

REAL-TIME ALIGNMENT OF THE LHCb VERTEX
DETECTOR AND OBSERVATION OF CHARMLESS
BARYONIC DECAYS $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$

A thesis submitted to the University of Manchester
for the degree of
DOCTOR OF PHILOSOPHY
(PhD)
in the Faculty of Science and Engineering

GIULIO DUJANY



The University of Manchester

Particle Physics Group
School of Physics and Astronomy
2017



CONTENTS

Introduction	21
1 THEORETICAL OVERVIEW	23
1.1 The Standard Model of particle physics	23
1.2 Lepton flavour violation	27
1.3 Charmless multibody baryonic B-meson decays	31
2 THE LHCb EXPERIMENT AT THE LHC	39
2.1 The Large Hadron Collider	39
2.1.1 LHC design and performances	39
2.1.2 The LHC experiments	40
2.1.3 LHC data taking periods: Run I and Run II	43
2.2 LHCb	45
2.2.1 Vertex Locator	46
2.2.2 Tracking system	49
2.2.3 Cherenkov detectors	49
2.2.4 Calorimeters	50
2.2.5 Muon system	51
2.2.6 Trigger	51
2.3 Simulation, reconstruction and other aspects of an analysis	53
2.3.1 Simulation	53
2.3.2 Track reconstruction	54
2.3.3 Particle identification	55
2.3.4 Selection	56
3 REAL-TIME ALIGNMENT OF THE LHCb VERTEX DETECTOR	59
3.1 Track-based alignment of a tracking detector	59
3.2 Real-time calibration and alignment	73
3.3 Real-time alignment of the VELO	79
3.4 Conclusion	94
4 SEARCH FOR THE LEPTON FLAVOUR VIOLATING DECAY $\tau \rightarrow \phi \mu$	97
4.1 Motivation	97

4.2	Dataset and simulated samples	98
4.3	Trigger and stripping selection	98
4.4	Background characterisation	100
4.5	MVA and PID selection	104
4.6	Fit and limit setting	105
4.7	Result on unblinded 2012 data	111
5	OBSERVATION OF CHARMLESS BARYONIC DECAYS $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$	115
5.1	Motivation	116
5.2	Analysis strategy	119
5.3	Dataset and simulated samples	120
5.4	Selection	121
5.4.1	Trigger and stripping selection and fiducial region	121
5.4.2	MVA selection	124
5.4.3	PID selection	128
5.4.4	Charm vetoes	132
5.4.5	Working point	133
5.4.6	Muon misidentification	137
5.4.7	Multiple candidates	138
5.5	Background studies	139
5.6	Efficiencies	142
5.6.1	Efficiencies for flat phase space decay models	143
5.6.2	Phase space dependence	144
5.6.3	Efficiencies of the charm vetoes	151
5.7	Fits to the data	153
5.7.1	Signal shapes	153
5.7.2	Parameterising the reflections	155
5.7.3	Simultaneous fit to the data including charm contributions	157
5.7.4	Substructures	158
5.7.5	Simultaneous fit to the data after applying the charm vetoes	161
5.7.6	Substructures after the charm vetoes	164
5.7.7	Normalisation mode	169
5.8	Systematic uncertainties	171
5.8.1	External contributions to the systematic uncertainties	171
5.8.2	Trigger	173

5.8.3	Tracking	174
5.8.4	Selection	176
5.8.5	B_s^0 lifetime	179
5.8.6	Fit model	180
5.8.7	Summary of systematic uncertainties	186
5.9	Results	190
Conclusions		195
A	CORRECTION OF τ AND D PRODUCTION RATES	197
B	CROSS CHECKS FOR $\tau \rightarrow \phi\mu$	201
B.1	Pseudoexperiments	201
B.2	Check on yield normalisation channel	202
C	FURTHER STUDIES ON BDT	203
C.1	On different BDTs for each $p\bar{p}h h'$ spectrum	203
C.2	Comparison between 2011 and 2012	207
C.3	BDT input variables	207
D	DATA-MC COMPARISONS	211
D.1	BDT input variables	211
D.2	PID variables	211
E	FIT VALIDATION WITH PSEUDOEXPERIMENTS	217
Bibliography		219

Word count: 37078

LIST OF FIGURES

Figure 1.1	Particles in the Standard Model	24
Figure 1.2	Main Standard Model diagram for $\tau \rightarrow \phi \mu$	27
Figure 1.3	Upper limits on LFV tau decays	30
Figure 1.4	Measured raw asymmetry in Dalitz plot bins of background-subtracted and acceptance-corrected events for $B^+ \rightarrow K^+ \pi^+ \pi^-$ decays [1].	34
Figure 1.5	Example of threshold enhancement	36
Figure 1.6	Illustration of threshold enhancement	37
Figure 2.1	LHC layout	41
Figure 2.2	CERN's accelerator complex	41
Figure 2.3	LHC's four main experiments	42
Figure 2.4	Luminosity delivered by LHC and recorded by LHCb	44
Figure 2.5	Evolution of the instantaneous luminosity at LHCb	44
Figure 2.6	LHCb's acceptance	45
Figure 2.7	Schematic of the LHCb detector	46
Figure 2.8	Schematic of the VELO layout	47
Figure 2.9	VELO R and Φ sensor geometry	48
Figure 2.10	Layout of the LHCb tracking system	49
Figure 2.11	Cherenkov bands for different particles in RICH1	50
Figure 2.12	Layout of the LHCb calorimeter system	52
Figure 2.13	Tracks types in LHCb	54
Figure 3.1	Improvement IP resolution with better alignment	61
Figure 3.2	Improvement IP resolution with better alignment	62
Figure 3.3	Degrees of freedom of a detector element	62
Figure 3.4	How the misalignment causes a bias in the track reconstruction	63
Figure 3.5	Definition of residual	63
Figure 3.6	Residual and misalignments	63
Figure 3.7	Millipede algorithm	67
Figure 3.8	Kalman filter	70
Figure 3.9	Trigger strategy for Run II	74

Figure 3.10	Rich mirror alignment	77
Figure 3.11	Estimation of the global time shift between the collision time and the LHCb clock	78
Figure 3.12	Misalignment between the two VELO halves in Run I	80
Figure 3.13	Comparison of the 2-halves T_x misalignment before and after the alignment procedure	81
Figure 3.14	Comparison of the track angular distribution, the number of hits per track and the vertex position for collision and beamgas events	82
Figure 3.15	Residual modules' misalignment fixing the position of some modules	85
Figure 3.16	Residual T_z misalignment	86
Figure 3.17	Convergence first VELO alignment of Run II	91
Figure 3.18	Stability of the alignment of the VELO halves during all fills of 2015	92
Figure 3.19	Stability of module alignment during the beginning of 2015	93
Figure 3.20	Stability of the alignment of the VELO halves during all fills of 2016	94
Figure 3.21	Fraction of updates triggered by each combination of degrees of freedom	95
Figure 3.22	Fraction of updates in which each degree of freedom is above threshold	95
Figure 4.1	Purity of the ϕ peak	101
Figure 4.2	Fit to $D^\pm \rightarrow \phi\pi^\pm$	101
Figure 4.3	Main sources of background to $\tau \rightarrow \phi\mu$	103
Figure 4.4	Variables used as input to the BDT classifier	106
Figure 4.5	ROC curve and output of the BDT classifier	106
Figure 4.6	Distribution of the data in the ProbNNmu-ProbNNpi plane	107
Figure 4.7	ROC curve for various PID variables	107
Figure 4.8	Value of the figure of merit for the different selections	108
Figure 4.9	Fit $\tau \rightarrow \phi\mu$ MC	108
Figure 4.10	Fit to 10% of the 2012 data	109
Figure 4.11	Distribution of the expected upper limit on $\tau \rightarrow \phi\mu$ using the full Run I dataset	111
Figure 4.12	Fit to the 2012 data	112
Figure 4.13	Upper limit on $\tau \rightarrow \phi\mu$ with the 2012 dataset	112

Figure 5.1	Feynman diagrams contributing to $B_s^0 \rightarrow p\bar{p}K^+K^-$, $B_s^0 \rightarrow p\bar{p}K^\pm\pi^\mp$, $B^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ and $B^0 \rightarrow p\bar{p}\pi^+\pi^-$ 117
Figure 5.2	Feynman diagrams contributing to $B_s^0 \rightarrow p\bar{p}\pi^+\pi^-$, $B^0 \rightarrow p\bar{p}K^+K^-$ and $B_{(s)}^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ 118
Figure 5.3	Mass spectrum of $p\bar{p}hh'$ after stripping 124
Figure 5.4	Comparison of BDT output 126
Figure 5.5	Comparison of ROC curves 127
Figure 5.6	Output of the BDT as a function of the $p\bar{p}hh'$ invariant mass 127
Figure 5.7	ROC curves comparing the performance of various PID variables 129
Figure 5.8	Binning scheme for proton calibration sample 131
Figure 5.9	Distributions of the number of tracks in the signal and the calibration samples 132
Figure 5.10	Invariant mass distributions for $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples after trigger and stripping 135
Figure 5.11	Significance vs BDT output 137
Figure 5.12	Muon retention 138
Figure 5.13	$\Lambda_b^0 \rightarrow p\bar{p}ph$ invariant mass in the $p\bar{p}hh'$ mass spectra 140
Figure 5.14	Data and signal-only MC reconstructing the final state as $p\bar{p}ph$ 141
Figure 5.15	Scatter plot of $m_{K\pi}$ vs $m_{pK\pi}$ for $B^0 \rightarrow p\bar{p}K\pi$ signal MC candidates 151
Figure 5.16	Fit to the <i>sWeight'</i> ed $B^0 \rightarrow p\bar{p}K\pi$ signal after the charm vetoes 152
Figure 5.17	Signal shapes obtained from MC 154
Figure 5.18	Invariant mass distributions for $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples after the full selection 156
Figure 5.19	Shapes of the reflections obtained from MC 157
Figure 5.20	Simultaneous fit to the three mass spectra including all charm contributions 159
Figure 5.21	Substructures in the <i>sPlot'</i> ed $B_s^0 \rightarrow p\bar{p}KK$ signal 161
Figure 5.22	Substructures in the <i>sPlot'</i> ed $B^0 \rightarrow p\bar{p}K\pi$ signal 162
Figure 5.23	Substructures in the <i>sPlot'</i> ed $B^0 \rightarrow p\bar{p}\pi\pi$ signal 162

Figure 5.24	Fit to the charmonium and charm substructures 163
Figure 5.25	Simultaneous fit to the three mass spectra excluding all charm contributions 166
Figure 5.26	Substructures in the <i>sPlot</i> 'ed $B_s^0 \rightarrow p\bar{p}KK$ signal after removing all charm contributions 167
Figure 5.27	Substructures in the <i>sPlot</i> 'ed $B^0 \rightarrow p\bar{p}K\pi$ signal after removing all charm contributions 167
Figure 5.28	Substructures in the <i>sPlot</i> 'ed $B^0 \rightarrow p\bar{p}\pi\pi$ signal after removing all charm contributions 168
Figure 5.29	Fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ MC 170
Figure 5.30	Fit to the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ 172
Figure 5.31	Simultaneous fit to the three mass spectra adding extra components to the fit 182
Figure 5.32	Distributions of the differences of the yield parameters obtained from pseudoexperiments 184
Figure 5.33	3-D fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ data using the alternative model in which the more components are added to the fit 187
Figure 5.34	Profile likelihood for the yields of the three decay modes in which the significance is less than 20 191
Figure 5.35	Distribution of the profile likelihood used to set the upper limit on $B_s \text{ToppPiPi}$ 194
Figure B.1	Toy study on fit uncertainty 201
Figure C.1	Comparison of the BDT output trained for the three different final states 204
Figure C.2	ROC curve for the BDTs trained used in all three final states 205
Figure C.3	Comparison of the ROC curves for the BDTs trained using different final states 206
Figure C.4	Comparison of the BDT output in the 2012 and 2011 dataset 207
Figure C.5	Comparison of BDT0 input variables for signal and background samples 208
Figure C.6	Comparison of BDT1 input variables for signal and background samples 209

Figure C.7	Linear correlations between the BDT input variables 210
Figure D.1	Comparison of all BDT input variables between data and MC for the $B_s^0 \rightarrow p\bar{p}KK$ mode. 212
Figure D.2	Comparison of all BDT input variables between data and MC for the $B^0 \rightarrow p\bar{p}K\pi$ mode. 213
Figure D.3	Comparison of all BDT input variables between data and MC for the $B^0 \rightarrow p\bar{p}\pi\pi$ mode. 214
Figure D.4	Effect of PID resampling on PIDk on kaons 214
Figure D.5	Effect of PID resampling on PIDk on pions 215
Figure D.6	Effect of PID resampling on ProbNNp 215
Figure E.1	Pull distributions of the signal yields obtained generating and fitting pseudoexperiments 218
Figure E.2	Pull distributions of the yield parameters of the normalisation mode obtained generating and fitting pseudoexperiments 218

LIST OF TABLES

Table 2.1	LHC parameters 40
Table 3.1	Residual two-halves' misalignments 84
Table 3.2	Residual two-halves' misalignments when also the modules are misaligned 84
Table 3.3	Residual two-halves' and modules' misalignments 87
Table 3.4	Residual two-halves' and modules' misalignments aligning also for subdominant degrees of freedom 88
Table 3.5	Precision and accuracy of the alignment procedure evaluated from MC and data 90
Table 3.6	Thresholds on the variation alignment constant used in 2015 90
Table 3.7	Thresholds on the variation alignment constants used in 2016 93
Table 4.1	Upper limits on $\mathcal{B}(\tau \rightarrow \phi\mu)$ 98
Table 4.2	Stripping selection for $\tau \rightarrow \phi\mu$ 99

Table 4.3	Main sources of expected background, with rejection rates of the selections not taken into account. 102
Table 4.4	Variables used as input to the BDT classifier 104
Table 4.5	Hyper-parameters of the BDT classifier 104
Table 4.6	Parameters of the fit to the MC signal shape 109
Table 4.7	Parameters of the fit to 10% of the 2012 data 110
Table 4.8	Inputs to the limit setting 110
Table 4.9	Parameters of the fit to the 2012 dataset 111
Table 5.1	Stripping selection criteria 123
Table 5.2	Hyper-parameters of the BDT classifier 125
Table 5.3	Input variables of BDT classifier 125
Table 5.4	Expectations on the relative yields before the selection of the $B_{(s)}^0 \rightarrow p\bar{p}h h'$ modes 133
Table 5.5	Final selection's working point 136
Table 5.6	Total phase-space-integrated efficiencies 145
Table 5.7	Phase-space-integrated efficiencies 145
Table 5.8	Phase-space-integrated efficiencies including cross-feed decays 146
Table 5.9	Phase-space-corrected efficiencies 150
Table 5.10	Efficiencies of the charm vetoes 152
Table 5.11	Parameters of the fits to the MC signal shapes 154
Table 5.12	Parameters of the fits to the MC misidentified decay modes 155
Table 5.13	Parameters of the simultaneous fit to the data including all charm contributions 160
Table 5.14	Yields of the $B_{(s)}^0$ peaks from fit with all the yields free to vary 161
Table 5.15	Parameters of the simultaneous fit to the data excluding all charm contributions 165
Table 5.16	Yields of the $B_{(s)}^0$ peaks from fit with all the yields free to vary after applying the charm vetoes 166
Table 5.17	Parameters of the 3-D fit to the MC shape of the normalisation mode 169
Table 5.18	Parameters of the 3-D fit to the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ 171
Table 5.19	Ratios of the efficiencies of the L0 hadron trigger computed using a data-driven method naively from MC samples 174

Table 5.20	Correction factors to the reconstruction efficiencies 175
Table 5.21	Correction factors to the reconstruction efficiencies, increased to take into account hadronic interactions 176
Table 5.22	Total relative systematic uncertainty on the PID selection efficiency 177
Table 5.23	Efficiency of the selection on PID variables 178
Table 5.24	Systematic uncertainties on the efficiencies due to the different B_s^0 lifetime between mass eigenstates 180
Table 5.25	Yields from the fit along with their statistical and systematic uncertainties 185
Table 5.26	Systematic uncertainties on the absolute branching fractions 188
Table 5.27	Systematic uncertainties on the relative branching fractions 189
Table 5.28	Significances of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ peaks 192
Table 5.29	Branching fractions of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decays 192
Table 5.30	Relative branching fractions among $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decays 193
Table 5.31	Efficiency-corrected yields of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decays 193
Table A.1	Production cross-sections at 8 TeV of D mesons and $b\bar{b}$ pairs 197
Table A.2	Correction of τ relative production 199
Table B.1	Quantities used to check the luminosity value 202

ABSTRACT

University name: The University of Manchester

Candidate name: Giulio Dujany

Degree title: Doctor of Philosophy (PhD)

Thesis title: Real-time alignment of the LHCb vertex detector and observation of charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$

Submission date: 24th March 2017

This thesis presents measurements of the branching fractions of the charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$, where $h^{(\prime)}$ denotes a kaon or a pion. Three new modes ($B^0 \rightarrow p\bar{p}\pi^+\pi^-$, $B_s^0 \rightarrow p\bar{p}K^+K^-$ and $B_s^0 \rightarrow p\bar{p}K^\pm\pi^\mp$) are observed for the first time and evidence is found for a fourth ($B^0 \rightarrow p\bar{p}K^+K^-$). The inclusive branching fraction of $B^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ is measured for the first time and the upper limit is set on the branching fraction of the $B_s^0 \rightarrow p\bar{p}\pi^+\pi^-$ decay. This represents the first observation of four-body charmless baryonic decays of a B_s meson and one of the first observations of baryonic B_s^0 decays. The implementation of the real-time alignment of LHCb's vertex detector is also described. The novel real-time alignment and calibration strategy adopted by LHCb is essential to allow more stable data taking conditions and an optimal "offline-quality" reconstruction to be performed at the trigger level, ensuring more efficient trigger selections and the possibility to perform physics analyses directly on the trigger output.

DECLARATION

Candidate name: Giulio Dujany

Faculty: Faculty of Science and Engineering

Thesis title: Real-time alignment of the LHCb vertex detector and observation of charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$

This work represents the combined efforts of the LHCb collaboration. Some of the content has been published elsewhere and/or presented to several audiences. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy¹, in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations² and in The University’s policy on Presentation of Theses.

¹ See <http://documents.manchester.ac.uk/display.aspx?DocID=24420>

² See <http://www.library.manchester.ac.uk/about/regulations/>

*On fait la science avec des faits comme une maison avec des pierres mais
une accumulation de faits n'est pas plus une science qu'un tas de pierres
n'est une maison.*

— Henri Poincaré

Many are the people to whom
I am indebted and who made this thesis possible,
from my family and my friends, the other members of the LHCb
experiment and the other members of the Manchester physics department.
It is a great pleasure to thank in particular my supervisors George Lafferty,
Silvia Borghi, Eduardo Rodrigues and Chris Parkes for their invaluable assis-
tance and guidance; my sincere thanks also to the other members of the LHCb
BnoC physics working group and other colleagues that helped improving the
quality of the analysis, in particular Stefano Perazzini, Daniel O'Hanlon, Al-
varo Gomes, Kristof De Bruyn, Matteo Palutan and Adlene Hitcher. I have
had invaluable help from the other members of the Tracking and Align-
ment working group and the LHCb Online computing team in partic-
ular Clara, Beat, Markus, Roel, Lucia, Patrick, and Barbara. Many
thanks are also due to the University of Manchester and
to the Science and Technology Facilities Council that
funded my PhD. A special thanks goes also to
my housemate Suzanne. I finally wish to
thank Andrzej Szec and Miriam
Watson for the enjoyable
PhD viva.



INTRODUCTION

The Standard Model of particle physics, the theory that describes the elementary particles and their interactions, is the most precisely tested theory in the history of science with, in some cases, an agreement between theoretical predictions and experimental results of ten parts in a billion. This theory is however believed to be incomplete and, to look for hints of a more fundamental theory, deviations from the Standard Model are searched for, looking for new particles or for decays not allowed, like the one discussed in Chapter 4, or measuring with unprecedented precision the parameters of this theory, as is done in the study of CP violation.

Even if the basis of the Standard Model is well known, it is not possible to make accurate predictions at every energy regime. The transition from quarks to hadrons is dominated by non perturbative quantum chromodynamics and analytical computations are not available. To constrain phenomenological models and deepen our understanding of the theory in these regimes the input of experiments is crucial, for example with the measurements of branching fractions like those presented in Chapter 5.

To probe such small distance scales and measure very rare processes with high accuracy, large collaborations are needed to build large and complex detectors, like the LHCb detector, which was used for all the work presented in this thesis. The analysis of massive amounts of data is also needed, posing unprecedented technical challenges that require new creative solutions. The real-time alignment and calibration of the LHCb detector, also presented in this thesis, goes along this direction allowing to exploit more efficiently the computing resources and making possible analyses otherwise too onerous.

The thesis is organised as follows.

The first chapter summarises the theoretical framework used in the following chapters. A brief introduction to the Standard Model of particle physics is given. Lepton flavour violation is discussed and, finally, the theoretical framework used to describe B-meson decays is given.

The second chapter is devoted to a description of the Large Hadron Collider and, in particular, of the LHCb experiment. Its sub-detectors are described, highlighting the characteristics that made possible the analyses exposed in the following chapters. Some details are also given about the software framework and some of the analysis techniques used in the LHCb collaboration.

The third chapter describes, after an introduction on the principles behind the track-based alignment of a tracking detector, the novel real-time alignment and calibration adopted by the LHCb experiment. The focus is given to the author's contributions: the preparation, the commissioning and the study of the performance of the real-time alignment of the vertex locator.

The fourth chapter presents a feasibility study for a search for the lepton flavour violating decay $\tau \rightarrow \phi\mu$ using the full Run I LHCb dataset.

The fifth chapter presents the search for the six inclusive charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$ ($h^{(\prime)} = \pi, K$) based on the 3 fb^{-1} dataset collected by LHCb during the Run I of the LHC. All the details of the analysis are illustrated, from the selection to the determination of the systematic uncertainties. Four modes are observed (three for the first time) and strong evidence is found of a fifth one. This represents the first observation of four-body charmless baryonic decays of B_s^0 mesons and one of the first observations of baryonic B_s^0 decays.

The author has been the main analyst of both the analyses described in the last two chapters. Throughout this thesis the convention $c = 1$ is used and charge conjugation is implied.

1

THEORETICAL OVERVIEW

The present chapter gives a short overview of the theoretical background underlying the research described in this thesis. In Section 1.1 a short introduction to the Standard Model of particle physics is given, Section 1.2 focuses on lepton flavour violation, in particular in tau decays, while Section 1.3 covers the physics of charmless multi-body baryonic B-meson decays.

1.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model of particle physics (SM) is the theory that describes the elementary particles and their interactions. This theory is developed in the framework of quantum field theory (QFT) and describes three of the four fundamental forces of nature: the electromagnetic, the weak and the strong interactions but not the gravitational force.

In QFT each particle type has an associated field. As the SM is a relativistic theory it is based on special relativity and each field transforms under the Poincaré transformations according to one of the representations of the Poincaré group; each representation can be labelled by the spin quantum number. There are thus spinor fields that describe the matter content of the theory (the fermions: leptons and quarks with spin $1/2$), vector fields (the bosons with spin 1) and a scalar field (the Higgs boson with spin 0). It is moreover important to note that for massless particles there are two different spinor representations, called chiral representations: fermions can be thus be right-handed or left-handed. A schematic representation of the particle content of the SM can be found in Figure 1.1.

Particle types

The SM is a *gauge theory* so the interactions between particles can be deduced from the internal local symmetries of the theory and ac-

Bosons

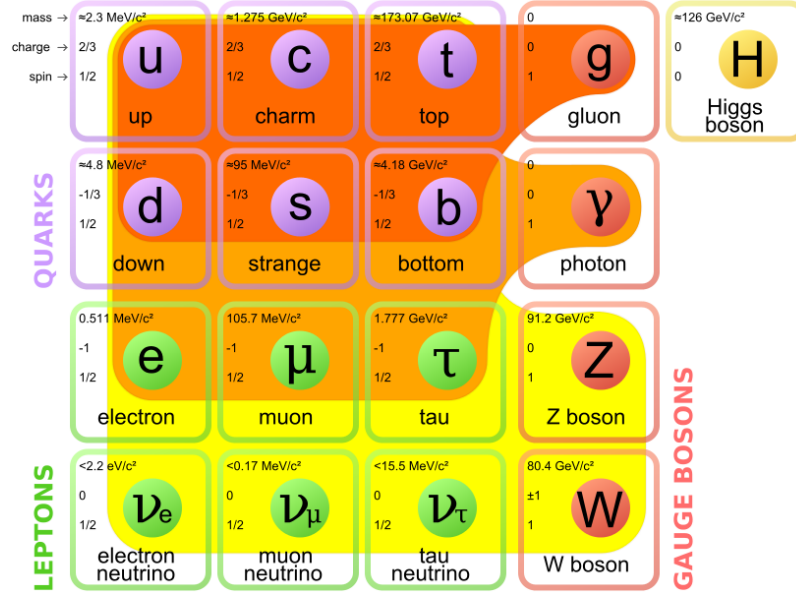


Figure 1.1: Illustration of the particle content of the Standard Model of particle physics. The different shades of orange and yellow indicate which bosons (red) couple to which fermions (purple and green) [2].

cording to which representation of the local gauge symmetry each particle field transforms. The local gauge symmetry of the SM is:

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$$

where $SU(3)_c$ gives rise to quantum chromodynamics (QCD) that describes the strong force and $SU(2)_L \otimes U(1)_Y$ is responsible for the electro-weak force. The gauge bosons are in the adjoint representation of the corresponding group: there are 8 gluons (g) for QCD which couple to colour charge, the gauge bosons of $SU(2)_L$ are the $W_{(1,2,3)}$ which couple to weak isospin J and B is the gauge boson of $U(1)_Y$ which couples to hypercharge Y . In the Higgs mechanism a complex scalar $SU(2)_L$ doublet with nonzero vacuum expectation value causes a spontaneous symmetry breaking

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_c \otimes U(1)_Q$$

Higgs mechanism

so that the observable electro-weak force mediators are a linear combination of the gauge bosons W_i and B as determined by the weak mixing angle, ϑ_w

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \vartheta_w & \sin \vartheta_w \\ -\sin \vartheta_w & \cos \vartheta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix},$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2).$$

The three bosons of the weak force ($W^\pm$ and Z^0) are massive while the residual symmetry $U(1)_Q$ gives rise to the electromagnetic force that is carried by the photon (γ) and couples to the electric charge. The electric charge is related to the $SU(2)_L$ and $U(1)_Y$ quantum numbers by the Gell-Mann-Nishijima relation:

$$Q = J_3 + \frac{Y}{2}.$$

A consequence of this symmetry breaking is also the existence of a electrically neutral massive scalar field, the Higgs boson.

The fermions can be divided in two families depending on their behaviour under $SU(3)_c$: the *quarks* that transform under the fundamental representation of $SU(3)_c$, so interact strongly and can have three *colour* charges; and the *leptons* that do not interact strongly because they are a singlet of $SU(3)_c$. All the fermions couple to $U(1)_Y$ although only the left-handed leptons are doublets of $SU(2)_L$ while the right-handed are singlets so do not couple to the W_i bosons. As the residual $U(1)_Q$ symmetry couples to both chiralities with the same charge it is possible to organise both quarks and leptons in doublets with a top and a bottom component. In the quark sector the top component, or up-like, has electric charge $+2/3$ and the bottom one, down-like, has electric charge $-1/3$. In the lepton sector the top component, called also charged lepton, has electric charge -1 and the bottom component, its associated neutrino, is electrically neutral. One should note that as the right-handed neutrinos are neutral for all three interactions they are unobservable and hence are not part of the SM. Each pair of quark doublet and lepton doublet forms a generation and in the SM there are three such generations.

Fermions

The mass of the fermions comes from the interaction of the left-handed $SU(2)_L$ doublets and the right-handed singlets with the Higgs doublet. The weak eigenstates of the quarks (flavour eigenstates) do

CKM matrix

not coincide with the mass eigenstate but are linked by a unitary matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{weak}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}_{\text{mass}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{mass}}.$$

The charged bosons of the weak interaction can couple to an up-like quark and a down-like quark of different generations while the strong and the electromagnetic interaction cannot. On the other hand in the SM there are no right handed neutrinos so, up to unobservable phases, the weak eigenstates of the charged leptons are also mass eigenstates. As a consequence it is possible to define a conserved quantum number called *lepton flavour* for each family of leptons that has value +1 for each particle and −1 for each antiparticle. In each reaction the three lepton numbers for the three lepton families are conserved. For example in the weak decay of a muon: $\mu \rightarrow e \bar{\nu}_e \nu_\mu$ the muon lepton number is +1 before and after the decay and the electron lepton number is always zero.

*Discrete
transformations*

Along with continuous space-time and gauge transformations also three discrete transformations are important to fully characterise a theory.

- The parity transformation P inverts the direction of the 3 space dimensions: it acts on 4-vectors as $(t, \vec{x}) \mapsto (t, -\vec{x})$, it inverts the momenta but not the spins and it is represented by a unitary transformation in the Hilbert space of states.
- The time reversal T inverts the direction of the time dimension, it acts on 4-vectors as: $(t, \vec{x}) \mapsto (-t, \vec{x})$, it inverts both the momenta and the spins and it is represented by an anti-unitary transformation in the Hilbert space of states.
- The charge conjugate C swaps particles with antiparticles and it is represented by a unitary transformation in the Hilbert space of states.

Individually each one of these three discrete transformations is a symmetry for the strong and electromagnetic interactions while the weak interaction is not invariant under these transformations because the W_i bosons couple only to one chirality. Also the combination CP is not a symmetry for the weak interaction; unitary matrices in three dimensions like the CKM matrix can be parameterised by three an-

gles and a complex phase and, in the SM, this complex phase is the only term responsible for CP violation (CPV). The combination CPT is a symmetry of the SM; this is expected because of the CPT theorem that states that CPT must be a symmetry for each QFT that conserves probabilities and is Lorentz invariant.

1.2 LEPTON FLAVOUR VIOLATION

As seen in the previous section in the SM the *lepton flavour* is conserved for each generation separately. There is however not a strong motivation for this conservation rule as there is not a fundamental symmetry associated with lepton family number but this accidental conservation arises because there is no right-handed neutrino.

The discovery of neutrino oscillations [3] was the first discovery of *lepton flavour violation* (LFV). However, with a Dirac mass term for the neutrinos in the SM Lagrangian, the typical branching fraction of lepton flavour violating decays are orders of magnitude too low to be experimentally observed. The main diagram contributing to the $\tau \rightarrow \phi\mu$ transition is the penguin diagram in figure 1.2 where the neutrino in the loop oscillates.

ν_{SM}

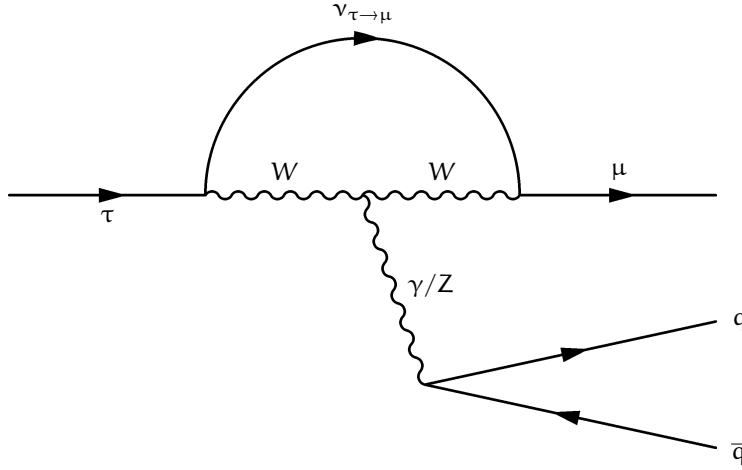


Figure 1.2: Main diagram contributing to the $\tau \rightarrow \phi\mu$ transition in the SM.

It is easy to see with a simple dimensional argument that this kind of diagram is very suppressed. The probability of a neutrino to oscillate between two flavours can be approximated by

$$\mathcal{P}(\nu_\tau \rightarrow \nu_\mu) = \sin^2(2\vartheta_{23}) \sin^2\left(\frac{\Delta m_{23}^2 L}{4E}\right),$$

where $\vartheta_{23} \sim \pi/4$ [4] is the mixing angle between flavour eigenstates:

$$\begin{pmatrix} \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} \cos \vartheta_{23} & \sin \vartheta_{23} \\ -\sin \vartheta_{23} & \cos \vartheta_{23} \end{pmatrix} \begin{pmatrix} \nu_2 \\ \nu_3 \end{pmatrix},$$

$\Delta m_{23}^2 = m_2^2 - m_3^2 \sim 2.4 \cdot 10^{-3} \text{ eV}^2$ [4], L is the distance from the neutrino emission to its observation and E is the neutrino energy. In the penguin diagram above the energy scale is given by the mass of the virtual W boson and, from the uncertainty principle, the length scale is the inverse of its mass. Assuming thus $E \sim L^{-1} \sim M_W$ and with the known approximation for the sine of small angles, the oscillation probability becomes:

$$\mathcal{P}(\nu_\tau \rightarrow \nu_\mu) \propto \left(\frac{\Delta m_{23}^2}{M_W^2} \right)^2 \sim 10^{-49}.$$

A factor of approximately this order of magnitude thus suppresses the diagram considered.

LFV decays in the SM extended to non-zero neutrino masses are therefore far beyond the experimental reach and an observation of such a decay will be a clear indication of new physics. Extensions of the SM predict LFV with branching fractions as high as 10^{-9} [5] by enhancing loop contributions or by directly introducing lepton flavour violating vertices. SUSY models, for example, allow penguin diagrams similar to the one in figure 1.2 where in the loop a slepton and a neutralino substitute the W and the neutrino and a neutral Higgs boson substitutes the neutral gauge boson.

Other examples are the models involving the see-saw mechanism to naturally explain the very light mass of the neutrinos [6]. These models are characterised by the introduction of right-handed sterile neutrinos, which interact with the left-handed SM neutrinos via a Yukawa coupling and also have a Majorana mass term. If the Majorana mass, M , is much larger than the Dirac one, m_D , acquired via the interaction with the Higgs field, the mass of the SM neutrinos goes as $m_\nu \sim m_D^2/M$ [7]. Weakly interacting neutrinos can oscillate into these heavy sterile neutrinos, which can thus enter in the diagram of figure 1.2 and in the other diagrams responsible for LFV in the SM, enhancing the branching fraction of lepton flavour violating decays.

Many other models predict lepton flavour violation: they are built as extensions of the SM adding new symmetries or new particles and

LFV arises naturally. For this reason setting limits on the branching fraction of charged lepton flavour violating decays can constrain their parameter space or exclude them completely. Conversely an observation will give hints on the origin of the new physics.

To move beyond the SM one must combine results from many different measurements: direct searches for new particles at the energy frontier, neutrino oscillation measurements, electric and magnetic dipole moment measurements as well as searches for LFV. The discovery of a single charged lepton flavour violating decay alone will not be enough to provide sufficient information on the underlying mechanism; in many models, for example, there are strong correlations between the expected rates of various channels. Moreover we do not know which lepton flavour violating decay mode will first be discovered, for instance in the unconstrained minimal supersymmetric model (MSSM), tau LFV branching fractions can be within the experimental reach even with the strong experimental bounds on muon LFV [8]. Due to their mass, taus can decay also via lepton flavour violating channels kinematically not allowed to muons.

LFV in τ decays

Historically LFV in the tau sector has been studied in e^+e^- storage rings operating at a centre-of-mass energy near the mass of the $\Upsilon(4S)$ meson, 10.58 GeV, in particular by the BaBar [9], Belle [10] and CLEO [11] experiments. Tau pairs are produced via the process $e^+e^- \rightarrow \tau^+\tau^-$. Using knowledge of the kinematics of the centre-of-mass system, the event is separated into two hemispheres, each one containing the decay products of a tau. One tau is tagged via its decay into a charged lepton plus two undetected neutrinos, while the other, the signal, is required to decay into the final state of the lepton flavour violating decay under study. Since in the decay of the signal tau no undetectable neutrinos are produced, in the centre of mass frame the invariant mass and the energy of the observed final state are equal to that of the tau: m_τ and $\sqrt{s}/2$. It is therefore possible to define a signal region in the mass-energy plane, estimate the number of expected events in the signal region and, comparing it with the number of observed signal candidates, establish an upper limit on the branching fraction of the lepton flavour violating decay under study. In figure 1.3 there is a summary of these upper limits.

Experimental status

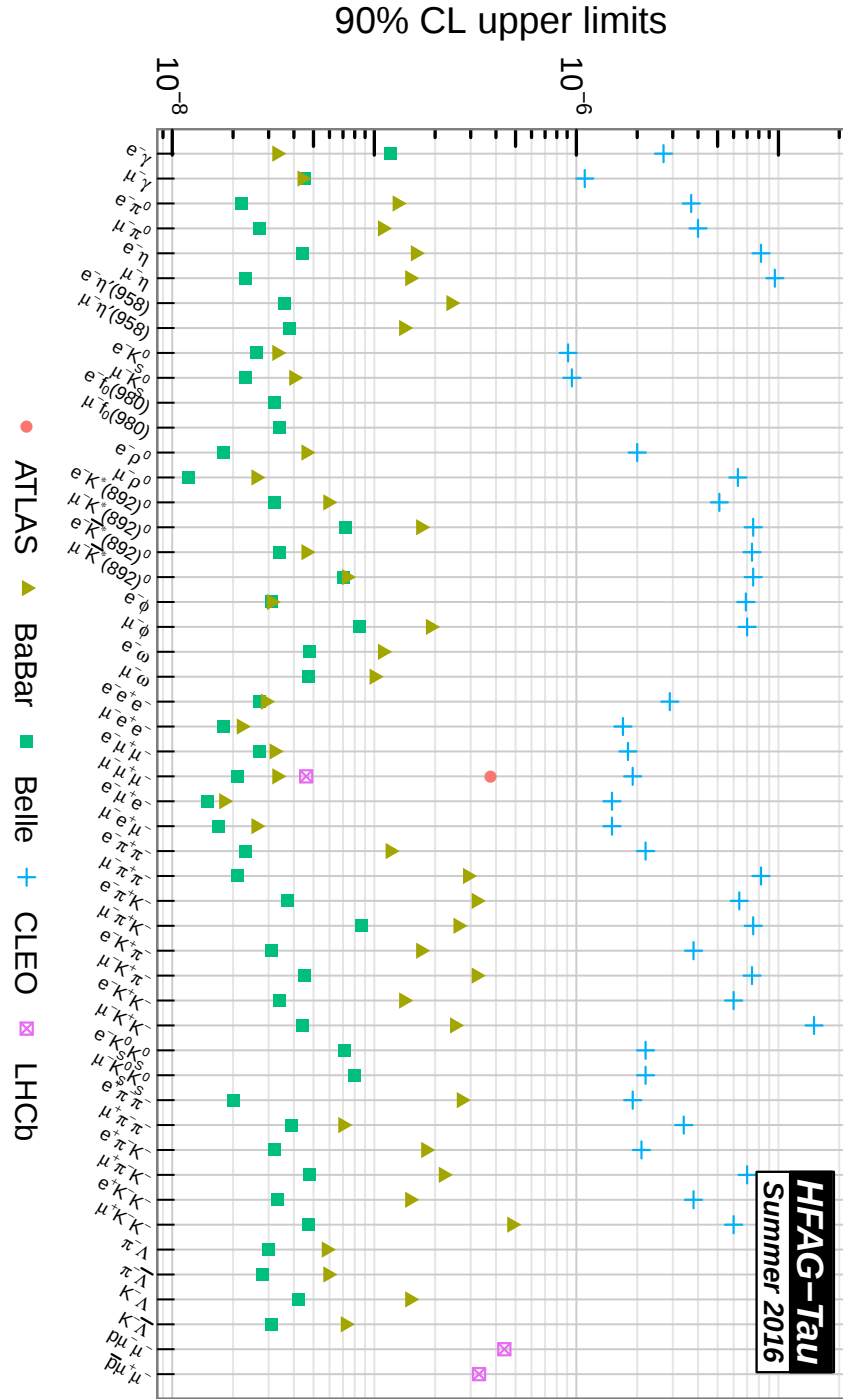


Figure 1.3: Summary of the 90% confidence level (CL) upper limits on tau lepton flavour violation branching fractions provided by the HFAG [12, 13].

1.3 CHARMLESS MULTIBODY BARYONIC B-MESON DECAYS

Mesons that contains a bottom antiquark and a lighter quark are called B mesons: the ground states are B^+ , B^0 , B_s^0 and B_c^+ depending on whether the lighter quark is an up, down, strange or charm quark respectively. These ground state mesons decay weakly as only the gauge bosons of the weak interaction can couple to two different kinds of quarks. This, together with the high mass that allows them to decay to a plethora of final states, makes B mesons an ideal laboratory to study CPV.

The violation of the CP symmetry is an essential ingredient to explain the matter-antimatter asymmetry that we observe in our current universe. As seen in Section 1.1 the only source of CPV in the SM is the phase of the CKM matrix; its contribution however is not sufficient to explain the matter-antimatter asymmetry observed [14]. Usually CP-violating effects in the quark sector are classified in three categories:

*CP violation in B
meson decays*

In decay when the rate of a particle M decaying into a final state f is different from the rate of its antiparticle $\bar{M}$ to decay into the final state $\bar{f}$, the charge conjugate of f .

In mixing involves only a neutral meson M and its antiparticle $\bar{M}$. As they have the same charge they can oscillate one into the other $M \leftrightarrow \bar{M}$. CPV can occur if the rate for $M \rightarrow \bar{M}$ is different from the rate for $\bar{M} \rightarrow M$.

In interference for neutral mesons M decaying to a final state f accessible also to its antiparticle $\bar{M}$, CPV can occur from the interference between the amplitudes for $M \rightarrow f$ and $\bar{M} \rightarrow f$.

To have CPV in decay it is necessary to have at least two amplitudes interfering: the decay rate is proportional to the square of the total amplitude so any overall phase of the total amplitude is not observable and only phase differences between amplitudes are. It is useful to separate the phase differences into two components: the difference of the so called *strong phases* $\Delta\delta$, that is CP-even, and the difference of the so called *weak phases* $\Delta\phi$, that is CP-odd. These names derive from the fact that in the SM $\Delta\phi$ originates from the phase of the CKM matrix and hence comes from the weak interactions while $\Delta\delta$ arises from final state interactions or from the presence of inter-

mediate resonances, both dominated by the strong interaction. The variable typically used to look for CPV in decay is

$$\mathcal{A}_{\text{CP}} = \frac{N(M \rightarrow f) - N(\bar{M} \rightarrow \bar{f})}{N(M \rightarrow f) + N(\bar{M} \rightarrow \bar{f})} \propto \sin \Delta\delta \sin \Delta\phi,$$

where N is the number of observed decays assuming the same initial number of M and $\bar{M}$ and no detection asymmetry. It is important to notice that, using this variable, CPV is observable ($\Delta\phi \neq 0$) only if a difference in strong phases is present ($\Delta\delta \neq 0$).

*Effective theories for
decays of B-mesons*

A good understanding of both weak and strong interactions is crucial to describe the decay of B mesons. Different energy scales are involved: $M_W \sim 80 \text{ GeV}$ is the energy scale typical of the weak interaction; the mass of the b quark $m_b \sim 4 \text{ GeV}$ is another important energy scale that sets the order of magnitude of the mass of the B meson and hence the energy available for the decay. It is important to consider also $\Lambda_{\text{QCD}} \sim 1 \text{ GeV}$, the energy scale at which QCD becomes non perturbative. As there is a clear hierarchy among these different energy scales, $M_W > m_b > \Lambda_{\text{QCD}}$, effective field theories (EFT) are the theoretical framework used to describe B meson decays. In this framework if a theory has two separate energy scales, a high-energy-low-distance scale and low-energy-large-distance one, it is possible to set an intermediate energy scale μ and write an effective Hamiltonian by “integrating out” the high energy fields [15],

$$\mathcal{H}_{\text{eff}} = \sum_i c_i(\mu) O_i(\mu).$$

The coefficients $c_i(\mu)$ are called *Wilson coefficients* and contain the high energy physics while O_i can be any operator compatible with the symmetries of the system. The sum is often infinite but, depending on their dependence on the energy μ , only a finite number of terms is relevant for describing low energy phenomena. This is the case, for example, of the *Fermi theory* of weak interactions where the heavy $W^\pm$ bosons are “integrated out” in the description of the weak decays of particles lighter than M_W and a four-fermion vertex, not present in the SM, is introduced. This is the starting point to describe B meson decays and the amplitude for a B meson to decay into a final state f can be expressed as

$$A(B \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | B \rangle = \frac{G_F}{\sqrt{2}} \sum_i \lambda_i^{\text{CKM}} c_i(m_B) \langle f | O_i(m_B) | B \rangle$$

where the dependence on the Fermi constant $G_F \simeq 1.166 \text{ GeV}^{-2}$ and on the CKM factors $\lambda_i^{\text{CKM}} = V_{q_1 q_2} V_{q_3 q_4}^*$ has been written explicitly. The Wilson coefficients c_i include the physics from the highest scales, including M_W , down to the scale m_B and can be computed perturbatively from the SM or from a beyond-the-SM theory. The hadronic matrix elements $\langle f | O_i(m_B) | B \rangle$ on the other hand contain the strong interaction effects and cannot be computed perturbatively as they involve energies lower than Λ_{QCD} . Two main approaches exist to estimate them. The first one exploits the approximate SU(3) symmetry of the SM under the exchange of the up, down and strange quarks [16]. The hadronic matrix elements of different decays are then connected to one another via Clebsh-Gordan coefficients so that it is possible to obtain the branching fraction of a decay if the branching fractions of its SU(3) partners are known. The accuracy of this approach depend on the size of SU(3)-breaking corrections, that depend on the ratio between the strange-quark mass and Λ_{QCD} . The second approach is based on the EFT paradigm [17, 18] and exploits the gap between the energy scales m_b and Λ_{QCD} by expressing the hadronic matrix elements in a double perturbative expansion in the QCD coupling constant α_s and in Λ_{QCD}/m_b . Three popular approaches are available in this regard: QCD factorisation (QCDF) [19], perturbative QCD (pQCD) [20] and soft-collinear effective theory (SCET) [21, 22]. They differ from each other in the details on how they separate the perturbative and non perturbative scales. These approaches have been successfully used to describe decays to two-body final states while a general description of multibody decays is still challenging.

Due to the their high mass, B mesons can decay to many different final states. These are classified depending on the particle content of the final state, for example depending on the presence or not of leptons in the final states, B decays can be divided into leptonic, semi-leptonic and hadronic. Focusing on hadronic decays, a further distinction can be made on the presence of particles containing a c quark. Charmless decays are not dominated by a $b \rightarrow c$ transition but proceed via $b \rightarrow u$ transitions, dominated by tree-level Feynman diagrams, and $b \rightarrow s$ and $b \rightarrow d$ transitions, dominated by penguin loop diagrams. Due to the small magnitude of the CKM matrix element V_{ub} , tree and penguin diagrams can have similar magnitude; their interference can thus give rise to CPV in decay. Another reason for interest in this class of decays is that beyond the SM particles could produce sizeable differences with respect to the SM expectations both

Charmless decays

in the branching fractions and in the amount of CP violation, by entering as virtual particles in the loops of the penguin diagrams.

Another classification of decays can be made depending on the number of final state particles. Decays can thus be divided into two-body or multibody decays. Multibody decays can proceed through different intermediate states and the paths from the initial to the final state interfere. Moreover while the kinematics of a two-body decay are totally determined by the mass of the initial and of the final state particles, the phase space of multibody decays is multidimensional: for example a three body decay is two-dimensional and a four-body one is five-dimensional. In the search for CPV this translates into the possibility to study not only global CPV but also one local in multibody decays as different regions of the phase space can present different levels of CPV. As an example, Figure 1.4 shows the raw asymmetry, which is the sum of the CP asymmetry, the production and the detection asymmetry, in $B^+ \rightarrow K^+ \pi^+ \pi^-$ where large variations are evident and the local CP asymmetries are as large as 60% [1]. In

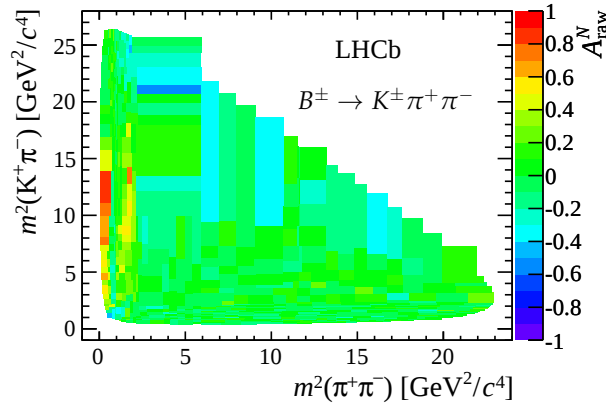


Figure 1.4: Measured raw asymmetry in Dalitz plot bins of background-subtracted and acceptance-corrected events for $B^+ \rightarrow K^+ \pi^+ \pi^-$ decays [1].

addition CPV can be searched for in multibody decays not only using the observable $\mathcal{A}_{CP}$ defined above but also $\hat{T}$ -odd observables. The *motion reversal* operator $\hat{T}$ [23, 24] is a unitary operator that reverses both momentum and spin three-vectors. Its action on helicities and momenta is identical to that of CP and differs from the anti-unitary operator T because it does not invert initial and final states. It is possible to define $\hat{T}$ -odd observables using triple products,

$$C_{\uparrow} = \vec{k}_1 \cdot (\vec{k}_2 \times \vec{k}_3),$$

where $\vec{k}_i$ are momentum or spin three-vectors of final state particles. By the properties of triple products, the three $\vec{k}_i$ must be different so it is possible to construct these observables in three body decays only if they involve non scalar particles so that it is possible to use spins¹. In four-body decays triple products can be constructed using only momenta. It is then possible to define the $\hat{T}$ -odd asymmetry for a given decay

$$\mathcal{A}_{\hat{T}} = \frac{N(C_{\hat{T}} > 0) - N(C_{\hat{T}} < 0)}{N(C_{\hat{T}} > 0) + N(C_{\hat{T}} < 0)}$$

and the corresponding asymmetry for its C-conjugate decay,

$$\overline{\mathcal{A}}_{\hat{T}} = \frac{N(-\overline{C}_{\hat{T}} > 0) - N(-\overline{C}_{\hat{T}} < 0)}{N(-\overline{C}_{\hat{T}} > 0) + N(-\overline{C}_{\hat{T}} < 0)}$$

where $\overline{C}_{\hat{T}}$ is defined analogously to $C_{\hat{T}}$ but with the final state particles of the C-conjugate decay; the minus sign in front of $\overline{C}_{\hat{T}}$ is a consequence of the parity transformation. The variable sensitive to CPV is then

$$a_{CP}^{\hat{T}\text{-odd}} = \frac{1}{2} (\mathcal{A}_{\hat{T}} - \overline{\mathcal{A}}_{\hat{T}}) \propto \cos \Delta\delta \sin \Delta\phi.$$

It is important to notice that this variable is sensitive to CPV ($\Delta\phi \neq 0$) even if no difference in strong phases is present ($\Delta\delta = 0$) so it is complementary to the variable $\mathcal{A}_{CP}$ defined above [25].

Due to their large mass, B mesons can also decay to baryons. In mesons' decays baryons are produced in baryon-antibaryon pairs as the difference between baryons and antibaryons is conserved. To decay into baryons a meson must thus have a mass larger than twice the mass of a proton. This means that, apart from B mesons and quarkonia, the only other meson baryonic decay possible is $D_s^+ \rightarrow p\bar{n}$. The first observation and studies of inclusive baryonic B decays was done by the ARGUS collaboration [26] in the late 1980s. The first observation of exclusive baryonic B-meson decays came with the B factories in the early 2000s with the first observation of $B^+ \rightarrow p\bar{p}K^+$ [27] and $B^0 \rightarrow p\bar{p}K^{*0}$ [28] by the Belle collaboration. The first observation of baryonic B_c^+ decays was made by LHCb in 2014 [29] which also recently reported the first observation of a baryonic B_s^0 decay [30], the last of the four B meson species for which a baryonic decay mode

Baryonic decays

¹ It is not possible to use three body decays involving only scalar particles because in the rest frame of the decaying particle the four-momentum of a final state particle can be expressed as a linear combination of the four-momenta of the other two particles due to energy-momentum conservation.

had yet to be observed. In contrast to their mesonic counterparts, two-body baryonic decays are more rare than multibody ones: the typical branching fraction of a B meson decaying into a baryon-antibaryon pair is around 10^{-8} – 10^{-7} while if a meson is also produced the branching fraction is $\mathcal{O}(10^{-6})$. The only two-body baryonic B meson decay observed to date is $B^0 \rightarrow p\bar{\Lambda}(1520)$ [31] while evidence has been obtained for $B^0 \rightarrow p\bar{p}$ [32] and $B^0 \rightarrow p\bar{\Lambda}$ [33].

Another interesting phenomenon is the *threshold enhancement* observed in the lower region of the baryon-antibaryon mass spectrum in many multibody baryonic B decays like $B^+ \rightarrow p\bar{p}\pi^+$ and $B^+ \rightarrow p\bar{p}K^{*+}$ shown in Figure 1.5, where the distributions obtained from data by Belle [28] are compared with a phase-space MC.

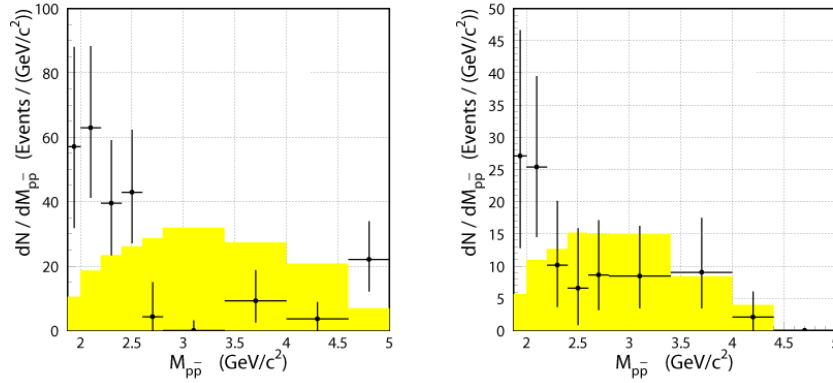


Figure 1.5: Experimental results from Belle for (left) $B^+ \rightarrow p\bar{p}\pi^+$ and (right) $B^+ \rightarrow p\bar{p}K^{*+}$. The shaded distribution is from the phase-space MC simulation with area normalised to signal yield [28]. In both cases a sharp peak is evident near the threshold of twice the proton mass at around 2 GeV.

phenomenon was first seen in $B^+ \rightarrow p\bar{p}K^+$ [27] and $B \rightarrow p\bar{p}D^{(*)}$ [34] and observed in other three-body B-meson decays [35] including decays with a baryon-antibaryon pair different from $p\bar{p}$ like $B^0 \rightarrow p\bar{\Lambda}\pi^-$ [36] and $B^+ \rightarrow \Lambda\bar{\Lambda}K^+$ [37] but also in other baryonic decays like $J/\psi \rightarrow p\bar{p}\gamma$ [38] and $\Upsilon(1S) \rightarrow p\bar{p}\gamma$ [39]. A simple short distance picture can explain this phenomenon and connect it to the branching fractions hierarchy in baryonic B decays [40]. As can be seen in Figure 1.6, to produce a baryon-antibaryon pair the two-body decay one energetic quark-antiquark pair must be emitted back to back by a gluon. That gluon must then be highly off mass shell and its emission is suppressed by the power of α_s/t , where t is the four-momentum square transferred through the gluon. In the three-body

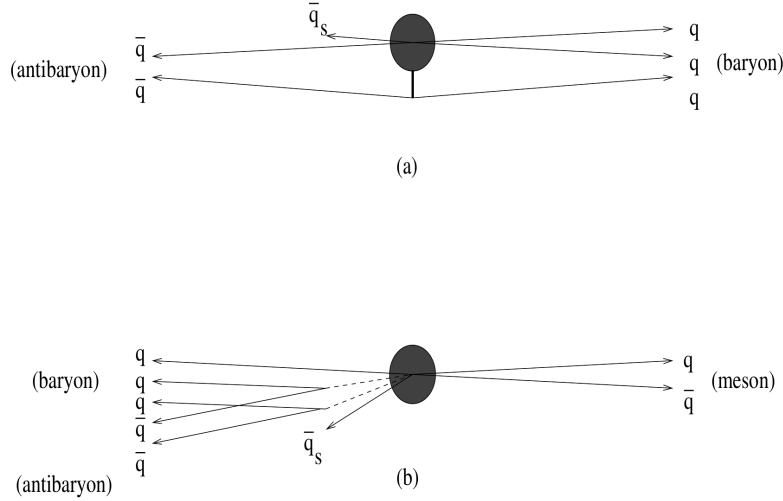


Figure 1.6: Short-distance picture in quarks and antiquarks for (a) two-body baryonic decay and (b) three-body baryonic decay. In the two-body decay (a) the virtual gluon (the thick vertical solid line) must split into $q\bar{q}$ while in the three-body decay (b) nearly on-shell gluons (broken lines) turn into $q\bar{q}$. The slow spectator antiquark is denoted by the short line q_s [40].

decay with an additional meson, a baryon-antibaryon pair can be emitted collinearly against the energetic meson in the final state. In this configuration a quark and an antiquark are emitted by a gluon nearly in the same direction so that the gluon is close to its mass shell and the short-distance suppression does not occur. A baryon-antibaryon pair is thus preferentially produced near to the threshold and the higher the mass the more its production is suppressed. This picture thus explains both the threshold enhancement observed in multibody decays and the suppression of two-body decays where the baryon-antibaryon pair cannot be produced close to the threshold because it must have the mass of the decaying B meson. This simple picture however does not explain all the aspects of this phenomenon, for example even if it explains the angular distribution of the protons in $B^+ \rightarrow p\bar{p}\pi^+$ decays it fails to do so for $B^+ \rightarrow p\bar{p}K^+$ [41]. Many theories have been proposed to explain this threshold enhancement. The first discussion was given by Hou and Soni using a pole model [42] followed by studies using the QCD naive-factorisation [43, 44]. Various interpretations include models with baryon-antibaryon bound states [45] or with a glueball as intermediate state [46]. The problem has been studied also with pQCD [47] or with a partial wave analysis [40]. A satisfactory explanation of all the aspects of this phe-

nomenon is however still missing and any new experimental input is welcome to guide the theoretical effort.

2

THE LHCb EXPERIMENT AT THE LHC

The present chapter gives a short description of the Large Hadron Collider (LHC) in Section 2.1 and focuses on the LHCb experiment in Section 2.2. Section 2.3 is devoted to a description of other important aspects common to any analysis done with data collected by the LHCb experiment such as the track reconstruction and the simulation.

2.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [48, 49, 50] is the world's largest and highest energy particle accelerator. Proposed and realised by the European Organization for Nuclear Research (CERN), it was designed to collide protons, as well as lead ions, at an unprecedented energy and rate, in order to address some of the most fundamental questions of physics.

2.1.1 LHC design and performances

The LHC lies in the 26.7 km long LEP [51] tunnel, situated at a depth of about 100 m underground near the border between Switzerland and France. The main design characteristics are listed in table 2.1.

Collisions occur between particles of the same charge and the tunnel contains two adjacent and parallel beam pipes where proton (or ion) beams travel in opposite directions and intersect at four points, where the main experimental halls are built and detectors are placed (see figure 2.1).

Some 1232 dipole magnets keep the beams on their circular path, while an additional 392 quadrupole magnets are used to keep the beams focused, in order to maximize the chances of interaction in the four intersection points, where the two beams cross. In total, over 1600 superconducting magnets are installed. Approximately 96 tonnes of liquid helium are needed to keep the superconducting magnets at their operational temperature of 1.9 K. The field in the dipole

Magnets

Table 2.1: LHC design parameters for p-p and Pb-Pb collisions [48, 49, 50].

Parameter	p - p	Pb - Pb
Circumference [km]		26.659
Beam radius at interaction point[μm]		15
Dipole peak field [T]		8.3
Design center of mass energy [TeV]	14	1148
Design luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	10^{34}	$2 \cdot 10^{27}$
Luminosity lifetime [h]	10	4.2
Number of particles per bunch	$1.1 \cdot 10^{11}$	$\sim 8 \cdot 10^7$
Number of bunches	2808	608
Bunch length [mm]	53	75
Time between collisions [ns]	24.95	$124.75 \cdot 10^3$
Bunch crossing rate [MHz]	40.08	0.008

magnets increases from 0.53 T to 8.3 T while the protons are accelerated from 450 GeV to 7 TeV.

Accelerating steps

Before being injected into the main accelerator, the protons are prepared by a series of systems that successively increase their energy (see figure 2.2). The first system is the linear particle accelerator (LINAC 2) generating 50 MeV protons, which feeds the Proton Synchrotron Booster (PSB). There the protons are accelerated to 1.4 GeV and injected into the Proton Synchrotron (PS), where they are accelerated to 26 GeV. Finally the Super Proton Synchrotron (SPS) is used to further increase their energy up to 450 GeV before they are injected into the LHC. Here the proton bunches are accumulated, accelerated (over a period of 20 minutes) to their peak energy, and finally circulated while collisions occur at the four intersection points (IP).

2.1.2 The LHC experiments

The Large Hadron Collider hosts seven different experiments; each experiment has a different composition and geometry of the subdetectors so that it is more specialised in a particular area of the research in particle physics.

ALICE

ALICE (A Large Ion Collider Experiment) [58] is a detector optimised for PbPb collisions, in particular for the study of the properties of matter at high temperature and high energy density generated by such collisions (Quark Gluon Plasma).

