

DISSERTATION

submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of

Doctor of Natural Sciences

Put forward by

Marianna Fontana

born in: Sorgono, Italy

Oral examination: 17 December 2014

Search for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays
at LHCb

Referees:

Prof. Werner Hofmann

Prof. Stephanie Hansmann-Menzemer

A mamma e babbo

ACKNOWLEDGMENTS

I would like to thank Prof. Michael Schmelling who gave me the possibility to do my PhD in this group and the International Max Planck Society for the fundings. I also thank the LHCb group at MPI, for the nice atmosphere, for our coffees, our Friday meetings and for not opening the windows during winter!

A very warm thanks goes to the K_s^{*0} microWG, and especially to Stephan Monteil and Thomas Latham: this work would not have been possible without your precious help and guidance.

And now I would like to thank all the people who supported me through these years. Francesca you are the first one, because I shared with you the feeling to be exile, the difficulties to live in a different country and the happiness to go back home for our holidays. Matteo for being such a bright and good physicist. Matteo and Francesca together for our Flamkuchen, our sushies and for the piece of Sardinia always here with me.

Jelenina, moja sestrica, you know more than others what this period have meant for us. Thanks for being always with me, for enjoying Stockholm and Heidelberg together, for our adventures in Machu Picchu, for our past and future ziveli, and for our Bella Ciao.

Dorothea, sorella, compagna, when I left home you wrote me "There will be a moment in which we will be together again", I could not imagine that moment was arriving so fast. You add another big piece of Sardinia here and made the last year so wonderful. Thanks for our tisanine, our salad and eggs dinners, our thai place, for our tears and difficult moments always supporting each other...I am still waiting that you write a book together with Valerio. Ah trucidato che te devo di? You are so brilliant, and you should remind to yourself more often. A part your qualities as a physicist and your innate attitude to art, you have been a precious and irreplaceable friend, always ready to make me laugh and sustain. You really deserve linguine alla bottarga. Ah and thanks for the help with the systematics!

Zizzo, we should regain the biers we lost in these years. Thanks for being always present and for the enormous help you gave me. I did not lose the hope in our red future.

Andre thanks a lot to you, because you will remain my tutor for ever probably.

A big thanks to all my friends, those who came to visit me and the others always ready to hug me when coming back home: Paola, Silvia, Fra Loi and Fabri, Charlettino and Giuli, Tina, Rob, Albe, Sonjetta, Richi and Ale, Babi, Moni.

Ringrazio mio fratello Nicola e sua moglie Jessica, per essermi stati vicini, per il tempo trascorso assieme ad Heidelberg e in Sardegna. E a te piccolo Daniele, che sei arrivato proprio alla fine di questo percorso e mi hai riempito di gioia. Quando leggerai questi ringraziamenti saprai perchè purtroppo non sono potuta essere lì quando sei nato.

A mamma e babbo, non vi ringrazierò mai abbastanza per l'amore con cui mi avete cresciuto. Per avermi dato la possibilità di istruirmi, per le enormi possibilità che mi avete sempre concesso e senza le quali ora non sarei qui e non avrei potuto fare questa esperienza.

A te Antonio, forse senza questo percorso non ti avrei mai conosciuto. Non credo ce l'avrei fatta senza di te in questi ultimi anni. E' stata dura stare lontano, fare dei viaggi infiniti per stare assieme anche un solo giorno. Non ho bisogno di elencare i motivi per cui ti ringrazio, sappi solo che rifarei ogni singola cosa assieme a te. Tranquillo l'anno prossimo le vacanze in Sardegna te le faccio fare :)

Heidelberg, November 2014

Marianna

ABSTRACT

The search for the previously unobserved $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays is performed at the LHCb experiment, using proton-proton collisions at a centre of mass energy of $\sqrt{s} = 7$ TeV. These decays provide a fertile ground in the search of New Physics in CP violation measurements. The data sample was recorded in 2011 and corresponds to an integrated luminosity of 1 fb^{-1} . The first observation of the $B_s^0 \rightarrow K^{*0} K_S^0$ decay is reported. A measurement of the limit and the branching fraction relative to the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ decay yields the following results:

$$\begin{aligned} \frac{\mathcal{B}(B^0 \rightarrow K^{*0} K_S^0)}{\mathcal{B}(B^0 \rightarrow K_S^0 \pi^+ \pi^-)} &< 0.12(0.15) \text{ at } 90\% \text{ (95\%) CL} \\ \frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_S^0)}{\mathcal{B}(B^0 \rightarrow K_S^0 \pi^+ \pi^-)} &= 0.22 \pm 0.06(\text{stat}) \pm 0.02(\text{syst}) \pm 0.01(f_s/f_d), \end{aligned}$$

where the first uncertainty is statistical, the second systematic and the third comes from the uncertainty on the hadronisation fraction.

ZUSAMMENFASSUNG

Eine Suche nach den bisher noch nicht entdeckten Zerfällen $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ wird am LHCb-Experiment mit Proton-Proton Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 7$ TeV durchgeführt. Diese Zerfälle stellen vielversprechende Kandidaten für die Suche nach Neuer Physik in CP -Verletzungsmessungen dar. Die verwendeten Daten wurden 2011 aufgenommen und entsprechen einer integrierten Luminosität von 1 fb^{-1} . Der Zerfallskanal $B_s^0 \rightarrow K^{*0} K_S^0$ wird erstmalig beobachtet. Für das Verzweigungsverhältnis relativ zum Zerfall $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ liefert die Analyse

$$\begin{aligned} \frac{\mathcal{B}(B^0 \rightarrow K^{*0} K_S^0)}{\mathcal{B}(B^0 \rightarrow K_S^0 \pi^+ \pi^-)} &< 0.12(0.15) \text{ at } 90\% \text{ (95\%) CL} \\ \frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_S^0)}{\mathcal{B}(B^0 \rightarrow K_S^0 \pi^+ \pi^-)} &= 0.22 \pm 0.06(\text{stat}) \pm 0.02(\text{syst}) \pm 0.01(f_s/f_d). \end{aligned}$$

Der erste Fehler ist die statische, der zweite ist die systematische Unsicherheit der Messung, und der dritte resultiert aus der Unsicherheit auf das Hadronisierungsverhältnis.

CONTENTS

1	INTRODUCTION	1
2	THEORETICAL OVERVIEW	5
2.1	The Standard Model of Particle Physics	5
2.2	The CKM mechanism	8
2.3	CP violation	11
2.4	Physics of $b \rightarrow s$ hadronic penguin decays	14
2.4.1	The $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decays	17
3	THE LHCb EXPERIMENT	21
3.1	The Large Hadron Collider	21
3.2	The LHCb experiment and detector	25
3.2.1	Detector design	25
3.2.2	The tracking system	27
3.2.3	Particle Identification	34
3.2.4	Trigger	41
3.2.5	Simulation in LHCb	44
3.2.6	Reconstruction and analysis	45
4	ANALYSIS STRATEGY AND SELECTION	47
4.1	Analysis strategy	47
4.2	Data samples	48
4.3	Background studies	50
4.3.1	Peaking backgrounds	50
4.3.2	Partially reconstructed decays	51
4.3.3	Misidentified background	53
4.3.4	Combinatorial background	54
4.4	Selection	54
4.4.1	Variable definitions	55
4.4.2	Multivariate analysis	57
4.4.3	Trigger strategy	58
4.4.4	Preselection	61
4.4.5	Fitting ranges	66

4.4.6	Boosted decision tree	67
5	FIT MODEL	77
5.1	Maximum likelihood fit method	77
5.1.1	Extended fits	78
5.1.2	Constraints on parameters	78
5.1.3	Implementation	79
5.2	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ model	79
5.2.1	Signal model	79
5.2.2	Non resonant $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$ background model	82
5.2.3	Misidentified-signal model	85
5.2.4	Partially reconstructed background model	88
5.2.5	Combinatorial background model	88
5.2.6	Summary of the fit model	91
5.2.7	Simulation studies	93
5.3	$B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$ model	96
5.3.1	Signal model	96
5.3.2	Misidentified signal model	96
5.3.3	Partially reconstructed background model	99
5.3.4	Combinatorial background model	101
5.3.5	Summary of the fit model	101
5.4	Fit results on data	103
5.4.1	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ fit results	103
5.4.2	$B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$ fit results	105
6	EFFICIENCIES	107
6.1	General considerations	107
6.2	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ efficiencies	108
6.2.1	Detector acceptance	108
6.2.2	Reconstruction and stripping efficiencies	108
6.2.3	Trigger efficiencies	109
6.2.4	Combined offline selection	111
6.2.5	PID efficiencies	111
6.2.6	Summary of the efficiencies	112
6.3	$B^0 \rightarrow K_s^0 \pi^+ \pi^-$ efficiencies	113
7	SYSTEMATICS	117
7.1	Event selection	117
7.1.1	Trigger requirements	117

7.1.2	Binning scheme for the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ efficiencies	118
7.1.3	PID requirements	120
7.2	Fit model	121
7.2.1	Fixed shape parameters	121
7.2.2	Shape model	122
7.2.3	Fit model: S-wave interference	123
7.2.4	Fit bias	126
7.3	Summary	128
8	RESULTS	131
8.1	Observation significance	131
8.2	Upper limit	133
8.3	Branching ratio measurement	134
9	SUMMARY AND CONCLUSIONS	137
	Appendix	139
A	DALITZ PLOT	139
B	MC DATA COMPARISON	141
	Bibliography	143
	List of Figures	149
	List of Tables	153

1 | INTRODUCTION

The *Standard Model* of Particle Physics (SM) is the theory which best describes the fundamental particles and their interactions, apart from gravity. However though being the most successful theory to date, the SM is inherently an incomplete theory [1]. There are fundamental physics phenomena in nature that deviate from the predictions of the SM. From cosmological observations we know that the SM accounts for only 4% of the matter present in the universe. Of the missing 96%, about 27% should be dark matter [2], weakly-interacting matter for which the SM does not supply any fundamental particles that are good candidates for it. Moreover the nature of neutrinos, which are proven experimentally to have mass [3,4], is not consistent with the minimal formulation of the SM and the mass terms have been added a posteriori to it. The universe is made out of mostly matter, and CP violation is a fundamental ingredient to explain the observed matter-antimatter asymmetry [5]. However the observed baryon asymmetry is too large to be generated by the SM mechanism. This discrepancy suggests additional sources of CP violation beyond the SM.

From the theoretical side there are some conceptual problems in the SM, like the presence of many free parameters, the puzzles of flavour hierarchies, the fact that gravity is not included in theory and the related hierarchy problem [6]. The evolution with the energy of the effective gauge couplings leads to an electro-weak and strong forces unification at a scale of $10^{15} - 10^{16}$ GeV. This value is close to the Planck scale of quantum gravity $\sim 10^{19}$ GeV, so it is possible to imagine a unified theory of all interactions, including gravity. Then the question arises whether the SM is valid up to such energies. Some problems like the explanation of the three generations of fermions and their masses can be postponed to a more fundamental theory. However there are other problems which have to be solved in the low-energy theory. In particular the SM has no solution for the hierarchy problem, meaning the relative smallness of the weak scale of mass, set by the Higgs mass (~ 125 GeV). This problem is related to the instability of the SM with respect to quantum corrections due to the presence of fundamental scalar fields with quadratic divergences. In this respect the infinities are not a problem since the theory is renormalizable, but the hierarchy problem is one of naturalness.

Radiative corrections to the Higgs mass amount to a sum of different terms, multiplied by the square of the cut-off energy of the theory. If that energy is as large as the Planck scale, then there are several correction terms which cancel out in an unexplained and accurate way, to give the actual small mass of the Higgs. These experimental and theoretical problems motivate the search for physics beyond the SM.

This *New Physics* (NP) can be searched either by increasing the available energy at colliders like the LHC at CERN to produce new particles and reveal them directly or by increasing the experimental precision on certain processes involving SM particles [7]. This latter indirect search for NP may be pursued using decays that are forbidden, very rare, or precisely calculable in the SM. Flavour-changing neutral current (FCNC) and CP violating processes are among the most powerful probes of NP. In fact in the SM FCNC cannot arise at tree level, and even at loop level they are strongly suppressed by the GIM mechanism [8]. Once CKM matrix elements are precisely determined, it is possible to search for NP contributions. In particular, penguin-dominated nonleptonic B decays can reveal the presence of NP in decay amplitudes. Thanks to the B -factories, CP violation has been measured with a high accuracy in several $b \rightarrow s$ penguin dominated channels. With LHC, if new particles are revealed, these processes can help in the identification of the flavour structure of the NP model. If no new particles are seen, still $b \rightarrow s$ transitions can indirectly reveal the presence of NP or at least push up the lower bound on the NP scale. However, to test the SM, it is necessary to know with high precision the parameters describing the hadronic interactions, that are difficult to determine theoretically as well as from data. Many efforts have been made to develop methods that allow to better control these parameters. The experimental inputs needed for these methods are the magnitudes and relative phases of the decay amplitudes. A common approach to determine the relative phases is to perform an amplitude analysis by studying the structure of a Dalitz plot. Though, this technique requires an high statistics data sample. A simplified approach instead can provide branching fraction measurements using a much smaller data sample. In all cases the progress in this kind of processes is crucial for our understanding of flavour physics.

The LHCb experiment is one of the main experiments at the LHC. It has been designed mainly for the study of b-hadron physics, by precise measurements of CP violation and the study of rare decays. The LHCb collaboration provided first results on inclusive $B_{d,s}^0 \rightarrow K_s^0 h^\pm h'^\mp$ ($h = \pi, K$) decays [9], and just recently the first measurements of B_s^0 meson decays to $K^{*-}\pi^+$ and $K^{*\pm}K^\mp$ final states have been published [10]. The analysis reported in this thesis was developed in the same

context.

During my PhD I was involved in the micro group at CERN, working on the inclusive measurement, just mentioned. Following the simplified approach, this work analysed the same data, looking specifically at the resonant structure in the low $K\pi$ invariant mass region and measuring the branching fractions of the decays which proceed through an intermediate $K^*(892)^0$ resonance. Some tools developed for the former analysis were used also for the results reported in this thesis and the publication is referenced to when studies are presented that have not been performed by the author. To document the work for the LHCb collaboration a so-called analysis note was written

- M.Fontana, "Search for the $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decays at LHCb", LHCb analysis note (LHCb-ANA-2014-012), 2014.

The analysis is currently under review and a subsequent paper is in preparation, for which I will be the contact author.

The material in this thesis is organised as follows. In the first part a theoretical introduction on the standard model of particle physics and the current status of the predictions for the relevant branching fractions is given (Chapter 2), followed by the description of the LHCb detector (Chapter 3). The second part is devoted to the experimental results on the $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decays. The background studies and the selection criteria are detailed in Chapter 4. In Chapter 5 the development and the validation of the fit to extract the yields is explained. Chapter 6 is dedicated to the measurement of the efficiencies, while in Chapter 7 the study of the systematic uncertainties are reported. The main results of the analysis, that is the branching fraction measurement of the decay, is presented in Chapter 8. At the end some implications of the results and prospects for further studies are given (Chapter 9).

2 | THEORETICAL OVERVIEW

This chapter details the theoretical background of the analysis presented in this thesis. First the Standard Model of particle physics is shortly introduced, followed by a description of the quark mixing mechanism and a summary of the different sources of CP violation. In addition the interest in the charmless B decays object of this analysis, the theoretical predictions and the experimental status for their branching ratios is given.

2.1 THE STANDARD MODEL OF PARTICLE PHYSICS

Practically all the experimental results from high energy physics experiments can be explained by the so-called *Standard Model* (SM) of particles and their interactions, formulated in the 1970s [11–14]. According to this theory, three families of fundamental spin 1/2 particles, called *fermions* constitute matter. There are six *quarks* and six *leptons*. The main properties of these particles are summarised in Tab. 1. The *leptons* carry integral electric charge. The *electron* e carries a unit negative charge by convention and the other charged leptons are the *muon* μ and the *tauon* τ , which are heavy versions of the electron. The neutral leptons are called *neutrinos* ν and each flavour of neutrino is paired with the corresponding flavour of a charged lepton. The *quarks* carry fractional charges, of $+2/3$ or $-1/3$. As for the leptons, the quarks are grouped into pairs differing by one unit of electric charge. The quark flavour is denoted by a symbol: u (*up*), d (*down*), s (*strange*), c (*charmed*), b (*beauty*), t (*top*). Quarks can form composite particles, mesons (quark and anti-quark) and baryons (three quarks). Quarks and charged leptons are massive, contrary to the neutrinos which are assumed to be massless in the SM. However the discovery of neutrino oscillations [3,4] leads to the conclusion that neutrinos have very small but non-zero masses. For each quark and lepton there exists an antiparticle with the same mass but opposite quantum numbers.

The SM comprises the *electromagnetic*, the *weak* and the *strong* interactions between particles, described in terms of the exchange of vector *bosons* of integer spin. The *gravitational* force does not play a role in particle physics experiments due to

its weakness and currently it is not included in the SM. A summary of the bosons can be found in Tab. 2. The mediators of the electromagnetic and the strong force, *photons* and *gluons*, are massless while the bosons carrying the weak force, the W^\pm and Z^0 bosons, are massive. There is in addition another boson, called Higgs (H^0), which is related to the appearance of mass in the SM. A candidate was discovered by the ATLAS and CMS collaboration, and further measurements favour the 0-spin hypothesis [15, 16].

Table 1.: Properties of the fundamental fermions in the Standard Model [17].

quarks				leptons		
generation	type	el. charge	mass	type	el. charge	mass
1	u	+2/3	$2.3^{+0.7}_{-0.5}$ GeV/ c^2	ν_e	0	< 2 eV/ c^2
	d	-1/3	$4.8^{+0.7}_{-0.3}$ GeV/ c^2	e	-1	0.511 GeV/ c^2
2	c	+2/3	1.275 ± 0.025 GeV/ c^2	ν_μ	0	< 2 eV/ c^2
	s	-1/3	95 ± 5 MeV/ c^2	μ	-1	105.7 GeV/ c^2
3	t	+2/3	$173.07 \pm 0.52 \pm 0.72$ GeV/ c^2	ν_τ	0	< 2 eV/ c^2
	b	-1/3	4.18 ± 0.03 GeV/ c^2	τ	-1	1776.82 ± 0.16 GeV/ c^2

Table 2.: Gauge bosons and Higgs particle in the Standard Model [17].

interaction	gauge boson	mass	spin
electromagnetic	photon (γ)	0	1
strong	gluon (g)	0	1
	W^\pm	80.4 GeV/ c^2	1
weak	Z^0	91.2 GeV/ c^2	1
	-	Higgs (H)	125.9 GeV/ c^2

The SM is a Relativistic Quantum Field Theory (RQFT) invariant under the gauge transformation of the combined $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ group. For each interaction the invariance of the Lagrangian under the respective transformation leads to the conservation of a quantum number, denoted in the above formula by the subscripts [18].

The $SU(3)_C$ gauge symmetry group of the Quantum ChromoDynamics (QCD) is a non abelian group and describes the strong interaction among quarks. The

conserved quantum number is called *colour* (C) and it can appear in three different flavours (labeled red, green, blue and the corresponding anti-colours). The mediators of the strong interaction are called *gluons*. They are massless and they carry colour themselves so they can self-interact, with triple and quadruple gluon vertices. Among the most important characteristics of QCD are *asymptotic freedom* and *confinement*. This means that interactions between quarks become weaker with decreasing distance and that free quarks cannot be seen in nature since they always occur in a bound state.

The electromagnetic and the weak force can be unified into the electroweak interaction. This unified interaction can be described by the $SU(2)_L \otimes U(1)_Y$ symmetry group. Associated to the weak $SU(2)_L$ group there are three gauge bosons W_i ($i = 1, 2, 3$) and the associated quantum number is the weak isospin T . An important feature of the theory is *chirality*: this implies that only left-handed fermions and right-handed anti-fermions have $T \neq 0$ and couple to the gauge bosons. The conserved quantum number of the $U(1)_Y$ group is the *hypercharge* Y and the mediator is a gauge boson A . The hypercharge of a fermion is given by $Y = Q - T_3$, where Q is the electric charge and T_3 is the third component of the weak isospin. The exchange bosons of the electromagnetic and weak interactions are given by linear combinations of the gauge bosons W_i and A . The boson which mediates the electromagnetic interaction, the *photon*, is a combination of the W_3 and A gauge bosons. It couples to the electric charge Q of the fermions. All fermions except the neutrinos have an electric charge and can interact electromagnetically. There are two ways to classify the weak interaction processes: the *neutral current* (NC) is carried by the Z^0 boson. This is again a combination of the W_3 and A gauge bosons. The *charged current* (CC) is mediated by the charged W^\pm bosons that are linear combinations of the W_1 and W_2 gauge bosons. In the SM it is the only process where fermions of different generations can participate. Flavour Changing Neutral Currents (FCNC) are suppressed in the SM at tree-level, due to the Glashow-Iliopoulos-Maiani (GIM) mechanism [8], and can only occur in so-called loop processes.

The W_i and A bosons are massless, while the bosons of the weak interaction observed experimentally are massive. Separate mass terms in the Standard Model Lagrangian would violate the invariance of the Lagrangian under continuous local transformations [19]. The introduction of these terms is achieved by the spontaneous symmetry breaking mechanism, which causes the Lagrangian to remain invariant. The simplest way of doing this is to introduce a potential, consisting of a doublet of two complex scalar fields, which gives rise to a non-zero vacuum expectation value. This theory predicts a massive spin-0 particle, the Higgs bo-

son H^0 . The Yukawa couplings of the fermions to the Higgs field determine the fermion masses. See [19] for a more detailed review of the Higgs mechanism. In the quark sector the resulting mass eigenstates are not the same as the eigenstates of the weak interaction. The weak eigenstates can be expressed as a superposition of the mass eigenstates. This phenomenon is described by the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [20]. A more detailed description of the CKM matrix will be given in the next section.

2.2 THE CKM MECHANISM

The weak interaction has the unique feature to allow change of the quark family in transitions between down-type quarks and up-type quarks. This phenomenon is called *quark mixing* and it can be described by a unitary matrix, the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} . The weak interaction eigenstates (d', s', b') are connected to the mass eigenstates as follows [20]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1)$$

Each element of the matrix represents the relative amplitude for the different quark–quark couplings to the W^\pm bosons. The square of each element $|V_{qq'}|^2$ is proportional to the probability of a transition from one quark to another quark. The SM does not predict the values of these elements, which need to be determined from weak decays of the relevant quarks.

This 3×3 matrix has in principle nine free complex terms (18 free parameters) which are reduced to 9 due to the unitarity constraint. In addition 5 degrees of freedom can be absorbed into non-observable quark phases. This leaves 4 degrees of freedom which can be parametrised by 3 Euler angles $\theta_{12}, \theta_{23}, \theta_{31}$ and 1 complex phase δ , which is the source of CP violation in the SM.

There are several parametrisations of the CKM matrix. One of them is the standard parametrisation used by the Particle Data Group [17]. Introducing the notation $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ where i and j run over the number of generations we can write the matrix as follows

$$\mathbf{V} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \quad (2)$$

The diagonal elements, that correspond to the transitions between quarks of the same generation, are of order $\mathcal{O}(1)$. The off-diagonal elements responsible for the transition between different generations are progressively smaller: $V_{us}, V_{cd} \sim \mathcal{O}(10^{-2})$ and $V_{ub}, V_{td} \sim \mathcal{O}(10^{-3})$. Wolfenstein proposed another parametrisation [21] which takes into account the observed hierarchy in the matrix elements. Defining

$$s_{12} = \lambda, \quad s_{23} = A\lambda^2, \quad s_{13}e^{i\delta} = A\lambda^3(\rho - i\eta) \quad (3)$$

the CKM matrix can be re-written as a power expansion of the parameter $\lambda = |V_{us}|$

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

with

$$\lambda \simeq 0.23, \quad A \simeq 0.81 \quad (5)$$

The unitarity condition of the CKM matrix implies some relations between its elements. Three are constraints on the diagonal elements of the $\mathbf{V}\mathbf{V}^\dagger$:

$$\sum_{i=1}^3 |V_{ij}|^2 = 1 \quad \text{with } j = 1, \dots, 3. \quad (6)$$

This implies that the sum of all couplings of any of the up-type quarks to all the down-type quarks is the same for all generations. N. Cabibbo studied this feature, known as *weak universality*, in 1967. The other six constraints can be written in the form

$$\sum_{i=1}^3 V_{ji}V_{ki}^* = 0 = \sum_{i=1}^3 V_{ij}V_{ik}^* \quad j, k = 1, \dots, 3, \quad j \neq k. \quad (7)$$

These relations can be presented in the complex plane as triangles. The area of each triangle is the same and is directly related to the size of CP violation in the SM [22]. The most commonly used triangle is the one coming from the relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (8)$$

since each term is of equal order in λ in the Wolfenstein parametrisation. This is shown in Fig. 1, where each side is normalized to $|V_{cd}V_{cb}^*|$. The internal angles of

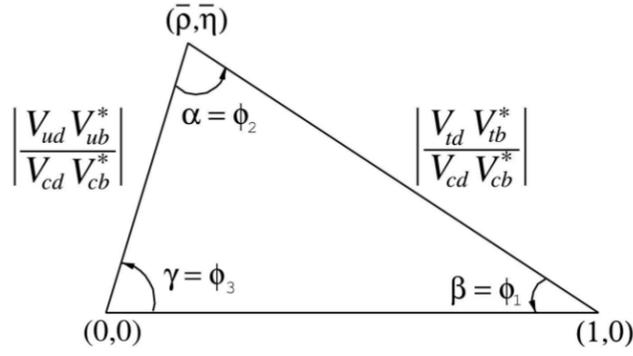


Figure 1.: A sketch of the unitarity triangle [17].

the unitarity triangle are related to the elements of the CKM matrix by

$$\begin{aligned}
 \alpha &= \phi_2 = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \\
 \beta &= \phi_1 = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \\
 \gamma &= \phi_3 = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right).
 \end{aligned} \tag{9}$$

An important measurement from the B_s^0 triangle, defined by the relation

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \tag{10}$$

is that one of the CP violation phase ϕ_s , related to the small angle

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{11}$$

Measuring the sides and angles of a unitarity triangle is a sensitive method for testing the SM. If the unitarity triangle can be over-constrained by precision measurements and does not close, the existence of physics beyond the SM can be determined. The experimental status of the measurements for the triangles relevant in the B^0 and B_s^0 systems, are shown in the $\bar{\rho}, \bar{\eta}$ plane in Fig. 2.

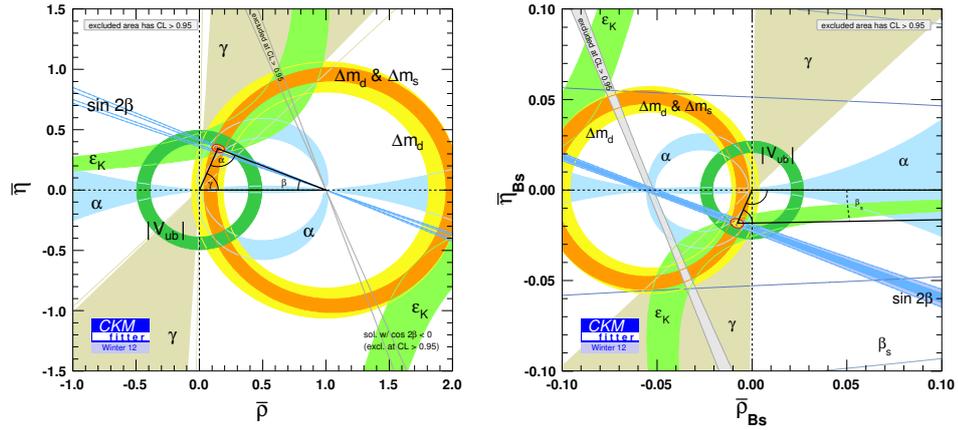


Figure 2.: Current experimental status of the unitarity triangles relevant for the B^0 (left) and B_s^0 (right) systems. The red dashed region depicts the 95% confidence region for the apex of the triangles that is determined from the measurements [23].

