Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN

FISICA

CICLO XXVIII

COORDINATORE Prof. Vincenzo Guidi

B-flavour tagging calibration for CP violation measurement in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays at LHCb

Settore Scientifico Disciplinare FIS/01

CERN-THESIS-2016-448

Dottorando

Dott. Giulia Tellarini

Tutori

Dott. Stefania Vecchi

Dott. Wander Baldini

ANNI 2013/2015

Giulia Tellarini: *B-flavour tagging calibration for CP violation measurement in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays at LHCb*, 18/07/2016

This thesis is licensed under a Creative Commons Attribution 3.0 License.

CONTENTS

1 Introduction 1

2 Physics Introduction 3
  2.1 The Standard Model of Particle Physics 3
  2.2 CP violation in the Standard Model 6
    2.2.1 The CKM formalism 7
    2.2.2 The Unitarity triangles 9
  2.3 *B* meson Physics 12
    2.3.1 Mixing of neutral B mesons 13
    2.3.2 Time evolution of flavour eigenstates 15
  2.4 Different types of CP violation 17
  2.5 Measurement of γ through *B* decays 18
    2.5.1 Tree-level decays to measure γ 20
  2.6 Measurement of the CKM angle γ in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays 23

3 The LHC*b* Experiment 25
  3.1 The Large Hadron Collider 25
  3.2 The LHC*b* Experiment 27
    3.2.1 The tracking system 30
    3.2.2 The LHCb magnet 36
    3.2.3 Mass and Momentum resolution 37
    3.2.4 Particle Identification 38
    3.2.5 The calorimeters system 41
    3.2.6 Muon detector 43
    3.2.7 The LHCb trigger 44
    3.2.8 Reconstruction algorithms 47
    3.2.9 The LHCb software 48

4 Flavour Tagging 51

4.1 Flavour Tagging definitions 51
4.2 Flavour Tagging algorithms 56
4.2.1 NNetSSK tagger 56
4.2.2 OS tagger 61
4.2.3 Mistag probabilities and combination of taggers 62
4.3 Flavour Tagging calibration 63

5 Flavour Tagging Calibration using $B_s^0 \to D_s^- \pi^+$ Run1 data 67
5.1 Fit strategy 68
5.1.1 Mass Fit 69
5.1.2 Time sFit 72
5.2 Calibration of the NNetSSK tagger 75
5.2.1 Flavour tagging asymmetries 78
5.2.2 Systematic uncertainties 81
5.2.3 Test in MC samples 85
5.2.4 NNetSSK calibration summary and performances 86
5.3 Calibration of OS tagger 89

6 Measurement of γ angle from the analysis of $B_s^0 \to D_s^\mp K^\pm$ using 1fb^{-1} 93
6.1 Data sample 94
6.1.1 Monte Carlo Simulation 94
6.2 Event selection 96
6.3 Multi-dimensional fit 100
6.3.1 Probability Density Functions 102
6.3.2 Summary of fixed yields 109
6.3.3 Results of the Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$ 109
6.3.4 Results of the Multidimensional fit to $B_s^0 \to D_s^\mp K^\pm$ 110
6.4 Flavour Tagging Calibration 122
6.4.1 Calibration of OS tagger 122
6.4.2 Calibration of NNetSSK tagger 125
6.4.3 Comparison of Tagging Performance in $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ 126
6.4.4 Tagging asymmetries 128

6.4.5 OS and NNetSSK combination 129

6.5 Time resolution and time acceptance 132

6.6 Time fit 135

6.7 Results of the sFit 136

6.7.1 sFit in $B_s^0 \to D_s^- \pi^+$ 136

6.7.2 sFit in $B_s^0 \to D_s^\mp K^\pm$ 138

6.8 Results of the cFit 139

6.8.1 $B_s^0 \to D_s^\mp K^\pm$ fit results 142

6.9 Systematics uncertainties 143

6.10 The γ angle determination 145

7 $B_s^0 \to D_s^\mp K^\pm$ analysis using Run1 data 151

7.1 Preliminary Studies 152

7.1.1 Correlations among observables 152

7.1.2 Effects of a time-dependent tagging efficiency 156

7.2 Data sample and MC simulation 163

7.3 Event selection 165

7.4 Multi-dimensional fit 169

7.4.1 Probability Density Functions 169

7.4.2 Summary of fixed yields 175

7.4.3 Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$ 177

7.4.4 Multi-dimensional fit to $B_s^0 \to D_s^\mp K^\pm$ 179

7.5 Time sFit 193

8 Conclusions 195

A Appendix 197

A.1 Study on CP sensitivities 197

A.2 Validation of the $B_s^0 \to D_s^- \pi^+$ fit with toys and bootstraps studies 202

A.2.1 Toys studies 202

A.2.2 Bootstrap studies 205

A.3 TGenPhaseSpace-based Emulator 207

A.3.1 Events generation 207

A.3.2 Events reconstruction and selection 208
A.3.3 Background emulation 212
A.3.4 Correlations among observables 213
A.4 Toys Study for ϵ_{tag} versus time 220
A.4.1 Pull distributions 220
A.4.2 Toy by toy difference 220

Bibliography 225
Index 230

LIST OF FIGURES

Figure 1 Representation of the Unitarity Triangle in the complex plane. 10

Figure 2 Current constraints on the Unitarity Triangle parameters in the complex plane: including direct measurements of angles (left), not including them (right) [9]. 11

Figure 3 Dominant box diagrams for the $B_q^0 - \overline{B}_q^0$ transition ($q = d$ or s). 13

Figure 4 Feynman diagrams of the $B \to DK$ tree-level decays. 19

Figure 5 Feynman diagrams contributing to the $B \to h^+ h^{'-}$ decays: Tree (T), Penguin (P), Penguin Annihilation (PA), Colour-suppressed Electroweak Penguin (P_{EW}^C) and Exchange (E). 19

Figure 6 Illustration of γ sensitivity in tree-level decays with direct CP violation. 21

Figure 7 Two-dimensional CL distributions of γ vs. r_B^{DK} (left) and γ vs. δ_B^{DK} for various $B^+ \to DK^+$ decays [10]. 22

Figure 8 Two-dimensional CL distributions of γ vs. r_B^{DK} for various B^0 decays [10]. 23

Figure 9 A schematic representation of the LHC accelerator system and the position of the four experiments ATLAS, CMS, ALICE and LHCb. 26

Figure 10 $b - \bar{b}$ production cross-section at $\sqrt{s}$= 8 TeV as a function of polar angles. In red the part of the distribution within the LHCb acceptance. Figure taken by [33]. 28

Figure 11 Layout of the LHCb detector in $y-z$ plane. From left to right: the VErtex LOcator, RICH1 and RICH2 indicate the Ring Imaging Cherenkov detectors, TT and T1-T3 denote the tracking stations, M1-M5 are the muons detectors. SPD and PS, as well as ECAL and HCAL compose the Calorimeters system. The coordinate system is a right-handed Cartesian system where the z axis corresponds to the beam direction, the y axis is contained in the non-bending plane while the x axis is oriented horizontally towards the bending plane. 29

Figure 12 IP_x and IP_y dependence on the momentum (left) and IP_x dependence on $1/P_T$ in data and simulation. Figure taken from [34]. 31

Figure 13 Dependence of the decay time resolution on the B momentum (left) and on the estimation of the decay time error (right) [34]. 32

Figure 14 Cross section of the VELO in the $x-z$ and $y-z$ plane showing the 21 modules [31]. 32

Figure 15 Schematic view of r and ϕ sensors [31]. 33

Figure 16 Layout of two TT detection layers (X and U)[36]. 34

Figure 17 Layout of a single layer (X) of the IT detector [36]. 34

Figure 18 Schematic view of the tracking system. The blue colored parts represent the OT while the purple ones represent the TT and IT trackers. Figure taken from[31]. 35

Figure 19 Perspective view of the LHCb dipole magnet. The interaction point lies behind the magnet [31]. 36

Figure 20 Relative momentum resolution versus momentum for long tracks in data obtained using J/ψ decays. Figure taken from [32]. 37

Figure 21 Relative mass resolution as a function of the mass of the dimuon resonance (left) and mass resolution values for the six dimuon resonances (right). Figures taken from [32]. 38

Figure 22 Right: Cherenkov angle versus particle momentum for RICH radiators. Figure taken from [31]. Left: Kaon identification and pion misidentification as a function of the track momentum measured on Run1 data. Plot for two different $DLL(K-\pi)$ requirements are shown. Figure taken from[32]. 39

Figure 23 Schematics of RICH1 and RICH2 detectors. Figure taken from [31]. 40

Figure 24 Schematic representation of modules and granularity of the ECAL (left) and HCAL (right). Only the top right quarter is shown. The black area corresponds to the empty space occupied by the beam pipe [31]. 42

Figure 25 Side view of the five stations of the muon detector. Figure taken from[31]. 43

Figure 26 Schematic representation of the muon chamber assembly showing the different granularities [31]. 44

Figure 27 LHCb trigger schema. Figure taken from [24]. 45

Figure 28 Schematic representation of the possible taggers to infer the B flavour. The image summarizes the case of the $B^0 \rightarrow J/\psi K^{*0}$ decay channel. 52

Figure 29 Schematic representation of the charge correlation between B^0_s , B^0_d and B^+ and the possible tagging particles. 53

Figure 30 The black curve represents the oscillated B mesons while the red curve represents the non-oscillated events. The left plot shows the case of an ideal measurement in case of perfect tagging. The right plot corresponds to a mistag of 0.3. 55

Figure 31 Plots for the verification of the training for the first NN 59

Figure 32 Multiplicity of tagging track candidates selected by NN1 for data (red) and simulation (black). The first bin represents untagged candidates. 60

Figure 33 Verification plots for the training of the second neural network. Shown are the response for B_s^0 and $\bar{B}_s^0$ in the simulation (a) and the comparison of the response of the NN2 in the data and the simulation (b). 61

Figure 34 Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the B_s^0 mass window [5100-5600] MeV/c^{2}used to determine the relative fraction for each background component. 70

Figure 35 Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the [5320-5600] MeV/c^{2}mass window used to extract the *sWeights*. 72

Figure 36 Predicted mistag distribution of signal events obtained applying sweights for the 2011+2012 sample. 76

Figure 37 Average mistag fraction ω_{cat} in bins of predicted mistag probability η. The solid line is the result of a linear fit to the data points and the dashed line is the calibration obtained from the unbinned fit. The shaded area corresponds to the 95% Confidence Level of the linear fit to data points. 76

Figure 38 Average mistag fraction ω_{cat} in bins of predicted mistag probability η. The red line is the result of a linear fit to the 2011+2012 data points and the black line is the calibration obtained from the unbinned fit. The shaded areas correspond to 68 and 95% Confidence Level of the linear fit to data points. 78

Figure 39 Asymmetry between the number of mixed and unmixed B candidates in one period of the mixing oscillations (folding the full decay time distribution), with the fit projection overlaid. The data are sWeighted to consider only the signal candidates. In the left plot, all the candidates tagged by the NNetSSK are considered, with their per-event mistag; in the right plot, only the events with $\eta < 0.35$. The offset $t_0 = 0.2\,\mathrm{ps}$ corresponds to the preselection requirement on the decay time. 79

Figure 40 Left: distribution of the $KK\pi$ invariant mass for selected prompt D_s candidates. Right: difference of the NNetSSK calibration plots for prompt D_s ($D_s^- - D_s^+$), the line correspond to the best fit of the data points with a linear funcion. 81

Figure 41 Distributions of (a) p_T (B_s^0), (b) ϕ and (c) track multiplicity for $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J\psi\phi$, $B_s^0 \to \phi\phi$ and $B_s^0 \to D_s D_s$ decays in simulation. 85

Figure 42 Calibration of NNetSSK tagger MC 88

Figure 43 Fit in category for the OS tagger in the 2011 + 2012 $B_s^0 \to D_s^- \pi^+$ data. The dashed line is the ideal calibration: $p_0 = \langle\eta\rangle$ and p_1 = 1. The shaded area is the 95% Confidence Level 90

Figure 44 BDTG response distributions for training (left) and test (right) samples of $B_s^0 \to D_s^- \pi^+$. The red histogram corresponds to signal weighted by the signal *sWeights*, while the blue shows background. 98

Figure 45 Signal B_s^0 mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated from MC on the joint magnet up and down samples, separately for each D_s final states. The right plots show the pull distributions. 103

Figure 46 Signal B_s^0 mass shapes of $B_s^0 \to D_s^\mp K^\pm$ as evaluated from MC on the joint magnet up and down samples, separately for each D_s final states. The right plots show the pull distributions 105

Figure 47 `RooBinned1DQuinticBase` templates for signal PIDK. Left: $B_s^0 \to D_s^- \pi^+$. Right: $B_s^0 \to D_s^\mp K^\pm$. 106

Figure 48 Combinatorial background B_s^0 mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated, separately for each D_s final state. The right plots show the pull distributions 107

Figure 49 Combinatorial background B_s^0 mass shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$ as evaluated, separately for each D_s final state in the upper B_s^0 sidebands in order to study the PDF in a region of pure combinatorial events. The right plots show the pull distributions 108

Figure 50 Combinatorial background D_s mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated (in the upper B_s^0 sidebands), separately for each D_s final state. The right plots show the pull distributions 109

Figure 51 Combinatorial background D_s mass shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$ as evaluated (in the upper B_s^0 sidebands), separately for each D_s final state. The right plots show the pull distributions. 110

Figure 52 Combinatorial background $PIDK$ templates used in the $B_s^0 \to D_s^- \pi^+$ fit (left) and $B_s^0 \to D_s^{\mp} K^{\pm}$ fit (right), combined for magnet polarities. 111

Figure 53 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$, $D_s^- \to K^- \pi^+ \pi^-$, $D_s^- \to \pi^- \pi^+ \pi^-$. 114

Figure 54 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$, $D_s^- \to K^- \pi^+ \pi^-$, $D_s^- \to \pi^- \pi^+ \pi^-$. 115

Figure 55 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of $-PIDK$ mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$, $D_s^- \to K^- \pi^+ \pi^-$, $D_s^- \to \pi^- \pi^+ \pi^-$. 116

Figure 56 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 118

Figure 57 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 119

Figure 58 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $log(PIDK)$ mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 120

Figure 59 The simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates for both polarities and all D_s final states combined, from top left to bottom: B_s^0 mass, D_s mass, log(PIDK). 121

Figure 60 The measured OS mistag from the time fit in $B_s^0 \to D_s^- \pi^+$ control mode against the average predicted mistag for each category. The error bars represent only the statistical uncertainties. The blue curve is the linear fit to the 7 data points, the light blue area defines the 68% confidence level region of the calibration function (statistical only). 123

Figure 61 The measured NNetSSK mistag from the time fit against the average predicted mistag for each category in the $B_s^0 \to D_s^- \pi^+$ decay mode. The blue curve is the linear fit to the 7 data points. 125

Figure 62 η distributions for the OS combination (left) and the SS kaon tagger (right) in simulated data. Black is $B_s^0 \to D_s^{\mp} K^{\pm}$ and red is $B_s^0 \to D_s^- \pi^+$. 127

Figure 63 Dependency of the observed mistag ω on calibrated mistag η, for selected $B_s^0 \to D_s^- \pi^+$ (black) and $B_s^0 \to D_s^\mp K^\pm$ (red) events in simulation. Left is OS tagging, right is NNetSSK. 128

Figure 64 η distributions for the OS (left) and the NNetSSK (right) taggers. Red is $B_s^0 \to D_s^\mp K^\pm$ data and black is $B_s^0 \to D_s^- \pi^+$ data. The Kolmogorov test gives 0.94 for the OS and 0.05 for the NNetSSK taggers. 128

Figure 65 Scatter plot of OS mistag (on y-axis) and NNetSSK mistag (on x-axis). 130

Figure 66 The mistag templates obtained from *sWeighted* $B_s^0 \to D_s^- \pi^+$ data. From left to right: only the OS, only the NNetSSK, both taggers. 131

Figure 67 MC Acceptances for $B_s^0 \to D_s^- \pi^+$ (left) and $B_s^0 \to D_s^\mp K^\pm$ (right). The acceptance function is shown in red, while the total fit curve is blue. 134

Figure 68 Time distribution of the $B_s^0 \to D_s^- \pi^+$ data sample. Bottom plot shows pulls. 137

Figure 69 Time *sFit* to the $B_s^0 \to D_s^\mp K^\pm$ data sample. Bottom plot shows pulls. 139

Figure 70 Result of the decay time (on the left side) *sFit* and (on the right side) *cFit* to the $B_s^0 \to D_s^\pm K^\mp$ candidates. 143

Figure 71 Profile likelihood contours of γ vs. r_{D_sK}, and γ vs. δ. The contours are the $n\sigma$ profile likelihood contours, where $\Delta\chi^2 = n^2$ with $n = 1, 2$. The markers denote the best-fit values. All errors (stat. and syst.) are included. 145

Figure 72 Graph showing 1-CL for γ using the *cFit* (green) and *sFit* (blue) CP parameters. All errors (stat. and syst.) are included. 146

Figure 73 1-CL curves for the combination of tree-level measurements with central values (solid vertical lines) and 1σ uncertainties (dashed vertical lines) labelled. 148

Figure 74 Profile likelihood contours from the combination of tree-level measurements with two-dimensional 1σ and 2σ boundaries. 149

Figure 75 From left to right, the fitted value of the C_f, S_f, $S_{\bar{f}}$, D_f, $D_{\bar{f}}$ observables, the corresponding error and pull distributions resulting from the time sFit to $B_s^0 \to D_s^{\mp}K^{\pm}$ pseudo-experiments are reported. 160

Figure 76 Emulated bakgrounds decay distributions for B_s mass, D_s mass and PIDK for the bachelor particle. 161

Figure 77 Emulated bakgrounds decay distributions for B_s decay time and its error. 161

Figure 78 Tagging efficiency dependence on the decay time for NNetSSK (left) and OS (right) taggers in $B_s^0 \to D_s^-\pi^+$ Run1 data. 162

Figure 79 BDTG response distributions for training and test samples. The red histogram corresponds to signal weighted by the signal sWeights, while the blue shows combinatorial background. 166

Figure 80 $B_s^0 \to D_s^{\mp}K^{\pm}$ (left) and $B_s^0 \to D_s^-\pi^+$ (right) events after the overall selection on the $3fb^{-1}$ data set with the superimposed B_s^0 mass distributions for the different background components described in Sec.7.4.1. 168

Figure 81 Fit to B_s^0 invariant mass in $B_s^0 \to D_s^-\pi^+$ signal MC events on the joint magnet up and down samples, separately for D_s final states. From top left to bottom right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. 170

Figure 82 Fit to B_s^0 invariant mass in $B_s^0 \to D_s^{\mp}K^{\pm}$ signal MC events on the joint magnet up and down samples and D_s final states. 174

Figure 83 Fit to D_s invariant mass in $B_s^0 \to D_s^-\pi^+$ signal MC events on the joint magnet up and down samples, separately for D_s final states. From top left to bottom right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. 176

Figure 84 Fit to D_s invariant mass in $B_s^0 \to D_s^{\mp} K^{\pm}$ signal MC events on the joint magnet up and down samples, separately for $D_s^- \to KK\pi$ (left), $D_s^- \to K^-\pi^+\pi^-$ (center) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes. 177

Figure 85 D_s mass distributions for $B_s^0 \to D_s^- \pi^+$ combinatorial (black) and signal (red) events for $D_s^- \to \phi\pi^-$ (left) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes. 178

Figure 86 D_s mass distributions for $B_s^0 \to D_s^- \pi^+$ combinatorial (black) and signal (red) events for $D_s^- \to \phi\pi^-$ (left) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes applying the old BDTG selection. 179

Figure 87 From top to bottom Kinematic distributions of p, p_T and nTracks of bachelor particles for $B_s^0 \to D_s^- \pi^+$ combinatorial events in the signal region [5300,5800]MeV/c^{2}splitted for the different D_s final states (red $D_s^- \to \phi\pi^-$, blue $D_s^- \to K^-\pi^+\pi^-$, green $D_s^- \to K^{*0}K^-$, cyan $D_s^- \to (KK\pi)_{nonres}$, magenta $D_s^- \to \pi^-\pi^+\pi^-$). 180

Figure 88 $PIDK$ distributions for pions (left) and kaons (right) for $B_s^0 \to D_s^- \pi^+$ combinatorial events split for the different D_s final states (red $D_s^- \to K^{*0}K^-$, magenta $D_s^- \to \phi\pi^-$, blue $D_s^- \to (KK\pi)_{nonres}$, black $D_s^- \to K^-\pi^+\pi^-$, cyan $D_s^- \to \pi^-\pi^+\pi^-$,). 181

Figure 89 Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the wide mass region [5100,5800]MeV/c^{2}in the different D_s final states from top left to bottom right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-\pi^+\pi^-$,$D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. 182

Figure 90 Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the wide mass region [5100,5800]MeV/c^2. 182

Figure 91 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 184

Figure 92 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 185

Figure 93 Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of $-log(PIDK)$ for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 186

Figure 94 The simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates for both polarities and all D_s final states combined for B_s^0 mass, D_s mass, $-log(PIDK)$ distributions. 187

Figure 95 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 189

Figure 96 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 190

Figure 97 Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $log(PIDK)$ for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$. 191

Figure 98 The simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates for both polarities and all D_s final states combined, from top to bottom and left to right: B_s^0 mass, D_s mass, log(PIDK)distributions. 192

Figure 99 The signal yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported. 198

Figure 100 The combinatorial yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported. 199

Figure 101 The $B^0 \to D_s K$ yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported. 200

Figure 102 The overall yields distribution of $B_s^0 \to D_s^{(*)}\rho(\pi)$, $\Lambda_b \to D_s^{(*)}p$, $B_s^0 \to D_s^-\pi^+$ backgrounds (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported. 201

Figure 103 Toys studies: distributions of the fitted values and of the pulls for $p_0 - \langle\eta\rangle$ and p_1. 204

Figure 104 Toys studies (signal only): distributions of the fitted values and of the pulls for $p_0 - \langle\eta\rangle$ and p_1. 205

Figure 105 Bootstrap studies (10K samples): distributions of the fitted values and of the fitted error of $p_0 - \langle\eta\rangle$ and p_1. 206

Figure 106 Flight distance error distribution in terms of the measured momentum and flight distance of the B_s meson in MC sample. 210

Figure 107 Kinematic distribution for B_s, D_s, $Kbach$, $K1$, $K2$ and Pi in red from MC and in blue from the output of the Emulator for signal $B_s^0 \to D_s^{\mp} K^{\pm}$ events. 212

Figure 108 Mass and decay time distributions for B_s and D_s, in red from MC and in blue from the output of the emulator for signal $B_s^0 \to D_s^{\mp} K^{\pm}$ events. 213

Figure 109 Comparison among B_s, B_d and Λ_b kinematic distributions from MC samples. Left: P, right: η. 214

Figure 110 Emulated bakgrounds decay distributions for B_s mass, D_s mass and PIDK for the bachelor particle. 214

Figure 111 Emulated bakgrounds decay distributions for B_s decay time and its error. 215

Figure 112 Scatter plots B_s mass versus δ_t for each emulated decay starting from the signal. 215

Figure 113 Scatter plots PIDK of bachelor particle versus δ_t for each emulated decay starting from the signal. 216

Figure 114 Scatter plots PIDK of bachelor particle versus t_R for each emulated decay starting from the signal. 216

Figure 115 Scatter plots t_R mass versus B_s for each emulated decay starting from the signal. 216

Figure 116 Scatter plots t_R mass versus D_s for each emulated decay starting from the signal. 217

Figure 117 Scatter plots δ_t mass versus D_s for each emulated decay starting from the signal. 217

Figure 118 Scatter plots D_s versus B_s mass for each emulated decay starting from the signal. 217

Figure 119 Scatter Plots PIDK of bachelor particle versus B_s mass for each emulated decay starting from the signal. 218

Figure 120 Scatter Plots PIDK of bachelor particle versus D_s mass for each emulated decay starting from the signal. 218
Figure 121 Scatter plots δ_t versus t_R for each emulated decay starting from the signal. 218
Figure 122 CP observables, error and pull distributions for the Toys set ϵ_{tag} (t). 221
Figure 123 CP observables, error and pull distributions for the Toys set ϵ_{tag} = const. 222
Figure 124 CP observables pull distributions for the toy by toy difference between Toys set with ϵ_{tag} (t) and Toys set with ϵ_{tag} = const. 223

LIST OF TABLES

Table 1 Main properties of elementary particles. 4
Table 2 Gauge bosons and Higgs particle in the Standard Model. Values taken from [3]. 5
Table 3 Current results for the CKM Unitarity Triangle angles from direct measurements and from the overall fit without including the direct measurements taken from [9]. 11
Table 4 Recent combinations of γ measurements by the LHCb, Belle and Babar collaborations. Results are taken from CKMFitter group [9] except LHCb* which is taken from [10] and still not used in the CKMfitter combination for γ. 12
Table 5 Integrated luminosity delivered by LHCb during 2010, 2011 and 2012 data taking. For each data taking period also the average number of interactions per bunch crossing (μ) and the peak luminosity are reported. 27

Table 6 Gas admixture, refractive index (and the wavelength which they correspond) and yield of photo-electrons for each track with $\beta = 1$. 39

Table 7 Typical L0 thresholds used in Run1. Table taken from [32]. 46

Table 8 Pre-selection requirements common to all tagging algorithms. All the cuts refer to the Reco14 optimisation. The track type are defined in Sec.3.2.1. The angle θ is defined with respect to the beam pipe; $|\Delta\phi|$ is the difference in the polar angle ϕ between the tagging track candidate and each reconstructed decay product of the signal B decay; The IP significance is computed with respect to the primary vertex associated to the signal B (PV) or to possible additional primary vertices (PU, pileup). The requirements on $\Delta\phi$ vetoes all reconstructed daughter particles of the signal B. 57

Table 9 Input variables to the NN1 for the discrimination between the fragmentation tracks and the UE tracks. The relative transverse momentum defined as $p_{\mathrm{T_{rel}}} = \sqrt{p^2_{\mathrm{track}} - p^2_{\mathrm{L_{rel}}}}$, where the longitudinal momentum of the tagging track is defined as $p_{\mathrm{L_{rel}}} = \vec{p}_{B^0_s} \cdot \vec{p}_{\mathrm{track}} / \left|\vec{p}_{B^0_s}\right|$, and the distance in the polar angle, ϕ, and the pseudorapidity, η, between the tagging track candidate and the B^0_s candidate, $\Delta\phi$ and $\Delta\eta$, respectively. 58

Table 10 OS tagging algorithms performance from optimization in $B^+ \to J/\psi K^+$ decay mode. 62

Table 11 Parameters of the PDF describing the signal (Double Crystal Ball) and the combinatorial background obtained from the fit to $B^0_s \to D^-_s \pi^+$ data for 2011, 2012 and 2011+2012 samples. Fixed parameters are obtained from a fit to simulated signal events. 71

Table 12 Yields of the background contribution in the mass range [5320:5600] MeV/c^{2}from the results of the fit to the [5100:5600] MeV/c^{2}mass range for 2011, 2012 and 2011+2012 samples. 72

Table 13 Parameters of the triple gaussian time resolution model determined from a fit to the $B_s^0 \to D_s^- \pi^+$ events. The scale factor represents the multiplicative factor applied to the sigmas in order to describe the resolution in data (see Ref. [59]). The average effective time resolution is reported in the last row. 74

Table 14 Results of the NNetSSK tagging calibration parameters p_0 and p_1 obtained by the unbinned and the binned fit to the decay time distribution of the 2011 and 2012 data samples. The tagging efficiencies ϵ_{tag} are computed before the calibration. 77

Table 15 Results of the simultaneous time fit on 2011+2012 data for the average mistag ω_{cat} and tagging efficiency ϵ_{tag} in bins of the predicted mistag probability η. 78

Table 16 Results of the NNetSSK tagging calibration parameters p_0 and p_1 obtained by the unbinned and the binned fit to the time distribution of the merged data sample (2011+2012). The result of the unbinned fit is considered as the reference calibration. The quoted tagging efficieiency is computed after the calibration. 79

Table 17 Selection cuts for prompt D_s candidates. 80

Table 18 Initial flavour asymmetry of calibration parameters and tagging efficiency, evaluated using MC samples of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to J/\psi\phi$ decays, and data sample of D_s decays re-weighted for the B_s^0 transverse momentum of $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J/\psi\phi$, $B_s^0 \to \phi\phi$ and $B_s^0 \to D_s D_s$ decays. We take as reference values the ones from the D_s data re-weighted for the $B_s^0 \to D_s^- \pi^+$ p_T distribution and obtained by fixing $\langle\eta\rangle$ to the reference value of $B_s^0 \to D_s^- \pi^+$ (0.4377); the first uncertainty is statistical; the second one is the systematic. The systematic uncertainty is defined as the maximum difference of the individual D_s results re-weighted to match the distribution of p_T or number of tracks observed in different channels with respect to the values corresponding to the re-weighting to the $B_s^0 \to D_s^- \pi^+$. 82

Table 19 Systematic uncertainties related to the fit model and calibration method of NNetSSK. 83

Table 20 Systematic uncertainties on the NNetSSK calibration related to the event properties and the B kinematics evaluated by means of a re-weighting procedure of simulated $B_s^0 \to D_s^- \pi^+$ events to match the $p_T(B_s^0)$, ϕ and the tracks multiplicity distributions of $B_s^0 \to J/\psi\phi$, $B_s^0 \to D_s^+ D_s^-$ and $B_s^0 \to \phi\phi$. 86

Table 21 NNetSSK Calibration in different MC samples. 87

Table 22 Final calibration parameters for the NNetSSK tagger in which the first uncertainty is statistical and the second one is systematic. Also Initial flavour asymmetry in the flavour tagging calibration is reported. 88

Table 23 Results of the simultaneous time fit on 2011+2012 $B_s^0 \to D_s^- \pi^+$ data for the average mistag ω_{cat} in bins of the predicted mistag probability η of the OS combination. 89

Table 24 Results of the time fit on 2011+2012 $B_s^0 \to D_s^- \pi^+$ data for the OS calibration parameters p_0 and p_1 using the unbinned and binned fit of the predicted mistag probability η. The reference calibration is the results of the unbinned fit. 90

Table 25 Systematic uncertainties of the OS combination related to the fit model for p_0 and p_1 from the $B_s^0 \to D_s^- \pi^+$ analysis. They are evaluated as the maximum difference with the reference values. 91

Table 26 Calibration parameters of the combined OS tagger extracted from different control channels. Statistic and systematic uncertainties are quoted. 91

Table 27 Final calibration parameters for the OS tagger in which the first uncertainty is statistical and the second one is systematic. Also Initial flavour asymmetry in the flavour tagging calibration is reported. 91

Table 28 MC samples used during the analysis for selection and background studies. The second column indicates the specific B-daughter decay and in the last column the main use of each sample is reported. 95

Table 29 Branching fractions of B_s^0 decays are taken from Ref. [73], those of D_s decays are taken from the PDG [62], while $f_s/f_d$is the ratio of hadronization factors which corresponds to the combination of the LHCb measurements using hadronic and semi-leptonic B meson decays [74]. 97

Table 30 Kinematic selection requirements and Vetos definition specific for each D_s mode. 99

Table 31 PID selection requirements. 100

Table 32 Parameters of the Double Crystal Ball function describing the signal B_s^0 shapes of $B_s^0 \to D_s^- \pi^+$, obtained from signal MC, separately for each D_s final state. Parameters without uncertainties are fixed in the fit. 104

Table 33 Parameters of the Double Crystal Ball function describing the signal B_s^0 shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC, separately for each D_s final state. 106
Table 34 Fully reconstructed backgrounds considered in the nominal fits to $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$. 112
Table 35 Constrained yields in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit. 112
Table 36 Constrained yields in the $B_s^0 \to D_s^- \pi^+$ MD fit. 112
Table 37 Definition of background groups in the $B_s^0 \to D_s^- \pi^+$ MD fit. 112
Table 38 Fitted values of the parameters for the $B_s^0 \to D_s^- \pi^+$ signal MD fit. The $slopes_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. 113
Table 39 Definition of the background groups in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit. 113
Table 40 Fitted values of the parameters for the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal MD fit. The $slopes_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. 117
Table 41 Calibration parameters of the combined OS tagger extracted from $B_s^0 \to D_s^- \pi^+$ decays. For a perfectly calibrated tagger one expects $p_1 = 1$ and $p_0 - \langle \eta \rangle = 0$. 123
Table 42 Calibration parameters of the OS tagger extracted from different control channels. Statistic and systematic uncertainties are quoted. 124
Table 43 The reference values for the OS tagger calibration are quoted, including the systematic uncertainties that account for differences measured in different channels. 124
Table 44 Calibration parameters of the NNetSSK tagger extracted from $B_s^0 \to D_s^- \pi^+$ decays. 125
Table 45 Calibration parameters of the NNetSSK tagger, based on the Neural Network classifier, extracted from $B_s^0 \to D_s^- \pi^+$ data. The first uncertainty is statistic and the second one is systematic. 126

Table 46 Initial flavour asymmetry for OS tagging extracted from $B^+ \to J/\psi K^+$ decays and for NNetSSK tagging extracted from prompt $D_s^\pm$ reweighting the p_T distribution to $B_s^0 \to D_s^- \pi^+$. 129

Table 47 The Flavour Tagging performances for only OS tagged, only SSK tagged and both OS and SSK tagged events for $B_s^0 \to D_s^- \pi^+$. 130

Table 48 Propagation of tagging calibrations uncertainties after combination of taggers. 131

Table 49 Tagging efficiencies and tagging efficiency asymmetries a^{tageff} for OS only, NNetSSK only and both OS and NNetSSK tagged events. 132

Table 50 Results of the time resolution study of the Δm_s analyses [59] and [61]. The effective mean time resolution, the scale factor with the relative uncertainties and the average dilution factor D_t are quoted. 133

Table 51 Parameters of the time acceptance from a fit to the time distribution of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ MC events and their ratio. 135

Table 52 Fitted values to the $B_s^0 \to D_s^- \pi^+$ time distribution for the *sFit* . 137

Table 53 Fitted values of the CP observables to the $B_s^0 \to D_s^\mp K^\pm$ time distribution for *sFit* . All parameters different than CP observables are constrained in the fit. 138

Table 54 Correlation matrix of the $B_s^0 \to D_s^\mp K^\pm$ CP *sFit* . Here we have restricted ourselves to correlations between the CP parameters for readability, as the tagging efficiencies all have negligible correlations with the CP parameters anyway. As can be seen the only large correlation is the 50% between D_f and $\overline{D}_f$, which is understandable because they are both sensitive to the slowly varying hyperbolic terms induced by $\Delta\Gamma_s$. 139

Table 55 Categorisation of the different physics backgrounds for the cFit. See text for an explanation of the categories. 140

Table 56 Fitted values of the CP observables to the $B_s^0 \to D_s^{\pm} K^{\mp}$ decay time distribution. The statistical and systematic uncertainties are also quoted. 143

Table 57 Systematic uncertainties related to the *sFit* results expressed as fractions of the statistical uncertainties. 144

Table 58 Systematic uncertainties related to the *cFit* results expressed as fractions of the statistical uncertainties. 144

Table 59 Pseudo-experiments for the 5 CP observables extracted from a simulated statistics of 3 and 1 fb^{-1} of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. 153

Table 60 Emulated backgrounds with the corresponding number of events and correlation factors (%) between MDFitter observables and time *sFit* ones. 155

Table 61 Emulated backgrounds with the corresponding number of events and correlation factors (%) among observables used in the MDFitter and in the time *sFit* separately. 156

Table 62 CP observables mean value and σ of the pull distributions for the toys set generated with a tagging efficiency which depends on the decay time and for the toys set generated with a constant tagging efficiency. 158

Table 63 CP observables mean value and σ of the pull distributions for the toy-by-toy difference among Toys set with ϵ_{tag} (t) and ϵ_{tag} = const. 158

Table 64 MC samples used during the analysis for selection, normalization, extraction of signal and background models for the different observables. The second column indicates the specific B-daughter decay. 164

Table 65 BDTG performance evaluated at the optimal cut value of $>$ 0.0. 167

Table 66 Comparison between signal yields and signal to background ratio in the updated analysis ($3fb^{-1}$) and in the previous one ($1fb^{-1}$). N_{bkg} corresponds to the sum of all the background contributions. 168

Table 67 Parameters of the Double Crystal Ball describing the signal B_s^0 shapes of $B_s^0 \to D_s^- \pi^+$, obtained from signal MC, separately for each D_s final state. 171

Table 68 Parameters of the Double Crystal Ball describing the signal B_s^0 shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC. Parameters without uncertainties are fixed in the fit to MC events. 172

Table 69 Parameters of the Double Crystal Ball describing the signal D_s shapes of $B_s^0 \to D_s^- \pi^+$, obtained from signal MC, separately for each D_s final state. 173

Table 70 Parameters of the Double Crystal Ball describing the signal D_s shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC, separately for each D_s final state. Parameters without uncertainties are fixed in the fit to MC events. 175

Table 71 Constrained yields in the $B_s^0 \to D_s^{\mp} K^{\pm}$ Multi-dimensional fit. 177

Table 72 Constrained yields in the $B_s^0 \to D_s^- \pi^+$ Multi-dimensional fit. 177

Table 73 Fitted values of the parameters for the $B_s^0 \to D_s^- \pi^+$ signal MD fit. Means are the parameters of the double Crystal Ball used to describe the signal in B_s^0 and D_s masses. The $slope_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. The fractions $fComb_{B_s}$ are the fractions between an exponential and flat distribution in B_s^0 mass, whereas $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. The $fComb_{PIDK}$ is a fraction between pion and kaon components in the PIDKcombinatorial shape. 183

Table 74 Fitted values of the parameters for the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal MD fit. Means are the parameters of the double Crystal Ball used to describe the signal in B_s^0 and D_s masses. The $slope_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. The fractions $fComb_{B_s}$ are the fractions between an exponential and flat distribution in B_s^0 mass, whereas $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. $fComb_{PIDK(1,2)}$ are a fraction between pion, kaon and proton components in the $PIDK$ combinatorial shape, shared among all subsamples. 188
Table 75 Results of the OS and NNetSSK tagging calibration parameters p_0 and p_1 obtained on the $B_s^0 \to D_s^- \pi^+$ Run1 data sample. 194
Table 76 Toys studies results for the calibration parameters; 500 samples generated. 203
Table 77 Toys studies results for the calibration parameters; 500 samples generated (signal only). 203
Table 78 Bootstrap studies results for the calibration parameters; 10K bootstrapped samples. 206
Table 79 Emulated Backgrounds with the corresponding expected number of events. 219

1

INTRODUCTION

The Standard Model successfully describes the fundamental interactions of elementary particles, except gravity. The violation of the CP symmetry, i.e. charge conjugation and parity, in the weak interactions is depicted by the Standard Model through the Cabibbo-Kobayashi-Maskawa (*CKM*) paradigm and the parameters which naturally accounts for the CP violation are the sides and the angles of the *CKM* Unitarity Triangle. The CP violation amount accounted in the SM description is not enough to explain the asymmetry between matter and anti-matter observed in the universe. For this reason important efforts are put in place to improve the knowledge and the precision of the CP violation measurements and to measure the relevant parameters of the *CKM* unitarity triangle which are not predicted by the theory. A deep investigation of the CP violation sector is particularly promising as it can indirectly reveal effects due to physics beyond the Standard Model. Through the production of virtual particles in loop transitions it is possible to test higher energy scales than the one of the particles which are directly produced. The LHCb experiment located at CERN is designed to study beauty and charm hadrons decays, in particular to perform precise measurements of CP violation effects and to search for physics beyond the Standard Model through B mesons rare decays.

Indeed weak decays containing heavy quarks, as b and c, are extremely suitable to measure Standard Model parameters and to test its validity. More specifically, B mesons decays represent the most direct way to measure the parameters of the theory which are related to CP violation effects since they predominantly decay through weak processes. One of the least well measured *CKM* parameter is the angle γ of the Unitarity Triangle (UT). Babar and Belle Collaborations have already measured it, and now LHCb is starting to perform the most precise measurements. Very recently the LHCb collaboration has measured γ for the first time, in the decay mode $B_s^0 \to D_s^{\mp} K^{\pm}$ by means of a tagged time-dependent analysis [1]. This decay is dominated by tree-level diagrams (loop contributions to this decay are negligible) allowing to perform precise measurements not being limited by the theoretical uncertainties related to loop processes. The sensitivity to γ arises from the interference between $b \to c$ and $b \to u$ transitions in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays since they show CP violation effect in the interference between decays with $B_s^0 - \overline{B}_s^0$ mixing and without. A crucial ingredient of the analysis is the need to tag the flavour of the B_s^0 and $\overline{B}_s^0$ mesons at production time. LHCb has developed several Flavour Tagging algorithms. One of these, the NNetSSK, is particularly relevant for B_s^0 decays and thus for the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis. Details on the development and calibration of such algorithms in the $B_s^0 \to D_s^- \pi^+$ decays are presented in this thesis, followed by the description of the first measurement of γ from a time-dependent analysis in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays using the dataset collected by LHCb during Run1.

This thesis is organized as follows: in Chapter 2 an introduction to the physics phenomena depicted in this dissertation is given, followed by the description of the LHCb experimental apparatus in Chapter 3. The Flavour Tagging tool is introduced in Chapter 4 and the study of the calibrations in the $B_s^0 \to D_s^- \pi^+$ is described in Chapter 5. In Chapter 6 the first measurement of γ in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays using the data set collected by LHCb in the 2011 is presented. The Chapter 7 is dedicated to the updated $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis for measuring γ exploiting the full Run1 data set. In Chapter 8 a summary of the results is given.

2

PHYSICS INTRODUCTION

Contents

2.1 The Standard Model of Particle Physics 3
2.2 CP violation in the Standard Model 6
2.2.1 The CKM formalism 7
2.2.2 The Unitarity triangles 9
2.3 B meson Physics 12
2.3.1 Mixing of neutral B mesons 13
2.3.2 Time evolution of flavour eigenstates 15
2.4 Different types of CP violation 17
2.5 Measurement of γ through B decays 18
2.5.1 Tree-level decays to measure γ 20
2.6 Measurement of the CKM angle γ in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays 23

The chapter aims to provide an introduction to the physics phenomena depicted in this dissertation. After a brief summary of the Standard Model concept of particle physics, a description of CP violation in the Standard Model is given. Particular relevance is given to the phenomenological model for beauty (B) mesons decays, which can manifest CP violation in three different ways. Finally an experimental overview on the CP violation measurements related to the angle γ is provided, focusing on the time-dependent measurement of γ in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays.

2.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model (SM) [2] of particle physics describes the fundamental entities, quarks and leptons, of which all matter is composed and it depicts the fundamental forces which governs nature, except gravity. Either quarks and leptons are fermions, i.e. half-integer spin (1/2) particles.

- Quarks are characterized by 6 flavours with fractional electric charge: up (u), charm (c), top (t): $+2/3e$ and down (d), strange (s), beauty (b): $-1/3e$, where e is the electric charge of the electron. Quarks carry color charge (red, blue or green) of *Quantum Chromodynamics* (QCD) and they are bound by the strong interaction in color singlet called hadrons. Hadrons are distinguished in *baryons*, bound states of three quarks (qqq), and *mesons*, composed by a couple of quark and anti-quark ($q\bar{q}$)[1].

- Leptons are distinguished in three families which relate a charged particle and a neutral one called neutrino: (electron e, ν_e), (muon μ, ν_μ) and (tau τ, ν_τ).

Each fermion has an anti-particle which has exactly the same mass and spin but it is characterized by opposite quantum numbers. The basic properties of the elementary particles in the Standard Model are summarized in Tab.1. Elementary particles exhibit an increasing mass moving from the first generation to the third one.

Table 1.: Main properties of elementary particles.

Name	Spin[$\hbar$]	Generation						Electric Charge	Colour Charge
		1^{st}	mass	2^{nd}	mass	3^{rd}	mass		
Quarks	$\pm\frac{1}{2}$	u	1.8-3.0 MeV	c	1.27 GeV	t	173 GeV	$\frac{2}{3}e$	red, blue, green
	$\pm\frac{1}{2}$	d	4.5-5.3 MeV	s	95 MeV	b	4.18 GeV	$-\frac{1}{3}e$	red, blue, green
Leptons	$\pm\frac{1}{2}$	e	0.511 MeV	μ	105.7 MeV	τ	1.78 GeV	$-e$	-
	$\pm\frac{1}{2}$	ν_e	< 2 eV	ν_μ	< 2 eV	ν_τ	< 2 eV	0	-

The Standard Model describes the interactions between the elementary particles merging the description of the strong, the electromagnetic and the weak interactions via the direct product of the $SU(3)_C \otimes SU(2)_I \otimes U(1)_Y$ gauge groups. The interactions between the elementary particles take place through the exchange of integer-spin particles called *Gauge bosons* while the mass of the particles is given by the interaction with the *Higgs boson* for both quarks and leptons. They are summarized in Tab.2.

1 More complex states such as tetra-quarks and penta-quarks are also allowed and recently discovered.

Table 2.: Gauge bosons and Higgs particle in the Standard Model. Values taken from [3].

Name	Spin [$\hbar$]	Mass [GeV/c^2]	Mediated Force
Gluon g	1	0	strong
Photon γ	1	0	electromagnetic
W^+, W^-	1	80.39 ± 0.02	weak (charged current)
Z^0	1	91.188 ± 0.002	weak (neutral current)
Higgs particle H^0	0	125.09 ± 0.24	-

Strong interactions are described by the $SU(3)_C$ gauge symmetry of the QCD, where C indicates the color charge and it defines the strong couplings among quarks. The generators of such non abelian group are 3×3 matrices and the 3 eigenvalues are labelled as red, blue and green. The gluons (see Tab.2) are mediators of the strong interaction and they are 8 bi-colored massless bosons. Hadrons are color singlet and this is due to color confinement. The confinement of the color charge, within a distance exceeding the nucleon radius, is due to the fact that gluons have color charge and thus self-interact. The electromagnetic and weak interactions are unified in a common theoretical description via the direct product of $SU(2)_I \otimes U(1)_Y$ gauge groups where I and Y stay for the weak Isospin and Hypercharge symmetries. The SU(2) invariance requires three vector boons, W_1, W_2 and W_3, and and U(1) one , B^0, boson. The $SU(2)_I \times U(1)_Y$ symmetry is spontaneously broken by the Higgs mechanism. The Spontaneous Symmetry Breaking (SSB) is operated by the Higgs field which is represented by a scalar, complex doublet, in which the ground state is not invariant under the $SU(2)_I$ and $U(1)_Y$ symmetries, but it is for the electric charge operator, described by a $U(1)_Q$ gauge group. According to the Goldstone's theorem 3 vector bosons gain a mass, while the generator of the $U(1)_Q$ remains massless. In this framework the three massive vector bosons are W^+ and W^- and the Z^0. The two charged bosons arise as a combination of W_1 and W_2. The neutral one is a linear combination of the third $SU(2)_I$ boson with the $U(1)_Y$ boson B^0. The remaining massless boson corresponds to the photon which is a particular linear combination of W_3 and B^0 bosons. The Higgs field permeates the whole space and it is associated to a particle called Higgs boson discovered in

2012 by ATLAS [4] and CMS [5]. The masses of fundamental fermions arise from the interaction with the Higgs boson.

2.2 CP VIOLATION IN THE STANDARD MODEL

The violation of the CP symmetry, i.e. charge conjugation and parity, is one of the SM sector which is currently under deep investigation. The first evidence of parity violation in weak interaction arose from an experiment in 1956 on beta decay of ^{60}Co nuclei (Wu et al.,1957). It implies that in nature only left-handed (right-handed) neutrino (anti-neutrino) exists, while a parity transformation, which means a space coordinates inversion, would lead a left-handed neutrino into a right-handed neutrino. The same kind of violation takes place under charge conjugate transformations which transforms a particle into its anti-particle. In 1957 L.D. Landau proposed that the combined C and P symmetries should be considered as an invariant for weak interactions. In 1964 James Cronin e Val Fitch found experimentally a CP violation effect in neutral kaons decays which was also observed for B mesons in 2001[2]. Provided that the CPT theorem [6] holds for a Lorentz invariant quantum field theory with a hermitian hamiltonian, the CP transformation is equivalent to T transformation. This states definitely that the weak interaction is not invariant neither for time reversal transformation T whose violation has been first observed in the K system. The direct observation of the violation of the T symmetry in the B^0 has been made by the Babar collaboration [7]. The Standard Model of particle physics includes this effects in the description of weak interaction. CP conjugates processes involving weak interactions occur through the exchange of charged vector bosons $W^{\pm}$, which are the interaction carriers of the weak charged current. The weak interaction acts in different way for left-handed and right-handed particles and weak charged current interaction couples only to

2 B meson are mesons that contain a b or $\bar{b}$ and another quark of the 1^{st} or 2^{nd} family.

left-handed particles (L label), in particular for quarks it can be described in the following way:

$$-\frac{g_w}{\sqrt{2}}\left(\bar{u}_L,\bar{c}_L,\bar{t}_L\right)\gamma^{\mu}W_{\mu}^{+}V_{CKM}\begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}+h.c. \tag{2.1}$$

plus the hermitian conjugate term.

The V_{CKM} term describes the mixing between quarks in the weak interaction. The g_w parameter is the coupling constant of weak interaction, W_{μ}^{+} corresponds to the field of the $W^{\pm}$ boson and γ^{μ} are the Dirac's matrices in four dimensional space-time.

2.2.1 *The CKM formalism*

The couplings between quarks are described by a 3×3 unitarity matrix V_{CKM} called the *Cabibbo-Kobayashi-Maskawa* matrix (CKM matrix). The CKM matrix acting on the physical quarks (d,s,b), the ones involved in the other fundamental interactions, represents the quarks as weak eigenstates (d',s',b'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM}\begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}\begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.2}$$

In other words, the weak eigenstates are linear combinations of the physical quarks. A 3×3 complex matrix has 18 parameters, which reduce to 9 thanks to the unitarity condition ($\mathbf{V}^{\dagger}\mathbf{V}$=$\mathbf{V}\mathbf{V}^{\dagger}$ = $\mathbb{1}$). The 6 quark phases are free and they can be factorized out as relative phases remaining with one global phase. In this way the independent parameters of the V_{CKM} can be described by four independent parameters: three real angles describe the quark mixing and an irreducible, non null, imaginary phase accounts for the CP-violation effect. Notice that in case of a model with two generation of quarks, no CP violation could be described by the CKM paradigm. Within the Standard Model framework the quark mixing, i.e. the possibility to change the quark flavour and the CP-violation effect are embedded in the weak charge current interaction. A common way to display the V_{CKM} elements dependence on the free parameters is the following:

$$\mathbf{V_{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2.3}$$

where c_{ij} and s_{ij} indicate the cosine and the sine of the three real angles θ_{12}, θ_{23} and θ_{13} which connect different quark generations. The δ parameter is the complex phase responsible of CP violation effects. The diagonal elements represent transition between quarks within the same family and their amplitudes are close to one. Moving from $1 \to 2$, $2 \to 3$ and $1 \to 3$ family transition the amplitude magnitude becomes smaller. This sort of rank can be visualized adopting a different parametrization called Wolfenstein parametrization to express the matrix element. Indeed expanding in powers of the small parameter $\lambda = s_{12} = \sin\theta_{12}$[3] $= 0.22$ the V_{CKM} matrix is the following:

$$\mathbf{V_{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.4}$$

defining:

$$s_{12} = \lambda = \frac{|\ V_{us}\ |}{\sqrt{|\ V_{ud}\ |^2 + |\ V_{us}\ |^2}}, \tag{2.5}$$

$$s_{23} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right|, \tag{2.6}$$

$$s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta). \tag{2.7}$$

The parameters of the CKM matrix are not predictable by the model but they are experimentally determined. In Eq.2.8 the best estimates of the moduli of CKM matrix elements are reported:

$$\mathbf{V_{CKM}} = \begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22520 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886^{+0.00033}_{-0.00032} & 0.0405^{0.0011}_{0.0012} & 0.99914 \pm 0.00005 \end{pmatrix} \tag{2.8}$$

values are taken from [3].

3 θ_{12} is the Cabibbo angle θ_c

2.2.2 *The Unitarity triangles*

The unitarity of the CKM matrix, translates into 9 orthogonality conditions among its matrix elements:

$$\sum_{i=0}^{3} |V_{ij}|^2 = 1 \quad \text{con} \quad j = 1,2,3 \tag{2.9}$$

$$\sum_{i=0}^{3} V_{ji} V_{ki}^* = 0 = \sum_{i=0}^{3} V_{ij} V_{ik}^* \quad j,k = 1,2,3, \quad j \neq k \tag{2.10}$$

For the off-diagonal terms 6 orthogonality conditions are:

$$\begin{aligned}
V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad & (\lambda^3, \lambda^3, \lambda^3)(db) \\
V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad & (\lambda^4, \lambda^2, \lambda^2)(sb) \\
V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad & (\lambda, \lambda, \lambda^5)(ds) \\
V_{ud}V_{td}^* + V_{cs}V_{ts}^* + V_{ub}V_{tb}^* = 0 \quad & (\lambda^3, \lambda^3, \lambda^3)(ut) \\
V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \quad & (\lambda^4, \lambda^2, \lambda^2)(ct) \\
V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \quad & (\lambda, \lambda, \lambda^5)(uc),
\end{aligned} \tag{2.11}$$

which can be represented as triangles in the complex plane. The area of all these triangles is the same and the strength of the CP-violation. It corresponds to half of the Jarlskog invariant:

$$|J_{CP}| = |Im[V_{ij}V_{kl}V_{ij}^*V_{kj}^*]|, \quad i \neq k, j \neq l \tag{2.12}$$

whose small size quantifies the strength of the CP violation in the SM, $J_{PC} \approx A^2\lambda^6\eta \approx O(10^{-5})$. Substituting to the matrix elements the corresponding Wolfenstein parametrization expression in Eq 2.11, immediately appears that only the triangles labelled (db) and (ut) have the sides of the same order, while all the others look substantially degenerates. The triangle involving transitions with the top quark, the (ut) one, is difficult to be experimentally accessible due to the fact that t quarks decay before forming an hadron. The Unitarity Triangle (UT) which is experimentally more interesting from the point of view of CP violation measurements is the first one, labelled as (db). The involved matrix elements are experimentally accessible in B mesons decays. Including also the λ^5 terms given the increased precision of current experiments, each side of the triangle can be expressed in terms of the Wolfenstein parameters:

- $V_{ub}^* V_{ud} = A\lambda^3(\rho + i\eta)(1 - \frac{\lambda^2}{2})$
- $V_{cb}^* V_{cd} = -A\lambda^3$
- $V_{tb}^* V_{td} = A\lambda^3((1 - \rho - i\eta) + \lambda^2(\rho + i\eta))$.

In general, each side is then normalized by the $A\lambda^3$ factor, leading to a representation in terms of the $\bar{\rho} = \rho(1 - \frac{\lambda^2}{2})$ and $\bar{\eta} = \eta(1 - \frac{\lambda^2}{2})$ parameters. The angles of this unitarity triangle are then defined as follows:

$$\alpha = \Phi_1 = arg(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}), \tag{2.13}$$

$$\beta = \Phi_2 = arg(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}), \tag{2.14}$$

$$\gamma = \Phi_3 = arg(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}). \tag{2.15}$$

Fig.1 shows the normalized Unitarity Triangle in term of the variables $(\bar{\rho}, \bar{\eta})$. The

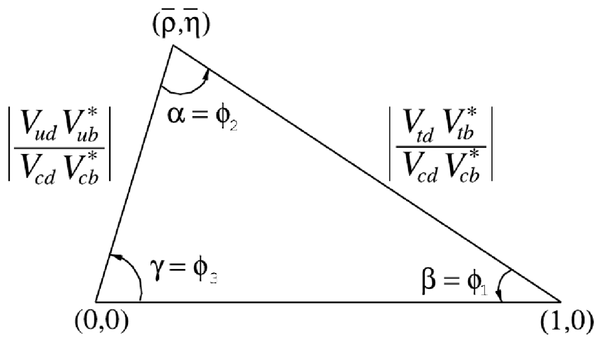


Figure 1.: Representation of the Unitarity Triangle in the complex plane.

direct measurement of sides and angles of the Unitarity Triangle is a crucial test of the SM and comparisons between the different evaluations can reveal the presence of new phenomena beyond the SM itself.

The combination of all the current experimental results and theoretical predictions (from Lattice QCD) related to the Unitarity Triangle parameters is depicted in Fig.2. Indeed, to determine the *CKM* parameters, several measurements of lifetimes and branching fraction ratios (in *B* and *K* decays) as well as the mixing

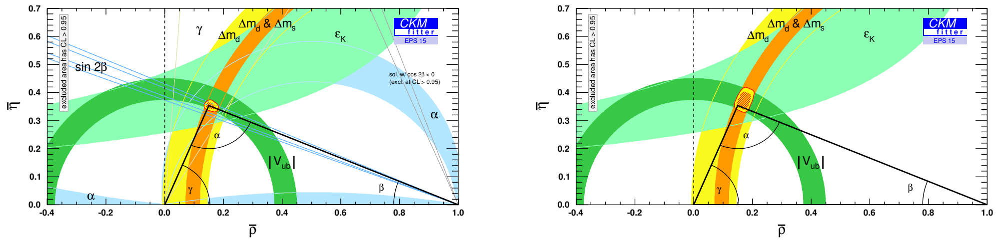


Figure 2.: Current constraints on the Unitarity Triangle parameters in the complex plane: including direct measurements of angles (left), not including them (right) [9].

frequencies ($\Delta m_{d(s)}$) and the mixing phases of B mesons are needed. In addition theoretical estimations from lattice QCD are also needed. The overall fit to determine the CKM parameters is performed using averaging the direct measurements of the Unitarity Triangle angles, α, β and γ (Fig.2 left) performed by different collaborations. The fit for the CKM parameters can be done letting the angles of the UT to be determined by the overall fit (Fig.2 right). In Table 3 the values from direct measurements of the CKM angles are reported. The poor precision of the direct measurement of γ makes the comparison with the overall fit determination not affordable.

Table 3.: Current results for the CKM Unitarity Triangle angles from direct measurements and from the overall fit without including the direct measurements taken from [9].

The CKM angle	Direct measurements [°]	Result from the overall fit [°]
α	$87.6^{+3.5}_{-3.3}$	$91.5^{+4.2}_{-1.3}$
β	$21.85^{+0.68}_{-0.67}$	
γ	$73.2^{+6.3}_{-7.0}$	$66.9^{+1.0}_{-3.7}$

Several collaboration as Belle at KEK (Japan), Babar at SLAC (USA) and LHCb at CERN (Switzerland) have contributed to the study of the SM in the B sector with excellent results. In particular they have measured γ and the average values are reported in Tab.4.

Table 4.: Recent combinations of γ measurements by the LHCb, Belle and Babar collaborations. Results are taken from CKMFitter group [9] except LHCb* which is taken from [10] and still not used in the CKMfitter combination for γ.

Experiment	Combination of γ measur. [°]
LHCb	$74.6^{+8.4}_{-9.2}$
LHCb*	$70.9^{+7.1}_{-8.5}$
Belle	70 ± 18
Babar	73 ± 13

The current knowledge of the Unitarity Triangle does not indicate evidence of discrepancy with the Standard Model but the CP violation sector can be the right place to get indirect evidence of Physics phenomena beyond the Standard Model. Indeed the CP violation amount accounted in the SM description is not enough to explain the asymmetry between matter and anti-matter observed in the universe. For this reason important efforts are put in place to improve the precision of the CP violation measurements and of CKM matrix. In particular being γ the least known angle, more precise measurements are needed. The combination of the LHCb measurements of γ is described in this thesis in Sec.6.10.

2.3 B MESON PHYSICS

Weak decays of hadrons containing heavy quarks, as b and c, are extremely suitable to measure Standard Model parameters and to test its validity. More specifically, they represent the a direct way to measure Unitarity Triangle parameters since they show CP violation effects. B mesons are particles made by a b or an $\bar{b}$ quark bounded with another lighter quark (c,u, d or s). They are relatively heavy ($m_B \simeq 5GeV$) and have relatively long lifetime ($\tau_B \simeq 1.5ps$) since they mainly decay through weak interactions. The following sections are dedicated to illustrate properties of neutral B mesons.

2.3.1 *Mixing of neutral B mesons*

Flavour oscillation of neutral B mesons has been observed in 2001 and it relies on the non conservation of flavour in the weak interaction. B^0_q and its antiparticle $\overline{B}^0_q$ $(q = d, s)$

$$B^0_q = |\bar{b}q>, \overline{B}^0_q = |b\bar{q}> \tag{2.16}$$

differ only by the flavour quantum number F, $\Delta F = 2$. The $B^0_q \leftrightarrow \overline{B}^0_q$ transition can take place through two $|\Delta F| = 1$ transitions:

$$B^0_q \xrightarrow{\Delta F=1} I \xrightarrow{\Delta F=1} \overline{B}^0_q \quad \text{or} \quad \overline{B}^0_q \xrightarrow{\Delta F=-1} I \xrightarrow{\Delta F=-1} B^0_q$$

where I is a common intermediate state. The mixing takes place through a second-order diagram with the exchange of two $W^{\pm}$ bosons where the intermediate state is composed mainly by the top quark (virtual state). In Fig.3 the dominant box diagrams responsible for $B^0_q - \overline{B}^0_q$ mixing are shown.

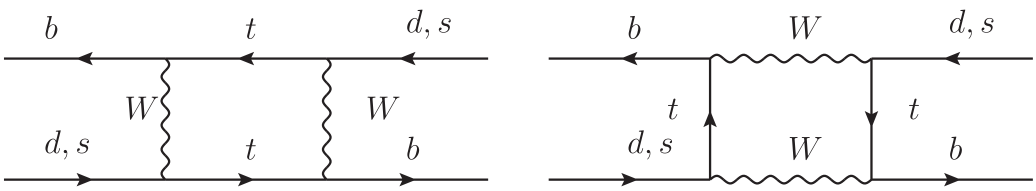


Figure 3.: Dominant box diagrams for the $B^0_q - \overline{B}^0_q$ transition ($q = d$ or s).

Let's consider the initial state $|\psi(t = 0)> = a(0)|B^0_q> + b(0)|\overline{B}^0_q>$. The time evolution of this initial state can introduce other possible final states in addition to $|B^0_q>$ and $|\overline{B}^0_q>$ but, since we are interested in the flavour mixing, we can focus ourselves on the values of $a(t)$ and $b(t)$ neglecting the other possible final states.

The time evolution is described by the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\psi(t) = H\psi(t) \tag{2.17}$$

where

$$\psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{2.18}$$

The hamiltonian operator can be written as

$$H = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} = \begin{pmatrix} M_0 - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_0 - \frac{i}{2}\Gamma \end{pmatrix}$$

where the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$ are 2×2 hermitian matrices.

The CPT invariance leads to:

- $M_{11} = M_{22} = M_0$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$.
- $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$.

The weak eigenstates are determined diagonalizing the H matrix and they can be generically written as linear combinations of the considered flavour eigenstates:

$$|B_L>= p|B_q^0> +q|\overline{B}_q^0>, \quad |B_H>= p|B_q^0> -q|\overline{B}_q^0>, \tag{2.19}$$

where $|p|^2 + |q|^2 = 1$ and

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}. \tag{2.20}$$

The eigenvalues are:

$$\lambda_i = \left(M_0 \pm \frac{\Delta m}{2}\right) - \frac{i}{2}\left(\Gamma \pm \frac{\Delta\Gamma}{2}\right) \tag{2.21}$$

with $i = H, L$. The two eigenvalues $\lambda_{L,H}$ are related to the mass and width differences as follows:

$$M_0 = \frac{M_H + M_L}{2} \tag{2.22}$$

$$\Delta m \equiv m_H - m_L = Re(\lambda_H - \lambda_L) \tag{2.23}$$

$$\Gamma = \frac{\Gamma_H + \Gamma_L}{2} \tag{2.24}$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L = -2Im(\lambda_H - \lambda_L). \tag{2.25}$$

The time evolution of weak mass eigenstates can be written as:

$$|B_L(t)>= e^{-i\lambda_L t}|B_L>, \quad |B_H(t)>= e^{-i\lambda_H t}|B_H> . \tag{2.26}$$

For B_d^0 mesons $\Delta\Gamma$ is approximately zero while for B_s^0 mesons $\Delta\Gamma_s \neq 0$, and LHCb measured it to be of $(0.106 \pm 0.013) ps^{-1}$ [72].

2.3.2 *Time evolution of flavour eigenstates*

Starting from a pure matter state:

$$|B_q^0(t=0)>=\frac{1}{2p}(|B_L(t=0)>+|B_H(t=0)>) \tag{2.27}$$

it evolves in time using Eq.2.26-2.19 as following:

$$\begin{aligned}|B_q^0(t)> &= \frac{1}{2p}(|B_L(t)>+|B_H(t)>) = \frac{1}{2p}(e^{-i\lambda_L t}|B_L(0)>+e^{-i\lambda_H t}|B_H(0)>) \qquad (2.28)\\ &= \frac{1}{2p}(e^{-\lambda_L t}(p|B_q^0>+q|\overline{B}_q^0>)+e^{-\lambda_H t}(p|B_q^0>-q|\overline{B}_q^0>))\\ &= g_+(t)|B_q^0>+\frac{q}{p}g_-(t)|\overline{B}_q^0>,\end{aligned}$$

where:

$$g_\pm(t)=\frac{1}{2}\left[e^{-i\lambda_L t}\pm e^{-i\lambda_H t}\right]. \tag{2.29}$$

An anti-matter component arises in the time evolution of a pure matter initial state. Due to mixing both the B_q^0 and the $\overline{B}_q^0$ initial states can decay into the same final state f and its charge conjugate $\bar{f}$. The decay amplitudes are given by:

$$A_f=<f|H|B_q^0>,\quad \overline{A}_f=<f|H|\overline{B}_q^0>,\quad A_{\bar{f}}=<\bar{f}|H|B_q^0>,\quad \overline{A}_{\bar{f}}=<\bar{f}|H|\overline{B}_q^0> \tag{2.30}$$

where the CPT invariance assures that:

$$|\overline{A_{\bar{f}}}|=|A_f|,\quad |\overline{A}_f|=|A_{\bar{f}}|. \tag{2.31}$$

All the necessary ingredients to describe the decay of a B_q^0 meson produced at time $t=0$ to a final state f at time t have been already provided, so that the corresponding decay rate can be written as:

$$\begin{aligned}\frac{d\Gamma_{B_q^0\to f}(t)}{dt} &= |\left\langle f|H|B_q^0(t)\right\rangle|^2 = |g_+(t)A_f+\frac{q}{p}g_-(t)\bar{A}_f|^2\\ &= |A_f|^2\left(|g_+(t)|^2+|\lambda_f|^2|g_-(t)|^2+\lambda_f^* g_+(t)g_-^*(t)+\lambda_f g_+^*(t)g_-(t)\right)\\ &= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Big[(1+|\lambda_f|^2)\cosh(\frac{\Delta\Gamma}{2}t)+\\ &\qquad (1-|\lambda_f|^2)\cos(\Delta m t)-2Re(\lambda_f)\sinh(\frac{\Delta\Gamma}{2}t)-2Im(\lambda_f)\sin(\Delta m t)\Big].\end{aligned} \tag{2.32}$$

Similarly the decay rate for $\overline{B}_q^0 \to f$ is:

$$\begin{aligned}\frac{d\Gamma_{\overline{B}_q^0\to f}(t)}{dt} &= |\left\langle f|H|\overline{B}_q^0(t)\right\rangle|^2 \\ &= \frac{1}{2}|A_f|^2|\frac{p}{q}|^2 e^{-\Gamma t}\left[(1+|\lambda_f|^2)\cosh(\frac{\Delta\Gamma}{2}t)-\right. \\ &\left.(1-|\lambda_f|^2)\cos(\Delta mt)-2Re(\lambda_f)\sinh(\frac{\Delta\Gamma}{2}t)+2Im(\lambda_f)\sin(\Delta mt)\right].\end{aligned} \tag{2.33}$$

The parameter λ_f is defined as:

$$\lambda_f = \frac{1}{\overline{\lambda}_f} = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{2.34}$$

The decay rates to the $\overline{f}$ final state, $B_q^0 \to \overline{f}$ and $\overline{B}_q^0 \to \overline{f}$, are described by the same equations of 2.32 and 2.33 substituting the parameter λ_f with:

$$\lambda_{\overline{f}} = \frac{1}{\overline{\lambda}_{\overline{f}}} = \frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}}. \tag{2.35}$$

Factorizing out the $(1+|\lambda_f|^2)$ in Eq.2.32 and 2.33, it is possible to highlight the CP asymmetry observables as coefficients of the trigonometric functions:

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

and for the $B_q^0(\overline{B}_q^0) \to \overline{f}$ decays:

$$C_{\bar{f}} = C_f \Rightarrow CPT\ invariance, \quad S_{\bar{f}} = \frac{2Im(\lambda_{\bar{f}})}{1+|\lambda_{\bar{f}}|^2}, \quad A_{\bar{f}}^{\Delta\Gamma} = \frac{2Re(\lambda_{\bar{f}})}{1+|\lambda_{\bar{f}}|^2}. \tag{2.37}$$

Flavour specific decays

A particular type of B meson decay is the so called *flavour specific decay*. In these decays the final states f can be reached only by the B and not by the $\overline{B}$; viceversa for the $\overline{f}$ final state. As a consequence, for flavour specific decays $\overline{A}_f = A_{\overline{f}} = 0$ and $\lambda_f = \overline{\lambda}_f = 0$; so that in the decay rate only the cosine and the hyperbolic cosine survive:

$$\frac{d\Gamma_{B_q^0\to f}(t)}{dt} = \frac{1}{2}|A_f|^2 e^{-\Gamma t}\left[\cosh(\frac{\Delta\Gamma}{2}t)+\cos(\Delta mt)\right] \tag{2.38}$$

$$\frac{d\Gamma_{\overline{B}_q^0\to \overline{f}}(t)}{dt} = \frac{1}{2}|A_f|^2|\frac{p}{q}|^2 e^{-\Gamma t}\left[\cosh(\frac{\Delta\Gamma}{2}t)-\cos(\Delta mt)\right]. \tag{2.39}$$

Assuming that the CP asymmetry in the B_q^0 mixing is negligible ($|q/p| = 1$) the mixing asymmetry can then be written as:

$$\mathcal{A}^{mix}(t) = \frac{d\Gamma/dt(B_q^0 \to f(t)) - d\Gamma/dt(\overline{B}_q^0 \to \overline{f}(t))}{d\Gamma/dt(B_q^0 \to f(t)) + d\Gamma/dt(\overline{B}_q^0 \to \overline{f}(t))} = \frac{cos(\Delta m_q t)}{cosh(\frac{\Delta\Gamma_q}{2}t)}. \tag{2.40}$$

This expression represents the theoretical mixing asymmetry in the hypothesis that the determination of the production B meson flavour given the final state f or $\overline{f}$ is perfect. In general this is not the case because of the B_q^0 mixing or because of wrong identification of the final state particles. The modifications of the mixing asymmetry due to experimental effects will be presented in the following Chapters.