ATLAS & CMS

ATLAS (A Toroidal Lhc ApparatuS) [59] and CMS (Compact Muon Solenoid) [60] are two general-purpose detectors; they are built with a

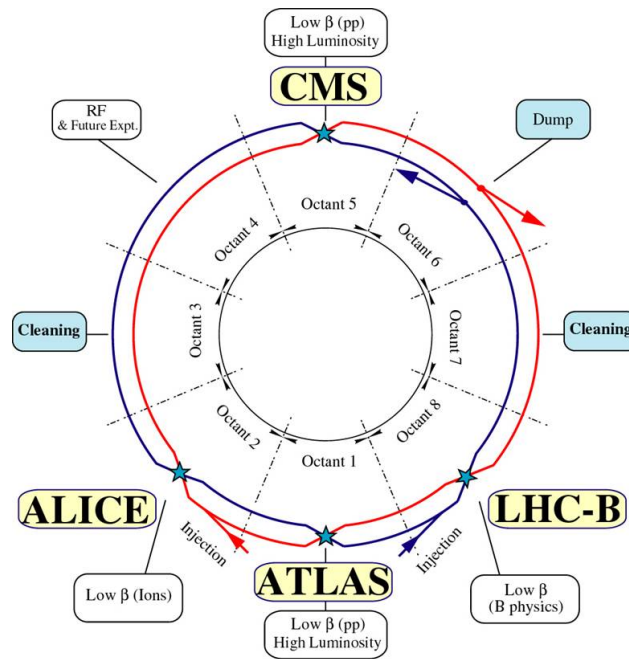


Figure 2.1: A schematic picture of the LHC layout [52].

CERN's Accelerator Complex

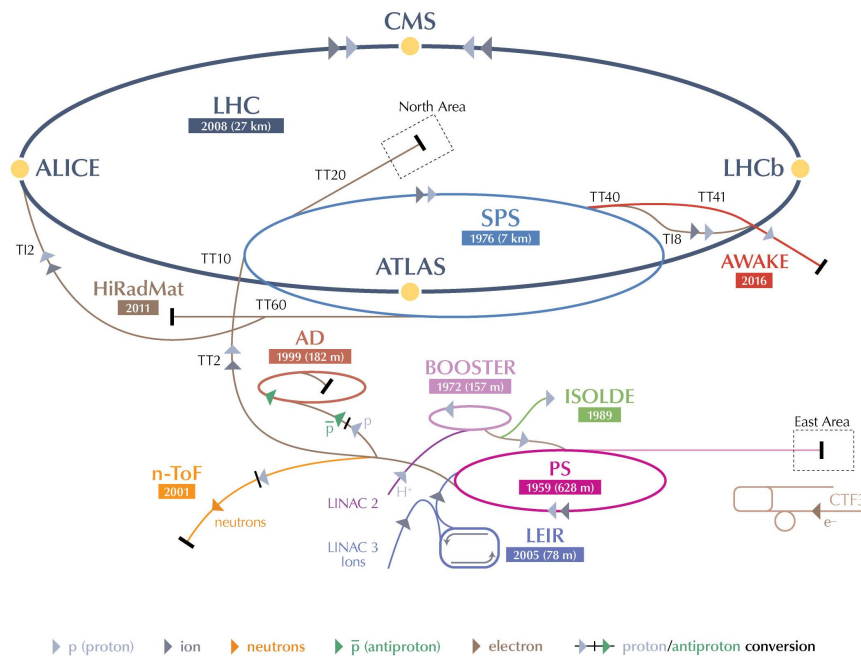


Figure 2.2: CERN's accelerator complex [53].

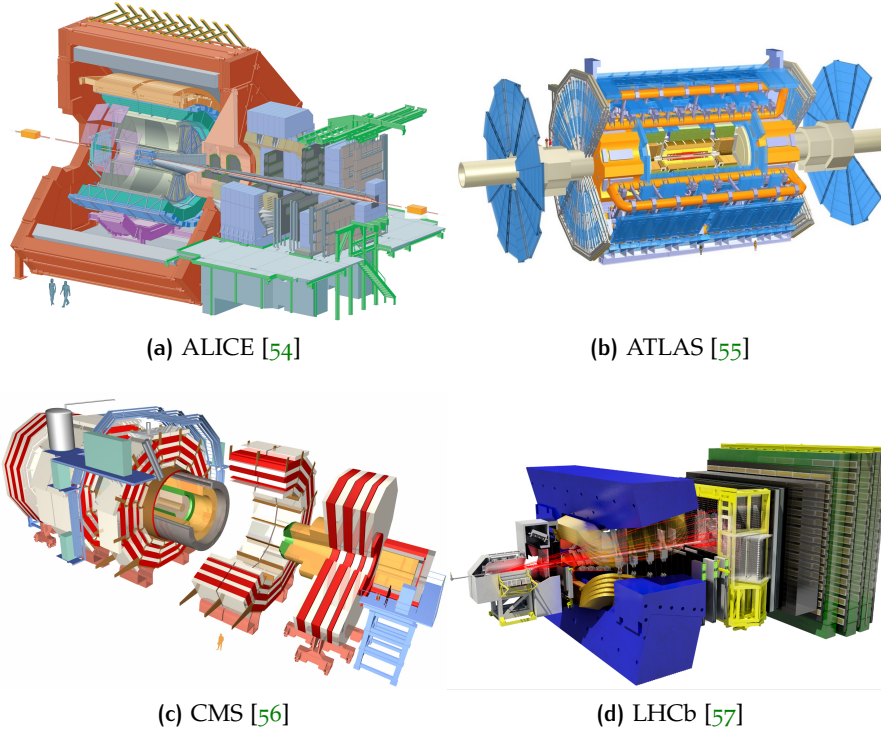


Figure 2.3: LHC's four main experiments

cylindrical geometry around the beamline. They have been designed for the discovery of new physics at the TeV scale, thus their subdetectors are optimised for the reconstruction of high energy objects with great efficiency and accuracy. These two detectors would be able to measure masses of new particles up to 3-4 TeV. While similar in their purposes, the design of the two detectors differs significantly, since different solutions were chosen for the configuration of the magnetic field. ATLAS uses a toroidal field produced by three sets of air-core toroids complemented by a small solenoid in the inner region, while CMS uses a solenoidal field generated by the world's largest superconducting solenoid.

LHCb LHCb [61] is specialised in studies regarding the physics of heavy quarks and heavy mesons with a particular attention to the b quark and its mesons; its detailed description can be found in Section 2.2.

TOTEM & LHCf TOTEM (TOTAl Elastic and diffractive cross section Measurement) [62] and LHCf ("f" stands for forward) [63] are forward detectors near CMS and ATLAS respectively; they are placed ~ 100 m from the interaction points of the main experiments to study diffractive physics happening in the very forward region of the collision. These detectors were put far from the interaction point so that the products of such

very forward (i.e. small angle with respect to the beam-line) inelastic or elastic collisions may exit the beam-pipe.

MoEDAL (Monopole and Exotics Detector at the Large Hadron Collider) [64] is a passive detector dedicated to the search for magnetic monopoles or other highly ionising stable or pseudo-stable massive particles. It is composed of plastic nuclear track detectors attached to the walls and ceiling around LHCb's vertex locator. The passage of a massive highly ionising particle would leave, in the plastic nuclear track detectors, damage that can be revealed by controlled etching in a hot sodium hydroxide solution.

MOEDAL

2.1.3 LHC data taking periods: Run I and Run II

LHC operations are organised in periods of data taking followed by long shutdowns during which maintenance and upgrade can be performed to the accelerators and the detectors. The first period of data taking, known as Run I, happened between 2010 and 2013. After an initial ramp up in energy during 2010, the nominal centre of mass collision energy was $\sqrt{s} = 7$ TeV during 2011 while in 2012 the collision energy was increased to 8 TeV. Between 2013 and 2015 the first long shutdown took place: many interventions were performed to the LHC to enable collisions at 14 TeV and also the other components of the accelerator chain were improved. Run II started in 2015 and will be concluded in 2018 with the beginning of the second long shutdown. In 2015 and 2016 the nominal centre of mass collision energy has been $\sqrt{s} = 13$ TeV.

Figure 2.4 shows the luminosity delivered by LHC to the general purpose detectors and the luminosity collected by LHCb. It can be noticed that the luminosity collected by LHCb is much smaller than that delivered to the general purpose detectors. The reason is that the beams are less focused at the LHCb interaction point; moreover they collide with an offset in the horizontal direction to reduce the luminous region and the offset is reduced during the fill to maintain a constant instantaneous luminosity and have a lower number of primary vertices per bunch crossing (1 or 2 in LHCb and approximately 20 in ATLAS and CMS). The development of the instantaneous luminosity in LHCb and in the general purpose detectors can be seen in Figure 2.5.

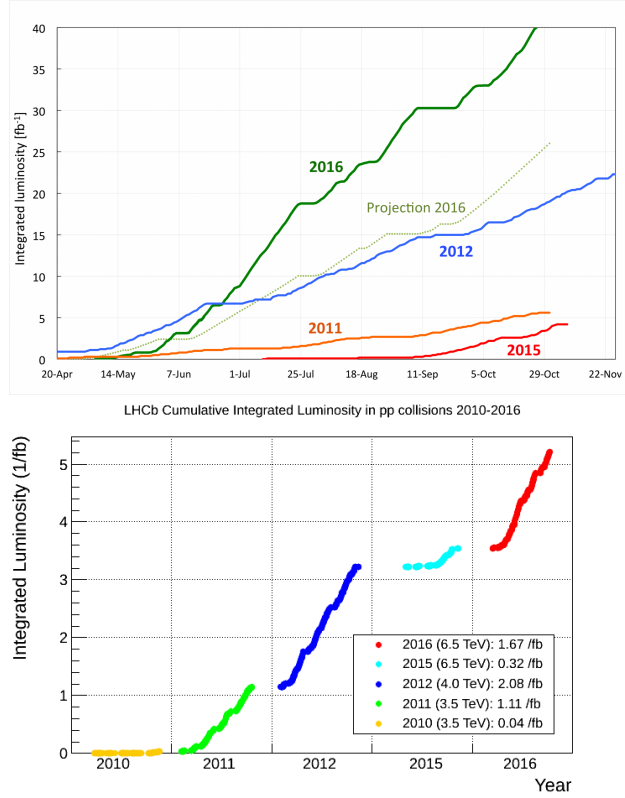


Figure 2.4: (*top*) Comparison of luminosity delivered by LHC in the different years of operations. Also the expected luminosity in 2016 is plotted to show how the luminosity delivered exceeded the expectations [65]. (*bottom*) Cumulative integrated luminosity collected by the LHCb experiment in the different years of data taking [66].

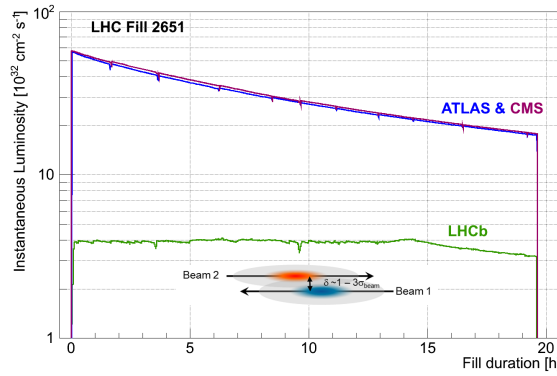


Figure 2.5: Development of the instantaneous luminosity for ATLAS, CMS and LHCb during LHC fill 2651. After ramping to the desired value of $4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ for LHCb, the luminosity is kept stable in a range of 5% for about 15 hours by adjusting the transversal beam overlap. The difference in luminosity towards the end of the fill between ATLAS, CMS and LHCb is due to the difference in the final focusing at the collision points [67].

2.2 LHCb

LHCb [61] is one of the four main experiments at the Large Hadron Collider (LHC); its main goal is the study of CP-violation and rare decays of beauty and charm hadrons. In proton-proton collisions $b\bar{b}$ and $c\bar{c}$ pairs¹ are produced mainly through gluon fusion and quark-antiquark annihilation with the two incident partons having very different momenta in the laboratory frame. For this reason the $b\bar{b}$ and $c\bar{c}$ pairs are preferentially produced around the beam axis in the same forward or backward direction, as shown in figure 2.6. The LHCb detector is thus built as a single arm forward spectrometer covering the pseudorapidity² range $2 < \eta < 5$.

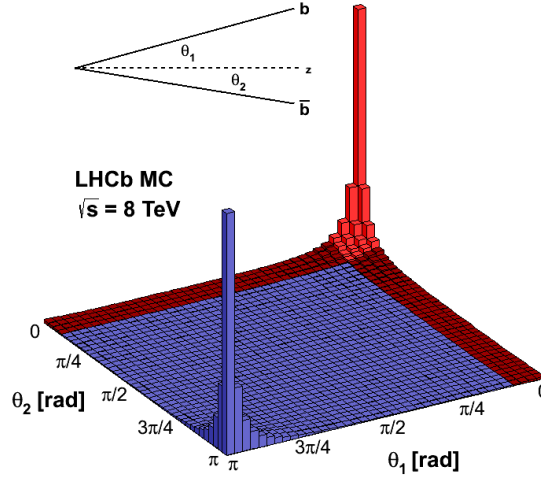


Figure 2.6: Simulated $b\bar{b}$ production angles in pp collisions at $\sqrt{s} = 7$ TeV. The region highlighted in red is inside the LHCb acceptance and ϑ_i are the angles from the beam direction [61].

In order to achieve its physics goals the LHCb detector must satisfy some experimental requirements:

Experimental requirements

- discriminate between primary and secondary vertices as b and c hadrons travel a distance of the order of a few mm before decaying,
- discriminate between particle species, both hadrons and leptons,

¹ Quark and antiquark pairs of the beauty and charm quarks.

² $\eta = -\ln\left(\tan\frac{\vartheta}{2}\right)$ where ϑ is the angle from the beam direction.

- trigger fast and efficiently for selecting interesting events among the large background.

In the following paragraphs a brief description of the various sub-detectors will be given, emphasising how they meet the experimental requirements. A schematic of the LHCb detector can be seen in figure 2.7.

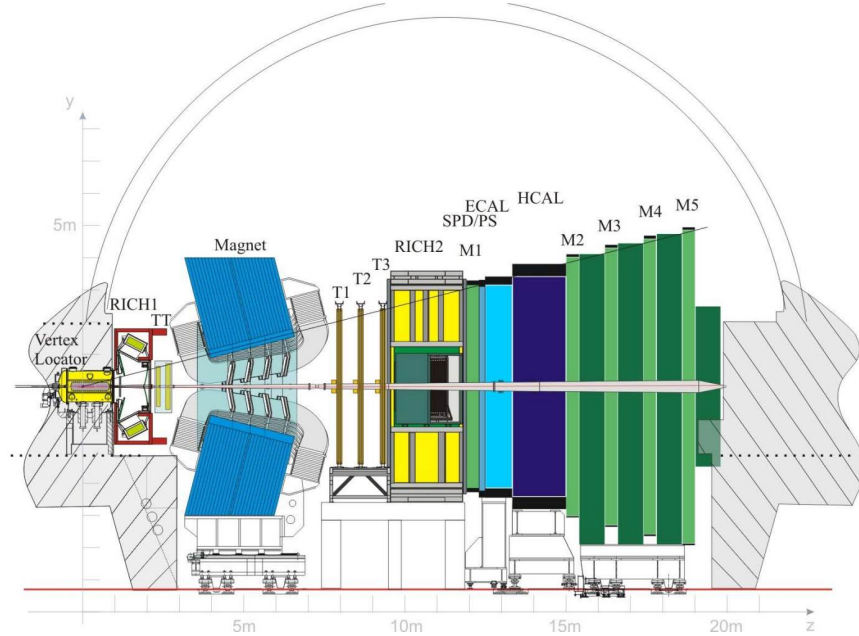


Figure 2.7: Schematic of the LHCb detector [61].

2.2.1 Vertex Locator

The VERtex LOcator (VELO) surrounds the interaction point of the proton beams and is designed to reconstruct tracks close to the interaction point in order to separate primary vertices, due to proton-proton interactions, from secondary ones, due to the decay of short-lived particles.

The VELO is made of two halves (A and C side), each half contains 21 modules composed of two semi-circular silicon strip sensors (R and Φ sensors). The modules are perpendicular to the beam line. The layout has been optimised to allow each track in the detector acceptance to traverse at least four modules. The modules are densely packed around the interaction point to reduce the extrapolation distance from the first measured hit to the vertex. In addition, four R sensors are placed upstream of the detector and can be used by the trigger. An overview of the VELO system can be seen in figure 2.8.

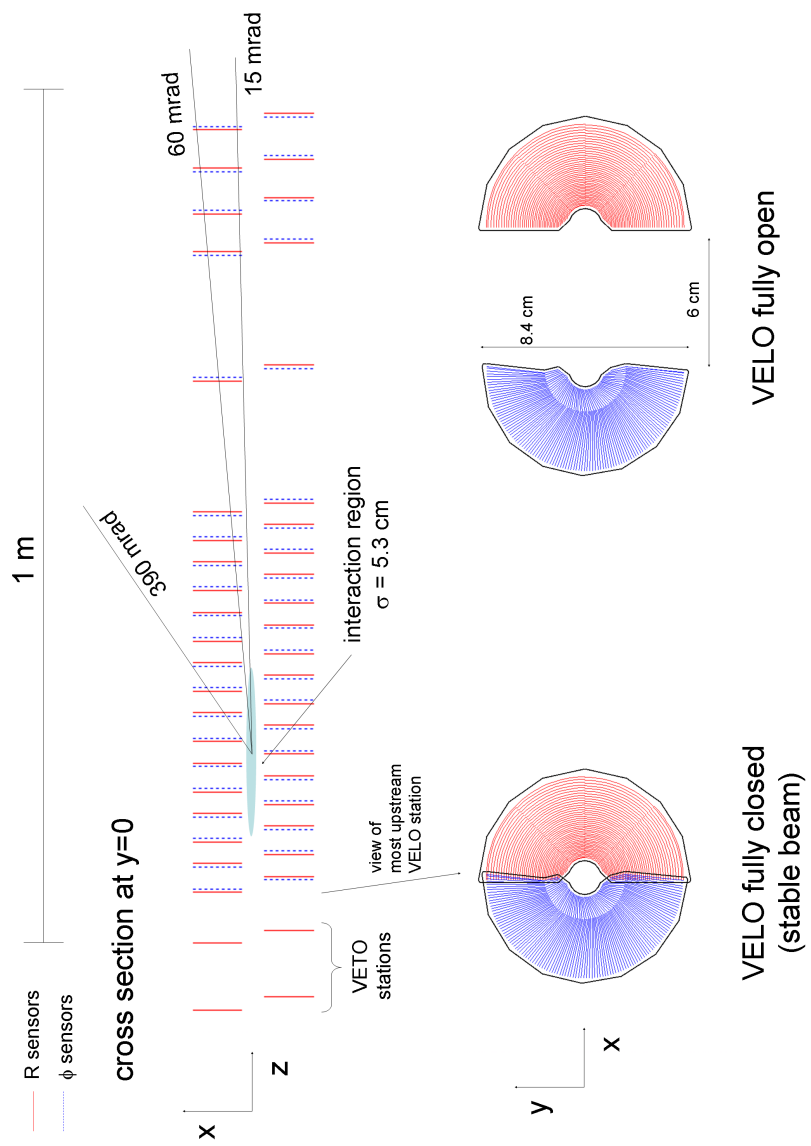


Figure 2.8: Schematic of the VELO layout [61].

During the data taking the active area starts at 8 mm from the nominal beam position, while during the beam injection the two halves are retracted for safety by 29 mm. When stable beam is declared the VELO is closed around the real position of the beam which can change from fill to fill. The corresponding modules in the two halves overlap by ~ 1.5 mm to cover the full azimuthal acceptance and to help with the alignment.

The VELO modules are placed inside an aluminium box in order to prevent RF pickup from the LHC beam; the side of the box is $300\text{ }\mu\text{m}$ thick to reduce the material budget, therefore the VELO must be operated in a secondary vacuum.

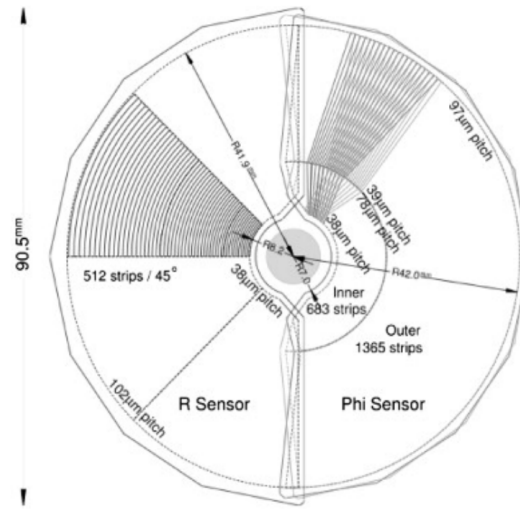


Figure 2.9: VELO R and Φ sensor geometry [61].

Figure 2.9 shows the details of the sensors. The R-sensors consist of circular silicon strips divided in four sectors each 45° wide. The strip pitch varies linearly from 38 to $102\text{ }\mu\text{m}$ to keep the strip occupancy approximately constant. The Φ -sensors contain straight silicon strips and are divided radially into two sections to reduce occupancy and to prevent too large a strip pitch at the outer edge. The strips are not perfectly radial but are inclined by a so-called stereo-angle of 10° (20°) for the inner (outer) part to improve the pattern recognition capability. The hit resolution depends on the strip pitch and the angle between the track and the normal to the sensor plane. The best hit precision measured is around $4\text{ }\mu\text{m}$ for an optimal projected angle of 8° and the smallest pitch of approximately $40\text{ }\mu\text{m}$. The minimum distance of a track to a primary vertex, the impact parameter, is measured with a resolution of $(15 + 29/p_T)\text{ }\mu\text{m}$ [67], where p_T is the component of the momentum transverse to the beam, in GeV.

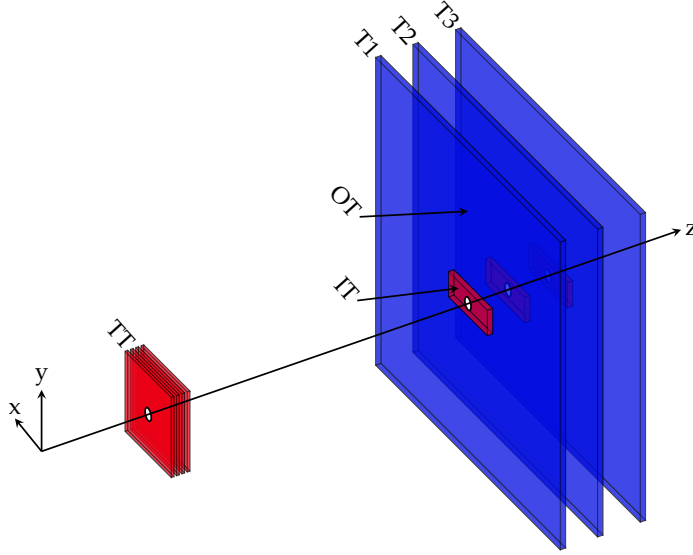


Figure 2.10: Layout of the LHCb tracking system, highlighting the IT and OT portions of T₁–T₃. The TT and IT collectively form the ST [68].

2.2.2 Tracking system

An overview of the tracking system can be seen in Figure 2.10. It consists of tracking subdetectors, positioned both downstream and upstream of the magnet in order to measure the momenta of the particles. LHCb uses a warm dipole magnet with a bending power of about 4 Tm and a peak strength of 1.1 T; its polarity can be switched in order to minimize systematic effects in the tracking. The upstream detector is the Tracker Turicensis (TT) composed of four layers of silicon strips. Three tracking stations (T₁–T₃) are located downstream of the magnet. As the particle flux is much higher in the central region than in the outer one, two different technologies are used: silicon microstrip for the Inner Tracker (IT) and straw tubes for the Outer Tracker (OT). The relative uncertainty on the momentum of a charged particle varies from 0.5% at low momentum to 1.0% at 200 GeV [67].

2.2.3 Cherenkov detectors

In a Cherenkov detector a particle travelling faster than the local speed of light in the detector emits a cone of radiation at a certain angle related to the speed of the particle and the refractive index of the medium. This information, combined with the momentum, allows for the mass to be determined and thus to discriminate between dif-

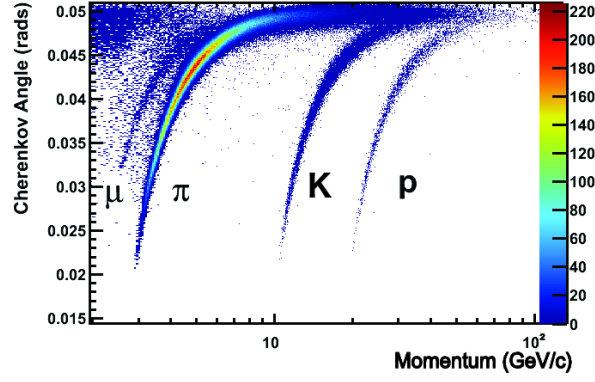


Figure 2.11: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator [69].

ferent particles such as pions, kaons and protons. This discrimination depends both on the radiator used and on the momentum of the particles. Different radiators should be used in order to achieve a good discrimination across the momentum spectra of interest. LHCb uses two Cherenkov detectors: RICH1, situated between the VELO and TT, which uses C_4F_{10} as radiator and covers the momentum range $1 < p < 60 \text{ GeV}$. RICH2 is situated after T_3 , uses CF_4 as radiator and covers the momentum range $15 < p < 100 \text{ GeV}$. RICH1 covers all LHCb acceptance and detects those low momentum particles which are bent out of the acceptance by the magnet, while RICH2 covers only the central region ($\eta < 3$) where high momentum particles are more abundant. Both RICH detectors use two sets of mirrors: spherical *primary* mirrors reflect the Cherenkov photons onto the plane *secondary* mirrors, that deflect the photons outside the LHCb acceptance where they are detected by Hybrid Photon Detectors (HPD). Figure 2.11 shows the different Cherenkov bands for muons, pions kaons and protons in RICH1.

2.2.4 Calorimeters

The calorimeter system provides particle identification, energy and position measurements for electrons, photons and hadrons. It is composed of four subdetectors: the Scintillating Pad Detector (SPD), the Pre-Shower detector (PS) the Electromagnetic CALorimeter (ECAL) and the Hadron CALorimeter (HCAL).

SPD/PS

The SPD and PS are two planes of rectangular scintillator pads between which there is a 15 mm thick lead converter which corresponds

to $2.5 X_0$. The SPD interacts only with charged particles allowing to discriminate between electrons and photons before showering occurs in ECAL. The lead converter is enough to make photons convert into electron-positron pairs that interact with the PS. This allow to discriminate between photons and neutral pions that do not interact with the PS.

Both the ECAL and HCAL consist of alternate layers of absorber material (lead for the ECAL and iron for the HCAL) and scintillator material. The ECAL provides a good energy resolution and fine transverse segmentation for electron, photon and neutral pion reconstruction. Its total radiation length ($25 X_0$) is enough to fully contain the electromagnetic showers. Its energy resolution is $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$ [67]. The HCAL's main purpose is to provide information to the hardware trigger. As a consequence it requires fast response even if not particularly high energy resolution: $\sigma_E/E \sim 70\%/\sqrt{E} \oplus 10\%$ [67]. For this reason and for space constraints its total thickness is not enough to contain the full hadronic shower. An overview of the calorimeter system can be seen in Figure 2.12.

ECAL

HCAL

2.2.5 Muon system

The muon system is composed of five stations, M1-M5. The first one is placed between RICH2 and SPD in order to give an improved evaluation of the muon p_T used in the hardware trigger not biased by the multiple scattering in the calorimeters. The other four modules are after the calorimeter system and are interleaved with 80 cm thick iron absorbers to select high momentum muons; a muon must have a momentum of at least 6 GeV to reach the final station. Two detector technologies have been used: the inner part of M1 is constructed from triple gas electron multiplier (GEM) detectors due to their radiation hardness while the rest of M1 and M2-M5 are composed of multi wire proportional chambers.

2.2.6 Trigger

The design bunch crossing rate at the LHC is 40 MHz³ but only a tiny fraction of the collisions produces a $b\bar{b}$ or $c\bar{c}$ pair with all the decay products inside the detector acceptance. Besides, only a tiny fraction

³ During Run I however LHC ran with a beam crossing rate of 20 MHz.

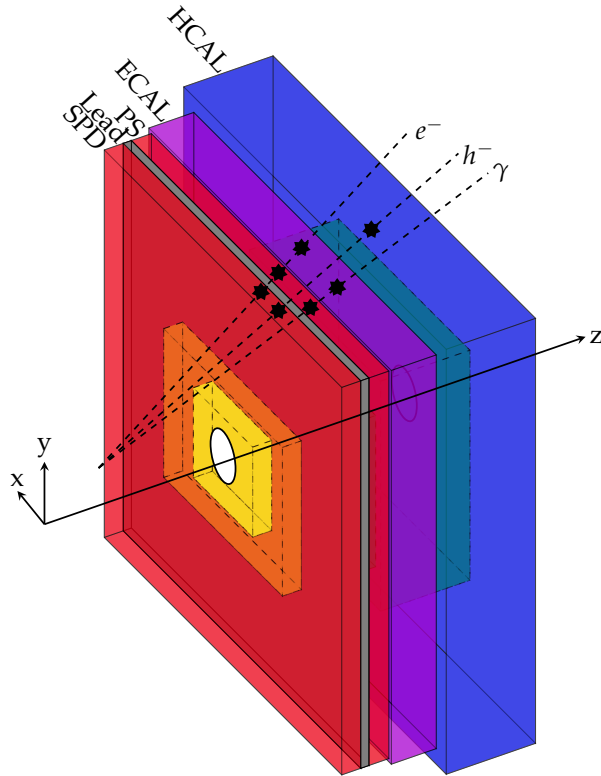


Figure 2.12: Layout of the SPD/PS, ECAL and HCAL showing the segmentation and the interactions of different particle species. The relative dimensions of the ECAL and HCAL are correct, but the z-scale of the SPD/PS is exaggerated [68].

of these contain a decay of interest⁴. Moreover only a fraction of the events (5 kHz in Run I and 12.5 kHz in Run II) can be stored onto disk for later analysis; thus it is necessary to have a trigger which identifies the signal of interest among the large amount of background. The LHCb trigger has three levels:

L0 is the hardware trigger which combines the information from the pile-up system, the calorimeters and the muon system. It selects events with high p_T and $E_T = \sqrt{m^2 + p_T^2}$ calorimeter clusters. It reduces the rate to 1 MHz.

HLT is implemented in C++ and runs on a dedicated processor farm, it has access to the full event information read out from the LHCb detector. It is divided into two stages:

HLT1 is composed of a set of inclusive selections requiring, for example, the presence of a track with high impact parameter or the presence of a muon with high p_T . It reduces the rate down to 40 – 80 kHz.

HLT2 can use the full event information to perform a complete reconstruction of the event and the reconstructed tracks are combined to select composite particles. There are many lines tailored for specific physics analysis.

More details about the LHCb experiment and its performance can be found in [67, 61].

2.3 SIMULATION, RECONSTRUCTION AND OTHER ASPECTS OF AN ANALYSIS

The ability to detect proton-proton collisions within LHCb is only the first step to perform a physics analysis. This section collects some information about the simulation, the reconstruction and other technicalities useful to perform analyses within the LHCb experiment.

2.3.1 Simulation

The production of Monte Carlo (MC) samples involves two phases: the generation and the simulation. In the first phase the proton proton interaction is simulated using an event generator, usually PYTHIA

⁴ The order of magnitude of their branching fraction is $10^{-3} - 10^{-5}$.

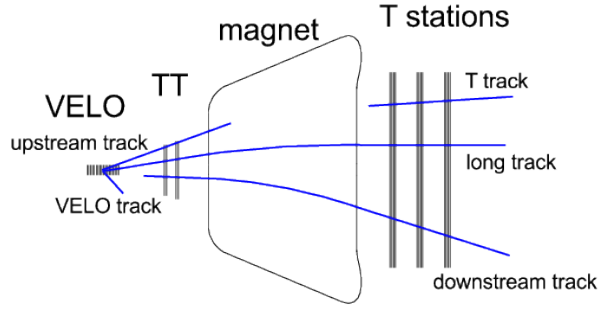


Figure 2.13: Tracking detectors and track types reconstructed by the track finding algorithms at LHCb. [77].

[70, 71], with a specific LHCb configuration [72]. The decays are simulated with EVTGEN [73] and the final-state radiation is generated using PHOTOS [74]. To avoid simulating events which cannot be reconstructed a selection is then usually applied to the generated particles before simulating the interaction with the detector. In the second step the interaction of the generated particles with the detector is simulated with GEANT4 [75, 76]; after that the simulation of the detector response to produce digitalised data is performed. The output of the digitalisation is then reconstructed using the same software used for real data.

2.3.2 Track reconstruction

The tracks – the trajectories of charged particles – are reconstructed in LHCb combining the hits – the measured positions of the intersections between tracks and detectors – coming from different tracking detectors (VELO, TT, IT and OT). Not all charged particles will leave hits in all subdetectors: some are produced in the backward direction and leave hits only in the VELO before flying outside the LHCb acceptance⁵, other low-momentum particles are deflected outside the acceptance by the magnet while particles coming from the decay of neutral long-lived particles like K_S^0 or Λ are often produced after the VELO and hence leave hits only in the tracker. The track reconstruction must accommodate these different possibilities, get a high efficiency in finding the tracks and, at the same time, have a low probability of reconstructing spurious tracks, also known as ghosts. In Figure 2.13 a schematic illustration of the various types of tracks is given.

⁵ Those tracks are crucial in the reconstruction of the primary vertices.

The first step of the track reconstruction is the *pattern recognition* in which independently in the VELO and in the T stations sequences of hits are grouped together and identified as coming from the same track. These VELO tracks and T tracks are then used as input to find long, upstream and downstream tracks. The long tracks are found either by extrapolating the VELO tracks into the T station and looking for matching hits, or by matching directly the VELO tracks with T tracks; TT hits are then added. Upstream tracks are found by extrapolating VELO tracks into the TT while the Downstream tracks combine T tracks with TT information.

Pattern recognition

The second step of the track reconstruction is the track *fitting* that, in LHCb, is done with a Kalman filter [78] that takes into account effects from multiple scattering and energy loss due to ionisation.

Kalman filter

Ghost tracks are then removed based on the *ghost probability*, the output of a neural network that takes as an input various quality variables like the track χ^2 and the number of hits in each subdetector. The ghost probability response is calibrated on simulated events to be the rejection rate of ghost tracks (eg. 30% of all the ghost tracks have a ghost probability of less than 0.3) The last step of the track reconstruction is the *clone killing* which consists of removing tracks that are also subtracks of other tracks, for example a VELO track used to build a long track.

Removal of ghosts and clones

The tracking efficiency for charged tracks that pass through the full tracking system varies as a function of the kinematics of the track and the occupancy of the detector and is above 95%.

2.3.3 Particle identification

Particle identification (PID) at LHCb is performed using information from the RICH detectors, the calorimeter system and the muon system. The PID of a single track can be obtained in a Cherenkov detector by measuring the Cherenkov angle and knowing the momentum of the track. In LHCb however this approach is not directly applied as the high number of tracks generated in each collision make it too computationally intensive to associate a Cherenkov cone to each track. What is done instead is to assign to each particle in the event a mass hypothesis and compute a likelihood taking into account the information from the various subdetectors. Default particles identities are then assigned, maximising this likelihood. One can then estimate if a particle is more likely to be a kaon or a pion by taking the differ-

PIDx

ence between the logarithms of the likelihood under the two different mass hypotheses for that particle. In LHCb these variables are called PIDK, PIDp etc. where the first mass hypothesis tested is that of a kaon or a proton while the second mass hypothesis is always a pion⁶.

ProbNNx

A second class of PID variables are commonly used: ProbNNx. They are the output of a neural network that takes as an input the various PIDx and also complementary information from all the subdetectors including the tracking detectors. The output is a variable between zero and one that can be interpreted as absolute probability of a particle to have a certain PID.

On average the efficiency to correctly identify a kaon is $\sim 95\%$ with a misidentification probability to identify a pion as a kaon of $\sim 5\%$ [67]. More detailed information on the performance of these variables will be given in the chapters devoted to the description of the physical analyses.

isMuon

The identification of a track as a muon is mainly based on the muon system. A boolean variable called *isMuon* is set to true based on the association of hits around the extrapolated trajectory of the track in the muon system. A search is performed within rectangular windows; the size of these windows in each muon station and the number of stations required to have hits are optimised to maximise the efficiency and, at the same time, provide low misidentification probabilities. On average the probability to correctly identify a muon is $\sim 97\%$ with a probability to misidentify a hadron (pion kaon or proton) as a muon an hadron between 1 and 3%.

2.3.4 Selection

In an LHCb analysis the selection usually involves three steps:

the trigger decides which events are interesting enough to be saved for a later analysis. Details of the LHCb trigger are given in Section 2.2.6;

the stripping is the centralised selection of interesting events after the reconstruction. A *stripping line* contains the instructions for reconstructing the particles of interest from the reconstructed stable particles and the corresponding selections to be applied. Usually the selection performed at this stage involves only independent selections on each variable even if using multivariate

⁶ All the other differences in log-likelihood can be obtained as differences of PIDx.

classifiers is also possible. A group of stripping lines selecting similar events is collected into a *stream* and their output is saved in a common file;

offline selection is the term used to refer to the end-user selection on the events which passed the trigger and stripping.

3

REAL-TIME ALIGNMENT OF THE LHCb VERTEX DETECTOR

The present chapter describes the novel real-time alignment and calibration adopted by LHCb for the Run II of data taking, focusing in particular on the real-time alignment of the VELO. In Section 3.1 a short introduction of the theory behind the track-based alignment of a tracking detector is given. Section 3.2 focuses on the real-time alignment and calibration strategy for the whole LHCb detector while Section 3.3 describes the preparation, the commissioning and the performance of the real-time alignment of the VELO.

3.1 TRACK-BASED ALIGNMENT OF A TRACKING DETECTOR

The spatial alignment of a detector is essential to achieve the best physics performance. For example at LHCb the correct alignment of the VELO is needed to discriminate between the primary vertices and the secondary vertices coming from the decay of particles containing b or c quarks; its accurate alignment allows excellent impact parameter and proper time resolution. A misalignment of the tracking system, on the other hand, would degrade the momentum and mass resolution. Figure 3.1 shows how the impact parameter resolution improves with a better VELO alignment while Figure 3.2 shows how an improved alignment greatly enhances the Υ mass resolution from 92 MeV with the first alignment to 49 MeV with the improved one.

To obtain a precise alignment of a particle detector three steps are necessary:

1. A precise assembly of the detector. This does not need to have the same level of precision as the desired resolution but should be within the same order of magnitude.
2. A precise metrology to have a first estimate of the real position of the detector elements, *i.e.* their displacement from the ideal position.

3. A software track-based alignment that allows the best determination of the real position of the detector elements. This is the topic of the rest of this section.

Alignment parameters

The position of each detector element, or *alignable*, is stored in a database used by the simulation and reconstruction software. This position is usually decomposed into two terms: the *design position*, which is stored in the database of the detector description, that does not change frequently, and a *misalignment* with respect to the design position, which is stored in a separate database together with all the parameters that can change frequently, such as the calibration ones. The aim of the track-based alignment is to estimate the misalignment of a set of alignables with respect to their expected positions, which might be their ideal ones or the result of the metrology or of a previous alignment. This misalignment can be parameterised by the six degrees of freedom shown in Figure 3.3: three translations T_x , T_y and T_z parameterised by the small quantities Δ_x , Δ_y and Δ_z and three rotations R_x , R_y and R_z around the corresponding axes that are parameterised by the small quantities Δ_α , Δ_β and Δ_γ . These alignment parameters are also known as *alignment constants*.

Reference frames

The global LHCb reference frame has the x axis toward the centre of the LHC ring, the y axis toward the “up” direction and the z axis along the beam direction forming a right-handed coordinate system; the origin is set to the design collision point. It is often useful to use also a local reference frame. For example in a semi-circular VELO module the origin of the reference frame is the centre of the circle.

Residuals

The principle behind the track-based alignment is illustrated in Figure 3.4: the misalignment of a detector element biases the reconstruction of the tracks. This also has the consequence that the average distance of the hits (the red dots in the figure) from the reconstructed track (in dotted blue in the figure) is larger than in the case without misalignment. It is thus useful to define for each hit of a track the *residual* (Figure 3.5) as the difference between the hit (the red dot in the figure) and the intersection between the track and the sensor plane (the green star in the figure). The residual can be *unbiased* if the hit under consideration is not used in the track fit (as is the case in the figure) or *biased* otherwise.

Residuals and misalignment

There is thus a correlation between the residuals and the misalignment parameters. Let’s consider the case of a planar alignable as illustrated in Figure 3.6. To derive the relation between the residual and the misalignment three positions should be considered: the intercept

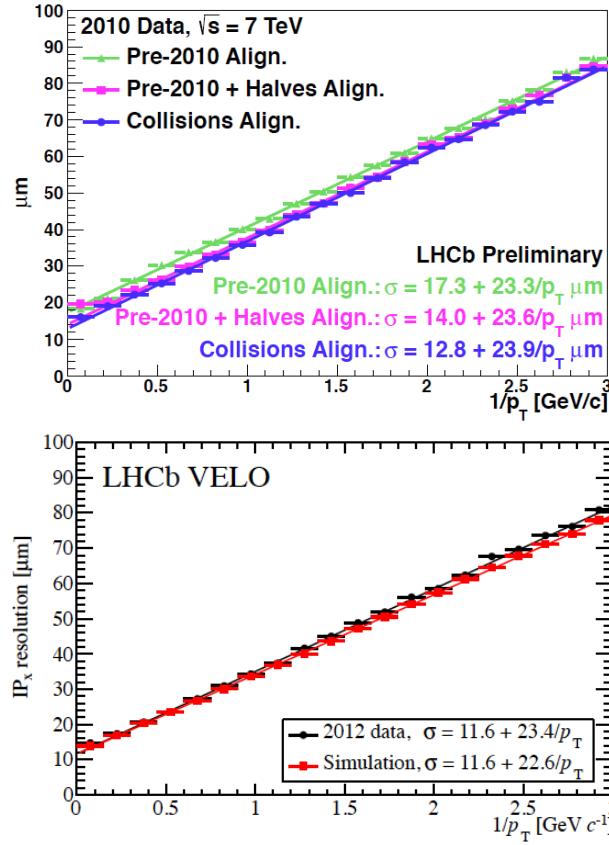


Figure 3.1: Impact parameter x resolution as a function of the $1/p_T$. The impact parameter x resolution at high p_T goes from (*top*) 17.3 μm before the 2010 alignment to (*bottom*) 11.6 μm after the 2012 alignment [67].

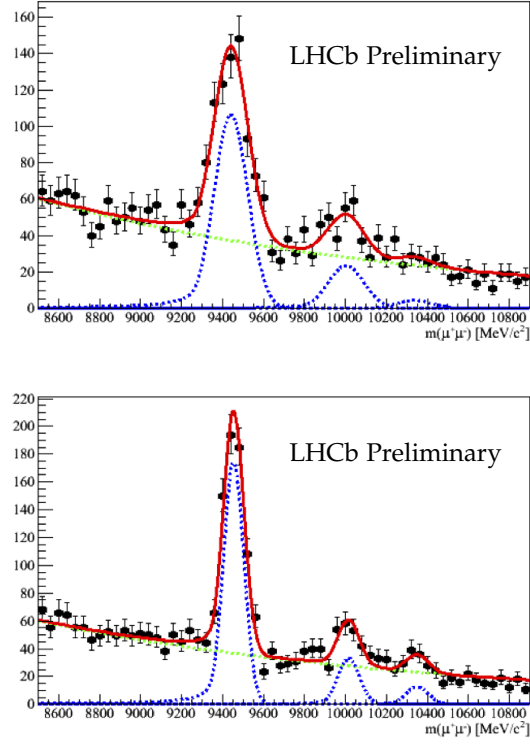


Figure 3.2: Invariant mass distribution for $\Upsilon \rightarrow \mu\mu$. The mass resolution is (top) 92 MeV with the first alignment and is enhanced to (bottom) 49 MeV with an improved alignment [79].

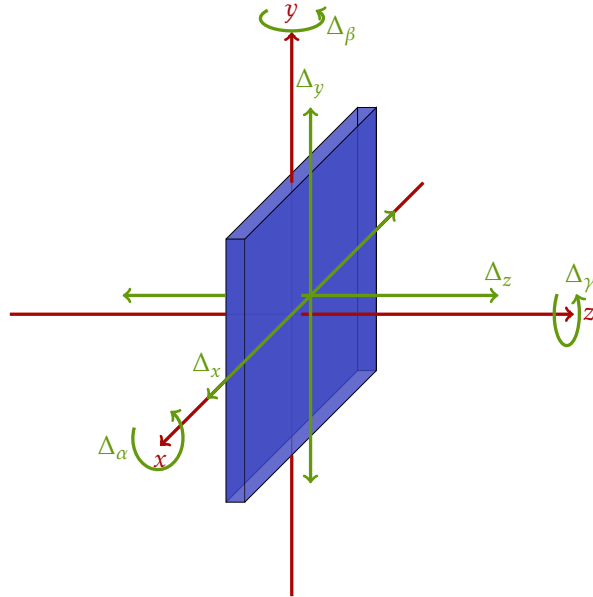


Figure 3.3: The six degrees of freedom for a detector element: the three translations Δ_x , Δ_y and Δ_z along the x , y and z axis, respectively, and the rotations Δ_α , Δ_β and Δ_γ around the x , y and z axis respectively [80].

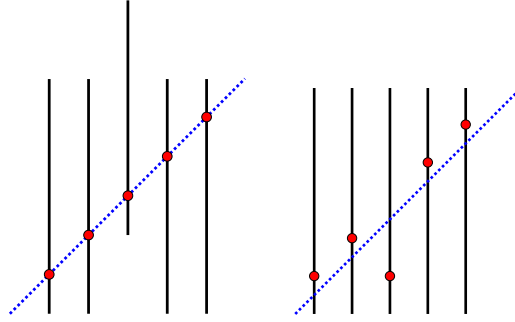


Figure 3.4: Schematic representation on how the misalignment of a detector element causes a bias in the track reconstruction. All the detector elements (*left*) in their real position with the central one misaligned and (*right*) in their design position. The hits (in red) are not aligned any more and this biases the track reconstruction.

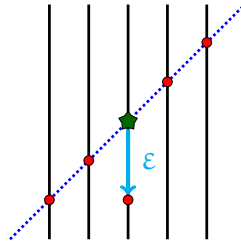


Figure 3.5: The unbiased residual ϵ of an hit is defined as the difference between the hit and the intersection between the track and the sensor plane.

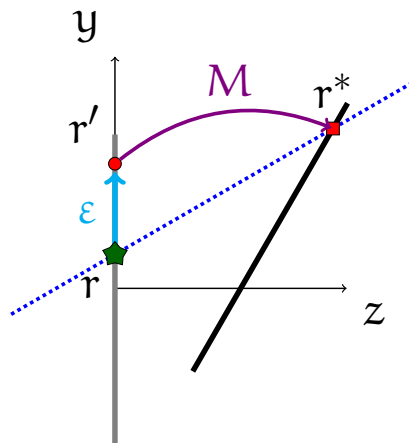


Figure 3.6: Representation of the expected position of a module in grey and its real position in black along with their intercepts with a track and the hit position. Refer to the text for details.

between the track (the blue dotted line) and the expected position of the sensor (the grey line) r , the intercept between the track and the real position of the sensor (the black line) r^* and the hit r' , the point corresponding to r^* if the sensor was in its expected position. The two points r' and r^* are thus linked by a roto-translation. Using homogeneous coordinates, and using the small-angle approximation, it is possible to express this with a matrix multiplication:

$$r^* = Mr' \implies \begin{pmatrix} x^* \\ y^* \\ z^* \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -\Delta_\gamma & \Delta_\beta & \Delta_x \\ \Delta_\gamma & 1 & -\Delta_\alpha & \Delta_y \\ -\Delta_\beta & \Delta_\alpha & 1 & \Delta_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 0 \\ 1 \end{pmatrix},$$

where $z' = 0$ because of the choice of the reference frame in Figure 3.6. As r and r^* are both on the track they are linked by the equation

$$r^* = r + th \implies \begin{pmatrix} x^* \\ y^* \\ z^* \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \\ 0 \end{pmatrix} \cdot h$$

where h is a parameter to be determined, t_x and t_y are the gradients of the track ($t_x = \tan \vartheta \cos \phi$, $t_y = \tan \vartheta \sin \phi$ with ϑ and ϕ the polar and the azimuthal angle) and $z = 0$. It follows that

$$r' = M^{-1}(r + th) \implies \begin{pmatrix} x' \\ y' \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & \Delta_\gamma & -\Delta_\beta & -\Delta_x \\ -\Delta_\gamma & 1 & \Delta_\alpha & -\Delta_y \\ \Delta_\beta & -\Delta_\alpha & 1 & -\Delta_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x + t_x h \\ y + t_y h \\ h \\ 1 \end{pmatrix}.$$

From the z component of this equation it is possible to obtain an expression for h

$$h = \frac{\Delta_z + \Delta_\alpha y - \Delta_\beta x}{1 + \Delta_\alpha t_x - \Delta_\alpha t_y} \simeq \Delta_z + \Delta_\alpha y - \Delta_\beta x;$$

hence, neglecting the second-order terms,

$$\begin{cases} \varepsilon_x &= x' - x \simeq -\Delta_x + \Delta_\gamma y + t_x(\Delta_z + \Delta_\alpha y - \Delta_\beta x) \\ \varepsilon_y &= y' - y \simeq -\Delta_y - \Delta_\gamma x + t_y(\Delta_z + \Delta_\alpha y - \Delta_\beta x) \end{cases} \quad (3.1)$$

where ε_x and ε_y are the components of the residual. In general it is thus possible to estimate the alignment parameters for the different elements of a detector by exploiting the correlation between residuals

and misalignments. In the example above only a single alignable is misaligned while in any practical application it is necessary to estimate the misalignment of several detector elements, with the consequence that the track parameters, and hence the residual distributions for one alignable, depend on the misalignment of the other alignables.

There are two classes of methods to estimate the misalignments from the distribution of residuals: local and global. In the local ones the misalignment of each alignable is estimated independently and the correlations between different alignables are not taken into account. An example is the *histogram based* alignment. In a perfectly aligned scenario the distribution of the residuals should be a Gaussian function centred on zero, with the detector resolution as width. A misalignment causes a shift in the distribution of the residuals. In Equation (3.1), neglecting the rotations and the terms proportional to t_x and t_y – in the approximation of tracks perpendicular to the plane of the alignables – the residual in x and y is the opposite of the corresponding misalignment translation. It is thus possible, for a simple detector, to estimate these misalignments just by looking at the residual distributions. For more complex geometries, where the simplifications adopted here are not valid, it is more difficult to extract several alignment parameters but it can partially be done by fitting with Equation (3.1), or a similar one, the distribution of the residuals as a function of the position of the residuals in the sensor's plane. This approach is, for example, adopted in the alignment of the relative position between R and Φ sensors in a VELO module [81]. Residuals histograms, even if not directly used to estimate the alignment parameters, are an invaluable monitoring tool to determine if a detector is well aligned. The main downside of a local approach is that it does not take into account the correlations between alignables introduced by the bias on the track parameters caused by the misalignment of other detector elements. The way to overcome this is to adopt an iterative procedure: perform a first alignment of all the alignables, then recompute the track parameters and perform the alignment again, repeating this procedure until convergence is reached.

*Histogram based
alignment*

In a global alignment method the correlations among the misalign-

Global Alignment

ment of different detector elements are taken into account by defining a global χ^2 ,

$$\begin{aligned}\chi^2 &= \sum_i \chi_i^2 = \sum_i \varepsilon(\xi_i, \alpha)^T V^{-1} \varepsilon(\xi_i, \alpha) \\ &= \sum_i [r'(\alpha) - r(\xi_i)]^T V^{-1} [r'(\alpha) - r(\xi_i)],\end{aligned}$$

where the index i runs over the tracks considered. A matrix notation is adopted so that ε is the vector with all the residuals of a track and V is the covariance matrix of the hit measurements, generally diagonal. The expected position of the hits r , depends on the track parameters ξ_i , which are in common only with the other residuals of the same track, while the hit positions r' depend on the alignment parameters α which are in common among all the residuals¹.

The χ_i^2 can be minimised individually to obtain the track parameters of each track if the alignment constants are maintained constant. Defining the matrix derivative of the residual with respect to the track parameters

$$H \equiv \frac{\partial \varepsilon}{\partial \xi_i},$$

the minimisation of χ_i^2 can be expressed as

$$0 \equiv \frac{d\chi^2}{d\xi_i} = 2H^T V^{-1} \varepsilon \simeq 2H^T V^{-1} (\varepsilon_0 + H\Delta\xi_i),$$

where ε has been linearised around the initial values of the track parameters $\xi_{i,0}$. The optimal track parameters are then

$$\xi_i = \xi_{i,0} - CH^T V^{-1} \varepsilon_0 = \xi_{i,0} - \left(\frac{d^2\chi^2}{d\xi_i^2} \right)^{-1} \bigg|_{\xi_{i,0}} \frac{d\chi^2}{d\xi_i} \bigg|_{\xi_{i,0}},$$

where the matrix

$$C \equiv (H^T V^{-1} H)^{-1} = 2 \left(\frac{d^2\chi^2}{d\xi_i^2} \right)^{-1} \bigg|_{\xi_{i,0}}$$

is, by error propagation, the covariance matrix of the parameters ξ_i . This method requires the inversion of a matrix the size of the square of the number of parameters. Another important point is that an exact solution is provided only if the residuals depend linearly on the parameters to determine; otherwise it can be seen as an application

¹ Assigning to r instead of to r' the dependence on the parameters α would also be possible and would lead to the same conclusions.

of the Newton-Raphson method to find the zero of the χ^2 derivative and several iterations may be needed.

It would be conceptually possible to adopt the same procedure to minimise the global χ^2 simultaneously for the track parameters ξ_i and the alignment parameters α ; however it would require the inversion of a matrix the size of the square of the number of tracks times the number of track parameters per track plus the number of alignment parameters. For a complex detector where samples of more than 10k tracks are needed, this matrix inversion becomes very challenging. However it is possible to exploit the fact that only the parameters α are common to all the tracks while the parameters ξ_i are not. It is thus possible, for a given value α_0 , to compute for each track the solution $\xi_i(\alpha_0)$ and use its derivative to predict how the tracks change as a function of α . In this way the problem will be reduced to first finding the track parameters, inverting a number of matrices equivalent to the number of tracks and having the size of the square of the number of track parameters per track, and then inverting a matrix the size of the square of the number of the alignment parameters. This approach is illustrated in Figure 3.7. To derive the results shown

Millipede

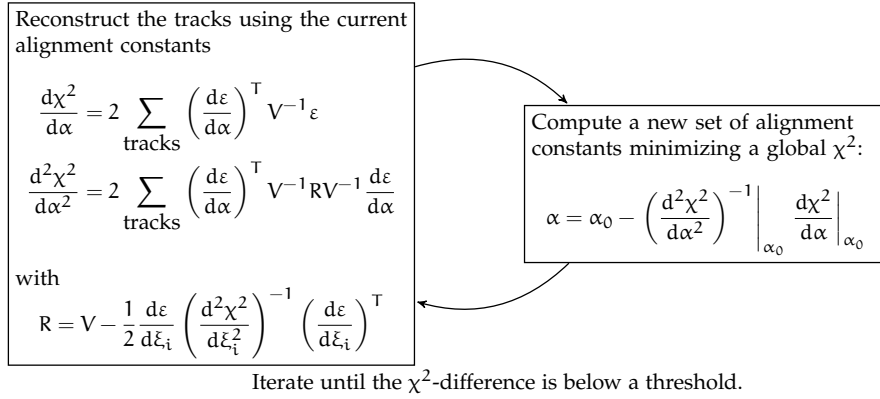


Figure 3.7: Representation of iterative χ^2 minimisation. The vector ε contains the residuals of a single track, V is the covariance matrix of the hits while R_{track} is the covariance matrix of the residuals. Refer to the text for details.

in Figure 3.7, the dependence on the alignment constants of the track parameters is made explicit: $\xi_i = \xi_i(\alpha)$. The total derivative with respect to alpha becomes then

$$\frac{d}{d\alpha} = \frac{\partial}{\partial\alpha} + \frac{d\xi_i}{d\alpha} \frac{\partial}{\partial\xi_i} \simeq \frac{\partial}{\partial\alpha} - A^T V^{-1} H C \frac{\partial}{\partial\xi_i},$$

where the partial matrix derivative of the residual with respect to the parameters α has been introduced,

$$A \equiv \frac{\partial \varepsilon}{\partial \alpha},$$

second order derivatives have been neglected, and the expression for $\frac{d\xi_i}{d\alpha}$ has been obtained by exploiting the fact that, once the χ^2 is at its minimum, there should be no change in the derivative of the track χ^2 with respect to ξ_i ,

$$\frac{d}{d\alpha} \frac{\partial \chi^2}{\partial \xi_i} = 0 \quad \Rightarrow \quad \frac{d\xi_i}{d\alpha} = -\frac{\partial^2 \chi^2}{\partial \alpha \partial \xi_i} \left(\frac{\partial^2 \chi^2}{\partial \xi_i^2} \right)^{-1}.$$

It is then possible to obtain the expressions

$$\begin{aligned} \frac{d\chi^2}{d\alpha} &= 2 \sum_{\text{tracks}} A^T V^{-1} (V - HCH^T) V^{-1} \varepsilon = 2 \sum_{\text{tracks}} A^T V^{-1} \varepsilon, \\ \frac{d^2 \chi^2}{d\alpha^2} &= 2 \sum_{\text{tracks}} A^T V^{-1} (V - HCH^T) V^{-1} A = 2 \sum_{\text{tracks}} A^T V^{-1} R V^{-1} A, \end{aligned}$$

where optimal track parameters have been used such that $H^T V^{-1} \varepsilon = 0$. The matrix $R \equiv V - HCH^T$, the residuals' covariance matrix, has also been introduced. The expressions for the χ^2 derivatives can then be used to obtain the alignment parameters

$$\alpha = \alpha_0 - \left(\frac{d^2 \chi^2}{d\alpha^2} \right)^{-1} \bigg|_{\alpha_0} \frac{d\chi^2}{d\alpha} \bigg|_{\alpha_0}.$$

This method to avoid the inversion of a prohibitively large matrix is known as the Millipede method [82]. It gives the same result as inverting the large matrix because all the correlations are taken into account in the residuals' covariance matrix. If the residuals depend only linearly on the track and alignment parameters, then an exact solution is obtained, otherwise some iterations are needed. The difference in χ^2 produced by one iteration is equivalent to the significance of the variation of the alignment parameters,

$$\Delta \chi^2 = (\alpha - \alpha_0)^T \frac{d\chi^2}{d\alpha} \bigg|_{\alpha_0} = -(\alpha - \alpha_0)^T \text{Cov}(\alpha_0)^{-1} (\alpha - \alpha_0),$$

where α_0 are the initial alignment parameters and $\text{Cov}(\alpha_0) \equiv \left(\frac{d^2\chi^2}{d\alpha^2} \right)^{-1} \Big|_{\alpha_0}$ is the covariance matrix for the alignment parameters. It is thus convenient to judge the convergence of the method by the χ^2 variation.

While finding the alignment parameters by minimising the bias in the distributions of residuals, either with a local or a global method, particular care must be taken to avoid *weak modes*. This term refers to misaligned configurations of the detector elements that allow good track fits with residual distributions centred on zero. A typical example is a global translation or a global rotation of all the detector elements; other common weak modes of a detector made of a series of planes, like the VELO, are the *shearing* ($\Delta_x(z) \propto z$ or $\Delta_y(z) \propto z$) and the *scaling* ($\Delta_z(z) \propto z$). Local alignment methods are particularly affected by weak modes but they also appear in the χ^2 of the global method. There they correspond to unconstrained degrees of freedom: the eigenvectors of the matrix $\left(\frac{d^2\chi^2}{d\alpha^2} \right) \Big|_{\alpha_0}$ with small or null eigenvalues that leads to poorly converging alignment. Unconstrained degrees of freedom can be removed in various ways. It is possible to fix the position of some alignable so that, for example, a global translation is no longer possible. An alternative way is to use *Lagrange constraints* adding to the χ^2 terms like $\lambda f(\alpha)$ where λ is known as a *Lagrange multiplier* and $f(\alpha)$ is a function of the alignment parameters that should be fixed to zero. Minimising the χ^2 with respect to λ then leads to the requested identity $f(\alpha) = 0$. It is possible to add to the χ^2 also *Gaussian constraints* of the form $[(\alpha - \tilde{\alpha})/\sigma_\alpha]^2$ where, for example, $\tilde{\alpha}$ and σ_α come from a survey. These kinds of constraints, besides fighting weak modes, are particularly useful when aligning different levels of granularity simultaneously, for example aligning one VELO half with respect to the other and simultaneously aligning the modules within their VELO half, because the position of different levels of granularity is known with different precisions.

