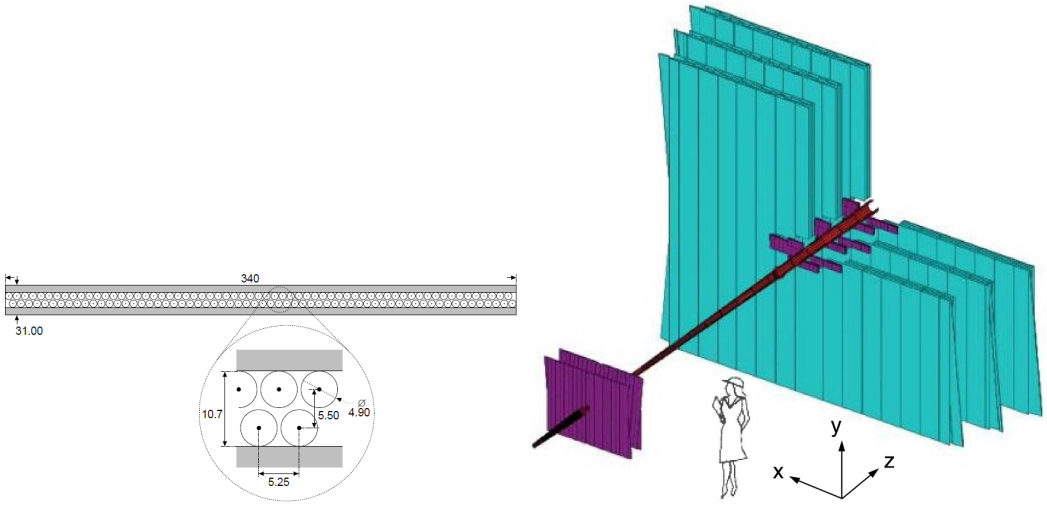

Figure 3.9: Cross-section of an OT module (left) and a full overview of the tracking systems excluding the VELO (right). Taken from Ref. [58].

OT sensor module. The OT geometry is similar to that of the IT, four layers of modules arranged in the same $x-u-v-x$ geometry where u, v planes are offset at -5° and $+5^\circ$ to provide stereo hit resolution. The design of the OT was driven by the need to achieve momentum resolution $\delta p/p$ close to 0.4 % to resolve the mass of reconstructed b hadrons to within around 10 MeV/c^2. This necessitated a rigid design so that the position of the straw-tubes could be guaranteed to a precision of 100 µm in the x-direction. Good electrical shielding of each tube was also required to ensure low cross-talk and noise. Each drift tube is made from two strips of thin foils twisted together in a spiral with an inner diameter of 4.9 mm filled with a mixture of Ar/CO_2 (70 : 30). The gases are chosen since they are easily ionised by

charged particles providing sub-50 ns drift times in order to restrict spillover to no more than two bunch crossings. The inner (cathode) foil is made of carbon-doped polymide (kapton) while the outer foil is a laminate of polymide and aluminium. Electrons from the gas drift towards the anode. As the drift velocity v is known ($v = \mu E$ where E is the applied electric field and μ is the mobility of the electrons), a measurement of the drift time with respect to the LHC clock allows the distance from the anode to be measured. The measurement of the drift time of the ionisation greatly improves the resolution. The aluminium ensures fast signal transmission and good shielding, which is crucial to prevent cross talk and noise, whilst the polymide is important to ensure the straw is gas-tight. Straw tube technology has a coarser resolution (about 0.2 mm) than the silicon detectors but is much cheaper and therefore very well suited to cover the large areas of the tracking system where particle densities are not as high. By the end of 2011, typical momentum resolutions were around 0.5 % for b hadrons.

3.3.3 Magnet

LHCb uses a warm dipole magnet which provides the bending of charged particles and can be used to measure the momentum of those particles. The application of the magnetic field causes charged tracks to bend in the xz plane. A warm magnet was chosen simply for the cost effectiveness and the time constraints involved in building a superconducting magnet. It was necessary to calibrate the strength of the magnetic field across the detector for use in software and this was extensively mapped using an array of Hall probes in three dimensions to a relative precision of about 4×10^{-4}. The magnetic field strength provides an integrated bending power of $\int Bdl \approx 4\text{Tm}$ within the region $z = 2.50 - 7.95\,\text{m}$. The main component of the B-field (B_y) is also displayed as a function of z-position in figure 3.10, it also shows the various types of tracks that are reconstructed at LHCb.

3.4 Particle identification

As will be seen in the analysis in this thesis, there are many ways for B mesons to decay to a final state. In some circumstances, the only way to differentiate one decay from another is by the correct identification of just one final state particle. For example, $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ is almost exactly the same as $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$, where the only difference is a pion interchanged for a kaon in the final state. If one is not careful and does not treat this appropriately it can lead to pollution, or cross-feed, as these decay modes appear in each other's invariant mass spectra and

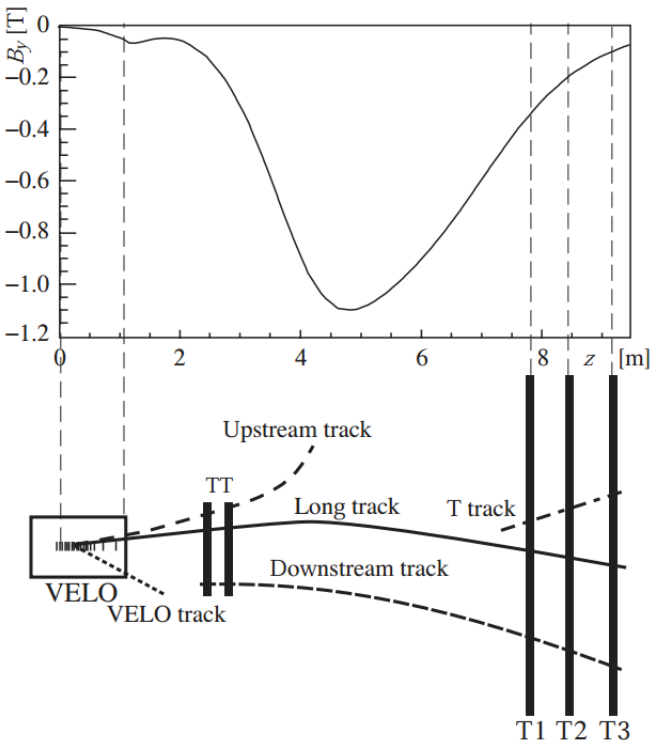

Figure 3.10: Schematic of the B_y-field component as a function of z-axis. The various types of reconstructible tracks are also displayed: long, upstream, downstream, VELO and T tracks. Figure is taken from Ref. [67].

need modelling. The need to identify these specific final states places an emphasis on charged particle identification (PID) in heavy flavour experiments. A prime example of the importance of particle identification comes from the two body $B \to h^+h'^-$ decays [68] (where $h^{(\prime)} = \pi, K$) as shown in figure 3.11. A handle on the backgrounds can be obtained when applying positive identification for kaons, pions and protons. It also provides a much better resolution of the prime signal mode which in this case was $B^0 \to \pi^+\pi^-$.

3.4.1 Particle identification - RICH

The primary role of the particle identification system is to distinguish between charged particles, π's, K's and p's (e's and μ's are identified by the calorimeter and muon systems respectively in sections 3.4.2 and 3.4.3). This separation is achieved using Ring Imaging Cherenkov (RICH) detectors. There are two RICH detectors at LHCb, RICH1 and RICH2. These both work by measuring emissions of Cherenkov radiation produced in a forward cone when a charged particle is travelling faster than the phase velocity of light in a medium of a certain refractive index. The shape

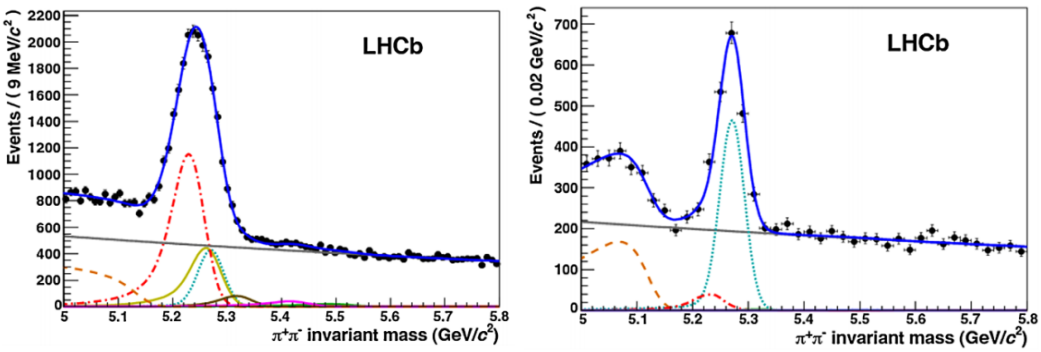

Figure 3.11: Invariant mass distribution for $B \to h^+h'^-$ decays [68] (where $h^{(\prime)} = \pi, K$) before the application of the RICH information (left) and after applying particle identification (right). The distributions seen here are the signal being investigated which is $B^0 \to \pi^+\pi^-$ (turquoise dotted line) then several background components: mis-identified $B^0 \to K^+\pi^-$ (red dashed-dotted line), partially reconstructed B three body modes (orange dashed line), double mis-identification from $B_s^0 \to K^+K^-$ (yellow line), single mis-identification from $B_s^0 \to \pi^+K^-$ (brown line), $\Lambda_b^0 \to pK$ (purple/magenta line) and $\Lambda_b^0 \to p\pi$ (green line). After applying particle identification only the signal and two small background contributions remain, all others are negligible. The solid grey line represents the combinatorial background in both figures. Taken from Ref. [69].

of the cone of light depends on the particle's velocity and is given by the formula

$$\cos(\theta_C) = \frac{1}{n\beta}\,, \tag{3.1}$$

where θ_C defines the half cone in which radiated photons are emitted (known as the Cherenkov angle), n is the refractive index of the medium and β is the usual velocity relative to the speed of light in a vacuum (v/c). Using the formula for relativistic momentum, $p = \gamma m_0 c\beta$, we can re-write Eq. 3.1 in terms of the rest mass of a particle.

$$\cos(\theta_C) = \frac{E}{npc} = \frac{1}{n}\sqrt{1 + \left(\frac{mc}{p}\right)^2} \tag{3.2}$$

From Eq. 3.2 it can be seen that if a measurement of the particle's momentum and the Cherenkov angle of the emitted light can be made then this leads directly to a measurement of the particle's rest mass and thus the particle type can be inferred.

The photodetectors are silicon-based pixel chips and these need to be outside the acceptance of charged particles. An advantage to this will minimise the degradation of the tracking systems from material interactions, the RICH photodetectors are located outside of the LHCb acceptance. One challenge (in both RICH1 and RICH2) was to reduce the material budget within the detector acceptance. Therefore, designs were considered that tilted focusing mirrors outwards slightly

and introduced secondary planar (flat) mirrors that direct the Cherenkov radiation out of the detector acceptance. This allows for a smaller photon-detector area and a more compact system, thus achieving a reduced material budget.

The gases and aerogel in the RICH detectors are contained within large aluminium boxes. The RICH1 box is 300 μm thick inside the acceptance. As RICH2 is larger, the pressure differential is greater so two 1 mm thick carbon fibre and aluminium skins are separated by 30 mm of foam in the entrance and exit windows respectively. The Cherenkov light is then collected and focused by spherical mirrors before being reflected onto the photodetectors using flat mirrors. The result of this layout are rings of radius, $r = f\theta_C$ where f is the focal length of the spherical mirror. Both RICH detectors use hybrid photon detectors (HPDs) to measure the positions of the emitted Cherenkov photons. The HPD is a vacuum photon detector in which a photoelectron, released when an incident photon converts within a photo cathode, is accelerated by a high voltage of typically 10 to 20 kV onto a reverse-biased, pixel silicon anode. The HPDs are read out by integrated pixel chips and arranged in a hexagonal pattern outside the two RICH detectors. Once these rings have been imaged, pattern recognition follows by the use of a likelihood fit to determine the ring that best matches the expectation from a kaon, pion, or proton mass hypothesis. A set of likelihoods describing how consistent data are with a particle being a kaon, pion or proton is then made available for analyses such as we will see used in this thesis and as an example we use the nomenclature $\mathrm{DLL}_{K\pi}$ for the kaon hypothesis as an example.

Particle identification - RICH1

RICH1 is located immediately after the VELO exit window as seen in figure 3.2. In order to limit its overall size it is placed as close to the PV as possible. RICH1 combines silica aerogel and fluorocarbon (C_4F_{10}) gas radiators and allows discrimination between charged hadrons in the momentum range $10 - 40\,\mathrm{GeV}/c$. Silica aerogel is a colloidal form of quartz, that is solid but very light with a relatively large refractive index which makes it ideally suited for low-momentum particle identification, providing positive kaon identification above $2\,\mathrm{GeV}/c$ and $\pi - K$ separation upto $10\,\mathrm{GeV}/c$.

A track passing through 5 cm of aerogel with refractive index $n = 1.03$ for light of wavelength 400 nm is expected to yield around 6.5 photoelectrons in a ring from a charged particle with $\beta = 1$. The expected yield for 95 cm of C_4F_{10} ($n = 1.0014$ at 400 nm) is 30 photoelectrons. The RICH1 detector covers an angular acceptance of $25 - 300$ mrad where the lower angle is limited by the reintroduction

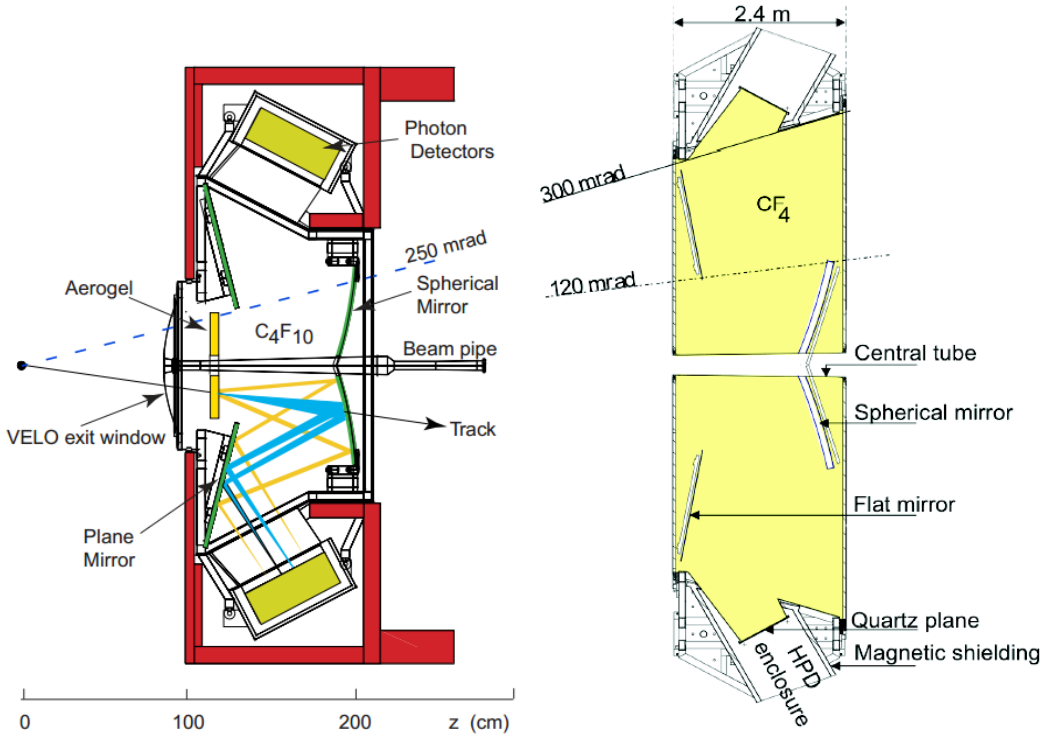

Figure 3.12: (Left) RICH1 detector as viewed from the side and (right) RICH2 as viewed from above. Spherical and flat mirrors are indicated as well as the size of each detector. Taken from Ref. [58].

of the beam pipe. As previously mentioned, minimising the material budget in the spectrometer acceptance was a key requirement for the RICH detectors. Since the spherical mirrors need to be within the acceptance, they are constructed using a light-weight carbon-fibre reinforced substrate. Cherenkov light is focused onto the photon detector planes using tilted spherical mirrors and secondary plane mirrors, as shown in figure 3.12 (left). Since the photon detector arrays of the RICH1 sit in the fringe field of the magnet, magnetic shielding is required in order to retain the photoelectron efficiency of the PMT's. The total amount of material falling within the detector acceptance corresponds to a material budget around $0.08X_0$.

Particle identification - RICH2

RICH2 is positioned downstream of the tracking system and magnet since the high-momentum tracks being measured by it will be less affected by the magnetic field. This in turn reduces the material budget of the tracking systems. In RICH2 the stray magnetic field is reduced by shielding boxes from 15 mT to around 0.5 mT. It has an acceptance that is limited to the low-angle region where there are mostly

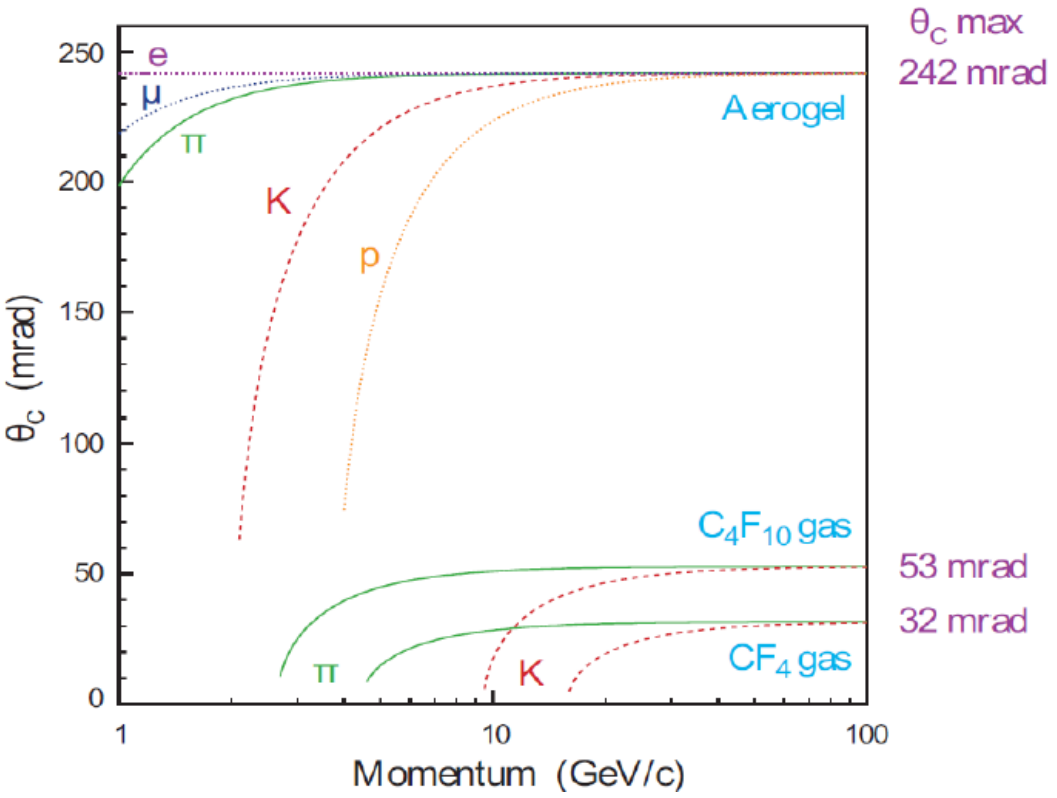

Figure 3.13: Cherenkov angle data for the reconstructed Cherenkov angle as a function of track momentum for different radiators. Taken from Ref. [69]

high-momentum particles. Its lower angular acceptance of 15 mrad is limited by the beam pipe and its upper acceptance angle is 120 mrad, highlighting that its main purpose is for the identification of high-momentum particles.

Figure 3.12 shows the top view of the RICH2 detector. It contains a single gas radiator, CF_4, which covers a wider momentum range of $16 - 100$ GeV which complements that of the RICH1 detector. The flat and spherical mirrors are manufactured from 6 mm thick glass. Including the gas radiator, the total material budget of RICH2 corresponds to a radiation length of $0.15X_0$. The expected yield for 180 cm of CF_4 ($n = 1.0005$ at 400 nm) is 22 photoelectrons. The reconstructed Cherenkov angle as a function of track momentum is shown for the various radiators used in the RICH system and how they overlap in figure 3.13.

3.4.2 Particle identification - calorimeters

The calorimeter's purpose is to determine the energy deposition that a particle produces as it "bumps" into absorber material, and in doing so, loses its own energy

to the detector. By collection of the cascade shower produced, one can infer the original energy of a particle. Thus calorimeters usually come at the end of a detector's layout since their purpose is to stop a particle escaping in order to measure its energy. To do this they typically have a large radiation length which would be detrimental to the tracking performance if placed elsewhere in the detector.

A classical structure of an electromagnetic calorimeter (ECAL) followed by a hadron calorimeter (HCAL) has been adopted at LHCb [70]. Calorimeters also provide a major contribution to the first level trigger system where they are used to select electron, photon and hadron candidates above specified transverse energy thresholds. The calorimeter system comprises of several main parts: the scintillator pad detector (SPD), preshower (PS), ECAL and HCAL in order of increasing z; each will be discussed further in turn. All of the calorimeter systems share a common technology: scintillation light is transmitted to a photo-multiplier (PMT) by wavelength-shifting (WLS) fibres. The single fibres for the SPD/PS cells are read out using multianode photo-multiplier tubes (MAPMT), while the fibre bunches in the ECAL and HCAL modules require individual phototubes.

SPS and PS

To separate electrons from the large pion background, two detectors are placed in front of the ECAL: the scintillator pad detector (SPD) and the preshower (PS) which consists of a 15 mm thick lead plate sandwiched between two layers of scintillator pads, before the ECAL. The PS indicates the electromagnetic character of the particle (i.e. whether it is an electron, if charged, or a photon, if neutral) and the SPD determines whether particles hitting the calorimeter system are charged or neutral before showering. They are used at the trigger level in association with the ECAL to indicate the presence of electrons, photons, and neutral pions. Both electrons and photons then interact in the thin lead plate and the deposited energy is collected by the second scintillator. Hadrons have a longer interaction length and therefore they pass through without depositing very much energy (although the hits they create are retained).

ECAL

The ECAL employs "shashlik" technology of alternating scintillating tiles and lead plates. Each ECAL module consists of 2 mm of lead that induces electromagnetic showers, followed by 4 mm of scintillator material. The ECAL consists of 66 layers of such modules. A drawing of a module can be seen on the right side of figure 3.14.

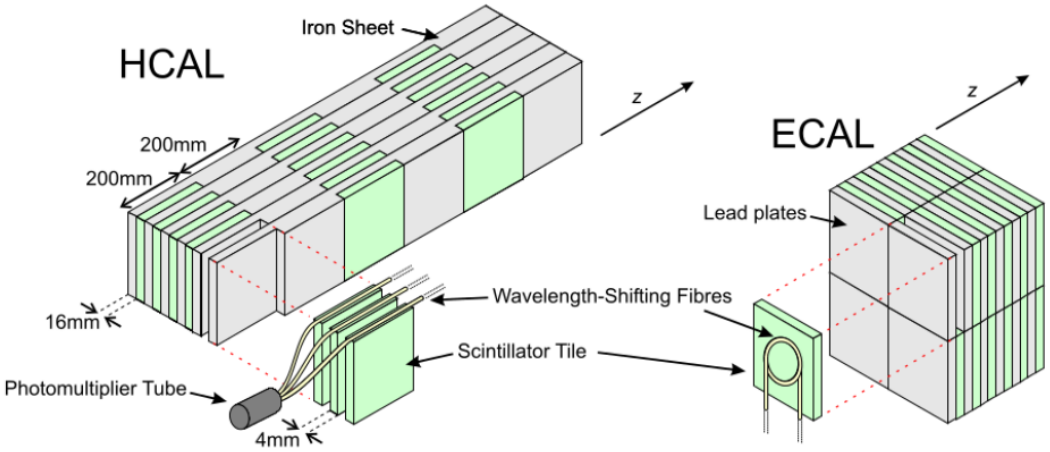

Figure 3.14: The internals of the HCAL (left) and ECAL (right). The main difference being the HCAL uses scintillating tiles parallel to z whilst the ECAL has its placed perpendicular to the z-axis. This causes the readout electronics to be directed differently under each scenario.

The scintillator pads have a similar design to the SPD and PS sensors, using WLS fibres to read out the light produced, except that the fibres are grouped in bundles, and each bundle is passed to a single PMT, giving a coarser granularity. The pad size is adjusted to three different values to achieve three zones with higher granularity around the beam-pipe as shown in figure 3.15 (left). This is because the hit density varies by over two orders of magnitude over the calorimeter surface. The ECAL, PS and SPD are scaled such that they have a projective channel geometry as seen from the interaction point. This makes combining measurements across the

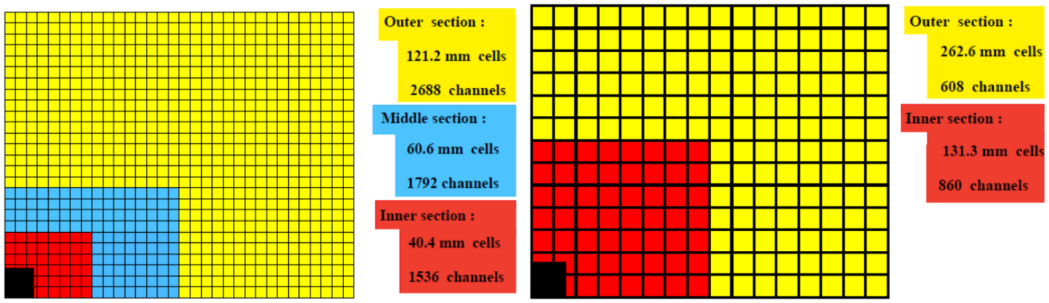

Figure 3.15: Diagrams showing the segmentation of the calorimeter components. Segmentation of the SPD, PS and ECAL (left) and the segmentation of the HCAL sensors (right). Taken from [58].

three subdetectors straight forward. The total amount of material falling within the detector acceptance corresponds to a material budget around $8\,\%X_0$. The energy resolution achieved is $\frac{\sigma_E}{E} \approx \frac{9\,\%\pm0.5\,\%}{\sqrt{E}} \oplus 0.8\,\%$ where E is in units of GeV and the

last term represents the contribution from electronics noise.

HCAL

In total, the HCAL weighs in at around 500 tonnes. It is a sampling device made from iron and scintillating tiles, as absorber and active material respectively. The main difference with the hadron calorimeter compared to the ECAL is the orientation of the scintillating tiles, that instead, run parallel to the beam axis. This can be highlighted when comparing the HCAL innards of figure 3.14 with the right side of the same figure. This design is more preferable since hadronic showers tend to be less collimated that those in the ECAL. Overall, it has 1500 cells divided in two regions, inner and outer, corresponding to the distance to the beam-pipe as shown in figure 3.15 (right). The energy resolution achieved is $\frac{\sigma_E}{E} \approx \frac{69\,\% \pm 5\,\%}{\sqrt{E}} \oplus (9\,\% \pm 2\,\%)$ where again E is in units of GeV.

3.4.3 Particle identification - muon system

Muons are an incredibly important source of identification since they leave behind extremely clean traces in our detector. The muon system provides fast information for the high-p_{T} muon trigger at the earliest level (Level-0), muon identification for the high-level trigger (HLT) and offline analysis. Due to the large mass of the muon and because it is a lepton it is very unlikely to interact with material. Since it is still a charged particle it will cause ionisation as it traverses the detector leaving a trail to be picked up by the tracking system. The energy loss (dE/dx) is so small that a muon can generally pass straight through the calorimeter system if it has sufficient momentum, so muon detectors need be placed at the rear of the detector. The minimum momentum of a muon to cross the five stations is approximately $6\,\mathrm{GeV}/c$ since the total absorber thickness, including the calorimeters, is approximately 20 interaction lengths [71]. The muon system is composed of five stations (labelled M1-M5) placed at increasing z as can be seen in figure 3.16 (left). The system as a whole comprises 1380 chambers, of 20 different sizes, and covers a total area of $435\,\mathrm{m}^2$. Station M1 is placed in front of the calorimeters and is used to improve the p_{T} measurement in the trigger since it provides an additional hit close to the downstream trackers. Stations M2-M5 are positioned downstream behind the calorimeters. The more upstream stations, M1-M3, have a higher spatial resolution along the x-axis (bending plane). Thus these can be used to define the track direction and the p_{T} of any candidate muon with a resolution of around 20 %. Stations M4 and M5 have a limited spatial resolution, their main purpose being

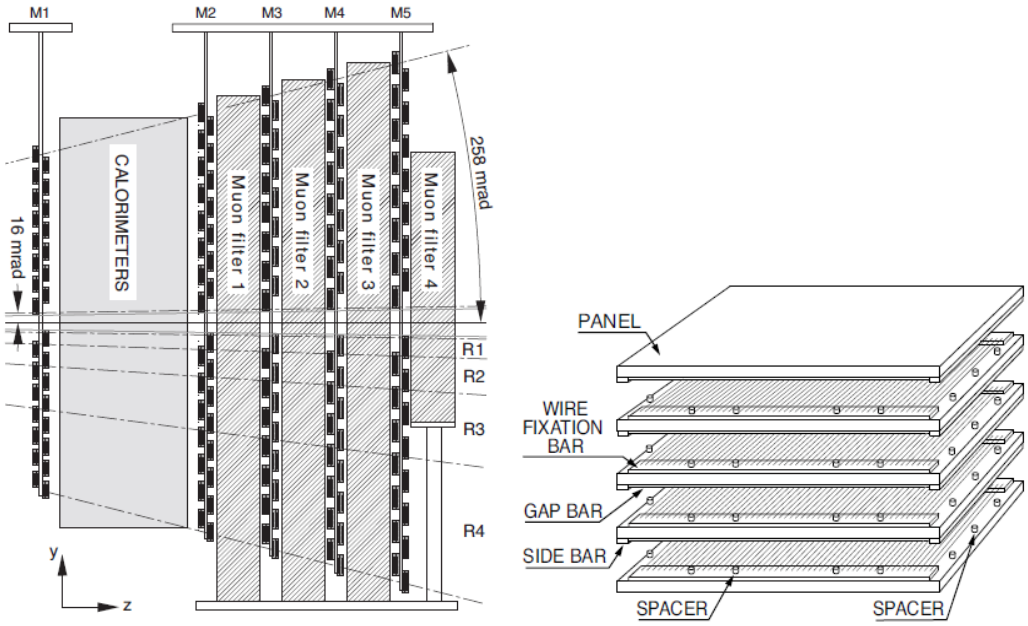

Figure 3.16: Side view of the muon system (left) and exploded view of a single muon chamber (right), notice the multiple gas gaps. Taken from [58].

the identification of penetrating particles and reducing backgrounds such as hadron punch-through.

In order for the muon trigger to fire, it requires aligned hits in all five muon stations. This means that the chambers must have an efficiency of $> 99\,\%$ within the 25 ns bunch crossing. The system is therefore equipped with Multi Wire Proportional Chambers (MWPC) with 2 mm wire spacing and a small gas gap (5 mm). The tight requirement on the time resolution is made possible by using a fast gas mixture and suitable charge-collection geometry, i.e. a gas mixture of $Ar/CO_2/CF_4(40 : 55 : 5)$ and a wire plane of 2 mm spacing, symmetrically placed in a 5 mm gas gap. A muon crossing the 5 mm MWPC gas gap will ionise around 50 electrons that will drift toward the wires due to the applied electric field. This will cause an avalanche effect that will induce a negative signal on the wire and a positive signal with the same shape (but half the magnitude) on each of the cathodes.

An exploded view of a standard muon chamber is presented in figure 3.16 (right). Triple-GEM detectors are used in the innermost region of station M1 covering approximately ($3\,m^2$) because the particle rate is at its highest. This choice was dictated by the better-ageing properties of this type of detector. From the point of view of this analysis the muon system is vital as it provides a high efficiency for triggering on our decays of interest.

3.5 Trigger system

The 40 MHz maximum collision rate of the pp bunch crossing is far too high to record every single event, so a selective trigger system is employed. During 2011 data taking, the bunch crossing rate at LHCb was about 10 MHz. The trigger system is designed to reduce this to a manageable storage rate of 3 kHz, keeping only the most interesting events for physics analyses in line with the specified physics programme. The main reason for the reduction is that the data can only be written to permanent storage at a certain rate, and a limited volume of raw storage space is available.

The LHCb trigger is composed of two stages. Stage one, level 0, is a hardware trigger synchronous with the LHC bunch crossing rate of 40 MHz and is designed to reduce the event rates to 1 MHz. The second is a flexible software, high level trigger, which uses the full detector information to further reduce the event rate to 3 kHz suitable for offline storage. Events that fail either of these stages are rejected.

3.5.1 Trigger - level 0

The first level trigger, level 0, has to make logical decisions in just 4 µs after a collision using only hit information *i.e.* there are no tracks reconstructed. Consequently it only uses information from the calorimeters, muon stations and pile-up stations of the VELO. Each detector's information is provided to the L0 decision unit (L0DU) where the decision to keep or discard each event is made. The mesons we are interested in typically produce daughter tracks with large p_{T} and E_{T} of several GeV/c (GeV) due to the large mass of the mother. This allows us to apply some simple logic via the following requirements:

- At least one cluster in the HCAL with $E_{\mathrm{T}} > 2.5$ GeV.
- At least one cluster in the ECAL with $E_{\mathrm{T}} > 2.5$ GeV.
- A muon candidate in the muon chambers with $ip_{\mathrm{T}} > 1.48$ GeV/c or two muons (dimuon trigger) with $p_{\mathrm{T}}{}_t^1 + p_{\mathrm{T}}{}^2 > 1.3$ GeV/c.

The calorimeter system performs several functions by selecting, based on their transverse energy, hadron, electron and photon candidates. Muons are interesting to various analyses, including this one, because of their natural particle identification that is implied by the very detection of muons so a relatively small p_{T} cut can be used. To cut down too complex events, the VELO pile-up stations as well as the SPD also provide a charged multiplicity veto. Together they provide an estimate on the number of tracks and thus remove events with potentially large backgrounds that are not worth keeping.

3.5.2 Trigger - high level trigger

The data from all detector elements are read out for events passing L0. These data are then passed, at a rate of $\approx 1\,\text{MHz}$, to the Event Filter computer Farm (EFF) consisting of 15200 computing cores, which runs the HLT algorithms. There are two parts to the HLT software trigger: HLT1 and HLT2. The HLT1 reduces the rate from 1 MHz to approximately 30 kHz using selection variables obtained from the partially reconstructed events. Events passing HLT1 then have HLT2 algorithms run on them. At this point in the filtering the rate is low enough for the remaining tracks to be reconstructed.

HLT1 begins by matching tracks reconstructed in the tracking system with the hits in the calorimeters and muon stations (or in the case of neutral particles to check that there is not a compatible track in the tracking system). Tracks must have large enough p_{T} and impact parameter with respect to all primary vertices in the event.

HLT1 reduces the retention rate to 30 kHz, and passes the events selected to HLT2. A more comprehensive reconstruction is run for HLT2 but it is still more simplified than the full offline reconstruction. For example, a simplified tracking algorithm is used whereby the RF-Foil is made of a few simple solids, and so looser track quality cuts are applied than would be used for offline data analysis. Multiple tracks can be considered under the hypothesis that they are the decay products (daughters) of a single mother particle. Their momenta are summed to give the momentum of the mother particle, and a vertexing algorithm run to find the most likely decay point of the mother. Track quality cuts such as χ^2 values can be applied and the overall result is a reduction in the rate to about 3 kHz, where these events are then stored permanently to be reconstructed fully by LHCb software offline. For the analysis in this thesis, the most important trigger selections come from those involving muon candidates. This is because we can seach for high p_{T} muon tracks coming from the J/ψ decay, and as we have discussed the muon efficiency is very high. In fact most LHCb analyses involving muons for triggering have very high efficiencies in the $90-99\,\%$ range. Further detail of the muon triggers can be found in Ref. [72].

3.6 Stripping and trigger

At LHCb, raw events are stored on disk, in order to make access to the dataset easier for those carrying out analyses and to avoid multiple accessing, the data is filtered for each analysis using tailored selection criteria. This analysis uses a

dedicated stripping line that selects likely signal events from the LHCb dataset. Where possible our stripping line uses similar cuts to those of the $B^0 \to J/\psi K^0_S$ stripping line, as it is used to obtain the normalisation channel in Part 1 of the analysis.

3.6.1 K^0_S reconstruction

The K^0_S channels are split into two categories depending on how the pions from the K^0_S decay are reconstructed. For decays where both pions have hits inside the VELO and the downstream tracking detectors the K^0_S candidates are classified as long (LL). If the daughter pions are reconstructed without VELO hits (but still with TT hits upstream of the magnet) they are classified as downstream (DD) K^0_S candidates. We use both LL and DD K^0_S categories in this analysis. Separate selections are applied to the LL and DD categories in order to maximise the sensitivity. The downstream tracks typically have a worse quality vertex and momentum resolution than long tracks. The pions are selected with momentum greater than $2\,\text{GeV}/c$ and the quality of the tracks is ensured by requiring their χ^2 per degree of freedom (ndf) to be less than 4. Pairs of pions with opposite charge are combined and required to come from a common vertex with $\chi^2 < 20$. For the downstream K^0_S we require a slightly tighter vertex constraint, $\chi^2 < 18$, to account for the absence of VELO hits. The mass of the pion pair must be within $\pm 30\,\text{MeV}/c^2$ of the true K^0_S mass to be accepted in the pre-selection. In the `BetaSBd2JpsiKsDetachedLine` the K^0_S candidate is required to have a decay length significance greater than 5, and the same requirement is made in our line.

3.6.2 J/ψ reconstruction

To reconstruct the $J/\psi \to \mu^+\mu^-$ decay, two long muon tracks with $\text{DLL}_{\mu\pi} > 0$ and $p_\text{T}(\mu^\pm) > 500\,\text{MeV}/c$ are combined. The two muon candidates must be reconstructed in the vertex detector and the tracking chambers downstream of the magnet, and must have hits in the muon stations. The p_T cut is mainly intended to reduce the combinatorics at an early stage of the analysis. It does not introduce any significant bias in the J/ψ angular distribution and enhances B events over prompt and non-B backgrounds. The pairs of muons are required to have opposite charge and to come from a common vertex for which the reconstruction gives $\chi^2/\text{ndf} < 16$ (a very loose cut). The $\mu^+\mu^-$ pair mass must have an invariant mass within $\pm 80\,\text{MeV}/c^2$ of the known J/ψ mass [15]. A final constraint that the muon combination has a χ^2 distance of closest approach below 20 is applied.

3.6.3 Bachelor pions and kaons

Candidates for the pions and kaons coming from the B decay (referred to as "bachelor" tracks) are only relevant to the `B2JpsiKshhLine`. Such candidates are selected if they are produced sufficiently far away from the primary vertex. The impact parameter (IP) is the minimum distance of approach of the track with respect to the primary vertex. We require that each bachelor track has IP χ^2 with respect to the primary vertex greater than 4, and transverse momentum satisfying $p_{\mathrm{T}} > 250\,\mathrm{MeV}/c$. It is important to note that `StdLooseKaons` has a particle identification cut from the outset, given at $\mathrm{DLL}_{K\pi} > -5.0$.

3.6.4 B^0 reconstruction

The reconstructed J/ψ and K^0_{S} (and bachelor π or K) particles are combined into a B meson candidate. The mass window is chosen to be very loose, $\pm 500\,\mathrm{MeV}/c^2$ around the true B^0_s mass, to allow the study of the full distribution including background. We further require that the bachelor particles form a vertex with $\chi^2/\mathrm{ndf} < 10$ and that this B candidate points back to the primary vertex with $\cos(\theta) > 0.9998$. For the `B2JpsiKshhLine` we also require that all two-particle subcombinations of the daughter tracks have a χ^2 distance of closest approach below 5. In the `BetaSBd2JpsiKsDetachedLine` there is an additional constraint such that the B candidate is detached from the primary vertex by requiring that the proper lifetime is greater than 0.2 ns but we do not impose such a cut that explicitly biases lifetime in our stripping line.

3.6.5 Trigger lines

LHCb has excellent trigger efficiency for muons. It is for this reason that we use the dimuons from the J/ψ decay to trigger on as it gives much higher efficiencies than the hadronic trigger lines.

At L0 the candidates are typically selected by requiring one or two high-p_{T} muons in the final state. At HLT1, those candidates are confirmed if they contain a single high-p_{T} muon and a track with large impact parameter. At HLT2, the candidates are selected topologically by requiring two-to-three tracks to be reconstructed with a large invariant mass and a vertex that is displaced from the primary vertex. For our analysis we use trigger decisions labelled as `TOS`, meaning trigger on signal, *i.e.* the trigger was caused by the muons from the J/ψ in our signal decay.

3.7 Dataset

We use the 2011 dataset totalling $\mathcal{L}_{\text{int}} = 1.0\,\text{fb}^{-1}$ of integrated luminosity. An offline selection is applied to the dataset, known as *stripping*. This produces reduced data files that can be easily processed for off-line analysis. Table 3.1 shows the representative number of candidates we obtain after the stripping selection applied to the full 2011 dataset. The use of simulated datasets is vital in extracting efficiencies

Table 3.1: Numbers of events from 2011 dataset passing the stripping selection. Inclusive samples of signal and background.

Decay mode	K^0_S (LL) type		K^0_S (DD) type	
	MagDown	MagUp	MagDown	MagUp
$B \to J/\psi\, K^0_S$	59486	39964	322896	220648
$B \to J/\psi\, K^0_S \pi^+\pi^-$	285629	180042	1699252	1058432
$B \to J/\psi\, K^0_S K^\pm \pi^\mp$	285454	181444	1895065	1169993
$B \to J/\psi\, K^0_S K^+ K^-$	62613	40098	465421	284969

and systematics for many analysis procedures. The details of the simulated datasets used in this analysis can be found in Table 3.2. LHCb uses a dedicated package called Gauss to create a virtual representation of detector events. Gauss consists of three main aspects, namely: pp collision and hadronisation of quarks where the kinematics are governed by Pythia, EvtGen which contains decay rate and decay kinematics information thus governing how particles decay down the chain and finally Geant4 which describes the transport (and interactions) of particles as they traverse the detector. All simulated samples are produced using conditions that represent the period of data taking and use the same stripping selection. The average number of interactions per bunch crossing for these simulated events was $\nu = 2$ and each beam had an energy of 3500 MeV. All simulated samples have a baseline generator level cut, `DaughtersInLHCb`, which constrains charged tracks to be in the LHCb acceptance ($0.01 \leq \theta \leq 0.4\,\text{rad}$), excluding daughters of any Λ or K^0_S hadron. For the $B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ modes we have an additional cut, `DaughtersInLHCbAndWithMinP`, which is the same as before except it also imposes a minimum momentum cut of $1600\,\text{MeV}/c$ for stable charged tracks, i.e. pions and/or kaons, and a separate $3000\,\text{MeV}/c$ for the muon tracks in this case, from the J/ψ. A systematic uncertainty to account for any efficiency variation across the phase space due to this cut will be looked at in more detail.

Table 3.2: Simulated datasets for all B decay modes including event type and both magnet polarisations. The $B^0 \to J/\psi K^0_S \pi^+\pi^-$ and $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ samples are composed of 50 % PHSP, with the remainder produced through resonances decaying to $K^0_S\pi^+\pi^-$: $K^0_1(1270)$ (18 %), ρK^0_S (17.5 %), ωK^0_S (14.5 %). The rest of the decays proceed through the PHSP model.

Decay mode	Decay model	Generated events	
		MagUp	MagDown
$B^0 \to J/\psi K^0_S$	SSD_CP	5047978	5048487
$B^0 \to \psi(2S) K^0_S$	SSD_CP; $\psi(2S)$ VVPIPI	523498	518998
$B^0 \to J/\psi K^0_S \pi^+\pi^-$	cocktail	500000	500000
$B^0 \to J/\psi K^0_S K^\pm \pi^\mp$	PHSP	407497	415499
$B^0 \to J/\psi K^0_S K^+ K^-$	PHSP	411498	405997
$B^0_s \to J/\psi K^0_S \pi^+\pi^-$	PHSP	500000	500000
$B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$	PHSP	399000	400000
$B^0_s \to J/\psi K^0_S K^+ K^-$	PHSP	411498	411498

Some of the samples contain resonances in the decay tree, for example the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode has the following structure: 50 % non-resonant final state $J/\psi K^0_S \pi^+\pi^-$ and 50 % via $J/\psi K^0_1(1270)$, with the kaon resonance proceeding 36 % to $K^0_S\pi^+\pi^-$ (PHSP), 35 % to ρK^0_S (VVS_PWAVE) and 29 % to ωK^0_S (VVS_PWAVE) where both ω and ρ decay to two charged pions (VSS model). Details of the various models and their implementation can be found in EvtGen Ref. [73].

4

Analysis overview

The main objective of the analysis is to measure the relative branching fractions of the $B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}$ decays. Data is preselected via dedicated stripping lines as outlined in Sec. 3.6, using the full 2011 dataset of $1.0\,\text{fb}^{-1}$. Our selection method is based on the **C**ut **R**ecursive **OP**timiser **CROP**. CROP is a high-speed rectangular cut optimisation tool[1] that recursively re-optimises an ensemble of cuts until a stable maximum in the signal significance is achieved, in this case using as the figure of merit $S/\sqrt{S+B}$, where S and B are the expected numbers of signal and background events, respectively. The values of S and B can be obtained in different ways depending on the analysis strategy. The analysis is carried out in two parts:

- Part I – Branching fraction ratio $\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)/\mathcal{B}(B^0 \to J/\psi K^0_S)$ where the primary objective is to improve the knowledge of $\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)$.
- Part II – Branching fraction ratio $\mathcal{B}(B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-})/\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)$ where the primary objective is to search for, and hopefully observe, new decay channels.

A general feature of the analysis is that the K^0_S candidates are reconstructed in two categories. The long lifetime of K^0_S mesons and the large boost of particles produced in LHC pp collisions means that some K^0_S decays occur inside the LHCb VELO detector but a significant fraction occur outside. As for other LHCb analyses of b hadron decays to final states containing K^0_S mesons [74, 75, 76, 77, 78], two

[1] The use of a Neural Network to account for non-linear correlations between our cut variables was investigated. It was found that with the statistical sample we currently have access to, a simpler rectangular cut based approach gives good results (to be discussed further in Part II).

categories are considered: "long" where both tracks from the $K^0_S \to \pi^+\pi^-$ decay produce hits in the VELO, we denote these K^0_S(LL), and "downstream" where neither does and use the notation K^0_S(DD). The long candidates have better mass, momentum and vertex resolution, so different selection requirements are imposed for candidates with the two types of K^0_S decay, and the ratios given in Eqs. (4.2) and (4.5) are determined independently for each. These are then combined taking into account the effects of systematic uncertainties that are correlated between the two categories. Finally, upper limits are set for modes where no significant signal is observed, and the absolute branching fractions are obtained by multiplying by the relevant normalisation factor.

In addition to the quantitative results obtained, in channels where significant signals are seen, the phase-space is inspected for resonant contributions from either exotic or conventional states. The presence or absence of resonances could guide future analyses. However, no attempt is made to determine the relative production rates of the different possible contributions. To avoid introducing a bias in the experimental procedures, the regions of the invariant mass distributions potentially containing previously unobserved decays were not inspected until after all analysis procedures were established. We now outline our method for each of these two parts in more detail.

Branching fraction ratio $\mathcal{B}(B^0 \to J/\psi\, K^0_S\pi^+\pi^-)/\mathcal{B}(B^0 \to J/\psi\, K^0_S)$

Since the most precise previous measurement of any of the $B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ branching fractions is $\mathcal{B}(B^0 \to J/\psi\, K^0\pi^+\pi^-) = (103 \pm 33 \pm 15) \times 10^{-5}$ [16], where the first uncertainty is statistical and the second is systematic, conversion of relative to absolute branching fractions would introduce a large normalisation uncertainty. Thus we need an alternative control channel to normalise to and so we use $B^0 \to J/\psi\, K^0_S$ for this purpose. The branching fraction of $B^0 \to J/\psi\, K^0_S$ is known to be [15]

$$\mathcal{B}(B^0 \to J/\psi\, K^0_S) = \frac{1}{2}(8.73 \pm 0.32) \times 10^{-4}\,, \tag{4.1}$$

where the factor of 2 comes from converting the $B^0 \to J/\psi K^0$ branching fraction to that with $K^0 \to K^0_S$. We aim (in "part I" of the analysis) to measure the $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ branching fraction relative to that for $B^0 \to J/\psi K^0_S$. This decay will have a similar trigger and muon identification efficiency to the signal but a different number of final-state particles. The following equation can be used to

calculate the respective branching fraction

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_S \pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_S)} = \frac{\epsilon_{B^0 \to J/\psi\, K^0_S}}{\epsilon_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}} \frac{N_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}}{N_{B^0 \to J/\psi\, K^0_S}}\,, \tag{4.2}$$

where ϵ represents the total efficiency, including acceptance, trigger, reconstruction and particle identification efficiencies and N gives the measured number of events. As with all analyses we wish to use data driven methods wherever possible. However, it is important not to optimise an analysis procedure on data that will be used for the measurement. Optimising a selection in this manner could introduce a bias as we have already maximised the very thing we wish to measure. To avoid introducing a bias we employ a selection strategy that is trained on $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ signal Monte Carlo. We also take some information from the side-bands of the data sample to provide a more representative description of the background.

As a by-product of "part I" of the analysis, we also study the contribution to $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ decays from $B^0 \to \psi(2S) K^0_S$ with $\psi(2S) \to J/\psi\, \pi^+\pi^-$. (The $\psi(2S)$ contribution is vetoed from the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ analysis, and therefore this study is done simply by reversing the veto.) The relevant branching fractions are known to be [15]

$$\mathcal{B}(B^0 \to \psi(2S) K^0_S) = (6.2 \pm 0.5) \times 10^{-4}/2\,, \tag{4.3}$$

$$\mathcal{B}(\psi(2S) \to J/\psi\, \pi^+\pi^-) = (33.6 \pm 0.4)\,\%\,, \tag{4.4}$$

and therefore this serves as a useful cross-check as well as providing the potential to improve the knowledge of $\mathcal{B}(B^0 \to \psi(2S) K^0_S)$.

Branching fraction ratio $\mathcal{B}(B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-})/\mathcal{B}(B^0 \to J/\psi\, K^0_S \pi^+\pi^-)$

As mentioned we can use the previously measured mode $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ as a proxy to optimise the rectangular cuts for both $B \to J/\psi\, K^0_S K^\pm \pi^\mp$ and $B \to J/\psi\, K^0_S K^+ K^-$ modes, and also as the normalisation channel. First we apply some very loose cuts to the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ mode in order to clean up the signal and remove some background. This allows us to perform a fit to the invariant mass spectrum $m(B^0 \to J/\psi\, K^0_S \pi^+\pi^-)$ where ${}_s\mathcal{W}eights$ can be computed. The ${}_s\mathcal{W}eights$ are calculated in the B candidate invariant mass range, $5245\,\mathrm{MeV}/c^2 < m(B^0 \to J/\psi\, K^0_S \pi^+\pi^-) < 5315\,\mathrm{MeV}/c^2$.