2.4 DIFFERENT TYPES OF CP VIOLATION

CP violation phenomena have been observed so far in Kaons and more recently in the B meson systems and they are driven by a phase which enter in the decay amplitude, as described in Sec.2.2.1. To be sensitive to the CP violation phase it is necessary to analyze a decay in which at least two amplitudes which behave differently under CP transformation contribute:

$$|A|^2 = |A_1 + e^{i\psi}A_2|^2 \xrightarrow{CP} |A_1 + e^{-i\psi}A_2|^2 = |\overline{A}|^2. \tag{2.41}$$

In this way the squared modulo of the decay amplitudes for the decay and its CP conjugate depends on the CP-violating phase:

$$|A|^2 = |A_1|^2(1 + \frac{|A_2|^2}{|A_1|^2} \pm \frac{|A_2|}{|A_1|}\cos(\psi)). \tag{2.42}$$

Such interference can occur in different process leading to three different categories of CP violating processes:

- in decays, also called direct CP-violation
- in mixing, also called indirect CP-violation
- in the interference between decay and mixing.

Direct CP violation occurs for neutral and charged mesons, while CP violation in mixing and in the interference between decay and mixing can take place only for neutral mesons. CP violation in decays occurs if the probability of $B_q^0 \to f$ decay is different from the $\overline{B}_q^0 \to \overline{f}$ one: $\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| \neq 1$. The CP asymmetry in decay can be measured in several decays ($B^+ \to DK^+$, $B^0 \to DK^{*0}$, $B \to K\pi$ etc) through:

$$A_{CP} = \frac{\Gamma(B_q^0 \to f) - \Gamma(\overline{B}_q^0 \to \overline{f})}{\Gamma(B_q^0 \to f) + \Gamma(\overline{B}_q^0 \to \overline{f})} = \frac{|A(B_q^0 \to f)|^2 - |A(\overline{B}_q^0 \to \overline{f})|^2}{|A(B_q^0 \to f)|^2 + |A(\overline{B}_q^0 \to \overline{f})|^2}$$
$$= \frac{1 - |\overline{A}_{\overline{f}}/A_f|^2}{1 + |\overline{A}_{\overline{f}}/A_f|^2} \neq 0. \tag{2.43}$$

CP violation in mixing takes place when the oscillation probability $B_q^0 \to \overline{B}_q^0$ is different from $\overline{B}_q^0 \to B_q^0$. This means that $|q/p| \neq 1$ and $|\lambda_f| \neq 1$ and it can be measured in flavour specific decays as the semi-leptonic decays:

$$A_{CP} = \frac{\Gamma(B_q^0 \to l^-\bar{\nu}X) - \Gamma(\overline{B}_q^0 \to l^+\nu X)}{\Gamma(B_q^0 \to l^-\bar{\nu}X) + \Gamma(\overline{B}_q^0 \to l^+\nu X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \neq 0. \tag{2.44}$$

CP violation in $B_{d(s)}^0$ mixing has not been observed yet[4] and it is expected to be small (O(10^{-4})). Current measurements have limited experimental precision and don't show any deviation from expectation.
The third type of **CP violation occurs in the interference between decay without mixing and decay with mixing**: $A(B \to f)$ and $A(B \to \overline{B} \to f)$ and this is possible when $Im(\lambda_f) = Im(\frac{q}{p}\frac{\overline{A}_f}{A_f}) \neq 0$. Final states which are accessible by both the B and $\overline{B}$ mesons (non flavour specific decays as $B_s^0 \to D_s^{\mp}K^{\pm}$,$B^0 \to D^-\pi^+$, $B \to \pi\pi$, $B_s^0 \to J\psi\Phi$ etc) may show this kind of CP violation effect . The CP asymmetry in the interference between mixing and decayis measured through time-dependent analyses which require the use of *Flavour tagging algorithm* (see Chapter 4) to infer the initial flavour of the B meson. One of such decays for the B_s^0 and $\overline{B}_s^0$ mesons is the $D_s^{\mp}K^{\pm}$, in which the CP violation measurement is described in this thesis.

2.5 MEASUREMENT OF γ THROUGH B DECAYS

The sensitivity to the angle γ in B mesons decays arises from tree-level $B \to DK$ decays and in loop-level measurements of $B \to h^+h'^-$ decays. Tree-level decays

4 A part the D0 result [14] which has not been confirmed by other measurements.

occur by the exchange of a $W^{\pm}$ vector boson and their theoretical prediction are very precise. The Feynman diagrams of the tree-level $B \to DK$ decays are shown in Fig.4.

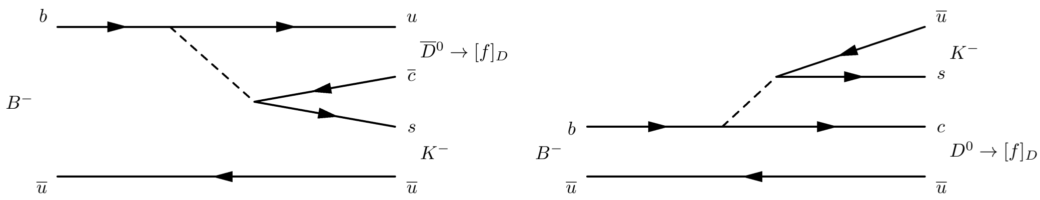


Figure 4.: Feynman diagrams of the $B \to DK$ tree-level decays.

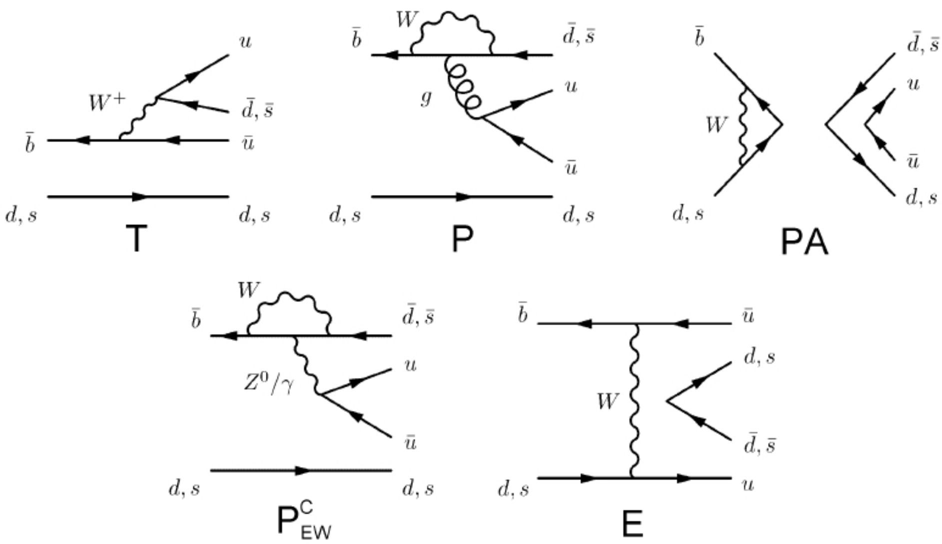


Figure 5.: Feynman diagrams contributing to the $B \to h^+h^{'-}$ decays: Tree (T), Penguin (P), Penguin Annihilation (PA), Colour-suppressed Electroweak Penguin (P^C_{EW}) and Exchange (E).

For $B \to h^+h^{'-}$ decays loop diagrams give sizeable contribution to the decay amplitude. The Feynman diagrams contributing to the amplitudes of $B \to h^+h^{'-}$ decays are shown in Fig.5. Assuming the validity of the U-spin symmetry among B^0_d and B^0_s decays which differ only by the spectator quark (d or s) and taking the values of the mixing phases ϕ_d and ϕ_s from independent measurements it is possible to measure the γ angle in time-dependent analyses. Loops diagram have the $W^{\pm}$ boson in the loop and a gluon or a photon or Z^0 is emitted from the loop. Within the loop any virtual particle of the correct quantum numbers can be exchanged, including high mass super-symmetric particles, provided they have a non negligible coupling. On the other hand γ measurements from tree-level decays have not contributions from next-order diagrams and they can be used as a Standard Model benchmark for γ that compared with the measurement from loop decays might indicate the effect of new physics. As reported in Sec.2.2.2, the precision on γ from direct measurements must be improved in order to compare it with the theoretical prediction, and LHCb is dominating the sensitivity on this CKM parameter exploiting both tree and loop level decays. In the following we will concentrate on γ angle measurements from tree-level decays.

2.5.1 *Tree-level decays to measure γ*

Measurement of γ from tree-level decays are considered clean since the eventual contribution from loops diagram to these decays is estimated to be very small: $\delta\gamma/\gamma \approx O(10^{-7})$ [20]. The sensitivity to γ from tree-level decays arises from the interference between $b \to u$ and $b \to c$ quarks transitions in the decay of a B meson to a final state with a charmed meson. These decays are characterized by small branching ratios $O(10^{-4}; 10^{-5})$ and this explains why the precision on γ is so far limited. Depending on the considered B decay, two kinds of CP violation process carry the dependence on γ: CP violation in decays and CP violation in the interference between mixing and decay. The former is measurable through time-integrated analysis in several $B \to Dh$ decays, where the B can be a $B^{\pm}$ or a B^0, h can be a $K^{\pm}$, K^{*0},$\pi^{\pm}$ and D is a D^0 or a $\overline{D}^0$ meson. For the first time LHCb has also exploited multi-body B decays as $B \to DK\pi\pi$ final states [15].

In this kind of decays the B meson decay can occur via a charmed meson or its anti-particle. The crucial point is that both the D^0 and the $\overline{D}^0$ mesons can decay in the same final state and the sensitivity to γ depends on the interference between these two paths where f_D is the final state of both the D^0 and the $\overline{D}^0$ mesons. The D^0-$\overline{D}^0$ mixing effect must be taken into account. An illustrative schematic view is given in Fig.6. The size of the interference depends on the weak phase γ, the r_B

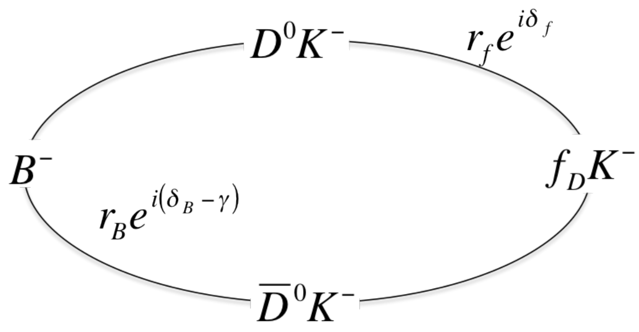


Figure 6.: Illustration of γ sensitivity in tree-level decays with direct CP violation.

parameter which is the amplitude ratio ($r_B = \Gamma(B \to \overline{D}K)/\Gamma(B \to DK)$) and on δ_B which is the CP-invariant strong phase difference. Of course r_B and δ_B are different for each B decay while γ is common, so that many independent measurements can be combined to improve the precision on this CKM angle. Three methods of extracting γ have been determined depending on the D meson final state:

- the Gronau-London-Wyler (GLW) method [17] for f_D being CP eigenstates ($D^0 \to K^+K^-$ or $D^0 \to \pi^+\pi^-$).

- the Atwood-Dunietz-Soni (ADS) method [18] for f_D being flavour-specific eigenstates ($D^0 \to K^+\pi^-$ or $D^0 \to K^+\pi^-\pi^+\pi^-$). The D meson decay parameters, r_D and δ_D, enters in the game and they usually are taken from other measurements.

- the Giri-Grossman-Soffer-Zupan (GGSZ) method [19] for f_D being self-conjugate three-body decays ($D^0 \to K^0_S K^+K^-$ or $D^0 \to K^0_S\pi^+\pi^-$). Being f_D a three-body decay r_D and δ_D vary across the phase space. Different analysis methods have been used also by LHCb to account for them (Model dependent [21] and Model independent measurements [22]).

In recent years, extensions of these methods have been successfully used by LHCb, like the quasi-GLW and quasi-ADS methods, adding a neutral pion to the D final states ($D \to h'^+h^-\pi^0$). Time-integrated analyses are performed through an invariant mass fit devoted to measure ratios and asymmetries of the decay widths of B^+ vs B^-, or B^0 vs $\overline{B}^0$. These observables bring the dependence on γ, but also on r_B and δ_B which are not known with great precision from theory and thus are measured as well by each analysis. Depending on the considered B and D decay, and thus on the adopted analysis method, γ and δ_B can be determined with more or less ambiguity. The interplay of the different $B^{\pm} \to DK^{\pm}$ measurements per-

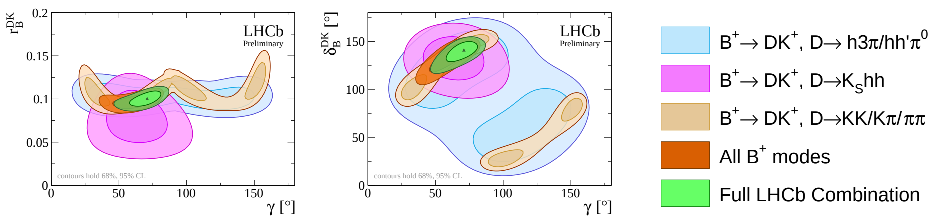


Figure 7.: Two-dimensional CL distributions of γ vs. r_B^{DK} (left) and γ vs. δ_B^{DK} for various $B^+ \to DK^+$ decays [10].

formed by LHCb for the γ (r_B, δ_B) determination is reported in figure 7. There isn't a golden mode, but each of them help to constrain γ. ADS and GLW decays provide very narrow solutions for γ but with the larger level of ambiguity while the GGSZ decay finds one solution with a larger uncertainty. Similarly to $B^+ \to DK^+$ decay, sensitivity to γ can be gained analysing $B^+ \to D\pi^+$ decays. In this case r_B is smaller ($r_B \approx 0.01$) than $B \to DK$ ($r_B \approx 0.1$), and for this reason they are less sensitive to γ. On the other hand these decays have a larger branching ratios (10^{-3}). Also neutral B^0 decays, as the $B^0 \to DK^{*0}$, are sensitive to γ. They have larger r_B values ($r_B \approx 0.3$) than the charged modes however they are affected by larger contribution of backgrounds as the ones from B_s^0 and also with the residual non resonant $K\pi$ events. Although they have lower statistics (BR 10^{-5} vs. 10^{-4}) with respect to the charged modes, the B^0 modes combined together are starting to constrain γ, as reported in Fig.8. Also for neutral B decays the ADS and GLW modes provide more precise γ and $\delta_B^{DK^{*0}}$ determinations but with the larger level of ambiguity while the GGSZ decay finds one broader solution.

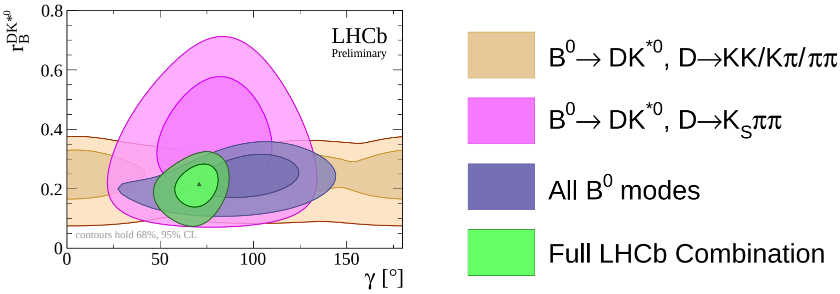


Figure 8.: Two-dimensional CL distributions of γ vs. r_B^{DK} for various B^0 decays [10].

2.6 MEASUREMENT OF THE CKM ANGLE γ IN $B_s^0 \to D_s^{\mp} K^{\pm}$ DECAYS

In addition to the time-integrated analyses, B decays which show CP-violation effects in the interference between decay and mixing can be used to determine γ through time-dependent analysis. Time-dependent measurements to extract γ have been done so far by Babar and Belle by using the $B^0 \to D^{(*)\mp}\pi^{\pm}$ decays (in which the magnitude of the interference was limited by the small value of the amplitude ratio $r_B \approx 0.1$). Very recently the LHCb collaboration has measured, for the first time, the angle γ in the decay mode $B_s^0 \to D_s^{\mp} K^{\pm}$ by means of a time-dependent analysis [1]. Such decay mode has the advantage to have a larger value for r_B ($\approx$0.5) since the $B_s^0 \to D_s^- K^+$ and the $\overline{B}_s^0 \to D_s^- K^+$ amplitudes are of the same order of magnitude allowing a large interference effect. The physical description of the decay rate is given in Sec.2.3.2 and the resulting CP-asymmetry for B_s^0 and $\overline{B}_s^0$ with $f = D_s^- K^+$ and $\overline{f} = D_s^+ K^-$ can be expressed as:

$$A_{CP}^{f}(t) = \frac{\frac{d\Gamma}{dt}(\overline{B}_s^0 \to f)(t) - \frac{d\Gamma}{dt}(B_s^0 \to f)(t)}{\frac{d\Gamma}{dt}(\overline{B}_s^0 \to f)(t) + \frac{d\Gamma}{dt}(B_s^0 \to f)(t)} = \frac{-C_f \cos(\Delta m_s t) + S_f \sin(\Delta m_s t)}{\cosh(\frac{\Delta\Gamma_s}{2} t) + A_f^{\Delta\Gamma} \sinh(\frac{\Delta\Gamma_s}{2} t)}. \tag{2.45}$$

and an analogous expression can be written for $\overline{f}$. Four decay rates $B_s^0 \to D_s^- K^+$, $B_s^0 \to D_s^+ K^-$, $\overline{B}_s^0 \to D_s^- K^+$ and $\overline{B}_s^0 \to D_s^+ K^-$ are used by the fit to determine the CP observables defined by equations 2.36-2.37. The CP observables can be written in terms of the weak phases difference (γ -2β_s), the amplitude ratio r_B and the strong phase difference δ_B as follows:

$$C_f = \frac{1 - r_B^2}{1 + r_B^2}, \quad S_f = \frac{2r_B \sin(\delta_B - (\gamma - 2\beta_s))}{1 + r_B^2}, \quad A_f^{\Delta\Gamma} = \frac{-2r_B \cos(\delta_B - (\gamma - 2\beta_s))}{1 + r_B^2} \tag{2.46}$$

$$S_{\bar{f}} = \frac{-2r_B \sin(\delta_B + (\gamma - 2\beta_s))}{1 + r_B^2}, \quad A_{\bar{f}}^{\Delta\Gamma} = \frac{-2r_B \cos(\delta_B + (\gamma - 2\beta_s))}{1 + r_B^2} \tag{2.47}$$

where r_B is the magnitude of the ratio of decay amplitudes: $|A_{\overline{B_s^0} \to f} / A_{B_s^0 \to f}|$, and β_s ($\simeq -2\phi_s$) is the mixing phase for B_s^0 mesons. Thanks to the possibility to measure the five CP observables from a tagged time-dependent fit, the ambiguity on γ is limited to two possible solutions (γ and $\gamma + \pi$). In this thesis it will be described one of the most important tool to perform time-dependent measurement, the Flavour Tagging, and also the first measurement of γ from a time-dependent analysis in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay mode using the dataset collected by LHCb in the 2011. The updated measurement of γ exploiting the full Run1 data sample of $B_s^0 \to D_s^{\mp} K^{\pm}$ events is on-going. The major improvements of this updated analysis will be described in the final chapter.

3

THE LHC*b* EXPERIMENT

Contents

3.1 The Large Hadron Collider **25**
3.2 The LHC*b* Experiment **27**
3.2.1 The tracking system 30
3.2.2 The LHCb magnet 36
3.2.3 Mass and Momentum resolution 37
3.2.4 Particle Identification 38
3.2.5 The calorimeters system 41
3.2.6 Muon detector 43
3.2.7 The LHCb trigger 44
3.2.8 Reconstruction algorithms 47
3.2.9 The LHCb software 48

After a brief description of the main features of the Large Hadron Collider (LHC), the LHCb experimental apparatus used for the studies presented in this thesis is described conferring particular relevance to the sub-detectors which play a crucial role for the analysis described in this dissertation.

3.1 THE LARGE HADRON COLLIDER

The LHC is a proton-proton collider [27] built by the European Organization for Nuclear Research (CERN) in the vicinity of Geneva (Switzerland). The LHC is located 100 m underground along a 27 km long tunnel. The LHC is designed to accelerate proton beams up to an energy of 7 TeV, leading to a center of mass energy of $\sqrt{s}$=14TeV and to a peak luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$. The two counter-rotating proton beams are bent by dipolar NbTi superconducting magnets that require to be maintained at a temperature of 1.9K. The peak dipole magnetic field is of 8.33T. Different bunch filling schemes can be produced corresponding to a different interaction frequency. For example, in the nominal configuration, the beams are structured in 2808 bunches containing each $\sim 10^{11}$ protons and spaced by 25ns. The beam collision frequency is then 40MHz. During 2011 and 2012 data taking the LHC worked with a lower interaction frequency.

In correspondence of the four interaction points, four experimental setup are located. ATLAS (A Toroidal LHC ApparatuS [28]) and CMS (Compact Muon Solenoid [29]) are general purposes detectors. The ALICE experiment (A Lead Ion Collision Experiment [30]) is devoted to study through Lead ions head-on collisions a state of matter wherein quarks and gluons are freed. The LHCb experiment (Large Hadron Collider beauty) will be described in details in the next section.

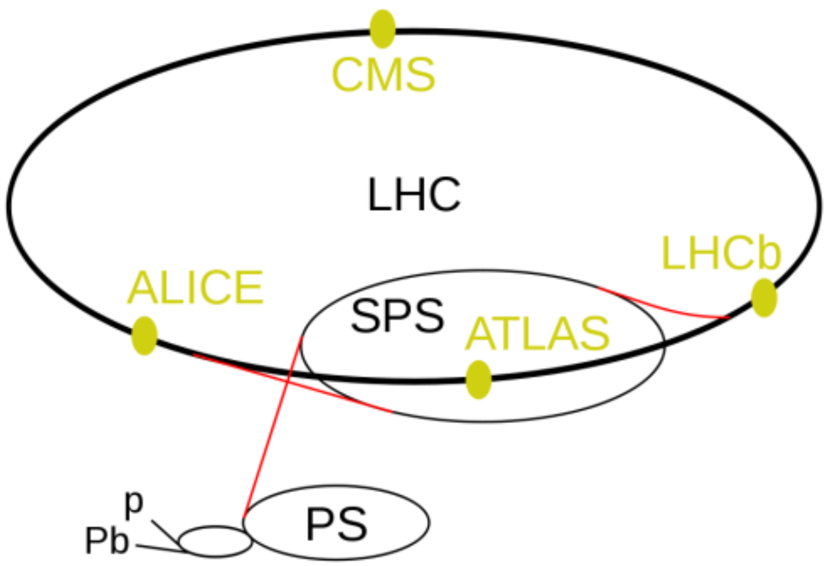


Figure 9.: A schematic representation of the LHC accelerator system and the position of the four experiments ATLAS, CMS, ALICE and LHCb.

A schematic view of the LHC accelerator system is reported in Fig. 9.

3.2 THE LHCb EXPERIMENT

The LHCb experiment is designed to study beauty and charm hadrons decays, in particular to perform precise measurements of CP violation effects and to search for physics beyond the Standard Model through rare decays. The optimal luminosity for the experiment is such that there are at maximum 2.5 pp collisions per bunch crossing. To decrease the instantaneous luminosity with respect to the one delivered to the other experiments which is of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, the two proton beams are tilted each others so that the beams do not collide head-on and the interaction area is larger. This technique is called *luminosity levelling* and it allows LHCb to maintain the required luminosity constant to a level of $10^{32}\text{cm}^{-2}\text{s}^{-1}$.

The integrated luminosity delivered by LHCb during 2010, 2011 and 2012 data taking are summarized in Tab.5. The center of mass energy has been increased from 7TeV to 8TeV in 2012.

Table 5.: Integrated luminosity delivered by LHCb during 2010, 2011 and 2012 data taking. For each data taking period also the average number of interactions per bunch crossing (μ) and the peak luminosity are reported.

Year	$\mathcal{L}_{int}$ [fb^{-1}]	$\sqrt{s}$ [TeV]	μ	$\mathcal{L}_{peak}$ [10^{32} cm^{-2}s^{-1}]
2010	0.037	7	$1-2.5$	1.6
2011	1.0	7	$1.5-2.5$	4.0
2012	2.2	8	$\simeq 1.8$	4.0

At the LHC the $b-\bar{b}$ quarks production cross-section is dominated by *gluon-gluon* fusion and *quark-antiquark* annihilation (hard QCD scattering), and thus it mainly occurs in two cones pointing in the forward and backward regions respect to the beams direction (Fig. 10). In contrast to the other detectors, LHCb is a single-arm forward spectrometer with a pseudorapidity η[1] between 1.8 and 4.9. The estimated $b-\bar{b}$ cross-section at $\sqrt{s}$ = 14 TeV is around 500 μb of which 200 μb within the LHCb acceptance. The physics program related to spectroscopy and study of B decays requires to have at least one $b(\bar{b})$ quark in the detector acceptance. Instead several of the CP violation measurements which involve B_s^0 or

1 $\eta = -\ln\left[\tan\frac{\theta}{2}\right]$ where θ is the angle of the particle's momentum respect to the beam axis.

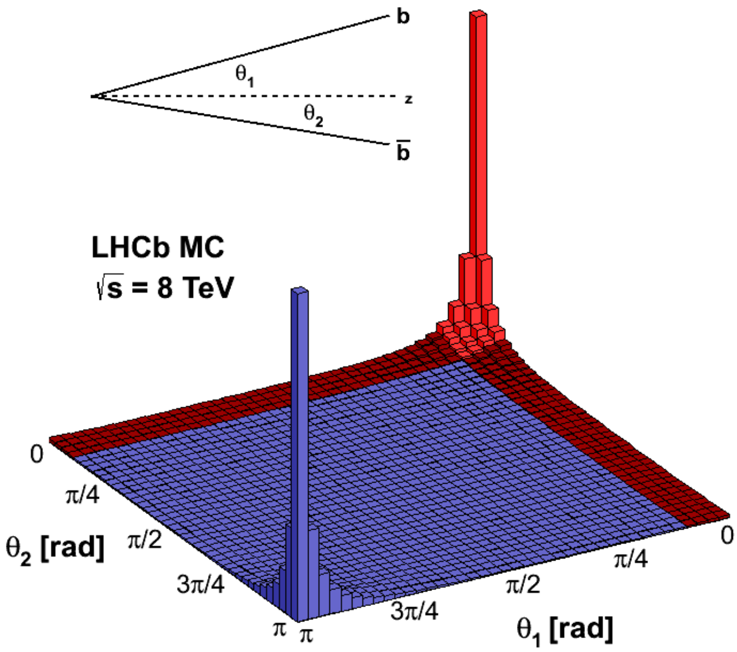


Figure 10.: $b-\bar{b}$ production cross-section at $\sqrt{s}$= 8 TeV as a function of polar angles. In red the part of the distribution within the LHCb acceptance. Figure taken by [33].

B_d^0 decays need to have both the b and $\bar{b}$ hadrons within the LHCb acceptance in order to determine the flavour at production time. Events containing both b and $\bar{b}$ are important for Flavour Tagging algorithms which reconstruct the non signal b hadron decay to infer the signal B meson flavour at production time. More details on Flavour Tagging algorithms will be given in the dedicated Chapter 4. The $c-\bar{c}$ cross-section is 20 times larger than the $\sigma_{b\bar{b}}$ allowing LHCb to perform interesting measurements also in the charm sector. Despite the hadronic environment of pp collision represents a challenge for event reconstruction, LHCb has demonstrated to be able to perform precise measurements with high statistics even beyond expectations.

The LHCb detector is composed of several sub-detectors as shown in Fig.11. The following subsections will describe each sub-detector with more details for those which have more impact on the analysis presented in this thesis.

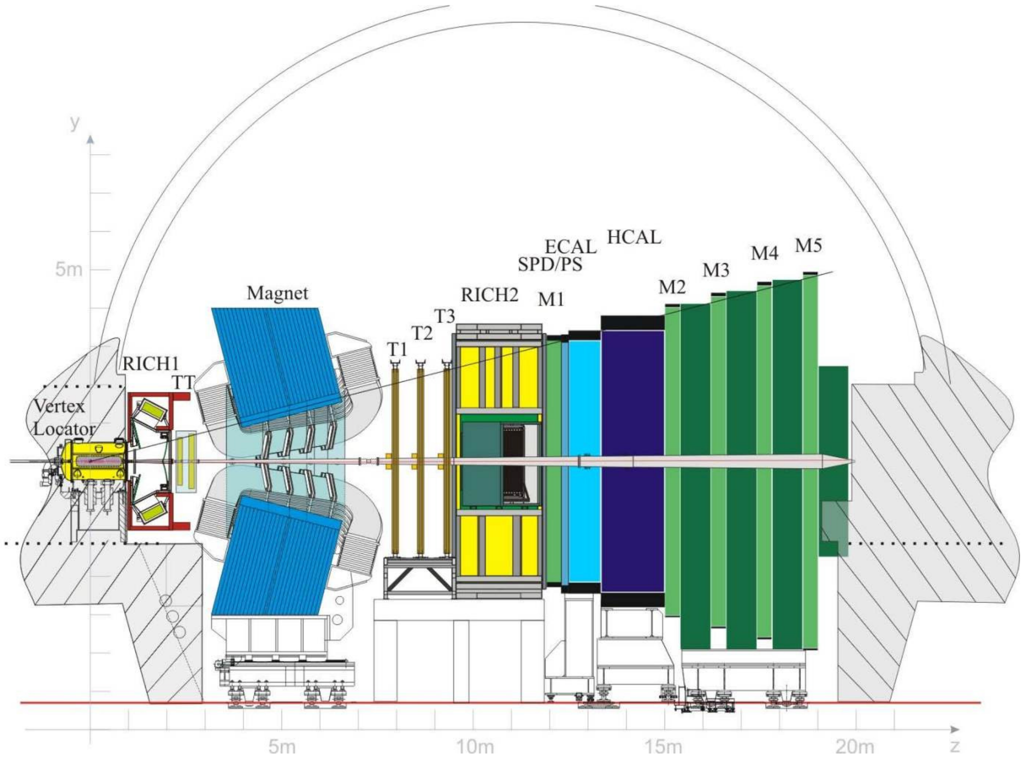


Figure 11.: Layout of the LHCb detector in $y - z$ plane. From left to right: the VErtex LOcator, RICH1 and RICH2 indicate the Ring Imaging Cherenkov detectors, TT and T1-T3 denote the tracking stations, M1-M5 are the muons detectors. SPD and PS, as well as ECAL and HCAL compose the Calorimeters system. The coordinate system is a right-handed Cartesian system where the z axis corresponds to the beam direction, the y axis is contained in the non-bending plane while the x axis is oriented horizontally towards the bending plane.

3.2.1 *The tracking system*

The tracking system provides information to reconstruct charged particle trajectories, to measure their momentum and to reconstruct the vertices of the decays. It consists of four sub-detectors employing different technologies and geometries: the VErtex LOcator (VELO), the Tracker Turicensis (TT), the Inner Tracker (IT) and the Outer Tracker (OT). These latter two are both placed in T1-T2-T3 stations.

The Vertex Locator

The VErtex LOcator is devoted to the vertices determination. Indeed it is designed to provide precise determination of the position of the charged tracks close to the interaction point. The Primary Vertex (PV) indicates the position of the pp collisions, while the Secondary Vertex (SV) relates to the position of the particle decay. Indeed b-hadrons have displaced production and decay vertex ($\sim$ 0.5 cm) because of their relatively long lifetimes. The precise measurement of b-hadrons flight distance ($\mid \vec{PV} - \vec{SV} \mid$) together with momentum information from other sub-detectors allows to reconstruct the b-hadrons decay time. This is a fundamental observable for each time-dependent CP violation measurement in b-hadrons decays as well as lifetime and $\Delta\Gamma_{(s)}$ measurements. The resolution of the primary vertex depends on the number of tracks used to reconstruct the vertex, on average it is 71 μm in the z direction and 13 μm in the perpendicular direction [34]. Another important variable to detect properly B meson decays is the *Impact Parameter* (IP) of a track, defined as the distance between the track and the PV at the track's point of closest approach to the PV. Long lifetime particles, such as B and D mesons, are characterized by large IP and they can be distinguished from particles originated directly from pp collisions requiring an IP larger than a given value. The IP resolution depends on the multiple scattering effect ($\approx 1/P_T$ where P_T is the transverse momentum) due to the detector material, on the resolution of the positions of track hits and on the distance of the first hit of a track from the PV. Figure 12 shows the IP_x and IP_y resolutions as a function of the momentum of the track extracted from data and the comparison between data and simulation for the IP dependence on the inverse of the transverse momentum. Time-dependent

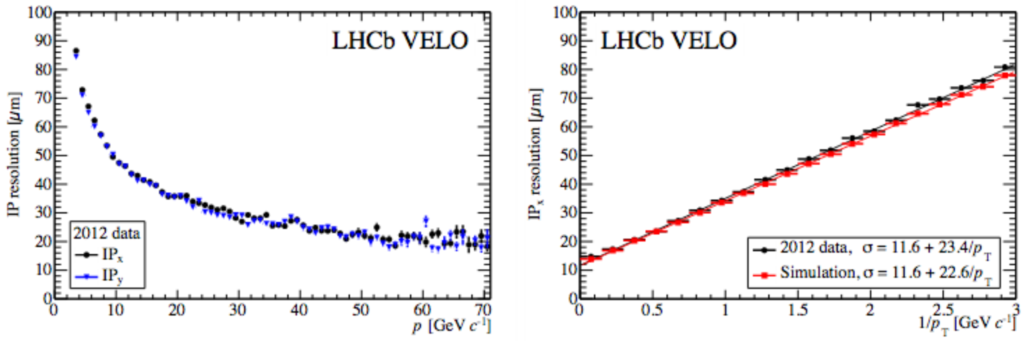


Figure 12.: IP_x and IP_y dependence on the momentum (left) and IP_x dependence on $1/P_T$ in data and simulation. Figure taken from [34].

measurements, in particular the ones related to B_s^0 decays require a good time resolution σ_t in order to resolve the fast $B_s^0 - \overline{B}_s^0$ oscillations and measure properly the decay amplitude. Decay time is computed by the following formula

$$t = \frac{m_B d_B}{P_B} \tag{3.1}$$

where d_B is the B meson flight distance: $d = |\vec{SV} - \vec{PV}|$, m_B and P_B are respectively the mass and the momentum of the B meson. Of course the position of the secondary vertex is determined with a poorer resolution with respect ot the PV due to the fact that the number of tracks is much lower and they all are in the forward region. The distribution of the time resolution and its dependence on are reported in Fig.13. The σ_t evaluation underestimates the real resolution on the decay time and for this reason it needs to be calibrated.
In order to reduce the distance from the beam and to use less material as possible the VELO has been placed in a secondary vacuum environment. To preserve the vacuum condition of the beams line, an intermediate layer, called RF foil is placed. This foil can be very thin and made by an aluminum alloy which minimize the multiple scattering effect between the PV and the first hit of the track allowing the VELO sensors to stay as close as possible to the beam line. The VELO is made by 21 semicircular tracking stations surrounding the beam axis at a distance of 8 mm in nominal working conditions (see Fig.14).

To protect the detector from beam interaction with the sensors during beam injection, each station has left and right parts which can be displaced during the

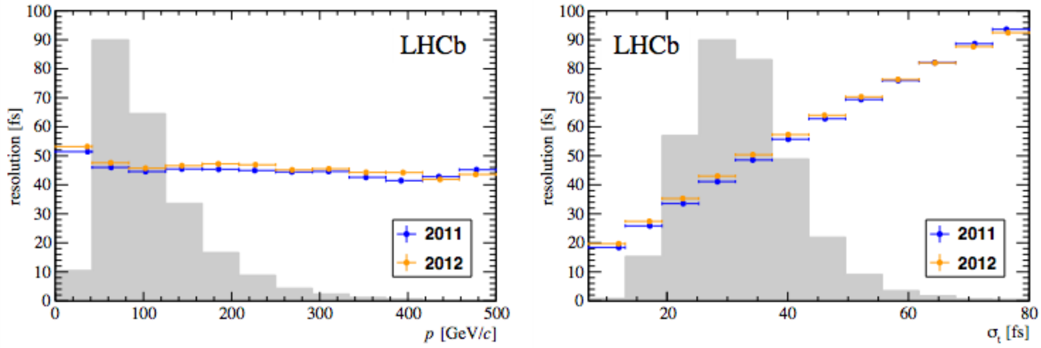


Figure 13.: Dependence of the decay time resolution on the B momentum (left) and on the estimation of the decay time error (right) [34].

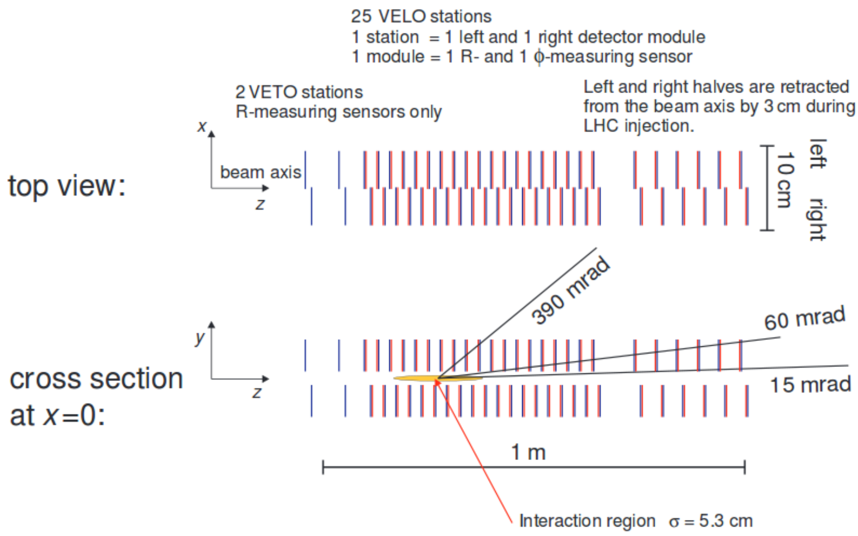


Figure 14.: Cross section of the VELO in the $x-z$ and $y-z$ plane showing the 21 modules [31].

beam injection and moved close to each other once the beam is stable. The half disk shape modules (Fig.15) are silicon wafers covered by aluminum strips. There are two kinds of VELO sensors, depending on which spatial coordinate they are sensitive to: r, which is the radial distance of the hit from the beam axis, and ϕ, which is the polar angle and it ranges from 15 mrad to 390 mrad. Each sensor has a diameter of 84 mm and a thickness of 300 μ m.

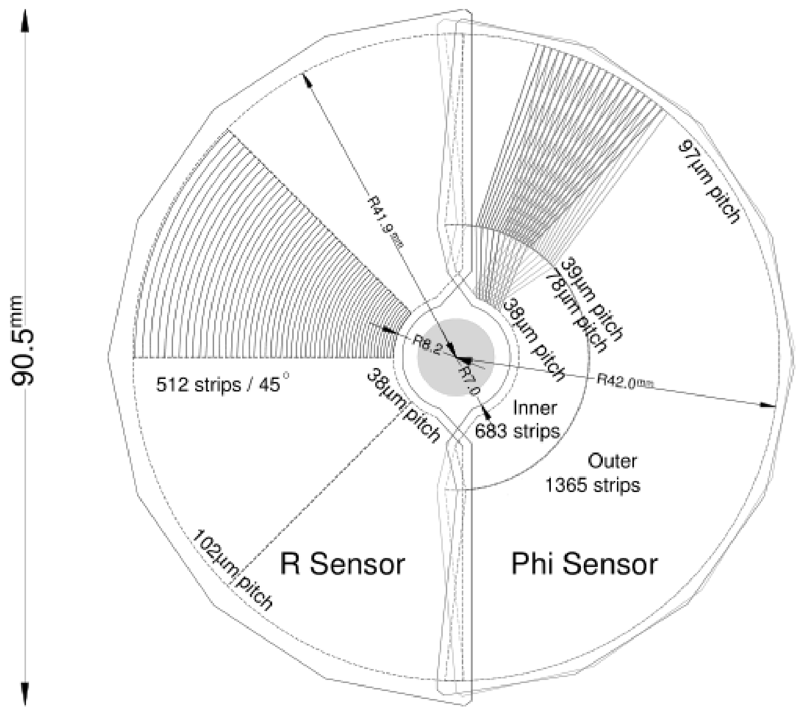


Figure 15.: Schematic view of r and ϕ sensors [31].

The Tracker Turicensis

The Tracker Turicensis (TT) is placed between the RICH1 detector and the magnet and it is designed to reconstruct long-lived particles. The TT is composed by four layers of silicon micro-strip detectors with a thickness of 500 μm grouped into two stations TTa and TTb. According to the particle occupancy, each layer is subdivided into sections. In order to determine the transverse component of the momentum, the TT is composed by 3 types of layers: the X layers have vertical readout strips, while U and V are stereo layers rotated by $+5°$ and $-5°$ respectively along the y axis. The TTa (TTb) stations are composed by the X-U (X-V) layers, respectively. Figure 16 outlines the layout of TTa. The TT covers an area of 8.4 cm^2 with a single hit resolution of 50μ m. Given the proximity to the magnet, the TT sees a residual magnetic field.

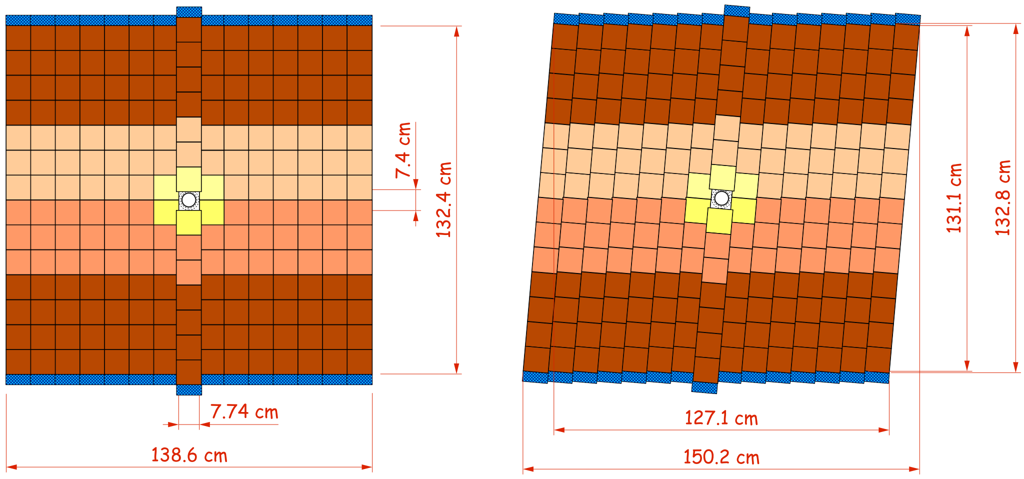


Figure 16.: Layout of two TT detection layers (X and U)[36].

The Inner Tracker

The Inner Tracker (IT) is the closest part of the T1-T3 stations to the beam pipe, located downstream the LHCb magnet. It is cross shaped and made by four detector boxes surrounding the beam axis (Fig.17). Each IT station is composed by 4 layers of silicon micro-strips, as the TT, following the X-U-V-X arrangement. In this acceptance region, 1.3% of the LHCb acceptance, the particle flux is extremely high and the silicon detectors must have radiation hardness properties. The pitch between the sensors is 200μm and the single hit resolution is 50 μm as for TT.

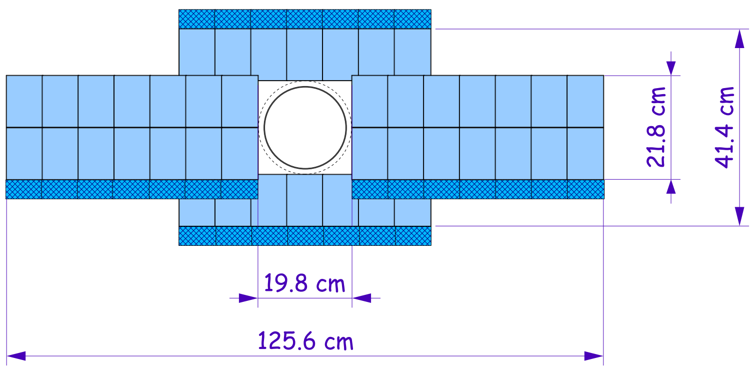


Figure 17.: Layout of a single layer (X) of the IT detector [36].

The Outer Tracker

The Outer Tracker (OT) covers the area outside the IT acceptance in the T1-T3 tracking stations downstream the LHCb magnet as shown in Fig.18. The OT consists of gaseous straw-tube detectors. The gas mixture is composed by Ar, Co_2 and O_2 in the following percentages: 70%, 28.5% and 1.5%. The OT covers an acceptance of 300 mrad (250 mrad) in the horizontal (vertical) plane. As for the IT, each of the three OT stations is composed by four layers of straw tubes in the X-U-V-X configuration. Each layer is made of a double layer of straw tubes.

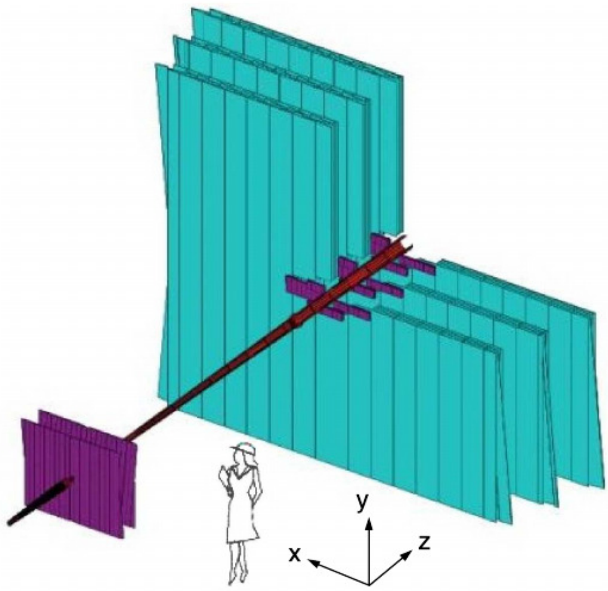


Figure 18.: Schematic view of the tracking system. The blue colored parts represent the OT while the purple ones represent the TT and IT trackers. Figure taken from[31].

The straws have an inner diameter of 4.9 mm and in the center of the straws there is a 24 μm diameter gold coated tungsten wire. The spatial resolution of the single straw tube is 200 μm . The position of the interaction is given by the position of the hit straw tube and by the drift time.

3.2.2 *The LHCb magnet*

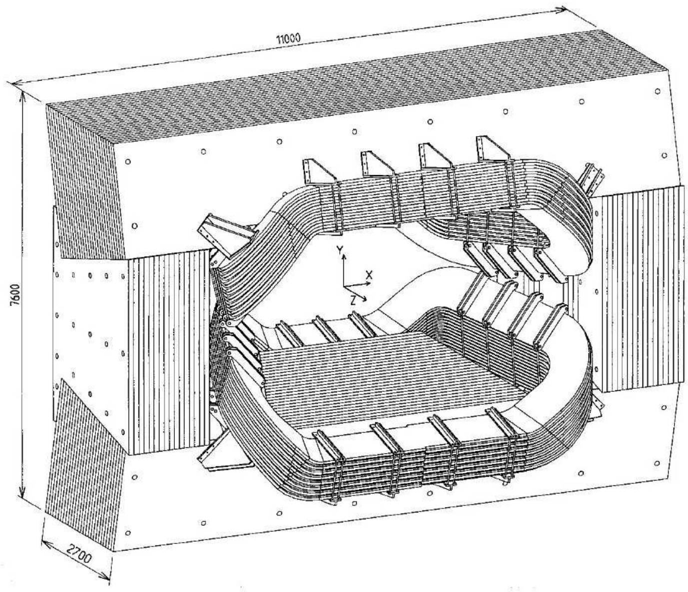


Figure 19.: Perspective view of the LHCb dipole magnet. The interaction point lies behind the magnet [31].

A dipole magnet placed in between the TT and the T1-T3 stations is used to bend charged tracks in order to measure the momentum of the charged particles. It is composed by two symmetric non-superconductive aluminum coils arranged above and below the beam axis, as it is represented in Fig. 19. A charged particle is bent when passing through a magnetic field depending on its momentum. The curvature of the trajectory allows to determine the charge and the momentum of the particle.

The magnetic field is along the y axis and its integrated value is $4Tm$ with a precision of about 4×10^{-4}. The polarity of the magnetic field can be switched during data taking to minimize possible detection asymmetries that can affect the accuracy of asymmetry measurements.

3.2.3 *Mass and Momentum resolution*

The momentum resolution for long tracks is measured using data samples of $J/\psi \to \mu^+\mu^-$ decays. In Fig.20 the relative momentum resolution versus the momentum for long tracks obtained using J/ψ decays is shown. The momentum resolution is of about 0.5 % for momentum particles below 20 GeV/c reaching the 0.8% for momentum particles around 100 GeV/c. The mass resolution is studied using six di-muons resonances which show a clear mass peak and share the same topology: J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ mesons plus the Z^0 boson. A loose selection is applied to get the invariant mass distribution that is fitted with a Double Gaussian plus a power low tail to account for final state radiation. The background component is modelled by an exponential shape. The fitted values of the mass resolution for the six resonances are reported in Fig. 21 (right). The relative mass resolution as a function of the mass of the dimuon resonances is shown in Figure 21 where the mass resolution is at the 0.5% level up to the Υ masses.

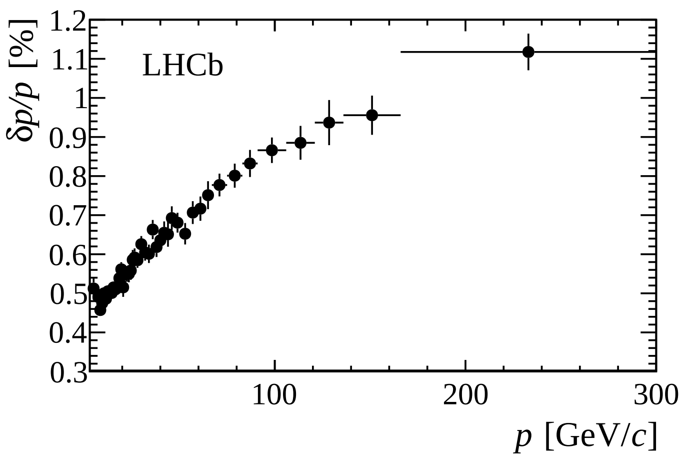
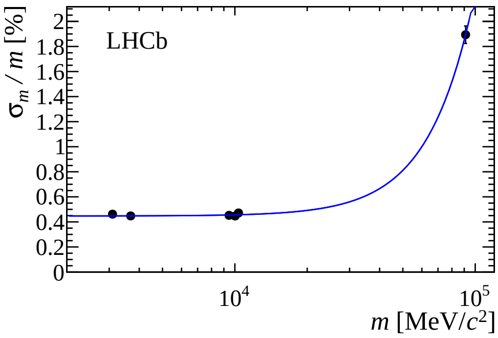


Figure 20.: Relative momentum resolution versus momentum for long tracks in data obtained using J/ψ decays. Figure taken from [32].



Resonance	Mass resolution (MeV/c^2)
J/ψ	14.3 ± 0.1
$\psi(2S)$	16.5 ± 0.4
$\Upsilon(1S)$	42.8 ± 0.1
$\Upsilon(2S)$	44.8 ± 0.1
$\Upsilon(3S)$	48.8 ± 0.2
Z^0	1727 ± 64

Figure 21.: Relative mass resolution as a function of the mass of the dimuon resonance (left) and mass resolution values for the six dimuon resonances (right). Figures taken from [32].

3.2.4 *Particle Identification*

In pp collisions it is crucial for the reconstruction of b-hadron decays to identify the different charged particles such as electrons, muons, protons, kaons and pions. The sub-detectors which contribute to the particles identification are: the Ring Imaging Cherenkov detectors, the calorimeters system and the Muon detector. In the following subsections a brief description of each of them will be presented.

The Ring Imaging Cherenkov detectors

The Ring Imaging CHerenkov detectors (RICH) employ the Cherenkov effect [37] to measure the particle's velocity. When a charged particle passing through a medium has a velocity which is larger than the speed of light in vacuum, $c' = c/n$, it emits the typical Cherenkov radiation. The Cherenkov radiation is the emission of photons in a cone centered along the propagation direction with a characteristic angle called θ_C which relies on the particle's velocity and on the refraction index n, accordingly to this formula:

$$\cos\theta_c = \frac{c'}{v} = \frac{1}{\beta n}. \tag{3.2}$$

Thus the measurement of the Cherenkov angle θ_C determines the velocity of the particle.

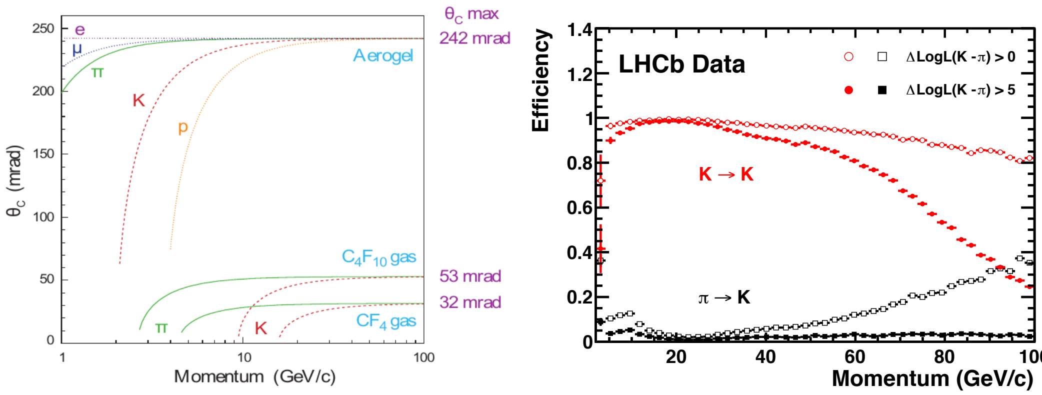


Figure 22.: Right: Cherenkov angle versus particle momentum for RICH radiators. Figure taken from [31]. Left: Kaon identification and pion misidentification as a function of the track momentum measured on Run1 data. Plot for two different $DLL(K-\pi)$ requirements are shown. Figure taken from[32].

This information, joined to the measurement of the momentum from the tracking system, allows to obtain the mass of the charged particle and consequently to define the particle identity. In the LHCb environment the momentum spectrum is quite large ranging from 1 to 100 GeV/c. In order to cover efficiently the whole momentum spectrum it is required to use different media, called radiators. In Fig. 22 the dependence of the Cherenkov angle on the particle momentum is shown using three kinds of radiators: Areogel, C_4F_{10} and CF_4. The Areogel has a larger discriminating power at low momentum while the C_4F_{10} and CF_4 media are more indicated to distinguish between high momentum particles.

Table 6.: Gas admixture, refractive index (and the wavelength which they correspond) and yield of photo-electrons for each track with $\beta = 1$.

	Gas	n	λ [nm]	Yield fo p.e.
RICH1	C_4F_{10}	1.0014	400	30
RICH1	Aerogel	1.03	400	6.5
RICH2	CF_4	1.0005	400	22

The RICH1 is located in front of the magnet with an acceptance ranging from 25 mrad to 250 mrad (300 mrad) in the vertical (horizontal) direction. It is filled with Areogel and C_4F_{10} radiators which makes the RICH1 sensitive to low momenta particles and to identify protons. The RICH2 is placed immediately after the T1-T3 tracking stations. It designed to identify high momentum particles. For this reason its acceptance is limited between 15 mrad and 120 mrad (100 mrad) along the x (y) direction and it uses CF_4 radiator. The refraction index and the photo-electrons yield are summarized in Tab. 6 for each medium.

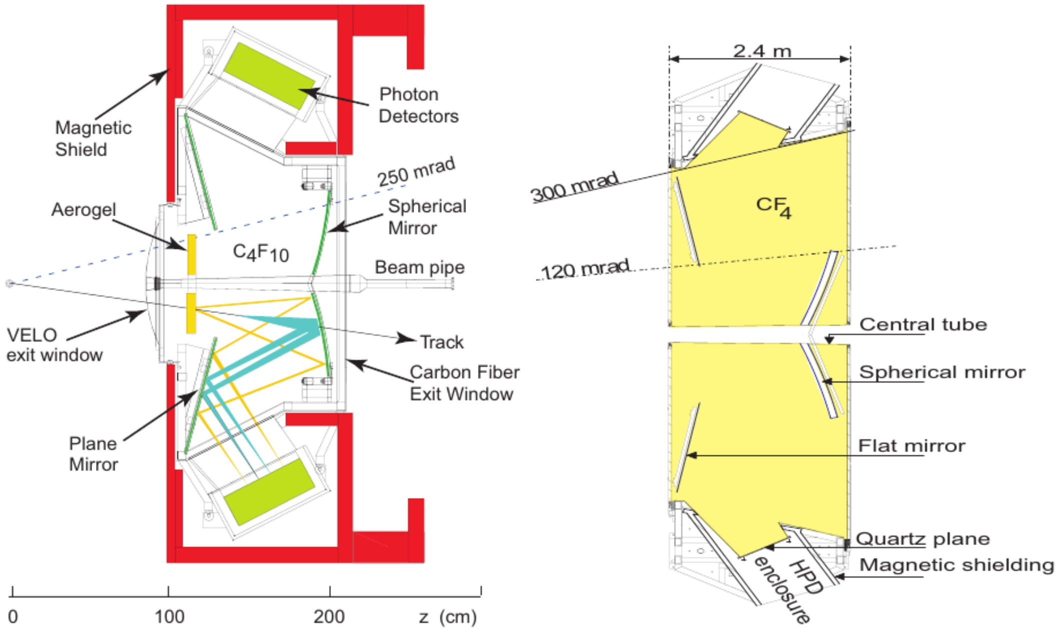


Figure 23.: Schematics of RICH1 and RICH2 detectors. Figure taken from [31].

The Cherenkov radiation is focused via spherical and plane mirrors to the Hybrid Photo Detectors (HPD) that convert photons in electrons (photo-electrons). On the focal plane the Cherenkov light forms rings and the radius determines the Cherenkov angle θ_C. Associating the Cherenkov ring image, which means a velocity, to a track, which means a momentum, a mass hypothesis can be extracted.

For each mass hypothesis a likelihood is calculated using the informations from RICH detectors, calorimeters and the muon system. The difference in logarithmic likelihood, defined as $DLL(A-B) = ln\mathcal{L}_A - ln\mathcal{L}_B$, is the observable used for the particle identification. Roughly speaking, in case of $DLL(K-\pi) < 0$ the particle is more likely to be a pion rather than a kaon. As an important example, in the $B_s^0 \to D_s^{\mp}K^{\pm}$ decays the final charged particles are mainly pions and kaons, furthermore one of the most relevant peaking background is the $B_s^0 \to D_s^- \pi^+$ which has a larger branching fraction ($\mathcal{O}(15)$). It becomes crucial for the analysis to distinguish between the two different bachelor particles and the related observable is PIDK=$DLL(K-\pi)$. Thus an extensive use of the PIDK is done both in the selection and in the fitter itself. As can be seen from Figs.22 choosing $DLL(K-\pi) > 0$ (kaon hypothesis better than pion hypothesis) the average kaon efficiency ($K \to K$) identification over the momentum spectrum is $\sim$ 95% while the pion misidentification ($\pi \to K$) is $\sim$ 10%. Choosing $DLL(K-\pi) > 5$ results in a pion misidentification of $\sim$ 3%. Similarly also for proton hypothesis choosing $DLL(p-\pi) > 5$ reduce the pion misidentification.

3.2.5 *The calorimeters system*

The calorimeters system is designed to identify hadrons, electrons and photons as well as to measure their energy and position. The system exploits absorber materials to trigger the interaction of the particles which is predominantly an electromagnetic or hadronic shower whose signal is collected by a layer of active material. The Scintillating Pad Detector (SPD), the Preshower (PS) detector and the Electromagnetic calorimeter are devoted to the identification of electrons and photons and are placed behind the first Muon station M1. The angular acceptance is between 30 mrad and 300 mrad (250 mrad) horizontally (vertically). The SPD and the PS are made of scintillating pads whose granularity varies depending on the distance from the beam axis to guarantee an optimal detector occupancy. In the SPD a photon doesn't produce any signal before triggering a shower while the electron does. In this way the SPD is able to distinguish between electrons and photons which passing trough the 15 mm of lead after the SPD generate a shower in the PS. In contrast, charged hadrons as pions do not produce a shower, allowing the PS to distinguish them from electrons. Each ECAL cell is composed by 2 mm plates of lead alternated by 4 mm layers of scintillating pads. Particles passing through the ECAL, which corresponds to 25 radiation lengths X_0, deposit their energy and electrons and photons are absorbed. The ECAL energy resolution is a function of energy:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/GeV}} \oplus 1\%, \tag{3.3}$$

where the symbol $\oplus$ indicates sum in quadrature of the two terms. The last element of the calorimeters system is the HCAL in which neutral and charged hadrons deposit their remaining amount of energy. It is made by alternating iron absorbers and scintillating tiles reaching 5.6 nuclear interaction lengths (λ_{int}) of hadrons in iron. The energy resolution is given by:

$$\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E/GeV}} \oplus 9\%. \tag{3.4}$$

The granularity of the calorimeters system is outlined in Fig. 24. The electron identification efficiency of the calorimeters system with a requirement of $DLL_{calo}(e-h) > 2$ is 91.9%, with a mis-identification rate of 4.5%. Including information from the RICH detectors the electron efficiency raises up to 97% for a mis-identification rate lower than 2% [38]. Large energy deposit in ECAL and HCAL are the first signature used by the L0 trigger to accept an event.

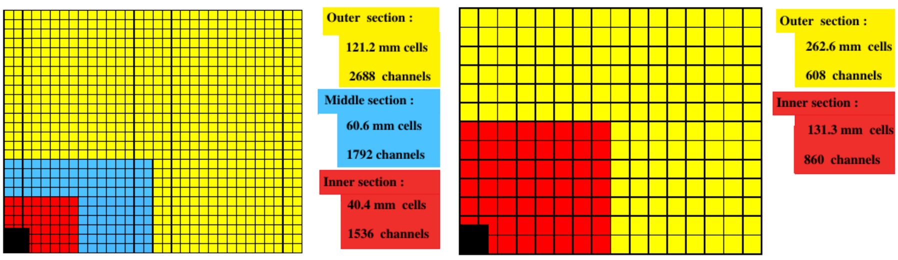


Figure 24.: Schematic representation of modules and granularity of the ECAL (left) and HCAL (right). Only the top right quarter is shown. The black area corresponds to the empty space occupied by the beam pipe [31].

3.2.6 *Muon detector*

The most downstream detector component of the LHCb apparatus is the Muon detector which is composed by 5 stations M1-M5. As for the other detectors, also in this case each station is subdivided in regions (R1-R4) characterized by different granularities. The first station M1 is located before the calorimeters system, while the remaining stations are placed behind the HCAL. Each Muon station is equipped with chambers which consist of gaseous detector: The innermost part of M1 ($M1R1$) uses gas electron multipliers (GEM) given the high particle fluence, while all other regions and stations are made by multiwire proportional chambers (MWPC). Between the muon stations 80 cm thick iron walls are used to absorb hadrons. The layout of the five stations is shown in Fig.25 and Fig.26.

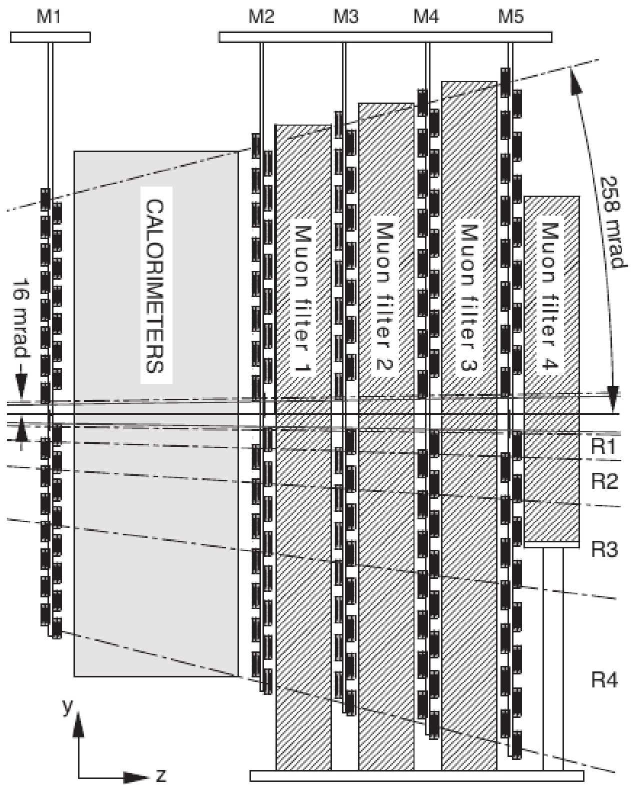


Figure 25.: Side view of the five stations of the muon detector. Figure taken from[31].

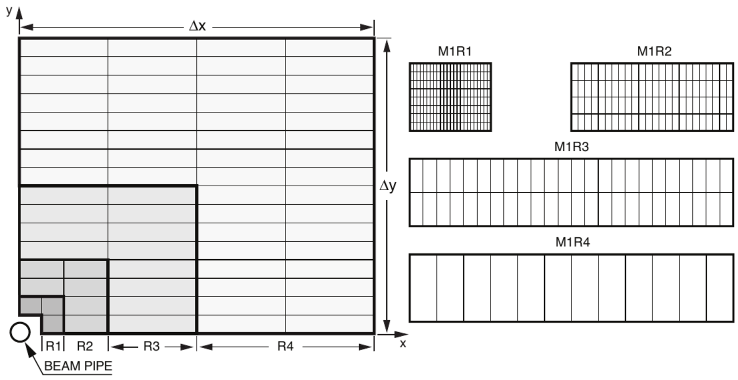


Figure 26.: Schematic representation of the muon chamber assembly showing the different granularities [31].

The Muon detector covers an acceptance between $20 - 300$ mrad in the horizontal plane and $16 - 258$ mrad in the vertical one. The average muon identification efficiency of 98% can be achieved with a pion and kaon misidentification level below 1%, the hadron misidentification probabilities are below 0.6%. The signatures in the muon detector are used by the L0 trigger to select an event.