2.3 CP VIOLATION

One of the most important topics in modern physics is the violation of the CP symmetry, that is the non-invariance of fundamental forces under the combined transformation of charge conjugation (C) and parity (P). Charge conjugation C transforms particles into antiparticles, by inversion of their internal quantum numbers. The parity transformation P changes the handedness of the reference frame by inverting the spatial coordinates. If CP were an exact symmetry of nature, then the laws of physics for matter and antimatter would be the same. Among the fundamental forces, gravitational, electromagnetic and strong interactions are invariant under C and P and therefore also their combination CP . The weak force instead violates C and P separately, but CP is still preserved in most of the processes. The CP symmetry is however violated in certain decays. In particular the main objective of the LHCb experiment is the precise measurement of CP violation in the B meson systems (bound states of the heavy \bar{b} antiquark and a lighter u , d , s or c quark). In this section the general formalism of the CP violation of a pseudoscalar neutral meson P is briefly described (see [17] for more details), where P can be a K , D , B or B_s^0 meson. Given the Hamiltonian \mathcal{H} governing the weak interaction, the decay amplitude of P and its CP conjugate \bar{P} to a final state f and its CP conjugate \bar{f} can be expressed as

$$A_f = \langle f | \mathcal{H} | P \rangle \quad , \quad \bar{A}_f = \langle f | \mathcal{H} | \bar{P} \rangle \quad (12)$$

$$A_{\bar{f}} = \langle \bar{f} | \mathcal{H} | P \rangle \quad , \quad \bar{A}_{\bar{f}} = \langle \bar{f} | \mathcal{H} | \bar{P} \rangle. \quad (13)$$

In the case CP is not violated $[CP, \mathcal{H}] = 0$, then A_f and $\bar{A}_{\bar{f}}$ have the same magnitude and an arbitrary unphysical relative phase. If a neutral meson is initially a superposition of P^0 and \bar{P}^0

$$|\psi(t=0)\rangle = a(0)|P^0\rangle + b(0)|\bar{P}^0\rangle$$

the state will evolve in time acquiring components of all the possible decay final states

$$|\psi(t)\rangle = a(t)|P^0\rangle + b(t)|\bar{P}^0\rangle + c_1(t)|f_1\rangle + c_2(t)|f_2\rangle + \dots$$

If we consider times t larger than the typical strong interaction scale, then the time evolution is determined by a simplified 2×2 effective Hamiltonian \mathbf{H}

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

The light (P_L) and heavy (P_H) mass eigenstates can be written as a linear combination of flavour eigenstates

$$\begin{aligned} |P_L\rangle &\propto p\sqrt{1-z}|P^0\rangle + q\sqrt{1+z}|\bar{P}^0\rangle \\ |P_H\rangle &\propto p\sqrt{1+z}|P^0\rangle - q\sqrt{1-z}|\bar{P}^0\rangle \end{aligned}$$

with $|p|^2 + |q|^2 = 1$ when $z = 0$. The real and imaginary parts of the difference of the two eigenvalues $\omega_{L,H}$ represent the mass and decay width differences

$$\begin{aligned} \Delta m &\equiv m_H - m_L = \mathcal{R}e(\omega_H - \omega_L) \\ \Delta\Gamma &\equiv \Gamma_H - \Gamma_L = -2\mathcal{I}m(\omega_H - \omega_L) \end{aligned}$$

The eigenvalue problem leads to the following solutions

$$\left(\frac{q}{p}\right)^2 = \frac{\mathbf{M}_{12}^* - \frac{i}{2}\mathbf{\Gamma}_{12}^*}{\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}} \quad (14)$$

and

$$z = \frac{\delta m - (i/2)\delta\Gamma}{\Delta m - (i/2)\Delta\Gamma} \quad (15)$$

where

$$\delta m = \mathbf{M}_{11} - \mathbf{M}_{22}, \quad \delta\Gamma = \mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{22} \quad (16)$$

are the mass and decay-rate differences for the flavour eigenstates P_0 and \bar{P}_0 . For meson decays it is possible to classify CP violation into three categories.

I. *CP violation in decay.* This type of *CP* violation occurs in both charged and neutral *B* meson decays if

$$|\bar{A}(\bar{f})/A(f)| \neq 1, \quad (17)$$

i.e. when the amplitude for a decay is not equal to the amplitude for its *CP* conjugate decay. In this case *CP* violation is said to have occurred directly and it can be observed measuring the time-integrated asymmetry

$$a = \frac{\Gamma(P^0 \rightarrow f) - \Gamma(\bar{P}^0 \rightarrow \bar{f})}{\Gamma(P^0 \rightarrow f) + \Gamma(\bar{P}^0 \rightarrow \bar{f})} \quad (18)$$

II. *CP violation in mixing.* In contrast to *CP* violation in decay, this type of *CP* violation occurs only in the neutral systems, when the two neutral mass eigenstates cannot be chosen to be *CP* eigenstates. When *CP* symmetry is conserved, the mass eigenstates must be *CP* eigenstates. The condition to have this type of *CP* violation is

$$|q/p| \neq 1, \quad (19)$$

that is the oscillation probability $P^0 \rightarrow \bar{P}^0$ is different from $\bar{P}^0 \rightarrow P^0$. For the neutral *B* system, this effect is very small ($\mathcal{O}(10^{-2})$).

III. *CP violation in interference between decay and mixing.* This kind of *CP* violation arises only in the neutral system. Even if *CP* symmetry would be conserved in the decay amplitudes and in the mixing, *CP* violation may still be observed in the interference between $P^0 \rightarrow f$ and $P^0 \rightarrow \bar{P}^0 \rightarrow f$. This occurs when there is a relative phase between the mixing parameter q/p and the decay amplitudes \bar{A}_f/A_f . This kind of *CP* violation is defined by

$$\text{Im}(\lambda_f) \neq 0 \quad (20)$$

with

$$\lambda_f \equiv \frac{q}{p} \frac{\bar{A}_f}{A_f}. \quad (21)$$

This form of *CP* violation is time-dependent and can be observed using decays of neutral mesons into a *CP* eigenstate f_{CP}

$$A_{CP}(t) = \frac{d\Gamma/dt[\bar{P}^0(t) \rightarrow f_{CP}] - d\Gamma/dt[P^0(t) \rightarrow f_{CP}]}{d\Gamma/dt[\bar{P}^0(t) \rightarrow f_{CP}] + d\Gamma/dt[P^0(t) \rightarrow f_{CP}]} \quad (22)$$

For B^0 mesons $q/p = 1$ and $\Delta\Gamma \approx 0$ holds with good approximation. In this case the asymmetry can be written as

$$A_{CP}(t) = S_f \sin(\Delta m t) - C_f \cos(\Delta m t) \quad (23)$$

with

$$S_f = \frac{2\mathcal{I}m(\lambda_f)}{1 + |\lambda_f|^2}, \quad C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}. \quad (24)$$

For B_s^0 mesons however $\Delta\Gamma_s$ has been measured to be different from zero [24]. In this case the asymmetry becomes

$$A_{CP}(t) = \frac{-C_f \cos(\Delta m_s t) + S_f \cos(\Delta m_s t)}{\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - A_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)} \quad (25)$$

where

$$A_f^{\Delta\Gamma} = -\frac{2\mathcal{R}e(\lambda_f)}{1 + |\lambda_f|^2}. \quad (26)$$

For final states which are not CP eigenstates, similar expressions hold.

2.4 PHYSICS OF $b \rightarrow s$ HADRONIC PENGUIN DECAYS

In this section a brief introduction about hadronic decays is given, with particular emphasis on so-called pure penguin processes. A generic two-body hadronic B decay is mediated by $b \rightarrow q_1 \bar{q}_2 d(s)$ transitions, where $q_{1/2}$ can be a u, d, c, s quark [25]. These processes can be divided in two groups according to the Feynman diagram: tree-level topology and penguin (loop) topology. The latter consists of gluonic (QCD) and electroweak (EW) penguins. Depending on the flavour content of the final state, the hadron decays can be classified into:

- $q_1 \neq q_2 \in u, c$: transitions mediated by tree-level topologies;
- $q_1 = q_2 \in d, s$: transitions mediated by penguin topologies;
- $q_1 = q_2 \in u, c$: transitions mediated by both tree-level and penguin topologies.

The diagrams for each of these topologies are sketched in Figs. 3 to 5.

In order to describe the weak decays of hadrons, it is necessary to take into account also the strong interactions binding together the constituent quarks. Thanks

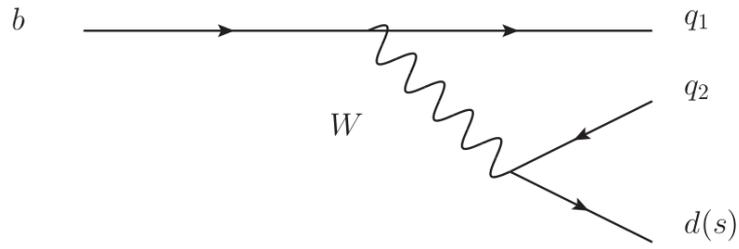


Figure 3.: Tree-level Feynman diagram governing the transition of a non-leptonic B hadron decay [25].

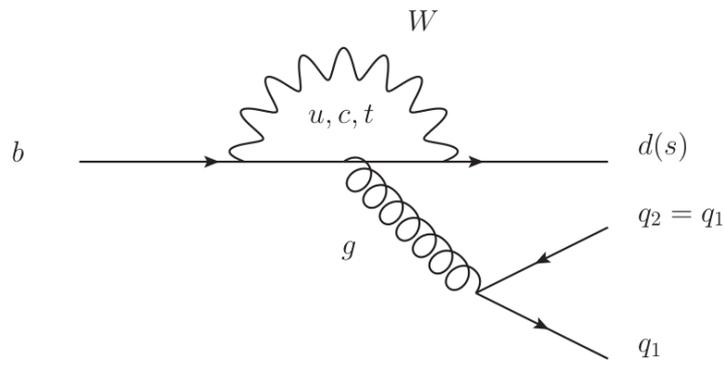


Figure 4.: QCD penguin Feynman diagram governing the transition of a non-leptonic B hadron decay [25].

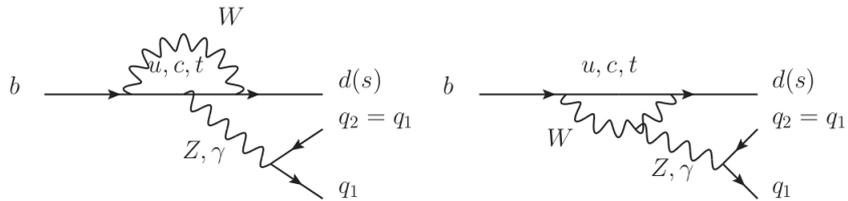


Figure 5.: EW Feynman diagrams governing the transition of a non-leptonic B hadron decay [25].

to the asymptotic freedom of QCD one can treat the short-distance corrections, that is the contribution of hard gluons at energies of the order $\mathcal{O}(M_W)$ down to hadronic scales $\geq 1\text{ GeV}$, in perturbation theory. The formal framework used to exploit such property is the so called Operator Product Expansion (OPE) [25, 26].

Operator product expansion allows one to write the amplitude for a generic non-leptonic B meson decay into a final state f as follows

$$\mathcal{A}(B \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | B \rangle = \left(\frac{G_F}{\sqrt{2}} \sum_{i=1}^{12} V_i^{\text{CKM}} C_i(\mu) + C_i^{\text{NP}}(\mu) \right) \langle f | Q_i(\mu) | B \rangle + \sum_{i=1}^{N_{\text{NP}}} \tilde{C}_i^{\text{NP}}(\mu) \langle f | \tilde{Q}_i(\mu) | B \rangle, \quad (27)$$

where \mathcal{H}_{eff} is the effective weak Hamiltonian. The Q_i represent the local operators that govern the decay within the SM and \tilde{Q}_i denote the ones possibly arising from New Physics. The short distance contributions are represented by perturbative Wilson coefficient functions $C_i(\mu)$, while long-distance contributions are represented by non-perturbative hadronic matrix elements $\langle f | Q_i(\mu) | B \rangle$. The terms $C_i^{\text{NP}}(\mu)$ and $\tilde{C}_i^{\text{NP}}(\mu)$ denote the Wilson coefficients arising within a given NP model, which can in general be complex. The parameter G_F is the Fermi constant, V_i^{CKM} is a factor arising from the CKM matrix element entering the weak interaction vertices and μ is the renormalization scale.

The effective weak Hamiltonian for non-leptonic $b \rightarrow s$ decays within the SM is given by [7]

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} \left\{ V_{ub} V_{us}^* [C_1(\mu)(Q_1^u(\mu) - Q_1^c(\mu)) + C_2(\mu)(Q_2^u(\mu) - Q_2^c(\mu))] - V_{tb} V_{ts}^* \left[C_1(\mu) Q_1^c(\mu) + C_2(\mu) Q_2^c(\mu) + \sum_{i=3}^{12} C_i(\mu) Q_i(\mu) \right] \right\} \quad (28)$$

with

$$\begin{aligned} Q_1^u &= (\bar{b}_L \gamma^\mu u_L^i) (\bar{u}_L^i \gamma_\mu s_L), & Q_2^u &= (\bar{b}_L \gamma^\mu s_L^i) (\bar{u}_L^i \gamma_\mu u_L^i), \\ Q_{3,5} &= \sum_q (\bar{b}_L \gamma^\mu s_L) (\bar{q}_{L,R} \gamma_\mu q_{L,R}), & Q_4 &= \sum_q (\bar{b}_L \gamma^\mu q_L) (\bar{q}_L \gamma_\mu s_L), \\ Q_6 &= -2 \sum_q (\bar{b}_L q_R) (\bar{q}_R s_L), & Q_{7,9} &= \frac{3}{2} \sum_q (\bar{b}_L \gamma^\mu s_L) e_q (\bar{q}_L \gamma_\mu q_L), \\ Q_8 &= -3 \sum_q e_q (\bar{b}_L q_R) (\bar{q}_R s_L), & Q_{10} &= \frac{3}{2} \sum_q e_q (\bar{b}_L \gamma^\mu q_L) (\bar{q}_{R,L} \gamma_\mu q_{R,L}), \\ Q_{11} &= \frac{e}{16\pi^2} m_b (\bar{b}_R \sigma^{\mu\nu} s_L) F_{\mu\nu}, & Q_{12} &= \frac{g}{16\pi^2} m_b (\bar{b}_R \sigma^{\mu\nu} T^a s_L) G_{\mu\nu}^a \end{aligned} \quad (29)$$

where $q_{L,R} \equiv (1 \mp \gamma_5)/2q$, $u_i = u, c$ and e_q denotes the quark electric charge. The sum over the quarks q runs over the active flavours at the scale μ . Here

Q_1 and Q_2 are the current-current operators, Q_{3-6} the QCD penguin operators, Q_{7-10} the electroweak penguin operators, and Q_{11-12} the (chromo)magnetic penguin operators. The Wilson coefficients $C_i(\mu)$ depend in general on the QCD renormalization scheme for the operators. Their order of magnitude at the renormalization scale $\mu = \mathcal{O}(m_b)$ is $C_1(\mu) = \mathcal{O}(10^{-1})$ and $C_2(\mu) = \mathcal{O}(1)$ for the current-current operators, and $\mathcal{O}(10^{-2})$ for the penguin operators. This method can be applied in general to all B decays governed by the same quark-level transitions ($b \rightarrow q_1 \bar{q}_2 d(s)$). The differences between different decay modes come only from the hadronic matrix elements associated with the operators. The way these elements can be computed represents the most challenging task and the principal source of theoretical uncertainties. Without giving any details, here the different approaches commonly used to deal with non-leptonic B decays are simply listed.

- *QCD factorization*: the idea is that factorization is valid for hadrons containing a quark Q with $m_Q \gg \Lambda_{QCD}$. The amplitudes are calculable as leading order Λ_{QCD}/m_Q expansions [27]- [28];
- *Perturbative QCD*: this method separates hard components from a QCD process, which can be treated by perturbation theory. Non-perturbative components are organized in the form of hadron wave functions, which can be extracted from experimental data [29];
- *Six-quark effective Hamiltonian*: the previous approaches are based on the four-fermion operator effective Hamiltonian. The short-distance QCD contributions are characterized by the Wilson coefficient functions of four-quark operators, whose calculation is well developed. The long-distance ones are in principle obtained by evaluating the hadronic matrix elements of relevant four-quark operators, but the calculation remains a hard task as it involves non-perturbative effects of QCD. For the B decays into two mesons, it involves three quark-antiquark pairs, that motivate the application of a six-quark (rather than four-quark) operator effective Hamiltonian [30].

2.4.1 The $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decays

In this paragraph some aspects of pure penguin $b \rightarrow s$ transitions are discussed, since the decays object of this thesis belong to this category. The amplitude for pure penguin decays can be written as [7]

$$\mathcal{A}(B \rightarrow f) = -V_{ub}^* V_{us} \sum P_i^{\text{GIM}} - V_{tb}^* V_{ts} \sum P_i, \quad (30)$$

where P_i contains penguin contractions of charmed current-current operators together with the matrix elements of $b \rightarrow s$ penguin operators, while P_i^{GIM} represents the GIM-suppressed difference of penguin contractions of current-current operators containing charm and up quarks respectively. Neglecting the CKM matrix elements governing the transitions between the third and the first family (doubly Cabibbo-suppressed terms), the decay amplitude has a vanishing weak phase and therefore there is no direct CP violation. In terms of the coefficients S_f and C_f of sine and cosine terms in the time-dependent CP asymmetry (for f being a CP eigenstate with eigenvalue η_f) in the SM one expects

$$S_f = \eta_f \mathcal{I}m \lambda_f = -\eta_f \sin 2\phi_M \quad (31)$$

which measures the phase of the mixing amplitude, and

$$C_f = 0, \quad (32)$$

where

$$\lambda_f \equiv \frac{q}{p} \frac{\bar{A}}{A} = e^{-2i\phi_M}, \quad (33)$$

with $A = \mathcal{A}(B \rightarrow f)$, $\bar{A} = \mathcal{A}(\bar{B} \rightarrow f)$ and $\phi_M = \beta(-\beta_s)$ the mixing angle for the B^0 (B_s^0) system in the SM. Comparing the measured S_f to the one obtained theoretically is a crucial ingredient to search for NP effects: a deviation from zero much larger than the estimated SM error would be a strong signal of NP.

The $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decays are a particular case of pure penguin decays. We assume that the resonance is the $K^*(892)^0$ if not stated differently. The Feynman diagram for the B_s^0 decay is shown in Fig. 6. All three up-type quarks can

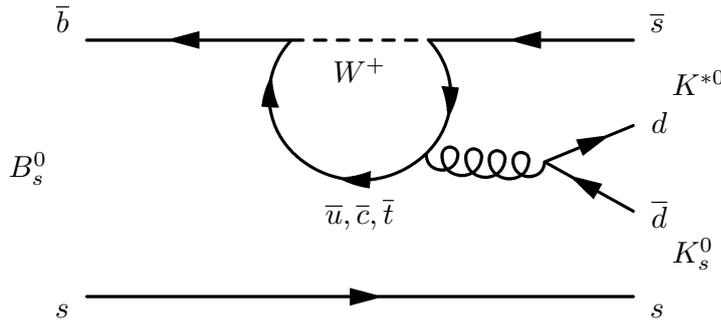


Figure 6.: Feynman diagram of the $B_s^0 \rightarrow K^{*0} K_s^0$ decay [31].

¹ Charge-conjugate modes are implicitly included throughout this thesis.

participate in the loop, but the top quark is dominant due to its much larger mass. Some theoretical predictions given in the approaches mentioned before, for the branching ratios of these decays are summarized in Tab. 3. The sizable branching ratios are due to the fact that the related CKM elements ($V_{tb}^* V_{ts}$) are relatively large. However, SM predictions are plagued by large uncertainties which have to be taken into account in order to probe the SM itself and to disentangle SM effects from NP.

Table 3.: Theoretical predictions of unmeasured $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decay branching fractions. The three theoretical models are QCD Factorization (QCDF), perturbative QCD (pQCD) and a six-quark Effective Hamiltonian method. Branching fractions given in units of ($\times 10^{-6}$).

Branching fraction	QCDF [32]	pQCD [33]	Eff. Hamiltonian [30]
$\mathcal{B}(B^0 \rightarrow \bar{K}^{*0} K^0)$	$0.70^{+0.18+0.28}_{-0.15-0.25}$	-	-
$\mathcal{B}(B^0 \rightarrow K^{*0} K^0)$	$0.47^{+0.36+0.43}_{-0.17-0.27}$	-	-
$\mathcal{B}(B_s^0 \rightarrow \bar{K}^{*0} K^0)$	$10.5^{+3.3+5.1}_{-2.8-4.5}$	$7.3^{+2.5+2.1+0.0}_{-1.7-1.3-0.0}$	$8.5^{+1.8+1.5}_{-2.1-1.6}$
$\mathcal{B}(B_s^0 \rightarrow K^{*0} K^0)$	$10.1^{+7.5+7.7}_{-3.6-4.8}$	$4.3^{+0.7+2.2+0.0}_{-0.7-1.4-0.0}$	$7.1^{+0.2+1.3}_{-0.4-1.1}$

From the experimental point of view many charmless 3-body decays of B mesons have been previously studied and their branching ratios measured. Nevertheless the decays addressed in this thesis have not yet been observed experimentally, although the inclusive decays $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$ with $K_s^0 \rightarrow \pi^+ \pi^-$ have already been studied by LHCb [9].

In Tab. 4 some world averages and measurements are presented for some of the charmless decays to similar final states. For the second and third row the first un-

Table 4.: Experimental measurements of charmless B decay branching fractions.

Measurement	Value ($\times 10^{-6}$)	Reference
$\mathcal{B}(B^0 \rightarrow K^0 \pi^+ \pi^-)$	49.6 ± 2.0	[17]
$\mathcal{B}(B^0 \rightarrow K^0 K^\pm \pi^\mp)$	$6.4 \pm 0.9 \pm 0.4 \pm 0.3$	[9]
$\mathcal{B}(B_s^0 \rightarrow K^0 K^\pm \pi^\mp)$	$73.6 \pm 5.7 \pm 6.9 \pm 3.0$	[9]

certainty is statistical, the second systematic and the last is due to the uncertainty on $\mathcal{B}(B^0 \rightarrow K^0 \pi^+ \pi^-)$.

The main goal of this analysis is to measure the branching ratios of the $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decays with respect to the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ decay. In the following the former is referred to as “signal” and the latter as “normalization”.

3 | THE LHCb EXPERIMENT

In this chapter the Large Hadron Collider (LHC) and its main experiments are briefly discussed, followed by a more detailed presentation about the LHCb experiment.

3.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [34] is the largest existing facility for the study of particle physics, placed 100 m underground across the Swiss and French borders, as shown in Fig. 7.

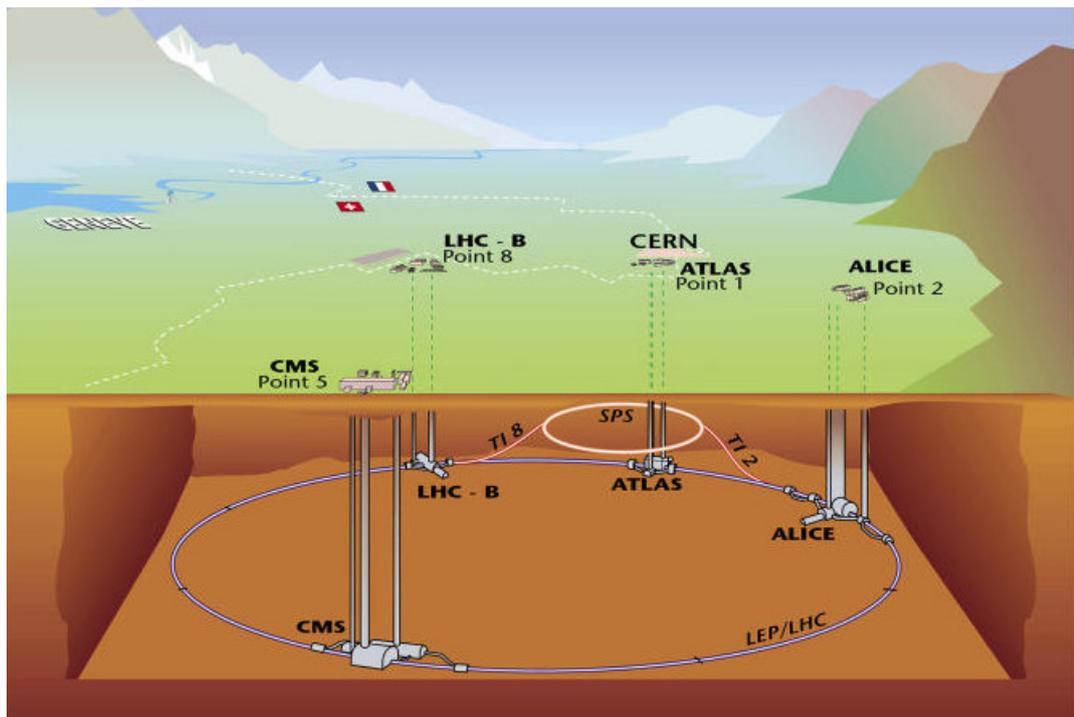


Figure 7.: A schematic view of the LHC collider and the position of the four experiments ATLAS, CMS, ALICE, LHCb [35].

The machine mainly collides protons accelerated in a tunnel of 27 km. Protons are injected into the LHC from the chain of preaccelerators Linac2 - Proton Synchrotron Booster (PSB) - Proton Synchrotron (PS) - Super Proton Synchrotron (SPS). Linac2 is a linear accelerator providing the PSB with proton bunches of 50 MeV energy. In the PSB protons are then accelerated up to 1 GeV before being injected into the PS. The PS raises their energy up to 26 GeV and passes them to the SPS. The SPS performs the last acceleration step, up to 450 GeV, before the injection of beams into the LHC via the two tunnels TI2 and TI8. The beams are then brought to collisions at four interaction point, where the main experiments are placed: the ATLAS (A Toroidal LHC ApparatuS [36]), CMS (Compact-Muon-Solenoid [37]), ALICE (A Large Ion Collider Experiment [38]) and LHCb (Large Hadron Collider beauty [39]) detectors. The former two are general purpose detectors, ALICE is investigating the properties of quark-gluon plasma in heavy ion collisions and LHCb is designed to study beauty and charmed hadrons. The complex of CERN accelerators is schematically shown in Fig. 8.

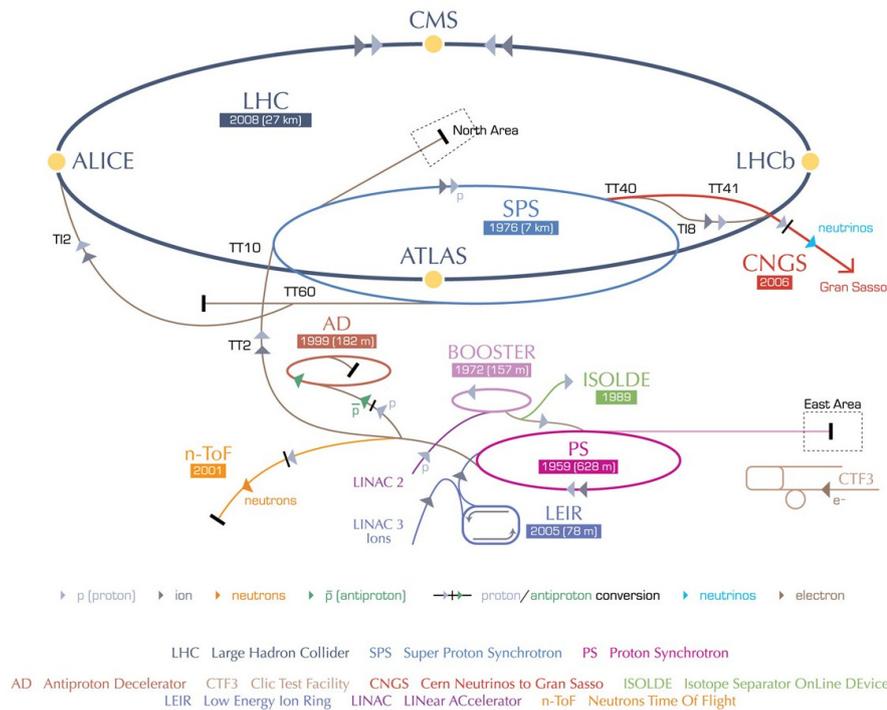


Figure 8.: Schematic view of the complex of CERN accelerators: Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), Super Proton Synchrotron (SPS) and LHC. The two tunnels to inject the proton beams into the LHC, TI2 and TI8, are also shown [40].