Using CROP, we optimise a selection on the control mode $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ using the ${}_s\mathcal{W}eights$ to correctly weight each event. The returned selection can then

be used as an initial estimator for all the other modes from which we derive a series of approximate signal and background scaling factors; based on the expected relative signal and background yields with respect to the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ control channel. These factors can be computed from simulation (signal) and data sidebands (background) and each in turn can be used to weight the ${}_s\mathcal{W}eight$-ed data and obtain approximate selections for each of the modes we are interested in, as discussed further in Part II Section 10.4. $B^0 \to J/\psi K^0_S \pi^+\pi^-$ is a good channel to tune the selections on as this has the same topology, thus we expect similar reconstruction, stripping and trigger efficiencies. All other branching fractions can then (in "part 2" of the analysis) be measured relative to that of $B^0 \to J/\psi K^0_S \pi^+\pi^-$, using

$$\frac{\mathcal{B}(B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-})}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} = \frac{\epsilon_{B^0 \to J/\psi K^0_S \pi^+\pi^-}}{\epsilon_{B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}}} \left(\frac{f_q}{f_d}\right)^{-1} \frac{N_{B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}}}{N_{B^0 \to J/\psi K^0_S \pi^+\pi^-}}, \tag{4.5}$$

where ϵ represents the total efficiency, including effects from acceptance, trigger, reconstruction, and selection and particle identification requirements and f_q/f_d is the relevant ratio of fragmention fractions ($f_s/f_d = 0.259 \pm 0.015$ [79, 80, 81]). The fragmentation fractions, f_q ($q = d, s$), denote the probability of a b-quark hadronising into a B^0_q meson. This process is intended to minimise the systematic uncertainties in each measurement. In particular, in Eq. 4.5 most sources of systematic uncertainties in the efficiencies cancel, apart from those due to particle identification requirements (in the case $h^+h^{(\prime)-} \neq \pi^+\pi^-$). This is where one, or both, of the bachelor kaons and/or pions is mis-identified and creates a shifted mass distribution that needs to be modelled. The particle misidentification, known as *cross-feed*, is reduced by choosing suitable $\mathrm{DLL}_{K\pi}$ and $\mathrm{DLL}_{p\pi}$ PID cuts for each bachelor which reduce the possibility of kaon candidates becoming pion candidates and vice versa. To get a handle on this we will use a simultaneous fit to all data sub-samples.

Part I

$B^0 \to J/\psi\, K^0_S \pi^+ \pi^-$ analysis

5

Backgrounds

In this section we discuss the possible sources of background in the samples, with the goal of identifying the dominant sources that need to be taken into account in the fit model. We break this section into two parts. The first deals with background channels for the $B^0 \to J/\psi\, K^0_S$ mode and the second looks at the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ mode. Recall that we fit the invariant mass spectrum in the range $5180 \leq M \leq 5500$ MeV/c^2.

5.1 $B^0 \to J/\psi\, K^0_S$ backgrounds

Several previous LHCb analyses have studied the $B^0 \to J/\psi\, K^0_S$ decay [82, 74, 75], and we benefit from their prior considerations of the backgrounds.

Table 5.1 shows the main background that we consider for $B^0 \to J/\psi\, K^0_S$ which comes from Λ baryons. This occurs when one of the Λ daughters is reconstructed under the pion mass hypothesis in which case the invariant mass can be consistent with that of the K^0_S meson. We use Monte Carlo to study this background, and a veto to remove it (discussed in Sec. 6.1). The left plot of Fig. 5.1 displays the number of events before applying the veto, then right plot after. Applying a Λ veto removes 3 % of the B^0 signal for the K^0_S(LL) mode, and similarly for the K^0_S(DD) sample just 6 % of the signal is removed. The number of events in the sidebands is reduced by over 40 % for each reconstructed K^0_S type.

Table 5.1: List of simulation decay files used for background characterisation of $B^0 \to J/\psi\, K^0_S$. The total number events consists of both magnet polarities.

Mode	Total events
$\Lambda^0_b \to J/\psi\, \Lambda$	7929474

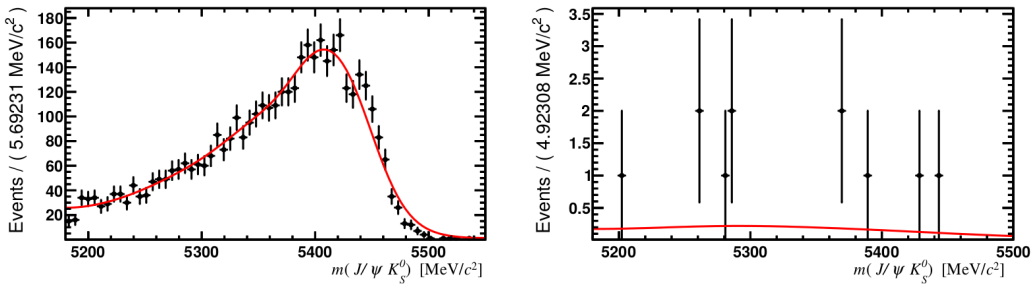

Figure 5.1: $\Lambda_b^0 \to J/\psi\,\Lambda$ simulated decays reconstructed in the $J/\psi\,K_S^0$ final state (left) before and (right) after a veto is applied. The top row shows the effect on the K_S^0(LL) reconstructed mode. The curves are only to guide the eye.

Following this veto, the remaining background is consistent with being combinatorial in nature. We note that in principle there should be a partially reconstructed background from $B_s^0 \to J/\psi\,K^{*0}(892)$ with $K^{*0}(892) \to K_S^0\pi^0$. The branching fraction has been measured to be [83]

$$\mathcal{B}(B_s^0 \to J/\psi\,K^{*0}(892)) = (4.4\,^{+0.5}_{-0.4} \pm 0.8) \times 10^{-5}\,, \tag{5.1}$$

of which 1/6 gives the $J/\psi\,K_S^0\pi^0$ final state.[1] Considering that only a fraction of this background will enter the fit window, it appears to be negligible (and indeed has been neglected in other LHCb analyses of the $J/\psi\,K_S^0$ final state). We note that contributions from B_s^0 decays to intermediate states other than $J/\psi\,K^{*0}(892)$ may also produce the $J/\psi\,K_S^0\pi^0$ final state, but these are expected to contribute even less to our fit window and are therefore neglected.

Another potential source of background which is rendered negligible by our selection requirements is from $B^0_{(s)} \to J/\psi\,\pi^+\pi^-$ decays. In particular, the K_S^0 flight distance significance requirement (in the LL category) removes any contribution from this decay or others with similar topologies.

5.2 $B^0 \to J/\psi\,K_S^0\pi^+\pi^-$ backgrounds

Since we expect $B^0 \to J/\psi\,K_S^0\pi^+\pi^-$ to give the largest yield of the various $B^0_{(s)} \to J/\psi\,K_S^0 h^+h^{(\prime)-}$ modes, potential backgrounds from cross-feed are less serious than they are for the other channels. The particle identification requirements on the bachelor pions further reduce the cross-feeds. Moreover, as mentioned in Sec. 6.1, we explicitly veto potential backgrounds from $\Lambda_b^0 \to J/\psi\,K_S^0 p\pi^-$ decays. We therefore

[1] This is the product of 1/3 for $K^* \to K^0\pi^0$ and 1/2 for $K^0 \to K_S^0$.

neglect residual cross-feed contributions at this stage.

Partially reconstructed backgrounds could also cause a significant contribution. We note that any that involve a $\psi(2S) \to J/\psi\pi^+\pi^-$ decay will be removed by the $\psi(2S)$ veto. Nonetheless, there are several observed decays that could cause backgrounds, such as $B^0 \to J/\psi K^0_S\omega$, $B^0 \to J/\psi K^*(892)^0\pi^+\pi^-$ and $B^+ \to J/\psi K^*(892)^+\pi^-\pi^+$ which have been measured by BaBar [84, 85], CDF [16] and CLEO [86], with branching fractions of $\sim 5 \times 10^{-4}$.

We study potential peaking backgrounds using MC. We use dedicated samples to study $B^0 \to J/\psi K^0_S\pi^+\pi^-\pi^0$ and $B^+ \to J/\psi K^0_S\pi^+\pi^-\pi^+$. In order to check that there are no other sources of peaking background, we analyse also the $b \to J/\psi X$ MC cocktail samples. The MC samples used in the study are listed in Table 5.2.

Table 5.2: List of simulation decay files used for background characterisation of $B^0 \to J/\psi K^0_S\pi^+\pi^-$. The total number events consists of both magnet polarities.

Mode	Total events
$B^0 \to J/\psi K^0_S\eta'$	541749
$B^0_s \to J/\psi K^0_S\eta'$	530249
$B^0 \to J/\psi K^0_S\eta$	534998
$B^0_s \to J/\psi K^0_S\eta$	537999
$B^0 \to J/\psi K^0_S\pi^+\pi^-\pi^0$	503996
$B^+ \to J/\psi K^0_S\pi^+\pi^-\pi^+$	514999
$B^0 \to J/\psi X$	$\approx$ 8M
$B^0_s \to J/\psi X$	$\approx$ 2M
$B^\pm \to J/\psi X^\pm$	$\approx$ 8M
$\Lambda^0_b \to J/\psi X$	$\approx$ 2M

The backgrounds from $B^0 \to J/\psi K^0_S\pi^+\pi^-\pi^0$ and $B^+ \to J/\psi K^0_S\pi^+\pi^-\pi^+$ decays are shown in Fig. 5.2, in the full mass range. It can be noted that the majority of the events have reconstructed mass below $5180\,\mathrm{MeV}/c^2$, which is the lower limit of our fit range. Above this value only a small tail remains, which can be absorbed into the combinatorial component.

Figures 5.3 display the simulated events from the $b \to J/\psi X$ cocktails that pass the full selection for the K^0_S(LL) category. A clear peak is seen in the $B^0 \to J/\psi X$ sample – but under further investigation this is caused by our signal channel. There is no evidence for any additional source of peaking background. In particular, for $\Lambda^0_b \to J/\psi X$ we see no candidates passing the selection criteria hence it does not appear in Figure 5.3.

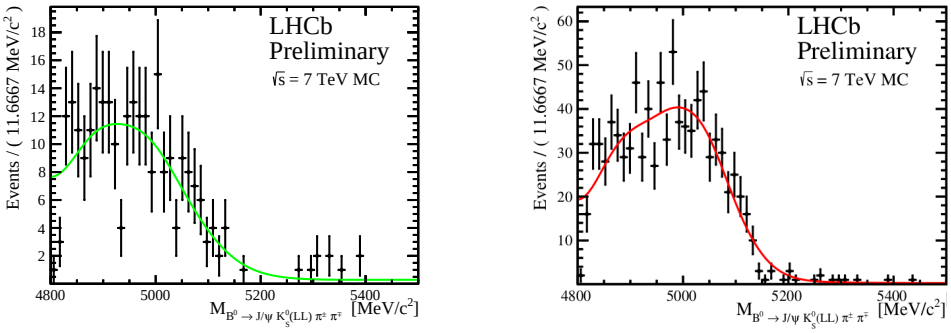

Figure 5.2: Simulated partially reconstructed background modes (left) $B^0 \to J/\psi K^0_S \pi^+\pi^-\pi^0$ and (right) $B^+ \to J/\psi K^0_S \pi^+\pi^-\pi^+$ when reconstructed as $B^0 \to J/\psi K^0_S \pi^+\pi^-$. Both samples above display the K^0_S(LL) category only.

We note that in principle there can be partially reconstructed backgrounds from $B^0_s \to J/\psi K^0_S \pi^+\pi^-\pi^0$ decays. The tail of the distribution of such events would be shifted up compared to that from B^0 and B^+ decays by the B^0_s–B^0 mass difference, and would therefore enter our fit range. However, using similar arguments as those used in the discussion of $B^0_s \to J/\psi K^0_S \pi^0$ backgrounds (Sec. 5.1), we consider that these are negligible in the current analysis and note that the branching fraction for $B^0_s \to J/\psi K^0_S \pi^+\pi^-\pi^0$ has not actually been measured.

5.3 Backgrounds from mis-reconstructed photons - $B^0 \to J/\psi K^0_S \eta'$

We consider decays of the form $B^0_{(s)} \to J/\psi K^0_S \eta^{(\prime)}$ where $\eta^{(\prime)} \to \pi^+\pi^-\gamma$. After reconstruction we can sometimes loose a photon from the final state and this loss can form a background component to our decay mode. By inspection of figures 5.4 and 5.5, which represent the B^0 and B^0_s channels respectively, we see that most of the these decays do not cause us much concern and lie outside the mass fitting range and as such there is no peaking background except one, $B^0_s \to J/\psi K^0_S \eta$ as seen in figure 5.5. First one should note that the decay $B^0_s \to J/\psi K^0_S \eta$ has not been observed however we can gauge an estimate of its relative contribution using the following formula

$$\frac{N_{\rm bkg}}{N_{B^0 \to J/\psi K^0_S \pi^+\pi^-}} = \frac{f_x}{f_d} \times \frac{\mathcal{B}({\rm bkg})}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} \times \frac{\epsilon_{\rm bkg}}{\epsilon_{B^0 \to J/\psi K^0_S \pi^+\pi^-}}\,, \qquad (5.2)$$

where $N_{\rm bkg}$ is the expected number of events that may peak in the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ dataset, f_x is the b-quark fragmentation fraction responsible for the peaking back-

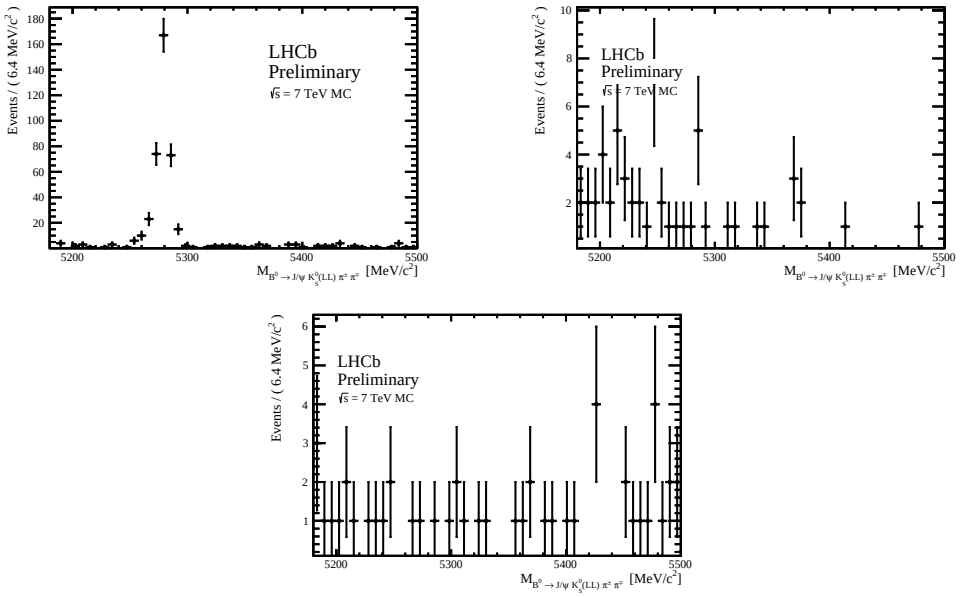

Figure 5.3: Analysis of various cocktail MC samples to see if there are any peaking backgrounds we did not consider. Reconstructed under the $B^0 \to J/\psi K^0_{\rm S}({\rm LL})\pi^+\pi^-$ invariant mass hypothesis. (Top left) $B^0 \to J/\psi X$, (top right) $B^0_s \to J/\psi X$ and (bottom) $B^+ \to J/\psi X^+$.

ground, $\mathcal{B}(\text{bkg})$ is the branching fraction taken from the PDG (or estimated in this case) and ϵ_{bkg} is the selection efficiency which we take simply as the number of events passing the cuts in the fitted range. Table 5.3 summaries the efficien-

Table 5.3: Estimates for potential peaking background coming from the mode $B^0_s \to J/\psi K^0_{\rm S}\eta$.

Mode	Branching fraction	Efficiency
$B^0 \to J/\psi K^0_{\rm S}({\rm LL})\pi^+\pi^-$	0.5×10^{-3}	2.2×10^{-3}
$B^0 \to J/\psi K^0_{\rm S}({\rm DD})\pi^+\pi^-$	0.5×10^{-3}	5.4×10^{-3}
$B^0_s \to J/\psi K^0_{\rm S}({\rm LL})\eta$	10^{-4}	2.6×10^{-4}
$B^0_s \to J/\psi K^0_{\rm S}({\rm DD})\eta$	10^{-4}	5.4×10^{-4}

cies and branching fractions for $B^0 \to J/\psi K^0_{\rm S}\pi^+\pi^-$ and $B^0_s \to J/\psi K^0_{\rm S}\eta$. The maximum value for the relative pollution of this potentially peaking background is $\frac{N_{B^0_s \to J/\psi K^0_{\rm S}({\rm LL})\eta}}{N_{B^0 \to J/\psi K^0_{\rm S}({\rm LL})\pi^+\pi^-}} \approx \frac{N_{B^0_s \to J/\psi K^0_{\rm S}({\rm LL})\eta}}{N_{B^0 \to J/\psi K^0_{\rm S}({\rm DD})\pi^+\pi^-}} \approx 0.1\,\%$ where we have assumed a conservative branching fraction[2] for $B^0_s \to J/\psi K^0_{\rm S}\eta$ at the order of 10^{-4}, given the CKM

[2] It is conservative since we would expect a branching fraction of the order $(8 \pm 4) \times 10^{-5}$ as found in the $B^0 \to J/\psi K^0_{\rm S}\eta$ mode.

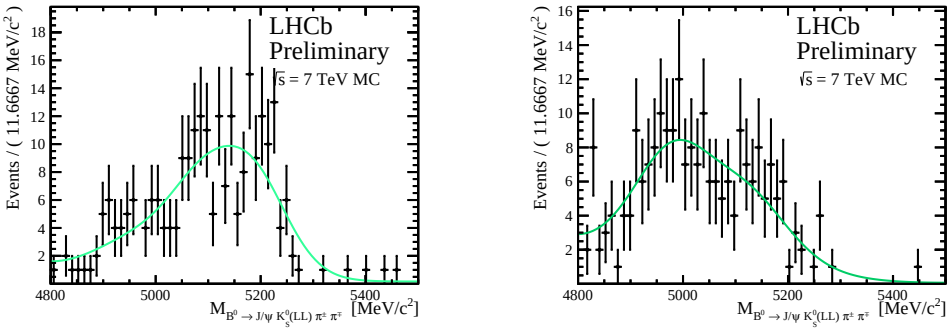

Figure 5.4: Simulated partially reconstructed backgrounds from (left) $B^0 \to J/\psi K^0_S \eta$ and (right) $B^0 \to J/\psi K^0_S \eta'$ events that pass all selection requirements and are in the mass range used in the fit. Both samples above display the K^0_S(LL) category only.

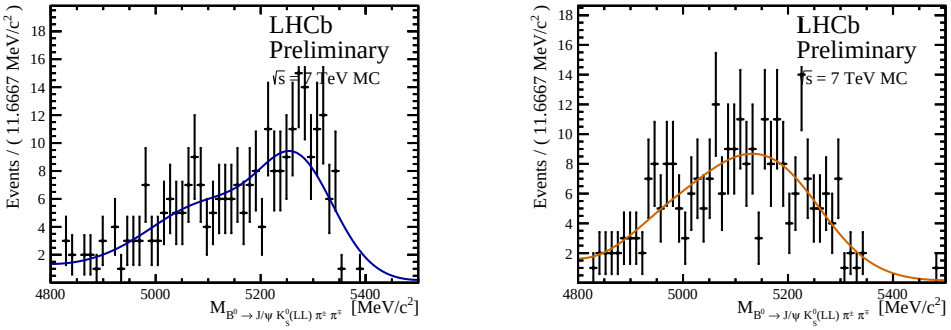

Figure 5.5: Simulated partially reconstructed backgrounds from (left) $B^0_s \to J/\psi K^0_S \eta$ and (right) $B^0_s \to J/\psi K^0_S \eta'$ events that pass all selection requirements and are in the mass range used in the fit. Both samples above display the K^0_S(LL) category only.

elements involved in the transition. The small value of the contamination means we can neglect this background, moreover since this mode is not observed it makes sense not to include it in the fit model.

6

Final selection requirements

Throughout the analysis we use Decay Tree Fitter [87] (DTF) which constrains the B candidate to originate from the primary vertex (PV) and that the decay products come from a common B vertex with the J/ψ and K^0_S masses fixed to their (PDG) values respectively. We remove candidates where the DTF fit does not converge. We find on data that for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes, $K^0_S(\mathrm{DD}) : 0.49\%$ and $K^0_S(\mathrm{LL}) : 0.15\,\%$ of events fail the DTF requirement. Likewise a similar story for $B^0 \to J/\psi K^0_S$ modes we find; $K^0_S(\mathrm{DD}) : 5.0\,\%$ and for $K^0_S(\mathrm{LL}) : 2.0\,\%$ fail this cut.

It is possible, in the LL sample, to have peaking backgrounds where the two tracks that form the K^0_S candidate are in fact bachelor tracks (*i.e.* they originate directly from the B decay) or come from intermediate charmed states with short, but non-negligible lifetimes. In order to remove such backgrounds a requirement is imposed (on the LL sample only) that the K^0_S vertex separation with respect to the B vertex satisfies $\chi^2 > 16$. We have an additional particle identification requirement on the bachelor pions in the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode, which is

- (π^+, π^-) PID cut: $\mathrm{DLL}_{K\pi} < 0$ and $\mathrm{DLL}_{p\pi} < 10$.

Since the performance of the RICH is optimal in the momentum range $0 < p < 100$ GeV/c, we set this as an upper bound for the momentum of any bachelor track. This corresponds to a loss of around 195 events for $K^0_S(\mathrm{LL})$ and 450 events for $K^0_S(\mathrm{DD})$ when applied after the $\psi(2S)$ veto (see Table 8.5).

6.1 Vetoes

As discussed in Sec. 5, there are potential backgrounds from b-baryon decays. These are removed by applying vetoes.

The first is misidentification of Λ decays as K^0_S candidates. These are removed with a cut of $|m(p\pi^-) - m_\Lambda| > 10\,\text{MeV}/c^2$ for the LL and $|m(p\pi^-) - m_\Lambda| < 25\,\text{MeV}/c^2$ for DD modes, as illustrated in Fig. 6.1. The veto window is wider for DD modes since the proton mass assignment at the tuple level leads to a smearing of the reconstructed $p\pi$ mass that is larger for DD candidates. This is applied to K^0_S candidates in both $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ samples. The remaining

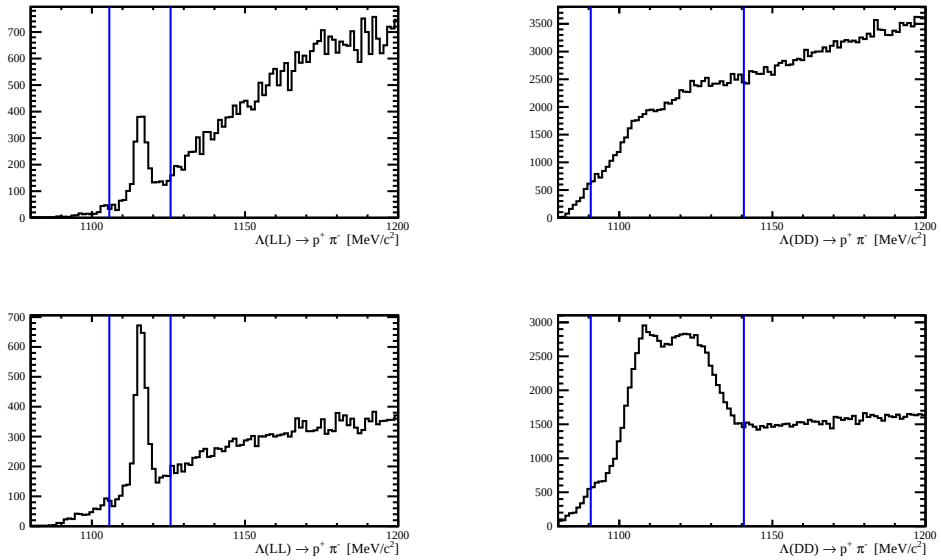

Figure 6.1: Candidate (top) $B^0 \to J/\psi K^0_S \pi^+\pi^-$ and (bottom) $B^0 \to J/\psi K^0_S$ decays in the whole invariant mass range reconstructed with either pion track from the K^0_S candidate assigned the proton mass hypothesis. (Left) K^0_S(LL) sample and (right) K^0_S(DD) type. The blue lines show $|m(p\pi^-) - m_\Lambda|$ where the cut is at $\pm 10\,\text{MeV}/c^2$ for K^0_S(LL) and $\pm 25\,\text{MeV}/c^2$ for K^0_S(DD)).

vetoes discussed below only apply to the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode.

The second potential source of baryonic background is from $\Lambda^0_b \to J/\psi K^0_S p\pi^-$ decays. Although such contributions should be suppressed due to the particle identification requirement applied to the bachelor pions, the branching fraction of this decay is unknown. In principle this background could be handled in a similar way to the cross-feed backgrounds, but since we are not considering baryonic decays in this analysis, we prefer instead to apply a veto. We remove events that have invariant mass after either bachelor track is assigned the proton mass hypothesis in the range $\left|m(J/\psi K^0_S p\pi^-) - m_{\Lambda^0_b}\right| < 25\,\text{MeV}/c^2$, as illustrated in Fig. 6.2. Several LHCb analyses [88] have used a veto to remove any resonant K^0_S contributions coming from the combined invariant mass of the bachelor pions. We also employ a K^0_S veto to

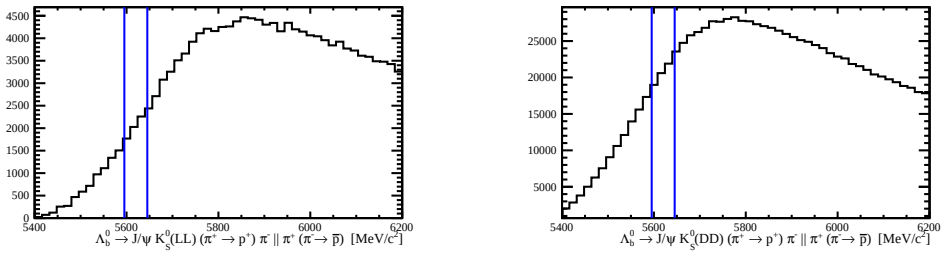

Figure 6.2: Candidate $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ decays in the whole invariant mass range reconstructed with either bachelor track assigned the proton mass hypothesis. (Left) LL sample, (right) DD sample. The blue lines show the vetoed region.

remove $B^0_{(s)} \to J/\psi\, K^0_S K^0_S$ contribution, $\left|m(\pi^+\pi^-) - m_{K^0_S}\right| < 25\,\text{MeV}/c^2$.[1] Finally, the decay chain $B^0 \to \psi(2S)K^0_S$ with $\psi(2S) \to J/\psi\,\pi^+\pi^-$ gives a large contribution to the $J/\psi\, K^0_S\pi^+\pi^-$ final state. This resonant contribution is well-known and is not included in our study of the $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ decay. It is removed with a $\psi(2S)$ veto within the mass range, $\left|m(J/\psi\,\pi^+\pi^-) - m_{\psi(2S)}\right| < 15\,\text{MeV}/c^2$. We will however use the corresponding anti-veto to make a measurement of the $B^0 \to \psi(2S)K^0_S$ branching fraction in this part of the analysis.

The veto requirements are summarised in Table 6.1.

Table 6.1: Veto requirements. The known values of all particles are taken from the PDG [15]. All cuts are the same for LL and DD samples except where indicated otherwise.

Cut	Value
Λ(LL)	$\lvert m(p\pi^-) - m_\Lambda\rvert > 10\,\text{MeV}/c^2$
Λ(DD)	$\lvert m(p\pi^-) - m_\Lambda\rvert > 25\,\text{MeV}/c^2$
Λ_b^0	$\left\lvert m(J/\psi\, K^0_S p\pi^-) - m_{\Lambda_b^0}\right\rvert > 25\,\text{MeV}/c^2$
$\psi(2S)$	$\left\lvert m(J/\psi\,\pi^+\pi^-) - m_{\psi(2S)}\right\rvert > 15\,\text{MeV}/c^2$
K^0_S	$\left\lvert m(\pi^+\pi^-) - m_{K^0_S}\right\rvert > 25\,\text{MeV}/c^2$

6.2 Data – Monte Carlo comparisons

To aid the reader, please note there is a table provided in Appendix A that outlines any variable alias's used in this analysis along with their definitions. We use signal

[1] It would be of interest to search for $B^0_{(s)} \to J/\psi\, K^0_S K^0_S$ decays, but this requires a dedicated analysis.

MC to optimise the selection for the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ channel. An almost identical selection is applied to the $B^0 \to J/\psi K^0_S$ mode (the normalisation channel), thus reducing systematic errors. In this section we validate this strategy by comparing data with MC for the variables of interest to us. Figure. 6.3 shows that for most variables the distributions are in very good agreement for the K^0_S(LL) catgeory The background has also been shown for comparison and a similar picture is found for the K^0_S(DD) distributions. Due to the relatively good agreement between signal MC and data, we are confident that optimisation on MC will allow us to obtain a reasonable set of cuts to apply to data.

6.3 CROP optimisation

The final selection requirements are obtained using CROP [89]. The signal sample is taken from $B^0 \to J/\psi K^0_S \pi^+\pi^-$ Monte Carlo (since using data here would risk causing a bias in the branching fraction measurement), and the background sample is taken from the upper-upper side band region mentioned in Sec. 4. Figures 6.4 and 6.5 show the correlation between variables that are used in the optimisation for the K^0_S (LL) and K^0_S (DD) modes respectively. Large correlations would indicate that some of the variables should not be used in the procedure (or alternatively could indicate that significant gains could be expected if a multivariate analyser such as a neural network or boosted decision tree was used). The correlations are generally found to be fairly small.

Table 6.2 shows the outcome of the CROP optimisation process. It provides the optimised selection criteria for all variables used as input. The variables are listed in descending order by their separation power the K^0_S(LL) (while their ranking is explicitly labelled for the K^0_S(DD)). The separation power is defined as

$$\delta(v) = \frac{1}{2}\int \frac{(f_s(v) - f_b(v))^2}{f_s(v) + f_b(v)}\,, \tag{6.1}$$

the K^0_S (LL) (while their ranking is explicitly labelled for the K^0_S (DD)) where f_s, f_b are the signal and background distributions of variable v, for distributions that well separated, $\delta(v)$. We use eleven variables in total for the K^0_S(LL) modes and 10 for the K^0_S(DD). The difference arises in that there is no point constraining the decay length of the K^0_S in the downstream kaons as we know that they are sufficiently displaced from the primary vertex or the B decay vertex. We also present the variable distributions in Figures 6.6 and 6.7 for each of the DD and LL K^0_S modes respectively.

Table 6.2: Cuts for the optimised selection on simulation for both DD and LL K^0_S types. The list is in order of significance for the LL mode, for the DD mode the numbering is labelled in open brackets next to the cut value.

Cut	$B^0 \to J/\psi K^0_S \pi^+\pi^-$	
	K^0_S(LL)	K^0_S(DD) (rank)
K^0_S significance of separation wrt related PV, $\chi^2_{\rm VD}$	> 6918	-
B^0 vertex probability, $P(\chi^2, \text{ndf})$	$> 5.6 \times 10^{-4}$	$> 1.29 \times 10^{-2}$ (3)
π minimum IP significance wrt related PV, $\min(\pi^+_{\chi^2_{\rm IP}}, \pi^-_{\chi^2_{\rm IP}})$	> 9	> 11 (2)
K^0_S IP significance with respect to PV, $\chi^2_{\rm IP}$	> 4.0	< 288 (6)
B^0 pointing angle, $\cos(\theta)$	> 0.999973	0.999891 (8)
J/ψ significance of separation wrt related PV, $\chi^2_{\rm VD}$	> 66	> 206 (5)
J/ψ IP significance with respect to PV, $\chi^2_{\rm IP}$	> 0.54	$> 1.5 \times 10^{-3}$ (10)
B^0 transverse momentum, $p_{\rm T}$,	> 18 MeV	> 17 MeV (9)
B^0 IP significance with respect to PV, $\chi^2_{\rm IP}$	-	< 4 (1)
K^0_S vertex probability, $P(\chi^2, \text{ndf})$	$> 2.5 \times 10^{-6}$	$> 2.2 \times 10^{-2}$ (4)
J/ψ vertex probability, $P(\chi^2, \text{ndf})$	$> 2.5 \times 10^{-5}$	$> 1.6 \times 10^{-3}$ (7)

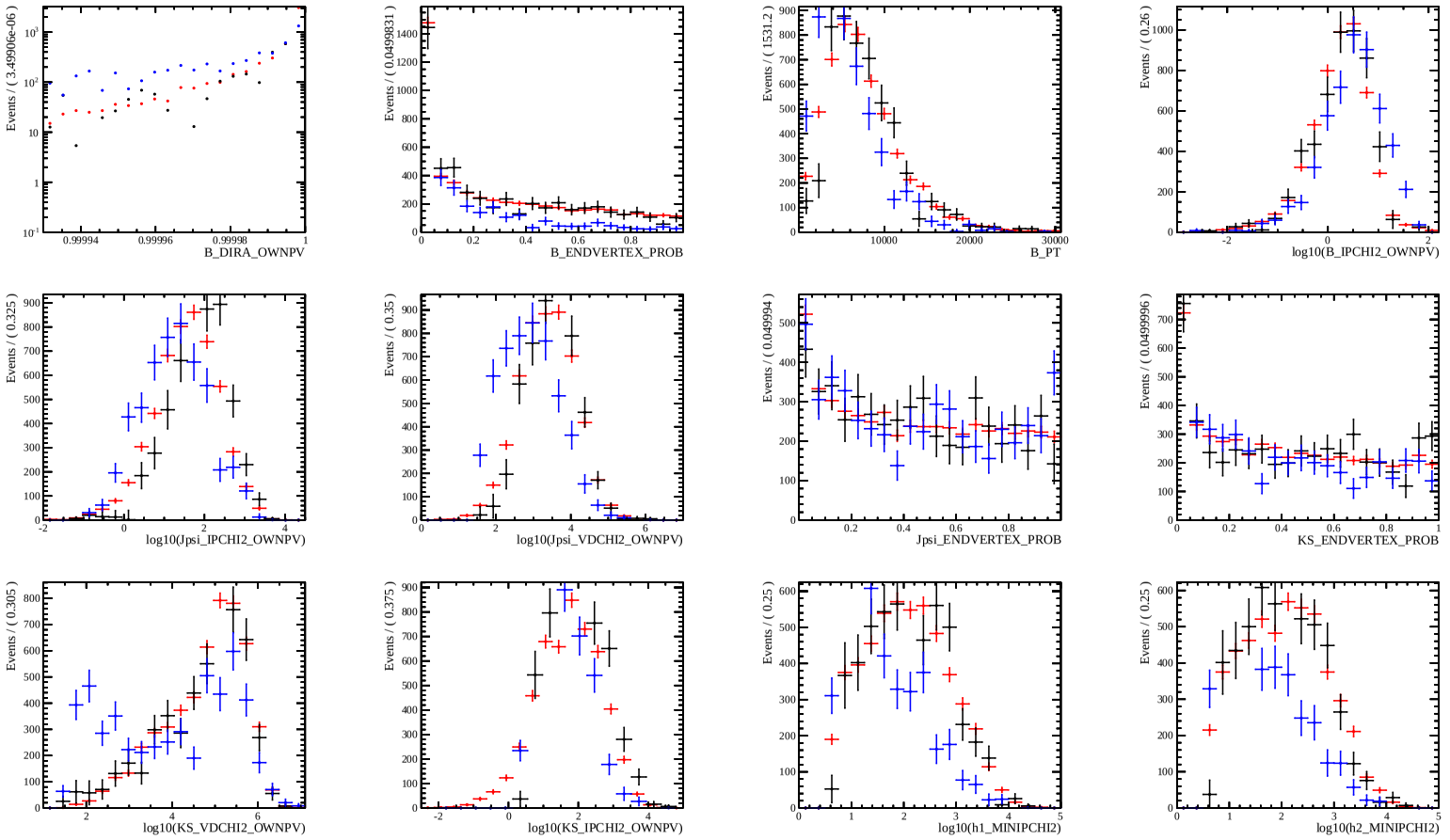

Figure 6.3: $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ normalised distributions comparing data and Monte Carlo, red line indicates MC and black is for signal ${}_s\mathit{Weighted}$ data and blue is for background ${}_s\mathit{Weighted}$ data. Units are omitted. There is a relatively good agreement between signal like variables.

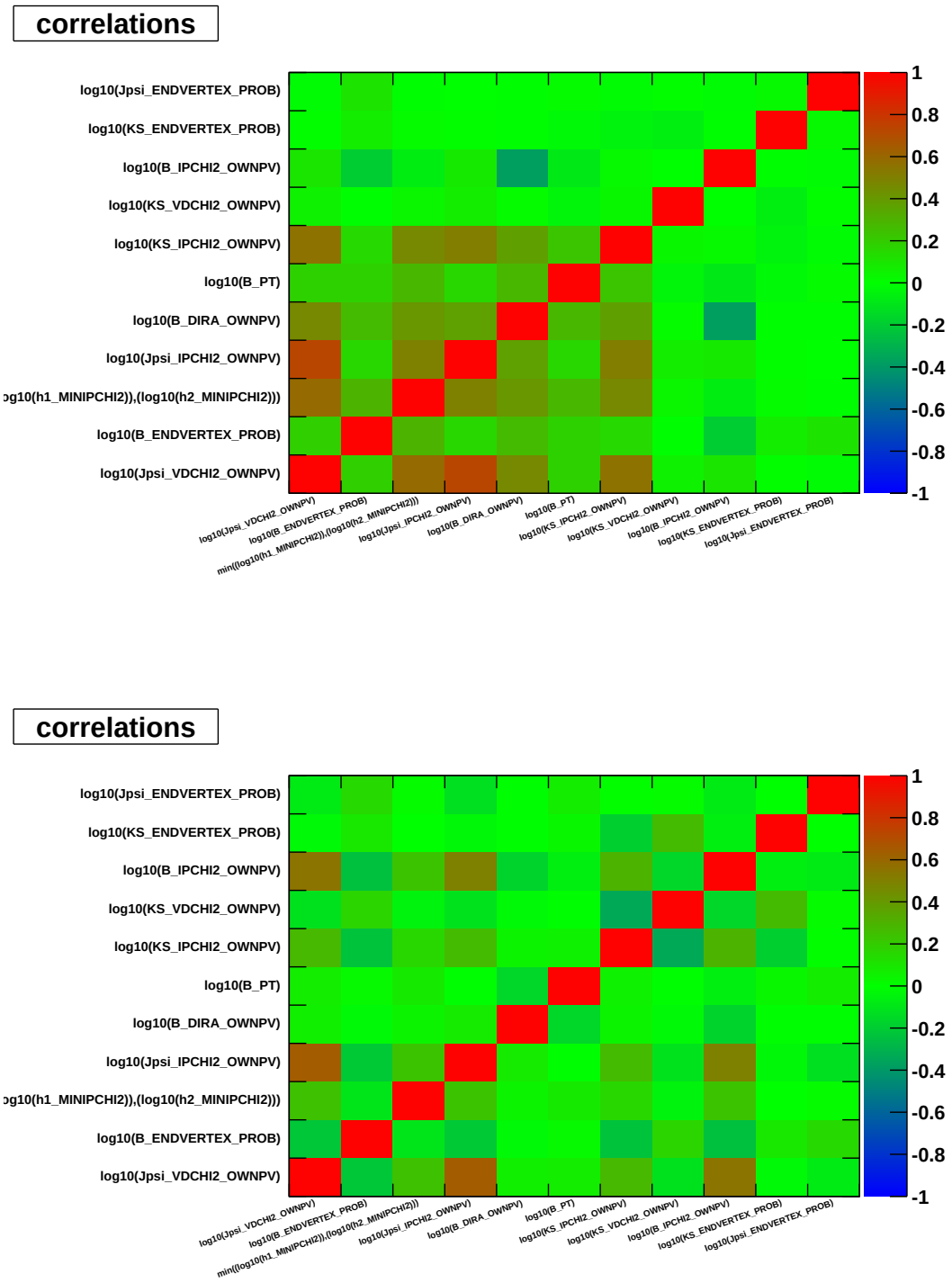

Figure 6.4: Correlations between training variables for $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$; (top) signal MC, (bottom) data sideband.

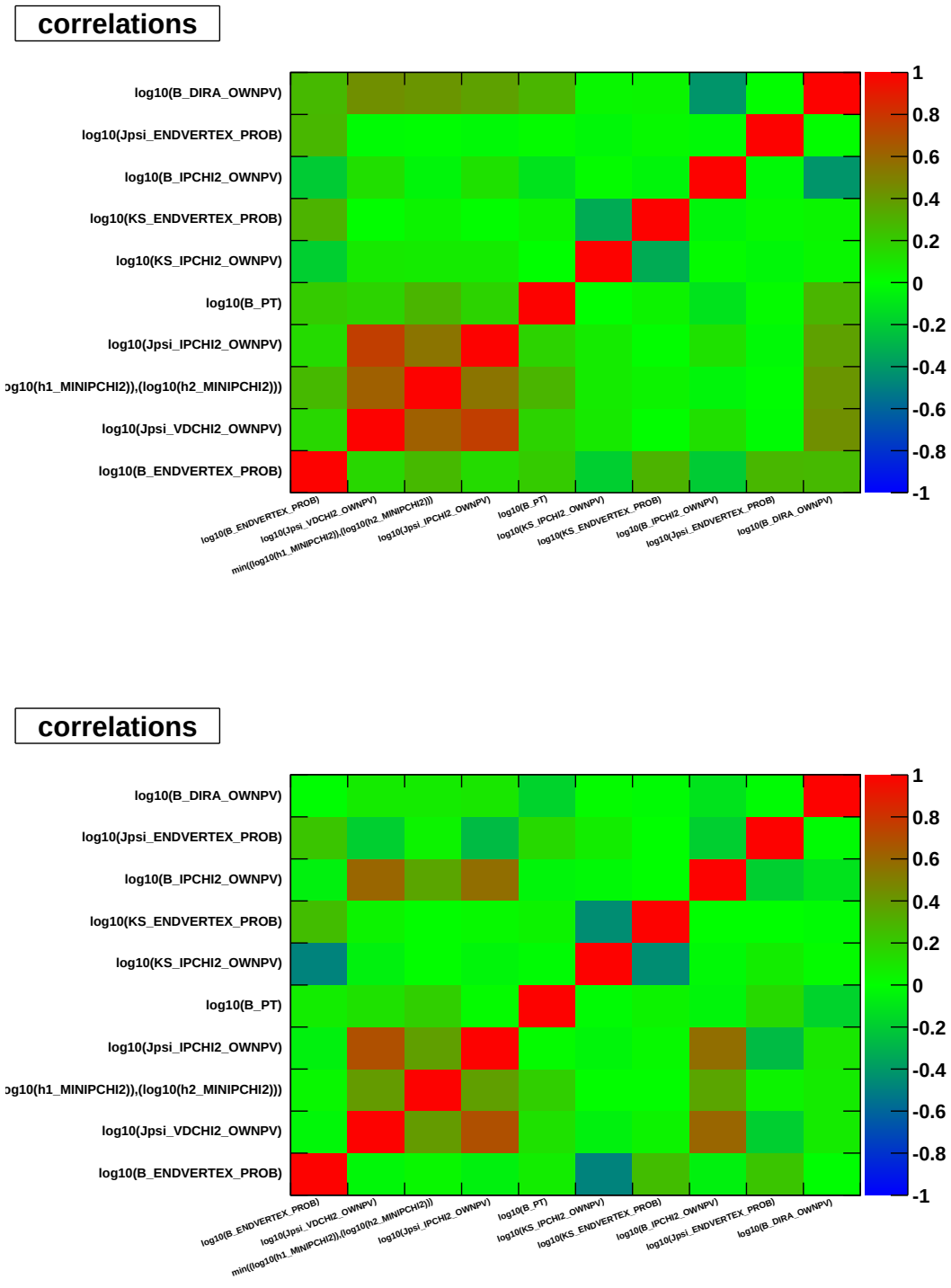

Figure 6.5: Correlations between training variables for $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$; (top) signal MC, (bottom) data sideband.

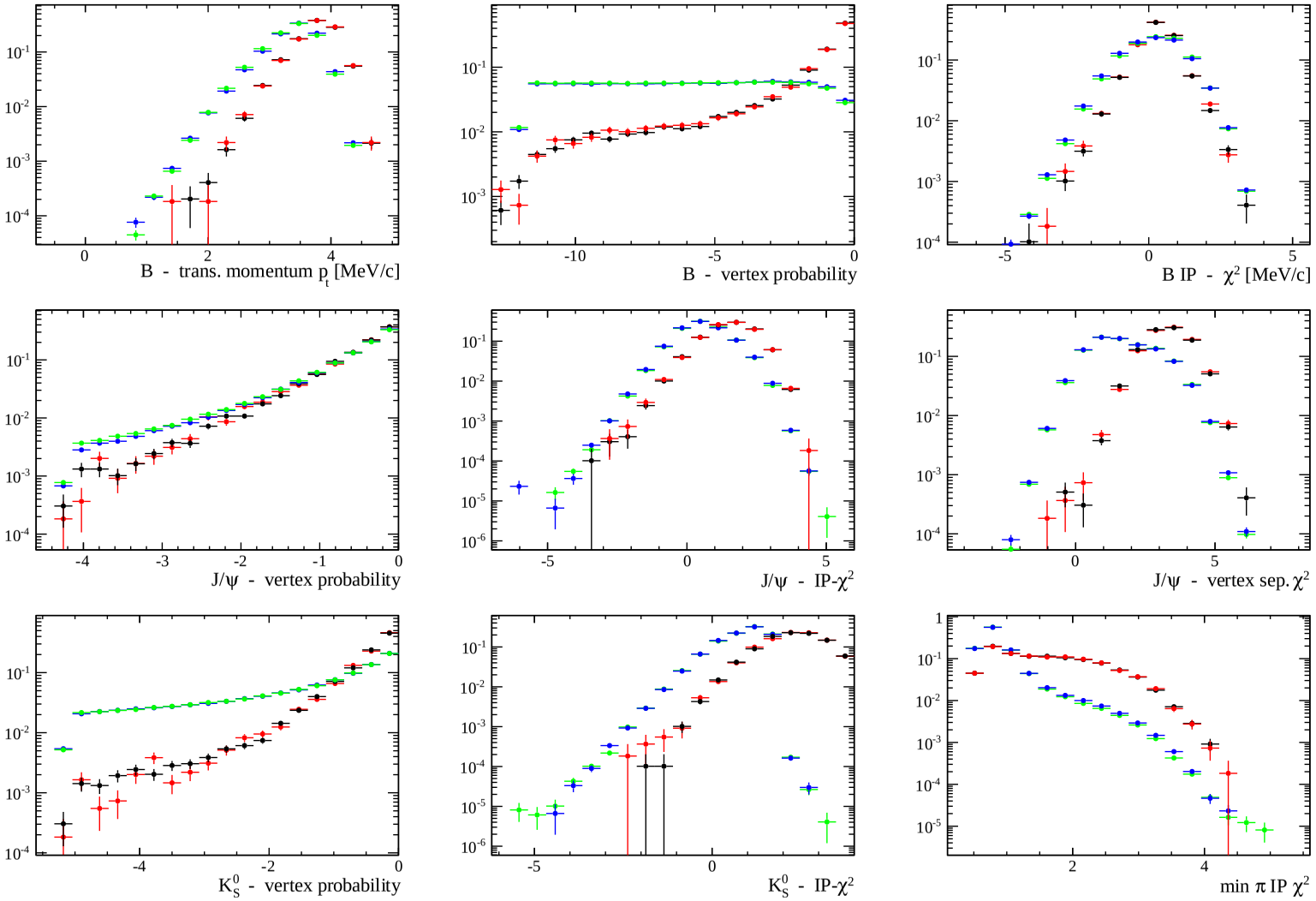

Figure 6.6: Separation of variables for the $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ mode where green and blue is signal for different magnet polarities, red and black are background from the two magnet polarities.

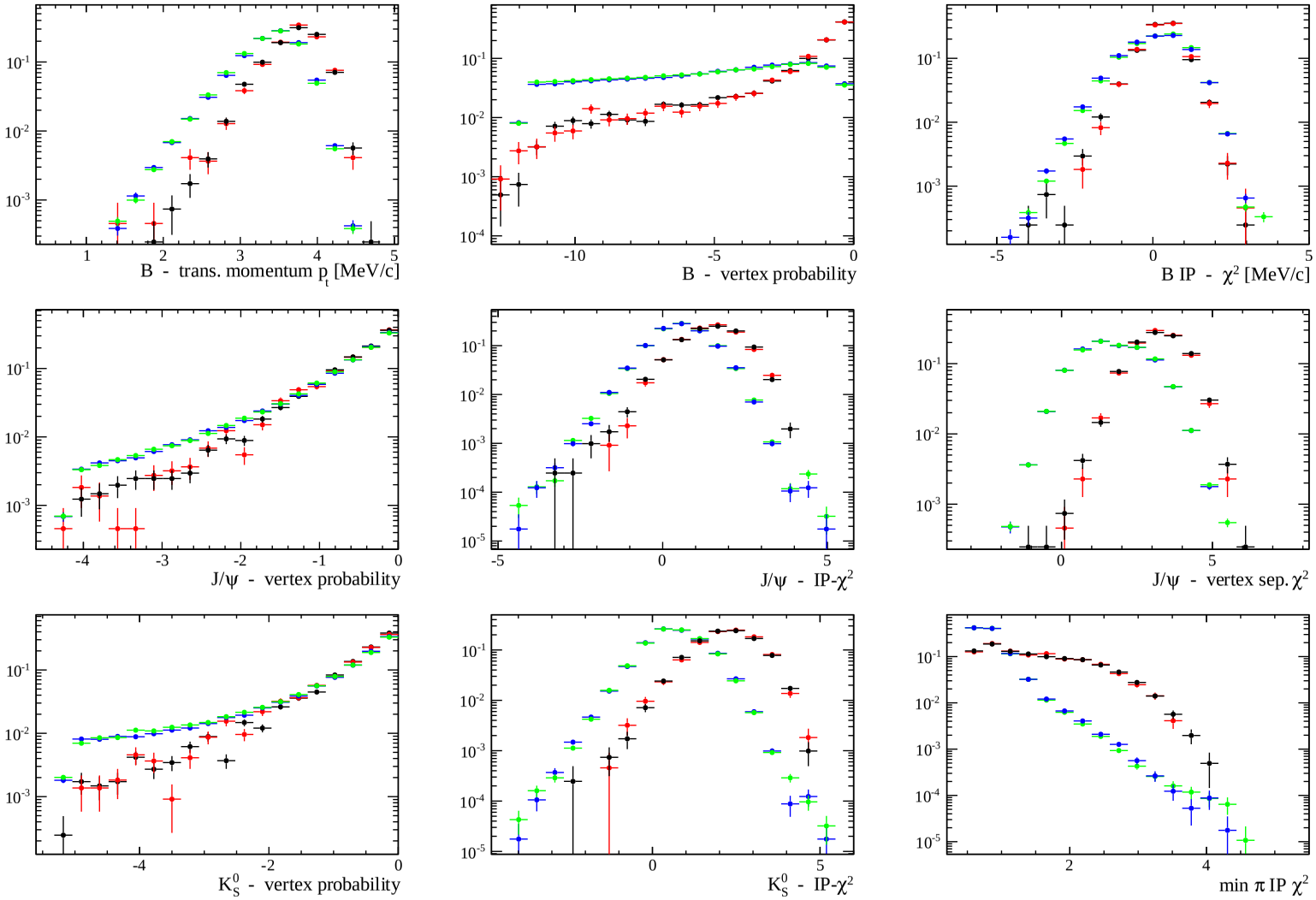

Figure 6.7: Separation of variables for the $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ mode where green and blue is signal for different magnet polarities, red and black are background from the two magnet polarities.

7

Fit model and yields

We have used similar models to fit each of the $B^0 \to J/\psi\, K^0_S$ and $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ modes. As described in earlier sections, after all selection requirements, there are no significant sources of peaking background, so the only components in the fits are due to B^0 and B^0_s signals and combinatorial background.