Weak modes

As seen in Section 2.3.2, during the reconstruction, the tracks are fitted using a Kalman filter [78] and not by minimising a χ^2 . The reason is that, using a Kalman filter, it is easy to take into account effects from multiple scattering and energy loss due to ionisation. Figure 3.8 shows a schematic representation of how this method works. In each

Kalman Filter

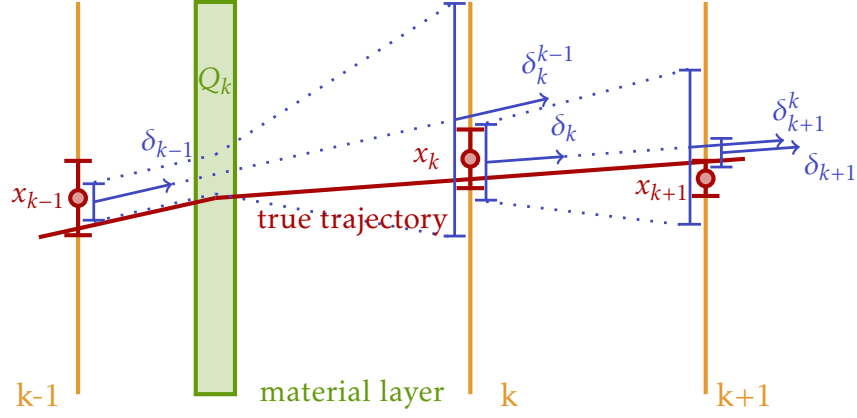


Figure 3.8: Illustration of the Kalman Filter fit showing the evolution of a track state δ_k from prediction to filtering [80].

position in z a track can be represented by a vector, called *state* and its covariance matrix. In LHCb, tracks are parameterised by the vector

$$\delta_z = \begin{pmatrix} x \\ y \\ t_x \\ t_y \\ q/p \end{pmatrix}$$

where t_x and t_y are the gradients of the track and q/p is the ratio between the electric charge and the momentum. Other parametrisations are also possible. In a perfectly deterministic scenario, in the absence of multiple scattering or energy loss, the knowledge of the magnetic field is enough to predict the state of the track in any position, once its initial value is known,

$$\delta_z = f(\delta_{z_0}),$$

where f is known as a *transport function*. For the application of the Kalman filter the above equation is discretised: as the detector is made of a discrete series of active detector elements, between which there can be some material that causes multiple scattering or energy loss, the state of the track is described only at the position of the active detector elements, labelled in the figure by the index k . It is thus possible to predict the state in the element k knowing the state

at the previous element $k - 1$; in particular the transport function can be linearised by introducing the *transport matrix* F ,

$$\delta_k^{k-1} = f(\delta_k) \simeq F_k^{k-1} \delta_k.$$

The prediction of the state does not take into account the presence of the random noise introduced by multiple scattering and energy loss. This information can be included in the covariance matrix. Calling C_k the covariance matrix of δ_k the predicted covariance matrix of δ_k^{k-1} is

$$C_k^{k-1} = F_k^{k-1} C_k (F_k^{k-1})^T + Q_{k-1}$$

where the first term is the usual transformation formula for a matrix and the second term represents the random noise whose effect is to increase the uncertainty on the prediction. This is the first step of the Kalman filter and it is known as *predicting*. The *predicted state*, which represents the information on the track parameters coming from the measurements on the preceding detector elements, can be combined with x_k , the *measured state* of the track on that detector element, and V_k , its covariance matrix; this step is known as *filtering*. Defining the residual

$$\varepsilon_k = x_k - \delta_k^{k-1},$$

and the *Kalman gain*

$$K_k = C_k^{k-1} (C_k^{k-1} + V_k)^{-1},$$

the state of the track at the element k , δ_k , and its covariance matrix C_k are

$$\begin{aligned} \delta_k &= \delta_k^{k-1} + K_k \varepsilon_k \\ C_k &= (1 - K_k) C_k^{k-1}. \end{aligned}$$

It is thus possible to start from the first detector element and apply the Kalman filter iteratively to the following ones up to the last. Only the state on the last element has been built using the information from all the measurements along the track. To propagate the information backward smoothing equations are applied or, alternatively, the Kalman filter can be run in the opposite direction and the weighted average can be considered; this last step is known as *smoothing* and allows to obtain δ_k^n and C_k^n , the state and its covariance that take into account the information from the measurement on all n detector ele-

ments. It is important to notice that the Kalman filter only uses the correlation matrix of a single state and does not compute a correlation matrix among all states δ_k^n .

Kalman filter and alignment

To be able to use the Kalman filter also in the alignment of the detector is desirable, both because it allows to take into account the multiple scattering and energy loss, but also for consistency as it is the method used for the track fitting. An essential element of the χ^2 based global alignment of Figure 3.7 is the covariance matrix of all the residuals of a track R . By default however the Kalman filter does not compute the correlation between the state of the track in different detector elements. For example $C_{k,l}^n$, the covariance between δ_k^n and δ_l^n is not computed, while the “diagonal” terms $C_{k,k}^n \equiv C_k^n$ are. It is possible however to compute also the “off-diagonal” terms [83]

$$\begin{aligned} C_{k,k+1}^n &= C_k F_k^T (C_{k+1}^k)^{-1} C_{k+1}^n \\ C_{k,l}^n &= C_{k,k+1}^n (C_{k+1}^n)^{-1} C_{k+1,l}^n \end{aligned}$$

and obtain an expression for the covariance of the residuals that can be used in the equations of Figure 3.7 to obtain the alignment parameters,

$$R_{k,l} = (V_k) \delta_{k,l} - H_k C_{k,l}^n H_l^T$$

where $\delta_{k,l}$ is the Kronecker delta and, as before, H is the derivative of the residuals with respect to the track parameters.

Vertex and mass constraints

The knowledge of the covariance of the residuals allows also to add additional constraints with respect to the ones discussed above; for example it is possible to include vertex constraints, requiring that all the tracks share a common vertex, or mass constraints, requiring that the invariant mass computed from a subset of tracks has a certain value. These constraints do not appear directly in the definition of the χ^2 but add additional correlations among tracks. These correlations cause a change in the residuals and the correlation matrix $R_{k,l}$ used to compute the alignment constants [84]. These kinds of constraints are very effective in fighting against weak modes because, for example, a scaling of the detector elements has a significant impact on the reconstructed masses. This alignment method based on the Kalman filter is the default one used in LHCb.

3.2 REAL-TIME CALIBRATION AND ALIGNMENT

In Run II the LHCb collaboration revolutionised its trigger and data acquisition strategy with the adoption of the real-time alignment and calibration and the possibility to perform full data analyses with offline-quality directly on the output of the trigger. *Real time* is defined as the interval between the collision in the detector and the moment the data are sent to permanent storage. The event reconstruction performed in the trigger farm is denoted as *online reconstruction* while the additional reconstruction performed after the data has already been sent to permanent storage is denoted *offline reconstruction*.

The rate of collisions doubled between Run I and Run II going from 15 MHz to 30 MHz while the output rate of events saved on disk changed from 5 kHz in Run I to 12.5 KHz in Run II. In Run I the online reconstruction performed by the trigger was simpler and quicker than the offline one in order to meet time constraints; the final detector calibration and an improved alignment were obtained offline on triggered data and the data used for most of the physics results was processed at the end of the year using the latest alignment and calibration parameters. During 2012 LHC spent only approximately 30% of its time in stable running due, for example, to planned technical stops, machine development phases and the time between data taking fills. In order to optimise the usage of the event filter farm where the software trigger is run, the farm nodes were equipped with local storage space and 25% of L0 output was buffered into the nodes and processed during the LHC downtime.

*Trigger conditions
in Run I and Run II*

In Run II this deferral strategy has been exploited even further as shown in Figure 3.9; after L0 and HLT1 all the selected events are buffered on disk allowing more time to process a single event (150 ms/event in Run I, 350 ms/event in Run II). The automatic calibration and alignment is performed in the trigger farm within a few minutes. The same offline reconstruction is run in HLT2 thanks to this larger time budget together with the reduced time required by the improved track reconstruction. When the new alignment and calibration are available, and if a significant variation is observed, a change of run² can be triggered. The new constants are updated for the new run to be used online by the two stages of the software trigger, and offline, for every subsequent reconstruction and selection. The events

*Trigger strategy in
Run II*

² A run is a unit of data taking with homogeneous conditions that last up to one hour and it is uniquely identified by a sequential run number.

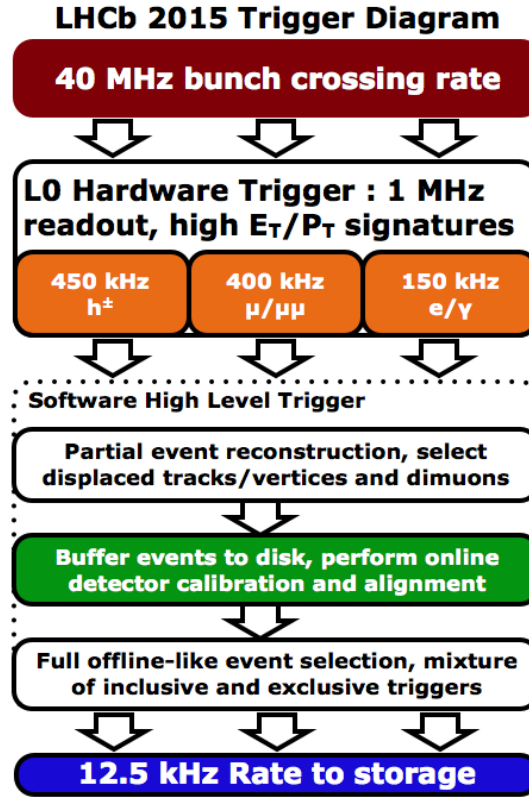


Figure 3.9: Trigger strategy for Run II: after the hardware stage and a first software stage based on a partial reconstruction, the selected events are buffered to disk while the real-time calibration and alignment are performed. The second stage of the software trigger performs the same reconstruction performed offline using the same calibration and alignment [85].

collected in the previous run, during the automatic alignment and calibration processing, are still reconstructed in HLT2 and offline with the previous constants for consistency with HLT1. As an exception the RICH calibration constants are updated for each stage after HLT1, both online, in HLT2, and offline as the previous trigger stages do not rely on hadron identification requirements.

Advantages of the new trigger strategy

This strategy has several advantages: firstly it minimises the difference between the online and offline performance, allowing to use a more effective trigger selection that, thanks to the complete calibration of the RICH detectors, can take advantage of the hadron identification information. For example charm physics is limited by trigger output rate constraints; using hadron identification in the trigger allows to a higher selection efficiency and purity for the doubly Cabibbo suppressed modes and, at the same time, satisfy the output rate constraints by pre-scaling the more abundant Cabibbo favoured

modes. Secondly it ensures the stability of the alignment quality and hence of the physics performance. Finally, since in Run II the last level of the trigger performs the same reconstruction as the one performed offline, it is possible to run some physics analyses, with the same offline performance, directly on the output of the trigger using the special stream of data known as *Turbo stream* [85], for example the J/ψ and prompt charm cross sections were measured with this method at the beginning of Run II [86, 87]. This approach also has the advantage that, for some events, it will be possible to save only the information on the signal candidate tracks (~ 5 kB per event) instead of all the electronic signals recorded from the detector (~ 70 kB per event). This decrease of more than one order of magnitude in the event size will allow a higher selection rate, for example in the charm analyses that used trigger lines which were pre-scaled due to output rate requirements during Run I. It is obvious that the real-time alignment and calibration becomes essential when the raw event is not saved.

Two main kinds of tasks are defined: alignment tasks and calibration tasks. The real-time evaluation of the alignment and the calibration uses for each task a dedicated data sample selected by HLT1, and is performed at regular intervals: either at the beginning of the run, fill, or less frequently depending on the task. They are performed in a few minutes using the nodes of the trigger farm and the alignment or calibration constants are updated only if the difference from the previous values is significant, in order to ignore random fluctuations due to the different input data samples used. In the following sections the different kinds of tasks are described.

Tracking alignment

As explained in Section 3.1, the track-based alignment of the tracking detectors at LHCb is based on an iterative procedure where the residuals of a Kalman filter are minimised. The relevant equations are summarised in Figure 3.7 from where it is clear that there are two distinct steps. The computation of the χ^2 derivatives can be parallelised by reconstructing the tracks and computing part of the sum over different events on different nodes. The partial sums can then be added and the new alignment parameters calculated on a single node. For this reason two different alignment tasks are defined.

- The *analyser* performs the track reconstruction based on the input set of alignment constants and evaluates the sums from the

equations in the left box of Figure 3.7. Many instances run in parallel within the ~ 1700 nodes of the HLT farm. In order not to compete with the HLT1 processes just one instance is run per node.

- The *iterator* collects the output of all the analysers, sums them and finds the new alignment constants using the equation in the right box of Figure 3.7.

After the initial configuration, the analysers read the initial alignment constants and run on the events assigned to them. The partial sums computed by the analysers are written on a fixed location of a shared file system when all the events have been analysed. The iterator then reads the output of the analysers, combines them and computes a new set of alignment constants. The analysers then start a new iteration of the alignment procedure starting from the new alignment constants. The iterations continue until the difference of the χ^2 between two successive iterations falls below a threshold.

The alignment procedure for the different components of the tracking system is performed at the beginning of each fill, but the frequency of updates is different. The alignment of the VELO is evaluated first and, as it can change for each fill, if a significant variation of alignment parameters is found, a change of run is triggered and the new alignment constants are used from the following run. An update of the VELO alignment parameters happens approximately every 3-4 fills. The alignment of the tracker is run after the alignment of the VELO. These alignment constants usually change every few weeks and, if a significant variation is found, the new alignment parameters are applied from the following fill. The alignment procedure of the muon stations is also run at the beginning of each fill after the tracker's but only serves as monitoring, as its alignment is not expected to vary except in case of hardware intervention. More details on the real-time alignment of the VELO are given in Section 3.3.

Rich mirror alignment

Both RICH detectors have two sets of mirrors: photons are reflected off a primary mirror onto a secondary mirror, from where they are deflected out of the LHCb acceptance onto the photon detection plane. The RICH mirror alignment follows the same general procedure of the tracking alignment: there is a task performed in parallel by the analysers while the computation of the alignment constants is per-

formed by the iterator on a single node. The RICH mirror alignment relies on the fact that a misalignment of the mirrors causes the Cherenkov ring on the photodetector plane to be shifted with respect to its expected position. The projected track coordinate is not at the centre of the ring, thus the Cherenkov opening angle (Θ) varies as a function of the azimuthal angle (ϕ) in the photodetector plane:

$$\Delta\Theta = \Theta - \Theta_0 = T_x \cos(\phi) + T_y \sin(\phi), \quad (3.2)$$

where T_x and T_y are the components of the misalignment of the projected track coordinate with respect to the expected position and Θ_0 is the Cherenkov angle calculated from the momentum of the track and the refractive index of the radiator using pion mass hypothesis [69]. The analysers perform the photon reconstruction and fill histograms of the $\Delta\Theta(\phi)$ distribution for each pair of mirrors on different events. The iterator collects all the histograms and combines them. It fits the combined histogram using Equation (3.2) and extracts the alignment constants. The procedure is iterated until the variations are below a threshold. Figure 3.10 shows for one mirror the distribution of $\Delta\Theta$ as a function of ϕ before and after the mirror alignment.

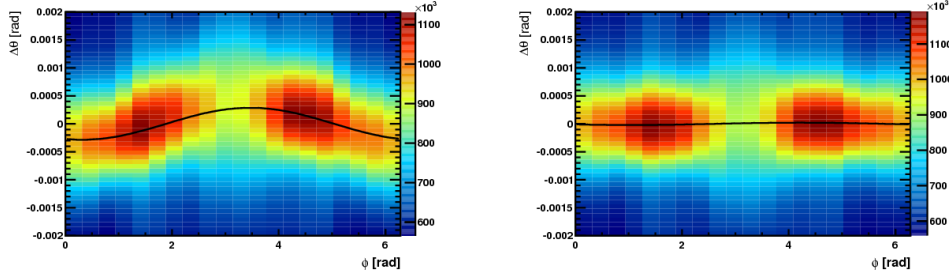


Figure 3.10: Difference between the measured and expected Cherenkov angle $\Delta\Theta$ as a function of the azimuthal angle ϕ before (left) and after (right) the mirror alignment for one mirror [69].

RICH calibration

The RICH automatic calibration consists of calibrating the RICH refractive index and the Hybrid Photon Detector (HPD) images. Both these calibrations are evaluated and updated every run. The refractive index of the gas radiators depends on the ambient temperature and pressure, and the exact composition of the gas mixture and so can change in time. These quantities are monitored to compute an expected refractive index, but this does not have a high enough precision and needs to be corrected. The distribution of the difference

between the reconstructed and expected Cherenkov angles is fitted to extract the scale factor to correct the expected refractive index.

HPDs are used to detect Cherenkov photons. They consist of vacuum tubes separated from the radiator gas by a quartz window and a photocathode. The photoelectrons produced are focused onto a silicon pixel array using an accelerating voltage [69]. The anode images are affected by magnetic and electric fields, and so are cleaned and a Sobel filter [88] is used. The evaluation of the new calibration constants does not require an iterative procedure and can be obtained by fitting the relevant distribution on a single node.

Global time alignment of the Outer Tracker

In the straw tubes that compose the Outer Tracker, the drift time measured may be different from the time estimated from the distance of the track to the wire. This is mainly due to the difference between the collision time and the LHCb clock and is common to all modules. A delay in the electronic readout, different for each module, may be partially responsible for this difference, but its contribution is small in comparison and constant in time, and thus it can be calibrated offline. The automatic procedure is performed every fill and computes a single parameter which accounts for the global time alignment by fitting on a single node the distribution of the difference between the measured and the estimated time (Figure 3.11); an update is needed every few days.

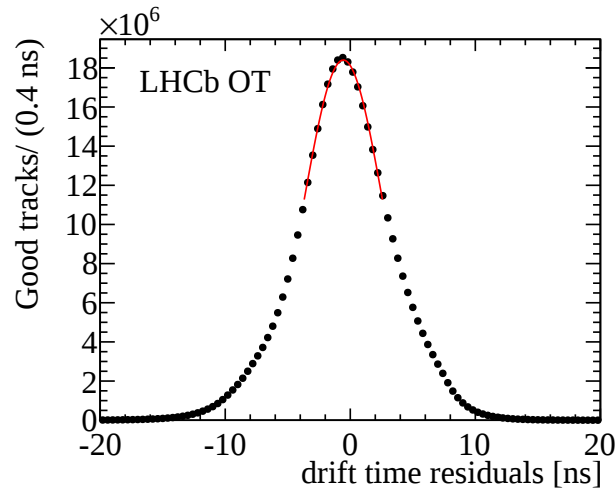


Figure 3.11: Fit to the distribution of the drift-time residuals of Outer Tracker hits used to estimate the global time shift between the collision time and the LHCb clock [79].

Calorimeter calibration

In order to have a constant L0 rate the calorimeter gain should not change with time. The variation of the gain can be estimated by the variation of the relative occupancy. The occupancy for a cell is defined as the fraction of events in which the ADC output is above a threshold. The ratio of the occupancies with respect to a reference sample is proportional to the changes in gain. A relative calibration is performed online on a single node for each fill using this occupancy method and when needed the high voltage is changed accordingly.

An absolute calibration can be obtained with the π^0 method: the reconstructed π^0 mass can be determined for each cell by fitting di-photon mass distributions where one of the photons has the cell as the seed. The calibration coefficients can be tuned to constrain the reconstructed π^0 mass to the nominal one. This fine calibration requires an iterative procedure and it is run on the HLT farm similarly to the tracking or RICH mirror alignment. This second method takes several hours to run and is performed a few times per year i.e. during technical stops.

3.3 REAL-TIME ALIGNMENT OF THE VELO

As seen in Section 2.2, the VELO is made of two halves that, during the data taking, have sensitive areas that starts approximately 8 mm away from the nominal beams position. For the safety of the detector, during the beam injection at the beginning of a fill, the halves are retracted by 29 mm; when stable beams are declared the VELO is closed around the beams. The VELO halves are moved using stepper motors and their position is read from resolvers mounted on the motor axes, with an accuracy of approximately 10 μm . An automated closure procedure positions the VELO halves around the beams using the response of the hardware and the positions of the beams measured using the VELO during the closing procedure. As the VELO is closed for each fill, its alignment may change with the same frequency. Figure 3.12 shows for a subsample of the Run I dataset the variation of the misalignment between the two halves in the horizontal direction perpendicular to the beam. This misalignment can be estimated by taking the difference of the positions of the primary vertices reconstructed separately with tracks in only one half of the VELO. In a perfectly aligned detector the mean of this difference should be zero.

The average variation in Run I is $\sim 4 \mu\text{m}$ while the maximum variation is $\mathcal{O}(10 \mu\text{m})$, which is more than the $\mathcal{O}(2 \mu\text{m})$ precision of the track based software alignment procedure [89].

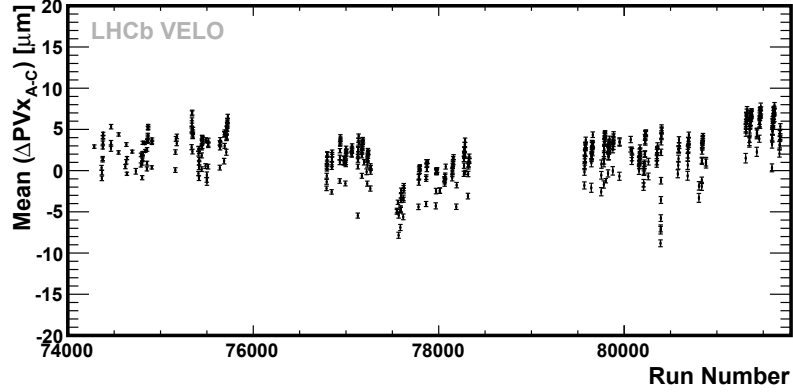


Figure 3.12: Misalignment between the two VELO halves along the main movement direction for the different runs, evaluated by fitting the primary vertices separately with tracks in each half of the VELO. The run numbers shown here span the period of the last four months of operations in 2010 [67].

During Run I for each year of data taking, a single VELO alignment was obtained using data collected during dedicated calibration runs. In Run II, instead, a new VELO alignment is obtained at the beginning of each fill and the alignment used is updated if it is significantly different from the new one. It is clear that in a real-time environment all the aspects of the alignment procedure must be decided before the data taking in contrast to when the procedure was offline and different strategies could be tested in parallel. In particular the following aspects have to be decided in advance:

1. which degrees of freedom to align for and which constraints to use;
2. which data sample to use as an input to the alignment procedure;
3. which criteria should be adopted to decide if two alignments are significantly different and hence an update is needed.

Moreover it is important to have in place an effective monitoring system to be able to timely detect any malfunctioning of the alignment procedure. An incorrect alignment would result in the need to reconstruct the data offline with a different alignment than the one used online, breaking the online-offline correspondence and hence adding

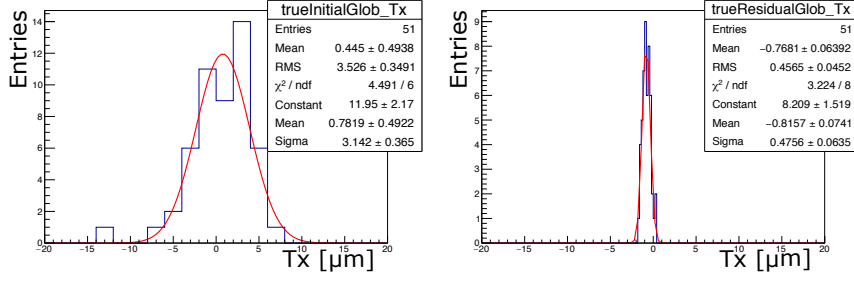


Figure 3.13: Comparison of the 2-halves T_x misalignment (*left*) before and (*right*) after the alignment procedure.

complexity to physics analyses. Furthermore analyses performed directly on the trigger output would not have the option to perform a new reconstruction of the data, potentially resulting in a loss of data.

Degrees of freedom and data sample determination

Pseudoexperiments are used to determine which degrees of freedom to align for. The general strategy is to produce 50 different datasets reconstructing the same MC data sample with random misalignments generated according to a Gaussian function, plus a dataset reconstructed without misalignment. These datasets are then aligned and the residual misalignments are fitted with Gaussian functions to extract the bias, the mean of the Gaussian function, and the precision, the width of the Gaussian function. An example of one of these distributions, before and after the alignment procedure, is shown in Figure 3.13. The datasets used in this study have approximately 15k events corresponding to approximately 6k vertices and 100k tracks. Two different kinds of events have been considered:

Collision events: collisions of the two proton beams in the interaction region; in this study a sample of $D^0 \rightarrow K^+\pi^-$ has been used to simulate collision events.

Beamgas events: collisions of a proton beam with the residual gas present in the VELO region; in this study a sample of collisions between one of the two beams with a mixture of Hydrogen and Oxygen in the same quantities has been used.

Figure 3.14 shows the different distributions of the polar angle of the tracks, the number of hits per track and the position of the primary vertices, for collision and beamgas events. Tracks from beamgas events are almost parallel to the beam and hence are less sensitive to

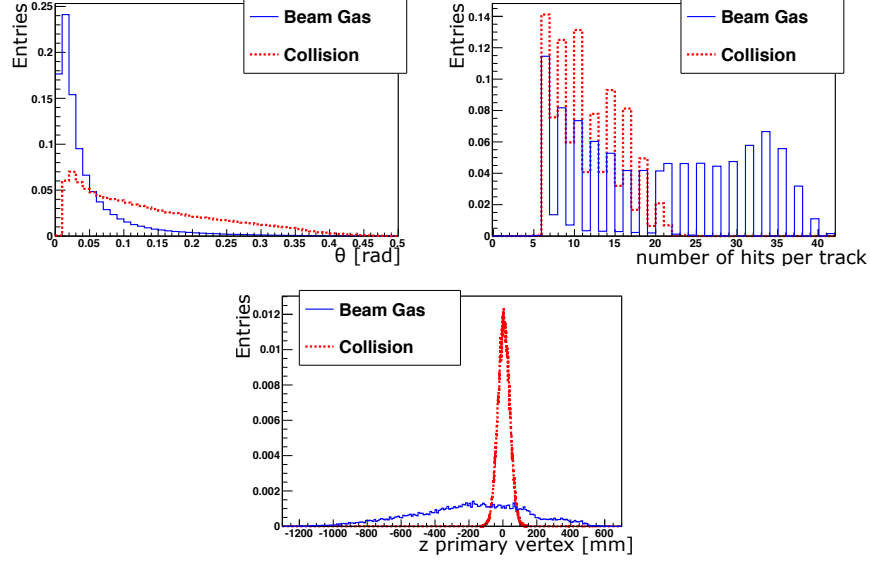


Figure 3.14: Comparison for collision and beamgas events of (top left) the track angular distribution, (top right) the number of hits per track and (bottom) the vertex position [90].

the misalignments Δ_z , Δ_α and Δ_β , as in Equation (3.1) the track gradients t_x and t_y are proportional to $\cos\vartheta$. On the other hand they traverse more modules and so they are more sensitive to the misalignment of a module inside its half.

*Two-halves'
alignment*

The first test performed, whose results are reported in Table 3.1, is to misalign only one half with respect to the other. As the VELO is the first detector element to be aligned, a global movement of both halves cannot be constrained so only the relative misalignment of one half with respect to the other is evaluated, fixing the average VELO position. The first line of the table reports the width of the Gaussian functions used to generate the random misalignments and both collision and beamgas events are considered. The precision achievable is $\mathcal{O}(1\,\mu\text{m})$ for T_x and T_y , a few microns for T_z and $\mathcal{O}(10\,\mu\text{rad})$ for the rotations. The beamgas events perform better for the rotations but worse for T_z . The first test probes an optimistic scenario as it is assumed that the modules are not misaligned. To study what effect the misalignment of the modules has on the alignment of the VELO halves, a second test has been performed where the modules are also misaligned according to a random Gaussian function. The results of this test are reported in Table 3.2 and the second line reports the width of the Gaussian functions used to generate the random misalignments of the modules; these values are compatible with the

residual module misalignment as evaluated in Ref. [90]. This level of misalignment of the modules leads to results compatible with the scenario without module misalignment. Aligning simultaneously for the two halves and the modules of the VELO is thus not needed if the modules have been already aligned and their misalignment is not expected to vary. Of course a big module misalignment would severely bias the alignment of the two halves; for example a random module misalignment of $\mathcal{O}(7\text{ }\mu\text{m})$ in T_x and T_y leads to a similar residual misalignment of the two halves after running the alignment procedure.

A similar study is performed, aligning also the modules. From Equation 3.1 it is clear that the alignment procedure is less sensitive to the misalignment of the modules in R_x , R_y and T_z so, as a baseline, only the main degrees of freedom T_x , T_y and R_z are aligned. As seen in Section 3.1 different kinds of constraints can be used to fight against weak modes. A common choice is to fix the position of a few modules. Figure 3.15 shows the result of this strategy on a randomly generated misalignment: the misalignment of the fixed modules, in this case the modules 10, 11, 32 and 33, can lead to a weak mode. A different choice is thus made: the average alignment constants of the modules inside the same half is zero in T_x , T_y and R_z and constrained to remain so, as a global misalignment would be corrected by the alignment of the two halves. Moreover there are Lagrangian constraints against the shearing in x and y by fixing to zero the sums of $\Delta_x \cdot z$ and $\Delta_y \cdot z$. Table 3.3 shows the results of this test: the residual misalignment of the two halves is compatible with the scenario in which the modules are left misaligned. The residual misalignment of the modules is $\mathcal{O}(1\text{ }\mu\text{m})$ for T_x and T_y and $\mathcal{O}(50\text{ }\mu\text{rad})$ for R_z .

Modules' alignment

A further test has been performed in which, starting from the result of the previous test, a new alignment is performed, this time aligning also for the subdominant degrees of freedom R_x , R_y and T_z and fixing the position of the modules 10, 11, 32 and 33. In this case it is necessary to fix the position of some modules as, for numerical reasons, the constraints applied before would not be applied correctly by the alignment procedure because strong correlations between different degrees of freedom are present. For example aligning for T_x and R_y , even if these are the only two degrees of freedom considered, does not allow to constrain the average T_x position of the modules inside one half. The results of this test are reported in Table 3.4 and it can be seen that the residual misalignments in R_x and R_y reduce to

Table 3.1: Residual two halves' misalignments from 50 samples obtained reconstructing the same simulated sample with random misalignments distributed according to a Gaussian function. The first line contains the width of Gaussian functions used to generate the random misalignments while the other two lines contain the residual misalignments for beamgas and collision events. The distribution of the residual misalignments is fitted with a Gaussian function and the table reports its mean $\pm$ its width. Refer to the text for more details.

	Tx [μm]	Ty [μm]	Tz [μm]	Rx [mrad]	Ry [mrad]	Rz [mrad]
σ misalignment	4	1	50	0.4	0.4	0.09
beamgas	-0.7 ± 0.2	0.40 ± 0.03	5.6 ± 0.9	0.0015 ± 0.0004	0.005 ± 0.002	-0.007 ± 0.005
$D^0 \rightarrow K^- \pi^+$	-1 ± 5	0.3 ± 0.6	1 ± 2	-0.05 ± 0.03	0.05 ± 0.10	0.2 ± 0.4

Table 3.2: Residual two halves' misalignments from 50 randomly misaligned samples as Table 3.1, the difference being that the modules are also randomly misaligned by Gaussian functions whose means are reported in the second line. Refer to the text for details.

Misalign 2-halves						
	Tx [μm]	Ty [μm]	Tz [μm]	Rx [mrad]	Ry [mrad]	Rz [mrad]
σ misalignment	4	2	50	0.6	0.5	0.07
σ misalignment modules	2	2	10	0.6	0.5	0.1
beamgas	-0.7 ± 0.6	0.4 ± 0.5	2 ± 3	0.001 ± 0.002	-0.000 ± 0.002	0.01 ± 0.03
$D^0 \rightarrow K^- \pi^+$	-1 ± 2	-0.0 ± 0.9	-2 ± 3	-0.002 ± 0.003	-0.07 ± 0.06	0.0 ± 0.2

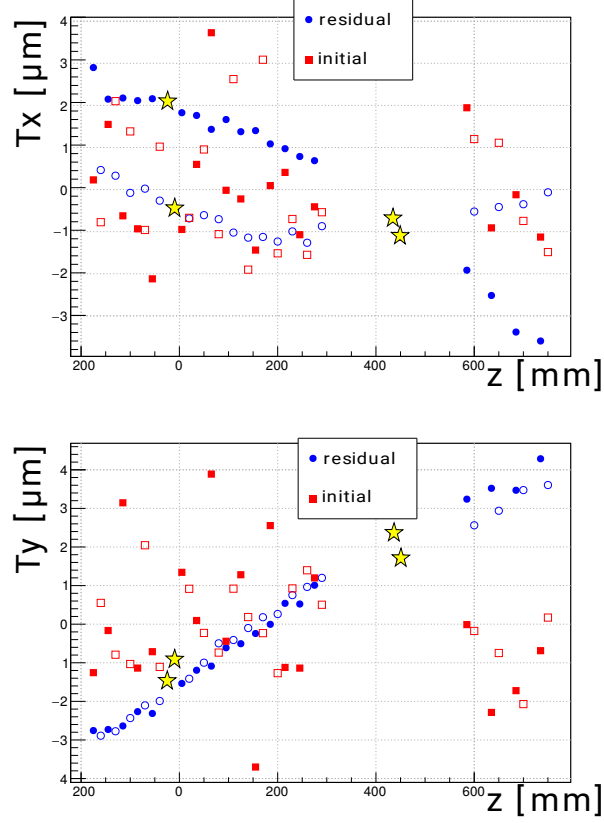


Figure 3.15: Residual modules' misalignment for a randomly misaligned sample fixing the position of the modules 10, 11, 32 and 33; (top) T_x and (bottom) T_y . Open (closed) markers represent modules in the left (right) half of the VELO and the stars represent the modules whose position is fixed. A shearing is particularly evident for T_y after the alignment as the modules whose position is kept fixed are significantly misaligned. Refer to the text for details.

$\mathcal{O}(200 \mu\text{rad})$. On the other hand the misalignment in T_z worsens. This is due to a weak mode of the alignment procedure as, for geometrical reasons, the distribution of residuals in the VELO is not sensitive to z -translation of the modules. This can be seen in Figure 3.16 where, on the left, the distribution of the residual misalignments of 50 different samples with random initial misalignments is shown. On the right instead different constraint strategies (applying the average constraint, fix the position of different modules, do not apply any constraint) are tried, without success. The conclusion is that an accurate survey is essential for the correct T_z alignment of the VELO modules, while the track-based alignment can be used for the main degrees of freedom T_x , T_y and R_z and, with an iterative approach, also for R_x and R_y .

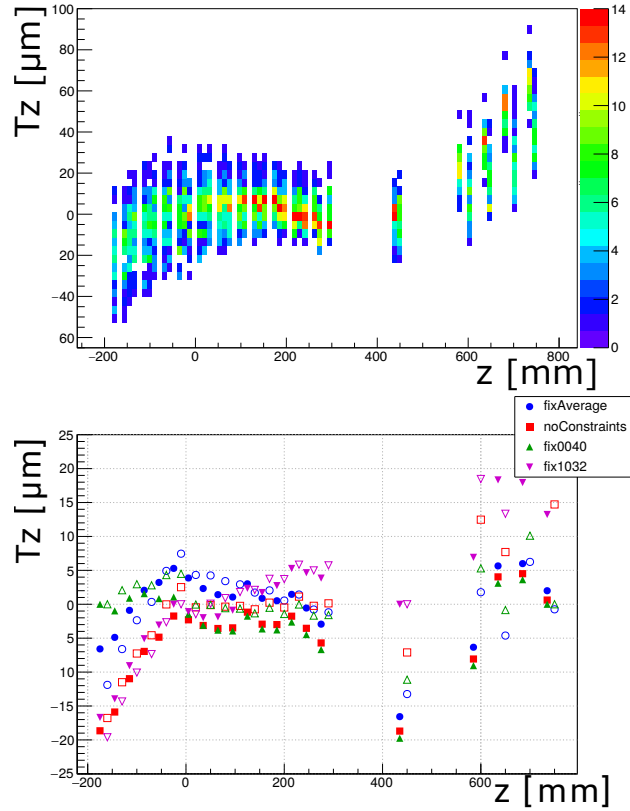


Figure 3.16: Residual T_z misalignment of (*top*) 50 different samples with a random initial misalignments and (*bottom*) the same sample using different constraints strategies. Refer to the text for details.

The conclusion of these studies is that, if the relative alignment of the two VELO halves can change from fill to fill while the module misalignment is more stable, the optimal strategy is to align in real

Table 3.3: Residual two halves' and modules' misalignments from 50 randomly misaligned samples as in Table 3.1. All six degrees of freedom are aligned for the two halves while only the main ones, T_x , T_y and R_z , are aligned for the modules. Refer to the text for details.

Misalign 2-halves						
	T_x [μm]	T_y [μm]	T_z [μm]	R_x [mrad]	R_y [mrad]	R_z [mrad]
σ misalignment	4	2	50	0.6	0.5	0.07
beamgas	-0.7 ± 0.6	0.4 ± 0.5	2 ± 2	0.002 ± 0.001	-0.001 ± 0.002	-0.004 ± 0.008
$D^0 \rightarrow K^- \pi^+$	1 ± 7	-1 ± 1	-1 ± 5	-0.001 ± 0.006	1.0 ± 0.2	0.3 ± 0.2
Misalign Modules						
	T_x [μm]	T_y [μm]	T_z [μm]	R_x [mrad]	R_y [mrad]	R_z [mrad]
σ misalignment	2	2	10	0.6	0.5	0.1
beamgas	0.0 ± 0.5	-0.0 ± 0.5	0 ± 10	-0.0 ± 0.6	-0.0 ± 0.5	-0.00 ± 0.03
$D^0 \rightarrow K^- \pi^+$	-1 ± 9	0 ± 2	0 ± 10	-0.0 ± 0.6	-0.0 ± 0.5	-0.01 ± 0.07

Table 3.4: Residual two halves' and modules' misalignments from starting from the aligned samples whose results are reported in Table 3.3. All six degrees of freedom are aligned for the two halves and the modules and four modules are fixed to their initial positions. Refer to the text for details.

Misalign 2-halves						
	Tx [μm]	Ty [μm]	Tz [μm]	Rx [mrad]	Ry [mrad]	Rz [mrad]
σ misalignment	4	2	50	0.6	0.5	0.07
beamgas	-0.8 ± 0.5	0.3 ± 0.3	0 ± 10	0.002 ± 0.001	-0.000 ± 0.002	-0.001 ± 0.007
$D^0 \rightarrow K^- \pi^+$	0 ± 4	-1.1 ± 0.9	-7 ± 9	-0.003 ± 0.003	-0.04 ± 0.01	0.1 ± 0.3
Misalign Modules						
	Tx [μm]	Ty [μm]	Tz [μm]	Rx [mrad]	Ry [mrad]	Rz [mrad]
σ misalignment	2	2	10	0.6	0.5	0.1
beamgas	-0.0 ± 0.5	0.0 ± 0.4	0 ± 20	-0.0 ± 0.2	-0.0 ± 0.2	0.00 ± 0.02
$D^0 \rightarrow K^- \pi^+$	-0 ± 4	-0.0 ± 1.0	0 ± 20	-0.0 ± 0.2	-0.0 ± 0.2	-0.01 ± 0.06

time only for the six degrees of freedom of the relative two-halves alignment and check, and eventually update, the alignment of the modules at the beginning of each data taking year and a few times per year. Moreover a combination of beamgas and collision events should be used as different degrees of freedom are better constrained by one or the other data kind of events.

Determination of the update criteria

As seen in Section 3.1, the alignment procedure is iterative and the convergence is reached when the χ^2 variation between two successive iterations falls under a threshold. In particular the condition adopted is:

$$\Delta\chi^2/\text{ndofs} < 4 \quad \text{or} \quad \max(\Delta\chi_{\text{mode}}^2) \lesssim 25$$

where ndofs is the number of degrees of freedom, χ_{mode}^2 is the contribution to the χ^2 for a mode³ and the $\lesssim$ is present because, depending on the number of modes, this value changes to take into account the look-elsewhere effect. These thresholds on the χ^2 variations are common to all the tracking elements; however it is important to define also thresholds on the variation of the individual degrees of freedom of the alignables: a minimum variation to decide if an alignment is significantly different from the previous one, and a maximum expected variation to avoid applying unphysically large corrections that may arise due to mistakes or an incorrect convergence in the alignment procedure.

To decide the minimum thresholds, the accuracy and the precision of the alignment procedure have to be estimated. Three different input are considered:

accuracy from MC evaluated with 50 different misalignments over the same data sample (this is the test with pseudoexperiments discussed above),

precision from MC evaluated with 8 different samples (starting from a misaligned or non misaligned scenario gives compatible results),

precision from data evaluated from 8 different samples taken from the same fill (it is assumed that the misalignment does not change inside the same fill).

³ A linear combination of the alignment parameters α and an eigenvector of the matrix $\left(\frac{d^2\chi^2}{d\alpha^2}\right)\bigg|_{\alpha_0}$.

These inputs are compared with the variation of the alignment constants between three different fills of 2012. The results of this comparison are reported in Table 3.5. From these the minimum and

Table 3.5: Precision and accuracy of the alignment procedure evaluated from MC and data and comparison with variation of the alignment between three different fills of 2012. Refer to the text for details.

2 halves				
	Accuracy from MC	Precision from MC	Precision from data	Variation 3 fills 2012
$T_x T_y$ [μm]	< 1	< 1	2	up to 6
T_z [μm]	3	4	1	0.5
R_x, R_y [μrad]	3	2	3	4
R_z [μrad]	10	15	40	up to 25
Modules				
$T_x T_y$ [μm]	< 1	< 1	2	2
R_z [μrad]	30	50	100	60

maximum threshold for the 2015 data taking were decided and are reported in Table 3.6. Moreover as the variations of the alignment constants of the modules in different fills of the year are comparable with the precision and accuracy of the alignment it was decided to align only the two VELO halves in 2015.

Table 3.6: Thresholds on the variation alignment constant used in 2015. Refer to the text for details.

2 halves		
	Min variation	Max variation fills 2012
$T_x T_y$ [μm]	2	15
T_z [μm]	4	15
R_x, R_y [μrad]	3.5	25
R_z [μrad]	15	100
Modules		
$T_x T_y$ [μm]	2	15
R_z [μrad]	100	300

As already seen in Section 3.2, the real-time alignment of the VELO works as follows. At the beginning of each fill a data sample of approximately 50k beamgas and collision events is selected by dedicated HLT1 trigger lines. The alignment procedure runs on the HLT trigger farm to provide a new alignment for the new fill. If only one iteration is needed to converge or if the variations of all the alignment constants are below the minimum thresholds, the new alignment is equivalent to the old one and an update is not needed. Otherwise, if the new VELO alignment is significantly different from the old one, a run change is triggered and the new alignment is used for the rest of the fill. For consistency with HLT1, the old alignment is used to reconstruct the previous runs of the fill also in HLT2 and offline. If the alignment procedure fails or the variation of any alignment constant is larger than its maximum threshold no update is triggered and experts are called to investigate. The full real-time alignment procedure, including the collection of the data sample, takes a few minutes.

This procedure has been commissioned and used at the start of Run II, in 2015; at first running the alignment and updating the constants was done manually and latterly in a fully automatic way. Figure 3.17 shows the convergence of the alignment of the first fill of Run II while Figure 3.18 shows the stability of the two halves T_x and T_y for the full 2015.

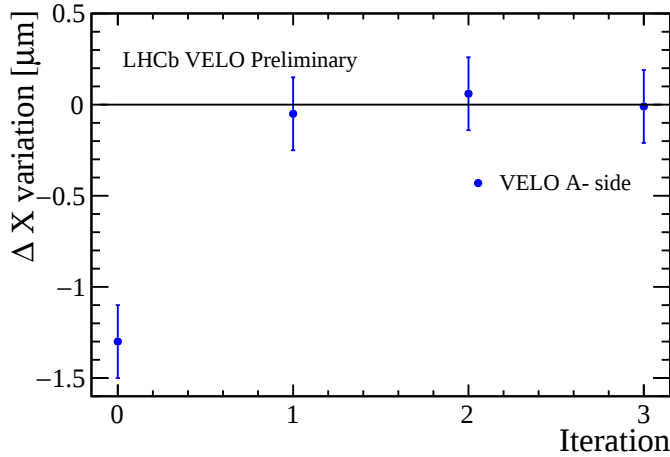


Figure 3.17: Convergence of the alignment of the two halves of the VELO obtained on the first fill of Run II starting from the 2012 VELO alignment. Each point shows the variation of the VELO left side x-translation with respect to the previous iteration [91].

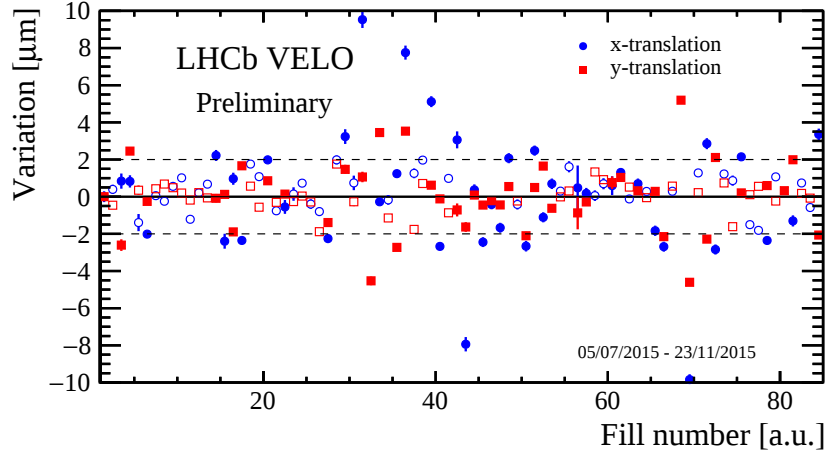


Figure 3.18: Stability of the alignment of the VELO halves during all fills of 2015. Each point is obtained running the online alignment procedure and shows the difference between the initial alignment constants (the ones used in the previous fill) and the new ones computed by the alignment. Closed (open) markers represent alignments that triggered (did not trigger) an update. The horizontal lines represent the thresholds to trigger an alignment update [91].

Lessons from 2015

Using the data collected for the alignment procedure in 2015 it is possible to check again the precision of the alignment on the actual data samples used and update the minimum thresholds. Five fills are considered and are divided into independent data samples of approximately 8k events each. An alignment is computed for each, starting from the optimal alignment obtained from the full data sample for that fill. The difference with respect to the initial alignment constants depends on how sensitive the alignment procedure is to the data sample used and hence is a measurement of the precision. The precision obtained is

$$\begin{array}{ll} T_x, T_y \sim 1 \mu\text{m} & T_z \sim 4.5 \mu\text{m} \\ R_x, R_y \sim 3.5 \mu\text{rad} & R_z \sim 25 \mu\text{rad} \end{array}$$

and the thresholds used in 2016 have been updated to the ones reported in Table 3.7.

It is possible also to check how stable the module alignment is. At the beginning of the 2015 data taking period the module alignment has been run in monitoring mode and it has been found that its varia-

Table 3.7: Thresholds on the variation alignment constants used in 2016. Refer to the text for details.

dof	Min variation	Max variation
T_x T_y [μm]	1.5	15
T_z [μm]	5	15
R_x, R_y [μrad]	4	25
R_z [μrad]	30	100

tion is compatible with the precision of the alignment procedure. As an example Figure 3.19 shows the variation with respect to the preceding alignment of T_x for four different modules over a period of 3 months.

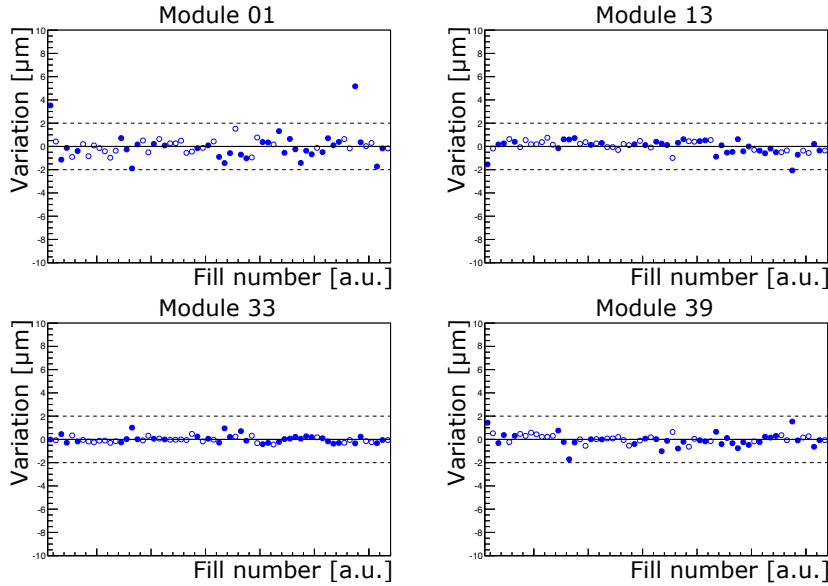


Figure 3.19: Stability of module T_x alignment during the beginning of 2015 for four different VELO modules as a function of increasing fill number. Closed (open) markers represent alignments that triggered (did not trigger) an update. See caption of Figure 3.18 for details.

Real-time alignment of the VELO in 2016

The real-time alignment of the VELO ran successfully also in 2016. Figure 3.20 shows the stability of T_x and T_y of the two halves for the full 2016. On a total of 198 fills there have been 68 updates. It is interesting also to see which degrees of freedom triggered the majority of the updates. Figure 3.21 shows the fraction of updates triggered by

each combination of degrees of freedom while Figure 3.22 shows the fraction for each degree of freedom, *i.e.* in which fraction of updates that degree of freedom was above threshold. As expected T_x is the degree of freedom that triggers the majority of the updates as x is the main direction along which the two halves move.

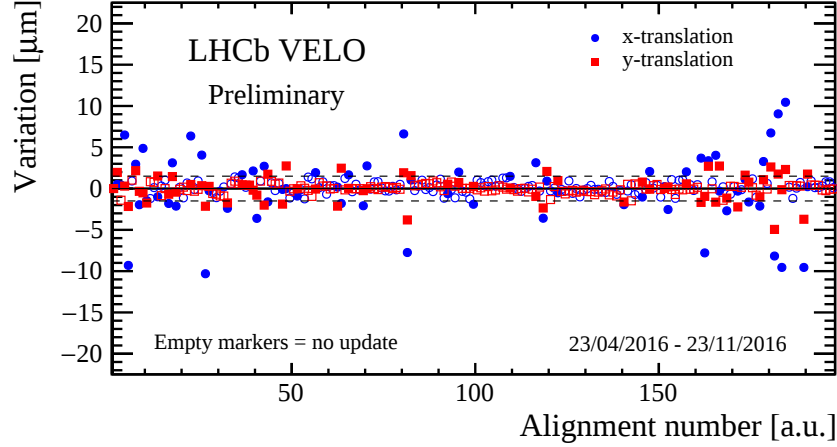


Figure 3.20: Stability of the alignment of the VELO halves during all fills of 2016. Closed (open) markers represent alignments that triggered (did not trigger) an update. See caption of Figure 3.18 for details [91].

3.4 CONCLUSION

In Run II the new scheme for the software trigger at LHCb allows the alignment and calibration to be performed in real time. A dedicated framework has been put in place to parallelise the alignment and calibration tasks on the multi-core farm infrastructure used for the trigger in order to meet the computing time constraints. This procedure allows a more stable alignment quality hence better performance, online-offline consistency, thanks also to the same online-offline reconstruction, and, consequently, more effective trigger selections. This represents a change of paradigm with physics analyses performed directly on the trigger output with the same online-offline performance. This allows to use more effectively the computing resources, expanding the physics reach of the experiment and allowing to perform analyses that require extremely large datasets. The full procedure has been successfully commissioned and tested in 2015 leading to the publication of the first analyses performed directly on

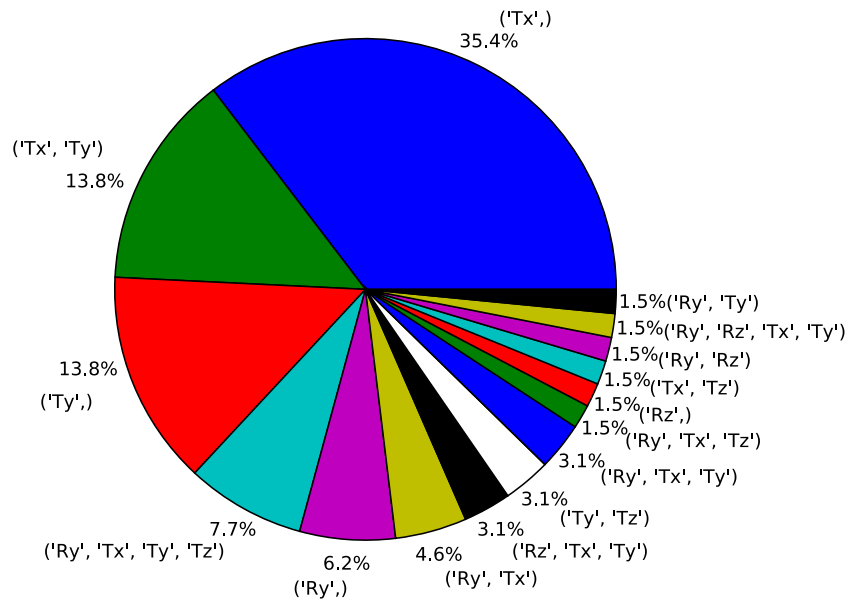


Figure 3.21: Fraction of updates triggered by each combination of degrees of freedom.

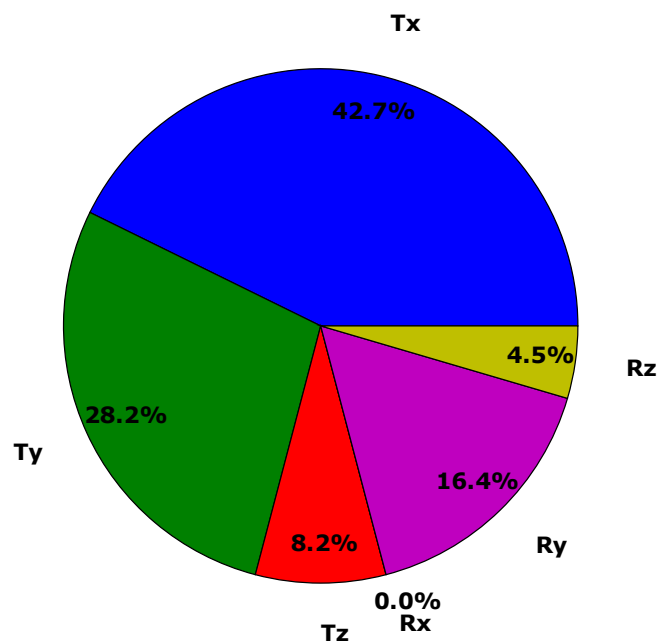


Figure 3.22: Fraction of updates in which each degree of freedom is above threshold.

the trigger output [86, 87] and ran smoothly also in 2016. More and more analyses are now adopting this novel paradigm of performing analyses directly on the trigger output that will become the default in the next run of data taking starting in 2021.

4

SEARCH FOR THE LEPTON FLAVOUR VIOLATING DECAY

$$\tau \rightarrow \phi \mu$$

The present chapter describes a feasibility study for a search for the lepton flavour violating decay $\tau \rightarrow \phi \mu$ using the full Run I LHCb dataset, 3.0 fb^{-1} collected at 7 and 8 TeV. An expected limit of $\mathcal{B}(\tau \rightarrow \phi \mu) \leq 3.8 \cdot 10^{-6}$ is obtained at 90% CL, which is approximately 50 times larger than the current limit set by Belle [92]. Section 4.1 gives the motivations for this search, the details of which are given in the subsequent sections: the selection in Sections 4.3 and 4.5, the background characterisation in Section 4.4 and the determination of the expected upper limit using the full Run I dataset in Section 4.6. Finally Section 4.7 describes the limit set using the 2012 dataset.

4.1 MOTIVATION

Section 1.2 contains a theoretical overview on LFV and on how tau LFV decays were searched for at e^+e^- storage rings. At the LHC the tau production cross-section is almost five orders of magnitude larger than in the e^+e^- storage rings ($\sim 85 \mu\text{b}$) so LHC can be considered a tau-factory. The taus are not produced in pairs, but come mainly from the leptonic decays of charmed mesons, in particular from the D_s^+ meson¹. The specific features of the LHCb detector needed for the study of rare b and c decays make this experiment suitable also for the study of LFV tau decays as proven by a competitive limit on the branching fraction for the $\tau \rightarrow 3\mu$ decay and the world's first limit on the $\tau \rightarrow \rho\mu\mu$ and $\tau \rightarrow \bar{\rho}\mu\mu$ branching fractions [93, 94]. It is thus natural to consider other LFV tau decays, for example those where the tau decays into a muon and a neutral meson. In Table 4.1 the upper limits on the branching fractions set by the other experiments for the $\tau \rightarrow \phi\mu$ channel are listed.

¹ The tag and probe method used in e^+e^- experiments and explained in Section 1.2 cannot be used in the selection of taus at LHCb.

Table 4.1: Upper limits on $\mathcal{B}(\tau \rightarrow \phi\mu)$. All experiments used $e^+ e^-$ collisions with centre-of-mass energy of $\sqrt{s} = 10.6 \text{ GeV}$.

Experiment	Upper limit 90% C.L.	Year	Integrated luminosity
CLEO [95]	$7.0 \cdot 10^{-6}$	1998	4.79 fb^{-1}
BABAR [96]	$1.9 \cdot 10^{-7}$	2009	451 fb^{-1}
Belle [92]	$8.4 \cdot 10^{-8}$	2011	854 fb^{-1}

4.2 DATASET AND SIMULATED SAMPLES

Dataset description

To estimate the expected upper limit on $\tau \rightarrow \phi\mu$, 10% of the 2012 LHCb dataset has been used, corresponding to an integrated luminosity of $\sim 0.2 \text{ fb}^{-1}$. According to the current limit, set by Belle (Table 4.1), no signal events are expected in this dataset. The full 2012 LHCb dataset, corresponding to an integrated luminosity of 2.0 fb^{-1} is used to set the observed limit, as described in section 4.7.

MC samples

For this analysis two MC samples have been produced: the signal sample $\tau \rightarrow \phi\mu$ and the normalisation sample $D \rightarrow \phi\pi$. In Section 2.3.1 details of the production of simulation samples within the LHCb collaboration are given. Also for these samples a selection on the generated particles has been applied before simulating the interaction with the detector: the two kaons and the muon are required to be inside the detector acceptance and to have $p_T > 250 \text{ MeV}$ and $p > 2.5 \text{ GeV}$; this selection has an efficiency (ε_{GEN}) around 16%. As the τ production rates in the MC samples do not match the ones measured by the LHCb experiment a correction is put in place as described in Appendix A.

4.3 TRIGGER AND STRIPPING SELECTION

Trigger selection

The candidate events are required to have been selected by specific trigger lines. To pass the L0 level trigger the candidate must have a muon with $p_T > 1.48 \text{ GeV}$; alternatively part of the event, independent from the signal, must have passed the requirements of any L0 line. The candidates are then required to contain two tracks coming from the same vertex, whose combined mass, under the kaon mass hypothesis, is inside a 20 MeV window from the ϕ mass or to have

Table 4.2: Details of the stripping selection for $\tau \rightarrow \phi\mu$.

KAON SELECTION	
track type	long
track fit χ^2/ndof	< 3
track probability to be a ghost	< 0.3
impact parameter divided by its error of the track	> 9
p_T	$> 300 \text{ MeV}$
PIDK	> 5
ϕ SELECTION	
$ m_{KK} - m_\phi $	$< 10 \text{ MeV}$
vertex fit χ^2	< 10
Particle trajectory minimum distance from a primary vertex	$> 4\sigma$
MUON SELECTION	
track fit χ^2/ndof	< 3
track probability to be a ghost	< 0.3
impact parameter divided by its error of the track	> 9
p_T	$> 300 \text{ MeV}$
τ SELECTION	
$ m_{\phi\mu} - m_\tau $	$< 150 \text{ MeV}$
$c\tau$	$> 200 \mu\text{m}$
vertex fit χ^2	< 10
χ^2 of the fit of the closest primary vertex adding this track	< 100

three tracks coming from a same vertex and whose combined mass is, under the kaon or pion hypothesis, between 1.8 and 2.04 GeV².

A dedicated stripping line³ selects $\tau \rightarrow \phi\mu$ events. Each selected event is required to have

Stripping selection

- at least two well reconstructed tracks, detached from any primary vertex, with $p_T > 300 \text{ MeV}$ and likely to be kaons,
- a well-reconstructed muon, detached from any primary vertex, and with $p_T > 300 \text{ MeV}$.

The tracks of each pair of kaons whose combined mass is compatible with the ϕ mass are re-fitted to a common decay vertex to form a ϕ -

² This selects D mesons decaying to three hadrons but, changing the mass hypothesis of a kaon into a muon, the selection for a D meson decaying to three kaons selects also $\tau \rightarrow \phi\mu$.

³ For reference the name of this stripping line is LVLinesTau2PhiMuLine.

candidate. The fit χ^2 is required to be less than 10 and the ϕ candidate must not come from a primary vertex. Finally a tau candidate is built from the muon and the ϕ if their combined mass is within a 150 MeV window around the tau mass. The tau candidate is selected if it comes from a primary vertex, its flight distance is greater than 200 μm and the fit χ^2 is less than 10. Further details of the stripping selection are given in Table 4.2.

Efficiencies

The efficiency corresponding to this stripping line has been evaluated using the simulated events described in section 4.2 and found to be

$$\varepsilon_{\text{STRIP}} = (4.85 \pm 0.02)\%.$$

The trigger efficiency is evaluated on the simulated events that passed the stripping selection and is found to be

$$\varepsilon_{\text{TRIG}} = (20.80 \pm 0.17)\%.$$

4.4 BACKGROUND CHARACTERISATION

Purity of the ϕ sample

To evaluate what fraction of background events contains a genuine ϕ , and not a random combinations of kaons from the same event, a fit to the invariant mass of the kaon pair is performed (Figure 4.1). The fit model consists of a Voigtian distribution⁴ for the ϕ peak and a linear function for the background. It can be seen that the ϕ candidates selected in the stripping have a low background contamination. Therefore the main sources of background in the $\phi\mu$ mass spectrum come from events which contain a genuine ϕ .