The collider is designed to operate at a centre-of-mass energy of 14 TeV and an instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. For the design energy the magnetic field of a single dipole, used to bend proton trajectories, needs to reach a magnitude of 8.33 T. This is achieved by super-conducting dipole magnets working at a temperature of 1.9 K. The protons can be stored in 2808 proton bunches per beam, with a spacing of 25 ns, each of them containing $\sim 10^{11}$ protons. The design luminosity of the LHCb experiment is lower and corresponds to $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, in order to limit the number of primary interactions per bunch crossing to have an efficient event reconstruction. Technically this is done by adjusting the transversal beams overlap. LHC started its physics operation on the 23rd November 2009 and has now an history of three years of data taking. The data were collected with a centre-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ in 2010 and 2011, that was increased to $\sqrt{s} = 8 \text{ TeV}$ in 2012. The bunch spacing was 50 ns. The instantaneous luminosity delivered to LHCb was $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ in 2010 and then it was increased above the design value to $4 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. The integrated luminosities delivered and recorded by LHCb during 2010, 2011 and 2012 data taking are summarized in Fig. 9. The efficiency of data taking, given by the ratio between recorded and delivered luminosity, is very high and reached 90% in 2011. The analysis presented in this thesis is performed using the data collected in 2011.

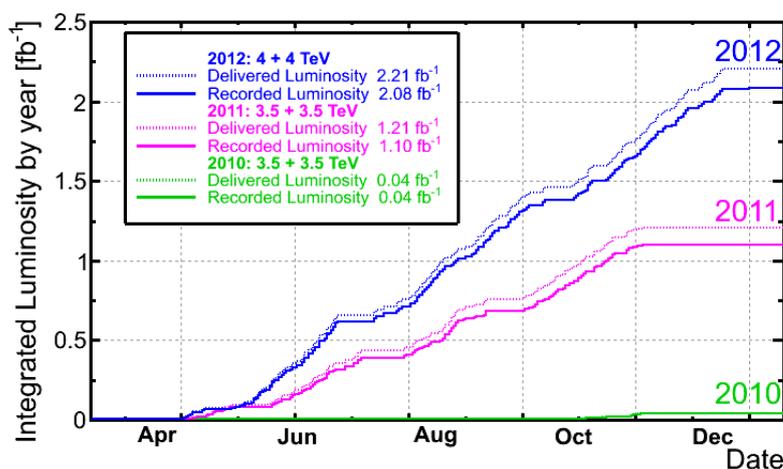


Figure 9.: Integrated luminosity delivered and recorded by LHCb in the three years of operation [41].

With the energy available at LHC the b -quarks are mainly produced by inelastic pp collisions. At the leading order the dominant production modes are the gluon-gluon fusion or the quark-antiquark annihilation processes, shown in Fig. 10.

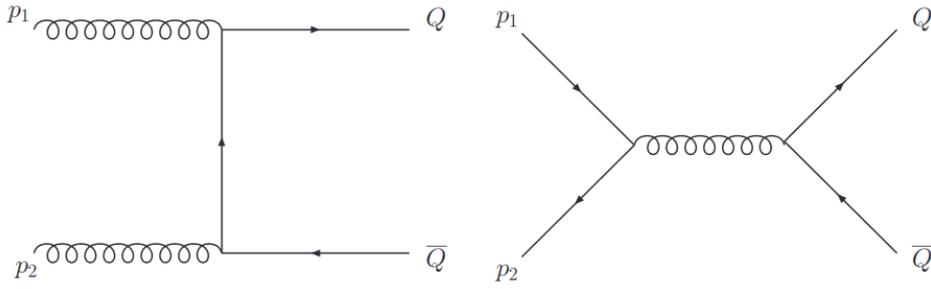


Figure 10.: Leading order Feynman diagrams of b -hadron production. Left: gluon-gluon fusion. Right: quark-antiquark annihilation [42].

The cross-section for inelastic pp collisions at $\sqrt{s} = 7$ TeV has been measured by LHCb [43] and the extrapolation to the full solid angle leads to

$$\sigma_{b\bar{b}} = \sigma(pp \rightarrow b\bar{b}X) = (288 \pm 4 \pm 48)\mu\text{b}. \quad (34)$$

It is unlikely that the gluons or the quarks producing the $b\bar{b}$ pairs carry the same fraction of proton momentum. As a result the $b\bar{b}$ pairs are boosted in the direction of the beam axis and hence produced predominantly in the forward or backward direction. This can be observed in Fig. 11, where the simulated distribution of the angle between the $b(\bar{b})$ quarks and the beam direction is shown.

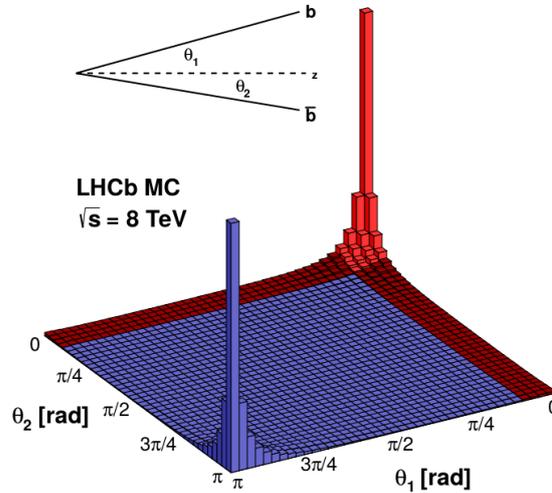


Figure 11.: Azimuthal angle distribution of the $b\bar{b}$ quark pairs. In red the part of the distribution in the LHCb acceptance [41].

The number of produced $b\bar{b}$ pairs can be calculated as $N_{b\bar{b}} = \sigma(pp \rightarrow b\bar{b}X) \cdot$

\mathcal{L}_{int} , where \mathcal{L}_{int} is the integrated luminosity recorded in 2011, and it corresponds to approximately $3 \cdot 10^{11}$. The fraction of them within the LHCb acceptance is approximately 25%.

3.2 THE LHCb EXPERIMENT AND DETECTOR

The LHCb experiment [39] is optimised for flavour physics studies. In particular precision measurements of CP violation and rare decays in *beauty* and *charmed* hadrons are performed for indirect searches of New Physics. To achieve such delicate measurements, LHCb has been designed with some important characteristics. A flexible and robust trigger is needed to select many different final states in the hadronic environment. An excellent vertex resolution is required to detect short lived resonances and to perform lifetime measurements. In addition good momentum and invariant mass resolution help to reduce combinatorial background. To distinguish among many different final states a good particle identification is an important feature.

3.2.1 Detector design

The LHCb detector is a single arm spectrometer, designed to exploit the forward characteristic of B hadron production. The layout of the detector is displayed in Fig. 12. The coordinate system is a right-handed cartesian system, defined such that the z axis corresponds to the beam axis, the y axis is in the vertical direction (non-bending plane) and the x axis is in the horizontal direction (bending plane). The geometrical acceptance in the x - z and y - z planes are respectively $10 - 300$ mrad and $10 - 250$ mrad. The pseudo-rapidity (η)¹ range for tracks inside the LHCb geometrical acceptance is approximately between 1.8 and 4.9. LHCb consists of several subdetectors.

- *VELO*: the Vertex Locator is placed around the interaction point and provides a very good spatial resolution of reconstructed vertices;
- *RICH1*: the first Ring Imaging Cherenkov detector is placed just after the VELO and it provides information for the identification of charged particles;
- *TT*: the Tracker Turicensis is a tracking detector placed in front of the magnet;

¹ The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the angle formed by the particle trajectory and the z axis.

- *Magnet*: this dipole magnet is used to bend particle trajectories in order to measure their momentum;
- *T1, T2, T3*: these are the main tracking stations placed behind the magnet;
- *RICH2*: the second Ring Imaging Cherenkov detector provides particle identification information for high momentum particles;
- *ECAL*: the electromagnetic calorimeter is used to trigger and for the identification of electrons and photons;
- *HCAL*: the hadronic calorimeter is placed behind the ECAL and it is used for the hadronic trigger;
- *Muon chambers*: the muon system is used to detect muons. One of them is placed before ECAL and the other four behind the HCAL.

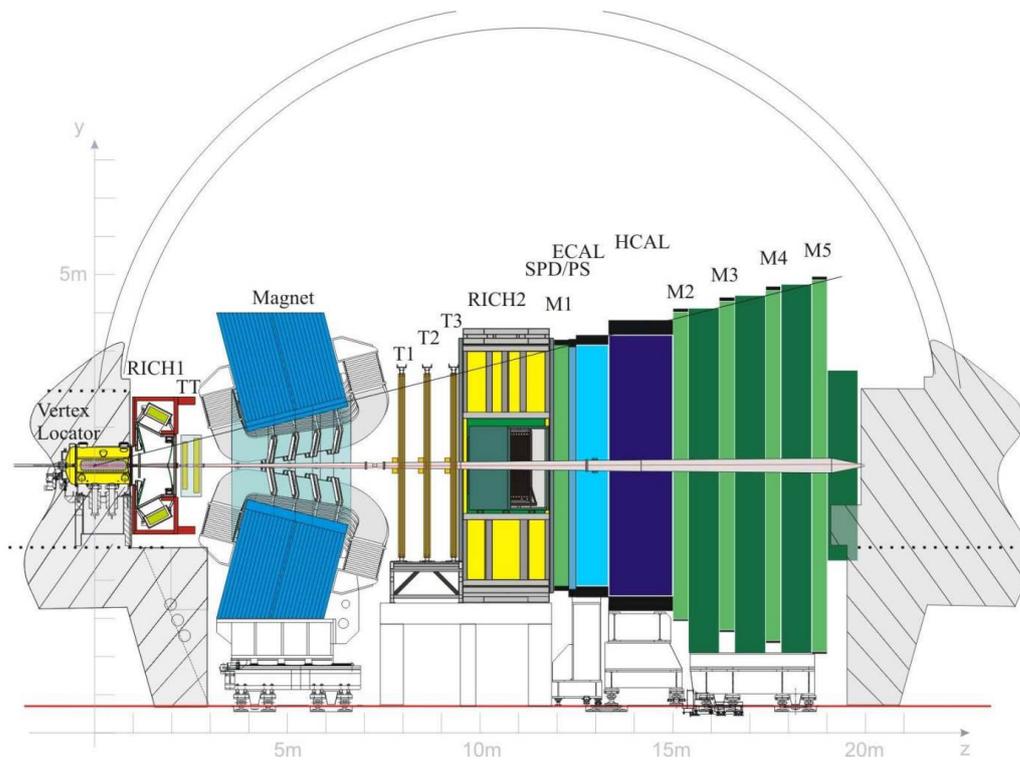


Figure 12.: Layout of the LHCb detector in the y - z section. From left to right the different subdetectors are visible: VELO, RICH1, TT, Magnet, main Tracking Stations, RICH2, Electromagnetic Calorimeter (ECAL), Hadronic Calorimeter (HCAL) and Muon stations [39].

3.2.2 The tracking system

The tracking system consists of the VELO and four tracking stations: the TT and T1-T3, located respectively upstream and downstream the magnet. These detectors provide information to reconstruct charged particle trajectories and to measure their momentum. VELO, TT and the inner part of T1-T3, called *Inner Tracker* (IT), are silicon microstrip detectors, while the outer part of T1-T3, called *Outer Tracker* (OT), employs straw-tubes.

The Vertex Locator

The ability to reconstruct primary and secondary vertices of hadrons containing b and c quarks in order to select in the most efficient way signal and reject the combinatorial background is achieved with the VELO detector [44]. It consists of 21 modules with silicon strip detectors, oriented parallel to the vertical (x - y) plane, along the beam direction, as shown in Fig. 13. Each module is subdivided into two halves, that can be moved apart to allow a safe LHC beam injection. The individual VELO sensors are mounted on a moveable device inside a vessel which maintains the vacuum and allows to move the two halves between 3 cm (fully open) and 8 mm (data taking conditions) from the beam. The sensors are placed inside two boxes in order to be under a secondary vacuum. The walls of the boxes on the beam side are called RF (Radio Frequency) foils. Their aim is to prevent Radio Frequency pickup of the VELO sensors from the LHC beams. Two additional sensor planes (pile-up veto system) are placed upstream of the VELO and are used in the trigger system. The polar angular acceptance of the halves is more than 180° allowing them to overlap during data taking. The modules are composed of two planes of $300\ \mu\text{m}$ thick silicon microstrip sensors allowing to measure radial (R sensors) and polar (ϕ sensors) coordinates of the hits generated by the particles. A schematic view of R and ϕ sensors is shown in Fig. 14.

The microstrips of the R sensor are modeled in a semi-circular shape, with their centres coinciding with the nominal LHC beam position. Each strip is divided in four sectors 45° wide. The aim is to ensure a low occupancy per readout channel and low strip capacitance. The strip pitch increases linearly from a minimum of $40\ \mu\text{m}$ (near the beam) to a maximum of $101.6\ \mu\text{m}$ (far from the beam), to take into account the higher particle occupancy near the interaction point. The ϕ sensors are subdivided into two regions at a radius of 17.25 mm. The pitch of the sensors in the inner part increases radially from $39.3\ \mu\text{m}$ to $78.3\ \mu\text{m}$, while in the outer one it goes from $39.3\ \mu\text{m}$ to $97\ \mu\text{m}$. Inner and outer regions are skewed with respect to

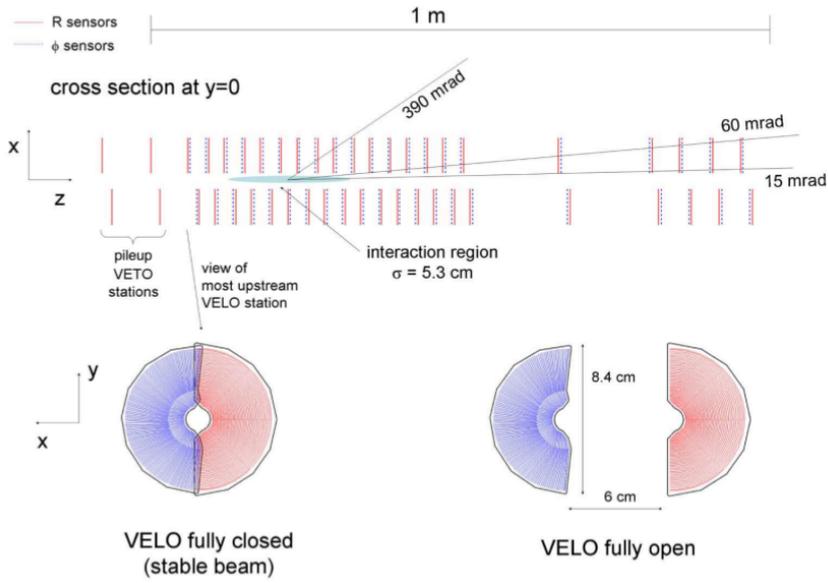


Figure 13.: Top view of the VELO silicon sensors, with the detector in the fully closed position (top). Front view of the modules in both the closed (bottom left) and open positions (bottom right) [39].

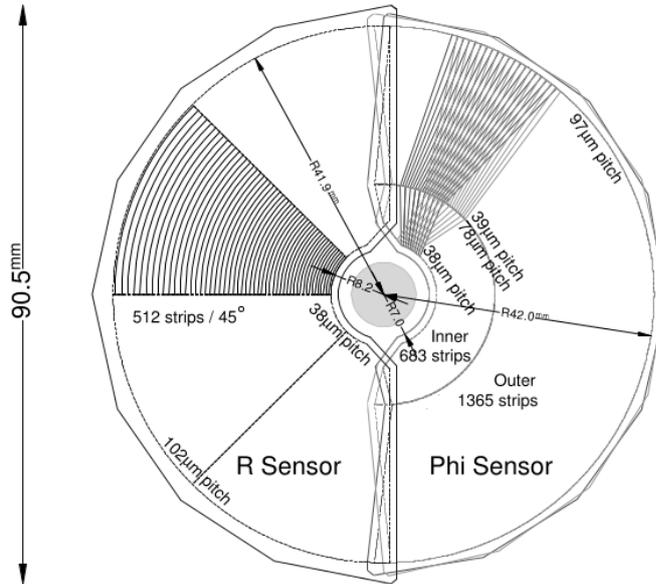


Figure 14.: Schematic view of R and ϕ sensors [39].

the radial direction by 20° (inner part) and 10° (outer part), in order to improve the pattern recognition. In addition, adjacent ϕ sensors have opposite skew. The VELO was designed in order to minimise the amount of material traversed by particles, at the same time with a good geometrical acceptance. The tracks coming from the primary interaction point, and in the LHCb acceptance pass through at least three modules. From simulation studies it has been possible to calculate the amount of material traversed by each of these tracks which in average is 17.5% of a radiation length. The largest contribution is coming from the RF-foil.

The performance of the VELO detector has been studied with the large sample of minimum bias events collected during the different data taking periods and has been compared with fully simulated Monte Carlo events. The resolutions in the reconstruction of vertices as a function of the number of tracks fitted into the vertex are summarized in Fig. 15, for 2011 data and the corresponding MC simulation, for one primary vertex (PV) in the event. The resolution on the x and y directions is in the range from about $10\ \mu\text{m}$ to about $40\ \mu\text{m}$, while the z direction varies between about $50\ \mu\text{m}$ and $250\ \mu\text{m}$. Fig. 16 shows the resolutions on the impact parameter of tracks with respect to the primary vertex in the x and y directions, as a function of the inverse of the transverse momentum of tracks.

Silicon Tracker

The Silicon Tracker (ST) comprises two detectors: the Tracker Turicensis (TT) and the Inner Tracker (IT) [45]. Both detectors use silicon microstrip sensors with a strip pitch of about $200\ \mu\text{m}$. Each ST station is composed of four layers of sensors. The two vertical layers are called x -layers, while the u and v -layers are rotated by an angle of $+5^\circ$ and -5° . The layout has been optimised for hit resolution in the bending plane and resolution of ambiguities.

The *Tracker Turicensis* consists of two $150\ \text{cm}$ wide and $130\ \text{cm}$ high planar stations placed upstream of the dipole magnet. It provides tracking information also for low momentum particles, which are bent out of the acceptance before reaching the main tracker. The first two layers (x - u) are separated by the other two (v - x) by $27\ \text{cm}$ along the beam axis. A sketch of the layers is shown in Fig. 17. They consist of different modules put side by side, each of them covering half of LHCb acceptance in height. In order to prevent acceptance gaps and to have a better alignment, the modules slightly overlap by a few millimeters. Each module is composed of seven silicon sensors placed in a row. These sensors are $500\ \mu\text{m}$ thick and each of them is made of 512 silicon strips with a pitch of $183\ \mu\text{m}$. The *Inner Tracker* covers a $120\ \text{cm}$ wide and $40\ \text{cm}$ high cross shaped in the inner part of the three

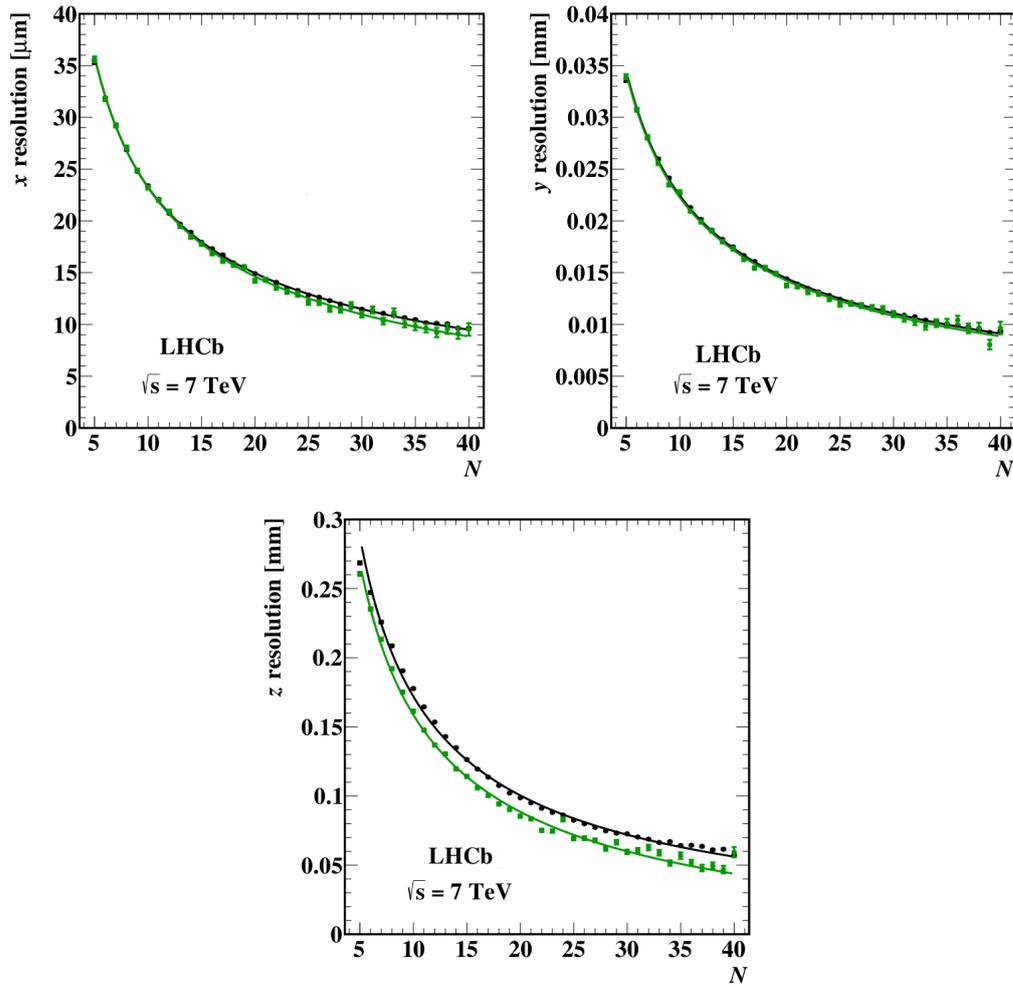


Figure 15.: Resolutions in the reconstruction of primary vertex coordinates as a function of the number of tracks in the event, obtained from 2011 data and Monte Carlo simulations. In the figures the resolutions for the x (top left), y (top right) and z (bottom) coordinates are reported. The performance shown refers to events where only one primary vertex has been reconstructed [41].

tracking stations (T1-T3) placed downstream of the magnet. Each station is made of four detector boxes placed around the beam-pipe as shown in Fig. 18. Each box contains 4 layers of detectors, consisting of 7-sensor modules. The layers are placed in the x - u - v - x configuration. To avoid acceptance gaps and to help the alignment, the adjacent modules overlap in the x direction by 3 mm and are staggered in z by a few millimeters. Each module consists of one or two silicon sensors plus a front-end readout hybrid. The thickness of the sensors is 320 μm for the boxes

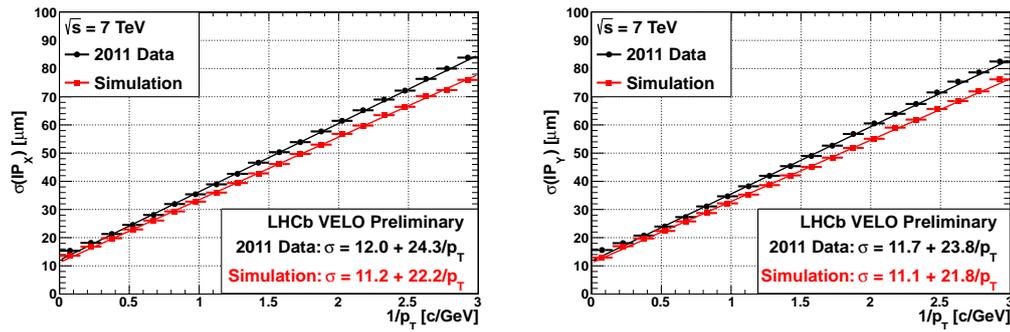


Figure 16.: Resolution achieved for the x (left) and y (right) components of the impact parameter (IP) of tracks with respect to the primary vertex as a function of the inverse of the transverse momentum of tracks $1/p_T$ [41].

above and below the beam line, and $410\ \mu\text{m}$ on the sides. The pitch between the strips is $200\ \mu\text{m}$ and the single hit resolution is $50\ \mu\text{m}$ as for the TT.

Outer Tracker

The Outer Tracker (OT) [46] is a gas detector, whose purpose is to measure the trajectory of charged particles in a large acceptance range. It is made of drift tubes with an inner diameter of $4.9\ \text{mm}$ arranged in modules. The straw tubes are filled with a gas mixture of Ar (70%), CO_2 (28.5%) and O_2 (1.5%) continuously flushed and purified in a closed loop. The drift time is below $50\ \text{ns}$ for the mean drift path of $2.45\ \text{mm}$, the spatial resolution is about $200\ \mu\text{m}$. The detector consists of three stations placed after the dipole magnet (Fig. 19). The three stations are of equal size with the outer boundary corresponding to an acceptance of $300\ \text{mrad}$ in the horizontal plane and $250\ \text{mrad}$ in the vertical one. The stations consist of four layers with the same layout described for the TT. Each layer contains two rows of tubes arranged in a honeycomb structure to maximise the sensitive area.

The magnet

In order to measure charged particle momenta a magnetic field is used to bend their trajectories. In LHCb this is achieved using a warm dipole magnet [47] placed between the first (TT) and the second group (T1-T3) of tracking stations. It consists of two coils of conical shape placed symmetrically with respect to the beam pipe, and with an angular opening following the LHCb acceptance, i.e. $\pm 250\ \text{mrad}$ in the non-bending plane and $\pm 300\ \text{mrad}$ in the bending plane. It is made of 15 layers of $Al - 99.7$ hollow conductor, grouped in five triplets. The magnet yoke surrounding

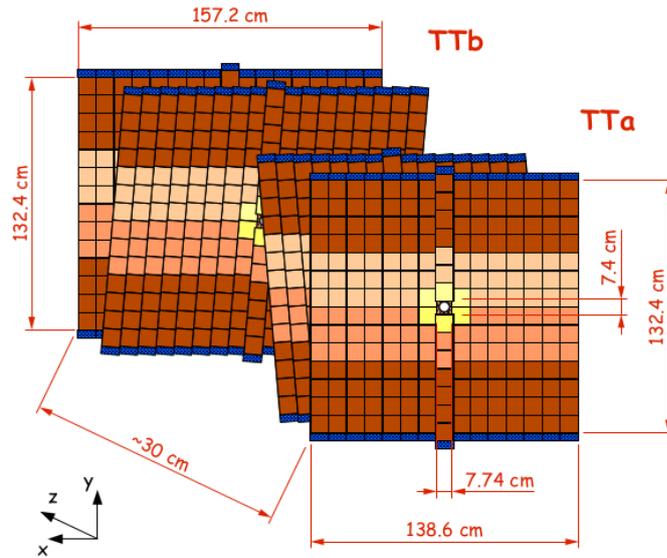


Figure 17.: Layout of the four TT layers. The front and rear planes have sensors vertically arranged, while the two planes in the middle represent the u -plane and v -plane tilted by $\pm 5^\circ$ respectively [39].

the coils are plates of laminated low carbon steel. In Fig. 20 a schematic layout is shown.

The magnet was designed to produce an integrated magnet field of 4 Tm mainly in the y direction and to leave just a residual field in the RICH1 detector. To achieve the required momentum resolution on charged particles the integrated field has to be known with a relative precision of 10^{-4} . An important feature of the magnet is the possibility to change the polarity of the field. This allows for the evaluation of systematics introduced by the (small) left-right asymmetry of the detector.