3.2.7 *The LHCb trigger*

In the first years of operation (2010-2012) LHC worked at reduced energy and a bunch crossing of 15 MHz instead of 40 MHz. The data recording and storage capabilities require a much lower frequency. The LHCb trigger system is designed to reduce such rate to an acceptable limit which is set to 5 kHz. Indeed in the pp collision, only a small fraction of the events are potentially interesting. The LHCb trigger makes this first level selection via three stages: an hardware trigger called Level 0 (L0), which is synchronous with the bunch crossing frequency, and two software triggers called High Level Trigger 1 (HLT1) and High Level Trigger 2 (HLT2). The trigger schema is summarized in Fig.27 where the two HLT are denoted with Software High Level Trigger. The L0 selection is based on the information related to particles transverse momentum or energy and its bandwidth is subdivided in different percentages among hadrons (h), muons (μ,$\mu\mu$) and electrons and photons

(e,γ). The Software High Level Trigger runs offline reconstruction using different inclusive and exclusive selection algorithms.

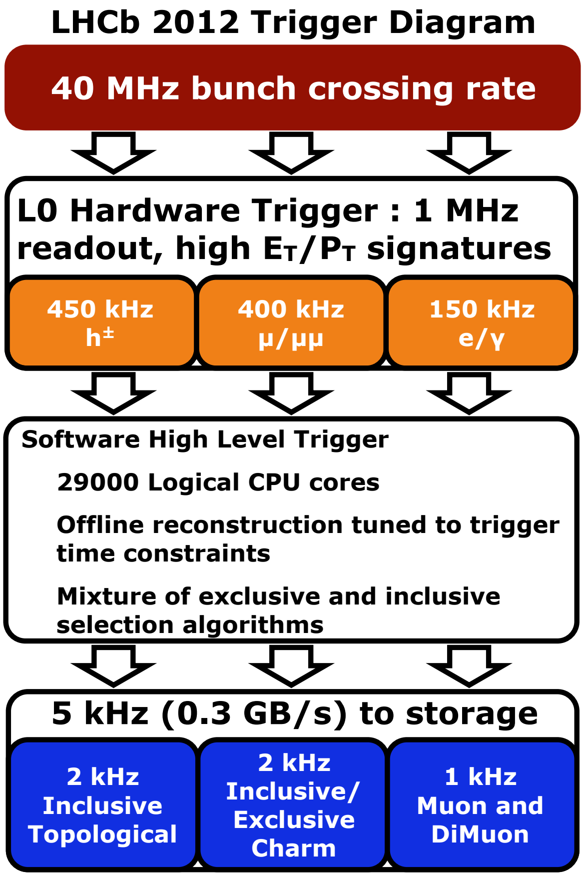


Figure 27.: LHCb trigger schema. Figure taken from [24].

Hardware Trigger

b-hadron decays are characterized by particles with high p_T spectra and larger impact parameter with respect to particles produced in the PV, respectively due to the large values of B mass and lifetime. The L0 trigger has to spot these general signatures of interesting decays and to reduce the rate of accepted events to 1 MHz. To reject events with multiple interactions, which need long times to be processed, the L0 trigger cuts on the maximum number of hits in the SPD. An event is accepted by the L0 trigger if at least one of this condition is satisfied:

- **L0-Muon**: a track in the Muon detector (i.e. one hit in each of the 5 stations) with transverse momentum (p_T) above a given threshold;
- **L0-DiMuon**: two tracks in the Muon detector with $p_T^1 \cdot p_T^2$ larger than a given threshold;
- **L0-Hadron**: a cluster in HCAL with transverse energy above a given threshold;
- **L0-Photon**: a cluster in ECAL with transverse energy above a given threshold;
- **L0-Electron**: a cluster in ECAL with transverse energy above a given threshold, anticipated by hits in the PS and at least one hit in the SPD.

Table 7.: Typical L0 thresholds used in Run1. Table taken from [32].

	p_T or E_T		SPD
	2011	2012	2011 and 2012
L0-Muon [GeV/c]	> 1.48	> 1.76	< 600
L0-DiMuon [(GeV/c)^2]	$> (1.296)^2$	$> (1.6)^2$	< 900
L0-Hadron [GeV]	> 3.5	> 3.7	< 600
L0-Photon [GeV]	> 2.5	> 3	< 600
L0-Electron [GeV]	> 2.5	> 3	< 600

The L0 thresholds are determined in order to maximize the trigger efficiencies of particular benchmark decays. The threshold values for 2011 and 2012 years of data taking are different (see Tab.7), given the different center of mass energy and so different $b - \bar{b}$ cross-section over the total one.

Software trigger

The events that pass the L0 trigger are processed by the software trigger which is subdivided in two stages HLT1 and HLT2. It runs on a large computing cluster called Event Filter Farm (EFF). The HLT1 performs a partial reconstruction of L0 candidates, thus the requirements are applied to a subset of the final particles. Exploiting the information from the VELO and the T1-T3 stations tracks with high transverse momentum and high displacement from the primary vertex are searched. Peculiar requirements are used depending on the L0 decision. Candidates which satisfies the requirements obtain a positive **1Track** decision. The HLT2 performs a more precise reconstruction using exclusive and inclusive algorithms. There are inclusive topological algorithms and they reconstruct b-hadrons decays looking at 2-, 3- or 4-body vertex and are labeled as **2-, 3-** or **4-body TopoBBDT**. Topological algorithms requires a secondary vertex with a large value of the scalar sum of the transverse momenta (p_T) of the tracks, and a significant displacement from the primary interaction vertex. At least one of the tracks used to form this vertex is required to have $p_T > 1.7$ GeV/c, an impact parameter (IP) χ^2 with respect to the primary interaction$>$ 16, and a track fit χ^2 per degree of freedom $\chi^2/ndf < 2$. The χ^2_{IP} is defined as the difference between the $\chi^2$of the PV reconstructed with and without the considered track. A multivariate algorithm is used for the identification of the secondary vertices The remaining trigger lines are inclusive or exclusive algorithms devoted to reconstruct c-hadrons decay. One of the inclusive trigger lines is the **InclPhi** that selects $\phi(1020) \to K^+K^-$ decays without any request on the mother particle. Either the topological algorithms and the **InclPhi** one are used by the B $_s \to D_s h$ analysis.

3.2.8 *Reconstruction algorithms*

The reconstruction of tracks and primary vertices is based on several algorithms of pattern recognition. Starting from the charged particle hits in the detector, a track is formed using different tracking strategy depending on the type of track. The tracks are classified as:

- **VELO tracks** : signatures only in the VELO detector. They are useful to determine the primary vertex and are reconstructed as straight line since they do not pass through the magnetic field;
- **Upstream tracks** : they contain hits in the VELO and in the TT detectors. They are usually low momentum tracks that are bent out of the LHCb detector acceptance by the magnetic field;

- **Downstream tracks** : tracks reconstructed with the TT and the T1-T3 tracking stations. They are useful to reconstruct long lived particles such as K_S and Λ which often decay after the VELO;

- **Long tracks** : they have hits in all sub-detectors, from the VELO to the T1-T3 stations. They have the most precise momentum measurement and thus they are the most important for b-physics measurements;

In order to find the best estimate of the track parameters a fit, called Kalman fit, is performed to all found tracks. Within the fit procedure a multiple scattering correction is applied as well as possible energy loss effects. Then the best track is created and all the other ones are discharged. Each track and its related uncertainty are defined by a state vector and its covariance matrix. The relevant track parameters are the momentum, the change of slope and the track point which is closest to the PV.

3.2.9 *The LHCb software*

The LHCb experiment makes use of several software packages to generate simulated data and to analyze collected data. The software is based on ROOT [39] and Gaudi [40] frameworks. The main packages are:

- **GAUSS** [41]: is responsible of the events simulation. It is structured in several packages which take care of different processes. PYTHIA is used to generate the pp interaction, while the hadrons decays are reproduced by EvtGen [42]. The simulation of the particle interaction with the LHCb detector is described by GEANT4.

- **BOOLE** [43]: reproduce the digitization of the front end electronics and the detector responses in simulated events.

- **MOORE** [44]: simulates the response of the software trigger stage

- **BRUNEL** [45]: performs the full events reconstruction starting from raw data. It is used for real and simulated events.

- **DaVinci** [46]: reconstructs the decay chain combining together the different particles. It can use the decay tree fitter (DTF) to refit all the vertices involved in the chain applying kinematical constraints, if required or useful. The DaVinci project can be used for simulated and real data.

- **URANIA** [47]: is a collection of packages devoted to the data analysis. The physics analysis presented in this thesis is a based on the B2DXFitters package, which is a part of the Urania project.

4

FLAVOUR TAGGING

Contents

4.1 Flavour Tagging definitions **51**
4.2 Flavour Tagging algorithms **56**
4.2.1 NNetSSK tagger 56
4.2.2 OS tagger 61
4.2.3 Mistag probabilities and combination of taggers 62
4.3 Flavour Tagging calibration **63**

As it has been described in chapter 2, CP violation and mixing measurements of neutral B mesons decays require the determination of the initial flavour of the reconstructed B meson, i.e. whether it contained a b ($\overline{B}^0$) o $\bar{b}$ (B^0) quark at production time. To infer the B initial flavour it is not enough the knowledge of its decay product charge because of the $B_q^0 - \overline{B}_q^0$ oscillations in time and the fact that the final state could be in common to both B_q^0 and $\overline{B}_q^0$ decay. The technique that allows to determine the production B flavour is referred to as *Flavour Tagging* (FT). Given the physics purposes of the LHCb experiment, the FT represents an important tool employed in all the time-dependent analysis. This chapter starts describing the Flavour Tagging principles and its related observables. Then two particular taggers are taken under consideration, the Neural Network Same Side Kaon (NNetSSK) [48] and the Opposite Side (OS) [49].

4.1 FLAVOUR TAGGING DEFINITIONS

At LHC, b and $\bar{b}$ quarks are produced in pairs by strong interaction in pp collisions. Each quark hadronise independently and can produce different species of b-hadrons: B^+, B_d^0, B_s^0, B_c^+ mesons or Λ_b baryon. FT algorithms determine if a given B signal decay contains a b or an $\bar{b}$ quark at production time by exploiting the properties of correlated charged particles (tagging particle) reconstructed in the same collision. At LHCb several FT algorithms exist and they are classified as *Opposite Side* (OS) taggers and *Same Side* (SS) taggers. In Fig.28 a schematic representation of the different taggers developed in LHCb is shown. Opposite Side taggers exploit the decay of the other b-hadron to infer the flavour of the signal B meson. Different OS algorithms exist: they exploit different decay processes of the opposite side b-hadron: OS Kaon inclusively reconstructs $b \to c \to s$ transitions, OS electron and OS muon inclusively reconstruct leptonic decay, while the OS charge vertex looks for the whole charge of the secondary vertex. An additional OS tagger recently developed, the OS charm [50], is based on the semi-exclusive reconstruction of secondary charged charm hadrons from the decay of the other b-hadron. In contrast, Same Side taggers aim to select charged particles that are produced in the same hadronization process of the signal B meson. Depending on the species of the B meson, different types of particle can be produced: for B^0_d SS pions and protons, while for B^0_s SS Kaons. In Fig.29 the charge correlation among the B^0_s , B^0_d and B^+ mesons with the SS tagging particles is shown.

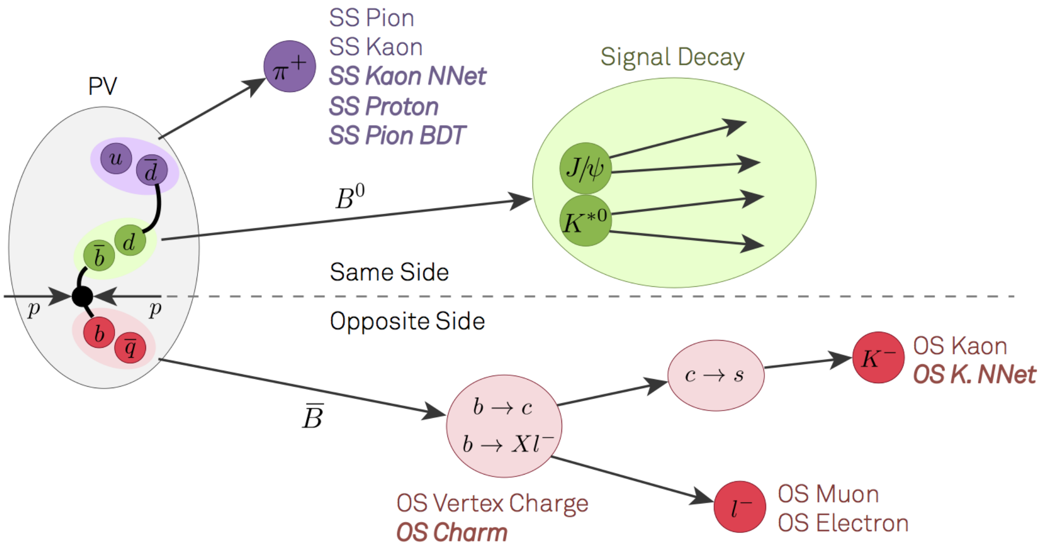


Figure 28.: Schematic representation of the possible taggers to infer the B flavour. The image summarizes the case of the $B^0 \to J/\psi K^{*0}$ decay channel.

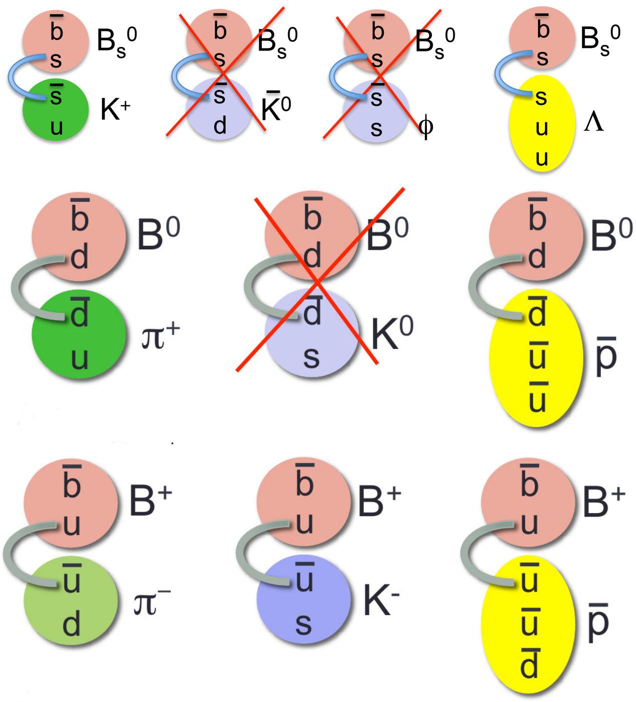


Figure 29.: Schematic representation of the charge correlation between B_s^0 , B_d^0 and B^+ and the possible tagging particles.

Given a reconstructed B signal candidate, which can be a B_d^0 , B_s^0 and a B^+, each FT algorithm provides a tag decision d based on the charge of the selected tagging particle. For initial B signal containing a $\bar{b}$ quark the tag decision is +1 (-1), while if the algorithm failed to find the tagging particle the tag decision is 0 and the B candidate results untagged. For each non-zero tag decision the algorithm estimates the probability of the provided tag decision to be wrong. The mistag probability, η, ranges from 0.0 (perfectly tagged) to 0.5 (randomly tagged). The performances of flavour tagging algorithms can be studied in $B^{\pm}$ decays and in B_d^0 and B_s^0 flavour specific decays and they are driven by two quantities: the tagging efficiency ϵ_{tag} and the mistag fraction ω. The tagging efficiency is defined as:

$$\epsilon_{tag} = \frac{N_T}{N_T + N_U}, \tag{4.1}$$

where N_T stands for the number of tagged events and N_U stands for the number of untagged events. The mistag ω is the fraction of events with the wrong tagging decision which corresponds to η if the tagger is calibrated:

$$\omega = \frac{N_W}{N_R + N_W} \tag{4.2}$$

where N_W and N_R are respectively the numbers of wrong and right tagged events. The mistag can be easily measured with $B^{\pm}$ meson decays comparing the tag decision with the charge of the B_q^0 meson itself. For neutral B decays, to flavour specific final states, due to $B_q^0 - \overline{B}_q^0$ mixing the mistag can be extracted only by a fit to the decay time distribution.

The fraction of wrong tagged events ω and the tagging efficiency ϵ_{tag} affect the measured decay rate with respect to what is theoretically expected. The decay rates of $B_q \to f$, $B_q \to \overline{f}$ and the respective CP conjugates decays, become:

$$\begin{aligned}
\Gamma_{obs}^{Tagged}(B_q(t) \to f) &= \epsilon_{tag}[(1-\omega)\Gamma(B_q(t) \to f) + \omega\Gamma(\overline{B}_q(t) \to f)] \\
\Gamma_{obs}^{Tagged}(\overline{B}_q(t) \to f) &= \epsilon_{tag}[(1-\omega)\Gamma(\overline{B}_q(t) \to f) + \omega\Gamma(B_q(t) \to f)] \\
\Gamma_{obs}^{Tagged}(\overline{B}_q(t) \to \overline{f}) &= \epsilon_{tag}[(1-\omega)\Gamma(\overline{B}_q(t) \to \overline{f}) + \omega\Gamma(B_q(t) \to \overline{f})] \\
\Gamma_{obs}^{Tagged}(B_q(t) \to \overline{f}) &= \epsilon_{tag}[(1-\omega)\Gamma(B_q(t) \to \overline{f}) + \omega\Gamma(\overline{B}_q(t) \to \overline{f})] \\
\Gamma_{obs}^{Untagged}(t) &= (1-\epsilon_{tag})[\Gamma(\overline{B}_q(t) \to f) + \Gamma(B_q(t) \to f)] \\
\Gamma_{obs}^{Untagged}(t) &= (1-\epsilon_{tag})[\Gamma(B_q(t) \to \overline{f}) + \Gamma(\overline{B}_q(t) \to \overline{f})]
\end{aligned} \tag{4.3}$$

where Γ_{obs} represents the observed decay rate. The first four expressions are the tagged decay rates while the last two represents the decays rates for the untagged events. For simplicity in these equations it is assumed the same mistag and ϵ_{tag} for initial B_q^0 and $\overline{B}_q^0$ mesons ($\omega = \overline{\omega}$ and $\epsilon_{tag} = \overline{\epsilon}_{tag}$). The time-dependent amplitude for the ϵ_{tag} fraction of events is reduced by the dilution factor, $(1 - 2\omega)$ that depends on the mistag:

$$\begin{aligned}
\mathcal{A}_{obs}^{f} &= \frac{\Gamma_{obs}(B_q(t) \to f) - \Gamma_{obs}(\overline{B}_q(t) \to f)}{\Gamma_{obs}(B_q(t) \to f) + \Gamma_{obs}(\overline{B}_q(t) \to f)} \\
&= (1-2\omega)\frac{\Gamma(B_q(t) \to f) - \Gamma(\overline{B}_q(t) \to f)}{\Gamma(B_q(t) \to f) + \Gamma(\overline{B}_q(t) \to f)} = (1-2\omega)\mathcal{A}^f = D\mathcal{A}^f
\end{aligned} \tag{4.4}$$

where $\mathcal{A}^f$ is the theoretical mixing amplitude. The effect of the dilution factor is to reduce the amplitude of the oscillation in the decay rate as can be seen in Fig.30

where each plot shows the oscillated and non-oscillated B decay rates in the ideal case and in case of $\omega = 0.3$. The sensitivity to $\mathcal{A}^f$ depends on the mistag as follows:

$$\sigma_{\mathcal{A}^f} = \frac{\sigma_{\mathcal{A}^f_{obs}}}{1 - 2\omega} \tag{4.5}$$

where the error on ω has been neglected.

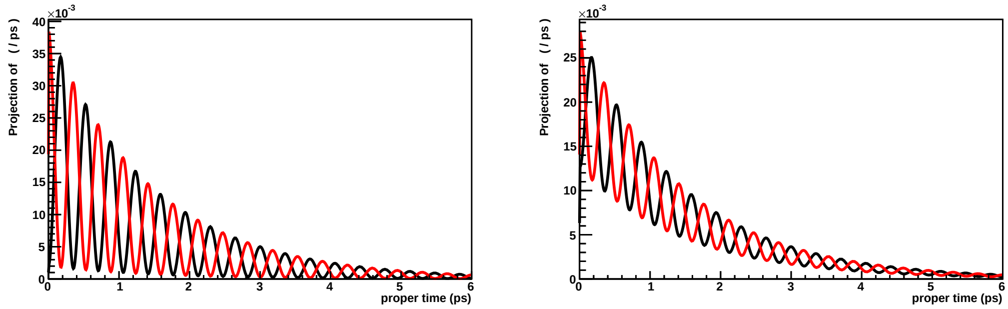


Figure 30.: The black curve represents the oscillated B mesons while the red curve represents the non-oscillated events. The left plot shows the case of an ideal measurement in case of perfect tagging. The right plot corresponds to a mistag of 0.3.

Determining the $\sigma_{\mathcal{A}^f_{obs}}$ through error propagation, one can express se sensitivity to the amplitude as:

$$\sigma_{\mathcal{A}^f} = \frac{\sqrt{1 - \mathcal{A}^{f\,2}_{obs}}}{\sqrt{\epsilon_{eff} N}} \tag{4.6}$$

which shows that to minimize the error on the amplitude the *effective efficiency* or *tagging power*

$$\epsilon_{eff} = \epsilon_{tag}(1 - 2\omega)^2 \tag{4.7}$$

must be maximized. Thus the figure of merit that measures the performance of flavour tagging algorithms is the tagging power. At LHCb, each tagging algorithm has been optimized in order to maximize the tagging power (Eq.4.7) using data samples or simulated events. The estimated mistag probability η needs to be calibrated to well represent the measured mistag ω. More details on the calibration will follows in Sec.4.3. The mistag probability can be used to assign larger weights to events with low mistag probability and thus to increase the overall significance of an asymmetry measurement. Indeed it has been proved that the ϵ_{eff} results improved when it is evaluated on a per-event basis:

$$\epsilon_{eff} = \frac{1}{N}\sum_{i=1}^{N} D_i^2 = \frac{1}{N}\sum_{i=1}^{N}(1-2\omega(\eta_i))^2 \tag{4.8}$$

where N is the total number events and $\omega = 0.5$ (D=0) for untagged events.

4.2 FLAVOUR TAGGING ALGORITHMS

The implementation of flavour tagging algorithms can be described in three steps: pre-selection, selection of tagging particles and evaluation of the tagging decision and of the mistag probability η.

The preselection is common to all tagging algorithms and it is devoted to select tracks which come from the same pp collision of the signal B but it should not be a decay product of the signal. Minimal requirements are imposed on the goodness of the track reconstruction and on the kinematics. More details on the pre-selection criteria are reported in Tab 8. The selection of the tagging tracks and the mistag probability estimation strictly depends on the specific taggers. For the purpose of the $B_s^0 \rightarrow D_s K$ analysis and for the research activity I performed, in the next sections two particular taggers are described with much more details. They are the NNetSSK and the combination of the different OS taggers.

4.2.1 *NNetSSK tagger*

The NNetSSK tagger is a new kind of tagger developed by LHCb in recent years [48]. Given the better performance, it replaces the previous SSK tagger which used to select tracks candidates for the tagging particles on a cut-based approach. The NNetSSK tagger employs two Neural Networks, referred as NN1 and NN2 for brevity. The aim of NN1 is to distinguish fragmentation tracks correlated to the signal B from tracks originating from the OS b-hadron decay and tracks originated from soft QCD processes in pp collisions. Both these two last categories of events

Table 8.: Pre-selection requirements common to all tagging algorithms. All the cuts refer to the Reco14 optimisation. The track type are defined in Sec.3.2.1. The angle θ is defined with respect to the beam pipe; $|\Delta\phi|$ is the difference in the polar angle ϕ between the tagging track candidate and each reconstructed decay product of the signal B decay; The IP significance is computed with respect to the primary vertex associated to the signal B (PV) or to possible additional primary vertices (PU, pileup). The requirements on $\Delta\phi$ vetoes all reconstructed daughter particles of the signal B.

Pre-selection
$p \in [2, 200]\mathrm{GeV/c}$
$p_T < 10$ GeV/c
long or upstream track
charge: $\pm$
$\theta > 12\mathrm{mrad}$
$\lvert\Delta\phi\rvert > 5$ mrad
min. $\mathrm{IP}_{\mathrm{PU}}$ significance > 3
$\notin$ signal decay chain
ghost probability < 0.5
not a clone particle

are referred as underlying events (UE) in the following. The output of NN1, $nn1$, is the probability, on a per-track base, that the selected charged track corresponds to the fragmentation kaon accompanying the signal B_s^0. For the successful training of the NNet classifier it is important that all the different categories of tracks are present in the data sample. NN2 provides the final tagging decision and the relative mistag probability combining the information from tracks selected by NN1. This approach has the advantage with respect to the old SSK tagging algorithm that multiple tracks with similar quality can be selected and handled. The $nn1$ output and the particle ID information of these tracks are inputs of the NN2 as well as the number of tracks in the events, the number of Primary Vertices (PVs), and also the p_T of the B_s^0 candidate. The NNetSSK tagger is developed using a large

Table 9.: Input variables to the NN1 for the discrimination between the fragmentation tracks and the UE tracks. The relative transverse momentum defined as $p_{\mathrm{T_{rel}}} = \sqrt{p^2_{\mathrm{track}} - p^2_{\mathrm{L_{rel}}}}$, where the longitudinal momentum of the tagging track is defined as $p_{\mathrm{L_{rel}}} = \vec{p}_{B^0_s} \cdot \vec{p}_{\mathrm{track}} / \left|\vec{p}_{B^0_s}\right|$, and the distance in the polar angle, ϕ, and the pseudorapidity, η, between the tagging track candidate and the B^0_s candidate, $\Delta\phi$ and $\Delta\eta$, respectively.

Variable
number of PVs
IP
$\mathrm{IP}/\sigma_{\mathrm{IP}}$
$\chi^2_{\mathrm{track}}/\mathrm{ndf}$
$p_{T_{rel}}$
$\Delta\phi$
$\Delta\eta$
p
p_T
number of tagging track candidates
p_{T,B_s}

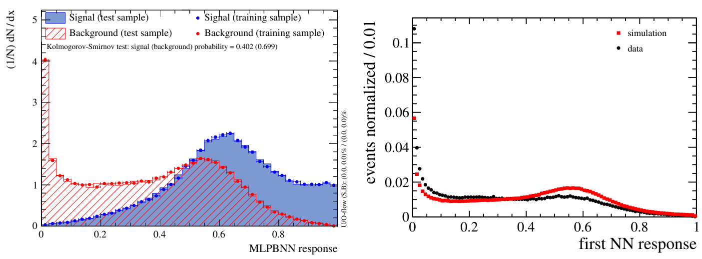
samples (160K) of MC $B^0_s \to D_s\pi$ reconstructed events. The MC sample has been reweighted to match the kinematic distribution of data (Number of PV, Number of tagging tracks, Impact parameter). The information at the generator level are used in the training of the neural networks. The relative tagging performance are then evaluated using $B^0_s \to D_s\pi$ data samples which are also used to calibrate the predicted mistag $\omega(\eta)$ (see Sec.4.3).

NNetSSK implementation

The first neural network NN1 has to disentangle fragmentation tracks from all the other tracks that can be globally referred to as underlying events (UE). In Tab.9 the input variables used to train NN1 in its final configuration are listed.

To facilitate the learning process of the NN1, we require that the tagging track candidates pass additionally to the common tagging preselection minimum kaon PID requirements, i. e.$DLL(K-\pi) > 0.75$ and $DLL(K-p) > -8.5$, enriching in Kaon candidates.

Figure 31a shows the NN1 response for fragmentation tracks (signal) and UE tracks (backgrounds), for both the training and the test samples. Indeed the original sample has been divided in a training sample used for the learning process of the neural network and in a test sample used to verify that the configured neural network does not suffer of overtraining effects. A good discrimination is visible between the background and the signal. The distributions of the training sample and the test sample are in good agreement, which serves as an indication that the NN1 is not overtrained. As the training procedure relies on simulated events, we check the NN1 output distribution in data. Figure 31b shows that the NN1 responses for all tagging track candidates in the simulation and in the data are similar. The visible remaining differences are likely to be due to residual discrepancies in the distributions of input variables, especially to the differences in the $\chi^2_{\text{track}}/\text{ndf}$ from the track fit.



(a) (b)

Figure 31.: Verification plots for the training of NN1. Shown are (a) the NN1 output distribution for fragmentation tracks (signal, blue) and UE tracks (background, red dashed area) in simulations and (b) the comparison of the NN1 response in the data and the simulation.

Tagging track candidates with NN1>0.65 are then considered as input to the NN2. This selection removes a large fraction of background tracks and corresponds to a tagging efficiency larger than 50%. This requirement is chosen such that for about 50% of the B_s^0 signal candidates there is at least one tagging track candidate left. In MC this corresponds to a tagging efficiency of about 50%, and in data of about 65%. Figure 32 shows the number of tagging track candidates per event that survive this cut for all B_s^0 candidates (tagged and untagged) in the data and in the simulation. For the tagged candidates, on average there are 1.6 tracks per-event in the data and 1.4 tracks in simulation, respectively. Before the cut, the sample of tagging track candidates in the MC contains 0.9 fragmentation tracks and 8.2 UE tracks; after the cut, 0.5 fragmentation tracks and 1.1 UE tracks.

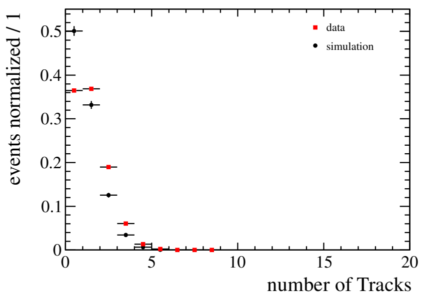
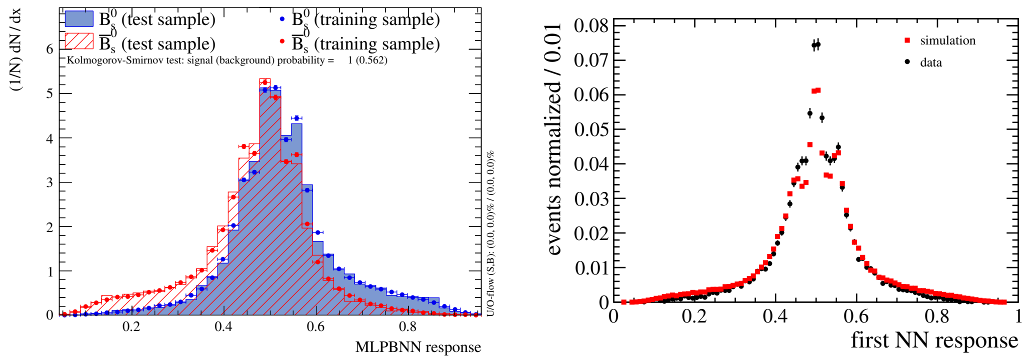


Figure 32.: Multiplicity of tagging track candidates selected by NN1 for data (red) and simulation (black). The first bin represents untagged candidates.

The second neural network, NN2, determines the tagging decision and the mistag probability for the flavour of B_s^0 and $\bar{B}_s^0$ among the tagging tracks candidates which pass the NN1 selection. Its input variables are: the number of tagging track candidates, the transverse momentum of the B_s^0, the number of PVs, and the $q \times DLL(K-\pi)$ and $q \times nn1$ variables for the individual tracks (q is the charge of the kaon tagging candidate). The final configuration accepts events with up to three tagging tracks candidate. If no second or third tagging track candidate is available the corresponding input variables to NN2 are set to zero. Figure 33a shows the NN2 output distribution for signal candidates produced as B_s^0 and signal candidates produced as $\overline{B}_s^0$, for both the training and test samples. The distributions feature a good discrimination between the two flavours, especially in the tails, where the mistag probability is lower. In absence of CP and K^+/K^- asymmetries the output of NN2 can be directly interpreted as the mistag probability where the $nn2$ distributions for B_s^0 and $\overline{B}_s^0$ are mirrored at $nn2 = 0.5$ as shown in Fig.33a. Thus $\eta = nn2$ and $d = -1$ for $\overline{B}_s^0$, $\eta = 1 - nn2$ and $d = +1$ for B_s^0.



(a) Distribution of NN2 response for B_s^0 and $\overline{B_s^0}$ events.

(b) Convergence of NN training.

Figure 33.: Verification plots for the training of the second neural network. Shown are the response for B_s^0 and $\bar{B}_s^0$ in the simulation (a) and the comparison of the response of the NN2 in the data and the simulation (b).

4.2.2 *OS tagger*

Opposite side tagging algorithms [49] exploit the flavour of the B signal meson looking at the decay products of the opposite b-hadron. In the case of single particle algorithms the particles are muon, electron and kaons, while vertex charge tagging algorithm calculates the weighted charge of the tracks that originates from a common opposite side secondary vertex. Single particle tagging algorithms select a $\mu^\pm$, $e^\pm$ from semi-leptonic decay of the opposite b-hadron or $K^\pm$ from the $b \to c \to s$ transition making some requests to ensure that they come from a b-hadron decay such as a large impact parameter significance with respect to the primary vertex (IP/σ_{IP}) and a large transverse momentum p_T. Also particle identification is used to define each tagger. In addition a cut on the Ghost Probability is applied. This is the probability for a track to be made from a random combination of hits. For each of the OS tagging algorithms the selection is optimized by means of an iterative procedure in which the value of ϵ_{eff} is plotted as a function of the cut value of a given observable. The cut value that maximises the ϵ_{eff} is chosen as a reference. Given that the OS is independent on the signal, the procedure has been applied on data sample of $B^+ \to J\psi K^+$ decays. The tagging performance for each OS tagger are reported in Tab.10.

Table 10.: OS tagging algorithms performance from optimization in $B^+ \to J/\psi K^+$ decay mode.

Algorithm	ϵ_{tag} [%]	ω [%]	ϵ_{eff} [%]
OSμ	5.43 ± 0.05	29.97 ± 0.39	0.87 ± 0.04
OSe	1.62 ± 0.03	29.46 ± 0.70	0.27 ± 0.02
OSK	17.08 ± 0.07	39.16 ± 0.24	0.80 ± 0.04
OSvtx	17.76 ± 0.08	39.59 ± 0.23	0.77 ± 0.03
OS	30.95 ± 0.09	-	2.75 ± 0.08

The estimation of per-event mistag probability for the decision to be wrong (η) is done using a multivariate classifier (neural networks) that exploits kinematic and geometric properties of the tagging track as long as properties of the event and of the signal B. The multivariate classifier is trained to identify as signal the tracks that give the right tag decision based on the comparison between the charge of the tagging particle and of the charge of the signal. In order to determine the optimal performance, the data sample used should contain only signal events. This is guaranteed by applying a selection that minimise the background and by using the *sWeights* technique to produce background-subtracted samples.

4.2.3 *Mistag probabilities and combination of taggers*

As we have seen before, for each tagger the probability of the tag decision to be wrong is estimated event by event by using a neural network (or in general a multivariate classifier) that combines the properties of the tagger and of the event itself. If there is more than one tagger available per event, their decisions and mistag probabilities are combined to provide a final decision and mistag probability on the initial flavour of the signal B. In this way the overall sample of tagged events gets increased. The combined probability $P(b)$ that the signal contains a b-quark is calculated as:

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \quad P(\bar{b}) = 1 - P(b) \tag{4.9}$$

where

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i(1 - \eta_i) \right), p(\bar{b}) = \prod_i \left(\frac{1 - d_i}{2} + d_i(1 - \eta_i) \right) \tag{4.10}$$

Here, d_i is the decision taken by the $i - th$ tagger based on the charge of the particle with the convention $d_i = 1(-1)$ for the signal B containing a $\bar{b}(b)$ quark and η_i the corresponding predicted mistag probability. The combined tagging decision and the corresponding mistag probability are $d = -1$ and $\eta = 1 - P(b)$ if $P(b) > P(\bar{b})$, otherwise $d = +1$ and $\eta = 1 - P(b)$. Due to the correlation among taggers, which is neglected in Eq.4.10, the combined probability is slightly overestimated. The largest correlation occurs between the vertex charge tagger and the other OS taggers, since the secondary vertex may include one of these particles. To correct for this overestimation, the combined OS probability is calibrated as well.

4.3 FLAVOUR TAGGING CALIBRATION

One important element of the tagging is the mistag probability η: it can be used to weight each event increasing the ϵ_{eff}. The last step regards the possibility to further improve the tagging performances calibrating the estimated mistag η in order to match the observed mistag. Each tagging algorithm has been developed and optimized using one particular decay mode which can be different from tagger to tagger. Indeed the NNetSSK has been optimized on $B_s^0 \to D_s^- \pi^+$ simulated events while the OS taggers on $B^+ \to J/\psi K^+$ data. To check the mistag estimation provided by the tagger and to improve the portability of each optimized tagger to other decay modes the measured mistag ω is determined as a function of the η estimate. Usually the calibration function is defined as:

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - < \eta >) \tag{4.11}$$

where p_0 and p_1 are the calibration parameters and $< \eta >$ is the measured value of η average. This is important also to make combination of tagging responses when more than one tagger is available. The calibration procedure is done on data of flavour specific decays. Indeed the charge of the final state f is correlated to the flavour of the B meson, thus in this kind of decays it is possible to perform a direct measurement of the mistag. In order to distinguish signal events of the flavour specific decay we are interested in from background events, an invariant mass fit of the B meson is performed extracting signal weights. By means of the *sPlot* technique [52], applying the *sWeights* event by event a statistical background subtraction is performed obtaining a sample equivalent to pure signal events. In $B^{\pm}$ decays (which are flavour specific by definition) the measured ω can be evaluated comparing the flavour of the reconstructed $B^{\pm}$ meson to the charge of the tagging particle. If they have the same charge the tagging decision is right, otherwise it is wrong. Using Eq.4.2, the η distributions for right and wrong tagged events are used to display the measured mistag fraction ω as a function of η and the tagging calibration parameters are extracted using Eq.4.11. For neutral flavour specific decay modes, after the invariant mass fit, a fit to the flavour oscillation asymmetry as a function of the B decay time is performed. The mistag ω measured fitting the oscillation amplitude as reported in equations 4.3. Of course the mixing frequency, Δm_q, as well as the B^0_q lifetime (q = d,s) and $\Delta\Gamma_s$ are taken from independent measurements. Examples of flavour-specific B decay modes are $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$, $B^0 \to D^{(*)-}\mu^+\nu_\mu$, $B^+ \to D^0\pi^+$ and $B^0_s \to D^-_s\pi^+$. A first calibration of the NNetSSK tagger using the $B^0_s \to D^-_s\pi^+$ 2012 data has been obtained from a preliminary study and the results of this calibration ($p_0 = 0.436$, $p_1 = 0.9$ and $\langle\eta\rangle = 0.430$) have been applied internally to the Flavour Tagging code. Also the OS tagger is pre-calibrated in the Flavour Tagging code applying the reference calibration parameters found in the $B^+ \to J/\psi K^+$ decays. The results presented in the next Chapter are obtained from the mistag derived after this internal pre-calibration of the NNetSSK tagger (which is applied in DaVinci). Systematic uncertainties on the flavour tagging calibration parameters must be considered by the time-dependent analyses for CP-violation. Systematics uncertainties are due to systematics effects related to the fit model used to extract the calibration parameters p_0 and p_1, but also to kinematic differences among the calibration channels and the other analysis channels. There could be a dependence of the flavour tagging calibration parameters on the B initial flavour. This initial flavour asymmetry in the tagging calibration is measured. Detailed and quantitative information on flavour tagging calibrations of NNetSSK and OS using the flavour specific decay $B_s^0 \to D_s^- \pi^+$ are given in the next chapter.

5

FLAVOUR TAGGING CALIBRATION USING $B_s^0 \to D_s^- \pi^+$ RUN1 DATA

Contents

5.1 Fit strategy **68**

5.1.1 Mass Fit 69

5.1.2 Time sFit 72

5.2 Calibration of the NNetSSK tagger **75**

5.2.1 Flavour tagging asymmetries 78

5.2.2 Systematic uncertainties 81

5.2.3 Test in MC samples 85

5.2.4 NNetSSK calibration summary and performances 86

5.3 Calibration of OS tagger **89**

In order to use the mistag event by event and to take advantage on the combination of different taggers, the predicted mistag η must be checked and calibrated with data. The Flavour Tagging strategy for the Run1 data is to compare calibrations from as many as possible control channels in order to verify and estimate the portability of the calibration from one channel to another one. Using simulated events the portability of the calibration is verified also for CP violating decays. The Flavour Tagging calibrations of the NNetSSK and the OS combination in the flavour specific decay mode $B_s^0 \to D_s\pi$ are described in this chapter. The relevance of this study finds its main reasons in the following topics:

- $B_s^0 \to D_s\pi$ is the main suitable flavour specific decay mode for calibrating the SSK taggers
- $B_s^0 \to D_s\pi$ is the main control channel for the $B_s^0 \to D_sK$ decay which allows to measure γ. This analysis uses the NNetSSK and the OS taggers to determine the flavour of the B_s^0 mesons.

The fit strategy to extract Flavour Tagging calibrations in the flavour-specific $B_s^0 \to D_s^- \pi^+$ decay mode is described in Sec.5.1, while Sec.5.2 and Sec.5.3 are dedicated respectively to the Flavour Tagging calibration of NNetSSK and OS taggers.

5.1 FIT STRATEGY

Since the $B_s^0 \to D_s^- \pi^+$ is a neutral channel, the mistag is measured performing a time-dependent measurement of the oscillation amplitude to measure the mistag. For this reason after the fit to the B_s^0 invariant mass, which discriminates signal events from backgrounds, a fit to the signal decay time distribution is performed. The analysis of $B_s^0 \to D_s^- \pi^+$ decays presented in this section is based on the full Run1 sample (3.0 fb^{-1}) of pp collisions data taken in 2011 ($\sqrt{s}$=7 TeV) and in 2012 ($\sqrt{s}$=8 TeV). The event selection and the fit procedure is based on the analysis of $B_s^0 \to D_s K$ and $B_s^0 \to D_s^- \pi^+$ decays for the measurement of the CKM angle γ performed using the 2011 sample (1fb^{-1}) [54] and it will be described in Chapter 6. The full Run1 analysis uses the events reconstructed by the Reco14 version of the software and that pass the B02DPiD2HHHBeauty2CharmLine stripping line selection of the Stripping $v20r0$ (2012) and $v20r1$ (2011) production. The offline event selection is based on the output of a BDT trained to recognize signal decays (BDT[1] output > 0.5) from background combinatorial events. In addition, a good identification of the pion from $B_s^0 \to D_s^- \pi^+$ (namely *bachelor* pion in the following) (DLL(K-π) < 0) is required. No specific requirements for the trigger are applied.

Firstly the B_s^0 mass distribution (m_{B_s}) of the selected B_s^0 candidates is fitted with a probability density function (PDF) to describe the signal and the different background contributions. Then, a fit to the B_s^0 decay time distributions of the events tagged as mixed, tagged as unmixed and untagged is performed, by using the sFit method [55]. This procedure exploits the cancellation of the background contribution that is obtained assigning to each event the signal *sWeights* determined by the mass fit. The requirement for this procedure to be applicable is that the mass

1 We use the BDT that optimised the selection for Reco12 sample, which might be not optimal for Reco14 data, but not wrong. Indeed the selected sample presents already a good background suppression and signal purity.

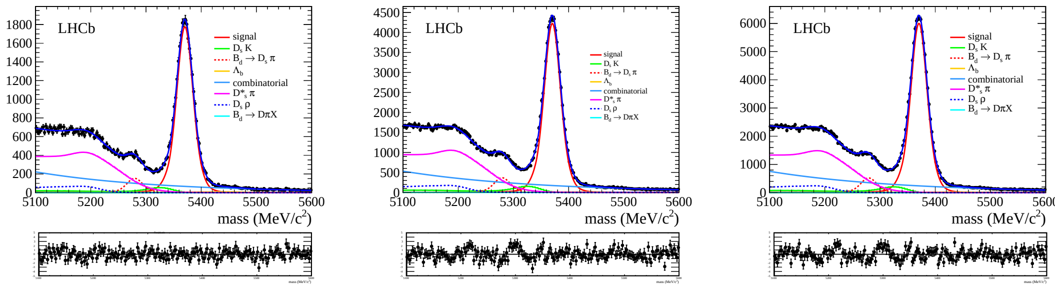
observable is not correlated from the observables of the sFit. This requirement was verified using simulated events both for signal and background decays.

5.1.1 *Mass Fit*

The B_s^0 mass distribution of the selected $B_s^0 \to D_s^- \pi^+$ events are reported in Fig.34 where the signal peak is clearly visible and at lower mass values partially reconstructed events are selected. The fit to the B_s^0 mass distribution has to distinguish signal events from backgrounds extracting the signal sWeights. The signal PDF is a Double Crystal Ball function, defined as follows:

$$PDF_{sig}(m) = fCB_1(\mu, \sigma_1, \alpha_1, n_1) + (1 - f)CB_2(\mu, \sigma_2, \alpha_2, n_2) \tag{5.1}$$

where the mean is common and left free in the fit to data as well as the relative fraction f. The tail parameters and the sigmas are fixed to the result of a fit to the $B_s^0 \to D_s \pi$ simulated data. The parameters values are listed in Table 11. The background contributions are of three types: the background due to mis-identified particles (mis-ID background) from wrongly reconstructed B_s^0 and Λ_b decays, the partially reconstructed B_s^0 decays and the combinatorial backgrounds. The $B^0 \to D_s \pi$, $B^0 \to D\pi$, $B_s^0 \to D_s K$ and the Λ_b decays rely on the mis-identification of one particle in the final state or in the mis-identification of the B meson. The $B_s^0 \to D_s \rho$ and $B_s^0 \to D_s^* \pi$ decays are partially reconstructed backgrounds. The parametrization of these backgrounds have been taken from simulated events, reweighted for the efficiency corrections due to the PID cut computed with the PIDCalib tool [56]. The combinatorial background is formed by those events in which the particles involved in the signal decay are randomly selected among the tracks of the pp collision. Combinatorial background is described by an exponential function. For simplicity, we don't split the data sample according to the D_s decay modes ($\phi\pi$, K^*K, non resonant $KK\pi$, $K\pi\pi$ and $\pi\pi\pi$) as it is done for the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis. It has been verified that the signal yields of the individual D_s decay sum up to the same value obtained fitting the mass distribution of the merged samples. For this reason we merge together the D_s decay modes in a single sample and both magnet polarities are included. The fit to the B_s^0 mass distribution is performed two times. A first time in the wide mass window [5100-5600] MeV/c^2 to determine the contributions of the different backgrounds and a second time in a narrow mass region [5320-5600] MeV/c^2 to extract the signal *sWeights*.



(a) 2011 (b) 2012 (c) 2011+2012

Figure 34.: Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the B_s^0 mass window [5100-5600] MeV/c^{2}used to determine the relative fraction for each background component.

The total mass PDF is:

$$PDF(m) = n_{sig} \cdot PDF_{sig}(m) + \sum_{k=1}^{N} n_k \cdot PDF_{bkg}^k(m) \tag{5.2}$$

where n_{sig} and n_k are respectively the signal and the individual background yields [2] while PDF_{sig} and PDF_{bkg} are the PDFs for signal and background events. Figure 34 show the fit to the mass distribution for 2011 34a, 2012 34b and 2011+2012 34c data samples.

All backgrounds yields are floating in the fit, except for the B^0 backgrounds different form $B^0 \to D_s^- \pi^+$ decays, which are set to zero in the fit to 2011+2012 data as it is found to be negligible in the fit to 2011 and 2012 samples. The background yields determined by the mass fit and rescaled to a narrow mass range of [5320;5600] MeV/c^{2}are reported in Tab.12 and they are used to determine the relative fraction $f_k = n_k/n_{tot}$ of each background component.

2 A poissonian term is implicitly considered in the fit to data by means of the extended maximum likelihood fit.

Table 11.: Parameters of the PDF describing the signal (Double Crystal Ball) and the combinatorial background obtained from the fit to $B_s^0 \to D_s^- \pi^+$ data for 2011, 2012 and 2011+2012 samples. Fixed parameters are obtained from a fit to simulated signal events.

PDF_{sig}(m)	fixed
n_1	1.1
n_2	5.8
α_1	2.13
α_2	-2.07
σ_1 [MeV/c^2]	12.69
σ_2 [MeV/c^2]	20.49

PDF_{sig}(m)	2011	2012	2011+2012
$\mu(B_s^0)$[MeV/c^2]	5370.5±0.1	5370.47±0.08	5370.46±0.07
f	0.530±0.015	0.51±0.01	0.514±0.009
n_{sig}	27 870±199	66960±311	94833±372

PDF_{comb}(m)	2011	2012	2011+2012
α [MeV/c^2]$^{-1}$	-0.0044±0.0002	-0.0042±0.0002	-0.0042±0.0001
n_{comb}	5355.4±191	14096±325	19462±376

The second mass fit in the narrow mass range [5320;5600] MeV/c^{2}determines the *sWeights* assigning to each event a weight corresponding to the hypothesis of signal or background. Only the signal and the total background yields are floating, while the fraction of each background contribution is fixed. The total PDF is described by Eq. 5.3-5.4:

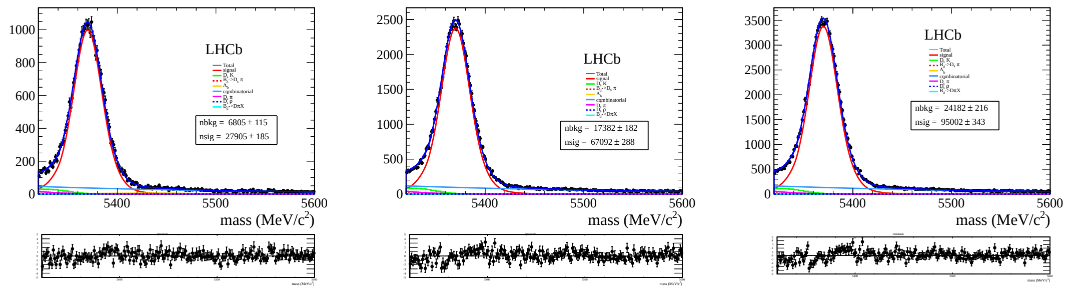
$$PDF_{tot}(m) = n_{sig} \cdot PDF_{sig}(m) + n_{bkg} \cdot PDF_{bkg}(m) \tag{5.3}$$

where PDF_{bkg}(m) is defined in the following way:

$$PDF_{bkg}(m) = \sum_{k=1}^{N-1} f_k \cdot PDF_{bkg}^{k} + (1 - \sum_{k=1}^{N-1} f_k) \cdot PDF_{bkg}^{N} \tag{5.4}$$

Table 12.: Yields of the background contribution in the mass range [5320:5600] MeV/c^{2}from the results of the fit to the [5100:5600] MeV/c^{2}mass range for 2011, 2012 and 2011+2012 samples.

background yield	2011	2012	2011+2012
$B^0 \to D_s \pi$	37.4±2.9	95.1±4.8	132.6±5.6
all other B^0 decays	0±121	0 ±98	0 (fixed)
$B_s^0 \to D_s K$	705±87	2001±141	2703±166
$B_s^0 \to D_s \rho$	2.9±0.7	7.6±1.2	10.5±1.4
$B_s^0 \to D_s^* \pi$	30.3±10.6	738±17	1040±20
$\Lambda_b \to \Lambda_c \pi$	478±135	710±227	1184±263



(a) 2011 (b) 2012 (c) 2011+2012

Figure 35.: Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the [5320-5600] MeV/c^{2}mass window used to extract the *sWeights*.

Figure 35 shows the result of the second mass fit where the total measured yields are n_{sig} = 95 002 ± 343 and n_{bkg} = 24 182 ± 216 in the merged sample 2011 + 2012.

5.1.2 *Time sFit*

The fit to the B_s^0 decay time, t, is performed using the sFit method [55]. The background events are subtracted applying the *sWeights* previously extracted from the mass fit, event by event. In this way the PDF is simplified to the signal component only. The signal PDF is based on the B_s^0 decay rate amplitudes to the different final states, $\Gamma(t)$, and accounts for several experimental factors, such as the mistag dilution, $1-2\omega$, the tagging efficiency ϵ_{tag} , the time resolution, $R(t-t')$ and the time acceptance, $a(t)$:

$$\mathrm{PDF}(t) = a(t)\left(\Gamma(t') \otimes R(t-t')\right) \ . \tag{5.5}$$

The decay rates are defined by Eq.2.38 where $|\lambda| = 0$ since the $B_s^0 \to D_s^- \pi^+$ is a flavour-specific decays and the equations can be simplified in this way:

$$\Gamma(t') \propto \epsilon_{tag}\, e^{-t'/\tau_s} \left(\cosh\left(\frac{\Delta\Gamma_s}{2}t'\right) + q^{\mathrm{mix}}\,(1-2\omega)\cos\left(\Delta m_s t'\right)\right) \tag{5.6}$$

is the amplitude for tagged events and

$$\Gamma(t') \propto \left(1-\epsilon_{tag}\right) e^{-t'/\tau_s} \cosh\left(\frac{\Delta\Gamma_s}{2}t'\right) \ , \tag{5.7}$$

the amplitude for untagged events; q^{mix} represents the mixing state, which is defined as the product of the tagging decision with the charge of the reconstructed final state (here q^{mix} assumes the meaning of $+1/-1$ for unmixed/mixed events). The CP violation in B_s^0 mixing ($|q/p| \neq 0$) is neglected.

The time resolution model $R(t-t')$ is a sum of three Gaussian functions, obtained by fitting the resolution distribution resulting from the sum of single Gaussian functions with widths sampled from the per-event decay-time error distribution of data. To correct for the measured underestimation of the decay-time error for data, an overall scale factor of 1.37 is applied to each widths [54]. Studies performed on different decay channels indicate that the time resolution has not undergone relevant changes between Reco12 and Reco14 data; thus, we assume the parametrization used for Reco12 data reported in Table 13. The average effective time resolution, evaluated as:

$$< \sigma_{eff} >= SF \cdot \sqrt{f_1\sigma_1^2 + f_2\sigma_2^2 + (1-f_1-f_2)\sigma_3^2} \tag{5.8}$$

is of 49 fs.

Table 13.: Parameters of the triple gaussian time resolution model determined from a fit to the $B_s^0 \to D_s^- \pi^+$ events. The scale factor represents the multiplicative factor applied to the sigmas in order to describe the resolution in data (see Ref. [59]). The average effective time resolution is reported in the last row.

parameter	value
μ [ps]	0
σ_1 [ps]	0.0215
σ_2 [ps]	0.0368
σ_3 [ps]	0.0638
f_1	0.372
f_2	0.565
Scale factor	1.37
$< \sigma_{eff} >$ [ps]	0.049

The selection of events (trigger, offline selection) introduces an acceptance function on the decay time distribution. Its shape $a(t)$ has been derived in MC and corresponds to the parametrization described in [60]:

$$a(t) = \begin{cases} 0 & : \text{ when } (\mathrm{at})^\mathrm{n} - \mathrm{b} < 0 \text{ or } \mathrm{t} < 0.2\mathrm{ps} \\ (1 - \frac{1}{1+(at)^n - b}) \cdot (1 - \beta t) & : \text{ otherwise} \end{cases} \tag{5.9}$$

The acceptance parameters are determined by fitting the $B_s^0 \to D_s^- \pi^+$ data with the B_s^0 lifetime fixed to the reference value of PDG 1.466 ps [62] in the nominal fit for the extraction of the tagging calibration. The value of Δm_s is fixed to the LHCb measurement of 17.768 ps^{-1} [64]. No differences have been seen in the calibration parameters if Δm_s is allowed to float within the measurement uncertainty ($\pm$0.0238 ps^{-1}), by means of a gaussian constraint in the fit. The $\Delta\Gamma_s$ parameter is fixed in the fit to: 0.1 ps^{-1} [61]. Their central values have negligible impact to the calibration parameter.

The tagging calibration parameters are determined by substituting Eq. 4.11 in place of ω in Eq. 5.6, where p_0 and p_1 are the only floating parameters of the nominal fit. In this case the PDF depends also on the predicted mistag η, whose distribution is taken from data. This method is referred to as *unbinned* and it has been chosen as the nominal one to extract the tagging calibration parameters. Alternatively, it is possible to determine the average mistag fraction ω by fitting the decay time distribution of candidates split into bins of the predicted mistag η. This method, commonly called *fit in categories* or *binned*, is particularly useful for checking the linearity of the calibration. The $(\langle\eta\rangle_{cat}, \omega_{cat})$ data points are fitted with the linear function in Eq. 4.11 to measure the calibration parameters p_0 and p_1.

5.2 calibration of the nnetssk tagger

The calibration of SSK taggers can be done only using B_s^0 decays as shown in Fig.29. Although also in the fragmentation from processes leading to B^+ SS kaons can be produced in addition to SS pions, the pions contamination might alter the SSK calibration results. Recently the $B_{s2}^* \to B^+K^-$ decay mode has been used for the first time to make a calibration of the SSK tagger even if at the moment it has limited impact due to low statistics. The NNetSSK calibrations in $B_s^0 \to D_s^-\pi^+$ and $B_{s2}^* \to B^+K^-$ have combined together [65]. The NNetSSK calibration has been extensively studied in the $B_s^0 \to D_s^-\pi^+$ decay mode. Semileptonic B_s^0 decays have large statistics, but they are not used for tagging calibration purposes due to the large, time dependent, time resolution. The η distribution is shown in Figure 36.

The results of the tagging calibration in $B_s^0 \to D_s^-\pi^+$ for the 2011 and 2012 samples are reported, separately, in order to verify the compatibility of the results. The results of the calibration performed with both fit approaches, unbinned and binned, are reported in Tab. 14, where p_0 and p_1 are the tagging parameters, $\langle\eta\rangle$ is the average value of the predicted mistag evaluated on data and ϵ_{tag} is the tagging efficiency before the calibration. In Fig. 37a and 37b the plots with the binned fit in categories of η are reported respectively for 2011 and 2012 data. The data points are the ω_{cat} values measured by means of a simultaneous time fit in 5(8) categories of η for the 2011(2012) sample. The plots well show the linearity between η and the measured mistag ω for both samples, showed also by the χ^2/NDF reported in Tab.14. The 2011 and 2012 values are compatible in less than 1 σ for $p_0 - \langle\eta\rangle$ and

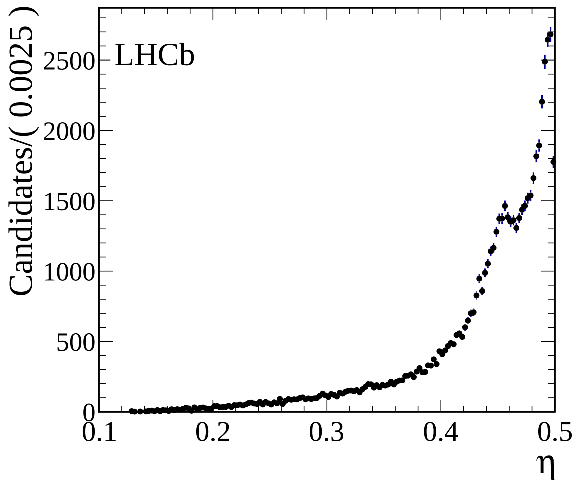
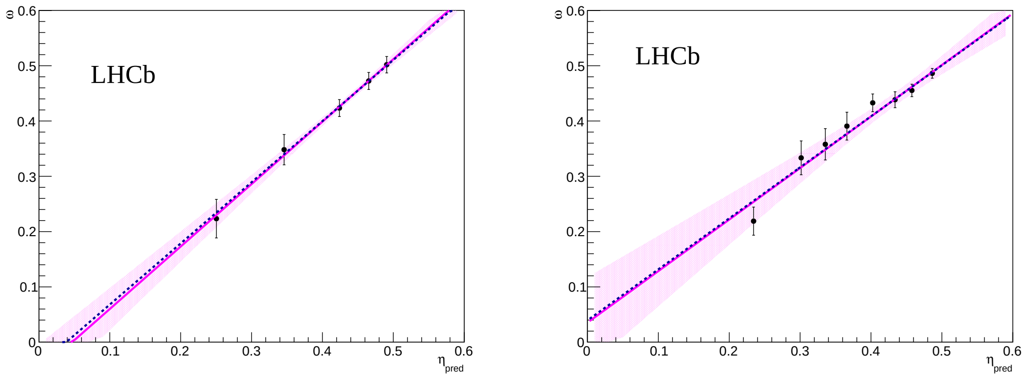


Figure 36.: Predicted mistag distribution of signal events obtained applying sweights for the 2011+2012 sample.



(a) 2011

(b) 2012

Figure 37.: Average mistag fraction ω_{cat} in bins of predicted mistag probability η. The solid line is the result of a linear fit to the data points and the dashed line is the calibration obtained from the unbinned fit. The shaded area corresponds to the 95% Confidence Level of the linear fit to data points.

Table 14.: Results of the NNetSSK tagging calibration parameters p_0 and p_1 obtained by the unbinned and the binned fit to the decay time distribution of the 2011 and 2012 data samples. The tagging efficiencies ϵ_{tag} are computed before the calibration.

2011	p_0	p_1	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	$\rho(p_0, p_1)$	ϵ_{tag} (%)	χ^2/NDF
unbinned	0.4440±0.0081	1.106±0.130	0.4399	0.004±0.008	0.040	64.6	-
binned	0.4439±0.0081	1.129±0.133	0.4399	0.004±0.008	0.012	64.6	0.07

2012	p_0	p_1	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	$\rho(p_0, p_1)$	ϵ_{tag} (%)	χ^2/NDF
unbinned	0.4424±0.0053	0.923±0.083	0.4368	0.006±0.005	0.040	63.5	-
binned	0.4423±0.0054	0.932±0.085	0.4368	0.006±0.005	0.015	63.5	0.84

in 1.2 σ for p_1. For this reason the two samples are merged providing a unique set of flavour tagging parameters for the Run1 data sample.

Figure 38 shows the calibration plot for the merged 2011+2012 data sample. Each data point in the plot corresponds to the measured values of ω_{cat} and $\langle\eta\rangle_{cat}$ reported in table 15 together with the definition of categories. The linearity between η and the measured mistag ω is satisfactory as the linear fit chisquare value is $\chi^2/NDF = 9.2/7$. In addition to the linear fit (full line), the calibration obtained from the unbinned approach (dashed line) is superimposed; the results are in agreement. Table 16 reports the results of the calibrations using the unbinned and binned methods. We consider the unbinned fit as the reference calibration; the difference between the two approaches is taken as a systematic uncertainty.

Figure 39 shows the mixing asymmetry for B candidates in one period of the mixing oscillations (folding the full decay time distribution), with the fit projection overlaid for events with predicted mistag η in different ranges. The data are sWeighted to consider only the signal candidates; all the candidates tagged by the NNetSSK are considered. The offset $t_0 = 0.2\,\text{ps}$ corresponds to the preselection requirement on the decay time. The two asymmetry plots show the impact of the mistag dilution on the mixing amplitude. Indeed selecting events with low mistag value ($\eta < 0.35$) increases the mixing amplitude.

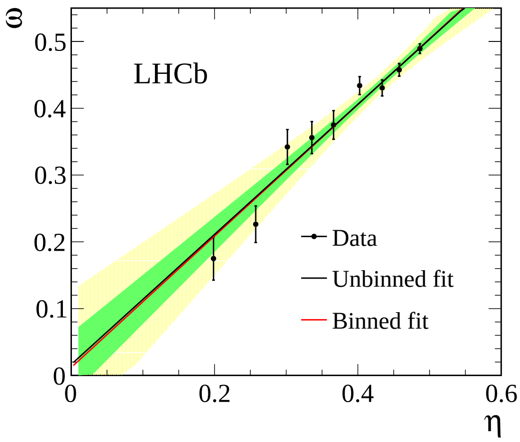


Figure 38.: Average mistag fraction ω_{cat} in bins of predicted mistag probability η. The red line is the result of a linear fit to the 2011+2012 data points and the black line is the calibration obtained from the unbinned fit. The shaded areas correspond to 68 and 95% Confidence Level of the linear fit to data points.

Table 15.: Results of the simultaneous time fit on 2011+2012 data for the average mistag ω_{cat} and tagging efficiency ϵ_{tag} in bins of the predicted mistag probability η.

category definition	$\langle\eta\rangle_{cat}$	ω_{cat}	ε_{tag} (%)
$\eta < 0.23$	0.20	0.175±0.032	0.9
$0.23 \leq \eta < 0.28$	0.26	0.226±0.027	1.6
$0.28 \leq \eta < 0.32$	0.30	0.342±0.026	2.1
$0.32 \leq \eta < 0.35$	0.34	0.356±0.024	2.3
$0.35 \leq \eta < 0.38$	0.37	0.375±0.021	3.0
$0.38 \leq \eta < 0.42$	0.40	0.434±0.014	7.1
$0.42 \leq \eta < 0.445$	0.434	0.430±0.012	8.8
$0.445 \leq \eta < 0.47$	0.458	0.458±0.0095	14.2
$0.47 \leq \eta < 0.5$	0.487	0.489±0.0072	24.2

5.2.1 *Flavour tagging asymmetries*

Performances and calibrations of the tagging algorithm can depend on the B_s^0 meson initial flavour, due to the different interaction cross sections of K^+/K^- with mat-

Table 16.: Results of the NNetSSK tagging calibration parameters p_0 and p_1 obtained by the unbinned and the binned fit to the time distribution of the merged data sample (2011+2012). The result of the unbinned fit is considered as the reference calibration. The quoted tagging efficieiency is computed after the calibration.

	p_0	p_1	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	$\rho(p_0, p_1)$	ε_{tag} (%)
unbinned	0.4429±0.0044	0.977±0.070	0.4377	0.0052±0.0044	0.042	60.38±0.16
binned	0.4427±0.0044	0.983±0.072	0.4377	0.0050±0.0045	0.025	60.38±0.16

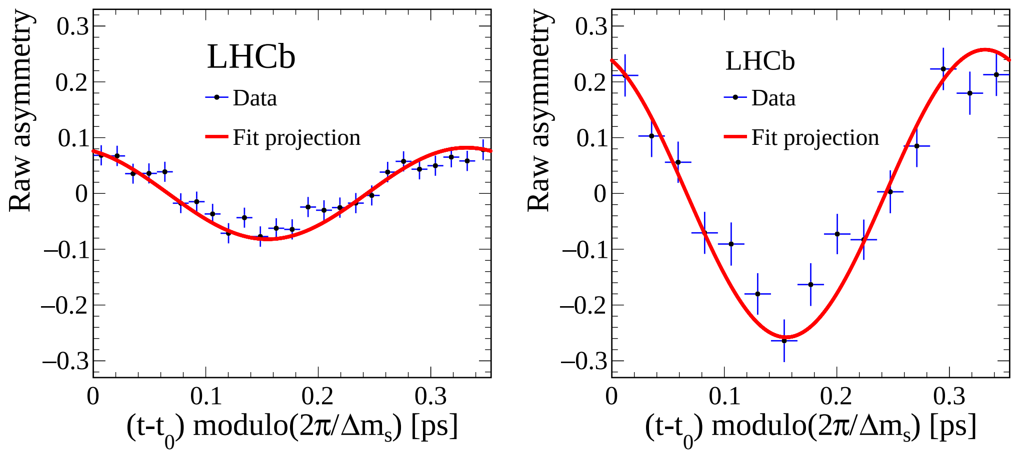


Figure 39.: Asymmetry between the number of mixed and unmixed B candidates in one period of the mixing oscillations (folding the full decay time distribution), with the fit projection overlaid. The data are sWeighted to consider only the signal candidates. In the left plot, all the candidates tagged by the NNetSSK are considered, with their per-event mistag; in the right plot, only the events with $\eta < 0.35$. The offset $t_0 = 0.2\,\mathrm{ps}$ corresponds to the preselection requirement on the decay time.

ter. Differences between the calibrations results according to the initial flavour, generally referred to as "flavour tagging asymmetries", are studied in data using a large statistics control sample of prompt $D_s \to \phi(K^+K^-)\pi^+$ decays which have the advantages to be charged particles. As well as B_s^0 mesons, the hadronization of prompt D_s involves a quark $\bar{c}$ and a pair of $s - \bar{s}$ which can produce a same side [66]. Even if the fragmentation process and the kinematic range are different, the asymmetry due to the different interaction between K^+/K^- can be evaluated and assumed to be valid also for B_s^0 mesons.

Table 17.: Selection cuts for prompt D_s candidates.

parameter	value
$DLL(K-p)$	> -10
isMuon	0
$\lvert m(K^+K^-) - m_{\phi(1020)} \rvert$	< 7 MeV/c^2
$m(K^+K^-\pi)$	[1920,2040] MeV/c^2

Prompt D_s decays are selected from candidates that pass the "D2hhhControl_-KKPiLine" stripping selection and the additional criteria listed in Table 17. Figure 40 shows the invariant mass distribution of the selected candidates. A fit to the invariant mass distribution $m(KK\pi)$ is performed to separate the signal from the background. A sum of two gaussians with a common mean is used for the signal, while an exponential shape is used to describe the background. In total $\sim 784\,000$ signal events are selected from the 2011 and 2012 samples. The contribution of background is $\sim 43\,000$ events. The signal *sWeights* from the mass fit are used to produce the η distributions of the wrong and right tagged signal events, with the *sPlot* method. The ratio of the "Wrong/Right+Wrong" histograms defines the calibration plot for the predicted mistag. Right and wrong tagged events are classified simply comparing the tagging decision with the charge of the reconstructed D_s candidate.[3] Figure 40 right, shows the difference of the calibration plots. A linear fit to the data points gives the following parameters: $\Delta p_0 = -0.015 \pm 0.003$, $\Delta p_1 = 0.016 \pm 0.030$, in perfect agreement to the difference of the parameters p_0 and p_1 of the linear fit to D_s^- and D_s^+: $\Delta p_0 = -0.014 \pm 0.003$, $\Delta p_1 = 0.020 \pm 0.030$.