For both $B^0 \to J/\psi\, K^0_S$ and $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ modes we are using a similar fit model to that used in similar LHCb analyses [75]; a double Crystal Ball probability density function (PDF) to approximate the signal shape, with the background parametrised by an exponential function. The $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ fit is discussed in more detail below. For all modes we extract the event yields from independent extended unbinned maximum likelihood fits to the mass distributions of the reconstructed candidates in the DD and LL samples. We fit candidates with masses in the range $5180 < m_{J/\psi\, K^0_S \pi\pi} < 5500\,\mathrm{MeV}/c^2$ to reduce contributions from partially reconstructed backgrounds.[1] Any region that is yet to be observed is blinded (this only affects $B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$ and $B^0_s \to \psi(2S)\, K^0_S$ signal regions). When obtaining goodness of fit variables such as the χ^2, the value is taken from the unblinded region only of an associated plot.

7.1 Signal model

Each individual signal shape is parametrised by a double-sided Crystal Ball (CB) function [90] with common mean value. The parameters describing the CB tails are taken from simulation; all other parameters of the CB are allowed to vary in the fit.

$$\mathcal{P} = (1-f) \times \mathrm{CB}(m; \mu_{B^0}, \alpha_1, n_1, \sigma_1) + f \times \mathrm{CB}(m; \mu_{B^0}, \alpha_2, n_2, \sigma_2) \tag{7.1}$$

[1] In fact we are limited to this range due to the mass cut that is imposed in the $B^0 \to J/\psi\, K^0_S$ stripping line selection, thus this also avoids any edge effects.

where the CB function is defined by

$$CB(m;\mu,\alpha,n,\sigma) = \mathcal{N}\cdot\begin{cases} \exp(-(m-\mu)^2/2\sigma^2) & \text{if } (m-\mu)/\sigma > -\alpha \\ (\frac{n}{|\alpha|})^n \exp(-\alpha^2/2)(\frac{n-\alpha^2}{|\alpha|} - \frac{m-\mu}{\sigma}) & \text{if } (m-\mu)/\sigma \leq -\alpha, \end{cases} \tag{7.2}$$

where μ and σ are the mean value and resolution of the Gaussian part of the function, respectively, and $\mathcal{N}$ the normalization of the function. The sign of the α parameter governs whether the tail is to the upper or lower side, while its magnitude describes the value at which it switches on and the parameter n drives the power law associated to the decrease of the tail. In the fit to data, the tail parameters α and n are fixed according to values extracted from a fit to the simulated signal events, with the same selection requirements as the data. The mean μ, widths of each CB σ_1 and σ_2 as well as the relative fractions of each CB f, are floated in the fit to data. A double Gaussian function was tested but did not provide a better description. The mean value of the B_s^0 signal PDFis fixed via the $m_{B_s^0}$-m_{B^0} mass difference given by the PDG value ($87.35 \pm 0.23\,\text{MeV}/c^2$) [15].

Figure 7.1 shows the result of the fits to the simulated samples, and Table 7.1 gives the fitted parameters as well as the goodness of fit.

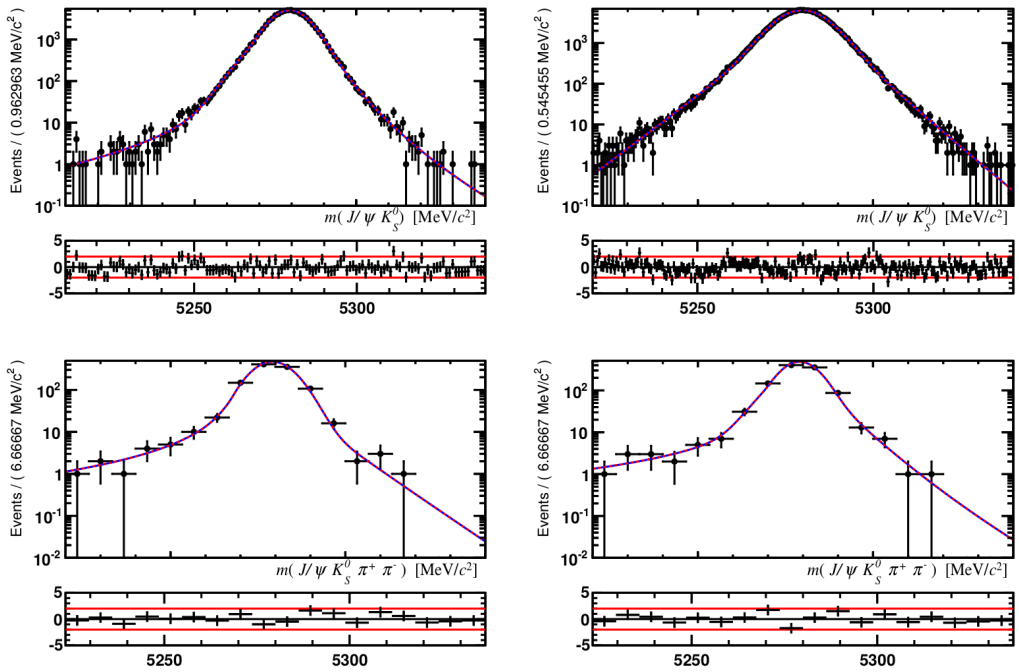

Figure 7.1: Results of the fits to the MC samples used to obtain the signal PDF parameters. (Top) $B^0 \to J/\psi K^0_S$ and (bottom) $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (Left) LL and (right) DD samples.

Table 7.1: Fitted parameters and goodness of fit obtained from fits to MC for both $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. The yield which was also floated is not included in this table. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations.

Parameter	$B^0 \to J/\psi K^0_S$		$B^0 \to J/\psi K^0_S \pi^+\pi^-$	
	LL	DD	LL	DD
$\mu_{B^0}(\,\text{MeV}/c^2)$	5279.4 ± 0.025	5279.8 ± 0.03	5279.2 ± 0.20	5279.0 ± 0.21
f	0.766 ± 0.027	0.473 ± 0.02	0.86 ± 0.12	0.57 ± 0.02
n_1	7.977 ± 0.071	76.2 ± 27.40	1.87 ± 0.13	1.30 ± 0.78
α_1	2.018 ± 0.029	1.116 ± 0.01	1.91 ± 0.43	1.85 ± 0.39
$\sigma_1(\,\text{MeV}/c^2)$	5.258 ± 0.112	7.028 ± 0.08	5.77 ± 0.31	4.58 ± 0.63
n_2	9.713 ± 0.368	128.5 ± 9.00	134.17 ± 4.72	19.18 ± 70.9
α_2	-1.968 ± 0.112	-1.168 ± 0.01	-0.80 ± 0.48	-1.41 ± 0.60
$\sigma_2(\,\text{MeV}/c^2)$	8.710 ± 0.245	6.618 ± 0.07	6.52 ± 3.77	7.22 ± 1.09
χ^2	106.7	213.7	6.5	5.6
# floated params.	9	9	9	9
# bins	135	200	18	18
$P(\chi^2, \text{ndf})$	0.89	0.13	0.69	0.78

7.2 Combinatorial background model

We use a single exponential function to describe the combinatorial background. This model is purely empirical – based on what we see in the data. We use an exponential shape to model the combinatorial background for both $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$, with the slope parameter labelled c.

7.3 Fit results

The full fit has been implemented for all of the $J/\psi K^0_S$ (Figs. 7.2 and 7.3), $J/\psi K^0_S \pi^+\pi^-$ (Figs. 7.4 and 7.5) and $\psi(2S)K^0_S$ modes (Figs. 7.6 and 7.7). In the $J/\psi K^0_S \pi^+\pi^-$ and $\psi(2S)K^0_S$ final states the B^0_s region is blinded. For the $J/\psi K^0_S$ mode we include the B^0_s signal which is modelled using the same parameters as the B^0 signal PDF. Tables 7.2, 7.3 and 7.4 summarise the fit results.

For the $B^0 \to \psi(2S)K^0_S$ modes, the selection (which, to reiterate, is the same as that for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ except for the inversion of a veto) leaves very little

Table 7.2: Results of the fits to the $B^0 \to J/\psi K^0_S$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations.

Parameter	$B^0 \to J/\psi K^0_S(\text{DD})$	$B^0 \to J/\psi K^0_S(\text{LL})$
N_{comb}	1926 ± 64	145 ± 17
c	$(-1.76 \pm 0.29) \times 10^{-3}$	$(-3.30 \pm 1.04) \times 10^{-3}$
N_{B^0}	9870 ± 107	4869 ± 71
$N_{B^0_s}$	115 ± 20	75 ± 10
$\mu_{B^0}(\text{MeV}/c^2)$	5281.9 ± 0.1	5281.1 ± 0.1
f	0.57 ± 0.09	0.67 ± 0.05
$\sigma_1(\text{MeV}/c^2)$	6.55 ± 0.37	6.16 ± 0.19
$\sigma_2(\text{MeV}/c^2)$	11.75 ± 0.75	8.16 ± 0.51
χ^2	141.8	45.1
# floated params.	8	8
# bins	145	60
$P(\chi^2, \text{ndf})$	0.37	0.74

Table 7.3: Results of the fit to the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations.

Parameter	$B^0 \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$	$B^0 \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$
N_{comb}	278 ± 23	95 ± 13
c	$(-7.20 \pm 7.30) \times 10^{-4}$	$(-7.65 \pm 12.1) \times 10^{-5}$
N_{B^0}	483 ± 26	269 ± 18
$\mu_{B^0}(\text{MeV}/c^2)$	5281.7 ± 0.4	5281.2 ± 0.49
f	0.14 ± 0.09	0.79 ± 0.16
$\sigma_1(\text{MeV}/c^2)$	16.75 ± 13.30	5.86 ± 0.52
$\sigma_2(\text{MeV}/c^2)$	6.09 ± 0.42	9.86 ± 4.58
χ^2	36.2	22.0
# floated params.	8	8
# bins	45	35
$P(\chi^2, \text{ndf})$	0.51	0.74

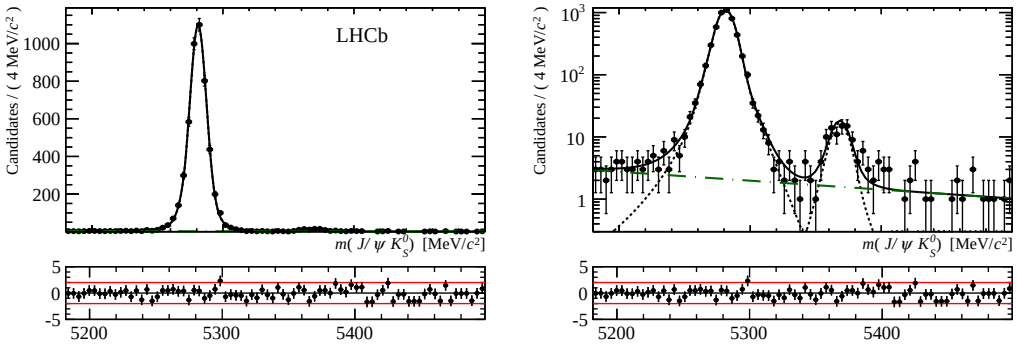

Figure 7.2: Fit results for the $B^0 \to J/\psi K^0_S$(LL) sample on (left) linear and (right) logarithmic scales.

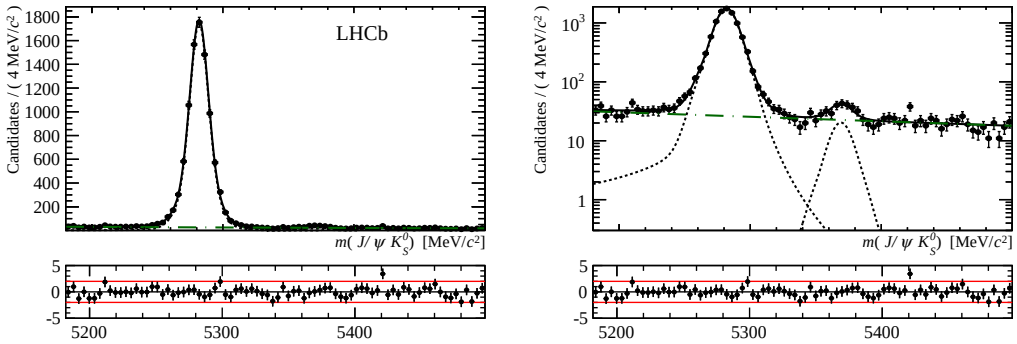

Figure 7.3: Fit results for the $B^0 \to J/\psi K^0_S$(DD) sample on (left) linear and (right) logarithmic scales.

background and also low signal yields. In order to have a stable fit, it is therefore necessary to reduce the number of free parameters. As shown in Table 7.4, we fix all parameters to the values found in data for $B^0 \to J/\psi K^0_S \pi^+\pi^-$. The background shape is also kept the same.

7.4 Correcting yields for non-resonant $\psi(2S)$ decays

As mentioned previously, we apply a veto of ± 15 MeV/c^2 about the $\psi(2S)$ mass to make a measurement of the $B^0 \to \psi(2S)K^0_S$ yield. Since the width of the $\psi(2S)$ is known to be narrow, we could have an additional contribution from non-resonant decays to the same $J/\psi\pi^+\pi^-$ final state that fall within this mass window. In an attempt to quantify this we perform a fit to the $J/\psi\pi\pi$ invariant mass having applied a signal ${}_s\mathcal{W}eight$ obtained from the fit to the $B^0 \to \psi(2S)K^0_S$ invariant mass (also referred to as *background subtraction*). This provides a clean B^0 yield from which we determine the number of $\psi(2S)$ decays and any non-resonant contribution using

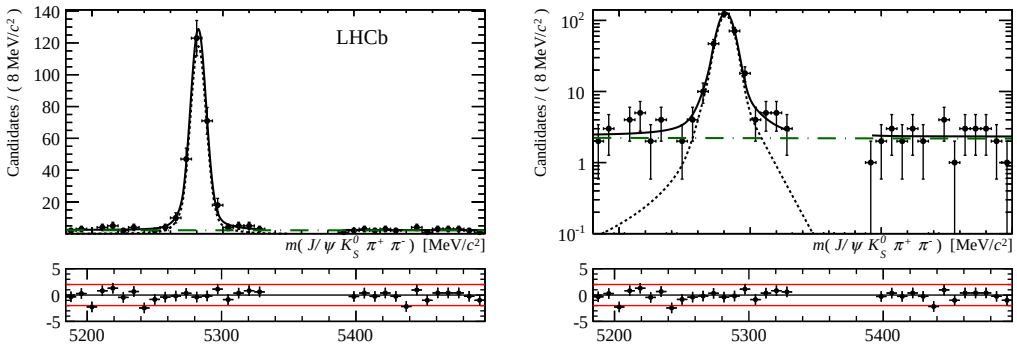

Figure 7.4: Fit results for the $B^0 \to J/\psi K^0_{\rm S}({\rm LL})\pi^+\pi^-$ sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind.

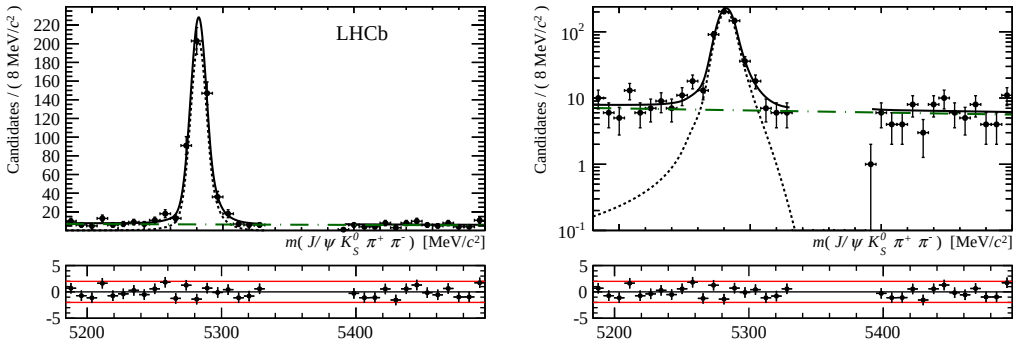

Figure 7.5: Fit results for the $B^0 \to J/\psi K^0_{\rm S}({\rm DD})\pi^+\pi^-$ sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind.

a simple fit model. The fit model is composed of a single Gaussian for the $\psi(2S)$ signal with floating mean ($\mu_{\psi(2S)}$) and width ($\sigma_{\psi(2S)}$), the non-resonant component is described using a first order Chebychev polynomial a_0. Table 7.5 contains the fit results for the $\psi(2S)$ fit and Fig. 7.8 show the $J/\psi\pi^+\pi^-$ invariant mass fit for the $K^0_{\rm S}$(DD) and $K^0_{\rm S}$(LL) categories. After performing the fit we find that the veto we apply (± 15 MeV/c^2) corresponds to $5\,\sigma$ coverage of the $\psi(2S)$ Gaussian peak. We also note that the non-resonant component of the fit has very large uncertainty. Since the ± 15 MeV/c^2 window cuts out only a tiny fraction of the total spectrum (figure 7.10 bottom left), the non-resonant component is consistent with being negligible, as expected, thus we do not correct the yield due to this effect.

7.5 Invariant mass combinations

There is a rich underlying resonance structure for these decay modes. Looking at the ${}_s\mathcal{W}eighted$ $B^0 \to J/\psi K^0_{\rm S}\pi^+\pi^-$ data (calculated from the results of the fit in

Table 7.4: Results of the fit to the $B^0 \to \psi(2S)K^0_S$ data samples. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. The measures of the goodness of fit do not make much sense with low statistics and arbitrary binning scheme.

Parameter	$B^0 \to \psi(2S)K^0_S(\text{DD})$	$B^0 \to \psi(2S)K^0_S(\text{LL})$
N_{comb}	11 ± 4	1 ± 1
N_{B^0}	51 ± 7	30 ± 5
χ^2	10.8	1.2
# floated params.	3	3
# bins	35	15
$P(\chi^2, \text{ndf})$	0.77	1.0

Table 7.5: Results of the $J/\psi\pi\pi$ invariant mass spectrum which has been weighted using the ${}_s\mathcal{W}eight$-ed fit from the $B^0 \to \psi(2S)K^0_S$ mode. The number of degrees of freedom (ndf) is estimated as the number of bins minus the number of floated parameters, both of which are given to allow alternative $P(\chi^2, \text{ndf})$ calculations. The measures of the goodness of fit do not make much sense with low statistics and arbitrary binning scheme. (Left) side show decays from $K^0_S(\text{DD})$ and (right) side for $K^0_S(\text{LL})$ type decays.

Parameter	$\psi(2S) \to J/\psi\pi^+\pi^-(K^0_S(\text{DD}))$	$\psi(2S) \to J/\psi\pi^+\pi^-(K^0_S(\text{LL}))$
$N_{\text{non-resonant}}$	10 ± 7	5 ± 4
a_0	$(9.1 \pm 4.1) \times 10^{-1}$	$(-2.6 \pm 5.5) \times 10^{-1}$
$N_{\psi(2S)}$	41 ± 9	25 ± 6
$\mu_{\psi(2S)}(\text{MeV}/c^2)$	3686.1 ± 0.6	3686.1 ± 0.9
$\sigma_{\psi(2S)}(\text{MeV}/c^2)$	3.0 ± 0.7	3.4 ± 1.1
χ^2	4.1	3.1
# floated params.	5	5
# bins	15	15
$P(\chi^2, \text{ndf})$	0.94	0.98

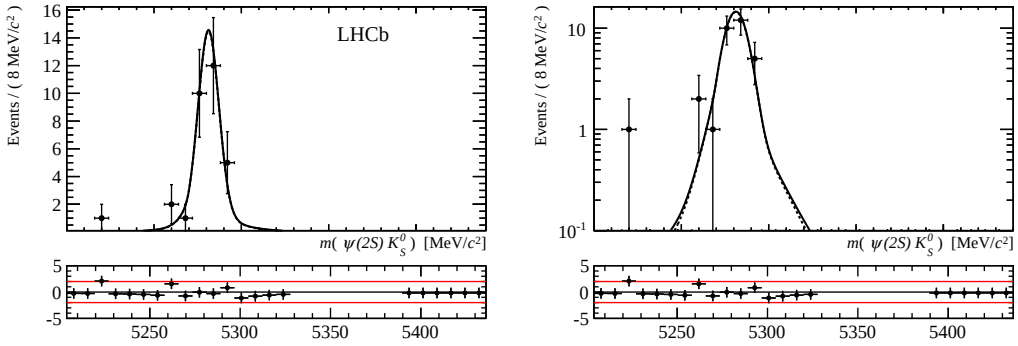

Figure 7.6: Fit results for the $B^0 \to \psi(2S)K^0_S$(LL) sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind.

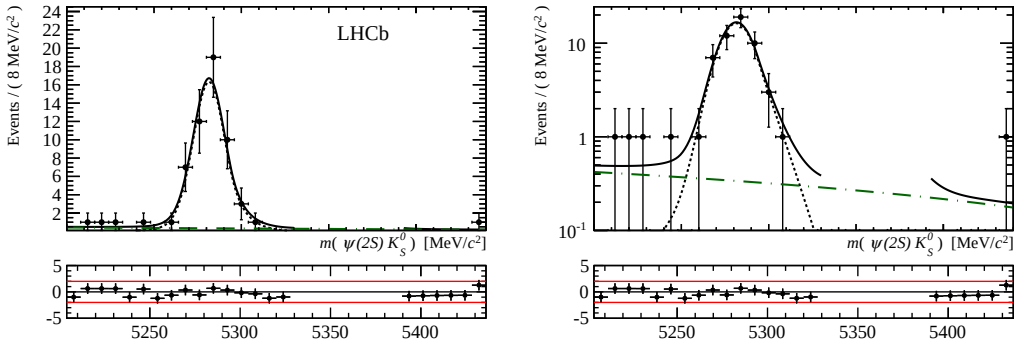

Figure 7.7: Fit results for the $B^0 \to \psi(2S)K^0_S$(DD) sample on (left) linear and (right) logarithmic scales. Note that the B^0_s region is kept blind.

Fig. 10.3) and comparing it with the equivalent simulated sample, we can check for any unexpected structures that may affect the analysis (*e.g.* by changing the average efficiency). Figure 7.9 and 7.10 show the invariant mass distributions for the various daughter particle combinations, we show only the K^0_S(DD) category but the K^0_S(LL) sample exhibits the same outlook. We can see that the simulation reproduces the main features present in the data. There are clear $K^*(892)$ resonant peaks in the data that are not included in the simulation. On the other hand, the $\rho(770)$ and $K_1(1270)$ peaks in the data are not as pronounced as in the MC. No exotic structures are seen in any combination containing the J/ψ. In particular, there is no significant excess corresponding to $B^0 \to X(3872)K^0_S$, $X(3872) \to J/\psi\pi^+\pi^-$. Overall, considering the relatively slow variation of the efficiency across the phase space (see Fig. B.2), the modelling in the MC seems sufficient at the current level of precision.

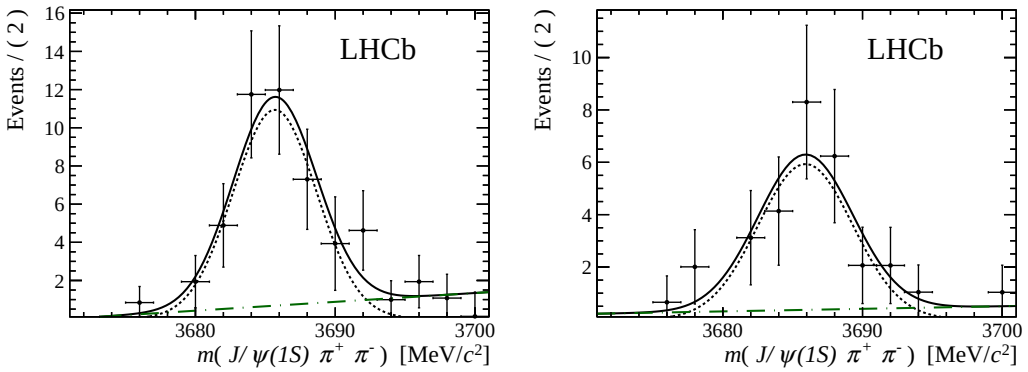

Figure 7.8: Fit results of the $\psi(2S) \to J/\psi\pi^+\pi^-$ invariant mass spectrum having applied ${}_s\mathcal{W}eights$ calculated from Figs. 7.7 and 7.6 for K^0_S(DD) and K^0_S(LL) types respectively. The sample on (left) K^0_S(DD) and (right) K^0_S(LL).

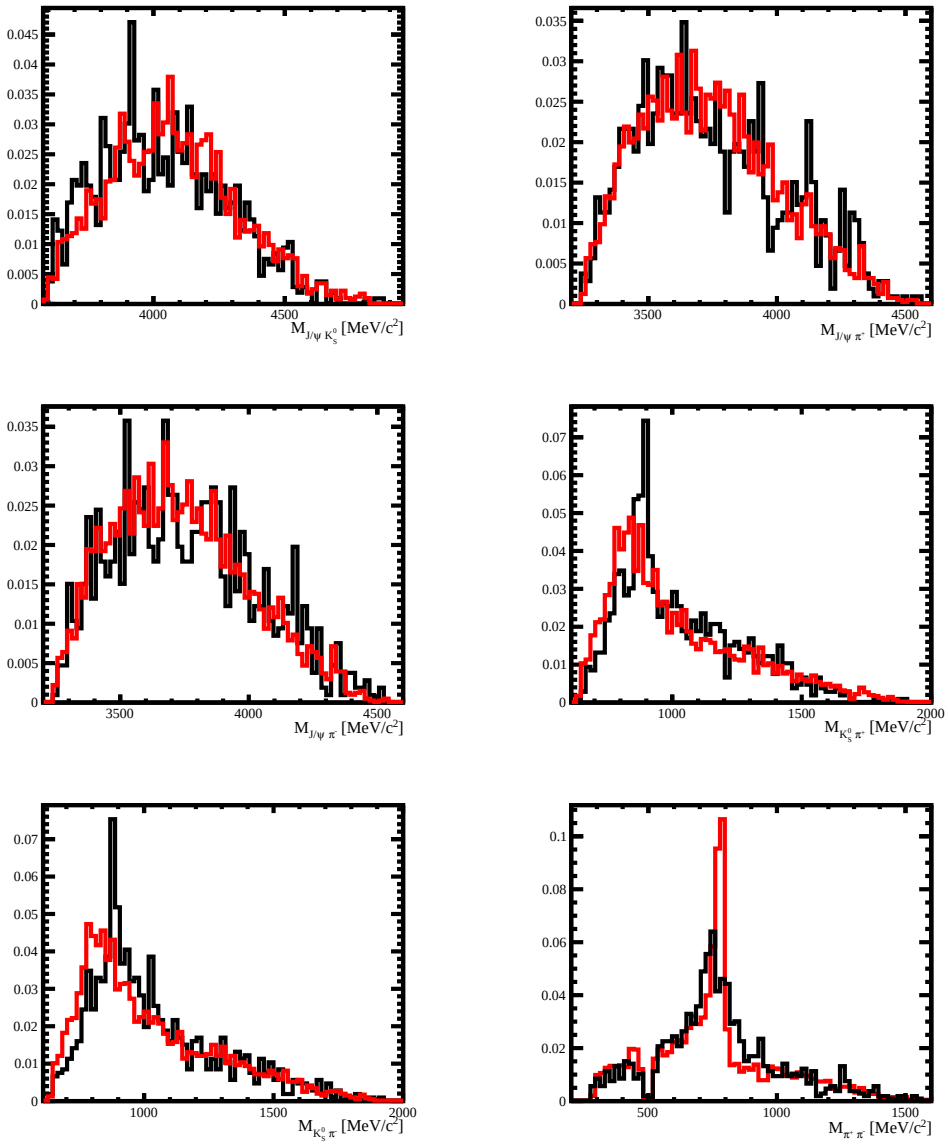

Figure 7.9: $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ two body combinations of the daughters, Red indicates simulation and black is for ${}_s\mathcal{W}eighted$ data. There is an excess of $\rho(770) \to \pi^+\pi^-$ in the simulation due to the forced branching fraction set in the decay file.

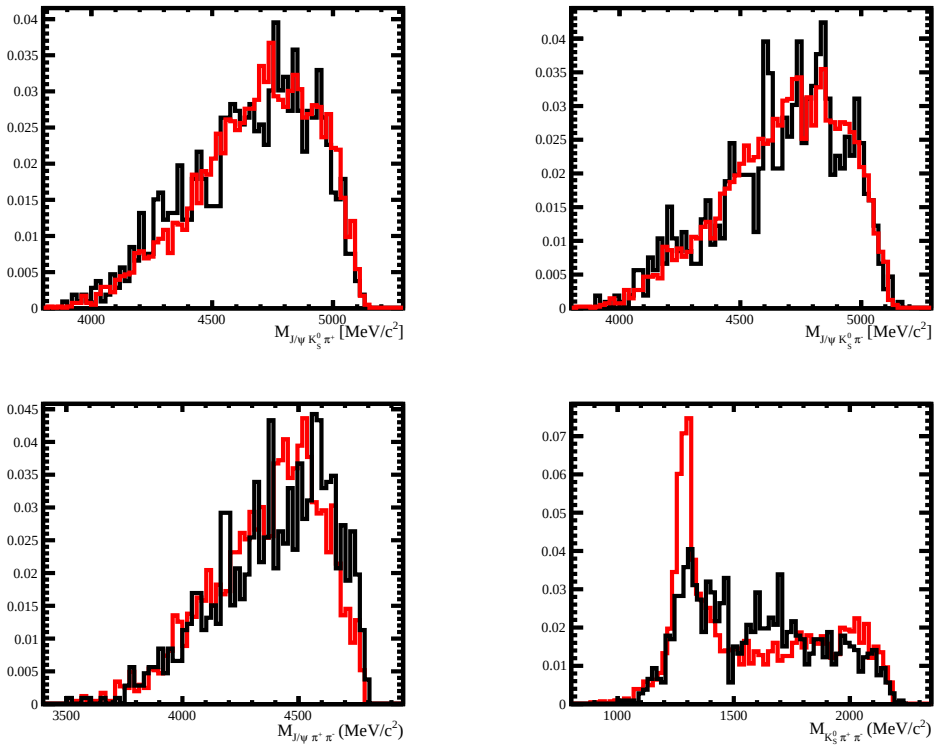

Figure 7.10: $B^0 \to J/\psi K^0_{\rm S}({\rm DD})\pi^+\pi^-$ three body combinations of the daughters, Red indicates simulation and black is for ${}_s\mathcal{W}eighted$ data. There is an excess of $K_1(1270) \to K^0_{\rm S}\pi^+\pi^-$ in the simulation due to the branching fraction set in the decay file. We see no obvious exotic charmonium states such as $X(3872) \to J/\psi\pi^+\pi^-$.

8

Efficiencies and systematic uncertainties

In this section we outline our handling of the various sources of inefficiency and systematic uncertainty that affect the analysis. This study includes a look at the dependence of the efficiency on the phase space of the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ decay. In addition, we have separately studied the efficiency of the $B^0 \to \psi(2S)K^0_S$ channel, using dedicated simulated samples, as presented in Appendix C. The results are consistent with the efficiency obtained in the $\psi(2S)$ region of $J/\psi\pi^+\pi^-$ invariant mass spectrum, and therefore provides confidence in our method of obtaining the branching fraction measurement for $\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)$.

8.1 Harmonic weighting of efficiencies

We evaluate various sources of inefficiency separately for candidates containing DD and LL K^0_S types, and also separately for MagUp and MagDown. The different evaluations are combined into a single efficiency value by taking an *average weighted* efficiency, otherwise known as the weighted harmonic mean. This average weighted efficiency is calculated as the inverse of the weighted average of the inverse efficiencies, shown below:

$$\bar{\varepsilon} = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} \frac{N_i}{\varepsilon_i}} \tag{8.1}$$

where N_i is the number of events for the i^{th} dataset and ϵ_i is the efficiency for that sample. We estimate the N_i from the numbers of signal events that pass the

full selection within a $50\,\mathrm{MeV}/c^2$ mass window, $|M_{B^0} - \langle m_{B^0}\rangle| < 50\,\mathrm{MeV}/c^2$, and perform an extended unbinned maximum likelihood fit in that region having fixed all the values to that which we find in Tables 7.2 and 7.3. Tables 8.1 and 8.2 show the

Table 8.1: Number of events used in weighting procedure for $B^0 \to J/\psi K^0_S$.

	LL	DD	Sum
MagUp	2195 ± 47	4410 ± 71	6605
MagDown	3191 ± 60	5971 ± 83	9162
Sum	5386	10381	15767

Table 8.2: Number of events used in weighting proceedure for $B^0 \to J/\psi K^0_S\pi^+\pi^-$.

	LL	DD	Sum
MagUp	127 ± 12	245 ± 17	372
MagDown	177 ± 14	330 ± 20	507
Sum	304	575	879

numbers extracted from data for the $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S\pi^+\pi^-$ modes respectively. Since there are relatively low statistics for each magnet polarity, we fix the fit parameters to those provided from the fit to the full K^0_S(LL) and K^0_S(DD) data samples. The relative ratios are in line with that from what we get from fitted yields to both magnet data samples.

8.1.1 Geometric efficiency

The LHCb experiment is a forward spectrometer in the pseudo-rapidity range $2 < \eta < 5$. The geometric acceptance is obtained using Monte Carlo studies, with a simple accept/reject counter at the event generation level. Please note that the $B^0 \to J/\psi K^0_S$ Monte Carlo has a different cut applied at the generator level to the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode as mentioned in Sec. 3.7. So we are required to study the effect of both the `DaughtersInLHCb` and `DaughtersInLHCbAndWithMinP` cuts for our nominal mode, as examined in detail in Appendix. B with peak-to-peak variation at the 10 % level for all invariant mass ranges.

Figure 8.1 presents the study based on a different simulated sample for the

$J/\psi\,\pi^+\pi^-$ invariant mass dependence.[1] In this case the efficiencies of both the `DaughtersInLHCb` requirement and as well as a generator level cut such that both pions have $p > 1600\,\text{MeV}/c$ and muons have $p > 3000\,\text{MeV}/c$ are examined. The

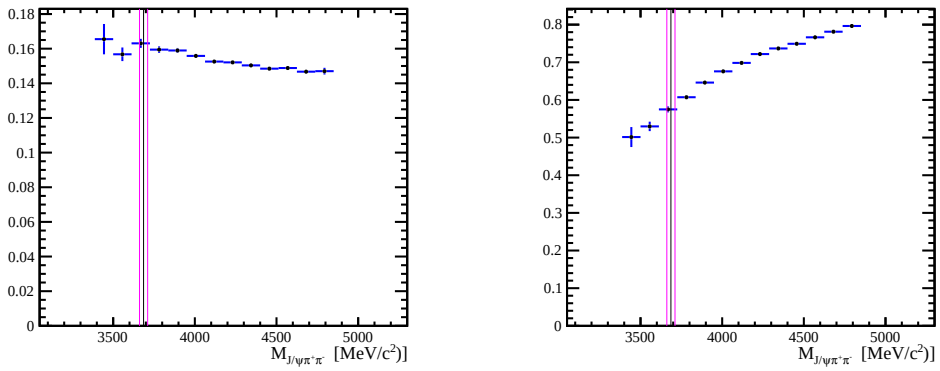

Figure 8.1: Generator level cut efficiencies of (left) `DaughtersInLHCb` requirement and (right) the additional requirement that pion and muon tracks have minimum momentum greater than $1600\,\text{MeV}/c$ and $3000\,\text{MeV}/c$ respectively (`DaughterInLHCbWithMinP` cut). The $\psi(2S)$ region is shown with vertical lines.

Table 8.3: Generator level cut efficiencies for $B^0 \to J/\psi\,K^0_S$ and $B^0 \to J/\psi\,K^0_S\pi^+\pi^-$ modes. The Monte Carlo productions for $B^0 \to J/\psi\,K^0_S$ and $B^0 \to J/\psi\,K^0_S\pi^+\pi^-$ were produced with the LHCb generator level cuts `DaughtersInLHCb` `DaughtersInLHCbWithMinP` respectively, they are outlined in more detail in Sec. 3.7.

Decay mode	Efficiency (%)
$\epsilon^{\text{Generator.}}_{B^0\to J/\psi\,K^0_S}$	17.97 ± 0.05
$\epsilon^{\text{Generator.}}_{B^0\to J/\psi\,K^0_S\pi^+\pi^-}$	11.06 ± 0.03

results, summarised in Table 8.3 show that the geometrical efficiency of the $B^0 \to J/\psi\,K^0_S$ normalisation channel is about a factor of 1.5 higher than that of the $B^0 \to J/\psi\,K^0_S\pi^+\pi^-$ mode. This is in line with expectation as there are more tracks to "lose" in the latter case.

8.1.2 Reconstruction, selection and trigger efficiency

The reconstruction and selection efficiency is calculated from the $B^0 \to J/\psi\,K^0_S$ and $B^0 \to J/\psi\,K^0_S\pi^+\pi^-$ simulation samples. The results are shown as functions

[1] It may seem trivial to take the $J/\psi\ \pi^+\ \pi^-$ invariant mass as an example but we originally intended to make a measurement of the $\mathcal{B}(B^0 \to \psi(2S)K^0)/\mathcal{B}(B^0 \to J/\psi\,K^0_S\pi^+\pi^-)$ relative branching fraction under the assumption that the efficiencies cancel. This however, turned out not to be the case as the efficiency at the $\psi(2S)$ mass is significantly smaller than the integrated efficiency.

of $J/\psi\pi^+\pi^-$ invariant mass in Figures 8.2–8.3, and are summarised in Tables 8.4 and 8.5 for $B^0 \to J/\psi K^0_S\pi^+\pi^-$ and $B^0 \to J/\psi K^0_S$ respectively, where the uncertainties given are statistical (from the size of the MC samples) only.

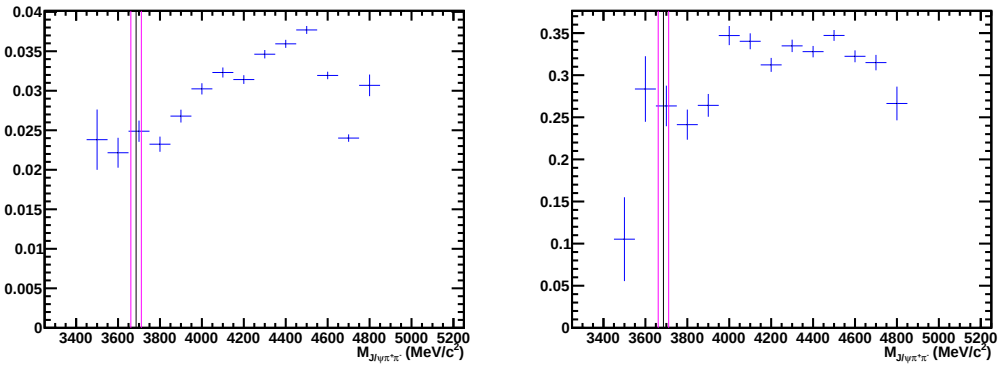

Figure 8.2: Reconstruction level (left) stripping line and (right) CROP selection efficiencies. These are for the DD K^0_S sample. The $\psi(2S)$ region is shown with vertical lines.

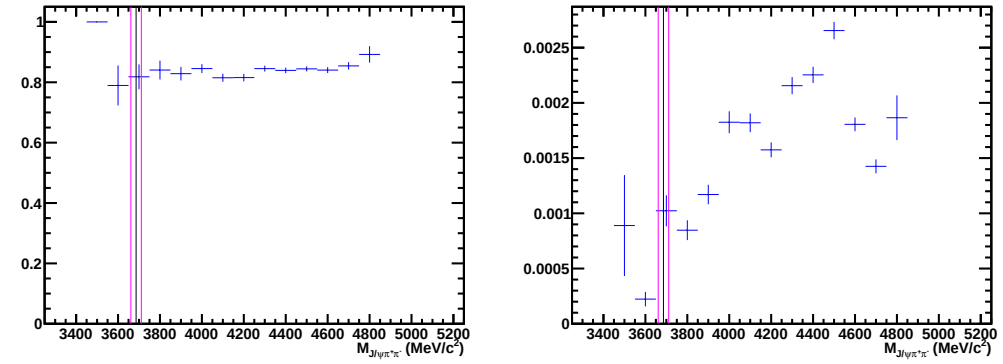

Figure 8.3: Reconstruction level (left) trigger and (right) total efficiency in bins of $m_{J/\psi\pi\pi}$. The total efficiency is the product of all the relative efficiency histograms that have been applied. These are for the DD K^0_S sample. The $\psi(2S)$ region is shown with vertical lines.

Tables 8.4 and 8.5 provide the integrated efficiencies, $\epsilon^{\text{Selection}}$, and their uncertainty, $\sigma_{\epsilon^{\text{Selection}}}$ (statistical only), where $\epsilon^{\text{Selection}} = \epsilon^{\text{reconstruction-stripping}} \times \epsilon^{\text{selection}} \times \epsilon^{\text{trigger}}$, where $\epsilon^{\text{trigger}} = \epsilon^{\text{L0}} \times \epsilon^{\text{HLT1}} \times \epsilon^{\text{HLT2}}$. The efficiency should not vary dramatically across the phase space yet from figure 8.3 we see that it does. This has been investigated and found be due to an ensemble of the cuts and vetoes we use, in particular the χ^2 discriminating variables and Λ and K^0_S vetoes.

Table 8.4: The integrated efficiency, $\epsilon^{\text{Selection}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S$.

	$B^0 \to J/\psi K^0_S$(LL) (%)		$B^0 \to J/\psi K^0_S$(DD) (%)	
	MagDown	MagUp	MagDown	MagUp
$\epsilon^{\text{Selection}}$	0.86	0.85	2.07	2.04
σ_ϵ(stat.)	$\pm$0.008	$\pm$0.08	$\pm$0.01	$\pm$0.01

Table 8.5: The integrated efficiency, $\epsilon^{\text{Selection}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$.

	$B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$ (%)		$B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
$\epsilon^{\text{Selection}}$	0.203	0.199	0.468	0.455
σ_ϵ (stat.)	$\pm$0.010	$\pm$0.010	$\pm$0.032	$\pm$0.031

8.1.3 Track reconstruction efficiency

There are two additional charged pion tracks in the final state $B^0 \to J/\psi K^0_S \pi^+\pi^-$ compared with the $B^0 \to J/\psi K^0_S$ normalisation channel. We can assign individual track efficiencies using track efficiency tables, obtained from $J/\psi \to \mu^+\mu^-$ decays [91]. This method not only provides the track efficiency data/simulation correction factor, but also its uncertainty, which enters our result as a source of systematic uncertainty. The method proceeds by re-weighting the simulated data according to the track multiplicity observed in data. This is important as the tracking efficiency shows variation with respect to track multiplicity, and also because the track multiplicity distributions observed in data and simulation are different. For a given track in the final state we compare the re-weighted $\eta - p$ values against an efficiency table, these are then averaged to give a correction factor for that particular track. The results for the π^+ and π^- hadrons are given in Table 8.6 together with the total correction, which is the product of the two. The values are all within statistical error of unity.

8.1.4 Particle identification efficiency

The particle identification efficiency observed in LHCb simulations is on average slightly larger than that which we find in data. This can be due to various inefficiencies such as variation over run number with HPDs that are disabled, detector

Table 8.6: Data/MC tracking corrections for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ as calculated using the tracking efficiency tables. Separated into LL/DD and magnet polarity categories. Note that the uncertainty comes from the size of the signal MC sample.

	$B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$		$B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$	
	MagDown	MagUp	MagDown	MagUp
$R(\pi^+)$	1.007 ± 0.021	1.009 ± 0.016	1.007 ± 0.021	1.007 ± 0.021
$R(\pi^-)$	1.007 ± 0.017	1.018 ± 0.017	1.008 ± 0.020	1.009 ± 0.022
$R^{\mathrm{Track}}_{\mathrm{DATA./MC.}}$	1.014 ± 0.027	1.028 ± 0.027	1.015 ± 0.029	1.016 ± 0.030

ageing, discrepancies between data and simulation track multiplicities as well as second order effects that are not fully simulated. Hence a data driven tool was devised in order to allow analysts access to a more accurate efficiency estimation. This tool is the `PIDCalib` package and incorporates a data driven method that utilises the decay $D^* \to D(K^-\pi^+)\pi_S$, taking advantage of the large number of charm mesons produced at LHCb leading to pure samples of pions and kaons with very high statistics. These samples can be used as calibration channels as the true DLL distributions are known.

To begin with, we make two-dimensional efficiency histograms in bins of $(p-\eta)$ from the 2011 calibration data sample. These give the efficiency of the PID cut (applied to either positive or negative tracks) in each bin. Here we are only considering the efficiency of the PID requirements applied to the bachelor pions in the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode (but the same method can be used to evaluate misidentification rates). The pion cuts, applied to each bachelor track, are

- (π^+, π^-) PID cut: $\mathrm{DLL}_{K\pi} < 0$ and $\mathrm{DLL}_{p\pi} < 10$.

Next we take our simulated data and apply the full selection excluding the PID requirement. As our decay has two bachelor charged pions we take each of them in turn and read the PID efficiency from the appropriate $(p-\eta)$ bin of the calibration histograms, provided that the pion came from our decay. This is done on an event-by-event basis, with the total PID efficiency per event obtained from the product of the two track efficiencies. The final efficiency used in the analysis is an average of the ensemble of the per-event efficiencies.

Table 8.7 shows the efficiencies determined with this method. Similar values are found for all magnet polarities and K^0_S types, as expected. The tiny uncertainty given in the table is from the statistics of the control sample only – systematic uncertainties associated to the PID efficiency will be discussed in Sec. 8.2.3.

Table 8.7: PID efficiency as calculated via the `PIDCalib` tool, for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ split by LL and DD and Up and Down magnet polarities, using the method described in the text. The uncertainty is from the statistics of the control sample only.

	$B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ (%)		$B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
$\epsilon^{\mathrm{PID.}}$	74.3 ± 0.01	72.2 ± 0.01	72.9 ± 0.01	71.0 ± 0.01

8.2 Systematic uncertainties

There are various systematic effects that must be studied for this analysis as they can cause uncertainty in the measurement of the branching ratio. Those that we have considered are listed below and discussed in turn. Most tables up to now have contained information regarding distinct magnet polarities and/or the K^0_S reconstructed types. Where appropriate, we will need to absorb these various components of the efficiencies and systematics alike into a final value. In order to do this we take an *average weighted* efficiency, otherwise known as the weighted harmonic mean, where the method and weightings are the same as explained in Sec. 8.1. The sources of systematic uncertainty are

- Track reconstruction;
- Trigger;
- Particle identification;
- Fit model;
- Monte Carlo sample size;
- Efficiency variation across the phase space.

8.2.1 Track reconstruction

The standard systematic error for 2011 data taking is 0.6% per track, see Ref. [92], therefore a total of $2 \times (0.006) = 1.2\%$ is assigned.

8.2.2 Trigger efficiency

The efficiency measurement comes straight from the information in Tables 8.4 and 8.5. We make a comparison of the relative trigger efficiencies at each stage

of the trigger between MC matched, reconstructed MC and data in order to validate the use of MC and more over, return an estimate of the systematic uncertainty.

LHCb has excellent trigger efficiency for muons, thus we use the dimuons from the J/ψ decay to trigger on, and expect a high trigger efficiency throughout the analysis. Since the final states of the two decay modes ($B^0 \to J/\psi K^0_S \pi^+\pi^-$ and $B^0 \to J/\psi K^0_S$) are not identical, the trigger efficiency ratio may not be unity but should give a consistent picture when comparing data and simulation due to differences in the dimuon momentum. The LHCb trigger has been characterised extensively in Ref. [93] and Ref. [94]. We look at the difference arising from truth matched MC, reconstructed MC and data to see if there are any discrepancies in the final ratios obtained after all decisions have been applied.

Table 8.8: Comparison of the relative trigger efficiencies for MC truth, reconstructed MC and reconstructed 2011 data. This table highlights the efficiency ratios for the K^0_S(LL) for both the $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. Errors are statistical only.

		L0		$L0 \times HLT1$		$L0 \times HLT1 \times HLT2$	
		MagDown	MagUp	MagDown	MagUp	MagDown	MagUp
MC	$B^0 \to J/\psi K^0_S$(LL)	0.8919 ± 0.0013	0.8939 ± 0.0013	0.9514 ± 0.0010	0.9501 ± 0.0010	0.9453 ± 0.0011	0.9442 ± 0.0011
matched	$B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$	0.9191 ± 0.0010	0.8949 ± 0.0111	0.9721 ± 0.0063	0.9457 ± 0.0087	0.9442 ± 0.0089	0.9492 ± 0.0056
$\epsilon^{\text{trg/tot}}$	ratio	0.9704 ± 0.0016	0.9989 ± 0.0112	0.9789 ± 0.0064	1.0047 ± 0.0088	1.0012 ± 0.0096	0.9947 ± 0.0057
MC	$B^0 \to J/\psi K^0_S$(LL)	0.8921 ± 0.0013	0.8941 ± 0.0013	0.9514 ± 0.0010	0.9500 ± 0.0010	0.9451 ± 0.0010	0.9437 ± 0.0010
	$B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$	0.9212 ± 0.0010	0.8509 ± 0.0110	0.9733 ± 0.0060	0.9453 ± 0.0080	0.9478 ± 0.0082	0.9492 ± 0.0083
$\epsilon^{\text{trg/tot}}$	ratio	0.9684 ± 0.0016	0.9990 ± 0.0111	0.9775 ± 0.0061	1.0050 ± 0.0081	0.9972 ± 0.0083	0.9942 ± 0.0084
2011	$B^0 \to J/\psi K^0_S$(LL)	0.9669 ± 0.0028	0.9698 ± 0.0032	0.9858 ± 0.0019	0.9876 ± 0.0021	0.9009 ± 0.0047	0.8936 ± 0.0058
data	$B^0 \to J/\psi K^0_S$(LL)$\pi^+\pi^-$	0.9748 ± 0.0054	0.9730 ± 0.0065	0.9791 ± 0.0050	0.9837 ± 0.0051	0.8869 ± 0.0112	0.8987 ± 0.0123
$\epsilon^{\text{trg/tot}}$	ratio	0.9919 ± 0.0061	0.9967 ± 0.0072	1.0068 ± 0.0053	1.0040 ± 0.0055	1.0158 ± 0.0121	0.9943 ± 0.0136

Table 8.9: Comparison of the relative trigger efficiencies for MC truth, reconstructed MC and reconstructed 2011 data. This table highlights the efficiency ratios for the K^0_S(DD) for both the $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes. Errors are statistical only.

		L0		$L0 \times HLT1$		$L0 \times HLT1 \times HLT2$	
		MagDown	MagUp	MagDown	MagUp	MagDown	MagUp
MC	$B^0 \to J/\psi K^0_S$(DD)	0.8853 ± 0.0008	0.8855 ± 0.0008	0.9356 ± 0.0007	0.9352 ± 0.0007	0.9319 ± 0.0007	0.9306 ± 0.007
matched	$B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$	0.9191 ± 0.0067	0.9052 ± 0.0070	0.9573 ± 0.0051	0.9570 ± 0.0051	0.9412 ± 0.0061	0.9492 ± 0.0056
$\epsilon^{\text{trg/tot}}$	ratio	0.9643 ± 0.0067	0.9782 ± 0.0070	0.9789 ± 0.0051	0.9772 ± 0.0051	0.9901 ± 0.0061	0.9804 ± 0.0056
MC	$B^0 \to J/\psi K^0_S$(DD)	0.8852 ± 0.0008	0.9351 ± 0.0007	0.9355 ± 0.0007	0.9351 ± 0.0007	0.9318 ± 0.0007	0.9305 ± 0.0007
	$B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$	0.9160 ± 0.0065	0.9538 ± 0.0050	0.9580 ± 0.0049	0.9538 ± 0.0050	0.9412 ± 0.0050	0.9510 ± 0.0053
$\epsilon^{\text{trg/tot}}$	ratio	0.9773 ± 0.0065	0.9804 ± 0.0050	0.9765 ± 0.0049	0.9784 ± 0.0050	0.9900 ± 0.0050	0.9784 ± 0.0053
2011	$B^0 \to J/\psi K^0_S$(DD)	0.9651 ± 0.0018	0.9660 ± 0.0021	0.9817 ± 0.0013	0.9826 ± 0.0015	0.8892 ± 0.0031	0.8987 ± 0.0036
data	$B^0 \to J/\psi K^0_S$(DD)$\pi^+\pi^-$	0.9657 ± 0.0038	0.9661 ± 0.0046	0.9767 ± 0.0032	0.9854 ± 0.0031	0.8813 ± 0.0069	0.8930 ± 0.008
$\epsilon^{\text{trg/tot}}$	ratio	0.9994 ± 0.0042	0.9999 ± 0.0051	1.0051 ± 0.0035	0.9972 ± 0.0034	1.0090 ± 0.0076	1.0064 ± 0.0088

Tables 8.8 and 8.9 provide comparisons of the relative efficiency at each stage of the trigger for K^0_S(LL) and K^0_S(DD) modes respectively. The tables comprise simulated truth information where the reconstructed B^0 candidate has been matched to a generated B^0 in the first row, second row is all reconstructed MC for a comparison, and finally in the third row we take the effect the trigger has on our data.