Track reconstruction

The tracking detectors (VELO, TT, IT and OT) are used to reconstruct tracks of charged particles and to determine their momentum. A schematic representation of the track types reconstructed by the tracking algorithm is shown in Fig. 21. The tracks can be classified as follows:

- *VELO tracks* contain only hits in the VELO detector. The associated particles are usually produced with a wide angle with respect to the beam-line, going

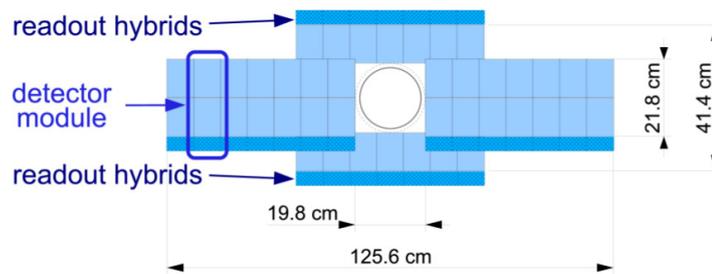


Figure 18.: Layout of one of the IT layers. The different modules placed just around the beam-pipe are shown [39].

out of the detector geometrical acceptance before traversing the TT. For these tracks the measurement of their momentum is not possible;

- *T tracks* are reconstructed from track segments appearing only in T1-T3 stations. They are typically produced in secondary interactions;
- *Upstream tracks* have hits on the VELO and TT detectors. They are usually low momentum tracks that are bent out of the detector acceptance by the magnetic field before they can reach T1. For these tracks it is possible to measure their momentum, thanks to the residual magnetic field in the region between VELO and TT, but with a relative uncertainty of about 20%;
- *Downstream tracks* are T-tracks reconstructed with additional hits from the TT detector;
- *Long tracks* have hits in the VELO, in the main tracker and possibly in the TT. The momentum measurement for these tracks is the most precise.

The relative momentum resolution is between $\delta p/p = 0.35\%$ for low momentum tracks ($\sim 10 \text{ GeV}/c$) and $\delta p/p = 0.55\%$ for high momentum tracks ($\sim 140 \text{ GeV}/c$).

For the analysis presented in this thesis *Downstream* and *Long* tracks are used to reconstruct K_s^0 decays into two charged pions, as they are neutral long-lived particles often decaying between the VELO and the TT. In the following we refer to Down-Down (DD) or Long-Long (LL) to indicate a decay in which is present a K_s^0 candidate created by using two Downstream or two Long tracks respectively.

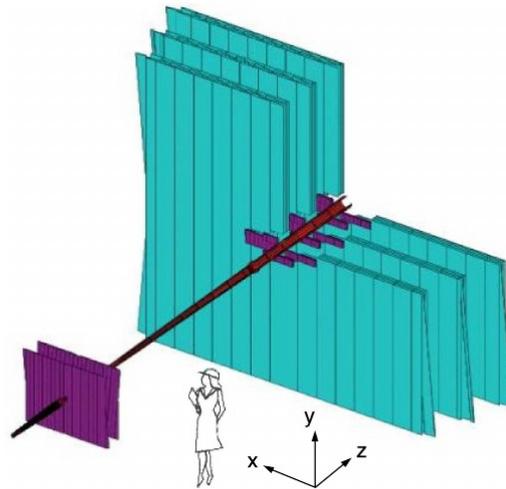


Figure 19.: Overall layout of the Outer Tracker (light blue) shown together with the TT, IT and the beam-pipe [39].

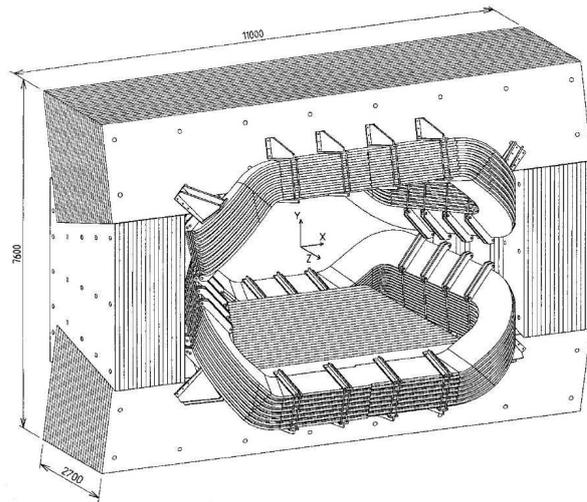


Figure 20.: Layout of the LHCb dipole magnet. In this figure the interaction point lies behind the magnet [39].

3.2.3 Particle Identification

Good particle identification (PID) is a crucial ingredient for the LHCb physics programme. In particular to reconstruct hadronic B decays, like those studied in this analysis, is very important to distinguish pions from kaons. Moreover to measure the energy of neutral hadrons and photons a good calorimetry system

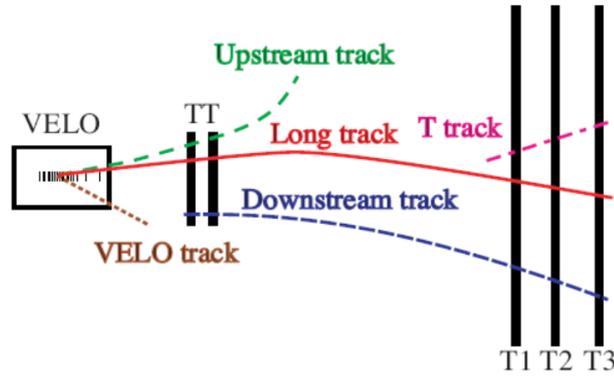


Figure 21.: Scheme of the different LHCb track types [39].

is essential. Finally the identification of muons is of primary importance for the analysis of leptonic decays. The different particle identification detectors used in LHCb are presented in the following: Ring Imaging Cherenkov detectors (RICH1 and RICH2), Hadronic and Electromagnetic Calorimeters (ECAL and HCAL) and muon detectors.

The RICH detectors

The LHCb particle identification system comprises two Cherenkov light detectors [48]. These detectors exploit the Cherenkov effect, which is explained as follows. When a charged particle traverses a medium with a velocity larger than the speed of light in that medium, it emits photons. The velocity of the particles can be found by measuring the angle, θ_C , between the Cherenkov photon and the direction of the particle using the relation

$$\theta_C = \frac{1}{n\beta} \quad (35)$$

where $\beta = v/c$, with v velocity of the particle, c velocity of the light in the vacuum and n the refraction index of the radiator. Thus combining this velocity measurement with a momentum measurement it is possible to determine the mass of a particle and as a result, its identity.

LHCb is instrumented with two Ring Imaging Cherenkov detectors (RICH1 and RICH2) which are able to efficiently discriminate charged hadrons in the momentum range between few GeV up to about 150 GeV. Most of the particles coming from B hadron decays have a momentum in this range. The RICH1 detector (Fig. 22 left) is located between the VELO and the TT. It covers the full LHCb

acceptance and uses aerogel with $n = 1.03$ and C_4F_{10} gas with $n = 1.0014$ as radiators. This provides identification of charged particles with low momenta from 1 to 60 GeV/c. The RICH2 detector (Fig. 22 right) is located directly behind the T stations and has a limited acceptance of ± 25 mrad to ± 120 mrad (horizontal) and ± 100 mrad (vertical). This detector allows to perform identification of charged particles with an higher momentum from around 15 GeV up to 150 GeV using CF_4 gas with $n = 1.0005$. Figure 23 shows the Cherenkov angle as a function of the momentum for different particles. The two detectors are equipped with an optical system made of spherical and flat mirrors able to image the emitted Cherenkov light on a lattice of photo-detectors (Hybrid Photon Detector, HPD). The HPDs are placed outside of the detector acceptance, where the particle multiplicity is low, and are shielded by iron to protect them from the residual magnetic field.

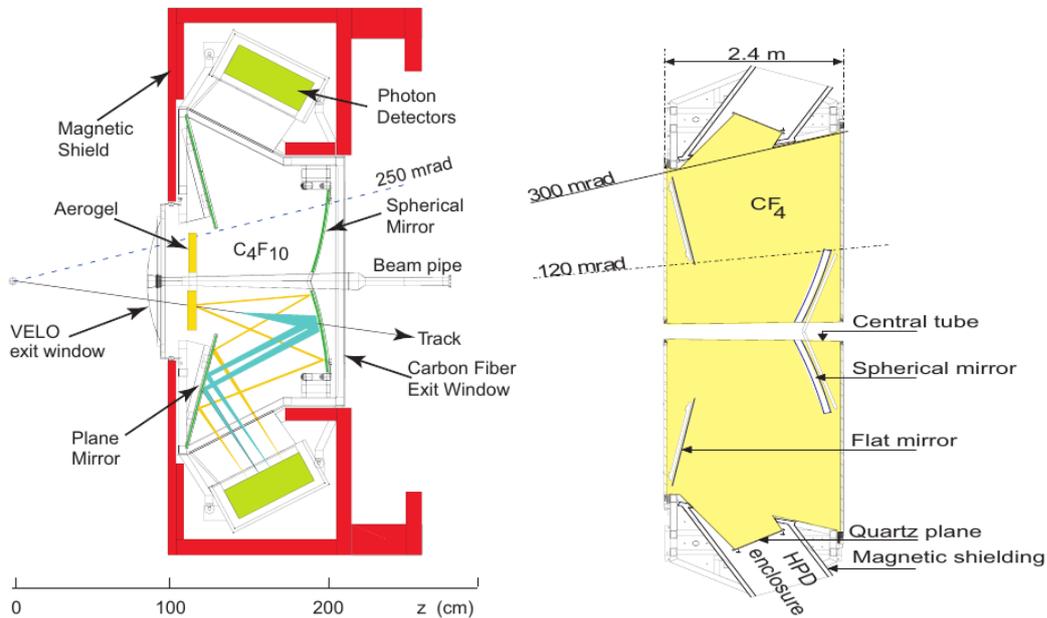


Figure 22.: Left: schematic layout from a side view of the RICH1 detector. The Cherenkov light as emitted by a charged track traversing the Aerogel tiles and the C_4F_{10} radiator is also drawn. Right: Schematic layout from a top view of the RICH2 detector [39].

The Cherenkov photons emitted by a particle are characterized by the same emission angle θ_C and form a ring on the HPD plane, with radius proportional to θ_C . Since the projection of the Cherenkov photons is outside the LHCb acceptance and because of imperfections of the imaging system, the images formed by photons on the photon detector planes are not perfect circles. Instead of fitting a

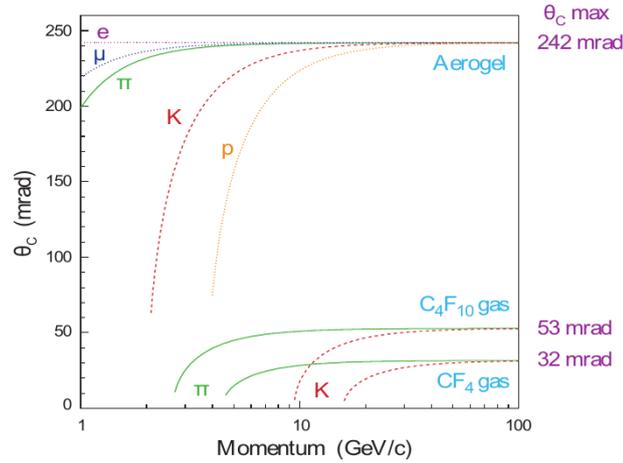


Figure 23.: Cherenkov angle θ_C as a function of particle momentum p . The curves corresponding to the various radiators used in RICH1 and RICH2 and to the various particle types as reported [39].

distorted ring to determine the Cherenkov angle, two independent pattern recognition techniques are used to assign particle hypotheses to tracks: the “local” and the “global” method. In the local approach each track is treated independently and a value of θ_C is determined for every hit in the photon detector plane. A likelihood function is calculated for each mass hypothesis. However in case of overlapping rings, hits from other tracks can form a background. The performance is especially poor in regions with high multiplicity of tracks. For this reason a global approach has been developed, which implements a simultaneous fit to all track hypotheses. This method takes into account the information from both RICH detectors. A likelihood function is calculated for the entire event by comparing the number of photoelectrons in each pixel with the number expected in that pixel from all possible sources, both signal and background. In this method the change of the likelihood depends only on the mass hypothesis assigned to the tracks. This quantity is considered for electrons, pions, kaons, protons and muons. However, in order to reduce the number of possible likelihoods, the pion mass hypothesis is assigned to all tracks and an initial value for the likelihood computed. Then for a given track, the mass-hypothesis is changed to electron, muon, kaon and proton, without changing all the other track hypotheses. The mass-hypothesis which gives the best likelihood is then set for that track. Then an iteration procedure which changes the mass-hypothesis for all tracks is performed until no improvements in the likelihood is found. The discriminating variable for the particle identification

is the so called $\Delta \log \mathcal{L}$. The pion mass hypothesis is taken as reference and the discrimination between the mass hypotheses is performed exploiting the difference of the likelihood logarithm under a generic hypothesis and the pion hypothesis. In this way the $\Delta \log \mathcal{L}$ for the pion hypothesis is always zero. In particular the variables used in this analysis are the $\Delta \log \mathcal{L}_{K\pi}$, that is the difference between the logarithm of the likelihood under the K hypothesis and under the π hypothesis

$$\Delta \log \mathcal{L}_{K\pi} = \log \mathcal{L}_K - \log \mathcal{L}_\pi \equiv DLL(K - \pi), \quad (36)$$

and $\Delta \log \mathcal{L}_{p\pi}$, the difference between the logarithm of the likelihood under the p hypothesis and under the π hypothesis

$$\Delta \log \mathcal{L}_{p\pi} = \log \mathcal{L}_p - \log \mathcal{L}_\pi \equiv DLL(p - \pi). \quad (37)$$

The performance of the RICH detectors has been studied using simulated and real data samples. In particular decays like $K_s^0 \rightarrow \pi^+\pi^-$, $\Lambda \rightarrow p\pi^-$ and $D^{*+} \rightarrow D^0(K^-\pi^+)\pi^+$ allow to select pure high statistics samples of pions, kaons and protons without using RICH information. Since θ_C depends on particle momentum also $\Delta \log \mathcal{L}$ depends on that. Fig. 24 shows the efficiency to identify or misidentify a particle, as a function of momentum, for pions, kaons and protons achievable with two PID requirements.

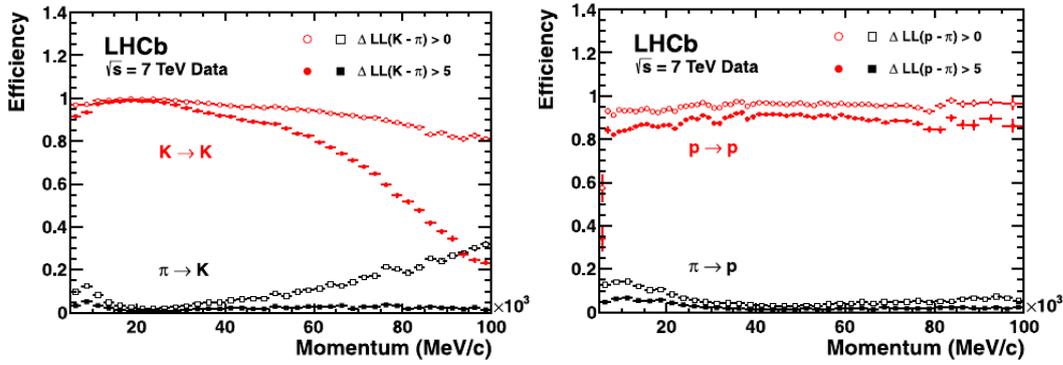


Figure 24.: Left: efficiency for the identification of kaons (red) and probability for the misidentification of pions as kaons (black) as a function of particle momentum, for different PID requirements (solid and empty points). Right: efficiency for the identification of protons (red) and probability for the misidentification of pions as protons (black) as a function of particle momentum, for different PID requirements (solid and empty points) [49].

The calorimeter system

The purpose of the calorimeter system [50] is the identification of neutral hadrons, electrons and photons and the measurement of their energy. In addition it provides information for the Level-0 trigger (see Sec. 3.2.4) by selecting high transverse energy (E_T) particles. Located downstream of RICH2, between the first two muon stations, the calorimeter system is divided into four sub-detectors:

- Scintillator Pad Detector (SPD);
- Pre-Shower (PS);
- Electromagnetic Calorimeter (ECAL);
- Hadronic Calorimeter (HCAL).

The entire calorimeter system is segmented into tiles in the x - y plane. The segmentation increases in dimension for different sections depending on the distance from the beam-pipe. This choice is motivated by a compromise between a good resolution in energy and position of the clusters and the need to limit the number of read-out channels. In particular the SPD/PS and ECAL are segmented in three sections with scaling widths in a projective design. The HCAL is segmented in only two sections with larger sizes. A schematic representation of this segmentation scheme is shown in Fig. 25. SPD and PS are auxiliary subdetectors of ECAL placed before it. The SPD provides electron and photon separation, while PS is used for a better discrimination between electrons and pions both at the trigger level and in the offline reconstruction. Both sub-detectors consist of scintillator

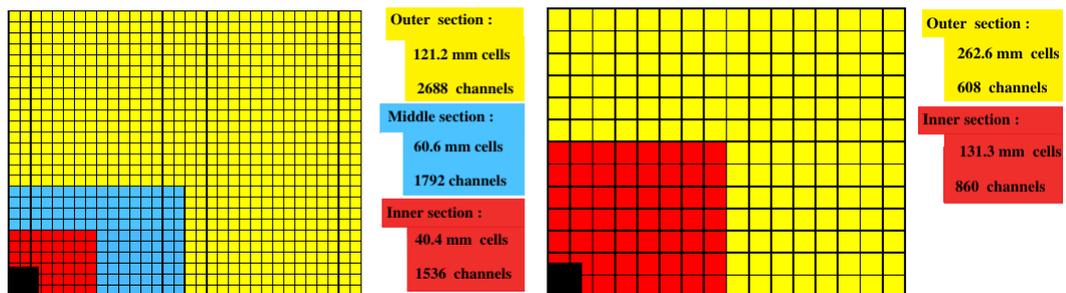


Figure 25.: Lateral segmentation of the SPD/PS and ECAL (left) and HCAL (right). One quarter of the detector front face is shown. In the left figure the cell dimensions are also given. The black area corresponds to the empty space occupied by the beam pipe [39].

material and they are separated by a 15 mm thick lead converter layer to convert photons to charged particles. The total material budget corresponds to about 2.5 – 3 radiation lengths. Wave length shifting (WLS) optical fibers are used to collect the light produced inside the scintillator, on multi-anode photo-multipliers.

The ECAL is a sampling calorimeter designed with the Shashlik technology. It consists of 66 alternating layers of lead absorber (2 mm thick) and scintillator pads (4 mm thick). To read out the scintillator light signals standard photomultipliers are used. The ECAL is segmented into cells, covering three regions and following the LHCb projectivity. The total material budget for each cell is about 25 radiation lengths and 1.1 nuclear interaction lengths. Its energy resolution is given by

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\% \quad (38)$$

where the energy is measured in GeV. The main purpose of the HCAL is the measurement of energies of hadronic showers needed by the Level-0 hadronic trigger. Its structure is similar to the ECAL, but each module is built by layers of scintillators 4 mm thick separated by layers of steel 16 mm thick. The total material budget corresponds to 5.6 nuclear interaction lengths. The energy resolution is given by

$$\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\% \quad (39)$$

where the energy is measured in GeV.

Muon System

Muon identification is vital to study rare and CP -sensitive semi-leptonic b -decays. The muon detector [51] consists of five stations (M1-M5), covering an angular acceptance of ± 300 mrad in the horizontal plane and ± 200 mrad in the vertical plane. To reduce the effect of multiple scattering effects from the calorimeter material, M1 is placed before the calorimeter system. The other stations are placed after the HCAL and they are interleaved by 80 cm thick iron walls to absorb hadronic particles. The layout of the five stations is shown in Fig. 26. Each muon station is divided into four regions with increasing distance and segmentation from the beam-pipe. This choice for the geometry ensures approximately equal charged particle occupancy in each region. Multi Wire Proportional Chambers (MWPC) technology was chosen for all regions, except for the M1 inner region where triple Gas Electron Multiplier (GEM) are used. The MWPCs are arranged in groups of four anode wire layers. The wires, spaced by 2 mm, have a length between 20 and

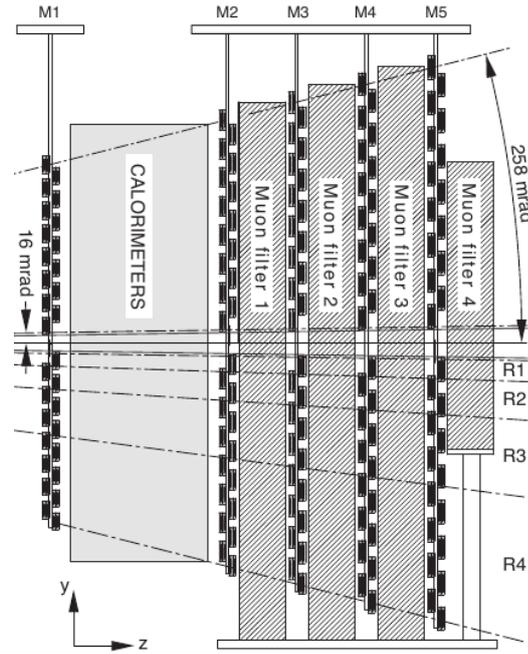


Figure 26.: Side view of the five stations of the muon detector [39].

30 cm and sit between a pair of cathode plates separated by 5 mm. The wire layers have a structure with four overlapping gaps, in order to increase the detection efficiency for single hit. This design allows for an high detection efficiency ($> 95\%$) and a fast response (~ 5 ns).

3.2.4 Trigger

The LHCb detector is designed to operate at an average luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, reduced by a factor 50 with respect to the LHC nominal luminosity. The aim is to limit the average number of pp collisions per event, which ensures a fast and efficient track reconstruction. During the 2011 running the average rate of visible interactions was ~ 10 MHz (the design value was 40 MHz). However the final output rate could not exceed ~ 3 kHz, due to technical limitations in the amount of data that can be stored and the processing time required. The rate reduction is achieved using a two-level trigger scheme [52]: the Level-0 (L0) hardware trigger, and the software-based High Level Trigger (HLT). The trigger layout, schematised in Fig. 27, satisfies the event rate constraint while efficiently selecting the events of interest.

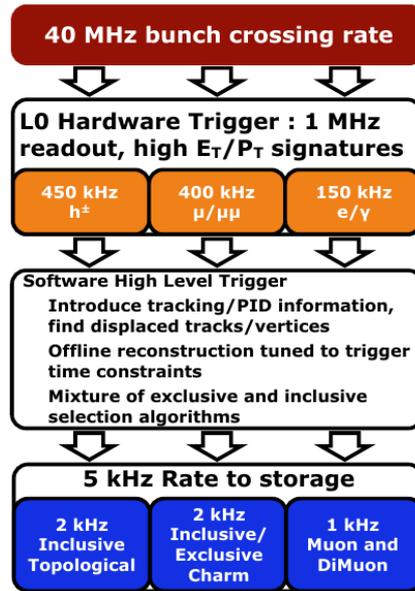


Figure 27.: LHCb trigger schema. The bunch crossing rate given here is only the nominal rate also including empty bunches [39].

The Level-0 Trigger

The Level-0 trigger (L0) is designed to reduce the event rate from 10 to ~ 1 MHz. It operates synchronously with the 40 MHz LHC clock: at this rate, only partial and prompt information from selected sub-detectors can be processed. Three sub-systems contribute to the L0 decision: the muon system, the pile-up system and the calorimeters. The collected information, for each event, is sent to the L0 decision unit (L0-DU) which performs the OR of the three subsystem decisions. Events are accepted when at least one of the subsystems gives a positive decision. The time between the pp interaction and the arrival of the L0 trigger decision to the front-end electronics is fixed to $4 \mu\text{s}$. The discriminating variables are estimates of the transverse momentum (p_T) of muons, the multiplicity of primary interactions and the transverse energy (E_T) deposited in the calorimeters.

The L0 muon trigger uses tracks reconstructed from hits in the muon system only, to allow for a fast momentum estimation. The processors in each sector use the hits in the pads and the nominal position of the pp interaction to form tracks. The trigger then selects the two muon candidates with the highest p_T in each quadrant. The event is triggered if at least one muon candidate has a transverse momentum greater than a threshold or if the product of the transverse momentum of a muon pair exceeds a given threshold.

The calorimeter trigger searches for high transverse energy particles using the information from the whole calorimeter system. Energy clusters are formed by adding the E_T of 2×2 calorimeter cells. The clusters are then associated to electron, photon or pion candidates. For each event only the cluster with highest E_T is considered.

The pile-up system, placed upstream of the VELO, is designed to detect events with multiple visible interactions. The system is equipped with two R sensor modules, used to estimate the primary vertex positions. For each pair of hits in the two sensors, an estimate of the primary vertex z position (PV_z) from which the particle originated is given.

The L0 trigger selectivity may depend on the running conditions. At the nominal conditions (pp interactions at $\sqrt{s} = 14$ TeV and $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$) typical requirements are at least one HCAL cluster with $E_T > 3.5$ GeV (hadrons), an ECAL cluster having $E_T > 2.5$ GeV (e^\pm , γ , π^0), a single muon with $p_T > 1.2$ GeV/ c or two muons with $p_{T,1} \times p_{T,2} > (1.3)^2$ (GeV/ c)².

The High level trigger

All the detector information of the events selected by the Level-0 trigger is passed to the Event Filter Farm (EFF). This is composed of about 1000 multi-core computing nodes, where the High Level Trigger (HLT) is run. The HLT is a software based application which performs a fast full reconstruction of events. The HLT is implemented in two steps: HLT1 and HLT2. HLT1 has to confirm the L0 decision and to further reduce the rate to 30 kHz, where the full pattern recognition can be performed. At this rate the HLT2 exploits the whole event information to reconstruct and select inclusive and exclusive final states.

HLT1

Events with high occupancy, especially in the OT, can take a processing time larger than 25 ms. Thus the HLT1 rejects events with an occupancy higher than 20%. For the remaining events the reconstruction is divided into two steps. In the first one VELO tracks are reconstructed asking for a large impact parameters with respect to the closest PV and for a minimum number of hits in the VELO. To kill possible ghost tracks, the expected number of hits in the VELO for each track is computed considering the track direction and its first hit in the detector. If the difference between this number of hits and the one used to reconstruct the tracks is greater than a certain threshold the track is rejected. The number of tracks selected by this first step allows to perform the full reconstruction. The tracks reconstructed in the

forward direction are then selected imposing minimal p and p_T requirements. The selected tracks are finally fitted using a bi-directional Kalman filter, to obtain an offline-quality value for the track χ^2 and an offline-quality covariance matrix.

HLT2

The HLT2 does a first full event reconstruction, based on a Kalman fitter algorithm, intended to be as close as possible to the final offline reconstruction. HLT2 requests are mainly based on inclusive selections, the so called *topological lines*, also used for this analysis. In addition there are some dedicated lines used by specific decay channels. The strategy used for the topological line is to build multibody candidates, starting from two input particles combined to form a two body object and adding one by one other candidates satisfying certain criteria on the distance of closest approach.

TIS and TOS

Before writing an event to tape it has to be necessarily selected by the trigger. However it is possible also to select events after the reconstruction of a decay candidate depending on which particle in the event caused the trigger to fire. If the trigger was fired by particles used in the decay candidate this is referred to Trigger On Signal (TOS). If instead the trigger was fired by particles not used to reconstruct the decay this is referred to Trigger Independent of Signal (TIS).

3.2.5 Simulation in LHCb

The software package used in LHCb to simulate events is called Gauss [53]. The simulation procedure uses first Pythia [54] to collide protons and then EvtGen [55] to create particle decays. After the generation Geant4 [56] is used to propagate the created particles through a simulated version of the LHCb detector. This step makes use of the so called Detector Description Data Base (DDDB), which stores information about the size, shape and materials which the detector is made of. The final step of the simulation is done using the Boole [57] software which simulates the digital output of the detector. By running the Moore [58] package it is also possible to re-apply the software trigger on the simulated events from Boole.