Differences between the D_s and B_s^0 phase spaces have been accounted for by re-weighting the D_s candidates to match the B_s^0 transverse momentum distribution of $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J/\psi\phi$, $B_s^0 \to \phi\phi$, and $B_s^0 \to D_s D_s$ decays. The differences

3 To match the decision of the B_s^0 case the tagging decision is reversed.

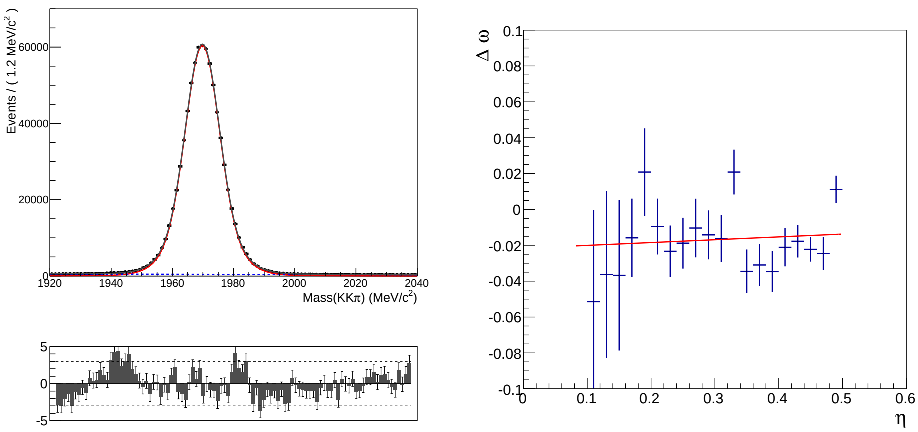


Figure 40.: Left: distribution of the $KK\pi$ invariant mass for selected prompt D_s candidates. Right: difference of the NNetSSK calibration plots for prompt D_s ($D_s^- - D_s^+$), the line correspond to the best fit of the data points with a linear funcion.

of the measured tagging parameters ($p_i(B) - p_i(\bar{B}, i = 0, 1)$ and tagging efficiency ($\Delta\epsilon_{tag}$) in the D_s^- and D_s^+ re-weighted samples allow the determination of the tagging asymmetry reported in Table 18.

For comparison, the evaluation of the initial flavour tagging asymmetry is also performed using samples of simulated events corresponding to the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to J/\psi\phi$ decays. The calibration parameters are evaluated exploiting the information of the true flavour present in MC. The results are reported in Table 18; they are in reasonable agreement with the results obtained using the D_s sample.

As reference values for Δp_0, Δp_1 and $\Delta\varepsilon_{tag}$ have been taken the ones obtained by the D_s re-weighted to match p_T of $B_s^0 \to D_s^- \pi^+$ sample, and as systematic uncertainties have been taken the maximum difference between the values obtained by the re-weighted D_s samples to match p_T spectra of the considered decays.

5.2.2 *Systematic uncertainties*

Usually in tagged, time-dependent fit the tagging calibration parameters are constrained through gaussian function where the sigma is equal to the sum of

Table 18.: Initial flavour asymmetry of calibration parameters and tagging efficiency, evaluated using MC samples of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to J/\psi\phi$ decays, and data sample of D_s decays re-weighted for the B_s^0 transverse momentum of $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J/\psi\phi$, $B_s^0 \to \phi\phi$ and $B_s^0 \to D_s D_s$ decays. We take as reference values the ones from the D_s data re-weighted for the $B_s^0 \to D_s^- \pi^+$ p_T distribution and obtained by fixing $\langle\eta\rangle$ to the reference value of $B_s^0 \to D_s^- \pi^+$ (0.4377); the first uncertainty is statistical; the second one is the systematic. The systematic uncertainty is defined as the maximum difference of the individual D_s results re-weighted to match the distribution of p_T or number of tracks observed in different channels with respect to the values corresponding to the re-weighting to the $B_s^0 \to D_s^- \pi^+$.

data sample	Δp_0	Δp_1	$\Delta\varepsilon_{tag}$ (%)
MC $B_s^0 \to D_s^- \pi^+$	-0.0126±0.0034	0.004±0.031	+0.40±0.30
MC $B_s^0 \to J/\psi\phi$	-0.0137±0.0007	0.016±0.008	+0.11±0.06
D_s re-weighted to match p_T of $B_s^0 \to D_s^- \pi^+$	−0.0163±0.0022	−0.031±0.025	+0.17±0.11
D_s re-weighted to match p_T of $B_s^0 \to J/\psi\phi$	−0.0141±0.0031	0.009±0.028	−0.49±0.11
D_s re-weighted to match p_T of $B_s^0 \to \phi\phi$	−0.0166±0.0025	−0.047±0.029	0.17±0.12
D_s re-weighted to match p_T of $B_s^0 \to D_s D_s$	−0.0193±0.0042	−0.048±0.038	0.12±0.12
systematic uncertainty	±0.0030	±0.040	0.66
D_s re-weighted to match nTrack of $B_s^0 \to D_s^- \pi^+$	−0.0152±0.0022	0.051±0.025	0.28±0.11
D_s re-weighted to match nTrack of $B_s^0 \to J/\psi\phi$	−0.0152±0.0031	0.053±0.028	0.39±0.12
D_s re-weighted to match nTrack of $B_s^0 \to \phi\phi$	−0.0154±0.0025	0.071±0.029	0.43±0.13
D_s re-weighted to match nTrack of $B_s^0 \to D_s D_s$	−0.0154±0.0042	0.064±0.038	0.29±0.12
systematic uncertainty	±0.0002	±0.020	0.15
Reference values	−0.0163±0.0022±0.0030	−0.031±0.025±0.045	0.17±0.11±0.68

statistical and systematic errors. The sources of systematic uncertainties can be divided in two categories: the first one is related to the fit method and calibration strategy; the second one is related to the use of a universal calibration. Indeed it accounts for the dependencies of the calibration parameters on the kinematics of the signal B or the event properties and the possible differences that can arise porting the calibration from the control decay to any other signal decay.

Table 19.: Systematic uncertainties related to the fit model and calibration method of NNetSSK.

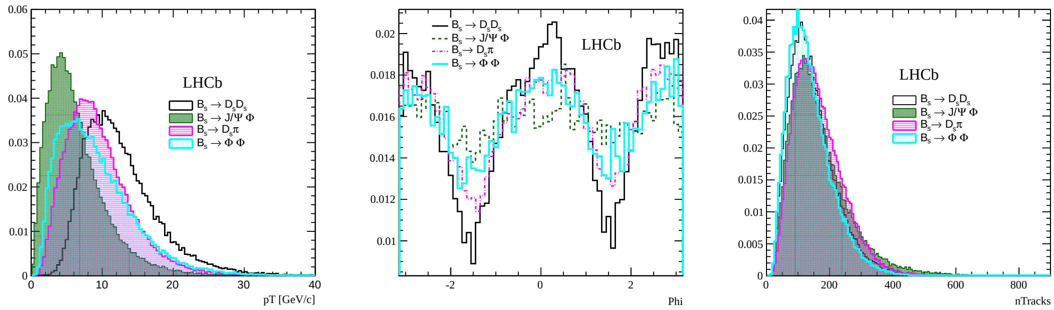
Source	σ_{p_0}	σ_{p_1}
calibration method	0.0002	0.0060
$B_s^0 \to D_s K$ yield	0.0001	0.0085
PDF mass model signal	0.0001	0.0022
PDF mass model background	0.0015	0.0245
time resolution	0.0033	0.0600
time acceptance	<0.0001	<0.001
values of $\Delta\Gamma_s$, Γ_s, and Δm_s	<0.0001	<0.001
sum in quadrature	0.0036	0.0657

fit model and calibration method. Table 19 summarizes the systematic uncertainties related to the fit model and calibration method. The σ_{p_0} and σ_{p_1} values represent the maximum deviation of the calibration parameters measured in each test with respect to the reference values quoted in Table 16. The total systematic uncertainty is the sum in quadrature of all the errors. As explained in section 5.1.2, the calibration parameters are determined by two methods, the unbinned fit and the fit in categories. The difference between the results of the two calibration methods is considered as a systematic uncertainty. Other sources of uncertainties are related to the fit model and to the background contamination. We study the impact on the calibration parameters of a variation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ background contribution, which peaks under the signal, by fixing in the mass fit the yield of the $B_s^0 \to D_s^{\mp} K^{\pm}$ background to be twice or half of the nominal value. In addition the signal PDF for the mass distribution has been changed from a sum of two Crystal Ball functions to a sum of two Gaussian functions, to evaluate the systematic related to the signal model applied. Its contribution to the uncertainty is negligible. By using generation of pseudo-experiments ("toys studies" see Appendix A.2), we compute the effect of choosing a wrong parametrisation of the combinatorial background: we generate the combinatorial events with a different model (exponential+constant PDF) with respect to the default model used to fit (single exponential PDF) the data and the generated toys. Similarly, we also address the systematic uncertainties due to choice of the shapes of the physics backgrounds in the mass fit: we change the shapes of the PDF in generation, and fit with the nominal ones. By doing these variations, a bias of about 35% of the statistical error is observed in both p_0 and p_1; this bias is taken as the systematic uncertainty due to the PDF mass model of backgrounds. The effect of the uncertainty in the time resolution has been studied varying the correction scale factor of 1.37 by $\sim \pm 10\%$ (1.28 and 1.48, similar values are used by the Δm_s analysis [59] in the systematic evaluation). Its contribution is the dominant part of the systematic uncertainty to the tagging calibration. We investigated the effect of the time acceptance function on the tagging calibration parameters, by producing several pseudo-experiments where the data are generated using the nominal acceptance function and have been fitted with an acceptance function in which each parameter is varied by $\pm 1\sigma$; no significant effect is observed (absolute variations on p_i are of the fraction of per mille). The effect on the calibration parameters of the variation of the parameters Δm_s, $\Delta\Gamma_s$ and B_s^0 lifetime is also negligible within their experimental accuracy.

EVENT PROPERTIES AND B KINEMATICS. The tagging calibration shows a dependency on the properties of the event (tracks and PV multiplicities) and on the kinematics of the signal B (transverse momentum p_T ,ϕ angle, and η). Therefore, the calibration parameters measured in a control sample of a given decay can be applied to another decay mode as long as the distributions of such variables are similar. This is not the case, given the different selection applied to identify the signal decays. Since we want to use a unique calibration for all the CP analyses, the systematic uncertainties need to account for these residual variations. Systematic uncertainties are then assigned to address the portability of the calibration derived in the control decay to another decay. For this reason the variations on p_0 and p_1 with respect to the nominal calibration have been evaluated by re-weighting the $B_s^0 \to D_s^- \pi^+$ candidates to match the distributions of the transverse momentum p_T, the ϕ angle and the track multiplicity for different decays.[4] The following

4 The pseudorapidity η is highly correlated with p_T, and the number of PV with the track multiplicity.

B_s^0 decay modes are considered: $B_s^0 \to J/\psi\phi$, $B_s^0 \to D_s D_s$ and $B_s^0 \to \phi\phi$. For a given variable, the largest variation observed in the re-weighting for the different decays is taken as the systematic uncertainty; we consider as variation the half-difference between the nominal calibration (no re-weighting applied), and the calibration after the re-weight. The re-weighting procedure is applied to one variable at the time, neglecting their correlations. To have larger statistical samples and simplify the analysis, the procedure is applied to $B_s^0 \to D_s^- \pi^+$ simulated events. The results are reported in Table 20; the final systematic uncertainty is computed as the sum in quadrature of the individual contributions. Figure 41 shows the distributions of the p_T (B_s^0), ϕ and of the number of tracks in the event for the different decays used in the re-weighting procedure.



(a) (b) (c)

Figure 41.: Distributions of (a) p_T (B_s^0), (b) ϕ and (c) track multiplicity for $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J\psi\phi$, $B_s^0 \to \phi\phi$ and $B_s^0 \to D_s D_s$ decays in simulation.

5.2.3 *Test in MC samples*

In order to prove the portability of the calibration obtained from $B_s^0 \to D_s^- \pi^+$ decays to other B_s^0 decay modes, a study is performed using simulated samples of $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s^+ D_s^-$ decays. In Figure 42 the calibrations obtained with the three samples are reported in Table 21. The calibrations are obtained exploiting the information on the true initial flavour available in the

Table 20.: Systematic uncertainties on the NNetSSK calibration related to the event properties and the B kinematics evaluated by means of a re-weighting procedure of simulated $B_s^0 \to D_s^- \pi^+$ events to match the $p_T(B_s^0)$, ϕ and the tracks multiplicity distributions of $B_s^0 \to J/\psi\phi$, $B_s^0 \to D_s^+ D_s^-$ and $B_s^0 \to \phi\phi$.

observable used for re-weighting	σ_{p_0}	σ_{p_1}
p_T	0.0011	0.030
ϕ	0.00001	0.00025
nTracks	0.00055	0.0061
sum in quadrature	0.0012	0.030

simulations. The value of $p_0 - \langle\eta\rangle$ is compatible with zero in the three cases and the curves are in agreement, although there is a small deviation in the p_1 value for the $B_s^0 \to D_s^+ D_s^-$ with respect to the other decays. Such deviation is understood as being related to differences in the phase space: in fact it is strongly reduced if the samples are re-weighted to match the $B_s^0 \to D_s^+ D_s^-$ transverse momentum distribution (Fig. 42b), proving the need of the systematic uncertainty calculated in the previous section.

5.2.4 *NNetSSK calibration summary and performances*

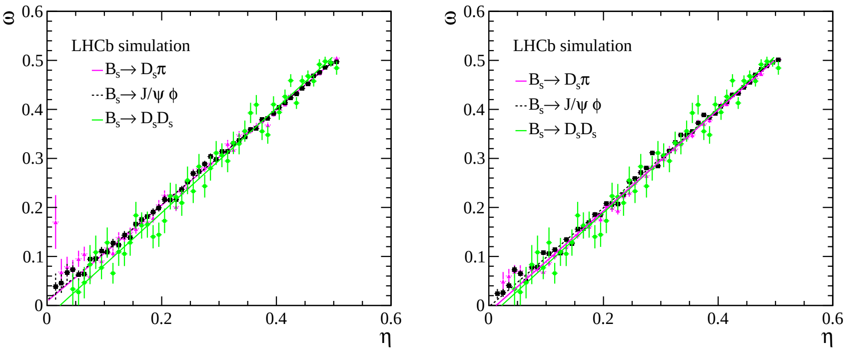
The final calibration parameters in the $B_s^0 \to D_s^- \pi^+$ decays with their statistical and total systematic uncertainties are reported in Table 22. The systematic uncertainties don't include the flavour tagging asymmetry, which should be added in quadrature to the systematic uncertainty in the case a given analysis doesn't correct for the asymmetries explicitly in the fit to the data. The calibrations parameters of $B_s^0 \to D_s^- \pi^+$ decays are combined together with the one obtained in the $B_{s2}^* \to B^{(*)+} K^-$ control channel [48], and the final tagging parameters for the NNetSSK taggers are:

$< \eta > = 0.4377$,

Table 21.: NNetSSK Calibration in different MC samples.

	p_1	$p_0 - \langle\eta\rangle$
$B_s^0 \to D_s^- \pi^+$	0.9814 ± 0.007	0.0005 ± 0.0009
$B_s^0 \to J/\psi\phi$	0.9754 ± 0.006	0.0000 ± 0.0005
$B_s^0 \to D_s^- D_s^+$	1.066 ± 0.021	0.0027 ± 0.0028
	After Reweighting	
$B_s^0 \to D_s^- \pi^+$	1.031 ± 0.007	-0.0010 ± 0.0009
$B_s^0 \to J/\psi\phi$	1.013 ± 0.007	0.0019 ± 0.0009
$B_s^0 \to D_s^- D_s^+$	1.066 ± 0.021	0.0027 ± 0.0028

$p_0 - < \eta >$ = 0.0070 ± 0.0039 ± 0.0035,
p_1 = 0.925 ± 0.061 ± 0.059,
Δp_0 = -0.0163 ± 0.0022 ± 0.0030,
Δp_1 = -0.031 ± 0.025 ± 0.045,
$\Delta\epsilon_{tag}$ = (0.17 ± 0.11 ± 0.68)%,
where the first uncertainty is statistic and the second one is systematic. The tagging efficiency and tagging power of the NNetSSK tagger (ϵ_{eff} = (1.80±0.19(stat)±0.17(syst))% evaluated event by event as in Eq.4.8, ϵ_{tag} =(60.38±0.16 (stat))%) improve with respect to the ones of the cut-based SSK algorithm (ϵ_{eff} = 1.19 ± 0.14, ϵ_{tag} = 15.7%). The development of NNetSSK represents an important progress that leads to an improvement for many CP physics analyses of B_s^0 decays. It is employed in measurement of time-dependent CP violation in several B_s^0 decays [68, 69, 1, 70, 71], confirming the good performances observed in $B_s^0 \to D_s^- \pi^+$ decays; in $B_s^0 \to J/\psi\phi$ decays, the NNetSSK tagging power obtained in the full Run I data set is $(1.26 \pm 0.17)\%$ [68], to be compared with $(0.66 \pm 0.13)\%$ in the same data set with the cut-based SSK tagger [72].



(a) (b)

Figure 42.: Calibration plots of the NNet SSK tagger in different MC samples: $B_s^0 \to D_s^- \pi^+$, $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s^+ D_s^-$. The left plot 42a corresponds to the distributions of the selected events, in the right plot 42b the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to J/\psi\phi$ samples have been reweighed to match the $B_s^0 \to D_s^+ D_s^-$ transverse momentum distribution.

Table 22.: Final calibration parameters for the NNetSSK tagger in which the first uncertainty is statistical and the second one is systematic. Also Initial flavour asymmetry in the flavour tagging calibration is reported.

p_0	p_1	$\langle\eta\rangle$
0.4429±0.0044±0.0035	0.977±0.070±0.068	0.4377

Δp_0	Δp_1	$\Delta\epsilon_{tag}$ [%]
−0.0159±0.0022±0.0016	−0.031±0.025±0.040	0.17±0.11±0.66

5.3 CALIBRATION OF OS TAGGER

The $B_s^0 \to D_s^- \pi^+$ decay mode is used to check the calibration of the OS taggers combination (OSe,OSμ,OSK and OSVtx defined in Tab.10) following the same procedure used for the NNetSSK tagger. The tagging calibration parameters have been evaluated using both the unbinned fit and the fit in category. The bins of η, the measured ω_{cat} and the $< \eta_{cat} >$ are reported in Tab.23 while the binned fit is shown in Fig.43. In Table 24 the calibration parameters of the merged sample 2011 + 2012 are reported using the unbinned and binned fit method. As for the other taggers, the result of the unbinned fit is considered as the reference calibration for the $B_s^0 \to D_s^- \pi^+$ decay mode.

Table 23.: Results of the simultaneous time fit on 2011+2012 $B_s^0 \to D_s^- \pi^+$ data for the average mistag ω_{cat} in bins of the predicted mistag probability η of the OS combination.

category	$\langle\eta\rangle_{cat}$	ω_{cat}	ϵ (%)
$\eta < 0.17$	0. 31	0.34 ± 0.02	6.78
$0.17 \leq \eta < 0.24$	0.37	0.41 ± 0.02	4.16
$0.24 \leq \eta < 0.31$	0.41	0.45 ± 0.01	9.70
$0.31 \leq \eta < 0.35$	0.44	0.47 ± 0.02	5.27
$0.35 \leq \eta < 0.38$	0.46	0.50 ± 0.01	7.13
$0.38 \leq \eta < 0.43$	0.48	0.50 ± 0.01	14.44
$0.43 \leq \eta < 0.5$	0.48	0.50 ± 0.01	14.44

The systematics related to the $B_s^0 \to D_s^- \pi^+$ fit strategy are reported in Tab. 25.

The time resolution uncertainty represents the dominant source of the systematic errors and globally they are of the same order of the statistical errors.

The OS tagger in the $B_s^0 \to D_s^- \pi^+$ decay mode has a tagging power of: ε_{eff}= $(3.15 \pm 0.23 \pm 0.24)\%$ evaluated using Eq.4.8, where the first uncertainty is statistical the second one is systematic. The values of the calibration here determined contribute to the final estimation of the OS tagger calibration which is evaluated as the weighted average of values corresponding to different control channels: $B^+ \to J/\psi K^+$, $B^+ \to D^0 \pi^+$, $B_d \to D^{*-} \mu^+ \nu_\mu$, $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to D_s^- \pi^+$.

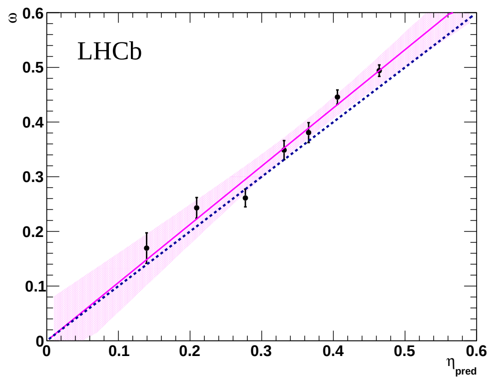


Figure 43.: Fit in category for the OS tagger in the 2011 + 2012 $B_s^0 \to D_s^- \pi^+$ data. The dashed line is the ideal calibration: $p_0 = \langle \eta \rangle$ and $p_1 = 1$. The shaded area is the 95% Confidence Level

Table 24.: Results of the time fit on 2011+2012 $B_s^0 \to D_s^- \pi^+$ data for the OS calibration parameters p_0 and p_1 using the unbinned and binned fit of the predicted mistag probability η. The reference calibration is the results of the unbinned fit.

	p_0	p_1	$\langle \eta \rangle$ (fixed)	Δp_0	ϵ_{tag} (%)
unbinned fit	0.3925 ± 0.0058	1.052 ± 0.054	0.3687	0.0238 ± 0.0058	37.2
binned	0.3925 ± 0.0059	1.065 ± 0.061	0.3687	0.0238 ± 0.0059	37.2

In such evaluation each decay mode provides: calibration parameters, statistical errors and systematics errors related to the implemented analysis.

In Table 26 for each control channel the calibration parameters are reported with the statistical and systematic uncertainty connected with that specific channel. As reference values for the OS tagger has been considered the weighted average of $p_0 - \langle \eta \rangle$ and p_1 on the five control channels, since the measurements are compatible. This strategy is aimed to define a coherent picture providing a universal calibration. For this reason a deep study has been done in order to evaluate which is the impact of a kinematic distribution in several physics channels which differs from the one of the calibration channel. The difference in the tagging calibration

Table 25.: Systematic uncertainties of the OS combination related to the fit model for p_0 and p_1 from the $B_s^0 \to D_s^- \pi^+$ analysis. They are evaluated as the maximum difference with the reference values.

Syst. source	σ_{p_0}	σ_{p_1}
calibration method	0.0	0.013
$B_s^0 \to D_s K$ yield	0.0004	0.019
time resolution	0.0055	0.055
PDF mass model signal	0.0003	0.001
sum in quadrature	0.0055	0.057

Table 26.: Calibration parameters of the combined OS tagger extracted from different control channels. Statistic and systematic uncertainties are quoted.

control ch.	p_0	p_1	$\langle\eta\rangle$
$B^+ \to J/\psi K^+$	$0.3927 \pm 0.0014 \pm 0.0015$	$0.978 \pm 0.012 \pm 0.005$	0.3919
$B^+ \to D^0\pi^+$	$0.3854 \pm 0.0016 \pm 0.0015$	$0.984 \pm 0.008 \pm 0.005$	0.3836
$B^0 \to J/\psi K^{*0}$	$0.399 \pm 0.003 \pm 0.0060$	$0.978 \pm 0.021 \pm 0.039$	0.390
$B^0 \to D^{*-}\mu^+\nu_\mu$	$0.3953 \pm 0.0019 \pm 0.0069$	$0.949 \pm 0.010 \pm 0.057$	0.3872
$B_s^0 \to D_s^-\pi^+$	$0.3972 \pm 0.0097 \pm 0.0071$	$1.061 \pm 0.063 \pm 0.049$	0.3813

parameters due to the difference in phase space between calibration and physics channels are considered as systematics effects. The final average values are quoted

Table 27.: Final calibration parameters for the OS tagger in which the first uncertainty is statistical and the second one is systematic. Also Initial flavour asymmetry in the flavour tagging calibration is reported.

$p_0 - \langle\eta\rangle$	$\langle\eta\rangle$	p_1
$0.0062 \pm 0.0019 \pm 0.0040$	$\langle\eta\rangle_{decay}$	$0.982 \pm 0.007 \pm 0.034$

Δp_0	Δp_1	$\Delta\epsilon_{tag}$ [%]
$+0.014 \pm 0.0012$	$+0.066 \pm 0.012$	-0.0008 ± 0.0007

in Tab. 27 with a statistical error in which are included the systematics related to each control channel and a systematic error which accounts for phase space differences among the calibration channel and different physics channels.

The initial flavour asymmetry has been evaluated in charged control channels $B^+ \to J/\psi K^+$ and $B^+ \to D^0\pi^+$, measuring the calibration parameters on B^+ and B^- samples separately. The values are reported in Tab.27.

6

MEASUREMENT OF γ ANGLE FROM THE ANALYSIS OF $B_s^0 \to D_s^\mp K^\pm$ USING 1FB^{-1}

Contents

6.1 Data sample **94**
 6.1.1 Monte Carlo Simulation 94
6.2 Event selection **96**
6.3 Multi-dimensional fit **100**
 6.3.1 Probability Density Functions 102
 6.3.2 Summary of fixed yields 109
 6.3.3 Results of the Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$ 109
 6.3.4 Results of the Multidimensional fit to $B_s^0 \to D_s^\mp K^\pm$ 110
6.4 Flavour Tagging Calibration **122**
 6.4.1 Calibration of OS tagger 122
 6.4.2 Calibration of NNetSSK tagger 125
 6.4.3 Comparison of Tagging Performance in $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ 126
 6.4.4 Tagging asymmetries 128
 6.4.5 OS and NNetSSK combination 129
6.5 Time resolution and time acceptance **132**
6.6 Time fit **135**
6.7 Results of the sFit **136**
 6.7.1 sFit in $B_s^0 \to D_s^- \pi^+$ 136
 6.7.2 sFit in $B_s^0 \to D_s^\mp K^\pm$ 138
6.8 Results of the cFit **139**
 6.8.1 $B_s^0 \to D_s^\mp K^\pm$ fit results 142
6.9 Systematics uncertainties **143**
6.10 The γ angle determination **145**

This chapter reports the first measurement of the CP violating observables in $B_s^0 \to D_s^\pm K^\mp$ decays depending on γ using a dataset corresponding to 1fb^{-1} of integrated luminosity collected at $\sqrt{s} = 7$ TeV of pp collisions by the LHCb experiment pubblished in Ref. [1]. To measure the CP violating observables it is necessary to perform a fit to the decay time distribution of the selected $B_s^0 \to D_s^{\pm} K^{\mp}$ candidates. The $B_s^0 \to D_s^- \pi^+$ decay is used as a control mode for the optimization of the selection, for studying and constraining some background contributions and for the calibration of flavour tagging algorithms. In addition the $B^0 \to D^- \pi^+$ decay mode is used to evaluate differences between data and MC samples that are used to improve the simulation of the events.

6.1 DATA SAMPLE

The dataset used in this analysis comprises two distinct sub-samples recorded with opposite direction of the magnetic field in the spectrometer dipole: 0.44 fb^{-1} with the magnet up and 0.59 fb^{-1} with the magnet down polarities, for a total of $\approx$ 1 fb^{-1} of integrated luminosity. The events are reconstructed using version Reco12 of the reconstruction algorithm. In addition to the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays two other control modes have been employed in the analysis: the $B_s^0 \to D_s^- \pi^+$ and the $B^0 \to D^- \pi^+$. The former differs from the signal decay only by the bachelor particle hypothesis (Kaon or Pion). The latter has a larger statistics and it allows to estimate the correction factors to make the MC samples more similar to data. The following D_s decays are considered: $D_s^- \to K^- K^+ \pi^-$, $D_s^- \to K^- \pi^+ \pi^-$, $D_s^- \to \pi^- \pi^+ \pi^-$. The $D_s \to KK\pi$ sample is further divided into three subsamples: $D_s \to \Phi\pi$, $D_s \to K^* K$ and the remaining $D_s \to (KK\pi)_{non\ resonant}$. In this way different optimized selections are applied to the three subsamples being looser in the case of intermediate resonances. These D_s^- candidates are subsequently combined with a companion particle to form the $B_s^0 \to D_s^- K^+$ or the $B_s^0 \to D_s^- \pi^+$ decay candidates.

6.1.1 *Monte Carlo Simulation*

Several Monte Carlo (MC) samples with the MC11a configuration were used in the analysis, notably for selection, normalisation and background studies. A full list is given in Tab. 28. All MC samples are selected using the same criteria as the data except for the requirements on the particle identification whose effect is emulated by means of a data driven method. MC samples of $B_s^0 \to D_s^- \pi^+$ are widely used to cross-check the flavour tagging results and to verify the portability to the $B_s^0 \to D_s^\mp K^\pm$ channel. In all the other case the MC samples are used to determine distributions of the observables which enter in the fit.

Table 28.: MC samples used during the analysis for selection and background studies. The second column indicates the specific B-daughter decay and in the last column the main use of each sample is reported.

Sample		Sample size	Use
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to KK\pi$	1052495	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^- \pi^+$	$D_s \to K\pi\pi$	3093985	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^- \pi^+$	$D_s \to \pi\pi\pi$	1053991	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^\mp K^\pm$	$D_s^- \to KK\pi$	1887293	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^\mp K^\pm$	$D_s \to K\pi\pi$	1008497	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^\mp K^\pm$	$D_s \to \pi\pi\pi$	1027996	Signal study, Flavour Tagging, Time acceptance
$B_s^0 \to D_s^{*-} \pi^+$	$D_s^- \to KK\pi$	524098	Background study
$B_s^0 \to D_s^- \rho^+$	$D_s^- \to KK\pi$	2019391	Background study
$\Lambda_b^0 \to D_s^- p$	$D_s^- \to KK\pi$	539994	Background study
$\Lambda_b^0 \to D_s^{*-} p$	$D_s^- \to KK\pi$	630598	Background study
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	$\Lambda_c \to pK\pi$	2033496	Background study
$\Lambda_b^0 \to \Lambda_c^+ K^-$	$\Lambda_c \to pK\pi$	519495	Background study
$B^0 \to D^- \pi^+$	$D \to K\pi\pi$	1016198	Obtain Data/MC corrections, estimate yields
$B^0 \to D^- K^+$	$D \to K\pi\pi$	958393	Background study

In addition to the signal samples listed in Tab. 28, we requested the production of additional Monte Carlo simulated events for the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ final states, 12.5 million of each for the $D_s^- \to K^- K^+ \pi^-$ final state and 2.5 million of each for the $D_s^- \to \pi^- \pi^+ \pi^-$final state. These samples are needed to reduce the uncertainty in the acceptance description and the associated systematic uncertainty.

6.2 EVENT SELECTION

Event selection for the $B_s^0 \to D_s^{\mp} K^{\pm}$, $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$ channels consists of several steps; they can be summarized as: trigger and stripping selection, offline selection which contains also the particle identification requirements.

TRIGGER STRATEGY All events are required to be triggered by the signal (TOS) on the `"1TrackAllL0"` line in the HLT1 and on the `"2-"," 3-"`, or `"4-body" "TopoBBDT"` lines or the "inclusive ϕ" trigger in HLT2.

STRIPPING SELECTION $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates are selected using the stripping line called *B02DKD2HHHBeauty2CharmLine* based on the Stripping17 processing. The $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$ candidates are taken from a different stripping line, *B02DPiD2HHHBeauty2CharmLine*, in which the bachelor track has the pion mass hypothesis.

At stripping level no PID selection is applied; loose pre-selection criteria on the particle's kinematics and displacement from the primary interaction are used to select events. To further reduce the rate of selected events a multi-variate based classifier is used. In all cases the $D_{(s)}$ meson is reconstructed on the output of the stripping from three particles originating from the same vertex with the appropriate mass hypotheses; a "bachelor" track is then added to form the B candidates. A kinematic fit based on the DecayTreeFitter tool to the B decay chain is performed to further improve the selection. When computing the B mass, the D mass is constrained to its nominal value (separately for the D and D_s hypotheses). When computing the B lifetime, the B is additionally constrained to point to the primary interaction. The final selection results in a small number of multiple candidates, below 2%, and all are kept in the final analysis. In addition the B_s^0 mass is evaluated constraining the D_s mass to the current PDG value.

OFFLINE SELECTION The offline selection consists of different requirements:

- kinematic requirements common to all modes: reconstructed D_s mass in [1930,2015]MeV/c^{2}and D_s lifetime > 0 ps . Specific requirements on the

Table 29.: Branching fractions of B_s^0 decays are taken from Ref. [73], those of D_s decays are taken from the PDG [62], while $f_s/f_d$is the ratio of hadronization factors which corresponds to the combination of the LHCb measurements using hadronic and semi-leptonic B meson decays [74].

Channel	Branching fraction
$B_s^0 \to D_s^{\mp} K^{\pm}$	$(1.90 \pm 0.12(stat) \pm 0.13(syst) \pm 0.14(f_s/f_d)) \times 10^{-4}$
$B_s^0 \to D_s^- \pi^+$	$(2.95 \pm 0.05(stat) \pm 0.17(syst) \pm 0.22(f_s/f_d)) \times 10^{-3}$
$D_s^- \to KK\pi$	$(5.49 \pm 0.27) \times 10^{-2}$
$D_s^- \to K^-\pi^+\pi^-$	$(6.9 \pm 0.5) \times 10^{-3}$
$D_s^- \to \pi^-\pi^+\pi^-$	$(1.10 \pm 0.06) \times 10^{-2}$
f_s/f_d	$0.259 \pm 0.015((stat)\&(syst))$

D_s separation vertex are used to suppress charmless backgrounds (listed in Tab.30 for each D_s final state)

- Multivariate selection to suppress combinatorial background. The TMWA tool used in this analysis is the Boosted Decision Tree Gradient (BDTG) [75].
- PID requirements on the final charged particles reported in Tab.31
- selection to suppress mis-reconstructed peaking backgrounds (D , Λ_c^+ and J/ψ decays) specific for each D_s final states. The vetoes are applied in terms of PID requirements and invariant mass range typical of the vetoed particles (full list in Tab.30)

Optimization of both the BDTG output cut and the PID cut of the bachelor particle to maximize the signal significance is performed. The reported values are the nominal ones, after the optimization.

The core of the offline selection relies on the BDTG selection and on the PID requirement. The BDTG selection is devoted to suppress combinatorial events and it follows a data-driven approach. The classifier is trained on $B_s^0 \to D_s^- \pi^+$, $D_s^- \to KK\pi$ data samples where combinatorial events are selected from upper B_s^0 mass sidebands (m(B_s^0)> 5450 MeV/c^2). Signal events are defined as data weighted by the *sWeights* extracted from a fit to the B_s^0 mass where a gaussian (exponential) function describes signal (background) component. Kinematic variables are combined as input information for the classifier. Half of the $B_s^0 \to D_s^- \pi^+$ ($D_s^- \to KK\pi$) data sample is used to train the BDTG, the other half is used to check that the BDTG output does not suffer of overtraining effects. In Fig.44 the BDTG response for signal (red histogram) and background (blue histogram) events is shown separately for the training sample (left) and the test sample (right).

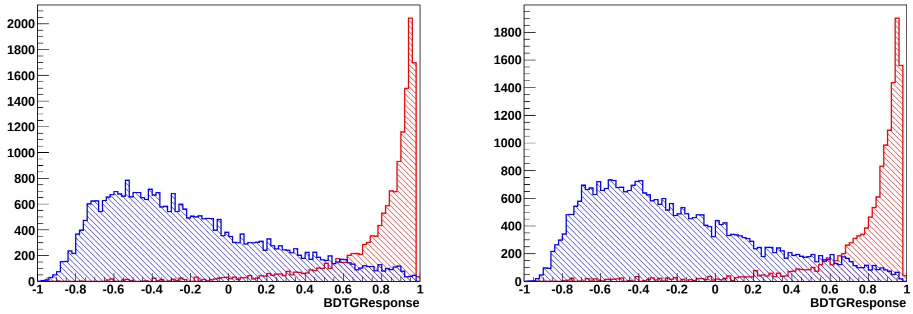


Figure 44.: BDTG response distributions for training (left) and test (right) samples of $B_s^0 \to D_s^- \pi^+$. The red histogram corresponds to signal weighted by the signal *sWeights*, while the blue shows background.

The BDTG selection is in common for all decay channels ($B_s^0 \to D_s^{\mp} K^{\pm}$, $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$)followed by the PID requirements on the final charged particles. The PID criteria are based on the difference of the logarithm of the likelihood (DLL) of a track to be of each particle type, a proton, a pion, a kaon, a muon or an electron. The observable is defined as $PIDK$ = DLL(K-π), similarly for $PIDp$ = DLL(p - π). The PID performance depends on the momentum track, so that a PID cut will change the momentum distribution of that final particle. As a consequence also the mass shape may change, and this is particularly relevant for those backgrounds which are selected thanks to the mis-identification of one or more particle. The PID performance, efficiency and mis-identification rate, are evaluated for each PID cut applied in the analysis using calibration samples of $D^* \to D\pi$ data from the PIDCalib tool [56]. The optimal PID cuts are reported in Tab.31.

Table 30.: Kinematic selection requirements and Vetos definition specific for each D_s mode.

D_s decay mode	Description	Requirement
$D_s^- \to KK\pi$	D_s vertex separation χ^2(wrt. B_s^0)	> 2
	D^0 veto:	
	$m(K^+K^-)$	< 1840 MeV/c^2
	D veto:	
	$PIDK$ of same charge K	> 10, or
	D_s under D hypothesis	not in $[1840, 1900]$ MeV/c^2
	$\Lambda_c$veto:	
	p veto, same charge K	$PIDK - PIDp > 5$, or
	D_s under $\Lambda_c$hypothesis	not in $[2255, 2315]$ MeV/c^2
$D_s^- \to \phi\pi^-$	$m(K^+K^-)$	in $[1000, 1040]$ MeV/c^2
$D_s^- \to K^{*0}K^-$	$m(K^+K^-)$	not in $[1000, 1040]$ MeV/c^2
	$m(\pi^-K^+)$	in $[842, 942]$ MeV/c^2
$D_s^- \to (KK\pi)_{nonres}$	$m(K^+K^-)$	not in $[1000, 1040]$ MeV/c^2
	$m(\pi^-K^+)$	not in $[842, 942]$ MeV/c^2
$D_s^- \to K^-\pi^+\pi^-$	D_s vertex separation χ^2(wrt. B_s^0)	> 9
	D^0 veto:	
	$m(K^+\pi^-)$	< 1750 MeV/c^2
	D veto:	
	$PIDK$ of opposite charge π	< -10, or
	$PIDK$ of opposite charge K	> 20, or
	D_s under D hypothesis	not in $[1839, 1899]$ MeV/c^2
	$\Lambda_c$veto:	
	p veto, same charge K	$PIDK - PIDp > 5$, or
	D_s under $\Lambda_c$hypothesis	not in $[2255, 2315]$ MeV/c^2
$D_s^- \to \pi^-\pi^+\pi^-$	D_s vertex separation χ^2(wrt. B_s^0)	> 9
	D^0 veto:	
	Both $m(\pi^+\pi^-)$	< 1700 MeV/c^2

Table 31.: PID selection requirements.

D_s decay mode	Description	Requirement
Bachelor track	$B_s^0 \to D_s^- \pi^+$	$PIDK < 0$
	$B_s^0 \to D_s^{\mp} K^{\pm}$	$PIDK > 5$
$D_s^- \to KK\pi$		
$D_s^- \to \phi\pi^-$	Both kaons	$PIDK > -2$
$D_s^- \to K^{*0}K^-$	Same charge kaon	$PIDK > 5$
	Opposite charge kaon	$PIDK > -2$
$D_s^- \to (KK\pi)_{nonres}$	Both kaons	$PIDK > 5$
$D_s^- \to K^-\pi^+\pi^-$	Kaon	$PIDK > 10$
	Both pions	$PIDK < 5$
	Both pions	$PIDp < 10$
$D_s^- \to \pi^-\pi^+\pi^-$	All pions	$PIDK < 10$
	All pions	$PIDp < 10$

6.3 MULTI-DIMENSIONAL FIT

The $B_s^0 \to D_s^{\mp} K^{\pm}$ decay mode is characterized by a relatively small branching fraction (10^{-4}) and by the presence of other similar hadronic decays which can mimic the signal. In order to get a better discrimination between signal and background events a fit (namely Multi-dimensional fit or MDfit) is performed to the distributions of B_s^0 and D_s reconstructed masses and of the PID of the bachelor kaon ($PIDK$). The $PIDK$ observable is able to disentangle decays with different bachelor particles as pions, kaons and protons. This strategy has allowed to use a more efficient selection, with looser cuts, since backgrounds event can be separated exploiting the different distributions of the observables used in the MDfit. The Multi-dimensional fit is common to the two different approaches that have been developed to extract the CP-observables. The first approach is a time *sFit* where the *sWeights* are obtained from the Multi-dimensional fit and applied to the time fit in order to subtract all the background events. The second approach is to perform a classical fit in four dimension: B_s^0 and D_s masses, $PIDK$ and decay time without any background subtraction. More details on the time fitters will be given in Sec.6.6.

The structure of the PDF used by the Multi-dimensional fit can be defined as follows:

$$PDF(m_{B_s}, m_{D_s}, PIDK) = \sum^{i} PDF^i(m_{B_s}) \cdot PDF^i(m_{D_s}) \cdot PDF^i(PIDK) \tag{6.1}$$

where i indicates signal and backgrounds components. Three kinds of backgrounds can be distinguished: decays in which one or more particles are not reconstructed (partially reconstructed backgrounds), fully reconstructed decays in which one or more particles could be mis-identified (Mis-ID backgrounds) and events which are randomly selected as signal events (combinatorial events).

The Multi-dimensional fit strategy can be illustrated through these key points:

- Simultaneous Maximum Unbinned likelihood fit to the five D_s final states in order to get advantage in the determination of common parameters while exploiting the different features of each decay mode
- The Multi-dimensional fit is performed for both $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ independently
- Observables fit ranges:
  - m(B_s^0) $\in$ [5300,5800] MeV/c^2
  - m(D_s) $\in$ [1939,2015] MeV/c^2
  - log($PIDK$) $\in$ [log(5),log(150)] for $B_s^0 \to D_s^{\mp} K^{\pm}$
  - $-PIDK \in$ [0,150] for $B_s^0 \to D_s^- \pi^+$
- Many of the probability density functions (PDFs) make use of predefined shapes that have as free parameter only the normalization. They are generally called templates and can be obtained through different models depending on what is more appropriate: `RooKeysPdf` and `RooBinned1DQuinticBase`.
- Probability density functions describing the $PIDK$ distributions are obtained using the PID calibration samples The $PIDK$ cuts for the bachelor particles are applied to these samples and then they are reweighted to match the

kinematic distributions of signal and backgrounds decays. The so obtained distributions are used to get the $PIDK$ templates for the Multi-dimensional fit.

- PID performance are included to correct mass shapes for Mis-ID backgrounds and to constrain the expected yields.
- All B_s^0 mass templates taken from MC are corrected for a mass shift due to imperfect detector alignment which has been evaluated as the difference between the signal mass peak in data and MC: $\Delta = 3.9 \pm 0.2$ MeV/c^2.
- No correlation has been assumed among the observables of the fit, so that the three PDFs (PDF(m(B_s^0)), PDF(m(D_s)) and PDF($PIDK$)) factorizes. Studies confirming this assumption have been done on MC.

6.3.1 *Probability Density Functions*

SIGNAL The shapes of the signal mass distributions are affected by the detector resolution and by the contribution of radiative decays of the B_s^0 and D_s decays. To describe the radiative tail in the lower mass region a gaussian model is not enough. It has been found that a Double Crystal Ball (Eq.5.1) function [58] well suite the signal mass distributions either for B_s^0 and D_s mass. The parameters of the Double Crystal Ball are extracted from $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ simulated events for each of the 5 subsamples: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$,$D_s^- \to K^{*0}K^-$,$D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$ and they are listed in Tab.32 and in Tab. 33 respectively for $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$. To avoid to consider the samll fraction of mis-reconstructed events, selected decays are required to match the correct signal decay chain (true ID). In the simulated sample the selection steps are applied as well as PID efficiencies. For brevity only the B_s^0 mass shapes parameters are reported. Signal shapes for $B_s^0 \to D_s^{\mp}K^{\pm}$ ($B_s^0 \to D_s^-\pi^+$) are reported in Fig.46 (in Fig.45) and the Double Crystal Ball parameters are listed in .

In the fit to data all signal B_s^0 and D_s mass parameters are fixed, except the mean and corrections to account for differences in the mass resolutions between data and MC (evaluated in the $B^0 \to D^-\pi^+$ decay mode).

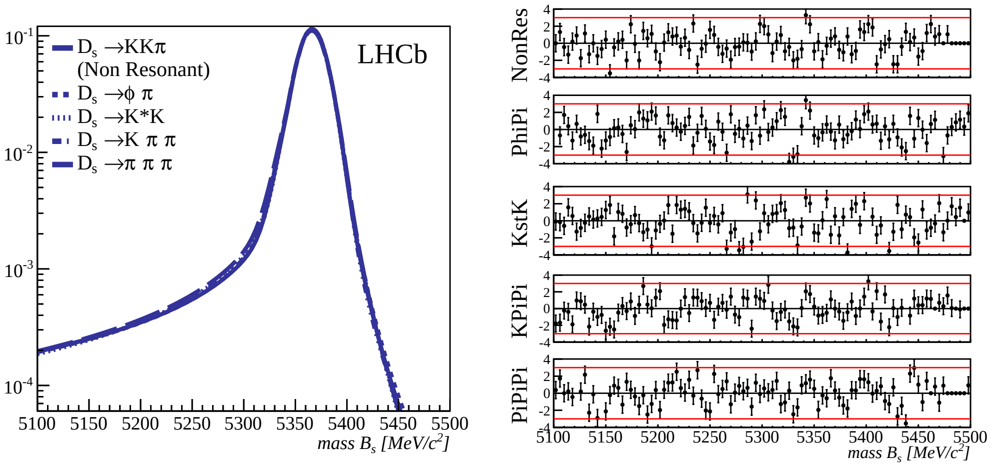


Figure 45.: Signal B_s^0 mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated from MC on the joint magnet up and down samples, separately for each D_s final states. The right plots show the pull distributions.

The signal $PIDK$ shapes are obtained as described in previous section and they are shown in Fig.47 for both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$.

COMBINATORIAL BACKGROUND Combinatorial is a combination of pure combinatorial events, where all the four final particles are selected randomly, and of true D_s accompanied by a random bachelor particle. The shape for combinatorial backgorund is determined on data looking at the events in the B_s^0 mass upper sideband within the nominal D_s mass range ([1939,2015] MeV/c^2).

The B_s^0 mass distribution of combinatorial events is described:

- in $B_s^0 \to D_s^\mp K^\pm$, all D_s final states by a single exponential
- in $B_s^0 \to D_s^- \pi^+$, all $D_s^- \to KK\pi$ final states by a single exponential plus a constant
- in $B_s^0 \to D_s^- \pi^+$, $D_s^- \to K^- \pi^+ \pi^-$ and $D_s^- \to \pi^- \pi^+ \pi^-$ final states by a single exponential

Combinatorial shapes for $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^\mp K^\pm$) events are shown in Fig.48 (49) The parameters are left free in the Multi-dimensional fit.

Table 32.: Parameters of the Double Crystal Ball function describing the signal B_s^0 shapes of $B_s^0 \to D_s^-\pi^+$, obtained from signal MC, separately for each D_s final state. Parameters without uncertainties are fixed in the fit.

Parameter	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu(B_s^0)$ [MeV/c^2]	5366.3 ± 0.04	5366.3 ± 0.02	5366.3 ± 0.02	5366.2 ± 0.04	5366.2 ± 0.05
σ_1 [MeV/c^2]	11.54 ± 0.13	16.60 ± 0.17	11.65 ± 0.14	11.43 ± 0.13	11.39 ± 0.10
σ_2 [MeV/c^2]	16.18 ± 0.09	11.49 ± 0.08	15.00 ± 0.09	16.87 ± 0.12	17.65 ± 0.10
α_1	1.91	-2.09	1.70	1.91	2.09
α_2	-2.04	1.89	-1.84	-2.26	-2.33
n_1	1.13 ± 0.02	5.27 ± 0.26	1.27 ± 0.02	1.16 ± 0.02	1.27 ± 0.02
n_2	6.14 ± 0.51	1.15 ± 0.02	9.66 ± 0.09	4.12 ± 0.33	4.02 ± 0.36
f	0.54 ± 0.04	0.44 ± 0.02	0.54 ± 0.03	0.54 ± 0.04	0.70 ± 0.02

To account for the two types of combinatorial events, the D_s mass shape is defined by the sum of the Double Crystal Ball used for signal and a single exponential for $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ final states. For the $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ final states a single exponential is enough since the true D_s are completely removed by the BDTG cut. In this way the fraction of events which are true D_s are treated as signal events in the D_s mass variable and the remaining events are considered as pure combinatorial events. The D_s mass distribution for combinatorial events are reported in Fig.50 for $B_s^0 \to D_s^-\pi^+$ and in Fig.51 for $B_s^0 \to D_s^{\mp}K^{\pm}$. Apart from the Double Crystal Ball signal parameters which are in common with the signal, the exponential parameters as well as the relative fractions of true D_s are floating in the fit to data.

Combinatorial events can be composed by events which have as bachelor particle pions, kaons or protons. In order to accurately describe the distribution of $PIDK$ for these events the PDFs is defined as follows:

$$PDF_{PIDK} = f_{PIDK1}PDF_{PIDK1} + (1 - f_{PIDK1})[f_{PIDK2}PDF_{PIDK2} + (1 - f_{PIDK2}PDF_{PIDK3})] \tag{6.2}$$

and the relative fractions of each contribution are let free in the fit to data. The different PDF_{PIDKi} shapes are obtained from the PID calibration samples for pions and kaons while for protons a different set of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ sample is used.

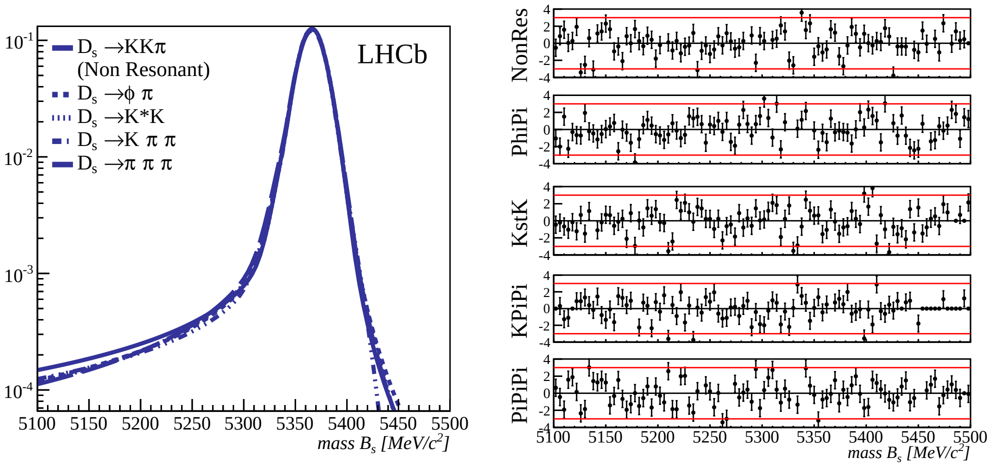


Figure 46.: Signal B_s^0 mass shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$ as evaluated from MC on the joint magnet up and down samples, separately for each D_s final states. The right plots show the pull distributions

As usual, $PIDK$ distributions from the calibration samples need to be reweighted to match the kinematic distributions of combinatorial events data. This is done through a data driven method which allows to extract the kinematic distributions (p_T,nTracks) for the bachelor particle of combinatorial events in the nominal mass range ([5300,5800]MeV/c^2) and uses these distributions to weight the $PIDK$ distributions from the calibration samples.

In Fig.52 $PIDK$ templates for combinatorial events are reported for both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ . .

partially reconstructed backgrounds Decays in which one or more particles are missed are called here partially reconstructed backgrounds. Typically the B_s^0 mass distribution of these events peak in the low mass range but they can extend till the signal region. Partially reconstructed backgrounds are considered:

- $B_s^0 \to D_s^{*-} \pi^+$for the $B_s^0 \to D_s^- \pi^+$ decay mode
- $B_s^0 \to D_s^- \rho^+$, $B_s^0 \to D_s^{*-} \pi^+$, $\Lambda_b^0 \to D_s^- p$, $\Lambda_b^0 \to D_s^{*-} p$ for the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay mode.

Table 33.: Parameters of the Double Crystal Ball function describing the signal B_s^0 shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC, separately for each D_s final state.

Parameter	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu(B_s^0)$ [MeV/c^2]	5366.3 ± 0.02	5366.5 ± 0.03	5366.4 ± 0.03	5366.0 ± 0.08	5366.2 ± 0.05
σ_1 [MeV/c^2]	10.72 ± 0.12	11.23 ± 0.08	10.77 ± 0.09	11.27 ± 0.30	11.39 ± 0.10
σ_2 [MeV/c^2]	16.01 ± 0.11	17.02 ± 0.10	15.34 ± 0.08	19.41 ± 0.93	17.65 ± 0.10
α_1	2.21	2.21	2.05	2.40	2.09
α_2	-2.42	-2.19	-2.03	-3.42	-2.33
n_1	1.00 ± 0.02	1.12 ± 0.02	1.21 ± 0.02	0.98 ± 0.05	1.27 ± 0.02
n_2	3.15 ± 0.22	3.61 ± 0.19	6.57 ± 0.46	0.52 ± 0.29	4.02 ± 0.36
f	0.62 ± 0.02	0.70 ± 0.02	0.58 ± 0.13	0.78 ± 0.05	0.70 ± 0.01

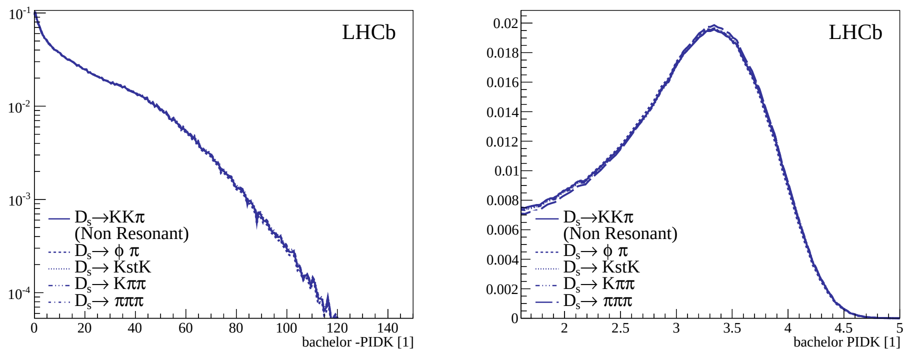


Figure 47.: `RooBinned1DQuinticBase` templates for signal PIDK. Left: $B_s^0 \to D_s^- \pi^+$. Right: $B_s^0 \to D_s^{\mp} K^{\pm}$.

The B_s^0 mass shapes for these partially reconstructed backgrounds are taken using simulated backgrounds events reconstructed under the wrong mass hypothesis and they are corrected for PID efficiencies and data/MC differences. Since all the considered partially reconstructed backgrounds decay through a D_s meson, the D_s mass shapes is taken to be as for the signal.

The $PIDK$ shapes of partially reconstructed backgrounds are determined in the usual way using the PID calibration samples.

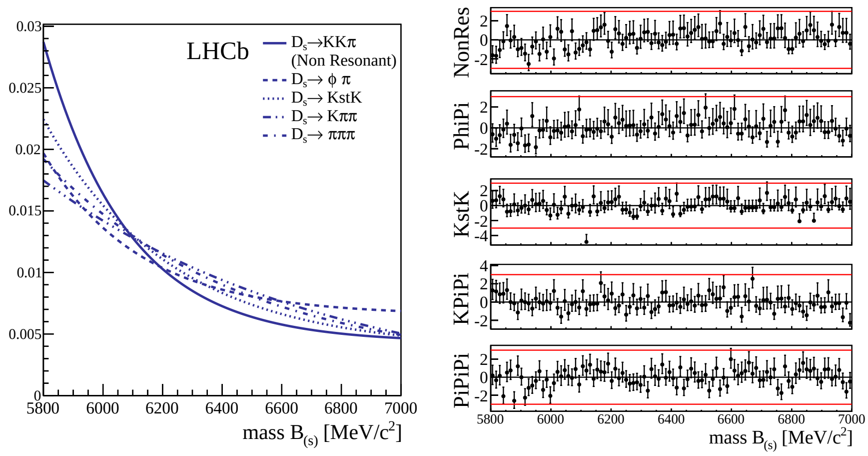


Figure 48.: Combinatorial background B_s^0 mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated, separately for each D_s final state. The right plots show the pull distributions

FULLY RECONSTRUCTED BACKGROUNDS Fully reconstructed backgrounds are considered those decays where particle misidentification takes place and/or where another mother particle decays exactly in the same final state as the signal one. In Tab.34 the fully reconstructed backgrounds included in this analysis are listed.

Different methods have been used to obtain the B_s^0 mass shapes for these backgrounds.

- Mass shapes for backgrounds which decay in the same signal final state use the Double Crystal Ball of the signal. The means are shifted by the well known m(B_s^0) - m(B) = 86.9 MeV/c^2 mass difference while the sigmas are corrected by the width ratio $B_s^0 \to D_s^{\mp} h^{\pm} / B^0 \to D_s^{\mp} h^{\pm}$ measured in MC samples.

- The $B^0 \to D^- \pi^+$ as background of the $B_s^0 \to D_s^- \pi^+$ and the $B_s^0 \to D_s^- \pi^+$ as background of the $B_s^0 \to D_s^{\mp} K^{\pm}$ are the most relevant fully reconstructed backgrounds (second row of Tab.34). For this reason their B_s^0 mass shapes are obtained by a data-driven approach which employs a pure data sample of the considered decay reconstructed under the signal mass hypothesis for the misidentified particle. Through the PID calibration samples, PID efficiency

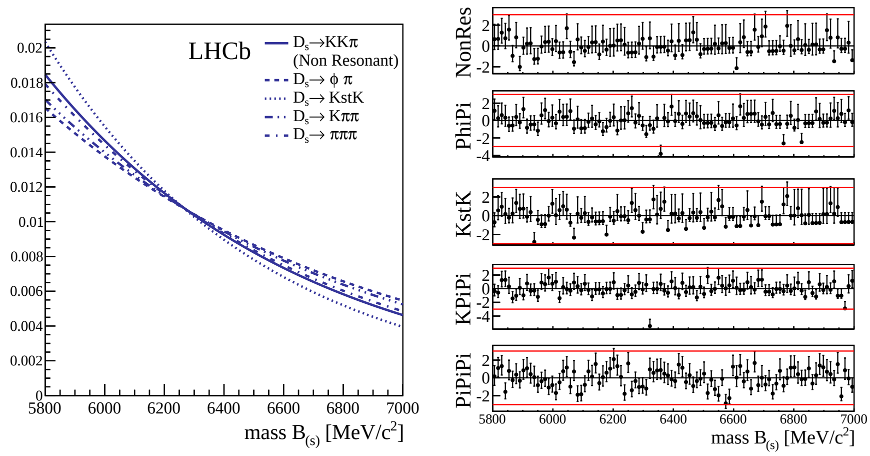


Figure 49.: Combinatorial background B_s^0 mass shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$ as evaluated, separately for each D_s final state in the upper B_s^0 sidebands in order to study the PDF in a region of pure combinatorial events. The right plots show the pull distributions

and Mis-ID rate are extracted as functions of the momentum and applied to obtain the correct B_s^0 mass shape.

- The shapes of the remaining fully reconstructed backgrounds are taken from MC, corrected by the momentum-dependent PID efficiency and by residual kinematic discrepancies between data and MC.

The D_s mass shape for the fully reconstructed backgrounds is considered as the signal one if it contains a D_s meson. In the case of the $B^0 \to D^- \pi^+$ decay the same data driven approach is used to get the D_s mass shapes. The $PIDK$ shapes are obtained from the PID calibration samples, which got reweighted to match the kinematics found in the relevant signal samples. When the final state is the same the one of the signal, the same $PIDK$ shape is employed. In the other cases a fixed shape is used.

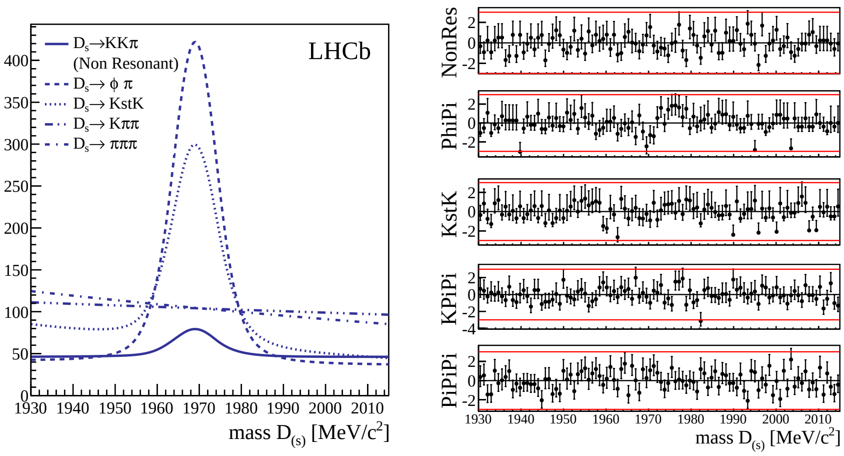


Figure 50.: Combinatorial background D_s mass shapes of $B_s^0 \to D_s^- \pi^+$ as evaluated (in the upper B_s^0 sidebands), separately for each D_s final state. The right plots show the pull distributions

6.3.2 *Summary of fixed yields*

All backgrounds yields are free in the fit except yields of those backgrounds which can be constrained by means of measured yields (as for $B^0 \to D^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$ decay modes), ratios of branching fractions and PID efficiencies. A summary of the yields is given in Tables 35 and 36.

6.3.3 *Results of the Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$*

The fit results to the $B_s^0 \to D_s^- \pi^+$ mass distribution, separately for each considered sample are shown in Fig. 53, Fig. 54, Fig. 55 for the B_s^0 and D_s masses, and $PIDK$, respectively. Finally, the fitted parameters are listed in Tab. 73. The N_i are the yields of the signal and background contributions. In the signal model only the means of the double Crystal Ball are floating for the B_s^0 and D_s masses. The parameters f_1 is the relative fraction between modes in the group 1: $B_s^0 \to D_s^{*-} \pi^+$, $B^0 \to D_s^- \pi^+$. The fractions $fComb_{B_s}$ are the fractions between an exponential and flat distribution in B_s^0 mass, whereas $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. The $fComb_{PIDK}$ is a fraction between

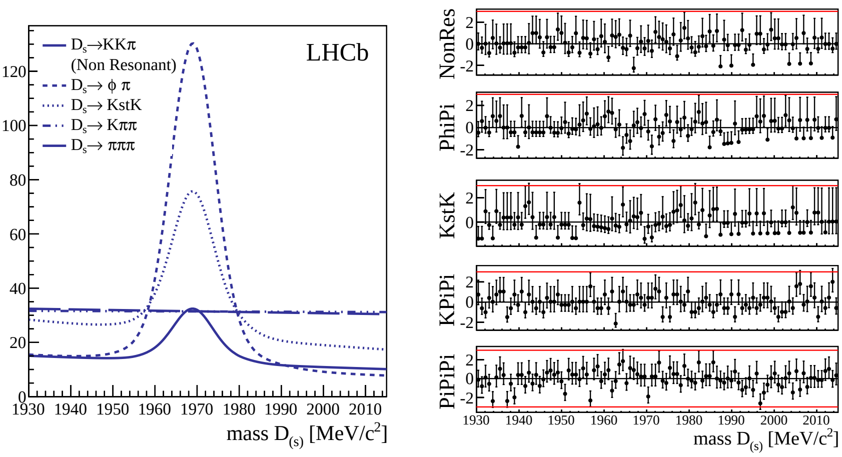


Figure 51.: Combinatorial background D_s mass shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$ as evaluated (in the upper B_s^0 sidebands), separately for each D_s final state. The right plots show the pull distributions.

pion and kaon components in the $PIDK$ combinatorial shape, shared among all subsamples. The total signal yield of $B_s^0 \to D_s^- \pi^+$ decays is found to be of $\approx$ 28 200 events.

6.3.4 *Results of the Multidimensional fit to* $B_s^0 \to D_s^{\mp} K^{\pm}$

The fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ distribution is constructed analogously to the $B_s^0 \to D_s^- \pi^+$ fit.

The backgrounds are arranged into five groups, defined in Tab. 39.

The expected yields are summarized in Tab. 35.

The fit results to the $B_s^0 \to D_s^{\mp} K^{\pm}$ separately for each considered sample, and combined, are shown in Figs. 56, 57, 58, and 59. Finally, the fitted parameters are listed in Tab. 40. The N_i are the yields of the signal and background contributions. The floating parameters of the signal model are only the means of the double Crystal Ball used to describe B_s^0 and D_s masses. The fractions $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. The $fComb_{PIDK(1,2)}$ are a fraction between pion, kaon and proton components in the PIDKcombinatorial shape,

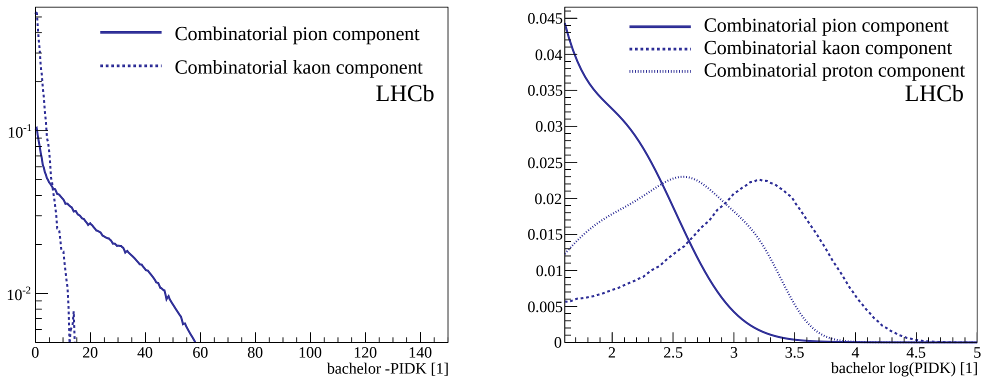


Figure 52.: Combinatorial background $PIDK$ templates used in the $B_s^0 \to D_s^- \pi^+$ fit (left) and $B_s^0 \to D_s^\mp K^\pm$ fit (right), combined for magnet polarities.

shared among all subsamples. The total signal yield of $B_s^0 \to D_s^\mp K^\pm$ events is of $\approx$ 1750 events.

As it can be seen by the plots, the result of the Multi-dimensional fitter provides a good description of the data allowing to discriminate signal and background events to extract the *sWeights* on a per-event basis. Such discrimination between signal and backgrounds represents a crucial ingredient to measure the CP observables of the $B_s^0 \to D_s^\mp K^\pm$ decay in both the two developed approaches:

- one is a time sFit which is based on a statistics-based backgrounds subtraction (sPlot techiniques) applying the *sWeights* event by event. As already seen in previous chapter, the sFit simplifies the analysis requiring only to describe PDF of the signal component

- the second approach is the cFit, i.e. a fit to the B_s^0 mass, D_s mass, $PIDK$ and B_s^0 decay time in which signal and background decays have to be correctly describe in all the four observables to measure the CP observables in the $B_s^0 \to D_s^\mp K^\pm$ decays. In the cFit the results of the Multi-dimensional fitter (yields and PDF's parameters) are fixed. As it will be shown in the next sections, this second approach requires many more ingredients with respect to the sFit.

Table 34.: Fully reconstructed backgrounds considered in the nominal fits to $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$.