Peaking backgrounds

The mass distribution of the tau candidates in the 10% of the data, in which only background would be expected, shows a peak at around 1860 MeV, which is due to the decay $D^\pm \rightarrow \phi\pi^\pm$ where the pion is misidentified as a muon. In Figure 4.2 the tau candidate mass spectrum is shown assuming the muon to be a pion. The D-peak is fitted with a Crystal Ball function [97] and the continuum background with a Chebyshev polynomial of third order.

Main sources of background

As seen above, the main expected sources of background come from decay channels with a genuine ϕ . In Table 4.3 these channels are listed with their expected yields before the selection. To evaluate their relative contributions it is assumed that the probability to misidentify a pion for a muon is 2% [67]. A simple phase-space generation has

⁴ The convolution of a Breit-Wigner with a Gaussian.

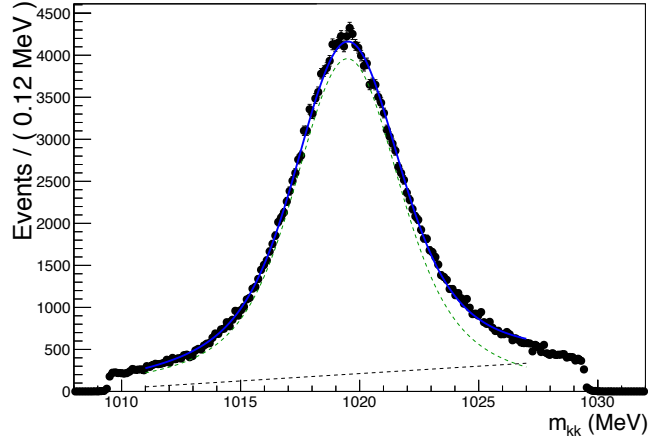


Figure 4.1: Invariant mass of the kaon pairs. The ϕ peak, in dashed green, is fitted with a Voigtian distribution and the background, in dashed black, with a linear function.

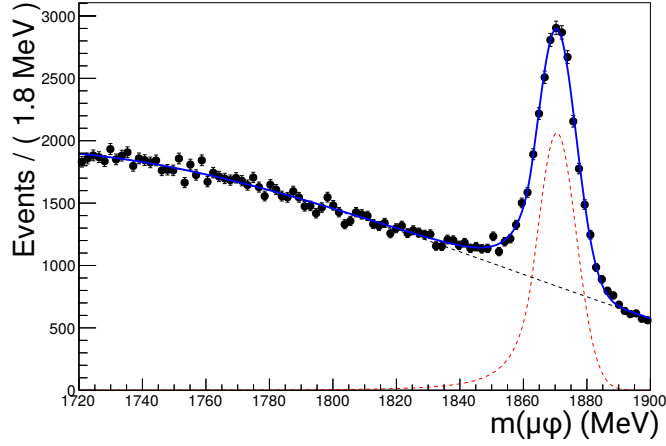


Figure 4.2: Mass spectrum of tau candidates with μ mass changed to π mass. The D peak, in dashed red, is fitted with a Crystal Ball distribution and the background, in dashed black, with a Chebyshev polynomial of third order.

Table 4.3: Main sources of expected background, with rejection rates of the selections not taken into account.

Decay channel	σ	Branching fraction	Prob misID	expected yield in 1 fb^{-1}
$D \rightarrow \phi \tau \tau^0$	$721 \text{ } \mu\text{b}$	$1.12 \cdot 10^{-2}$	0.02	$1.62 \cdot 10^8$
$D \rightarrow \phi \pi$	$721 \text{ } \mu\text{b}$	$2.65 \cdot 10^{-3}$	0.02	$3.82 \cdot 10^7$
$D_s \rightarrow \phi \mu \nu$	$222 \text{ } \mu\text{b}$	$1.22 \cdot 10^{-2}$	1	$2.70 \cdot 10^9$
$D_s \rightarrow \phi \rho (\rightarrow \tau \tau^0)$	$222 \text{ } \mu\text{b}$	$4.11 \cdot 10^{-2}$	0.02	$1.82 \cdot 10^8$
$D_s \rightarrow \phi \pi$	$222 \text{ } \mu\text{b}$	$2.28 \cdot 10^{-2}$	0.02	$1.01 \cdot 10^8$
$D_s \rightarrow \phi \pi \rho^0$	$222 \text{ } \mu\text{b}$	$6.6 \cdot 10^{-3}$	0.02	$2.93 \cdot 10^7$
$D_s \rightarrow \phi a 1 (\rightarrow \pi \rho^0)$	$222 \text{ } \mu\text{b}$	$7.5 \cdot 10^{-3}$	0.02	$3.33 \cdot 10^7$
$D_s \rightarrow \phi 3 \pi$	$222 \text{ } \mu\text{b}$	$5.92 \cdot 10^{-3}$	0.02	$2.63 \cdot 10^7$

been performed to give an estimate of the shape of the expected background. These distributions are shown in Figure 4.3; they have been normalised with the yields of Table 4.3 and for the signal an unrealistic branching fraction of 10^{-4} has been assumed for the sake of visualisation. It can be seen that the dominant expected source of continuum background is from $D_s \rightarrow \phi\mu\nu$, with small contributions from $D_s \rightarrow \phi\pi\pi^0$ and $D \rightarrow \phi\pi\pi^0$.

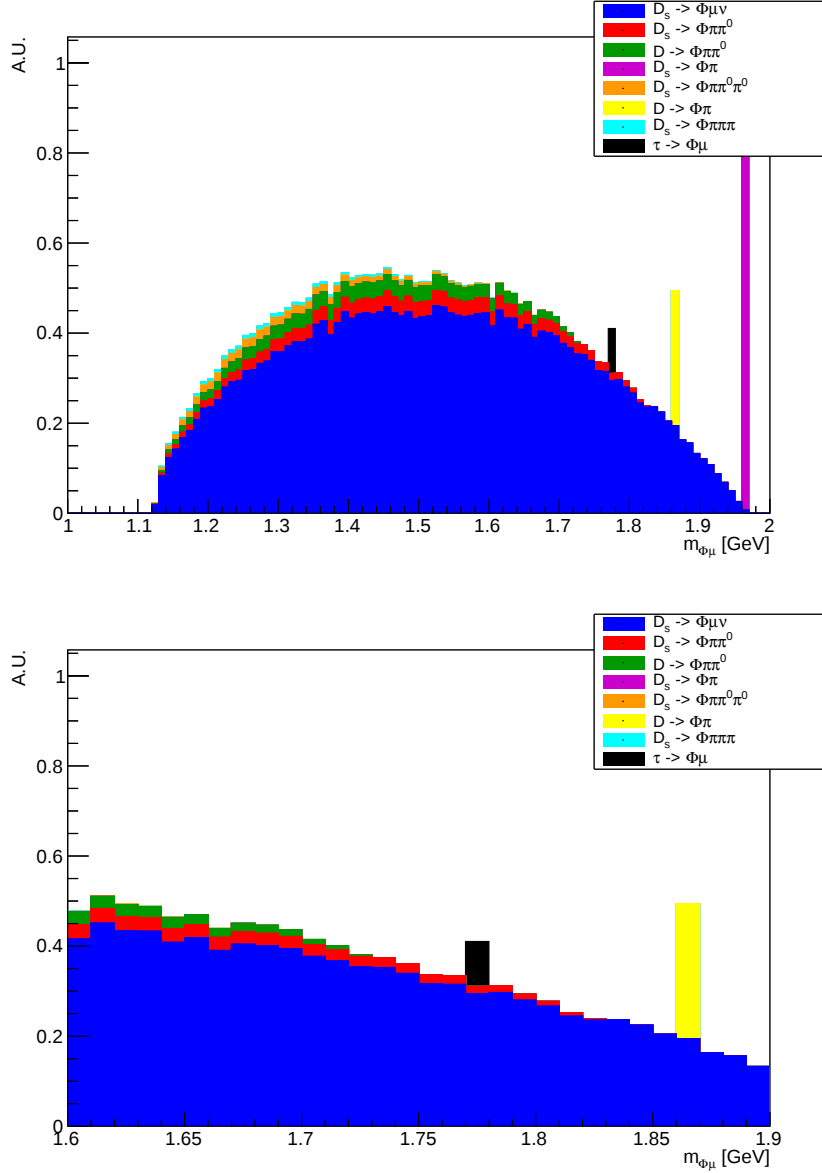


Figure 4.3: Distributions of the reconstructed tau candidate mass obtained with a phase-space generation for the main sources of expected background containing a ϕ . (*top*) In a wide mass range, (*bottom*) in the region of interest.

4.5 MVA AND PID SELECTION

MVA selection

To improve discrimination between signal and background a multi variate analysis (MVA) selection has been performed. The main differences expected between $\phi\mu$ candidates from a τ and from a D_s are a larger lifetime and a different distribution in the IP, as the majority of the taus originate in a D_s decay. A boosted decision tree (BDT) classifier [98] has been trained using as signal the $\tau \rightarrow \phi\mu$ MC sample and as background the lower mass sideband ($m_{\phi\mu} < 1750$ MeV). The variables used as input to the BDT are listed in Table 4.4 and their distributions in the signal and background samples are compared in Figure 4.4. A tuning of the hyper-parameters of the BDT has been

Table 4.4: Variables used as an input of the BDT classifier.

VARIABLE	DEFINITION
Tau_IP_OWNPV	Impact parameter of the τ candidate
Tau_IPCHI2_OWNPV	IP- χ^2 of the τ candidate
Phi_PT	p_T of the ϕ meson
Mu_PT	p_T of the muon
Mu_IPCHI2_OWNPV	IP- χ^2 of the muon
Tau_sinDira	Sine of the angle between the direction of flight and the momentum of the τ candidate
K_minPT	Smaller p_T of the two kaons
K_minIPCHI2	Smaller IP- χ^2 of the two kaons
Pointing	$\frac{P \sin \vartheta}{P \sin \vartheta + \sum p_T}$ where P is the momentum of the τ candidate, ϑ the angle between its direction of flight and its reconstructed momentum and the sum is over the p_T of the muon and the ϕ meson

performed, and the ones chosen are listed in Table 4.5. A multi-layer perceptron [99] has also been considered as an alternative MVA classifier but did not provide a better performance. The receiver operating characteristic (ROC) curve and the output of the BDT classifier for the signal and background samples are shown in Figure 4.5.

Table 4.5: Hyper-parameters of the BDT classifier used for the training.

VARIABLE	VALUE
Boost Type	Adaptive
AdaBoosBeta	0.5
N _{trees}	850
Separation Type	Gini Index
Min number of events required in a leaf node	50
Max depth	3

In the stripping line no explicit selection is performed on the PID of the muon to discriminate the signal from backgrounds such as $D \rightarrow \phi\pi$ or $D_{(s)} \rightarrow \phi\pi\pi^0$. Various alternatives were considered:

PID selection

- 1-ProbNNpi
- weighted sum: $7/10 \cdot \text{ProbNNmu} + 3/10 \cdot (1 - \text{ProbNNpi})$
- PIDmu
- simple sum: $1/2 \cdot \text{ProbNNmu} + 1/2 \cdot (1 - \text{ProbNNpi})$
- ProbNNmu
- prod: $\text{ProbNNmu} \cdot (1 - \text{ProbNNpi})$

where PIDmu and ProbNNmu are defined in Section 2.3.3. The weighted sum was inspired by Figure 4.6 as a way to better separate real and fake muons. These alternatives were tested on MC taking $\tau \rightarrow \phi\mu$ as the signal and $D \rightarrow \phi\pi$ as the background. The ROC curves are shown in Figure 4.7 from which it is clear that ProbNNmu is the most suitable variable.

To choose the value of the cuts to apply to the output of the BDT classifier and to ProbNNmu a combined optimisation has been performed using as figure of merit the product of signal efficiency and background rejection $\varepsilon \cdot (1 - \varepsilon_B)$. Figure 4.8 shows how this value change as functions of the selection in BDT output and ProbNNmu. The values which optimise this figure of merit are: $\text{BDT} > -0.02$, $\text{ProbNNmu} > 0.59$. The signal efficiency is $\varepsilon_{\text{CUTS}} = 71.4 \pm 0.4\%$ and the background rejection rate is $\sim 65\%$.

Working point

4.6 FIT AND LIMIT SETTING

The expected mass shape for the signal is obtained using the simulated sample described in section 4.2. Since the tau has almost zero decay width, the core of the distribution is expected to be Gaussian; however the tails of the distribution could go to zero more slowly than in a Gaussian, for example following a power law distribution. As the shape is not expected to be asymmetric, the function chosen to describe the signal distribution is a *double-sided symmetric Crystal Ball* function: the central part is Gaussian (with mean μ and standard deviation σ) while the tails are described by a power law with exponent ν ; the parameter α is the distance, in numbers of standard deviations

Signal model

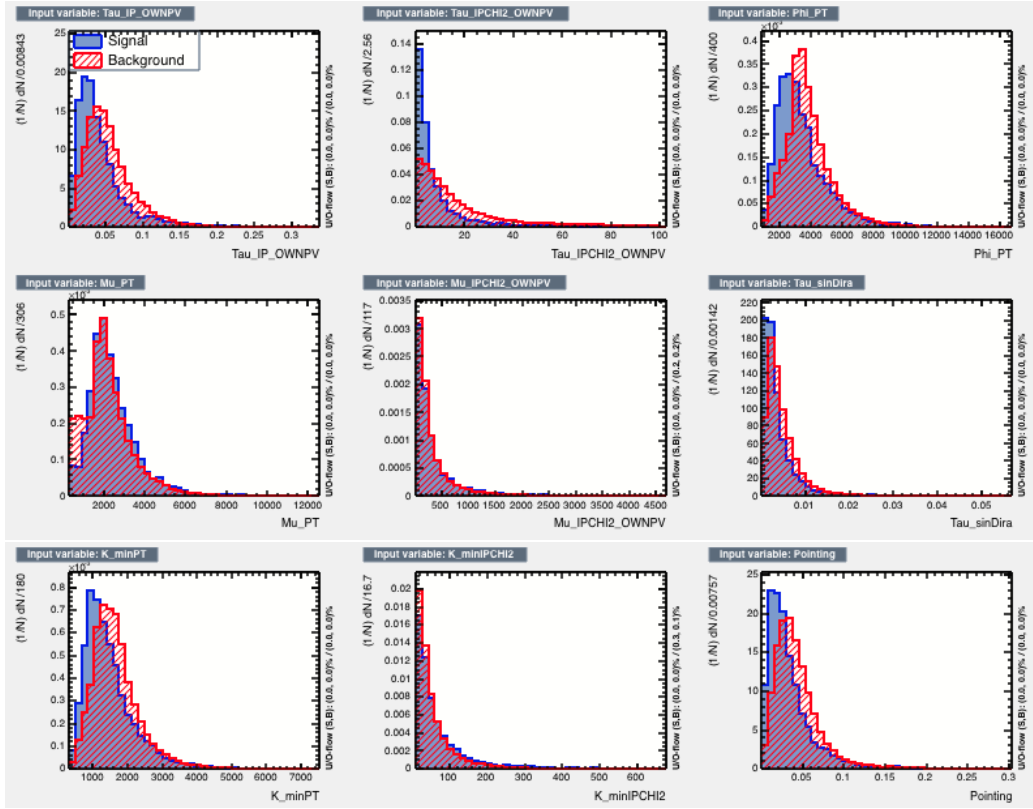


Figure 4.4: Variables used as input to the BDT classifier.

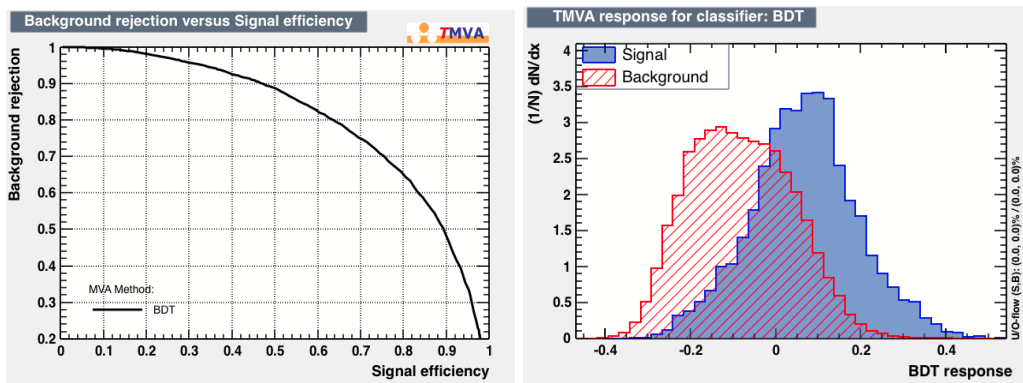


Figure 4.5: (left) ROC curve and (right) output of the BDT classifier for the signal and background samples.

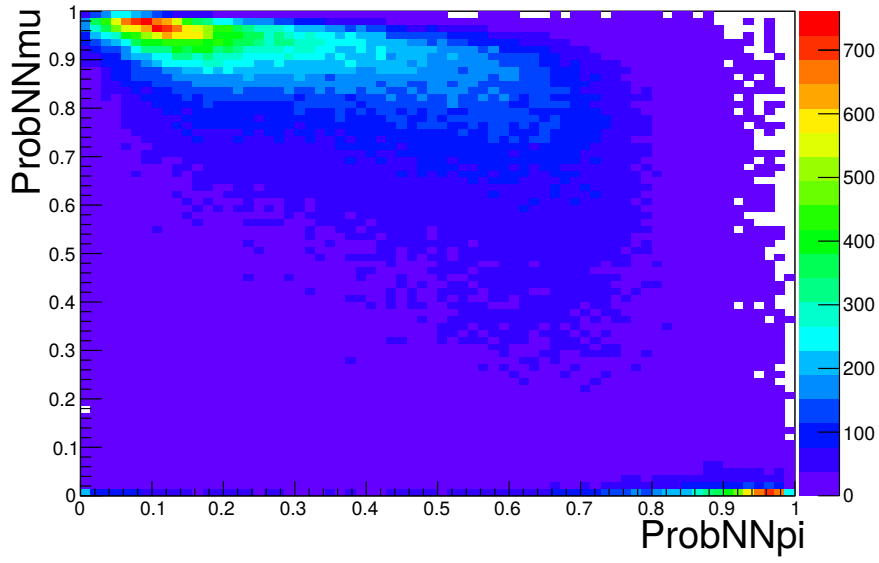


Figure 4.6: Distribution of the data in the ProbNNmu-ProbNNpi plane; the two peaks corresponding to real and fake muons in the upper-left and lower-right corners are clearly visible.

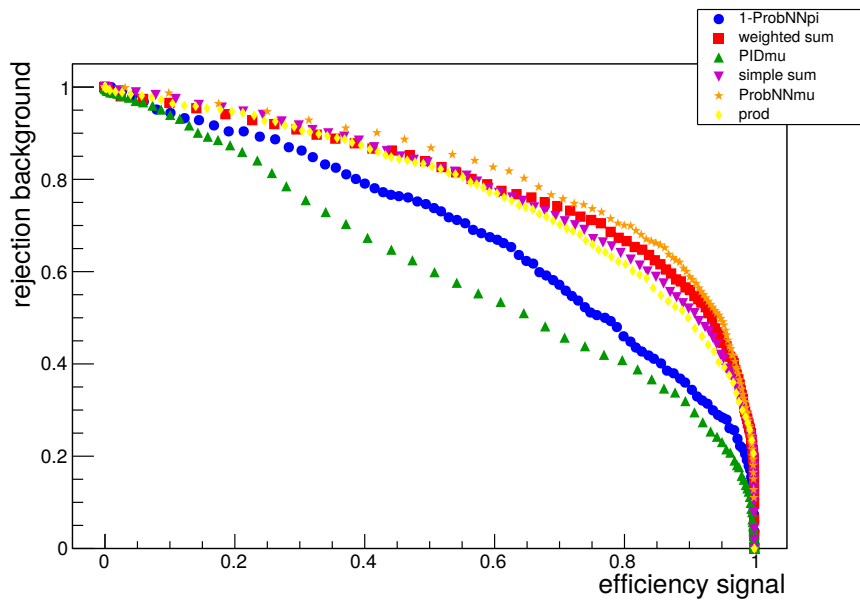


Figure 4.7: ROC curve for various PID variables. Refer to the text for details.

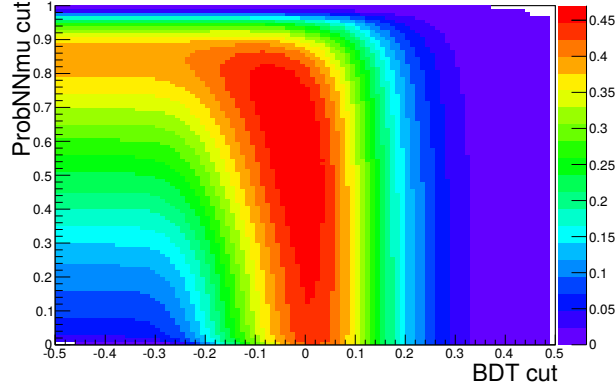


Figure 4.8: Value of the figure of merit for the different values of the cut in BDT output and ProbNNmu.

from the mean, where the tails start. The two tails of a double-sided Crystal Ball function can in general have different values for the parameters ν and α , but a symmetric function is used in this fit. The fit to the simulated sample is shown in Figure 4.9 and its parameters are reported in Table 4.6; In this figure and in others in this thesis the pulls⁵ are also shown to prove the goodness of the fit. As the parameters ν and α are highly correlated, the value of α is fixed to 1.2.

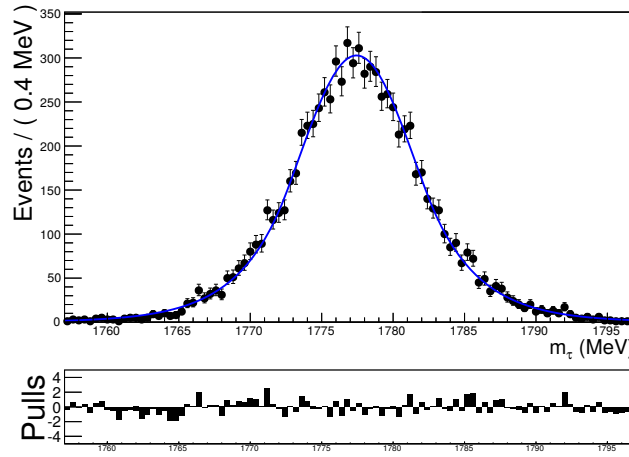


Figure 4.9: Mass spectrum of tau candidates from the simulated sample fitted with a double-sided symmetric Crystal Ball function. Refer to the text for details.

Data fit

To obtain an estimate of the order of magnitude of the limit on the

⁵ A pull is the difference between the y-position of the data point and that of the fitting function at the same x-position, divided by the error on the data point

Table 4.6: Parameters of the fit to the MC signal shape.

μ [MeV]	1777.438 ± 0.056
σ [MeV]	4.083 ± 0.014
ν	145.78 ± 0.16

branching fraction of $\tau \rightarrow \phi\mu$, an unbinned maximum likelihood fit has been performed to the 10% of the 2012 dataset:

- the signal shape is taken from the fit to the simulated data described above, with the normalisation as the only parameter left free in the fit,
- the D peak is modelled with a Crystal Ball function with the parameter ν fixed to 3,
- the continuum background is modelled with a third-order Chebyshev polynomial.

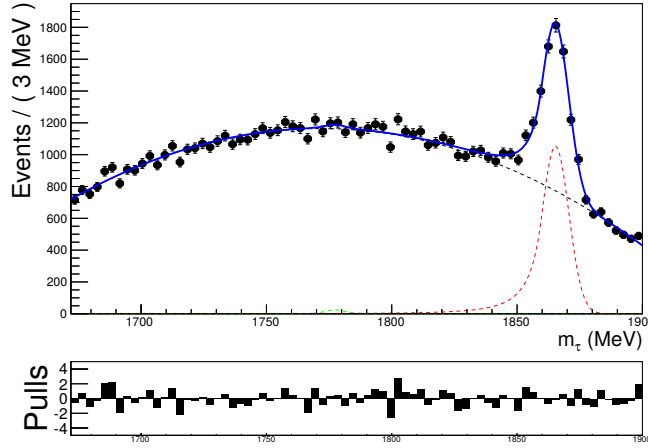


Figure 4.10: Mass distribution of tau candidates passing the full selection in 0.2 fb^{-1} of data. The signal, in dashed green, is fitted with a double-sided symmetric Crystal Ball distribution, the D-peak, in dashed red with a Crystal Ball distribution and the background, in dashed black, with a Chebyshev polynomial of third order. Refer to the text for details.

The fit is shown in Figure 4.10 and its parameters are reported in Table 4.7.

The expected signal yield is given by the formula

Limit setting

$$n_{\text{sgn}} = \mathcal{L} \cdot \sigma(\tau) \cdot \varepsilon \cdot \mathcal{B}(\tau \rightarrow \phi\mu) \cdot \mathcal{B}(\phi \rightarrow K^+K^-)$$

where $\mathcal{L}$ is the integrated luminosity, $\sigma(\tau)$ is the tau production cross-section, ε is the selection efficiency and the symbol $\mathcal{B}$ represents the

Table 4.7: Parameters of the fit to 10% of the 2012 data. The quantities μ_D , σ_D and α_D are parameters of the Crystal Ball function describing the D peak while n_{sgn} , n_D and n_{bkg} are the yields. Refer to the text for details.

n_{sgn}	89 ± 90
n_{bkg}	73061 ± 354
n_D	5412 ± 213
μ_D [MeV]	1865.30 ± 0.19
σ_D [MeV]	5.32 ± 0.19
α_D	1.11 ± 0.11

branching fraction. The values of these quantities are given in Table 4.8. The observed signal, as expected, is compatible with zero.

Table 4.8: Quantities used as input to the limit setting.

n_{sgn}	89 ± 90
$\mathcal{L}$	$205.2 \pm 1.0 \text{ pb}^{-1}$
$\sigma(\tau)$	$88 \pm 15 \text{ }\mu\text{b}$
$\mathcal{B}(\phi \rightarrow K^+K^-)$	0.4890 ± 0.0050
ε	0.001137 ± 0.000019
ε_{GEN}	0.1579 ± 0.0020
$\varepsilon_{\text{STRIP}}$	0.04853 ± 0.00020
$\varepsilon_{\text{TRIG}}$	0.2080 ± 0.0017
$\varepsilon_{\text{CUTS}}$	0.7136 ± 0.0041

We assume that if the fit is repeated on equivalent but independent datasets, n_{sgn} will be Gaussian-distributed with mean zero and standard deviation 90. This hypothesis has been checked with pseudo experiments as explained in Section B.1. The 90% CL upper limit can thus be estimated as

$$\mathcal{B}(\tau \rightarrow \phi\mu) < \frac{1.64 \cdot 90}{\mathcal{L} \cdot \sigma(\tau) \cdot \varepsilon \cdot \mathcal{B}(\phi \rightarrow K^+K^-)} \approx 1.5 \cdot 10^{-5}.$$

If the limit scales as $\mathcal{L}^{-1/2}$, in the full 2012 and 2011 dataset the 90% CL upper limit would be

$$\mathcal{B}(\tau \rightarrow \phi\mu) < 3.8 \cdot 10^{-6},$$

which is below the CLEO limit [95] but almost 50 times larger than the Belle limit [92]. The distribution of the expected upper limit using the full Run I dataset is shown in Figure 4.11 as a function of the assumed branching fraction. The yellow and green bands cover the regions of 68% and 95% containment of compatible observations.

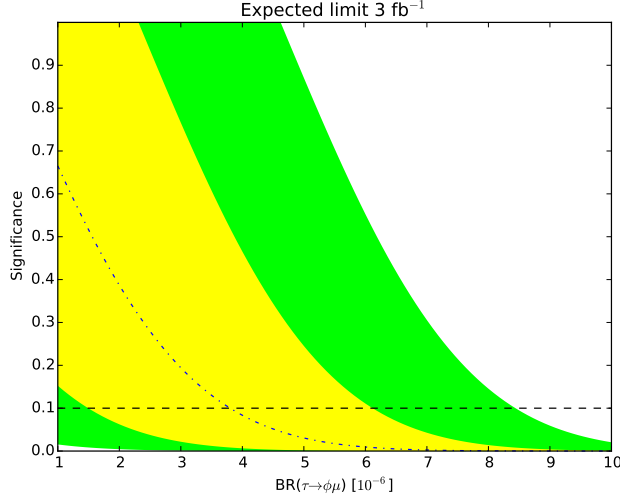


Figure 4.11: Distribution of the expected upper limit using the full Run I dataset as a function of the assumed branching fraction. The yellow (green) bands covers the region of 68% (95%) containment of compatible observations.

4.7 RESULT ON UNBLINDED 2012 DATA

To extract the limit in the whole 2012 dataset the procedure described in section 4.6 is applied. Figure 4.12 shows the result of the fit to the mass spectrum, Table 4.9 reports its parameters and Figure 4.13 shows the distribution of the upper limit. The signal yield is 236 ± 283 , which is compatible with zero. The associated error is compatible with that on the 10% of the data, scaled to account for the ten times larger luminosity, $90 \cdot \sqrt{10}$. The expected 90% CL upper limit is

Table 4.9: Parameters of the fit to the 2012 dataset. The quantities μ_D , σ_D and α_D are parameters of the Crystal Ball function describing the D peak while n_{sgn} , n_D and n_{bkg} are the yields. Refer to the text for details.

n_{sgn}	236 ± 283
n_{bkg}	731960 ± 1065
n_D	50892 ± 578
μ_D [MeV]	1865.300 ± 0.055
σ_D [MeV]	5.357 ± 0.057
α_D	1.258 ± 0.041

$4.7 \cdot 10^{-6}$ and the observed limit is

$$\mathcal{B}(\tau \rightarrow \phi\mu) < \frac{236 + 1.64 \cdot 283}{\mathcal{L} \cdot \sigma(\tau) \cdot \varepsilon \cdot \mathcal{B}(\phi \rightarrow \text{KK})} = 7.1 \cdot 10^{-6}.$$

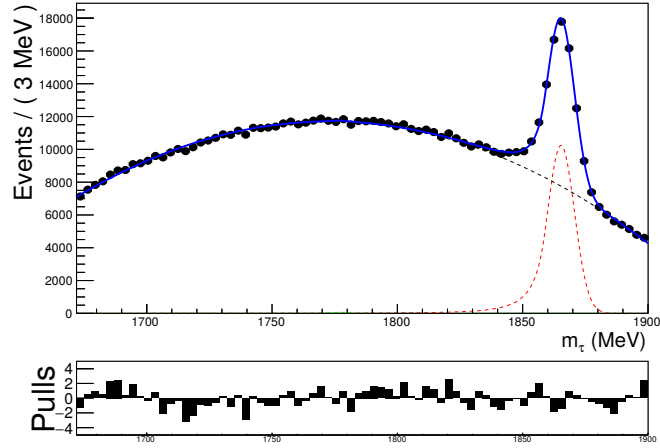


Figure 4.12: Mass distribution of tau candidates in the full 2012 dataset (2 fb^{-1}). The signal, in dashed green, is fitted with a double-sided symmetric Crystal Ball distribution, the D-peak, in dashed red, with a Crystal Ball distribution and the background, in dashed black, with a Chebyshev polynomial of third order. Refer to the text for details.

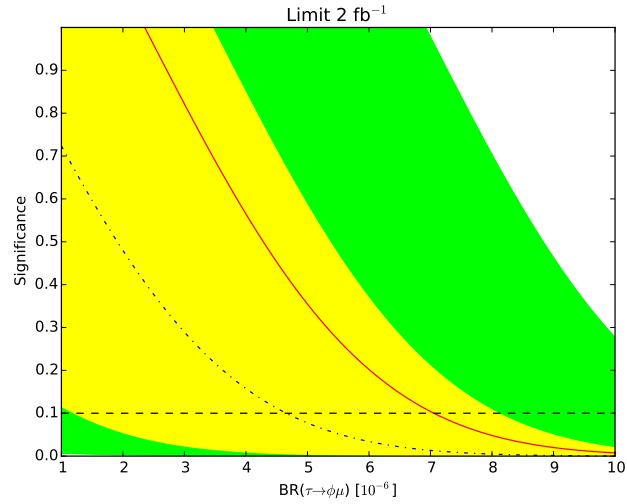


Figure 4.13: Distributions of the expected – blue dash-dot line – and observed – red line – upper limits using the full 2012 dataset, as a function of the assumed branching fraction. The yellow (green) bands covers the region of 68% (95%) containment of compatible observations.

As the expected limit obtainable with the full Run I dataset is not competitive with the World best limit [92] it has been decided not to analyse [92] the full datasample and pursue the result to a publication.

5

OBSERVATION OF CHARMLESS BARYONIC DECAYS

$$B_{(s)}^0 \rightarrow p \bar{p} h^+ h'^-$$

The present chapter describes the search for the six inclusive charmless baryonic decays $B_{(s)}^0 \rightarrow p \bar{p} h^+ h'^-$ ($h^{(\prime)} = \pi, K$) based on the 3 fb^{-1} dataset collected by LHCb during the Run I of the LHC. Contributions from intermediate charmonium and open-charm resonances in the $p \bar{p} h^+ h'^-$ final state are explicitly excluded. Four-body charmless baryonic B_s^0 decays are observed for the first time. The decays $B_s^0 \rightarrow p \bar{p} K^+ K^-$, $B^0 \rightarrow p \bar{p} K^\pm \pi^\mp$ and $B^0 \rightarrow p \bar{p} \pi^+ \pi^-$ are observed with significances greater than 20 standard deviations; the decay $B_s^0 \rightarrow p \bar{p} K^\pm \pi^\mp$ is observed with a significance of 6.5 standard deviations; and evidence at the level of 4.1 standard deviations is found for the $B^0 \rightarrow p \bar{p} K^+ K^-$ decay. Branching fractions are measured relative to the $B^0 \rightarrow J/\psi(\rightarrow p \bar{p}) K^*(892)^0$ normalisation channel in the kinematic region $m(p \bar{p}) < 2.85 \text{ GeV}$. The branching fractions of the $B^0 \rightarrow p \bar{p} K^+ K^-$, $B^0 \rightarrow p \bar{p} \pi^+ \pi^-$ and $B_s^0 \rightarrow p \bar{p} K^+ K^-$ decays relative to the $B^0 \rightarrow p \bar{p} K^\pm \pi^\mp$ mode and of the $B_s^0 \rightarrow p \bar{p} K^+ \pi^-$ decay relative to the $B_s^0 \rightarrow p \bar{p} K^+ K^-$ mode are also provided.

Section 5.1 gives the motivation for the analysis while Section 5.2 describes the general analysis strategy adopted. In Section 5.3 the datasets used are detailed while in Section 5.4 the description of the various stages of the selection is given. Possible sources of background are investigated in Section 5.5. The efficiencies, inputs to the determination of the branching fractions, are computed in Section 5.6, and the fit procedure to extract the signal yields is described in Section 5.7. The sources of possible systematic uncertainties are explored in Section 5.8. The results of the measurements are given in Section 5.9.

For simplicity, in the rest of the chapter, the charges of the $h^+ h'^-$ are omitted unless necessary.

5.1 MOTIVATION

As seen in Section 1.3, the field of charmless baryonic B decays is relatively young and unexplored. In particular very little is known about the family of $B_{(s)}^0 \rightarrow p\bar{p}hh'$ modes¹: in the Particle Data Book [4] (PDG), the only listed decays are

- $\mathcal{B}(B^0 \rightarrow p\bar{p}K^{*0}) = (1.24_{-0.25}^{+0.28}) \times 10^{-6}$, measured by both BaBar [100] and Belle [101];
- $\mathcal{B}(B^0 \rightarrow p\bar{p}\pi^+\pi^-) < 2.5 \times 10^{-4}$ at 90% confidence level (CL), from CLEO [102],

where K^{*0} denotes here and in the following $K^*(892)^0$. Furthermore the field of baryonic B_s^0 decays is even more in its infancy. The decay $B_s^0 \rightarrow p\bar{\Lambda}K^-$ [30] and the decays discussed in this thesis are the first decays in this family to be found. The large data sample collected by LHCb allows either the first observation of various $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decays and the measurement of their branching fractions, or else the setting of stringent upper limits.

Describing these decays theoretically is challenging due to the soft QCD processes involved. In fact, no literature has targeted these modes so far. Figure 5.1 shows typical Feynman diagrams contributing to the decays $B_s^0 \rightarrow p\bar{p}K^+K^-$, $B_s^0 \rightarrow p\bar{p}K^\pm\pi^\mp$, $B^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ and $B^0 \rightarrow p\bar{p}\pi^+\pi^-$. The two $K\pi$ charge combinations are always summed over when considering the $p\bar{p}K\pi$ final states, given the poor theoretical knowledge of the decays, and in spite of the fact that one charge combination should in principle be suppressed relative to the other.

The decays $B^0 \rightarrow p\bar{p}K^+K^-$ and $B_s^0 \rightarrow p\bar{p}\pi^+\pi^-$ proceed via more suppressed diagrams such as the ones shown in Figure 5.2.

A naive estimate of the relative branching fractions can be obtained from the $B_{(s)}^0 \rightarrow hh'$ modes [4]. From each $B_{(s)}^0 \rightarrow hh'$ diagram it is possible to obtain its $B_{(s)}^0 \rightarrow p\bar{p}hh'$ counterpart by radiating a $p\bar{p}$ pair from anywhere in the diagram. Of course this estimate is naive because it does not consider other possible diagrams and the different kinematics, but it provides an estimate that will be used in the optimisation of the selection described in Section 5.4.5.

The present analysis aims at the first observation of the unobserved charmless baryonic $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decay modes and the measurement

¹ Charge conjugation is implied throughout this chapter unless stated explicitly otherwise. Not to over-crowd the notation of the various decay modes, the charges of the $h^+h^{(\prime)-}$ combinations will often be omitted.

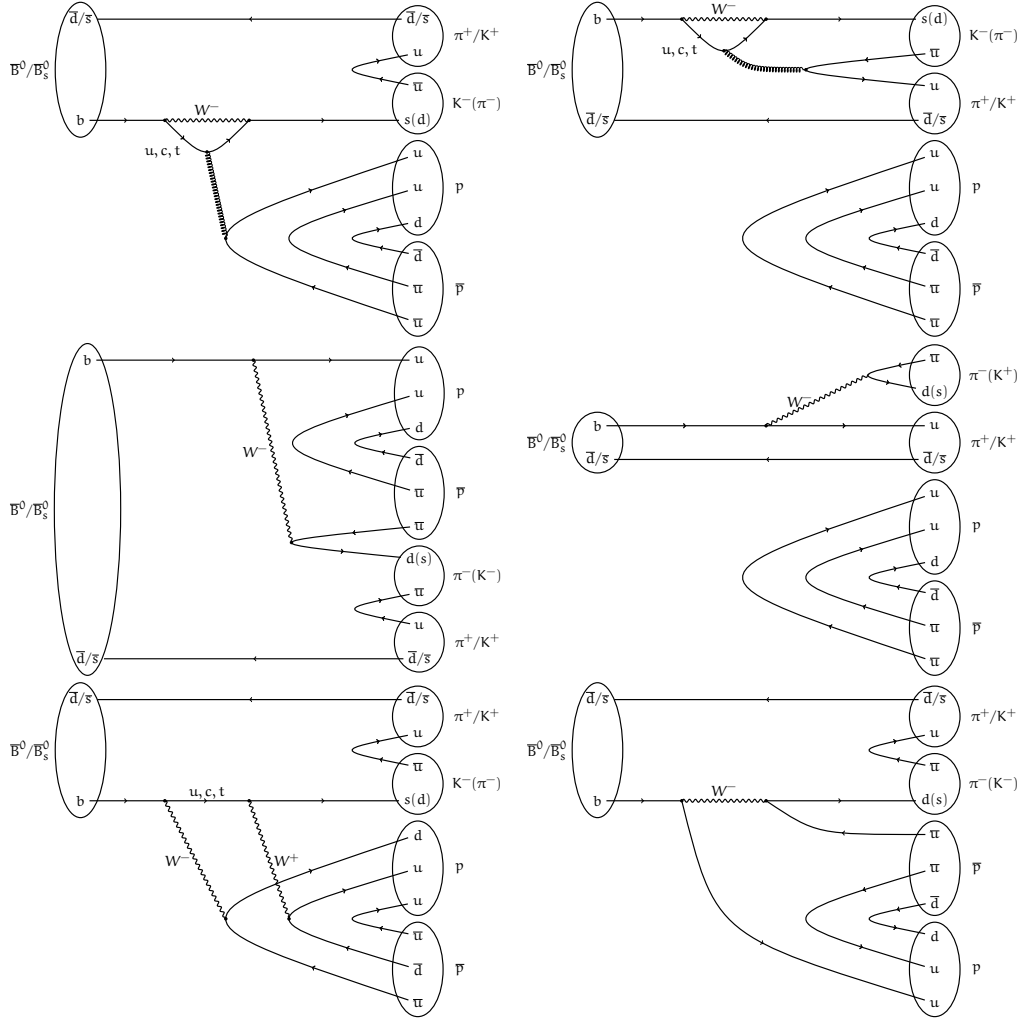


Figure 5.1: Feynman diagrams contributing to $B_s^0 \rightarrow p\bar{p}K^+K^-$, $B_s^0 \rightarrow p\bar{p}K^\pm\pi^\mp$, $B^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ and $B^0 \rightarrow p\bar{p}\pi^+\pi^-$.

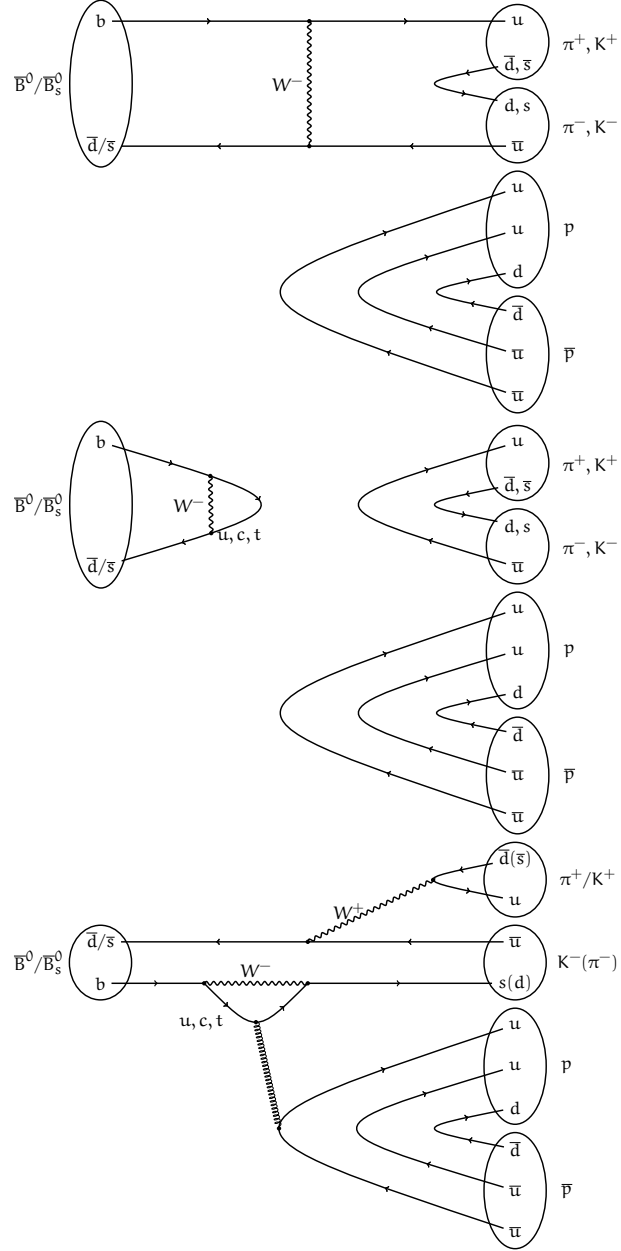


Figure 5.2: Feynman diagrams contributing to $B_s^0 \rightarrow p\bar{p}\pi^+\pi^-$, $B^0 \rightarrow p\bar{p}K^+K^-$ and $B_{(s)}^0 \rightarrow p\bar{p}K^\pm\pi^\mp$.

of their inclusive branching fractions. A natural extension of the analysis could, in the future, focus on the study of CP violation and, in particular, the measurement of triple-product asymmetries.

5.2 ANALYSIS STRATEGY

The main aim of this analysis is the first observation of the charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}hh'$ and the measurement of their branching fractions. Charmful final states are excluded in the various 2- and 3-body sub-systems: charmonium resonances in the $p\bar{p}$ mass spectrum, open charm mesons in the hh' mass spectra, and charmed baryons in the $\bar{p}hh'$ mass spectra. To reduce systematic uncertainties, for example in the determination of the luminosity or the efficiencies, the branching fraction measurements use as a normalisation channel $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$. The branching fractions for the $B_s^0 \rightarrow p\bar{p}hh'$ modes are obtained as

$$\frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}hh')}{\mathcal{B}_{\text{vis}}(B^0 \rightarrow J/\psi K^{*0})} = \frac{N^{\text{corr}}(B_s^0 \rightarrow p\bar{p}hh')}{N^{\text{corr}}(B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-))} \frac{f_d}{f_s}, \quad (5.1)$$

where $N^{\text{corr}} = N/\varepsilon$, N being the yields and ε the total selection efficiencies; f_d/f_s is the ratio of b-quark fragmentation probabilities into B^0 over B_s^0 mesons ($f_s/f_d = 0.259 \pm 0.015$ [103]); and $\mathcal{B}_{\text{vis}}(B^0 \rightarrow J/\psi K^{*0})$ is the $B^0 \rightarrow J/\psi K^{*0}$ visible branching fraction² defined as $\mathcal{B}_{\text{vis}}(B^0 \rightarrow J/\psi K^{*0}) = \mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) \times \mathcal{B}(J/\psi \rightarrow p\bar{p}) \times \mathcal{B}(K^{*0} \rightarrow K^+\pi^-) = (1.68 \pm 0.12) \times 10^{-6}$ where the $B^0 \rightarrow J/\psi K^{*0}$ branching fraction is taken from Ref. [104] and the others from Ref. [4].

Analogously, the branching fraction for $B^0 \rightarrow p\bar{p}hh'$ is expressed as

$$\frac{\mathcal{B}(B^0 \rightarrow p\bar{p}hh')}{\mathcal{B}_{\text{vis}}(B^0 \rightarrow J/\psi K^{*0})} = \frac{N^{\text{corr}}(B^0 \rightarrow p\bar{p}hh')}{N^{\text{corr}}(B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-))}. \quad (5.2)$$

As the normalisation mode has the same final state as the signals (ignoring pion-kaon differences), and the only difference is that it contains a charmonium resonance in the $p\bar{p}$ mass spectrum, the vetoes on the charmonium resonances, open-charm mesons and charmed baryons will be applied only at the end of the selection, in order to minimise the differences in the selection between signal and normali-

² $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) = (1.19 \pm 0.08) \times 10^{-3}$, $\mathcal{B}(J/\psi \rightarrow p\bar{p}) = (2.120 \pm 0.029) \times 10^{-3}$ and $\mathcal{B}(K^{*0} \rightarrow K^+\pi^-) = 2/3 \cdot (99.754 \pm 0.021)\%$ [104, 4].

sation modes. Moreover, such a procedure allows to keep the charm sub-structures in a first version of the mass spectrum fit and hence verify that the charm vetoes applied effectively remove all the charmonium resonances and charm mesons and baryons.

The following ratios of branching fractions among family mode are also measured,

$$\frac{\mathcal{B}(B^0 \rightarrow p\bar{p}KK)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)}, \frac{\mathcal{B}(B^0 \rightarrow p\bar{p}\pi\pi)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)}, \frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)} \text{ and } \frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)}{\mathcal{B}(B_s^0 \rightarrow p\bar{p}KK)}.$$

The significances of the signals will be evaluated from the negative log-likelihood ratios using Wilks' theorem. For channels where no significant signal is observed, *i.e.* in cases where the significance is found to be below 3σ , a 90% confidence level upper limit on the branching fraction will be determined by integrating the likelihood (the upper limit will be the value of the branching fraction that will correspond to 90% of the integral of the likelihood from 0 to ∞).

5.3 DATASET AND SIMULATED SAMPLES

Dataset description

The data used for this analysis consists of approximately 1 fb^{-1} collected during the 2011 run and 2 fb^{-1} collected during the 2012 run at centre-of-mass energies of 7 TeV and 8 TeV, respectively.

MC samples

Several MC samples are used to assess efficiencies and acceptances, and to study background contributions. Samples of 3 million events each have been used for the signal modes $B^0 \rightarrow p\bar{p}K\pi$, $B_s^0 \rightarrow p\bar{p}KK$ and for the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ while 1.5 million events have been used for each of the $B^0 \rightarrow p\bar{p}KK$, $B^0 \rightarrow p\bar{p}\pi\pi$, $B_s^0 \rightarrow p\bar{p}K\pi$ and $B_s^0 \rightarrow p\bar{p}\pi\pi$ decays. Each MC sample contains approximately the same number of events per magnet polarity and twice as many events reconstructed with 2012 conditions than with 2011 conditions. The production of simulation samples within the LHCb collaboration is described in Section 2.3.1. All $B_{(s)}^0$ decays are generated with a 4-body phase-space distribution of the final state particles except the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ where the correct angular distribution derived from the spin of the particles involved in the decay is used. To avoid simulating events which cannot be reconstructed, a selection is applied to the generated particles before simulating the interaction with the detector: all particles are required to be inside the detector ac-

ceptance, with the transverse momentum (p_T) of the protons greater than 250 MeV and the p_T of the pions and the kaons greater than 150 MeV. Finally, the $B_{(s)}^0$ decay time must be greater than 0.1 ps. In the following sections the expression *truth-matched candidates* refers to candidates where both the reconstructed B meson and the reconstructed final state particles can be matched with a corresponding generated particle.

5.4 SELECTION

The various stages of the selection are detailed in the next sub-sections. All events are selected by dedicated trigger and stripping lines (Section 5.4.1). Next, a multivariate classifier is chosen and trained on the kinematical and topological variables (Section 5.4.2), and particle identification (PID) variables are identified (Section 5.4.3) to further separate the signal from the background. Finally, the working point is chosen by maximising a suitable figure of merit (Section 5.4.5) function of both the multivariate classifier output and the PID discriminating variables. The rejection of backgrounds containing muons in the final state is the topic of Section 5.4.6 while multiple candidates are discussed at the end, Section 5.4.7.

5.4.1 Trigger and stripping selection and fiducial region

At the hardware trigger stage, events are required to have a muon with high transverse momentum or a hadron, photon or electron with high transverse energy in the calorimeters. For hadrons, the transverse energy threshold is 3.5 GeV [105]. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from the primary pp interaction vertices (PVs). At least one charged particle must have $p_T > 1.7$ GeV and be inconsistent with originating from a PV. A multivariate algorithm [106] is used for the identification of secondary vertices consistent with the decay of a b hadron. As in the final selection, trigger signals are associated with reconstructed particles, selection requirements can therefore be made not only on which trigger caused the event to be recorded, but also on whether the decision was due to the signal candidate or other particles produced in the pp collision [105]. Only candidates from events with a hardware trigger caused by deposit of the signal in the

Trigger selection

calorimeter, or caused by other particles in the event, are retained. It is also required that the software trigger decision must have been caused by the signal candidate. Increasing the trigger signal efficiency by adding additional trigger lines has also been investigated by evaluating their efficiency, relative to that for candidates passing the stripping selection, using the $B^0 \rightarrow p\bar{p}K\pi$ MC sample. The increase of total trigger efficiency is less than 2%, and hence not worth the increased complexity of the analysis.

Stripping selection

Three dedicated stripping lines select the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ events³. The three lines differ in particular in the loose PID requirements imposed in order to differentiate the three different final states. The $B_{(s)}^0$ candidates are formed combining four charged hadrons. Requirements are made on the quality, the p , p_T , and χ_{IP}^2 of the tracks, loose PID requirements and a lower limit on the $p\bar{p}$ invariant mass are also set. The χ_{IP}^2 is defined as the difference between the vertex-fit χ^2 of a PV reconstructed with and without the particle in question. Each $B_{(s)}^0$ candidate must have a good quality vertex that is displaced from the closest PV (*i.e.* that with which it forms the smallest χ_{IP}^2), must satisfy p and p_T requirements, and must have a reconstructed invariant mass loosely consistent with that of a $B_{(s)}^0$ meson under the signal mass hypothesis. A requirement is also imposed on the angle ϑ_{dir} between the candidate momentum vector and the line between the PV and the candidate decay vertex. Further details of the stripping selection are given in Table 5.1. Figure 5.3 shows the mass spectrum of each of the three final states coming out of the stripping, without further selection. Emerging peaks are just about visible for $B_s^0 \rightarrow p\bar{p}KK$ and $B^0 \rightarrow p\bar{p}K\pi$.

Fiducial region

A further preselection is applied after the stripping selection to unambiguously define the kinematic region which contains the final-state particles. Each particle in the final state (any of the protons, kaons and pions) is required to have momentum smaller than 110 GeV and pseudorapidity between 2 and 5. This will be important in the calibration of the PID selection in Section 5.4.3.

³ For reference the names of these stripping lines are B2pphh_kkLine, B2pphh_kpiLine and B2pphh_pipiLine.

Table 5.1: List of stripping selection criteria.

PROTON SELECTION	
Track fit χ^2/ndof	< 3
Has RICH information	True
p	$> 1500 \text{ MeV}$
p_T	$> 300 \text{ MeV}$
χ^2_{IP}	> 2
Track ghost probability	< 0.35
ProbNNp	> 0.05
PION SELECTION	
Track fit χ^2/ndof	< 3
Has RICH informations	True
p	$> 1500 \text{ MeV}$
p_T	$> 300 \text{ MeV}$
χ^2_{IP}	> 6
Track ghost probability	< 0.35
ProbNNpi	> 0.05
KAON SELECTION	
Track fit χ^2/ndof	< 3
Has RICH informations	True
p	$> 1500 \text{ MeV}$
p_T	$> 300 \text{ MeV}$
χ^2_{IP}	> 4
Track ghost probability	< 0.35
ProbNNk	> 0.05
$p\bar{p}$ SELECTION	
$m_{p\bar{p}}$	$< 4700 \text{ MeV (} p\bar{p}KK \text{)}$
	$< 5000 \text{ MeV (} p\bar{p}K\pi \text{)}$
	$< 5350 \text{ MeV (} p\bar{p}\pi\pi \text{)}$
Sum p	$> 7 \text{ GeV}$
Sum p_T	$> 750 \text{ MeV}$
Max p	$> 4 \text{ GeV}$
Max p_T	$> 400 \text{ MeV}$
Product of ProbNNp	> 0.05
$B_{(s)}^0$ SELECTION	
Mass	$\in [5050, 5550] \text{ MeV}$
Vertex fit χ^2	< 30
Sum p_T of daughters	$> 3000 \text{ MeV}$
p_T	$> 1000 \text{ MeV}$
Particle trajectory minimum distance from a primary vertex	$< 0.2 \text{ mm}$
Distance of closest approach between any two daughters divided by its error	< 20
Cosine of the angle between the particle momentum and its direction of flight	> 0.9999

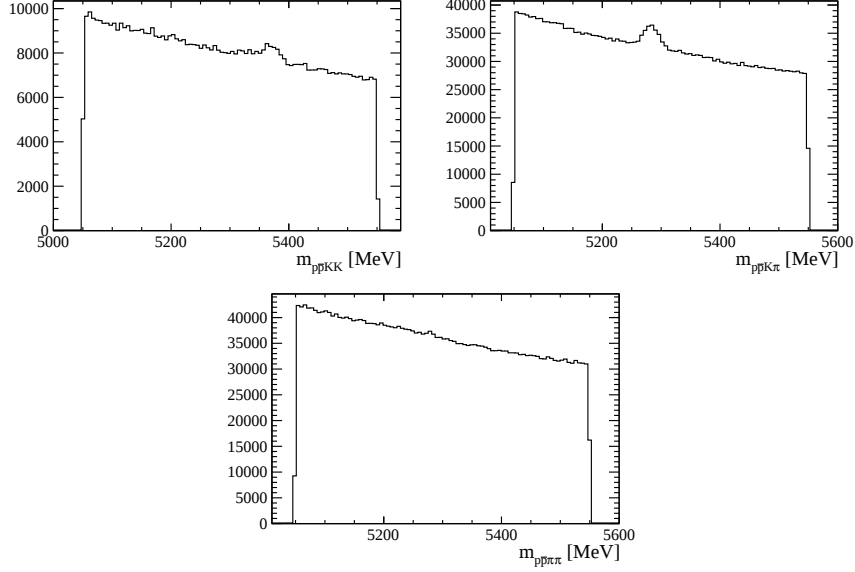


Figure 5.3: Mass spectra obtained after the dedicated stripping selection for *(top left)* the $p\bar{p}KK$ final state, *(top right)* the $p\bar{p}K\pi$ final state and *(bottom)* the $p\bar{p}\pi\pi$ final state.

5.4.2 MVA selection

To increase the signal to background separation, several multivariate classifiers, as implemented in the TMVA package [107], were tried. A BDT was found to give the best performance.

The BDT is trained using as signal sample all the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples described in Section 5.3 together, with both 2011 and 2012 conditions, with candidates truth-matched coming from both B^0 and B_s^0 mesons and decaying in all three modes $p\bar{p}KK$, $p\bar{p}K\pi$ and $p\bar{p}\pi\pi$. As background sample the upper mass sideband of the $p\bar{p}K\pi$ data is used ($5.45 < m_{p\bar{p}K\pi} < 5.55$ GeV) as at this stage of the selection all three $p\bar{p}hh'$ spectra contain many of the same events. A tuning of the hyper-parameters of the BDT was first performed, and the ones chosen are listed in Table 5.2. The decision to combine all $B_{(s)}^0 \rightarrow p\bar{p}hh'$ samples and both years was studied carefully to make sure that no significant loss of performance results from the simplification. Appendix C.1 presents all studies and justifications.

The signal and the background samples were each divided into two independent samples based on the event number; the approach is pseudo-random and reproducible. Two BDTs – denoted hereafter BDT0 and BDT1 – with the same hyper-parameters have thus been trained, each one trained on one of the two independent samples and tested on and applied to the other sample. The discriminating

Table 5.2: Hyper-parameters of the BDT classifier used for the training, after their optimisation. Refer to [107] for a description of all parameters.

VARIABLE	VALUE
Number of trees	300
Max. depth	3
Boost type	Adaptive
AdaBoost β	0.5
Separation type	Gini Index
Min. fraction of events per node	5%

variables used as input are listed in Table 5.3. No attempt has been made to reweight any input variable distribution given that the level of agreement between data and MC was found to be reasonable, see Appendix D.1. The comparison of the distributions of the input variables in the signal and background samples as well as their linear correlations are given in Appendix C.3.

Table 5.3: Variables used as input to the BDT classifier.

VARIABLE	DEFINITION
B_FD_0WNPV	Flight distance of the $B_{(s)}^0$ candidate from the primary vertex
B_DOCAMAX	Largest distance of closest approach between all pairs of decay daughters
B_logDira	Logarithm of the angle between the direction of flight and the momentum of the $B_{(s)}^0$ candidate
Pointing	$\frac{P \sin \vartheta}{P \sin \vartheta + \sum p_T}$ where P is the momentum of the $B_{(s)}^0$ candidate, ϑ is the angle between its direction of flight and its momentum and the sum is over the p_T of the decay daughters
B_ETA	Pseudorapidity of the $B_{(s)}^0$ candidate
B_ENDVERTEX_CHI2NDOF	χ^2/ndof of the vertex fit of the $B_{(s)}^0$ candidate
B_PT	Transverse momentum of the $B_{(s)}^0$ candidate
p_sumPT	Sum of the p_T of the two protons
h_sumPT	Sum of the p_T of the two mesons
h_minPT	Smaller p_T of the two mesons
B_IPCHI2_0WNPV	IP- χ^2 of the $B_{(s)}^0$ candidate
p_sumIPCHI2	Sum of the IP- χ^2 of the two protons
h_sumIPCHI2	Sum of the IP- χ^2 of the two mesons
p_minIPCHI2	Smaller IP- χ^2 of the two protons
h_minIPCHI2	Smaller IP- χ^2 of the two mesons

Figure 5.4 compares the BDT output of both testing and training samples for both BDTs. The agreement between training and testing

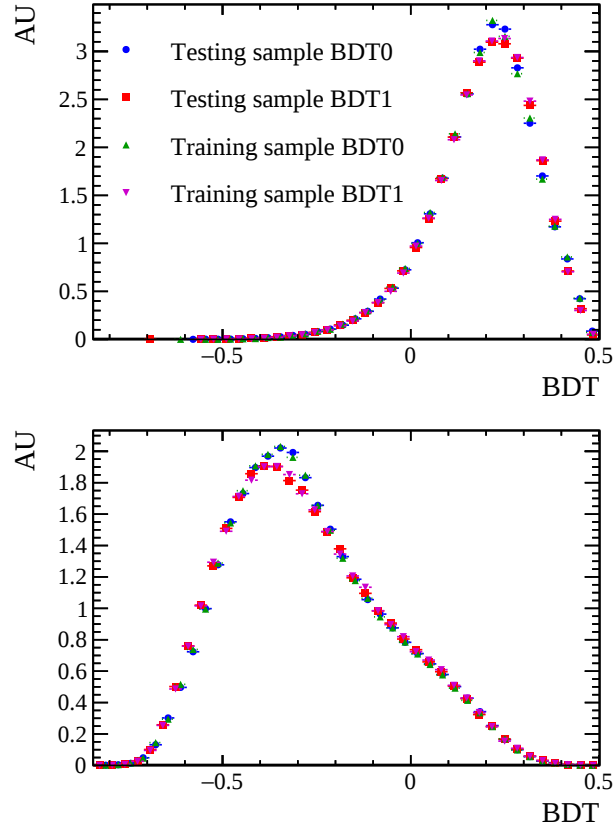


Figure 5.4: Comparison of the BDT output for (*top*) the signal and (*bottom*) the background. The two BDTs trained with the two different halves of the signal and background samples were used on the other halves as testing sample. The plots show the BDT output for both BDTs on both training and testing samples.

samples of a given BDT shows that there is no overtraining while the agreement between BDT outputs of the same sample with different BDTs (training sample for one and testing sample for the other) shows that the two BDTs are equivalent and hence applying one BDT to half the data sample and the other BDT to the other half will give coherent results.