3.2.6 Reconstruction and analysis

The capability to store the enormous amount of data delivered by LHC and collected by the experiments constitutes one of the key factors in High Energy Physics today. The computing resources to store, distribute and analyse the data are provided by a global infrastructure called *Worldwide LHC Computing Grid* (WLCG) [59]. The WLCG is organized in a multi-tier regional centre structure where each tier provides a specific set of services. The events filtered by the EFF are first stored at the Tier-0 at CERN. At this level the data contain information on the detector, like the number of hits in the tracking system or the read-out response of some sub-detectors. The file format is called RAW data. These data are then copied to different Tier-1s where they are further processed. The next step is the full event reconstruction, done in LHCb using the Brunel [60] package. The output of the reconstruction consists in physics quantities like primary vertex coordinates, track trajectories and momentum and energy of calorimeter clusters. This information is stored in a type of data called SDST. As far as simulated data are concerned the output from `Boole` is used from Brunel to apply the full reconstruction in the same way as data.

Only a small fraction of the data recorded by LHCb is stored on disk, since the amount of disk space required by the full data sample would be too expensive. Moreover the fraction of events useful for each single analysis is very small. Therefore the next stage of the analysis chain consists of a preselection of events named “Stripping”. The Stripping procedure applies loose selections on the collected data sample, filtering the sample and creating the candidates that will be used in the final analysis. The output files are named DSTs and contain the possible candidates, the information of the reconstruction phase and also the RAW data relative to each event. Stripping lines for similar physics programmes are grouped into “Streams”. The aim is to decrease the amount of time required to extract the information about the signal decays. Final DSTs can be used for user analysis and they are stored on disk for a faster and more efficient access. To save disk space, after the creation of the corresponding DST, the RAW and SDSTs files are transferred to magnetic tapes.

4

ANALYSIS STRATEGY AND SELECTION

The aim of this chapter is first to give a brief overview of the analysis strategy adopted to measure or to determine a limit on the branching fraction of the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ with respect to the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ channel. Then the data and Monte Carlo samples used for the analysis are outlined. Finally the studies on the background sources and the adopted selection criteria are reported.

4.1 ANALYSIS STRATEGY

The branching ratio studied in this thesis can be expressed with the following formula

$$\frac{\mathcal{B}(B_{d,s}^0 \rightarrow K^{*0} K_S^0)}{\mathcal{B}(B^0 \rightarrow K_S^0 \pi^+ \pi^-)} = \frac{f_d}{f_{d,s}} \cdot \frac{\varepsilon_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}^{\text{Acc}}}{\varepsilon_{B_{d,s}^0 \rightarrow K^{*0} K_S^0}^{\text{Acc}}} \cdot \frac{\varepsilon_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}^{\text{Sel}}}{\varepsilon_{B_{d,s}^0 \rightarrow K^{*0} K_S^0}^{\text{Sel}}} \cdot \frac{\varepsilon_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}^{\text{PID}}}{\varepsilon_{B_{d,s}^0 \rightarrow K^{*0} K_S^0}^{\text{PID}}} \cdot \frac{N_{B_{d,s}^0 \rightarrow K^{*0} K_S^0}}{N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}}, \quad (40)$$

where f_d and f_s are the fragmentation fractions of a b quark in a B^0 and B_s^0 meson respectively [61]. Here $N_{B_{d,s}^0 \rightarrow K^{*0} K_S^0}$ and $N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}$ represent the number of signal candidates for the main and the normalization decay, while ε^{Acc} is the LHCb acceptance efficiency, ε^{Sel} is the combined trigger, reconstruction, stripping and offline selection efficiencies and ε^{PID} is the PID efficiency.

The main challenge in this analysis is to establish a selection strategy which allows one to observe the relatively small number of signal yields compared to the background. This is achieved using standard cut-based methods and multivariate algorithms.

In order not to bias the final results it has been decided to perform a *blind analysis* for the main signals. This means that the development of a fit model to extract the yields relies largely on simulation. Before extracting the signal yields, the fit strategy has to be validated using pseudo-experiments, and the main systematic uncertainties have to be studied. The following rules were followed:

- The signal yield in the fit is kept blind;

- From the graphical point of view the data points are masked in the signal regions of the B and K^{*0} spectra.

Once the fit is unblinded and the number of signal candidates is determined, the strategy is the following:

- Quote the central value and the 1σ uncertainty of the $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ yield and of the ratio of branching ratios;
- Calculate the significance as the absolute difference in Delta Log Likelihood, including the systematic uncertainties, when the signal yield is floating or fixed to zero;
- Compute the 90% (95%) confidence level upper limit for the ratio of branching ratios from the integral of the likelihood function in the positive region.

4.2 DATA SAMPLES

The analysis is performed using proton-proton collision data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV recorded during 2011 operations with the LHCb detector. These correspond to an integrated luminosity of $\int L = 1025 \pm 36$ pb. The recorded integrated and instantaneous luminosity for the 2011 period can be seen in Fig. 28.

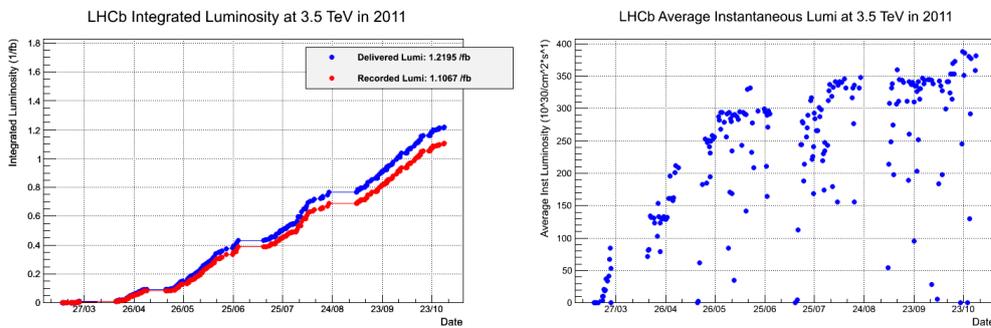


Figure 28.: Left: LHCb integrated luminosities for 2011. Right: LHCb average instantaneous luminosity during 2011 [41].

Many studies like the determination of the detector acceptances rely on simulated events. Proton-proton collisions are simulated using the Monte Carlo (MC) event generator Pythia [54] in a specific LHCb configuration [53]. Each event contains at least one B candidate. Only events with the candidates emitted into the

angular range from 0 to 400 mrad around the beam pipe (roughly corresponding to the detector acceptance) are further processed. To allow for comparisons between data and simulation, the trigger criteria and the stripping requirements are also applied to the simulation. The signal and background samples were generated for both magnet polarities and the number of events is reported in Tab. 5.

Table 5.: Number of generated signal and background Monte Carlo events for each decay mode for each magnet polarity.

Mode	Number of events	
	MagUp	MagDown
$B^0 \rightarrow K^{*0} K_s^0$	1006996	1021494
$B_s^0 \rightarrow K^{*0} K_s^0$	1029994	1007496
$B^0 \rightarrow K_s^0 K^\pm \pi^\mp$ (phSpace)	1012496	1025995
$B_s^0 \rightarrow K_s^0 K^\pm \pi^\mp$ (phSpace)	1012492	1025996
$B^+ \rightarrow D^0(K_s^0 \pi^+ \pi^-) K^+$	1506491	1515990
$B^+ \rightarrow D^0(K_s^0 K^+ K^-) \pi^+$	507996	554499
$B^+ \rightarrow \bar{D}^{*0}(\bar{D}^0(K_s^0 \pi^+ \pi^-) \pi^0) K^+$	54999	54499
$B_s^0 \rightarrow K^*(K\pi) K^*(K_s^0 \pi^0)$	100000	100000
$B^0 \rightarrow K_s^0 \pi^+ \pi^-$ (phSpace)	1026497	1012493
$B_s^0 \rightarrow K_s^0 \pi^+ \pi^-$ (phSpace)	936997	921096
$B^0 \rightarrow K_s^0 \pi^+ \pi^-$ (sqDalitz) ¹	921096	936997
$B_s^0 \rightarrow K_s^0 \pi^+ \pi^-$ (sqDalitz)	931392	918198
$B^0 \rightarrow \eta'(\rho^0 \gamma) K_s^0$	1012995	1010997
$B^0 \rightarrow \eta'(\eta \pi^+ \pi^-) K_s^0$	1007497	1009994
$B^0 \rightarrow K_s^0 \pi^+ \pi^- \gamma$	769496	762497
$B^+ \rightarrow D^0(K_s^0 \pi^+ \pi^-) \pi^+$	1013993	1013991
$B^+ \rightarrow \bar{D}^{*0}(\bar{D}^0(K_s^0 \pi^+ \pi^-) \pi^0) \pi^+$	102499	101500
$B^0 \rightarrow K^{*0}(K_s^0 \pi^0) \rho^0(\pi^+ \pi^-)$	773495	766495
$B^+ \rightarrow K^{*+}(K_s^0 \pi^+) \pi^+ \pi^-$	1040492	1014496

¹ The Square Dalitz model has the effect of spreading the edges of the Dalitz plot towards the centre (see App. A).

4.3 BACKGROUND STUDIES

In order to define an efficient selection strategy it is very important to understand which background pollutions might affect the signal purity. Many different types of potential background can be found for the $B_{d,s}^0 \rightarrow K^{*0}K_s^0$ and $B^0 \rightarrow K_s^0\pi^+\pi^-$ decays. They can be classified in four main categories:

- PEAKING background: fully reconstructed decays which have the same final state as the signal or the normalization;
- PARTIALLY RECONSTRUCTED background: decays where a pion or a kaon in the final state is not reconstructed;
- MISIDENTIFIED background: decays where a pion or a kaon is misidentified respectively as a kaon or a pion;
- COMBINATORIAL background: random combination of tracks not necessarily coming from the same particle decay which can form a vertex.

In the following a more detailed classification of the backgrounds is given.

4.3.1 Peaking backgrounds

- *Non resonant* $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$: the signal decays are reconstructed in the $K_s^0 K \pi$ final state. This means that any B decay to the correct final state can be potentially a signal B candidate. However only those which also contain a true $K^*(892)^0$ decay will peak in the $K\pi$ spectrum at the nominal mass of the $K^*(892)^0$. We define as "non resonant" those intermediate states which decay into the correct final state, but do not proceed via a real $K^*(892)^0$.

The previous background is present only in the $K_s^0 K \pi$ spectrum, while the following backgrounds instead affect both the signal and the normalization.

- *Charmonia transitions*: the decay of $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_s^0$ and those proceeding via higher mass charmonium states such as $B^0 \rightarrow \psi(2S)(\mu^+\mu^-)K_s^0$ have very similar topology to the signal. Due to the similar μ and π masses, the reconstruction of these decays does not shift the B invariant mass peak significantly, so they can appear in the mass window. The most important contribution comes from decays to muonic final state but there are also possible contributions from the hadronic decays.

- *Baryonic decays*: the difference in mass between B_s^0 and Λ_b is approximately $250 \text{ MeV}/c^2$. The Λ_b can decay to a $pK_s^0 h$ final state and, if the proton is misidentified as a pion or a kaon, reflections can appear in the signal mass window. The decay can proceed through charmed transitions ($\Lambda_c \rightarrow pK_s^0$) or charmless ones (not yet observed). The mass distribution of $K_s\pi\pi$ candidates (with the K_s^0 reconstructed as Down-Down) when one of the pions is reconstructed under the proton hypothesis, is shown in Fig. 29 on the left. The data mass distribution of the pK_s combinations belonging to the Λ_b is shown on the right of the same figure, where a clear peak at the Λ_c mass is visible. In both plots an additional requirement on the proton identification was used.

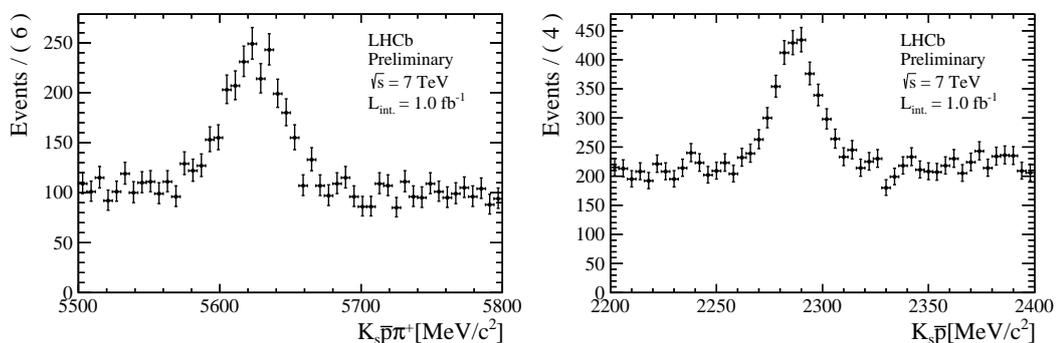


Figure 29.: Left: data mass of $K_s^0 \pi^- p$ combinations with Down-Down K_s^0 candidates. Right: pK_s^0 combinations with Down-Down K_s^0 candidates for $K_s^0 \pi^+ p$ combinations consistent with Λ_b mass.

- Several B decays to D intermediate states can peak in the $K_s^0 K \pi$ and $K_s^0 \pi \pi$ spectra. A list of them is given in Tab. 6, where also a summary of the previously discussed backgrounds is given.

4.3.2 Partially reconstructed decays

These decays contain photons or hadrons which are not used to reconstruct the B decay vertex. The missing track causes the measured B invariant mass to be shifted to lower values, potentially giving a contribution in the left part of the signal region. In this section a discussion of the relevant partially reconstructed backgrounds is done according to the spectrum which is affected.

Table 6.: Charmonium and charm background contributions.

Description	Decay
Charmonia with muons in the final state	$B^0 \rightarrow (J/\psi, \chi_c, \psi(2S) \rightarrow (\mu\mu))K_s$
Charmonia with hadrons in the final state	$B^0 \rightarrow (J/\psi, \chi_c, \psi(2S) \rightarrow (\pi\pi))K_s$
Baryonic decays	$\Lambda_b \rightarrow (\Lambda_c \rightarrow pK_s^0)\pi$
	$B^0 \rightarrow (D^0 \rightarrow K^-\pi)h$
	$B^0 \rightarrow (D^0 \rightarrow \pi^+\pi^-)h$
	$B^0 \rightarrow (D^0 \rightarrow K^+K^-)h$
Two body charm decays	$B^0 \rightarrow (D^+ \rightarrow K_s^0K^+)h$
	$B^0 \rightarrow (D^+ \rightarrow K_s^0\pi^+)h$
	$B^0 \rightarrow (D_s^+ \rightarrow K_s^0K^+)h$
	$B^0 \rightarrow (D_s^+ \rightarrow K_s^0\pi^+)h$

- Background of the $B_{d,s}^0 \rightarrow K^{*0}K_s^0$ signal decays:
we can classify the backgrounds into two categories.
 - *Charmless decays*: the $B_s^0 \rightarrow K^*(892)^0K^*(892)^0$ decay of B_s^0 into two vector mesons where one K^{*0} decays into a $K^+\pi^-$ final state and the other K^{*0} decays to the $K_s^0\pi^0$ final state with the π^0 not being reconstructed. In this background a real K^{*0} may be present and hence it can peak in the $K\pi$ spectrum. The LHCb experiment observed this channel in the charged final state [62].
 - *Charmed decays*: the $B^+ \rightarrow D^0(K_s^0K^+K^-)\pi^+$, $B^+ \rightarrow D^0(K_s^0\pi^-\pi^+)K^+$, $B^+ \rightarrow D^*(D^0(K_s^0\pi^-\pi^+)\pi^0)K^+$ decays where a D meson decays to three tracks, one of which is not reconstructed.

A summary of these backgrounds is shown in Tab. 7, where also their branching fractions are given.

- Background of the $B^0 \rightarrow K_s^0\pi^+\pi^-$ normalisation decay:
we can classify the backgrounds for this mode into four categories, of which the first two are the same as for the $B_{d,s}^0 \rightarrow K^{*0}K_s^0$ decays.
 - *Charmless decays*: $B^0 \rightarrow K^{*0}(K_s^0\pi^0)\rho^0(\pi^+\pi^-)$, $B^+ \rightarrow K^{*+}(K_s^0\pi^+)\pi^+\pi^-$, $B^0 \rightarrow \eta'(\eta\pi^+\pi^-)K_s^0$ where a pion is not reconstructed;
 - *Charmed decays*: $B^+ \rightarrow D^0(K_s^0\pi^+\pi^-)\pi^+$, $B^+ \rightarrow \bar{D}^{*0}(\bar{D}^0(K_s^0\pi^+\pi^-)\pi^0)\pi^+$ where a charged or neutral pion is not reconstructed;

- Resonant radiative charmless decays: $B^0 \rightarrow \eta'(\rho^0\gamma)K_s^0$ where a γ is not reconstructed;
- Non resonant radiative charmless decays: $B^0 \rightarrow K_s^0\pi^+\pi^-\gamma$ where a γ is not reconstructed;

These backgrounds are listed in Tab. 8, together with their branching fractions.

Table 7.: Partially reconstructed background for the $B_{d,s}^0 \rightarrow K^{*0}K_s^0$ decays.

Description	Decay	BF [17]
Charmless	$B_s^0 \rightarrow K^*(892)^0 K^*(892)^0$	$6.2 \cdot 10^{-6}$
	$B^+ \rightarrow D^0(K_s K^+ K^-)\pi^+$	$2.1 \cdot 10^{-7}$
Charmed	$B^+ \rightarrow D^0(K_s\pi^+\pi^-)K^+$	$1.1 \cdot 10^{-6}$
	$B^+ \rightarrow D^*(D^0(K_s\pi^+\pi^-)\pi^0)K^+$	$7.8 \cdot 10^{-9}$

Table 8.: Partially reconstructed background for the $B^0 \rightarrow K_s^0\pi^+\pi^-$ decay.

Description	Decay	BF [17]
Charmless	$B^0 \rightarrow K^{*0}(K_s^0\pi^0)\rho^0(\pi^+\pi^-)$	$1.1 \cdot 10^{-6}$
	$B^+ \rightarrow K^{*+}(K_s^0\pi^+)\pi^+\pi^-$	$1.4 \cdot 10^{-5}$
	$B^0 \rightarrow \eta'(\eta\pi^+\pi^-)K_s^0$	$2.5 \cdot 10^{-5}$
Charmed	$B^+ \rightarrow D^0(K_s^0\pi^+\pi^-)\pi^+$	$1.4 \cdot 10^{-5}$
	$B^+ \rightarrow \bar{D}^{*0}(\bar{D}^0(K_s^0\pi^+\pi^-)\pi^0)\pi^+$	$9.6 \cdot 10^{-6}$
Resonant radiative	$B^0 \rightarrow \eta'(\rho^0\gamma)K_s^0$	$9.7 \cdot 10^{-6}$
Non resonant radiative	$B^0 \rightarrow K_s^0\pi^+\pi^-\gamma$	$9.6 \cdot 10^{-6}$

4.3.3 Misidentified background

Decays with a $K_s hh'$ final state where the hh' can proceed also through resonances can potentially affect the signal yields. The background can be classified according to the signal or the normalization decay as follows:

- Background of the $B_{d,s}^0 \rightarrow K^{*0}K_s^0$ signal decays:
 - The normalization mode, $B^0 \rightarrow K_s^0\pi^+\pi^-$ can be a background for the signal,

if a pion is misidentified as a kaon;

- The same final state as the normalization decay may be reached through resonances like $f_0(980)$ and $\rho(770)$ which are rather broad and hence can be present in the K^{*0} mass window;
 - The $B_{d,s}^0 \rightarrow K_s^0 K^+ K^-$ can also be a background if a kaon is misidentified as a pion.
- Background of the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ normalization decay:
 - The inclusive $B^0 \rightarrow K_s^0 K^\pm \pi^\mp$ decay can be a background if a kaon is misidentified as a pion;
 - The $B_{d,s}^0 \rightarrow K_s^0 K^+ K^-$ is disfavored since it would require a double misidentification of kaons in pions.

4.3.4 Combinatorial background

This background is due to the reconstruction of a vertex from random tracks. The two kind of tracks used in this analysis to build the K_s^0 candidates (Long and Downstream) are very different in this respect. In the Downstream track sample, an higher percentage of fake combinations is observed.

4.4 SELECTION

In order to reduce the different sources of background, the offline analysis strategy consists of two steps. In the first step a preselection is used to mostly reduce the physics background, by exploiting rectangular cuts, while for the second step a multivariate analysis (MVA) has been designed to reject the combinatorial background, which is expected to be the dominant contribution under the signal peak.

4.4.1 Variable definitions

Before discussing the cuts applied in the different stages of the selection in detail, in this section the most important variables are defined, and some of them are explained using diagrams.

- TRANSVERSE MOMENTUM (p_T)

It is the momentum of a particle in the x - y plane as shown in Fig. 30.

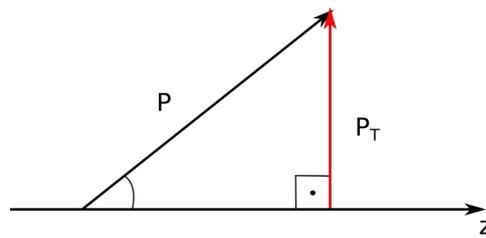


Figure 30.: Diagram for the definition of the transverse momentum.

- REDUCED χ^2 OF THE VERTEX (χ^2/ndf)

The decay vertex of a particle is reconstructed via a fit procedure which uses the daughter tracks. To this fit is associated a χ^2 which expresses how confident one can be that the particles indeed come from the same vertex. To have a good reconstructed vertex a $\chi^2/ndf \sim \mathcal{O}(1)$ is required.

- IMPACT PARAMETER (IP)

It is the minimum distance between the reconstructed track and the primary vertex (Fig. 31). It is useful to distinguish between particles coming from the primary vertex or short-lived resonances, and daughter particles coming from long-lived particles like B mesons. On average, a B daughter particle has a larger IP than a particle produced at the primary interaction. Usually the χ^2 of this parameter ($IP\chi^2$), defined as the increase in the vertex-fit χ^2 when including this track, shows a better separation than the IP itself.

- FLIGHT DISTANCE (FD)

It is the distance between the decay vertex of a particle and the origin vertex (Fig. 32), that is strongly connected with the lifetime of the particle. Usually a better separation between signal and background is achieved with the χ^2 of this parameter ($FD\chi^2$), that is the FD-significance squared.

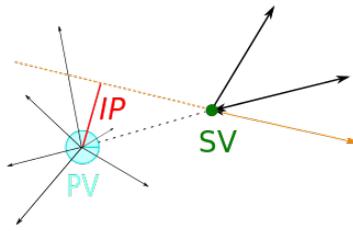


Figure 31.: Diagram for the definition of the impact parameter.

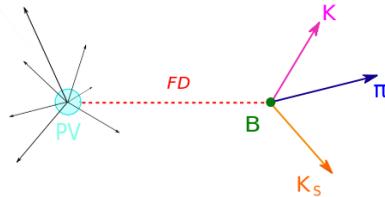


Figure 32.: Diagram for the definition of the flight distance.

- DIRA

It is the cosine of the angle between the flight direction and the reconstructed momentum of a particle (Fig. 33). The flight direction is identified by the vector connecting the production vertex and the decay vertex of the particle.

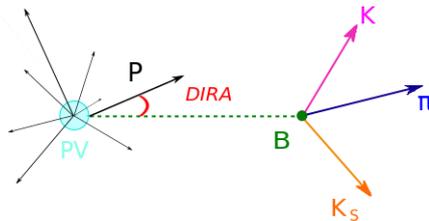


Figure 33.: Diagram for the definition of the DIRA.

- DISTANCE OF CLOSEST APPROACH (DOCA)

It is the distance in space between two tracks, which can be used in rejecting particles which are not coming from a certain decay vertex.

4.4.2 Multivariate analysis

Multivariate analysis is a powerful tool, used more and more in the recent measurements in high energy physics. This technique allows one to combine multiple variables into a single discriminant, instead of studying them independently. The signal-background separation can be obtained by selecting regions with linear or non-linear cuts, and in general this flexibility to choose the regions is more efficient than simple rectangular cuts. The concept is exemplified in Fig. 34.

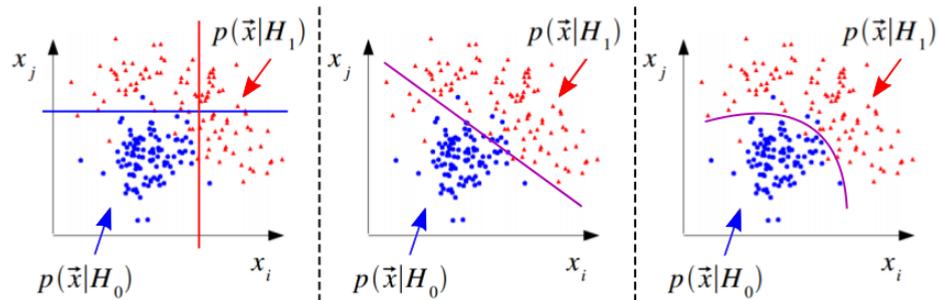


Figure 34.: Illustration of multivariate selections. The two data species are H_0 and H_1 (for example signal and background), while the cut variables are x_i and x_j . Left: classic selection using rectangular cuts. Middle: multivariate linear discriminant. Right: multivariate non-linear discriminant.

Among many different methods the Boosted Decision Tree (BDT) is of particular relevance for this analysis. Here only a short explanation about its functioning is given, for a more detailed review, also on the other multivariate methods, see Ref. [63].

A decision tree is a classifier able to identify different categories (for example signal and background) by making repeated yes/no decisions of a single variable, chosen from a list, as shown in Fig. 35. At each node, starting from the "root" node, the cut on the variable which gives the best signal-background separation is applied. At every split a certain number of cuts on each of the variable is tested to choose the most discriminant, according to a certain criterion. The criterion used for this analysis is the "Gini index", defined as $p \times (1 - p)$, where the purity $p = S/(S + B)$, with S the number of signal events and B the number of background events. In this way the variable space is splitted in different regions, which are classified in a certain category. The decision tree continues until a stopping criterion is fulfilled.

Decision trees are powerful but unstable classifiers, since their separation power depends on the statistical fluctuations of the input samples. A method used to make a BDT more robust and to enhance the performance is the so called "boosting": this allows one to build a forest of different decision trees combined recursively by reweighting misidentified events from a given tree and to classify an event on a majority vote of the classifications done by each tree in the forest. Among the different boosting procedures, the two most commonly used are the GradientBoost [64] and the AdaBoost [65]. See Ref. [63] for more details. The gradient boosting is found to give the better performance for this analysis.

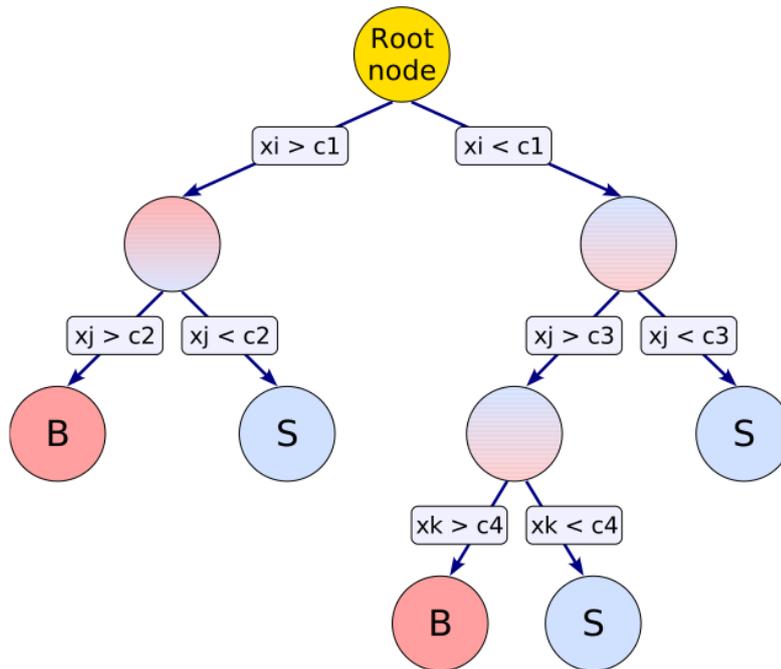


Figure 35.: Schematic view of a decision tree. At each node the most discriminant variable to separate different categories is cut on. A set of discriminant variables x_i, x_j, \dots is tested, and the best cut c_1, c_2, \dots is determined at every split [63].