Channel	Background to	How
$B_s^0 \to D_s^- K^+$	$B_s^0 \to D_s^- \pi^+$	$\pi^+ \to K^+$ (bachelor)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\pi^- \to K^-$ (D)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (\pi^+K^-\pi^-)\pi^+$	$K^+ \to \pi^+$ and $\pi^- \to K^-$ (D)
$B^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- \pi^+$	Same final state
$\Lambda_b \to \Lambda_c^- \pi^+ \to (\bar{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\bar{p} \to K^-$ (D)
$B_s^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- K^+$	$\pi^+ \to K^+$ (bachelor)
$B^0 \to D_s^- K^+$	$B_s^0 \to D_s^- K^+$	Same final state
$B^0 \to D^- K^+ \to (K^+\pi^-\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\pi^- \to K^-$ (D)
$\Lambda_b \to \Lambda_c^- K^+ \to (\bar{p}K^+\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\bar{p} \to K^-$ (D)

Table 35.: Constrained yields in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit.

Par.	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$B^0 \to D^- K^+$	17.0	0.0	5.0	0.0	0.0
$B^0 \to D^- \pi^+$	14.0	0.0	3.0	3.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ K^-$	15.0	2.0	4.0	0.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	11.0	1.0	3.0	0.0	0.0

Table 36.: Constrained yields in the $B_s^0 \to D_s^- \pi^+$ MD fit.

Background	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$B^0 \to D^- \pi^+$	374.0	6.0	93.0	30.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	290.0	36.0	69.0	1.0	0.0
$B_s^0 \to D_s^{\mp} K^{\pm}$	40.0	47.0	40.0	8.0	21.0

Table 37.: Definition of background groups in the $B_s^0 \to D_s^- \pi^+$ MD fit.

Group	Channels
1	$B_s^0 \to D_s^{*-} \pi^+$, $B^0 \to D_s^- \pi^+$,
2	$B^0 \to D^- \pi^+$
3	$\Lambda_b^0 \to \Lambda_c^+ \pi^-$
4	$B_s^0 \to D_s^{\mp} K^{\pm}$
5	Combinatorial background

Table 38.: Fitted values of the parameters for the $B_s^0 \to D_s^- \pi^+$ signal MD fit. The $slopes_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units.

Parameter	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s \to D_s\pi}$	4667 ± 76	10284 ± 108	7304 ± 91	1671 ± 44	4338 ± 73
N_{Group1}	88 ± 19	112 ± 22	89 ± 18	31 ± 11	54 ± 16
N_{Comb}	2587 ± 67	1153 ± 57	1173 ± 51	1096 ± 38	3021 ± 65
$N_{B_s \to D_s\pi}$	total yield: 28264 ± 182				
$slope_{B_s}$	-6.20 ± 0.66	-8.75 ± 1.19	-8.41 ± 0.91	-1.98 ± 0.25	-2.19 ± 0.16
$slope_{D_s}$	-4.53 ± 0.93	-2.75 ± 1.77	-4.81 ± 1.50	-4.24 ± 1.29	-2.58 ± 0.78
$fComb_{B_s}$	0.78 ± 0.06	0.66 ± 0.05	0.77 ± 0.05	only exponential	only exponential
$fComb_{D_s}$	0.97 ± 0.02	0.62 ± 0.03	0.84 ± 0.03	only exponential	only exponential
$\mu(B_s^0)$ [MeV/c^2]	5369.9 ± 0.1				
$\mu(D_s)$ [MeV/c^2]	1969.8 ± 0.1				
f_1	0.807 ± 0.787				
$fComb_{PIDK}$	0.913 ± 0.009				

Table 39.: Definition of the background groups in the $B_s^0 \to D_s^\mp K^\pm$ MD fit.

Group	Type	Channels
1	Bachelor K	$B^0 \to D_s^- K^+$
2	Bachelor π	$B_s^0 \to D_s^- \pi^+$, $B_s^0 \to D_s^{*-} \pi^+$, $B_s^0 \to D_s^- \rho^+$
3	Λ_b decays	$\Lambda_b^0 \to D_s^- p$, $\Lambda_b^0 \to D_s^{*-} p$
4	–	$B^0 \to D^- K^+$, $B^0 \to D^- \pi^+$
5	–	$\Lambda_b^0 \to \Lambda_c^+ K^-$, $\Lambda_b^0 \to \Lambda_c^+ \pi^-$
6	Combinatorial	–

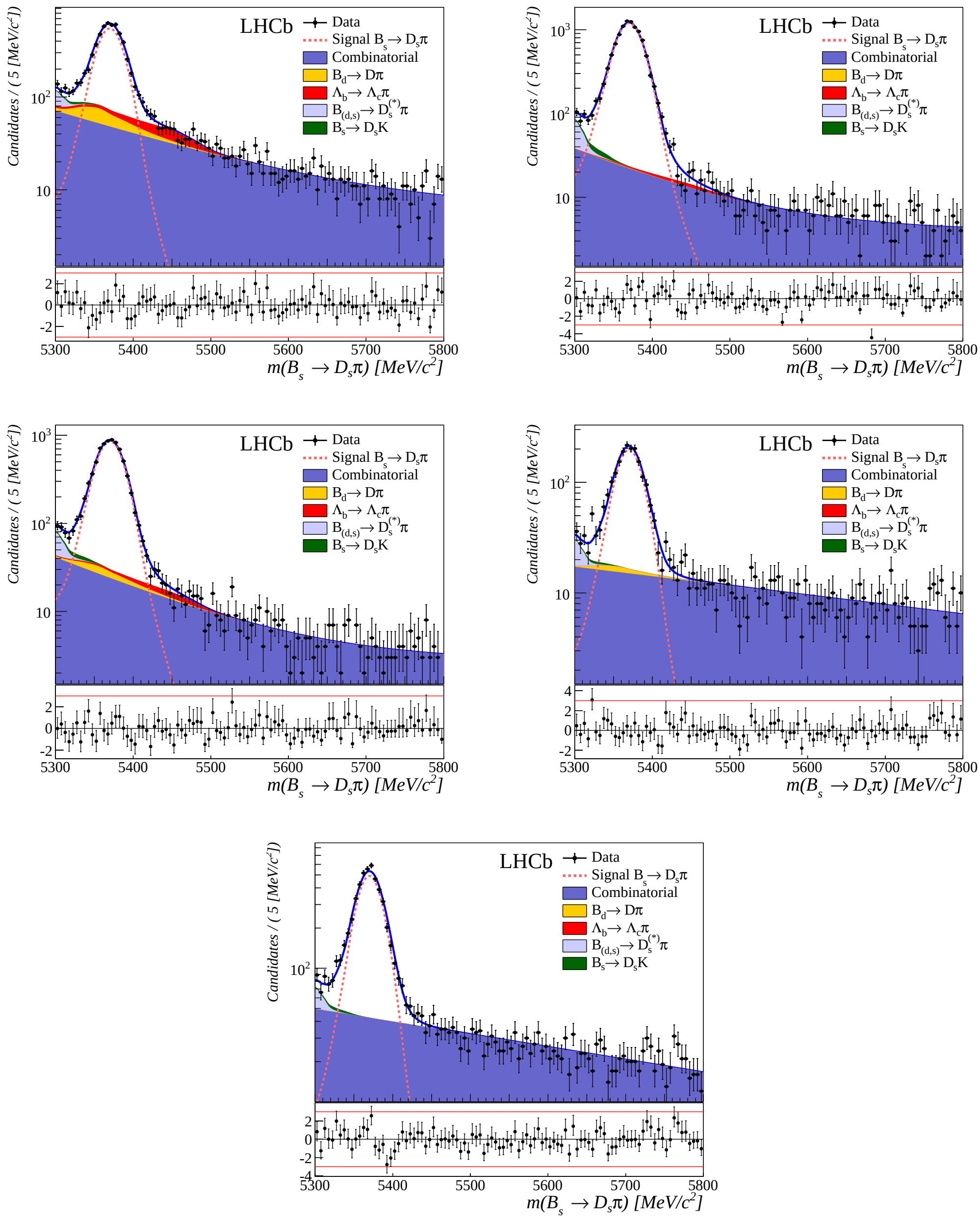


Figure 53.: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

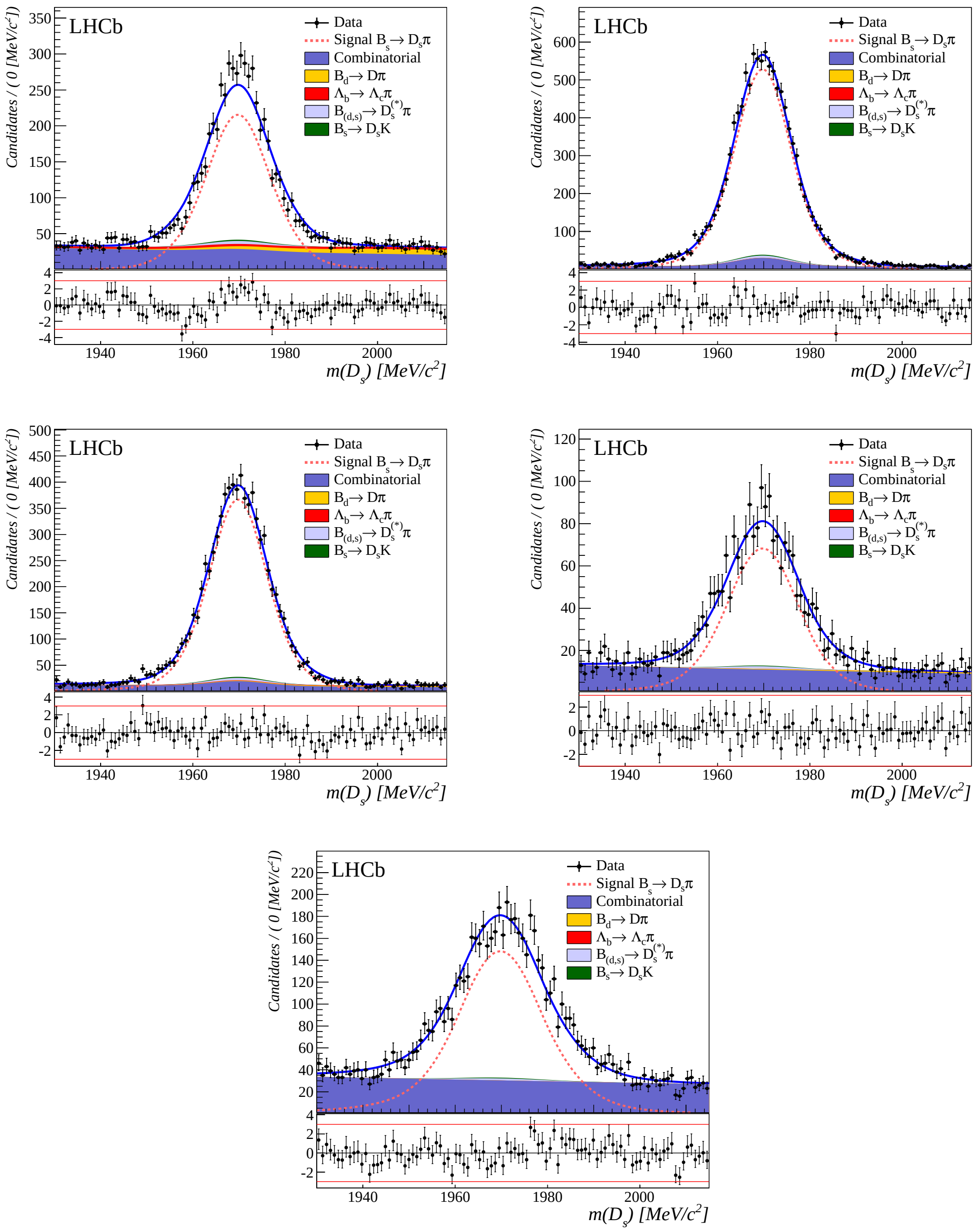


Figure 54.: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

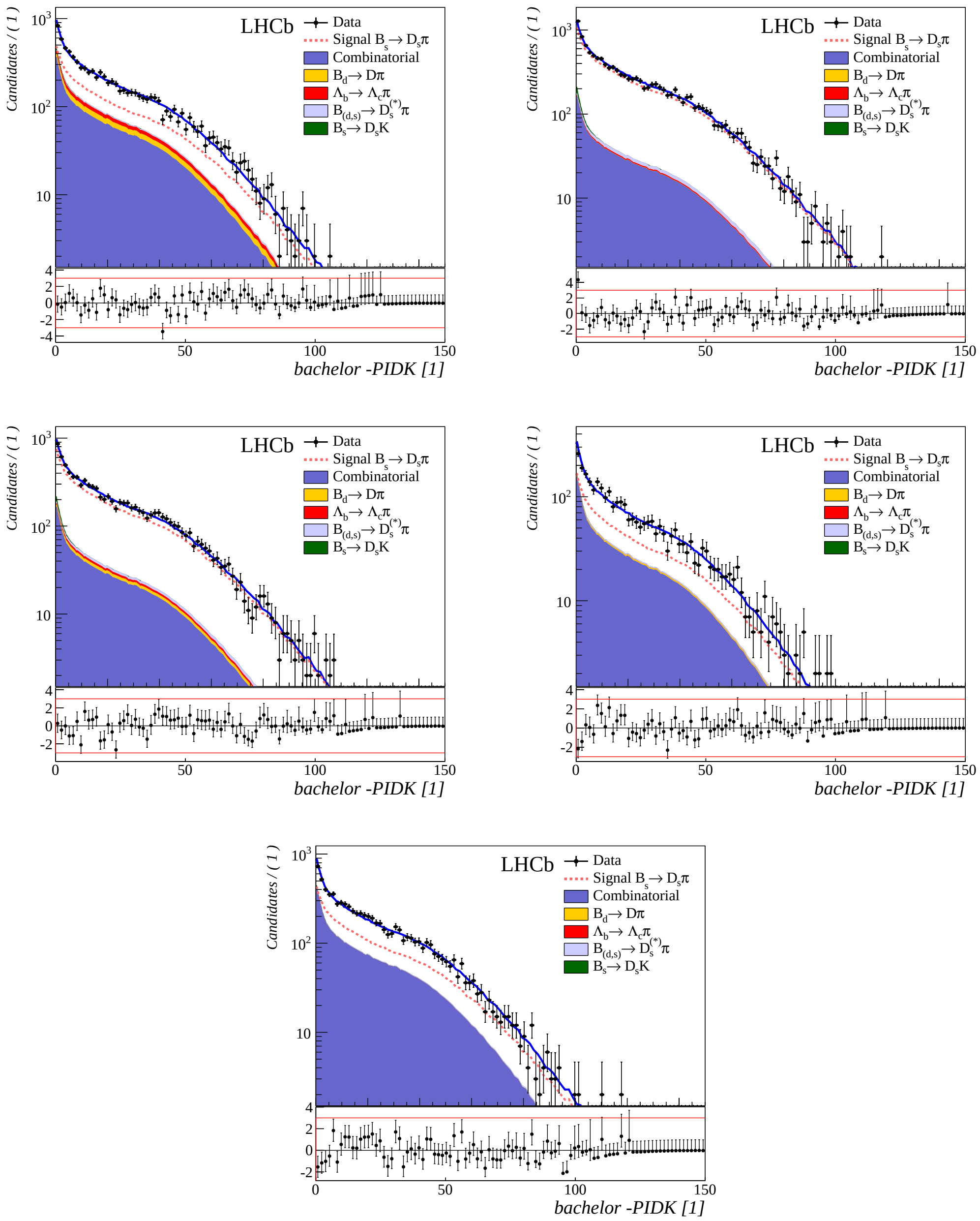


Figure 55.: Result of the simultaneous fit to the $B_s^0 \rightarrow D_s^- \pi^+$ candidates. Distributions of $-PIDK$ mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \rightarrow (KK\pi)_{nonres}$, $D_s^- \rightarrow \phi\pi^-$, $D_s^- \rightarrow K^{*0}K^-$, $D_s^- \rightarrow K^-\pi^+\pi^-$, $D_s^- \rightarrow \pi^-\pi^+\pi^-$.

Table 40.: Fitted values of the parameters for the $B_s^0 \to D_s^{\mp}K^{\pm}$ signal MD fit. The $slopes_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units.

Parameter	$D_s^- \to (KK\pi)_{nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s \to D_s K}$	309 ± 21	576 ± 27	475 ± 24	107 ± 13	301 ± 22
N_{Group1}	18 ± 7	34 ± 8	39 ± 9	9 ± 5	27 ± 8
$N_{Group23}$	225 ± 21	498 ± 30	327 ± 24	89 ± 16	258 ± 25
N_{Comb}	487 ± 27	311 ± 26	258 ± 22	428 ± 24	946 ± 37
$N_{B_s \to D_s K}$	total yield: 1768 ± 49				
$slope_{B_s}$	-3.17 ± 0.41	-1.82 ± 0.63	-2.91 ± 0.61	-1.09 ± 0.38	-1.55 ± 0.27
$slope_{D_s}$	-4.29 ± 2.05	-2.92 ± 3.54	-3.39 ± 3.10	-0.00 ± 1.02	-1.97 ± 1.36
$fComb_{D_s}$	0.95 ± 0.05	0.54 ± 0.07	0.80 ± 0.08	only exponential	only exponential
$\mu(B_s^0)$ [MeV/c^2]	5370.4 ± 0.5				
$\mu(D_s)$ [MeV/c^2]	1969.8 ± 0.1				
$fComb_{\mathrm{PIDK1}}$	0.504 ± 0.032				
$fComb_{\mathrm{PIDK2}}$	0.346 ± 0.079				

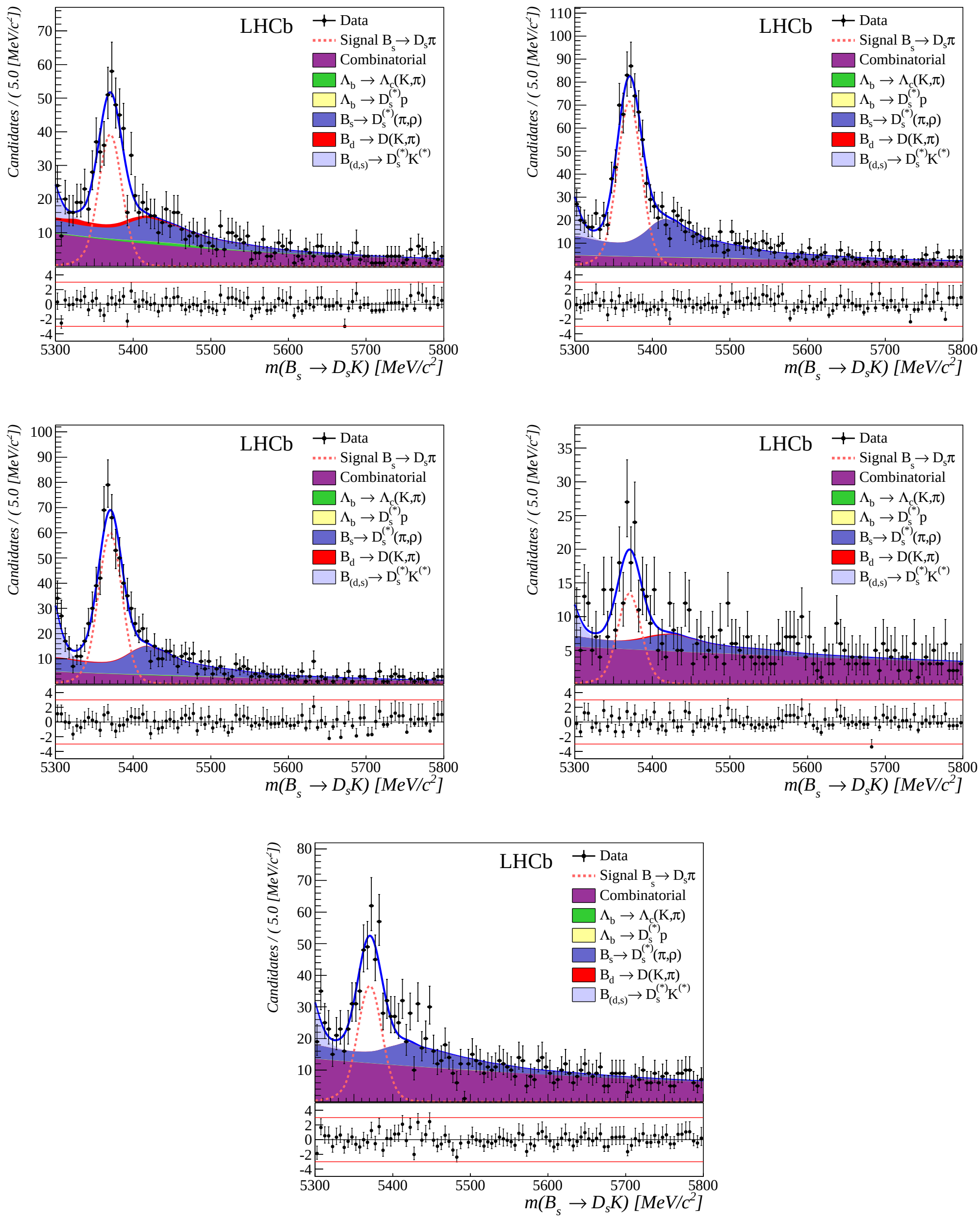


Figure 56.: Result of the simultaneous fit to the $B_s^0 \to D_s^\mp K^\pm$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

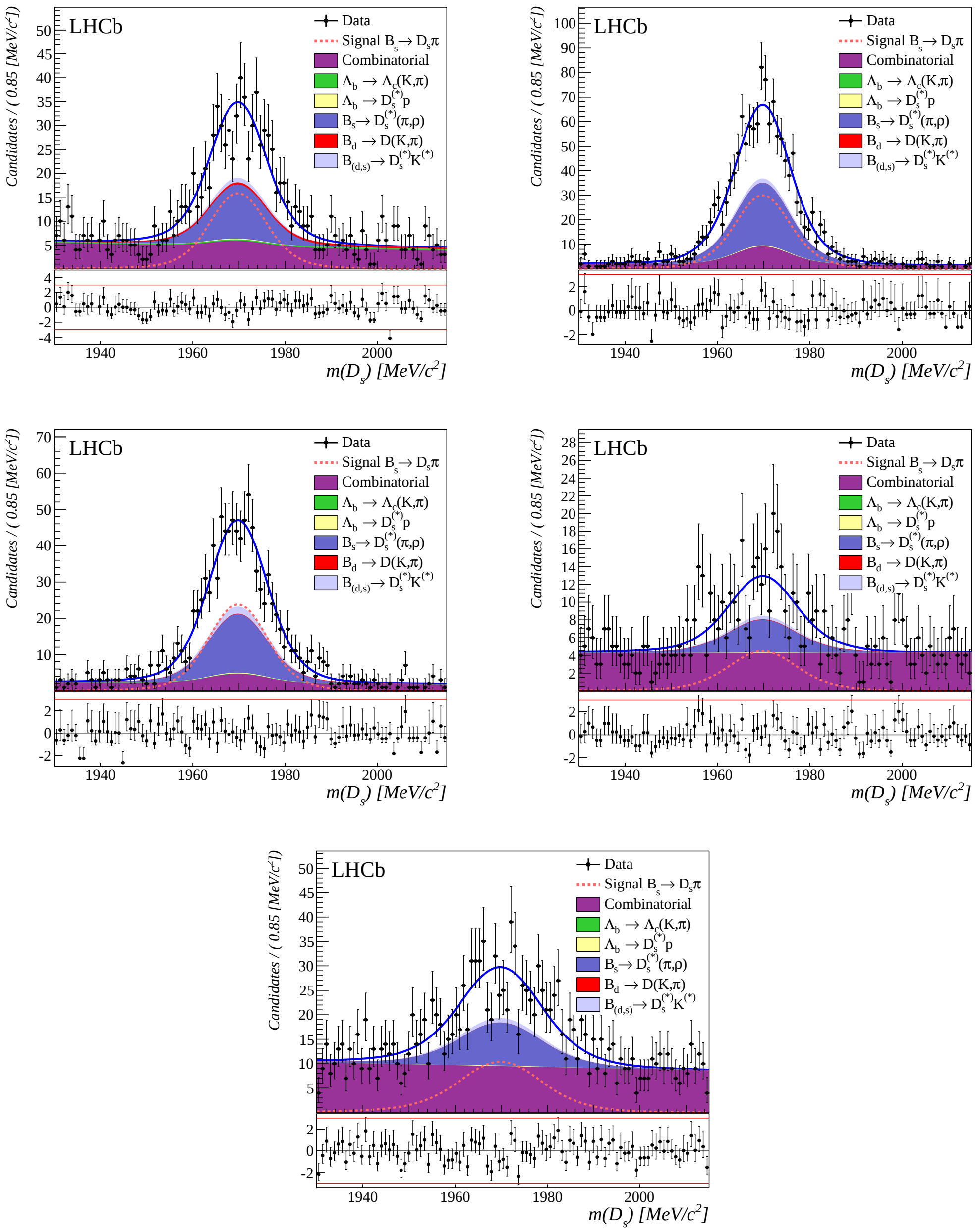


Figure 57.: Result of the simultaneous fit to the $B_s^0 \to D_s^\mp K^\pm$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

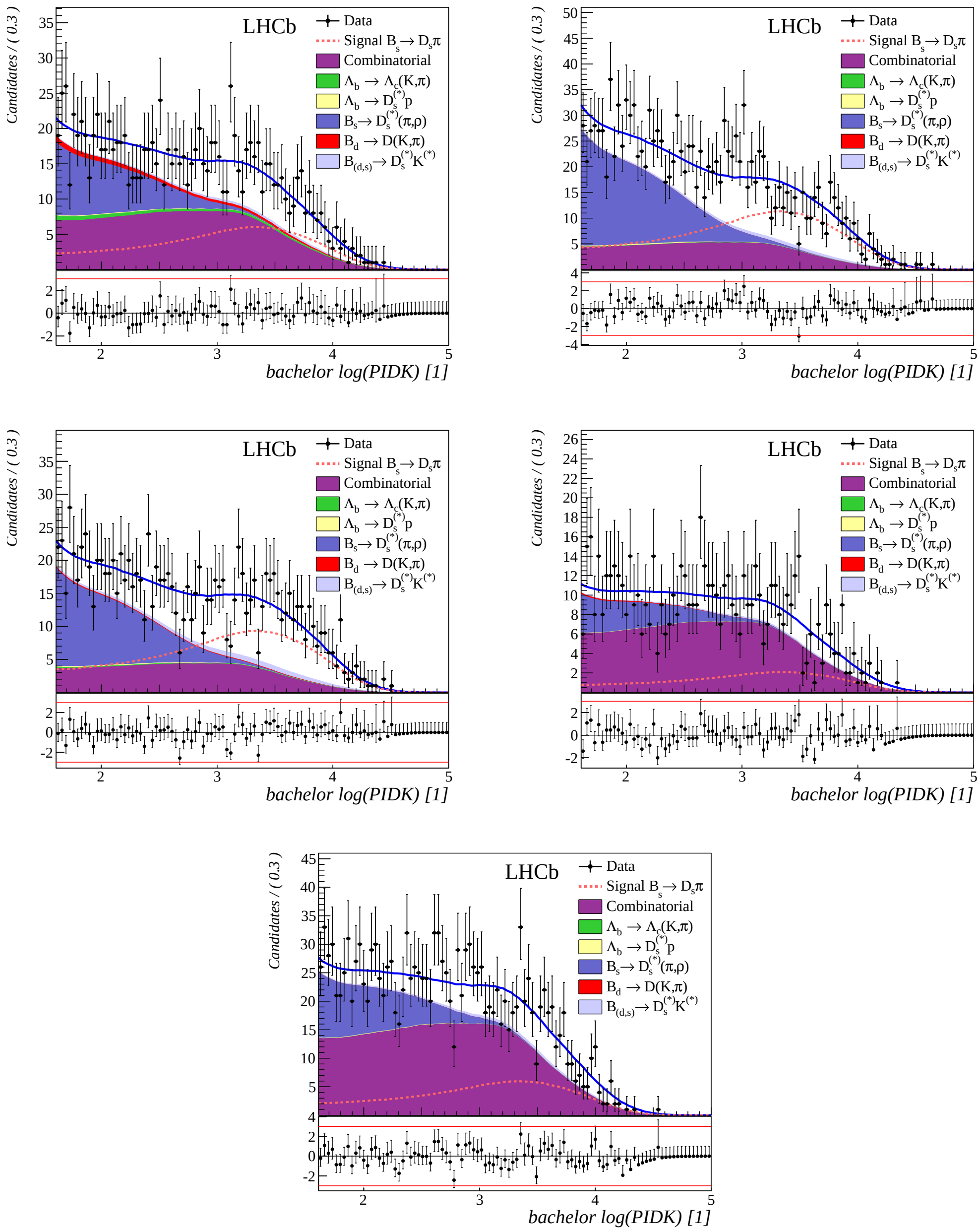


Figure 58.: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $log(PIDK)$ mass for D_s final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

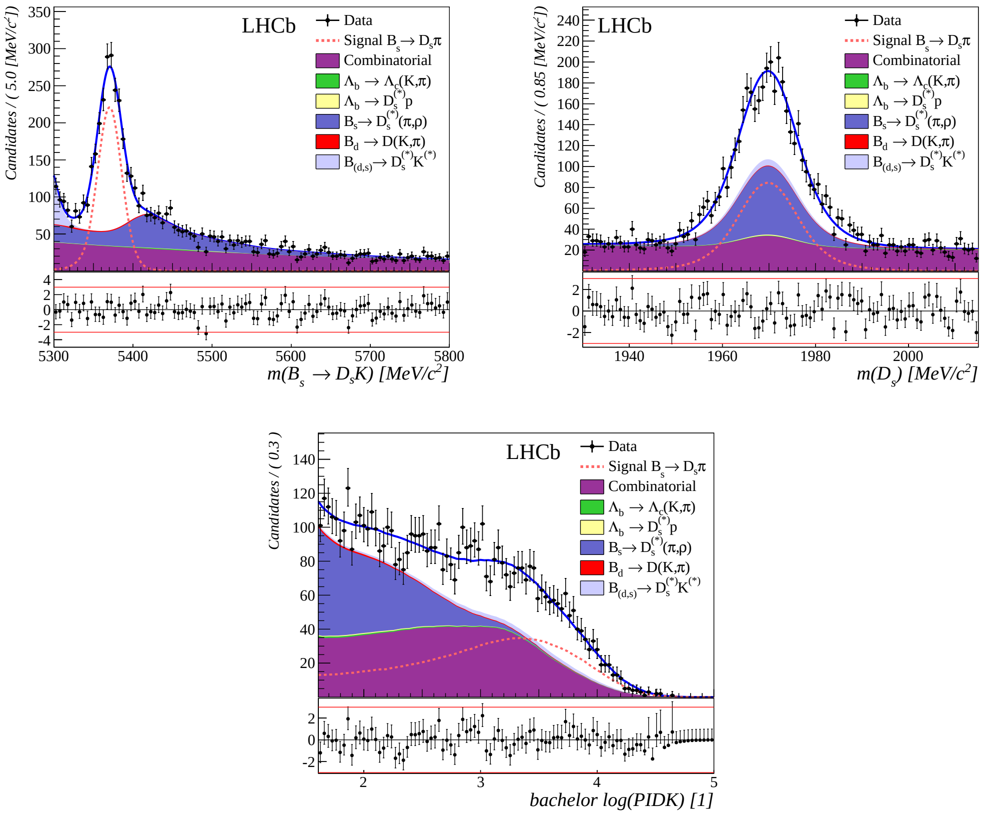


Figure 59.: The simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates for both polarities and all D_s final states combined, from top left to bottom: B_s^0 mass, D_s mass, log(PIDK).

6.4 FLAVOUR TAGGING CALIBRATION

The CP violating coefficients can be measured in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays and the sensitivity of this measurement depends on the dilution D due to the mistag probability ω,

$$D = 1 - 2\omega \; . \tag{6.3}$$

The statistical precision is proportional to the inverse square root of the effective tagging efficiency ε_{eff} as shown by Eq.4.6. The analysis uses the OS and the NNetSSK taggers whose development and optimization have been described in Chapter 4.

It must be clarified here that the 1 fb^{-1} data sample used for this analysis has been obtained with a different stripping version (Stripping17 instead of Stripping20) and a different reconstruction algorithm (Reco12 instead of Reco14) and a different tuning of the FT algorithms with respect to what described in Chapter 5. For this reason a dedicated study of the FT performances and calibration for the 1 fb^{-1} $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis has been performed.

To extract the calibration parameters in the $B_s^0 \to D_s^- \pi^+$ decay the *sFit* approach has been used where the corresponding *sWeights* are determined by the Multi-dimensional fit described in Sec.7.4.3 and the Δm_s parameter is fixed to the LHCb measured value of 17.768 ps $^{-1}$.

The already mentioned two calibration approaches, unbinned and binned, have been considered to extract the flavour tagging calibration parameters. The binned approach split the the data sample in 7 categories of η containing comparable statistics of signal events for both taggers. A simultaneous fit to the time distribution of the different samples determines the average mistag ω_{cat} in each category.

6.4.1 *Calibration of OS tagger*

In Fig. 60 the measured mistag in the $B_s^0 \to D_s^- \pi^+$ data sample is plotted against the mean predicted mistag for each category. The data points show a clearly linear dependency, and nicely fit the function of Eq. 4.11. The results of the tagging calibration parameters using both the approaches are compatible, as shown in Tab. 41.

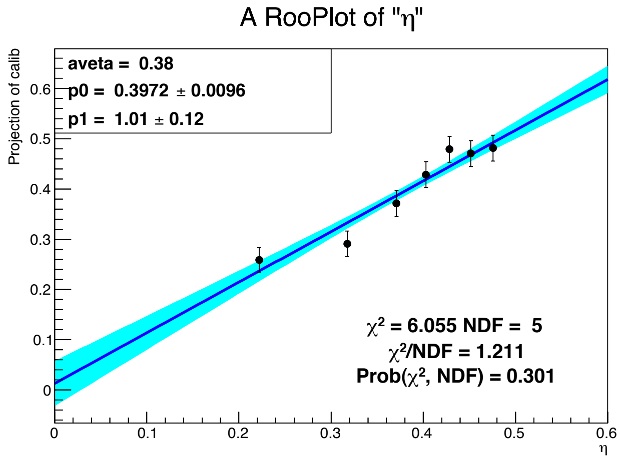


Figure 60.: The measured OS mistag from the time fit in $B_s^0 \to D_s^- \pi^+$ control mode against the average predicted mistag for each category. The error bars represent only the statistical uncertainties. The blue curve is the linear fit to the 7 data points, the light blue area defines the 68% confidence level region of the calibration function (statistical only).

Table 41.: Calibration parameters of the combined OS tagger extracted from $B_s^0 \to D_s^- \pi^+$ decays. For a perfectly calibrated tagger one expects $p_1 = 1$ and $p_0 - \langle\eta\rangle = 0$.

	p_0	p_1	$\langle\eta\rangle$	ϵ_{tag} [%]	ϵ_{eff} [%]
unbinned	0.3972 ± 0.0096	0.991 ± 0.110	0.3813	38.7 ± 0.3	2.70
binned	0.3972 ± 0.0097	1.009 ± 0.116	0.3813		

The average of the two values is taken as a reference for $B_s^0 \to D_s^- \pi^+$ channel and quoted in Tab. 42. Indeed the calibration procedure for the OS tagger using Reco12 2011 data has been done on several control channels: $B^+ \to J/\psi K^+$, $B^+ \to D^0 \pi^+$, $B_d \to D^{*-} \mu^+ \nu_\mu$, $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to D_s^- \pi^+$. The possibility to use several control channels is helpful to evaluate the calibration but also to check how similar are the calibrations from different decay modes (portability) and to estimate systematics effects. The corresponding values of the calibration parameters are summarized in Tab. 42. For each control channel the calibration

Table 42.: Calibration parameters of the OS tagger extracted from different control channels. Statistic and systematic uncertainties are quoted.

control channel	p_0	p_1	$\langle\eta\rangle$
$B^+ \to J/\psi K^+$	$0.3927 \pm 0.0014 \pm 0.0015$	$0.9818 \pm 0.017 \pm 0.005$	0.3919
$B^+ \to D^0\pi^+$	$0.3854 \pm 0.0016 \pm 0.0015$	$0.9722 \pm 0.017 \pm 0.005$	0.3836
$B^0 \to J/\psi K^{*0}$	$0.399 \pm 0.003 \pm 0.0060$	$0.882 \pm 0.043 \pm 0.039$	0.390
$B^0 \to D^{*-}\mu^+\nu_\mu$	$0.3953 \pm 0.0019 \pm 0.0069$	$0.946 \pm 0.019 \pm 0.061$	0.3872
$B_s^0 \to D_s^-\pi^+$	$0.3972 \pm 0.0097 \pm 0.0071$	$1.000 \pm 0.116 \pm 0.047$	0.3813

parameters are reported with the statistical and systematic uncertainty connected with that specific channel. As reference values for the OS tagger has been considered the weighted average of $p_0 - \langle\eta\rangle$ and p_1 on the five control channels, since the measurements, even if quite different, are compatible given the current statistics. This strategy is aimed to define a coherent picture providing a universal calibration which at the moment is reasonably assumed with the current precision. The final average values employed in the $B_s^0 \to D_s^{\mp}K^{\pm}$ analysis are quoted in Tab. 43 with a statistical error in which are included the systematics related to each control channel and a systematic error which accounts for phase space differences among different channels. As average mistag $< \eta >$ is taken the one of the $B_s^0 \to D_s^-\pi^+$

Table 43.: The reference values for the OS tagger calibration are quoted, including the systematic uncertainties that account for differences measured in different channels.

p_0	p_1	$\langle\eta\rangle$	ϵ_{eff} [%]
$0.3834 \pm 0.0014 \pm 0.0040$	$0.972 \pm 0.012 \pm 0.035$	0.3813	$3.13 \pm 0.06 \pm 0.16$

decay mode since it is the more similar channel to the $B_s^0 \to D_s^{\mp}K^{\pm}$ one.

6.4.2 *Calibration of NNetSSK tagger*

With the 2011 statistics the NNetSSK tagger can be calibrated only in $B_s^0 \to D_s^- \pi^+$ channel. In this analysis the new NNetSSK tagging algorithm described in Sec.4.2.1 has been used.

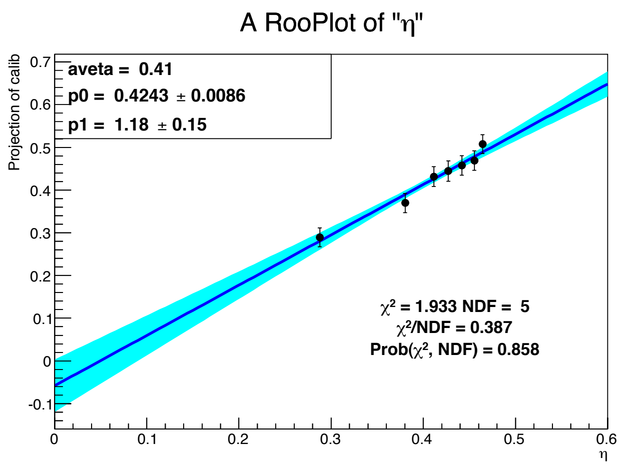


Figure 61.: The measured NNetSSK mistag from the time fit against the average predicted mistag for each category in the $B_s^0 \to D_s^- \pi^+$ decay mode. The blue curve is the linear fit to the 7 data points.

In Tab. 44 the calibration parameters and the measured performance obtained from the sFit to the decay time distribution of $B_s^0 \to D_s^- \pi^+$ decays are reported.

Table 44.: Calibration parameters of the NNetSSK tagger extracted from $B_s^0 \to D_s^- \pi^+$ decays.

tagger		p_0	p_1	$\langle \eta \rangle$	ϵ_{tag} [%]	ϵ_{eff} [%]
NNetSSK	unbinned	0.4244 ± 0.0086	1.255 ± 0.140	0.4097	47.7 ± 0.3	2.17
NNetSSK	binned	0.4244 ± 0.0086	1.18 ± 0.15	0.4097	-	-

For reference, also the performance of the former SSK tagger (referred as *cut based*) have been evaluated and found much lower (ϵ_{tag} = 15.15 % and ϵ_{eff} = 1.37 %) than the ones of the NNetSSK. Fig. 61 shows the dependency of the measured mistag on the average predicted mistag obtained by the fit in category approach for the new SSK tagger.

The improvements in the SSK algorithm based on Neural Network classifiers are visible in the enhancement of the tagging performances with respect to the cut based SSK tagging algorithm. This makes straightforward the decision to use the new tagger in the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis.

Similarly to what discussed in Sec.5.2.2, systematics uncertainties on the tagging calibration parameters for the NNetSSK have been evaluated by using different time resolution model (single gaussian based on the event-by-event time error and fixed triple gaussian of equivalent mean resolution), by varying the time resolution (by $\pm 10\%$), by using different fit models and by varying the fraction of the $B_s^0 \to D_s^{\mp} K^{\pm}$ background (of $+100\%$ or -50% the fraction). The reference parameters for the analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ are the average values between the unbinned fit and the fit in category, as they are quoted in Tab. 45. The correlation between p_0 and p_1 is small, less than 0.03.

Table 45.: Calibration parameters of the NNetSSK tagger, based on the Neural Network classifier, extracted from $B_s^0 \to D_s^- \pi^+$ data. The first uncertainty is statistic and the second one is systematic.

p_0	p_1	$\langle\eta\rangle$	$p0 - \langle\eta\rangle$	ϵ_{eff} [%]
$0.4244 \pm 0.0086 \pm 0.0071$	$1.218 \pm 0.150 \pm 0.104$	0.4097	0.0147	$2.11 \pm 0.35 \pm 0.27$

6.4.3 *Comparison of Tagging Performance in $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$*

In order to use $B_s^0 \to D_s^- \pi^+$ as a tagging calibration channel for $B_s^0 \to D_s^{\mp} K^{\pm}$, it is mandatory to confirm that the taggers perform identically on the two channels, or at least they have the same calibration. The distributions of η for MC signal candidates corresponding to the two channels is compared in Fig. 62 and are in good agreement (quantitatively the Kolmogorov Test evaluates a compatibility probability of 0.994 for the OS and 0.44 for the NNetSSK).

While the fact that the distributions of the calibrated mistag probability η agree between the two channels is promising, this element alone is not sufficient to demonstrate that the calibration parameters will be portable between the two chan-

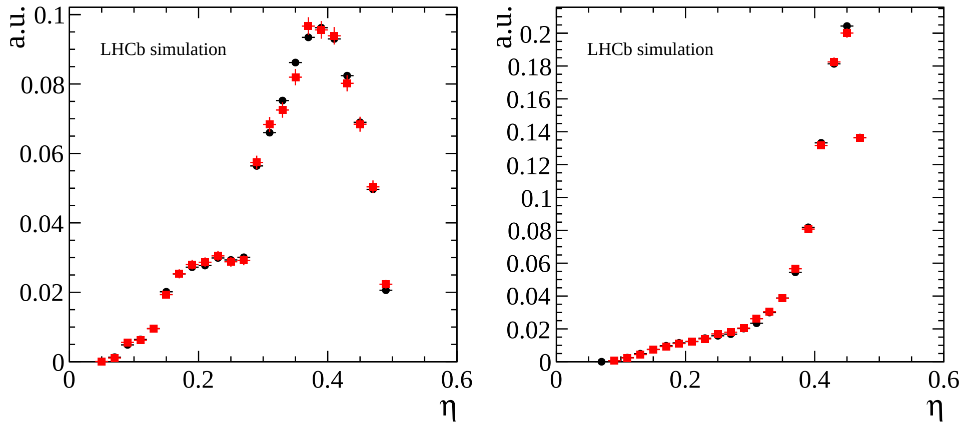


Figure 62.: η distributions for the OS combination (left) and the SS kaon tagger (right) in simulated data. Black is $B_s^0 \to D_s^{\mp}K^{\pm}$ and red is $B_s^0 \to D_s^-\pi^+$
.

nels. We rely on calibration results on simulated events between the two decay modes to justify this assumption.

The distribution of ω in bins of η for selected $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ events in simulation is shown in Fig. 63. It can be seen that the agreement between the two channels is very good, and also that the dependence of ω on η appears to be linear. This dependence is then fitted using Eq. 4.11. The results of the fit using Eq. 4.11 are overlaied in Fig. 63. The fitted values of p_0 are in very good agreement between the two channels, and are close to the ideal value of $\langle\eta\rangle$. The fitted values of p_1 agree with each other to better than one standard deviation, and are close to the ideal value of 1. These results motivate our assumption that the calibration of $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ is the same.

Once the tagger is shown to be consistent between the calibration channel and analysis channel in simulation, the same check was performed in data. Taking the $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ samples and weighting each event with its signal *sWeight* we plot the predicted mistag distributions as before in Fig. 64.

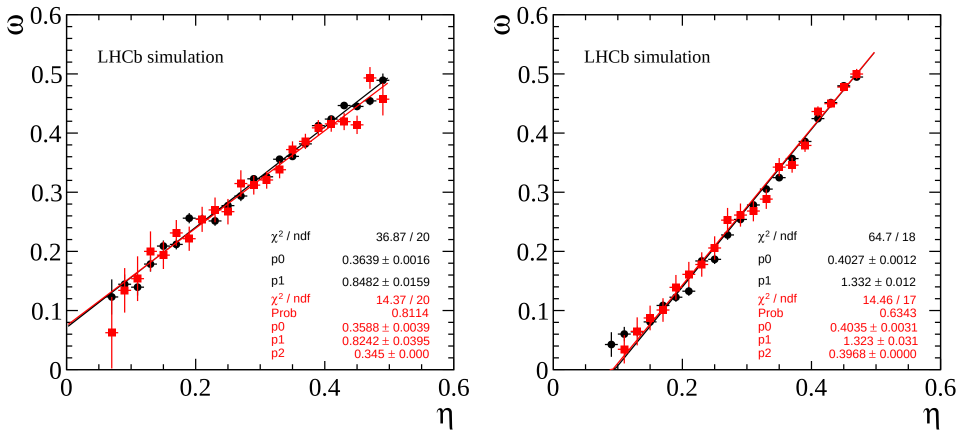


Figure 63.: Dependency of the observed mistag ω on calibrated mistag η, for selected $B_s^0 \to D_s^- \pi^+$ (black) and $B_s^0 \to D_s^{\mp} K^{\pm}$ (red) events in simulation. Left is OS tagging, right is NNetSSK.

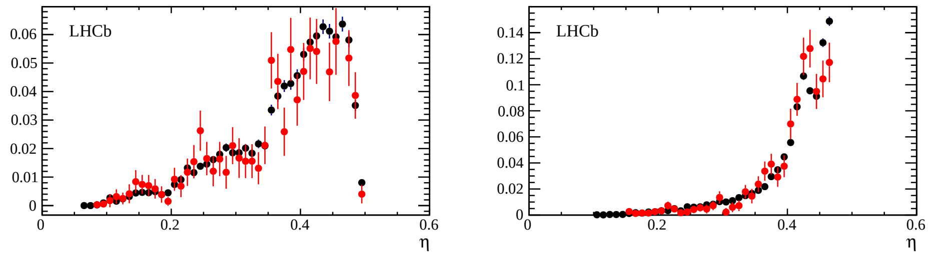


Figure 64.: η distributions for the OS (left) and the NNetSSK (right) taggers. Red is $B_s^0 \to D_s^{\mp} K^{\pm}$ data and black is $B_s^0 \to D_s^- \pi^+$ data. The Kolmogorov test gives 0.94 for the OS and 0.05 for the NNetSSK taggers.

6.4.4 *Tagging asymmetries*

In addition to the calibrations discussed in the previous sections, each tagger also has been calibrated on samples split by initial charge, to address a possible tagging asymmetry. Each tagger has been calibrated applying the following relations that allow for different calibrations on samples split by the initial flavour:

$$\omega = p_0 + \frac{\Delta p_0}{2} + (p_1 + \frac{\Delta p_1}{2}) \cdot (\eta - \langle \eta \rangle) for B_s^0 \tag{6.4}$$

$$\bar{\omega} = p_0 - \frac{\Delta p_0}{2} + (p_1 - \frac{\Delta p_1}{2}) \cdot (\eta - \langle \eta \rangle) for \overline{B}_s^0. \tag{6.5}$$

Here, p_0 and p_1 are the calibration constants from the previous sections, and Δp_0 and Δp_1 parameterize the differences. For the OS tagger the initial flavour asymmetry is easily evaluated through the charge of the B meson in the decay $B^+ \to J/\psi K^+$. For the NNetSSK tagger in principle is not straightforward to evaluate p_0 and p_1 separately for B_s^0 and $\bar{B}_s^0$ due to oscillations. Since the initial flavour asymmetry describes the effect on the tagging parameters of the different interactions between matter and anti-matter it can be evaluated between K^+ and K^- tagging particles. Similarly to what is described in Sec.5.2.1 ,this evaluation has been done on simulated events and on real data using prompt $D_s^{\pm}$. In Tab. 46 the inital flavour asymmetry for p_0, p_1 and ϵ_{tag} are reported.

Table 46.: Initial flavour asymmetry for OS tagging extracted from $B^+ \to J/\psi K^+$ decays and for NNetSSK tagging extracted from prompt $D_s^{\pm}$ reweighting the p_T distribution to $B_s^0 \to D_s^- \pi^+$.

tagger	Δp_0	Δp_1	$\Delta\epsilon_{tag}$ [%]
OS	0.0124 $\pm$ 0.0021	0.095 $\pm$ 0.024	-0.197 $\pm$ 0.126
NNetSSK	-0.020 $\pm$ 0.004	-0.01 $\pm$ 0.03	0.022 $\pm$ 0.004

6.4.5 *OS and NNetSSK combination*

The use of both the NNetSSK and the OS tagger information, provides an increase in the tagging performance for the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis: the effective available statistics is more than doubled. The NNetSSK and OS taggers are virtually independent as they rely on different physics processes. Indeed no correlation is visible between the predicted mistag for events that are tagged by both the taggers, as demonstrated in Fig. 65. The combination of the tagging decisions and the associated mistag prediction is trivial for events where only one of the two taggers fired. To gain further in tagging power, for the events that have decisions from both taggers the combination of tagging responses is done following the standard formula reported in Sec.4.2.3 where it is crucial to note that the combination assumes that the taggers are calibrated and uncorrelated. The data are split in three

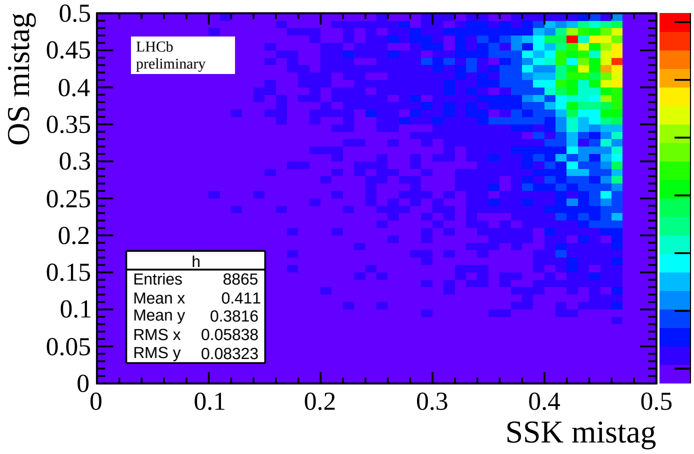


Figure 65.: Scatter plot of OS mistag (on y-axis) and NNetSSK mistag (on x-axis).

different samples depending on the tagging decision: OS only tagged, NNetSSK only tagged and OS and NNetSSK tagged events. The tagging performances in the three sub-samples are reported in Tab. 47 for $B_s^0 \to D_s^- \pi^+$.

Table 47.: The Flavour Tagging performances for only OS tagged, only SSK tagged and both OS and SSK tagged events for $B_s^0 \to D_s^- \pi^+$.

$B_s^0 \to D_s^- \pi^+$	ϵ_{tag} [%]	ϵeff [%]
OS only	19.80 ± 0.23	$1.61 \pm 0.03 \pm 0.08$
SSK only	28.85 ± 0.27	$1.31 \pm 0.22 \pm 0.17$
both OS-SSK	18.88 ± 0.23	$2.15 \pm 0.05 \pm 0.09$
total	67.53	5.07
total (previous result [76])	40	1.9

The total tagging power is about 5.1%, which represents an important enhancement with respect to the 1.9% of the previous analysis [60], performed on the same data sample but using a previous version of the OS tagger only and a different strategy for selecting and fitting data.

The mistag templates for only the OS, only the NNetSSK and both taggers obtained from *sWeighted* $B_s^0 \to D_s^- \pi^+$ data are shown in Fig. 66.

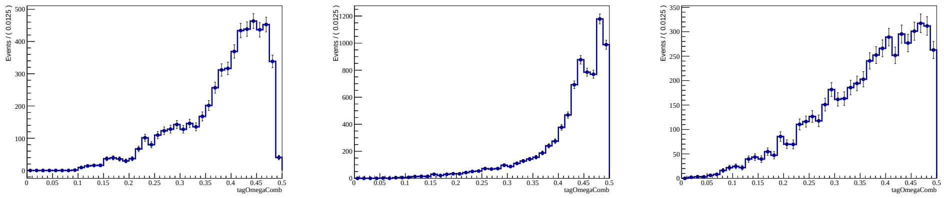


Figure 66.: The mistag templates obtained from *sWeighted* $B_s^0 \to D_s^- \pi^+$ data. From left to right: only the OS, only the NNetSSK, both taggers.

The uncertainties on the calibration of the OS and NNetSSK taggers are then propagated to the combination of both taggers (from Tables 43, 45 and 46). This is most significant for the $B_s^0 \to D_s^{\mp} K^{\pm}$ fits due to their sensitivity to these uncertainties (because C_f , S_f and $\overline{S}_f$ could be directly influenced by any slight miscalibration). It also allows to treat the small asymmetries in tagging calibration after the combination of taggers by "recalibrating" $\eta_{comb.}$ depending on true initial B flavour and type of tagged event (OS only, NNetSSK only, both). Propagating the uncertainties of the OS and NNetSSK tagger calibrations to the result of the combination gives rise to the following fluctuations (see Tab.48) about the point of perfectly calibrated taggers including initial flavour asymmetries for the three categories of tagged events.

Table 48.: Propagation of tagging calibrations uncertainties after combination of taggers.

	$\langle \eta_{comb.} \rangle$	p_0	p_1
B OS only	0.37	0.377 ± 0.004	1.05 ± 0.04
B NNetSSK only	0.415	0.405 ± 0.007	1.00 ± 0.15
B both	0.339	0.338 ± 0.006	1.03 ± 0.04
$\overline{B}$ OS only	0.371	0.366 ± 0.004	0.950 ± 0.040
$\overline{B}$ NNetSSK only	0.415	0.425 ± 0.011	1.00 ± 0.15
$\overline{B}$ both	0.339	0.339 ± 0.006	0.971 ± 0.040

Similarly, it is possible to propagate the tagging efficiencies $\epsilon_{OS}, \epsilon_{SSK}$ and their variations with initial B flavour $\Delta\epsilon_{OS}, \Delta\epsilon_{SSK}$ through the combination (assuming uncorrelated taggers). This will yield the tagging efficiencies and tagging efficiency asymmetries for the categories of tagged events reported in Tab.49.

Table 49.: Tagging efficiencies and tagging efficiency asymmetries a^{tageff} for OS only, NNetSSK only and both OS and NNetSSK tagged events.

	ϵ_{tag} [%]	$\Delta\epsilon_{tag}$ [%]
OS only	0.202 ± 0.002	-0.0028 ± 0.0016
NNetSSK only	0.292 ± 0.002	0.0018 ± 0.0010
both	0.185 ± 0.002	-0.0023 ± 0.0016

6.5 TIME RESOLUTION AND TIME ACCEPTANCE

The decay time is reconstructed with a limited resolution, the main reason being the limited spatial resolution of the B decay vertex (referred as Secondary Vertex). The time resolution affects the capability to properly solve the observable oscillation in the B_s^0 decays because it introduces a dilution term in the decay time amplitude, in addition to the already mentioned dilution factor due to flavour tagging:

$$A^{obs}(t) \approx (1 - 2\omega) e^{-\frac{(\Delta m_s \sigma_t)^2}{2}} A^{th}(t). \tag{6.6}$$

To get unbiased results and optimize the sensitivity to the CP observables contained in the $A^{th}(t)$ term of previous equation it is important to have a correct description of the decay time resolution model. In addition also the determination of the flavour tagging calibration parameters discussed in previous section (6.4) is strictly correlated to the dilution related to the time resolution. In this case, as in many other LHCb analyses, it has been used a time resolution function which has a different width for each event. This is possible because the reconstruction algorithm provides, in addition to the decay time measurement, also an estimation of the decay time error δ_t based on the uncertainity on the SV and B momentum. Such estimation needs to be calibrated in order to match the actual time resolution. The analysis for Δm_s measurement [59] performed a calibration study between the estimated δ_t and the time resolution seen in data. This has been done using prompt D_s associated to a random π tracks which form a fake B_s^0 candidate produced at time equal 0. Using this data sample a scale factor between the per-event time error δ_t, that was estimated by the lifetime fit, and the true time resolution σ_t was measured. The result for the scale factor SF is summarized in Tab. 50. While this is an old result, we feel that the size of the systematic uncertainty (in line with what other time dependent analyses, e. g.$B \to h^+h^-$ took), will cover any potential issues with it. Considering that the time resolution is used on a per-event basis it

Table 50.: Results of the time resolution study of the Δm_s analyses [59] and [61]. The effective mean time resolution, the scale factor with the relative uncertainties and the average dilution factor D_t are quoted.

sample	time resolution	SF	syst. SF	D_t
fake B_s^0	44 fs	1.37	$[1.25, 1.45]$	0.737

becomes a variables of the fit which needs to be described by the appropriate PDF. The dilution factor corresponding to the effective time resolution is evaluated to be of 0.737, larger than the dilution due to imperfect tagging (from values in Tab.47 $\sqrt{\frac{\epsilon_{eff}}{\epsilon_{tag}}} = 0.274$).

The selection requirements may introduce some acceptance effects in the decay time distribution which needs to be accounted by in the PDF. The strategy is to measure the decay time acceptance function fitting the decay time distribution of the values of $B_s^0 \to D_s^- \pi^+$ decay fixing Γ_s and $\Delta\Gamma_s$ and floating the acceptance parameters. The $B_s^0 \to D_s^\mp K^\pm$ time acceptance is then determined by correcting the $B_s^0 \to D_s^- \pi^+$ acceptance by the ratio of $B_s^0 \to D_s^\mp K^\pm$ $/B_s^0 \to D_s^- \pi^+$ acceptances determined in MC samples.

The acceptance function is described by splines, which are polynomial functions which can be implemented in an analytic way in the decay time fit following the method described in [77]. These splines have six free parameters. The spline boundaries ("knots") were chosen in an ad- hoc manner to produce a smooth acceptance shape, and placed at $0.5, 1.0, 1.5, 2.0, 3.0, 12.0$ ps for the nominal fit configuration. Essentially there are more knots at smaller lifetimes where the function varies more rapidly, and fewer at high lifetimes where we expect a flat behaeviour with a possible slow drop due to an upper acceptance effect from the VELO reconstruction. We studied the effect of the choice of knot positions by doubling the number of knots and repeating the fits, but this results in negligible changes.

We assume the same acceptance function for all the 5 D_s decay modes, while determining the ratio on MC, we rescale the samples according to the statistics observed in data. The fits to the simulated events are shown in Fig. 67 and Tab. 51.

The ratio of two splines is given simply by the ratio of its coefficients, which is also given in Tab. 51. The correlation matrices are considered for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ fit and for ratios.

We observe the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ acceptances to be significantly different. Although all ratios are within 2σ of unity, the proper computation of the p-value for the compatibility hypothesis gives 3.8×10^{-5} (4.1σ). A systematic is assigned due to our imperfect knowledge of the time acceptance.

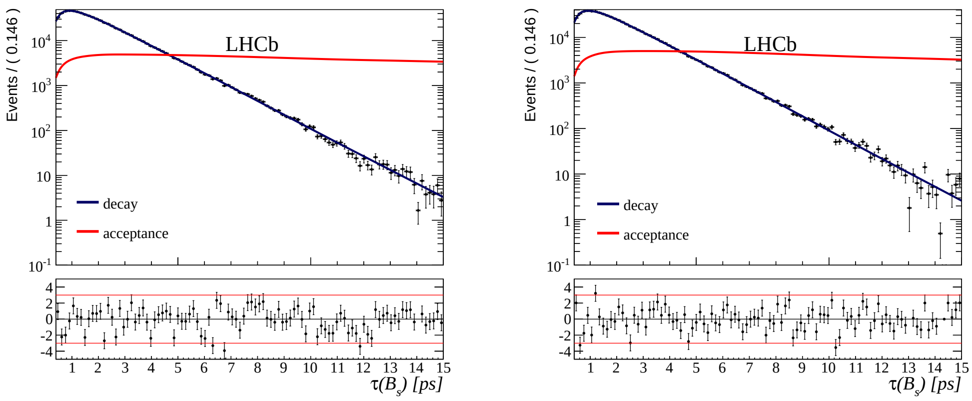


Figure 67.: MC Acceptances for $B_s^0 \to D_s^- \pi^+$ (left) and $B_s^0 \to D_s^{\mp} K^{\pm}$ (right). The acceptance function is shown in red, while the total fit curve is blue.

Parameters	Fit to $B_s^0 \to D_s^- \pi^+$	Fit to $B_s^0 \to D_s^\mp K^\pm$	$B_s^0 \to D_s^\mp K^\pm$ / $B_s^0 \to D_s^- \pi^+$
v_1	0.512 ± 0.008	0.498 ± 0.008	0.973 ± 0.022
v_2	0.745 ± 0.012	0.742 ± 0.013	0.996 ± 0.024
v_3	0.996 ± 0.015	0.981 ± 0.017	0.985 ± 0.023
v_4	1.131 ± 0.018	1.163 ± 0.020	1.028 ± 0.024
v_5	1.231 ± 0.017	1.242 ± 0.019	1.009 ± 0.021
v_6	1.227 ± 0.027	1.285 ± 0.030	1.047 ± 0.034

Table 51.: Parameters of the time acceptance from a fit to the time distribution of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ MC events and their ratio.

6.6 TIME FIT

As mentioned at the beginning of this chapter, two fit approaches have been used to measure the CP observables:

- sFit: decay time fit of the events weighted by the Multi-dimensional fitter *sWeights*, in which a statistical backgrounds subtraction allow to manage only the signal component

- cFit: standard fit of the events in which all backgrounds component needs to be described for each observable of the fit. The results of the Multi-dimensional fitter, i.e. yields and shapes parameters, are fixed in the cFit.

The observables involved in the decay time fit are the B_s^0 decay time and its error estimation δ_t, the tagging decision d, the predicted mistag η and the charge of the bachelor particle q which is used to identify the final state. The cFit in addition has the three observables of the Multi-dimensional fitter: B_s^0 mass, D_s mass and $PIDK$ for the bachelor particle.

The tagging calibration parameters and uncertainties of the OS and NNetSSK taggers can be found in Section 6.4.5. The tagging calibration parameters in the $B_s^0 \to D_s^\mp K^\pm$ fits on data are thus constrained in the fit to the central values and uncertainties define the Gaussian constraints. We fix the values of the following parameters to the central values:

$$\Gamma_s = (0.661 \pm 0.007) ps^{-1} , \tag{6.7}$$

$$\Delta\Gamma_s = (0.106 \pm 0.013) ps^{-1} , \tag{6.8}$$

$$\rho(\Gamma_s, \Delta\Gamma_s) = -0.39, \tag{6.9}$$

they are taken from current LHCb measurements [72]. In addition, the $B_s^0 \to D_s^{\mp} K^{\pm}$ fit is executed fixing Δm_s to the LHCb measurement value $\Delta m_s = 17.768 \pm 0.024$ ps^{-1} [64] and the acceptance parameters to the values found in the $B_s^0 \to D_s^- \pi^+$ fit multiplied by the correction given in Tab. 51. We assume no production asymmetry and constrain the signal detection asymmetry to $1\% \pm 0.5\%$ (in our sign convention this means a 1% higher detection asymmetry for positive kaons). The decay time fit (both sFit and cFit) described so far, is also performed to the $B_s^0 \to D_s^- \pi^+$ sample to measure the acceptance parameters in data (the one reported in Sec.6.5) where the mixing frequency Δm_s is left free as a cross-check of the fit procedure.

6.7 RESULTS OF THE SFIT

The *sFit* approach is based on the sPlot technique where the sWeights are extracted from the Multi-dimensional fit on the five D_s final states.

6.7.1 *sFit in* $B_s^0 \to D_s^- \pi^+$

For the $B_s^0 \to D_s^- \pi^+$ time fit, the PDF is obtained convoluting the B_s^0 decay rate (described in Sec.2.3.2) with a single Gaussian time resolution and multiplied by the acceptance function (Sec. 6.5):

$$PDF = \frac{d\Gamma_{B_s \to D_s \pi}(t, d, q)}{dt} \otimes G(0, \delta_t) \times a(t) \ . \tag{6.10}$$

where d and q represent the tagging decision and the charge of the bachelor particle.

The free parameters in the fit are Δm_s and the six parameters describing the time acceptance function. The distributions of δ_t and η are taken from the sweighted data. The data distribution superimposed with the time sFit result is presented in Fig. 68. The fitted values are listed in Tab. 52 and the correlation matrix is considered. The fitted Δm_s agrees extremely well with the existing LHCb measurement $\Delta m_s = (17.768 \pm 0.024) ps^{-1}$ [64].

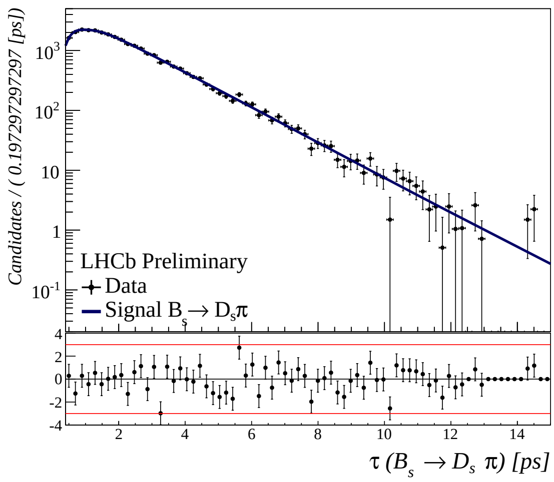


Figure 68.: Time distribution of the $B_s^0 \to D_s^- \pi^+$ data sample. Bottom plot shows pulls.

Table 52.: Fitted values to the $B_s^0 \to D_s^- \pi^+$ time distribution for the *sFit* .

Parameter	Fitted value
Δm_s	17.772 ± 0.0215 ps^{-1}
acceptance function	
v_1	0.459 ± 0.032
v_2	0.690 ± 0.052
v_3	0.885 ± 0.065
v_4	1.130 ± 0.081
v_5	1.223 ± 0.076
v_6	1.228 ± 0.121

6.7.2 *sFit in* $B_s^0 \to D_s^{\mp} K^{\pm}$

For the $B_s^0 \to D_s^{\mp} K^{\pm}$ fit the decay amplitude depends on 5 additional CP parameters (see Chapter 2), which are limited to the fit range $[-4, 4]$, corresponding to about 6σ based on the expected uncertainties seen in the Toy-MC studies. The fitted values for the CP observables are listed in Tab. 53 and in Tab. 54 the correlations are reported. The B_s^0 decay time distribution with the superimposed fit result are plotted in Fig. 69. The other parameters of the time sFit are constrained to the values reported in Tab.48 and in Tab.49.

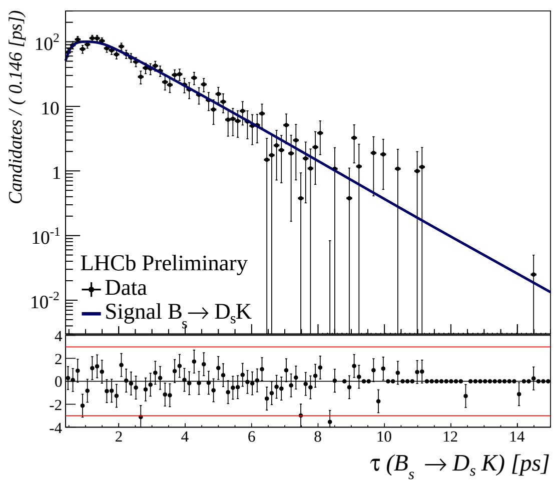
Table 53.: Fitted values of the CP observables to the $B_s^0 \to D_s^{\mp} K^{\pm}$ time distribution for *sFit* . All parameters different than CP observables are constrained in the fit.

Parameter	Fitted value
C_f	0.52 ± 0.25
S_f	-0.90 ± 0.31
$\overline{S}_f$	0.36 ± 0.34
D_f	-0.29 ± 0.42
$\overline{D}_f$	-0.14 ± 0.41

As is discussed in detail in Sec. 6.9, most systematics are obtained from large scale toy studies. Nevertheless the stability of the data fit when varying the fixed parameters (whether in the mass or time fits), when smearing the acceptance function correction, or when varying the decay time resolution or reconstruction bias is checked.