Figure 5.5 shows the good agreement between the receiver operating characteristic (ROC) curves of the two BDTs. It can be seen that more than 85% rejection is obtained with a signal efficiency of 90%.

To avoid biases in the selection, the BDT output should not be correlated with the reconstructed mass. To check this assumption the BDT output is plotted as a function of the $p\bar{p}hh'$ invariant mass in the right-hand sideband of the data. Figure 5.6 shows a flat profile in all three decay modes, which proves that the BDT output and the mass are not strongly correlated.

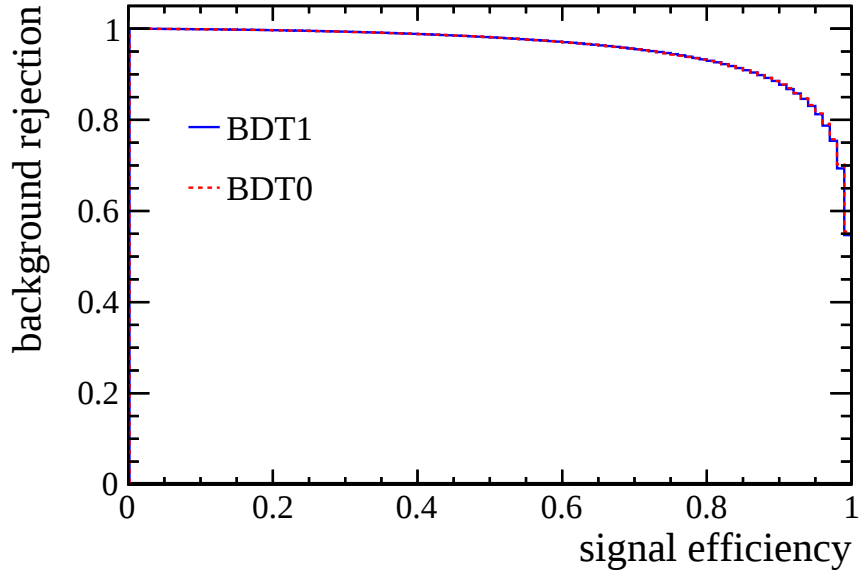


Figure 5.5: Comparison of the ROC curves obtained for both BDTs.

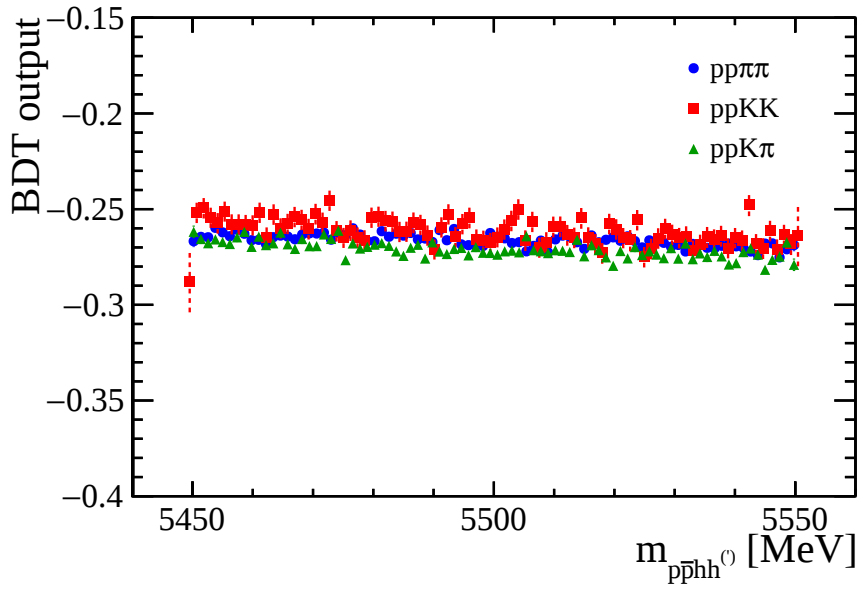


Figure 5.6: Profile of the BDT output as a function of the $p\bar{p}h h'$ invariant mass in the right-hand sideband data.

5.4.3 PID selection

Choice of variables

PID variables are used to further distinguish between signal and background, and in particular in order to reduce reflections, *e.g.* $B^0 \rightarrow p\bar{p}K\pi$ decays reconstructed in the $p\bar{p}KK$ mass spectrum. Various PID variables can be used, and the choice of the most suitable set of variables for this analysis can be made studying ROC curves similarly to what is done to choose between available MVA classifiers. The performance of the proton selection is hereby compared among the following variables:

- ProbNNp;
- PIDp;
- PIDpmu = (PIDp - PIDmu);
- PIDpk = (PIDp - PIDk).

Likewise, the discrimination between kaons and pions is studied with the possible set of variables

- PIDk;
- ProbNNk;
- ProbNNkpiDiff = (ProbNNk - ProbNNpi);
- ProbNNkpi = (ProbNNk · (1 - ProbNNpi)).

The ROC curves are drawn by scanning different values of cuts on each PID variable and evaluating the background retention rate versus the signal efficiency. For this purpose the $B^0 \rightarrow p\bar{p}KK$ MC sample is taken as the signal whereas the right-hand sideband of the data ($5.45 < m_{p\bar{p}KK} < 5.55$ GeV) in the $p\bar{p}KK$ mass spectrum is taken as the background. A preselection cut $BDT > 0$ is applied to both samples. The resulting curves are collected in Figure 5.7. The top figure shows the proton discriminating variables. No variable seems to be notably better than the others; ProbNNp is chosen as proton PID selection variable. The figure at the bottom left compares the variables discriminating between kaons and pions. The variable ProbNNk seems to perform best. However, alone, it does not allow to select disjointly kaons from pions, meaning that the same particle can satisfy at the same time $ProbNNk > x$ and $ProbNNpi > x$; the same can be said for ProbNNkpi. Having a selection that does not allow the same particle

to be identified both as a kaon and a pion allows to reduce multiple candidates in the $p\bar{p}K\pi$ spectrum and guarantees to select different candidates in the three mass spectra. The figure at the bottom right shows the comparison between the same variables as the ones in the bottom-left figure, but for ProbNNkpi and ProbNNk the extra requirement $\text{ProbNNk} > \text{ProbNNpi}$ is added to disjointly select kaons from pions. With this additional condition PIDk and ProbNNkpiDiff perform better; PIDk is chosen as kaon-pion PID selection variable. It should be noted that these studies have been performed with the PID variables directly coming from the MC. It is known that the MC does not provide a good description of the PID variables in data but, as this first study was mainly qualitative, also a low level of agreement was sufficient.

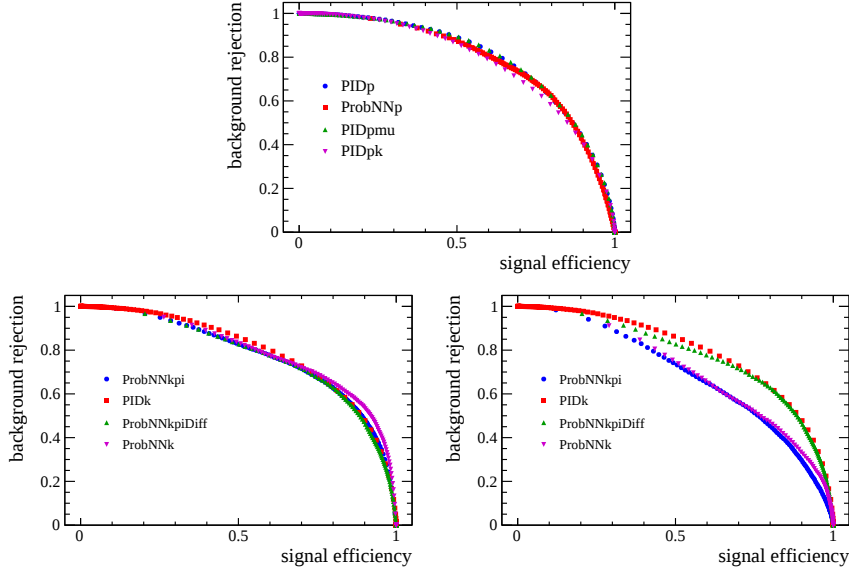


Figure 5.7: ROC curves comparing the performance of various PID variables in discriminating (*top*) protons, (*bottom left*) kaons and (*bottom right*) kaons requiring disjoint samples selecting kaons and pions. Refer to the text for details.

The MC used in LHCb does not reproduce correctly the distribution of the PID variables. Thus a calibration using control samples, where the identity of a particle does not rely on PID selection criteria, is required to compute the efficiencies. A pure sample of pions and kaons is obtained from the decay $D^{*+} \rightarrow (D^0 \rightarrow K^-\pi^+)\pi^+$ while protons are kinematically selected from the decay of a Λ ($\Lambda \rightarrow p\pi^-$) or of a Λ_c ($\Lambda_c^+ \rightarrow pK^-\pi^+$). The *sPlot* technique [108] is used to separate the signal from the background in the calibration samples. It is assumed that the distribution of the PID variables associated to a

Correcting the PID variables in MC

final state track depends only on the kinematics of that track and, to a lesser extent, on the occupancy of the detector. It is thus possible to divide a MC signal sample in bins, for example, of momentum and pseudorapidity of a final state track and assume that the distribution of any PID variable associated with that track is the same across each bin and can be obtained from the calibration samples. Two methods are then possible to compute the efficiency of a selection based on PID variables:

1. Resampling the values of the PID variables in the signal MC sample used to calculate the efficiency. This is done by substituting for each event and for each track in the event the value of each PID variable with a random value extracted from the distribution that the PID variable has in the bin in which the candidate falls. This distribution can be obtained from the calibration samples by fitting in each bin the value of that PID variable using a kernel estimation function.
2. Reweighting the MC samples and computing directly the efficiency of a cut on a PID variable as $\sum_i \varepsilon_i / N$ where the sum is over the candidates in the MC signal sample, N is the total number of candidates in the MC signal sample and ε_i is the per-candidate weight: the efficiency obtained from the calibration samples knowing in which bin the candidate falls. This method is easily extended for cuts on PID variables of several tracks by taking ε_i as the product of the efficiencies of the cuts on PID variables of the different tracks.

Optimal binning

Both methods require a division of the MC signal sample in bins narrow enough that the efficiency of a cut on a PID variable can be considered constant over the bin, but wide enough to have enough tracks such that the accuracy of the method is not compromised by large statistical fluctuations. In this analysis a 2-D binning scheme in momentum and pseudorapidity is adopted for both protons and mesons. The dependence on the event occupancy is discussed below. For calibrating the PID_k variable of the mesons 10 uniform bins in momentum and 4 in pseudorapidity are used while for calibrating the ProbNN_p variable for protons the binning scheme shown in Figure 5.8 is used. Using bins of different width is needed because of the small number of protons with large momentum in the calibration sample. For the purpose of exploring the effect of different cuts it is easier to resample the PID variables and evaluate the efficiencies

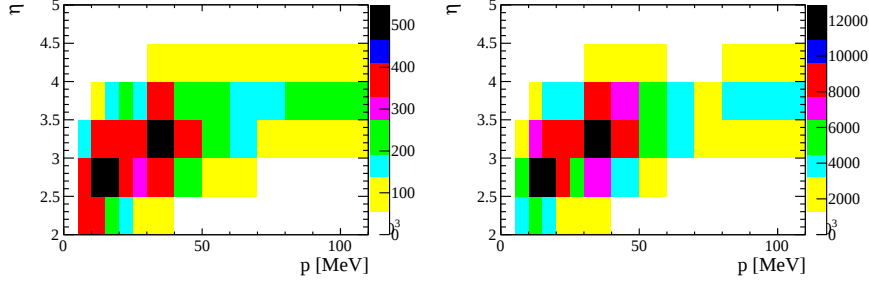


Figure 5.8: Distributions of events for (left) the signal MC and (right) the calibration sample, using the chosen binning scheme.

directly from the corrected MC. Once the selection optimisation has been performed, however, the direct computation of the efficiency of the PID selection is used because it is less affected by the finite size of the MC signal samples. The results of the resampling are given in Appendix D.2.

The efficiency of a cut on a PID variable may depend not only on the kinematic properties of the particle under study, but also on the event occupancy and hence on the number of tracks in the event. It is possible to use the method detailed above and use a binning in 3-D (momentum, pseudorapidity and number of tracks); however, this requires the MC sample to correctly reproduce the distribution of the number of tracks in the signal events. An alternative approach, if one knows the distribution of the numbers of tracks in the signal events under study, is to compute the efficiencies using the 3-D binning scheme and then integrate out the dependence on the number of tracks. Assuming the required analytical expressions to be known, this would be

*Effect of the event
occupancy*

$$\langle \varepsilon(p, \eta) \rangle_f = \int \varepsilon(p, \eta, n_{\text{Tr}}) \cdot f(n_{\text{Tr}}) \, dn_{\text{Tr}},$$

where $\varepsilon(p, \eta, n_{\text{Tr}})$ $\langle \varepsilon(p, \eta) \rangle_f$ are the 3-D and averaged 2-D analytical distributions of the efficiencies, and $f(n_{\text{Tr}})$ is the unity-normalised distribution of the track multiplicity in signal events. The case where the analytical distributions are not known and are replaced by histograms is analogous except that the integral becomes a sum and p , η and n_{Tr} become discrete indices:

$$\langle \varepsilon_{p, \eta} \rangle_f = \sum_{n_{\text{Tr}}} \varepsilon_{p, \eta, n_{\text{Tr}}} \cdot f_{n_{\text{Tr}}}.$$

In this analysis the distribution of the number of tracks is assumed to be the same for all six modes and it is obtained from the *sWeight*'ed

$B^0 \rightarrow p\bar{p}K\pi$ events. In Figure 5.9 the distributions of the number of tracks in the signal and the calibration samples are shown. The distributions are similar between the signal and the calibration samples, as seen in Figure 5.9. Therefore we consider the dependence of PID efficiencies on n_{Tr} to be a second-order effect that does not need to be taken into account in the resampling procedure used during the optimisation, but only for the final evaluation of the PID efficiencies performed by reweighting the MC samples.

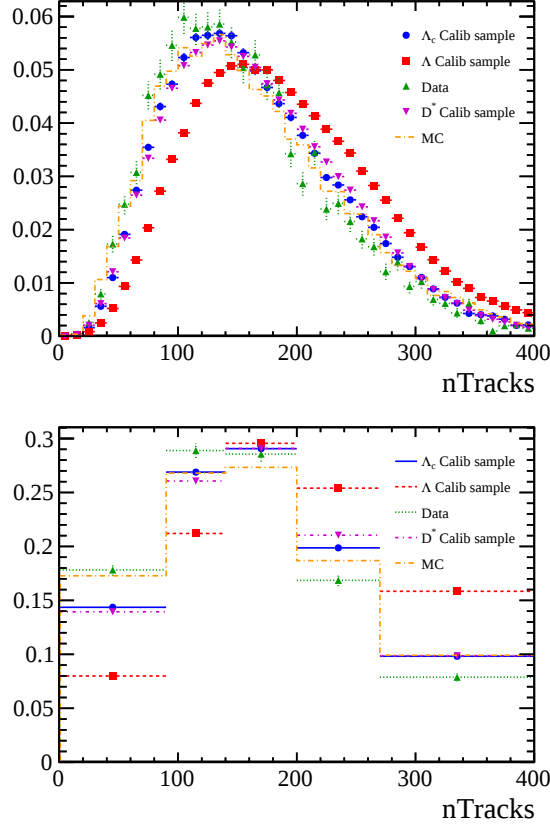


Figure 5.9: Distributions of the number of tracks in the signal and the calibration samples, on the top with a fine binning and on the bottom with the binning used in the procedure.

5.4.4 Charm vetoes

It will be shown in Section 5.7.4 that various charmonium resonances, open-charm mesons and charmed baryons are present in the $m_{p\bar{p}}$, $m_{hh'}$ and $m_{\bar{p}hh'}$ mass spectra. To remove them the following vetoes are applied in order to obtain the charmless spectra:

- Selection of $m_{p\bar{p}} < 2.85 \text{ GeV}$ to remove the all charmonium resonances. This is a commonly used cut, used for example in [109].
- Veto the $\pm 40 \text{ MeV}$ window around the D^0 mass in the $m_{hh'}$ mass spectrum, $1825 < m_{hh'} < 1905 \text{ MeV}$, to remove the D^0 meson contribution.
- Veto the $\pm 25 \text{ MeV}$ window around the Λ_c^+ mass in the $m_{\bar{p}hh'}$ mass spectrum, $2261 < m_{\bar{p}hh'} < 2311 \text{ MeV}$, to remove the Λ_c^+ baryon contribution.

5.4.5 Working point

To find the optimal selection, the following variables are optimised simultaneously: ProbNNp for protons and antiprotons, PIDk for kaons and pions, and BDT output. This implies a sampling in a 3-D space for the $p\bar{p}KK$ and $p\bar{p}\pi\pi$ modes, and in a 4-D space for the $p\bar{p}K\pi$ mode. As we expect to see clear signals in most decay modes, the significance

$$\text{Significance} = \frac{S}{\sqrt{S + B_{\text{tot}}}} = \frac{S}{\sqrt{S + M + B}} \quad (5.3)$$

is an appropriate figure of merit to use. In this equation, S is the signal yield, and two main sources of the total background (B_{tot}) have been separated: the combinatorial background (B) and the misidentified background, *i.e.* the reflection (M).

Table 5.4: Naive expectations of the relative yields before the selection based on the different production fractions and the estimates of the ratios between branching fractions (*cf.* Section 5.1).

Signal	Scaling	Value
$B^0 \rightarrow p\bar{p}KK$	$\frac{\mathcal{B}(B^0 \rightarrow KK)}{\mathcal{B}(B^0 \rightarrow K\pi)}$	0.007
$B^0 \rightarrow p\bar{p}K\pi$	1	1
$B^0 \rightarrow p\bar{p}\pi\pi$	$\frac{\mathcal{B}(B^0 \rightarrow \pi\pi)}{\mathcal{B}(B^0 \rightarrow K\pi)}$	0.26
$B_s^0 \rightarrow p\bar{p}KK$	$\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(B_s^0 \rightarrow KK)}{\mathcal{B}(B_s^0 \rightarrow K\pi)}$	0.33
$B_s^0 \rightarrow p\bar{p}K\pi$	$\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(B_s^0 \rightarrow K\pi)}{\mathcal{B}(B_s^0 \rightarrow K\pi)}$	0.07
$B_s^0 \rightarrow p\bar{p}\pi\pi$	$\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(B_s^0 \rightarrow \pi\pi)}{\mathcal{B}(B_s^0 \rightarrow K\pi)}$	0.01

The dominant reflection in each of the three $p\bar{p}hh'$ spectra is identified in the first instance. Figure 5.10 presents the two real signals (B^0 and B_s^0) and the four reflections for each spectrum, obtained from

the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples after the trigger and the stripping selections. The relative normalisations take into account the naive expected yields before the selection – as discussed in Section 5.1 and summarised in terms of (relative) scaling factors in Table 5.4 – and the (phase-space-integrated) selection efficiencies evaluated from MC, which are detailed in Section 5.6.

For simplicity only the signal mode with highest expected yield is considered for each final state and it is assumed that just one reflection is significant and peaks approximately under the signal peak. Such an assumption is good enough for the purposes of the optimisation at hand. In the $p\bar{p}\pi\pi$ spectrum two reflections seem to have similar expected yields, namely $B^0 \rightarrow p\bar{p}K\pi$ and $B_s^0 \rightarrow p\bar{p}KK$; but only $B^0 \rightarrow p\bar{p}K\pi$ is considered in the optimisation for simplicity and also because $B_s^0 \rightarrow p\bar{p}KK$ has a double misidentification and will therefore be reduced further by cuts on the PID variables. Moreover, only a fifth of the naive expected yield ratio is used for this $p\bar{p}\pi\pi$ spectrum because only approximately a fifth of the misidentified component actually lies under the signal peak. The signal and misidentified decays used in each spectrum are detailed in Table 5.5.

The significance cannot be optimised directly on the data to avoid enhancing artificially the signals, which would bias the measurements. A method based on MC is thus used:

1. Before any selection (also trigger and stripping) one expects to have S_I signal candidates and M_I misidentified candidates. One can define $f = M_I/S_I$, the ratio of misidentified background over the signal produced prior to any selection. This ratio has to be estimated and reasonable values are inferred from Table 5.4 and also reported, for each mode, in Table 5.5.
2. Fit the real data after applying the stripping and trigger selections, and the loose cuts of Table 5.5. Fit with a Gaussian function for the peak and a straight line for the background.
3. From this fit extract the number of background candidates under the signal peak region (B_0) and the normalisation of the peak ($P_0 = S_0 + M_0$). One can define $g = B_0/P_0$, the ratio of continuum (combinatorial) background under the peak over the yield of the peak.
4. Under the peak there are S_0 signal candidates and M_0 misidentified candidates. One can define $\varepsilon_{S0} = S_0/S_I$ and $\varepsilon_{M0} =$

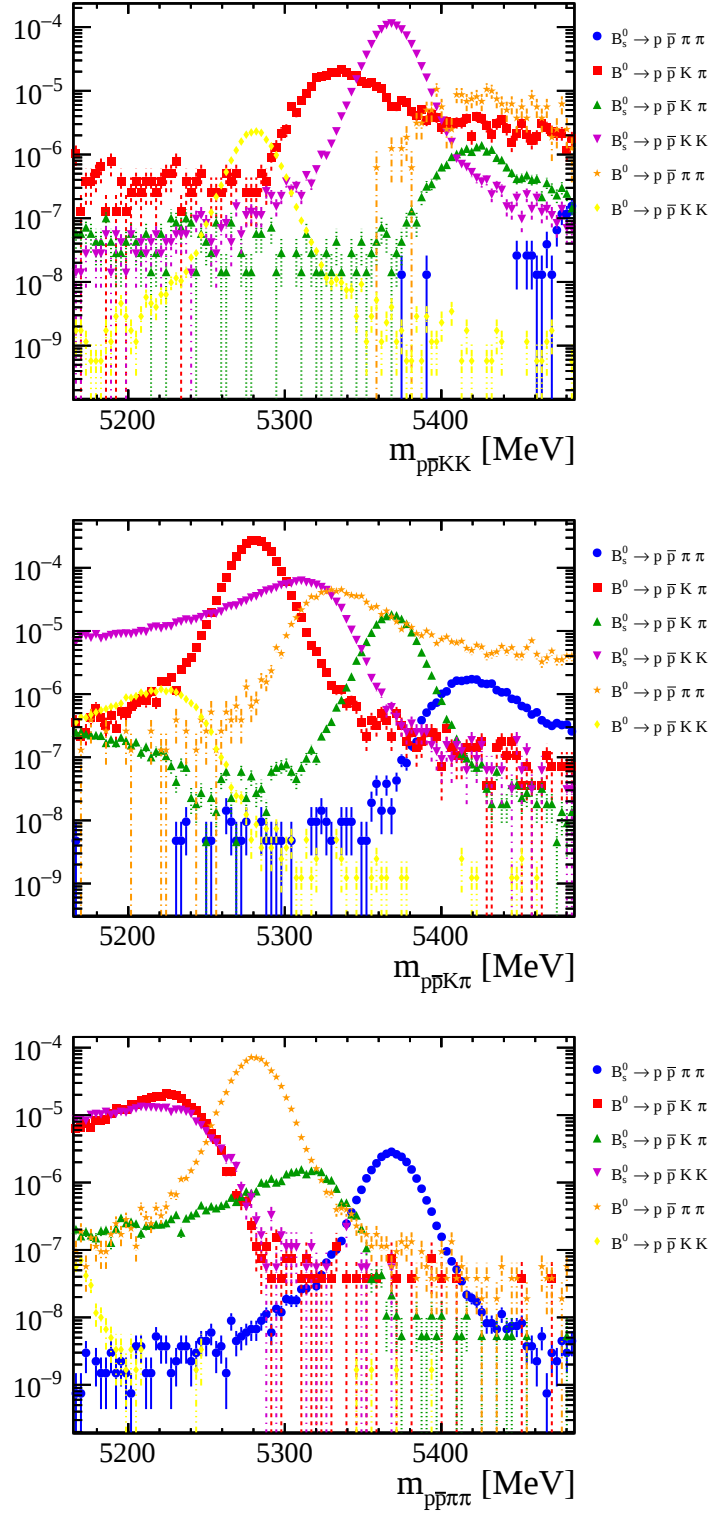


Figure 5.10: Invariant mass distributions for all $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples after trigger and stripping taking into account the naively expected relative branching fractions as detailed in Table 5.4 and the different selection efficiencies. The $p\bar{p}KK$ final state is at the top, $p\bar{p}K\pi$ in the middle and $p\bar{p}\pi\pi$ at the bottom.

Table 5.5: Selection chosen for the various modes together with the MC samples used as signal and misidentified samples in the optimisation. The estimate used as production fraction between the two decays is also given. The XCut variables refer to the cut values for the BDT cut and the PID cuts on the protons, kaons and pions. Refer to the text for details.

Mode	Signal	misID	f	BDTCut	pCut	KCut	piCut
$p\bar{p}K K$	$B_s^0 \rightarrow p\bar{p}K K$	$B^0 \rightarrow p\bar{p}K\pi$	$\frac{f_d}{f_s} \cdot \frac{\mathcal{B}(B^0 \rightarrow K\pi)}{\mathcal{B}(B_s^0 \rightarrow KK)}$	0.1	0.45	5	-
$p\bar{p}K\pi$	$B^0 \rightarrow p\bar{p}K\pi$	$B_s^0 \rightarrow p\bar{p}K K$	$\frac{f_s}{f_d} \cdot \frac{\mathcal{B}(B_s^0 \rightarrow KK)}{\mathcal{B}(B^0 \rightarrow K\pi)}$	0.05	0.35	3	-0.5
$p\bar{p}\pi\pi$	$B^0 \rightarrow p\bar{p}\pi\pi$	$B^0 \rightarrow p\bar{p}K\pi$	$\frac{\mathcal{B}(B^0 \rightarrow K\pi)}{5 \cdot \mathcal{B}(B^0 \rightarrow \pi\pi)}$	0.15	0.5	-	-0.5
Initial loose cuts				0	0.1	0.1	-0.1

M_0/M_I the total efficiencies of the selection up to the loose cuts, and compute them with the MC samples.

5. For each set of MVA and PID cuts applied after the loose selection one can estimate S , M and B . The efficiencies of tightening the selection $\varepsilon_S = S/S_0$ and $\varepsilon_M = M/M_0$ can be computed from the MC while B , the number of combinatorial background events under the peak, can be estimated by fitting the sidebands of the data with a straight line and interpolating it under the peak region. From that it is possible to define $\varepsilon_B = B/B_0$.

It is then possible to rewrite Equation (5.3), up to a global multiplication factor, using the variables defined above, noting that everything can be expressed relative to the initial signal yield: $S = \varepsilon_S \cdot \varepsilon_{S0} \cdot S_I$, $M = \varepsilon_M \cdot \varepsilon_{M0} \cdot f \cdot S_I$ and $B = \varepsilon_B \cdot B_0 = \varepsilon_B \cdot g \cdot (S_0 + M_0)$ with $S_0 + M_0 = S_I \cdot (\varepsilon_{S0} + \frac{M_0}{M_I} \cdot \frac{M_I}{S_I})$, and hence

$$\text{Significance} = \frac{\varepsilon_S}{\sqrt{\varepsilon_S + f \cdot \frac{\varepsilon_{M0}}{\varepsilon_{S0}} \cdot \varepsilon_M + g \cdot \frac{\varepsilon_{S0} + f \cdot \varepsilon_{M0}}{\varepsilon_{S0}} \cdot \varepsilon_B}}. \quad (5.4)$$

In Table 5.5 the result of the optimisation is given together with additional details. Figure 5.11 shows as an example the significance for various values of the BDT cut for the $p\bar{p}KK$ final state.

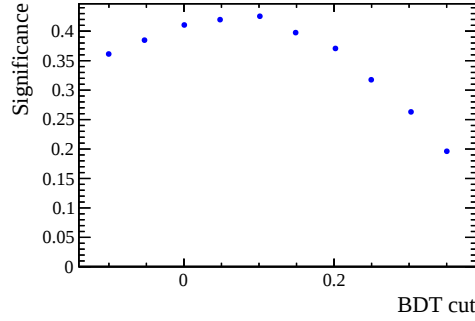


Figure 5.11: Significance for different values of the cut on the output of the BDT classifier for the $p\bar{p}KK$ final state.

5.4.6 Muon misidentification

In the optimisation of the selection the case when (at least) a muon is misidentified as a particle of the final state under study has not been considered. The two bottom plots of Figure 5.12 show that the looser cuts used to select protons and kaons are enough to reject effectively the muons. This is not true for pions as it can be seen

from the top left plot where the muon retention is comparable to the pion efficiency. Adding the request `isMuon==0` (see Section 2.3) as done in the top right plot drastically reduces the muon retention while keeping a high pion efficiency. This extra requirement is thus added to the selection of all the pions in the $p\bar{p}K\pi$ and $p\bar{p}\pi\pi$ final states. Section 5.5 discusses possible sources of background from decays with muons in the final state and argues that these sources are insignificant in spite of the relatively large misidentification rates observed in certain ranges of momentum.

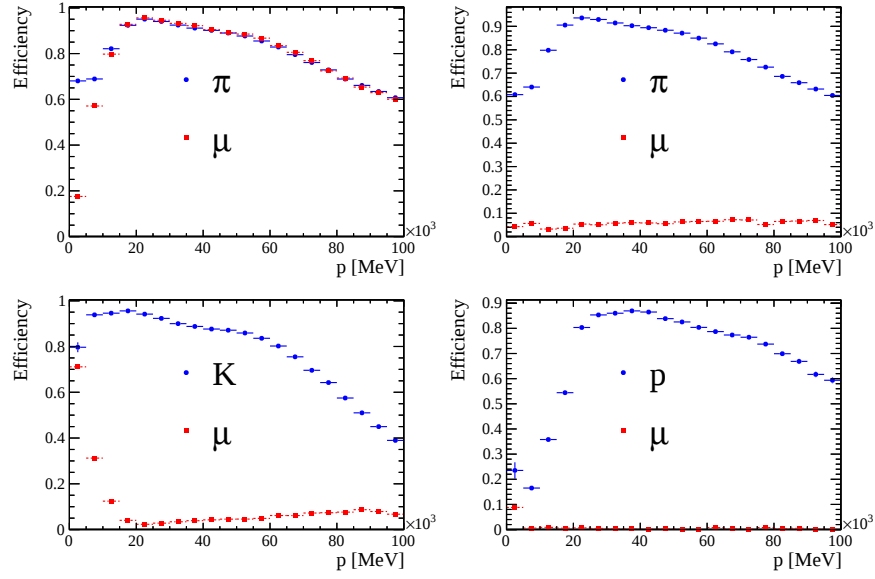


Figure 5.12: Muon retentions versus momentum compared to the efficiencies of the final state particles considered. For each particle the looser cut among the three selection optimisations is considered and the efficiencies are evaluated from calibration samples (muons from $J/\psi \rightarrow \mu\mu$, pions and kaons and protons from the samples described in Section 5.4.3). At the top left pions and muons requiring $PIDk < -0.5$, at top right the extra requirement `isMuon==0` is added. Bottom left kaons and muons requiring $PIDk > 3$ and bottom right protons and muons requiring $ProbNNp > 0.35$.

5.4.7 Multiple candidates

After the full selection approximately 2 – 3% of events contain multiple candidates. A single candidate is then chosen in a pseudo-random way in order to make this selection reproducible and avoid biasing the distribution of background candidates.

5.5 BACKGROUND STUDIES

Different sources of background to the $p\bar{p}hh'$ spectra have been considered. Background can originate from similar final states with at least one misidentified daughter particle, from partially reconstructed backgrounds or from decays to final states containing charm hadrons.

Considering the three possible final states $p\bar{p}KK$, $p\bar{p}K\pi$ and $p\bar{p}\pi\pi$, the signal in one mass spectrum can be present also in another mass spectrum as a reflection due to the misidentification of a meson; for example $B^0 \rightarrow p\bar{p}K\pi$, the prominent signal in the $p\bar{p}K\pi$ mode, is a reflection in the $p\bar{p}KK$ mode if the pion is misidentified as a kaon. These reflections are modelled in the simultaneous fit as explained in Section 5.7.

Internal reflections

The inclusive decay of a B meson to a $p\bar{p}hh'$ final state can proceed via charm hadron intermediate states, which are of no interest in the present analysis of charmless final states. These backgrounds comprise (1) charmonium resonances in the $p\bar{p}$ mass spectrum, (2) open-charm mesons in the hh' spectrum, and (3) charmed baryons in the $\bar{p}hh'$ spectrum. All these charm states are effectively removed with the veto cuts detailed in Section 5.4.4. Light resonances such as the ϕ or the K^{*0} , which do not contain charm, are kept in our inclusive study of charmless $p\bar{p}hh'$ final states.

Final states with charm hadrons

Partially reconstructed backgrounds are assumed to be dominated by the yet unobserved $B_{(s)}^0 \rightarrow p\bar{p}hh'\pi^0$ decay modes: as the π^0 is not reconstructed, these decays have the same final states as the signals. The mass distributions are reasonably smooth and only the $B_s^0 \rightarrow p\bar{p}hh'\pi^0$ decays can peak under the B^0 signal region. In the $p\bar{p}\pi\pi$ final state also contributions from the yet unobserved $B_{(s)}^0 \rightarrow p\bar{p}\eta$ with $\eta \rightarrow \pi\pi\gamma$ are expected. As the photon is not reconstructed these decays have the same final state as the signal and present a smooth mass distribution. The contribution from these decays is hereafter ignored given that the data do not require it. A systematic uncertainty coming from this simplification will be assigned.

Partially reconstructed backgrounds

The decay modes $B_{(s)}^0 \rightarrow p\bar{p}p\bar{p}$ have not yet been studied experimentally, though it is reasonable to expect branching fractions of the order of $10^{-8} - 10^{-6}$ [110]. They give the same final states as the signals under study if a proton-antiproton pair is misidentified as two mesons. However, the wrong-mass hypothesis will result in an expected mass considerably smaller than the mass of the $B_{(s)}^0$ meson –

$B_{(s)}^0 \rightarrow p\bar{p}p\bar{p}$

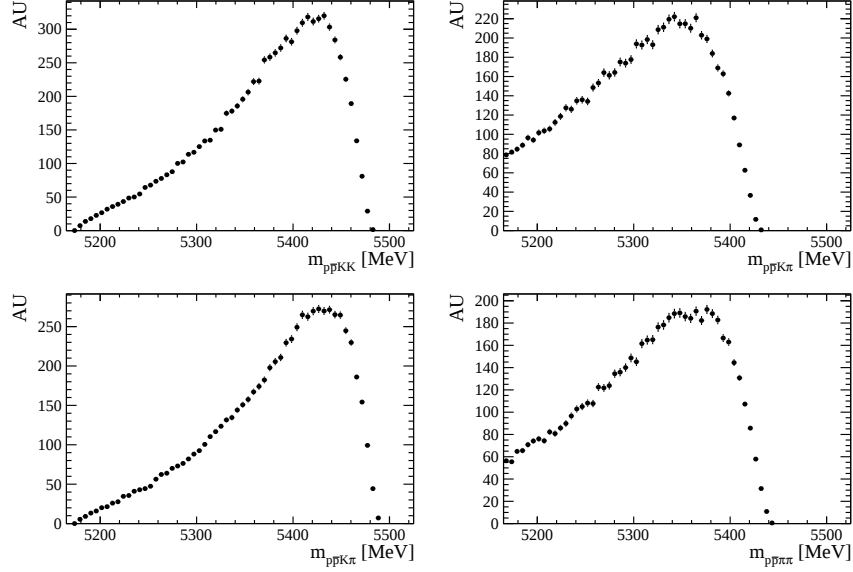


Figure 5.13: Fast simulation of the $\Lambda_b^0 \rightarrow p\bar{p}hh'$ invariant mass in the $p\bar{p}hh'$ mass spectra if a proton is misidentified as a pion or kaon. The mass range is restricted to the region considered in the data fit. (top) $\Lambda_b^0 \rightarrow p\bar{p}pK$, (bottom) $\Lambda_b^0 \rightarrow p\bar{p}p\pi$; (left) $p \rightarrow K$ misidentification, (right) $p \rightarrow \pi$ misidentification.

below 4.9 (4.8) GeV for the B_s^0 (B^0) – and outside the mass range considered in the fit. These decay modes can therefore be safely ignored.

$\Lambda_b^0 \rightarrow p h h h$

The family of decays $\Lambda_b^0 \rightarrow p h h h$ is under study in LHCb, and large branching fractions are observed for various modes. They can only contribute to the $p\bar{p}h h'$ spectra if one of the mesons is misidentified as a proton, which results in a reconstructed mass greater than the mass of the Λ_b^0 , therefore outside the fit region. These decays may also be safely ignored.

$\Lambda_b^0 \rightarrow p\bar{p}p h$

The decays $\Lambda_b^0 \rightarrow p\bar{p}p h$, so far unobserved, can in principle contribute to the $p\bar{p}h h'$ spectra if a proton is misidentified as a pion or kaon. A simplified simulation can be used to examine the expected shapes of these components in the $p\bar{p}h h'$ mass spectra. In this simulation the distribution of the decay products is flat in the phase space and the generated events do not undergo the full reconstruction using the same software used to reconstruct the data. The detector resolution is included by smearing the momentum of the final state particles with a Gaussian distribution with sigma of 0.6% [67]. Figure 5.13 shows the resulting invariant mass distributions for the two decay modes and the two possible misidentifications. It is however evident from the fits to the signal samples, shown in Figure 5.25, that these components are not present after the full selection. To further

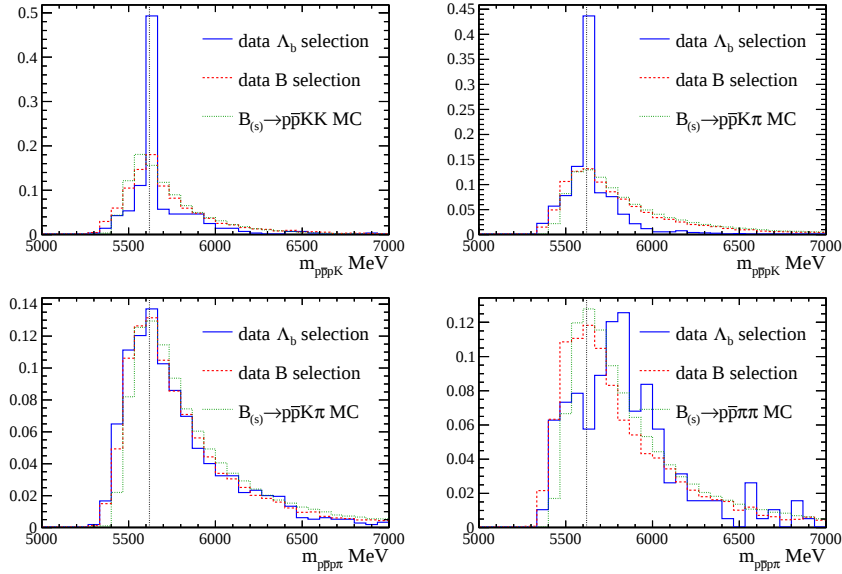


Figure 5.14: Data and signal-only MC reconstructing the final state as $p\bar{p}h$. The data are selected using PID variables: the dashed red histograms are selected with the nominal signal selection while for the solid blue ones the PID selection is optimised to select Λ_b^0 . The black-dotted vertical line indicates the nominal Λ_b^0 mass. The green dotted histogram shows the signal-only MC, with selections as for the data shown as the red dashed histogram. (top) $\Lambda_b^0 \rightarrow p\bar{p}K$, (bottom) $\Lambda_b^0 \rightarrow p\bar{p}\pi$; (left) $p \rightarrow K$ misidentification, (right) $p \rightarrow \pi$ misidentification. Refer to the text for details.

check for the possible presence of $\Lambda_b^0 \rightarrow p\bar{p}h$ in the data, the default selected events are shown in Figure 5.14 after changing the mass hypotheses, *i.e.* $K \rightarrow p$ and $\pi \rightarrow p$. The red dashed histogram represents the candidates selected from the $p\bar{p}hh'$ stripping lines falling in the signal region considered in the fit ($5165 < m_{p\bar{p}K\pi} < 5525$ MeV) after requiring the nominal offline selections on BDT and PID. It can be seen that there is a reasonably good agreement with the signal-only MC selected similarly (the dotted green histogram). Conversely, the blue solid histograms represent the same data as the red dashed line but with the selection on one of the mesons changed to $\text{ProbNNp} > 0.5$ to better select $\Lambda_b^0 \rightarrow p\bar{p}h$. As a reference, the black dotted vertical line is drawn at the Λ_b^0 mass. It can be seen that in the data, with the selection optimised for Λ_b^0 , the presence of $\Lambda_b^0 \rightarrow p\bar{p}K$ is clear while the nominal selection effectively removes this component. On the other hand no $\Lambda_b^0 \rightarrow p\bar{p}\pi$ peak is clearly visible. We thus conclude that the contamination of $\Lambda_b^0 \rightarrow p\bar{p}h$ after the nominal selection can be neglected in the baseline fit model. The effect of this component

will, however, be evaluated when assessing systematic uncertainties in Section 5.8.6.

Decays with muons

As detailed in Section 5.4.6, the muon to hadron misidentification rates can be up to the level of a few percent in certain ranges of momentum. Decays with muons in the final state are hence a potential source of background. No decay with muons in the final state is however of concern after the PID selection. As fully reconstructed decays must contain at least two muons to preserve lepton flavour, a double misidentification would be required. Moreover decays where at least one muon must be misidentified as a proton can be neglected because of the low misidentification probability. This leaves decays with $p\bar{p}\mu\mu$ as final state that should be dominated by $B_{(s)}^0 \rightarrow (J/\psi \rightarrow \mu^+\mu^-)p\bar{p}$. These decays are however harmless since the visible branching fractions – not accounting for misidentifications – are smaller than 3×10^{-8} and 3×10^{-7} at 90% CL for the B^0 and B_s^0 decays, respectively [111]. Partially reconstructed decays, such as $\Lambda_b^0 \rightarrow \Lambda_c^+ \mu \nu_\mu$, with $\Lambda_c^+ \rightarrow pK^-\pi^+$ are also not possible sources of background as they require a muon to proton misidentification or two misidentifications and in all cases the 4-body invariant mass gets shifted to masses above the Λ_b^0 mass.

5.6 EFFICIENCIES

The total selection efficiencies appearing in Equations 5.1 and 5.2 can be factorised in terms of the efficiencies for the various steps of the selection,

$$\varepsilon = \varepsilon_{\text{GEN}} \cdot \varepsilon_{\text{RECO+STRIP}} \cdot \varepsilon_{\text{TRIG}} \cdot \varepsilon_{\text{FIDUCIAL}} \cdot \varepsilon_{\text{BDT}} \cdot \varepsilon_{\text{PID}} \cdot \varepsilon_{\text{noMult}}.$$

The exact meaning of the various terms of this product is given in the next section. In this analysis all the efficiencies are evaluated trivially with truth-matched candidates from the MC samples described in Section 5.3, except for the PID efficiency, which is evaluated using the calibration samples detailed in Section 5.4.3. Section 5.6.1 presents the phase-space-integrated efficiencies: efficiencies integrated over phase space assuming a constant matrix element. Corrections for the observed distributions of events across the phase space are detailed in Section 5.6.2. For all the efficiencies reported in this section only the statistical uncertainty is reported.

In all subsequent sections the efficiencies are evaluated separately for 2011 and 2012, and combined and reported as the average

$$\langle \varepsilon \rangle = \frac{\int \mathcal{L}_{2011} \cdot \varepsilon_{2011} + \int \mathcal{L}_{2012} \cdot \varepsilon_{2012}}{\int \mathcal{L}_{2011} + \int \mathcal{L}_{2012}},$$

where $\int \mathcal{L}_{2011} = 974.6 \pm 0.8 \text{ pb}^{-1}$ and $\int \mathcal{L}_{2012} = 1978.7 \pm 0.8 \text{ pb}^{-1}$ are the values of the integrated luminosity considered for this analysis.

5.6.1 Efficiencies for flat phase space decay models

Table 5.6 shows the total phase-space-integrated efficiencies for the different possible final states (including the efficiencies of misidentified decay modes). Table 5.7 details these efficiencies for all stages of the selection. For convenience Table 5.8 shows all per-selection-level efficiencies including the efficiencies of misidentified decay modes. The errors in the tables are statistical only. The associated systematic uncertainties are discussed in Section 5.8.

ε_{GEN} is the efficiency of the generator level cuts, mainly on the p_T of the final-state particles, the minimum lifetime of the $B_{(s)}^0$ meson and the requirement that all the final-state particles are in the LHCb acceptance. It can be seen that between the different $B_{(s)}^0 \rightarrow p\bar{p}hh'$ modes the generator level efficiencies are approximately the same. It can also be seen that there is a hierarchy with $\varepsilon_{\text{GEN}}^{p\bar{p}KK} > \varepsilon_{\text{GEN}}^{p\bar{p}K\pi} > \varepsilon_{\text{GEN}}^{p\bar{p}\pi\pi}$, which can be explained by the larger average angle between the final-state particles: when the Q-value is larger the probability to have all the final-state particles inside the acceptance is smaller. The generator-level efficiency for the normalisation mode is not of the same order because of the looser generator level cuts; however, the combined efficiency for generator level cuts and reconstruction and stripping agrees with the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ modes.

$\varepsilon_{\text{RECO+STRIP}}$ is the efficiency of the reconstruction and stripping selection when the generator-level cuts have already been applied. It can be seen that they are of the same order between the different $B_{(s)}^0 \rightarrow p\bar{p}hh'$ modes, with a hierarchy $\varepsilon_{\text{RECO+STRIP}}^{p\bar{p}\pi\pi} > \varepsilon_{\text{RECO+STRIP}}^{p\bar{p}K\pi} > \varepsilon_{\text{RECO+STRIP}}^{p\bar{p}KK}$ due to the looser selection for pions over kaons. The efficiencies for B_s^0 are moderately bigger than for the B^0 due to the larger Q-value and harder distributions of momentum and p_T . The efficiency for $B^0 \rightarrow (J/\psi \rightarrow$

$p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ is lower because of the looser generator-level cuts but, as said before, the combined efficiency agrees well with the other modes. As the PID efficiency is computed later from calibration samples the PID cuts have been removed from the stripping before computing this efficiency.

$\varepsilon_{\text{TRIG}}$ is the trigger efficiency of the candidates passing the stripping selection. As expected, as there are no PID requirements in the trigger, the efficiencies are all similar. It is possible to see that final states with higher Q-value have a slightly bigger efficiency.

$\varepsilon_{\text{FIDUCIAL}}$ is the efficiency of the fiducial cuts in pseudo-rapidity and momentum. Again, all efficiencies are roughly the same, with a hierarchy favouring lower Q-values and hence smaller angles between final-state particles so a higher probability for them to be in the fiducial η -region.

ε_{BDT} is the efficiency of the cut on the BDT output. The hierarchy is the one expected by the different values of the cuts.

ε_{PID} is the efficiency of the cuts on the PID variables, including the one on `isMuon`. The hierarchy is the one expected by the different values of the cuts. These efficiencies are evaluated from calibration samples as detailed in Section 5.4.3.

$\varepsilon_{\text{noMult}}$ is the efficiency of the requirement that there is only one candidate per event. As expected, the looser the preceding selections the higher the probabilities to have multiple candidates.

5.6.2 Phase space dependence

To compute the efficiencies directly from MC, taking the ratio of selected and generated events, is equivalent to assuming the absence of substructures in $B_{(s)}^0 \rightarrow p\bar{p}hh'$, as all signal samples are generated with a constant matrix element across the phase space. This is an issue if the efficiencies are not uniform across the phase space. To take into account the presence of substructures that will be shown to be present in Figures 5.21 to 5.23 one can consider the efficiencies in

Table 5.6: Total phase-space-integrated selection efficiencies in % for all decay modes considered. The errors are statistical only (binomial).

Decay channel	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}K\bar{K}$	0.1098 ± 0.0012	0.00714 ± 0.00027	0.000055 ± 0.000024
$B_s^0 \rightarrow p\bar{p}K\bar{K}$	0.11166 ± 0.00080	0.00771 ± 0.00020	0.000064 ± 0.000018
$B^0 \rightarrow p\bar{p}K\pi$	0.001205 ± 0.000073	0.12926 ± 0.00076	0.000952 ± 0.000065
$B_s^0 \rightarrow p\bar{p}K\pi$	0.000853 ± 0.000086	0.1317 ± 0.0012	0.001011 ± 0.000093
$B^0 \rightarrow p\bar{p}\pi\pi$	0.000025 ± 0.000014	0.00907 ± 0.00027	0.06442 ± 0.00075
$B_s^0 \rightarrow p\bar{p}\pi\pi$	0.0000083 ± 0.0000081	0.00697 ± 0.00024	0.06606 ± 0.00076
$B^0 \rightarrow J/\psi K^*$	0.00115 ± 0.00012	0.1249 ± 0.0013	0.00081 ± 0.00010

Table 5.7: Phase-space-integrated selection efficiencies in % for all stages of the selection for all decay modes considered. The errors are statistical only (binomial).

Decay channel	ϵ_{GEN}	$\epsilon_{\text{RECO+STRIP}}$	ϵ_{TRIG}	$\epsilon_{\text{FIDUCIAL}}$	ϵ_{BDT}	ϵ_{PID}	ϵ_{noMult}	ϵ_{tot}
$B^0 \rightarrow p\bar{p}K\bar{K}$	15.746 ± 0.090	10.087 ± 0.025	27.57 ± 0.11	84.52 ± 0.18	77.74 ± 0.22	38.79 ± 0.29	98.37 ± 0.14	0.1098 ± 0.0012
$B_s^0 \rightarrow p\bar{p}K\bar{K}$	15.631 ± 0.041	10.282 ± 0.018	28.007 ± 0.081	84.19 ± 0.13	77.71 ± 0.16	38.55 ± 0.20	98.38 ± 0.10	0.11166 ± 0.00080
$B^0 \rightarrow p\bar{p}K\pi$	13.891 ± 0.016	10.033 ± 0.018	28.872 ± 0.082	83.72 ± 0.13	85.34 ± 0.14	46.29 ± 0.20	97.14 ± 0.11	0.12926 ± 0.00076
$B_s^0 \rightarrow p\bar{p}K\pi$	13.809 ± 0.045	10.360 ± 0.025	29.10 ± 0.11	83.07 ± 0.18	84.96 ± 0.18	46.01 ± 0.27	97.43 ± 0.14	0.1317 ± 0.0012
$B^0 \rightarrow p\bar{p}\pi\pi$	12.283 ± 0.043	10.165 ± 0.025	29.84 ± 0.12	82.98 ± 0.18	65.20 ± 0.24	32.19 ± 0.29	99.28 ± 0.11	0.06442 ± 0.00075
$B_s^0 \rightarrow p\bar{p}\pi\pi$	12.240 ± 0.042	10.383 ± 0.025	30.12 ± 0.12	82.59 ± 0.18	65.24 ± 0.24	32.22 ± 0.29	99.43 ± 0.10	0.06606 ± 0.00076
$B^0 \rightarrow J/\psi K^*$	42.661 ± 0.059	3.0071 ± 0.0097	30.46 ± 0.15	82.43 ± 0.23	87.86 ± 0.22	45.43 ± 0.34	97.11 ± 0.19	0.1249 ± 0.0013

Table 5.8: Phase-space-integrated efficiencies in % for all stages of the selection for all decay modes considered, including the cross-feed decays. The errors are statistical only (binomial). A 100 without error means that all candidates were selected in the signal MC sample considered.

Decay channel	ϵ_{GEN}	$\epsilon_{\text{RECO+STIRP}}$				$\epsilon_{\text{TRIGGER}}$	
		$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}K\bar{K}$	15.746 ± 0.090	10.087 ± 0.025	16.946 ± 0.031	4.424 ± 0.017	27.57 ± 0.11	27.657 ± 0.088	26.89 ± 0.17
$B_s^0 \rightarrow p\bar{p}K\bar{K}$	15.631 ± 0.041	10.282 ± 0.018	18.831 ± 0.024	7.089 ± 0.015	28.007 ± 0.081	28.169 ± 0.060	27.840 ± 0.096
$B^0 \rightarrow p\bar{p}K\pi$	13.891 ± 0.016	4.304 ± 0.012	10.033 ± 0.018	4.269 ± 0.012	27.80 ± 0.12	28.872 ± 0.082	28.85 ± 0.12
$B_s^0 \rightarrow p\bar{p}K\pi$	13.809 ± 0.045	3.753 ± 0.015	10.360 ± 0.025	4.795 ± 0.017	28.09 ± 0.18	29.10 ± 0.11	29.31 ± 0.17
$B^0 \rightarrow p\bar{p}\pi\pi$	12.283 ± 0.043	5.373 ± 0.018	17.488 ± 0.032	10.165 ± 0.025	27.49 ± 0.15	29.052 ± 0.088	29.84 ± 0.12
$B_s^0 \rightarrow p\bar{p}\pi\pi$	12.240 ± 0.042	2.452 ± 0.013	15.237 ± 0.030	10.383 ± 0.025	27.18 ± 0.23	29.173 ± 0.095	30.12 ± 0.12
$B^0 \rightarrow J/\psi K^*$	42.661 ± 0.059	1.1946 ± 0.0061	3.0071 ± 0.0097	1.2976 ± 0.0064	27.64 ± 0.23	30.46 ± 0.15	30.21 ± 0.23

Decay channel	$\epsilon_{\text{FIDUCIAL}}$				ϵ_{BDT}	
	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}K\bar{K}$	84.52 ± 0.18	85.65 ± 0.14	88.38 ± 0.24	77.74 ± 0.22	85.36 ± 0.15	60.60 ± 0.38
$B_s^0 \rightarrow p\bar{p}K\bar{K}$	84.19 ± 0.13	84.952 ± 0.099	86.91 ± 0.14	77.71 ± 0.16	85.57 ± 0.10	64.32 ± 0.21
$B^0 \rightarrow p\bar{p}K\pi$	85.54 ± 0.19	83.72 ± 0.13	84.75 ± 0.19	77.19 ± 0.24	85.34 ± 0.14	65.66 ± 0.26
$B_s^0 \rightarrow p\bar{p}K\pi$	86.20 ± 0.27	83.07 ± 0.18	84.00 ± 0.25	75.40 ± 0.36	84.96 ± 0.18	65.07 ± 0.35
$B^0 \rightarrow p\bar{p}\pi\pi$	87.84 ± 0.22	84.71 ± 0.14	82.98 ± 0.18	73.96 ± 0.31	84.56 ± 0.15	65.20 ± 0.24
$B_s^0 \rightarrow p\bar{p}\pi\pi$	88.49 ± 0.32	84.95 ± 0.14	82.59 ± 0.18	68.87 ± 0.49	84.13 ± 0.16	65.24 ± 0.24
$B^0 \rightarrow J/\psi K^*$	85.72 ± 0.34	82.43 ± 0.23	84.72 ± 0.33	82.34 ± 0.40	87.86 ± 0.22	70.14 ± 0.45

Decay channel	ϵ_{PID}				ϵ_{noMult}	
	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}K\bar{K}$	38.79 ± 0.29	1.330 ± 0.050	0.055 ± 0.024	98.37 ± 0.14	99.54 ± 0.33	100
$B_s^0 \rightarrow p\bar{p}K\bar{K}$	38.55 ± 0.20	1.293 ± 0.033	0.037 ± 0.010	98.38 ± 0.10	99.11 ± 0.30	100
$B^0 \rightarrow p\bar{p}K\pi$	1.098 ± 0.066	46.29 ± 0.20	1.002 ± 0.068	100	97.14 ± 0.11	100
$B_s^0 \rightarrow p\bar{p}K\pi$	0.902 ± 0.090	46.01 ± 0.27	0.955 ± 0.088	100	97.43 ± 0.14	100
$B^0 \rightarrow p\bar{p}\pi\pi$	0.021 ± 0.012	2.031 ± 0.059	32.19 ± 0.29	100	100	99.28 ± 0.11
$B_s^0 \rightarrow p\bar{p}\pi\pi$	0.017 ± 0.016	1.805 ± 0.060	32.22 ± 0.29	100	99.39 ± 0.61	99.43 ± 0.10
$B^0 \rightarrow J/\psi K^*$	1.15 ± 0.12	45.43 ± 0.34	0.82 ± 0.10	100	97.11 ± 0.19	100

bins of phase space and compute the average efficiency taking into account the distribution observed in data across the phase space as⁴

$$\langle \varepsilon \rangle = \frac{\sum_i n_i^{\text{sel}}}{\sum_i \frac{n_i^{\text{sel}}}{\varepsilon_i}}, \quad (5.5)$$

where ε_i and n_i^{sel} are, respectively, the efficiency and the number of selected events in the i -th bin of the phase space. The n_i^{sel} can be obtained directly from data with the *sPlot* technique as explained in Section 5.7.4.

Description of the phase space of $B_{(s)}^0 \rightarrow p\bar{p}hh'$

A four-body decay such as $B_{(s)}^0 \rightarrow p\bar{p}hh'$ can be described with five variables given that the angles describing the orientation of the whole decay with respect to the reference frame can be integrated out since the $B_{(s)}^0$ is a scalar meson. The following variables are used to study the 5-D efficiency distributions and 5-D distributions of events in the phase space:

$m_{p\bar{p}}$, the invariant mass of the $p\bar{p}$ pair.

$m_{hh'}$, the invariant mass of the hh' pair.

$\cos \vartheta_{p\bar{p}}$, the cosine of the angle between the momentum of the proton and the plane defined by the hh' pair in the $p\bar{p}$ pair rest frame (if the $p\bar{p}$ pair was a resonance this would be the proton's helicity angle).

$\cos \vartheta_{hh'}$, the cosine of the angle between the momentum of one of the h and the plane defined by the $p\bar{p}$ pair in the hh' pair rest frame.

χ , the angle between the $p\bar{p}$ and the hh' planes.

The baseline description makes use of four bins in $m_{p\bar{p}}$ and $m_{hh'}$ and three bins in $\cos \vartheta_{p\bar{p}}$, $\cos \vartheta_{hh'}$ and χ , for a total of 432 bins.

Statistical error on the efficiencies

The statistical error on an efficiency evaluated on MC as $\varepsilon = n/N$, where N (n) is the number of events before (after) the selection, can be obtained from the variance of the binomial distribution as $\sigma_\varepsilon = \sqrt{\varepsilon \cdot (1 - \varepsilon)/N}$. This formula however needs to be generalised if

⁴ This formula is simply the weighted average of $1/\varepsilon_i$ and it reduces to the usual $\varepsilon = n^{\text{sel}}/n^{\text{gen}}$ if the n_i^{sel} weights are the same used in $\varepsilon_i = n_i^{\text{sel}}/n_i^{\text{gen}}$.

1. The MC is reweighted: for each event of the MC a weight α_i is applied. This is useful for example if for some kinematical distribution the MC sample available does not describe the data reasonably well.
2. Each event has an associated acceptance factor $A_i \in [0, 1]$. This is for example the case with the evaluation of the PID efficiencies detailed in Section 5.4.3 where each event has an associated probability to be selected depending on its kinematical properties.

The generalised expression for σ_ε depends on

$$\begin{aligned} n &= \sum_i A_i \cdot \alpha_i, & N &= \sum_i \alpha_i, \\ w^2 &= \sum_i A_i^2 \cdot \alpha_i^2, & W^2 &= \sum_i \alpha_i^2, \end{aligned}$$

where the sums run over all the candidates. The efficiency and the corresponding uncertainty can be evaluated [112] as

$$\begin{aligned} \varepsilon &= \frac{n}{N}, \\ \sigma_\varepsilon^2 &= \frac{1}{N^2} \left[\left(1 - 2 \cdot \frac{n}{N} \right) w^2 + \frac{n^2}{N^2} \cdot W^2 \right]. \end{aligned}$$

One can observe that this formula reduces to the binomial one if no weights are applied (i.e. $\alpha_i = 1$ and $A_i \in \{0, 1\} \forall i$, hence $w^2 = n$ and $W^2 = N$).

Efficiencies from MC

The values of the efficiencies in bins of phase space are obtained from the MC samples described in Section 5.3 for all the selection steps except for ε_{GEN} and $\varepsilon_{\text{noMult}}$. To evaluate the efficiencies of the generator-level cuts in bins of phase space, other MC samples are used with no generator cuts on the final state particles so that they do not depend on the phase-space distribution of the decay products. The $B_{(s)}^0$ instead is required to be inside the LHCb acceptance and to have a decay time greater than 0.1 ps. These dedicated samples have 20k (10k) events passing the loose generator-level cuts for the 2012 (2011) conditions and the same number of events has been sim-

ulated for each magnet polarity. For $\varepsilon_{\text{noMult}}$ only the phase-space-integrated efficiencies are evaluated because it does not depend on the phase-space distribution of the signal.