4.4.3 Trigger strategy

As described in Sec. 3.2.4, LHCb uses a two-level trigger scheme, one hardware and two software trigger levels. The topology of a three body hadronic decay has a distinctive signature. Therefore, the same trigger requirements can be used

for the signal and the normalization. The trigger requirements at each level are summarized in Tab. 9 and described in detail below.

Table 9.: Trigger selection requirements after reconstruction of decays.

Trigger Level	Requirement
L0	L0Hadron TOS OR L0Global TIS
HLT1	Hlt1TrackAllL0 TOS
HLT2	(2,3,4)-BBDT TOS OR (2,3,4)-Simple TOS

The first stage corresponds to the L0 trigger, where the events that fired either L0Hadron TOS or L0Global TIS decisions are collected [66]. As already mentioned in Sec. 3.2.4, L0 is divided into three independent triggers (L0-Calorimeter, L0-Muon and L0-PileUp), according to the subdetector giving the information. Of particular interest in this analysis is the L0-Calorimeter trigger which compares the transverse energy of the candidates deposited in a clusters to a fixed threshold. If an event contains at least one candidate above threshold then it is retained.

At HLT1 level, the Hlt1TrackAllL0 TOS line is required [67]. The extrapolation of all the track segments in the VELO to the tracking stations would be very time expensive. For this reason some requirements are made on the impact parameter of the VELO track, on the number of hits on the track and the number of missing hits, defined as the difference between the number of hits on the track and the number of expected hits, given the first measured point on the track and its direction. In order to reduce the rate, the final tracks are further selected, applying cuts on the their momentum and transverse momentum. The remaining tracks are fitted and then selected using their track χ^2/ndf and $IP\chi^2$. The cuts applied in this trigger line are listed in Tab. 10.

The HLT2 makes use of the 2-, 3- or 4-body topological triggers named as Hlt2Topo[2,3,4]BBDT and Hlt2Topo[2,3,4]Simple [68], which use either a multivariate selection algorithm (Bonsai Boosted Decision Tree) or a “Simple” decision respectively. The candidates selected as TOS in those lines are accepted in the analysis. The HLT2 trigger lines start applying a preselection on the momentum, the transverse momentum, the χ^2 and the $IP\chi^2$ of the tracks. The tracks passing further requirements on the sum of their transverse momenta, their invariant mass, their distance of closest approach (DOCA) and their flight distance χ^2 with respect to the PV, can be combined to form vertices. The cuts applied at this stage are shown in Tab. 11.

Table 10.: Hlt1 trigger selection requirements.

Variable definition	Selection requirement
VELO tracks	
Impact parameter	100 μm
Number of hits	> 10
Number of missing hits	< 2
Final tracks	
P	$> 10 \text{ GeV}/c$
p_T	$> 1.7 \text{ GeV}/c$
χ^2/ndf of track fit	< 2.5
IP χ^2	> 16

Table 11.: Hlt2 trigger selection requirements on the candidates.

Variable definition	Selection requirement
Preselection	
p_T	$> 1700 \text{ MeV}/c$
P	$> 10 \text{ GeV}/c$
χ^2/ndf of track fit	< 2
IP χ^2	> 16
2-3-4 body combinations	
Σp_T	$> 3, 4, 4 \text{ GeV}/c$
Invariant mass	$< 7 \text{ GeV}/c^2$
DOCA	$> 0.2\text{mm}$
FD χ^2	> 100

An important feature for the analysis presented in this thesis is the possibility to include the K_s^0 as input particle when forming the candidates. Two pions passing requirements on the track χ^2 and the $\text{IP}\chi^2$ are combined to form a vertex. The candidate is accepted as a proto- K_s^0 if the vertex χ^2 and the invariant mass pass certain criteria. Finally the candidates are further selected by requirements on the momentum, the transverse momentum, the $\text{IP}\chi^2$ and the $\text{FD}\chi^2$.

Table 12.: Hlt2 trigger selection requirements on the K_s^0 candidates.

Variable definition	Selection requirement
Pions	
track χ^2	< 3
$\text{IP}\chi^2$	> 16
K_s^0	
vertex χ^2	< 10
$ m_{\text{Rec}} - m_{\text{PDG}} $	$< 30 \text{ MeV}/c^2$
p_T	$> 500 \text{ MeV}/c$
P	$> 5 \text{ GeV}/c$
$\text{IP}\chi^2$	> 4
$\text{FD}\chi^2$	> 1000

In the last step the topological lines are constructed by applying a multivariate selection, based on a boosted decision tree (BDT), to the previous filtered candidates.

4.4.4 Preselection

The preselection consists of two phases, the first one is the stripping, which is controlled centrally in LHCb, and the second one which consists of some additional offline requirements.

Stripping

The events used for this analysis are those passing the B2KShh stripping line. This stripping line was implemented for the inclusive measurement of the branching fractions of the six $B_{d,s} \rightarrow K_s^0 hh'$ decay modes [9]. The requirements contain different steps:

- *Global event cut (GEC)*

This is the fastest possible cut which assumes only minimal track reconstruction and retains events with fewer than 250 Long tracks.

- *Daughter track requirements*

The B decay vertex is formed by combining a K_s^0 candidate with two oppositely charged tracks and applying a selection on the following variables, associated with the candidate daughter tracks:

- The total momentum: necessary for the separation of the Cherenkov opening angles in the RICH subdetectors;
- The transverse momentum;
- The particle identification requirement of the kaon with respect to the pion;
- The Clone distance²: characterised by the Kullback-Leibler distance [69];
- The smallest impact parameter of the track computed with respect to all the reconstructed primary vertexes;
- The reduced χ^2 of the track reconstruction.

The cuts are listed in Tab. 13. A particular feature of the B2KShh stripping line is that it applies only mild cuts on the daughter momenta, in order to not bias the phase space. This would not allow the analysis of certain regions and resonances close to the kinematic boundaries of the Dalitz plot.

Table 13.: Requirements on the charged pion and kaon tracks.

Variable definition	Pion cut	Kaon cut
Total P	$< 100 \text{ GeV}$	$< 100 \text{ GeV}$
p_T	$> 250 \text{ MeV}/c$	$> 250 \text{ MeV}/c$
Kaon vs. pion PID ($DLL_{K\pi}$)	-	> -5
"Clone distance"	> 5000	> 5000
Minimum $IP\chi^2$ w.r.t. PVs	> 4	> 4
Track fit χ^2/ndf	< 4	< 4

² During pattern recognition it is possible that tracks providing the same information are created. These clones can be removed by comparing the hits shared by the two tracks and excluding the one with the lower quality.

- *K_s^0 candidate requirements*

The K_s^0 candidates are created combining two oppositely charged pions, which can be reconstructed either as Long or Downstream tracks. The two samples show differences in the distributions of the mass variables, in the mass resolutions and in the reconstruction efficiencies. Thus different selections are required. The requirements on the K_s^0 candidates are selection criteria on:

- The K_s^0 daughter track momentum;
- The reduced χ^2 of the K_s^0 daughter track reconstruction;
- The smallest impact parameter of K_s^0 daughter track computed with respect to all the reconstructed primary vertices;
- The difference between the reconstructed and the nominal K_s^0 mass;
- The χ^2 of the K_s^0 vertex coming from the fitting procedure;
- The χ^2 of the distance between the K_s^0 vertex and the associated primary vertex;
- The K_s^0 momentum;

The cut values are reported in Tab. 14.

Table 14.: Requirements for the Long-Long and Down-Down K_s^0 candidates.

Variable definition	Long-Long requirement	Down-Down requirement
K_s^0 daughter tracks P	$> 2 \text{ GeV}/c$	$> 2 \text{ GeV}/c$
K_s^0 daughter tracks χ^2/ndf	< 4	-
K_s^0 daughter minimum IP χ^2 w.r.t. PVs	> 9	> 4
Mass difference w.r.t. nominal K_s^0 mass	$< 20 \text{ MeV}/c^2$	$< 30 \text{ MeV}/c^2$
χ^2 of K_s^0 vertex fit	< 12	< 12
χ^2 separation of K_s^0 vertex and associated PV	> 80	> 50
$K_s^0 P$	-	$> 6 \text{ GeV}/c$

- *B candidate requirements*

The B candidates are formed first by four-momentum addition. Some loose "Combination" cuts, applied to reduce the number of the candidates, will be considered for the vertex fit. The quantities related to these candidates are:

- Transverse momentum of the B candidate;

- Sum of the transverse momentum of the B daughters;
- Transverse momentum of at least two B daughters;
- Mass window of the B candidate;
- Impact parameter with respect to the primary vertex of the B daughter with the highest transverse momentum;
- Maximum χ^2 of the distance of closest approach of any of two daughters;

As last step a vertex fit is performed and further quality "Mother" cuts are made on:

- Transverse momentum of the B candidate;
- The χ^2 of the B vertex coming from the fitting procedure;
- Cosine of B pointing angle;
- Minimum impact parameter χ^2 of the B with respect to the primary vertex;
- Minimum distance between the B and associated primary vertex;
- Significance of the separation between the B and associated primary vertex;

The "Combination" and "Mother" requirements are listed in Tab. 15.

Some additional requirements are applied to reject the combinatorial background:

- Due to the high track densities, the pattern recognition step during the reconstruction can form tracks using random combinations of hits, called *ghost* tracks. By the reconstruction software to each track is assigned a certain probability of it being a ghost, and a requirement on this quantity is applied to reduce the rate of these tracks.
- A cut on the isolation of the B vertex: this is calculated as the difference between the nominal vertex χ^2 and the one calculated with the addition of the closest track in the event. This cut helps also in removing partially reconstructed backgrounds;
- A cut on the B/K_s^0 vertex separation in the z direction;

The values of the applied cuts are shown in Tab. 16.

Table 15.: Requirements on the B candidates from four-momentum combination.

Variable Definition	Selection requirement
Combination cuts	
p_T of the B candidate	$> 1000 \text{ MeV}/c$
Sum of the daughters' p_T	$> 3000 \text{ MeV}/c$
p_T of at least two B daughters	$> 800 \text{ MeV}/c$
Mass of the B candidate	$4779 \text{ MeV}/c^2 < m_{K_S^0 h^+ h^-} < 5866 \text{ MeV}/c^2$
IP <i>w.r.t.</i> PV of highest p_T B daughter	$> 0.05 \text{ mm}$
Maximum DOCA χ^2 of any 2 daughters	< 5
Mother cuts	
p_T of the B candidate	$> 1500 \text{ MeV}/c$
χ^2 of B vertex fit	< 12
Cosine of B pointing angle	> 0.9999
Minimum B IP χ^2 <i>w.r.t.</i> PVs	< 8
Minimum vertex distance <i>w.r.t.</i> PVs	$> 1 \text{ mm}$
χ^2 Separation of B vertex and associated PV	> 50
Separation between K_S^0 and B vertices in the positive z direction	$> 10 \text{ mm}$

Table 16.: Preselection cuts applied to the stripped events.

Variable Definition	Selection requirement
Ghost probability	< 0.5
Isolation of the B vertex	> 4
B/K_S^0 vertex separation	> 30

Particle identification

To remove possible contributions from misidentified backgrounds of B mesons with a final state like $K_S \pi \pi$, $K_S K K$, $K_S \pi p$ etc, particle identification requirements are applied to the B daughter charged tracks. The $DLL_{K\pi}$ and $DLL_{p\pi}$ delta log-likelihood variables produced by the RICH and calorimeters sub-detectors are used. The set of PID selection requirements are listed in Tab. 17.

In addition to remove possible pollution from charmonium particle decays into a muon pair misidentified as pions or kaons, a muon identification requirement ($isMuon == 0$) is made on both the K_S^0 and the B charged daughter tracks.

Table 17.: PID cuts applied to the particles reconstructed as pions or kaons.

Track	π	K
Kaon PID requirement	$DLL_{K\pi} < 0$	$DLL_{K\pi} > 5$
Proton PID requirement	$DLL_{p\pi} < 10$	$(DLL_{p\pi} - DLL_{K\pi}) < 10$

Charm vetoes

The fully reconstructed and peaking background can be reconstructed giving the proper mass hypothesis to the final state particles. Therefore it is possible to veto these particles by making a requirement on the difference between the reconstructed mass and the PDG nominal value. The specific modes considered and the requirements made are listed in Tab. 18. These cuts were sufficient to reduce these backgrounds to a negligible level. Hence they will not be modeled in the final fit.

Table 18.: Charm vetoes.

Charmed particle and decay mode	Selection requirement
$J/\psi \rightarrow \pi^- \pi^+$	$ m_{\text{Rec}} - m_{\text{PDG}} > 30 \text{ MeV}/c^2$
$J/\psi \rightarrow K^- K^+$	"
$\chi_c \rightarrow \pi^- \pi^+$	"
$\chi_c \rightarrow K^- K^+$	"
$D^0 \rightarrow K^- \pi^+$	"
$D^0 \rightarrow \pi^+ \pi^-$	"
$D^0 \rightarrow K^+ K^-$	"
$D^+ \rightarrow K_s^0 K^+$	"
$D^+ \rightarrow K_s^0 \pi^+$	"
$D_s^+ \rightarrow K_s^0 K^+$	"
$D_s^+ \rightarrow K_s^0 \pi^+$	"
$\Lambda_c \rightarrow p K_s^0$	"

4.4.5 Fitting ranges

Since in this analysis it is essential to distinguish between the inclusive final state and the resonant one, a two-dimensional fit to the $K_s^0 K \pi$ and $K \pi$ spectra is

performed. For this reason a cut in the mass window of the two observables is applied. The chosen values are shown in Tab. 19.

Table 19.: Fitting ranges of the m_B and $m_{K\pi}$ variables.

Observable	Range (MeV/ c^2)
m_B	5000 – 5800
$m_{K\pi}$	600 – 1200

The m_B range was chosen in order to fully contain both the B^0 and B_s^0 mass peaks. It is also wide enough to contain the partially reconstructed backgrounds in the lower mass range and enough events in the upper side to model the combinatorial background. The $m_{K\pi}$ range instead should contain the $K^*(892)^0$ resonance ($m = 895.81$ MeV/ c^2 , $\Gamma = 47.4$ MeV/ c^2 [17]) but it has to exclude as much as possible the contamination from other high energy resonances such as the $K^*(1430)^0$ ($m = 1425$ MeV/ c^2 , $\Gamma = 270$ MeV/ c^2 [17]).

4.4.6 Boosted decision tree

On top of the preselection a multivariate discriminator based on a Boosted Decision Tree (BDT) [63] is used against the combinatorial background. The following steps are performed to use this method and they are explained in detail in the next sections:

- Decision on which data samples are used for the signal and the background;
- Decision on which variables are used to build the classifier;
- Decision on which cut to apply on the classifier to maximize a certain figure of merit.

Training of the BDT

The inclusive $B_{d,s}^0 \rightarrow K_S^0 h^\pm h'^\mp$ analysis developed and optimized a very efficient BDT which is also used for this analysis. In order to train a BDT without introducing potential bias it is necessary to use a data sample independent of the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ sample. A good candidate is the normalization decay sample which contains the largest number of preselected signal and combinatorial background events. The use of the BDT, trained on the $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ normalization mode,

for the $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ signal decays is justified by the very similar topology of these channels.

To train the classifier, samples of simulated $B_{d,s}^0 \rightarrow K_S^0\pi^+\pi^-$ events reconstructed, stripped and preselected as the data were used as signal. This decision assumes that the distribution of the input variables in the $K_S^0\pi\pi$ data sample are correctly modeled by the simulation sample. If this is not the case, it is necessary to account for the possible differences in the systematic error. This was checked a posteriori and the comparison of the variables for the data and the simulated samples are shown in App. B.

The background sample is constructed from stripped $K_S^0\pi\pi$ events from the right-hand sideband ($m_B > 5420 \text{ MeV}/c^2$) of data distribution. In addition the baryonic background of the Λ_b is removed by applying a veto on the $pK_S^0\pi$ and pK_S^0 combinations. The Downstream and Long tracks behave in a different way regarding the combinatorial background, and an higher contribution is expected from the first category. For this reason two different BDTs were built, one for each category.

A strategy which makes use of independent samples for the training and the test phases was developed in order to not introduce a bias in the procedure. The full statistics is randomly split into two sets. The first sample is used to train a first BDT (BDT1) which is tested on the second sample. The second sample, instead, is used to train a second BDT (BDT2) which is tested on the first sample as shown in Fig. 36.

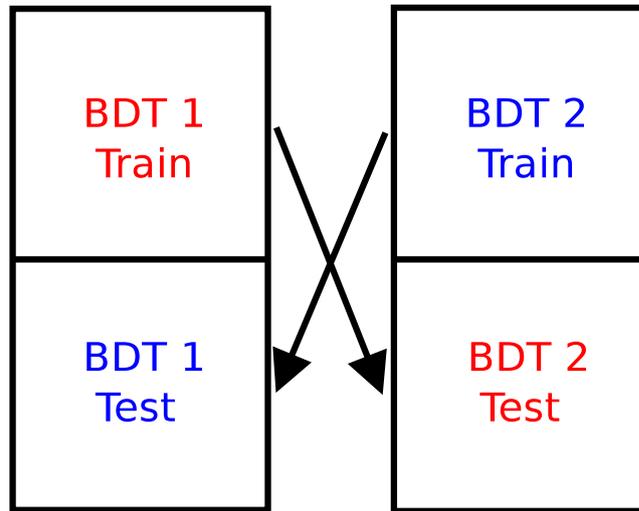


Figure 36.: Scheme explaining the procedure adopted to train and test the BDTs.

About the same number of events was used to train and to test each BDTs. This is reported in Tab. 20.

Table 20.: Number of events used to train the BDT.

	K_S^0 Down-Down	K_S^0 Long-Long
Signal	8525	1960
Background	6050	2000

The variables used to build the BDT are given in Tab. 21. These are in general estimators of the pointing of the B and its daughters with respect to the primary vertex, of the goodness of the B and the K_S^0 vertex reconstruction, and variables related to the lifetime of the B and the K_S^0 . The variable distributions for the signal and the background are shown in Figs. 37 and 38 for the two K_S^0 categories.

Table 21.: List of variables used to build the BDT discriminant for the two K_S^0 categories.

The variables labeled as Down-Down or Long-Long are used only for that specific category.

Variable	Description	Down-Down	Long-Long
B_PT	B transverse momentum	✓	✓
B_ETA	B pseudorapidity	✓	✓
B_IPCHI2_OWNPV	B IP significance w.r.t PV	✓	✓
B_VDCHI2_OWNPV	B flight distance significance w.r.t PV	✓	✓
B_DIRA_OWNPV	B pointing angle	✓	✓
B_ENDVERTEX_CHI2	B vertex fit χ^2	✓	✓
$\sum_i h_i$ -IPCHI2_OWNPV	sum of the hadron IP significance w.r.t PV	✓	✓
KSDD_ENDVERTEX_CHI2	K_S^0 vertex fit χ^2	✓	-
KSSL_VDCHI2_OWNPV	K_S^0 flight distance significance w.r.t PV	-	✓
KSSL_IPCHI2_OWNPV	K_S^0 IP significance w.r.t PV	-	✓

The package in which the BDT is implemented defines a ranking of the variables according to their discriminative power. The list of the variables in order of their importance is reported in Tab. 22.

The BDT outputs for the two K_S^0 categories are shown in Fig. 39, where a good agreement between the samples can be observed. In order not to bias the analysis a final BDT output is obtained through the combination of the two, using a random number to chose with equal probability one of the two.

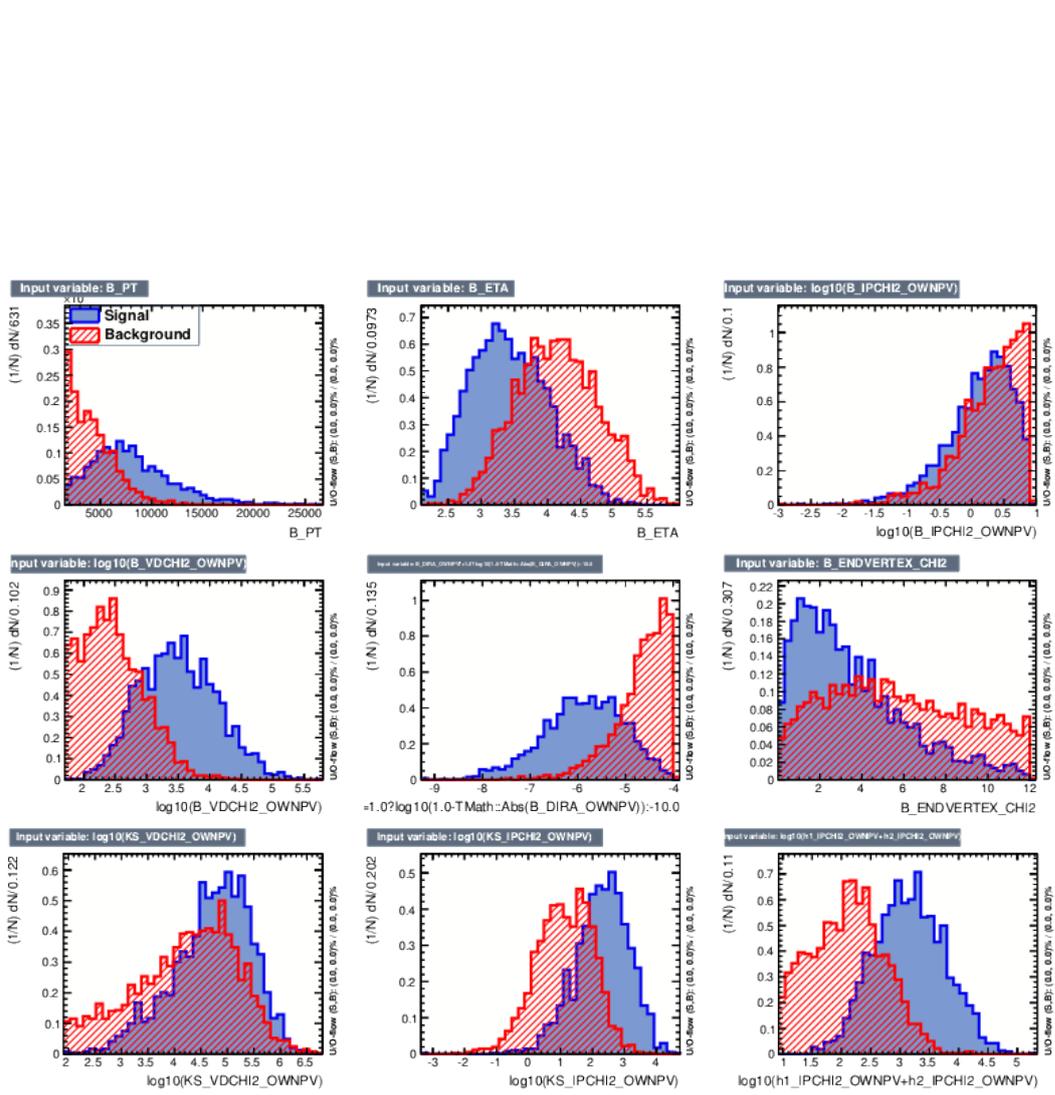


Figure 37.: Distribution of the variables used to build the BDT for the Long-Long K_S^0 . Blue: signal, Red: background.

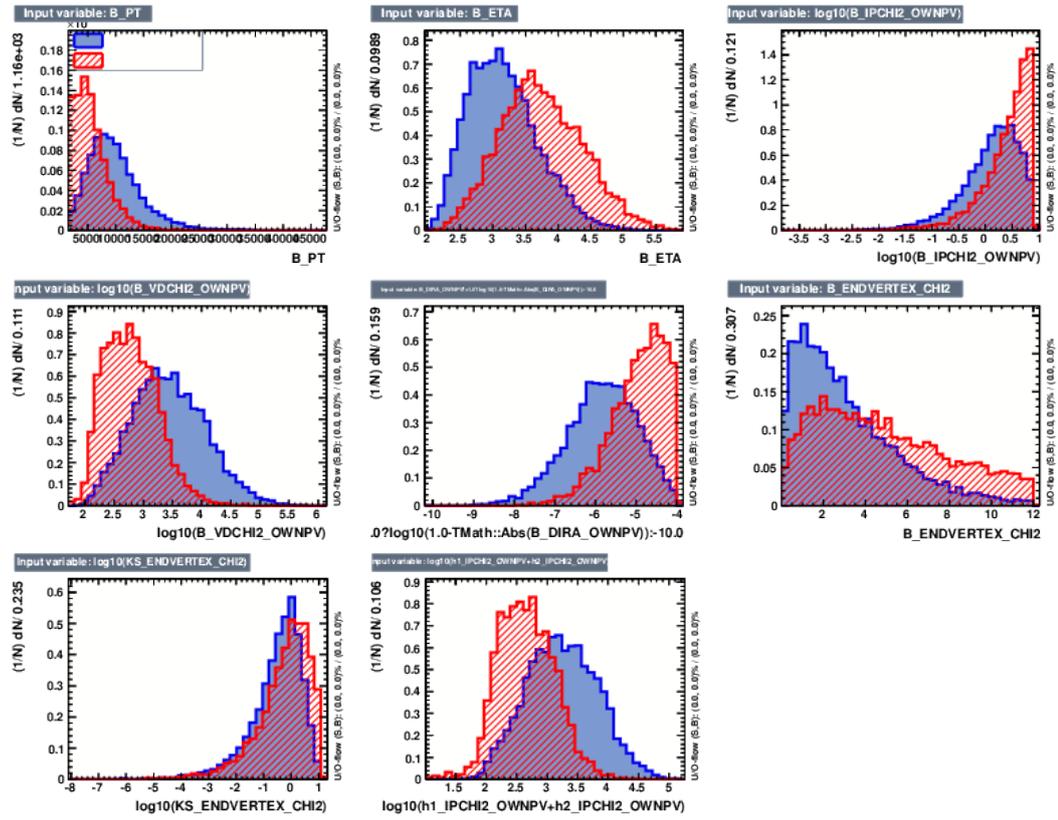


Figure 38.: Distribution of the variables used to build the BDT for the Down-Down K_s^0 . Blue: signal, Red: background.

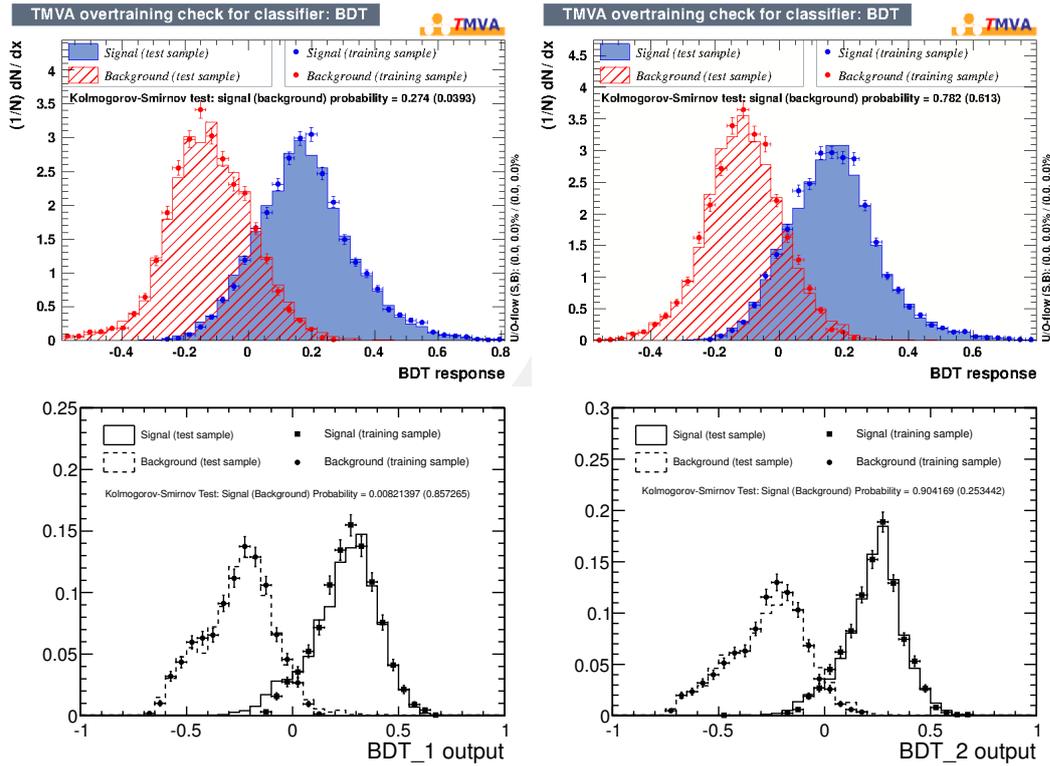


Figure 39.: Output from the training and test sample for the BDT₁ (left) and BDT₂ (right) discriminants for Down-Down (top) and Long-Long (bottom) MC signal events and background from the right-handed side-bands of the 2011 data [9].