The final ratios are given in the bottom row of the table. All are in agreement with unity within statistical uncertainty, as expected since we only trigger on the J/ψ in the signal decays. To estimate the systematic uncertainty, we take the largest

Table 8.10: Difference in the final ratios of $\epsilon^{\text{trg/tot}}_{B^0\to J/\psi\, K^0_S}/\epsilon^{\text{trg/tot}}_{B^0\to J/\psi\, K^0_S\pi^+\pi^-}$ between MC truth and 2011 Data for the full trigger selection, L0 × HLT1 × HLT2. The K^0_S type and magnet polarity are secified for each mode.

	K^0_S(LL) (%)		K^0_S(DD) (%)	
	MagDown	MagUp	MagDown	MagUp
$\left\lvert\frac{\text{ratio}_{\text{data}}-\text{ratio}_{\text{MCMatched}}}{\text{ratio}_{\text{data}}}\right\rvert$	1.46	0.04	1.89	2.6

difference between the central values of the final ratio (L0 × HLT1 × HLT2) between data and the truth matched simulation (the last end column in tables 8.8 and 8.9). Table 8.10 summarises the largest differences where we find the K^0_S(DD) magnet up sub-sample, gives an uncertainty of $\Delta^{\text{trg.}} = 2.6\,\%$. For an additional cross check we also follow the same analysis procedure without the trigger decision requirement and see how this changes the resulting yields. We find a difference of 0.2 % which is deemed negligible.

8.2.3 Particle identification efficiency

The systematic induced by the calibration procedure is discussed. We have used simulation to determine the source of systematic error with the following formula.

$$\Delta_{\text{sys,calib}} = \left| \frac{\epsilon_{\text{MC,sig}} - \epsilon_{\text{MC,cal}}}{\epsilon_{\text{MC,sig}} + \epsilon_{\text{MC,cal}}} \times \epsilon_{\text{data,calib}} \right| , \tag{8.2}$$

where $\epsilon_{\text{data,calib}}$ and $\epsilon_{\text{MC,cal}}$ are efficiencies obtained using the same calibration procedure (as outlined in Sec. 8.1.4) but using the data and MC calibration samples, respectively, and $\epsilon_{\text{MC,sig}}$ is the true efficiency of the cuts on signal simulation. Using this method we obtain the following systematic errors where each pion has an applied cut of $\text{DLL}_{K\pi} < 0$ and $\text{DLL}_{p\pi} < 10$. For the total systematic error we take a weighted average of the values given in Table 8.11, this yields a value of $\bar{\Delta}^{\pi^+\pi^-}_{\text{sys,calib}} = 3.6\%$.

Table 8.11: Particle identification systematic error for combined π^+ and π^- tracks.

	$B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ (%)		$B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
$\Delta^{\pi^+\pi^-}_{\mathrm{sys,calib}}$	4.2	3.9	3.4	3.4

8.2.4 Fit model – $B^0 \to J/\psi K^0_S\pi^+\pi^-$

The shape of the PDFs used in the fits can be a source of systematic uncertainty as they may not represent the true distributions perfectly. The fact that we do not know the combinatoric shape or the true signal PDF shape means we must assign a systematic uncertainty for our chosen parameterisation. In order to validate this we replace the signal model (originally a double Crystal Ball function) with a double Gaussian function. Separately, we use a second order Chebychev polynomial instead of an exponential to describe the background shape. The changes in the yields provide estimates of the systematic uncertainties. Table 8.12 shows the results of

Table 8.12: Variation of the signal and background PDFs for $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ and $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ modes.

Parameter	Signal variation		Background variation	
	$B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$	$B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$	$B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$	$B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$
χ^2	9.6	37.5	8.2	35.7
# floated params.	8	8	9	9
# bins	20	35	20	35
$P(\chi^2, \mathrm{ndf})$	0.65	0.09	0.69	0.10
N_{B^0}	260 ± 17	478 ± 25	261 ± 18	585 ± 26
$\Delta^{\mathrm{Fit.}}_{\mathrm{var}}(\%)$	3.4	1.1	3.0	0.4

varying the PDF models for both background and signal. The different signal model causes the yields to change by about 1.5 % (recall a weighting procedure is used and we have roughly twice as many $K^0_S(\mathrm{DD})$ candidates), while the change associated with using a different background model is almost negligible at 0.6 %. We assign a systematic uncertainty that combines the background (0.6 %) and signal (1.5 %) systematic errors summed in quadrature such that a final error of 1.6 % is obtained. Another approach was taken whereby we use the alternative double-Gaussian model to generate toys in which we fit our nominal double-CB shape to and look at the change in yield. This method gives an uncertainty of 0.04 %, which we believe to be an underestimate and thus use the procedure above.

8.2.5 Simulated sample size

Since we rely on various efficiencies that we have derived from simulation we must assign a systematic due to a finite statistical sample. In order to do this we take the total binomial error from the simulated sample before and after the full selection is applied to both $B^0 \to J/\psi K^0_S$ and $B^0 \to J/\psi K^0_S \pi^+\pi^-$ modes, the total statistical error we assign due to the MC statistics is the sum in quadrature of these values, and gives a total of $\bar{\Delta}^{\rm MC} = 1.3\%$.

8.2.6 Efficiency variation across the phase space

Since the phase space distributions of the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ decays do not exactly match those of the MC (see Figure 7.10), and since the efficiency varies across the phase space (see Figures 8.1–8.3), it is possible that our MC-driven estimate of the efficiency is not correct. We estimate the size of the possible bias by reweighting the simulation[2] to obtain a distribution similar to that in data, and re-evaluate the average efficiency. As seen in table 8.13, we find the total sytematic due to the full reconstruction, stripping, trigger and selection changes by 2.0%, which we assign as a systematic uncertainty. A separate efficiency due to the fact we may not have

Table 8.13: Number of events used in weighting procedure for $B^0 \to J/\psi K^0_S \pi^+\pi^-$.

Resonance	Scale	Systematic (%)
$\rho^0(779) \to \pi^+\pi^-$	0.5	1.0
$K^0_1(1270) \to K^0_S\pi^+\pi^-$	0.5	1.2
$K^*(892) \to K^0_S\pi^+$	1.9	0.5
$K^*(892) \to K^0_S\pi^-$	1.9	0.5
Total		2.0

generated the events with the correct angular distribution and kinematics is also computed (see Appendix B) and we found a 0.3 % systematic effect. Therefore after taking all the values in quadrature the total systematic is rounded up to 2.0 %.

[2] In practice we simply bin the efficiency as a function of the desired invariant mass combination, then increase the number of simulated events in the resonant region and compute the weighted average over all the bins.

8.3 Total systematic uncertainty

For the total systematic error we sum all the individual contributions that we have considered above in quadrature after taking the harmonic mean value. These results are summarised in Table 8.14.

Table 8.14: Summary of the weighted-average values of all considered sources of systematic error for the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. The final value is the sum of the contributions in quadrature.

	$B^0 \to J/\psi K^0_S \pi^+\pi^-$ (%)
$\bar{\Delta}^{\text{Trig.}}$	2.6
$\bar{\Delta}^{\text{PID.}}$	3.6
$\bar{\Delta}^{\text{Track.}}$	1.2
$\bar{\Delta}^{\text{Fit.}}$	1.6
$\bar{\Delta}^{\text{MC}}$	1.3
$\bar{\Delta}^{\text{Eff. var.}}$	2.0
$\bar{\Delta}^{\text{Tot.}}$	5.4

9

Measurement of $\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)/\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})$

In this section we detail the calculation of the branching fraction, using

$$\begin{aligned}\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})} &= \frac{\epsilon^{\rm Generator.}_{B^0\to J/\psi\, K^0_{\rm S}}}{\epsilon^{\rm Gen.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-}} \times \frac{\epsilon^{\rm Selection.}_{B^0\to J/\psi\, K^0_{\rm S}}}{\epsilon^{\rm Selection.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-}} \\ &\times \frac{1}{\epsilon^{\rm PID}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-}} \times \frac{1}{\epsilon^{\rm Track}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-}} \times \frac{N_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-}}{N_{B^0\to J/\psi\, K^0_{\rm S}}}\,,\end{aligned} \tag{9.1}$$

where

$$\epsilon^{\rm Selection.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-} = \epsilon^{\rm reconstruction.-stripping.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-} \times \epsilon^{\rm selection.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-} \times \epsilon^{\rm trigger.}_{B^0\to J/\psi\, K^0_{\rm S}\pi^+\pi^-} \tag{9.2}$$

As mentioned, where necessary, we absorb the various magnet polarities and $K^0_{\rm S}$ reconstructed types into a global average efficiency. This procedure is described in Sec. 8.1. Table 9.1 summarises the individual average efficiencies that are obtained.

Using the yields from Sec. 7, the relative efficiencies given above, and the systematic uncertainties discussed in Sec. 8.2, then gives

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})} = 0.493 \pm 0.034\ (\text{stat.}) \pm 0.027\ (\text{syst.}) \tag{9.3}$$

where we have statistical and systematic contributions to the uncertainty of 7.0 % and 5.4 % respectively. We substitute in the value of the normalisation channel branching fraction as found in Ref. [15], $\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}) = 0.5\times(8.73\pm0.32)\times10^{-4}$,

Table 9.1: Relative trigger, selection and reconstruction efficiencies from simulation. Based on simulation optimised selection for $B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-$.

	$B^0 \to J/\psi\, K^0_{\rm S}$	$B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-$
$\bar{\epsilon}^{\rm Generator.}$	0.1797 ± 0.0006	0.1106 ± 0.0005
$\bar{\epsilon}^{\rm Selection.}$	$(1.391 \pm 0.006) \times 10^{-2}$	$(0.317 \pm 0.013) \times 10^{-2}$
$\bar{\epsilon}^{\rm PID.}$	-	0.725 ± 0.025
$\bar{R}^{\rm Track.}_{\rm Data./MC.}$	-	1.016 ± 0.027
$\bar{\epsilon}^{\rm Tot.}$	$(2.499 \pm 0.013) \times 10^{-3}$	$(0.259 \pm 0.011) \times 10^{-3}$
N_{B^0}	14739	752

and find

$$\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-) = (21.52 \pm 1.50\ (\text{stat.}) \pm 1.17\ (\text{syst.}) \pm 0.79\ (\text{PDG})) \times 10^{-5}\,, \tag{9.4}$$

This is less than 1.4σ from the previous measurement carried out by CDF, $\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-) = (50 \pm 20) \times 10^{-5}$ [16]. However, the CDF analysis included the $\psi(2S)K^0_{\rm S}$ intermediate contribution, which we do not, and consequently this is not an "apples-with-apples" comparison.

9.1 Cross check measurement of $\mathcal{B}(B^0 \to \psi(2S)K^0_{\rm S})/\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})$

We now give the measurement of $\mathcal{B}(B^0 \to \psi(2S)K^0_{\rm S})$ relative to $\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})$. The ratios of efficiencies are given in Table 9.2 (see also Appendix C). Using a modified version of Eq. (9.1), and the yields from Sec. 7, then gives

$$\frac{\mathcal{B}(B^0 \to \psi(2S)K^0_{\rm S}) \times \mathcal{B}(\psi(2S) \to J/\psi\,\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S})} = 0.183 \pm 0.027\ (\text{stat.}) \pm 0.015\ (\text{sys.}) \tag{9.5}$$

where we have statistical and systematic contributions to the uncertainty of 15.0 % and 8.4 % respectively. (The sources of systematic uncertainty to this ratio are discussed in Appendix C.1.) Substituting the world average branching fraction

Table 9.2: Relative efficiencies of $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$.

	$B^0 \to J/\psi\, K^0_S$	$B^0 \to \psi(2S)K^0_S$
$\bar{\epsilon}^{\text{Generator.}}$	0.1797 ± 0.0006	0.1595 ± 0.0006
$\bar{\epsilon}^{\text{Selection.}}$	$(1.5344 \pm 0.0054) \times 10^{-2}$	$(0.0746 \pm 0.0052) \times 10^{-2}$
$\bar{\epsilon}^{\text{PID.}}$	-	0.678 ± 0.009
$\bar{R}^{\text{Track.}}_{\text{Data./MC.}}$	-	1.013 ± 0.028
$\bar{\epsilon}^{\text{Tot.}}$	$(2.76 \pm 0.005) \times 10^{-3}$	$(8.300 \pm 0.633) \times 10^{-5}$
N_{B^0}	14739	81

value for $\mathcal{B}(B^0 \to J/\psi\, K^0_S)$,

$$\mathcal{B}(B^0 \to \psi(2S)K^0_S) \times \mathcal{B}(\psi(2S) \to J/\psi\,\pi^+\pi^-) = (8.0 \pm 1.2\ (\text{stat.}) \pm 0.7\ (\text{sys.}) \pm 0.3\ (\text{PDG})) \times 10^{-5}$$

$$\mathcal{B}(B^0 \to \psi(2S)K^0_S) = (23.5 \pm 3.5\ (\text{stat.}) \pm 2.1\ (\text{sys.}) \pm 2.9\ (\text{PDG})) \times 10^{-5}$$

As shown in Table 9.3, we expect a branching fraction of $\mathcal{B}(B^0 \to (\psi(2S) \to J/\psi\,\pi^+\pi^-)K^0_S) = (10.42 \pm 0.85) \times 10^{-5}$ with an associated uncertainty of 8 %. Our value is consistent with the PDG expectation within 1.3σ.

Table 9.3: Branching fraction measurements found in Ref. [15] and used in the calculation of the expected branching fraction for $B^0 \to \psi(2S)K^0_S$.

Branching fraction	Value
$\mathcal{B}(B^0 \to \psi(2S)K^0)$	$(6.2 \pm 0.5) \times 10^{-4}$
$\mathcal{B}(\psi(2S) \to J/\psi\,\pi^+\pi^-)$	$(33.6 \pm 0.4) \times 10^{-2}$
$\mathcal{B}(K^0 \to K^0_S)$	0.5
Expected $\mathcal{B}(B^0 \to (\psi(2S) \to J/\psi\,\pi^+\pi^-)K^0_S)$	$(10.42 \pm 0.85) \times 10^{-5}$

Part II

$B^0_{(s)} \to J/\psi\, K^0_{\rm S} h^+ h^{(\prime)-}$ analysis

10

Event Selection

In this chapter we discuss the selection of candidates after the stripping for the $B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ modes.

Additional requirements

As in Part 1 of this analysis we use the DTF constraints on the B candidate (removing candidates where the DTF fit does not converge). Table 10.1 shows the efficiency of the DTF convergence requirement in simulated events – at most, a few percent of signal is lost.

Table 10.1: Efficiency when applying the Decay Tree Fit (DTF) constraint such that the K^0_S and J/ψ are mass constrained as well as the PV. Taken from signal simulated events where the errors are statistical only.

Decay mode	DTF efficiency (%)	
	K^0_S(LL)	K^0_S(DD)
$B^0 \to J/\psi\, K^0_S \pi^+\pi^-$	98.47 ± 0.12	99.51 ± 0.04
$B^0 \to J/\psi\, K^0_S K^\pm\pi^\mp$	98.34 ± 0.13	99.79 ± 0.03
$B^0 \to J/\psi\, K^0_S K^+K^-$	97.96 ± 0.19	99.75 ± 0.04
$B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$	99.70 ± 0.07	99.91 ± 0.02
$B^0_s \to J/\psi\, K^0_S K^\pm\pi^\mp$	98.42 ± 0.20	99.61 ± 0.06
$B^0_s \to J/\psi\, K^0_S K^+K^-$	97.83 ± 0.19	99.96 ± 0.02

We also impose the requirement on the K^0_S(LL) vertex separation with respect to the B vertex, $\chi^2_{\rm VS} > 16$. We apply the following set of particle identification requirements on the bachelor tracks,

- pion PID cut: $\mathrm{DLL}_{K\pi} < 0$ and $\mathrm{DLL}_{p\pi} < 10$,
- kaon PID cut: $\mathrm{DLL}_{K\pi} > 2$ and $(\mathrm{DLL}_{p\pi} - \mathrm{DLL}_{K\pi}) < 10$.

Note that in the list above, the pion requirement is the same as in Part I. The cuts used generate distinct datasets with no overlap within a given population (cross-feed is addressed as a separate issue later). Again, due to the performance of the RICH being optimal in the momentum range $p < 100\,\mathrm{GeV}/c$, we impose this upper bound for the momentum of any bachelor track.

10.1 Vetoes

In Part I of the analysis we specifically applied vetoes to the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ mass spectrum, as discussed in Sec. 6.1. We now consider a similar set of vetoes, in particular to remove possible baryonic contributions. Figure 10.1 shows the invariant mass range from data when substituting the π daughters from the K^0_S candidate with the proton mass hypothesis.

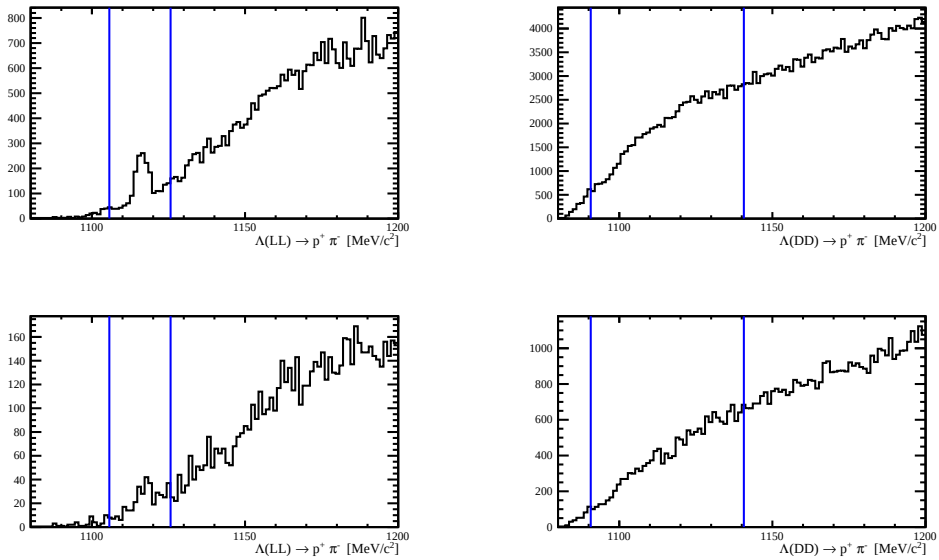

Figure 10.1: Candidate (top) $B^0_{(s)} \to J/\psi\, K^0_S K^\pm \pi^\mp$ and (bottom) $B^0_{(s)} \to J/\psi\, K^0_S K^+ K^-$ decays in the whole invariant mass range reconstructed with either daughter of the K^0_S candidate reconstructed under the proton mass hypothesis for 2011 stripped data. The blue lines show $|m(p\pi^-) - m_\Lambda| > 10\,\mathrm{MeV}/c^2$ (the cut is at $\pm 10\,\mathrm{MeV}/c^2$ for K^0_S(LL) and $\pm 25\,\mathrm{MeV}/c^2$ for K^0_S(DD).

Since we have already, in Part I, shown the distributions for the $B^0_{(s)} \to J/\psi\, K^0_S \pi^+\pi^-$ mode we now show the distribution for the other two modes, $B^0_{(s)} \to$

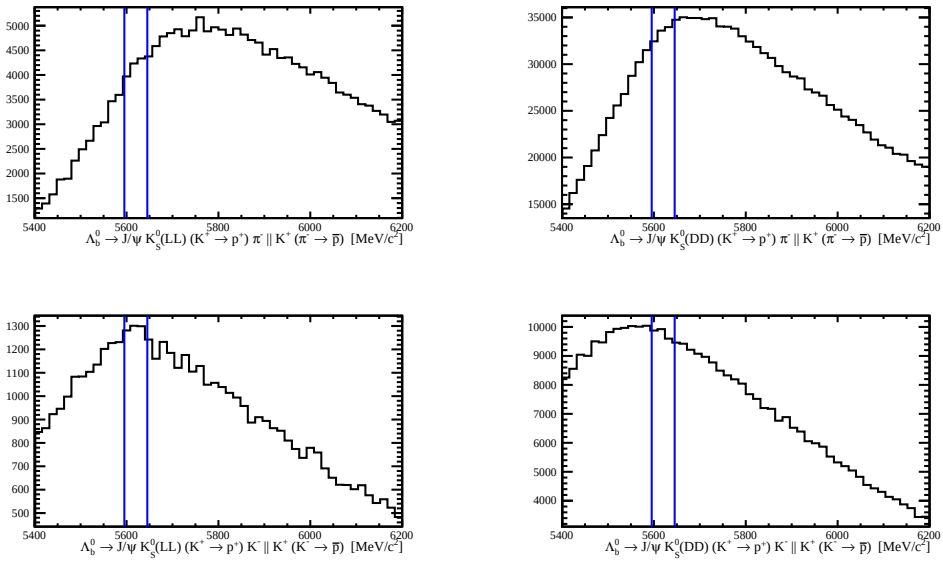

Figure 10.2: Candidate (top) $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and (bottom) $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ decays in the whole invariant mass range reconstructed with either bachelor track assigned the proton mass hypothesis for 2011 stripped data. The blue lines show the vetoed region. (Left) LL sample, (right) DD sample.

$J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$. Figures 10.1 and 10.2 highlight these distributions for the Λ and Λ_b^0 vetoes that we apply respectively. There are clearly peaking structure in the Λ regions but it is a little more difficult to ascertain visually the impact of the Λ_b^0 veto from the rest of the plots in Fig. 10.2. However we keep the cut for consistency. Table 10.2 gives the vetoes in detail, note the asterisk marked vetoes are remnants from Part I of the analysis and therefore are only applied to the $B^0_{(s)} \to J/\psi K^0_S \pi^+ \pi^-$ mode.

10.2 Initial cuts

A series of loose cuts is applied to clean up the sample after stripping so we can utilise the *sPlot* technique [95]. The *sWeights* were obtained from a fit to the data after nine common pre-selection cuts (plus two additional cuts on the K^0_S(LL) candidates) are applied to clean up the signal peak and obtain a stable fit. The cuts are listed in Table 10.3.

As seen in Fig. 10.3, we use a Gaussian signal peak and a first order polynomial to model the background shape in each K^0_S category. The *sWeights* are used to inform CROP how signal-like an event is, given that the variables used in the

Table 10.2: Veto requirements. The masses values of all particles are taken from the PDG [15]. All cuts are the same for LL and DD samples except where indicated otherwise. Those cuts marked with a (*) are only applied to the $B^0_{(s)} \to J/\psi\, K^0_S \pi^+\pi^-$ mode.

Cut	Value
Λ(LL) veto	$\lvert m(p\pi^-) - m_\Lambda\rvert > 10\,\text{MeV}/c^2$
Λ(DD) veto	$\lvert m(p\pi^-) - m_\Lambda\rvert > 25\,\text{MeV}/c^2$
Λ^0_b veto	$\left\lvert m(J/\psi\, K^0_S ph) - m_{\Lambda^0_b}\right\rvert > 25\,\text{MeV}/c^2$
*$\psi(2S)$ veto	$\left\lvert m(J/\psi\,\pi^+\pi^-) - m_{\psi(2S)}\right\rvert > 15\,\text{MeV}/c^2$
*K^0_S veto	$\left\lvert m(\pi^+\pi^-) - m_{K^0_S}\right\rvert > 25\,\text{MeV}/c^2$

selection are independent of the B^0 mass. The correlation of the B^0 mass with the other variables is plotted in Figs. 10.4 and 10.5 for K^0_S(LL) and K^0_S(DD) categories respectively. There is no strong dependence of any variable with the mass, implying that the *sWeight* procedure is valid. There is, however, some correlation between the training variables for both signal and background, which is ignored in this analysis.[1]

10.3 Optimisation of $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$

Using the K^0_S(DD) and K^0_S(LL) *sWeights*, we can begin training a full selection for the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$. The variables chosen are outlined in Table 10.4. The

[1] This provides some argument that an MVA-based analysis could provide better discrimination.

Table 10.3: Set of loose cuts applied to the data passing the stripping selection. These cuts clean up the signal enough so that we can perform a rough fit to the data. An asterisk (*) indicates a cut only applied to K^0_S(LL).

Particle	Variable	Cut
B^0	Pointing angle of B, $\cos(\theta_{\text{PV}})$	B_DIRA_OWNPV > 0.99993
	Transverse Momentum p_T	B_PT > 50 MeV
	Vertex fit, χ^2_{vertex}/ndf	B_ENDVERTEX_CHI2NDOF < 4
	IP significance wrt related PV, χ^2_{IP}	B_IPCHI2_OWNPV < 64
K^0_S	Vertex fit, χ^2_{vertex}/ndf	KS_ENDVERTEX_CHI2NDOF < 16
	IP significance with respect to PV, χ^2_{IP}	KS_IPCHI2_OWNPV > 3
*	IP significance with respect to PV, χ^2_{IP}	KS_IPCHI2_OWNPV > 5
*	Significance of separation wrt related PV, χ^2_{VD}	KS_VDCHI2_OWNPV > 10
*	Significance of separation wrt B^0 vertex, χ^2_{VD}	KS_VTX_SEPCHI2 > 9
J/ψ	Vertex fit, χ^2_{vertex}/ndf	Jpsi_ENDVERTEX_CHI2NDOF < 16
	Significance of separation wrt related PV, χ^2_{VD}	Jpsi_VDCHI2_OWNPV > 25
π^+,π^-	Minimum IP significance wrt related PV, χ^2_{IP}	min(h1_MINIPCHI2, h2_MINIPCHI2) > 5

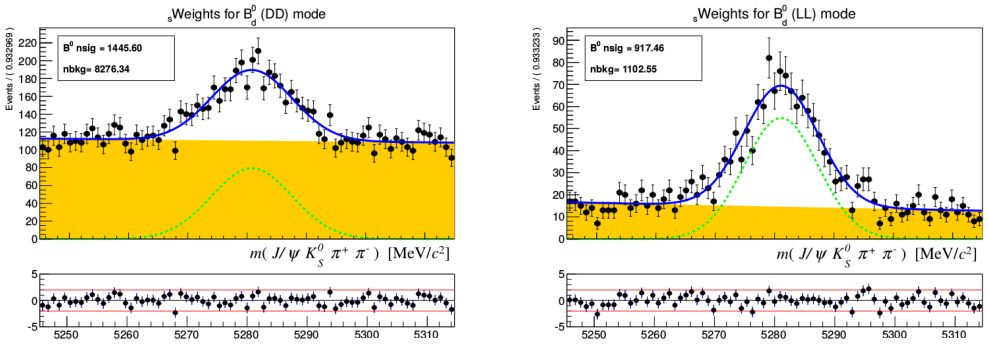

Figure 10.3: Gaussian with linear background mass fit after applying loose cuts outlined in Table 10.3. (Left) $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ data, (right) $B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ data.

CROP tool-kit is then used to maximise the signal yield whilst reducing the overall background using the figure of merit $N_S/\sqrt{N_S+N_B}$, where N_S and N_B are the numbers of signal and background decays in the mass range shown in Fig. 10.3.

Table 10.4: List of cut variables used by CROP. Candidates' vertex reconstruction, B pointing direction and lifetime-related variables are considered.

Particle	Variable
B^0	Pointing angle of B, $\cos(\theta_{\mathrm{PV}})$, B_DIRA_OWNPV
	Transverse Momentum p_{T}, B_PT
	Vertex fit, χ^2_{vertex}/ndf, B_ENDVERTEX_CHI2
	IP significance with respect to PV, χ^2_{IP}, B_IPCHI2_OWNPV
K^0_{S}	Vertex fit, χ^2_{vertex}/ndf, KS_ENDVERTEX_CHI2
	IP significance with respect to PV, KS_IPCHI2_OWNPV
K^0_{S} (LL)	Vertex separation χ^2, KS_VDCHI2_OWNPV
J/ψ	Vertex fit, χ^2_{vertex}/ndf, Jpsi_ENDVERTEX_CHI2
	IP significance with respect to PV, Jpsi_IPCHI2_OWNPV
	Vertex separation χ^2, Jpsi_VDCHI2_OWNPV
π	Minimum IP significance, min{h1_MINIPCHI2, h2_MINIPCHI2}

10.4 Tuning the selection for other modes

The relative amounts of signal and background decays will differ between the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ channel and other modes of interest. In principle, therefore, CROP should be tuned separately for each mode. However, the benefit from a fine-tuned

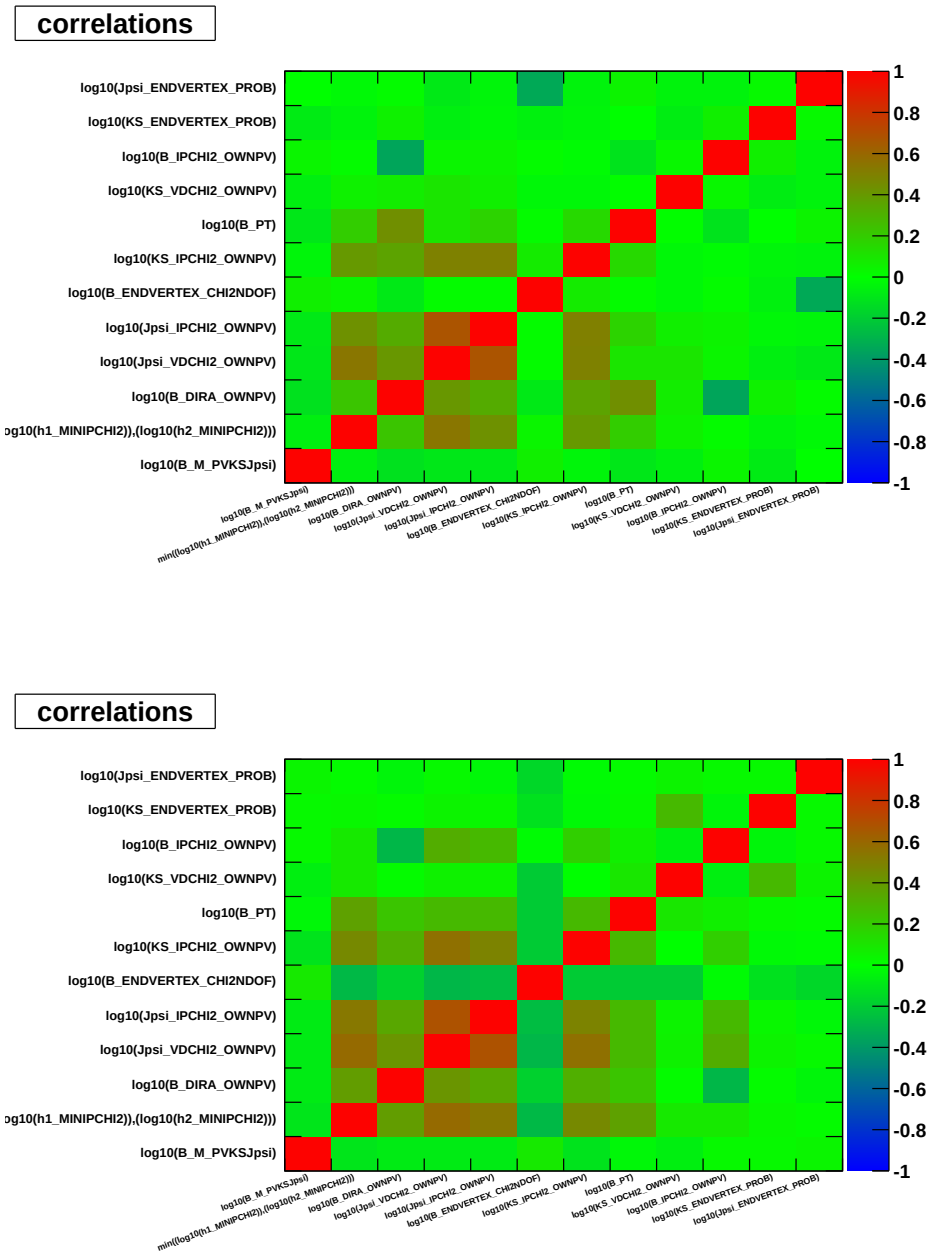

Figure 10.4: Correlations between variables using *sWeights* technique applied to the $B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-$ training data; (top) signal weighted, (bottom) background weighted.

optimisation is outweighed by the advantages of keeping the selection simple. In particular, we prefer to have a single selection for each final state (rather than

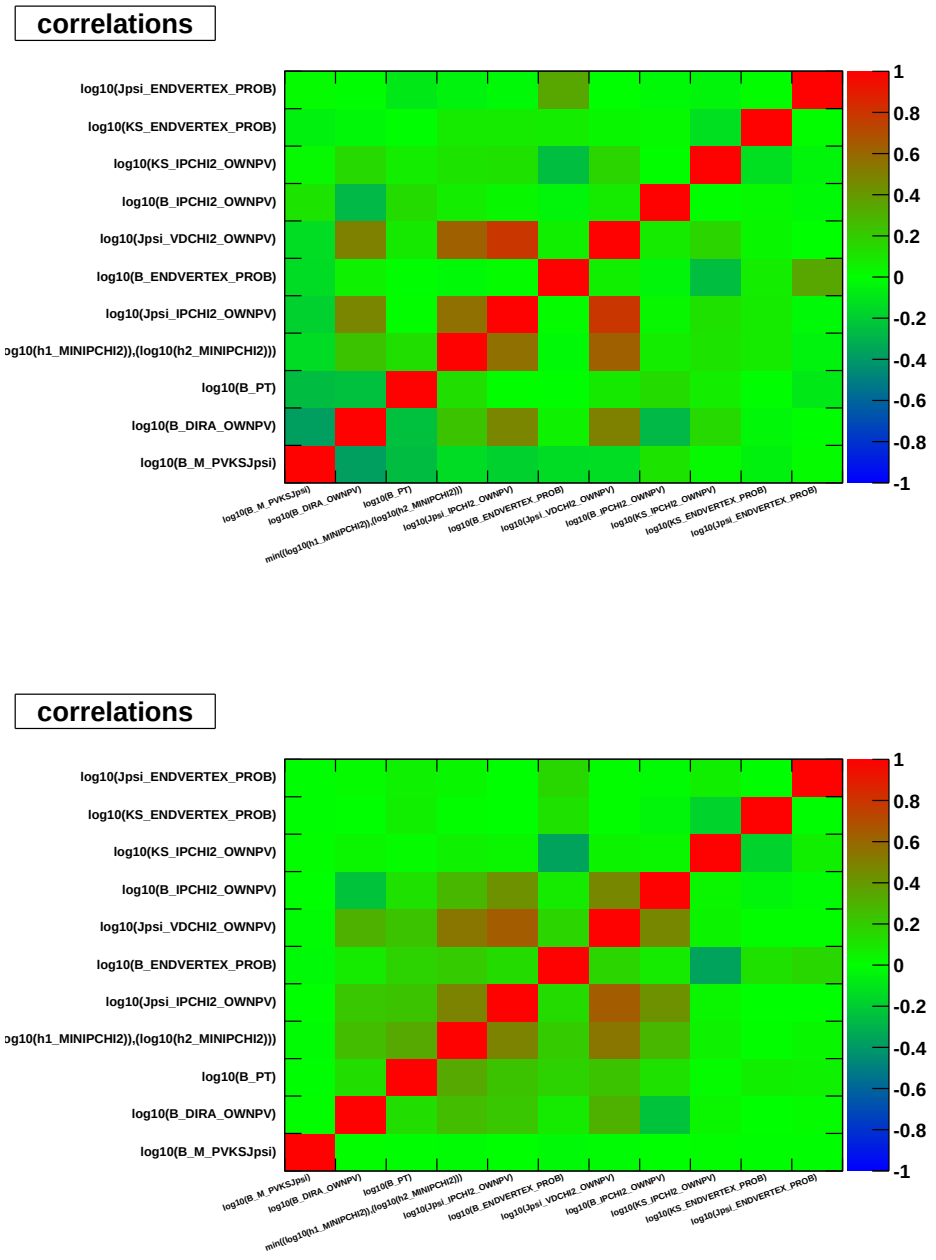

Figure 10.5: Correlations between variables using *sWeights* technique applied to the $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ training data; (top) signal weighted, (bottom) background weighted.

separate optimisations for both B^0 and B^0_s decays).

We make a rough tuning of the selection requirements for each final state with

the assumption that the total number of signal and background events will be proportional under some global scaling factors, $f_{\rm sig}$ and $f_{\rm bkg}$ for signal and background respectively. As an example, for the decay mode $B \to J/\psi\, K^0_{\rm S} K^+ K^-$,

$$N_S(B \to J/\psi\, K^0_{\rm S} K^+ K^-) = N_S(B \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-) \cdot f_{\rm sig}\,, \tag{10.1}$$

$$N_B(B \to J/\psi\, K^0_{\rm S} K^+ K^-) = N_B(B \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-) \cdot f_{\rm bkg}\,, \tag{10.2}$$

where the number of signal (N_S) and background (N_B) events for the $B \to J/\psi\, K^0_{\rm S} K^+ K^-$ mode are related to the numbers of signal and background events in the $B \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-$ channel by the signal and background scaling factors $f_{\rm sig}$ and $f_{\rm bkg}$ respectively. The procedure to obtain the signal and background scaling factors are discussed in Sec. 10.4.1 and Sec. 10.4.2, respectively. Once $f_{\rm sig}$ and $f_{\rm bkg}$ have been obtained, they are fed into CROP to as global re-weightings of the control channel ${}_s\mathcal{W}eights$. We emphasise that the intention is not to obtain a precise optimisation, but simply to get a close-to-optimal approximation of the selection for these modes without spending too long on optimisation. This will ultimately leave us with 6 cut based selection (for 3 final states and 2 $K^0_{\rm S}$ categories) as outlined in table 10.7.

10.4.1 Calculation of signal scale factor $f_{\rm sig}$

We estimate the relative yield $f^x_{\rm sig}$ for each B^0_q signal mode from

$$f^x_{\rm sig}(h'h/\pi\pi) = \frac{\epsilon_{\rm PID}(h'h)}{\epsilon_{\rm PID}(\pi\pi)} \times \frac{\mathcal{B}(B^0_x \to J/\psi\, K^0_{\rm S} h^\pm h^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-)} \times \frac{f_x}{f_d} \tag{10.3}$$

where $\frac{\epsilon_{\rm PID}(h'h)}{\epsilon_{\rm PID}(\pi\pi)}$ is the ratio of the particle identification efficiencies (we assume that all other efficiencies cancel to a fair approximation), and $\frac{f_x}{f_d}$ is the appropriate b-quark fragmentation ratio (we take $f_s/f_d = 0.256 \pm 0.020$ [80]). The ratio of branching fractions $\frac{\mathcal{B}(B^0_x \to J/\psi\, K^0_{\rm S} h^\pm h^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S} \pi^+ \pi^-)}$ is taken to be unity for simplicity.[2] The final equation then simplifies to

$$f^x_{\rm sig}(h'h/\pi\pi) = \frac{\epsilon_{\rm PID}(h'h)}{\epsilon_{\rm PID}(\pi\pi)} \times \frac{f_x}{f_d}\,. \tag{10.4}$$

We use the particle identification cut efficiencies given in Sec. 12.1.3. To try to strike a happy medium between optimising for a Cabibbo favoured B^0 mode or a suppressed B^0_s mode (or vice versa), we take an average value, $f_{\rm sig} = (f^d_{\rm sig} + f^s_{\rm sig})/2$. We obtain the six scaling factors given in Table 10.5.

[2] This is a reasonable, though crude, assumption for the Cabibbo favoured modes.

Table 10.5: Calculated values for the signal scale factors, based on the fraction of the signal yields.

Mode	Scale factor
$f_{\rm sig}(\pi\pi)$ (LL)	0.63
$f_{\rm sig}(\pi\pi)$ (DD)	0.63
$f_{\rm sig}(K\pi)$ (LL)	0.69
$f_{\rm sig}(K\pi)$ (DD)	0.69
$f_{\rm sig}(KK)$ (LL)	0.77
$f_{\rm sig}(KK)$ (DD)	0.77

10.4.2 Calculation of background scale factor $f_{\rm bkg}$

The background scale factor is found by taking the upper mass side-band where the distribution is most flat. An upper mass range of $5425 \rightarrow 5550\,\text{MeV}/c^2$ is used. Four ratios are found (two for each $K^0_{\rm S}$ category) by comparing the number of events passing the selection cuts described in Sec. 10.3, where the number of events surviving in each $B \rightarrow J/\psi\, K^0_{\rm S} K^+\pi^-$ and $B \rightarrow J/\psi\, K^0_{\rm S} K^+K^-$ mode is taken relative to the $B \rightarrow J/\psi\, K^0_{\rm S}\pi^+\pi^-$ mode in the denominator,

Table 10.6 shows the results for each mode (the $\pi\pi$ mode background scale factor is unity by definition).

Table 10.6: Calculated values for the background scale factor, based on the fraction of the background yields in the upper side band $5425 < M_{J/\psi\, K^0_{\rm S} h'h} < 5550\,\text{MeV}/c^2$.

Mode	Scale factor
$f_{\rm bkg}(K\pi)$ (LL)	4.62
$f_{\rm bkg}(K\pi)$ (DD)	2.36
$f_{\rm bkg}(KK)$ (LL)	1.33
$f_{\rm bkg}(KK)$ (DD)	0.68

10.5 Final selection

The final kinematic selection for the $B^0_{(s)} \rightarrow J/\psi\, K^0_{\rm S} h^+ h^{(\prime)-}$ mode for the DD and the LL scenarios are presented. The scale factors are inserted into CROP and the resulting figure of merit is maximised, $f_{\rm sig}.N_S/\sqrt{f_{\rm sig}.N_S + f_{\rm bkg}.N_B}$. A full list of all

relevant cuts can be found in Table 10.7 where the significance of each cut value is given in brackets next to the actual cut value.

In any one collision at LHCb there is a finite probability of observing more than one candidate in an event after the full selection. The frequency of this occurrance for the different modes is recorded in Table 10.8 with a maximum rate of 4.15 %. A single candidate is selected using a predetermined algorithm.

Table 10.7: Cuts for the data optimised selection both K^0_S(LL) and K^0_S(DD) types. The list is in order of significance for the LL mode for the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ selection, for the DD mode the numbering is labelled in open brackets next to the cut value and the other modes are also indexed appropriately when required.

Cut	$B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$		$B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$		$B^0_{(s)} \to J/\psi K^0_S K^+ K^-$	
	K^0_S(LL)	K^0_S(DD)	K^0_S(LL)	K^0_S(DD)	K^0_S(LL)	K^0_S(DD)
K^0_S significance of separation wrt related PV, $\chi^2_{\rm VD}$	> 776 (1)	–	> 1603 (1)	–	> 776 (1)	–
J/ψ significance of separation wrt related PV, $\chi^2_{\rm VD}$	> 299 (2)	> 89 (8)	> 299 (3)	282 (9)	> 299 (3)	159 (8)
B^0 pointing angle, $\cos(\theta)$	> 0.999974 (3)	> 0.999983 (4)	> 0.999961 (7)	> 0.999983 (5)	> 0.999961 (6)	0.999978 (4)
B^0 transverse momentum, $p_{\rm T}$ (MeV/c)	> 2126 (4)	> 1585 (5)	> 2126 (6)	> 1585 (7)	> 2126 (5)	> 1585 (7)
B^0 IP significance with respect to PV, $\chi^2_{\rm IP}$	< 13 (5)	< 11 (1)	< 13 (5)	< 11 (1)	< 13 (4)	< 11 (1)
K^0_S IP significance with respect to PV, $\chi^2_{\rm IP}$	> 3.7 (6)	< 513 (9)	> 3.7 (8)	< 513 (6)	> 3.7 (7)	< 1259 (9)
π min IP significance wrt related PV, $\min(\pi^+_{\chi^2_{\rm IP}}, \pi^-_{\chi^2_{\rm IP}})$	> 5 (7)	> 5.7 (7)	> 2.2 (9)	> 9.6 (3)	> 2.2 (9)	> 9.6 (2)
J/ψ IP significance with respect to PV, $\chi^2_{\rm IP}$	> 0.01 (8)	> 11 (2)	> 7.8 (4)	> 11 (4)	> 11 (2)	> 3.9 (6)
B^0 vertex probability, $P(\chi^2, {\rm ndf})(\%)$	> 0.158 (9)	> 6.310 (3)	> 1.585 (2)	> 6.310 (2)	0.251 (8)	> 6.310 (3)
K^0_S vertex probability, $P(\chi^2, {\rm ndf})(\%)$	> 0.009 (10)	> 2.163 (6)	> 0.009 (10)	> 2.163 (8)	> 0.009 (10)	> 1585 (5)
J/ψ vertex probability, $P(\chi^2, {\rm ndf})(\%)$	No cut (11)	No cut (10)	No cut (11)	No cut (10)	No cut (11)	No cut (10)

Table 10.8: Number of multiple candidates in each of the data samples.

	$J/\psi K^0_S \pi\pi$	$J/\psi K^0_S K\pi$	$J/\psi K^0_S KK$
Number of Candidates (LL)	297	517	122
Number of multiple cands (LL)	3 (1.01 %)	14 (2.71 %)	2 (1.64 %)
Number of Candidates (DD)	689	1435	434
Number of multiple cands (DD)	6 (0.87 %)	22 (1.53 %)	18 (4.15 %)

11

Fit model

We now outline the fit model and its implementation. The fit contains different categories of events, both K^0_S(DD) and K^0_S(LL), as well as several possible final states in which potential cross-feed can leak through. Therefore, we decided that a simultaneous unbinned extended maximum likelihood fit to the B-candidate invariant mass distributions of all decay channels to extract the relevant yields and to constrain the levels of cross-feed would best suit our needs. The simultaneous fit is to 12 modes: $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ (6 decays) in each of the DD and LL K^0_S reconstruction categories.

This section will outline the various contributions to the full probability density function used to fit data. We use a similar method as in Part I of the analysis, summing the probability density function (PDF) for signal and backgrounds but including both B^0 and B^0_s signals, and with additional signal cross-feeds. We generally use simulation to fix parameters when we can't take values from data (such as background shapes), as discussed further below.

The full fit procedure has been developed and tested whilst blinding the signal yields as well as masking those regions on any plots, excluding the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. We also test the stability of the full model by means of "Toy Monte Carlo" (toy MC) studies, which have indicated the constraints on the model parameters in order to obtain a stable fit and are discussed more in Sec. 11.4.

11.1 Signal model

As done in Part 1 of the analysis, the signal is modelled with the sum of two CB functions (Eq. 7.2) that share common values for the peak position and width but have independent tails on opposite sides of the peak. We now add an additional

"complication" in this set-up so that we can fit the modes simultaneously. This is done by constraining several parameters related to the widths of the double-CB function as follows

- the widths of the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ and $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ signals are shared;
- the relative widths of $K^0_S K^+\pi^-$ and $K^0_S K^+ K^-$ to the $K^0_S \pi^+\pi^-$ width, are the same for B^0 and B^0_s decay modes;
- the ratio of widths between the LL and DD samples are the same for all decay modes;
- the peak of the B^0 component, m_{B^0}, is floated whilst the B^0_s peak, $m_{B^0_s}$, is constrained by the mass difference, $m^{\text{float}}_{B^0_s} = m^{\text{float}}_{B^0} + (m_{B^0_s} - m_{B^0})_{\text{PDG}}$ where we take the difference from the PDG value, $(m_{B^0_s} - m_{B^0})_{\text{PDG}} = 87.35$ MeV/c^2 [15].

We need to fix the CB tail parameters n and α from simulation because the tails overlap with combinatorial background and cross-feeds and we have limited statistics to determine these from a fit to data. It also makes sense given that a large proportion of our modes were blinded. Table 11.1 summarises the fit parameters to simulated signal datasets for both the B^0 and B^0_s modes. A consistent result is obtained when fitting to the B^0_s modes hence we are confident in the choice of using the same PDF for describing both B^0_s and B^0 signal shapes.[1]

As in Part 1 of the analysis the parameters α an n govern the shape of the tails of the two CB functions. Since this model is fit simultaneously to all the modes we can parameterise the values of the tail parameters in one mode relative to another, using scale factors. The parameters $kn_1(J/\psi K^0_S K^+\pi^-)$ and $kn_1(J/\psi K^0_S K^+ K^-)$ as well as $k\alpha_1(J/\psi K^0_S K^+\pi^-)$ and $k\alpha_1(J/\psi K^0_S K^\pm K^\mp)$ are such multiplicative factors for the $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ modes respectively to account for difference in PDF shape for each mode, these are then fixed in the fit to data to the values obtained in simulation. We can make similar parametrisations between the different datasets and in fact as mentioned above we do so for the widths, by the ratio relative to the $B \to J/\psi K^0_S \pi^+\pi^-$ mode.

[1] The only slight difference is the width, σ_1 parameter. In this case the Q value is smaller in the B^0 decay meaning the resolution is slightly better and thus smaller width. One could add a Gaussian constraint to the B^0_s width but since we do not even know if any modes exist we felt it better to keep the model relatively simple in the case of unobserved modes. A systematic uncertainty will be assigned due to the choice of PDF but the systematic due to the difference in PDF between B^0_s or B^0 shapes should be small since most parameters agree very well.

Table 11.1: Fitted parameters from fits to MC for $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ modes.

Parameter	B^0 values	B^0_s values
α_1	1.03 ± 0.24	1.10 ± 0.21
f_2	0.714 ± 0.089	0.66 ± 0.10
$k\alpha_1(J/\psi K^0_S K^+ K^-)$	1.23 ± 0.22	1.32 ± 0.21
$k\alpha_1(J/\psi K^0_S K^+ \pi^-)$	1.03 ± 0.20	1.38 ± 0.28
α_2/α_1	-1.57 ± 0.43	-1.47 ± 0.35
$kn_1(J/\psi K^0_S K^+ K^-)$	0.98 ± 0.41	0.83 ± 0.35
$kn_1(J/\psi K^0_S K^+ \pi^-)$	1.03 ± 0.48	0.46 ± 0.25
n_2/n_1	1.86 ± 0.77	1.58 ± 0.67
$\sigma(J/\psi K^0_S K^+ K^-)/\sigma(J/\psi K^0_S \pi^+\pi^-)$	0.700 ± 0.015	0.702 ± 0.015
$\sigma(J/\psi K^0_S K^+ \pi^-)/\sigma(J/\psi K^0_S \pi^+\pi^-)$	0.874 ± 0.020	0.873 ± 0.024
$\sigma(\mathrm{LL})/\sigma(\mathrm{DD})$	0.923 ± 0.016	0.902 ± 0.016
m_{B^0} (MeV/c^2)	5279.437 ± 0.054	5366.127 ± 0.058
n_1	3 ± 1	4 ± 1
σ_1 (MeV/c^2)	5.88 ± 0.11	6.33 ± 0.12

11.2 Signal cross-feed model

By construction of the particle identification cuts, every selected candidate belongs to a distinct independent data sample, unique to one reconstructed final state. However, misidentified decays result in some cross-feed to other samples. These can be characterised using simulation. In Part 1 of the analysis for the measurement of the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ branching fraction we assumed that cross-feed was negligible, which seems a reasonable assumption (and can be checked when the fit in Part 2 of the analysis is unblinded). However, the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ channel may be the source of some cross-feed to other channels, in particular to the $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ mass spectrum.