With the *sPlot* technique only the distribution of the number of events after the full selection can be obtained; hence in Equation (5.5) the binned efficiency of the full selection should be used. However, to better understand the efficiencies of the various steps of the selection, one can estimate the distribution of events before a selection (N_i) knowing the binned efficiency (ε_i) and the distribution of events after (n_i) as $N_i = n_i / \varepsilon_i$ and hence iteratively apply Equation (5.5) for each step of the selection from the last one to the first. The results of this procedure are collected in Table 5.9. The maximum variation of these phase-space-corrected total selection efficiencies with respect to the phase-space-integrated efficiencies is of 4 – 8%; hence, to avoid introducing interdependencies and the need for an iterative procedure, the phase-space-corrected efficiencies are used only in the final determination of the branching fractions while the selection optimisation and the fit rely on the phase-space-integrated efficiencies. As the number of bins into which the phase space is divided is large, there are bins where just a few MC events fall, yielding large errors in those bins (which can also be larger than the value of the efficiency itself). However, this is not a problem because the efficiency tables are not used to reweight the data before performing the fit but only, in the last step, the ratios of total efficiencies are used to compute the branching fractions. Even if a flat phase space does not model perfectly these decays, the distribution of events over the phase space is not too dissimilar so that in the bins with very large efficiency errors there are also very few events from the data so the final impact on the total error is small as can be seen in Table 5.9.

Efficiencies for each selection step

To further cross check the validity of the phase-space-corrected efficiencies computed with the method described above, the efficiencies have been recomputed with Equation (5.5) taking n_i^{sel} directly from the special MC used to study the generator-level efficiencies. As expected the efficiencies obtained are almost identical to the ones in Table 5.8. A difference of 0.1% is observed due to the bins in which there are no selected events but some generated ones so that Equation (5.5) does not reduce to the simple $n^{\text{sel}}/n^{\text{gen}}$. This will be taken into account in assessing the systematic uncertainty in Section 5.8.4.

Cross check

Table 5.9: Phase-space-corrected efficiencies in % of the various stages of the selection for the various decays considered in the final state with no misidentification. Errors are statistical only. No phase-space correction has been applied to the efficiencies of $B^0 \rightarrow J/\psi K^{*0}$ as it is a two-body decay that is well described in the MC.

Decay channel	ϵ_{GEN}	$\epsilon_{\text{RECO+STRIP}}$	ϵ_{TRIG}	$\epsilon_{\text{FIDUCIAL}}$	ϵ_{BDT}	ϵ_{PID}	ϵ_{noMult}	ϵ_{tot}
$B^0 \rightarrow p\bar{p}KK$	15.42 ± 0.31	10.24 ± 0.11	27.14 ± 0.35	84.63 ± 0.62	77.77 ± 0.84	38.96 ± 0.66	98.37 ± 0.14	0.1079 ± 0.0033
$B_s^0 \rightarrow p\bar{p}KK$	15.28 ± 0.22	10.314 ± 0.035	27.85 ± 0.14	84.21 ± 0.20	76.64 ± 0.31	37.58 ± 0.29	98.38 ± 0.10	0.1047 ± 0.0017
$B^0 \rightarrow p\bar{p}K\pi$	13.71 ± 0.17	10.013 ± 0.027	28.80 ± 0.11	83.56 ± 0.16	84.65 ± 0.18	46.11 ± 0.23	97.14 ± 0.11	0.1252 ± 0.0017
$B_s^0 \rightarrow p\bar{p}K\pi$	13.84 ± 0.65	10.72 ± 0.35	28.85 ± 0.67	83.3 ± 1.1	81.7 ± 1.2	45.7 ± 1.2	97.43 ± 0.14	0.1306 ± 0.0086
$B^0 \rightarrow p\bar{p}\pi\pi$	11.99 ± 0.18	10.087 ± 0.068	29.77 ± 0.19	83.06 ± 0.28	61.82 ± 0.43	32.33 ± 0.35	99.28 ± 0.11	0.0592 ± 0.0012
$B_s^0 \rightarrow p\bar{p}\pi\pi$	12.20 ± 0.39	10.06 ± 0.23	30.20 ± 0.65	81.98 ± 0.91	65.0 ± 1.7	32.60 ± 0.93	99.43 ± 0.10	0.0640 ± 0.0032
$B^0 \rightarrow J/\psi K^*$	42.661 ± 0.059	3.0071 ± 0.0097	30.46 ± 0.15	82.43 ± 0.23	87.86 ± 0.22	45.43 ± 0.34	97.11 ± 0.19	0.1249 ± 0.0013

5.6.3 Efficiencies of the charm vetoes

To evaluate the efficiencies of the charm vetoes outlined in Section 5.4.4 a data-driven method is used. The efficiency of the $m_{p\bar{p}} < 2.85 \text{ GeV}$ selection is not considered since this cut defines the fiducial region where the branching fractions are measured. On the other hand the vetoes around the D^0 and Λ_c^+ hadrons remove also part of the signal that lies under these vetoed resonances. To evaluate how many signal candidates are removed by the vetoes an unbinned maximum likelihood fit to the sidebands of the vetoed region is done. An exponential function is fitted to the *sWeight*'ed dataset for each decay mode after applying the vetoes. As no events are at the same time in the D^0 vetoed region in the hh' mass spectrum and in the Λ_c^+ region in the $\bar{p}hh'$ one (Figure 5.15), it is enough to perform two independent 1-D fits to the two mass spectra and obtain by interpolation, for each decay mode, the number of charmless B signal candidates under the D^0 peak ($\tilde{n}_{D^0}$) and the number of signal candidates under the Λ_c^+ peak ($\tilde{n}_{\Lambda_c^+}$). The efficiency of the veto is then

$$\frac{n_{\text{sel}}}{n_{\text{sel}} + \tilde{n}_{D^0} + \tilde{n}_{\Lambda_c^+}},$$

where n_{sel} is the number of selected events after the vetoes.

Figure 5.16 shows the fits to the sidebands of the charm vetoes for the $B^0 \rightarrow p\bar{p}K\pi$ decay. Table 5.10 contains the efficiencies evaluated for each decay mode along with their statistical errors and the systematic errors that will be described in Section 5.8.4.

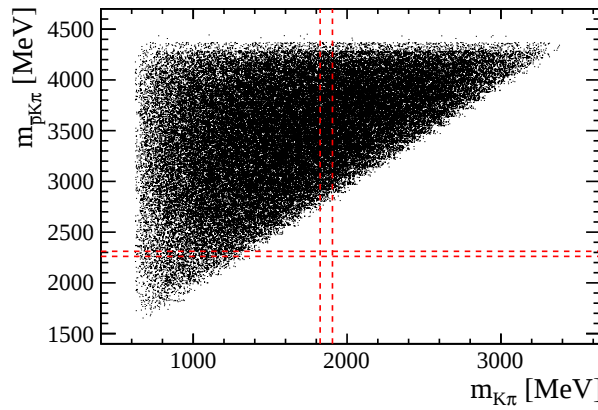


Figure 5.15: Scatter plot of $m_{K\pi}$ vs $m_{pK\pi}$ for $B^0 \rightarrow p\bar{p}K\pi$ signal MC candidates. The red dashed lines delimit the regions removed by the veto.

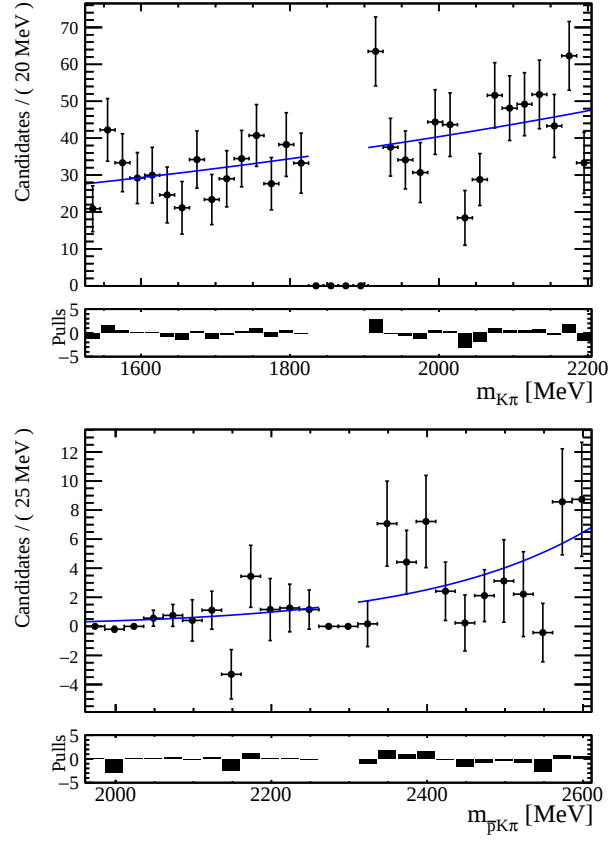


Figure 5.16: Fit to the $sWeight'$ ed $B^0 \rightarrow p\bar{p}K\pi$ signal after the charm vetoes; (top) $K\pi$ mass spectrum and (bottom) $pK\pi$ mass spectrum.

Table 5.10: Efficiencies of the charm vetoes in %; the first uncertainty is statistical and the second systematic. Refer to the text for details.

Decay channel	$\epsilon_{\text{Charm veto}}$
$B^0 \rightarrow p\bar{p}KK$	$96.1 \pm 2.0 \pm 0.4$
$B_s^0 \rightarrow p\bar{p}KK$	$97.0 \pm 0.7 \pm 0.3$
$B^0 \rightarrow p\bar{p}K\pi$	$96.56 \pm 0.28 \pm 0.08$
$B_s^0 \rightarrow p\bar{p}K\pi$	$96.3 \pm 0.9 \pm 0.4$
$B^0 \rightarrow p\bar{p}\pi\pi$	$96.11 \pm 0.61 \pm 0.05$
$B_s^0 \rightarrow p\bar{p}\pi\pi$	$96.8 \pm 1.8 \pm 0.2$

5.7 FITS TO THE DATA

To extract the yields of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ signal modes the following strategy is adopted: firstly, the line shapes of the signal peaks (Section 5.7.1) and of the reflections (Section 5.7.2) are determined from MC. Then, before applying the charm vetoes, a simultaneous fit to the three spectra is performed in Section 5.7.3. This fit is used to obtain the *sWeights* [108] to inspect the $m_{p\bar{p}}$, $m_{hh'}$ and $m_{\bar{p}hh'}$ mass spectra under the signal peaks (Section 5.7.4) and validate the charm vetoes of Section 5.4.4. In Section 5.7.5 the simultaneous fit is performed again after having applied the charm vetoes; this is the nominal data fit from which the analysis results will be derived. Finally, in Section 5.7.7, the fit to extract the yield of the normalisation mode $B^0 \rightarrow J/\psi K^{*0}$ is presented.

5.7.1 Signal shapes

The signal shapes are obtained from MC. The whole selection chain excluding the PID criteria is first applied. The distributions are subsequently weighted with the efficiencies of the PID variables evaluated as detailed in Section 5.4.3. All shapes are fitted with a double-sided Crystal Ball (DSCB) function [97]. The parameters are hence the following: the Gaussian core mean μ and resolution σ ; and the two parameters of each power-law tail, namely ν , which denotes the exponent of the tail, and α , which is the distance between the mean and the beginning of the tail divided by σ . The parameters ν and α are highly correlated, so that there is a region in the parameter space where their values change without a significant change in the goodness of the fit. The parameters α are thus fixed to increase the stability of the fits. Furthermore, if both ν parameters of the two tails are left free to vary independently or are constrained to have the same value, the results of the fits are compatible. It is therefore decided to share the parameter α and ν between the two tails.

The results of the fits are shown in Figure 5.17 whereas all fit parameters are collected in Table 5.11. The means are all consistent among B^0 and B_s^0 decays, as expected. The core resolutions follow the expected pattern of an increasingly better resolution as the Q-value of the decay mode gets smaller.

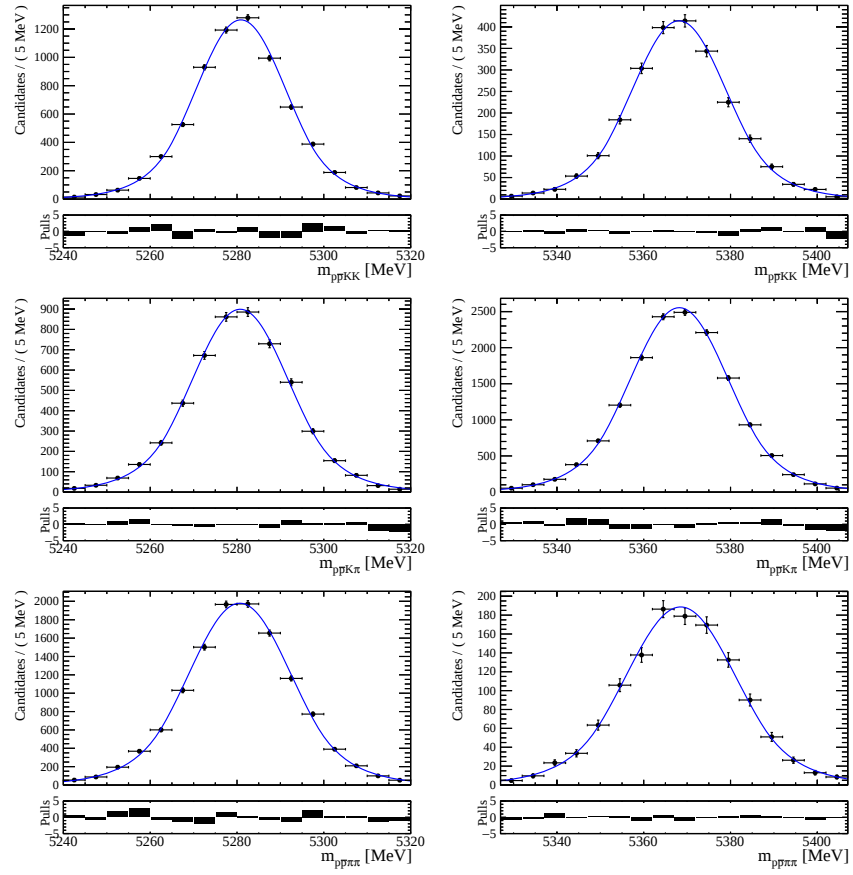


Figure 5.17: Signal shapes obtained from MC, after the full selection, and fitted with a double-sided Crystal Ball function: B^0 to the left and B_s^0 to the right. At the top the $p\bar{p}KK$ final state, $p\bar{p}K\pi$ at the middle and $p\bar{p}\pi\pi$ at the bottom.

Table 5.11: Parameters of the fits to the MC signal shapes.

Parameter	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
μ_B [MeV]	5280.88 ± 0.14	5280.74 ± 0.17	5280.65 ± 0.12
σ_B [MeV]	10.44 ± 0.11	11.18 ± 0.14	11.85 ± 0.11
α_B	1.50	1.50	1.50
ν_B	150.70 ± 0.85	151.83 ± 0.18	151.91 ± 0.16
μ_{B_s} [MeV]	5368.13 ± 0.25	5368.19 ± 0.10	5368.51 ± 0.40
σ_{B_s} [MeV]	10.92 ± 0.20	11.387 ± 0.084	12.74 ± 0.34
α_{B_s}	1.50	1.50	1.50
ν_{B_s}	137.37 ± 0.80	153.22 ± 0.16	151.82 ± 0.50

5.7.2 Parameterising the reflections

All six decay modes may be present as a signal or as a reflection in each of the three spectra. However, due to the low misidentification rates observed, it is expected that many of the modes will be negligible.

Figure 5.18 is analogous to Figure 5.10, except that it is obtained after the full selection. It shows the (relative) expected yields and shapes of the six modes in each final state after the full selection (including the PID weighting, as above). The contributions from the reflections are highly suppressed. For this reason, in the fit to the data, only the dominant misidentified channel is considered per final state: $B^0 \rightarrow p\bar{p}K\pi$ in the $p\bar{p}KK$ and $p\bar{p}\pi\pi$ spectra, and $B^0 \rightarrow p\bar{p}\pi\pi$ in the $p\bar{p}K\pi$ spectrum. For the $p\bar{p}K\pi$ spectrum, the relevant misidentified decay after the full selection, $B^0 \rightarrow p\bar{p}\pi\pi$, differs from the decay mode $B_s^0 \rightarrow p\bar{p}KK$ already considered in the optimisation of the selection and presented in Table 5.5. This is because from Fig. 5.18, both $B^0 \rightarrow p\bar{p}\pi\pi$ and $B_s^0 \rightarrow p\bar{p}KK$ seem to have a similar yield in different mass regions in the $p\bar{p}K\pi$ spectrum, but after the final fit it is clear that $B^0 \rightarrow p\bar{p}\pi\pi$ is dominant. The systematic uncertainty associated with the simplification of considering a single misidentified decay mode per spectrum is evaluated in Section 5.8.6. The MC samples of the per-spectrum misidentified modes are processed in the same way as the signal modes, *cf.* Section 5.7.1. The largely asymmetric line shapes are again fitted with a DSCB function. The results of these fits can be seen in Figure 5.19; the parameters are given in Table 5.12. For some modes the fit is not of the highest quality but it is preferred not to add complexity choosing different functions for different modes as the yields are small. The effect of this choice is assessed in Section 5.8.6 along with the other systematic uncertainties on the fit model.

Table 5.12: Parameters of the fits to the MC misidentified decay modes considered per spectrum: $B^0 \rightarrow p\bar{p}K\pi$ in the $p\bar{p}KK$ and $p\bar{p}\pi\pi$ mass spectra, and $B^0 \rightarrow p\bar{p}\pi\pi$ in the $p\bar{p}K\pi$ mass spectrum.

Parameter	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
μ [MeV]	5338 ± 10	5312.4 ± 1.5	5231.2 ± 3.9
σ [MeV]	18 ± 11	17.1 ± 1.2	12.4 ± 3.0
α_1	2.00	0.80	0.40
ν_1	125.6 ± 2.1	1.70 ± 0.30	123.157 ± 0.036
α_2	1.00	2.50	2.00
ν_2	1.7 ± 4.2	154.01 ± 0.14	1.9 ± 3.3

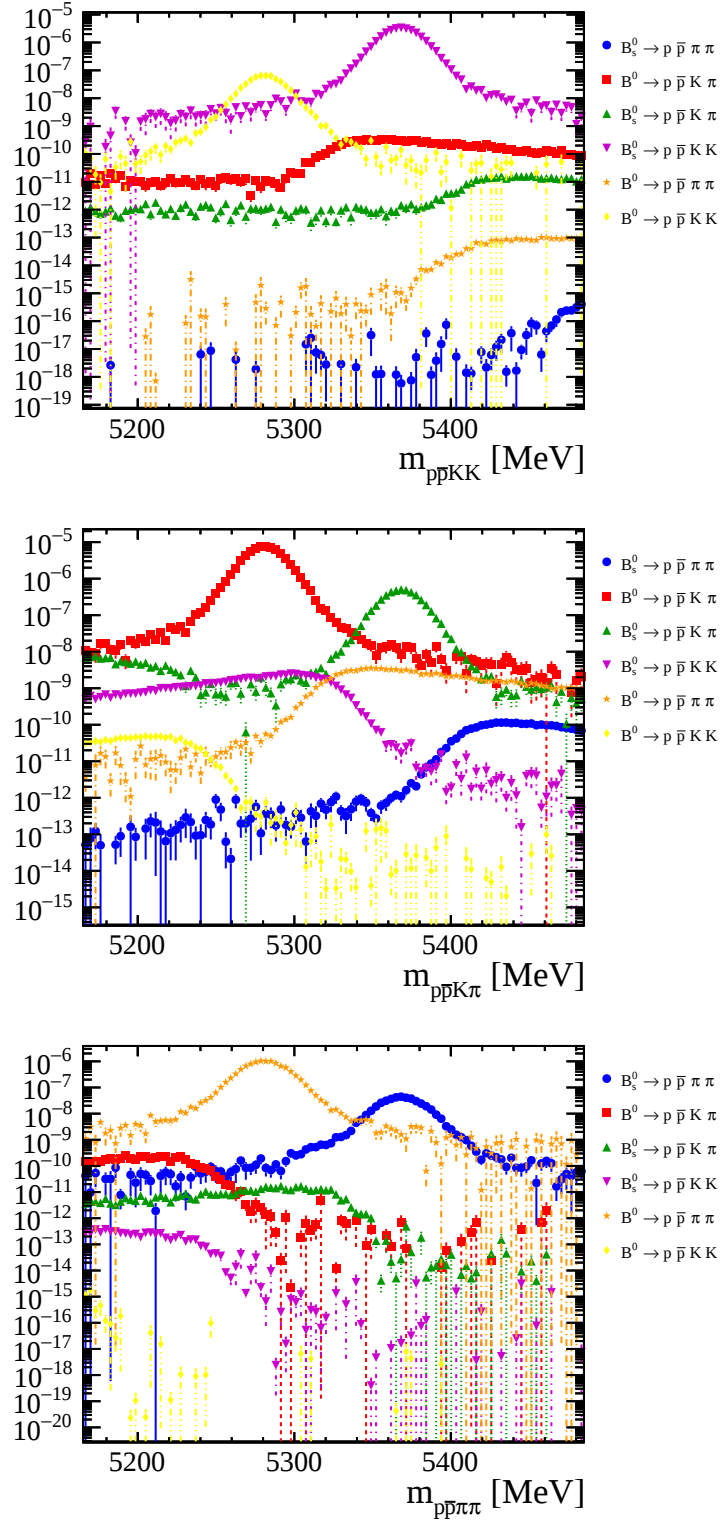


Figure 5.18: Invariant mass distributions for all $B_{(s)}^0 \rightarrow p\bar{p}hh'$ MC samples after the full selection taking into account the naive expected relative branching fractions as detailed in Table 5.4 and the different selection efficiencies. The $p\bar{p}KK$ final state is at the top, $p\bar{p}K\pi$ in the middle and $p\bar{p}\pi\pi$ at the bottom.

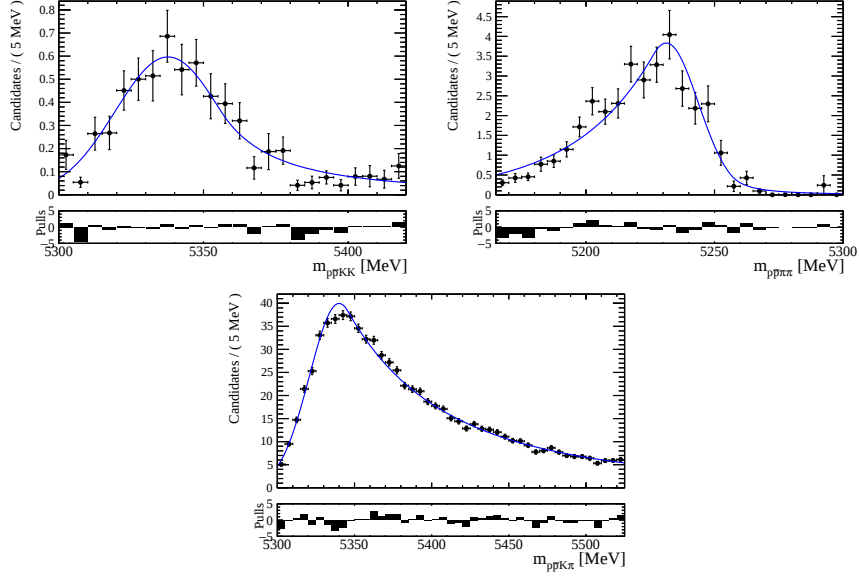


Figure 5.19: The reflections are fitted with a double-sided Crystal Ball function: (top left) $B^0 \rightarrow p\bar{p}K\pi$ in the $p\bar{p}KK$ mass spectrum and (top right) $p\bar{p}\pi\pi$ mass spectrum, and (bottom) $B^0 \rightarrow p\bar{p}\pi\pi$ in the $p\bar{p}K\pi$ mass spectrum.

5.7.3 Simultaneous fit to the data including charm contributions

As the same decay channel can be present in more than one spectrum, once as a real signal, and once as a reflection, a simultaneous fit allows to neatly constrain the yields of the reflections. All three $p\bar{p}hh'$ spectra are hereafter fitted simultaneously.

In each spectrum the two signal peaks of the B^0 and the B_s^0 are parameterised with a DSCB function. The B^0 mass is a floating parameter common to all three spectra and the difference between the B^0 and the B_s^0 masses is constrained with a Gaussian constraint to 87.33 ± 0.23 MeV [4]. Two core resolutions are allowed to float in the fit, namely σ_B describing the core resolution of the $B^0 \rightarrow p\bar{p}K\pi$ decay, and σ_{B_s} describing the core resolution of the $B_s^0 \rightarrow p\bar{p}KK$ decay. The core resolutions of the two other B^0 (B_s^0) modes are fixed in the fit and related to σ_B (σ_{B_s}) with a scaling factor given by the ratio of core resolutions obtained from the fits to the MC line shapes of Table 5.11. The parameters of the signal DSCB tails are fixed to the values from the MC fits.

Signal peaks

The combinatorial background is parameterised with an exponential, $n_{\text{bkg}} \cdot \exp(-c \cdot m_{p\bar{p}hh'})$, with parameters c and yield n_{bkg} different for the three spectra, while the shapes of the reflections are fixed to the MC shapes described in Section 5.7.2. The yields of the two

Background

signal PDFs (n_{B^0} and $n_{B_s^0}$) are free while the yield of the misidentified decay (n_{misID}) is fixed to the one of the same decay in the mass spectrum where it is not misidentified, the yield being scaled by the ratio of the efficiencies of that decay to be selected in the two signal samples. For example $B^0 \rightarrow p\bar{p}K\pi$ appears in the $p\bar{p}K\pi$ mass spectrum as a signal and in the $p\bar{p}KK$ mass spectrum as a cross-feed. Its yield in the $p\bar{p}KK$ mass spectrum is set to be

$$N_{B^0 \rightarrow p\bar{p}KK} = \frac{\epsilon_{B^0 \rightarrow p\bar{p}KK}^{p\bar{p}KK}}{\epsilon_{B^0 \rightarrow p\bar{p}K\pi}^{p\bar{p}K\pi}} \cdot N_{B^0 \rightarrow p\bar{p}K\pi} ,$$

and *mutatis mutandis* for the other two misidentified modes.

Figure 5.20 shows the simultaneous fit to the data including charm components while Table 5.13 gives the fit parameters. As expected, significant mass peaks are observed for the B^0 and B_s^0 decays to all three final states.

The fit model has been successfully validated against possible biases with pseudoexperiments, as detailed in Appendix E. No systematic uncertainty is taken from these studies given that the fit is seen to provide unbiased signal yields with uncertainties correct in magnitude.

5.7.4 Substructures

From the fit of the previous section it is possible to extract *sWeights* and inspect the $m_{p\bar{p}}$, $m_{hh'}$ and $m_{\bar{p}hh'}$ invariant mass spectra under the signal peaks. The *sPlot* procedure however requires that the yields of all the components of the fit are allowed to vary while all model parameters are fixed to the fitted values. The yields of the misidentified components of all spectra fixed in the nominal fit are hence allowed to vary freely for producing the *sWeights*. Table 5.14 collects the yields of the $B_{(s)}^0$ peaks from the fit without the constraints. It can be seen that they are perfectly compatible with the yields for the nominal fit of Table 5.13, as expected.

Figures 5.21 to 5.23 show the $m_{p\bar{p}}$, $m_{hh'}$ and $m_{\bar{p}hh'}$ invariant mass distributions and scatter plots under the main signal peaks for the $p\bar{p}KK$, $p\bar{p}K\pi$ and $p\bar{p}\pi\pi$ spectra. For all three final states the charmonium resonances are present in the $m_{p\bar{p}}$ mass spectrum and the D^0 is present in the $m_{hh'}$ mass spectrum. Moreover, the Λ_c^+ is present in the $m_{\bar{p}K\pi}$ mass spectrum of $B^0 \rightarrow p\bar{p}K\pi$. Other non-charm reso-

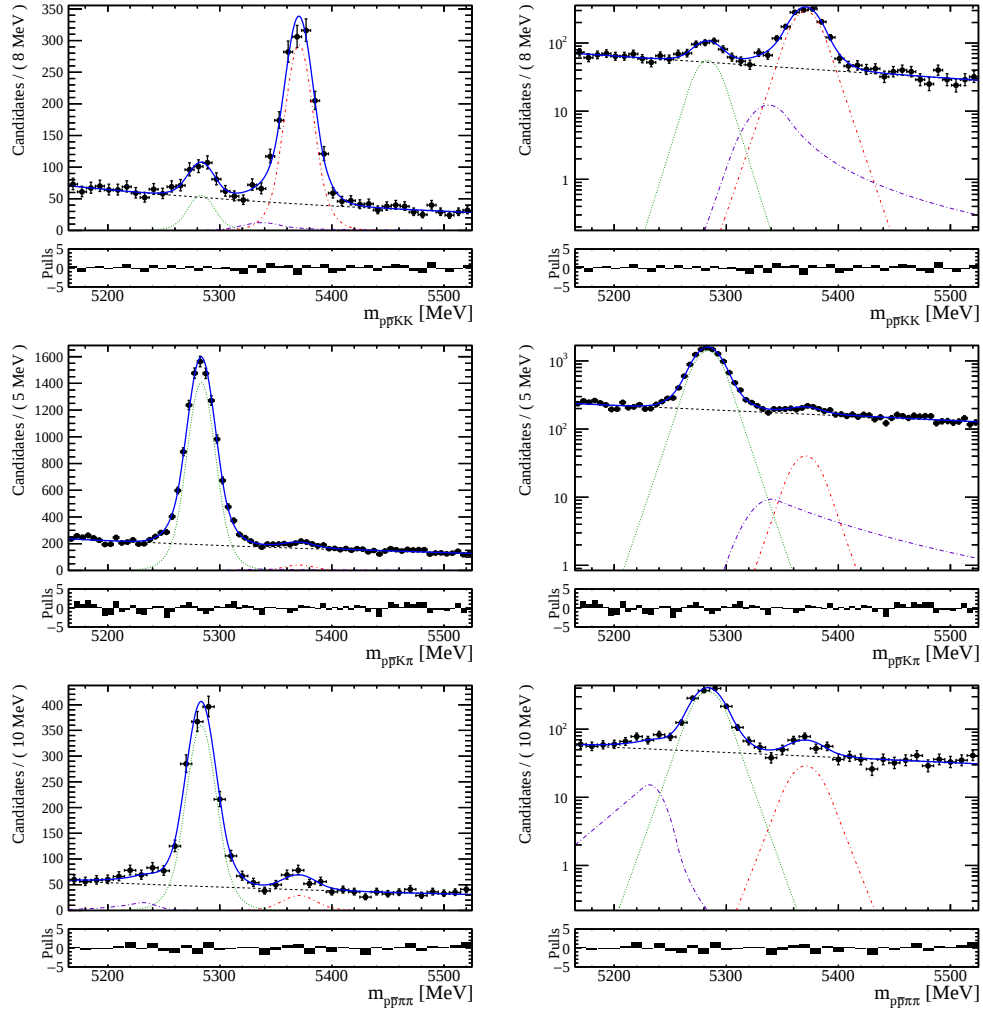


Figure 5.20: Simultaneous fit to the three mass spectra including all charm contributions: (*top*) $p\bar{p}KK$, (*middle*) $p\bar{p}K\pi$ and (*bottom*) $p\bar{p}\pi\pi$; a Logarithmic scale is used in the right column for the axis of ordinates. In each fit the following convention is used to identify the various model components: a green dotted line for the B^0 signal, a red dot-dashed line for the B_s^0 signal, a black dashed line for the combinatorial background, and a violet dot-dashed line for the relevant reflection. Refer to the text for further details on the fit model templates and parameters.

Table 5.13: Parameters of the simultaneous fit to the data including all charm contributions. Refer to the text for a description of the fit model.

Parameter	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
μ_B [MeV]		5283.40 ± 0.16	
$\mu_{B_s} - \mu_B$ [MeV]		87.33 ± 0.21	
σ_B [MeV]	12.18	13.05 ± 0.16	13.83
σ_{B_s} [MeV]	13.12 ± 0.47	13.68	15.30
c	-0.00255 ± 0.00023	-0.001737 ± 0.000091	-0.00171 ± 0.00027
n_B	220 ± 25	9560 ± 120	1293 ± 42
n_{B_s}	1254 ± 44	288 ± 51	115 ± 21
n_{bkg}	2080 ± 58	12670 ± 150	1539 ± 52
n_{misID}	89	182	70

Table 5.14: Yields of the $B_{(s)}^0$ peaks from the simultaneous fit to the data including all charm contributions when the yields of the misidentified components are all allowed to vary.

Parameter	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
n_B	220 ± 26	9560 ± 120	1315 ± 44
n_{B_s}	1277 ± 51	312 ± 64	127 ± 22

nances are clearly present as well, notably $\phi \rightarrow KK$, $K^{*0} \rightarrow K\pi$ and $\rho^0 \rightarrow \pi\pi$. The threshold enhancement is clearly visible in the $p\bar{p}$ mass spectra of all three decay modes considered.

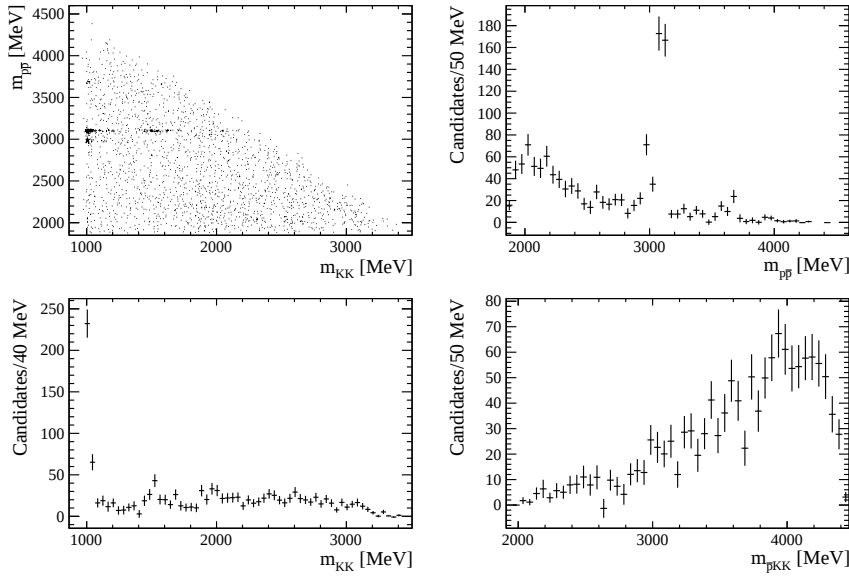


Figure 5.21: Substructures in the $sPlot'$ ed $B_s^0 \rightarrow p\bar{p}KK$ signal of the fit to the data including all charm contributions: scatter plot of $m_{p\bar{p}}$ versus m_{KK} at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, m_{KK} mass spectrum at the bottom left, and $m_{p\bar{p}KK}$ mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

It is possible to fit the substructures in the $sWeight'$ ed plots with simple Gaussians over an exponential background, see Figure 5.24. From the fits the core resolution, σ , of the D^0 peak is less than 10 MeV, while the σ of the Λ_c^+ peak is less than 6 MeV. The charm vetoes of Section 5.4.4 are thus adequate to remove the charm components.

5.7.5 Simultaneous fit to the data after applying the charm vetoes

The spectra for the charmless decay modes are obtained removing the charm components with the vetoes of Section 5.4.4. The same fit

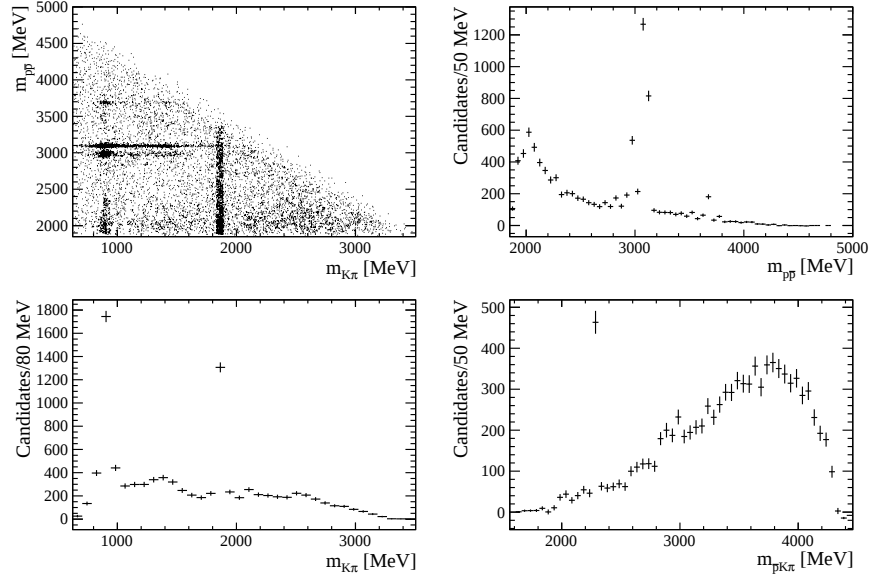


Figure 5.22: Substructures in the *sPlot'*ed $B^0 \rightarrow p\bar{p}K\pi$ signal of the fit to the data including all charm contributions: scatter plot of $m_{p\bar{p}}$ versus $m_{K\pi}$ at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, $m_{K\pi}$ mass spectrum at the bottom left, and $m_{p\bar{p}K\pi}$ mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

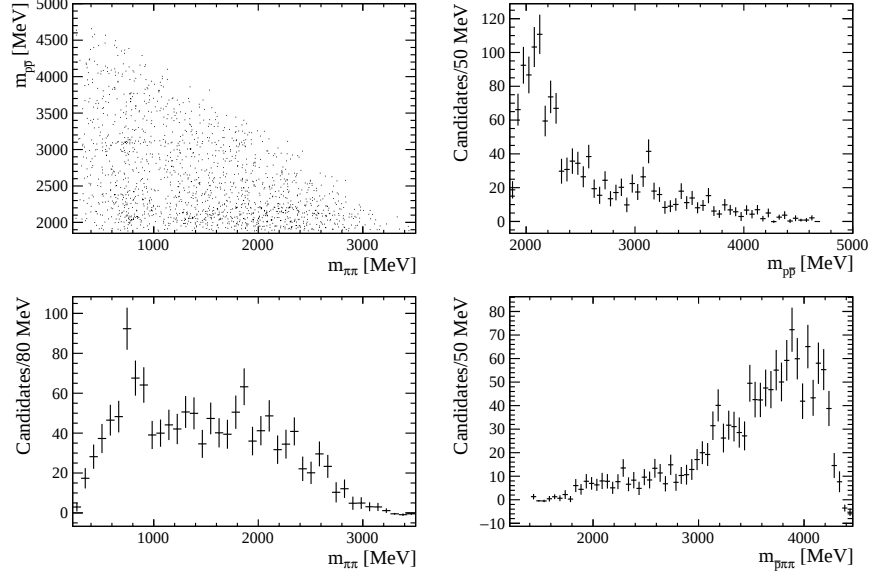


Figure 5.23: Substructures in the *sPlot'*ed $B^0 \rightarrow p\bar{p}\pi\pi$ signal of the fit to the data including all charm contributions: scatter plot of $m_{p\bar{p}}$ versus $m_{\pi\pi}$ at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, $m_{\pi\pi}$ mass spectrum at the bottom left, and $m_{p\bar{p}\pi\pi}$ mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

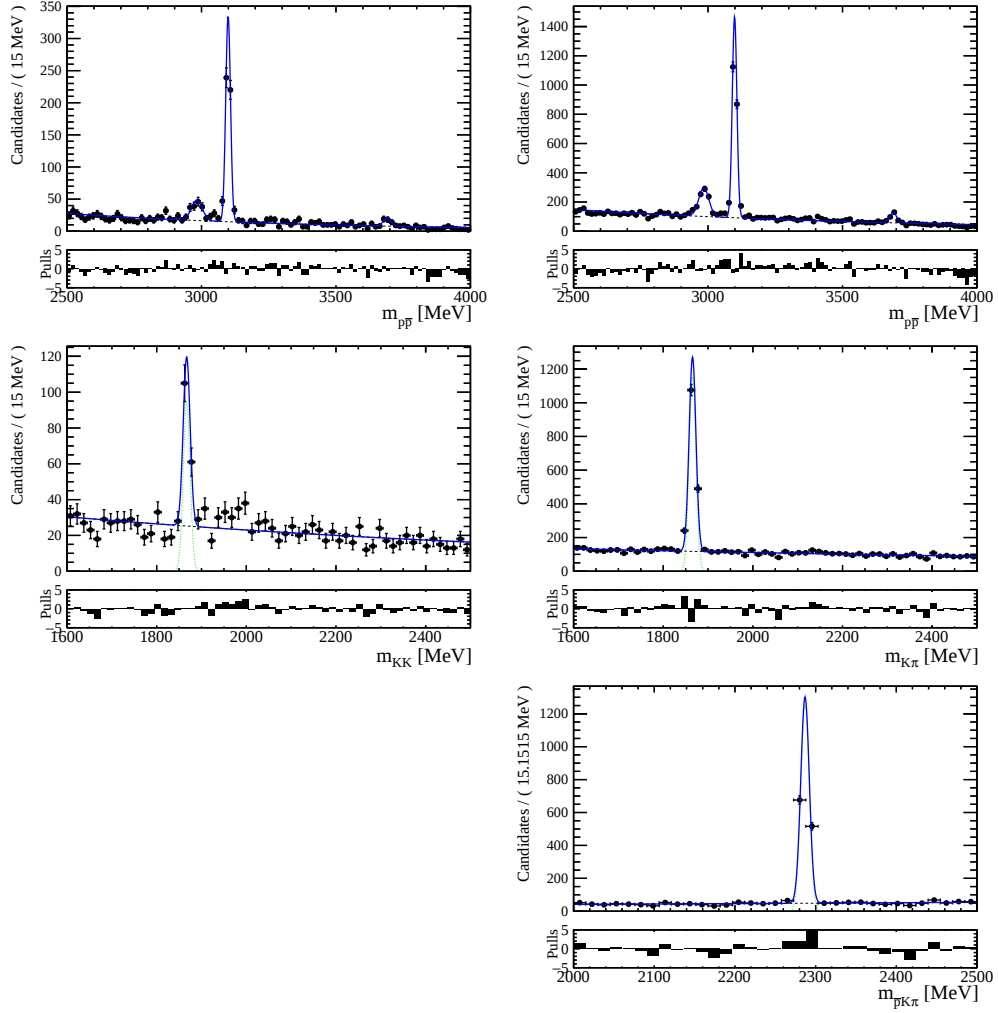


Figure 5.24: Fit to the charmonium and charm substructures in the *sWeight'*ed (left) $B_s^0 \rightarrow p\bar{p}KK$ and (right) $B^0 \rightarrow p\bar{p}K\pi$ spectra of the simultaneous fit to the data including all charm contributions. At the top $m_{p\bar{p}}$, $m_{hh'}$ at the centre, and $m_{p\bar{p}h'}$ at the bottom.

model strategy as that described in Section 5.7.3 is reused. Figure 5.25 shows the simultaneous fit while Table 5.15 gives the fit parameters.

Four out of the six $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decay modes, $B^0 \rightarrow p\bar{p}K\pi$, $B^0 \rightarrow p\bar{p}\pi\pi$, $B_s^0 \rightarrow p\bar{p}KK$ and $B_s^0 \rightarrow p\bar{p}K\pi$ are clearly observed. There is strong evidence for the $B^0 \rightarrow p\bar{p}KK$ decay mode and a hint for the $B_s^0 \rightarrow p\bar{p}\pi\pi$ mode. The final signal significances are evaluated with systematic effects in Section 5.9 and this is followed by the determination of the branching fractions.

It can be noticed that the yields of the B^0 and B_s^0 candidates scale differently once the charm vetoes are applied (i.e. comparing the yields from Table 5.13 and 5.15), for example in the $p\bar{p}KK$ mass spectrum the number of B_s^0 candidates reduced by approximately a factor 2 while the number of B^0 candidates by more than a factor 3. This is due to the different substructures in the two decays, for example in the $p\bar{p}KK$ final state the D^0 peak is much larger for B^0 than for B_s^0 .

5.7.6 Substructures after the charm vetoes

Analogously to what was done in Section 5.7.4, *sWeights* are extracted from the previous (nominal) fit leaving the yields of the misidentified modes free to float. Table 5.16 contains the yields of the signal peaks in this configuration of the fit, which, again, are perfectly compatible with the ones of the nominal fit in Table 5.15.

Figures 5.26 to 5.28 show the *sWeight'*ed $m_{p\bar{p}}$, $m_{hh'}$ and $m_{\bar{p}hh'}$ invariant mass spectra for the main signal peaks of each spectrum. All three $m_{p\bar{p}}$ distributions exhibit a clear threshold enhancement, as expected. In the case of the $B_s^0 \rightarrow p\bar{p}KK$ inclusive charmless spectrum, a prominent contribution from the so-far unobserved $B_s^0 \rightarrow p\bar{p}\phi$ decay (with $\phi \rightarrow KK$) is seen. As for the $B^0 \rightarrow p\bar{p}K\pi$ spectrum, a contribution from the well-known $B^0 \rightarrow p\bar{p}K^{*0}$ decay is clearly observed in the $m_{K\pi}$ spectrum. Finally, a $B^0 \rightarrow p\bar{p}\rho^0$ component (with $\rho^0 \rightarrow \pi\pi$), and possibly also $B^0 \rightarrow p\bar{p}f_0(980)$ (with $f_0(980) \rightarrow \pi\pi$), appears in the $m_{\pi\pi}$ distribution of the $B^0 \rightarrow p\bar{p}\pi\pi$ inclusive spectrum. Extracting branching fraction measurements for these resonant final states goes beyond the scope of this work.

Table 5.15: Parameters of the simultaneous fit to the data excluding all charm contributions. Refer to the text for a description of the fit model.

Parameter	$p\bar{p}K\bar{K}$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
μ_B [MeV]		5283.46 ± 0.24	
$\mu_{B_s} - \mu_B$ [MeV]		87.36 ± 0.22	
σ_B [MeV]	12.22	13.09 ± 0.25	13.87
σ_{B_s} [MeV]	12.88 ± 0.64	13.43	15.02
c	-0.00236 ± 0.00031	-0.00184 ± 0.00012	-0.00133 ± 0.00031
n_B	68 ± 17	4155 ± 83	902 ± 35
n_{B_s}	635 ± 32	246 ± 39	39 ± 16
n_{bkg}	1117 ± 42	7000 ± 110	1084 ± 43
n_{misID}	39	127	31

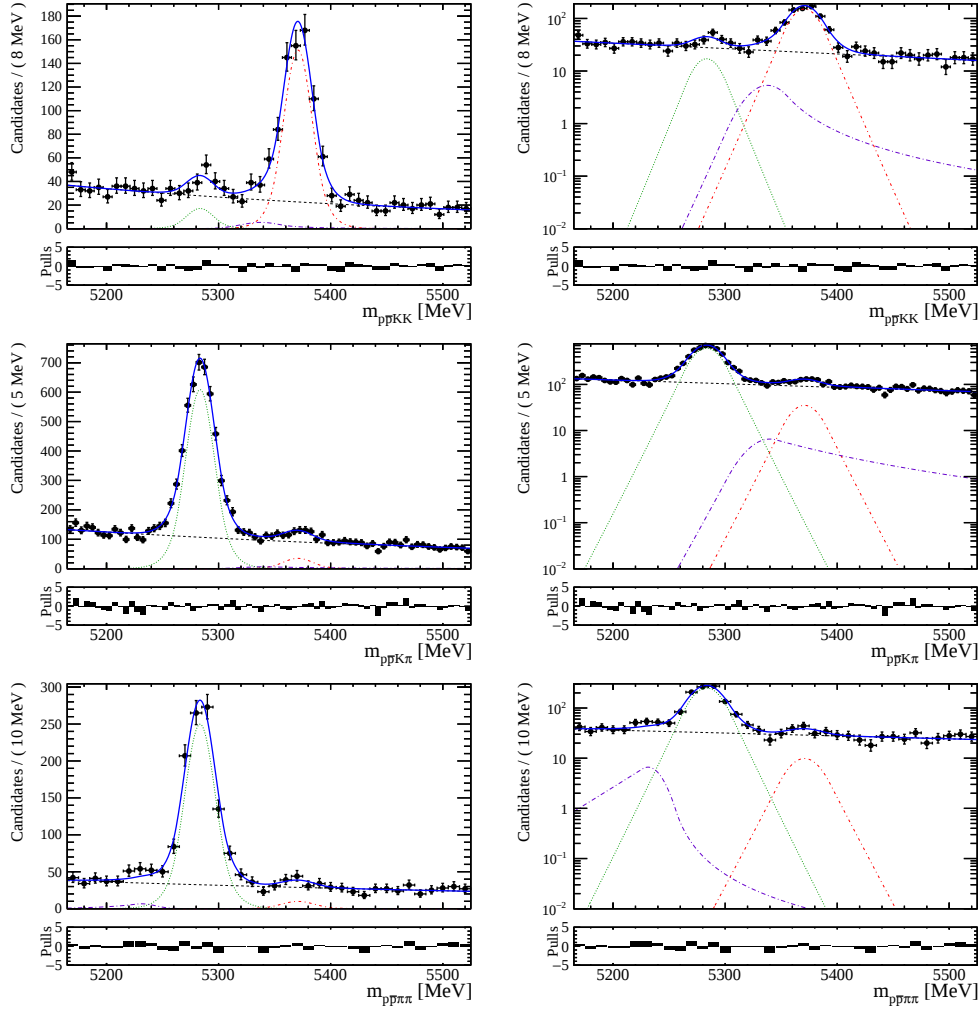


Figure 5.25: Simultaneous fit to the three mass spectra excluding all charm contributions: (top) $p\bar{p}KK$, (middle) $p\bar{p}K\pi$ and (bottom) $p\bar{p}\pi\pi$; a Logarithmic scale is used in the right column for the axis of ordinates. In each fit the following convention is used to identify the various model components: a green dotted line for the B^0 signal, a red dot-dashed line for the B_s^0 signal, a black dashed line for the combinatorial background, and a violet dot-dashed line for the relevant reflection. Refer to the text for further details on the fit model templates and parameters.

Table 5.16: Yields of the $B_{(s)}^0$ peaks from the simultaneous fit to the data excluding all charm contributions when the yields of the misidentified components are all allowed to vary.

Parameter	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
n_B	69 ± 17	4149 ± 83	925 ± 37
n_{B_s}	638 ± 35	264 ± 49	50 ± 17

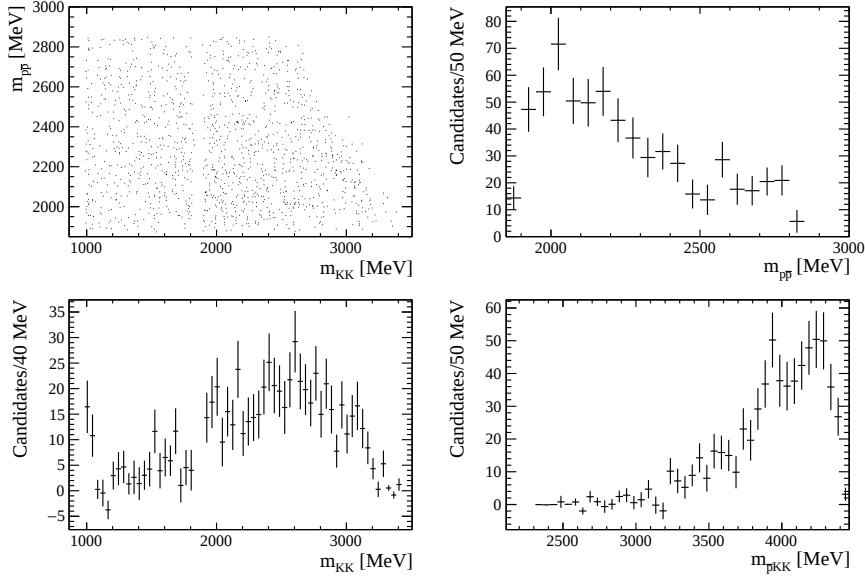


Figure 5.26: Substructures in the *sPlot*'ed $B_s^0 \rightarrow p\bar{p}KK$ signal after removing all charm contributions: scatter plot of $m_{p\bar{p}}$ versus m_{KK} at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, m_{KK} mass spectrum at the bottom left, and m_{pKK} mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

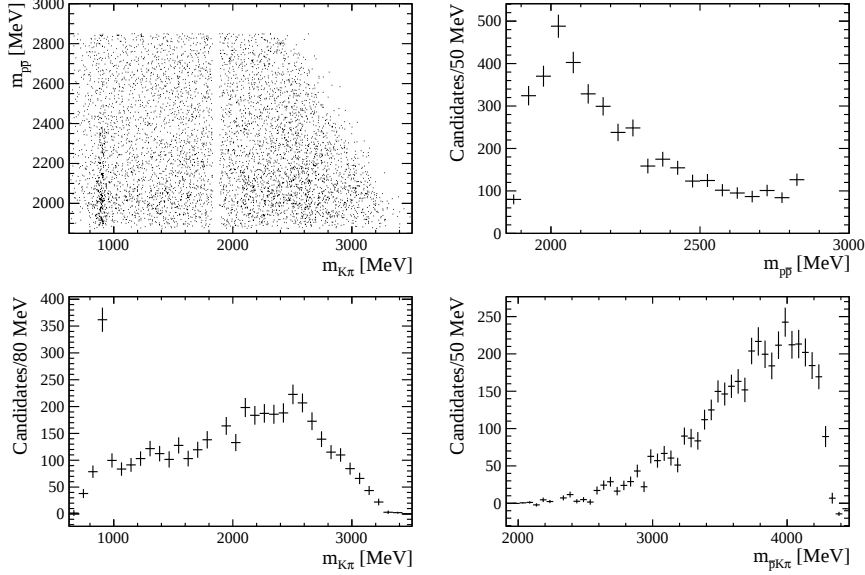


Figure 5.27: Substructures in the *sPlot*'ed $B^0 \rightarrow p\bar{p}K\pi$ signal after removing all charm contributions: scatter plot of $m_{p\bar{p}}$ versus $m_{K\pi}$ at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, $m_{K\pi}$ mass spectrum at the bottom left, and $m_{pK\pi}$ mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

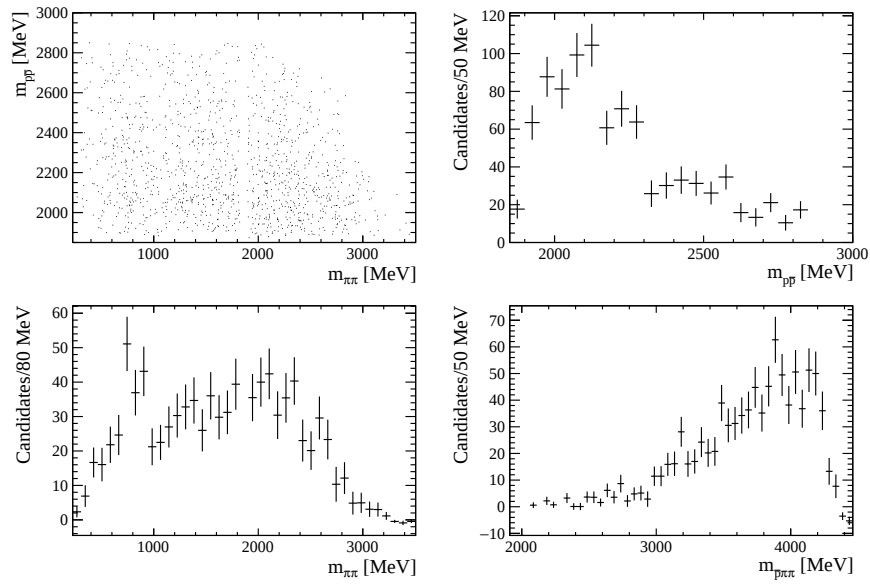


Figure 5.28: Substructures in the *sPlot'*ed $B^0 \rightarrow p\bar{p}\pi\pi$ signal after removing all charm contributions: scatter plot of $m_{p\bar{p}}$ versus $m_{\pi\pi}$ at the top left, $m_{p\bar{p}}$ mass spectrum at the top right, $m_{\pi\pi}$ mass spectrum at the bottom left, and $m_{p\pi\pi}$ mass spectrum at the bottom right. Refer to the text for a discussion of the visible structures.

5.7.7 Normalisation mode

To evaluate the yield (n_{sgn}) of the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ a 3-D fit in the $m_{p\bar{p}K\pi}$, $m_{p\bar{p}}$ and $m_{K\pi}$ invariant masses is performed on the data after the full selection except the charm vetoes, selecting the mass ranges of the J/ψ and the K^{*0} mesons. Also in this case some parameters of the signal line shape are derived from the MC. The B^0 and the J/ψ are parameterised with symmetric double-sided Crystal Ball functions (with parameters μ and σ , α and ν) while the K^{*0} is parameterised with a relativistic Breit-Wigner function (with parameters μ and Γ). Figure 5.29 shows the 3-D fit to the MC while Table 5.17 collects the values of all fit model parameters.

Table 5.17: Parameters of the 3-D fit to the MC shape of the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$. Refer to the text for further details on the fit model templates and parameters.

Parameter	Value
μ_B [MeV]	5281.12 ± 0.14
σ_B [MeV]	11.22 ± 0.13
α_B	1.50
ν_B	11.6 ± 1.7
$\mu_{J/\psi}$ [MeV]	3097.745 ± 0.094
$\sigma_{J/\psi}$ [MeV]	7.540 ± 0.089
$\alpha_{J/\psi}$	1.50
$\nu_{J/\psi}$	18.8 ± 5.3
μ_{K^*} [MeV]	899.64 ± 0.38
Γ_{K^*} [MeV]	49.06 ± 0.85

In the fit to the data the tail parameters α and ν of the double-sided Crystal Ball functions are fixed to the values obtained from the MC while all the other parameters are free to float. In each mass spectrum the combinatorial background is parameterised with an exponential function (with parameter c and yield n_{bkg}). The $K\pi$ S-wave component is taken to have exactly the same shape as the signal in the $p\bar{p}K\pi$ and $p\bar{p}$ mass spectra while in the $K\pi$ mass spectrum it is parameterised by a modified Breit-Wigner distribution that takes into account both the $K_0^*(1430)$ contribution and the non-resonant component and was first introduced by the LASS experiment [113, 114]. All the parameters are fixed to the ones obtained in the LHCb analysis [115] and only the yield ($n_{\text{S-wave}}$) is left free. This component is not present in the MC sample used in the previous fit.

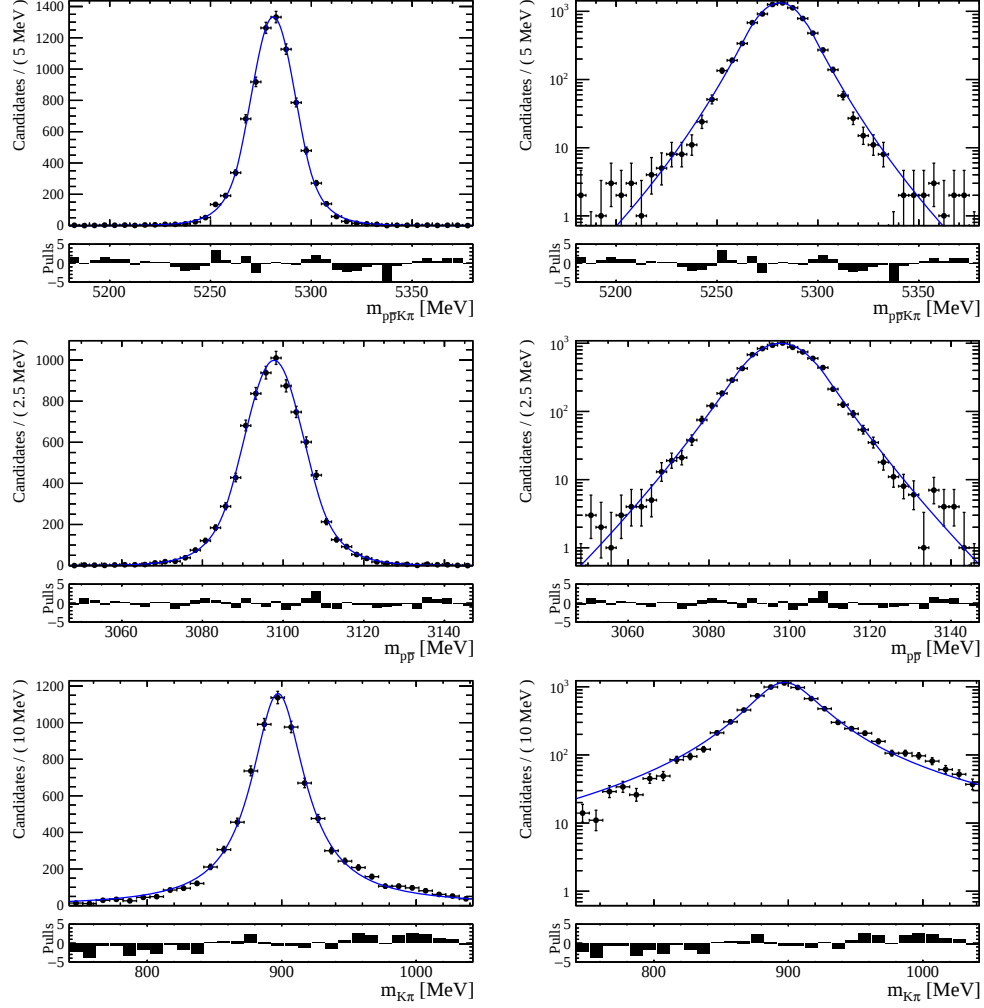


Figure 5.29: 3-D fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ MC to extract the line shapes. At the top the $m_{p\bar{p}K\pi}$ mass spectrum, at the middle the $m_{p\bar{p}}$ mass spectrum, and at the bottom $m_{K\pi}$ mass spectrum. A Logarithmic scale is used in the right column for the axis of ordinates.

Figure 5.30 shows the 3-D fit to the data while Table 5.18 collects the values of all fit model parameters.

Table 5.18: Parameters of the 3-D fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ normalisation mode data. Refer to the text for further details on the fit model templates and parameters.