Table 54.: Correlation matrix of the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP sFit* . Here we have restricted ourselves to correlations between the *CP* parameters for readability, as the tagging efficiencies all have negligible correlations with the *CP* parameters anyway. As can be seen the only large correlation is the 50% between D_f and $\overline{D}_f$, which is understandable because they are both sensitive to the slowly varying hyperbolic terms induced by $\Delta\Gamma_s$.

Parameter	C_f	D_f	$\overline{D}_f$	S_f	$\overline{S}_f$
C_f	1.000	−0.071	−0.097	0.117	0.042
D_f	−0.071	1.000	0.500	0.044	−0.003
$\overline{D}_f$	−0.097	0.500	1.000	0.013	−0.005
S_f	0.117	0.044	0.013	1.000	−0.007
$\overline{S}_f$	0.042	−0.003	−0.005	−0.007	1.000



.

Figure 69.: Time *sFit* to the $B_s^0 \to D_s^{\mp} K^{\pm}$ data sample. Bottom plot shows pulls.

6.8 RESULTS OF THE CFIT

The cFit is a full fit to all the observables of the Multi-dimensional fitter (m(B_s^0), m(D_s) and $PIDK$) and of the time sFit. The main difference with the Multi-dimensional fitter described in Sec.6.3 is that the cFit uses a restricted B_s^0 mass range ([5320,5420]MeV/c^2). The Multi-dimensional fit result provides:

- all the yields for signal and backgrounds decays, scaled in a narrower mass range [5320,5420]MeV/c^2
- the observables templates needed by the cFit for the B_s^0 , D_s masses and $PIDK$

The main difference with respect to the time sFit is that the data sample contains signal and background events which need to be accounted for in the time-dependent part of the PDF for all the different background components. The following template formula fits for signal and backgrounds decays:

$$\begin{aligned} PDF(m_{B_s},m_{D_s},log(PIDK),t,\omega,\sigma_t,d,q) &= a(t)\cdot PDF(m_{B_s},m_{D_s},log(PIDK))\cdot \\ & PDF(\delta_t)\cdot\int k\,dk\,F(k)\cdot \\ & \left(\int dt'\Gamma_{theo}(t',k\Gamma,k\Delta\Gamma,k\Delta_m;\epsilon_{tag},\omega,d,q)\cdot G(t-t',0,\sigma_t)\right) \end{aligned} \tag{6.11}$$

The description of $P(m_{B_s},m_{D_s},log(PIDK))$ is taken straight from the MD fit in Sec. 6.3 where the mass templates and all yields are split by D_s final states, while all other cFit components are common for all D_s decay modes. The theoretical decay rate distribution of $\Gamma_{theo}(t)$ is essentially discussed in Chapter 2, Sec.2.6. $PDF(\delta_t)$ represents the distribution of decay time estimated errors seen in data. From the point of view of the decay time distribution, the several backgrounds component can be grouped as depicted in Tab.55 since they have different behaviors in time and physics properties requiring different treatments. All modes (signal and back-

Table 55.: Categorisation of the different physics backgrounds for the cFit. See text for an explanation of the categories.

$D_s\pi$-like	D_sK-like	non-oscillating
$B_s \to D_s^- \pi^+$	$B_s \to D_s^{\pm} K^{\mp}$	$\Lambda_b \to D_s^- p$
$B_s \to D_s^{*-} \pi^+$	$B_d \to D\pi$	$\Lambda_b \to D_s^{*-} p$
$B_s \to D_s^- \rho^+$		$\Lambda_b \to \Lambda_c^+ K^-$
$B_d \to D_s^- K^+$		$\Lambda_b \to \Lambda_c^+ \pi^-$
$B_d \to D^- K^+$		
$B_d \to D_s^- \pi^+$		

grounds) have the same spline-based acceptance function.

The cFit approach needs to correct for the wrong estimation of the decay time of background decays by means of a multiplicative factor generally called *k-factor*:

$$k = \frac{(m/p)_{true}}{(m/p)_{erec}} \tag{6.12}$$

where "true" refers to the true 4-momentum of the mother particle (B_s^0) and "erec" emulates the mistakes of the event reconstruction (mis-identifying or losing a particle). The k-factors are evaluated simulated samples of backgrounds events.

DECAY TIME DISTRIBUTIONS To determine the decay time distribution of each backgrounds it is useful to underline which is the time evolution of the considered b-hadron:

1. Λ_b^0: non-oscillating backgrounds $\rightarrow$ simple exponential is used setting lifetime to the current world averages Specifically, we constrain to $\Gamma_{\Lambda_b} = 0.700 \pm 0.012\,\mathrm{ps}^{-1}$. Terms which have to do with tagging are dropped for these backgrounds.

2. $B_s \rightarrow D_s\pi$-like: B_s at decay time can only decay to one final state f and a $\overline{B_s}$ can only decay to the charge-conjugated final state $\overline{f}$. The CP parameters are fixed by physics ($C_f = C_{\bar{f}} = 1$, $S_f = S_{\bar{f}} = D_f = D_{\bar{f}} = 0$) and $\Gamma_{theo}(t)$ looks like the one used for the $B_s^0 \rightarrow D_s^- \pi^+$ decay.

3. $B_s \rightarrow D_s K$-like: B_s at decay time can decay to either f or its charge-conjugate $\overline{f}$, and a $\overline{B_s}$ at decay time can likewise decay to both f and $\overline{f}$. The CP parameters are assumed to be unknown and are thus floated in the fit with gaussian constraints to enhance fit stability.

Depending on the background, a specific treatment of the flavour tagging parameters is adopted.

CP PARAMETERS For the signal $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decay the set of C_f , D_f , $\overline{D}_f$, S_f , $\overline{S}_f$ parameters is floating and no constraints are applied. For the backgrounds, we float an effective $|\lambda_f|$, weak and strong phase in the fit which is common for all $D_s K$-like modes, and calculate C_f , D_f , $\overline{D}_f$, S_f , $\overline{S}_f$ based on these floating parameters. This approach keeps the CP parameters of the background physical at all times, helping fit stability.

COMBINATORIAL BACKGROUND LIFETIME Since in our analysis the combinatorial background is among the most prominent sources of background, a good description of its lifetime is an important cFit task.

For the combinatorial background, the (apparent) lifetime depends on the kinematics of random tracks and their distribution in phase space since also the efficiency of different taggers depends on kinematical properties of the (apparent) B candidate, combinatorial events can have different lifetimes depending on the tagging category of that event. We thus fit the lifetime of the combinatorial backgrounds separately for untagged, OS only, NNetSSK only and OS+NNetSSK events on data with the full selection applied, using a the high mass sideband of $m_{B_s} >$ 5700MeV/c^2.

6.8.1 $B_s^0 \to D_s^{\mp} K^{\pm}$ *fit results*

In figure 70 the fit to the decay time distribution in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays is shown: on the left plot is reported the *sFit* while on the right plot is reported the *cFit*. In Tab. 56 the results of the decay time fit are reported for both the fitters. The statistical and systematic uncertainties are quoted for each CP observable. The two time fit methods give compatible results and at the moment we are dominated by the low statistics.

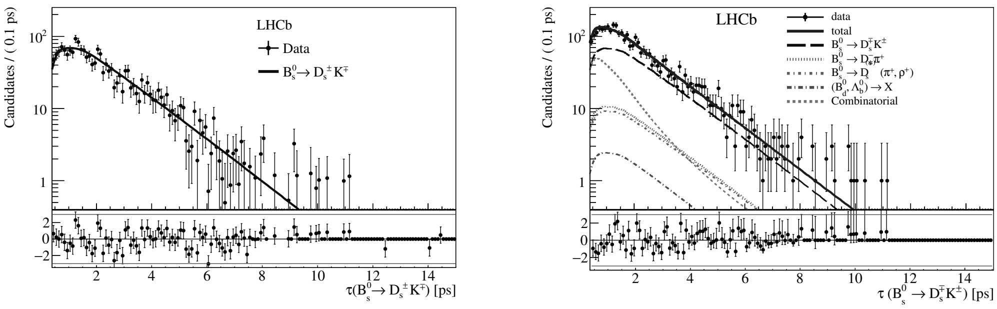


Figure 70.: Result of the decay time (on the left side) *sFit* and (on the right side) *cFit* to the $B_s^0 \to D_s^{\pm} K^{\mp}$ candidates.

Table 56.: Fitted values of the CP observables to the $B_s^0 \to D_s^{\pm} K^{\mp}$ decay time distribution. The statistical and systematic uncertainties are also quoted.

Parameter	*sFit* fitted value	*cFit* fitted value
C_f	0.52 ± 0.25 ± 0.04	0.53 ± 0.25 ± 0.04
$A_f^{\Delta\Gamma}$	0.29 ± 0.42 ± 0.17	0.37 ± 0.42 ± 0.20
$A_{\bar{f}}^{\Delta\Gamma}$	0.14 ± 0.41 ± 0.18	0.20 ± 0.41 ± 0.20
S_f	-0.90 ± 0.31 ± 0.06	-1.09 ± 0.33 ± 0.08
$S_{\bar{f}}$	-0.36 ± 0.34 ± 0.06	-0.36 ± 0.34 ± 0.08

6.9 systematics uncertainties

The main systematic uncertainties are due to the fixed parameters of the fit: Δm_s, Γ_s and $\Delta\Gamma_s$ and from the limited knowledge of the decay time resolution and acceptance. In order to estimate these uncertainties a large set of pseudo-experiments (1000) is used with the statistics seen in data. The pseudo-experiments are generated with the average of the *cFit* and *sFit* fit values and subsequently they are processed by the full data fitting procedure where each parameter (Δm_s, Γ_s, $\Delta\Gamma_s$ and the SF of time resolution) is varied according to its systematic uncertainty. The difference with the nominal fit is considered as systematic uncertainty. Splitting the samples in bins of BDTG reveals a small systematic effect which is taken into account. The uncertainties related to the tagging parameters are already included

Table 57.: Systematic uncertainties related to the *sFit* results expressed as fractions of the statistical uncertainties.

Source	C_f	S_f	$S_{\bar{f}}$	$A_f^{\Delta\Gamma}$	$A_{\bar{f}}^{\Delta\Gamma}$
Δm_s	0.062	0.104	0.100	0.013	0.013
time resolution	0.104	0.092	0.096	0.004	0.004
acceptance Γ_s, $\Delta\Gamma_s$	0.043	0.039	0.038	0.427	0.437
sample splits	0.124	0.072	0.071	0.000	0.000
total	0.179	0.161	0.160	0.427	0.437

Table 58.: Systematic uncertainties related to the *cFit* results expressed as fractions of the statistical uncertainties.

Source	C_f	S_f	$S_{\bar{f}}$	$A_f^{\Delta\Gamma}$	$A_{\bar{f}}^{\Delta\Gamma}$
Δm_s	0.068	0.131	0.126	0.014	0.011
time resolution	0.131	0.101	0.103	0.004	0.004
acceptance Γ_s, $\Delta\Gamma_s$	0.050	0.050	0.043	0.461	0.464
sample splits	0.102	0.156	0.151	0.000	0.000
total	0.187	0.234	0.226	0.466	0.470

in the statistical errors since for these parameters gaussian constraints are used and so are not considered further.

The decay time acceptance mostly affects the CP coefficients related to the hyperbolic terms of the decay rate ($A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$). An increase in the statistics of the $B_s^0 \to D_s^- \pi^+$ data would improve the description of the acceptance function. Further studies have been done splitting the data sample into two subsets according to the two magnet polarities and the hardware trigger decision. In Tables 57 and 58 we report the systematic uncertainties as fractions of the statistical uncertainties respectively for the *sFit* and for the *cFit* results. The *cFit* and *sFit* results are in good agreement and differences between the two fitters are evaluated from the distributions of the per-pseudoexperiment differences of the fitted values. Both fitters return compatible results.

6.10 THE γ ANGLE DETERMINATION

In order to extract the value of the γ angle the method consists in maximizing a likelihood including the physics parameters: γ, ϕ_s (-2β_s), r_{D_sK}, δ and the CP violating observables expressed through Eq.2.46 of Chapter 2. The value of γ as well as that of r_B and δ_B are determined by this likelihood maximization while the value of ϕ_s is fixed to an independent measurement performed by LHCb [72]. The profile likelihood contours in two dimensions are reported in Fig.71 showing the agreement between the *sFit* and *cFit* results.

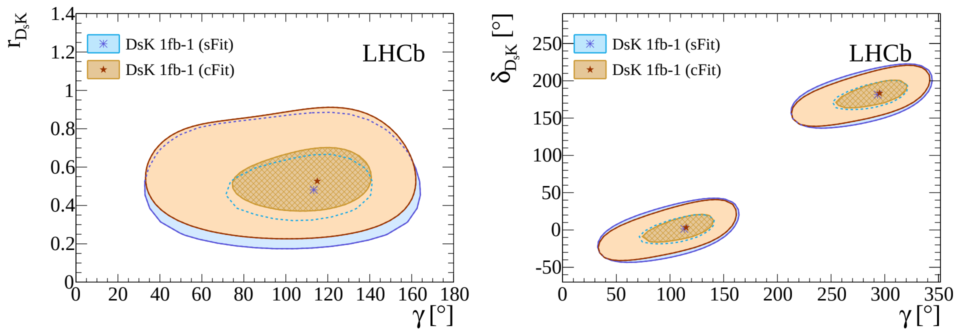


Figure 71.: Profile likelihood contours of γ vs. r_{D_sK}, and γ vs. δ. The contours are the $n\sigma$ profile likelihood contours, where $\Delta\chi^2 = n^2$ with $n = 1, 2$. The markers denote the best-fit values. All errors (stat. and syst.) are included.

The extraction of γ is reported in Fig.72. The *cFit* results have been arbitrarily chosen being compatible with the sFit ones and we quote γ =(115$^{+26}_{-35}$(stat)$^{+8}_{-25}$(syst)$\pm$ 4 (ϕ_s))$^\circ$ at 68% Confidence Level (CL), where the intervals are expressed modulo 180°.

The LHCb experiment has measured the γ angle also in other tree-level processes, as it has been discussed in Chapter 2 using time-integrated analysis. The measurement of γ in the $B_s^0 \to D_s^{\mp}K^{\pm}$ decay has never been done by other collaboration and it represents the first time-dependent measurement done by LHCb to determine γ. The advantage of this decay is that it has a larger interference

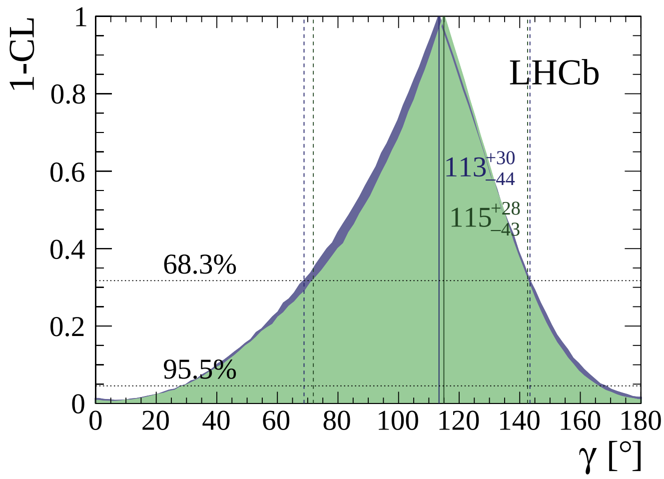


Figure 72.: Graph showing 1-CL for γ using the *cFit* (green) and *sFit* (blue) CP parameters. All errors (stat. and syst.) are included.

factor governed by the value of $r_B \approx 0.5$ with respect to the other tree-level decays already employed by LHCb. The time-dependent analysis allows to solve the fast oscillation of B_s^0 mesons, given also the good precision with which Δm_s is already measured. Measuring also the CP observables related to the hyperbolic terms of the decay rate ($A_f^{\Delta\Gamma}$) reduces the ambiguity on the γ measurement. Another important advantage with respect to time-integrated measurements is that the external parameters needed by this time-dependent analysis are already well established. On the other hand the analysis is limited by the low statistics of this channel and by the fact that the actual sensitivity to the CP observables gets reduced by the tagging performance.

Recently the LHCb collaboration has presented an overall combination of $B \to DK$ tree-level measurements for determining γ [10]. The results are obtained from time-integrated analyses of:

- $B^+ \to DK^+$, with $D \to h^+h^-$, GLW/ADS, $3fb^{-1}$
- $B^+ \to DK^+$, with $D \to h^+\pi^-\pi^+\pi^-$, quasi-GLW/ADS, $3fb^{-1}$
- $B^+ \to DK^+$, with $D \to h^+h^-\pi^0$, quasi-GLW/ADS, $3fb^{-1}$
- $B^+ \to DK^+$, with $D \to K_S^0 h^+h^-$, model independent GGSZ, $3fb^{-1}$

- $B^+ \to DK^+$, with $D \to K^0_S K^+\pi^-$, GLS, $3fb^{-1}$
- $B^0 \to DK^+\pi^-$, with $D \to h^+h^-$, GLW Dalitz, $3fb^{-1}$
- $B^0 \to DK^{*0}$, with $D \to K^+\pi^-$, ADS, $3fb^{-1}$
- $B^0 \to DK^{*0}$, with $D \to K^0_S\pi^+\pi^-$, model dependent GGSZ, $3fb^{-1}$
- $B^+ \to DK^+\pi^+\pi^-$, with $D \to h^+h^-$, GLW/ADS, $3fb^{-1}$

The measurement of γ using $B^0_s \to D^{\mp}_s K^{\pm}$ data sample ($1fb^{-1}$) described in this chapter is included in the combination. The results presented here represents an update which supersedes the previous combination by adding more decay channels and updating the most of measurements to the full Run1 data sets. The input measurements provide sensitivity to γ in the interference between $b \to c$ and $b \to u$ transition, as described with more detail in Chapter 2. The $B^+ \to DK^+$ modes offer the best sensitivity interplaying measurements with multiple narrow solutions for γ (ADS/GLW modes $\delta\gamma \approx 8°$) with other which have single broad solutions (GGSZ modes $\delta\gamma \approx 15°$). The overall combination accounts for $D^0 - \overline{D}^0$ mixing corrections and for CP violation effects in the $D \to h^+h^-$ decays. In Fig.73 1-CL curves for r_B^{DK}, δ_B^{DK}, $r_B^{DK^*}$, $\delta_B^{DK^*}$ and γ are shown with the 1σ and 2σ levels. In Fig.73 the profile likelihood two-dimensional contours for r_B^{DK}, δ_B^{DK}, $r_B^{DK^*}$, $\delta_B^{DK^*}$ and γ are shown.

LHCb has measured the γ angle from tree-level decays to be: $\gamma = (70.9^{+7.1}_{-8.5})°$ at 68% CL modulo 180° and it represents the most precise determination of γ from a single experiment to date. The contribution of the $B^0_s \to D^{\pm}_s K^{\mp}$ measurement to the combination is small for the moment but it represents an additional independent way to measure γ and it will contribute mostly given the enhancement in statistics of the next analysis with $\int \mathcal{L}$ = 3 fb^{-1}.

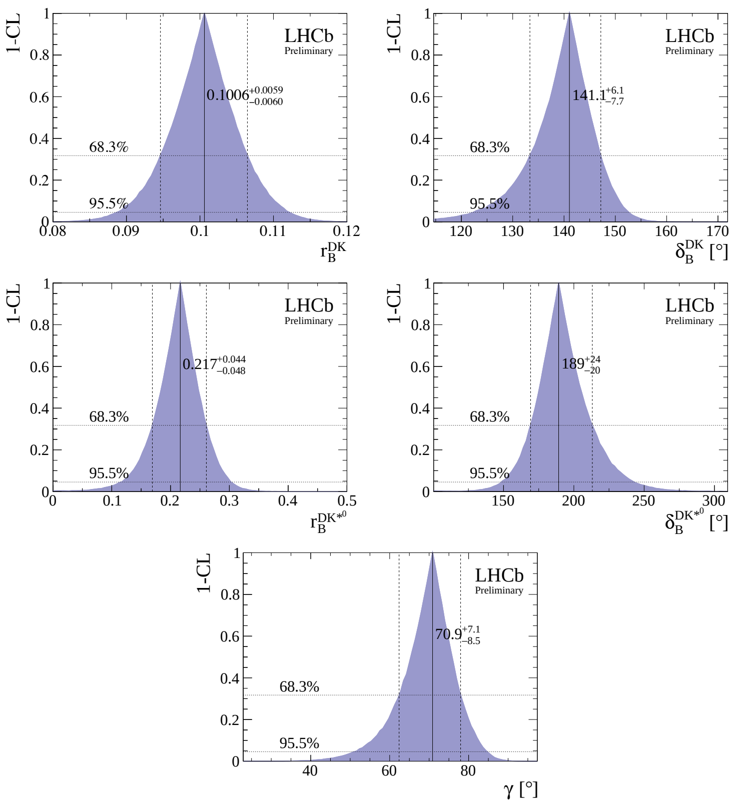


Figure 73.: 1-CL curves for the combination of tree-level measurements with central values (solid vertical lines) and 1σ uncertainties (dashed vertical lines) labelled.

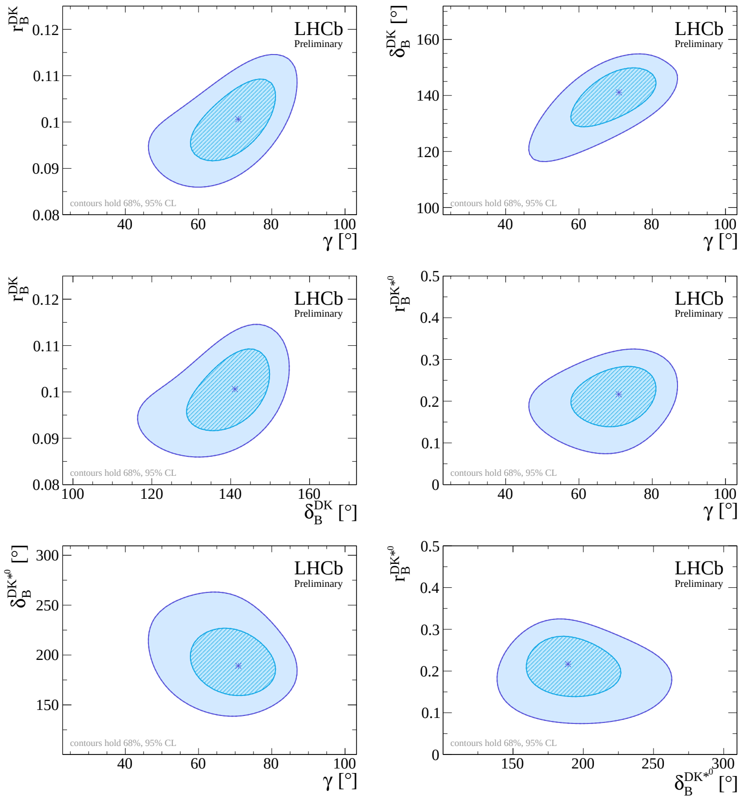


Figure 74.: Profile likelihood contours from the combination of tree-level measurements with two-dimensional 1σ and 2σ boundaries.

7

$B_s^0 \to D_s^{\mp} K^{\pm}$ ANALYSIS USING RUN1 DATA

Contents

7.1 Preliminary Studies **152**
  7.1.1 Correlations among observables 152
  7.1.2 Effects of a time-dependent tagging efficiency 156
7.2 Data sample and MC simulation **163**
7.3 Event selection **165**
7.4 Multi-dimensional fit **169**
  7.4.1 Probability Density Functions 169
  7.4.2 Summary of fixed yields 175
  7.4.3 Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$ 177
  7.4.4 Multi-dimensional fit to $B_s^0 \to D_s^{\mp} K^{\pm}$ 179
7.5 Time sFit **193**

The chapter describes the studies performed so far for the analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays with the full Run1 data sample (2011 at $\sqrt{s} = 7$ TeV and 2012 at $\sqrt{s} = 8$ TeV), which corresponds to an integrated luminosity of 3 fb^{-1}. The strategy follows the main steps of the 1 fb^{-1} analysis described in the previous chapter with some updates and improvements. The selection has been updated with respect to the one used for 1 fb^{-1} analysis. A good quality of the Multi-dimensional fit to the data distribution of the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ channels was already achieved. The results of the Multi-dimensional fitter for both $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ are satisfactory.

Considering the agreement between the results obtained by the time sFit and the cFit of the previous analysis, this analysis update will use only the time sFit given its major simplicity. The time sFit has still to be finalized. Several studies to investigate additional systematics effects with respect to the past analysis have been done. The following sections document the status of the analysis and its major improvements.

7.1 PRELIMINARY STUDIES

The enhancement of statistics may reveal minor systematic deviation introduced by the fit strategy that weren't observed in the analysis of 1 fb^{-1}. A first test has been performed by means of pseudo-experiments in which 3 times the 1 fb^{-1} statistics has been generated and fit with the default fit strategy (Multi-dimensional fit, time sFit). The aim of this test is double: test if any bias arises due to the increased statistics and evaluate the expected sensitivity on the CP parameters. The event generator uses the PDFs and the CP values of the observables corresponding to the nominal fit of the 1 fb^{-1} $B_s^0 \to D_s^{\mp} K^{\pm}$ data sample. The result of the Multi-dimensional fitter on the 930 pseudo-experiments shows that all the floating parameters (signal and background yields) are correctly determined.

Figures 75 show pull distributions[1] of the 5 CP-observables extracted from the time sFit corresponding to the pseudo-experiments samples. The pull distributions have the expected normal Gaussian distributions, indicating that both the values and the errors are correctly determined.

In Tab.59 the results of the pseudo-experiment test with the increased statistics is compared with the results of a similar test performed for the 1 fb^{-1} analysis. The estimated errors scale as expected due to the increase in statistics of a factor three. These results assures that the sophisticated description we used so far, even if the number of events is increased, does not introduce any bias.

7.1.1 *Correlations among observables*

Correlations among the observables involved in the fitter are possible and in principle they can be different from signal to any other background decays. In previous analysis the correlations among the observables were neglected, since in

1 $\mathrm{pull}_i = \frac{x^i_{\mathrm{gen}} - x^i_{\mathrm{fit}}}{\sigma^i_{\mathrm{fit}}}$

Table 59.: Pseudo-experiments for the 5 CP observables extracted from a simulated statistics of 3 and 1 fb^{-1} of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays.

	Parameter	μ of pull	σ of pull	average uncertainty
3 fb^{-1}	C_f	-0.052 ± 0.034	0.931 ± 0.024	0.143
1 fb^{-1}	C_f	0.052 ± 0.034	0.925 ± 0.028	0.249
3 fb^{-1}	S_f	-0.083 ± 0.035	0.975 ± 0.027	0.198
1 fb^{-1}	S_f	-0.095 ± 0.032	0.860 ± 0.026	0.345
3 fb^{-1}	$S_{\bar{f}}$	0.117 ± 0.036	1.005 ± 0.026	0.202
1 fb^{-1}	$S_{\bar{f}}$	0.052 ± 0.037	0.976 ± 0.027	0.350
3 fb^{-1}	D_f	0.029 ± 0.036	0.990 ± 0.029	0.237
1 fb^{-1}	D_f	-0.010 ± 0.040	1.056 ± 0.030	0.407
3 fb^{-1}	$D_{\bar{f}}$	0.026 ± 0.038	1.051 ± 0.029	0.241
1 fb^{-1}	$D_{\bar{f}}$	0.001 ± 0.039	1.054 ± 0.032	0.415

MC they were found to be small. In the current analysis a more accurate estimation is desired in order to define the best fit strategy. In case the impact of such correlation on the CP observables are large we can decide to improve the PDF description, otherwise we will have to evaluate the corresponding systematic uncertainty. Given the limitations in producing large samples of MC events that are required by a deeper study of such effect, an original approach has been developed and implemented. The idea is to produce samples of the main backgrounds decays with a simplified but fast generator, to emulate the detector reconstruction, to apply the core of the selection cuts and to evaluate the correlation between each pair of observables used in the fitter. A very simple way to generate large samples of n-body decays is to use the TGenPhaseSpace class [78] provided by the root package [39]. This class allows to generate any particle decay accounting for the phase space distributions knowing the the particle's kinematics and the masses of the decay products. The momentum P, the pseudorapidity η and the Φ angle of the particle are extracted randomly from the distributions observed in MC. In this way the four-vector of each particle is defined, $\vec{P}_i = (p_{T,i}, \eta_i, \Phi_i, M_i)$, and used together with the masses of the decay products generate the particle decays. The

TGenPhaseSpace returns the four-vector of each daughter particles which can be used to emulate another decay, simulating in this way a full decay chain. The tool is used for generating signal events $B_s^0 \to D_s K$ and the output of the first decay is used to generate $D_s \to KK\pi$. The same procedure has been repeated to generate samples of all relevant background decays:

- $B_s \to D_s\rho$, $D_s \to KK\pi$, $\rho \to \pi\pi^0$
- $B_d \to DK$, $D \to K\pi\pi$
- $\Lambda_b \to D_s^* p$, D $_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $\Lambda_b \to D_s p$, $D_s \to KK\pi$
- $\Lambda_b \to \Lambda_c\pi$, $\lambda_c \to Kp\pi$
- $\Lambda_b \to \Lambda_c K$, $\lambda_c \to Kp\pi$
- $B_s \to D_s^*(\gamma)K$ $D_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $B_s \to D_s^*(\gamma)\pi$, $D_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $B_s \to D_s\pi$, $D_s \to KK\pi$
- $B_d \to D\pi$, $D \to K\pi\pi$
- $B_d \to D_s K$, $D_s \to KK\pi$

Dynamic effects due to intermediate resonances and effects related to the spin of the particles might change the angular distributions. For simplicity, at this stage of the study, they are not considered. The generation tool only provides information on the considered decay and cannot reproduce fully a pp collision event, as it is characterized by tens of additional particles.

Parametric functions are used to emulate the effect of the detector reconstruction and of the signal selection. The effect of the particle reconstruction by the LHCb detector, such as the momentum smearing, has been already studied by LHCb [81] and it can be considered common to different decays and analyses. On the other hand the emulation of the event selection is specific of the considered analysis and it is tuned on the considered channel. Details are given in the appendix A.3. The observables used for the $B_s^0 \to D_s^{\mp}K^{\pm}$ time-dependent analysis for the correlations study are: m(B_s^0), m(D_s), $PIDK$ of the bachelor particle, the B_s^0 decay time t and its error δ_t. The whole procedure has been developed for $B_s^0 \to D_s^{\mp}K^{\pm}$ events and tuned with the signal MC sample comparing several distributions (kinematics of the involved particles and observables) till they match. Then, the main backgrounds which contribute to the $B_s^0 \to D_s^{\mp}K^{\pm}$ selected events have been generated and selected applying exactly the same procedure used for signal events. In Fig. 111 the B_s^0 and D_s mass distributions as well as the $PIDK$ of the bachelor particle are reported for signal and backgrounds decays. Scatter plots among each pair of observables of the Multi-dimensional fitter and of the time sFit, shown in Fig.112, 113, 114, 115, 116, 117, 118, 119, **??** and 121 of appendix A.3, do not highlight large correlations. The corresponding correlation factors, reported in Tab. 60 and Tab. 61: are lower than 10% between MDFitter and time sFit observables, supporting our assumptions of independent observables which makes possible the sFit method.

Table 60.: Emulated backgrounds with the corresponding number of events and correlation factors (%) between MDFitter observables and time *sFit* ones.

		Correlations (%)					
Channel	Entries	δ_t,m_{B_s}	δ_t,PIDK	t_R,PIDK	t_R,m_{B_s}	t_R,m_{D_s}	δ_t,m_{D_s}
$B_s \to D_s K$	84268	0.85	0	0	0.28	-0.14	0.31
$B_s \to D_s \rho$	593627	1.61	4.86	1.86	1.15	-0.29	-0.29
$B_d \to DK$	66001	3.01	4.87	1.10	2.44	2.58	3.00
$\Lambda_b \to D_s^* p$	14001	1.00	6.87	2.06	-0.094	0.29	0.85
$\Lambda_b \to D_s p$	44001	3.18	7.73	1.26	2.68	-0.73	0.19
$\Lambda_b \to \Lambda_c \pi$	45001	1.83	5.40	1.68	1.09	-0.26	-0.23
$\Lambda_b \to \Lambda_c K$	63001	2.44	5.89	0.85	0.07	-0.22	-0.36
$B_s \to D_s^*(\gamma)K$	238535	0.21	5.20	1.99	-1.80	0.16	0.26
$B_s \to D_s^*(\gamma)\pi$	331049	2.34	3.46	1.26	3.03	-0.27	0.51
$B_s \to D_s \pi$	500001	5.12	4.58	2.40	3.91	-0.15	0.27
$B_d \to D\pi$	60001	1.83	4.19	0.63	2.21	2.40	2.62
$B_d \to D_s K$	301275	4.10	5.14	-3.70	-3.54	0.57	0.20

A common feature to all the emulated decays, seen also in MC, is the correlation among the decay time and its error which ranges between 17-23%. For several de-

Table 61.: Emulated backgrounds with the corresponding number of events and correlation factors (%) among observables used in the MDFitter and in the time *sFit* separately.

		Correlations (%)			
Channel	Entries	m_{B_s},m_{D_s}	m_{B_s},PIDK	m_{D_s},PIDK	t_R,δ_t
$B_s \to D_s K$	84268	15.08	4.90	3.05	23.77
$B_s \to D_s \rho$	593627	0.30	1.64	-0.16	21.56
$B_d \to DK$	66001	12.23	-27.61	11.97	23.50
$\Lambda_b \to D_s^* p$	14001	-0.28	-8.39	-1.40	17.08
$\Lambda_b \to D_s p$	44001	1.09	-11.91	-1.05	18.75
$\Lambda_b \to \Lambda_c \pi$	45001	-0.75	8.16	-0.31	20.58
$\Lambda_b \to \Lambda_c K$	63001	0.73	20.85	0.48	20.77
$B_s \to D_s^*(\gamma) K$	238535	6.74	-10.31	1.74	22.66
$B_s \to D_s^*(\gamma)\pi$	331049	1.77	12.96	-0.53	21.66
$B_s \to D_s \pi$	500001	4.90	17.64	-0.30	21.68
$B_d \to D\pi$	60001	7.19	4.85	1.09	20.94
$B_d \to D_s K$	301275	2.44	-21.10	5.98	20.06

cays, the correlation among the B_s^0 mass and the $PIDK$ is found to be larger than 10%, but in any case lower than 28%. Given that the size of correlations is not so large, we are still allowed to neglect the correlations in the fit. The impact of such approximation will be evaluated by means of pseudo-experiment and accounted in the systematic uncertainties. The samples of events generated with the TGenPhaseSpace will be used to model the background contribution with the correlation among observables. Bootstrap method will be eventually used to enlarge further the statistics if needed. Then the nominal fitter (with neglected correlations) is used to see what it would be the impact of such neglected correlations on the CP observables. The pseudo-experiments study has to be finalized.

7.1.2 *Effects of a time-dependent tagging efficiency*

Studies of Flavour tagging performances have shown that the tagging efficiency slightly depends on the B decay time. For SS taggers this can be due to a correlation between the signal B and the tagging particle kinematics, introducing a modulation of the ϵ_{tag} from the momentum of the tagging particle and of the B momentum. For the OS taggers such effect should be smaller. Such a dependency is commonly neglected fit to the time distributions of tagged events and the tagging efficiency is always assumed as a constant. With the following study we wanted to evaluate the impact of such approximation neglecting the time dependence of the tagging efficiency on the CP observables in the $B_s^0 \to D_s^{\mp} K^{\pm}$ fit.

The $B_s^0 \to D_s^- \pi^+$ data are used to provide all the relevant flavour tagging information for the $B_s^0 \to D_s^{\mp} K^{\pm}$ time-dependent analysis and for both the NNetSSK and the OS taggers the tagging efficiency versus the decay time has been investigated. Figure 78 reports Run1 data the decay time dependence of the ϵ_{tag} for the OS and NNetSSK taggers. A linear fit of the decay time dependence of the ϵ_{tag} returns a slope of: α_{OS} = (-0.0029 $\pm$ 0.0018) ps^{-1} and $\alpha_{NNetSSK}$ = (-0.0080 $\pm$ 0.0026) ps^{-1}. While α_{OS} is consistent with zero, $\alpha_{NNetSSK}$ is about 3 σ away from zero. A priori it is not clear how this dependence would affect the CP observables if neglected, so a pseudo-experiments test is done generating the tagging efficiency with the decay time-dependence of NNetSSK ϵ_{tag} (Sample A) and fitting for the CP observables neglecting it. In addition to the sample of pseudo-experiment with the time-dependent tagging efficiency, a second set of toys with constant tagging efficiency is produced to have a reference unbiased result to compare with. Such sample (Sample B) largely overlaps with the corresponding sample with the tagging dependency on time simulated (Sample A). The fitted values of the CP observables do not show any bias in the pull distribution (see Fig.122) as it is reported in Tab.62. The "toy-by-toy" comparison of the result obtained from the two samples (A and B) does not show any bias (pull plots in Fig.124), the larger effect is around 3% (see Tab.63). Since the number of generated pseudo-experiment is not huge (100 toys), an analogous test is done enhancing the time dependence effect by a factor 10 with respect to what we see in data. Also in this case the pull distributions of the CP observables are reasonable and the toy by toy difference shows a maximum effect of the 6%. The conclusion is that the dependence of the NNetSSK ϵ_{tag} on the decay time can be reasonably neglected in the fit. Once the fit strategy to data is in its final shape further studies with a larger number of toys will be performed in order to finalize the evaluation of residual systematic uncertainities.

Table 62.: CP observables mean value and σ of the pull distributions for the toys set generated with a tagging efficiency which depends on the decay time and for the toys set generated with a constant tagging efficiency.

Sample	Parameter	μ of pull	σ of pull	average uncertainty
	C_f	0.23 ± 0.17	1.11 ± 0.24	0.15
	$D_{\bar{f}}$	-0.16 ± 0.18	1.13 ± 0.25	0.24
A ϵ_{tag} (t)	D_f	0.07 ± 0.14	1.02 ± 0.17	0.24
	$S_{\bar{f}}$	0.21 ± 0.15	1.10 ± 0.20	0.20
	S_f	0.17± 0.15	0.98 ± 0.24	0.20
	C_f	0.24 ± 0.13	1.02 ± 0.15	0.15
	$D_{\bar{f}}$	-0.16 ± 0.18	1.12 ± 0.26	0.24
B ϵ_{tag} const.	D_f	0.21 ± 0.16	0.99 ± 0.21	0.23
	$S_{\bar{f}}$	0.03 ± 0.23	1.35 ± 0.30	0.20
	S_f	0.16 ± 0.24	1.42 ± 0.40	0.20

Table 63.: CP observables mean value and σ of the pull distributions for the toy-by-toy difference among Toys set with ϵ_{tag} (t) and ϵ_{tag} = const.

Parameter	μ of pull	σ of pull
C_f	0.023 ± 0.015	0.150± 0.011
$D_{\bar{f}}$	-0.0012± 0.0021	0.0172 ± 0.0018
D_f	-0.0002 ± 0.0020	0.0194 ± 0.0014
$S_{\bar{f}}$	**0.030** ± 0.014	0.139 ± 0.011
S_f	-0.014 ± 0.016	0.153 ± 0.011

For the moment the response of these preliminary studies allow to conclude that we don't need to change the fit strategy. Thus the core of the updated analysis is very similar to the one with the 1 fb^{-1} integrated luminosity.

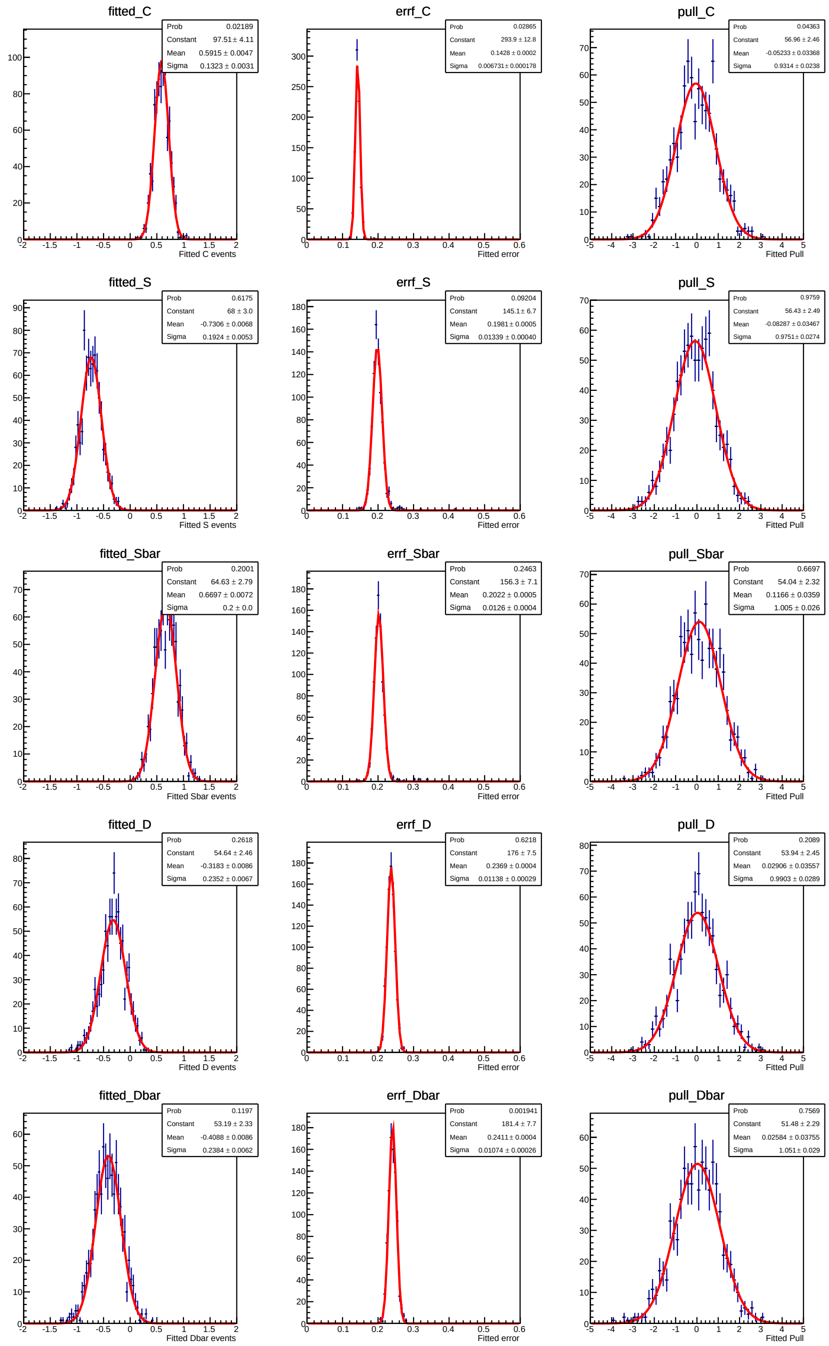


Figure 75.: From left to right, the fitted value of the C_f, S_f, $S_{\overline{f}}$, D_f, $D_{\overline{f}}$ observables, the corresponding error and pull distributions resulting from the time sFit to $B_s^0 \to D_s^{\mp} K^{\pm}$ pseudo-experiments are reported.

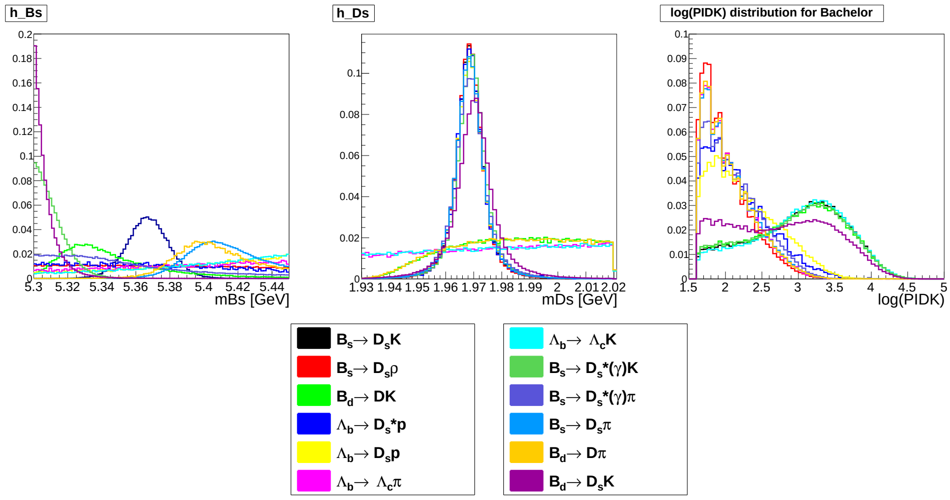


Figure 76.: Emulated bakgrounds decay distributions for B_s mass, D_s mass and PIDK for the bachelor particle.

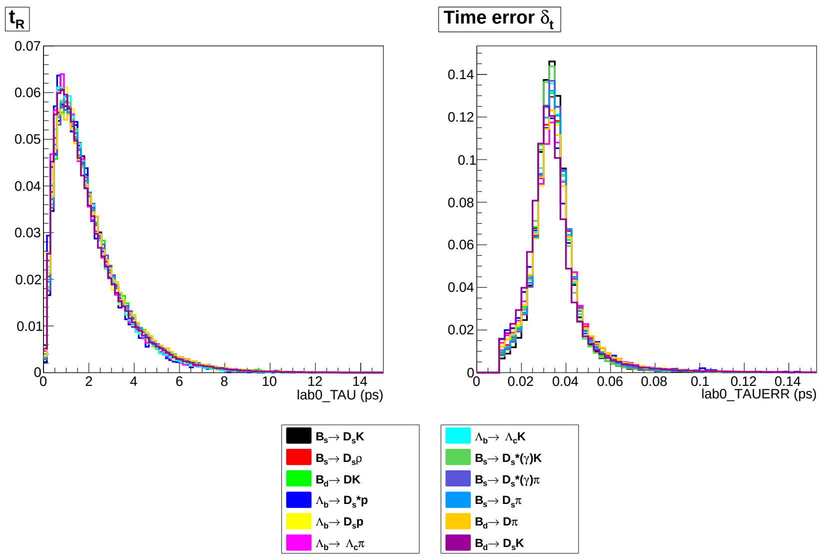


Figure 77.: Emulated bakgrounds decay distributions for B_s decay time and its error.

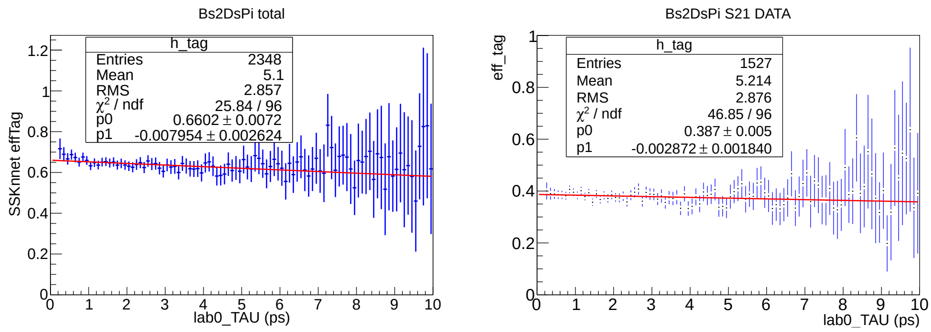


Figure 78.: Tagging efficiency dependence on the decay time for NNetSSK (left) and OS (right) taggers in $B_s^0 \to D_s^- \pi^+$ Run1 data.

7.2 DATA SAMPLE AND MC SIMULATION

The analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays presented in this chapter is based on the full Run1 sample (3.0 fb^{-1}) of pp collisions data taken in 2011 ($\sqrt{s}$=7 TeV) and in 2012 ($\sqrt{s}$=8 TeV). The events are reconstructed using version Reco14 of the reconstruction algorithm. For the 3 fb^{-1} analysis the same stripping lines of the 1 fb^{-1} analysis have been used (see Sec.6.2). As for the analysis of 1 fb^{-1} (Chapter 6), in addition to the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays the $B_s^0 \to D_s^- \pi^+$ decay mode is largely employed in the analysis. Starting from the reconstructed charged particles, the following 5 D_s decays are selected: $D_s^- \to K^- K^+ \pi^-$, $D_s^- \to K^- \pi^+ \pi^-$, $D_s^- \to \pi^- \pi^+ \pi^-$ where in the $D_s \to KK\pi$ selected sample two resonant structures are determined forming the three final samples:$D_s \to \Phi\pi$, $D_s \to K^* K$ and the remaining $D_s \to (KK\pi)_{non\ resonant}$. These D_s^- candidates are subsequently combined with a companion particle to form the $B_s^0 \to D_s^- K^+$ or the $B_s^0 \to D_s^- \pi^+$ decay modes. In all cases the D_s meson is reconstructed on the output of the stripping from three particles with the appropriate mass hypotheses; a "bachelor" track is then added to form the B candidates. When computing the B mass, the D mass is constrained to its nominal value (separately for the D and D_s hypotheses) using the standard `DecayTreeFitter` tools. When computing the B lifetime, the B is additionally constrained to point to the primary interaction, using again the `DecayTreeFitter` tools. The final selection results in a small number of multiple candidates, below 2%, and all are kept in the final analysis. Offline PID requirements are designed in such a way that all ten $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ samples are orthogonal to each other ($PIDK > 5$ and $PIDK < 0$ respectively).

Several Monte Carlo (MC) samples with the updated MC2012 configuration are used in the analysis, notably for selection, normalization and for the study and extraction of signal and background models in the different observables. A full list is given in Tab. 64. All MC samples are selected using the same criteria as the data.

The statistics of the MC samples for $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ decays are lower since for this analysis the strategy to determine the time acceptance will be based on a data driven method (swimming see Ref.[79]).

Table 64.: MC samples used during the analysis for selection, normalization, extraction of signal and background models for the different observables. The second column indicates the specific B-daughter decay.

Sample		Sample size	Use
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to KK\pi$	5188118	Signal study, Flavour Tagging
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to K^-\pi^+\pi^-$	529804	Signal study, Flavour Tagging
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to \pi^-\pi^+\pi^-$	1024641	Signal study, Flavour Tagging
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to KK\pi$	5121168	Signal study
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to K^-\pi^+\pi^-$	562885	Signal study
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to \pi^-\pi^+\pi^-$	1018581	Signal study
$B_s^0 \to D_s^{*-} \pi^+$	$D_s^- \to KK\pi$	2101381	Background study
$B_s^0 \to D_s^- \rho^+$	$D_s^- \to KK\pi$	2067161	Background study
$\Lambda_b^0 \to D_s^- p$	$D_s^- \to KK\pi$	506641	Background study
$\Lambda_b^0 \to D_s^{*-} p$	$D_s^- \to KK\pi$	572773	Background study
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	$\Lambda_c \to pK\pi$	2101092	Background study
$\Lambda_b^0 \to \Lambda_c^+ K^-$	$\Lambda_c \to pK\pi$	542701	Background study
$B^0 \to D^- \pi^+$	$D \to K\pi\pi$	5141477	Obtain Data/MC corrections, Background study
$B^0 \to D^- K^+$	$D \to K\pi\pi$	1052943	Background study

7.3 EVENT SELECTION

The recorded data samples pass subsequent selection requirements, each one being tighter than the previous one. They can be summarized as: trigger strategy, stripping selection, offline selection which contains also the particle identification requirements.

TRIGGER STRATEGY The trigger strategy does not undergo relevant changes but only minor adjustment related to the employment of swimming method. All events are required to be TOS (Triggered on Signal) on the `1TrackAllL0` line in HLT1 and on the 2-,3-body. Unlike the published analysis, the trigger line 4-body is excluded because difficult to treat with the swimming method (losing only 2% of the events).

STRIPPING SELECTION The stripping selection has been updated using the last version of stripping processing (Stripping21), already validated and employed by several analysis. There aren't significative difference in the stripping selection with respect to the $1fb^{-1}$ analysis.

OFFLINE SELECTION The offline selection is composed by different ingredients which do not undergo relevant changes from what has been done for the $1fb^{-1}$ analysis:

- kinematic requirements common to all modes: D_s mass in [1930,2015]MeV/c^2, D_s lifetime > 0ps . Specific requirements on the D_s separation vertex are used to suppress charmless backgrounds and they are reported in Tab.30 for each D_s final state

- gradient boosted decision tree (BDTG) to suppress combinatorial events

- optimization of the BDTG output cut to maximized the signal significance

- PID requirements on the final charged particles reported in Tab.31

- definition of vetoes for D , Λ_c^+ and J/ψ decays specific for each D_s final states (see Tab.30)

The offline selection is very similar to the one adopted by the previous analysis, despite the fact that the studies on the BDTG selection has been redone using the current dataset in order to be more effective in discriminating signal from backgrounds given the larger statistics and the different reconstruction algorithm used. Indeed the main goal of the BDTG selection is to suppress combinatorial events and it follows a data-driven approach. The classifier is trained on $B_s^0 \to D_s^- \pi^+$, $D_s^- \to KK\pi$ data samples and the kinematic variables used as input information for the classifier are the same of the old BDTG. Both signal and combinatorial samples are extracted from $B_s^0 \to D_s^- \pi^+$ data. As it has been done for the previous BDTG, the signal events are extracted through a B_s^0 mass fit in a narrow region around the signal peak (5310MeV/c^2< m(B_s^0)<5430 MeV/c^2) using sWeights while the combinatorial events are selected from upper B_s^0 mass sidebands (m(B_s^0)> 5450 MeV/c^2). Half of the $B_s^0 \to D_s^- \pi^+$ data sample is used to train the BDTG, the other half is used to check that the BDTG output does not suffer of overtraining effects. In Fig.79 the BDTG response for signal (red histogram) and background (blue histogram) events is shown for the training sample and the test sample. The new BDTG has better performance with respect to the old one in terms of background to signal ratio and background rejection. This BDTG is used in all the data samples: $B_s^0 \to D_s^\mp K^\pm$, $B_s^0 \to D_s^- \pi^+$ and also $B^0 \to D^- \pi^+$.

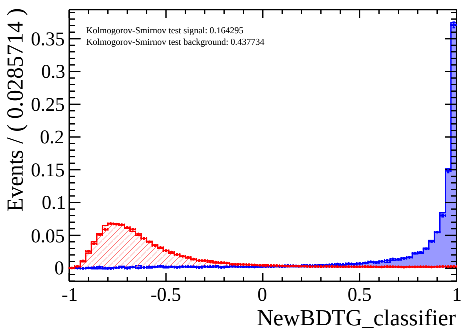


Figure 79.: BDTG response distributions for training and test samples. The red histogram corresponds to signal weighted by the signal sWeights, while the blue shows combinatorial background.

The BDTG selection is followed by the PID requirements on the final charged particles depending on the D_s decay mode and on the bachelor particle. The PID criteria are exactly the same of the ones applied in the $1fb^{-1}$ analysis. The PID performance, efficiency and mis-identification rate, for each PID cut applied in the analysis have been determined using the D^* data samples from the PIDCalib tool in which the reconstruction algorithm and stripping version are the same as for the analyzed data samples. PID cuts are the same as the ones applied in the 1 fb^{-1} analysis (see Tab.31). The optimization of the BDTG selection cut has been determined looking at the Multi-dimensional fitter result to $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ data corresponding to different cut values. The optimal value has been empirically decided looking at the signal significance[2] and signal purity[3] distributions with respect to the combinatorial backgrounds. In Tab.65 the performance of the BDTG at the optimal cut value are reported. Both the signal significance and signal purity values increase by increasing the cut value of the BDTG, reaching a plateau for BDTG $>$ 0.0, which is assumed as reference cut value.

Table 65.: BDTG performance evaluated at the optimal cut value of > 0.0.

BDTG cut	ϵ_{sig} [%]	Purity $\cdot\ \epsilon_{sig}$ [%]	Signal Significance $\cdot\ \epsilon_{sig}$ [%]
BDTG $>$ 0.0	96.1	82.4	93.8

Although the current BDTG has very similar inputs as the previous one, a different impact on the distributions of combinatorial events is visible, in particular this will be evident looking at the D_s mass distribution in Sec.7.4.1. The overall selection turns out to be improved with respect to the previous one: the observed signal yield is larger than the expected one (N_{sig} = 3 N_{sig}(1fb^{-1}) for both $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ and $B_s^0 \rightarrow D_s^{-} \pi^{+}$ decay modes as reported in Tab.66, and also the signal to background ratio is increased in this updated analysis.

The selected events are reported in Fig.80 for $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ (left) and $B_s^0 \rightarrow D_s^{-} \pi^{+}$ (right) decay modes. The signal peak is clearly visible on top of the combinatorial events. A contribution from partially reconstructed backgrounds in the lower mass region is also evident. In the following section details on the Multi-dimensional

2 Signal Significance $S/\sqrt{S+B}$.

3 Signal Purity S/S+B.

Table 66.: Comparison between signal yields and signal to background ratio in the updated analysis ($3fb^{-1}$) and in the previous one ($1fb^{-1}$). N_{bkg} corresponds to the sum of all the background contributions.

Channel	$3fb^{-1}$	$1fb^{-1}$
N_{sig} ($B_s^0 \to D_s^{\mp}K^{\pm}$)	5825 ± 95	1 768 ± 49
N_{sig} ($B_s^0 \to D_s^{-}\pi^{+}$)	91 409 ± 346	28 264 ± 182
N_{sig}/N_{bkg} ($B_s^0 \to D_s^{\mp}K^{\pm}$)	0.54	0.43
N_{sig}/N_{bkg} ($B_s^0 \to D_s^{-}\pi^{+}$)	3.17	2.7

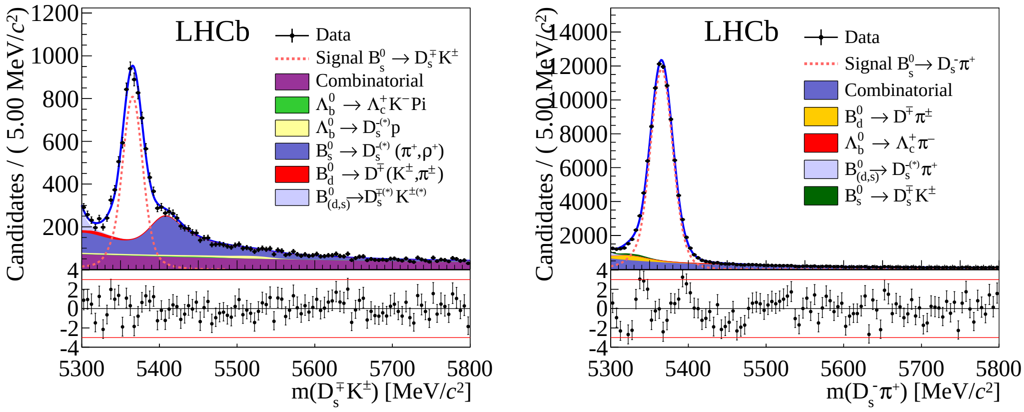


Figure 80.: $B_s^0 \to D_s^{\mp}K^{\pm}$ (left) and $B_s^0 \to D_s^{-}\pi^{+}$ (right) events after the overall selection on the $3fb^{-1}$ data set with the superimposed B_s^0 mass distributions for the different background components described in Sec.7.4.1.

fit and its results to extract signal weights in $B_s^0 \to D_s^{\mp}K^{\pm}$ and $B_s^0 \to D_s^{-}\pi^{+}$ decay mode will be given.

7.4 MULTI-DIMENSIONAL FIT

The aim of the Multi-dimensional fit to both $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ and $B_s^0 \rightarrow D_s^- \pi^+$ is to extract signal weights in order to perform a decay time fit in which only signal events are considered thanks to the statistics-based background subtraction applying signal sWeights to the sample of selected events. The fit strategy of the Multi-dimensional fitter of the 3 fb^{-1} data sample is very similar to the one of the 1 fb^{-1} data sample (see Sec.6.3) except for the following changes:

- $B_s^0 \rightarrow D_s^- \pi^+$ fit uses the logarithm of the $PIDK$ observable: -log($PIDK$) $\in$ [-7.0,5.0]
- Updated description of the masses and $PIDK$ PDFs of the combinatorial backgorund (related to the fratures of the new BDT)
- Updated mass and $PIDK$ PDFs based on the updated MC samples, PID and Mis-ID efficiencies
- Updated data/MC corrections obtained using the result of the fit to the $B^0 \rightarrow D^- \pi^+$ data sample. A fit to the B^0 invariant mass distribution split by year and magnet polarity

Samples corresponding to the two magnet polarities and years of data taking are merged together and the corresponding PDFs are added with relative weights defined according to the measure luminosities [fb^{-1}]: 0.546 (2011, Down), 0.408 (2011, Up), 0.932 (2012, Down) and 0.954 (2012, Up). Like for the 1 fb^{-1} analysis, a maximum likelihood fit of the 5 sub-samples ($D_s^- \rightarrow (KK\pi)_{nonres}$, $D_s^- \rightarrow \phi\pi^-$,$D_s^- \rightarrow K^{*0}K^-$,$D_s^- \rightarrow K^-\pi^+\pi^-$, $D_s^- \rightarrow \pi^-\pi^+\pi^-$) is performed simultaneously.

7.4.1 *Probability Density Functions*

SIGNAL The procedure to determine the PDF of the signal strictly follows what described for 1 fb^{-1} analysis (see Sec.6.3.1). The parameters of the PDF(m(B_s^0)) and PDF(m(D_s)) are determined from the fit to the current MC sample.

Given the high similarity among the signal shapes in the different D_s final states, also seen in the previous analysis, and the lower statistics with respect to the $B_s^0 \to D_s^- \pi^+$ sample, the signal parametrization for the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample is common to all the 5 final states, while for the D_s mass it is split in three components: $D_s^- \to KK\pi$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Figure 82 (in Fig.81) shows the result for $B_s^0 \to D_s^{\mp} K^{\pm}$ ($B_s^0 \to D_s^- \pi^+$) fit to the B_s^0 mass distribution. The corresponding parameters are listed in Tab.67 and in Tab. 68 respectively for the B_s^0 invariant mass in $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ decays.

The parameters of the fit to the D_s mass are reported in Tab.69 for $B_s^0 \to D_s^- \pi^+$ and in Tab.70 for $B_s^0 \to D_s^{\mp} K^{\pm}$ samples while the fit result and the data samples are plotted in Fig.83 for $B_s^0 \to D_s^- \pi^+$ and in Fig.84 for $B_s^0 \to D_s^{\mp} K^{\pm}$ events.

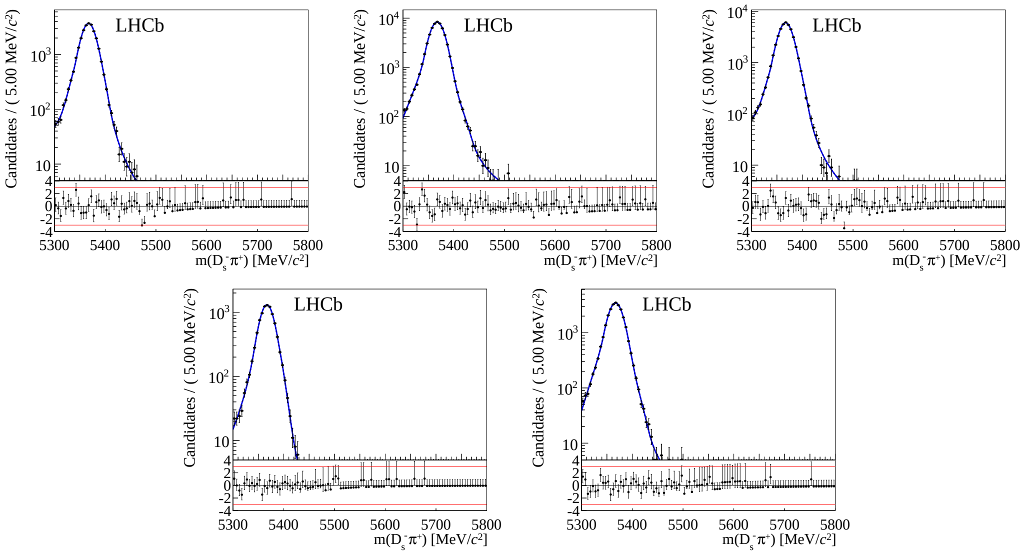


Figure 81.: Fit to B_s^0 invariant mass in $B_s^0 \to D_s^- \pi^+$ signal MC events on the joint magnet up and down samples, separately for D_s final states. From top left to bottom right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$.

The signal parametrizations are satisfactory as also visible from the pulls. In the fit to data, all signal B_s^0 and D_s mass parameters are fixed, except the mean values which are floating. Moreover, in order to account for eventual differences in the mass resolutions between data and MC, a scale parameter R is introduced ($R = \sigma_{data}/\sigma_{MC}$) and left free in the fit to data.

Table 67.: Parameters of the Double Crystal Ball describing the signal B_s^0 shapes of $B_s^0 \to D_s^- \pi^+$, obtained from signal MC, separately for each D_s final state.

	D_s decay mode				
Parameter	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$\mu(B_s^0)$ [MeV/c^2]	5367.2 ± 0.1	5367.2 ± 0.07	5367.0 ± 0.08	5366.9 ± 0.18	5367.0 ± 0.11
σ_1 [MeV/c^2]	11.75 ± 0.48	13.31 ± 0.12	12.02 ± 0.55	16.57 ± 0.70	13.40 ± 0.29
σ_2 [MeV/c^2]	18.92 ± 1.05	28.55 ± 1.19	19.64 ± 1.51	10.99 ± 0.63	26.15 ± 2.22
α_1	2.15 ± 0.18	1.90 ± 0.10	2.39 ± 0.13	1.45 ± 0.11	1.77 ± 0.15
α_2	-2.20 ± 0.10	-2.39 ± 0.08	-2.33 ± 0.09	-2.05 ± 0.22	-2.44 ± 0.16
n_1	0.64 ± 0.31	2.62 ± 0.69	0.39 ± 0.24	4.58 ± 3.86	3.14 ± 1.47
n_2	2.51 ± 0.31	1.31 ± 0.17	2.10 ± 0.28	1.96 ± 0.35	1.48 ± 0.37
f	0.56 ± 0.08	0.86 ± 0.02	0.60 ± 0.10	0.61 ± 0.08	0.81 ± 0.04

The signal $PIDK$ shapes for both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ are determined from PID calibration data samples reweighted by the kinematic distributions from signal simulated samples as for the $1fb^{-1}$ analysis.

COMBINATORIAL Combinatorial events are composed by combinatorial events in which all the final charged tracks are selected randomly plus a fraction of combinatorial events in which a true D_s is associated to a random bachelor particle. The shape for combinatorial events is determined on data looking at the B_s^0 mass upper sideband but in the nominal D_s mass range ([1930,2015]MeV/c^2). The B_s^0 mass distribution of combinatorial events is described:

- by a single exponential in $B_s^0 \to D_s^\mp K^\pm$
- by a single exponential plus a constant in $B_s^0 \to D_s^- \pi^+$.

In the fit to each D_s final states the combinatorial parameters are left free.

In the published analysis combinatorial events in the D_s mass observable were described by an exponential for pure combinatorial events and only for $D_s^- \to KK\pi$ modes the Double Crystal Ball used for the signal events is added to describe combinatorial events which are formed by a true D_s . It turns out that this descrip-

Table 68.: Parameters of the Double Crystal Ball describing the signal B_s^0 shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC. Parameters without uncertainties are fixed in the fit to MC events.

Parameter	All
$\mu(B_s^0)$ [MeV/c^2]	5367.4 $\pm$ 0.05
σ_1 [MeV/c^2]	12.92 $\pm$ 0.16
σ_2 [MeV/c^2]	19.94 $\pm$ 0.41
α_1	1.9
α_2	-1.56
n_1	0.69 $\pm$ 0.08
n_2	4.47 $\pm$ 0.12
f	0.49 $\pm$ 0.03

tion is no more reasonable for the updated analysis. Indeed two modifications wereneeded to get a good agreement with data:

- to add the contribution of a Double Crystal Ball function also for the $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ final states,

- to introduce an independent free scale factor R for the Double Crystal Ball resolution describing the D_s peak in combinatorial background.

The need of this extended parametrization of the combinatorial background is a consequence of the new BDTG selection and it is particularly important for $B_s^0 \to D_s^-\pi^+$ given the larger statistics with respect to $B_s^0 \to D_s^{\mp} K^{\pm}$ In Figure 85, the D_s mass distribution for signal (under the B_s^0 peak) and for backgrounds (in B_s^0 sidebands) events are plotted for $D_s^- \to \phi\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ final states, while the D_s component is clearly visible in both D_s decay modes, the width of the peak is similar to the one of the signal for $D_s^- \to \phi\pi^-$, while it is much broader for $D_s^- \to \pi^-\pi^+\pi^-$ decay mode. In Figure 86, the same data set are selected applying the old BDTG. In this case the D_s components in $D_s^- \to \pi^-\pi^+\pi^-$ decay mode is not present

The exponential parameters, as well as the relative fractions of true D_s and the R parameter of the combinatorial Double Crystal Ball, are floating in the fit to data.