All possible cross-feeds are studied using MC, as shown in Figures 11.1 and 11.2 for the B^0 and B^0_s modes respectively. Events shown in these plots have passed the selection and trigger requirement only, no particle identification require-

ment has been applied to retain sufficient statistics to gauge the remaining distributions shape. The PDF chosen to model these shapes is a one-dimensional kernel estimator known as a RooKeysPdf. The RooKeysPdf sums together a series of Gaussian kernels forming a spanning pdf of the probability space [96]. The user can specify a smoothing parameter which defines the cohesion to the data points. We set the smoothing parameter to 0.82 for all modes; this provides good agreement with the data points especially in the nominal fit region. In Figures 11.1 and 11.2 red lines have been drawn to show the nominal fitting range. It can be seen that none of these misidentified decays peak directly underneath our signal region, though some give tails close to the signal peaks and therefore need to be modelled appropriately.

For every possible cross-feed, the yield is constrained to be equal to the number of signal candidates in its corresponding true spectrum, multiplied by the relevant misidentification efficiency (see Table 12.9). Here, the misidentification efficiency includes contributions from both PID and the probability that the misidentified decay enters the fit range; the latter values are given in Table 11.2. Due to the low efficiencies outlined in Table 12.9 for doubly mis-reconstructed modes, $K^+K^- \to \pi^+\pi^-$ or vice versa, these are neglected. The constraints on the cross-feed yields are implemented in the fit as Gaussian terms multiply the likelihood function. The central values of the Gaussian are obtained from the ratios of yields and selection efficiencies, and the widths are based on the systematic uncertainties of the PID efficiencies. These Gaussian constraints essentially provide some "wiggle" room for each of the total cross-feed yields due to our uncertainty on values taken from MC. The list of constraints can be found in Table 11.3 which presents the mean values and the uncertainties for all modes.[2]

11.3 Combinatorial background model

Since we expect no peaking background structure other than that which we have discussed for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ in Sec. 5, we employ a freely floating linear shape for the combinatorial backgrounds in each channel (separate for DD and LL). We use a linear shape rather than the exponential model for combinatorial background used in Part I (Sec. 7.2) since the former is seen to be more stable in channels with low background yields, such as $B \to J/\psi K^0_S K^+K^-$, and it is preferable to use the same shape for all channels in the simultaneous fit.

[2] In the simultaneous fit we do not fit to the magnet up and down polarities separately, thus for ease of calculation we take the largest of the magnet up and down values (the values are in agreement for both magnet polarities).

Table 11.2: Efficiencies of B candidate reconstructed under a different final state to that which was generated. The full selection and trigger have been applied here but not particle identification. The values are obtained using the total sample combining both magnet up and down polarities, the errors are statistical only.

	K^0_S(LL) (%)	K^0_S(DD) (%)
$\epsilon^{\text{selection}}(B^0 \to J/\psi K^0_S \pi^+\pi^-$ as $J/\psi K^0_S K\pi)$	0.040 ± 0.002	0.077 ± 0.003
$\epsilon^{\text{selection}}(B^0 \to J/\psi K^0_S K^\pm\pi^\mp$ as $J/\psi K^0_S \pi\pi)$	0.032 ± 0.002	0.071 ± 0.003
$\epsilon^{\text{selection}}(B^0 \to J/\psi K^0_S K^\pm\pi^\mp$ as $J/\psi K^0_S KK)$	0.019 ± 0.002	0.032 ± 0.002
$\epsilon^{\text{selection}}(B^0 \to J/\psi K^0_S K^+K^-$ as $J/\psi K^0_S K\pi)$	0.036 ± 0.002	0.079 ± 0.002
$\epsilon^{\text{selection}}(B^0_s \to J/\psi K^0_S \pi^+\pi^-$ as $J/\psi K^0_S K\pi)$	0.021 ± 0.002	0.033 ± 0.002
$\epsilon^{\text{selection}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$ as $J/\psi K^0_S \pi\pi)$	0.105 ± 0.005	0.234 ± 0.005
$\epsilon^{\text{selection}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$ as $J/\psi K^0_S KK)$	0.006 ± 0.001	0.013 ± 0.001
$\epsilon^{\text{selection}}(B^0_s \to J/\psi K^0_S K^+K^-$ as $J/\psi K^0_S K\pi)$	0.184 ± 0.005	0.410 ± 0.007

11.4 Toy studies

In order to test the stability of our fit model and that the associated statistical errors on parameters are sensible we perform toy experiments. With the full fit model as defined in Sec. 11 we fit the model to real data using the fitted values to then generate 500 toy experiments. Note that the toy studies are performed with the results of the blind fit to data – *i.e.* the correct central values are used, but are not examined. All variable pull distributions should be normally distributed $\mathcal{N}\ (0, 1)$ thus the means are approximately zero and the widths are approximately unity which informs us that the variation is due to statistical fluctuations.

The results are shown in Table 11.4. Figures 11.3 and 11.4 show the pull distribution plots for the B^0 and B^0_s signal yields.

The results of the toy studies show that all biases on the fitted yields are small, less than 20 % of the statistical uncertainty. However, the biases are found to be significant for $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ and $B^0_s \to J/\psi K^0_S K^+K^-$. The pull widths are found to be larger than unity (by about 10 %) for these channels. Note that these are two channels where low signal yields are expected, and in the case of low background, biases are anticipated since the total yield is non-negative. The biases

are understood as by running toys with the yield set to an arbitrary large value (100) leaves a very stable result. To have a uniform procedure, we treat the biases on all yields in the same way, namely the central values are corrected according to the biases found in the toy experiments. In addition, a systematic uncertainty is assigned – it is evaluated as the sum in quadrature of the uncertainty on the bias and half the correction. The inclusion of half the correction accounts for the fact that we cannot trust our toy experiments to determine the bias perfectly. We do not apply any scaling of the statistical error, since we expect the effect on the pull width to be correlated with the bias.

11.5 Unblinded fit results

Upon unblinding the results of the simultaneous fit described in section 11, we obtain the distributions found in Figures 11.5, 11.6 and 11.7 for the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ modes respectively. Each mode is presented for both K^0_S(LL) and K^0_S(DD) categories with both linear and logarithmic y-axis scales. A summary of the various yields and floating parameters after the fit is given in Table 11.5. No correlations were observed above 10 % for any signal yield relative to the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ normalisation yield. There are some correlations (at the -40 % level) between background parameters, this is to be expected as the shapes of cross-feed and combinatorial background components are rather similar and prevent the total pdf yield from going negative.

The significance values listed for the signal yields in Table 11.5 are found from the change in twice the negative log-likelihood values obtained from a fit to the background only and signal and background hypotheses, via the formula $S_{\text{stat}} = \sqrt{-2\ln(\frac{\mathcal{L}_{S+B}}{\mathcal{L}_B})}$. Although the significance is only statistical (*i.e.* it does not include the effects of systematic uncertainty) at this point, there is clear evidence for the $B^0 \to J/\psi K^0_S \pi^+\pi^-$, $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0 \to J/\psi K^0_S K^+ K^-$ modes. The combined significance for DD and LL modes is obtained in an analogous manner by setting both yields to zero. Table 11.6 shows the statistical significance of each of the modes.

11.5.1 Bias correction

As shown in Sec. 11.4, prior to unblinding a strategy for bias correction was established. Specifically, although all biases are at most 20 % of the statistical error, all fitted central values are corrected according to the biases found in the toy experiments, and a systematic uncertainty is assigned. Table 11.7 presents the central

values from the fits and the corrected yields. The corrections, obtained from the mean of a Gaussian fit to the residual distributions in the pseudo-experiments, are generally negligible for significant yields, but we still follow the procedure set before unblinding. From here on, whenever we refer to signal yields we take the corrected value.

11.6 Invariant mass distributions

Clear signals are seen for $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ and $B^0 \to J/\psi K_S^0 K^+ K^-$ decays. The distributions of the signal decays in the available phase-space are examined, using the *sPlot* technique [97], with the B candidate invariant mass as the discriminating variable.

None of the channels show any structures in any invariant mass combinations involving the J/ψ meson (recall that the $\psi(2S)$ contribution in $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ decays is vetoed; in the same decay a small but not significant excess is seen around the $X(3872)$ in $m(J/\psi\pi^+\pi^-)$). In $B^0 \to J/\psi K_S^0 \pi^+\pi^-$, excesses from $K^*(892)$ and $\rho(770)$ mesons can be seen in $m(K_S^0\pi^\pm)$ and $m(\pi^+\pi^-)$ respectively, and there is a possible enhancement from the $K_1(1400)$ state in $m(K_S^0\pi^+\pi^-)$, as shown in Figs. 11.8 and 11.9. In $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ (Figs. 11.10 and 11.11), excesses from $K^*(892)$ resonances can be seen in $m(K_S^0\pi^\pm)$ and $m(K^\pm\pi^\mp)$, but no narrow structures are seen in $m(K_S^0 K^\pm \pi^\mp)$. Figure 11.12 shows a zoomed in region of the low invariant mass of the $m(K_S^0 K^\pm \pi^\mp)$ spectra. Upon closer inspection it seems to show hints of the $f_1(1420)$ resonance and with more data this will become clear, but is outside the scope of this thesis. In $B^0 \to J/\psi K_S^0 K^+ K^-$ (Figs. 11.13 and 11.14), the $\phi(1020)$ peak can be seen in $m(K^+K^-)$, but no other narrow structures are evident in any combination.

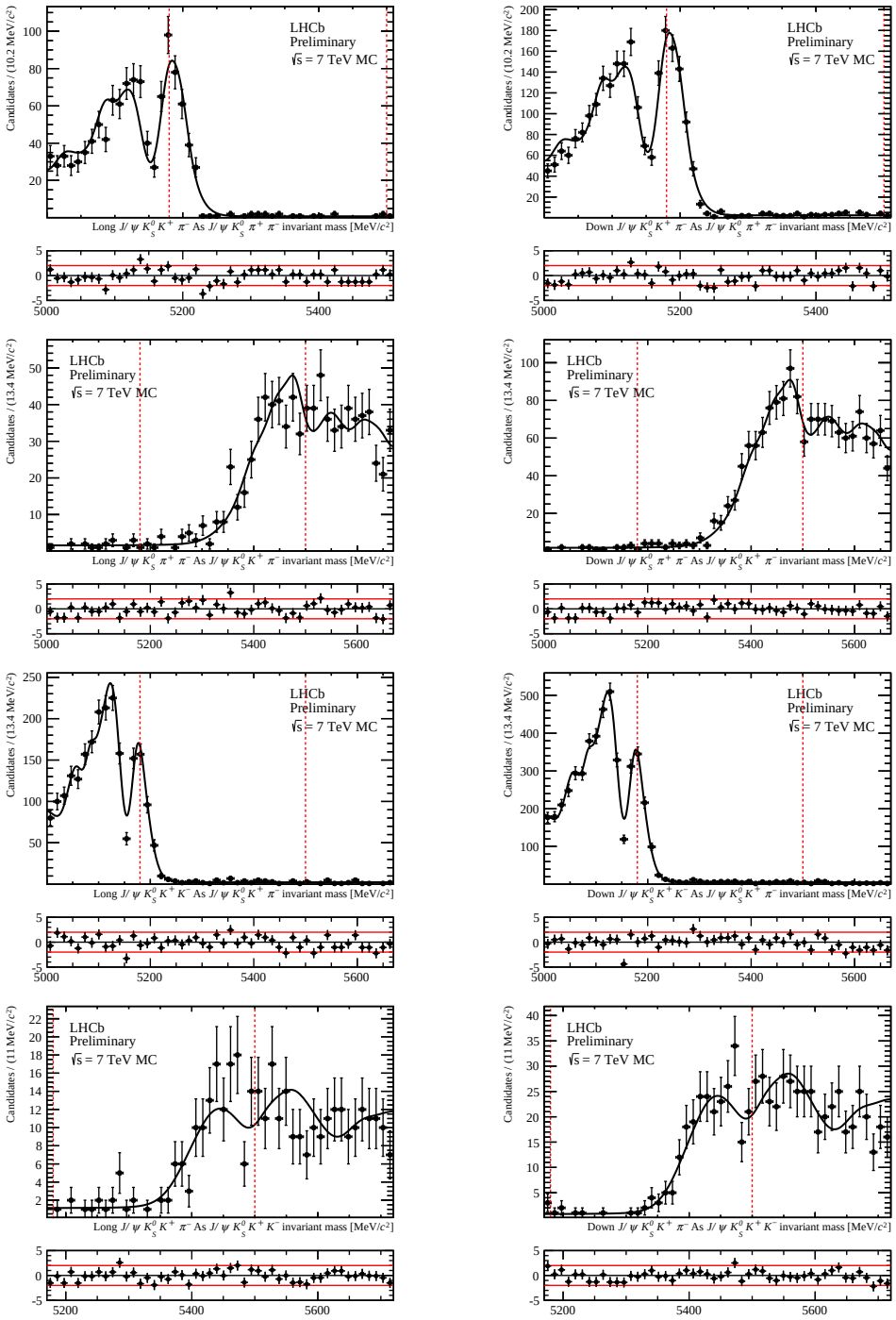

Figure 11.1: Various mis-identification invariant mass distributions for $B^0 \to J/\psi K^0_S h^+ h^{(\prime)-}$ where a $\pi \Leftrightarrow K$ is interchanged under the wrong mass hypothesis. The left side column are K^0_S(LL) and right column K^0_S(DD) KS types. From top to bottom we have: $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi K^0_S \pi^+ \pi^-$, $B^0 \to J/\psi K^0_S \pi^+ \pi^-$ reconstructed as $J/\psi K^0_S K^+ \pi^-$, $B^0 \to J/\psi K^0_S K^+ K^-$ reconstructed as $J/\psi K^0_S K^+ \pi^-$ and finally $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi K^0_S K^+ K^-$. The red lines are for visual purposes and indicate our nominal fitting range.

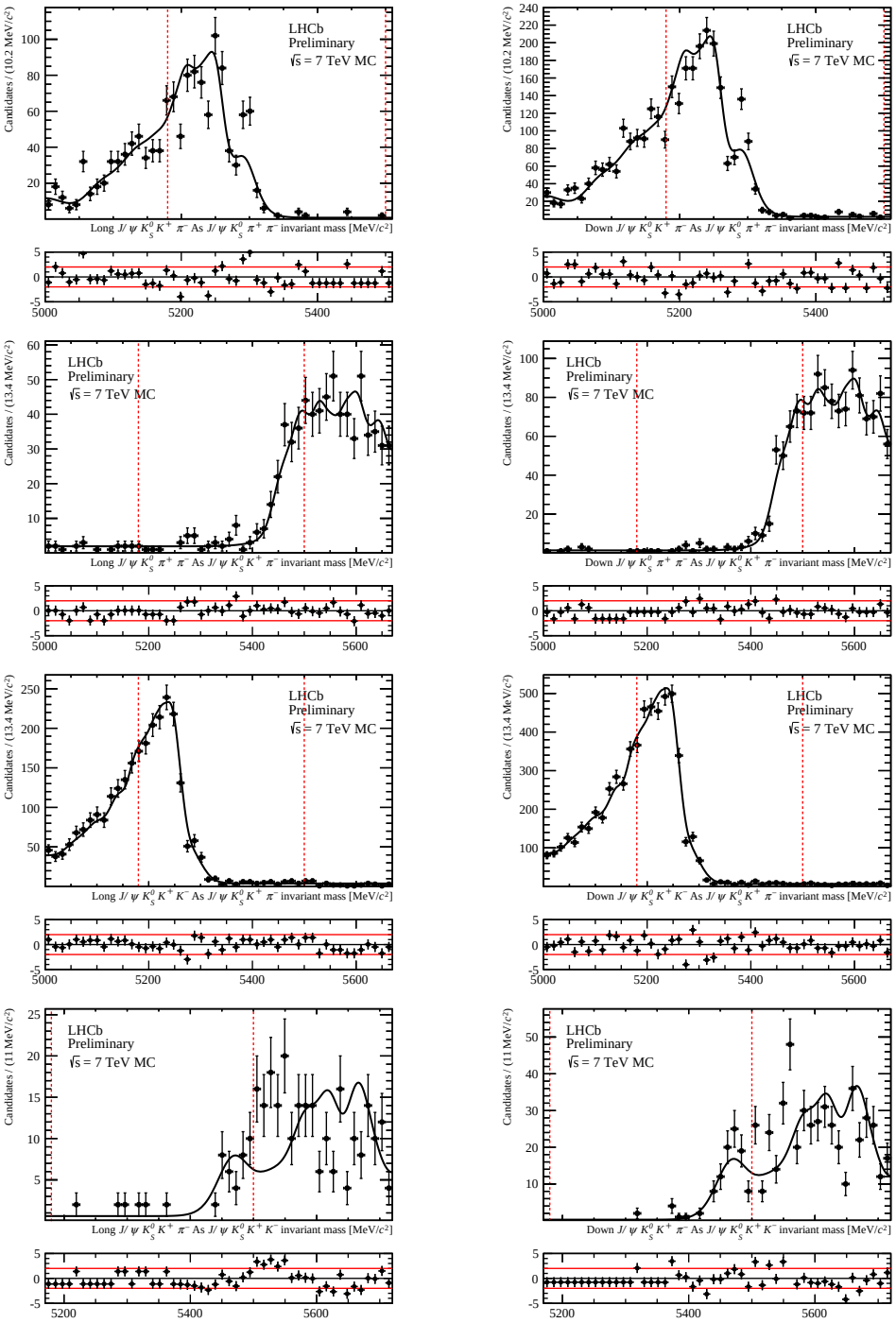

Figure 11.2: Various mis-identification invariant mass distributions for $B^0_s \to J/\psi K^0_S h^+ h^{(\prime)-}$ where a $\pi \Leftrightarrow K$ is interchanged under the wrong mass hypothesis. The left side column are K^0_S(LL) and right column K^0_S(DD) KS types. From top to bottom we have: $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi K^0_S \pi^+ \pi^-$, $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$ reconstructed as $J/\psi K^0_S K^+ \pi^-$, $B^0_s \to J/\psi K^0_S K^+ K^-$ reconstructed as $J/\psi K^0_S K^+ \pi^-$ and finally $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ reconstructed as $J/\psi K^0_S K^+ K^-$. The red lines are for visual purposes and indicate our nominal fitting range.

Table 11.3: Gaussian constraints that are applied to the cross-feed yields for all modes. If a mode is not present then its yield has been deemed negligible and thus omitted. The numerator and denominator are taken in an analagous fashion; $\epsilon^{\text{Generator}} \times \epsilon_X^{\text{Selection}} \times \epsilon_X^{\text{PIDCalib}}$ where X corresponds to whether it is a mis-identified (numerator) or truth (denominator) efficiency.

	$K^0_{\rm S}$(LL)(%)	$K^0_{\rm S}$(DD)(%)
$\frac{N(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-\text{ as }J/\psi\, K^0_{\rm S}K\pi)}{N(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-\text{ as }J/\psi\, K^0_{\rm S}\pi\pi)}$	2.71 ± 2.11	2.35 ± 0.47
$\frac{N(B^0 \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}\pi\pi)}{N(B^0 \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}K\pi)}$	0.48 ± 0.45	0.52 ± 0.52
$\frac{N(B^0 \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}KK)}{N(B^0 \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}K\pi)}$	1.47 ± 1.12	1.03 ± 0.76
$\frac{N(B^0 \to J/\psi\, K^0_{\rm S}K^+K^-\text{ as }J/\psi\, K^0_{\rm S}K\pi)}{N(B^0 \to J/\psi\, K^0_{\rm S}K^+K^-\text{ as }J/\psi\, K^0_{\rm S}KK)}$	0.50 ± 0.26	0.43 ± 0.18
$\frac{N(B^0_s \to J/\psi\, K^0_{\rm S}\pi^+\pi^-\text{ as }J/\psi\, K^0_{\rm S}K\pi)}{N(B^0_s \to J/\psi\, K^0_{\rm S}\pi^+\pi^-\text{ as }J/\psi\, K^0_{\rm S}\pi\pi)}$	1.18 ± 0.78	0.77 ± 0.55
$\frac{N(B^0_s \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}\pi\pi)}{N(B^0_s \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}K\pi)}$	0.86 ± 0.45	0.99 ± 0.99
$\frac{N(B^0_s \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}KK)}{N(B^0_s \to J/\psi\, K^0_{\rm S}K^\pm\pi^\mp\text{ as }J/\psi\, K^0_{\rm S}K\pi)}$	0.46 ± 0.33	0.35 ± 0.24
$\frac{N(B^0_s \to J/\psi\, K^0_{\rm S}K^+K^-\text{ as }J/\psi\, K^0_{\rm S}K\pi)}{N(B^0_s \to J/\psi\, K^0_{\rm S}K^+K^-\text{ as }J/\psi\, K^0_{\rm S}KK)}$	1.71 ± 0.81	1.61 ± 0.22

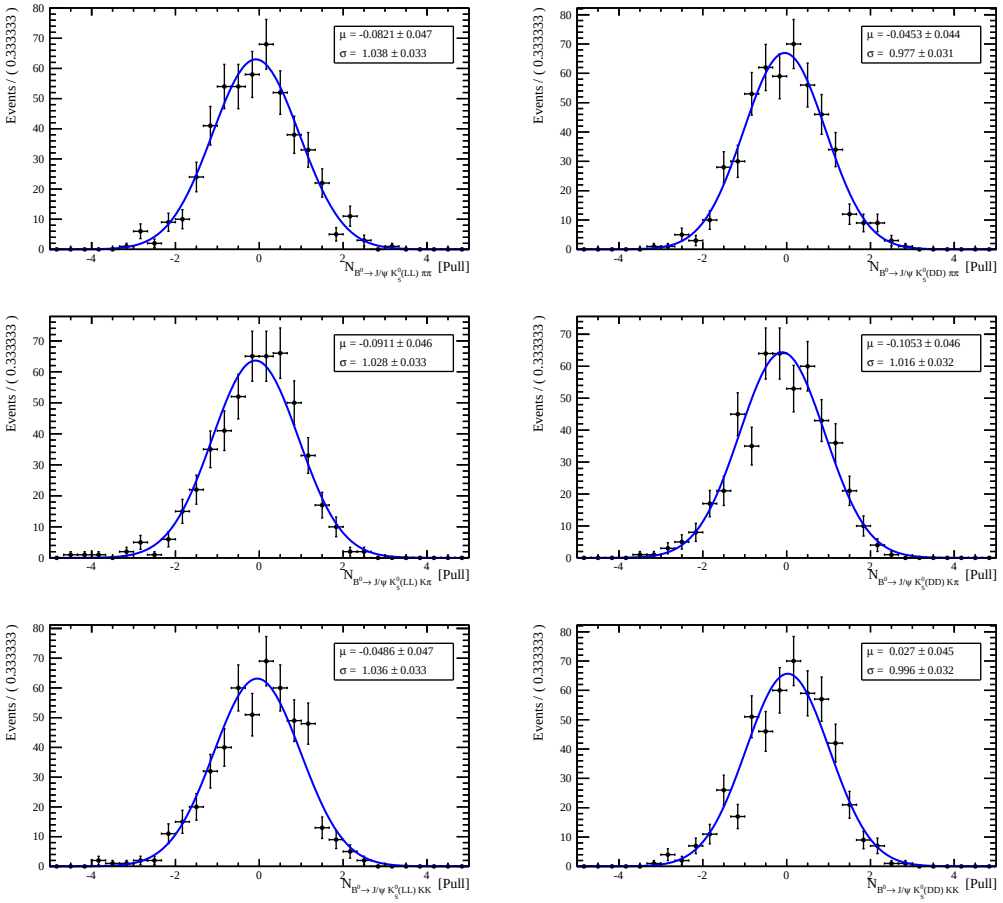

Figure 11.3: Unblinded pull distribution plots for B^0 signal yields. (Top) $B^0 \to J/\psi K^0_S \pi^+\pi^-$, (middle) $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ and (bottom) $B^0 \to J/\psi K^0_S K^+ K^-$. (Left) LL and (right) DD samples respectively.

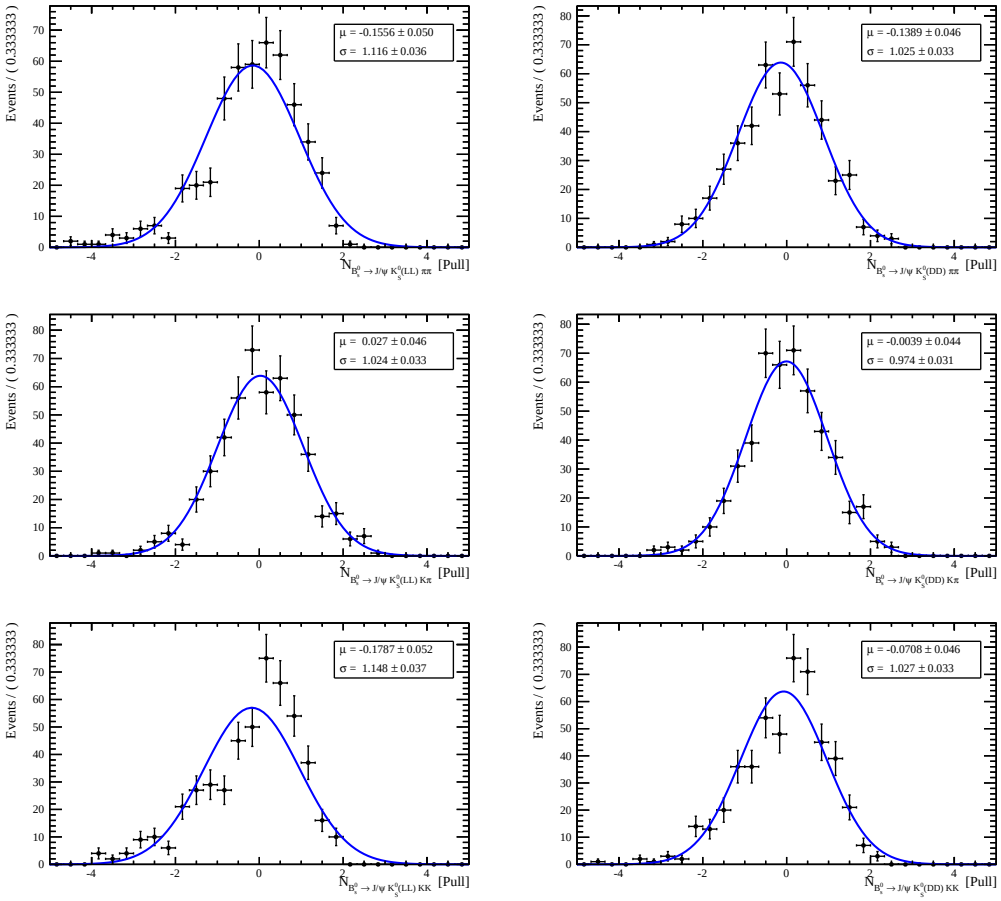

Figure 11.4: Pull distribution plots for B_s^0 signal yields. (Top) $B_s^0 \to J/\psi K_S^0 \pi^+\pi^-$, (middle) $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ and (bottom) $B_s^0 \to J/\psi K_S^0 K^+ K^-$. (Left) LL and (right) DD samples respectively.

Table 11.4: Pull mean and width values for parameters of the fit to data. The yields for the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm\pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ modes include goodness of fit information and correspond to figures 11.3 and 11.4 for B^0 and B^0_s modes respectively.

Parameter	Pull Mean	Pull Width	Probability (χ^2, ndf)
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	-0.082 ± 0.047	1.04 ± 0.03	$0.69\,(22.9, 27)$
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	-0.045 ± 0.044	0.98 ± 0.03	$0.95\,(16.4, 27)$
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	-0.091 ± 0.046	1.03 ± 0.03	$0.25\,(31.5, 27)$
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	-0.105 ± 0.046	1.02 ± 0.03	$0.95\,(16.4, 27)$
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	-0.049 ± 0.047	1.04 ± 0.03	$0.60\,(24.6, 27)$
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	0.027 ± 0.045	1.00 ± 0.03	$0.28\,(30.8, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	-0.156 ± 0.050	1.12 ± 0.04	$1.16 \times 10^{-9}\,(96, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	-0.139 ± 0.046	1.02 ± 0.03	$0.99\,(12.9, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	0.027 ± 0.046	1.02 ± 0.03	$0.24\,(31.7, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	-0.004 ± 0.044	0.97 ± 0.03	$0.99\,(11.6, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	-0.179 ± 0.052	1.15 ± 0.04	$1.00 \times 10^{-3}\,(56.3, 27)$
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	-0.071 ± 0.046	1.03 ± 0.03	$0.22\,(32.4, 27)$
$\sigma(J/\psi K^0_S K^+K^-)/\sigma(J/\psi K^0_S\pi^+\pi^-)$	0.029 ± 0.047	1.05 ± 0.03	–
$\sigma(J/\psi K^0_S K^+\pi^-)/\sigma(J/\psi K^0_S\pi^+\pi^-)$	-0.031 ± 0.046	1.02 ± 0.03	–
$\sigma(\mathrm{LL})/\sigma(\mathrm{DD})$	-0.050 ± 0.044	0.98 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	-0.034 ± 0.044	0.99 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	-0.048 ± 0.042	0.94 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	-0.013 ± 0.046	1.03 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	0.019 ± 0.043	0.97 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})K^+K^-)$	-0.094 ± 0.045	1.01 ± 0.03	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})K^+K^-)$	-0.110 ± 0.047	1.06 ± 0.03	–
$\mu_{B^0}(\,\mathrm{MeV}/c^2)$	0.003 ± 0.044	1.07 ± 0.03	–
$\sigma_1(\,\mathrm{MeV}/c^2)$	0.024 ± 0.045	0.99 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	0.018 ± 0.045	1.01 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	0.031 ± 0.047	1.05 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	-0.053 ± 0.047	1.04 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	-0.035 ± 0.046	1.02 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{LL})K^+K^-)$	0.036 ± 0.048	1.06 ± 0.03	–
$\mathrm{slope}(J/\psi K^0_S(\mathrm{DD})K^+K^-)$	-0.012 ± 0.043	0.96 ± 0.03	–

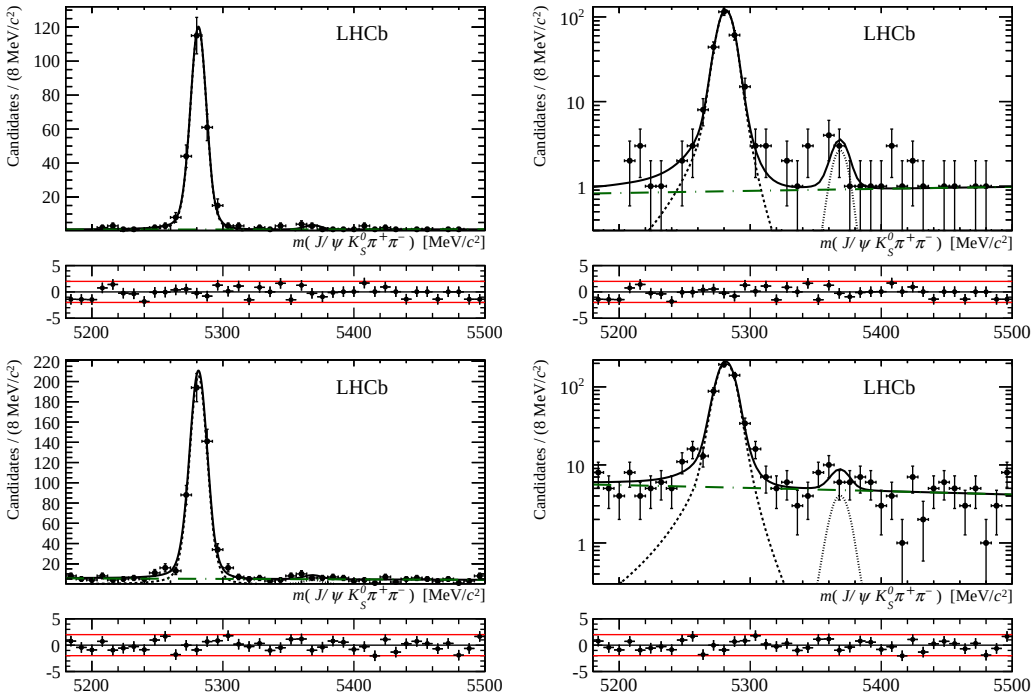

Figure 11.5: Invariant mass fits to the $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively.

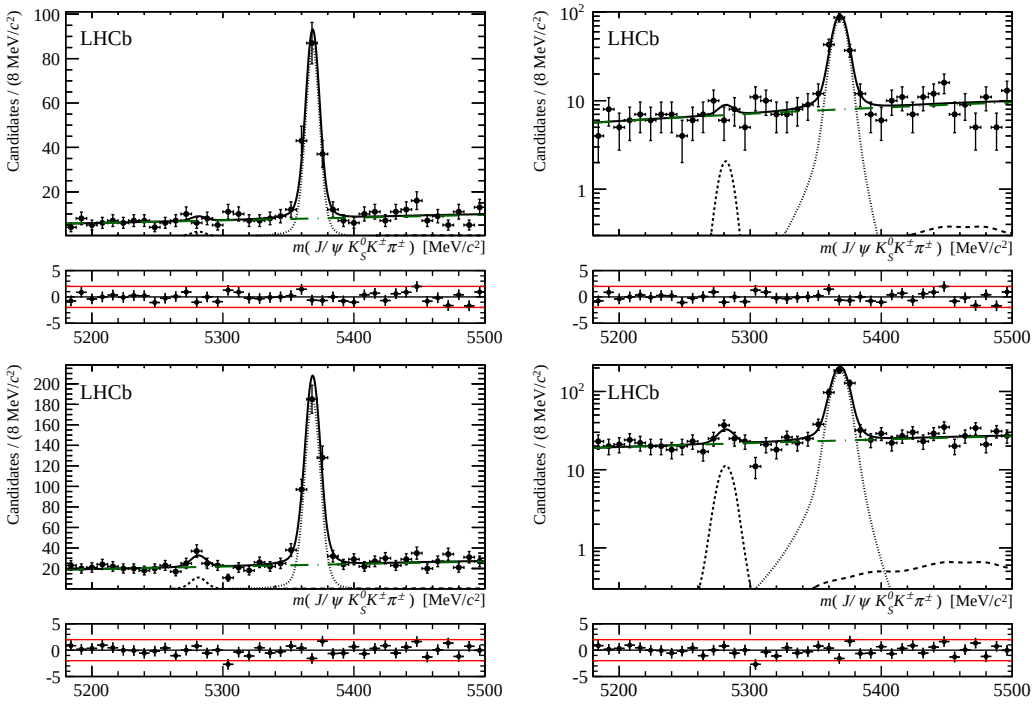

Figure 11.6: Unblinded invariant mass fits to the $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively.

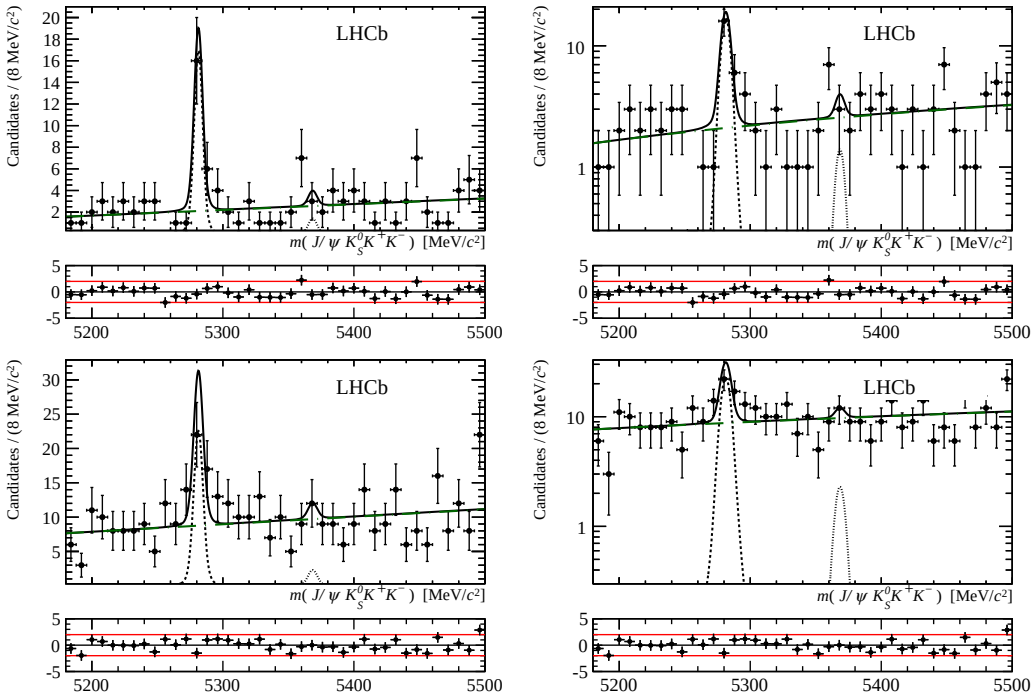

Figure 11.7: Unblinded invariant mass fits to the $B^0_{(s)} \to J/\psi K^0_S K^+ K^-$ 2011 data invariant mass spectra with (right) plots displayed on a logarithmic scale. (Top) and (bottom) display LL and DD samples respectively.

Table 11.5: Fitted parameters from unblinded simultaneous fit to 2011 dataset. The significance of the signal yields are statistical only and do not include systematic corrections.

Parameter	Value	Significance
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	$246.3^{+16.5}_{-15.9}$	28.2
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	$471.0^{+23.7}_{-23.0}$	35.4
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	$3.7^{+5.5}_{-4.8}$	0.8
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	$22.6^{+10.4}_{-9.6}$	2.5
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	$18.4^{+5.2}_{-4.5}$	6.2
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	$27.2^{+7.9}_{-7.1}$	5.1
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	$5.5^{+3.5}_{-2.8}$	2.3
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	$9.3^{+6.1}_{-5.3}$	1.9
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	$154.3^{+14.5}_{-13.8}$	17.9
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	$370.9^{+23.4}_{-22.6}$	25.8
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	$2.0^{+3.0}_{-2.2}$	0.7
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	$2.8^{+5.1}_{-4.3}$	0.6
$\mu_{B^0}(\,\mathrm{MeV}/c^2)$	5281.33 ± 0.22	–
$\sigma_1(\,\mathrm{MeV}/c^2)$	6.85 ± 0.33	–
$\sigma(J/\psi K^0_S K^+K^-)/\sigma(J/\psi K^0_S\pi^+\pi^-)$	0.53 ± 0.11	–
$\sigma(J/\psi K^0_S K^+\pi^-)/\sigma(J/\psi K^0_S\pi^+\pi^-)$	0.88 ± 0.06	–
$\sigma(\mathrm{LL})/\sigma(\mathrm{DD})$	0.90 ± 0.06	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	36 ± 8	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	195 ± 18	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})K^+K^-)$	96 ± 10	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})K^+K^-)$	376 ± 21	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	304 ± 20	–
$N_{\mathrm{comb.}}(J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	915 ± 34	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S K^+K^-/B^0 \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{LL})$	0.005 ± 0.005	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S K^+K^-/B^0 \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{DD})$	0.010 ± 0.010	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S K^+K^-/B^0_s \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{LL})$	0.005 ± 0.004	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S K^+K^-/B^0_s \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{DD})$	0.003 ± 0.004	–
$r_{\mathrm{CrossFeed}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp/B^0_s \to J/\psi K^0_S K^+K^-;\,\mathrm{LL})$	0.017 ± 0.008	–
$r_{\mathrm{CrossFeed}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp/B^0_s \to J/\psi K^0_S K^+K^-;\,\mathrm{DD})$	0.016 ± 0.005	–
$r_{\mathrm{CrossFeed}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp/B^0_s \to J/\psi K^0_S \pi^+\pi^-;\,\mathrm{LL})$	0.012 ± 0.008	–
$r_{\mathrm{CrossFeed}}(B^0_s \to J/\psi K^0_S K^\pm\pi^\mp/B^0_s \to J/\psi K^0_S \pi^+\pi^-;\,\mathrm{DD})$	0.008 ± 0.006	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S \pi^+\pi^-/B^0 \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{LL})$	0.005 ± 0.005	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S \pi^+\pi^-/B^0 \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{DD})$	0.005 ± 0.005	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S \pi^+\pi^-/B^0_s \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{LL})$	0.008 ± 0.005	–
$r_{\mathrm{CrossFeed}}(B^0 \to J/\psi K^0_S \pi^+\pi^-/B^0_s \to J/\psi K^0_S K^\pm\pi^\mp;\,\mathrm{DD})$	0.010 ± 0.010	–

Table 11.6: Combined statistical significance of the DD and LL yields from the simultaneous fit to 2011 dataset. Note that no systematic correction is applied.

Yield	Significance
$N(B^0 \to J/\psi K^0_S \pi^+\pi^-)$	44.3
$N(B^0 \to J/\psi K^0_S K^\pm \pi^\mp)$	2.6
$N(B^0 \to J/\psi K^0_S K^+ K^-)$	8.1
$N(B^0_s \to J/\psi K^0_S \pi^+\pi^-)$	3.0
$N(B^0_s \to J/\psi K^0_S K^\pm \pi^\mp)$	31.0
$N(B^0_s \to J/\psi K^0_S K^+ K^-)$	1.0

Table 11.7: Fitted and corrected yields for $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$, $B^0_{(s)} \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0_{(s)} \to J/\psi K^0_S \pi^+\pi^-$ modes.

Parameter	Fitted Yield	Bias	Corrected Yield
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	246.3	-1.0 ± 1.2	245.3
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	471.0	-2.0 ± 1.7	469.0
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	3.7	-0.3 ± 0.3	3.4
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	22.6	-1.6 ± 0.7	21.0
$N(B^0 \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	18.4	-0.7 ± 0.4	17.7
$N(B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	27.2	1.4 ± 0.5	28.6
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)$	5.5	-0.1 ± 0.3	5.4
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)$	9.3	-0.4 ± 0.4	8.9
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp)$	154.3	0.4 ± 1.1	154.7
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp)$	370.9	-3.5 ± 1.6	367.4
$N(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^+K^-)$	2.0	-0.8 ± 0.2	1.2
$N(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^+K^-)$	2.8	-0.9 ± 0.2	1.9

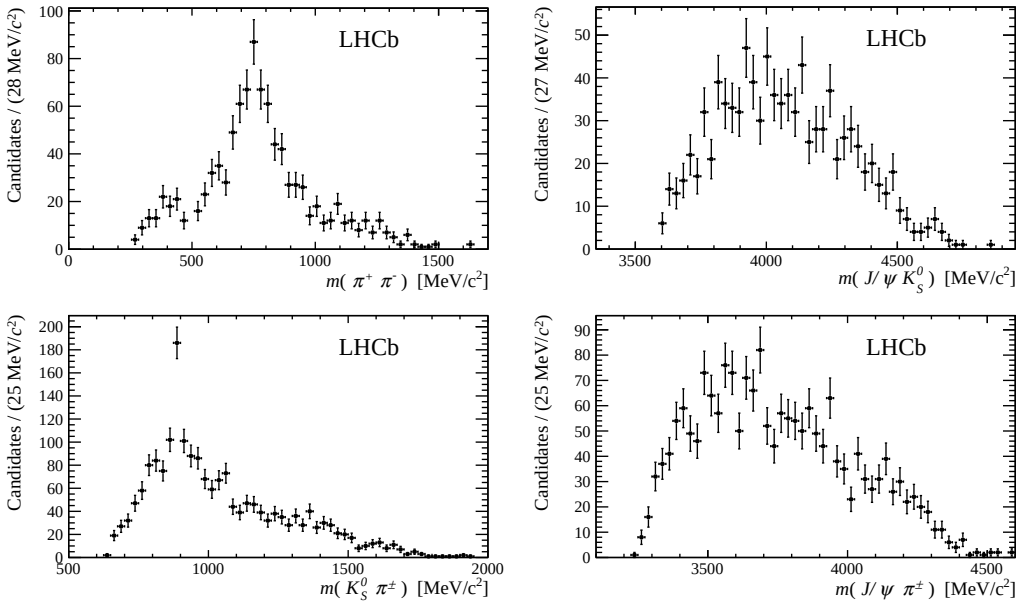

Figure 11.8: Background-subtracted distributions of the possible two-body invariant mass combinations in $B^0 \to J/\psi K^0_S \pi^+\pi^-$ decays. Contributions from the $\rho(770)^0$ and $K^*(892)^\pm$ mesons can be seen in the $m(\pi^+\pi^-)$ and $m(K^0_S\pi^\pm)$ distributions respectively.

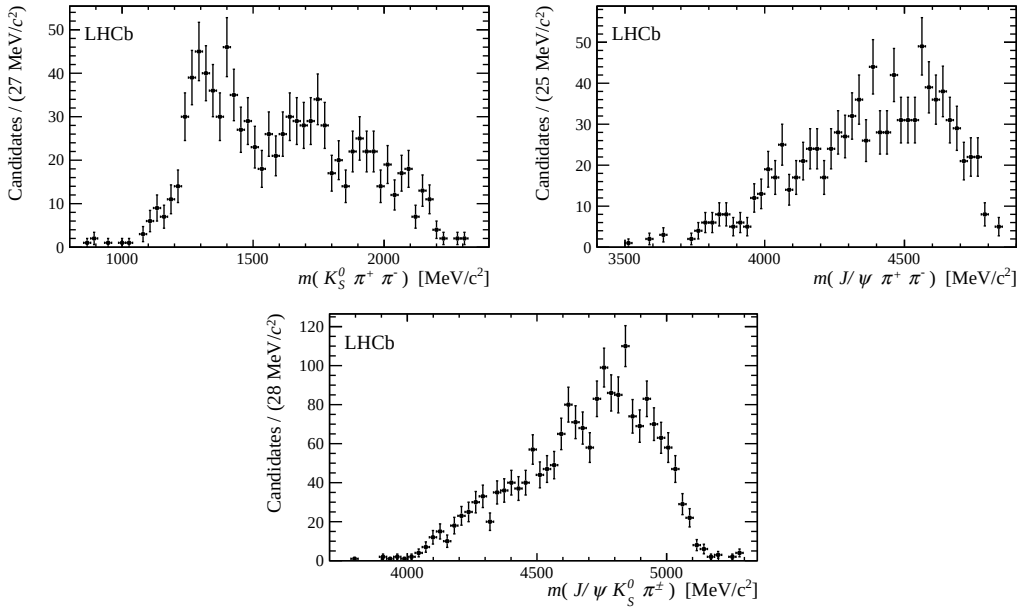

Figure 11.9: Background-subtracted distributions of the possible three-body invariant mass combinations in $B^0 \to J/\psi K^0_S \pi^+\pi^-$ decays. A possible enhancement from the $K_1(1400)$ state can be seen in the $m(K^0_S\pi^+\pi^-)$ distribution.

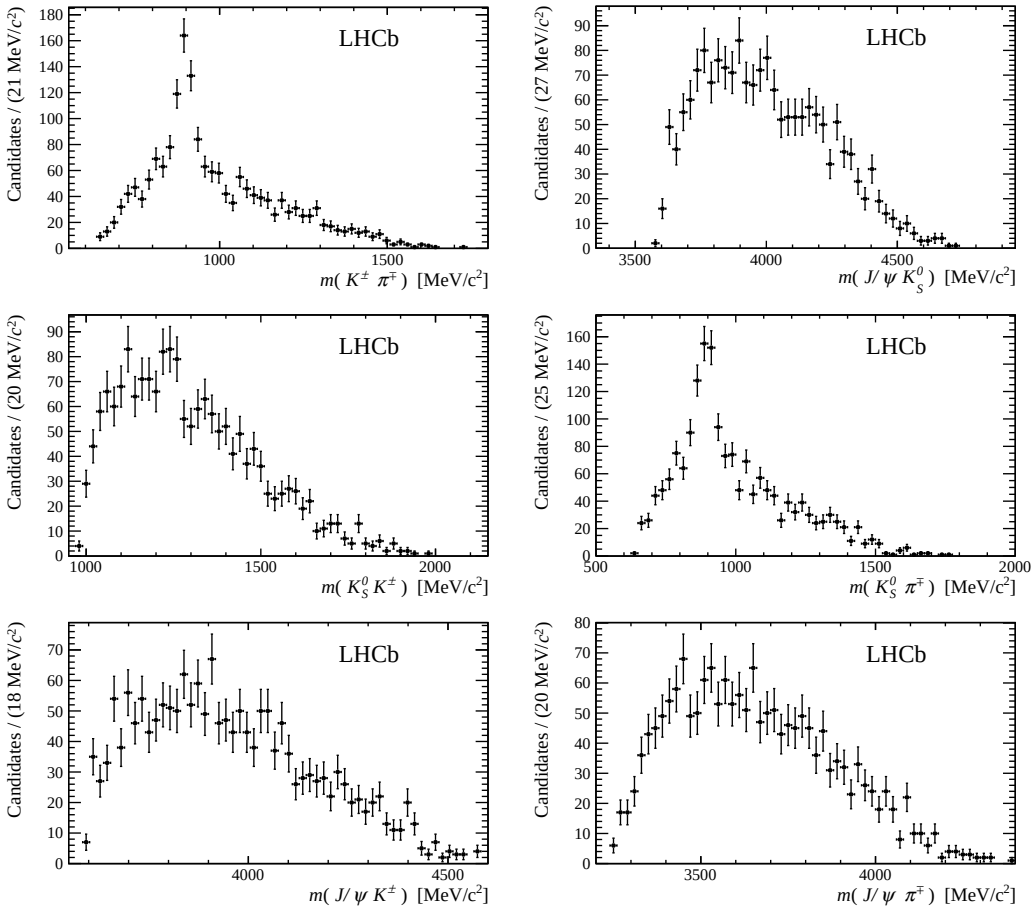

Figure 11.10: Background-subtracted distributions of the possible two-body invariant mass combinations in $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ decays. Contributions from the $K^*(892)^0 + \overline{K}^*(892)^0$ and $K^*(892)^\pm$ mesons can be seen in the $m(K^\pm\pi^\mp)$ and $m(K_S^0\pi^\pm)$ distributions respectively.

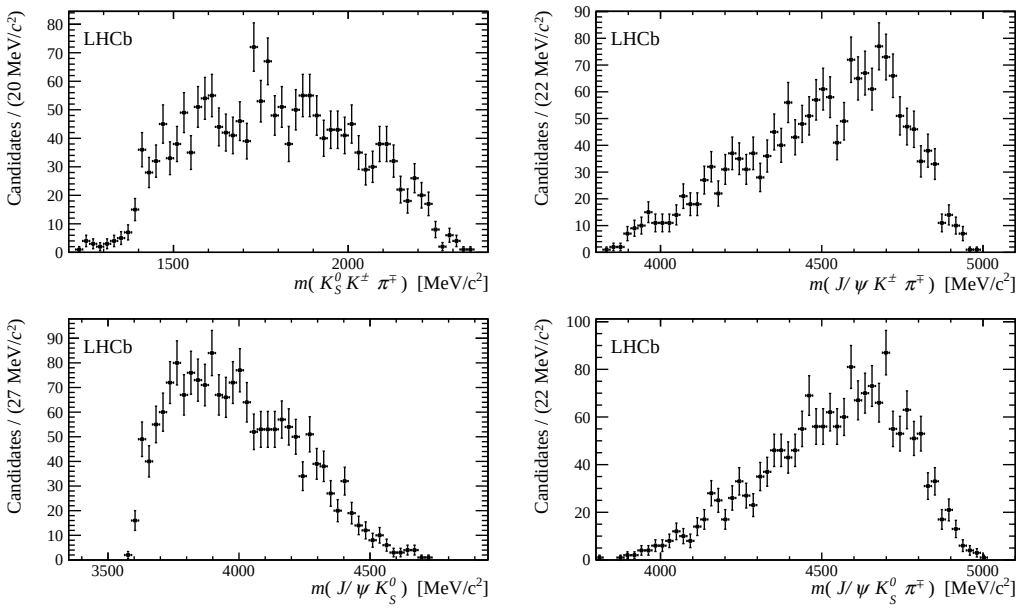

Figure 11.11: Background-subtracted distributions of the possible three-body invariant mass combinations in $B_s^0 \to J/\psi K_S^0 K^{\pm}\pi^{\mp}$ decays. No clear signatures of narrow resonances are observed.