Parameter	Value
μ_B [MeV]	5283.15 ± 0.39
σ_B [MeV]	12.29 ± 0.33
c_B	-0.0000 ± 0.0014
$\mu_{J/\psi}$ [MeV]	3098.51 ± 0.25
$\sigma_{J/\psi}$ [MeV]	8.00 ± 0.21
$c_{J/\psi}$	-0.0012 ± 0.0030
μ_{K^*} [MeV]	898.53 ± 1.00
Γ_{K^*} [MeV]	46.3 ± 2.6
c_{K^*}	0.00187 ± 0.00067
n_{sgn}	1216 ± 45
n_{bkg}	164 ± 16
$n_{\text{s-wave}}$	137 ± 31

5.8 SYSTEMATIC UNCERTAINTIES

This section discusses the various sources of systematic uncertainties. Three main categories of systematic uncertainties are present. External contributions like the branching fraction of the normalisation mode and the ratio of production cross sections between B and B_s^0 mesons is the first category of systematic uncertainty. Systematic uncertainties associated to the determination of the efficiencies and of the yields from the fits are the other two categories.

5.8.1 External contributions to the systematic uncertainties

The visible branching fraction of $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$, $\mathcal{B}_{\text{vis}}(B^0 \rightarrow J/\psi K^{*0}) = (1.68 \pm 0.12) \times 10^{-6}$, entering the expression used to calculate the absolute branching fractions, is given by the product of $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) = (1.19 \pm 0.08) \times 10^{-3}$ [104], $\mathcal{B}(J/\psi \rightarrow p\bar{p}) = (2.120 \pm 0.029) \times 10^{-3}$ [4] and $\mathcal{B}(K^{*0} \rightarrow K^+\pi^-) = 2/3 \cdot (99.754 \pm$

*Branching fraction
normalisation mode*

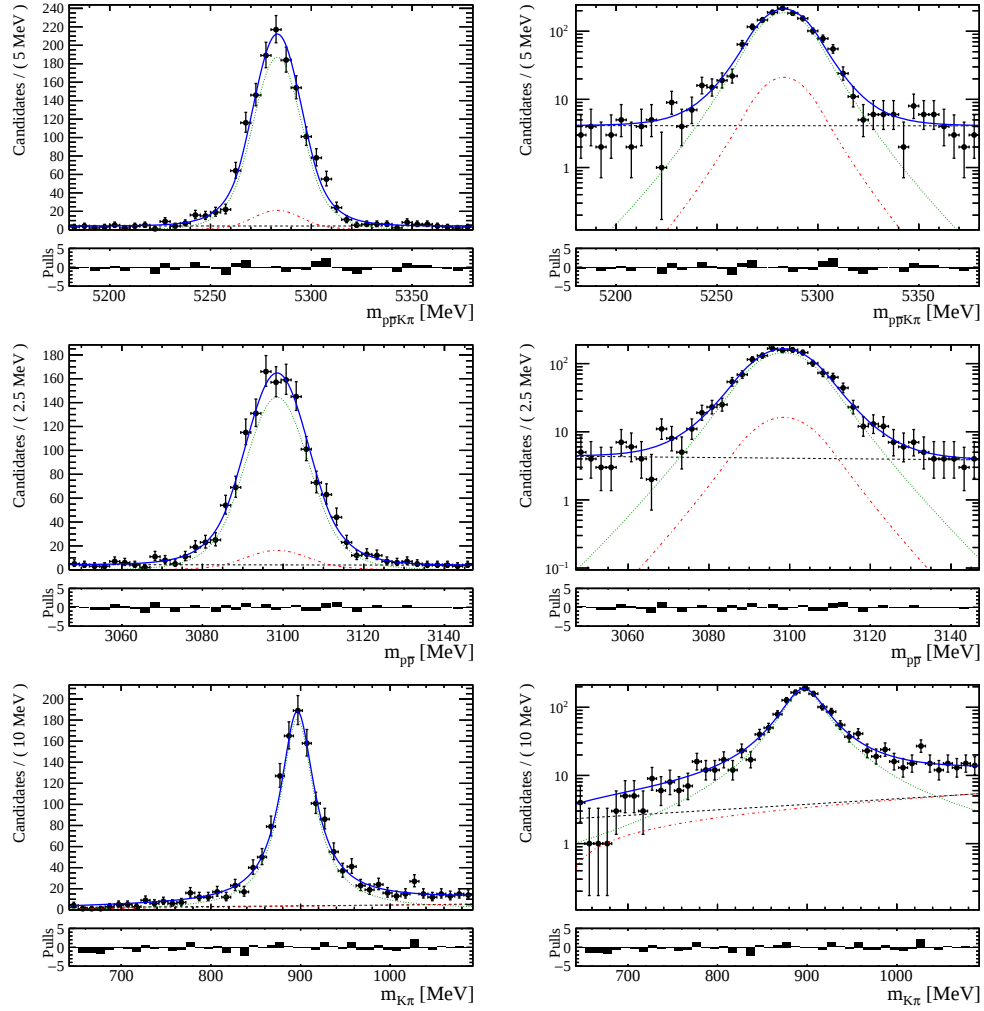


Figure 5.30: 3-D fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ data. At the top the $m_{p\bar{p}K\pi}$ mass spectrum, at the middle the $m_{p\bar{p}}$ mass spectrum, and at the bottom $m_{K\pi}$ mass spectrum. A Logarithmic scale is used in the right column for the axis of ordinates. In each fit the following convention is used to identify the various model components: a green dotted line for the $B^0 \rightarrow J/\psi K^{*0}$ signal, a red dot-dashed line for the $K\pi$ S-wave and a black dashed line for the combinatorial background. Refer to the text for further details on the fit model templates and parameters.

0.021)% [4]⁵. Its uncertainty of 6.9% is taken as a systematic uncertainty where relevant.

The factor $\frac{f_d}{f_s}$ enters in the expression of the ratios of branching fractions of B^0 relative to B_s^0 decays. Its uncertainty of 5.8% [103] is taken as a systematic uncertainty where appropriate.

*Fragmentation
probability*

5.8.2 Trigger

As stated in Section 5.4.1 the events are selected in two independent ways by the hardware trigger:

- *TOS*: The signal candidate leaves a deposit in the hadronic calorimeter with transverse energy higher than 3.5 GeV triggering the L0 hadron trigger line (trigger on signal, TOS).
- *TIS*: Any particle in the event, irrespective if it is part of the signal candidate or not, passes any of the L0 trigger lines (trigger independent of signal, TIS).

Modelling accurately the behaviour of the hadronic calorimeter, especially in the presence of multibody final states, is challenging and a mismodelling of the L0 hadron trigger line can occur. To estimate the systematic effect of this mismodelling a data-driven technique is used. As the TIS selection is independent from the signal, to first order the efficiency of the L0 hadronic trigger line can be computed as

$$\epsilon_{TOS} = \frac{n_{TIS\&TOS}}{n_{TIS}},$$

where n_{TIS} is the number of candidates selected by the TIS trigger line and $n_{TIS\&TOS}$ is the number of events selected by both. The efficiencies of the L0 hadronic trigger line can be evaluated on the candidates passing the full selection in two ways:

1. naively from the MC as n_{trig}/n_{strip} asking that the simulated candidate are selected by the hadronic trigger line;
2. from the *sWeight'*ed signal data using the data-driven method above.

Table 5.19 collects the ratios of the efficiencies of the L0 Hadron trigger computed using the data-driven method on *sWeight'*ed data and

⁵ In [4] only $\mathcal{B}(K^{*0} \rightarrow (K\pi)^0)$ is available. From isospin considerations $\mathcal{B}(K^{*0} \rightarrow K^+\pi^-)$ is 2/3 of that value.

from the MC after applying the full selection. The efficiencies evaluated with the data-driven method are approximately 20% lower than the ones computed using the MC. For the final branching fraction measurements, however, an overall bias largely cancels in the ratio of efficiencies. What gives rise to a systematic uncertainty is instead the difference between different decay modes. Even if all correction factors are statistically compatible, a systematic uncertainty of 3.7% is assigned. This value is half of the average difference between correction factors for all possible combinations of two signal or normalisation modes (21 combinations). Half the difference is considered instead of the difference under the assumption that the ratio of efficiencies of the TIS trigger selection is well modelled by the MC so the correction would be applied only to the candidates selected by hadronic trigger line, so approximately to half. No systematic uncertainties are

Table 5.19: Ratios of the efficiencies of the L0 hadron trigger computed using a data-driven method and naively from MC samples.

Decay channel	Efficiency ratio
$B^0 \rightarrow p\bar{p}KK$	0.852 ± 0.099
$B_s^0 \rightarrow p\bar{p}KK$	0.753 ± 0.039
$B^0 \rightarrow p\bar{p}K\pi$	0.788 ± 0.016
$B_s^0 \rightarrow p\bar{p}K\pi$	0.874 ± 0.074
$B^0 \rightarrow p\bar{p}\pi\pi$	0.783 ± 0.040
$B_s^0 \rightarrow p\bar{p}\pi\pi$	0.86 ± 0.12
$B^0 \rightarrow J/\psi K^*$	0.708 ± 0.048

considered for the HLT efficiencies given that the dominant source of systematic uncertainty is the reconstruction, which is the subject of the following section.

5.8.3 Tracking

Computing the reconstruction efficiencies directly on the MC samples can lead to a bias as the tracking efficiencies present some discrepancies between data and MC. These discrepancies have been evaluated in bins of momentum and pseudorapidity using the decay $J/\psi \rightarrow \mu\mu$ with a data-driven method [77]. Table 5.20 collects the multiplicative corrections to the reconstruction and stripping efficiencies. These correction factors are obtained under the assumption that the per-track corrections tabulated from the $J/\psi \rightarrow \mu\mu$ decay are applicable to the four final state tracks of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decay. The corrections are of the order of 2%. To determine the final results only ratios of ef-

ficiencies are relevant, and mismodelling effects cancel to first order. What remains are the differences between the correction factors of the different decay modes:

$$\frac{r_1}{r_2} = \frac{1 + \delta_1}{1 + \delta_2} \simeq (1 + \delta_1) \cdot (1 - \delta_2) \simeq 1 + (\delta_1 - \delta_2) = 1 + (r_1 - r_2)$$

A conservative estimate of the systematic uncertainty is the maximum difference between corrections factors of the different decay modes, hence 0.7%.

Table 5.20: Multiplicative factors to correct the reconstruction and stripping efficiencies to take into account the mismodelling of the tracking efficiencies in the MC. Values obtained from the data-driven method using $J/\psi \rightarrow \mu\mu$ decays. Refer to the text for details.

Decay channel	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}KK$	1.015	1.015	1.014
$B_s^0 \rightarrow p\bar{p}KK$	1.016	1.015	1.014
$B^0 \rightarrow p\bar{p}K\pi$	1.016	1.017	1.017
$B_s^0 \rightarrow p\bar{p}K\pi$	1.016	1.017	1.017
$B^0 \rightarrow p\bar{p}\pi\pi$	1.016	1.018	1.019
$B_s^0 \rightarrow p\bar{p}\pi\pi$	1.016	1.018	1.019
$B^0 \rightarrow J/\psi K^*$	1.018	1.021	1.021

The correction factors obtained from the muons of the $J/\psi \rightarrow \mu\mu$ decay however do not take into account the hadronic interactions. Following the work presented in the analysis of the $B^+ \rightarrow p\bar{\Lambda}$ decay [116] it is assumed that the corrections could vary by 1.5% for kaons and pions and 4% for protons with respect to the values obtained from muons. These estimates include both the uncertainty due to the hadronic interactions and the systematic uncertainties associated to the method used to estimate the correction factors from $J/\psi \rightarrow \mu\mu$ decays. To estimate the systematic uncertainty on the measurements the correction factors for $B_{(s)}^0 \rightarrow p\bar{p}hh'$ are recomputed using again the tables with the correction factors in bins of momentum and pseudorapidity, but increasing all correction factors accounting for the data-MC disagreement by the total systematic uncertainty. For example if in a certain bin the correction factor is 1.02 then the value used for proton tracks is 1.06 and for pions and kaons is 1.035. If the correction factor is less than one the systematic uncertainty is subtracted. Table 5.21 collects the multiplicative corrections obtained with these “worst case scenario” correction tables. Now the corrections are of the order of 6% and the maximum difference between correction factors is 1.1%. A systematic uncertainty of 1.1% is then

Effect of hadronic interactions

Table 5.21: Multiplicative factors to correct the reconstruction and stripping efficiencies to take into account the mismodelling of the tracking efficiencies in the MC. Values obtained from the data-driven method using $J/\psi \rightarrow \mu\mu$ decays and increasing the difference from unity to take into account the systematic uncertainties of the method and the hadronic interactions. Refer to the text for details.

Decay channel	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}KK$	1.057	1.058	1.059
$B_s^0 \rightarrow p\bar{p}KK$	1.056	1.057	1.057
$B^0 \rightarrow p\bar{p}K\pi$	1.058	1.059	1.06
$B_s^0 \rightarrow p\bar{p}K\pi$	1.057	1.059	1.06
$B^0 \rightarrow p\bar{p}\pi\pi$	1.058	1.06	1.061
$B_s^0 \rightarrow p\bar{p}\pi\pi$	1.056	1.058	1.06
$B^0 \rightarrow J/\psi K^*$	1.06	1.065	1.068

assigned for the data-MC disagreement in the ratio of tracking efficiencies. The reason that even with a large absolute disagreement, the relative disagreement, and hence the systematic uncertainty, is relative small, is that the kinematics among the different decay modes and the normalisation mode are similar.

5.8.4 Selection

The selection chain is identical for all $B_{(s)}^0 \rightarrow p\bar{p}hh'$ modes modulo the differences in PID; likewise for the selection of the normalisation channel, which differs by the presence of reversed charm vetoes. Systematic uncertainties arising from the selection hence reduce to PID systematics, assuming that any mismodelling cancels to first order when considering ratios of branching fractions. The effects of the charm vetoes and of the procedure to take into account the distribution of signal events across the phase space are also taken into account.

PID selection

Different sources of systematic uncertainty are evaluated for the data-driven evaluation of the efficiency of the selection based on PID information.

1. The statistical uncertainty coming from the finite size of the calibration samples propagated to the efficiencies. Simple error propagation is used to evaluate this.
2. The choice of the binning scheme used to divide the calibration and MC samples in the evaluation of the efficiency of the

PID selection as detailed in Section 5.4.3 is arbitrary. It is therefore necessary to evaluate the change of the evaluated efficiency when varying the binning scheme. The efficiency of the PID selection is evaluated with a finer binning scheme (double the number of bins in p , η and number of tracks) and the difference with respect to the baseline one is taken as the systematic uncertainty.

3. The systematic uncertainty coming from the use of the *sPlot* technique to separate the signal from the background in the calibration samples. A relative uncertainty of 0.1% is assigned following [117].

Table 5.22 summarises the relative systematic uncertainty (not including the statistical one) for the decay modes considered in the various final states. In Table 5.23 more details are given about the uncertainties coming from the finite size of the calibration samples and from the choice of the binning scheme.

Table 5.22: Total relative systematic uncertainty on the efficiency of the selection on PID variables [%]. Refer to the text for details.

Decay channel	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}KK$	1.32	0.88	0.67
$B_s^0 \rightarrow p\bar{p}KK$	1.46	0.27	0.64
$B^0 \rightarrow p\bar{p}K\pi$	0.55	1.11	0.44
$B_s^0 \rightarrow p\bar{p}K\pi$	0.53	1.11	0.93
$B^0 \rightarrow p\bar{p}\pi\pi$	1.75	0.25	0.45
$B_s^0 \rightarrow p\bar{p}\pi\pi$	3.58	0.82	0.50
$B^0 \rightarrow J/\psi K^*$	0.56	1.34	1.09

Two sources of systematic uncertainties have been considered associated with the correction of the efficiencies to account for the event distribution across the phase space:

*Phase-space
correction of
efficiencies*

1. The effect of the bins with no selected events but some generated ones that are not handled correctly by Equation (5.5), evaluated to be 0.1% comparing the same efficiencies computed with Equation (5.5) and simply as $n_{\text{sel}}/n_{\text{gen}}$.
2. The effect of the choice of the binning used to divide the phase space. To assess this the efficiencies have been computed using several binning schemes to divide the phase space:
 - the nominal one with four bins in $m_{p\bar{p}}$ and $m_{hh'}$ and three bins in $\cos \vartheta_{p\bar{p}}$, $\cos \vartheta_{hh'}$ and χ ,

Table 5.23: Efficiency of the selection on PID variables [%]. The first uncertainty is statistical, the second comes from the finite size of the calibration samples and the third is the systematic uncertainty from the choice of the binning scheme. The relative systematic uncertainty of 0.1% coming from the use of the *sPlot* technique is not included in this table. Refer to the text for details.

Decay channel	$p\bar{p}KK$	$p\bar{p}K\pi$	$p\bar{p}\pi\pi$
$B^0 \rightarrow p\bar{p}KK$	$38.79 \pm 0.29 \pm 0.036 \pm 0.51$	$1.330 \pm 0.050 \pm 0.0019 \pm 0.012$	$0.055 \pm 0.024 \pm 0.00023 \pm 0.00028$
$B_s^0 \rightarrow p\bar{p}KK$	$38.55 \pm 0.20 \pm 0.037 \pm 0.56$	$1.293 \pm 0.033 \pm 0.0019 \pm 0.0026$	$0.037 \pm 0.010 \pm 0.00019 \pm 0.00014$
$B^0 \rightarrow p\bar{p}K\pi$	$1.098 \pm 0.066 \pm 0.0011 \pm 0.0058$	$46.29 \pm 0.20 \pm 0.059 \pm 0.51$	$1.002 \pm 0.068 \pm 0.0040 \pm 0.0016$
$B_s^0 \rightarrow p\bar{p}K\pi$	$0.902 \pm 0.090 \pm 0.00091 \pm 0.0046$	$46.01 \pm 0.27 \pm 0.063 \pm 0.50$	$0.955 \pm 0.088 \pm 0.0041 \pm 0.0078$
$B^0 \rightarrow p\bar{p}\pi\pi$	$0.021 \pm 0.012 \pm 0.000020 \pm 0.00037$	$2.031 \pm 0.059 \pm 0.0027 \pm 0.0037$	$32.19 \pm 0.29 \pm 0.13 \pm 0.047$
$B_s^0 \rightarrow p\bar{p}\pi\pi$	$0.017 \pm 0.016 \pm 0.000018 \pm 0.00060$	$1.805 \pm 0.060 \pm 0.0024 \pm 0.014$	$32.22 \pm 0.29 \pm 0.14 \pm 0.073$
$B^0 \rightarrow J/\psi K^*$	$1.15 \pm 0.12 \pm 0.0013 \pm 0.0062$	$45.43 \pm 0.34 \pm 0.072 \pm 0.61$	$0.82 \pm 0.10 \pm 0.0039 \pm 0.0080$

- one with one bin more for each variable (hence 5,5,4,4,4),
- five, each one with double the number of bins in one variable (hence (8,4,3,3,3), (4,8,3,3,3), ...).

The differences between efficiencies computed with the nominal binning and the alternative ones are not statistically significant so no systematic uncertainty is assigned.

Two sources of systematic uncertainties are evaluated for the data-driven evaluation of the efficiencies of the charm vetoes:

Charm vetoes

1. The choice of the function used to fit the sidebands of the vetoed region is arbitrary. An exponential function is used as a baseline and a linear function as alternative. The number of events in the vetoed regions is obtained with both choices and the difference in the resulting efficiencies is taken as systematic uncertainty.
2. The parameter of the baseline exponential function has an uncertainty coming from the fit. The efficiencies corresponding to this parameter varied by $\pm 1\sigma$ are also obtained and half their difference taken as systematic uncertainty.

The sum in quadrature of the uncertainties coming from these two sources is reported in Table 5.10 as the systematic uncertainty associated to the efficiency of the charm vetoes.

5.8.5 B_s^0 lifetime

The two B_s^0 mass eigenstates have significantly different lifetimes ($\tau_L = 1.422 \pm 0.008$ ps and $\tau_H = 1.610 \pm 0.012$ ps) while in the simulation used to estimate the efficiencies a single average value $\tau = 1.51$ ps is used. This can lead to a bias in the evaluation of the efficiencies. To estimate the systematic uncertainty associated to this it is possible to reweight the MC samples so that the lifetime distribution of the B_s^0 has a different mean value. More precisely the B_s^0 lifetime distribution of the generated events is:

$$p(t; \tau) = \frac{e^{-t/\tau}}{\tau \cdot e^{-t_{\min}/\tau}},$$

where $t_{\min} = 0.1$ ps is the lower bound of the distribution, coming from the selection at the generator level. The weight to apply to the simulated events to change the mean lifetime from τ_{old} to τ_{new} is

Table 5.24: Systematic uncertainties on the efficiencies due to the different B_s^0 lifetime between mass eigenstates. Refer to the text for more details.

Decay channel	Syst. excluding ϵ_{GEN}	Syst. including ϵ_{GEN}
$B_s^0 \rightarrow p\bar{p}KK$	2.55%	2.58%
$B_s^0 \rightarrow p\bar{p}K\pi$	2.48%	2.51%
$B_s^0 \rightarrow p\bar{p}\pi\pi$	2.66%	2.69%

thus $p(t; \tau_{\text{new}})/p(t; \tau_{\text{old}})$. It is then possible to evaluate the selection efficiency in the two extreme scenarios where the mean lifetime is either τ_L or τ_H and take as systematic uncertainty the maximum difference with respect to the nominal efficiency. In the previous step the systematic uncertainty of the generator level selection is missing. It is possible to evaluate it analytically by assuming that the selection most affected by the possible bias is the one on the B_s^0 lifetime and neglect the bias on the other generator level cuts. The relative uncertainty is then

$$\frac{|e^{-t/\tau_{\text{new}}} - e^{-t/\tau_{\text{old}}}|}{e^{-t/\tau_{\text{old}}}},$$

which leads to a systematic uncertainty of 0.4%. Table 5.24 collects the systematic uncertainty on the selection efficiencies, both including and excluding the generator level selection.

5.8.6 Fit model

The systematic uncertainties coming from the mass fits to the various spectra are the last source of systematic uncertainties. Two sources of systematic uncertainties are considered: the freedom in choosing the fit model and the uncertainty on the misID efficiencies used to fix the yield of the misID components.

*Arbitrariness in
choosing the fit
model*

Some degree of arbitrariness is present in choosing the fit model, in particular:

1. using a double-sided Crystal Ball function as signal shape,
2. using an exponential function for the background,
3. neglecting some components in the fit.

To estimate the relative systematic uncertainty for each of the previous choices the following procedure is used:

- the dataset is fitted with an alternative model in which a different choice is made:
 1. In the first alternative model the signal and misidentified decay shapes are not double-sided Crystal Ball functions but kernel estimation functions obtained from the MC.
 2. In the second alternative model the background is modelled with a linear function instead of an exponential.
 3. In the third alternative model more components are added to the fit:
 - A component for $B^0 \rightarrow p\bar{p}hh'\pi^0$ is added in each mass spectrum and components for $B^0 \rightarrow p\bar{p}\eta$ and $B_s^0 \rightarrow p\bar{p}\eta$ with $\eta \rightarrow \pi\pi\gamma$ are added in the $p\bar{p}\pi\pi$ mass spectrum. Their shapes are obtained using kernel estimation functions on simplified simulations similar to the ones described in Section 5.5 and their normalisations are left free in the fit.
 - The second leading misidentified component is introduced for each mass spectrum. Its shape is derived from MC and its normalisation is fixed similarly to that of the leading misidentified decay mode.
 - Components are introduced for $\Lambda_b^0 \rightarrow p\bar{p}pK$ in the $p\bar{p}KK$ and $p\bar{p}K\pi$ mass spectra. Their shapes are obtained using kernel estimation functions on the simplified simulation of Section 5.5 and their normalisations are left free.
- Each alternative model, with the values obtained from the fit is used to generate 1000 toy datasets.
- These toy datasets are fitted with the nominal model and the distribution of the differences of signal yields between the fitted nominal model and alternative model used to generate the toy datasets is fitted with a Gaussian function.
- The absolute value of the mean of that Gaussian function is taken as the systematic uncertainty associated to the arbitrary choice made in the nominal fit.

Figure 5.31 shows the alternative model where more components are added to the fit and Figure 5.32 shows the distributions of the differ-

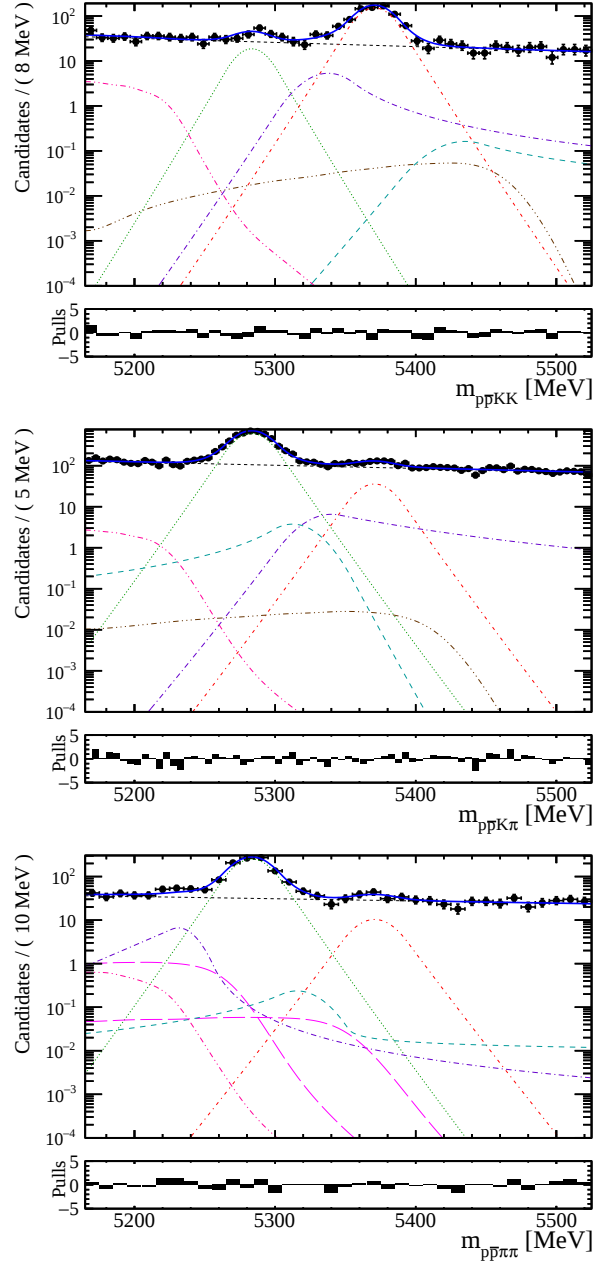


Figure 5.31: Simultaneous fit to the three mass spectra excluding all charm contributions using the third alternative fit model described in Section 5.8.6 (more components added to the fit): (top) $p\bar{p}KK$, (middle) $p\bar{p}K\pi$ and (bottom) $p\bar{p}\pi\pi$. In each fit the following convention is used to identify the various model components: a green dotted line for the B^0 signal, a red dot-dashed line for the B_s^0 signal, a black dashed line for the combinatorial background, a violet dot-dashed line for the dominant reflection, a cyan dashed line for the second dominant reflection, a brown triple-dot-dashed line for $\Lambda_b^0 \rightarrow p\bar{p}pK$, a pink double-dot-dashed line for the partially reconstructed background and magenta long-dashed lines for the partially reconstructed backgrounds $B_{(s)}^0 \rightarrow p\bar{p}\eta$ with $\eta \rightarrow \pi\pi\gamma$. Refer to the text for further details on the fit model templates and parameters.

ences to the fitted values using the nominal model with respect to the generated values using that alternative model.

The yields of the misidentified components are fixed in the fit knowing the ratio of efficiencies of the misidentified components between the mass spectrum where they are misidentified and where they are signal. This ratio has an associated uncertainty that should be propagated in the fit. To do so the fit is repeated 27 times fixing independently the value of the three ratios of efficiencies to the nominal value or to the nominal value plus or minus the associated uncertainty (where the uncertainty contains both the statistical and the PID systematic uncertainty of Section 5.8.4). Half of the maximum difference among the possible combinations is taken as the systematic uncertainty.

*Yields of the
misidentified
components*

Table 5.25 collects the values of the various components of the systematic uncertainties on the fit model together with the total systematic uncertainty associated to the fit obtained by summing in quadrature the various components.

To estimate the systematic uncertainty due to the fit model of the normalisation mode a similar method is used. Alternative models are obtained as follows:

Normalisation mode

1. substituting the DSCB functions and the relativistic Breit-Wigner function with kernel estimation functions obtained from the MC;
2. substituting the exponential function with a linear function to describe the background;
3. additional components are added to the fit and their normalisation is left free:
 - $B^0 \rightarrow p\bar{p}K^{*0}$ modelled in the same way as the signal in the $p\bar{p}K\pi$ and $p\bar{p}$ mass spectra and with an exponential function in the $K\pi$ mass spectrum;
 - $B^0 \rightarrow p\bar{p}K\pi$ modelled in the same way as the signal in the $p\bar{p}K\pi$ mass spectrum, an exponential function in the pp mass spectrum and a linear function in the $K\pi$ mass spectrum.
 - $\Lambda_b^0 \rightarrow J/\psi pK$ where the proton is misidentified as a pion; It is modelled in the same way as the signal in the $p\bar{p}$ mass spectrum and with kernel estimation functions obtained

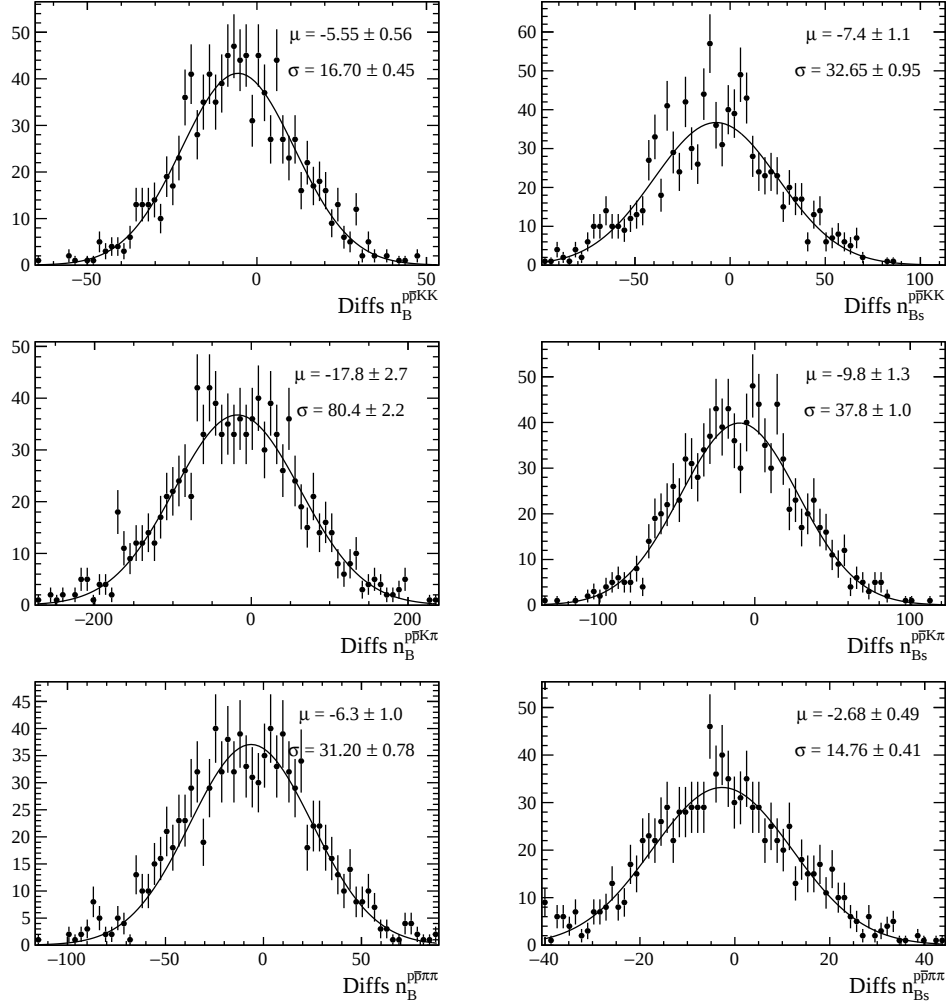


Figure 5.32: Distributions of the differences of the yield parameters obtained generating pseudoexperiments with the third alternative model described in Section 5.8.6 (more components added to the fit) and fitted with the nominal fit model described in Section 5.7. B^0 on the left and B_s^0 on the right; $pp̄KK$ at the top, $pp̄K\pi$ in the middle and $pp̄\pi\pi$ at the bottom.

Table 5.25: Yields from the fit along with their statistical and systematic uncertainties. Refer to the text for details.

Decay channel	yield	stat. uncert.	sgn. shape syst.	bkg. shape syst.	add compo- nents syst.	misID effs. syst.	total syst.	total syst. [%]
$B^0 \rightarrow p\bar{p}KK$	68.2	16.8	1.1	0.5	5.6	0.1	5.7	8.3
$B_s^0 \rightarrow p\bar{p}KK$	635.4	31.5	1.6	0.9	7.4	2.3	7.9	1.2
$B^0 \rightarrow p\bar{p}K\pi$	4154.8	82.6	0.4	1.4	17.8	1.7	18.0	0.4
$B_s^0 \rightarrow p\bar{p}K\pi$	246.35	39.32	0.02	0.99	9.84	1.63	10.02	4.1
$B^0 \rightarrow p\bar{p}\pi\pi$	902	35	4	2	6	1	8	0.8
$B_s^0 \rightarrow p\bar{p}\pi\pi$	38.54	15.87	0.73	0.02	2.68	0.65	2.85	7.4
$B^0 \rightarrow J/\psi K^*$	1216	45	2	2	27	-	27	2.2

from simplified simulations like the ones of Section 5.5 in the $p\bar{p}$ and $K\pi$ mass spectra.

The dataset is fitted with these alternative models to fix their free parameters; Figure 5.33 shows an example of one of these fits. A thousand toy datasets are generated for each alternative model and fitted with the nominal one. The assigned systematic uncertainty for each alternative model is the mean of the distribution of the difference in the number of signal events between the fitted nominal model and the alternative one used to generate the toys. Figure 5.33 shows one of these distributions while the values of the systematic uncertainties are collected in Table 5.25.

5.8.7 Summary of systematic uncertainties

Table 5.26 summarises all systematic uncertainties on the ratios of $B_{(s)}^0 \rightarrow p\bar{p}hh'$ branching fractions while Table 5.27 summarises all systematic uncertainties on the absolute $B_{(s)}^0 \rightarrow p\bar{p}hh'$ branching fractions obtained from their ratios with the branching fraction of the normalisation channel. The statistical uncertainty on the efficiency is taken as the systematic uncertainty due to the finite size of the MC and of the statistical uncertainties on the *sWeight*'ed distributions of events in the phase space. In Tables 5.26 and 5.27 the uncertainties due to the finite MC sample, the PID calibration and the fit model are the sums in quadrature of the relative uncertainties of the numerator and the denominator (that in the measurement of absolute branching fractions is the normalisation mode $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$). The uncertainties due to the phase space corrections, the charm vetoes and the normalisation of the reflections in the fits do not affect the normalisation mode; hence in the measurement of the absolute branching fractions they are only due to the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decay while in the ratios of branching fractions of different $B_{(s)}^0 \rightarrow p\bar{p}hh'$, the sums in quadrature of the relative uncertainties of the numerator and the denominator are considered. Finally the uncertainties on the trigger, tracking efficiencies and the external ones (branching fraction of the normalisation mode where appropriate and f_s/f_d) have been assigned directly to the ratios.

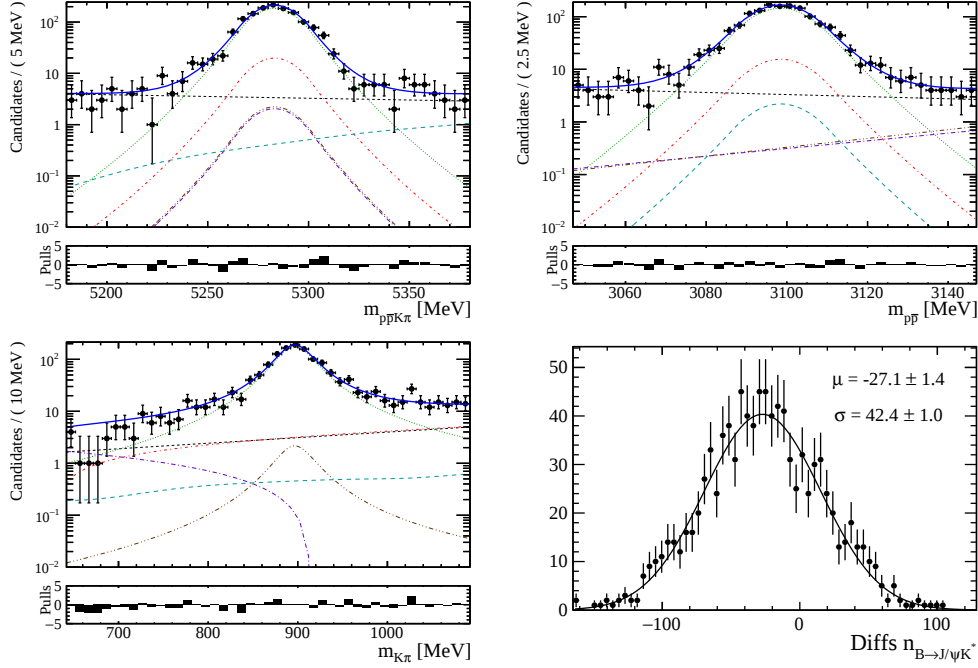


Figure 5.33: 3-D fit to the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ data using the third alternative model in which the more components are added to the fit. At the top left the $m_{p\bar{p}K\pi}$ mass spectrum, at the top right the $m_{p\bar{p}}$ mass spectrum, and at the bottom left $m_{K\pi}$ mass spectrum. In each fit the following convention is used to identify the various model components: a green dotted line for the normalisation mode $B^0 \rightarrow J/\psi K^{*0}$, a red dot-dashed line for the $K\pi$ S-wave, a black dashed line for the combinatorial background, a violet dot-dashed line for the non resonant $B^0 \rightarrow p\bar{p}K^{*0}$, a cyan dashed line for non resonant $B^0 \rightarrow p\bar{p}K\pi$ and a brown triple-dot-dashed line for $\Lambda_b^0 \rightarrow J/\psi pK$. Refer to the text for further details on the fit model templates and parameters. At the bottom right: distributions of the differences of the $B^0 \rightarrow (J/\psi \rightarrow p\bar{p})(K^{*0} \rightarrow K^+\pi^-)$ yield obtained generating pseudoexperiments with the alternative model in which the more components are added to the fit described in Section 5.8.6 and fitted with the nominal fit model described in Section 5.7.

Table 5.26: Relative contributions in percentage of the different components of the uncertainties on the ratio of branching fractions among $B_{(s)}^0 \rightarrow p\bar{p}h'h'$ decays. Refer to the text for details.

Decay Channel	$\frac{\mathcal{B}(B^0 \rightarrow p\bar{p}KK)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)}$	$\frac{\mathcal{B}(B^0 \rightarrow p\bar{p}\pi\pi)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)}$	$\frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)}{\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)}$	$\frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}KK)}{\mathcal{B}(B_s^0 \rightarrow p\bar{p}KK)}$
Statistical error efficiency	3.1	2.3	4.8	4.8
Lo trigger	3.7	3.7	3.7	3.7
Tracking efficiencies	1.1	1.1	1.1	1.1
PID calibration	1.7	1.2	1.6	1.8
B_s lifetime	-	-	2.5	0.1
Charm vetoes	0.5	0.1	0.4	0.5
Phase-space correction	0.1	0.1	0.1	0.1
Sgn. shape	1.6	0.4	0.0	0.2
Bkg. shape	0.7	0.2	0.4	0.4
Extra components	8.2	0.8	4.0	4.2
Normalisation reflections in fit	0.2	0.1	0.7	0.8
f_d/f_s	-	-	5.8	-
Total systematic uncertainty	9.9	4.7	7.9	7.7
Statistical uncertainty fit	24.8	4.5	16.3	17.0

Table 5.27: Relative contributions in percentage of the different components of the uncertainties on the absolute branching fractions of $B_{(s)}^0 \rightarrow p\bar{p}h h'$ decays. Refer to the text for details.

Decay Channel	$B^0 \rightarrow p\bar{p}KK$	$B^0 \rightarrow p\bar{p}K\pi$	$B^0 \rightarrow p\bar{p}\pi\pi$	$B_s^0 \rightarrow p\bar{p}KK$	$B_s^0 \rightarrow p\bar{p}K\pi$	$B_s^0 \rightarrow p\bar{p}\pi\pi$	$B_s^0 \rightarrow p\bar{p}K\pi$	$B_s^0 \rightarrow p\bar{p}\pi\pi$
Statistical error efficiency	3.0	1.7	2.1	1.9	4.7	4.1		
Lo trigger	3.7	3.7	3.7	3.7	3.7	3.7		
Tracking efficiencies	1.1	1.1	1.1	1.1	1.1	1.1		
PID calibration	1.9	1.7	1.4	2.0	1.7	1.4		
B_s lifetime	-	-	-	2.6	2.5	2.7		
Charm vetoes	0.5	0.1	0.1	0.4	0.4	0.2		
Phase-space correction	0.1	0.1	0.1	0.1	0.1	0.1		
Sgn. shape	1.7	0.2	0.4	0.3	0.2	1.9		
Bkg. shape	0.7	0.2	0.3	0.2	0.4	0.2		
Extra components	8.4	2.3	2.3	2.5	4.6	7.3		
Normalisation reflections in fit	0.2	0.0	0.1	0.4	0.7	1.7		
Branching fraction normalisation mode	6.9	6.9	6.9	6.9	6.9	6.9		
f_d/f_s	-	-	-	5.8	5.8	5.8		
Total systematic uncertainty	10.1	5.1	5.2	6.0	8.2	10.0		
Statistical uncertainty fit	25.0	4.2	5.4	6.2	16.4	41.3		

5.9 RESULTS

Significances To estimate the significance of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ peaks the profile likelihood method has been used. The profile likelihood is defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\nu})}{L(\hat{\mu}, \hat{\nu})}$$

where μ is the parameter of interest (the yield whose significance is determined) and ν are the other free parameters of the fit; $L(\hat{\mu}, \hat{\nu})$ is the likelihood value obtained from the best fit to the data while $L(\mu, \hat{\nu})$ is the likelihood value obtained from the best fit with μ fixed and ν left free. Asymptotically $-2 \ln \lambda(\mu)$ is distributed as a χ^2 distribution as proven by the Wilks's theorem so the significance can be defined as:

$$\sqrt{-2 \ln \lambda(0)}. \quad (5.6)$$

In each mass spectrum one decay mode has overwhelming significance (more than 20σ); hence computing the significance taking into account the systematic uncertainties from the fit is not necessary. For the other modes the profile likelihood is convolved with a Gaussian function of mean zero and sigma equal to the fit systematic uncertainty. The significance is then computed with Equation (5.6), using in place of the profile likelihood its convolution with the Gaussian function. Finally as the significance of $B_s^0 \rightarrow p\bar{p}\pi\pi$ is below 3σ an upper limit at 90% CL is computed. Defining

$$I(x) = \frac{\int_0^x \tilde{\lambda}(\xi) d\xi}{\int_0^{+\infty} \tilde{\lambda}(\xi) d\xi} \quad (5.7)$$

where $\tilde{\lambda}$ is the convolution of the profile likelihood with the Gaussian function, the upper limit is the value of x where $I(x) = 0.9$. Table 5.28 collects the significances of the various decay modes as well as the significances taking into account the fit systematic uncertainties where appropriate and the upper limit for $B_s^0 \rightarrow p\bar{p}\pi\pi$. Figure 5.34 shows the profile likelihood for the three decay modes in which the significance is less than 20.

Branching fractions To compute the branching fraction of the charmless $B_{(s)}^0 \rightarrow p\bar{p}hh'$ decays collected in Table 5.29, Equations (5.1) and (5.2) are used while

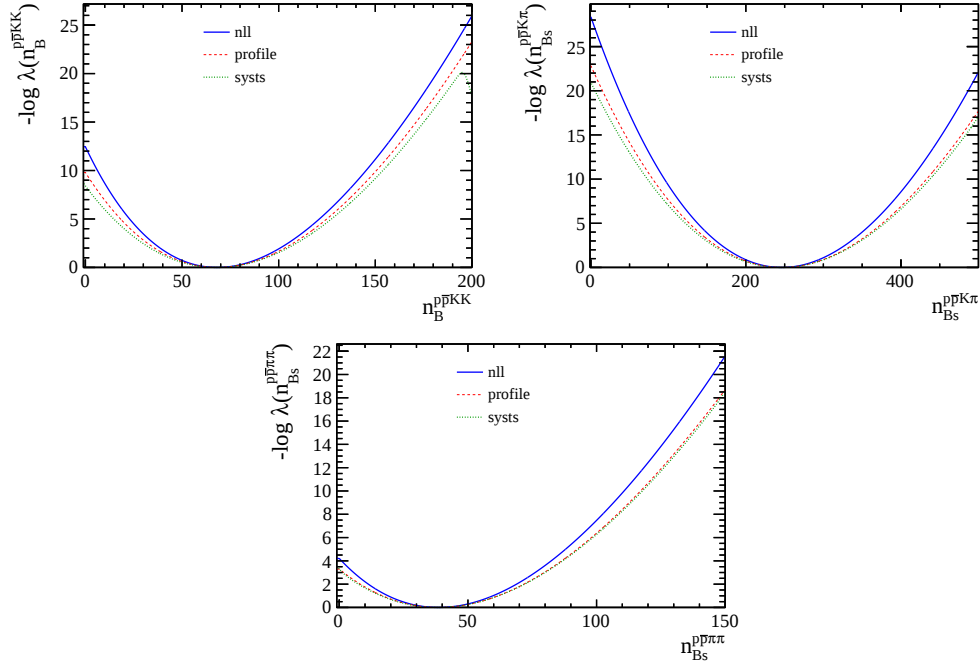


Figure 5.34: Profile likelihood for the yields of the three decay modes in which the significance is less than 20: (*top left*) $B^0 \rightarrow p\bar{p}KK$, (*top right*) $B_s^0 \rightarrow p\bar{p}K\pi$ and (*bottom*) $B_s^0 \rightarrow p\bar{p}\pi\pi$. In all the plots the blue solid line is the negative log likelihood (parabolic approximation of the negative profile log likelihood around its minimum), the dashed red line is the negative profile log likelihood and the dotted green line is the convolution of this last one with a Gaussian function centred in zero and with the systematic uncertainty on the yield as σ . Refer to the text for more information.

Table 5.28: Significances of the $B_{(s)}^0 \rightarrow p\bar{p}hh'$ peaks without considering the fit systematic uncertainties. When the significance is less than 20 also the value considering fit systematic uncertainties is given. If the significance is less than 3 also the upper limit at 90% CL on the yield is reported.

Parameter	Significance	Significance (syst)	Upper limit (90% CL)
$n_B^{p\bar{p}KK}$	4.4	4.1	-
$n_{B_s}^{p\bar{p}KK}$	30.0	-	-
$n_B^{p\bar{p}K\pi}$	74.3	-	-
$n_{B_s}^{p\bar{p}K\pi}$	6.8	6.5	-
$n_B^{p\bar{p}\pi\pi}$	36.2	-	-
$n_{B_s}^{p\bar{p}\pi\pi}$	2.6	2.6	60.6

Table 5.29: Branching fractions computed using Equations (5.1) and (5.2). The first uncertainty is statistical, the second systematic, the third comes from the uncertainty on the branching fraction of the normalisation mode and the fourth one, if present, comes from the uncertainty on f_d/f_s .

Decay Channel	Branching fraction / 10^{-6}				
$B^0 \rightarrow p\bar{p}KK$	0.113	± 0.028	± 0.011	± 0.008	
$B^0 \rightarrow p\bar{p}K\pi$	5.9	± 0.3	± 0.3	± 0.4	
$B^0 \rightarrow p\bar{p}\pi\pi$	2.7	± 0.1	± 0.1	± 0.2	
$B_s^0 \rightarrow p\bar{p}KK$	4.2	± 0.3	± 0.2	± 0.3	± 0.2
$B_s^0 \rightarrow p\bar{p}K\pi$	1.30	± 0.21	± 0.11	± 0.09	± 0.08
$B_s^0 \rightarrow p\bar{p}\pi\pi$	0.41	± 0.17	± 0.04	± 0.03	± 0.02

the ratios between branching fractions collected in Table 5.30 are computed with the similar formulae,

$$\frac{\mathcal{B}(B_{(s)}^0 \rightarrow X)}{\mathcal{B}(B_{(s)}^0 \rightarrow Y)} = \frac{N(B_{(s)}^0 \rightarrow X)}{N(B_{(s)}^0 \rightarrow Y)} \cdot \frac{\varepsilon(B_{(s)}^0 \rightarrow Y)}{\varepsilon(B_{(s)}^0 \rightarrow X)}, \quad (5.8)$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)}{\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)} = \frac{N(B_s^0 \rightarrow p\bar{p}K\pi)}{N(B^0 \rightarrow p\bar{p}K\pi)} \cdot \frac{\varepsilon(B^0 \rightarrow p\bar{p}K\pi)}{\varepsilon(B_s^0 \rightarrow p\bar{p}K\pi)} \cdot \frac{f_d}{f_s}. \quad (5.9)$$

Table 5.31 shows the comparison between the yields of the decay modes considered, corrected for the efficiencies, with the naive expectations used for the optimisation of the selection.

Upper limit on
 $B_s^0 \rightarrow p\bar{p}\pi\pi$

To compute the upper limit on the branching fraction of $B_s^0 \rightarrow p\bar{p}\pi\pi$ is analogous to computing the limit on its yield,

$$\mathcal{B}(B_s^0 \rightarrow p\bar{p}\pi\pi) < 6.6 \times 10^{-7} \text{ at 90\% CL.}$$

The difference is that now to compute $I(x)$ of Equation (5.7) the profile likelihood is calculated directly on the branching fraction and not on

Table 5.30: Ratio of branching fractions. The first uncertainty is statistical, the second systematic and the third one, if present, comes from the uncertainty on f_d/f_s .

$\mathcal{B}(B^0 \rightarrow p\bar{p}KK)/\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)$	$0.019 \pm 0.005 \pm 0.002$
$\mathcal{B}(B^0 \rightarrow p\bar{p}\pi\pi)/\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)$	$0.46 \pm 0.02 \pm 0.02$
$\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)/\mathcal{B}(B^0 \rightarrow p\bar{p}K\pi)$	$0.22 \pm 0.04 \pm 0.02 \pm 0.01$
$\mathcal{B}(B_s^0 \rightarrow p\bar{p}K\pi)/\mathcal{B}(B_s^0 \rightarrow p\bar{p}KK)$	$0.31 \pm 0.05 \pm 0.02$

Table 5.31: Comparison between the yields corrected for the efficiencies and the naive expectations used for the optimisation of the selection. The first uncertainty is statistical and the second is systematic.

Decay Channel	Estimated from $B_{(s)}^0 \rightarrow hh^{(\prime)}$	From Data
$B^0 \rightarrow p\bar{p}KK$	0.007	$0.019 \pm 0.005 \pm 0.002$
$B^0 \rightarrow p\bar{p}K\pi$	1	1
$B^0 \rightarrow p\bar{p}\pi\pi$	0.26	$0.46 \pm 0.02 \pm 0.02$
$B_s^0 \rightarrow p\bar{p}KK$	0.33	$0.182 \pm 0.010 \pm 0.010$
$B_s^0 \rightarrow p\bar{p}K\pi$	0.07	$0.057 \pm 0.009 \pm 0.005$
$B_s^0 \rightarrow p\bar{p}\pi\pi$	0.01	$0.018 \pm 0.007 \pm 0.002$

the yield, using the appropriate multiplicative factor taken from Equation (5.1). Moreover the sigma of the Gaussian function convolved to the profile likelihood to take into account the systematic uncertainties is the sum in quadrature of the uncertainties on the yield and on the multiplicative factor used to convert the yield into a branching fraction. Figure 5.35 shows the distribution of $I(BR)$ and the distribution of $\tilde{\lambda}(BR)$ with the region where $I(BR) < 0.9$ highlighted.

In summary, in this chapter the measurements of the branching fractions of the charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}hh'$ has been reported. Three of these modes are observed for the first time and evidence is found for a fourth. In particular the decays $B_s^0 \rightarrow p\bar{p}KK$ and $B_s^0 \rightarrow p\bar{p}K\pi$ are, together with the $B_s^0 \rightarrow p\bar{\Lambda}K^-$ decay, the first baryonic B_s^0 decays to be observed.

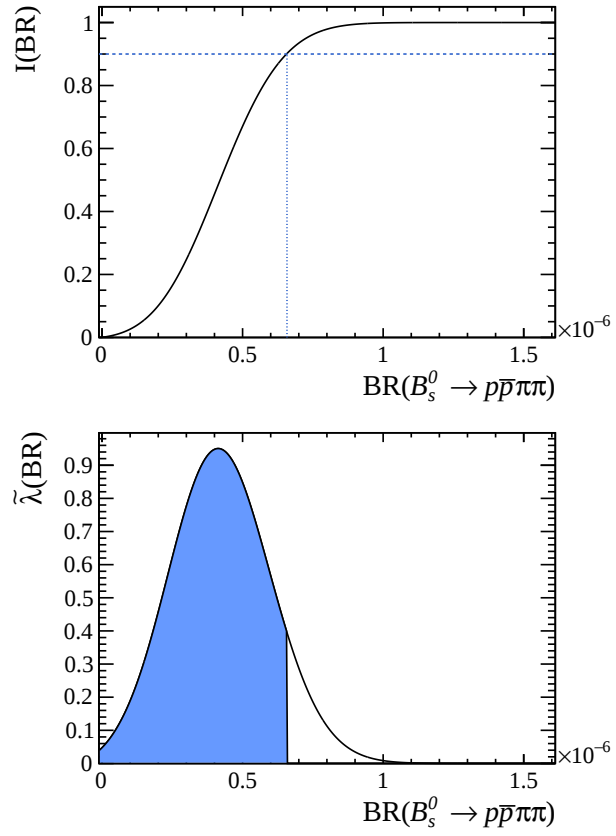


Figure 5.35: Distribution of (top) $I(BR)$ with a dashed blue horizontal line at 0.9 and (bottom) distribution of $\tilde{\lambda}(BR)$ with the region where $I(BR) < 0.9$ highlighted. Refer to the text for details.

CONCLUSIONS

The focus of this thesis is the search for the four-body charmless baryonic decays $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$ using the full Run I collision data from the LHCb detector, presented in Chapter 5. Three new decay modes are observed for the first time ($B_s^0 \rightarrow p\bar{p}K^+K^-$, $B_s^0 \rightarrow p\bar{p}K^\pm\pi^\mp$ and $B^0 \rightarrow p\bar{p}\pi^+\pi^-$), strong evidence is reported for $B^0 \rightarrow p\bar{p}K^+K^-$ while, as the significance of $B_s^0 \rightarrow p\bar{p}\pi^+\pi^-$ is below 3σ , an upper limit is set ($\mathcal{B}(B_s^0 \rightarrow p\bar{p}\pi^+\pi^-) < 6.6 \times 10^{-7}$ at 90% CL). The inclusive charmless branching fractions have been measured (Table 5.29) and also ratios of branching fractions among various $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$ decays have been obtained (Table 5.30). This work represents the first observation of four-body charmless baryonic B_s^0 decays and also the first observation of baryonic B_s^0 decays, together with the observation of the $B_s^0 \rightarrow p\bar{\Lambda}K^-$ mode [30]. In the future these $B_{(s)}^0 \rightarrow p\bar{p}h^+h'^-$ modes could be used for studies of CP violation in baryonic B decays. The outcome of this study can also be found in the publication [118].

A feasibility study for a search for the lepton flavour violating decay $\tau \rightarrow \phi\mu$ using the full Run I LHCb dataset has been presented in Chapter 4. The expected upper of $\mathcal{B}(\tau \rightarrow \phi\mu) \leq 3.8 \cdot 10^{-6}$ is obtained at 90% CL. This is approximately 50 times larger than the current World's best limit set by Belle.

The novel real-time alignment and calibration adopted in Run II by the LHCb collaboration has been discussed in Chapter 3, focusing in particular on the preparation, the commissioning and the performance of the real-time alignment of the vertex detector in the first two years of Run II. This real-time alignment and calibration strategy is essential to allow more stable data taking conditions and to perform at the trigger level the same optimal offline reconstruction. This ensures more efficient trigger selections and the possibility to perform physics analyses directly on the trigger output.

A

CORRECTION OF τ AND D PRODUCTION RATES

As detailed in [119, 120] the τ production rates, predominantly from D_s , D and B mesons, need to be corrected to agree with the latest LHCb measurements of the charm and beauty production cross-sections and the PDG values of appropriate branching fractions. The same must be done for the D production rates for the $\tau \rightarrow \phi\mu$ normalisation sample. This is done by generating a different Monte Carlo sample for each production channel and then combining them together proportionally to the production rates. The production cross-sections used in these combinations are shown in Table A.1. The details are given in Table A.2. The quantities $\text{Calc}_{4\pi}$ are the production fractions expected in LHCb, with which the Monte Carlo samples should agree. In the Monte Carlo samples produced for this combination the D and D_s are forced to decay into a τ and the corresponding neutrino while for taus coming from B mesons only the lepton flavour violating decay is forced. The quantities ε_{GEN} are the original production rates in the Monte Carlo generator and so represent the probabilities to produce a D, a D_s - either prompt or from a B meson - or a τ from a B meson. The sum of these production rates is not 100% as different decays are forced to occur in the different channels. The measured efficiency of the generator level cuts includes also the production rate as some generator level cuts are applied to the generated samples:

$$\varepsilon_{\text{GEN|CUT}} = \varepsilon_{\text{GEN}} \cdot \varepsilon_{\text{CUT}}.$$

Table A.1: Production cross-sections at 8 TeV used to calculate the Monte Carlo mixing.

$\sigma(\text{pp} \rightarrow D^+ X)$	$3606 \pm 415 \mu\text{b}$
$\sigma(\text{pp} \rightarrow D_s^+ X)$	$1113 \pm 175 \mu\text{b}$
$\sigma(\text{pp} \rightarrow b\bar{b})$	$2 \times (298 \pm 36) \mu\text{b}$

In the mixed Monte Carlo each channel must be included with a number of events proportional to $\text{Calc}_{4\pi} \cdot \varepsilon_{\text{CUT}}$. Explicitly,

$$f_{\text{Gauss},j} = \frac{\text{Calc}_{4\pi,j} \cdot \varepsilon_{\text{CUT},j}}{\sum_i \text{Calc}_{4\pi,i} \cdot \varepsilon_{\text{CUT},i}}.$$

N_{prod} is the number of event in each production channel Monte Carlo and N_{mix} is the number of events actually used.

Table A.2: MC mixing method for (top) $\tau \rightarrow \phi\mu$ and (bottom) $D \rightarrow \phi\pi$ at 8 TeV. Refer to the text for more details.

Decay chain	Calc _{4π} 8 TeV (%)	$\epsilon_{\text{GEN CUT}}$ (%)	ϵ_{GEN} (%)	ϵ_{CUT} (%)	f_{Gauss} (%)	N _{prod}	N _{mix}
$D_s \rightarrow \tau$	70.2 ± 4.0	10.292 ± 0.077	63.111 ± 0.077	16.31 ± 0.12	71.1 ± 4.1	874997	839677
$B_x \rightarrow D_s \rightarrow \tau$	9.3 ± 2.0	1.567 ± 0.018	10.063 ± 0.019	15.57 ± 0.18	9.0 ± 1.9	110750	106557
$D^+ \rightarrow \tau$	4.10 ± 0.75	10.643 ± 0.055	65.202 ± 0.077	16.322 ± 0.087	4.16 ± 0.76	78999	49082
$B_x \rightarrow D^+ \rightarrow \tau$	0.190 ± 0.039	1.2164 ± 0.0081	7.948 ± 0.015	15.30 ± 0.11	0.181 ± 0.037	91499	2132
$B_x \rightarrow \tau$	16.2 ± 2.8	4.579 ± 0.013	29.691 ± 0.079	15.423 ± 0.061	15.5 ± 2.7	183250	183250

Decay chain	Calc _{4π} 8 TeV (%)	$\epsilon_{\text{GEN CUT}}$ (%)	ϵ_{GEN} (%)	ϵ_{CUT} (%)	f_{Gauss} (%)	N _{prod}	N _{mix}
D	96.0 ± 1.0	13.4 ± 1.2	89.130 ± 0.030	15.1 ± 1.3	96.5 ± 8.7	5010056	5010056
$B_x \rightarrow D$	4.0 ± 1.0	1.420 ± 0.011	10.870 ± 0.030	13.06 ± 0.11	3.49 ± 0.87	240924	181070

B

CROSS CHECKS FOR $\tau \rightarrow \phi \mu$

In this appendix are presented two cross checks to the $\tau \rightarrow \phi \mu$ feasibility study presented in Chapter 4.

B.1 PSEUDOEXPERIMENTS

A toy study has been performed to check if the error associated to the normalisation of the signal in the fit can be correctly interpreted as the standard deviation of a Gaussian distribution obtained by repeating the fit over different datasets. Three hundred different toy datasets have been produced according to the background distribution obtained in the fit to the original dataset. These toy datasets were then fitted to signal plus background; the distribution of the normalisation of the signal component is shown in Figure B.1, with a Gaussian fit. It can be seen that the standard deviation (RMS) of the distribution is comparable with the error, ± 90 , on the signal in the original fit to the data.