Table 69.: Parameters of the Double Crystal Ball describing the signal D_s shapes of $B_s^0 \to D_s^- \pi^+$, obtained from signal MC, separately for each D_s final state.

	D_s decay mode				
Parameter	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$\mu(D_s)$ [MeV/c^2]	1969.2 ± 0.06	1969.2 ± 0.01	1969.2 ± 0.05	1969.3 ± 0.09	1969.5 ± 0.09
σ_1 [MeV/c^2]	5.81 ± 0.43	5.47 ± 0.15	5.78 ± 0.13	12.62 ± 1.61	9.06 ± 0.44
σ_2 [MeV/c^2]	5.35 ± 0.23	5.69 ± 0.15	5.72 ± 0.20	6.65 ± 0.30	8.16 ± 0.18
α_1	1.09 ± 0.12	1.10 ± 0.07	1.25 ± 0.07	1.45 ± 0.13	0.80 ± 0.07
α_2	-1.15 ± 0.15	-1.01 ± 0.05	-1.09 ± 0.08	-3.68 ± 0.43	-1.20 ± 0.10
n_1	8.49 ± 4.14	6.16 ± 1.16	7.88 ± 1.69	16.15 ± 3.63	91.50 ± 91.50
n_2	12.77 ± 8.04	20.07 ± 6.33	31.1 ± 16.2	0.0 ± 0.004	8.16 ± 0.18
f	0.46 ± 0.09	0.46 ± 0.04	0.55 ± 0.06	0.27 ± 0.08	0.41 ± 0.06

Combinatorial events can be composed by events which have as bachelor particle pions, kaons or protons. In order to accurately describe these events $PIDK$ shapes are evaluated for the three particle hypothesis:

$$PDF_{PIDK} = f_{PIDK1} PDF_{PIDK1} + (1 - f_{PIDK1})[f_{PIDK2} PDF_{PIDK2} + (1 - f_{PIDK2} PDF_{PIDK3})] \tag{7.1}$$

and the relative fractions are let free in the fit to data. The different PDF_{PIDKi} shapes are obtained from the PID calibration samples for pions, kaons and protons, the data driven procedure discussed in Sec.6.3.1 is applied here to the updated data samples. A more detailed description is now required to model the $PIDK$ distributions for the combinatorial in $B_s^0 \to D_s^- \pi^+$ decays. The $PIDK$ observable is very correlated to the kinematics of the bachelor particle and the bachelor pion in $B_s^0 \to D_s^- \pi^+$ decays show different kinematic distributions among the D_s final states as reported in Fig.87. Then the $PIDK$ shapes are split on the D_s final states (see Fig.88).

This further splitting is used in $B_s^0 \to D_s^- \pi^+$ fit but not in $B_s^0 \to D_s^{\mp} K^{\pm}$.

PARTIALLY RECONSTRUCTED BACKGROUNDS The description of partially reconstructed backgrounds does not change with respect to the previous analysis:

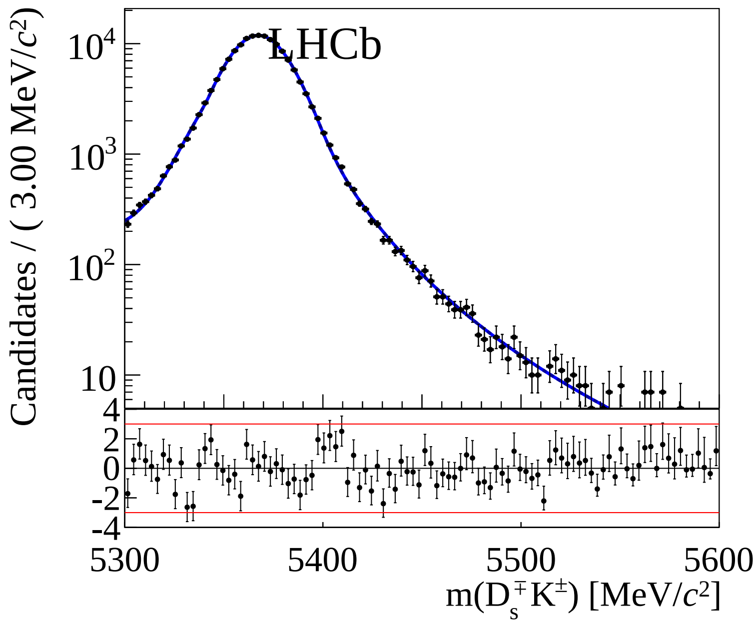


Figure 82.: Fit to B_s^0 invariant mass in $B_s^0 \to D_s^{\mp} K^{\pm}$ signal MC events on the joint magnet up and down samples and D_s final states.

all the shapes have been updated to the new data sets of MC and PID calibration samples and efficiencies.

FULLY RECONSTRUCTED BACKGROUNDS The description of fully reconstructed backgrounds undergoes only minor changes and tunings with respect to what is done for the $1fb^{-1}$ analysis. The fully reconstructed backgrounds included in this analysis are listed in Tab.34 of the previous chapter. Different methods have been used to obtain the B_s^0 mass shapes for these backgrounds.

- Mass shapes for backgrounds which decay in the same signal final state adopt the Double Crystal Ball of the signal. The means are shifted by the well known [?] m(B_s^0) - m(B) = 86.9MeV/c^{2}mass difference while the sigmas are corrected by the width ratio $B_s^0 \to D_s^{\mp} h^{\pm} / B^0 \to D_s^{\mp} h^{\pm}$ measured in MC samples.

- The $B^0 \to D^- \pi^+$ background of the $B_s^0 \to D_s^- \pi^+$ and the $B_s^0 \to D_s^- \pi^+$ background of the $B_s^0 \to D_s^{\mp} K^{\pm}$ are the most relevant fully reconstructed

Table 70.: Parameters of the Double Crystal Ball describing the signal D_s shapes of $B_s^0 \to D_s^{\mp} K^{\pm}$, obtained from signal MC, separately for each D_s final state. Parameters without uncertainties are fixed in the fit to MC events.

	D_s decay mode		
Parameter	$KK\pi$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$\mu(D_s)$ [MeV/c^2]	1969.2 $\pm$ 0.02	1969.3 $\pm$ 0.08	1969.4 $\pm$ 0.09
σ_1 [MeV/c^2]	7.95 $\pm$ 0.11	10.75 $\pm$ 0.55	8.56 $\pm$ 0.19
σ_2 [MeV/c^2]	4.75 $\pm$ 0.06	6.17 $\pm$ 0.22	7.82 $\pm$ 0.17
α_1	1.83	2.13	0.90
α_2	-2.22	-3.55	-0.93
n_1	2.72 $\pm$ 0.15	0.65 $\pm$ 0.27	23.0 $\pm$ 16.5
n_2	1.57 $\pm$ 0.08	0.0 $\pm$ 0.10	80.0 $\pm$ 79.9
f	0.48 $\pm$ 0.02	0.39 $\pm$ 0.06	0.54 $\pm$ 0.02

backgrounds (second row of Tab.34). In this updated analysis the B_s^0 mass shapes for these two backgrounds are obtained from MC, since the previous approach does not provide satisfactory results in the fit.

- The shapes of the remaining fully reconstructed backgrounds are taken from MC, corrected by the momentum-dependent PID efficiency and by residual kinematic discrepancies between data and MC.

The D_s mass shape for the fully reconstructed backgrounds is considered as the signal one for backgrounds with true D_s mesons. In the case of the $B^0 \to D^-\pi^+$ decay it is taken from MC. The $PIDK$ shapes are obtained from the PID calibration samples, which are reweighted to match the kinematics found in the relevant signal samples. When the final state is the same as the one of the signal, the same $PIDK$ shape is employed. In the other cases the `RooBinned1DQuinticBase` is used to obtain the $PIDK$ template.

7.4.2 *Summary of fixed yields*

The yields of some backgrounds contribution (in particular fully reconstructed backgrounds) need to be fixed because difficult to determine due to the fact that

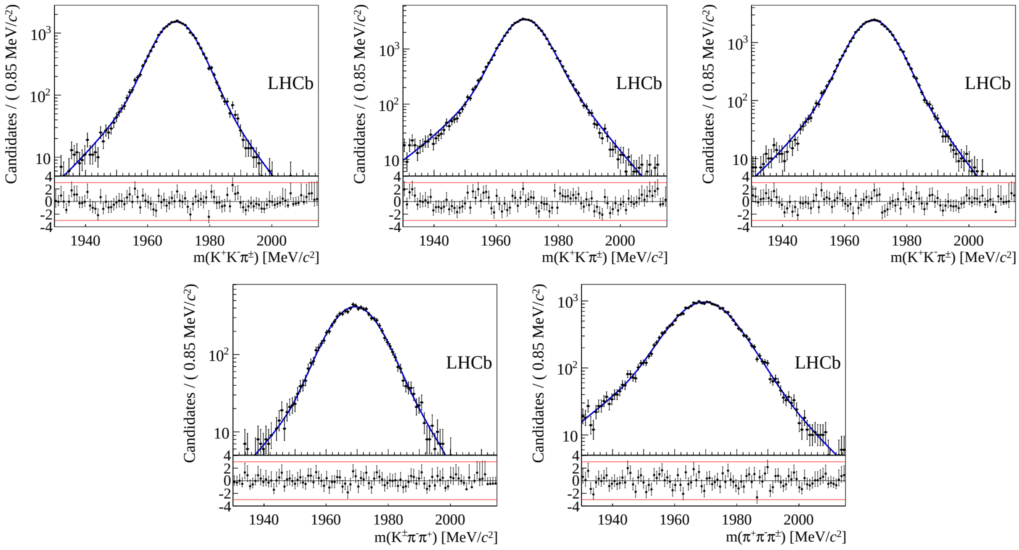


Figure 83.: Fit to D_s invariant mass in $B_s^0 \to D_s^- \pi^+$ signal MC events on the joint magnet up and down samples, separately for D_s final states. From top left to bottom right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$.

they peak under the signal and it would be difficult for the fitter to recognize them. The $B^0 \to D^-\pi^+$ yield in the $B_s^0 \to D_s^-\pi^+$ sample is estimated by multiplying the yield of the $B^0 \to D^-\pi^+$ obtained from the mass fit to this control mode to the ratio of selection efficiencies for $B^0 \to D^-\pi^+$ in B_d^0 and B_s^0 data sets evaluated from MC. A similar procedure is also used to estimate the yields of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ in the $B_s^0 \to D_s^-\pi^+$ fit. In this case the the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ data sample is obtained reconstructing the $B_s^0 \to D_s^-\pi^+$ data sample under the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ mass hypothesis. The $B_s^0 \to D_s^{\mp}K^{\pm}$ background yield in the $B_s^0 \to D_s^-\pi^+$ sample is related to the signal yield through a factor which accounts for the relative branching fraction between the two decay modes and the ratio of the PID Mis-ID and efficiency for the PIDK cut since the rest of the selection and the kinematic distributions are expected to be identical. This factor has been evaluated in MC and the $B_s^0 \to D_s^{\mp}K^{\pm}$ background yields have been determined in a first MDFitter, and then fixed in the nominal one. All the fixed yields are reported in Tables 71 and 72.

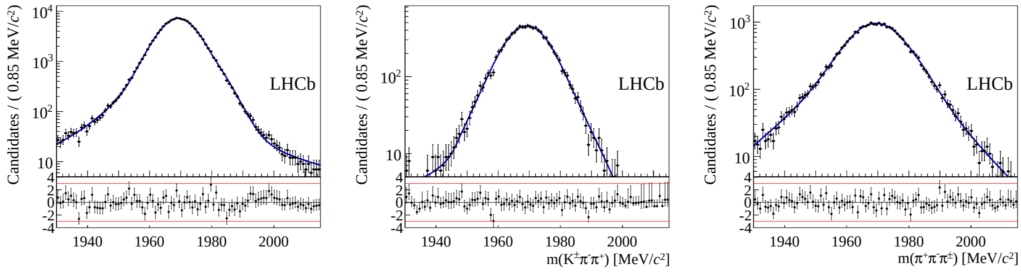


Figure 84.: Fit to D_s invariant mass in $B_s^0 \to D_s^{\mp} K^{\pm}$ signal MC events on the joint magnet up and down samples, separately for $D_s^- \to KK\pi$ (left), $D_s^- \to K^-\pi^+\pi^-$ (center) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes.

Table 71.: Constrained yields in the $B_s^0 \to D_s^{\mp} K^{\pm}$ Multi-dimensional fit.

	D_s decay mode				
Background	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$B^0 \to D^-K^+$	83.1	4.6	65.6	0.0	0.0
$B^0 \to D^-\pi^+$	50.0	2.6	38.6	9.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ K^-$	48.8	7.5	14.4	0.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	29.2	4.3	8.5	0.0	0.0

7.4.3 *Multi-dimensional fit to $B_s^0 \to D_s^- \pi^+$*

In order to make a first validation of the signal and backgrounds descriptions and also to see the complete shapes of partially reconstructed backgrounds, a fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass has been done fitting the B_s^0 mass distribution only in a wide range: [5100,5800]MeV/c^2. All the B_s^0 mass templates as well as signal parametrization have been updated taking into account the larger mass range. To fit the wider mass range an additional partially reconstructed background, which

Table 72.: Constrained yields in the $B_s^0 \to D_s^- \pi^+$ Multi-dimensional fit.

	D_s decay mode				
Background	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$B^0 \to D^-\pi^+$	1618.0	86.5	1269	0.0	0.0
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	961.0	151.0	284.0	8.0	0.0
$B_s^0 \to D_s^{\mp} K^{\pm}$	157.0	229	222	116	211

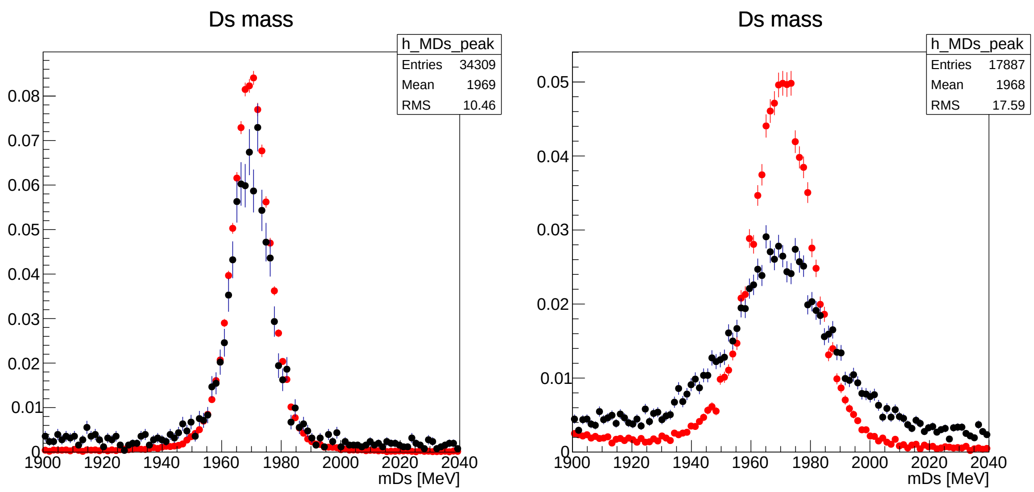


Figure 85.: D_s mass distributions for $B_s^0 \to D_s^- \pi^+$ combinatorial (black) and signal (red) events for $D_s^- \to \phi\pi^-$ (left) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes.

is negligible in the nominal mass window, is required. It corresponds to the $B_s^0 \to D_s^- \rho^+$decay mode. The fit result in the five D_s final states is reported in Fig.89, and merged result is in Fig.90. The plots show a reasonable agreement between data and our fit result. The number of signal yield is of $\approx$ 94 000 events, which is a little bit higher than the one obtained by the nominal Multi-dimensional fit. Notice that the fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass is mono-dimensional, and for this reason has a lower discriminating power between signal and backgrounds than the Multi-dimensional one. Thus it is possible that some backgrounds events are not well discriminated from signal events.

The results of the nominal $B_s^0 \to D_s^- \pi^+$ Multi-Dimendional fitter fits the B_s^0 mass, D_s mass, and the $PIDK$ variable of the 5 D_s final states are shown in Fig. 91, Fig. 92, Fig. 93 for the B_s^0 and D_s masses, and $PIDK$, respectively.

The combined result for all subsamples are presented Fig. 94. The corresponding fitted parameters are listed in Tab. 73. All the parameters related to combinatorial events are floating in the fit. The yields for signal, combinatorial and partially reconstructed events are quoted for each D_s final states. The total signal yields is of $\approx$ 91 400 events, that compared to the signal yield obtained in the 1 fb^{-1} analysis is roughly a 20% larger than what expected from the increase in statistics. The

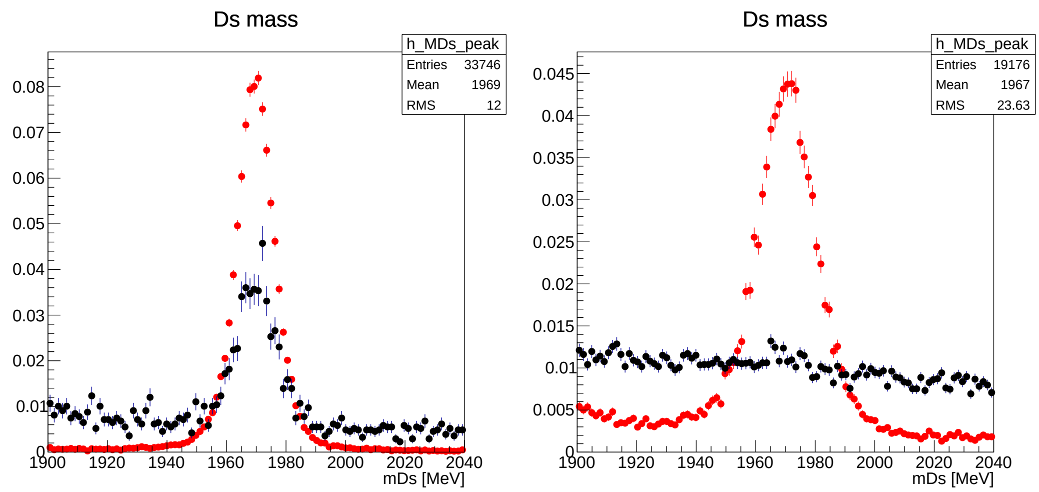


Figure 86.: D_s mass distributions for $B_s^0 \to D_s^- \pi^+$ combinatorial (black) and signal (red) events for $D_s^- \to \phi\pi^-$ (left) and $D_s^- \to \pi^-\pi^+\pi^-$ (right) decay modes applying the old BDTG selection.

combinatorial yield is of $\approx$ 23 000 events while the contribution of partially reconstructed events is smaller, around 315 events. We see an improved background rejection with respect to the corresponding background yields observed in 1 fb^{-1} analysis (see Tab.66). As expected, the R parameters are larger than 1.0 indicating that the mass resolution in MC is too optimistic with respect to what we see in data. A great improvement of the fit result is given by the use of independent R parameters for the Double Crystal Ball of combinatorial events with respect to signal ones in the D_s invariant mass: $RComb_{D_s}$ are significantly larger than the values found for the R_{D_s} of signal. There is a good agreement between Multi-dimensional fit result and the data distribution.

7.4.4 *Multi-dimensional fit to $B_s^0 \to D_s^{\mp} K^{\pm}$*

The fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ distribution is constructed analogously to the $B_s^0 \to D_s^- \pi^+$ fit.

The fit results are shown in Figs. 95, 96, 97, and 98. The corresponding fitted parameters are listed in Tab. 74. The number of signal (combinatorial) events is found to be of $\approx$ 6 000 (5 000) events and as for the $B_s^0 \to D_s^- \pi^+$ observed signal

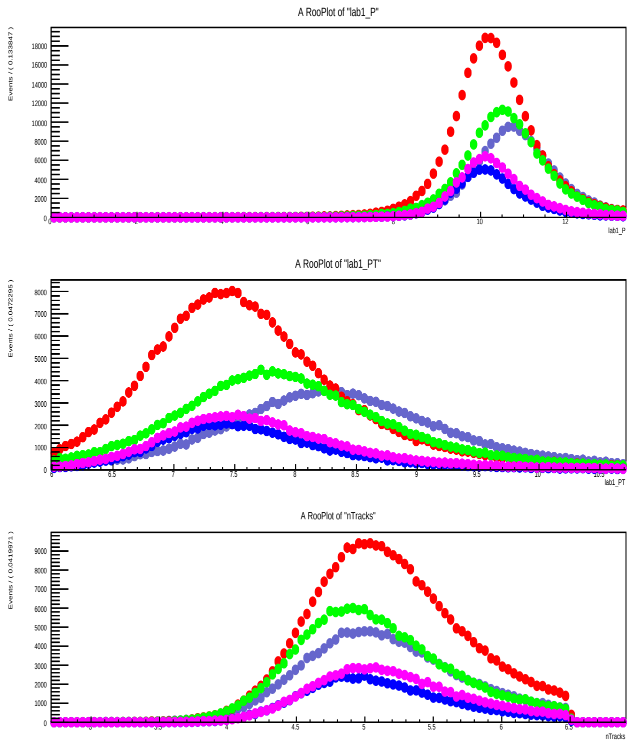


Figure 87.: From top to bottom Kinematic distributions of p, p_T and nTracks of bachelor particles for $B_s^0 \to D_s^- \pi^+$ combinatorial events in the signal region [5300,5800]MeV/c^{2}splitted for the different D_s final states (red $D_s^- \to \phi\pi^-$, blue $D_s^- \to K^-\pi^+\pi^-$, green $D_s^- \to K^{*0}K^-$, cyan $D_s^- \to (KK\pi)_{nonres}$, magenta $D_s^- \to \pi^-\pi^+\pi^-$).

yield, it is a 20% larger (30% smaller) than what expected from the increase in statistics with respect to the 1 fb^{-1} analysis. Such an improved level of background rejection and signal selection is a feature of the new BDTG optimization.

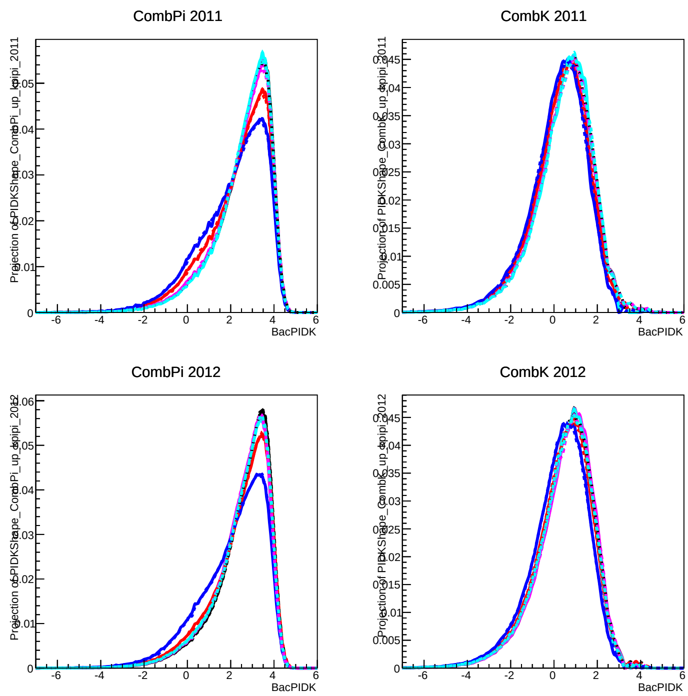


Figure 88.: $PIDK$ distributions for pions (left) and kaons (right) for $B_s^0 \to D_s^- \pi^+$ combinatorial events split for the different D_s final states (red $D_s^- \to K^{*0}K^-$, magenta $D_s^- \to \phi\pi^-$, blue $D_s^- \to (KK\pi)_{nonres}$, black $D_s^- \to K^-\pi^+\pi^-$, cyan $D_s^- \to \pi^-\pi^+\pi^-$,).

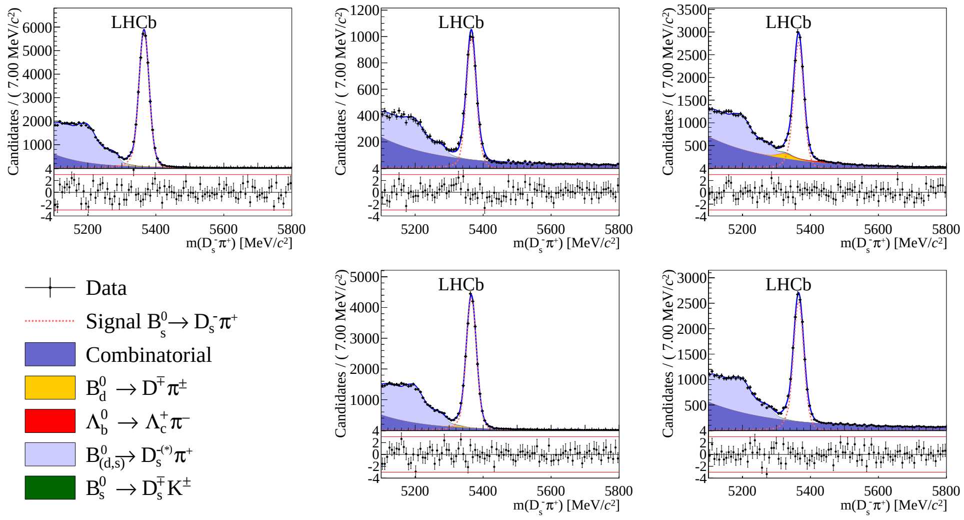


Figure 89.: Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the wide mass region [5100,5800]MeV/c^{2}in the different D_s final states from top left to bottom right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-\pi^+\pi^-$,$D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$.

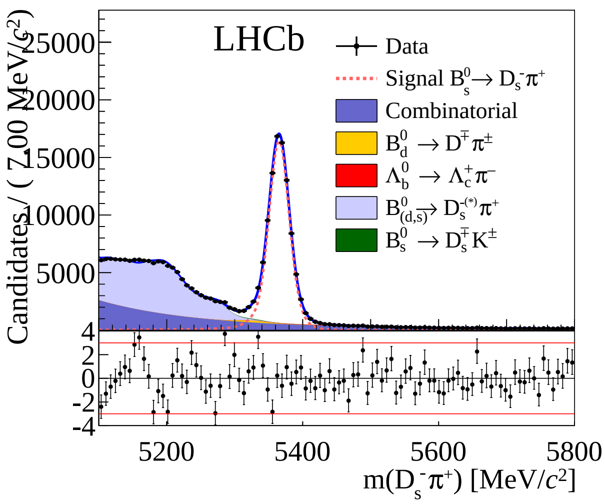


Figure 90.: Fit to the $B_s^0 \to D_s^- \pi^+$ invariant mass distribution in the wide mass region [5100,5800]MeV/c^2.

Table 73.: Fitted values of the parameters for the $B_s^0 \to D_s^- \pi^+$ signal MD fit. Means are the parameters of the double Crystal Ball used to describe the signal in B_s^0 and D_s masses. The $slope_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. The fractions $fComb_{B_s}$ are the fractions between an exponential and flat distribution in B_s^0 mass, whereas $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. The $fComb_{PIDK}$ is a fraction between pion and kaon components in the PIDKcombinatorial shape.

	D_s decay mode				
Parameter	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$N_{B_s \to D_s\pi}$	15 091 ± 140	32 354 ± 200	24 158 ± 168	5341 ± 92	14 465 ± 154
Tot. $N_{B_s \to D_s\pi}$	total yield: 91 409 ± 346				
$N_{partReco}$	39 ± 25	67 ± 31	143 ± 26	8 ± 27	60 ± 29
N_{Comb}	5 789 ± 119	3676 ± 119	2 248 ± 90	3 137 ± 95	8 381 ± 144
$\mu(B_s^0)$ [MeV/c^2]	5 365.50 ± 0.06				
R_{B_s}	1.057 ± 0.004			1.073 ± 0.002	1.049 ± 0.001
$\mu(D_s)$ [MeV/c^2]	1 969.8 ± 0.02				
R_{D_s}	1.051 ± 0.004			1.04 ± 0.02	1.05 ± 0.01
$slope_{B_s}$ [(MeV/c^2)$^{-1}$]	-4.76 ± 0.30	-12.41 ± 0.57	-4.35 ± 0.85	-9.57 ± 1.47	0.53 ± 0.04
$fComb_{B_s}$	0.97 ± 0.04	0.76 ± 0.02	0.86 ± 0.11	0.39 ± 0.04	0.52 ± 0.04
$slope_{D_s}$ [(MeV/c^2)$^{-1}$]	-4.61 ± 0.96	-5.45 ± 0.41	-14.55 ± 2.14	-4.88 ± 1.57	-3.46 ± 1.43
$fComb_{D_s}$	0.47 ± 0.02	0.72 ± 0.02	0.65 ± 0.02	0.67 ± 0.03	0.75 ± 0.04
$RComb_{D_s}$	1.528 ± 0.043			1.611 ± 0.077	1.99 ± 0.092
$fComb_{PIDK}$	0.73 ± 0.01	0.69 ± 0.02	0.69 ± 0.02	0.92 ± 0.01	0.90 0.01

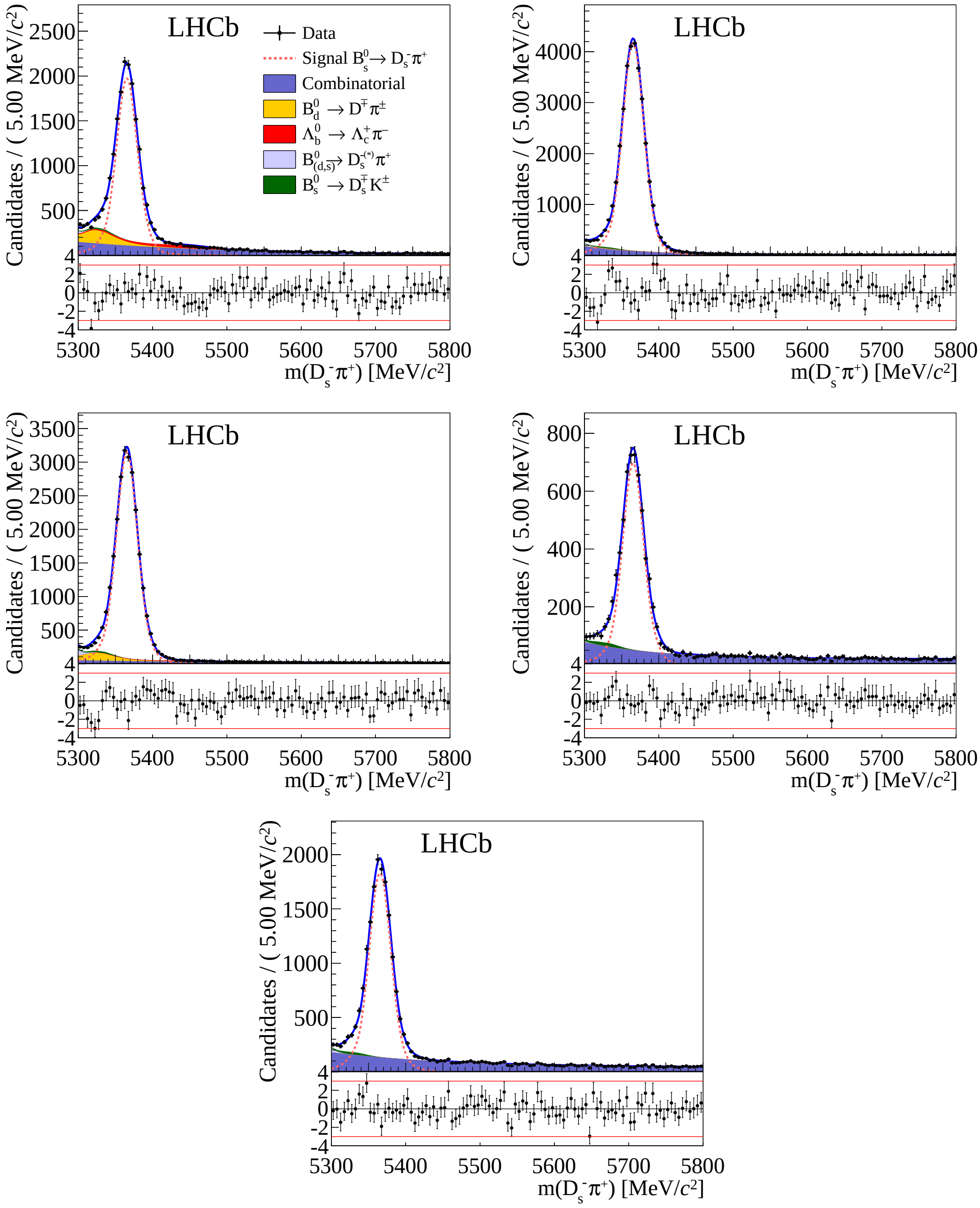


Figure 91.: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

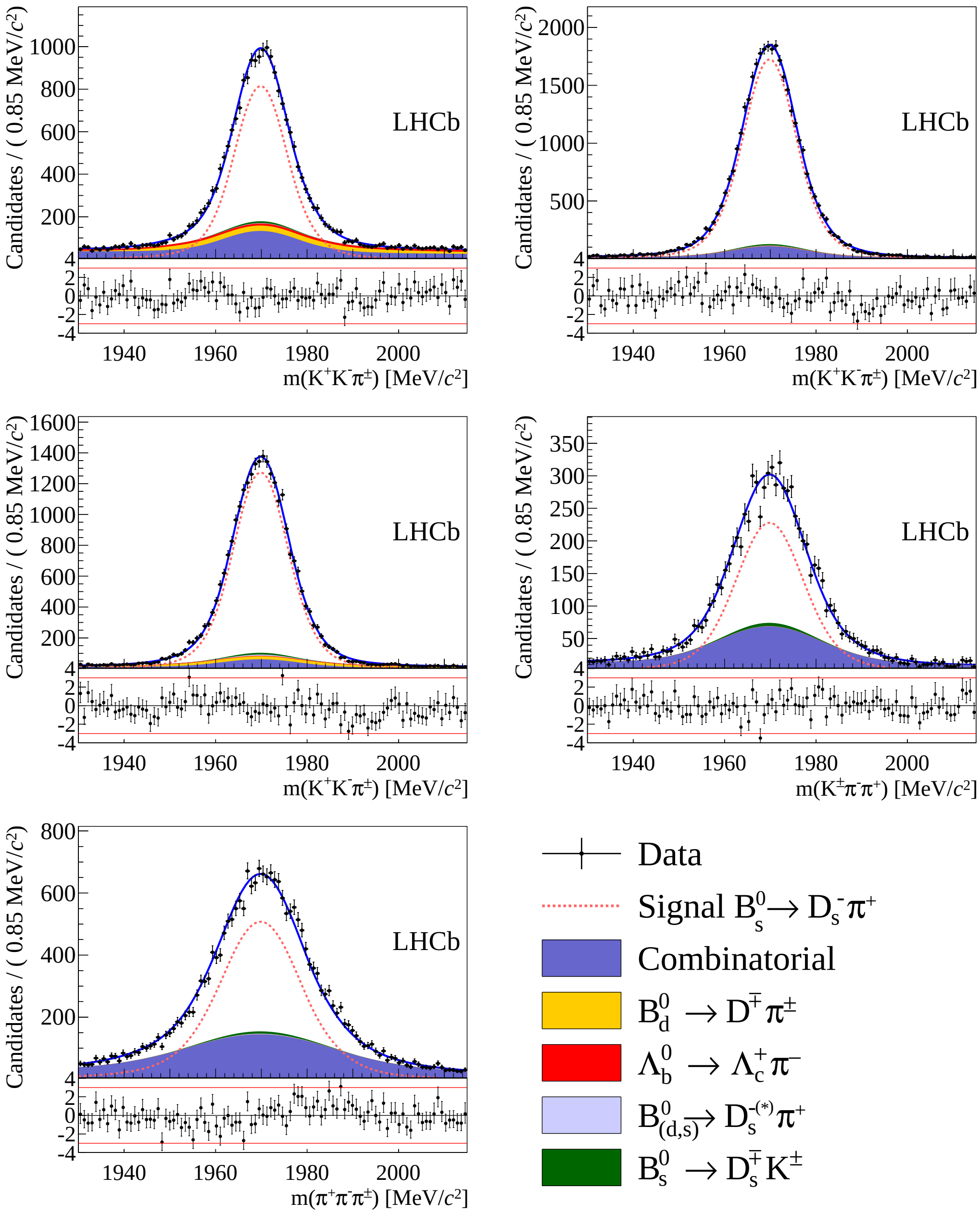


Figure 92.: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

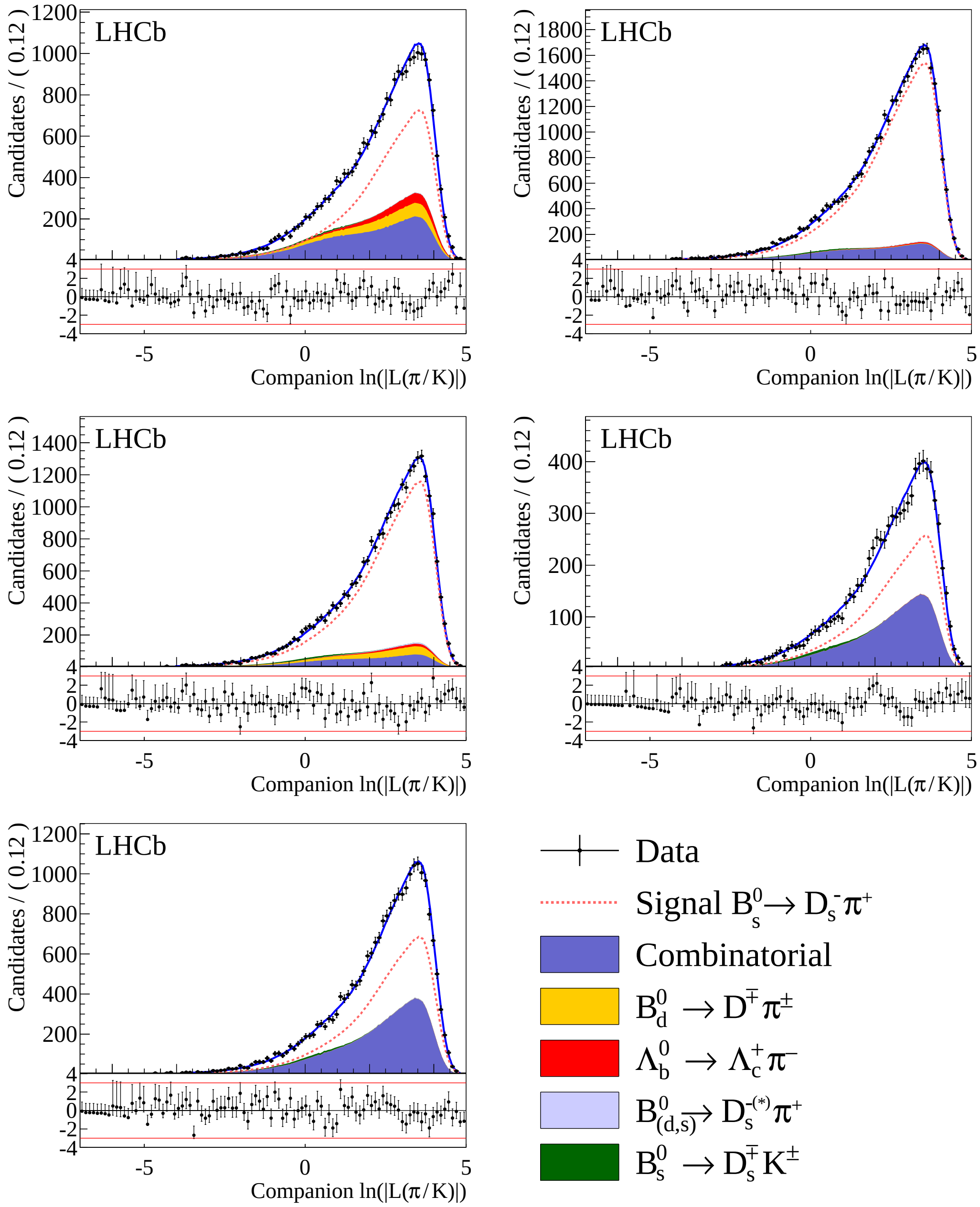


Figure 93.: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of $-log(PIDK)$ for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

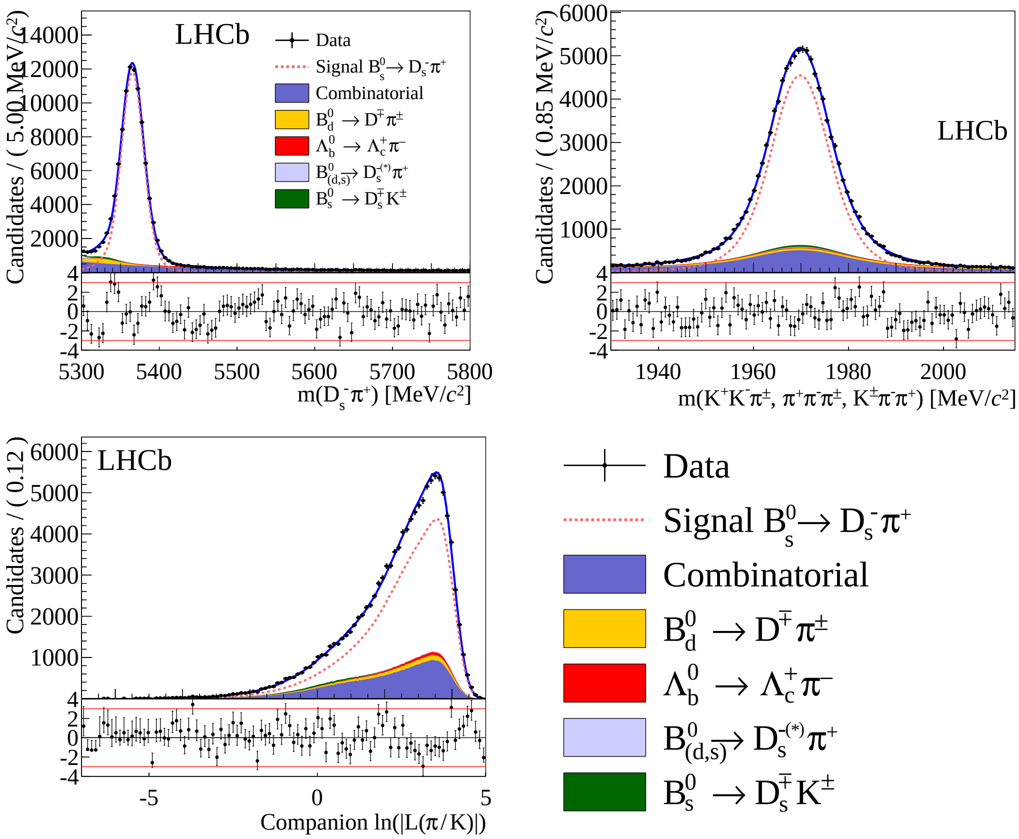


Figure 94.: The simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates for both polarities and all D_s final states combined for B_s^0 mass, D_s mass, $-log(PIDK)$ distributions.

Table 74.: Fitted values of the parameters for the $B_s^0 \to D_s^\mp K^\pm$ signal MD fit. Means are the parameters of the double Crystal Ball used to describe the signal in B_s^0 and D_s masses. The $slope_{B_s(D_s)}$ are the slopes of combinatorial background in 10^{-3} units. The fractions $fComb_{B_s}$ are the fractions between an exponential and flat distribution in B_s^0 mass, whereas $fComb_{D_s}$ are the fractions between an exponential and signal shape in D_s mass. $fComb_{PIDK(1,2)}$ are a fraction between pion, kaon and proton components in the $PIDK$ combinatorial shape, shared among all subsamples.

	D_s decay mode				
Parameter	$KK\pi_{nonres}$	$\phi\pi^-$	$K^{*0}K^-$	$K^-\pi^+\pi^-$	$\pi^-\pi^+\pi^-$
$N_{B_s\to D_sK}$	1 007 ± 39.0	1 873 ± 51	1 566 ± 46	407 ± 30	973 ± 45
$N_{B_s\to D_sK}$	total yield: 5825 ± 95				
$N_{B^0\to D_sK}$	21 ± 11	62 ±14	50 ± 12	13 ± 8	41 ± 13
$N_{bach\pi-p}$	820 ± 46	1 732± 60	1 268±52	263±32	1 040 ± 660
N_{Comb}	878 ± 55	634 ± 56	468 ± 50	968 ± 51	2 055± 66
$\mu(B_s^0)$ [MeV/c^2]	5 365.4 ± 0.2				
R_{B_s}	0.851 ± 0.016			1.014 ± 0.081	0.892 ± 0.046
$\mu(D_s)$ [MeV/c^2]	1 969.7 ± 0.1				
R_{D_s}	1.10 ± 0.01			1.31 ± 0.05	1.26 ± 0.03
$slope_{B_s}$ [(MeV/c^2)$^{-1}$]	-3.50±0.37	-1.79±0.51	-1.79±0.63	-1.15±0.34	-0.59±0.20
$slope_{D_s}$ [(MeV/c^2)$^{-1}$]	-2.49±1.34	-11.77±2.55	-6.04±1.99	-1.30±1.22	-4.41 ±1.28
$fComb_{D_s}$	0.57±0.04	0.32±0.04	0.30±0.04	0.41±0.04	0.47±0.03
$fComb_{\text{PIDK1}}$	0.13±0.03				
$fComb_{\text{PIDK2}}$	0.57±0.04				

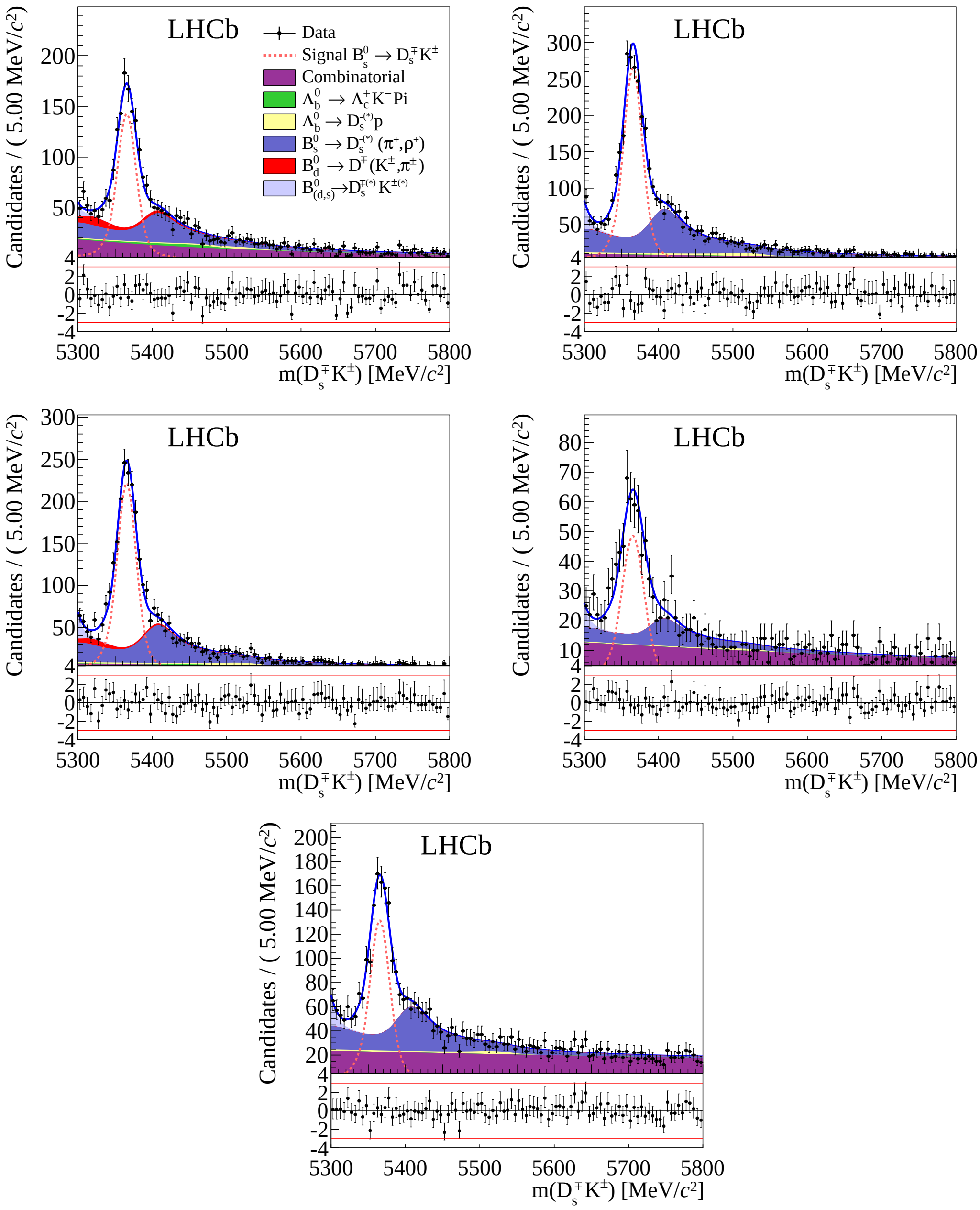


Figure 95.: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp}K^{\pm}$ candidates. Distributions of B_s^0 mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

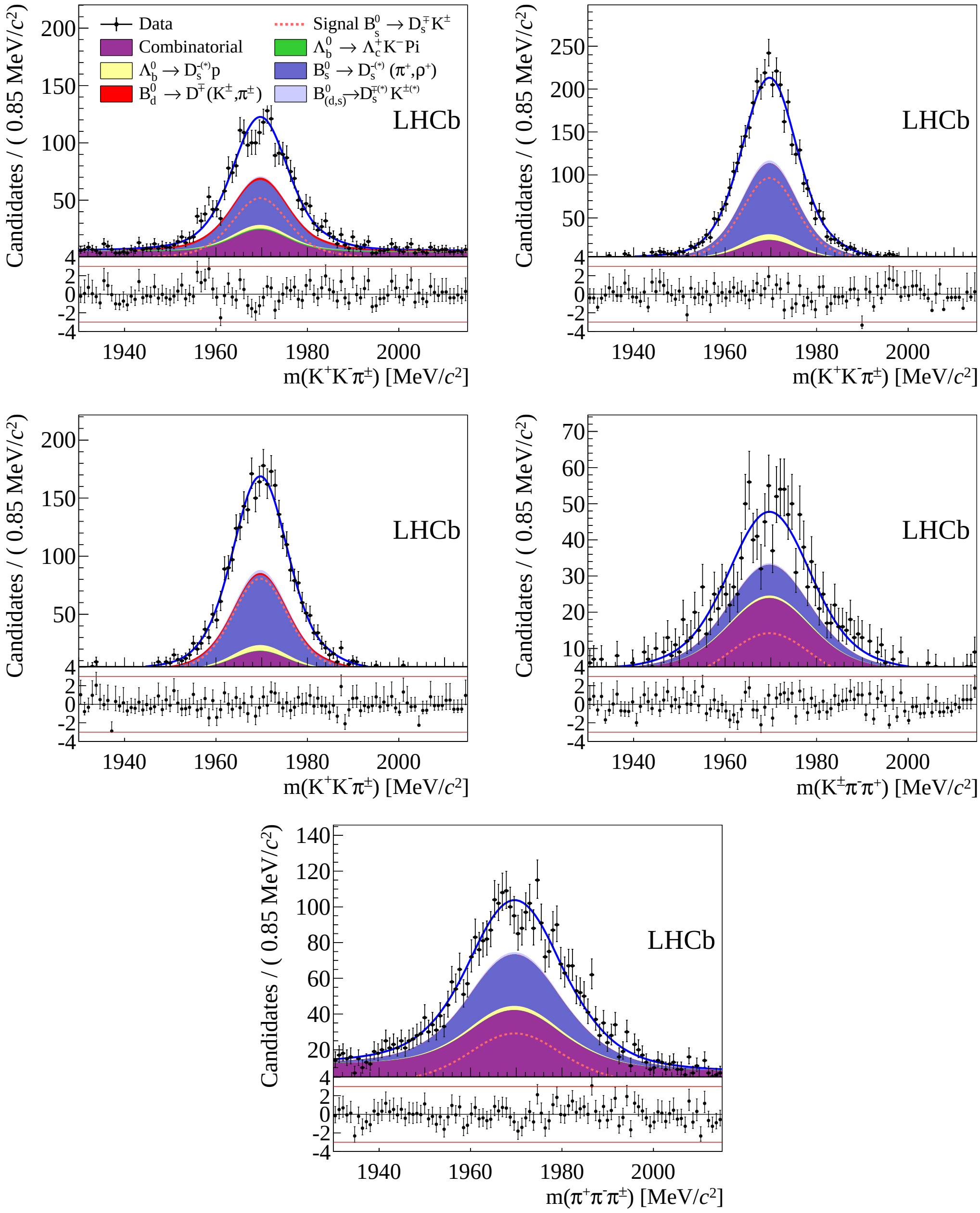


Figure 96.: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of D_s mass for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

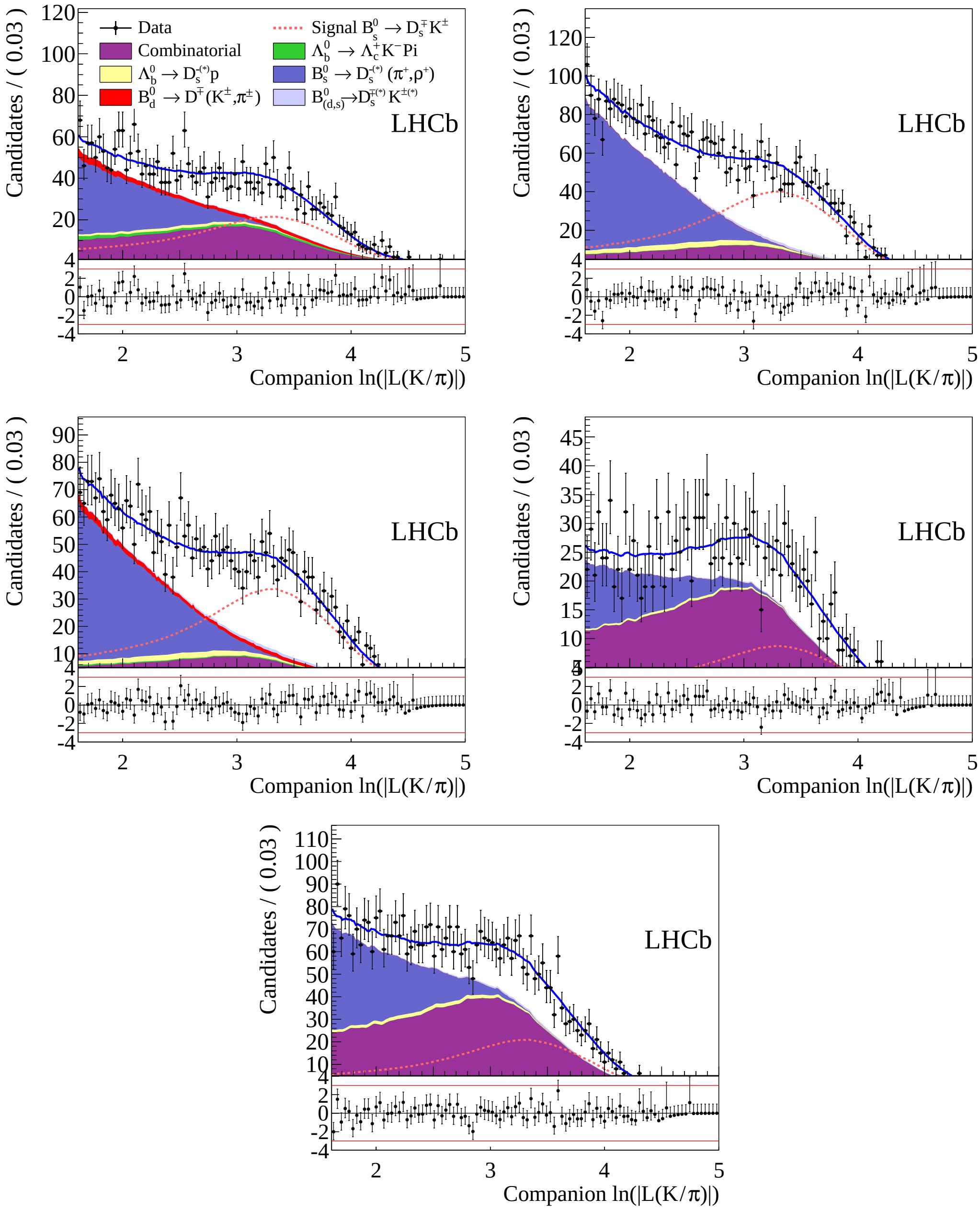


Figure 97.: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $log(PIDK)$ for D_s final states with combined magnet polarities, from top to bottom and left to right: $D_s^- \to (KK\pi)_{nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$, $D_s^- \to \pi^-\pi^+\pi^-$.

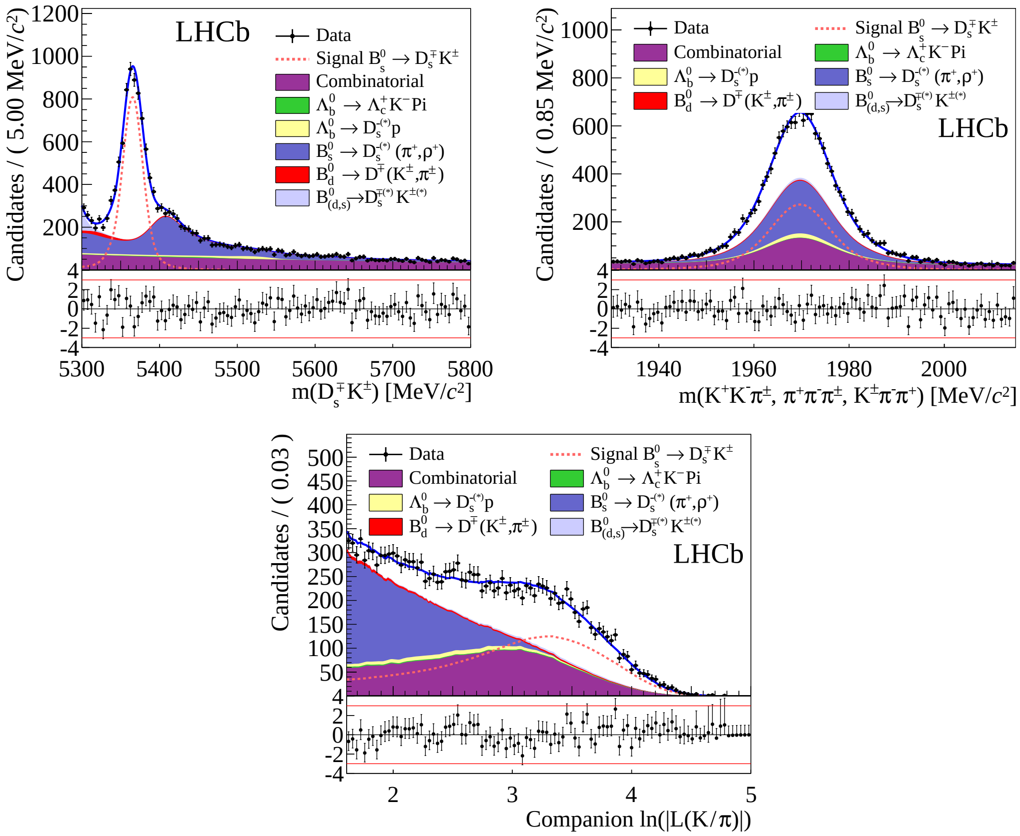


Figure 98.: The simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates for both polarities and all D_s final states combined, from top to bottom and left to right: B_s^0 mass, D_s mass, log(PIDK)distributions.

7.5 TIME SFIT

Several studies are on going to finalize the time-dependent part of the analysis whose main ingredients are the ones of the $1fb^{-1}$ analysis. Several improvements are foreseen:

- Time Resolution: it is used on a per-event basis like for the 1 fb^{-1} analysis. A scale factor that must be applied to correct the error estimate to match the proper time resolution. Updated studies on a consistent data samples of Run1 of prompt D_s decays are on-going. The reconstructed prompt D_s mesons associated with a random tracks to form fake B_s^0 candidates of which the true decay time is known and it is 0. The scale factor between estimated decay time error and measure time resolution is computed.

- Time acceptance: We plan to determine the time acceptance from data using a data-driven method that it is expected to reduce the systematics uncertainties related to the acceptance, the lifetime and $\Delta\Gamma_s$ value. This method is called swimming [79]. It is a data driven method to model per-event decay time acceptance and it is already used by other analysis of charmless decays. For each candidate which pass the selection (trigger, stripping and offline), the primary vertex position is shifted recursively ($\pm$ 600 mm) along the beam axis (z-axis) determining a different decay time. Then it is checked whether the candidate still pass the selection. From this procedure acceptance intervals in decay time are determined for each selected event and introduced as "observables" in the PDF description. In this way the determination of the acceptance function is not correlated to the knowledge of the B_s^0 lifetime and of the $\Delta\Gamma_s$ value, reducing the correlated systematics uncertainty which for $A^{\Delta\Gamma}_{f(\bar{f})}$ observables are $\simeq$ 0.43 the statistical uncertainty.

- Flavour Tagging: for the updated analysis we plan to calibrate and combine the tagging decision and predicted mistag of the NNetSSK and OS taggers during fitting. This approach is used also in other time-dependent analysis and has the advantage of simplify by a lot the error propagation.

Table 75.: Results of the OS and NNetSSK tagging calibration parameters p_0 and p_1 obtained on the $B_s^0 \to D_s^- \pi^+$ Run1 data sample.

	p_0	p_1	$\langle\eta\rangle$	ε_{tag} (%)
OS	0.3898 ± 0.0040	0.9907 ± 0.043	0.3688	37.4
NetSSK	0.4459 ± 0.0034	0.9617 ± 0.0509	0.4377	63.6

The tagging calibration parameters need to be updated with respect to the ones obtained in previous analysis since the data set is larger and different reconstruction algorithms are used so that also the tagging performance changes. Studies on the tagging calibration have been already done using the $B_s^0 \to D_s^- \pi^+$ data sample (see Chapter 5) by using a slightly different trigger, stripping and offline selection. Also the mass fit to extract *sWeights* is different from the one discussed in Chapter 5 indeed It has been used the Multi-dimensional fitter to the $B_s^0 \to D_s^- \pi^+$ data described in this chapter. To calibrate the NNetSSK and OS taggers it has been used a time sFit with preliminary time resolution Scale Factor while the time acceptance is still based on splines.

In Tab.75 the results of the tagging calibration for both the taggers are quoted: they are in agreement with what is reported in Chapter 5. The study on the systematic uncertainties of the calibration parameters can be inherited from what has been already discussed in Chapter 5. The final calibration will be re-evaluated once the time sFit will be in its final configuration. A first fit (sFit) to the decay time distribution of tagged $B_s^0 \to D_s^{\mp} K^{\pm}$ decays with preliminary time resolution, tagging calibration and time acceptance was performed. The results show that the statistical uncertainties on the γ angle scales according to expectations[4] with the increased statistics.

4 a factor $\sigma_{3fb^{-1}}(\gamma)/\sigma_{1fb^{-1}}(\gamma) \approx 1/\sqrt(3)$.

8 CONCLUSIONS

LHCb is successfully contributing to the measurements of CP violation in the sector of B-meson decays, even beyond the initial expectations. This thesis has documented the relevance of the LHCb contribution to the measurement of the γ angle of the Unitarity Triangle. In addition to established reference decay modes for measuring γ, several new channels and methods have been used for the first time. Among these are multi-body B decays and time-dependent tagged analysis of $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays. A combination of the available measurements of γ with Run1 data already provides the most accurate determination from a single experiment. The success of LHCb in this field is due to the high $b - \bar{b}$ cross-section available at LHC energies and to the specific features of the LHCb experiment in reconstructing B meson decays. Among these are the dedicated trigger, mass and vertex resolutions and the excellent particle identification capabilities. For the tagged time-dependent analysis a crucial aspect is the possibility to identify the initial flavour of the neutral B mesons. LHCb has developed several Flavour Tagging algorithms which show good performances despite difficulties in identifying tagging particles related to the hadronic environment of the pp collider. Particular relevance for the CP violation measurements in B_s^0 decays is brought by the Same Side Kaon tagger. Such tagging algorithm is based on the use of Neural Networks classifiers and it has been developed and optimized in the $B_s^0 \rightarrow D_s^- \pi^+$ decay mode. An original contribution of this thesis is the study of the performance and the calibration of the NNetSSK tagger in $B_s^0 \rightarrow D_s^- \pi^+$ decays through a time-dependent analysis of the $B_s^0 - \overline{B}_s^0$ mixing. The measured tagging power is $\epsilon_{eff} = (1.80 \pm 0.19(stat) \pm 0.17(syst))\%$.

The results of the $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ analysis using 1 fb^{-1} of data are presented in this thesis. The analysis is based on a sophisticated multi-dimensional fitter to discriminate signal from background and two fit approaches to determine the

CP observables from the decay time distribution of tagged events exploiting the NNetSSK tagger. The value of γ is found to be $(115^{+28}_{-43})^{\circ}$ and contributes to the combination of γ measurements evaluated by LHCb [10], [9]. The updated analysis of full 3fb^{-1}, Run1 data set of this decay mode has started. The expected precision improvements due to larger statistics required detailed preparatory studies and tunings of the analysis strategy documented in this thesis. The analysis needs to be finalized and several improvements are on-going.

A

APPENDIX

Contents

A.1 Study on CP sensitivities **197**
A.2 Validation of the $B_s^0 \to D_s^- \pi^+$ fit with toys and bootstraps studies **202**
A.2.1 Toys studies 202
A.2.2 Bootstrap studies 205
A.3 TGenPhaseSpace-based Emulator **207**
A.3.1 Events generation 207
A.3.2 Events reconstruction and selection 208
A.3.3 Background emulation 212
A.3.4 Correlations among observables 213
A.4 Toys Study for ϵ_{tag} versus time **220**
A.4.1 Pull distributions 220
A.4.2 Toy by toy difference 220

A.1 STUDY ON CP SENSITIVITIES

Results from the MDFitter.

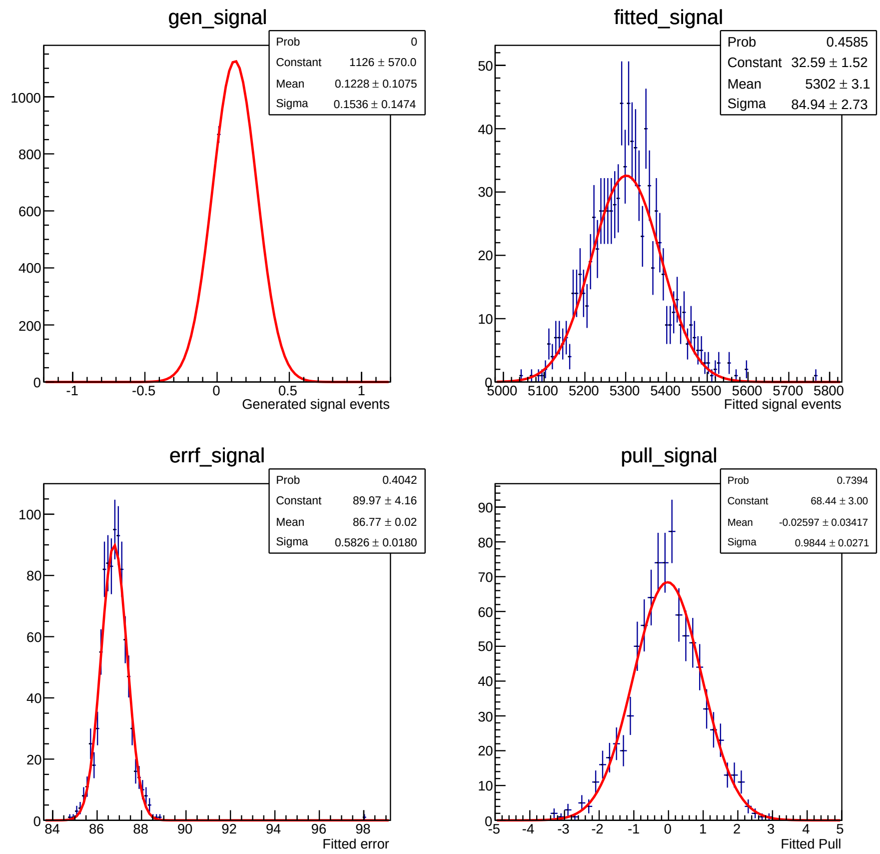


Figure 99.: The signal yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported.