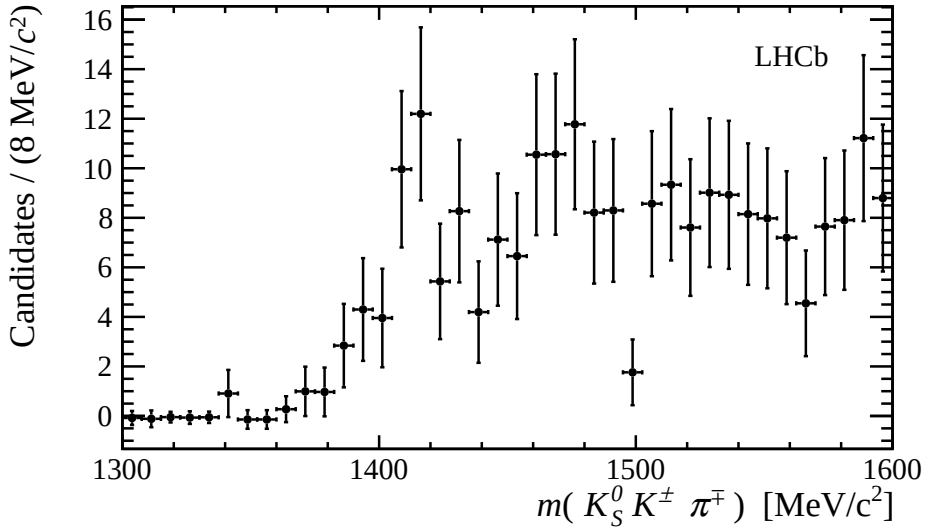

Figure 11.12: A zoom in of the lower invariant mass range of the background-subtracted $m(K_S^0 K^+\pi^-)$ invariant mass combination in $B_s^0 \to J/\psi K_S^0 K^{\pm}\pi^{\mp}$ decays. There are hints of the $f_1(1420)$ resonance but no significant observation is observed.

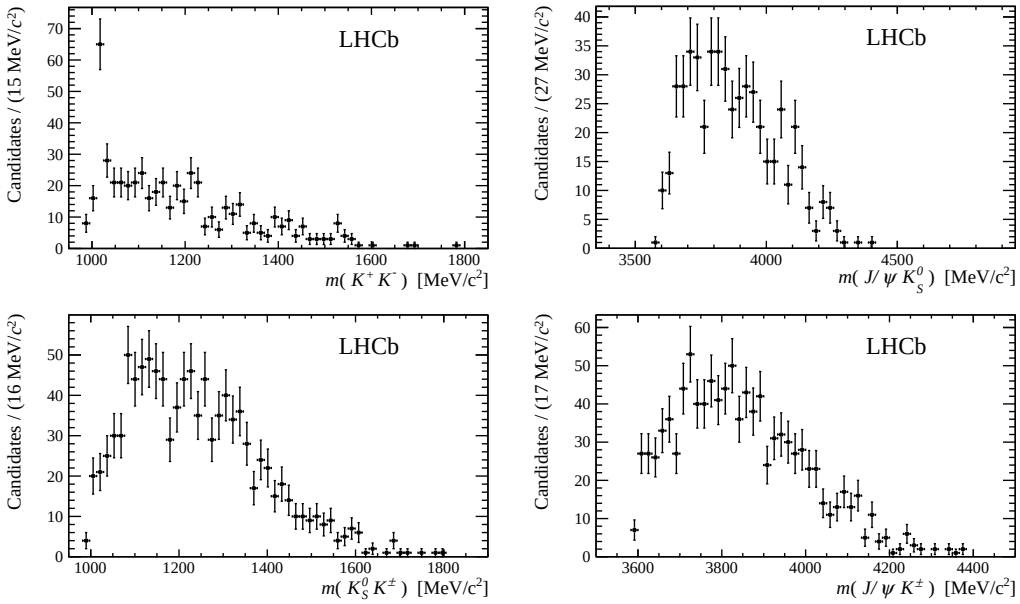

Figure 11.13: Background-subtracted distributions of the possible two-body invariant mass combinations in $B^0 \to J/\psi K^0_S K^+ K^-$ decays. The $\phi(1020)$ resonance is clearly seen in the $m(K^+K^-)$ distribution.

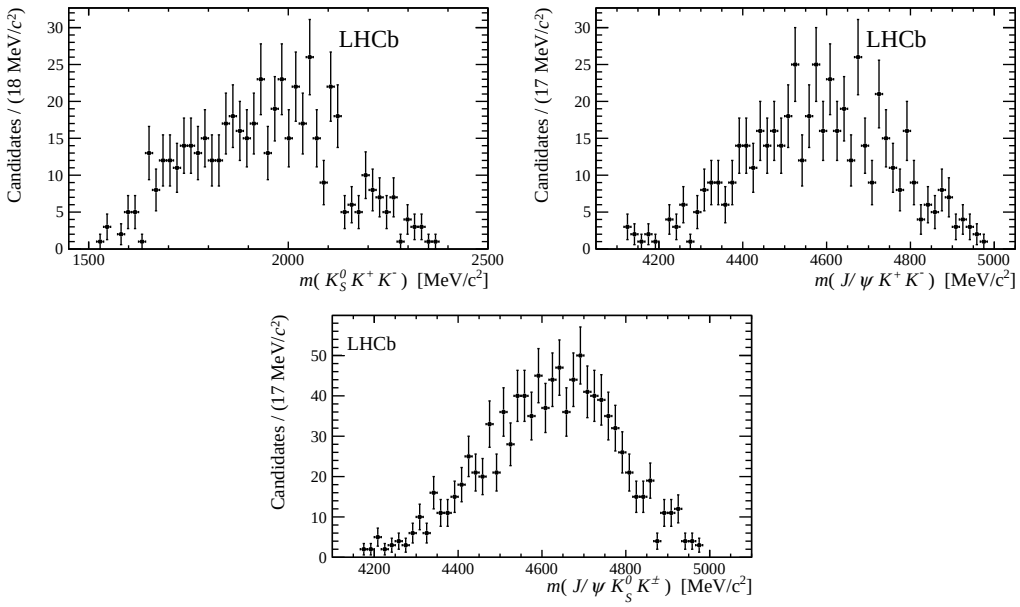

Figure 11.14: Background-subtracted distributions of the possible three-body invariant mass combinations in $B^0 \to J/\psi K^0_S K^+ K^-$ decays. No clear signatures of narrow resonances are observed.

12

Efficiencies and systematic uncertainties

12.1 Discussion of efficiencies

We start by considering the master formula for the relative branching fractions.

$$\frac{\mathcal{B}(B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-})}{\mathcal{B}(B^0 \to J/\psi\, K^0_S \pi^+\pi^-)} = \frac{\epsilon^{\text{Generator}}_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}}{\epsilon^{\text{Generator}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}}} \times \frac{\epsilon^{\text{Selection}}_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}}{\epsilon^{\text{Selection}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}}} \times \frac{\epsilon^{\text{PID}}_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}}{\epsilon^{\text{PID}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}}} \times \frac{N_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}}}{N_{B^0 \to J/\psi\, K^0_S \pi^+\pi^-}} \times \frac{f_d}{f_{d,s}}\,, \tag{12.1}$$

where

$$\epsilon^{\text{Selection}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}} = \epsilon^{\text{reco-strip}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}} \times \epsilon^{\text{trig}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}} \times \epsilon^{\text{selection}}_{B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}}\,. \tag{12.2}$$

The parameters $\epsilon^{\text{Generator}}_X$, $\epsilon^{\text{reco-strip}}_X$, ϵ^{trig}_X, $\epsilon^{\text{selection}}_X$, ϵ^{PID}_X are respectively the generator level cut, reconstruction and stripping, trigger, offline selection, and particle identification efficiencies of the decay mode X, N_X is the number of reconstructed events in this mode and f_d and f_s are the fragmentation fractions of a b quark to hadronise into a B^0 and B^0_s meson respectively. The order of appearance of each ratio of efficiencies in this formula reflects what is assumed to be known to estimate a given ratio: for instance, the estimation of the PID efficiencies assumes that all the other efficiencies are known and thus is computed for events that pass all the rest of the selection, while the estimation of reconstruction and stripping efficiencies only

assumes that the generator level cuts are applied. In total, the efficiency is split into three separate contributions due to: (i) geometric acceptance, (ii) stripping, trigger and selection efficiencies and (iii) particle identification efficiency.

12.1.1 Geometric efficiency

The efficiencies are obtained in an analogous manner as in Part I (Sec. 8.1.1). The generation level statistics for all modes can be found in Table 12.1. It is clear from

Table 12.1: Generator level cut efficiencies for all six modes. Both magnet polarities are presented but they fall well within the statistical limitation of one another. The simulation productions listed here were all run with the `DaughtersInLHCbWithMinP` requirement.

Decay mode	Event Type	Efficiency (%)	
		MagDown	MagUp
$B^0 \to J/\psi K^0_S \pi^+\pi^-$	11246104	11.07 ± 0.03	11.06 ± 0.03
$B^0 \to J/\psi K^0_S K^\pm \pi^\mp$	11246102	12.19 ± 0.05	12.15 ± 0.05
$B^0 \to J/\psi K^0_S K^+ K^-$	11146102	14.19 ± 0.05	14.23 ± 0.05
$B^0_s \to J/\psi K^0_S \pi^+\pi^-$	13246102	10.93 ± 0.03	10.95 ± 0.03
$B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$	13246101	12.30 ± 0.04	12.26 ± 0.04
$B^0_s \to J/\psi K^0_S K^+ K^-$	13146112	14.23 ± 0.05	14.22 ± 0.05

Table 12.1 that the B^0 and B^0_s decays have the same acceptance within the statistical errors. Decays into final states where the remaining phase-space is larger do not lie in the acceptance as often.

12.1.2 Reconstruction, selection and trigger efficiency

We now characterise the full reconstruction, selection and trigger efficiency as determined from simulated data. The simulation produces N events in total for a given Monte Carlo sample, the full processing proceeds under the following order: reconstruction, stripping and finally the optimised selection. The statistical uncertainly on the selection efficiencies is taken to be binomial, $\sigma(\epsilon) = \sqrt{\epsilon(1-\epsilon)/N}$, where N is the number of generated events, $\epsilon = m/N$ is the efficiency of the cut whereby m is the number of events that survived. Tables 12.2, 12.4 and 12.6 highlight the integrated efficiencies as applied to simulated B^0 decays for $B^0 \to J/\psi K^0_S \pi^+\pi^-$, $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0 \to J/\psi K^0_S K^+ K^-$ respectively. Likewise the B^0_s mode efficiencies are summarised in Tables 12.3, 12.5 and 12.7 for each of $B^0_s \to J/\psi K^0_S \pi^+\pi^-$,

$B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ and $B_s^0 \to J/\psi K_S^0 K^+ K^-$ mode respectively. The B^0 and B_s^0 efficiencies show good agreement across the board.

Table 12.2: Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K_S^0 \pi^+\pi^-$. (left) $B^0 \to J/\psi K_S^0(\text{LL})\pi^+\pi^-$ and (right) $B^0 \to J/\psi K_S^0(\text{DD})\pi^+\pi^-$. All uncertainties are statistical only.

	$B^0 \to J/\psi K_S^0(\text{LL})\pi^+\pi^-$ (%)		$B^0 \to J/\psi K_S^0(\text{DD})\pi^+\pi^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	2991(0.60)	2980(0.60)	7179(1.44)	7221(1.44)
Λ_b^0 veto	2650(89)	2661(89)	5900(82)	5809(80)
Λ veto	2650(100)	2661(100)	5900(100)	5809(100)
$\pi\pi$ as K_S^0 veto	2510(95)	2534(95)	5583(94)	5555(96)
$\psi(2S)$ veto	2505(99.8)	2529(99.8)	5575(99.7)	5541(99.8)
π^+ π^- momentum+isMuon	2311(90)	2344(92)	5126(92)	5119(90.9)
Selection	1056(45)	1025(44)	2407(47)	2421(47)
L0 decision	953(93)	954(93)	2235(93)	2256(93)
HLT1 decision	914(96)	926(97)	2154(96)	2185(97)
HLT2 decision	870(95)	885(96)	2039(95)	2062(94)
$\epsilon^{\text{Selection.}}$	0.174	0.180	0.408	0.412
σ_ϵ (stat.)	$\pm$0.007	$\pm$0.007	$\pm$0.026	$\pm$0.026

12.1.3 Particle identification efficiency

We use a data driven method that utilises the decay $D^{\star\pm} \to D^0(K^-\pi^+)\pi^\pm$, which takes advantage of the large number of charm mesons produced at the LHCb leading to pure samples of pions and kaons with very high statistics. These samples can be used as calibration channels as the true DLL distributions are known. The PID efficiency is characterised using the same `PIDCalib` tool as discussed in Sec. 8.1.4. The numbers in Table 12.8 vary according to the particle reconstructed in the final state, the more kaons the more efficient the selection becomes. This due to a less constraining kaon requirement than that for the pion.

From table 12.9 we infer that cross-feed arising from a double misidentification can be ignored in full since the efficiency is well below 1 % and we can expect

Table 12.3: Phase-space integrated relative trigger, selection and reconstruction efficiencies for $B_s^0 \to J/\psi K_S^0 \pi^+\pi^-$ modes. Based on data optimised selections using the control channel $B^0 \to J/\psi K_S^0 \pi^+\pi^-$. All uncertainties are statistical only. (left) $B_s^0 \to J/\psi K_S^0(\mathrm{LL})\pi^+\pi^-$ and (right) $B_s^0 \to J/\psi K_S^0(\mathrm{DD})\pi^+\pi^-$.

	$B_s^0 \to J/\psi K_S^0(\mathrm{LL})\pi^+\pi^-$ (%)		$B_s^0 \to J/\psi K_S^0(\mathrm{DD})\pi^+\pi^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	3089(0.62)	3152(0.63)	7576(1.51)	7483(1.50)
Λ_b^0 veto	2841(91)	2926(92)	6442(86)	6251(83)
Λ veto	2841(100)	2926(100)	6440(100)	6251(100)
$\pi\pi$ as K_S^0 veto	2708(95)	2758(94)	6166(97)	6007(96)
$\psi(2S)$ veto	2707(100)	2757(99.9)	6152(99.8)	6000(99.9)
π^+ π^- momentum+isMuon	2515(91)	2743(94)	5821(94)	5656(94)
Selection	1101(39)	1155(40)	2754(45)	2688(46)
L0 decision	1008(91)	1049(92)	2540(92)	2481(92)
HLT1 decision	985(97)	1012(96)	2444(96)	2390(96)
HLT2 decision	923(94)	961(94)	2313(95)	2242(94)
$\epsilon^{\text{Selection.}}$	0.190	0.212	0.463	0.450
σ_ϵ (stat.)	$\pm$0.008	$\pm$0.008	$\pm$0.031	$\pm$0.030

less than one event to contribute to the final yield.

12.2 Systematic uncertainties

The various sources of systematic uncertainty we consider in addition to those in Part I of the analysis are:-

- Fit bias (Sec. 12.2.2);
- Fragmentation fraction (Sec. 12.2.7).
- Statistical uncertainty on $B^0 \to J/\psi K_S^0 \pi^+\pi^-$ yields (Sec. 12.2.8).

12.2.1 Choice of PDF shapes (fit model)

The systematic uncertainty is evaluated as in Part I, by changing the PDF shapes and then refitting the data distribution. This is done for both signal and background PDFs.

Table 12.4: Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (left) $B^0 \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$ and (right) $B^0 \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$. All uncertainties are statistical only.

	$B^0 \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$ (%)		$B^0 \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	2286(0.55)	2178(0.54)	5524(1.33)	5172(1.27)
Λ^0_b veto	1920(83)	1830(83)	4340(78)	3979(78)
Λ veto	1920(100)	1830(100)	4339(100)	3977(100)
$K^\pm$ $\pi^\pm$ momentum+isMuon	1791(92)	1705(92)	4173(93)	3806(92)
Selection	940(52)	867(51)	1968(47)	1841(46)
L0 decision	861(91)	765(92)	1819(93)	1712(93)
HLT1 decision	836(97)	771(97)	1766(97)	1635(96)
HLT2 decision	785(94)	729(94)	1677(94)	1517(93)
$\epsilon^{\text{Selection.}}$	0.190	0.180	0.404	0.392
σ_ϵ (stat.)	$\pm$0.008	$\pm$0.008	$\pm$0.025	$\pm$0.023

Signal PDF

The signal PDF is changed from a double CB function to a double Gaussian function and the difference in the fitted yield is assigned as the associated systematic uncertainty. In addition, all fixed parameters of the signal PDF are varied within their uncertainties, and corresponding uncertainties assigned. A further uncertainty arises from the assumption that the B^0_s signal PDF width is the same as the B^0 width – this is evaluated from fits with an alternative assumption that the widths scale according to the Q-value in the decay.

Background PDF

The background PDF is changed from a linear function to a second order polynomial. The differences in the fitted signal yields are assigned as the associated systematic uncertainties.

Table 12.5: Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (left) $B^0_s \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$ and (right) $B^0_s \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$. All uncertainties are statistical only.

	$B^0_s \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$ (%)		$B^0_s \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	2307(0.58)	2313(0.58)	5404(1.35)	5489(1.38)
Λ^0_b veto	2004(83)	2004(98)	4318(81)	4420(80)
Λ veto	2003(100)	2004(100)	4318(100)	4420(100)
$K^\pm$ $\pi^\pm$ momentum+isMuon	1863(91)	1863(92)	4012(93)	4020(91)
Selection	943(48)	940(48)	2041(48)	1809(46)
L0 decision	863(93)	874(94)	1904(93)	1682(93)
HLT1 decision	832(97)	821(97)	1844(96)	1631(97)
HLT2 decision	819(95)	796(95)	1739(94)	1533(94)
$\epsilon^{\text{Selection.}}$	0.205	0.200	0.422	0.391
σ_ϵ (stat.)	$\pm$0.009	$\pm$0.009	$\pm$0.029	$\pm$0.030

12.2.2 Fit bias

The procedure to assign systematic uncertainty due to fit bias was established prior to unblinding, and is described in Sec. 11.4 (see also Sec. 11.5.1). Specifically, the systematic uncertainty is the sum in quadrature of the uncertainty on the bias and half the correction. The inputs to this calculation are given in Table 11.7, and the systematic error arising is summarised in Table 12.11.

12.2.3 Efficiencies

There are several source of systematic uncertainties related to the relative efficiency determination. Many other, potentially large, uncertainties cancel, however, due to the similar topologies of the decay modes in the ratios of branching fractions that are being measured.

Table 12.6: Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample,) following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (left) $B^0 \to J/\psi K^0_S(\text{LL})K^+K^-$ and (right) $B^0 \to J/\psi K^0_S(\text{DD})K^+K^-$. All uncertainties are statistical only.

	$B^0 \to J/\psi K^0_S(\text{LL})K^+K^-$ (%)		$B^0 \to J/\psi K^0_S(\text{DD})K^+K^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	2104(0.52)	1939(0.47)	4793(1.18)	4741(1.15)
Λ^0_b veto	1631(77)	1522(78)	3449(72)	3386(71)
Λ veto	1630(100)	1522(100)	3449(100)	3382(100)
K^+ K^- momentum+isMuon	1482(90)	1407(92)	3186(90)	3119(92)
Selection	798(54)	724(51)	1803(56)	1756(55)
L0 decision	751(94)	678(93)	1720(95)	1631(93)
HLT1 decision	723(96)	683(97)	1666(97)	1571(96)
HLT2 decision	678(94)	647(95)	1574(94)	1491(95)
$\epsilon^{\text{Selection.}}$	0.167	0.158	0.390	0.362
σ_ϵ (stat.)	$\pm$0.007	$\pm$0.007	$\pm$0.024	$\pm$0.022

12.2.4 Particle identification

We utilise the exact same method for calculating PID efficiency related systematic uncertainties as outlined in Part I (Sec. 8.1.4) of the analysis. Table 12.12 shows a summary of the final systematic values obtained.

12.2.5 Efficiency variation over the phase space

As discussed for Part I of the analysis in Sec. 8.2.6, the phase space distributions of the signal channels can not exactly match those of the MC, and since the efficiency varies across the phase space it is possible that our MC-driven estimate of the efficiency is not correct. For modes with a significant signal (*i.e.* $B^0 \to J/\psi K^0_S \pi^+\pi^-$, $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ and $B^0 \to J/\psi K^0_S K^+K^-$), the size of this effect can be estimated by reweighing the simulation, as done in Part I. To evaluate the systematic uncertainties the MC is reweighted according to the observed structures, as discussed in Sec. 11.6 and in detail in Appendix B.2.[1]

[1] Appendix B.2 discusses the variation across the phase space of the generator level efficiency. Exactly the same method is used to obtain the uncertainty corresponding to the variation across

Table 12.7: Numbers of events surviving each cut, with values in parentheses being the efficiencies calculated relative to the preceding requirement (or, for the first row, relative to the total number of events in the sample), following the order of the table. The integrated efficiency, $\epsilon^{\text{Selection.}}$, and its uncertainty, σ_ϵ (statistical only) are also provided. This is the total efficiency of the reconstruction, trigger and selection cuts for $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode. Based on data optimised selections using the control channel $B^0 \to J/\psi K^0_S \pi^+\pi^-$. (left) $B^0_s \to J/\psi K^0_S(\text{LL})K^+K^-$ and (right) $B^0_s \to J/\psi K^0_S(\text{DD})K^+K^-$. All uncertainties are statistical only.

	$B^0_s \to J/\psi K^0_S(\text{LL})K^+K^-$ (%)		$B^0_s \to J/\psi K^0_S(\text{DD})K^+K^-$ (%)	
	MagDown	MagUp	MagDown	MagUp
Reconstruction-Stripping	2114(0.51)	2047(0.49)	5016(1.24)	5092(1.24)
Λ^0_b veto	1822(83)	1683(84)	3960(76)	3793(74)
Λ veto	1822(100)	1683(100)	3960(100)	3793(100)
K^+ K^- momentum+isMuon	1712(92)	1548(92)	3550(91)	3566(93)
Selection	845(49)	863(54)	1950(55)	1925(54)
L0 decision	779(92)	790(91)	1813(93)	1790(93)
HLT1 decision	745(96)	739(96)	1756(96)	1736(97)
HLT2 decision	709(95)	695(95)	1656(94)	1649(95)
$\epsilon^{\text{Selection.}}$	0.172	0.169	0.403	0.401
σ_ϵ (stat.)	$\pm$0.007	$\pm$0.007	$\pm$0.026	$\pm$0.028

For channels without a significant signal, we consider the range of efficiency variation across the phase space as an estimate of the associated uncertainty, following the procedure used in other LHCb analyses of three-body b decays [76, 77]. Due to the high dimensionality of the final state, we simply consider the largest variation among the one-dimensional projections of the efficiency with invariant mass combinations.

Table 12.13 summarises the systematic uncertainty due to efficiency variation across the phase space. As in Part I of the analysis (Sec. 8.2.6 and Appendix B), we give separate uncertainties for generator level and selection (due to the full reconstruction, stripping, trigger and offline selection requirements) effects. The uncertainties for the observed channels are relatively small, which reflects the fact that the distributions of events in data and in our simulated samples are not too different. The uncertainties for the unobserved modes are, on the other hand, quite large (though still small compared to the large statistical uncertainties for these channels), which reflects the fact that the efficiency variation over the phase space is not small.

the phase space of the selection efficiency.

Table 12.8: PID efficiencies for all modes obtained using the `PIDCalib` procedure. The ± 0.01 uncertainty due to MC statistics is neglected.

	K^0_S (LL) (%)		K^0_S (DD) (%)	
	MagDown	MagUp	MagDown	MagUp
$B^0 \to J/\psi K^0_S \pi^+\pi^-$	$74.4 \times (100.0 \pm 4.1)\%$	$72.2 \times (100.0 \pm 4.5)\%$	$73.5 \times (100.0 \pm 2.9)\%$	$71.6 \times (100.0 \pm 1.9)\%$
$B^0 \to J/\psi K^0_S K^\pm \pi^\mp$	$82.4 \times (100.0 \pm 4.4)\%$	$81.0 \times (100.0 \pm 5.0)\%$	$82.2 \times (100.0 \pm 3.7)\%$	$81.1 \times (100.0 \pm 3.3)\%$
$B^0 \to J/\psi K^0_S K^+ K^-$	$93.2 \times (100.0 \pm 1.1)\%$	$92.7 \times (100.0 \pm 1.2)\%$	$93.3 \times (100.0 \pm 1.3)\%$	$93.0 \times (100.0 \pm 0.7)\%$
$B^0_s \to J/\psi K^0_S \pi^+\pi^-$	$74.5 \times (100.0 \pm 3.5)\%$	$73.1 \times (100.0 \pm 2.9)\%$	$73.1 \times (100.0 \pm 3.4)\%$	$73.1 \times (100.0 \pm 2.7)\%$
$B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$	$82.7 \times (100.0 \pm 2.6)\%$	$81.1 \times (100.0 \pm 4.5)\%$	$82.5 \times (100.0 \pm 2.0)\%$	$81.4 \times (100.0 \pm 2.8)\%$
$B^0_s \to J/\psi K^0_S K^+ K^-$	$92.8 \times (100.0 \pm 1.5)\%$	$92.7 \times (100.0 \pm 1.0)\%$	$93.0 \times (100.0 \pm 0.6)\%$	$92.9 \times (100.0 \pm 0.8)\%$

12.2.6 Simulation sample size

Various components of our result rely upon the usage of simulated events. The number of events within that sample and available to use after cuts are thus a source of systematic uncertainty in the final measurement we present. These are presented in Table 12.14.

12.2.7 Fragmentation fraction

As shown in Eq. (12.1), the determination of the branching fraction for B^0_s decay modes requires knowledge of the ratio of fragmentation fractions f_s/f_d. We take the latest value, $f_s/f_d = 0.259 \pm 0.015$ [81], which has 6 % uncertainty.

12.2.8 Statistical uncertainty on $B^0 \to J/\psi K^0_S \pi^+\pi^-$ yields

The systematic uncertainties have so far been given separately for each final state. However, the measurements that will result are of branching fractions determined relative to that of $B^0 \to J/\psi K^0_S \pi^+\pi^-$. Therefore the statistical uncertainty on the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ yields also enters the results. We separate this source of uncertainty (6.9 % for LL and 5.1 % for DD) from the signal channel statistics and from systematic uncertainties to allow a more straightforward evaluation of the absolute branching fractions (see Sec. 14).

12.3 Summary of systematic uncertainties

In Table 12.15 the systematic uncertainties are summarised. The total systematic uncertainty is the sum in quadrature of all contributions.

Table 12.16 translates the values from Table 12.15 into the uncertainties that are relevant for the measurements of the branching fraction relative to that for

Table 12.9: Mis-identification rates for all modes. The first number is the rate, second the statistical uncertainty and third is the systematic uncertainty due to the `PIDCalib` procedure.

	K^0_S (LL) (%)		K^0_S (DD) (%)	
	MagDown	MagUp	MagDown	MagUp
Truth $B^0 \to J/\psi K^0_S \pi^+\pi^-$				
as $J/\psi K^0_S K\pi$	$7.2 \pm 0.005 \pm 5.2$	$8.8 \pm 0.008 \pm 6.2$	$8.0 \pm 0.004 \pm 5.1$	$8.4 \pm 0.005 \pm 6.2$
as $J/\psi K^0_S KK$	$0.3 \pm 0.006 \pm 0.3$	$0.2 \pm 0.006 \pm 0.3$	$0.5 \pm 0.012 \pm 1.5$	$1.5 \pm 0.002 \pm 1.5$
Truth $B^0 \to J/\psi K^0_S K^\pm\pi^\mp$				
as $J/\psi K^0_S \pi\pi$	$2.3 \pm 0.007 \pm 2.3$	$2.7 \pm 0.008 \pm 1.3$	$2.4 \pm 0.005 \pm 2.6$	$2.8 \pm 0.006 \pm 2.8$
as $J/\psi K^0_S KK$	$7.9 \pm 0.0001 \pm 6.1$	$12.4 \pm 0.0001 \pm 7.6$	$8.9 \pm 0.008 \pm 5.9$	$9.1 \pm 0.010 \pm 6.6$
Truth $B^0 \to J/\psi K^0_S K^+K^-$				
as $J/\psi K^0_S \pi\pi$	$0.02 \pm 0.001 \pm 0.02$	$0.02 \pm 0.001 \pm 0.02$	$0.04 \pm 0.003 \pm 0.04$	$0.03 \pm 0.002 \pm 0.03$
as $J/\psi K^0_S K\pi$	$2.1 \pm 0.007 \pm 1.3$	$2.6 \pm 0.009 \pm 1.3$	$2.1 \pm 0.004 \pm 0.9$	$2.2 \pm 0.005 \pm 0.4$
Truth $B^0_s \to J/\psi K^0_S \pi^+\pi^-$				
as $J/\psi K^0_S K\pi$	$7.6 \pm 0.012 \pm 3.8$	$7.4 \pm 0.012 \pm 4.9$	$6.9 \pm 0.007 \pm 4.7$	$7.0 \pm 0.008 \pm 4.9$
as $J/\psi K^0_S KK$	0.0	0.0	0.0	0.0
Truth $B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$				
as $J/\psi K^0_S \pi\pi$	$1.6 \pm 0.005 \pm 0.8$	$1.5 \pm 0.006 \pm 0.4$	$1.5 \pm 0.004 \pm 0.1$	$1.6 \pm 0.004 \pm 1.6$
as $J/\psi K^0_S KK$	$11.5 \pm 0.054 \pm 8.0$	$7.5 \pm 0.035 \pm 6.9$	$5.8 \pm 0.018 \pm 4.6$	$7.7 \pm 0.025 \pm 5.4$
Truth $B^0_s \to J/\psi K^0_S K^+K^-$				
as $J/\psi K^0_S \pi\pi$	$0.01 \pm 0.001 \pm 0.01$	$0.02 \pm 0.002 \pm 0.02$	$0.02 \pm 0.001 \pm 0.02$	$0.03 \pm 0.007 \pm 0.03$
as $J/\psi K^0_S K\pi$	$1.7 \pm 0.003 \pm 0.7$	$1.7 \pm 0.003 \pm 0.8$	$1.7 \pm 0.002 \pm 0.2$	$1.6 \pm 0.002 \pm 0.1$

$B^0 \to J/\psi K^0_S \pi^+\pi^-$. The fit systematics are treated as 100 % correlated between signal and normalisation channels (*i.e.* the change in the normalisation yield cancels part of the change in the signal yield), as are the pion PID systematics but the kaon PID systematics are uncorrelated, and therefore the contributions for signal and normalisation channel are combined in quadrature. The uncertainties due to fit bias, simulation sample statistics and efficiency variation over the Dalitz plot are also treated as uncorrelated (for the latter, this is possibly a conservative treatment).

Table 12.10: Relative systematic uncertainties (in %) due to the change of functional form of the signal PDF, varying fixed parameters within their uncertainty, varying the B_s^0 width in accordance with the relative difference between the B^0-B_s^0 Q-values, and changing the functional form of the background PDF. The final column is the sum in quadrature of all these contributions.

Mode	Signal function	Fixed parameters	B_s^0 width (Q-value)	Background PDF	Total $\bar{\Delta}^{\text{Fit.}}$
$B^0 \to J/\psi K_S^0(\text{LL})\pi^+\pi^-$	±3.6	±0.4	±0.2	±0.0	±3.6
$B^0 \to J/\psi K_S^0(\text{DD})\pi^+\pi^-$	±5.1	±0.9	±0.5	±0.0	±5.2
$B^0 \to J/\psi K_S^0(\text{LL})K^\pm\pi^\mp$	±12.5	±0.0	±0.1	±2.5	±12.8
$B^0 \to J/\psi K_S^0(\text{DD})K^\pm\pi^\mp$	±7.0	±0.0	±0.3	±2.5	±7.4
$B^0 \to J/\psi K_S^0(\text{LL})K^+K^-$	±0.5	±2.7	±0.2	±1.7	±3.2
$B^0 \to J/\psi K_S^0(\text{DD})K^+K^-$	±4.1	±3.7	±0.1	±1.5	±5.7
$B_s^0 \to J/\psi K_S^0(\text{LL})\pi^+\pi^-$	±8.0	±15.0	±0.6	±10.0	±19.7
$B_s^0 \to J/\psi K_S^0(\text{DD})\pi^+\pi^-$	±20.0	±10.0	±1.5	±3.7	±22.7
$B_s^0 \to J/\psi K_S^0(\text{LL})K^\pm\pi^\mp$	±3.9	±0.7	±1.2	±0.5	±4.2
$B_s^0 \to J/\psi K_S^0(\text{DD})K^\pm\pi^\mp$	±4.3	±0.8	±1.3	±0.1	±4.6
$B_s^0 \to J/\psi K_S^0(\text{LL})K^+K^-$	±30.0	±0.0	±11.1	±18.4	±37.0
$B_s^0 \to J/\psi K_S^0(\text{DD})K^+K^-$	±3.3	±0.0	±4.0	±3.7	±6.4

Table 12.11: Summary of the systematic error induced by the fit model bias for all modes separated into K_S^0(LL) and K_S^0(DD) categories.

	K_S^0(DD) (%)	K_S^0(LL) (%)
$B^0 \to J/\psi K_S^0\pi^+\pi^-$	±0.4	±0.5
$B^0 \to J/\psi K_S^0K^\pm\pi^\mp$	±7.3	±9.1
$B^0 \to J/\psi K_S^0K^+K^-$	±3.2	±2.9
$B_s^0 \to J/\psi K_S^0\pi^+\pi^-$	±4.8	±5.6
$B_s^0 \to J/\psi K_S^0K^\pm\pi^\mp$	±0.6	±0.7
$B_s^0 \to J/\psi K_S^0K^+K^-$	±18.0	±22.0

Table 12.12: Summary of the PID systematics for all modes obtained using the `PIDCalib` procedure separated into K^0_S(LL) and K^0_S(DD) categories.

	K^0_S(DD) (%)	K^0_S(LL) (%)
$B^0 \to J/\psi\, K^0_S \pi^+\pi^-$	± 2.9	± 4.5
$B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$	± 3.7	± 5.0
$B^0 \to J/\psi\, K^0_S K^+ K^-$	± 1.3	± 1.2
$B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$	± 3.4	± 3.5
$B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$	± 2.8	± 4.5
$B^0_s \to J/\psi\, K^0_S K^+ K^-$	± 0.8	± 1.5

Table 12.13: Systematic uncertainty due to efficiency variation over the phase space. The column ϵ^{gen} is due to the generator level efficiency, and the column $\epsilon^{\text{Selection}}$ is due to the selection efficiency.

Mode	ϵ^{gen} (%)	$\epsilon^{\text{Selection}}$ (%)	Total (%)
$B^0 \to J/\psi\, K^0_S$(LL)$\pi^+\pi^-$	0.3	2.7	2.7
$B^0 \to J/\psi\, K^0_S$(DD)$\pi^+\pi^-$	0.3	2.6	2.6
$B^0 \to J/\psi\, K^0_S$(LL)$K^\pm\pi^\mp$	9.0	28.5	30.0
$B^0 \to J/\psi\, K^0_S$(DD)$K^\pm\pi^\mp$	9.0	25.0	27.0
$B^0 \to J/\psi\, K^0_S$(LL)K^+K^-	0.3	3.0	3.0
$B^0 \to J/\psi\, K^0_S$(DD)K^+K^-	0.3	3.7	3.7
$B^0_s \to J/\psi\, K^0_S$(LL)$\pi^+\pi^-$	16.0	29.0	33.0
$B^0_s \to J/\psi\, K^0_S$(DD)$\pi^+\pi^-$	16.0	25.0	30.0
$B^0_s \to J/\psi\, K^0_S$(LL)$K^\pm\pi^\mp$	0.2	3.0	3.0
$B^0_s \to J/\psi\, K^0_S$(DD)$K^\pm\pi^\mp$	0.2	3.9	3.9
$B^0_s \to J/\psi\, K^0_S$(LL)K^+K^-	5.0	32.0	32.4
$B^0_s \to J/\psi\, K^0_S$(DD)K^+K^-	5.0	36.0	36.3

Table 12.14: Summary of the systematics induced by the limitation in the simulation sample size, separated into K^0_S(LL) and K^0_S(DD) categories.

	K^0_S(DD) (%)	K^0_S(LL) (%)
$B^0 \to J/\psi\, K^0_S \pi^+\pi^-$	±1.0	±1.2
$B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$	±1.4	±1.8
$B^0 \to J/\psi\, K^0_S K^+ K^-$	±1.7	±2.0
$B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$	±1.4	±1.8
$B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$	±1.1	±1.2
$B^0_s \to J/\psi\, K^0_S K^+ K^-$	±1.9	±2.1

Table 12.15: Summary of the systematic uncertainties for each mode. Not included in the table are the uncertainties due to the fragmentation fraction.The total systematic uncertainty is the sum in quadrature of all contributions.

	$\bar{\Delta}^{\text{Fit+Bias.}}$ (%)	$\bar{\Delta}^{\text{PID.}}$ (%)	$\bar{\Delta}^{\text{MC}}$ (%)	$\bar{\Delta}^{\text{Eff. var.}}$ (%)	$\bar{\Delta}^{\text{Tot.}}$ (%)
$B^0 \to J/\psi\, K^0_S$(LL)$\pi^+\pi^-$	±3.6	±4.5	±1.2	±2.7	±6.5
$B^0 \to J/\psi\, K^0_S$(DD)$\pi^+\pi^-$	±5.2	±2.9	±1.0	±2.6	±6.6
$B^0 \to J/\psi\, K^0_S$(LL)$K^\pm\pi^\mp$	±15.7	±6.7	±1.8	±30.1	±34.7
$B^0 \to J/\psi\, K^0_S$(DD)$K^\pm\pi^\mp$	±10.1	±4.7	±1.4	±27.0	±29.2
$B^0 \to J/\psi\, K^0_S$(LL)K^+K^-	±4.3	±4.7	±2.0	±3.0	±7.3
$B^0 \to J/\psi\, K^0_S$(DD)K^+K^-	±6.9	±3.2	±1.7	±3.7	±8.6
$B^0_s \to J/\psi\, K^0_S$(LL)$\pi^+\pi^-$	±20.5	±5.7	±1.8	±33.0	±39.3
$B^0_s \to J/\psi\, K^0_S$(DD)$\pi^+\pi^-$	±23.2	±4.5	±1.4	±30.0	±38.2
$B^0_s \to J/\psi\, K^0_S$(LL)$K^\pm\pi^\mp$	±4.3	±6.4	±1.2	±3.0	±8.4
$B^0_s \to J/\psi\, K^0_S$(DD)$K^\pm\pi^\mp$	±4.6	±4.0	±1.1	±3.9	±7.3
$B^0_s \to J/\psi\, K^0_S$(LL)K^+K^-	±43.0	±4.7	±2.1	±32.4	±54.1
$B^0_s \to J/\psi\, K^0_S$(DD)K^+K^-	±20.0	±3.0	±1.9	±36.3	±41.6

Table 12.16: Summary of the systematic uncertainties that effect the measurement of the branching fraction relative to that of $B^0 \to J/\psi K^0_S \pi^+\pi^-$ for each mode. Not included in the table are the uncertainties due to the fragmentation fraction. The total systematic uncertainty is the sum in quadrature of all contributions.

	$\bar{\Delta}^{\text{Fit.}}$ (%)	$\bar{\Delta}^{\text{Bias.}}$ (%)	$\bar{\Delta}^{\text{PID.}}$ (%)	$\bar{\Delta}^{\text{MC}}$ (%)	$\bar{\Delta}^{\text{Eff. var.}}$ (%)	$\bar{\Delta}^{\text{Tot.}}$ (%)
$B^0 \to J/\psi K^0_S(\text{LL}) K^\pm \pi^\mp$	±8.9	±9.1	±6.7	±2.2	±30.2	±33.5
$B^0 \to J/\psi K^0_S(\text{DD}) K^\pm \pi^\mp$	±2.1	±7.3	±4.7	±1.7	±27.1	±28.6
$B^0 \to J/\psi K^0_S(\text{LL}) K^+ K^-$	±0.4	±2.9	±6.5	±2.3	±4.0	±8.5
$B^0 \to J/\psi K^0_S(\text{DD}) K^+ K^-$	±0.5	±3.2	±4.3	±2.0	±4.5	±7.3
$B^0_s \to J/\psi K^0_S(\text{LL}) \pi^+ \pi^-$	±15.5	±5.6	0.0	±2.2	±33.1	±37.0
$B^0_s \to J/\psi K^0_S(\text{DD}) \pi^+ \pi^-$	±16.6	±4.8	0.0	±1.7	±30.1	±34.8
$B^0_s \to J/\psi K^0_S(\text{LL}) K^\pm \pi^\mp$	±0.6	±0.9	±6.4	±1.7	±4.0	±7.8
$B^0_s \to J/\psi K^0_S(\text{DD}) K^\pm \pi^\mp$	±0.6	±0.7	±4.0	±1.6	±4.7	±6.4
$B^0_s \to J/\psi K^0_S(\text{LL}) K^+ K^-$	±32.2	±22.0	±6.5	±2.4	±32.5	±51.2
$B^0_s \to J/\psi K^0_S(\text{DD}) K^+ K^-$	±1.2	±18.0	±4.2	±2.2	±36.4	±40.9

13

Relative branching fraction measurements

We use Equation 12.1 to calculate each of the five relative branching fraction measurements. As was done in Part I, we take weighted efficiencies to take into account the different magnet polarities. Table 13.1 summarises the efficiencies for the K^0_S(DD) and K^0_S(LL) categories to be used in turn to convert the negative log likelihood curves as function of yield to a function of branching fraction. The relevant systematic uncertainties are provided in Table 12.16 with the exception of $(f_s/f_d) = 0.259 \pm 0.015$ which applies only to the B^0_s measurements. The statistical error on the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode is 5.0 % and 6.6 % for K^0_S(DD) and K^0_S(LL) respectively.

13.1 Results

We first obtain relative branching ratios for K^0_S(DD) and K^0_S(LL) final state categories separately. Since only small asymmetries are seen in the (statistical) uncertainties on the yields (Table 13.1), for the sake of simplicity we symmetrise the uncertainties. All systematic uncertainties are also treated as symmetric.

We then combine the results taking into account any correlated systematic uncertainties. For channels where the signal is not significant, we report upper limits in addition to central values and uncertainties. To determine the DD + LL combined significance of the signal, we follow a procedure used in previous LHCb measurements [76, 77]: [1]

[1] Note that the method to determine the combined significance does include the effect of any asymmetry in the likelihood function.

Table 13.1: Relevant efficiencies, fragmentation fractions and yields entering in the relative branching fractions measurements for $B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}$ decays for Down-Down and Long-Long K^0_S categories.

Mode	$\epsilon^{\text{Generator}}(\%)$	$\epsilon^{\text{Selection}}(\%)$	$\epsilon^{\text{PIDCalib}}(\%)$	Corrected yield
$B^0 \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$	11.07 ± 0.03	0.176 ± 0.007	73.5 ± 0.01	$245\,^{+17}_{-16}$
$B^0 \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$		0.410 ± 0.026	72.8 ± 0.01	$469\,^{+24}_{-23}$
$B^0 \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$	12.17 ± 0.05	0.187 ± 0.008	81.9 ± 0.01	$3\,^{+6}_{-5}$
$B^0 \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$		0.400 ± 0.024	81.8 ± 0.01	21 ± 10
$B^0 \to J/\psi K^0_S(\text{LL})K^+K^-$	14.21 ± 0.05	0.164 ± 0.007	93.0 ± 0.01	$18\,^{+5}_{-4}$
$B^0 \to J/\psi K^0_S(\text{DD})K^+K^-$		0.381 ± 0.024	93.2 ± 0.01	$29\,^{+8}_{-7}$
$B^0_s \to J/\psi K^0_S(\text{LL})\pi^+\pi^-$	10.94 ± 0.03	0.200 ± 0.008	74.0 ± 0.01	$5\,^{+4}_{-3}$
$B^0_s \to J/\psi K^0_S(\text{DD})\pi^+\pi^-$		0.459 ± 0.030	73.1 ± 0.01	$9\,^{+6}_{-5}$
$B^0_s \to J/\psi K^0_S(\text{LL})K^\pm\pi^\mp$	12.28 ± 0.04	0.203 ± 0.009	82.2 ± 0.01	$155\,^{+15}_{-14}$
$B^0_s \to J/\psi K^0_S(\text{DD})K^\pm\pi^\mp$		0.412 ± 0.030	82.1 ± 0.01	367 ± 23
$B^0_s \to J/\psi K^0_S(\text{LL})K^+K^-$	14.23 ± 0.05	0.171 ± 0.007	92.8 ± 0.01	$1\,^{+3}_{-2}$
$B^0_s \to J/\psi K^0_S(\text{DD})K^+K^-$		0.403 ± 0.027	93.0 ± 0.01	$2\,^{+5}_{-4}$

- Obtain the profile negative log likelihood by fitting the dataset many times a specific yield fixed to a range of value (referred to as "scanning" for a fit parameter). Scans are performed in each of the K^0_S categories.
- Plot the change in the negative log likelihoods for the different values of the scan, $\Delta \ln \mathcal{L} = -\ln \mathcal{L}_{\text{Max}} + \ln \mathcal{L}(N_{\text{sig}})$, where $\mathcal{L}_{\text{Max}}$ is the likelihood value obtained from the nominal fit where N_{sig} should be optimal, and $\mathcal{L}(N_{\text{sig}})$ is the signal yield when fixed to a set value.
- Transform the negative log likelihood to a likelihood function.
- Convolve the (statistical) likelihood function with a Gaussian ($\mathcal{N}(0, \sigma_{\text{systematic}})$) where the width is given by the systematic uncertainties that affect the yield only (*i.e.* the fit model and bias systematic uncertainties).
- Convert back to negative log likelihood functions and add the K^0_S categories together. This can be interpreted as a χ^2 for two degrees of freedom and the corresponding p-value calculated.

The scans are presented as a function of branching fraction values in Figures 13.1, 13.3, 13.4, 13.6 and 13.7 for $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$, $B^0 \to J/\psi\, K^0_S K^+ K^-$, $B^0_s \to J/\psi\, K^0_S \pi^+ \pi^-$, $B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$ and $B^0_s \to J/\psi\, K^0_S K^+ K^-$, respectively.

The combined central value is also taken from the scans. With this value in hand, we determine the various sources of uncertainty using a variation of the Best Linear Unbiased Estimate (BLUE) method [98, 99], as previously used in other LHCb analyses [77]. From the branching ratio value obtained by the convoluted log-likelihoods, we determine the weight to be used in a linear combination of the input uncertainties. Explicitly, we write

$$\bar{\mathcal{B}} = \alpha \mathcal{B}_{\rm DD} + (1-\alpha)\mathcal{B}_{\rm LL}\,, \tag{13.1}$$

$$\longrightarrow \alpha = \left(\bar{\mathcal{B}} - \mathcal{B}_{\rm LL}\right) / \left(\mathcal{B}_{\rm DD} - \mathcal{B}_{\rm LL}\right)\,. \tag{13.2}$$

With this value of α we can then combine each set of uncertainties using trivial error propagation. The exception is for the PID and fit model systematics, which are treated (conservatively) as being 100 % correlation between K^0_S categories.

The procedure to set upper limits, for channels where the combined signal is not significant, is the same strategy used in several previous LHCb measurements [100, 76].

- Quote the central values for the branching fractions $\pm 1\sigma$ in any case (even for negative central values);
- Obtain the significance from the change in twice the negative log-likelihood values obtained with and without each of the signal components in the fit in turn (*i.e.* using Wilks' theorem [101]), incorporating systematic uncertainties;
- Quote both the 90 % and 95 % confidence level upper limits obtained from integrating the likelihood in the positive region in case the significance is $< 3\sigma$. Note that this corresponds to a Bayesian treatment, where the prior is taken to be uniform in the physical region of non-negative branching fraction.

We convolve the combined statistical likelihood curve with a Gaussian to account for systematic uncertainties (taking correlations between the systematic uncertainties for DD and LL into account). We then obtain the 90 % and 95 % confidence level upper limits by integrating the likelihood in the region of positive branching fraction, *i.e.* we find the values that satisfy

$$\int_0^{\mathcal{U}_{90\,(95)}} \mathcal{L}\, d\mathcal{B} = 0.90\,(0.95) \times \int_0^{\infty} \mathcal{L}\, d\mathcal{B}\,. \tag{13.3}$$

This is a Bayesian treatment, where the prior is taken to be uniform in the physical region of non-negative branching fraction (*i.e.* a step function).

13.2 $B^0 \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp$ branching fraction

The likelihood scans are shown in Figure 13.1. The relative branching fraction values are

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm LL}) K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm LL})\pi^+\pi^-)} = 0.011 \pm 0.016\ ({\rm stat.}) \pm 0.004\ ({\rm syst.}) \pm 0.001\ (\pi^+\pi^-{\rm stat.})\,, \tag{13.4}$$

and

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm DD}) K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm DD})\pi^+\pi^-)} = 0.040 \pm 0.018\ ({\rm stat.}) \pm 0.012\ ({\rm syst.}) \pm 0.002\ (\pi^+\pi^-{\rm stat.})\,, \tag{13.5}$$

for the $K^0_{\rm S}({\rm LL})$ and $K^0_{\rm S}({\rm DD})$ modes repectively. The combined significance of these modes is 2.4σ. The combined branching fraction is

$$\begin{aligned} \frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)} &= 0.026 \pm 0.012\ ({\rm stat.}) \pm 0.007\ ({\rm syst.}) \pm 0.001\ (\pi^+\pi^-{\rm stat.})\,, \\ &< 0.048 \text{ at } 90\,\%\ {\rm CL}\,, \\ &< 0.055 \text{ at } 95\,\%\ {\rm CL}\,, \end{aligned} \tag{13.6}$$

where the posterior likelihood is illustrated in Fig. 13.2.