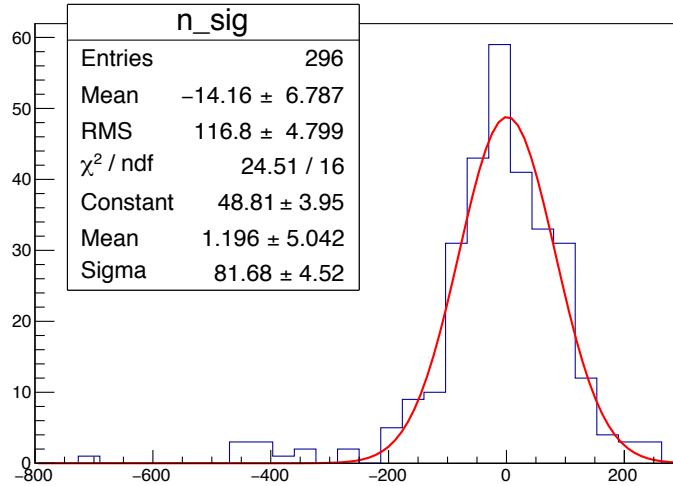


Figure B.1: Distribution of the fitted values of the normalisation of the signal over 300 toys fits.

To check that the limit scales as $\mathcal{L}^{-1/2}$ another toy study has been performed. Ten different toy datasets have been produced following

the distribution obtained in the fit to the original dataset, including only the background function but with a number of events 15 times larger than in the previous toy studies, to simulate the effect of using the whole Run I dataset. These simulated data have then been fitted and the average error on the signal yield is found to be 347, which is compatible with the expected scaling with luminosity ($90 \cdot \sqrt{15} = 349$).

B.2 CHECK ON YIELD NORMALISATION CHANNEL

To check the various steps of the analysis it is possible to compute the luminosity ($\mathcal{L}$) from the yield of $D \rightarrow \phi\pi$ in Figure 4.10 (N), its branching fraction ($\mathcal{B}$) and its efficiency (ϵ),

$$\mathcal{L} = \frac{N}{\sigma \mathcal{B} \epsilon} = (0.18 \pm 0.06) \text{ fb}^{-1}.$$

The values of the input quantities can be found in Table B.1 and the result is compatible with the value of the integrated luminosity in Table 4.8 used to compute the expected limit.

Table B.1: Quantities used to check the luminosity value.

N	5412 ± 213
$\sigma(D)$	$3606 \pm 396 \text{ } \mu\text{b}$
$\mathcal{B}(D \rightarrow \pi\phi(\rightarrow KK))$	$(2.65 \pm 0.09) \cdot 10^{-3}$
ϵ	$(3.08 \pm 0.12) \cdot 10^{-6}$
ϵ_{GEN}	0.101 ± 0.0007
ϵ_{STRIP}	0.001384 ± 0.000016
ϵ_{TRIG}	0.0972 ± 0.0035
ϵ_{CUTS}	0.2269 ± 0.0027



FURTHER STUDIES ON BDT

The multivariate selection of $B_{(s)}^0 \rightarrow p\bar{p}hh'$ events is presented in Section 5.4.2. This appendix collects some complementary material.

C.1 ON DIFFERENT BDTs FOR EACH $p\bar{p}hh'$ SPECTRUM

In this analysis a single MVA is trained for the three spectra. This is reasonable as the input variables are mainly topological and kinematical, and should not change much when substituting pions with kaons. However, in order to prove this assumption more rigorously, a different BDT has been trained for each $p\bar{p}hh'$ final state using the same hyper-parameters used in the baseline BDT (Table 5.2).

Figure C.1 compares the outputs of the three BDTs for signals and background. For each final state two BDTs were trained (BDT₀ and BDT₁) with the two different halves of the signal and background samples; the other halves were used as testing samples. The figure shows the agreement on the outputs for the two BDTs trained for each final state, similarly to Figure 5.4.

The ROC curves are evaluated on the respective testing samples. Figure C.2 compares the ROC curves of the various BDTs.

For a fair comparison one should however compare ROC curves evaluated over the same samples, as in Figure C.3. Each sub-figure of Figure C.3 presents four curves: three trained on the separate final states and one combining them. From this comparison it is clear that training a different BDT for each final state does not significantly improve the performance.

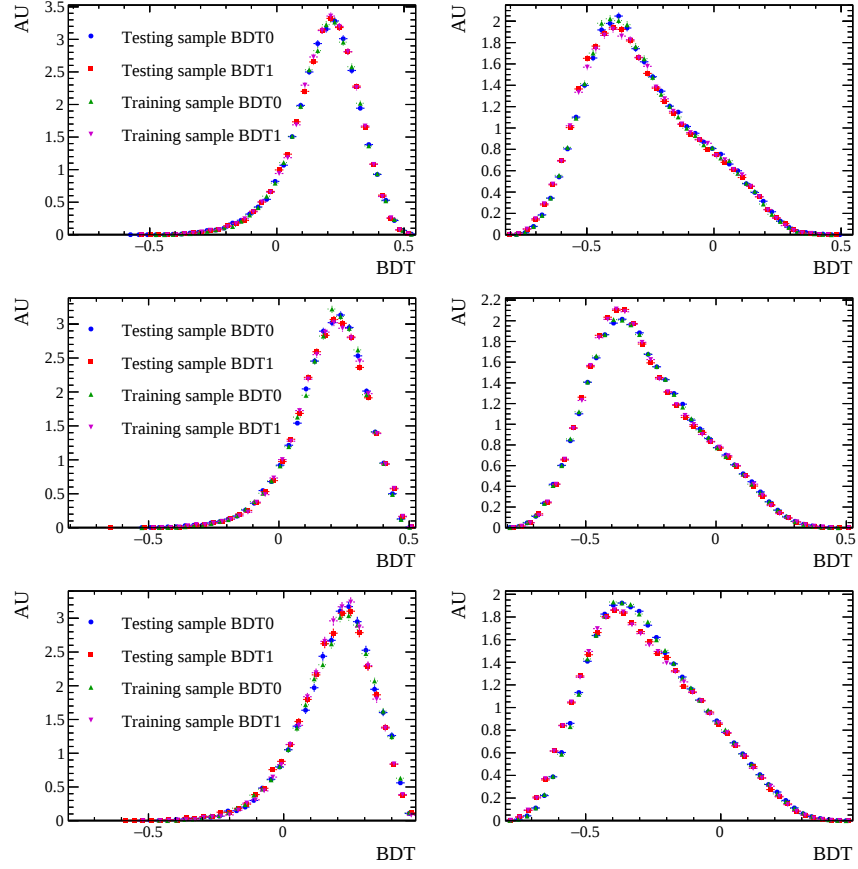


Figure C.1: Comparison of the BDT output for (*left column*) the signal and (*right column*) the background for the three final states: $p\bar{p}KK$ at the top, $p\bar{p}K\pi$ in the middle and $p\bar{p}\pi\pi$ at the bottom. For each final state two BDTs were trained (BDT₀ and BDT₁) with the two different halves of the signal and background samples and used the other halves as testing sample. The plots shows the BDT output for both BDTs on both training and testing samples.

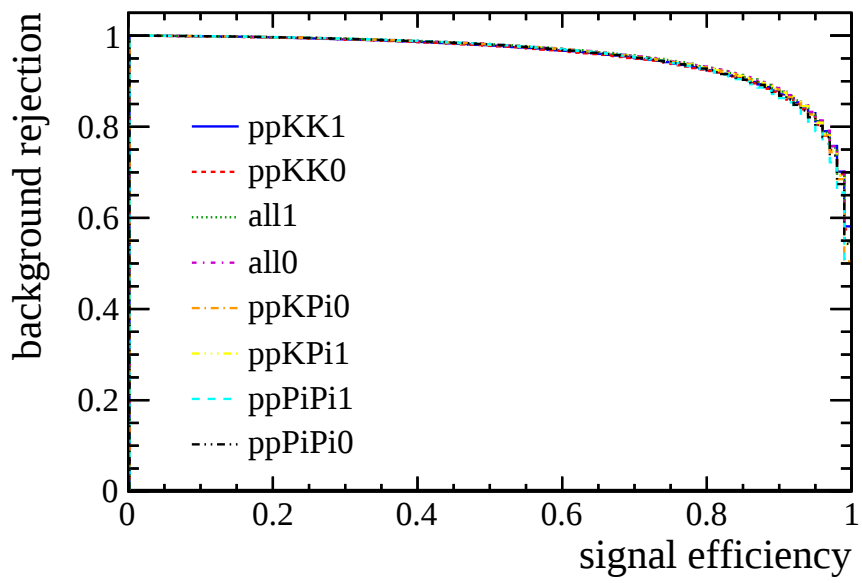


Figure C.2: Comparison of the ROC curves for the BDTs used in all three final states.

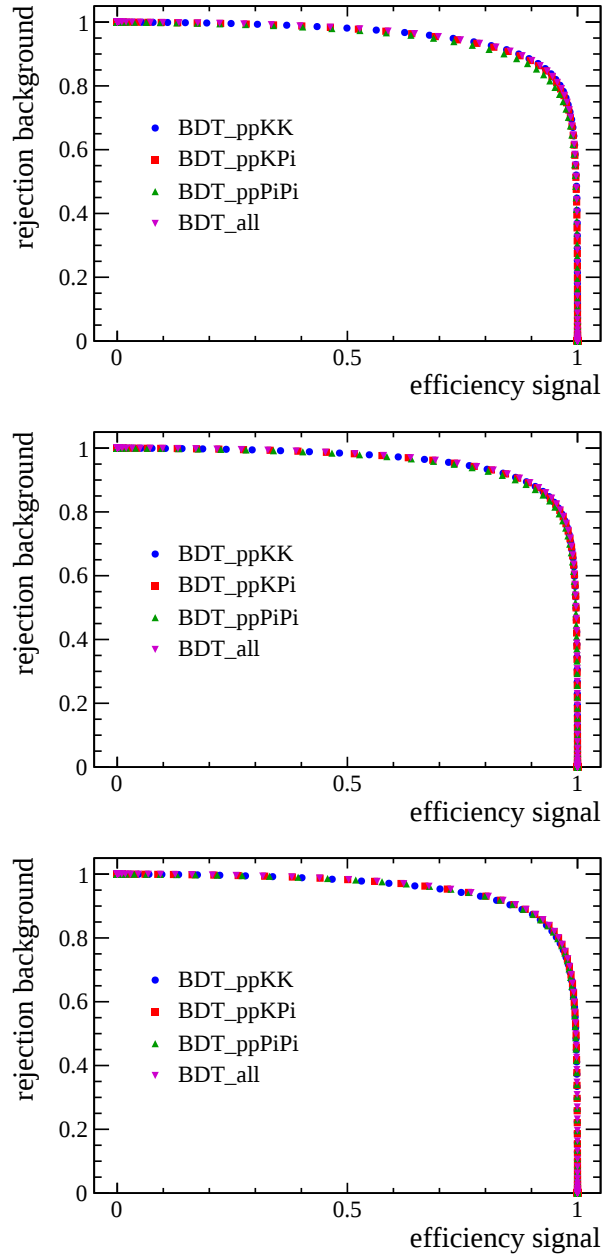


Figure C.3: Comparison of the ROC curves for the BDTs used in all three final states using as signal sample the MC samples of (*top*) $B_{(s)}^0 \rightarrow p\bar{p}KK$, (*middle*) $B_{(s)}^0 \rightarrow p\bar{p}K\pi$ and (*bottom*) $B_{(s)}^0 \rightarrow p\bar{p}\pi\pi$.

C.2 COMPARISON BETWEEN 2011 AND 2012

This analysis uses the combined Run I dataset comprising 2011 and 2012 data, and a single BDT is trained with this combined dataset. Figure C.4 shows the agreement between the BDT output for both signal and background samples of the two years, confirming that it is not necessary to train two different BDTs for the two years of data taking.

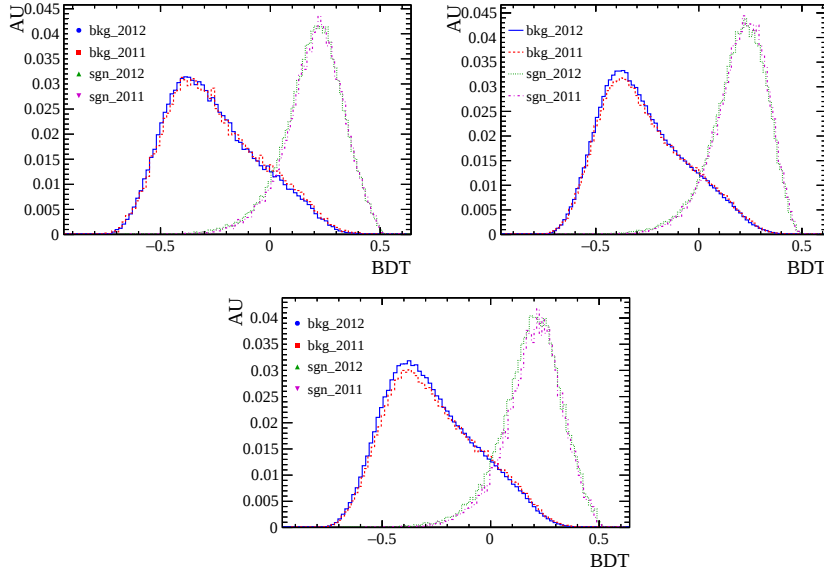


Figure C.4: Comparison of the BDT response for the signal and background samples for 2011 and 2012 conditions for the three final states: $p\bar{p}KK$ in the top left, $p\bar{p}K\pi$ in the top right, and $p\bar{p}\pi\pi$ at the bottom.

C.3 BDT INPUT VARIABLES

The distributions of the input variables in the signal and background samples are compared in Figures C.5 and C.6 and the linear correlations between the different variables are shown in Figure C.7. Some of the variables exhibit a high degree of correlation but are nevertheless kept as the differences between signal and background still contain discriminating power.

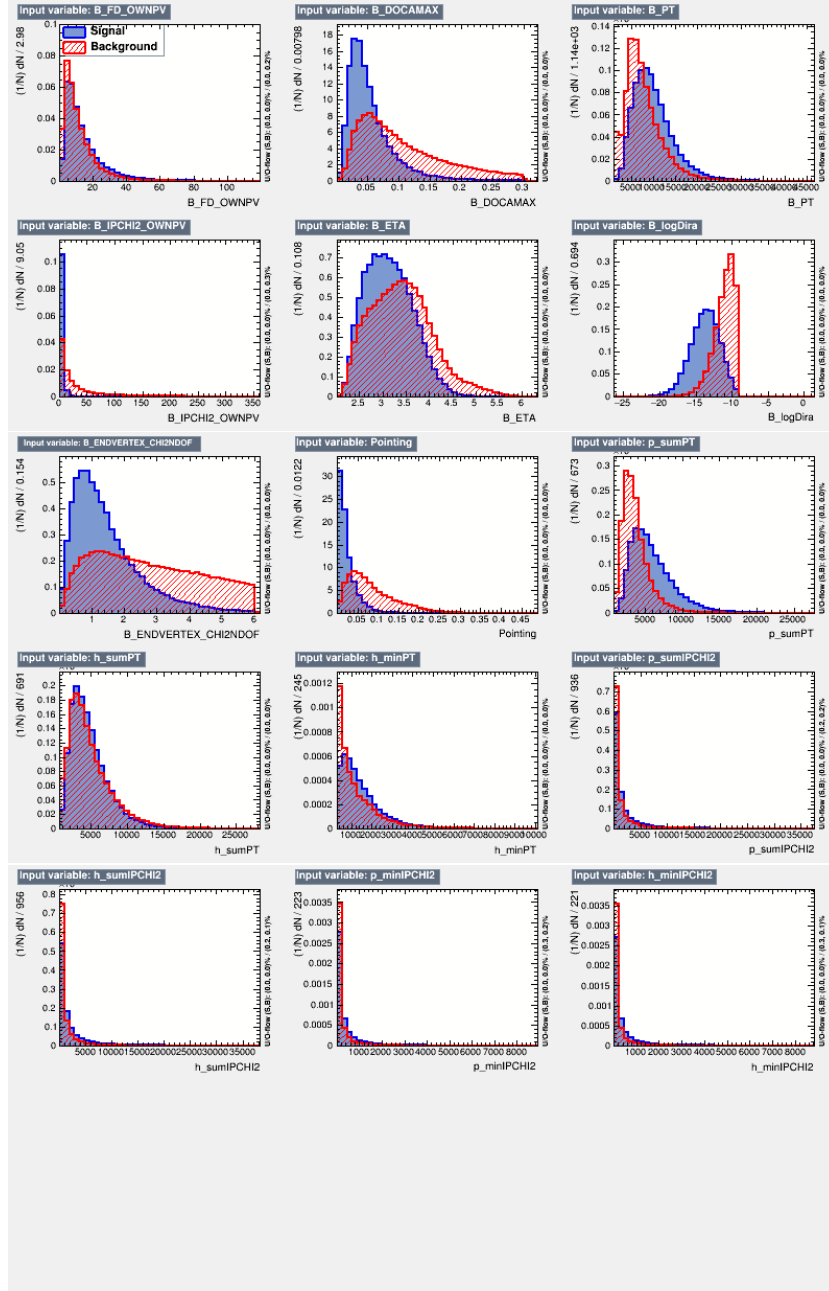


Figure C.5: Comparison of BDT0 input variables for (blue) the signal sample and (red) the background sample.

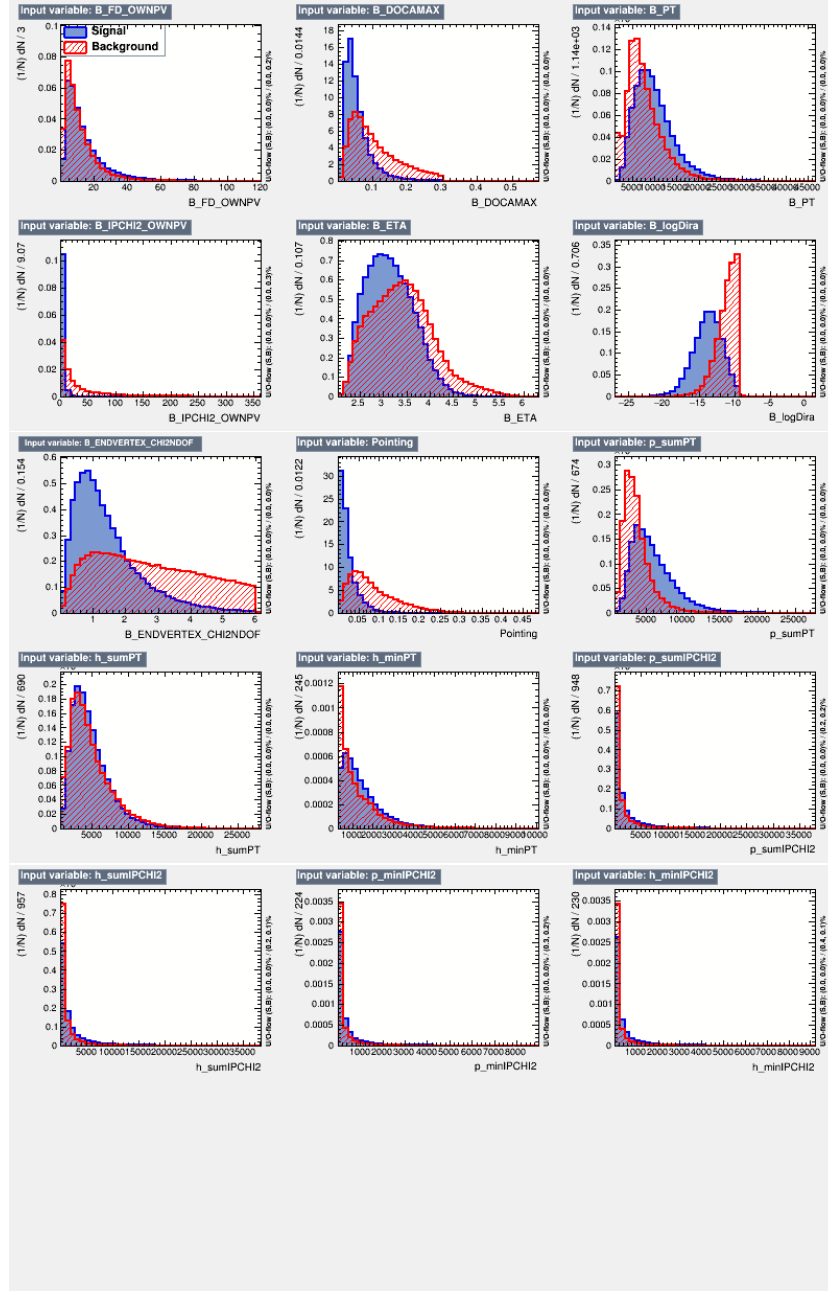


Figure C.6: Comparison of BDT1 input variables for (blue) the signal sample and (red) the background sample.

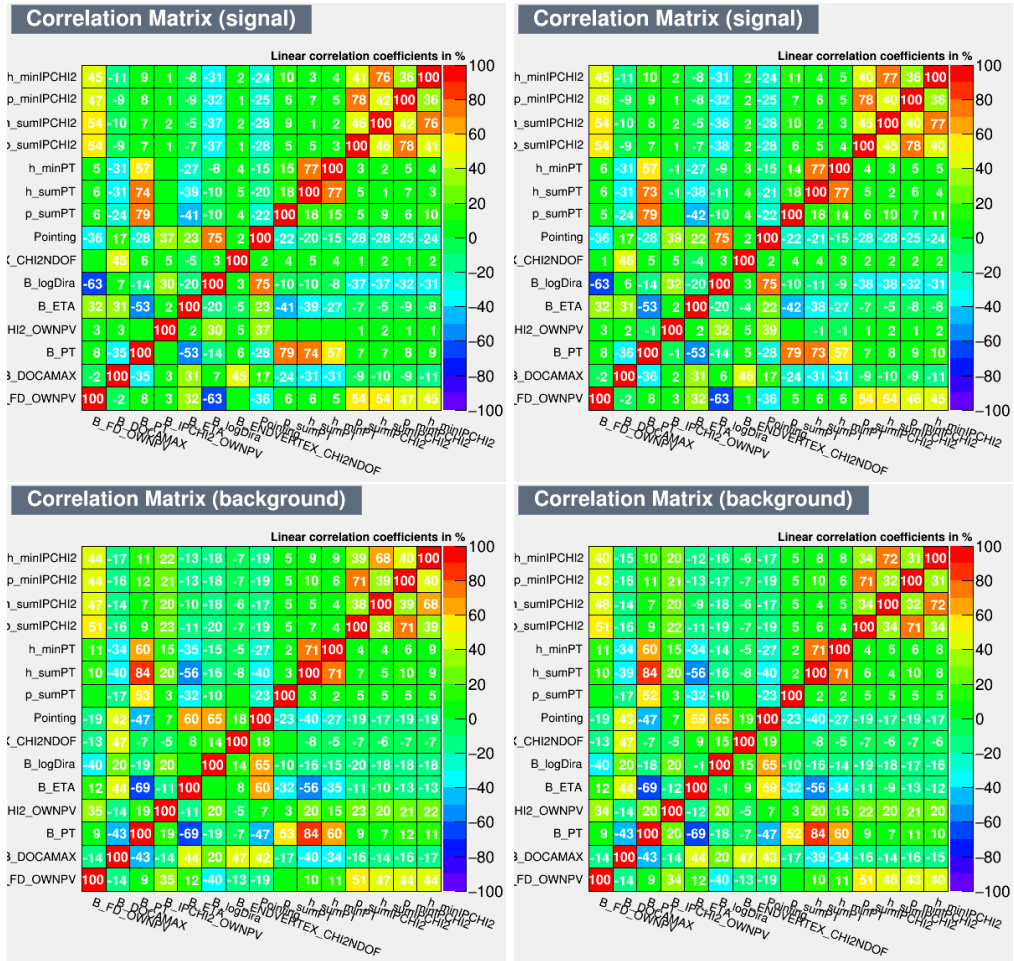


Figure C.7: Linear correlations between the variables used as input to the BDTs – BDT0 on the left and BDT1 on the right.

D | DATA-MC COMPARISONS

This appendix collects the comparison between data and MC for the variables used to train the BDT and the variables used in the PID selection.

D.1 BDT INPUT VARIABLES

To verify that it is possible to compute the BDT efficiency on the MC, the agreement between data and MC is checked for all BDT input variables. Figures [D.1](#) to [D.3](#) show this comparison for the three $p\bar{p}hh'$ final states. The data are *sWeight'*ed after the fit so, for a fair comparison, the full selection has been applied also to the MC samples. The agreement is overall good.

D.2 PID VARIABLES

The PID variables are not well described in the MC so, in the optimisation of the selection in Section [5.4.5](#), a resampling of those variables is used as detailed in Section [5.4.3](#). Figures [D.4](#) to [D.6](#) compare the PID variables used in the selection of protons and mesons before and after the resampling. Kaon PID_k is compared both when the meson is correctly identified and when it is misidentified. The MC sample $B^0 \rightarrow p\bar{p}K\pi$ is used and, for the correctly identified particles, the *sWeight'*ed data is also compared.

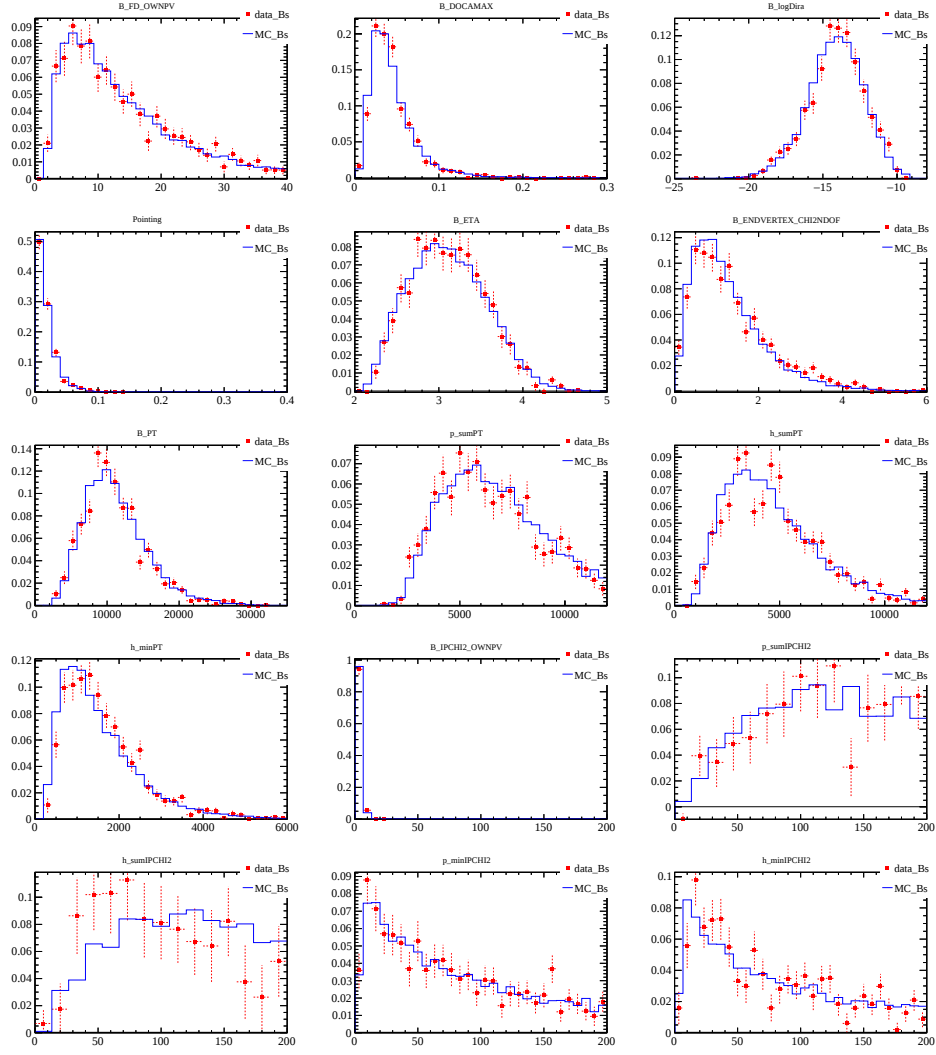


Figure D.1: Comparison of all BDT input variables between data and MC for the $B_s^0 \rightarrow p\bar{p}KK$ mode.

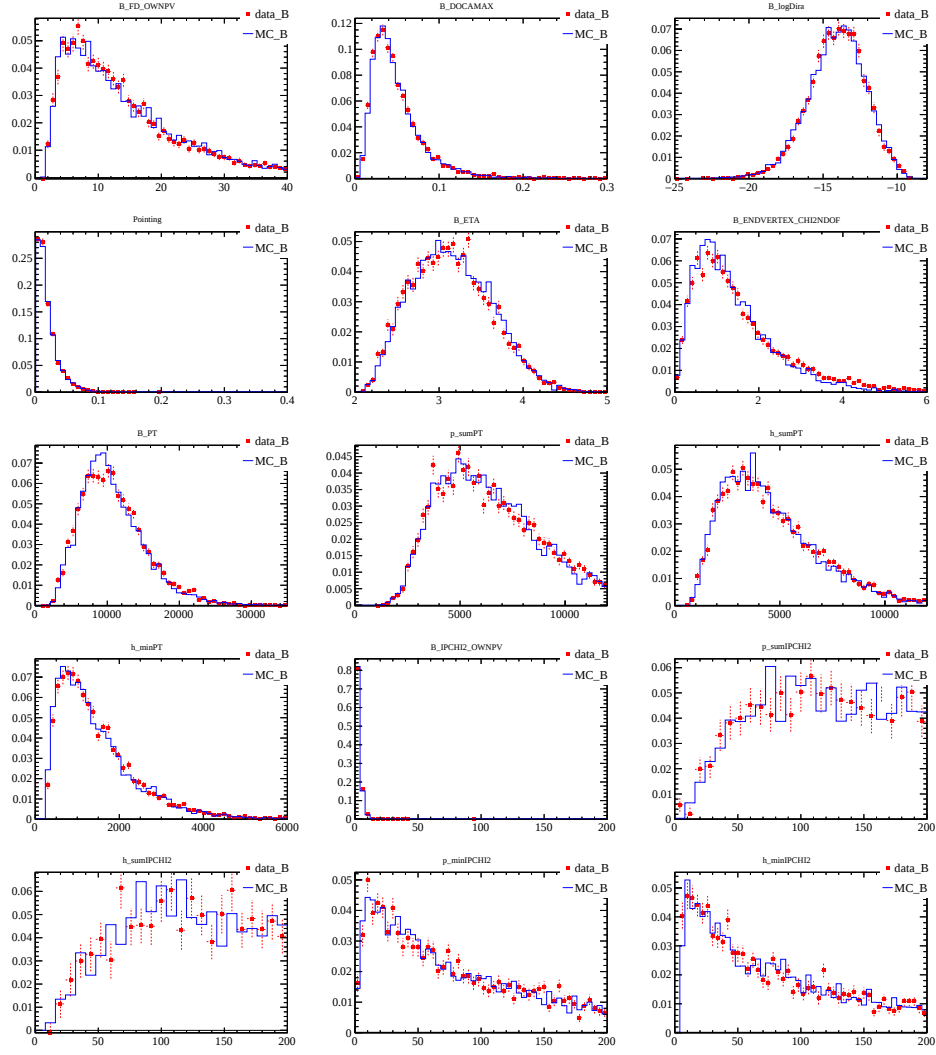


Figure D.2: Comparison of all BDT input variables between data and MC for the $B^0 \rightarrow \rho^+ \rho^- K \pi$ mode.

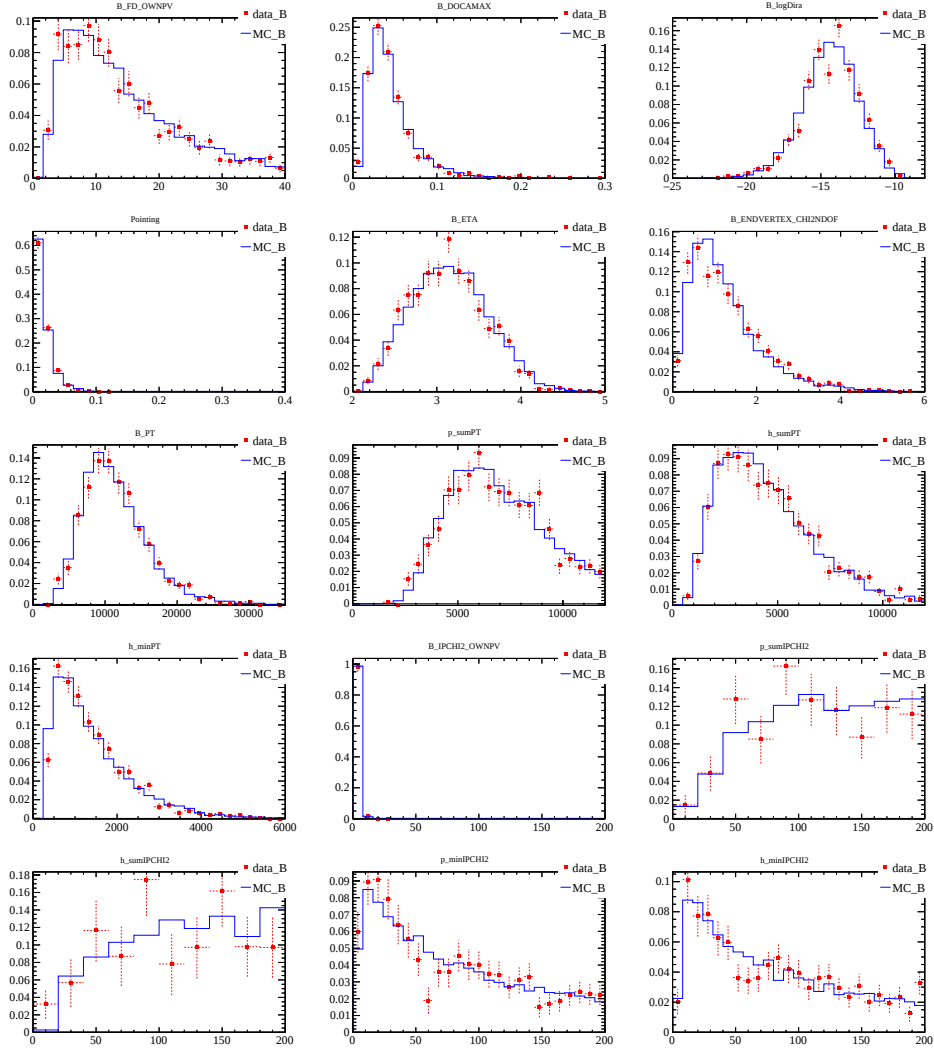


Figure D.3: Comparison of all BDT input variables between data and MC for the $B^0 \rightarrow \rho\pi\pi$ mode.

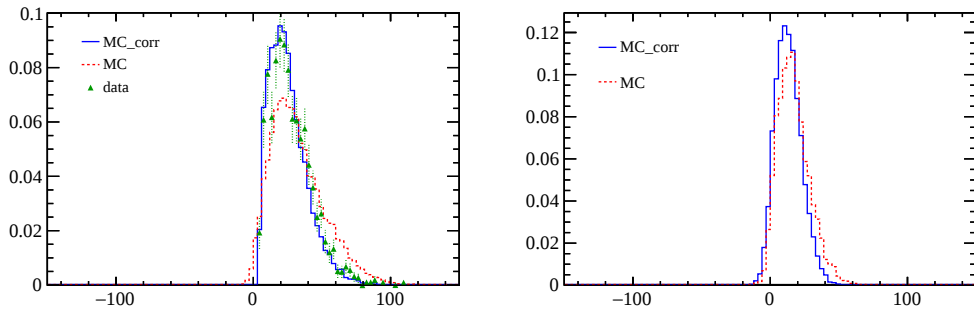


Figure D.4: Comparison of the MC PID distributions for (left) real kaons and (right) kaons reconstructed as pions in the stripping, before and after the resampling. The distribution of real kaons is compared also with $sWeight'$ ed data.

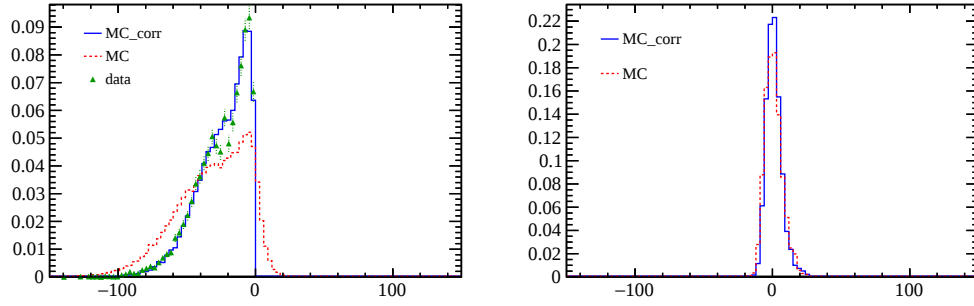


Figure D.5: Comparison of the MC PIDk distributions for (*left*) real pions and (*right*) pions reconstructed as kaons in the stripping, before and after the resampling. The distribution of real pions is compared also with *sWeight*'ed data.

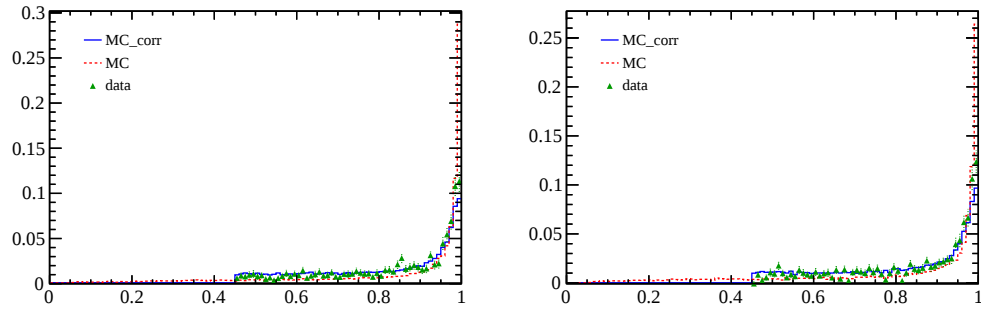


Figure D.6: Comparison of the ProbNNp distributions for (*left*) protons and (*right*) antiprotons, before and after the resampling, between MC and *sWeight*'ed data.

E | FIT VALIDATION WITH PSEUDOEXPERIMENTS

To validate the fit model 1000 pseudoexperiments were generated and fitted with the nominal fit model. Figure [E.1](#) shows the pulls of the yields of B^0 and B_s^0 ; it can be seen that they are compatible with Gaussian functions centered at zero and with standard deviation equal to one. It can be noticed that the statistical uncertainty for $B_s^0 \rightarrow p\bar{p}K\pi$ seems slightly overestimated by the fit. The same procedure has been used for the normalisation mode as shown in Figure [E.2](#).

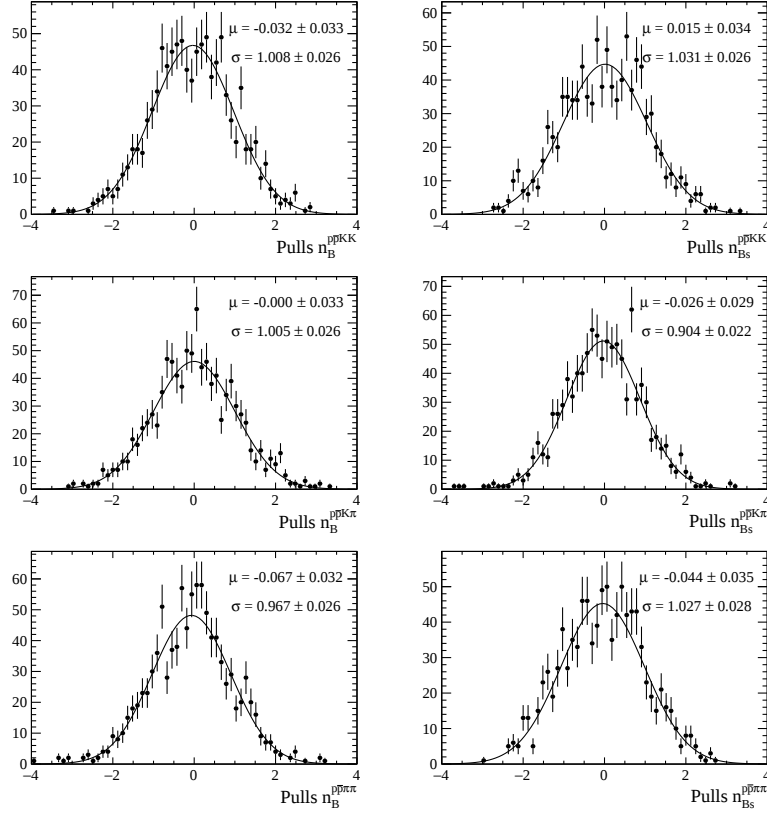


Figure E.1: Pull distributions of the yield parameters obtained generating and fitting pseudoexperiments obtained from the nominal fit model described in Section 5.7. B^0 on the left and B_s^0 on the right; $p\bar{p}KK$ at the top, $p\bar{p}K\pi$ in the middle and $p\bar{p}\pi\pi$ at the bottom.

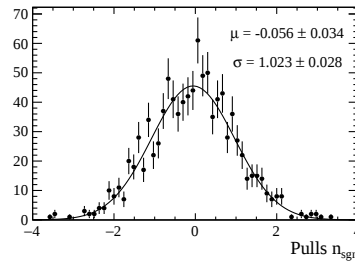


Figure E.2: Pull distributions of the yield parameters obtained generating and fitting pseudoexperiments obtained from the nominal fit model described in Section 5.7 for the normalisation mode $B^0 \rightarrow J/\psi K^{*0}$.

BIBLIOGRAPHY

- [1] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the three-body phase space of charmless $B^\pm$ decays*, *Phys. Rev. D* **90** (2014) 112004, CERN-PH-EP-2014-203, LHCb-PAPER-2014-044, [arXiv:1408.5373](#).
- [2] MissMJ. <https://commons.wikimedia.org/w/index.php?curid=428696>.
- [3] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562, [arXiv:hep-ex/9807003](#).
- [4] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, *Chin. Phys.* **C40** (2016) 100001.
- [5] J. R. Ellis, J. Hisano, M. Raidal, and Y. Shimizu, *A New parametrization of the seesaw mechanism and applications in supersymmetric models*, *Phys. Rev. D* **66** (2002) 115013, [arXiv:hep-ph/0206110](#).
- [6] A. Ilakovac, B. A. Kniehl, and A. Pilaftsis, *Semileptonic lepton-number and/or lepton-flavor-violating τ decays in Majorana neutrino models*, *Phys. Rev. D* **52** (1995) 3993, [arXiv:hep-ph/9503456](#).
- [7] S. F. King, *Neutrino mass*, *Contemp. Phys.* **48** (2007) 195, [arXiv:0712.1750](#).
- [8] W. J. Marciano, T. Mori, and J. M. Roney, *Charged lepton flavor violation experiments*, *Ann. Rev. Nucl. Part. Sci.* **58** (2008) 315.
- [9] BaBar collaboration, B. Aubert *et al.*, *The BaBar detector*, *Nucl. Instrum. Meth.* **A479** (2002) 1, [arXiv:hep-ex/0105044](#).
- [10] Belle collaboration, A. Abashian *et al.*, *The Belle Detector*, *Nucl. Instrum. Meth.* **A479** (2002) 117.
- [11] CLEO collaboration, D. Andrews *et al.*, *The CLEO Detector*, *Nucl. Instrum. Meth.* **211** (1983) 47.

- [12] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2014*, [arXiv:1412.7515](#).
- [13] Heavy Flavor Averaging Group (HFAG), S. Banerjee *et al.*, *HFAG-tau summer 2016 report*, <http://www.slac.stanford.edu/xorg/hfag/tau/summer-2016/>, 2016.
- [14] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pene, *Standard model CP violation and baryon asymmetry*, *Mod. Phys. Lett.* **A9** (1994) 795, [arXiv:hep-ph/9312215](#).
- [15] K. G. Wilson, *The renormalization group and strong interactions*, *Phys. Rev.* **D3** (1971) 1818.
- [16] C.-W. Chiang *et al.*, *Charmless $B \rightarrow VP$ decays using flavor SU(3) symmetry*, *Phys. Rev.* **D69** (2004) 034001, [arXiv:hep-ph/0307395](#).
- [17] S. Weinberg, *Phenomenological lagrangians*, *Physica* **A96** (1979) 327.
- [18] D. B. Kaplan, *Effective field theories*, in *Beyond the standard model 5. Proceedings, 5th Conference, Balholm, Norway, April 29-May 4, 1997*, 1995. [arXiv:nucl-th/9506035](#).
- [19] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for $B \rightarrow \pi\pi$ decays: Strong phases and CP violation in the heavy quark limit*, *Phys. Rev. Lett.* **83** (1999) 1914, [arXiv:hep-ph/9905312](#).
- [20] Y.-Y. Keum, H.-n. Li, and A. I. Sanda, *A novel PQCD approach in charmless B meson decays*, *AIP Conf. Proc.* **618** (2002) 229, [arXiv:hep-ph/0201103](#).
- [21] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An effective field theory for collinear and soft gluons: heavy to light decays*, *Phys. Rev.* **D63** (2001) 114020, [arXiv:hep-ph/0011336](#).
- [22] D. Pirjol, *Theory of hadronic B decays*, *Nucl. Phys. Proc. Suppl.* **156** (2006) 81, [arXiv:hep-ph/0502141](#).
- [23] R. G. Sachs, *The physics of time reversal*, Chicago, Univ. Pr. (1987) 309p, 1987.
- [24] G. Branco, L. Lavoura, and J. Silva, *CP violation*, International series of monographs on physics, Clarendon Press, 1999.

- [25] G. Durieux and Y. Grossman, *Probing CP violation systematically in differential distributions*, *Phys. Rev. D* **92** (2015), no. 7 076013, [arXiv:1508.0305](#).
- [26] ARGUS collaboration, H. Albrecht *et al.*, *Measurement of inclusive B meson decays into baryons*, *Z. Phys.* **C42** (1989) 519.
- [27] Belle collaboration, K. Abe *et al.*, *Observation of $B^\pm \rightarrow p\bar{p}K^\pm$* , *Phys. Rev. Lett.* **88** (2002) 181803, [arXiv:hep-ex/0202017](#).
- [28] Belle collaboration, M. Z. Wang *et al.*, *Observation of $B^+ \rightarrow p\bar{p}\pi^+$, $B^0 \rightarrow p\bar{p}K^0$, and $B^+ \rightarrow p\bar{p}K^{*++}$* , *Phys. Rev. Lett.* **92** (2004) 131801, [arXiv:hep-ex/0310018](#).
- [29] LHCb collaboration, R. Aaij *et al.*, *First observation of a baryonic B_c^+ decay*, *Phys. Rev. Lett.* **113** (2014) 152003 CERN-PH-EP-2014-187, LHCb-PAPER-2014-039, [arXiv:1408.0971](#).
- [30] LHCb collaboration, E. Rodrigues, *First observation of a baryonic B_s^0 decay*, LHCb-CONF-2016-016, presentation at the CKM2016 International Workshop.
- [31] LHCb collaboration, R. Aaij *et al.*, *Studies of the decays $B^+ \rightarrow p\bar{p}h^+$ and observation of $B^+ \rightarrow \bar{\Lambda}(1520)p$* , *Phys. Rev. D* **88** (2013) 052015 CERN-PH-EP-2013-125, LHCb-PAPER-2013-031, [arXiv:1307.6165](#).
- [32] LHCb collaboration, R. Aaij *et al.*, *First evidence for the two-body charmless baryonic decay $B^0 \rightarrow p\bar{p}$* , *JHEP* **10** (2013) 005 LHCb-PAPER-2013-038, CERN-PH-EP-2013-138, [arXiv:1308.0961](#).
- [33] LHCb collaboration, R. Aaij *et al.*, *Evidence for the two-body charmless baryonic decay $B^+ \rightarrow p\bar{\Lambda}$* , [arXiv:1611.0780](#) LHCb-PAPER-2016-048, CERN-EP-2016-275, [arXiv:1611.0780](#), submitted to JHEP.
- [34] Belle collaboration, K. Abe *et al.*, *Observation of $\bar{B}^0 \rightarrow D^{0(*)}p\bar{p}$* , *Phys. Rev. Lett.* **89** (2002) 151802, [arXiv:hep-ex/0205083](#).
- [35] Belle and BaBar collaborations, A. J. Bevan *et al.*, *The physics of the B factories*, *Eur. Phys. J.* **C74** (2014) 3026, [arXiv:1406.6311](#).
- [36] Belle collaboration, M. Z. Wang *et al.*, *Study of the baryon-antibaryon low-mass enhancements in charmless three-body baryonic B decays*, *Phys. Lett. B* **617** (2005) 141, [arXiv:hep-ex/0503047](#).

- [37] Belle collaboration, Y. J. Lee *et al.*, *Observation of $B^+ \rightarrow \Lambda \bar{\Lambda} K^+$* , *Phys. Rev. Lett.* **93** (2004) 211801, [arXiv:hep-ex/0406068](#).
- [38] BES collaboration, J. Z. Bai *et al.*, *Observation of a near threshold enhancement in the p anti- p mass spectrum from radiative $J/\Psi \rightarrow \gamma p \bar{p}$ decays*, *Phys. Rev. Lett.* **91** (2003) 022001, [arXiv:hep-ex/0303006](#).
- [39] CLEO collaboration, S. B. Athar *et al.*, *Radiative decays of the $\Upsilon(1S)$ to a pair of charged hadrons*, *Phys. Rev.* **D73** (2006) 032001, [arXiv:hep-ex/0510015](#).
- [40] M. Suzuki, *Partial waves of baryon-antibaryon in three-body B meson decay*, *J. Phys.* **G34** (2007) 283, [arXiv:hep-ph/0609133](#).
- [41] Belle collaboration, J. T. Wei *et al.*, *Study of $B^+ \rightarrow p \bar{p} K^+$ and $B^+ \rightarrow p \bar{p} \pi^+$* , *Phys. Lett.* **B659** (2008) 80, [arXiv:0706.4167](#).
- [42] W.-S. Hou and A. Soni, *Pathways to rare baryonic B decays*, *Phys. Rev. Lett.* **86** (2001) 4247, [arXiv:hep-ph/0008079](#).
- [43] H.-Y. Cheng and K.-C. Yang, *Charmless exclusive baryonic B decays*, *Phys. Rev.* **D66** (2002) 014020, [arXiv:hep-ph/0112245](#).
- [44] C.-K. Chua, W.-S. Hou, and S.-Y. Tsai, *Charmless three-body baryonic B decays*, *Phys. Rev.* **D66** (2002) 054004, [arXiv:hep-ph/0204185](#).
- [45] J. P. Dedonder, B. Loiseau, B. El-Bennich, and S. Wycech, *On the structure of the $X(1835)$ baryonium*, *Phys. Rev.* **C80** (2009) 045207, [arXiv:0904.2163](#).
- [46] J. L. Rosner, *Low mass baryon anti-baryon enhancements in B decays*, *Phys. Rev.* **D68** (2003) 014004, [arXiv:hep-ph/0303079](#).
- [47] C. Q. Geng and Y. K. Hsiao, *Angular distributions in three-body baryonic B decays*, *Phys. Rev.* **D74** (2006) 094023, [arXiv:hep-ph/0606141](#).
- [48] O. S. Bruning *et al.*, *LHC design report. 1. The LHC main ring*, CERN-2004-003-V-1, CERN-2004-003 (2004).
- [49] O. Buning *et al.*, *LHC design report. 2. The LHC infrastructure and general services*, CERN-2004-003-V-2, CERN-2004-003 (2004).
- [50] M. Benedikt *et al.*, *LHC design report. 3. The LHC injector chain*, CERN-2004-003-V-3, CERN-2004-003 (2004).

- [51] S. Myers, *The LEP collider, from design to approval and commissioning*, 1991.
- [52] <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/lhc-schematic.jpg>.
- [53] F. Marcastel, *CERN's Accelerator Complex. La chaîne des accélérateurs du CERN*, OPEN-PHO-CHART-2013-001, <https://cds.cern.ch/record/1621583> General Photo.
- [54] <http://aliceinfo.cern.ch/Public/Objects/Chapter2/ALICE-SetUp-NewSimple.jpg>.
- [55] <https://atlas.cern/resources/multimedia/detector>.
- [56] <http://cms-project-ecal-p5.web.cern.ch/cms-project-ECAL-P5/approved/detector.htm>.
- [57] <http://lhcb-public.web.cern.ch/lhcb-public/en/lhcb-outreach/multimedia/>.
- [58] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, **JINST** **3** (2008) So8002.
- [59] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, **JINST** **3** (2008) So8003.
- [60] CMS collaboration, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, **JINST** **3** (2008) So8004.
- [61] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, **JINST** **3** (2008) So8005.
- [62] TOTEM, G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, **JINST** **3** (2008) So8007.
- [63] LHCf, O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, **JINST** **3** (2008) So8006.
- [64] MoEDAL collaboration, B. Acharya *et al.*, *The Physics Programme Of The MoEDAL Experiment At The LHC*, **Int. J. Mod. Phys. A** **29** (2014) 1430050, [arXiv:1405.7662](https://arxiv.org/abs/1405.7662).
- [65] <https://home.cern/cern-people/updates/2016/12/lhc-report-far-beyond-expectations>.

- [66] https://indico.cern.ch/event/563813/contributions/2278462/attachments/1385965/2109084/161210_Plenary_LHCbWeek_RunCoordination_v2.pdf.
- [67] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022, [arXiv:1412.6352](#).
- [68] J. Harrison, *Radiation damage studies at the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, Univ. Manchester, 2014.
- [69] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J. C* **73** (2013) 2431, [arXiv:1211.6759](#).
- [70] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [71] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, [arXiv:0710.3820](#).
- [72] I. Belyaev *et al.*, *Handling of the generation of primary events in GAUSS, the LHCb simulation framework*, *Nuclear Science Symposium Conference Record (NSS/MIC) IEEE* (2010) 1155.
- [73] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.
- [74] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur. Phys. J. C* **45** (2006) 97, [arXiv:hep-ph/0506026](#).
- [75] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [76] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, *IEEE Trans. Nucl. Sci.* **53** (2006) 270.
- [77] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, *JINST* **10** (2015) P02007, [arXiv:1408.1251](#).
- [78] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444.

- [79] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, **J. Phys. Conf. Ser.** **664** (2015), no. 8 082010.
- [80] C. F. Hombach, *Search for the rare baryonic $B^+ \rightarrow p\bar{\Lambda}$ decay with the LHCb detector and alignment of pixel detectors*, PhD thesis, The University of Manchester, 2016.
- [81] M. Gersabeck, C. Parkes, and S. Viret, *LHCb VELO software alignment. Part III. The Alignment of the relative sensor positions*, **arXiv:0807.3532**.
- [82] V. Blobel and C. Kleinwort, *A New method for the high precision alignment of track detectors*, in *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*, pp. URL-STR(9), 2002. **arXiv:hep-ex/0208021**.
- [83] W. Hulsbergen, *The Global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, **Nucl. Instrum. Meth. A** **600** (2009) 471, **arXiv:0810.2241**.
- [84] J. Amoraal *et al.*, *Application of vertex and mass constraints in track-based alignment*, **Nucl. Instrum. Meth. A** **712** (2013) 48, **arXiv:1207.4756**.
- [85] R. Aaij *et al.*, *Tesla : an application for real-time data analysis in high energy physics*, **Comput. Phys. Commun.** **208** (2016) 35, **arXiv:1604.0559**.
- [86] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, **JHEP** **10** (2015) 172 LHCb-PAPER-2015-037, CERN-PH-EP-2015-222, **arXiv:1509.0077**.
- [87] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, **JHEP** **03** (2016) 159 LHCb-PAPER-2015-041, CERN-PH-EP-2015-272, **arXiv:1510.0170**.
- [88] I. Sobel, *An isotropic 3 x 3 image gradient operator*, Machine vision for three-dimensional scenes (1990) 376.
- [89] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, **JINST** **9** (2014) P09007, **arXiv:1405.7808**.

- [90] G. Dujany, P. Poier, S. Borghi, and C. Parkes, *A method for monitoring the stability of the alignment of the LHCb VELO*, LHCb internal note (2012).
- [91] <https://twiki.cern.ch/twiki/bin/view/LHCb/ConferencePlots>.
- [92] Belle Collaboration, Y. Miyazaki *et al.*, *Search for lepton-flavor-violating tau decays into a lepton and a vector meson*, **Phys. Lett. B** **699** (2011) 251, [arXiv:1101.0755](#).
- [93] LHCb collaboration, R. Aaij *et al.*, *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, **Phys. Lett. B** **724** (2013) 36, [arXiv:1304.4518](#).
- [94] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton flavour violating decay $\tau^- \rightarrow \mu^- \mu^+ \mu^-$* , **JHEP** **02** (2015) 121, [arXiv:1409.8548](#).
- [95] CLEO Collaboration, D. Bliss *et al.*, *New limits for neutrinoless tau decays*, **Phys. Rev. D** **57** (1998) 5903, [arXiv:hep-ex/9712010](#).
- [96] BaBar Collaboration, B. Aubert *et al.*, *Improved limits on lepton flavor violating tau decays to $l\phi$, $l\rho$, lK^* and $l\bar{K}^*$* , **Phys. Rev. Lett.** **103** (2009) 021801, [arXiv:0904.0339](#).
- [97] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, [DESY-F31-86-02](#).
- [98] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [99] S. Haykin, *Neural Networks and Learning Machines*, 3rd edn, Prentice Hall, 2009.
- [100] BaBar collaboration, B. Aubert *et al.*, *Evidence for the $B^0 \rightarrow p\bar{p}K^{*0}$ and $B^+ \rightarrow \eta_c K^{*+}$ decays and study of the decay dynamics of B meson decays into $p\bar{p}h$ final states*, **Phys. Rev. D** **76** (2007) 092004, [arXiv:0707.1648](#).
- [101] Belle collaboration, J. H. Chen *et al.*, *Observation of $B^0 \rightarrow p\bar{p}K^{*0}$ with a large K^{*0} polarization*, **Phys. Rev. Lett.** **100** (2008) 251801, [arXiv:0802.0336](#).

- [102] CLEO collaboration, C. Bebek *et al.*, *Search for the charmless decays $B \rightarrow p\bar{p}\pi$ and $p\bar{p}\pi\pi$* , **Phys. Rev. Lett.** **62** (1989) 8.
- [103] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, **JHEP** **04** (2013) 001, [arXiv:1301.5286](#), f_s/f_d value updated in LHCb-CONF-2013-011.
- [104] Belle, K. Chilikin *et al.*, *Observation of a new charged charmonium-like state in $\bar{B}^0 \rightarrow J/\psi K^- \pi^+$ decays*, **Phys. Rev.** **D90** (2014), no. 11 112009, [arXiv:1408.6457](#).
- [105] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, **JINST** **8** (2013) P04022, [arXiv:1211.3055](#).
- [106] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, **JINST** **8** (2013) P02013, [arXiv:1210.6861](#).
- [107] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, *The toolkit for multivariate data analysis, TMVA 4*, **J. Phys. Conf. Ser.** **219** (2010) 032057.
- [108] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, **Nucl. Instrum. Meth.** **A555** (2005) 356, [arXiv:physics/0402083](#).
- [109] LHCb collaboration, R. Aaij *et al.*, *Probing matter-antimatter asymmetries in beauty baryon decays*, [arXiv:1609.0521](#) LHCb-PAPER-2016-030, CERN-EP-2016-212, [arXiv:1609.0521](#), [arXiv:1609.05216 published online in Nature Physics](#).
- [110] Geng, Chao-Qiang and Hsiao, Yu-Kuo. Private communication.
- [111] LHCb collaboration, R. Aaij *et al.*, *Searches for $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ and $B^+ \rightarrow J/\psi p\bar{p}\pi^+$ decays*, **JHEP** **09** (2013) 006 LHCb-PAPER-2013-029, CERN-PH-EP-2013-099, [arXiv:1306.4489](#).
- [112] F. E. James, *Statistical methods in experimental physics; 2nd ed.*, World Scientific, Singapore, 2006.
- [113] D. Aston *et al.*, *A study of $K^-\pi^+$ scattering in the reaction $K^-p \rightarrow K^-\pi^+n$ at 11 GeV/c*, **Nucl. Phys.** **B296** (1988) 493.
- [114] BaBar collaboration, B. Aubert *et al.*, *Time-dependent and time-integrated angular analysis of $B \rightarrow \phi K_s^0 \pi^0$ and $\phi K^\pm \pi^\mp$* , **Phys. Rev.** **D78** (2008) 092008, [arXiv:0808.3586](#).

- [115] LHCb collaboration, R. Aaij *et al.*, *First observation of the decay* $B_s^0 \rightarrow \phi \tilde{K}^{*0}$, **JHEP** **11** (2013) 092, [arXiv:1306.2239](#).
- [116] J. Beddow *et al.*, *Search for the Rare two-body charmless baryonic decay* $B^+ \rightarrow p \bar{\Lambda}$, **LHCb-ANA-2014-094**.
- [117] S. Malde *et al.*, *PIDCalib packages*, <https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage>.
- [118] LHCb collaboration, R. Aaij *et al.*, *Observation of charmless baryonic decays* $B_{(s)}^0 \rightarrow p \bar{p} h^+ h'^-$, **LHCb-PAPER-2017-005**, CERN-EP-2017-052, [arXiv:1704.08497](#) submitted to Phys. Rev. Lett.
- [119] J. Albrecht *et al.*, *Search for the lepton flavour violating decay* $\tau^- \rightarrow \mu^+ \mu^- \mu^-$, **LHCb-ANA-2014-005**.
- [120] J. Harrison, *The inclusive τ cross-section at LHCb at $\sqrt{s} = 7$ TeV*, **LHCb-INT-2011-041**.