Table 22.: BDT variable importance ranking (0, 1). Variables are presented in the hierarchical order as obtained for Down-Down and Long-Long K_S^0 .

Variable	Importance DD		Importance LL	
	BDT1	BDT2	BDT1	BDT2
B_VDCHI2_0WNPV	0.173	0.174	0.239	0.240
$\Sigma_i h_i$ _IPCHI2_0WNPV	0.168	0.164	0.230	0.230
B_DIRA_0WNPV	0.166	0.164	0.224	0.219
B_ETA	0.154	0.151	0.171	0.181
B_PT	0.147	0.146	0.171	0.166
B_ENDVERTEX_CHI2	0.086	0.082	0.095	0.100
B_IPCHI2_0WNPV	0.069	0.064	0.026	0.034
KSDD_ENDVERTEX_CHI2	0.020	0.016	-	-
KSLI_IPCHI2_0WNPV	-	-	0.192	0.188
KSLI_VDCHI2_0WNPV	-	-	0.068	0.079

Sensitivity studies

To optimize the cut value on the BDTs for the best sensitivity to the signal, the Punzi figure of merit [70] F_{Punzi} was chosen. This figure of merit is particularly useful when the branching ratio for the signal is unknown or has a significant theoretical error. It is given by the formula

$$F_{\text{Punzi}} = \frac{\epsilon_{\text{Sig}}}{\frac{a}{2} + \sqrt{N_{\text{Bkg}}}} \quad (41)$$

where ϵ_{Sig} is the efficiency of the cut on the signal events and N_{Bkg} is the number of expected combinatorial background events. Finally, a corresponds to the statistical significance (in units of gaussian standard deviation σ) which is aimed to be placed on the observation. The value for this quantity has been chosen to be equal to 3. The signal efficiency was estimated using simulated signal samples and counting the number of events for each BDT cut. The value of N_{Bkg} can be estimated from an exponential fit in the upper side-band of the $K_S^0 K\pi$ data sample and extrapolating this number into the signal mass window. In order to have a more realistic estimate of the number of background events in the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays an additional requirement on the $K\pi$ mass combination should be imposed before applying the BDT cut. This request is tighter with respect to the fit mass range in Tab. 19, in order to take only those events near to the $K^*(892)^0$ mass region. The variable chosen for the $K\pi$ mass combination is the one coming from a refitting of the

$K_S^0 K \pi$ final state and constraining the B_S^0 mass to its nominal value. The B_S^0 and K^{*0} signal mass window are shown in Tab. 23 together with the right-sideband interval. The same mass windows were used for the two K_S^0 categories. As an example the B upper side-band for the Long-Long before and after applying this $K \pi$ mass requirement are shown in Fig. 40.

Table 23.: Definition of the signal and background mass intervals used to optimize the final offline selection cuts.

$K_S^0 K \pi$ signal window (MeV/ c^2)	$K \pi$ signal window (MeV/ c^2)	right-sideband (MeV/ c^2)
[5319-5425]	[800-1000]	[5550-5800]

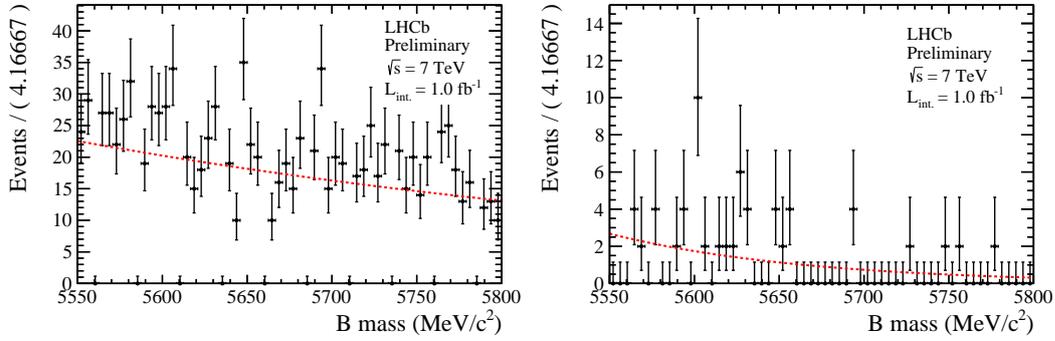


Figure 40.: B invariant mass of $K_S^0 K \pi$ collision data for the Long-Long K_S^0 sample. Upper sideband modeled with an exponential PDF (red). Left: No $K \pi$ mass requirement applied. Right: $K \pi$ mass requirement applied.

However, the number of events after the $K \pi$ mass window cut is small, yielding an estimate with a large uncertainty on the number of background events N_{Bkg} . The number of events in the right side-band before and after the $K \pi$ mass requirement and before any BDT cut was determined and their ratio calculated. Then the data sample before the $K \pi$ mass requirement was fitted to determine the number of combinatorial background events for different BDT cut values and to give estimates of the background yields in the signal region. Each yield was then normalised to the previously measured ratio, to give an estimate of N_{Bkg} after the $K \pi$ mass cut was applied. The values of F_{Punzi} were finally calculated for different BDT cut values as shown in Fig. 41, which also shows the corresponding signal efficiencies. The BDT cut value was chosen approximately in correspondence with the F_{Punzi} maximum.

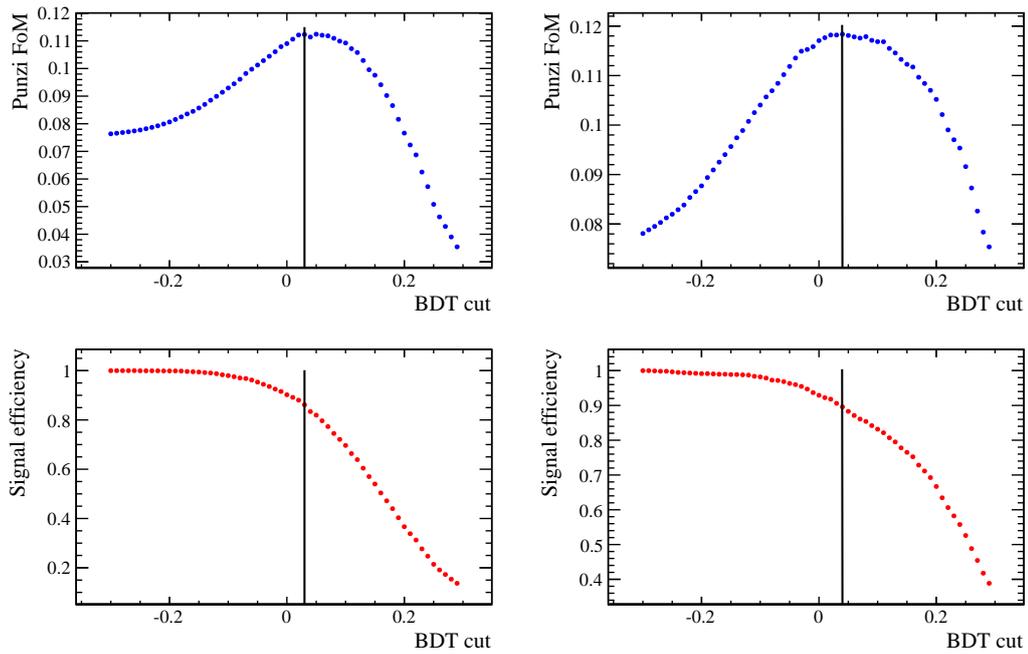


Figure 41.: Punzi figure of merit (top) and signal efficiency (bottom) for different BDT requirements. Left: Down-Down K_s^0 sample. Right: Long-Long K_s^0 sample.

The optimal BDT cuts and the corresponding efficiencies for the signal and the background are summarized in Tab. 24.

Table 24.: BDT optimal cut and efficiencies for signal and background.

K_s^0 mode	Optimisation	BDT Cut	ϵ Signal	ϵ Background
Down-Down	$B^0 \rightarrow K^{*0} K_s^0$	0.03	0.83 ± 0.01	0.12 ± 0.01
	$B_s^0 \rightarrow K^{*0} K_s^0$	0.03	0.86 ± 0.01	0.12 ± 0.01
Long-Long	$B^0 \rightarrow K^{*0} K_s^0$	0.04	0.92 ± 0.01	0.05 ± 0.01
	$B_s^0 \rightarrow K^{*0} K_s^0$	0.04	0.90 ± 0.01	0.05 ± 0.01

It has to be stressed that the Punzi FoM is optimized only for the B_s^0 case, since it is the main mode of interest and the same cut is applied in the B^0 case.

5 | FIT MODEL

The physics observables, that is the signal yields, are extracted using a maximum likelihood fit. In this chapter first the general concepts of the method are explained, followed by a detailed description of the model used for the signal and the normalization channels. The validation procedure by means of toy Monte Carlo studies is reported and finally the results of the fit are shown.

5.1 MAXIMUM LIKELIHOOD FIT METHOD

The maximum likelihood method is a technique to estimate the value of a set of parameters $\vec{\theta}$ from a datasample \vec{x} with N number of events. The data are assumed to be described by a Probability Density Function (PDF), a theoretical function $f(\vec{x}|\vec{\theta})$ which depends on the set of the unknown parameters. The integral of f over the entire range of \vec{x} is normalised to unity

$$\int_D f(\vec{x}; \vec{\theta}) dx = 1 \quad (42)$$

where D is the domain of \vec{x} . The meaning of $f(\vec{x})d\vec{x}$ is the probability of a measurement in the $[\vec{x} + d\vec{x}]$ interval, for each value of the parameters. The likelihood function is defined as the product of the PDFs for the N events

$$\mathcal{L}(\vec{\theta}) = \prod_i^N f(\vec{x}_i; \vec{\theta}). \quad (43)$$

The method consists in finding the set of parameters $\vec{\theta}$ which maximizes the likelihood. A more detailed review of the method can be found in [71]. In practice it is often more convenient to work with the negative logarithm of the likelihood function, called the log-likelihood

$$L(\vec{\theta}) = -\ln \mathcal{L}(\vec{\theta}; \vec{x}) = -\sum_{i=1}^n \ln f(\vec{x}_i | \vec{\theta}). \quad (44)$$

which has to be minimized. From the technical point of view in this analysis this step is done via the MINUIT program [72].

The maximum likelihood estimate is asymptotically unbiased, i.e. for samples with the number of events $n \rightarrow \infty$ the expectation value of the parameters is equal to their true value. The estimated parameters then are distributed like a Gaussian function and their error is equal to

$$\sigma(\theta_j) = \left(-\frac{d^2L}{d\theta_j^2} \right)^{-1/2} \quad (45)$$

5.1.1 Extended fits

The method of extended fits differs from the standard maximum likelihood technique through a relaxation of the normalisation condition. The function $f(\vec{x}; \vec{\theta})$, normalised according to Eq. 42, is replaced by a function $F(\vec{x}; \vec{\theta})$, normalised in the following way

$$\int_D F(\vec{x}; \vec{\theta}) d\vec{x} = \mathcal{N}(\vec{\theta}) \quad (46)$$

The interpretation of $F(\vec{x}; \vec{\theta})$ is that it describes not only the shape of the expected distribution, but also its integral. Events occur randomly in the range of \vec{x} , governed by Poisson statistics, and $F(\vec{x}; \vec{\theta}) d\vec{x}$ gives the number of events expected in the $[\vec{x} + d\vec{x}]$ interval. The parameter \mathcal{N} is thus the total number of events expected over the whole range of observation. The observed number N will in general be different from this due to the fluctuations of Poisson statistics. With the inclusion of the Poisson term the total probability of a sample (Eq. 43) is thus modified into

$$\mathcal{L}(\vec{\theta}) = \prod_i^N f(\vec{x}_i; \vec{\theta}) e^{-\mathcal{N}} \frac{\mathcal{N}^N}{N!} = \prod_i^N F(\vec{x}_i; \vec{\theta}) \frac{e^{-\mathcal{N}}}{N!}. \quad (47)$$

5.1.2 Constraints on parameters

Sometimes a parameter θ_j has a previous measurement so it is possible to constrain its value to be $\mu \pm \sigma$, where μ is the mean value and σ its error. Assuming a Gaussian distribution for the probability to measure a certain value of θ_j

$$p(\theta_j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\theta_j - \mu)^2}{2\sigma^2}} \quad (48)$$

the likelihood can be written as

$$\mathcal{L}(\vec{\theta}) = \prod_i^N f(\vec{x}_i; \vec{\theta}) p(\theta_j). \quad (49)$$

These so-called *Gaussian constraints* will be applied to some of the parameters in the fit.

5.1.3 Implementation

To extract the signal yields for the main $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ channels and the normalization mode $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ two independent fits were implemented:

- $B_{d,s}^0 \rightarrow K^{*0} K_s^0$: two dimensional extended unbinned maximum likelihood fit performed in the $K_s^0 K \pi$ and the $K \pi$ mass spectra;
- $B^0 \rightarrow K_s^0 \pi^+ \pi^-$: one dimensional extended unbinned maximum likelihood fit performed in the $K_s^0 \pi \pi$ mass spectrum.

Both fits are unbinned in order to exploit the full information available from the events. In addition they are conducted simultaneously in the two K_s^0 categories, this means that certain parameters are shared between the two models. The invariant masses used in the fits are described in the following:

- The B invariant mass is built using candidates constrained to come from the PV and with the K_s^0 mass forced to have the nominal value.
- The K^{*0} invariant mass is the $K \pi$ invariant mass, with the B candidate mass constrained to the B_s^0 mass, in order to improve the resolution of the K^{*0} mass.

Some parts of the implementation of the fitting algorithm were developed for the $B_{d,s}^0 \rightarrow K_s^0 h^\pm h'^\mp$ analysis and modified and extended for this analysis.

5.2 $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ MODEL

5.2.1 Signal model

In order to use a 2D fit to extract the signal yields it is necessary to ensure that there are no correlations between the two terms or to take the correlation explicitly into account. The correlation plots for the observables after the entire selection are displayed in Fig. 42, while the correlation coefficients are shown in Tab. 25. There is no significant correlation and therefore the correlation terms are neglected in the fit.

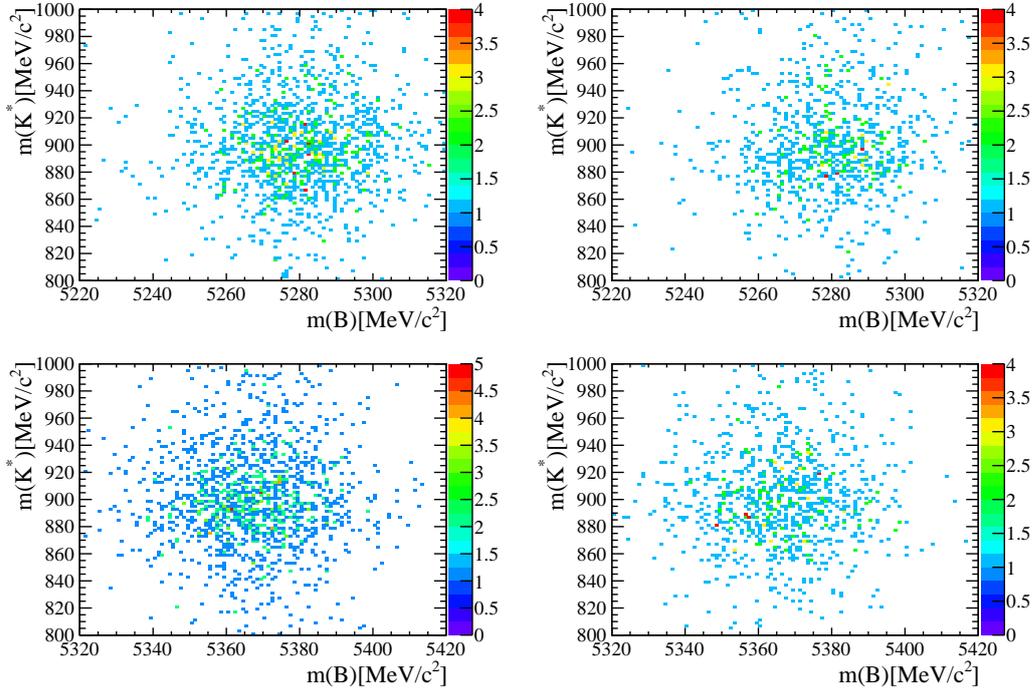


Figure 42.: Correlation plots between the fit variables in simulated $B_{d,s}^0 \rightarrow K^*K$ decays after the selection. Top: $B^0 \rightarrow K^*K$. Bottom: $B_s^0 \rightarrow K^*K$. Left: Down-Down K_s^0 sample. Right: Long-Long K_s^0 sample.

Table 25.: Correlation coefficients between the $m_{K_s^0 K \pi}$ and $m_{K \pi}$ fit variables for simulated signal events.

Mode	Correlation coefficient
$B^0 \rightarrow K^*K$ (Down-Down K_s^0)	0.010
$B^0 \rightarrow K^*K$ (Long-Long K_s^0)	0.036
$B_s^0 \rightarrow K^*K$ (Down-Down K_s^0)	-0.002
$B_s^0 \rightarrow K^*K$ (Long-Long K_s^0)	-0.056

As general strategy most of the signal model parameters are extracted after the whole selection, and fixed to the values obtained from the simulation except for the width of the B_s^0 and K^{*0} . For these parameters gaussian constraints are applied with a mean value equal to the value obtained from the fit to simulated samples. The signal in the $m_{K_s^0 K \pi}$ variable is modeled by the sum of two Crystal Ball (CB) functions [73] sharing the mean and the width but with different tail parameters. The CB function of a variable x , is defined as

$$CB(x) = \mathcal{N} \cdot \begin{cases} \exp(-x^2/2\sigma^2) & \text{if } x/\sigma > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \exp(-\alpha^2/2) \left(\frac{n-\alpha^2}{|\alpha|} - \frac{x}{\sigma}\right) & \text{if } x/\sigma \leq -\alpha, \end{cases} \quad (50)$$

with $x = m - \mu$, where m is the reconstructed mass and μ is the mean value of the gaussian part of the function. The parameter σ is the resolution of the gaussian and \mathcal{N} the normalization of the function. The sign of the α parameter governs the right-handed or left-handed location of the tail and the parameter n drives the power law associated to the decrease of the tail. The signal in the $m_{K \pi}$ variable is modeled with a spin-1 Relativistic Breit-Wigner amplitude [74]

$$M_1^{K\pi}(m) \propto \frac{m_1}{m_1^2 - m^2 - i m_1 \Gamma_1(m)}, \quad (51)$$

with

$$\Gamma_1(m) = \Gamma_1 \left(\frac{m_1}{m}\right) \left(\frac{q}{q_1}\right)^{2L+1} B_L^2(q, q_1) \quad (52)$$

where Γ_1 is the K^{*0} width (48.7 MeV/ c^2) and m_1 its nominal mass (896 MeV/ c^2). The value of L is the spin of the resonance ($L = 1$ in this case) and q is the momentum of the kaon in the K^{*0} rest frame, given by:

$$q(m, M_K, M_\pi) = \frac{\sqrt{(m^2 - (M_K + M_\pi)^2)(m^2 - (M_K - M_\pi)^2)}}{2m} \quad (53)$$

where M_K and M_π are the kaon and pion masses respectively. The functions $B_L(q, q_1)$ are the Blatt-Weisskopf barrier factors. For $L = 1$ we have:

$$B_1(q, q_1) = \sqrt{\frac{1+z_0}{1+z}}, \quad (54)$$

with $z = (rq)^2$, $z_0 = (rq_1)^2$ and $r = 3.4 \text{ GeV}^{-1}$ (in natural units) being the interaction radius.

In order to decrease the number of parameters in the fit and reach a more stable configuration, additional requirements are imposed:

- The difference between the B^0 and B_s^0 mean is constrained to the PDG value;
- The parameters modeling the radiative tails (α and n) are the same for the B^0 , B_s^0 , Long-Long and Down-Down categories;
- The fraction of the second CB function (f_2) is the same for the B^0 , B_s^0 Long-Long and Down-Down categories;
- The mean and the width of the two CB functions are identical;
- The width is the same for the B^0 and the B_s^0 ;
- The ratio of the B widths for the Long-Long and Down-Down categories (σ_{LL}/σ_{DD}) is the same for the B^0 and the B_s^0 ;
- The ratio of the CB function tail parameters for the Long-Long and Down-Down categories (α_{LL}/α_{DD} and n_{LL}/n_{DD}) is the same for the B^0 and the B_s^0 ;
- The mean and the width of the K^{*0} coming from the B^0 and the B_s^0 are assumed to be the same.
- The ratio of the K^* widths for the Long-Long and Down-Down categories (Γ_{LL}/Γ_{DD}) is the same for the K^* coming from the B^0 and the B_s^0 ;

A 2D fit of all the simulated samples is performed simultaneously to extract the various parameters after the whole selection. The parameters are summarised in Tab. 26 and the mass plots are shown in Figs. 43 and 44¹.

5.2.2 Non resonant $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$ background model

The final state $K_s^0 K \pi$ considered for this analysis can be the same for different decays. Nevertheless some of them differ from the signal by having a not-peaking resonant structure in the $K\pi$ spectrum. The contribution given by these non resonant decays has to be taken into account in the fit model. Therefore its shape is taken from simulated samples, generated uniformly in phase space. The $K_s^0 K \pi$ spectrum is modeled as for the signal using a double CB function. For this PDF

¹ Note that the error bars for entries at low statistics are not symmetric. In fact at low statistics symmetric Gaussian errors of magnitude \sqrt{N} are only an approximation of the actual statistical uncertainty on a bin with N entries. In these plots the 68% confidence interval for Poisson statistics are shown (see Ref. [75]).

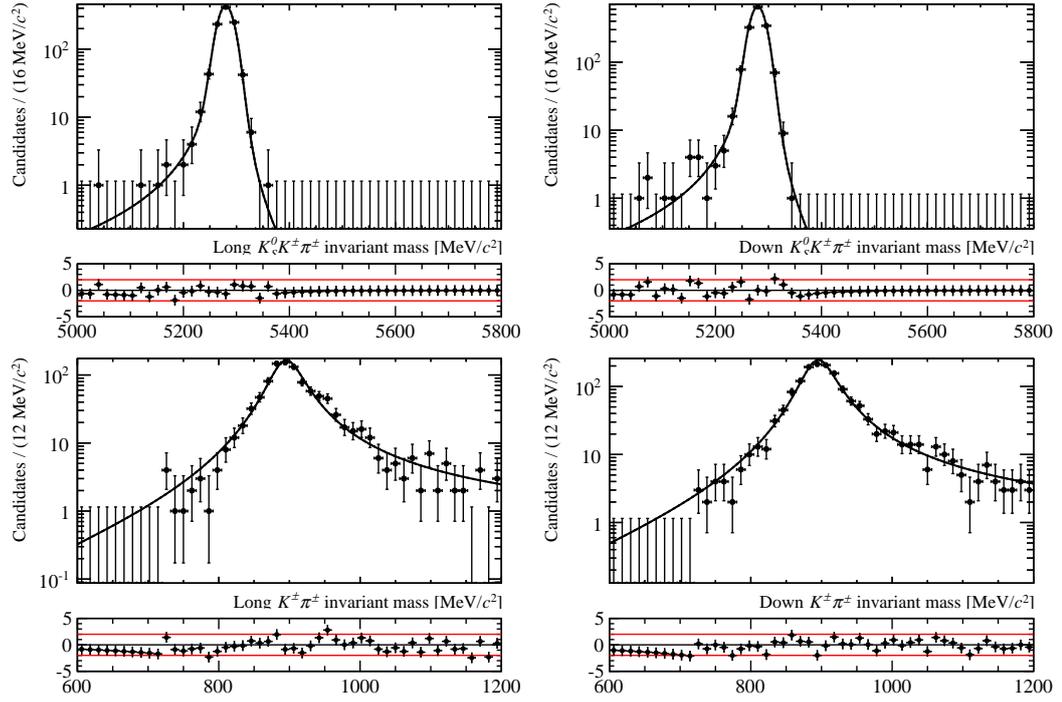


Figure 43.: $B^0 \rightarrow K^* K$ signal fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_S^0 K \pi}$ and $m_{K \pi}$ observables with the pull distributions shown underneath. Top: $m_{K_S^0 K \pi}$. Bottom: $m_{K_S^0 \pi}$. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

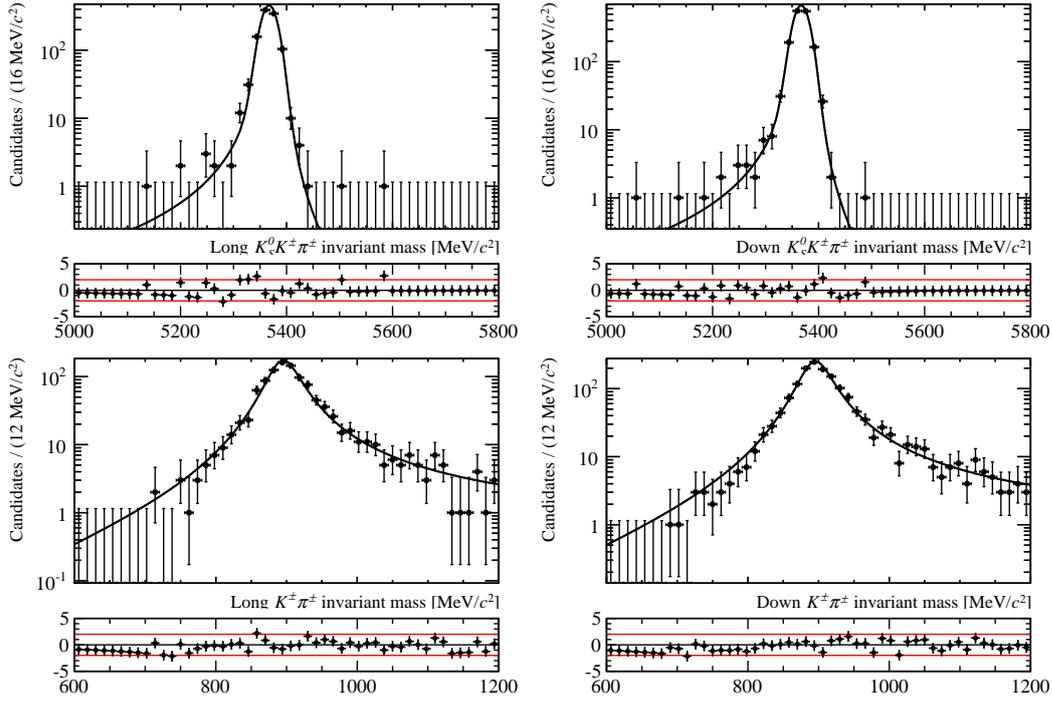


Figure 44.: $B_s^0 \rightarrow K^* K \pi$ signal fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_S^0 K \pi}$ and $m_{K \pi}$ observables with the pull distributions shown underneath. Top: $m_{K_S^0 K \pi}$. Bottom: $m_{K_S^0 \pi}$. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

Table 26.: Fitted parameter values for the fit to $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ simulated samples.