13.2.1 $B^0 \to J/\psi\, K^0_{\rm S} K^+ K^-$ branching fraction

The likelihood scans are shown in Figure 13.3. The relative branching fraction values are

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm LL}) K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm LL})\pi^+\pi^-)} = 0.051 \pm 0.015\ ({\rm stat.}) \pm 0.005\ ({\rm syst.}) \pm 0.003\ (\pi^+\pi^-{\rm stat.})\,, \tag{13.7}$$

corresponding to statistical error of 29 % and a systematic error of 9 % for the $K^0_{\rm S}({\rm LL})$ mode, and

$$\frac{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm DD}) K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}({\rm DD})\pi^+\pi^-)} = 0.041 \pm 0.010\ ({\rm stat.}) \pm 0.003\ ({\rm syst.}) \pm 0.002\ (\pi^+\pi^-{\rm stat.})\,, \tag{13.8}$$

corresponding to statistical error of 25 % and a systematic error of 7 % for the $K^0_{\rm S}({\rm DD})$ mode. The combined significance of these modes is 7.7σ. The combined

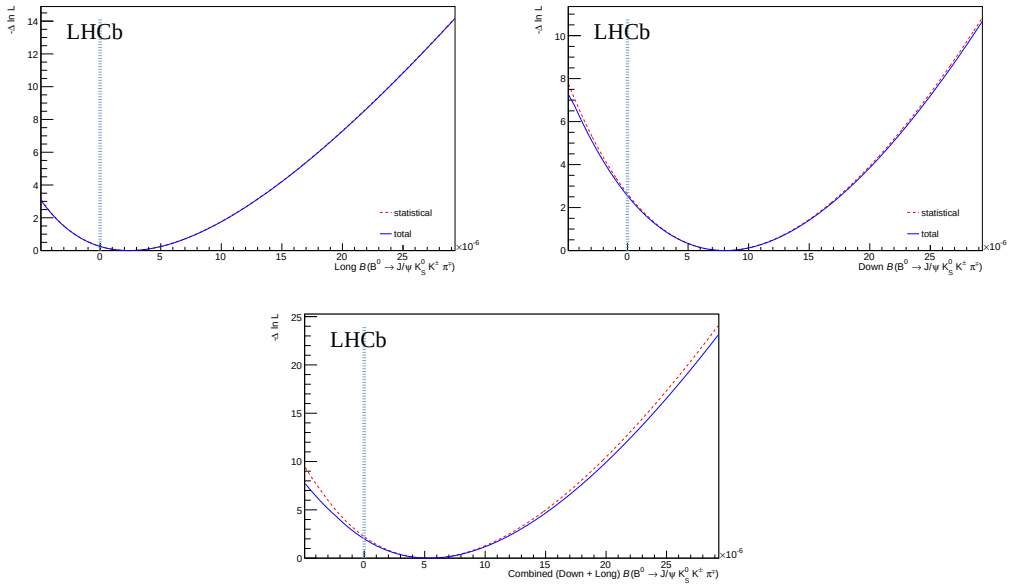

Figure 13.1: Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0 \to J/\psi K^0_S(\mathrm{LL})K^\pm\pi^\mp$ and (top right) $B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included.

branching fraction is

$$\frac{\mathcal{B}(B^0 \to J/\psi K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} = 0.047 \pm 0.010\ (\mathrm{stat.}) \pm 0.004\ (\mathrm{syst.}) \pm 0.002\ (\pi^+\pi^-\mathrm{stat.})\,, \tag{13.9}$$

13.2.2 $B^0_s \to J/\psi K^0_S \pi^+\pi^-$ branching fraction

The likelihood scans are shown in Figure 13.4. The relative branching fraction values are

$$\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)} = 0.054 \pm 0.030\ (\mathrm{stat.}) \pm 0.020\ (\mathrm{syst.}) \pm 0.003\ (f_s/f_d) \pm 0.004\ (\pi^+\pi^-\mathrm{stat.})\,, \tag{13.10}$$

and

$$\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)} = 0.052 \pm 0.035\ (\mathrm{stat.}) \pm 0.018\ (\mathrm{syst.}) \pm 0.003\ (f_s/f_d) \pm 0.003\ (\pi^+\pi^-\mathrm{stat.})\,, \tag{13.11}$$

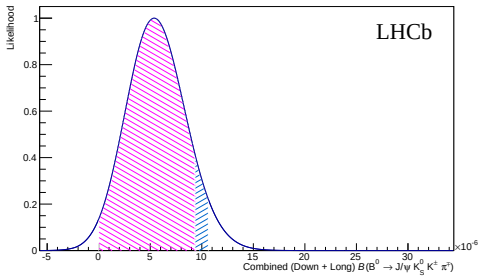

Figure 13.2: Posterior likelihood for $B^0 \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %.

for the $K^0_{\rm S}$(LL) and $K^0_{\rm S}$(DD) modes repectively. The combined significance of these modes is 2.7σ. The combined branching fraction is

$$\begin{aligned}\frac{\mathcal{B}(B^0_s \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)} &= 0.054 \pm 0.031\ (\text{stat.}) \pm 0.020\ (\text{syst.}) \pm 0.003\ (f_s/f_d) \pm 0.004\ (\pi^+\pi^-\text{stat.})\,, \\ &< 0.103 \text{ at } 90\,\%\ \text{CL}\,, \\ &< 0.115 \text{ at } 95\,\%\ \text{CL}\,,\end{aligned} \tag{13.12}$$

where the posterior likelihood is illustrated in Fig. 13.5.

13.2.3 $B^0_s \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp$ branching fraction

The likelihood scans are shown in Figure 13.6. The relative branching fraction values are

$$\frac{\mathcal{B}(B^0_s \to J/\psi\, K^0_{\rm S}(\text{LL}) K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}(\text{LL})\pi^+\pi^-)} = 1.70 \pm 0.17\ (\text{stat.}) \pm 0.14\ (\text{syst.}) \pm 0.10\ (f_s/f_d) \pm 0.11\ (\pi^+\pi^-\text{stat.})\,, \tag{13.13}$$

corresponding to statistical error of 10 % and a systematic error of 8 %, and

$$\frac{\mathcal{B}(B^0_s \to J/\psi\, K^0_{\rm S}(\text{DD}) K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}(\text{DD})\pi^+\pi^-)} = 2.42 \pm 0.23\ (\text{stat.}) \pm 0.19\ (\text{syst.}) \pm 0.14\ (f_s/f_d) \pm 0.12\ (\pi^+\pi^-\text{stat.})\,, \tag{13.14}$$

corresponding to statistical error of 9 % and a systematic error of 6 %. The combined significance is 30.0σ. The combined branching fraction is

$$\frac{\mathcal{B}(B^0_s \to J/\psi\, K^0_{\rm S} K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi\, K^0_{\rm S}\pi^+\pi^-)} = 2.12 \pm 0.15\ (\text{stat.}) \pm 0.14\ (\text{syst.}) \pm 0.08\ (f_s/f_d) \pm 0.08\ (\pi^+\pi^-\text{stat.})\,, \tag{13.15}$$

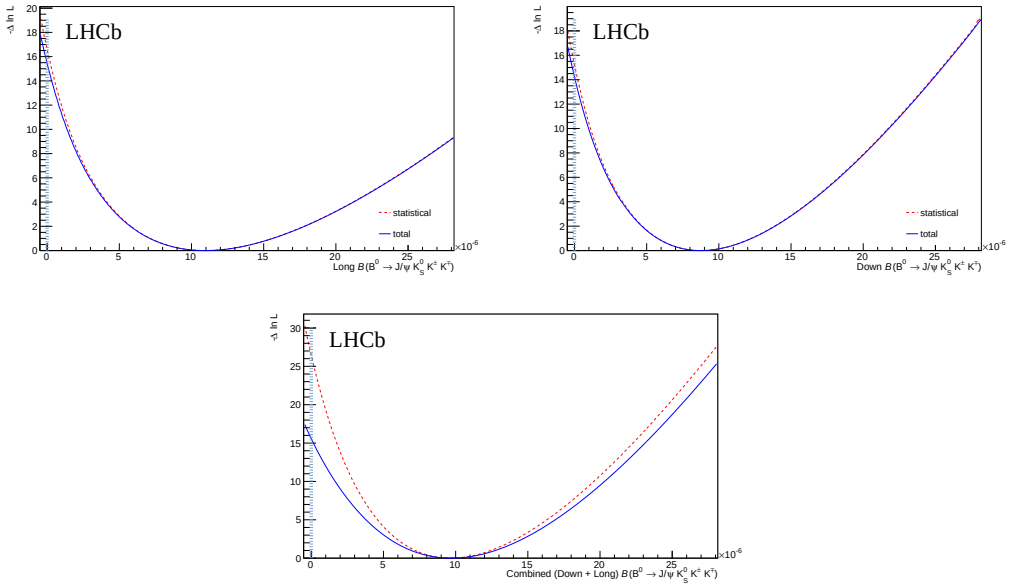

Figure 13.3: Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B^0 \to J/\psi K^0_S(\mathrm{LL})K^+K^-$ and (top right) $B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included.

13.2.4 $B^0_s \to J/\psi K^0_S K^+ K^-$ branching fraction

The likelihood scans are shown in Figure 13.7. The relative branching fraction values are

$$\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S(\mathrm{LL})K^+K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S(\mathrm{LL})\pi^+\pi^-)} = 0.012 \pm 0.031\ (\mathrm{stat.}) \pm 0.006\ (\mathrm{syst.}) \pm 0.001\ (f_s/f_d) \pm 0.001\ (\pi^+\pi^-\mathrm{stat.})\,, \tag{13.16}$$

and

$$\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S(\mathrm{DD})K^+K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-)} = 0.010 \pm 0.026\ (\mathrm{stat.}) \pm 0.004\ (\mathrm{syst.}) \pm 0.001\ (f_s/f_d) \pm 0.001\ (\pi^+\pi^-\mathrm{stat.})\,, \tag{13.17}$$

for the $K^0_S(\mathrm{LL})$ and $K^0_S(\mathrm{DD})$ modes repectively. The combined significance is 0.5σ. The combined branching fraction is

$$\begin{aligned} \frac{\mathcal{B}(B^0_s \to J/\psi K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 0.011 \pm 0.020\ (\mathrm{stat.}) \pm 0.006\ (\mathrm{syst.}) \pm 0.001\ (f_s/f_d) \pm 0.001\ (\pi^+\pi^-\mathrm{stat.})\,, \\ &< 0.027 \text{ at } 90\,\%\ \mathrm{CL}\,, \\ &< 0.033 \text{ at } 95\,\%\ \mathrm{CL}\,, \end{aligned} \tag{13.18}$$

where the posterior likelihood is illustrated in Fig. 13.8.

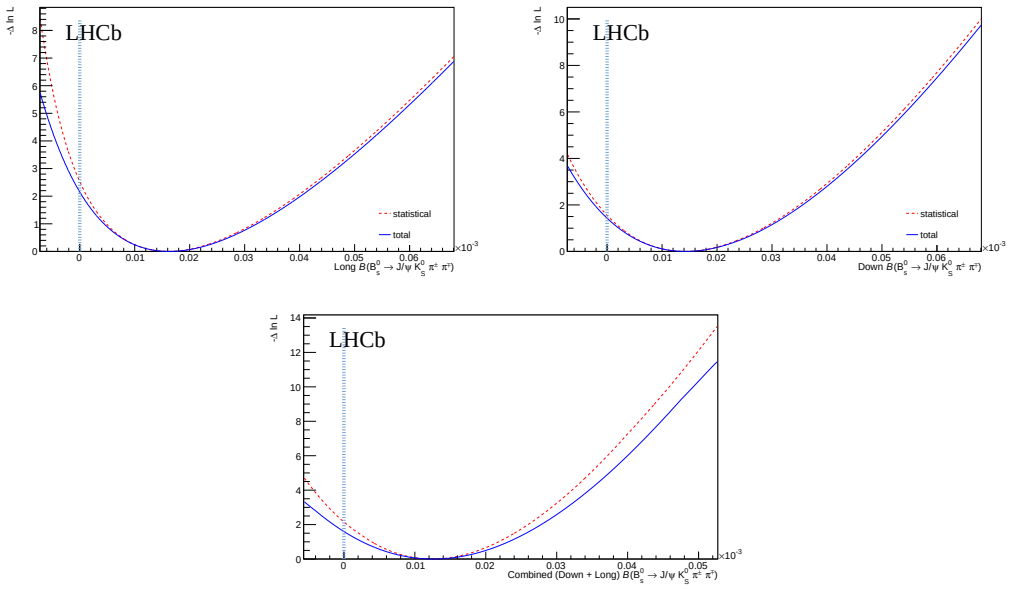

Figure 13.4: Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B_s^0 \to J/\psi K_S^0(\mathrm{LL})\pi^+\pi^-$ and (top right) $B_s^0 \to J/\psi K_S^0(\mathrm{DD})\pi^+\pi^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included.

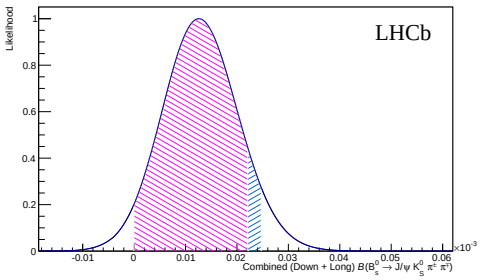

Figure 13.5: Profile likelihood for $B_s^0 \to J/\psi K_S^0\pi^+\pi^-$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %.

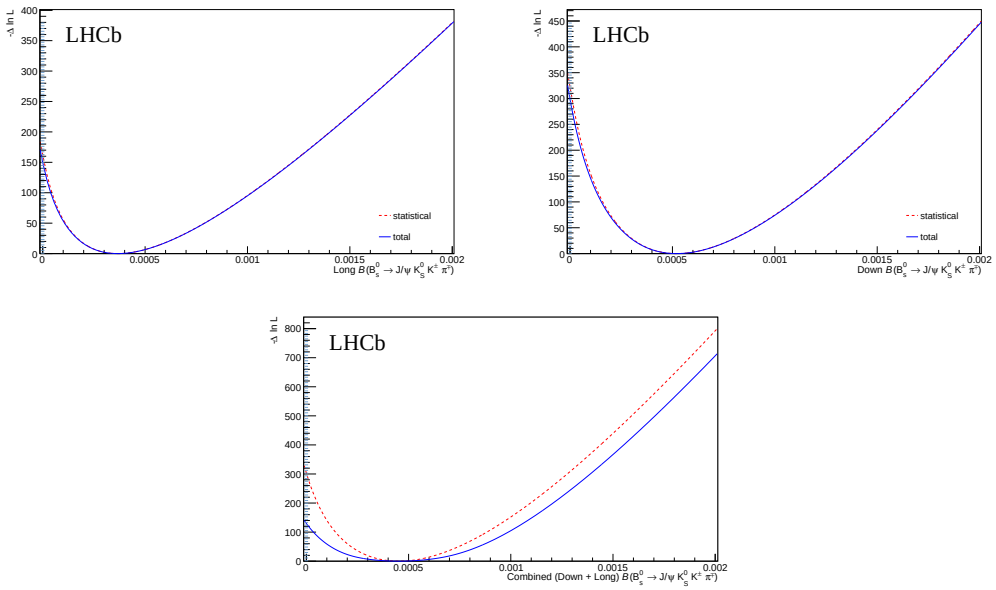

Figure 13.6: Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B_s^0 \to J/\psi K_S^0(\mathrm{LL})K^\pm\pi^\mp$ and (top right) $B_s^0 \to J/\psi K_S^0(\mathrm{DD})K^\pm\pi^\mp$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included.

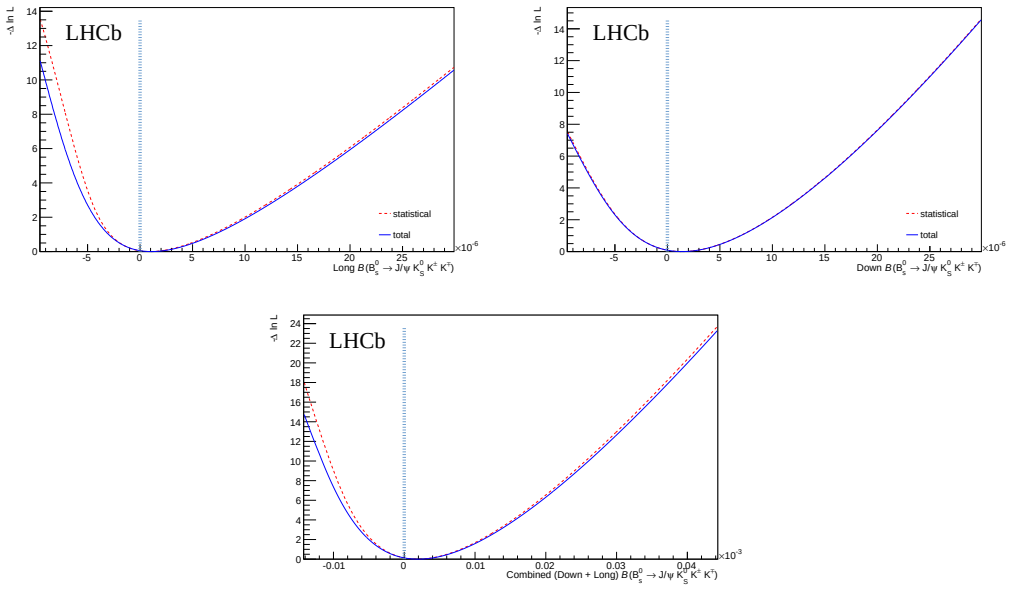

Figure 13.7: Likelihood profile using the negative log likelihood scan as a function of branching fraction for (top left) $B_s^0 \to J/\psi K_S^0(\mathrm{LL})K^+K^-$ and (top right) $B_s^0 \to J/\psi K_S^0(\mathrm{DD})K^+K^-$ modes. The bottom plot shows the combined likelihood curve, with both statistical and systematic uncertainties included.

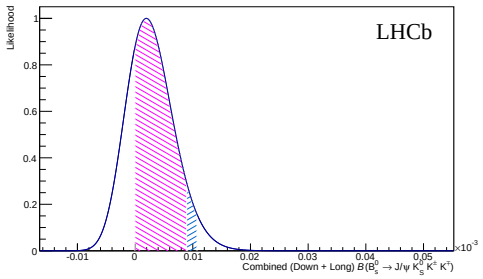

Figure 13.8: Posterior likelihood for $B_s^0 \to J/\psi K_S^0 K^+ K^-$. The red area covers 90 % of the likelihood in the region of positive branching fraction; the blue area takes this up to 95 %.

Part III

Summary

14

Summary and absolute branching fractions

Using the data sample collected in 2011 by the LHCb experiment corresponding to an integrated luminosity of $1.0\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 7\,\text{TeV}$, searches for the decay modes $B^0_{(s)} \to J/\psi\, K^0_S h^+ h^{(\prime)-}$ are performed. The process of a two-step analysis procedure has allowed the most precise measurement to date of the $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$ branching fraction, made with 4 times greater statistical precision than the previous world average:

$$\begin{aligned} \frac{\mathcal{B}(B^0 \to J/\psi\, K^0_S \pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi\, K^0_S)} &= 0.493 \pm 0.034\ (\text{stat.}) \pm 0.027\ (\text{syst.}) \\ \mathcal{B}(B^0 \to J/\psi\, K^0_S \pi^+\pi^-) &= (21.5 \pm 1.5\ (\text{stat.}) \pm 1.2\ (\text{syst.}) \pm 0.8\ (\text{PDG})) \times 10^{-5} \end{aligned}$$

A cross check using the $B^0 \to \psi(2S) K^0_S$ mode not only provides good agreement with a well known measurement but also supports the analysis method. A simultaneous fit to six spectra, $J/\psi\, K^0_S \pi^+\pi^-$, $J/\psi\, K^0_S K^+\pi^-$ and $J/\psi\, K^0_S K^+ K^-$ reconstructed with Down-Down and Long-Long K^0_S categories, has been performed. The first observations of the $B^0 \to J/\psi\, K^0_S K^+ K^-$ and $B^0_s \to J/\psi\, K^0_S K^\pm \pi^\mp$ decays have been reported and the first limits on the branching fractions of $B^0_s \to J/\psi\, K^0_S \pi^+\pi^-$, $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$ and $B^0_s \to J/\psi\, K^0_S K^+ K^-$ are set. The yields have been turned into branching fraction measurements relative to $B^0 \to J/\psi\, K^0_S \pi^+\pi^-$:

$$
\begin{aligned}
\frac{\mathcal{B}(B^0 \to J/\psi K^0_S K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 0.026 \pm 0.012\ (\text{stat.}) \pm 0.007\ (\text{syst.}) \pm 0.001\ (\pi^+\pi^-\text{stat.}) \\
&< 0.048 \text{ at } 90\,\% \text{ CL} \\
&< 0.055 \text{ at } 95\,\% \text{ CL} \\
\frac{\mathcal{B}(B^0 \to J/\psi K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 0.047 \pm 0.010\ (\text{stat.}) \pm 0.004\ (\text{syst.}) \pm 0.002\ (\pi^+\pi^-\text{stat.}) \\
\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S \pi^+\pi^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 0.054 \pm 0.031\ (\text{stat.}) \pm 0.020\ (\text{syst.}) \pm 0.003\ (f_s/f_d) \pm 0.004\ (\pi^+\pi^-\text{stat.}) \\
&< 0.103 \text{ at } 90\,\% \text{ CL} \\
&< 0.115 \text{ at } 95\,\% \text{ CL} \\
\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S K^\pm \pi^\mp)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 2.12 \pm 0.15\ (\text{stat.}) \pm 0.14\ (\text{syst.}) \pm 0.08\ (f_s/f_d) \pm 0.08\ (\pi^+\pi^-\text{stat.}) \\
\frac{\mathcal{B}(B^0_s \to J/\psi K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to J/\psi K^0_S \pi^+\pi^-)} &= 0.011 \pm 0.020\ (\text{stat.}) \pm 0.006\ (\text{syst.}) \pm 0.001\ (f_s/f_d) \pm 0.001\ (\pi^+\pi^-\text{stat.}) \\
&< 0.027 \text{ at } 90\,\% \text{ CL} \\
&< 0.033 \text{ at } 95\,\% \text{ CL}
\end{aligned}
$$

The $B^0 \to J/\psi K^0_S \pi^+\pi^-$ decay appears to have a large contribution from $B^0 \to J/\psi K_1(1400)$, but no clear signals for resonant structures in the $K^0_S K^\pm \pi^\mp$ ($K^0_S K^+ K^-$) system are seen for $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ ($B^0 \to J/\psi K^0_S K^+ K^-$) decays. Neither are exotic resonances seen.

By combining the results from Part I and Part II of the analysis, we also obtain results for the absolute branching fractions of the five $B^0_{(s)} \to J/\psi K^0_S h^+ h^{(\prime)-}$ modes other than $B^0 \to J/\psi K^0_S \pi^+\pi^-$ (already reported above). Here we treat the statistics of the $B^0 \to J/\psi K^0_S \pi^+\pi^-$ mode as being 100 % correlated between the measurements in Part I and Part II, neglecting small effects caused as a result of to the different selection requirements. We neglect the effect of systematic correlations between the results in Part I and Part II, since these uncertainties are small enough that the correlations have no impact on the result.

Finally, because the results for B decays in the PDG [15] are quoted with either K^0 or $\overline{K}^0$ (not K^0_S) in the final state, we divide by $\mathcal{B}(K^0 \to K^0_S) = \mathcal{B}(\overline{K}^0 \to$

$K^0_S) = \frac{1}{2}$ to put our results in the same format as shown below

$$
\begin{aligned}
&\mathcal{B}(B^0 \to J/\psi\, K^0 K^- \pi^+ + B^0 \to J/\psi\, \overline{K}^0 K^+ \pi^-) \\
&\qquad = (11 \pm 5\ (\text{stat.}) \pm 3\ (\text{syst.}) \pm 1\ (\text{PDG})) \times 10^{-6} \\
&\qquad < 21 \times 10^{-6} \text{ at } 90\,\%\ \text{CL} \\
&\qquad < 24 \times 10^{-6} \text{ at } 95\,\%\ \text{CL} \\
&\mathcal{B}(B^0 \to J/\psi\, K^0 K^+ K^-) = (20.2 \pm 4.3\ (\text{stat.}) \pm 1.7\ (\text{syst.}) \pm 0.8\ (\text{PDG})) \times 10^{-6} \\
&\mathcal{B}(B^0_s \to J/\psi\, \overline{K}^0 \pi^+ \pi^-) = (2.4 \pm 1.4\ (\text{stat.}) \pm 0.8\ (\text{syst.}) \pm 0.1\ (f_s/f_d) \pm 0.1\ (\text{PDG})) \times 10^{-5} \\
&\qquad < 4.4 \times 10^{-5} \text{ at } 90\,\%\ \text{CL} \\
&\qquad < 5.0 \times 10^{-5} \text{ at } 95\,\%\ \text{CL} \\
&\mathcal{B}(B^0_s \to J/\psi\, K^0 K^- \pi^+ + B^0_s \to J/\psi\, \overline{K}^0 K^+ \pi^-) \\
&\qquad = (91 \pm 6\ (\text{stat.}) \pm 6\ (\text{syst.}) \pm 3\ (f_s/f_d) \pm 3\ (\text{PDG})) \times 10^{-5} \\
&\mathcal{B}(B^0_s \to J/\psi\, \overline{K}^0 K^+ K^-) = (5 \pm 9\ (\text{stat.}) \pm 2\ (\text{syst.}) \pm 1\ (f_s/f_d)) \times 10^{-6} \\
&\qquad < 12 \times 10^{-6} \text{ at } 90\,\%\ \text{CL} \\
&\qquad < 14 \times 10^{-6} \text{ at } 95\,\%\ \text{CL}
\end{aligned}
$$

where the contribution from the PDG uncertainty to the last result is negligible.

Appendix A

Variable alias dictionary

Table. A.1 outlines the alias variable name used in the analysis and the approiate description of that variable. The general rule of thumb for this naming convention is; PARTICLE-CUT-RELATEDVERTEX with the exception of any vertex quailty cuts.

Table A.1: Dictionary of variable names used in the analysis and their definition.

Cut definition	Alias
K^0_S significance of separation wrt related PV, $\chi^2_{\rm VD}$	KS_VDCHI2_OWNPV
K^0_S significance of separation wrt B vertex, $\chi^2_{\rm SEP}$	KS_VTX_SEPCHI
K^0_S IP significance with respect to PV, $\chi^2_{\rm IP}$	KS_IPCHI2_OWNPV
K^0_S vertex quality (probability), $P(\chi^2, {\rm ndf})$	KS_ENDVERTEX_PROB
K^0_S vertex quality, $\chi^2/{\rm ndf}$	KS_ENDVERTEX_CHI2NDOF
B^0 vertex quaility (probability), $P(\chi^2, {\rm ndf})$	B_ENDVERTEX_PROB
B^0 vertex quality, $\chi^2/{\rm ndf}$	B_ENDVERTEX_CHI2NDOF
B^0 pointing angle, $\cos(\theta)$	B_DIRA_OWNPV
B^0 transverse momentum, $p_{\rm T}$,	B_PT
B^0 IP significance with respect to PV, $\chi^2_{\rm IP}$	B_IPCHI2_OWNPV
J/ψ significance of separation wrt related PV, $\chi^2_{\rm VD}$	Jpsi_VDCHI2_OWNPV
J/ψ IP significance with respect to PV, $\chi^2_{\rm IP}$	Jpsi_IPCHI2_OWNPV
J/ψ vertex quailty (probability), $P(\chi^2, {\rm ndf})$	Jpsi_ENDVERTEX_PROB
J/ψ vertex quality, $\chi^2/{\rm ndf}$	Jpsi_ENDVERTEX_CHI2NDOF
π minimum IP significance wrt related PV, $\min(\pi^+_{\chi^2_{\rm IP}}, \pi^-_{\chi^2_{\rm IP}})$	h1_MINIPCHI2, h2_MINIPCHI2

Appendix B

Generator level efficiencies

It is useful to check that the generator level cuts do not drastically bias the phase-space over the various invariant mass combinations. As mentioned in Part I of the analysis, Sec. 8.1.1, we now investigate the variation of the efficiency as a function of different invariant mass combinations. We can fully describe the kinematics of our four-body decay by a generic set of five invariant masses, we have chosen; $K^0_S\pi^+$, $K^0_S\pi^-$, $J/\psi\pi\pi$, $K^0_S\pi\pi$ and $\pi\pi$. In order to validate the uniformity of our generator level cuts we generated 1M events for the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode without any generator level cut applied. We then apply the `DaughtersInLHCb` cut which checks that all daughter particles lie in the LHCb acceptance ($0.01 \leq \theta \leq 0.4\,\text{rad}$), and the `DaughtersInLHCbWithMinP` which requires that the minimum momentum of daughter tracks is $p > 1600$ MeV/c or $p > 3000$ MeV/c for muon tracks. Figure B.1 shows the effect of the `DaughtersInLHCb` cut on the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode as a function of different invariant mass combinations. The plots show point-to-point variation at the 10 % level implying that this generator cut does not bias the phase-space too much.

Figure B.1 presents the efficiency of the `DaughtersInLHCbWithMinP` cut which is the one used in our analysis. The variation is slightly more prominent at the high and low mass edges in certain areas, in particular for the $m_{J/\psi\pi^+\pi^-}$ plot we see a drop in efficiency for low invariant mass.

B.1 Systematic due to generator level variation

A systematic can be deduced from the generator level information above since we know certain resonances are not modelled correctly in the simulation, this can be seen in Figures 7.9 and 7.10 which compare the invariant mass distributions for

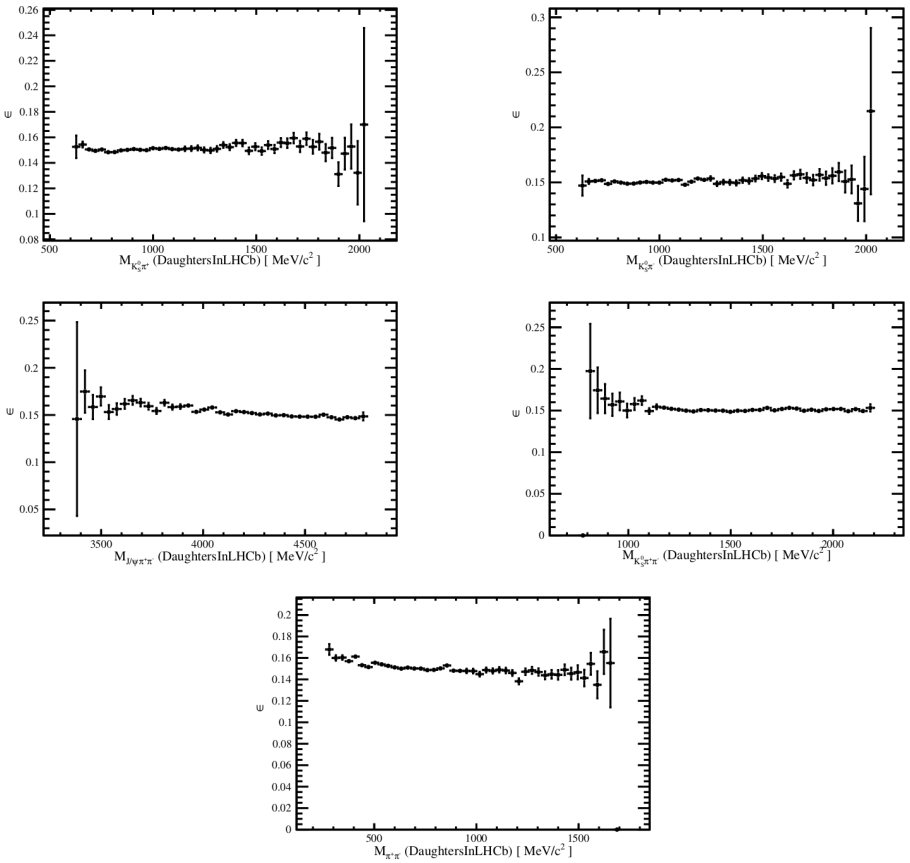

Figure B.1: Efficiency of `DaughtersInLHCb` cut across various invariant mass combinations. From left to right and top to bottom: $m(K^0_S\pi^+)$, $m(K^0_S\pi^-)$, $m(J/\psi\,\pi^+\pi^-)$, $m(K^0_S\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2).

background subtracted data and signal MC. We take the resonances with the largest discrepancy between data and MC, these are the $\rho^0(770)$, $K^0_1(1270)$ and $K^*(892)$ modes. The average efficiency can be obtained using Eq 8.1 for the average weighted efficiency where the N_i is the number of events and ϵ_i the efficiency for the i^{th} bin. We can obtain a resonance corrected efficiency by finding the bin in which the mass of the resonance falls into, then scaling that bin entry so that it matches what we have in data, and repeating the summation with this newly scaled bin. This allows us to validate just how important this region of the phase-space is to the overall efficiency. A systematic is obtained by looking at the percentage difference from the two procedures.

Table B.1 shows the efficiency variation[1]. The total systematic from this

[1]Note that the values in the table agree well with the central integrated efficiency given in Table 8.3 11.06 ± 0.03, providing confidence in the method.

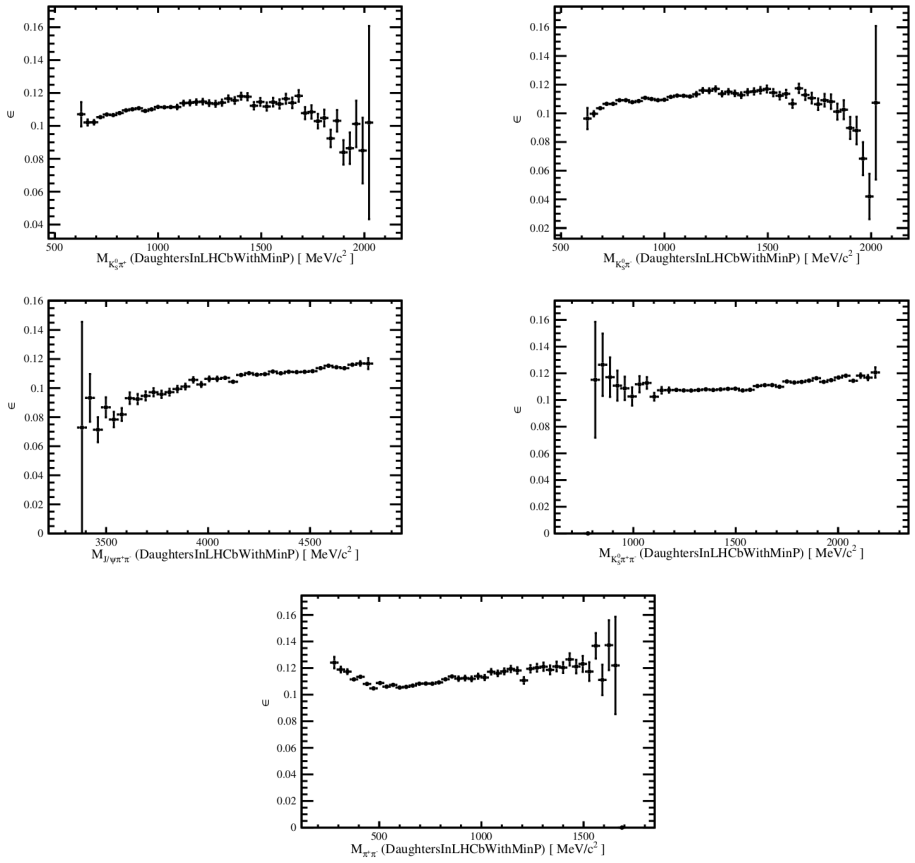

Figure B.2: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations. From left to right and top to bottom: $m(K^0_S\pi^+)$, $m(K^0_S\pi^-)$, $m(J/\psi\,\pi^+\pi^-)$, $m(K^0_S\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2).

contribution is then 0.3 % which is very small.

B.2 Generator level efficiencies for Part II modes

The same procedure outlined above can be applied to the modes observed in Part II of the analysis as mentioned in Sec. 12.2.5. For modes with a significant signal, the size of this effect can be estimated by reweighting the MC, as done in Part I. For channels without a significant signal, we consider the range of efficiency variation across the phase space as an estimate of the associated uncertainty, following the procedure used in other LHCb analyses of three-body b decays [76]. Due to the high dimensionality of the final state, we simply consider the largest variation among the one-dimensional projections of the efficiency with invariant mass combinations.

Since we can reduce the four-body decay kinematics to five invariant mass

Table B.1: Number of events used in weighting procedure for $B^0 \to J/\psi K^0_S \pi^+\pi^-$.

Resonance	Scale	Original ϵ (%)	Corrected ϵ (%)	Systematic (%)
$\rho^0(779) \to \pi^+\pi^-$	0.5	11.01 ± 0.11	11.02 ± 0.11	0.1
$K^0_1(1270) \to K^0_S\pi^+\pi^-$	0.5	11.01 ± 0.10	11.04 ± 0.11	0.23
$K^*(892) \to K^0_S\pi^+$	1.9	11.011 ± 0.110	11.014 ± 0.108	0.03
$K^*(892) \to K^0_S\pi^-$	1.9	11.011 ± 0.110	11.014 ± 0.108	0.03
Total				0.25

combinations and in keeping with Sec B, we choose the following combinations; $K^0_S h'^+$, $K^0_S h^-$, $J/\psi h'^+ h^-$, $K^0_S h'^+ h^-$ and $h'^+ h^-$. In the next two sections we proceed by first looking at the observed modes and the efficiency re-weighting, and secondly looking at the unobserved modes and the average variation uncertainty across the phase-space.

The generator level efficiencies as a function of different invariant mass combinations are shown in Figures B.3 and B.4 for the $B^0 \to J/\psi K^0_S K^+K^-$ and $B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$ modes respectively.

B.2.1 Observed modes and their efficiency corrections

There are three modes with significances greater than 5σ, $B^0 \to J/\psi K^0_S\pi^+\pi^-$, $B^0 \to J/\psi K^0_S K^+K^-$ and $B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$. Applying the correction to the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode provides no addition systematic over what was found in Part I of the analysis, therefore the same value is assigned. We now evaluate the efficiency when considering the resonances in the invariant mass distributions for the other two modes. The generator level efficiencies as a function of different invariant mass combinations are shown in Figures B.3 and B.4 for the $B^0 \to J/\psi K^0_S K^+K^-$ and $B^0_s \to J/\psi K^0_S K^\pm\pi^\mp$ modes respectively.

Using figure B.12 we see a clear peak at the ϕ resonance in the $\phi \to K^+K^-$ invariant mass. Due to the low statistics of these modes this is the only significant correction we apply. Table B.2 shows the value of the re-weighting to be small at 0.3 %.

Using same procedure by examining figure B.10 we see two clear $K^*(892)$ resonances; one for $K^*(892) \to K^0_S\pi^-$ and another for $K^*(892) \to K^+\pi^-$ invariant mass final states. These are the only clearly visible resonances and thus the only significant correction we apply. Table B.3 shows the value of the re-weighting to be

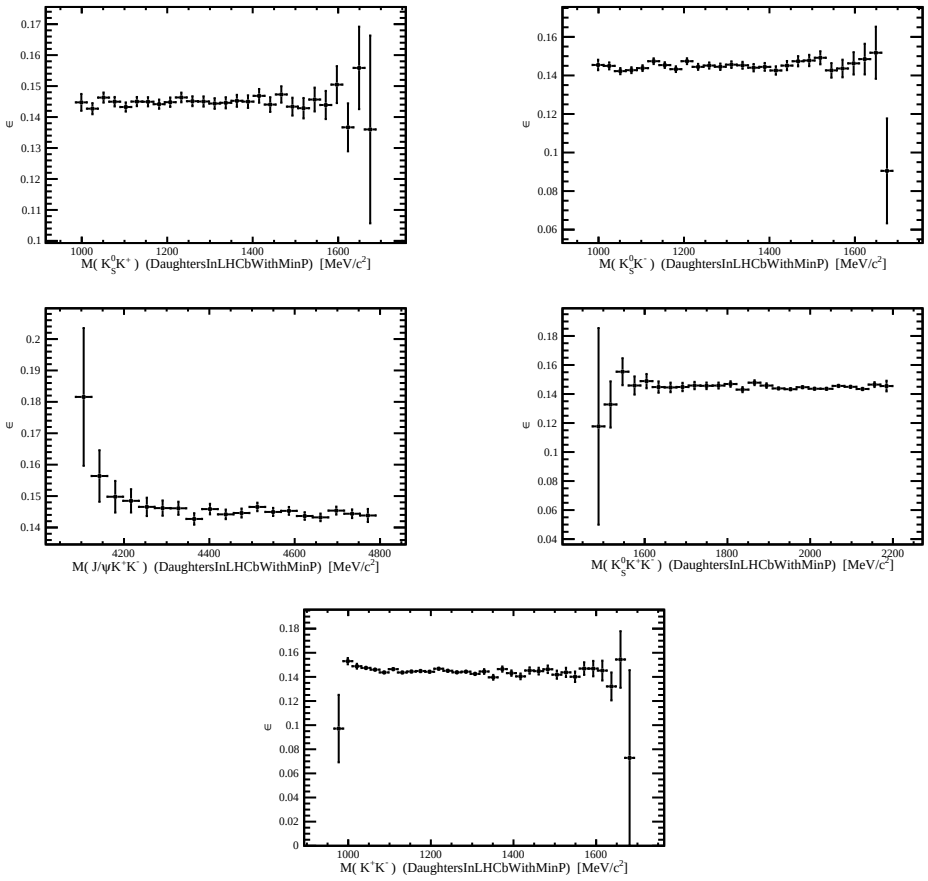

Figure B.3: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0 \to J/\psi\, K^0_S K^+ K^-$ mode. From left to right and top to bottom: $m(K^0_S K^+)$, $m(K^0_S K^-)$, $m(J/\psi\, K^+ K^-)$, $m(K^0_S K^+ K^-)$, $m(K^+ K^-)$ (in MeV/c^2).

small at 0.22 %.

B.2.2 Unobserved mode and associated uncertainties

There are three unobserved modes, $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$, $B^0_s \to J/\psi\, K^0_S \pi^+ \pi^-$ and $B^0_s \to J/\psi\, K^0_S K^+ K^-$. Since there is no way of knowing how well our simulation describes the data as we must make a conservative estimate of the systematic uncertainty due to efficiency variation over the phase space of our decay modes. The systematic error we assign comes directly by taking the standard deviation of the efficiency variation across the phase space, for each invariant mass combination, then assigning the maximal deviation as a conservative uncertainty.

Figures B.5, B.6 and B.7 provide the efficiency as a function of our chosen invariant mass combinations for the $B^0 \to J/\psi\, K^0_S K^\pm \pi^\mp$, $B^0_s \to J/\psi\, K^0_S \pi^+ \pi^-$ and

Table B.2: Number of events used in weighting procedure for $B^0 \to J/\psi K^0_S K^+ K^-$.

Resonance	Scale	Original ϵ (%)	Corrected ϵ (%)	Systematic (%)
$\phi \to K^+K^-$	3.0	14.49 ± 0.18	14.52 ± 0.19	0.3

Table B.3: Number of events used in weighting procedure for $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$.

Resonance	Scale	Original ϵ (%)	Corrected ϵ (%)	Systematic (%)
$K^*(892) \to K^0_S \pi^-$	2.5	12.55 ± 0.17	12.53 ± 0.17	0.2
$K^*(892) \to K^+ \pi^-$	2.7	12.54 ± 0.18	12.54 ± 0.18	0.1
Total				0.22

$B^0_s \to J/\psi K^0_S K^+ K^-$ modes respectively. The results of the standard deviation of the efficiency variations are provided in Tables B.4, B.5 and B.6 where we find a maximum uncertainty of 9 %, 16 % and 5 % for the $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$, $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$ and $B^0_s \to J/\psi K^0_S K^+ K^-$ modes respectively.

Table B.4: $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ mode invariant mass standard deviation of efficiency variation.

Combination	Systematic (%)
$K^\pm \pi^\mp$(MeV)	± 5.5
$K^0_S \pi^\mp$(MeV)	± 7.3
$K^0_S K^\pm$(MeV)	± 8.3
$J/\psi K^\pm \pi^\mp$(MeV)	± 9.0
$K^0_S K^\pm \pi^\mp$(MeV)	± 6.6

B.3 Observed invariant mass combinations

We now show observed distributions for invariant mass yields passing the selection criteria. Figures B.9 (for $B^0 \to J/\psi K^0_S \pi^+ \pi^-$), B.11 (for $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$) and B.13 (for $B^0 \to J/\psi K^0_S K^+ K^-$) compare the observed daughter invariant mass combinations in simulated events (in red) and data (black) for the K^0_S(DD) modes.

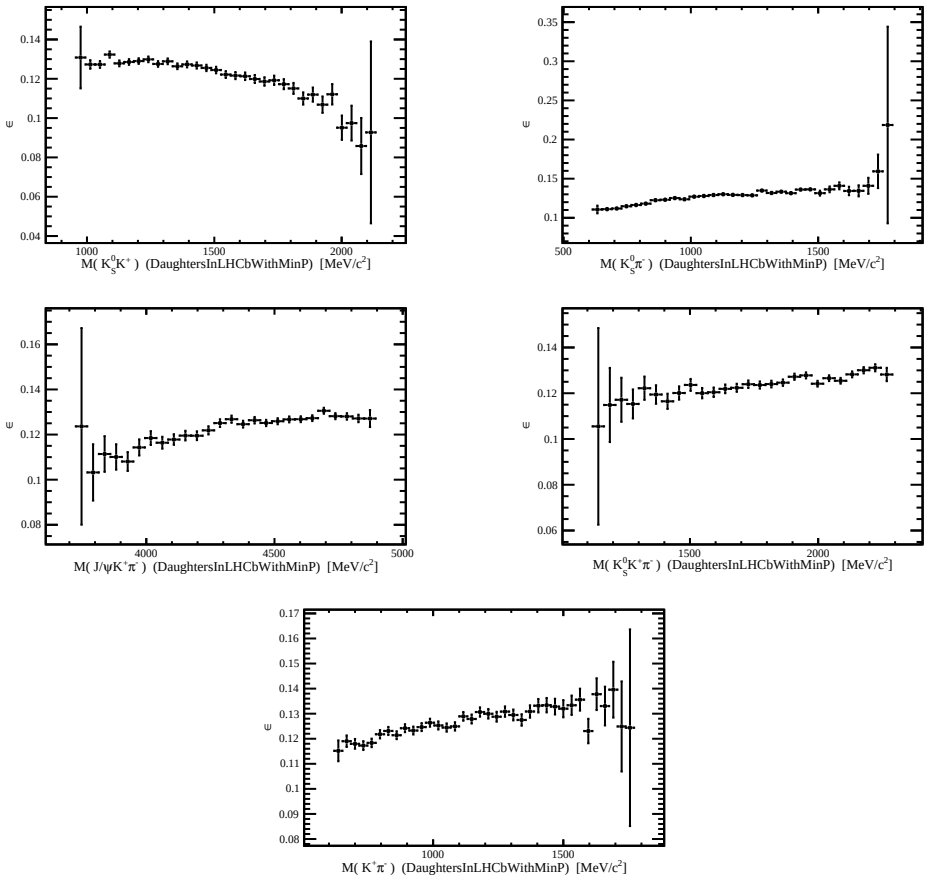

Figure B.4: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0_s \to J/\psi K^0_S K^\pm \pi^\mp$ mode. From left to right and top to bottom: $m(K^0_S K^+)$, $m(K^0_S \pi^-)$, $m(J/\psi K^+ \pi^-)$, $m(K^0_S K^+ \pi^-)$, $m(K^+ \pi^-)$ (in MeV/c^2).

The data distributions are the same as those shown in Sec. 11.6. The $B^0 \to J/\psi K^0_S \pi^+ \pi^-$ plots are very similar to those shown in Sec. 8.2.6 (small differences are possible due to the slight difference in selection criteria).

We simply give example plots for the three unobserved modes, $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$ (Fig. B.14), $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ (Fig. B.15) and $B^0_s \to J/\psi K^0_S K^+ K^-$ (Fig. B.16).

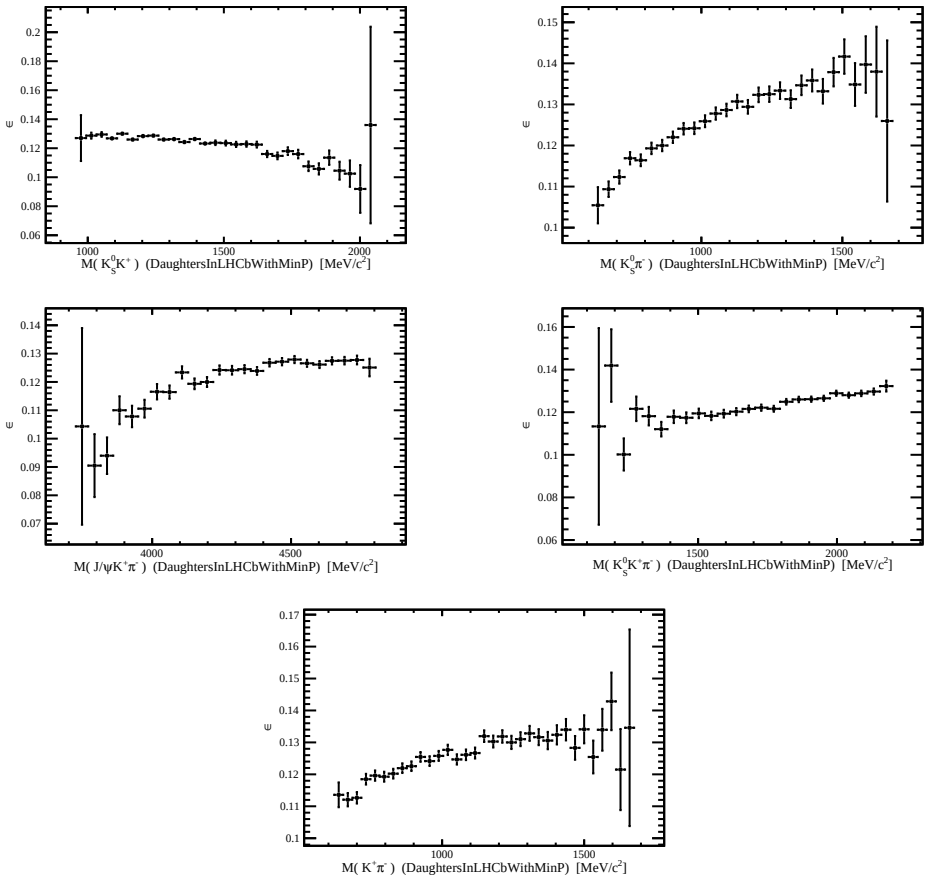

Figure B.5: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0 \to J/\psi K^0_S K^\pm \pi^\mp$ mode. From left to right and top to bottom: $m(K^0_S K^+)$, $m(K^0_S \pi^-)$, $m(J/\psi K^+ \pi^-)$, $m(K^0_S K^+ \pi^-)$, $m(K^+ \pi^-)$ (in MeV/c^2).

Table B.5: $B^0_s \to J/\psi K^0_S \pi^+ \pi^-$ mode invariant mass standard deviation of efficiency variation.

Combination	Systematic (%)
$\pi^+\pi^-$(MeV)	$\pm$6.8
$K^0_S\pi^-$(MeV)	$\pm$9.2
$K^0_S\pi^+$(MeV)	$\pm$7.8
$J/\psi K^+\pi^-$(MeV)	$\pm$16.0
$K^0_S K^+\pi^-$(MeV)	$\pm$5.2

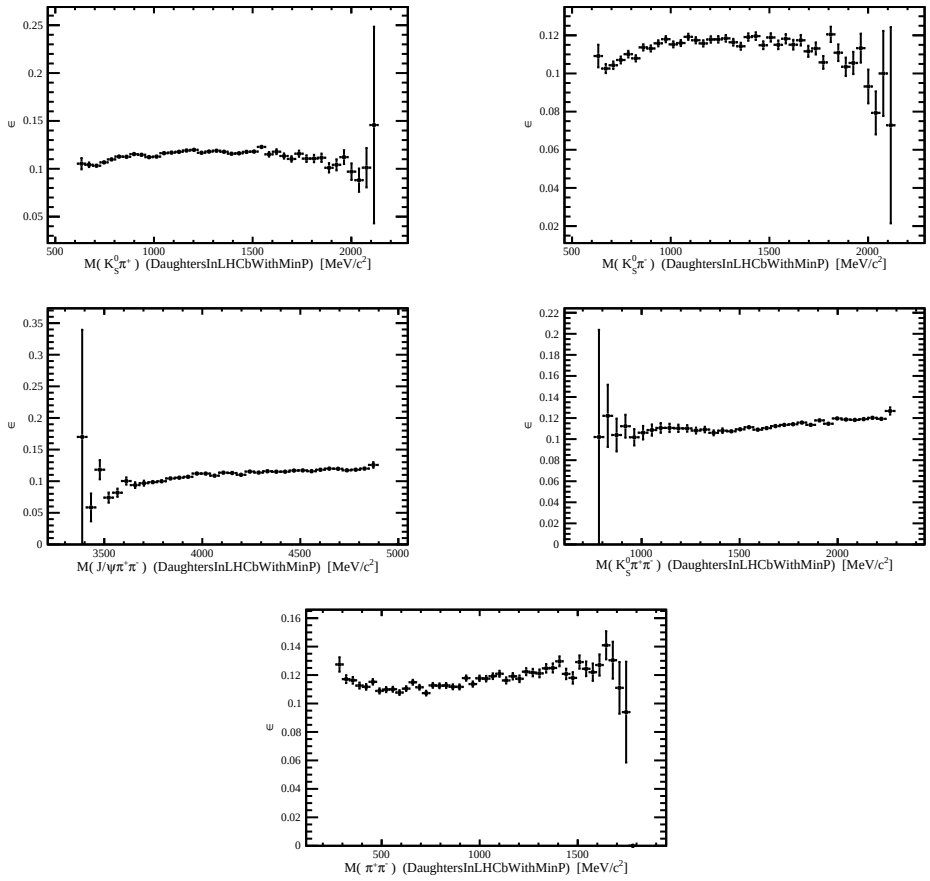

Figure B.6: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0_s \to J/\psi K^0_S\pi^+\pi^-$ mode. From left to right and top to bottom: $m(K^0_S\pi^+)$, $m(K^0_S\pi^-)$, $m(J/\psi\pi^+\pi^-)$, $m(K^0_S\pi^+\pi^-)$, $m(\pi^+\pi^-)$ (in MeV/c^2).

Table B.6: $B^0_s \to J/\psi K^0_S K^+K^-$ mode invariant mass standard deviation of efficiency variation.