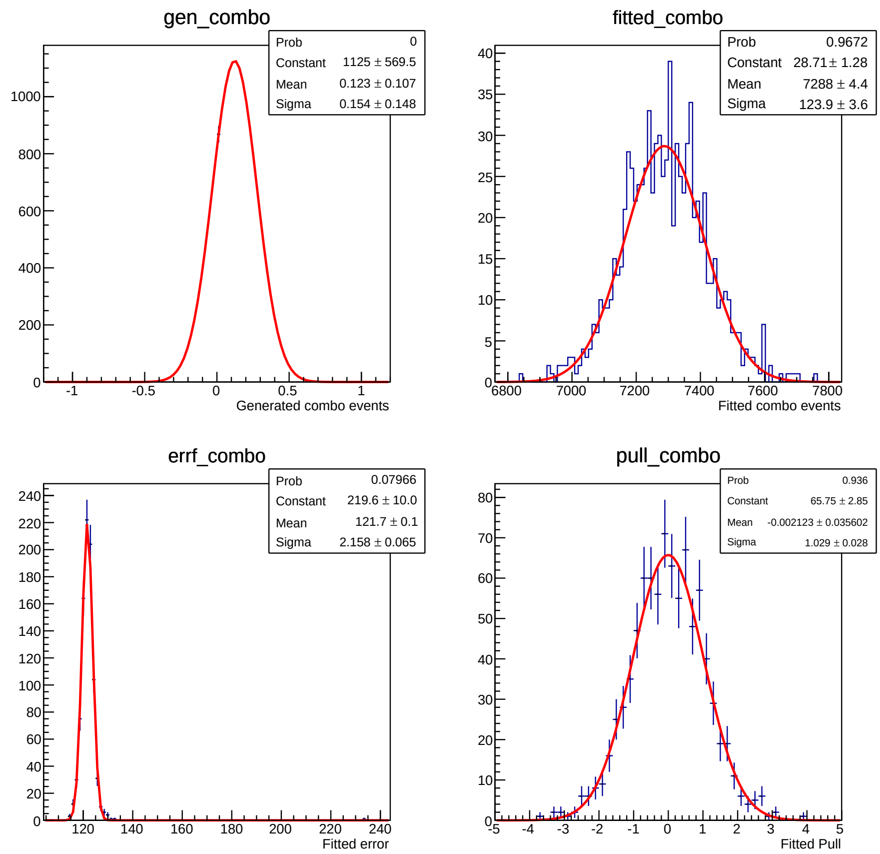


Figure 100.: The combinatorial yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported.

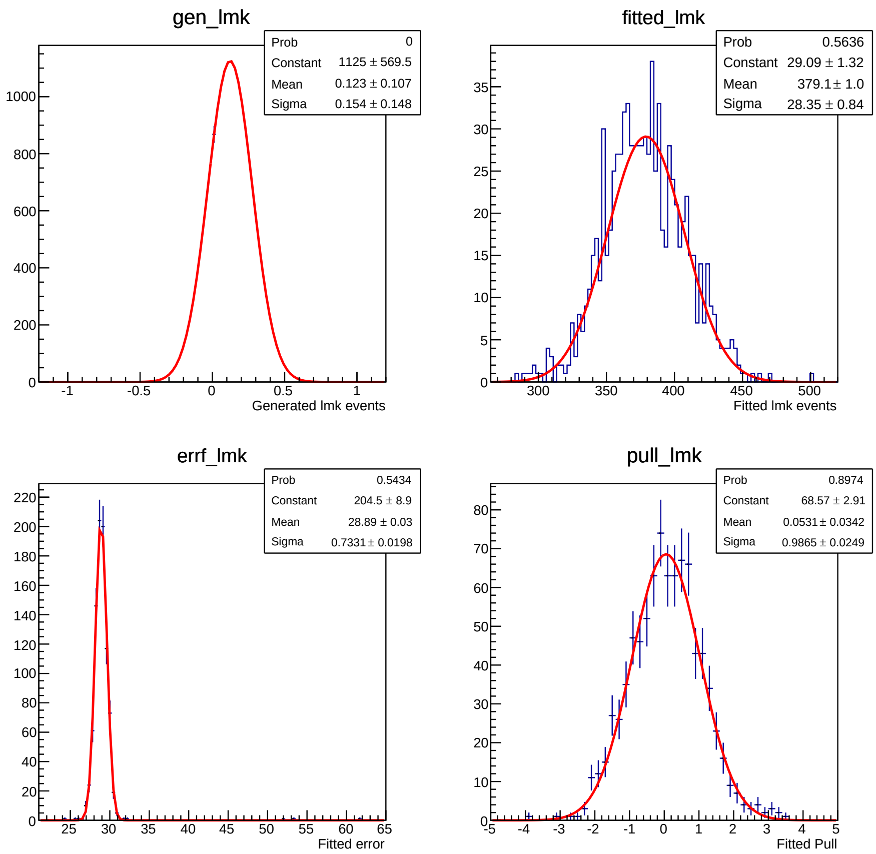


Figure 101.: The $B^0 \rightarrow D_s K$ yields distribution (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported.

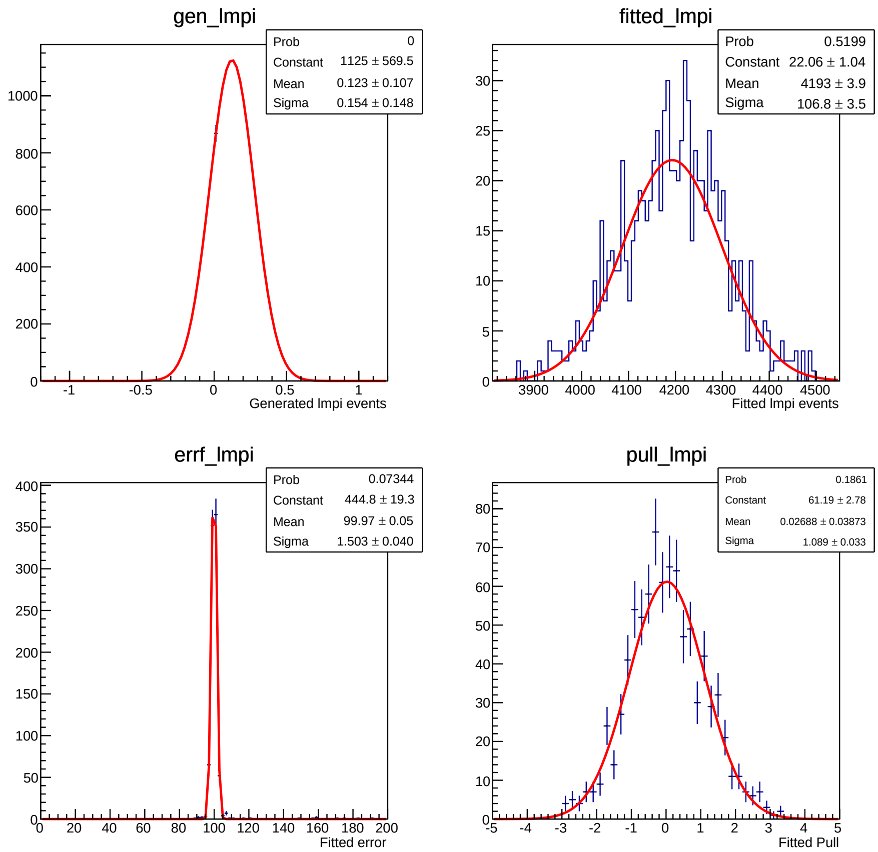


Figure 102.: The overall yields distribution of $B_s^0 \to D_s^{(*)}\rho(\pi)$, $\Lambda_b \to D_s^{(*)}p$, $B_s^0 \to D_s^-\pi^+$ backgrounds (top-right), the corresponding errors distribution (bottom-left) and the corresponding pull distribution obtained from the 930 pseudo-experiments are reported.

A.2 VALIDATION OF THE $B_s^0 \to D_s^- \pi^+$ FIT WITH TOYS AND BOOTSTRAPS STUDIES

In order to verify that the sFit to $B_s^0 \to D_s^- \pi^+$ data provides the correct values of the flavour tagging parameters and that the fit to data is under control describing all the relevant sources of backgrounds further studies have been done.

A.2.1 *Toys studies*

We check that the fit of the $B_s^0 \to D_s^- \pi^+$ decay provide a reliable determination of the calibration parameters, and a proper estimation of the parameters' uncertainties, by performing studies with fast simulations (toys), as follows:

- we generate 500 samples, each simulating the full statistics of Run I data. For the generation of the samples, we use the *B2DXFitter* code developed for the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis [54], in order to be able to produce signal and backgrounds candidates for all the variables used in the fit (mass, decay time, decay time error, tag decision, per-event mistag). We make sure that the signal and background templates of the mass distributions used in the generation of the candidates are the same used in the maximum likelihood fit to extract the sWeights.
- We fit the samples by applying the same procedure used for fitting the data:
  - a mass fit in the [5100, 5600] MeV range to determine the fractions of background in the [5300, 5600] MeV range
  - a mass fit in the [5300, 5600] MeV range with fractions of background fixed (signal and total background can float) to determine sWeights
  - a time fit of the sWeighted events to extract the calibration parameters.

We check then the distributions of the fitted parameters (θ_i^{fit}), of their estimated uncertainties (σ_{θ_i}), and of the pull, defined as:

$$p_{\theta_i} = \frac{\theta_i^{\mathrm{input}} - \theta_i^{\mathrm{fit}}}{\sigma_{\theta_i}}, \tag{A.1}$$

	$p_0 - \langle\eta\rangle$	p_1
input values	0	1
average fit value and error	0.0001 ± 0.0049	1.006 ± 0.068
Pull mean	-0.048 ± 0.045	0.057 ± 0.050
Pull sigma	0.936 ± 0.049	1.025 ± 0.042

Table 76.: Toys studies results for the calibration parameters; 500 samples generated.

	$p_0 - \langle\eta\rangle$	p_1
input values	0	1
average fit value and error	0.000 ± 0.0046	0.999 ± 0.053
Pull mean	0.03 ± 0.05	0.08 ± 0.05
Pull sigma	0.94 ± 0.04	0.92 ± 0.04

Table 77.: Toys studies results for the calibration parameters; 500 samples generated (signal only).

where θ_i^{input} is the value of the parameter used in the generation of the candidates. The results of the toys studies for the calibration parameters are reported in Tab. 76 and in Fig. 103: the pull distribution have Gaussian shapes, centred at zero, with unitary widths; the estimates of the calibration parameters are unbiased.

As a further check, we repeat the toys studies by simulating only signal events (thus avoiding the sWeights procedure). The results are still unbiased, as reported in Tab. 77 and Fig. 104

As expected the errors reported in Tab.77 are lower than the errors in Tab.76 since they are evaluated without the contamination of backgorund events.

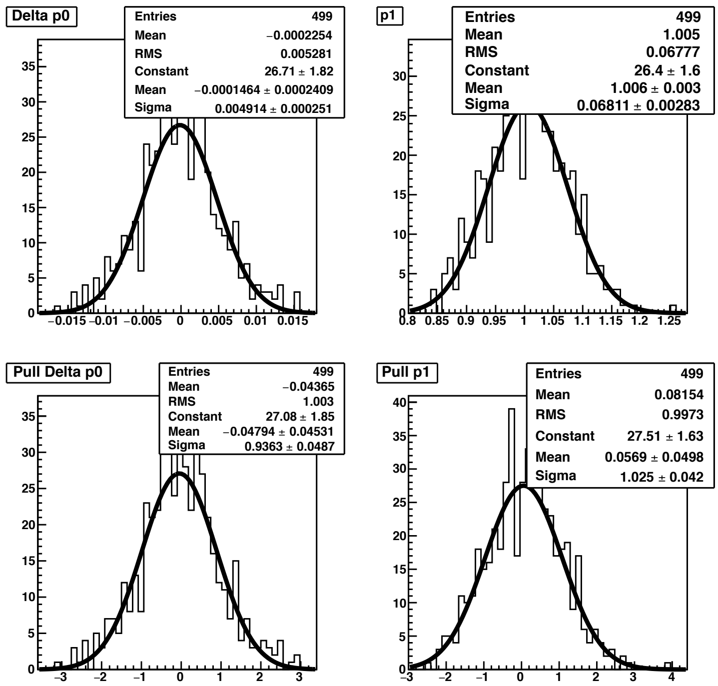


Figure 103.: Toys studies: distributions of the fitted values and of the pulls for $p_0 - \langle\eta\rangle$ and p_1.

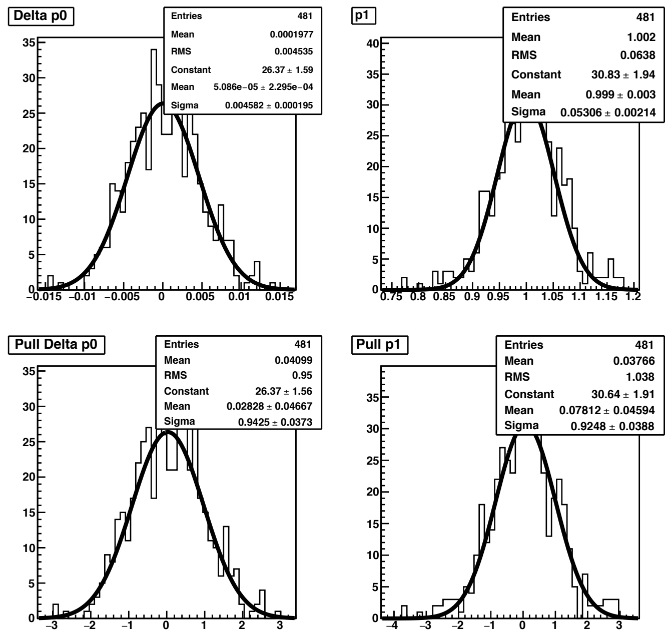


Figure 104.: Toys studies (signal only): distributions of the fitted values and of the pulls for $p_0 - \langle\eta\rangle$ and p_1.

A.2.2 *Bootstrap studies*

In order to verify that the fit to $B_s^0 \to D_s^- \pi^+$ data is under control, in particular that there aren't sources of backgrounds which are not considered it has been decided to use the bootstrap [80] method to produce 10K sample of $B_s^0 \to D_s^- \pi^+$ events with the same statistic of the full Run I data set. Each sample is fitted with the same procedure used in the analysis. The distribution of the fitted parameters, $p_0 - \langle\eta\rangle$ and p_1, and of their fitted errors are shown in Fig. 105. The results are summarised in Tab. 78: the RMS of the distribution of the fitted parameters is in agreement with the mean of the distribution of the fitted error. Thus we can conclude that the fit to the $B_s^0 \to D_s^- \pi^+$ data set is reliable and every background component has its correct description in the fit model.

	$p_0 - \langle \eta \rangle$	p_1
nominal result	0.0052 ± 0.0044	0.977 ± 0.070
average fitted value	0.0061	0.967
RMS of fitted value	0.0044	0.068
average fitted error	0.0044	0.069

Table 78.: Bootstrap studies results for the calibration parameters; 10K bootstrapped samples.

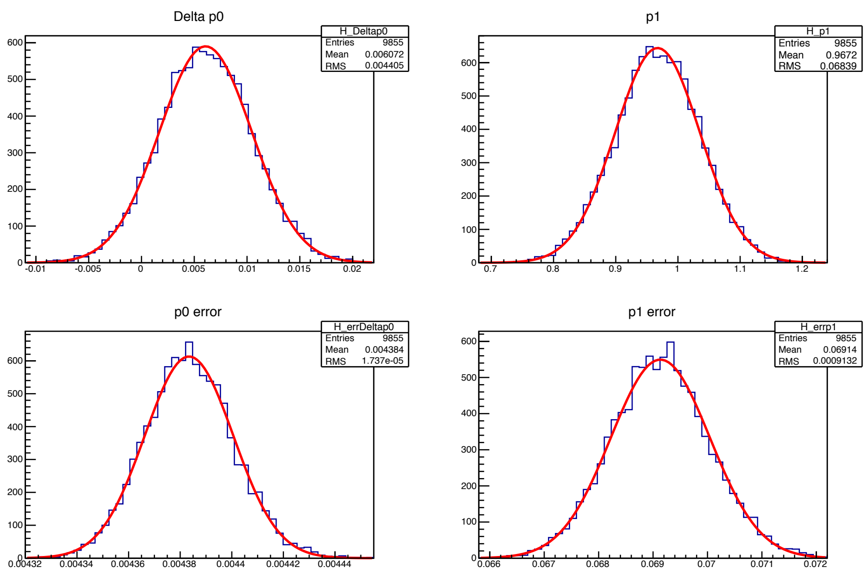


Figure 105.: Bootstrap studies (10K samples): distributions of the fitted values and of the fitted error of $p_0 - \langle \eta \rangle$ and p_1.

A.3 TGENPHASESPACE-BASED EMULATOR

Correlations among the observables involved in the fitter are possible and in principle they can be different from signal to any other background decays. In previous analysis the correlations among the observables have been neglected, since in MC they are found to be small. In the current analysis a more accurate estimation is desired in order to account for the correlations in the fitter itself if they are found to be relevant or in the systematics uncertainties if it turns out they are of small size. Given the limitations in producing large samples of MC events that would be required by a deeper study of such effect, an original approach has been developed and implemented. The idea is to generate samples of the main backgrounds decays, to emulate the detector reconstruction, to apply the core of the selection cuts and to evaluate the correlation between each pair of observables used in the fitter.

The reconstruction of the events is done under the $B_s^0 \to D_s^{\mp} K^{\pm}$ hypothesis, in the the $D_s \to KK\pi$ non resonant final states, but can be easily extended to other decay modes.

A.3.1 *Events generation*

A very simple way to generate large samples of n-body decays is to use the TGenPhaseSpace class [78] provided by the ROOT package [39]. This class allows to generate in the phase space each kind of decay starting from the knowledge of mother particle's kinematics. The momentum P, the pseudorapidity η and the Φ angle of the initial particles are extracted randomly from the relative MC distributions. In this way the four-vector of each mother particle is defined, $\vec{P}_i = (p_{T,i}, \eta_i, \Phi_i, M_i)$, and used together with the masses of the daughter particles to generate any kind of n-body decays. The TGenPhaseSpace returns the four-vector of each daughter particles and the weight of the event, and this information are used as kinematic input to generate daughter particles decays. The class is used for generating signal events $B_s^0 \to D_s K$ and the output of the first decay is used to generate $D_s \to KK\pi$. Moreover the same procedure has been repeated to generate samples of several other decays:

- $B_s \to D_s\rho$, $D_s \to KK\pi$, $\rho \to \pi\pi^0$
- $B_d \to DK$, $D \to K\pi\pi$
- $\Lambda_b \to D_s^* p$, D $_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $\Lambda_b \to D_s p$, $D_s \to KK\pi$
- $\Lambda_b \to \Lambda_c\pi$, $\lambda_c \to Kp\pi$
- $\Lambda_b \to \Lambda_c K$, $\lambda_c \to Kp\pi$
- $B_s \to D_s^*(\gamma)K$ $D_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $B_s \to D_s^*(\gamma)\pi$, $D_s^* \to D_s\gamma$, $D_s \to KK\pi$
- $B_s \to D_s\pi$, $D_s \to KK\pi$
- $B_d \to D\pi$, $D \to K\pi\pi$
- $B_d \to D_s K$, $D_s \to KK\pi$

which are the main backgrounds of the $B_s^0 \to D_s^{\mp}K^{\pm}$ decays. Intermediate resonances and effects related to the spin of the particles that might change the angular distributions are not considered in the generation of the decays.

A.3.2 *Events reconstruction and selection*

The generation tool is simple and easy to be implemented but it only provides information on the considered decay and doesn't fully reproduce $p-p$ collision event that is characterized by tens of additional particles. Parametric functions are used to emulate the effect of the detector reconstruction and of the signal selection. Only the combinatorial background is not investigated with this method since the information of the whole event is not available. For the same reason, the observables related to the flavour tagging tool can not be emulated from first principles but only reproduced according to the distributions found in data. In this way they are generated independently from all the other observables involved in the fitter and so it is not possible to evaluate correlations for them. The observables used for the $B_s^0 \to D_s^{\mp} K^{\pm}$ time-dependent analysis and emulated for the correlations study are: $m(D_s K)$, $m(D_s(KK\pi))$, PIDK of the bachelor, the B_s decay time t and its error δ_t.

The effect of the detector reconstruction is emulated by smearing the kinematic variables of the particles in the final state, namely the candidate bachelor K and the K, K, π from the D_s decays. Several effects are considered:

1. The smearing of the modulo of the momentum is evaluated following this formula [81] extracted from previous studies on data:
$$p \cdot (0.35 + 0.0017p)/100. \tag{A.2}$$

2. The smearing of the direction of the momentum has been extracted from the signal MC in the distributions of $P_x/P_z - P_x^{true}/P_z^{true}$ and $P_y/P_z - P_y^{true}/P_z^{true}$ for each final particle. The smearing of the momentum direction turns out to be crucial to correctly reproduce the mass resolution of B_s and D_s distributions.

3. The information on the particle ID, which is widely used in the analysis, has been emulated using a data-driven method based on control samples of kaons, pions and protons and the tools provided by the PIDCalib package [56]. From the PIDK distributions for kaons, pions and protons as function of the momentum and η of the considered particle a value of PIDK is extracted.

The D_s candidate is reconstructed from the three final particles. After a kinematic fit in which the mass constraint on the D_s is applied, the B_s candidate is formed.

A second step of the reconstruction regards the information related to the B_s decay time as well as the information on the decay time error. The decay time is generated using an exponential function with the B_s lifetime. From the generated decay time is possible to evaluate the true flight distance:
$$d = \frac{t \cdot P \cdot c}{M_{B_s}} \tag{A.3}$$
where t and P are respectively the generated values of the decay time and momentum, while for M_{B_s} it is used the PDG value. The uncertainty on the flight

distance measurement is estimated using MC. Figure 106 shows the dependence of the flight distance error on the measured momentum and flight distance and it is used to extract a first estimation of the error on the flight distance. The corre-

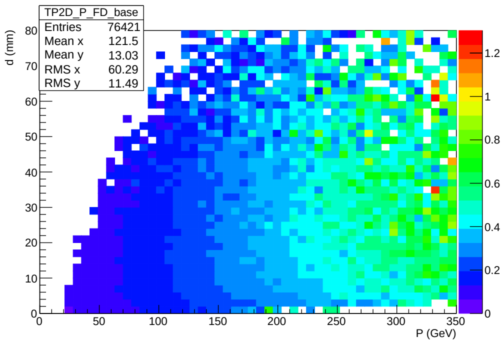


Figure 106.: Flight distance error distribution in terms of the measured momentum and flight distance of the B_s meson in MC sample.

lations between the momentum and the flight distance are included in the flight distance error estimation which has been corrected to match the flight distance resolution. The obtained flight distance resolution is used to smear the flight distance emulating the effect of the detector on the reconstructed value of the decay time:

$$t_R = \frac{M_{B_s}^R \cdot d_R}{P_R c} \tag{A.4}$$

where d_R, P_R and $M_{B_s}^R$ are rispectively the reconstructed flight distance, momentum and mass of the B meson. Propagating the error on the observable beforehand mentioned, the error estimation on the decay time is obtained:

$$\delta t_R = t_R \sqrt{(\frac{\sigma d_R}{d_R})^2 + (\frac{\sigma P_R}{P_R})^2 + \rho \frac{\sigma d_R}{d_R} \frac{\sigma P_R}{P_R}} \tag{A.5}$$

where $\sigma P_R = P_R - P_G$ and ρ = 0.1 taken from MC.

In the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis, the core of the selection consists of a BDTG selection followed by PID requirements and vetoes depending on the final state. To each final particle a cut for the geometrical acceptance is applied with the pseudorapidity η which ranges from 1.8 to 5.0. For the B_s (D_s) mass distributions events which fall in the mass window 5300-5800 (1930;2020) MeV are accepted. The BDTG response can not be reproduced as a whole because many of the input variables related to the p-p collision are not provided by the TGenPhaseSpace generator. For this reason parametric selection functions have been obtained fitting the ratio between the reconstructed momentum (transverse momentum) distribution in MC and the momentum (transverse momentum) distribution produced as output of the emulator. The selection function is applied to one particle at a time and recursively applied to all of them. The level of agreement between the MC distributions and the emulated ones can be improved repeating this procedure several times. In fact a correlation among the D_s daughters momentum distributions is found so that applying the selection function to one particle affects the momentum distribution of the other D_s daughter. In Fig. 107 the momentum and transverse momentum distributions are reported applying selection functions depending on the momentum to K bachelor and the two Kaons from D_s decay. The level of agreement shown by these plots is considered acceptable and reasonable for the purpose of this study. This aspect is very channel-dependent and in case of a cut-based selection it could be emulated in easier ways. The last part of the selection consists in the PID requirements. PID criteria for the final particles have been applied according to what is done in the analysis. The PID requirements for the bachelor Kaon and for the two Kaons of the $D_s \to KK\pi$ non resonant final state is PIDK > 5. Particles which do not satisfy the PIDK > 5 requirement are considered Pions and the event is rejected if two Kaons and one Pion are not found. To account for the effects of the selection efficiency functions and of the PID requirements on the decay time, an acceptance function depending on the flight distance of the B_s meson has been applied.

The whole procedure has been devoloped and cross-checked with the signal MC sample comparing the invariant mass distributions of B_s and D_s as well as the decay and the decay time and its error. In Fig. 108 the invariant mass distribution for the B_s and D_s hadrons are shown with the corresponding distribution from MC. Fitting the mass distribution with a simple gaussian the mass resolution for B_s (D_s) is 0.01391 (0.00545) GeV for the emulated events and 0.01389 (0.00606) GeV in the MC sample. The level of agreement between the emulated and the MC distributions is satisfactory in particular for the masses resolution and the decay time distributions. Considering the level of agreement between signal distributions of emulated and simulated events satisfactory, all the background

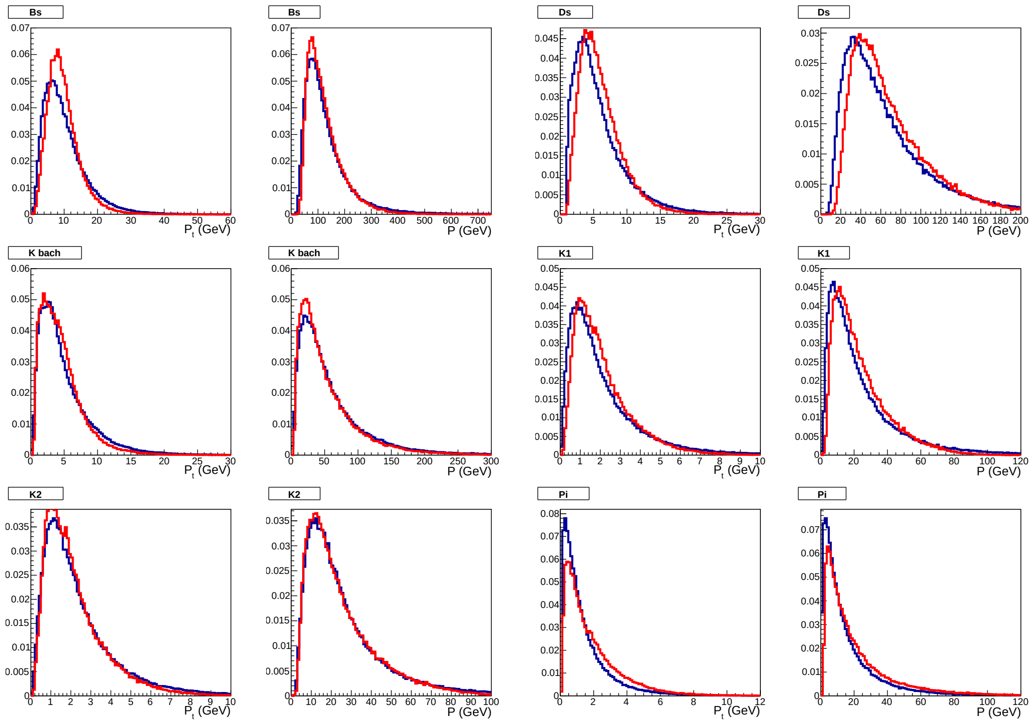


Figure 107.: Kinematic distribution for B_s, D_s, $Kbach$, $K1$, $K2$ and Pi in red from MC and in blue from the output of the Emulator for signal $B_s^0 \to D_s^{\mp} K^{\pm}$ events.

decays can be generated and selected with the described emulator, modifying only the generation step.

A.3.3 *Background emulation*

The main backgrounds which contribute to the $B_s^0 \to D_s^{\mp} K^{\pm}$ selected events have been generated and selected applying exactly the same procedure used for signal events. The kinematic input used to generate B_d decays are the same used for B_s while for Λ_b devays there are some differences and the proper Λ_b kinematic distributions are extracted from MC. In Fig. 109 the momentum and η distributions obtained from MC are compared for B_s, B_d and Λ_b hadrons. In addition to the Flavour Tagging observables, generated independently and calibrated with proper calibration parameters for B_d and B_s, also the mixing effect is described. In Tab.

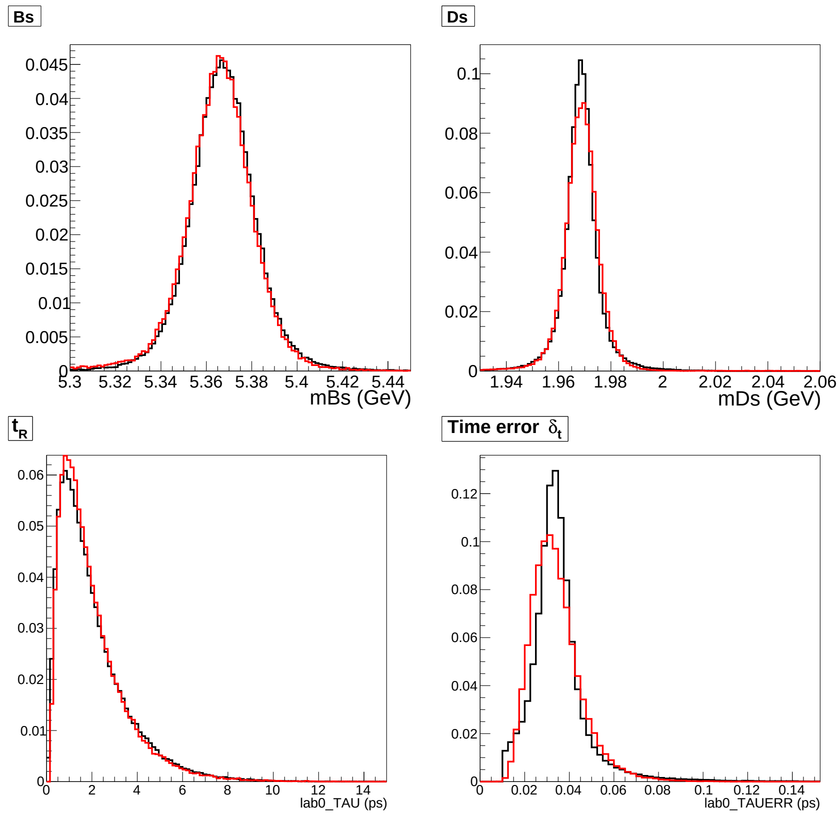


Figure 108.: Mass and decay time distributions for B_s and D_s, in red from MC and in blue from the output of the emulator for signal $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ events.

79 the background decays with the expected contribution to the 3 fb^{-1} analysis are listed. In Fig. 111 the B_s and D_s mass distribution as well as the PIDK of the bachelor particle are reported for signal and backgrounds decays.

A.3.4 *Correlations among observables*

Scatter plots among each pair of observables are reported from Fig.112-121. There aren't evidences of strong correlations.

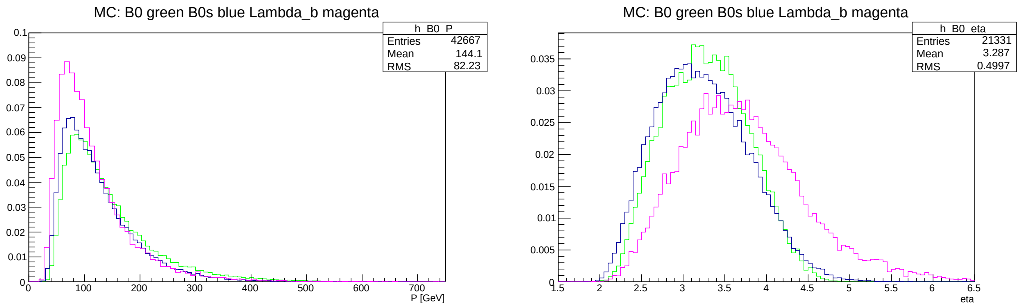


Figure 109.: Comparison among B_s, B_d and Λ_b kinematic distributions from MC samples. Left: P, right: η.

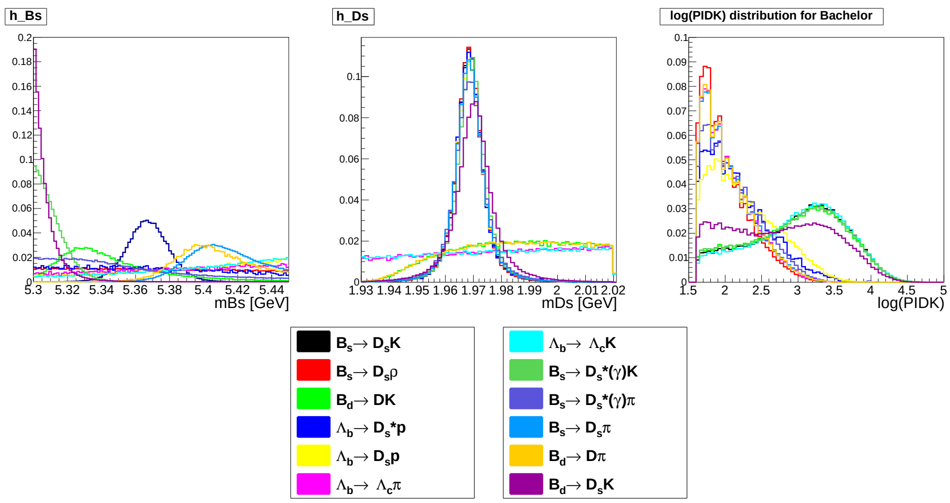


Figure 110.: Emulated bakgrounds decay distributions for B_s mass, D_s mass and PIDK for the bachelor particle.

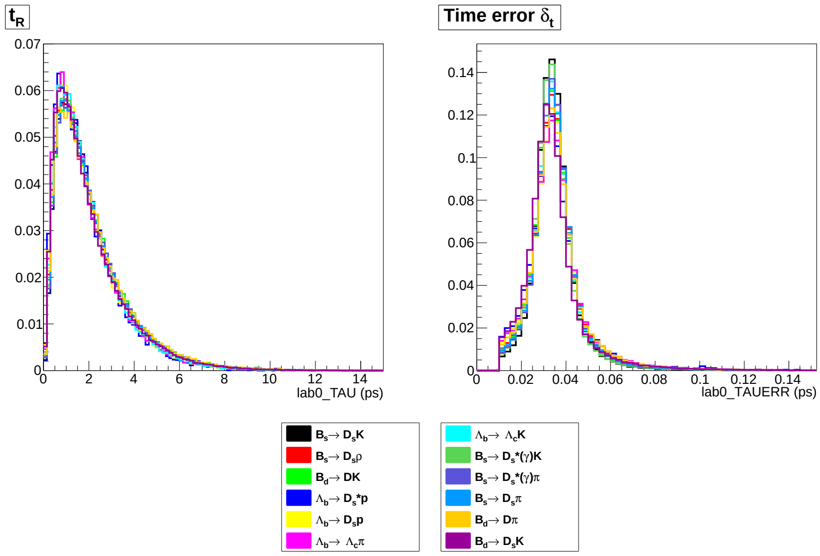


Figure 111.: Emulated bakgrounds decay distributions for B_s decay time and its error.

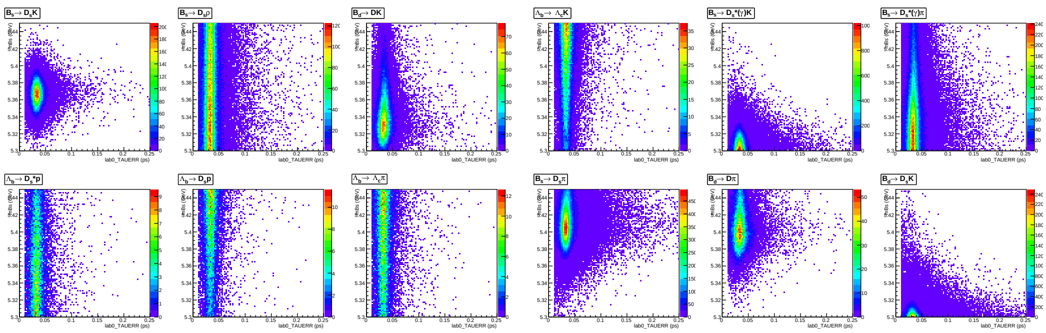


Figure 112.: Scatter plots B_s mass versus δ_t for each emulated decay starting from the signal.

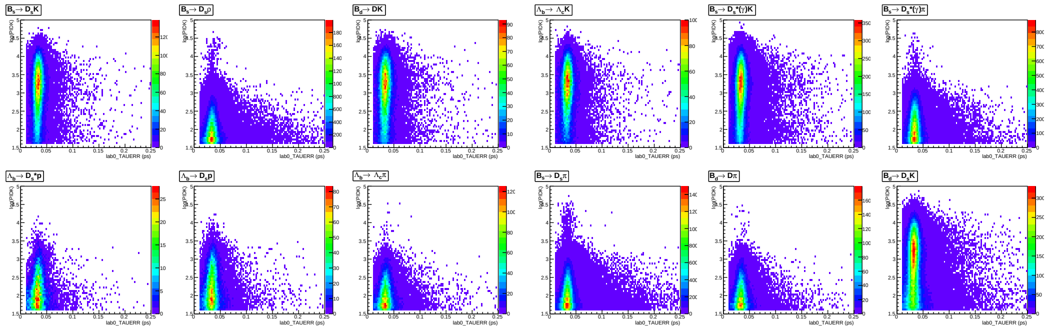


Figure 113.: Scatter plots PIDK of bachelor particle versus δ_t for each emulated decay starting from the signal.

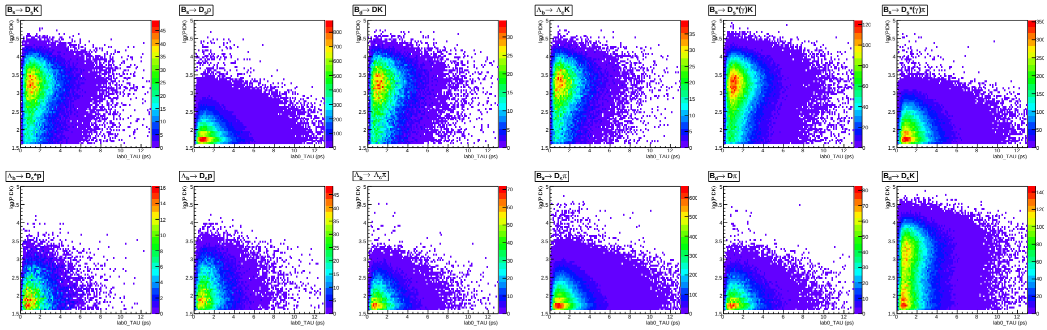


Figure 114.: Scatter plots PIDK of bachelor particle versus t_R for each emulated decay starting from the signal.

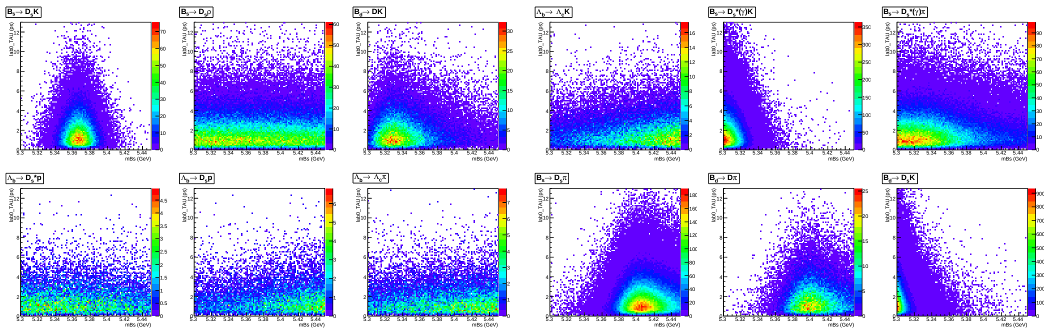


Figure 115.: Scatter plots t_R mass versus B_s for each emulated decay starting from the signal.

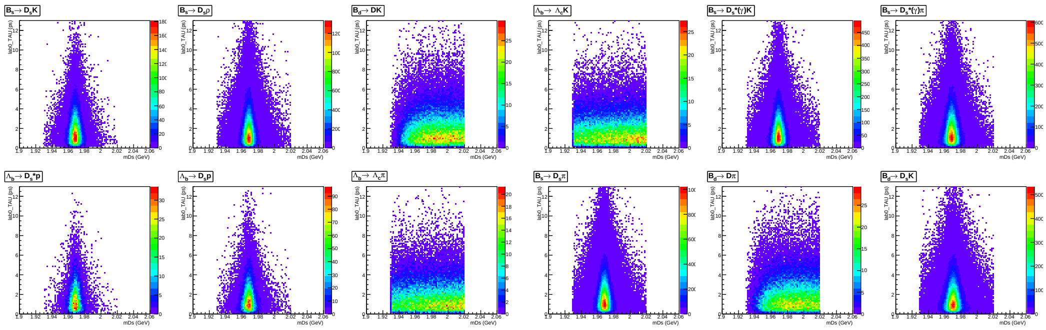


Figure 116.: Scatter plots t_R mass versus D_s for each emulated decay starting from the signal.

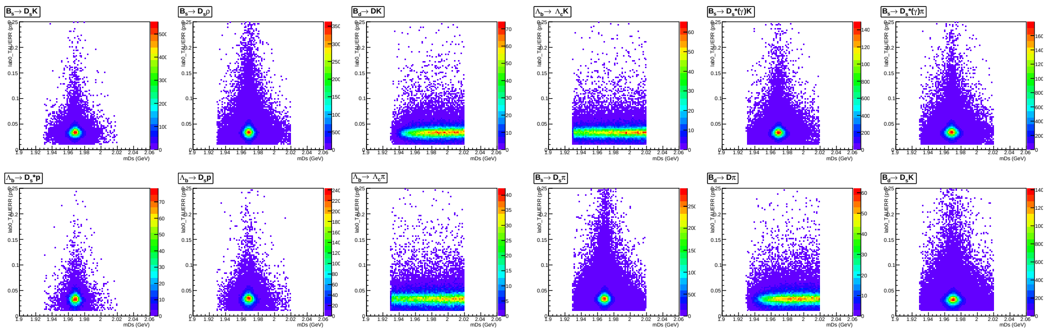


Figure 117.: Scatter plots δ_t mass versus D_s for each emulated decay starting from the signal.

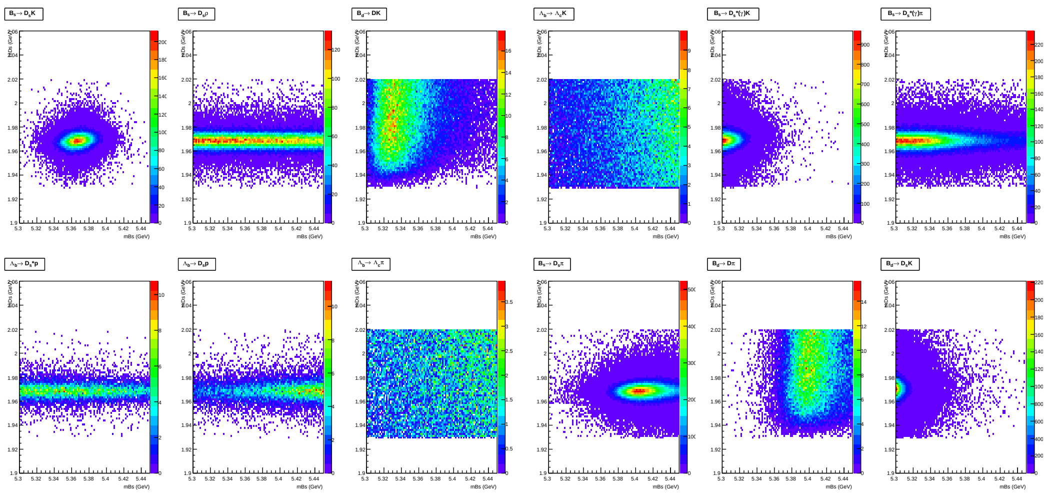


Figure 118.: Scatter plots D_s versus B_s mass for each emulated decay starting from the signal.

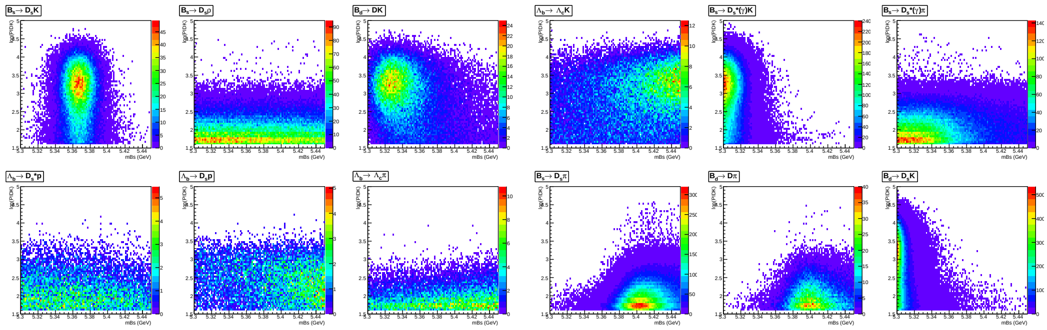


Figure 119.: Scatter Plots PIDK of bachelor particle versus B_s mass for each emulated decay starting from the signal.

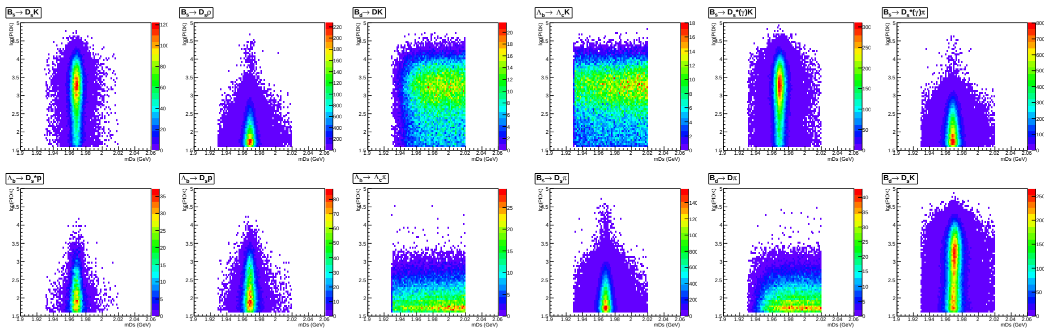


Figure 120.: Scatter Plots PIDK of bachelor particle versus D_s mass for each emulated decay starting from the signal.

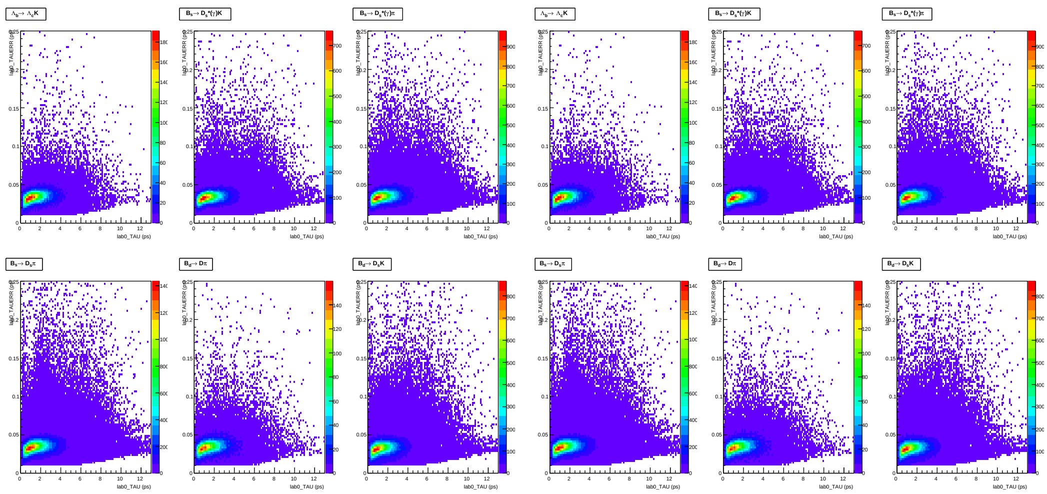


Figure 121.: Scatter plots δ_t versus t_R for each emulated decay starting from the signal.

Table 79.: Emulated Backgrounds with the corresponding expected number of events.

Decay	Expected Yields
$B^0 \to D^- \pi^+$	60
$B^0 \to D^- K^+$	66
$B_d \to D_s K$	381
$B_s^0 \to D_s^- \pi^+$	2678
$\Lambda_b \to \Lambda_c K$	63
$\Lambda_b \to \Lambda_c \pi$	45
$\Lambda_b \to D_s p$	44
$\Lambda_b \to D_s^* p$	14
$B_s \to D_s^*(\gamma)\pi$	727
$B_s \to D_s \rho$	727

A.4 TOYS STUDY FOR ϵ_{tag} VERSUS TIME

A.4.1 *Pull distributions*

A.4.2 *Toy by toy difference*

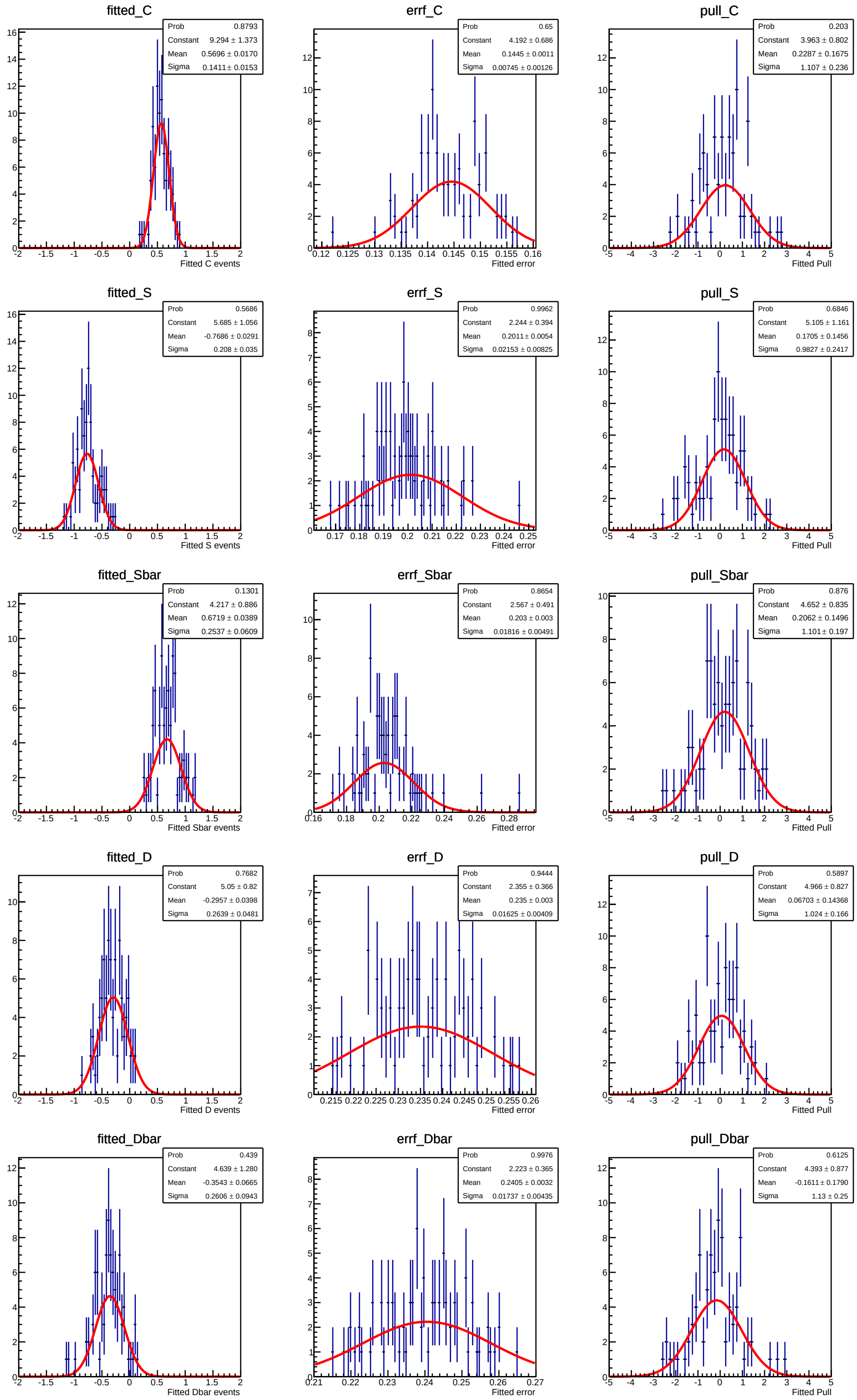


Figure 122.: CP observables, error and pull distributions for the Toys set ϵ_{tag} (t).

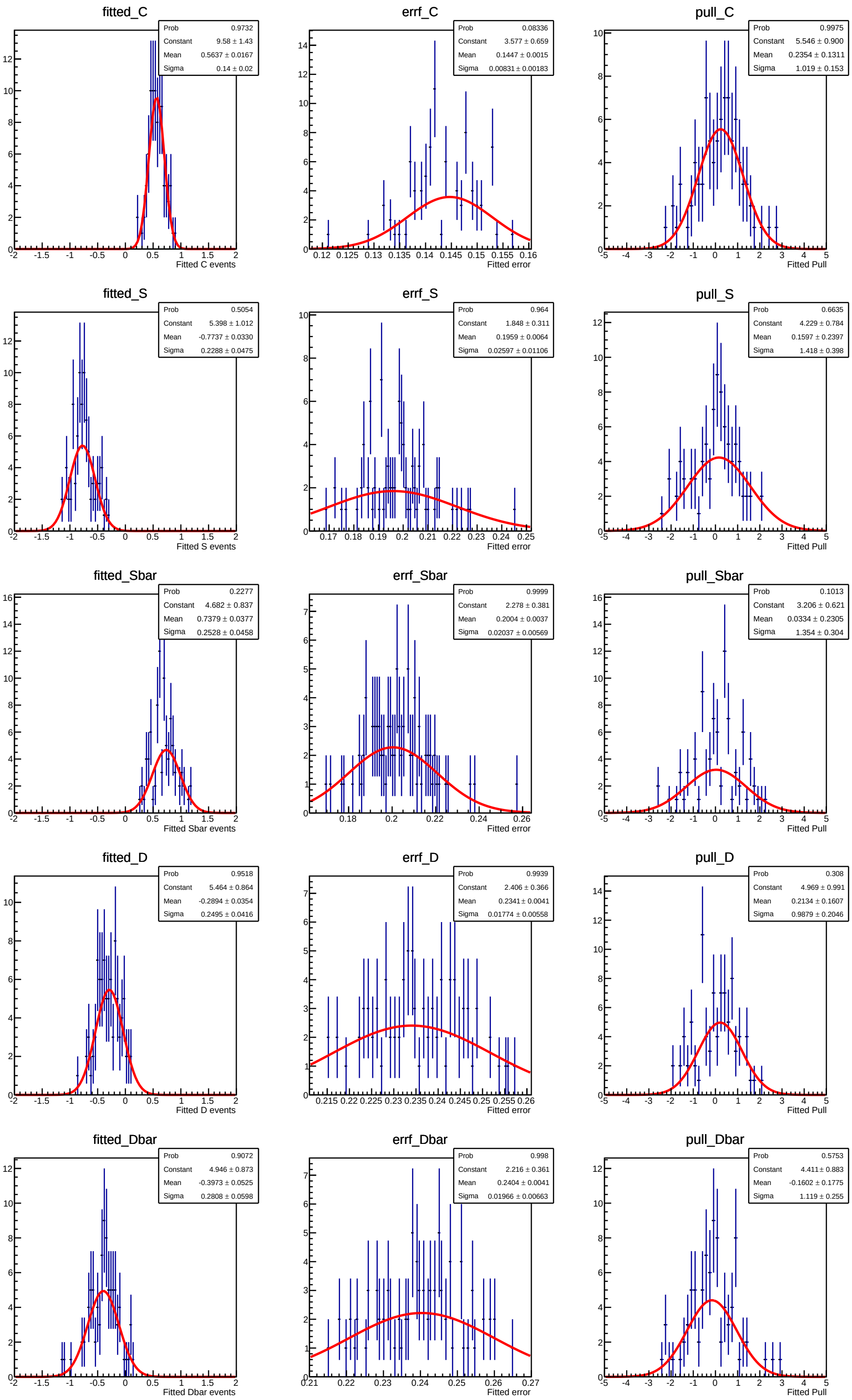


Figure 123.: CP observables, error and pull distributions for the Toys set ϵ_{tag} = const.

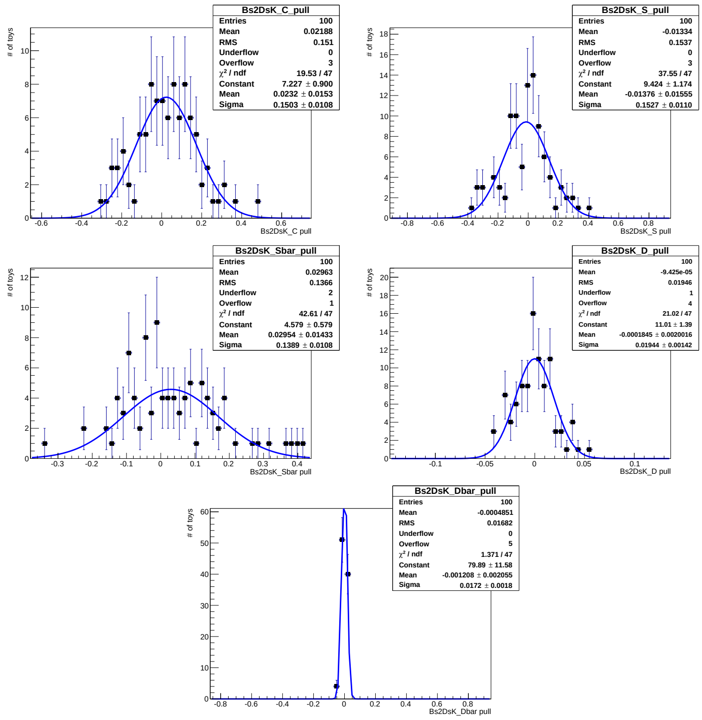


Figure 124.: CP observables pull distributions for the toy by toy difference between Toys set with ϵ_{tag} (t) and Toys set with ϵ_{tag} = const.

BIBLIOGRAPHY

[1] LHCb Collaboration, *Measurement of CP asymmetry in* $B_s^0 \to D_s^{\mp} K^{\pm}$ *decays* , J. High Energy Phys. 11 (2014) 060

[2] Gordon Kane, *Moder Elementary Particle Physics*, Addison-Wesley Publishing Company

[3] Particle Data Group web site.

[4] ATLAS Collaboration, G. Aad *et al., Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC* Phys.Lett. B716 (2012) 1-29.

[5] 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. B 716 (2012) 30.

[6] M.S. Sozzi, *Discrete Symmetries and CP violation*, Oxford University Press (2008)

[7] The Babar collaboration, *Observation of Time Reversal Violation in the* B^0 *Meson System*, Phys. Rev. Lett. 109, 211801 (2012), arXiv:1207.5832

[8] B.I. Bigi, A.I. Sanda, *CP violation, Cambridge monographs on particle physics*, nuclear phyiscs and cosmology 28 - Cambridge University Press(2009)

[9] CKMFitter group web page http://ckmfitter.in2p3.fr/www/results/plots_-eps15/ckm_res_eps15.html

[10] LHCb Collaboration, *Measurement of the CKM angle* γ *from a combination of* $B \to DK$ *analyses* , LHCb-CONF-2016-001

[11] LHCb Collaboration, *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*

[12] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. 51, 1945–1947 (1983).

[13] M. Bona et al. [UTfit Collaboration], http://www.utfit.org/UTfit

[14] D0 Collaboration, *Evidence for anomalous like-sign Dimuon charge asymmetry*, Phys. Rev. Lett. 105, 081801, 2010.

[15] LHCb Collaboration, *Study of $B^- \to DK^-\pi^+\pi^-$ and $B^- \to D\pi^-\pi^+\pi^-$ decays and determination of the CKM angle γ* , PRD 92 (2015) 112005

[16] LHCb Collaboration, *Measurement of the CKM angle γ using $B^\pm \to DK^\pm$ with $D \to K^0_S\pi^+\pi^-$, $K^0_SK^+K^-$ decays*, JHEP 10 (2014) 097

[17] Gronau & London, PLB 253 (1991) 483, Gronau & Wyler PLB 265 (1991) 172

[18] Atwood, Dunietz & Soni PRL 78 (1997) 3257, PRD 63 (2001) 036005 (ADS)

[19] Giri, Gronau, Soffer & Zupan, (GGSZ) PRD 68 (2003) 054018

[20] J.Brod, J. Zupan *The ultimate theoretical error on γ from $B \to DK$ decays*, JHEP 1401 (2014) 051

[21] LHCb Collaboration, *Model-independent measurement of the CKM angle γ in $B^+ \to DK^+$ with $D \to K^0_S\pi^+\pi^-$ decays*, Nucl. Phys. **B888** (2014) 169

[22] LHCb Collaboration, *Model-independent measurement of the CKM angle γ using $B^0 \to DK^{*0}$ decays with $D \to K^0_S\pi^+\pi^-$ and $K^0_SK^+K^-$*, JHEP 06 (2016) 131

[23] http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html

[24] http://lhcb.web.cern.ch/lhcb/speakersbureau/excel/default.html

[25] LHCb Collaboration, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7TeV$ in the forward region, Physics Letters B 694, 209 (2010), arXiv:1009.2731*

[26] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard Interactions of Quarks and Gluons: a Primer for LHC Physics* Rept. Prog. Phys.70 (2007) 89, arXiv:hep-ph/0611148

[27] L.Evans *et al., The LHC machine*, JINST **3** (2008) S08001

[28] ATLAS collaboration, G. Aad *et al., The ATLAS experiment at the CERN LHC* JINST **3** (2008) S08003

[29] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC* JINST **3** (2008) S08004

[30] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002

[31] LHCb Collaboration et al., Journal of Instrumentation 3, 8005 (2008)

[32] LHCb collaboration, R. Aaij *et al.*, *LHCb Detector Performance* , Int. J. Mod. Phys. A 30 (2015) 1530022

[33] LHCb collaboration, web page: http://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material for_Presentations.html

[34] LHCb collaboration, *Performance of the LHCb Vertex Locator*, J. Instrum. 9 P09007

[35] LHCb Collaboration et al.,*Performance of the LHCb RICH detector at LHC*, Eur. Phys. J. C (2013) 73:2431, DOI 10.1140/epjc/s10052-013-2431-9

[36] LHCb Silicon Tracker Website, http://lhcb.physik.uzh.ch/ST/public/

[37] P. Cherenkov, *Visible emission of clean liquids by action of γradiation*, Doklady Akademij Nauk SSSR **2** (1934) 451

[38] LHCb Collaboration, *First years of running for the LHCb Calorimeter system*, Conference Proceedings *Technology and Instrumentation in Particle Physics 2014*, June 2014, Amsterdam.

[39] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework* Nucl. Instrum. Meth. **A389** (1997) 81.

[40] The Gaudi project, web page: http://proj-gaudi.web.cern.ch/proj-gaudi/.

[41] The GAUSS project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/.

[42] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[43] The BOOLE project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[44] The MOORE project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[45] The BRUNEL project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[46] The DaVinci project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[47] The Urania project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/urania/.

[48] LHCb Collaboration, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, arXiv:1602.07252.

[49] LHCb Collaboration, Opposite-side flavour tagging of B mesons at the LHCb experiment, Eur. Phys. J. C (2012) 72:2022 DOI 10.1140/epjc/s10052-012-2022-1

[50] LHCb Collaboration, *B flavour tagging using charm decays at the LHCb experiment* JINST 10 (2015) P10005.

[51] LHCb Collaboration, *Observation of Orbitally Excited $B_{(s)}^{**}$ Mesons*, LHCb-CONF-2011-053

[52] LHCb Collaboration, *sPlot: a statistical tool to unfold data distributions*,arXiv:physics/0402083

[53] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," PoS A CAT 040 (2007) [physics/0703039]

[54] LHCb Collaboration, *Measurements of time-dependent CP-violation observables in* $B_s^0 \to D_s^{\mp} K^{\pm}$, LHCb-ANA-2012-123

[55] Yuehong Xie, *sFit: a method for background subtraction in maximum likelihood fit*, arXiv physics.data-an/0905.0724

[56] The Particle ID web page: https://twiki.cern.ch/twiki/bin/view/LHCb/GlobalParticleID

[57] The RooKeysPdf web page: https://root.cern.ch/doc/master/classRooKeysPdf.html

[58] Skwarnicki, Tomasz *A study of the radiative CASCADE transitions between the* Υ' *and* Υ *resonances* , DESY-F-31-86-02

[59] LHCb Collaboration, *Measurement of* Δm_s *in the decay* $B_s^0 \to D_s(K^+K^-\pi^-)\pi^+$ *and* $B_s^0 \to D_s(K^+K^-\pi^?)\pi^+\pi^+\pi^-$ *in 2011 data*, LHCb-ANA-2011-043

[60] LHCb Collaboration, *Measurement of time-dependent CP-violation observables in* $B_s^0 \to D_s^{\mp}K^{\pm}$, LHCb-ANA-2012-068

[61] LHCb Collaboration, *Measurement of* Δm_s *in the decay* $B_s^0 \to D_s^-\pi^+$ *using* $1fb^{?1}$, LHCb-ANA-2012-053

[62] J.Beringer et al.(Particle Data Group), Phys. Rev. D86, 010001 (2012)

[63] The RooBDecay web page: https://root.cern.ch/doc/master/classRooBDecay.html

[64] LHCb Collaboration, *Precision measurement of the* $B_s^0 - \overline{B}_s^0$ *oscillation frequency with the decay* $B_s^0 \to D_s^-\pi^+$

[65] LHCb Collaboration, *A new algorithm for identifying the flavour of* B_s^0 *mesons at LHCb*, arXiv:1602.07252

[66] PhD thesis, *Development and calibration of a same side kaon tagging algorithm and measurement of the* B_s^0-$\bar{B}_s^0$ *oscillation frequency* Δm_s *at the LHCb experiment*

[67] , PhD thesis,

[68] LHCb Collaboration, *Precision measurement of CP violation in* $B_s^0 \to J/\psi K^+K^-$ *decays*, Phys. Rev. Lett. 114 (2015) 041801

[69] LHCb Collaboration, *Measurement of the CP-violating phase* ϕ_s *in* $\overline{B}_s^0 \to D_s^+D_s^-$ *decays*, Phys. Rev. Lett. 113 (2014) 211801

[70] LHCb Collaboration, *Measurement of CP violation in* $B_s^0 \to \phi\phi$ *decays* , Phys. Rev. D 90 (2014) 052011

[71] LHCb Collaboration, *Measurement of the CP-violating phase* ϕ_s *in* $\overline{B}_s^0 \to J/\psi\pi^+\pi^-$ *decays*, Phys. Lett. B 736 (2014) 186

[72] LHCb Collaboration, *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*

[73] LHCb Collaboration, *Measurement of the branching fractions of the decays* $B_s^0 \to D_s^{\mp}K^{\pm}$ *and* $B_s^0 \to D_s^-\pi^+$, JHEP 06 (2012) 115

[74] LHCb Collaboration, *Updated average* f_s/f_d *b-hadron production fraction ratio for 7 TeV pp collisions*, EPS-HEP 2013 European Physical Society Conference on High Energy Physics, Stockholm, Sweden, 18 - 24 Jul 2013

[75] TMWA package web-site: http://tmva.sourceforge.net/

[76] LHCb Collaboration, *Measurement of time-dependent CP-violation observables in* $B_s^0 \to D_s^{\mp}K^{\pm}$ LHCb-CONF-2012-029

[77] Karbach, Till Moritz and Raven, Gerhard and Schiller Manuel, *Decay time integrals in neutral meson mixing and their efficient evaluation*, LHCb-INT-2013-043

[78] TGenPhaseSpace web page: https://root.cern.ch/doc/master/classTGenPhaseSpace.html

[79] Alexander, Michael T. Phd thesis, *Constraints on mixing and CP-violation in the neutral charmed meson system at LHCb*, University of Glagow

[80] Efron, B. and Tibshirani, R.J. , *An Introduction to the Bootstrap*, 1993, pubblished by Chapman & Hall

[81] LHCb Collaboration, *The LHCb detector at the LHC*, 2008 Jinst 3 S08005

INDEX

Control channel, 23

mistag, 22