Resonance	Systematic (%)
K^+K^-(MeV)	$\pm$3.9
$K^0_S K^-$(MeV)	$\pm$3.3
$K^0_S K^+$(MeV)	$\pm$4.9
$J/\psi K^+K^-$(MeV)	$\pm$2.6
$K^0_S K^+K^-$(MeV)	$\pm$2.0

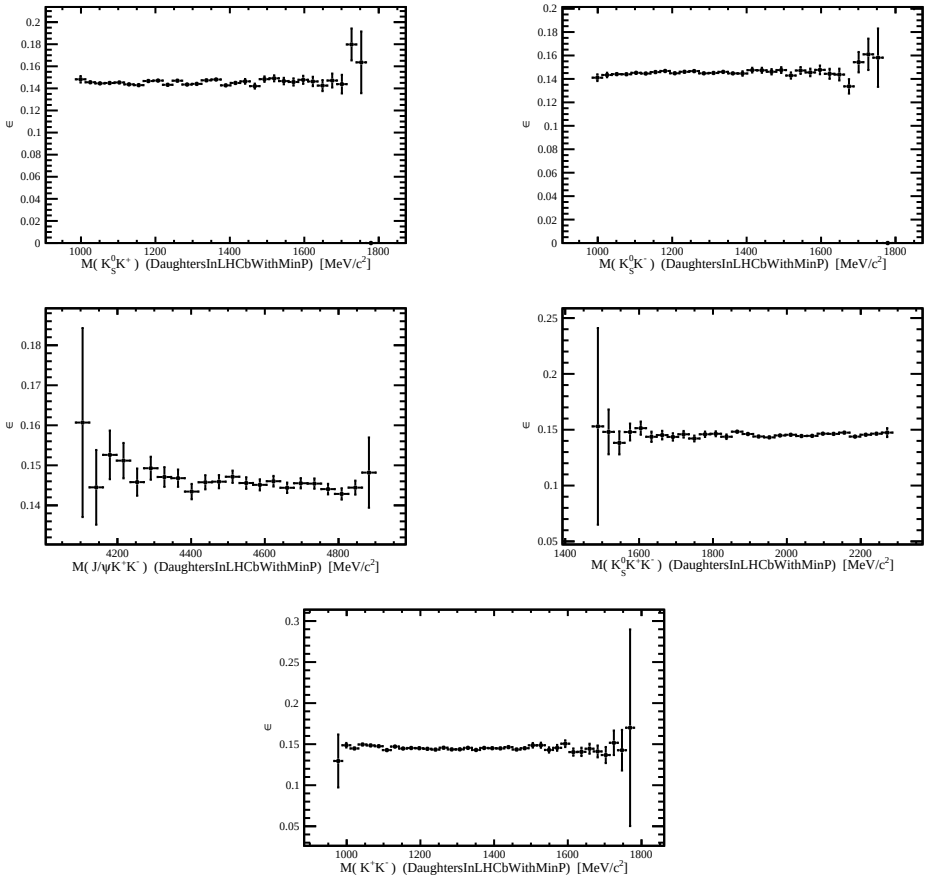

Figure B.7: Efficiency of `DaughtersInLHCbWithMinP` cut across various invariant mass combinations for the $B^0_s \to J/\psi\, K^0_S K^+ K^-$ mode. From left to right and top to bottom: $m(K^0_S K^+)$, $m(K^0_S K^-)$, $m(J/\psi\, K^+ K^-)$, $m(K^0_S K^+ K^-)$, $m(K^+ K^-)$ (in MeV/c^2).

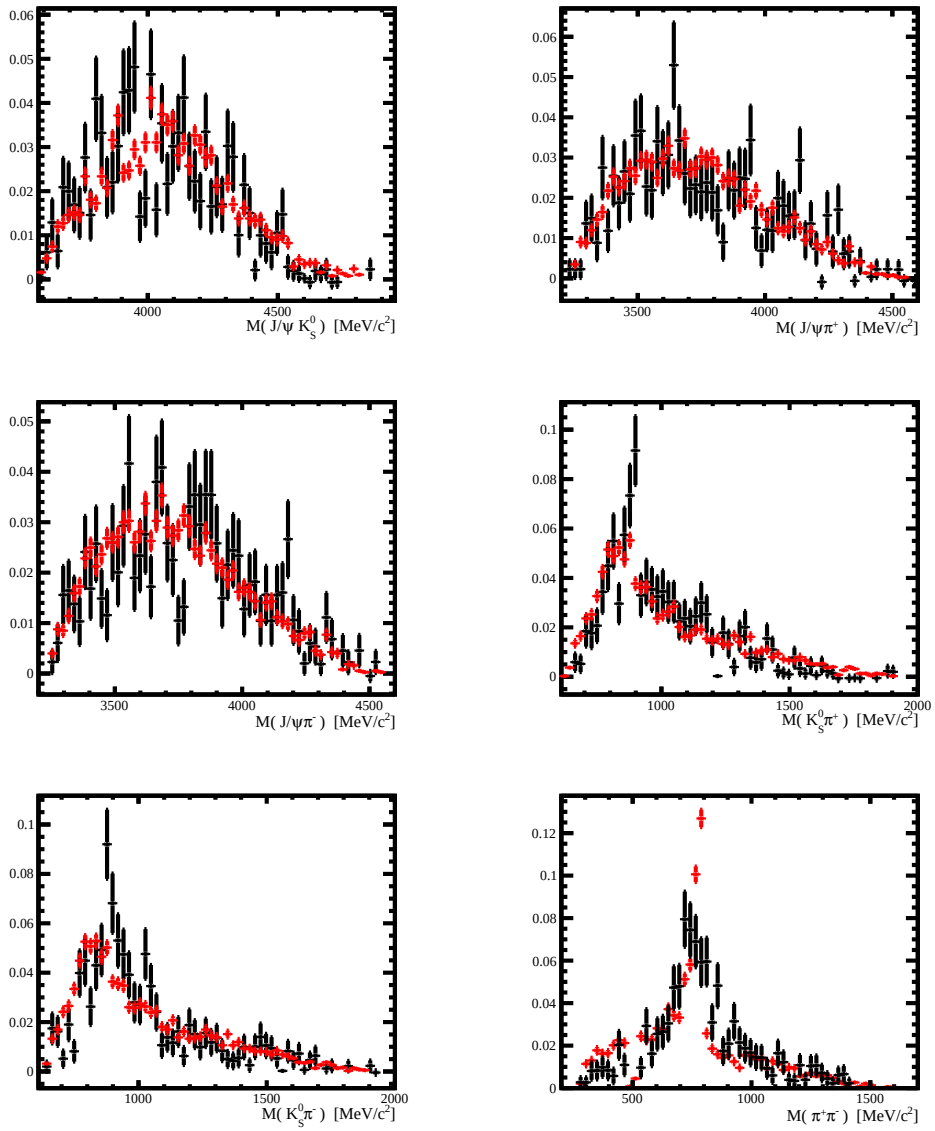

Figure B.8: $B^0 \to J/\psi K^0_S(\mathrm{DD})\pi^+\pi^-$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

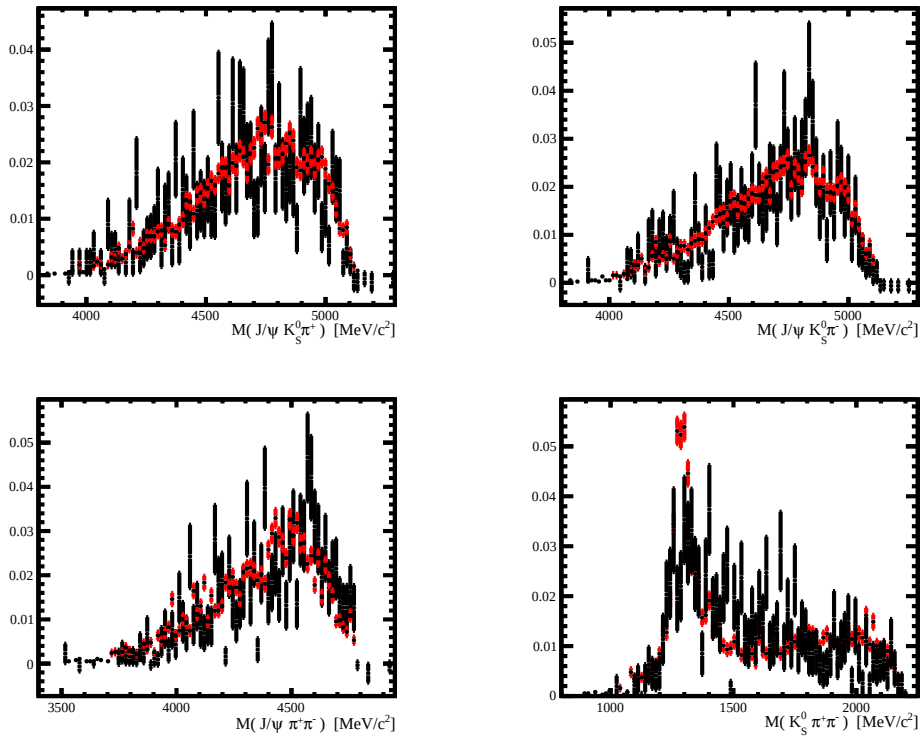

Figure B.9: $B^0 \to J/\psi\, K^0_S(\mathrm{DD})\pi^+\pi^-$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

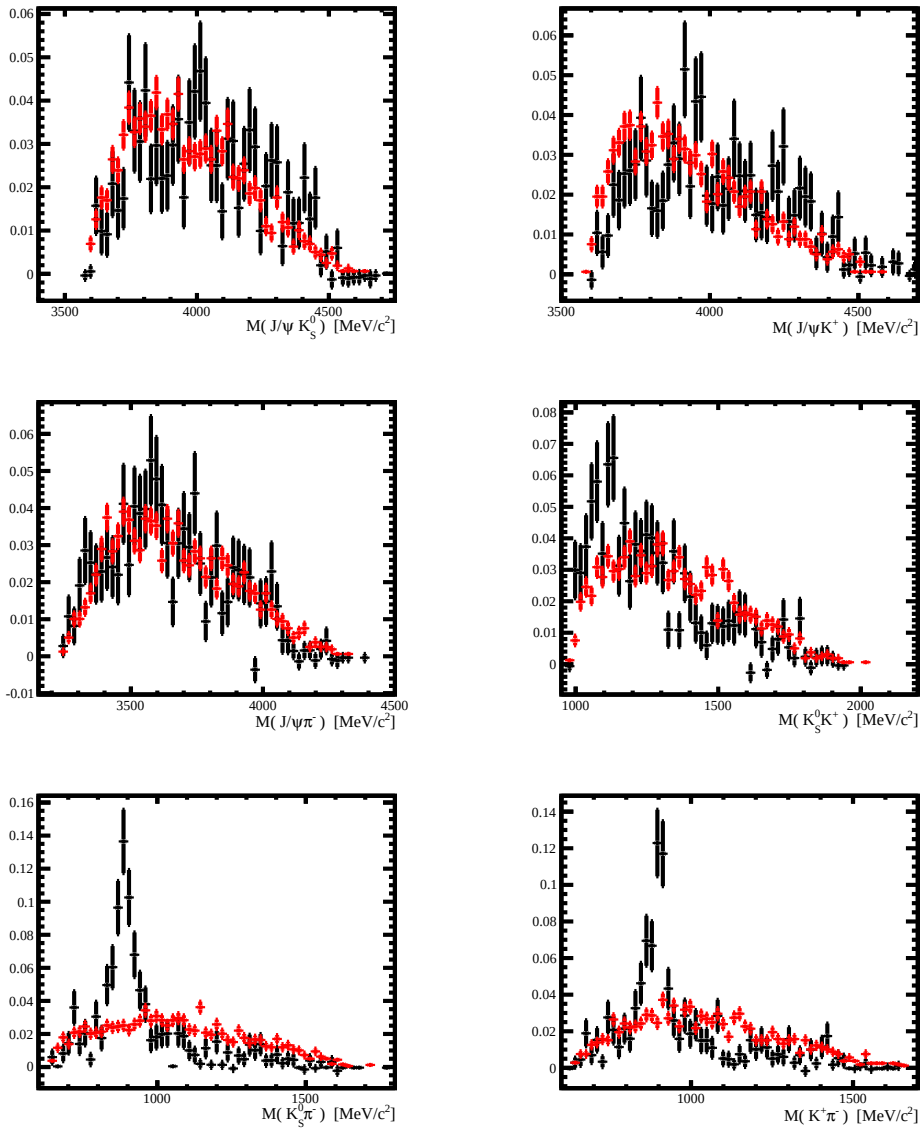

Figure B.10: $B_s^0 \to J/\psi\, K_S^0(\mathrm{DD})K^{\pm}\pi^{\mp}$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

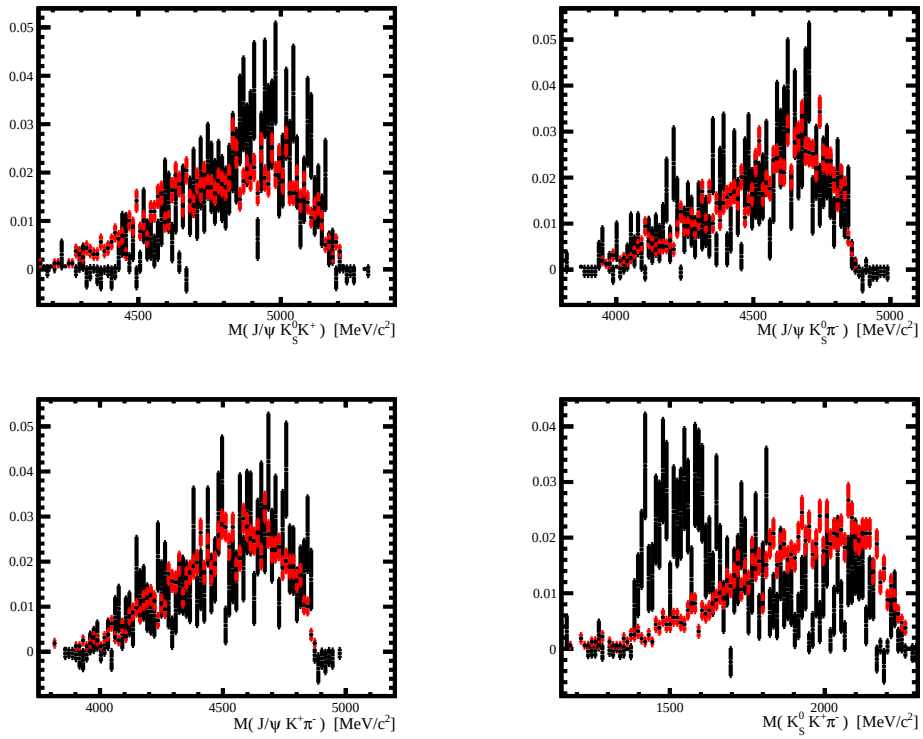

Figure B.11: $B^0_s \to J/\psi\, K^0_S(\mathrm{DD}) K^\pm \pi^\mp$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

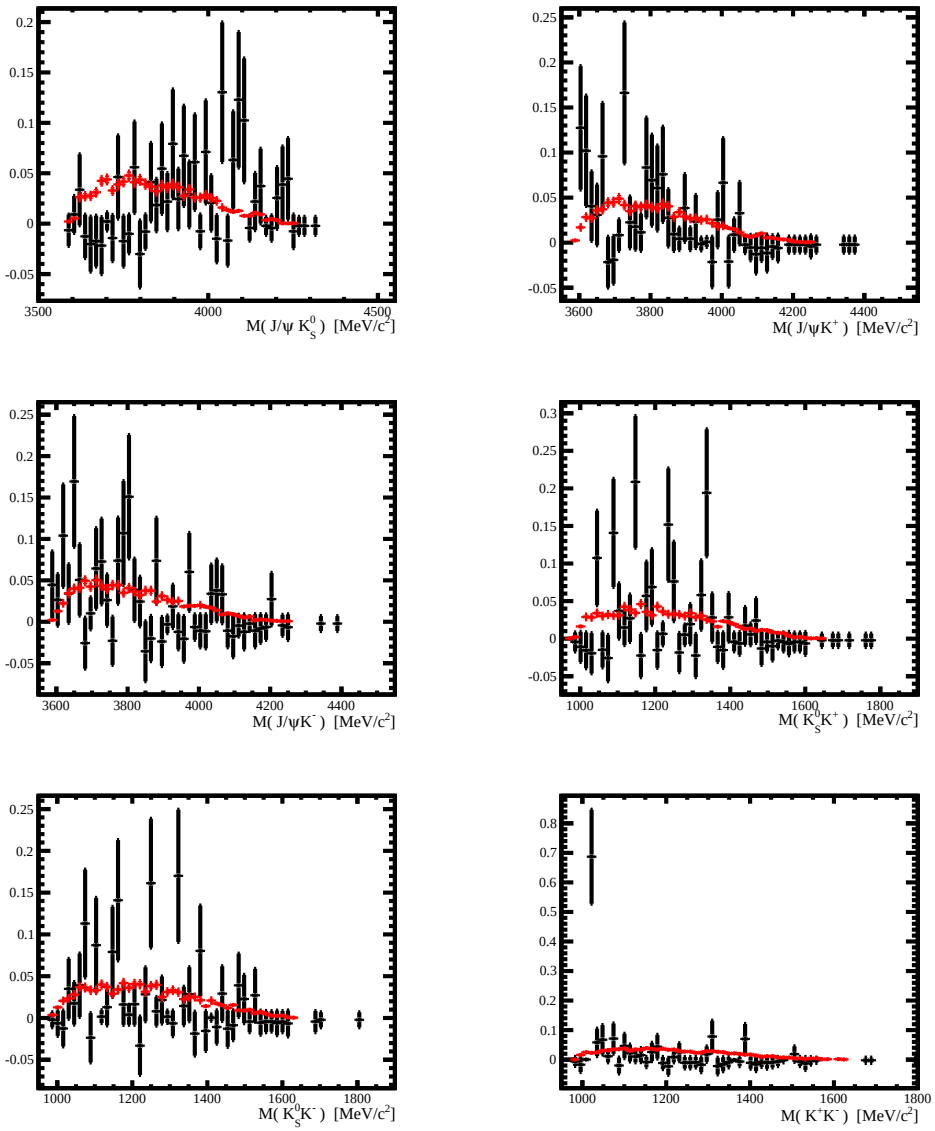

Figure B.12: $B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ two body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

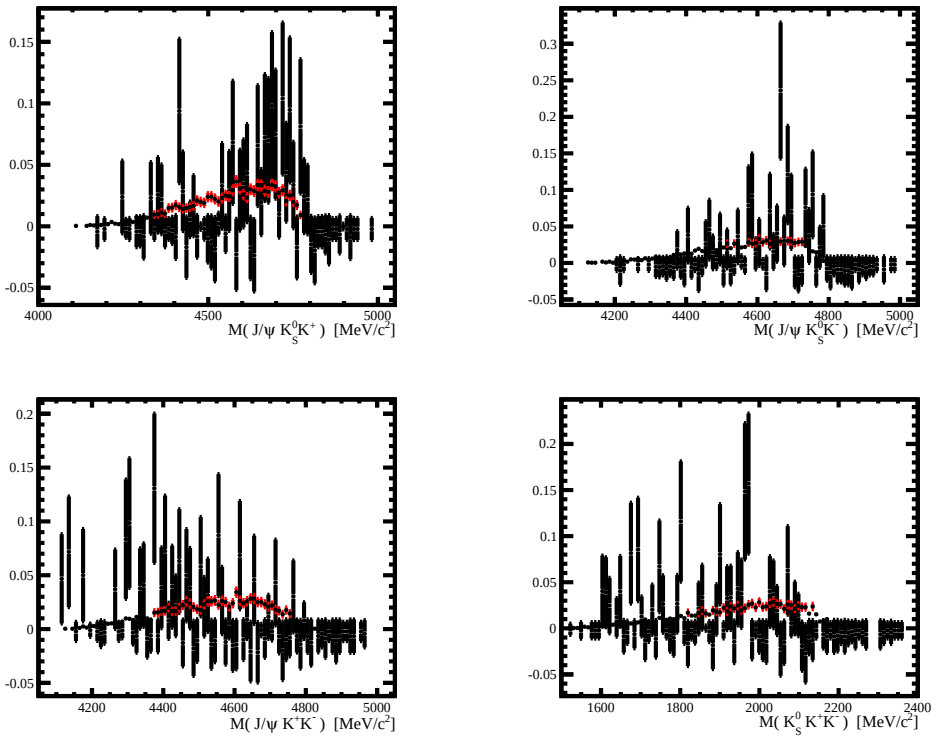

Figure B.13: $B^0 \to J/\psi K^0_S(\mathrm{DD})K^+K^-$ three body combinations of the daughters, black is for ${}_s\mathcal{W}eighted$ data, red represent simulated events. Each distributions is normalised with respect to total number of events in each sample thus the y-axis has arbitary units.

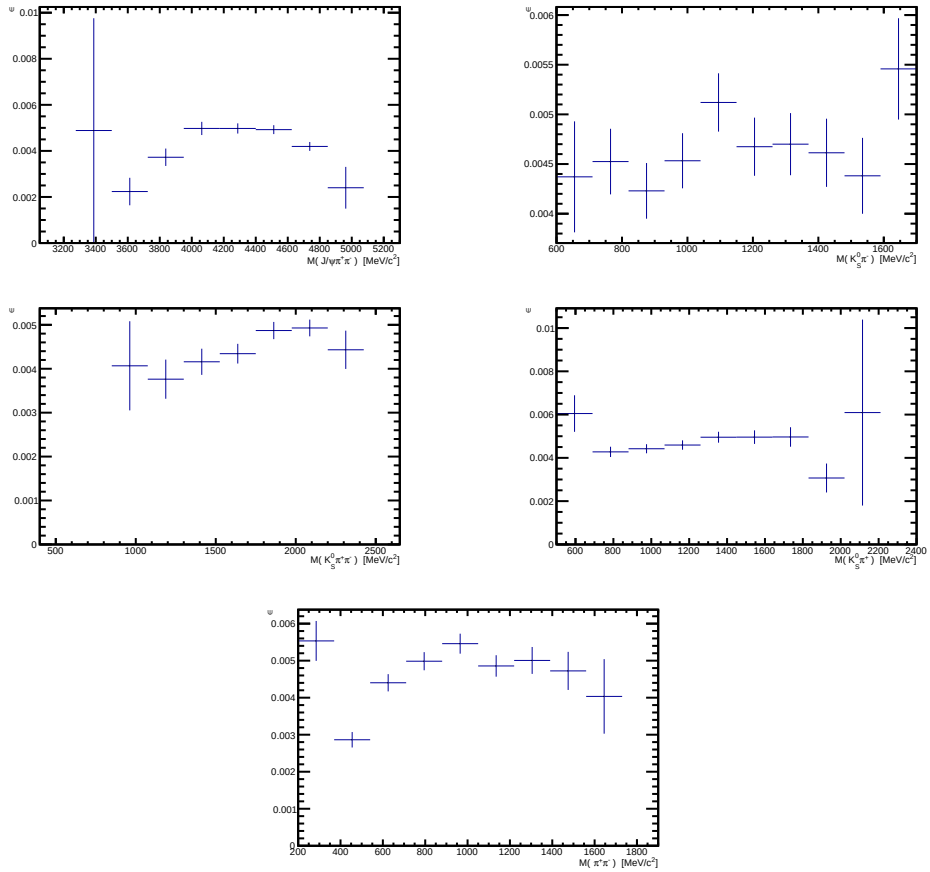

Figure B.14: $B_s^0 \to J/\psi\, K^0_{\rm S}({\rm DD})\pi^+\pi^-$ example showing invariant mass combinations as a function of $\epsilon^{\rm Selection}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements.

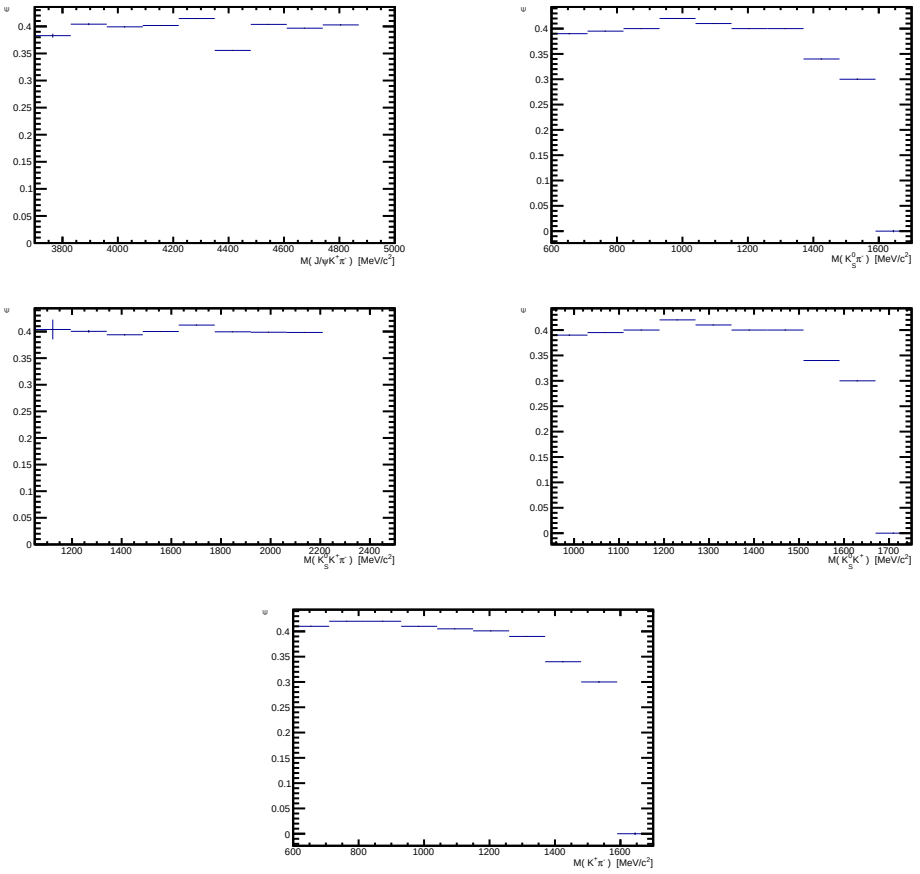

Figure B.15: $B^0 \to J/\psi K^0_S(\mathrm{DD})K^\pm\pi^\mp$ example showing invariant mass combinations as a function of $\epsilon^{\mathrm{Selection}}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements.

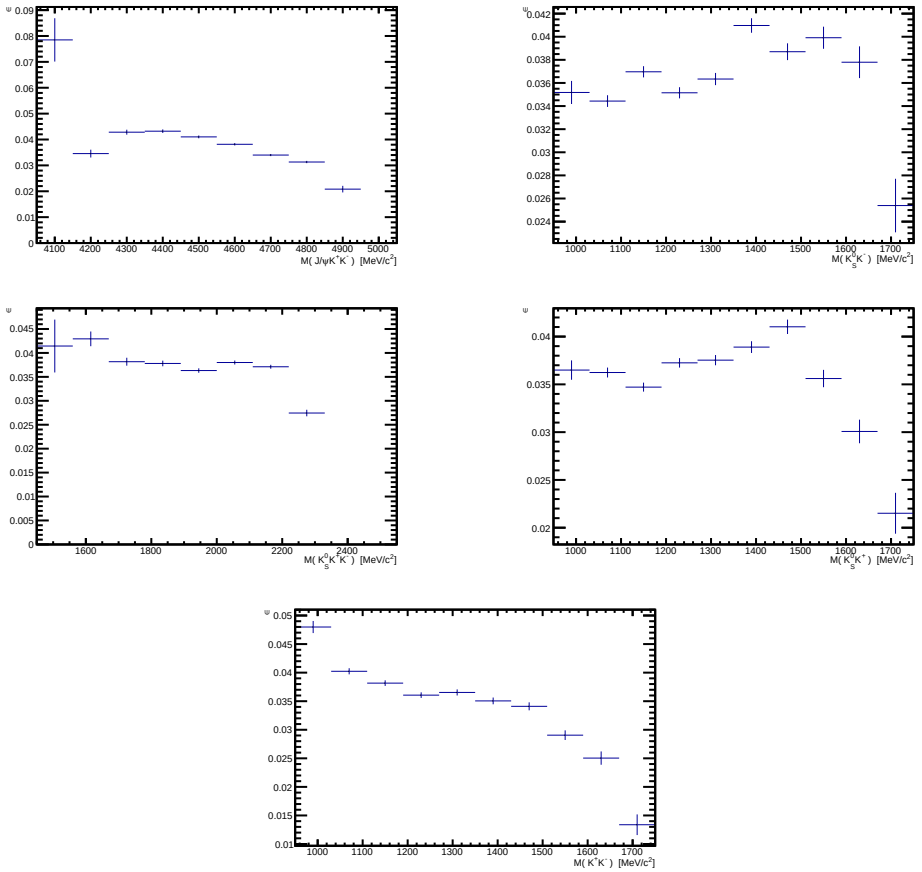

Figure B.16: $B_s^0 \to J/\psi K_{\rm S}^0({\rm DD})K^+K^-$ example showing invariant mass combinations as a function of $\epsilon^{\rm Selection}$, which is the efficiency for the event to pass the reconstruction, stripping, off-line selection and trigger requirements.

Appendix C

Efficiencies for $B^0 \to \psi(2S)K^0_S$ decays

The efficiencies in this section are all obtained using $B^0 \to \psi(2S)K^0_S$ simulated events, with $\psi(2S) \to J/\psi\pi^+\pi^-$. The MC samples are summarised in Tab. C.1. Since the $B^0 \to \psi(2S)K^0_S$ decay branching fraction has been measured to about 8 % precision, the consistency of the results in this channel provides a useful cross-check of the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ analysis.

Table C.1: Signal Monte Carlo datasets for $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)\ K^0_S$ decay modes including event type and both magnet polarisations.

Decay mode	Magnet polarity	
	MagUp	MagDown
$B^0 \to J/\psi K^0_S$	5047978	5048487
$B^0 \to \psi(2S)\ K^0_S$	523498	517998

C.0.1 Geometric efficiency

Both generated samples utilise the `DaughtersInLHCb` cut and the efficiency of this cut is outlined in Table C.2 for both the $\psi(2S)$ and J/ψ modes.

C.0.2 Reconstruction, selection and trigger efficiency

Tables C.3 and C.4 show the efficiencies as determined from MC11a simulated events for $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$ respectively The "Stripping" efficiency for

Table C.2: `DaughtersInLHCb` cut efficiency.

Decay mode	Geometric Efficiency (%)
$\epsilon^{\text{Gen.}}_{B^0\to J/\psi K^0_S}$	17.97 ± 0.06
$\epsilon^{\text{Gen.}}_{B^0\to \psi(2S) K^0_S}$	15.95 ± 0.06

Table C.3: Trigger, selection and reconstruction efficiencies for $B^0 \to \psi(2S)K^0_S$. The numbers of events surviving each cut are given, with the values in parentheses being the efficiencies calculated relative to the preceding requirement, following the order of the table. The overall efficiency and its uncertainty (statistical only) is also given.

	$B^0 \to \psi(2S)K^0_S$(LL)(%)		$B^0 \to \psi(2S)K^0_S$(DD)(%)	
	MagDown	MagUp	MagDown	MagUp
Stripping	1699(0.328)	1861(0.356)	6303(1.212)	6210(1.186)
Truth Matched	735(43)	829(45)	2533(40)	2458(40)
Selection	228(31)	266(32)	741(29)	706(29)
L0	216(95)	249(94)	697(94)	662(94)
Hlt1	202(94)	239(95)	679(97)	641(97)
Hlt2	194(96)	232(97)	637(94)	609(95)
$\epsilon^{\text{Selection.}}$ (%)	0.0375	0.0443	0.143	0.135
σ_ϵ(stat.) (%)	$\pm$0.0027	$\pm$0.0029	$\pm$0.0049	$\pm$0.0047

$B^0 \to J/\psi K^0_S$ is not the same as in Sec. 8. There are two differences in that we use MC matching and previously we had already constrained the mass window to be that of the fitted range, now the stripping numbers are all taken directly from what comes out of stripping over the whole mass range (the reasoning was merely internal comparison cross-check with numbers).

C.0.3 Track reconstruction efficiency $\psi(2S)$

Similarly we need to account for the tracking efficiency since we have two more charged tracks in the final state. Table C.5 clearly shows that the data/MC tracking efficiency corrections are consistent with unity.

C.0.4 Particle identification efficiency

The particle identification efficiency is characterised explicitly for the $\psi(2S)$ mode to see if there is a substantial difference from that of the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode. Since the decay is topologically similar the only difference is the available phase

Table C.4: Trigger, selection and reconstruction efficiencies for $B^0 \to J/\psi\, K^0_S$. The numbers of events surviving each cut are given, with the values in parentheses being the efficiencies calculated relative to the preceding requirement, following the order of the table. The overall efficiency and its uncertainty (statistical only) is also given.

	$B^0 \to J/\psi\, K^0_S$(LL)(%)		$B^0 \to J/\psi\, K^0_S$(DD)(%)	
	MagDown	MagUp	MagDown	MagUp
Stripping	119537(2.4)	119113(2.4)	352043(7)	351628(7)
Truth Matched	111841(94)	111302(93)	323299(92)	323117(92)
Selection	51564(46)	51030(46)	142754(44)	141441(44)
L0	46003(89)	45651(89)	126416(89)	125294(89)
Hlt1	43790(95)	43385(95)	118297(94)	117227(94)
Hlt2	41453(95)	40993(94)	110380(93)	109222(93)
$\epsilon^{\text{Selection.}}$ (%)	0.821	0.812	2.827	2.802
σ_ϵ(stat.) (%)	$\pm$0.004	$\pm$0.004	$\pm$0.0065	$\pm$0.0065

Table C.5: Data/MC tracking efficiency corrections for $B^0 \to \psi(2S)\ K^0_S$ as calculated using the tracking efficiency tables. Separated by LL/DD and magnet polarity. Note that the uncertainty comes from the size of the signal MC samples.

	$B^0 \to J/\psi\, K^0_S$(LL)$\pi^+\pi^-$		$B^0 \to J/\psi\, K^0_S$(DD)$\pi^+\pi^-$	
	MagDown	MagUp	MagDown	MagUp
$R(\pi^+)$	1.006 ± 0.020	1.007 ± 0.020	1.007 ± 0.018	1.007 ± 0.019
$R(\pi^-)$	1.007 ± 0.019	1.016 ± 0.017	1.007 ± 0.019	1.006 ± 0.017
$R^{\text{Track}}_{\text{DATA./MC.}}$	1.013 ± 0.028	1.013 ± 0.026	1.014 ± 0.026	1.013 ± 0.026

space for the bachelor pions which is limited to $m_{B^0} - m_{J/\psi} - m_{K^0_S} \approx 1685\,\text{MeV}/c$ for the $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ mode and for the $B^0 \to \psi(2S)K^0_S$ decay, $m_{\psi(2S)} - m_{J/\psi} \approx 590\,\text{MeV}/c$. The efficiencies outlined in Table C.6 show there is a drop of around 7 % when comparing with the $B^0 \to J/\psi\, K^0_S\pi^+\pi^-$ PID efficiency.

C.1 Systematic uncertainties for $\mathcal{B}(B^0 \to \psi(2S)K^0_S)/\mathcal{B}(B^0 \to J/\psi\, K^0_S)$

To obtain the total systematic error on $\mathcal{B}(B^0 \to \psi(2S)K^0_S)/\mathcal{B}(B^0 \to J/\psi\, K^0_S)$ we follow the same procedure as described in Sec. 8.2. Most of the contributions are described in Appendix C, while that from the fit model is given below. The contribution due to efficiency variation across the phase space is not relevant in this

Table C.6: PID efficiency, for $B^0 \to \psi(2S)K^0_S$ split by K^0_S(LL), K^0_S(DD) and magnet polarities, using the method described in the text. The uncertainties are given in order as; statistical due from the control sample and induced systematic using the outlined method.

	$B^0 \to \psi(2S)\ K^0_S$ (LL) (%)		$B^0 \to \psi(2S)\ K^0_S$ (DD) (%)	
	MagDown	MagUp	MagDown	MagUp
$\epsilon^{\text{PID.}\vert\text{Sel}}$	68.958 ± 0.013	67.090 ± 0.013	68.329 ± 0.008	66.612 ± 0.008
$\Delta^{\text{PID.}}$	4.128	5.254	5.418	4.505

case.

C.1.1 Fit model – $B^0 \to \psi(2S)K^0_S$

To fit the $B^0 \to \psi(2S)K^0_S$ mode, we fix the parameters of the double Gaussian function used in the fit to those obtained from the $B^0 \to J/\psi K^0_S\pi^+\pi^-$ mode.[1] The background PDF is varied in the same way as before, using a second order polynomial. Table. C.7 presents the variation of the fit model and the background where we get a total systematic summing to 7.2 %.

Table C.7: Variation of the signal and background PDFs for $B^0 \to \psi(2S)K^0_S$(LL) and $B^0 \to \psi(2S)K^0_S$(DD) modes.

Parameter	Signal variation		Background variation	
	$B^0 \to \psi(2S)K^0_S$(LL)	$B^0 \to \psi(2S)K^0_S$(DD)	$B^0 \to \psi(2S)K^0_S$(LL)	$B^0 \to \psi(2S)K^0_S$(DD)
χ^2	2.0	18.5	3.4	18.7
# floated params.	5	4	5	5
# bins	29	25	25	25
$P(\chi^2, \text{ndf})$	1.0	0.62	0.99	0.54
N_{B^0}	25 ± 4	57 ± 8.1	30 ± 6	59 ± 8
$\Delta^{\text{Fit.}}_{\text{var}}$(%)	3.3	3.3	0.0	0.0

C.1.2 Total systematic uncertainty

We sum all the individual contributions that we have considered above in quadrature after taking the harmonic mean value. These results are summarised in Table 8.14 with the final systematic result provided.

[1] Again the mean value is floated.

Table C.8: Summary of the weighted-average values of all considered sources of systematic error for $B^0 \to \psi(2S)K^0_S\pi^+\pi^-$. The final value is the sum of the contributions in quadrature.

	$B^0 \to \psi(2S)K^0_S$ (%)
$\bar{\Delta}^{\text{Trig.}}$	2.5
$\bar{\Delta}^{\text{PID.}}$	4.8
$\bar{\Delta}^{\text{Track.}}$	1.2
$\bar{\Delta}^{\text{Fit.}}$	3.3
$\bar{\Delta}^{\text{MC.}}$	3.0
$\bar{\Delta}^{\text{Tot.}}$	7.2

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Observation of the $B_s^0 \to J/\psi K_S^0 K^\pm \pi^\mp$ decay*, arXiv:1405.3219.

[2] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[3] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973) 652.

[4] T. Gershon, *Flavour physics in the LHC era*, arXiv:1306.4588.

[5] LHCb collaboration, R. Aaij *et al.*, and A. Bharucha *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013) 2373, arXiv:1208.3355.

[6] D0 collaboration, V. M. Abazov *et al.*, *Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions*, Phys. Rev. **D85** (2012) 032006, arXiv:1109.3166.

[7] CDF collaboration, T. Aaltonen *et al.*, *Measurement of the bottom-strange meson mixing phase in the full CDF data set*, Phys. Rev. Lett. **109** (2012) 171802, arXiv:1208.2967.

[8] ATLAS collaboration, G. Aad *et al.*, *Time-dependent angular analysis of the decay $B_s^0 \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase ϕ_s by ATLAS*, JHEP **12** (2012) 072, arXiv:1208.0572.

[9] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, Phys. Rev. Lett. **108** (2012) 101803, arXiv:1112.3183.

[10] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-violation and the B_s^0-meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[11] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violating phase ϕ_s in $\overline{B}^0_s \to J/\psi f_0(980)$*, Phys. Lett. **B707** (2012) 497, arXiv:1112.3056.

[12] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in $\overline{B}^0_s \to J/\psi\pi^+\pi^-$ decays*, Phys. Lett. **B713** (2012) 378, arXiv:1204.5675.

[13] LHCb collaboration, R. Aaij *et al.*, *First measurement of the CP-violating phase in $B^0_s \to \phi\phi$ decays*, Phys. Rev. Lett. **110** (2013) 241802, arXiv:1303.7125.

[14] LHCb collaboration, R. Aaij *et al.*, *Observation of $\overline{B}^0_{(s)} \to J/\psi f_1(1285)$ decays and measurement of the $f_1(1285)$ mixing angle*, Phys. Rev. Lett. **112** (2014) 091802, arXiv:1310.2145.

[15] J. B. et al., *2012 Review of Particle Physics*, Physics Review **D 86** (2012), no. 1.

[16] CDF collaboration, T. Affolder *et al.*, *A study of $B^0 \to J/\psi K^{(*)0}\pi^+\pi^-$ decays with the Collider Detector at Fermilab*, Phys. Rev. Lett. **88** (2002) 071801, arXiv:hep-ex/0108022.

[17] Belle collaboration, K. Abe *et al.*, *Observation of $B \to J/\psi K_1(1270)$*, Phys. Rev. Lett. **87** (2001) 161601, arXiv:hep-ex/0105014.

[18] Belle collaboration, S.-K. Choi *et al.*, *Bounds on the width, mass difference and other properties of $X(3872) \to \pi^+\pi^- J/\psi$ decays*, Phys. Rev. **D84** (2011) 052004, arXiv:1107.0163.

[19] BaBar collaboration, B. Aubert *et al.*, *Rare B decays into states containing a J/ψ meson and a meson with $s\bar{s}$ quark content*, Phys. Rev. Lett. **91** (2003) 071801, arXiv:hep-ex/0304014.

[20] CLEO collaboration, C. Jessop *et al.*, *First observation of the decay $B \to J/\psi\phi K$*, Phys. Rev. Lett. **84** (2000) 1393, arXiv:hep-ex/9908014.

[21] Belle collaboration, S. Choi *et al.*, *Observation of a narrow charmonium-like state in exclusive $B^+ \to K^\pm\pi^+\pi^- J/\psi$ decays*, Phys. Rev. Lett. **91** (2003) 262001, arXiv:hep-ex/0309032.

[22] BaBar collaboration, B. Aubert *et al.*, *Study of the $B \to J/\psi K^-\pi^+\pi^-$ decay and measurement of the $B \to X(3872)K^-$ branching fraction*, Phys. Rev. **D71** (2005) 071103, arXiv:hep-ex/0406022.

[23] LHCb collaboration, R. Aaij *et al.*, *Determination of the $X(3872)$ quantum numbers*, Phys. Rev. Lett. **110** (2013) 222001, arXiv:1302.6269.

[24] CDF collaboration, T. Aaltonen *et al.*, *Evidence for a narrow near-threshold structure in the $J/\psi\phi$ mass spectrum in $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **102** (2009) 242002, arXiv:0903.2229.

[25] D0 collaboration, V. M. Abazov *et al.*, *Search for the $X(4140)$ state in $B^+ \to J/\psi\,\phi K^+$ decays with the D0 detector*, Phys. Rev. **D89** (2014) 012004, arXiv:1309.6580.

[26] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a peaking structure in the $J/\psi\,\phi$ mass spectrum from $B^\pm \to J/\psi\,\phi K^\pm$ decays*, arXiv:1309.6920.

[27] LHCb collaboration, R. Aaij *et al.*, *Search for the $X(4140)$ state in $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. **D85** (2012) 091103(R), arXiv:1202.5087.

[28] R. Fleischer, R. Knegjens, and G. Ricciardi, *Anatomy of $B^0_{s,d} \to J/\psi f_0(980)$*, Eur. Phys. J. **C71** (2011) 1832, arXiv:1109.1112.

[29] R. Fleischer, R. Knegjens, and G. Ricciardi, *Exploring CP violation and η–η' mixing with the $B^0_{s,d} \to J/\psi\eta^{(\prime)}$ systems*, Eur. Phys. J. **C71** (2011) 1798, arXiv:1110.5490.

[30] S. Stone and L. Zhang, *Use of $B \to J/\psi\, f_0$ decays to discern the $q\bar{q}$ or tetraquark nature of scalar mesons*, Phys. Rev. Lett. **111** (2013) 062001, arXiv:1305.6554.

[31] A. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[32] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[33] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[34] T2K Collaboration, K. Abe *et al.*, *Measurement of Neutrino Oscillation Parameters from Muon Neutrino Disappearance with an Off-axis Beam*, Phys. Rev. Lett. **111** (2013), no. 21 211803, arXiv:1308.0465.

[35] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[36] M. Neubert, *B decays and CP violation*, Int. J. Mod. Phys. **A11** (1996) 4173, arXiv:hep-ph/9604412.

[37] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[38] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[39] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[40] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[41] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[42] CDF collaboration, F. Abe, *Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab*, Phys. Rev. Lett. **74** (1995) 2626.

[43] D0 collaboration, S. Abachi, *Search for high mass top quark production in $p\overline{p}$ collisions at $\sqrt{s} =$ 1.8 tev*, Phys. Rev. Lett. **74** (1995) 2422.

[44] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal <math><mi>cp</mi></math> nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[45] UTfit collaboration, D. Derkach, *Unitarity Triangle Fitter Results for CKM Angles*, arXiv:1301.3300.

[46] G. Branco, L. Lavoura, and J. Silva, *CP Violation*. International series of monographs on physics. Clarendon Press, 1999.

[47] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the phase space of $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}$ and $B^{\pm} \to K^{\pm}K^{+}K^{-}$ decays*, Phys. Rev. Lett. **111** (2013) 101801, arXiv:1306.1246.

[48] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the phase space of $B^{\pm} \to K^{+}K^{-}\pi^{\pm}$ and $B^{\pm} \to \pi^{+}\pi^{-}\pi^{\pm}$ decays*, arXiv:1310.4740.

[49] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013.

[50] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[51] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K^0_S$ decays*, Phys. Lett. B **721** (2012) 24, Comments: 16 pages, 3 figures.

[52] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of B-Hadron, C-Hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[53] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in $\overline{B}^0_s \to J/\psi\pi^+\pi^-$ decays*, Phys. Lett. **B713** (2012) 378, arXiv:1204.5675.

[54] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, arXiv:0901.0002.

[55] B. Aubert and others., *The {BABAR} detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **479** (2002), no. 1 1 , Detectors for Asymmetric B-factories.

[56] A. Abashian and others., *The belle detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **479** (2002), no. 1 117 , Detectors for Asymmetric B-factories.

[57] *LHC*, http://bigscience.web.cern.ch/bigscience/en/lhc/lhc2.html.

[58] LHCb collaboration, J. Alves, A. Augusto *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[59] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} =$ 7 TeV in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

[60] CMS collaboration, S. Chatrchyan *et al.*, *Measurement of the B(s) to mu+ mu- branching fraction and search for B0 to mu+ mu- with the CMS Experiment*, Phys. Rev. Lett. **111** (2013) 101804, arXiv:1307.5025.

[61] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0_s \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, arXiv:1307.5024.

[62] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.

[63] A. Affolder *et al.*, *Radiation damage in the LHCb vertex locator*, JINST **8** (2013) P08002, arXiv:1302.5259.

[64] D. Dossett, *Performance of the {LHCb} {VELO}*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **718** (2013), no. 0 310 , Proceedings of the 12th Pisa Meeting on Advanced Detectors La Biodola, Isola d'Elba, Italy, May 20 âĂŞ 26, 2012.

[65] *ST Material for Publications.*, http://lhcb.physik.uzh.ch/ST/public/material/index.php.

[66] M. Tobin, *Performance of the LHCb Silicon Tracker in pp Collisions at the LHC*, .

[67] LHCb collaboration, Antunes-Nobrega *et al.*, *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.

[68] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **1210** (2012) 037, arXiv:1206.2794.

[69] LHCb RICH group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. C **73** (2012) 2431. 25 p.

[70] LHCb collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.

[71] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.

[72] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[73] *EvtGen*, `http://evtgen.warwick.ac.uk/`.

[74] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K^0_S$ decays*, Phys. Lett. **B721** (2013) 24, `arXiv:1211.6093`.

[75] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0_s \to J/\psi K^0_S$ effective lifetime*, Nucl. Phys. **B873** (2013) 275, `arXiv:1304.4500`.

[76] LHCb collaboration, R. Aaij *et al.*, *Study of $B^0_{(s)} \to K^0_S h^+ h'^-$ decays with first observation of $B^0_s \to K^0_S K^\pm \pi^\mp$ and $B^0_s \to K^0_S \pi^+ \pi^-$*, JHEP **10** (2013) 143, `arXiv:1307.7648`.

[77] LHCb collaboration, R. Aaij *et al.*, *Searches for Λ^0_b and Ξ^0_b decays to $K^0_S p\pi^-$ and $K^0_S pK^-$ final states with first observation of the $\Lambda^0_b \to K^0_S p\pi^-$ decay*, Tech. Rep. arXiv:1402.0770. LHCB-PAPER-2013-061. CERN-PH-EP-2014-012, CERN, Geneva, Feb, 2014. Comments: 27 pages, 4 figures.

[78] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetry of $B \to K^{(*)}\mu^+\mu^+$ decays*, LHCb-PAPER-2014-006, in preparation.

[79] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, `arXiv:1111.2357`.

[80] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, `arXiv:1301.5286`.

[81] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011.

[82] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0_s \to J/\psi K^0_S$ branching fraction*, Phys. Lett. **B713** (2012) 172, `arXiv:1205.0934`.

[83] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s \to J/\psi \overline{K}^{*0}$ branching fraction and angular amplitudes*, Phys. Rev. **D86** (2012) 071102(R), `arXiv:1208.0738`.

[84] BaBar collaboration, B. Aubert *et al.*, *Measurement of branching fractions and charge asymmetries for exclusive B decays to charmonium*, Phys. Rev. Lett. **94** (2005) 141801, `arXiv:hep-ex/0412062`.

[85] BaBar collaboration, P. del Amo Sanchez *et al.*, *Evidence for the decay* $X(3872) \to J/\psi\omega$, Phys. Rev. **D82** (2010) 011101, arXiv:1005.5190.

[86] CLEO collaboration, S. Richichi *et al.*, *Study of* $B \to \psi(2S)K$ *and* $B \to \psi(2S)K^*(892)$ *decays*, Phys. Rev. **D63** (2001) 031103, arXiv:hep-ex/0010036.

[87] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[88] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in* $\overline{B}^0 \to J/\psi\pi^+\pi^-$, Phys. Rev. **D87** (2013) 052001, arXiv:1301.5347.

[89] C. Fitzpatrick, *simpletools: Handy command line tools for ntuple manipulation and analysis*, Tech. Rep. LHCb-INT-2009-029. CERN-LHCb-INT-2009-029, CERN, Geneva, Nov, 2009.

[90] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances.* PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[91] A. Jaeger, P. Seyfert, M. De Cian, J. van Tilburg, and S. Hansmann-Menzemer, *Measurement of the track finding efficiency*, LHCb-PUB-2011-025.

[92] *LHCb tracking efficiency*, https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbTrackingEfficiencies.

[93] R. Aaij and J. Albrecht, *Muon triggers in the High Level Trigger of LHCb*, LHCb-PUB-2011-017.

[94] J. A. Hernando Morata, E. Lopez Asamar, D. Martinez Santos, H. Ruiz-Pérez, and F. Teubert, *Measurement of trigger efficiencies and biases*, LHCb-2008-073.

[95] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research A **555** (2005) 356, physics/0.

[96] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, arXiv:hep-ex/0011057.

[97] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[98] L. Lyons, D. Gibaut, and P. Clifford, *How to combine correlated estimates of a single physical quantity*, Nucl. Instrum. Meth. **A270** (1988) 110.

[99] A. Valassi, *Combining correlated measurements of several different physical quantities*, Nucl. Instrum. Meth. **A500** (2003) 391.

[100] LHCb collaboration, R. Aaij and others., *Searches for $B^0_{(s)} \to J/\psi p\bar{p}$ and $B^+ \to J/\psi p\bar{p}\pi^+$ decays*, J. High Energy Phys. **09** (2013) 006. 12 p, Comments: 12 pages, 3 figures.

[101] S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, The Annals of Mathematical Statistics **9** (1938) 60.