Parameter	Value
$m(B_s)$	5367.17 ± 0.23
σ	14.35 ± 0.27
$\sigma(\text{LL})/\sigma(\text{DD})$	1.012 ± 0.024
α_1	1.67 ± 0.24
α_2/α_1	-1.28 ± 0.34
n_1	2.000 ± 0.097
n_2/n_1	1.91 ± 0.69
f_2	0.56 ± 0.22
$m(K^*)$	895.14 ± 0.48
Γ	49 ± 1
$\Gamma(\text{LL})/\Gamma(\text{DD})$	0.992 ± 0.040

the mean and the width parameters are shared with the signal, while the tail parameters are different. For the $K\pi$ spectrum a linear function is used. The latter is parametrised in order to assume only positive values

$$L(x) = \begin{cases} m(x - t) & \text{if } x \geq t \\ 0 & \text{if } x < t. \end{cases} \quad (55)$$

Since the slope m determines the normalization it has been fixed to positive values, while t is a threshold value where the function assumes positive values. A 2D fit of all the simulated samples is performed simultaneously to extract the various parameters, after the whole selection. The parameters are summarised in Tab. 27 and the mass plots are shown in Figs. 45 and 46.

It is very important to notice that interference between signal and non resonant components may arise modifying both the signal and background PDF shapes. This can modify the shape of the $K\pi$ spectrum and consequently the number of signal yields. This effect will be accounted in the systematic uncertainties evaluation in Sec. 7.2.3.

5.2.3 Misidentified-signal model

The misidentification of a pion as kaon or vice versa, coming for example from decays like $B_{d,s}^0 \rightarrow K_S^0\pi^+\pi^-$ and $B_{d,s}^0 \rightarrow K_S^0K^+K^-$ or decays with resonant interme-

Table 27.: Fitted parameter values for the fit to $B_{d,s}^0 \rightarrow K_S^0 K^\pm \pi^\mp$ simulated samples.

Parameter	Value
$\alpha_{1NonReso}$	1.51 ± 0.34
$\alpha_{2NonReso} / \alpha_{1NonReso}$	-0.18 ± 0.21
$n_{1NonReso}$	4 ± 2
$n_{2NonReso} / n_{1NonReso}$	3 ± 2
$t_{NonReso}$	631 ± 12

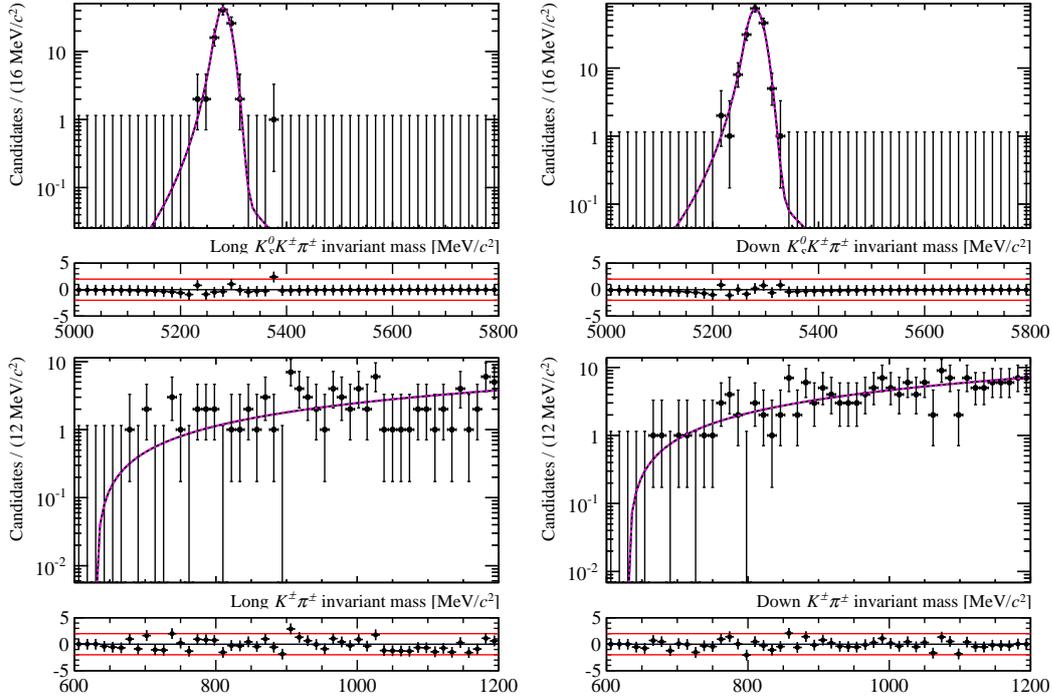


Figure 45.: $B^0 \rightarrow K_S^0 K^\pm \pi^\mp$ non resonant fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_S^0 K^\pm \pi^\mp}$ and $m_{K^\pm \pi^\mp}$ observables with the pull distributions shown underneath. Top: $m_{K_S^0 K^\pm \pi^\mp}$. Bottom: $m_{K^\pm \pi^\mp}$. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

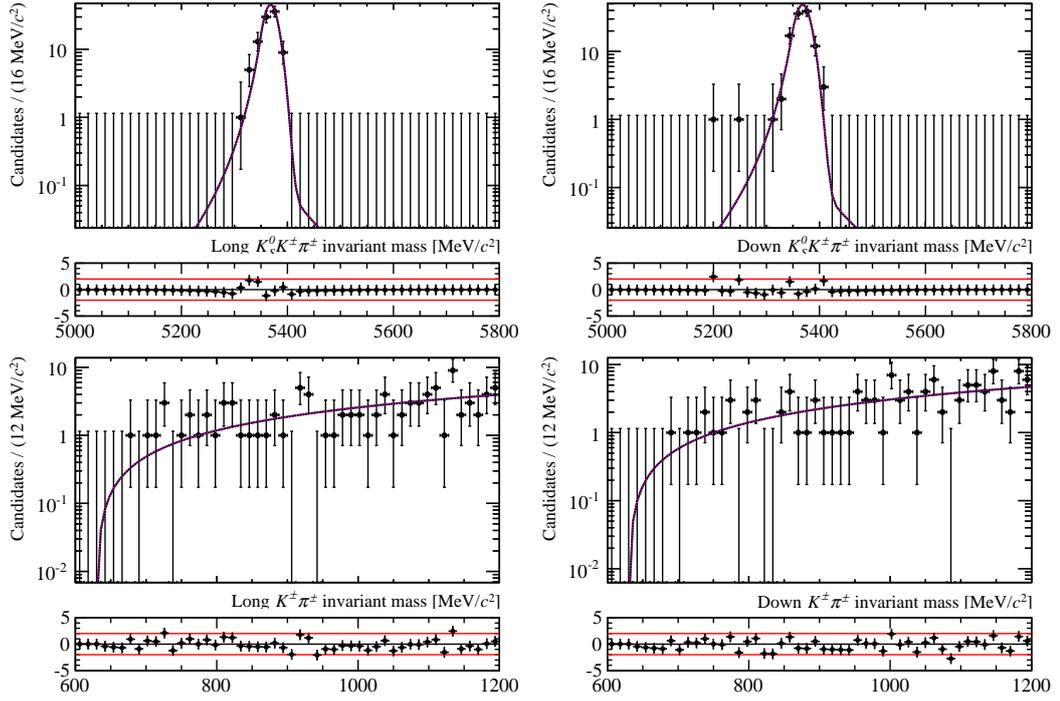


Figure 46.: $B_s^0 \rightarrow K_s^0 K^\pm \pi^\mp$ non resonant fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_s^0 K \pi}$ and $m_{K \pi}$ observables with the pull distributions shown underneath. Top: $m_{K_s^0 K \pi}$. Bottom: $m_{K \pi}$. Left: Long-Long K_s^0 sample. Right: Down-Down K_s^0 sample.

diate state like $B_d^0 \rightarrow K_s^0 \rho(770)^0$, could potentially affect the signal yields. However the studies using simulated data show that this background is present only at a negligible level due to the cut on the $K\pi$ mass window. Therefore this component will be neglected in the final fit to data.

5.2.4 Partially reconstructed background model

Partially reconstructed background can come either from charmless or charmed decays of B hadrons, in which one of the final state particle is not reconstructed. The shapes and the parameters of the mass distribution are derived from dedicated simulated samples. Among the partially reconstructed decays listed in Tab. 7 only two of them are used to parametrise this background: the $B^+ \rightarrow D^0(K_s^0 \pi^- \pi^+) K^+$ and the $B_s^0 \rightarrow K^{*0}(K^\pm \pi^\mp) K^{*0}(K_s^0 \pi^0)$ decays. In both cases the $K_s^0 K\pi$ spectrum is satisfactorily modeled using a convolution of an ARGUS function

$$A(m; m_t, c, p) = \frac{2^{-p} c^{2(p+1)}}{\Gamma(p+1) - \Gamma(p+1, c^2/2)} \cdot \frac{m}{m_t^2} \left(1 - \frac{m^2}{m_t^2}\right)^p \exp\left[-\frac{1}{2} c^2 \left(1 - \frac{m^2}{m_t^2}\right)\right] \quad (56)$$

with a Gaussian. This PDF introduces three parameters: m_t is the threshold mass value (fixed to the difference between the B and the missing particle masses), c is the curvature of the function and p controls the power law behaviour of the tails. The $K\pi$ spectrum is instead modeled using in the charmed case a linear function very similar to the one of the non resonant background, and in the charmless case a spin-1 relativistic Breit-Wigner sharing all the parameters with the signal one.

Figures 47-48 show the reconstructed mass distribution of the stripped and truth matched simulated events. The values of the fitted parameters are reported in Tab. 28. These parameters will be fixed in the fit to data. The partially reconstructed backgrounds suffer from a large statistical uncertainty due to the limited MC statistics. This could potentially lead to a systematic uncertainty which will be treated in Sec. 7.2.

5.2.5 Combinatorial background model

To model the combinatorial background the right-hand sideband from the $m_{K_s^0 K\pi}$ spectrum after the entire selection is used. In the $m_{K_s^0 K\pi}$ spectrum it is modeled using an exponential function. The slope is left free in data and only a scale factor to account for possible differences between Long-Long and Down-Down slopes is

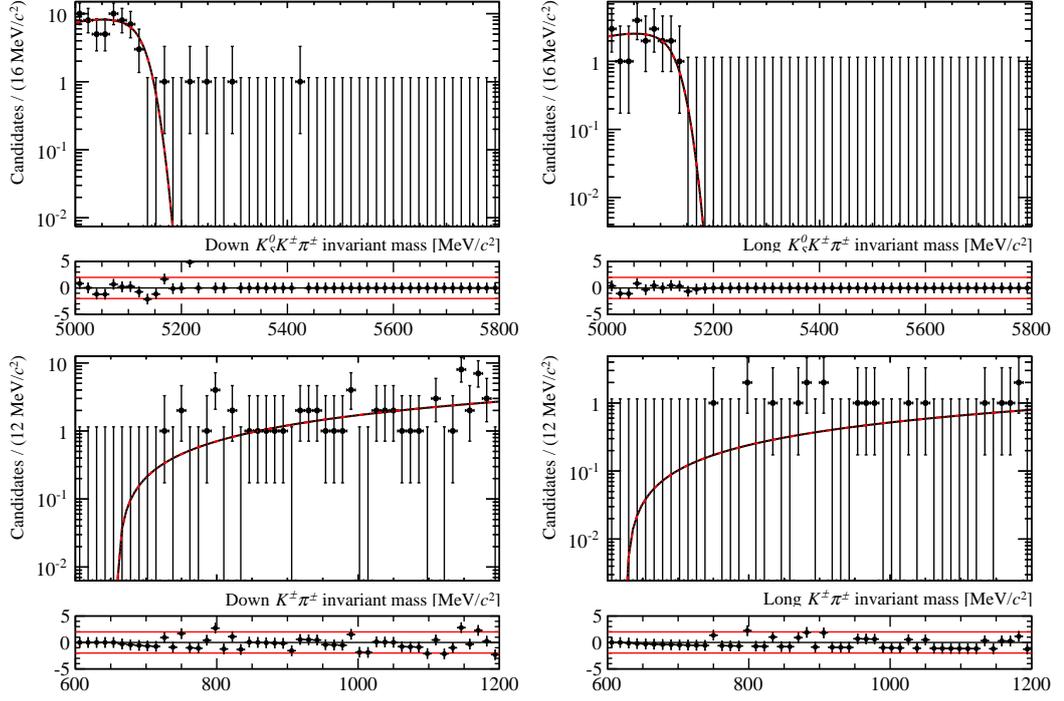


Figure 47.: $B^+ \rightarrow D^0(K_s^0 \pi^- \pi^+) K^+$ background fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_s^0 K \pi}$ and $m_{K \pi}$ observables with the pull distributions shown underneath. Top: $m_{K_s^0 K \pi}$. Bottom: $m_{K \pi}$. Left: Long-Long K_s^0 sample. Right: Down-Down K_s^0 sample.

Table 28.: The fit parameters of the partially reconstructed decays.

Parameter	Value
$p(B^+ \rightarrow D^0 K^+)$	$0.71^{+0.49}_{-0.38}$
$s(B^+ \rightarrow D^0 K^+)$	-21^{+18}_{-20}
$t(B^+ \rightarrow D^0 K^+)$	642^{+14}_{-14}
$p(B_s^0 \rightarrow K^{*0} K^{*0})$	$0.83^{+0.37}_{-0.25}$
$s(B_s^0 \rightarrow K^{*0} K^{*0})$	-33^{+16}_{-21}

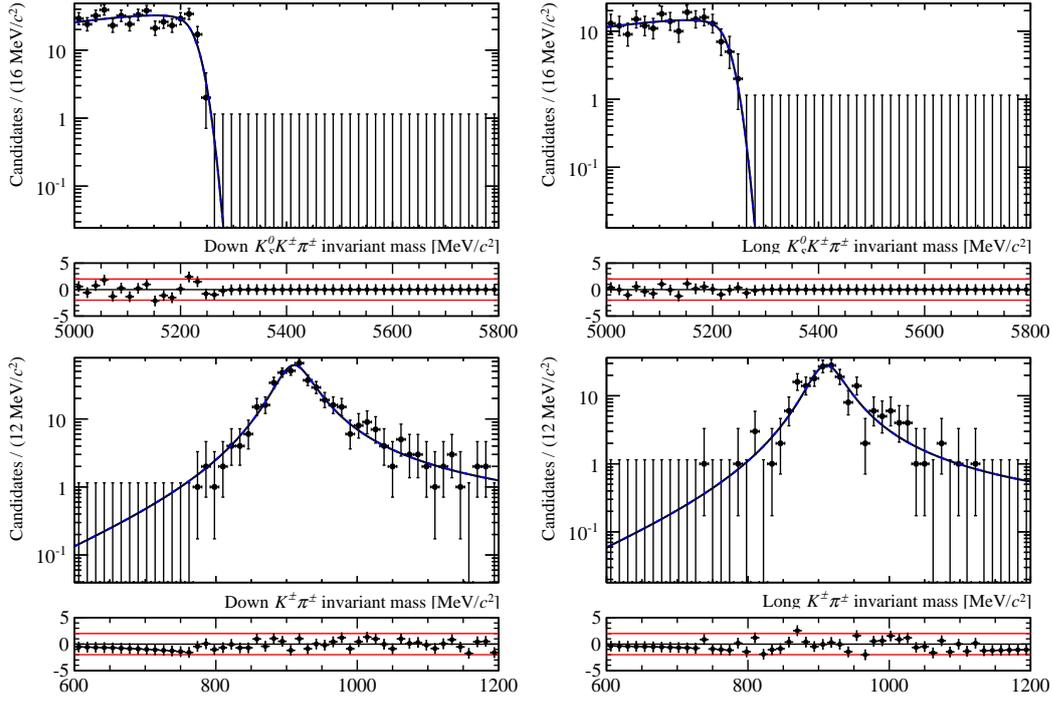


Figure 48.: $B_s \rightarrow K^*(892)^0(K^\pm\pi^\mp)K^*(892)^0(K_S^0\pi^0)$ background fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K_S^0 K\pi}$ and $m_{K\pi}$ observables with the pull distributions shown underneath. Top: $m_{K_S^0 K\pi}$. Bottom: $m_{K\pi}$. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

introduced. For the $m_{K\pi}$ spectrum instead a linear function very similar to the one defined for the non resonant background is used. The threshold parameter for the Long-Long and Down-Down categories is different and is left free in data. As an example, Fig. 49 shows the $m_{K_S^0 K\pi}$ and $m_{K\pi}$ mass distributions for events coming from the $m_{K_S^0 K\pi}$ right-hand sideband for the Down-Down category.

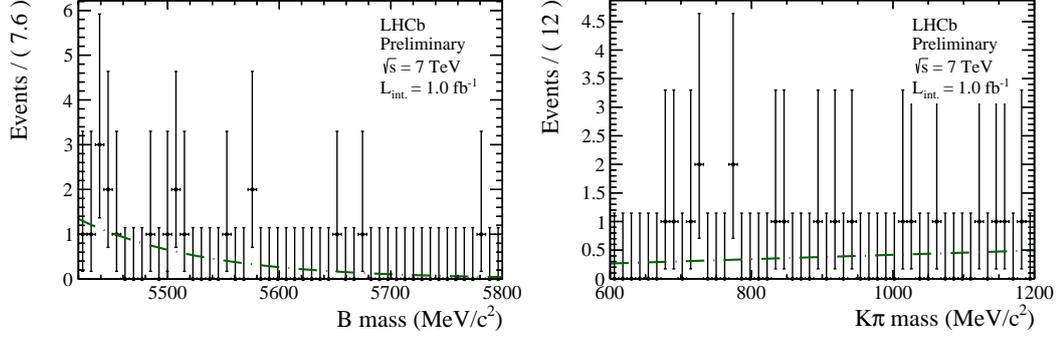


Figure 49.: $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ combinatorial background fit model for the Down-Down category. Left: $m_{K_S^0 K\pi}$. Right: $m_{K\pi}$.

5.2.6 Summary of the fit model

A summary of the PDF shapes used for the signal and the background fit model is given in Tab. 29.

Table 29.: Summary of PDF shapes used for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ fit model.

Component	m_B PDF	$m_{K\pi}$ PDF
$B^0 \rightarrow K^{*0} K_S^0$	Double Crystal Ball	Relativistic Breit-Wigner
$B_S^0 \rightarrow K^{*0} K_S^0$	Double Crystal Ball	Relativistic Breit-Wigner
$B^0 \rightarrow K_S^0 K^\pm \pi^\mp$	Double Crystal Ball	Linear
$B_S^0 \rightarrow K_S^0 K^\pm \pi^\mp$	Double Crystal Ball	Linear
$B^+ \rightarrow D^0 K^+$	ARGUS	Linear
$B_S^0 \rightarrow K^{*0} K^{*0}$	ARGUS	Relativistic Breit-Wigner
Combinatorial	Exponential	Linear

The free parameters in the final fit model for the $K_S^0 K\pi$ final state are given in Tab. 30. Any parameter without a K_S^0 reconstruction category label (DD or LL) is shared between the models during the simultaneous fit.

Table 30.: Free fit model parameters for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays.

Free parameter symbol	Description
$N_{B^0 \rightarrow K^{*0} K_S^0}(\text{LL})$	Yield of Long-Long $B^0 \rightarrow K^{*0} K_S^0$
$N_{B^0 \rightarrow K^{*0} K_S^0}(\text{DD})$	Yield of Down-Down $B^0 \rightarrow K^{*0} K_S^0$
$N_{B_s^0 \rightarrow K^{*0} K_S^0}(\text{LL})$	Yield of Long-Long $B_s^0 \rightarrow K^{*0} K_S^0$
$N_{B_s^0 \rightarrow K^{*0} K_S^0}(\text{DD})$	Yield of Down-Down $B_s^0 \rightarrow K^{*0} K_S^0$
$R_{B^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	Ratio of $B^0 \rightarrow K_S^0 K^\pm \pi^\mp / B^0 \rightarrow K^{*0} K_S^0$ yield
$R_{B_s^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	Ratio of $B_s^0 \rightarrow K_S^0 K^\pm \pi^\mp / B_s^0 \rightarrow K^{*0} K_S^0$ yield
$N_{B^+ \rightarrow D^0 K^+}$	Yield of $B^+ \rightarrow D^0 K^+$
$N_{B_s^0 \rightarrow K^{*0} K^{*0}}$	Yield of $B_s^0 \rightarrow K^{*0} K^{*0}$
$N_{\text{Comb}}(\text{DD})$	Yield of Down-Down Combinatorial
$N_{\text{Comb}}(\text{LL})$	Yield of Long-Long Combinatorial
$\bar{m}_{B_s^0}$	Mean of the $B_s^0 \rightarrow K^{*0} K_S^0 B_s^0$ mass peak
\bar{m}_{K^*}	Mean of the $B_s^0 \rightarrow K^{*0} K_S^0 K^*$ mass peak
σ_B	Width of the $B_s^0 \rightarrow K^{*0} K_S^0 B_s^0$ mass peak
Γ_{K^*}	Width of the $B_s^0 \rightarrow K^{*0} K_S^0 K^*$ mass peak
s	Exponential slope for Down-Down Combinatorial of m_B
r	Ratio of the LL and DD exponential slopes for Combinatorial of m_B
$t_{K^*}(\text{DD})$	Threshold for the Down-Down Combinatorial of $m_{K\pi}$
$t_{K^*}(\text{LL})$	Threshold for the Long-Long Combinatorial of $m_{K\pi}$

5.2.7 Simulation studies

Monte Carlo pseudo-experiments ("toy samples") are used to validate the final fit model and to check its stability. In these studies samples of similar content to the measured data samples are generated and fitted. Data is generated from the nominal PDF for each component. Because of the large uncertainty on the expected signal yields, samples with different configurations of the signal yields were generated. The number of background events as well as the other background parameters were generated with the value extracted directly from the blind fit to data shown in Fig. 50. It is worth to note that in the fit to data gaussian constraints are applied to the width of the B and the K^* mass. The mean values for the constraints are taken from the fit results to simulated samples reported in Tab. 26 and are also used as generated values for these parameters. Each study has been performed with 500 toy samples for each signal yield configuration.

A useful quantity to check the stability and the absence of bias for a certain parameter θ in the fit is the *pull* distribution, defined as

$$p_i = \frac{\theta_i^{\text{fit}} - \theta_i^{\text{gen}}}{\sigma^{\theta^{\text{fit}}}}, \quad (57)$$

where i is the index of the particular pseudo-experiment, θ_i^{fit} and θ_i^{gen} are the fitted and generated value of the parameter, and $\sigma^{\theta^{\text{fit}}}$ is the error on the fitted value. The value of $\theta_i^{\text{fit}} - \theta_i^{\text{gen}}$ can be taken as the bias of the fit. For an unbiased fit the pull distribution should approach a Gaussian distribution with mean zero and unit width. The values of the generated and fitted values, followed by the error, the bias and the pull mean and width are reported in Tab. 31. It can be observed that a small bias is present, reflecting the problems of the likelihood fit with low statistics. A systematic uncertainty is assigned for this bias in Sec. 7.2.4.

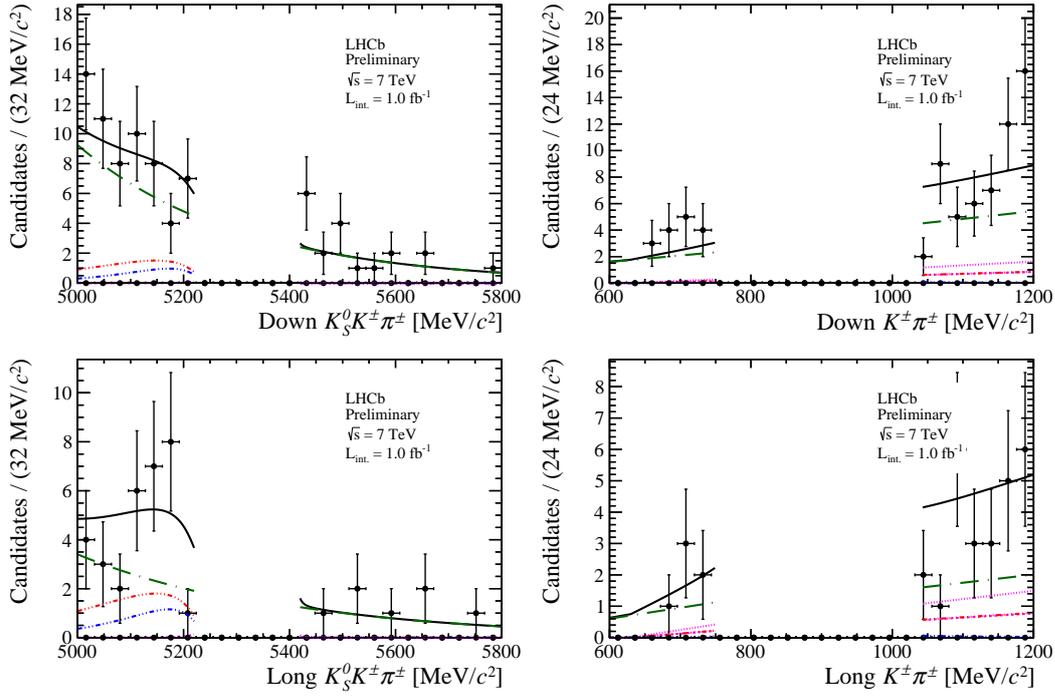


Figure 50.: $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ blind spectra. Top: Down-Down category. Bottom: Long-Long category Left: $m_{K_S^0 K \pi}$. Right: $m_{K \pi}$. The 2011 data are described by the points. Each component of the fit model is displayed on the plot except the signal. The non resonant background is in magenta, combinatorics is the green dashed line, partially reconstructed charmed transitions are drawn in red and partially reconstructed charmless transition are in blue. The overall fit is represented by the black line.

Table 31.: Signal yields obtained from fits to 500 toy samples with gaussian constraints on signal shape parameters applied. The number of generated events (Gen), the mean value of the distribution of the fitted yield (Fit), the mean value of the distribution of the error on the fitted yield (Err), the mean bias on the fitted yield (Bias, computed as Fit-Gen), the mean value of the pull and the sigma of the pull are shown for different signal scenarios.

Component	Gen	Fit	Err	Bias	Pull μ	Pull σ
B^0 Long-Long	5	4.66 ± 0.26	4.23	-0.34 ± 0.26	-0.02 ± 0.06	0.91 ± 0.06
B^0 Down-Down	5	5.96 ± 0.20	2.79	0.96 ± 0.20	0.38 ± 0.04	0.78 ± 0.05
B_s Long-Long	15	14.13 ± 0.46	6.36	-0.87 ± 0.46	0.14 ± 0.06	0.97 ± 0.06
B_s Down-Down	20	17.81 ± 0.52	6.33	-2.19 ± 0.52	-0.07 ± 0.06	0.95 ± 0.06
B^0 Long-Long	5	4.40 ± 0.24	4.07	-0.6 ± 0.24	0.10 ± 0.06	0.93 ± 0.06
B^0 Down-Down	5	6.18 ± 0.21	2.95	1.18 ± 0.21	0.50 ± 0.05	0.73 ± 0.04
B_s Long-Long	20	18.41 ± 0.52	6.29	-1.59 ± 0.52	0.04 ± 0.06	0.99 ± 0.05
B_s Down-Down	25	22.37 ± 0.61	6.85	-2.63 ± 0.61	-0.06 ± 0.06	0.94 ± 0.06
B^0 Long-Long	5	4.76 ± 0.22	3.88	-0.24 ± 0.22	0.01 ± 0.06	1.03 ± 0.06
B^0 Down-Down	5	6.57 ± 2.69	2.69	1.33 ± 0.17	0.52 ± 0.04	0.71 ± 0.05
B_s Long-Long	25	25.17 ± 0.45	6.57	0.17 ± 0.45	0.10 ± 0.05	0.82 ± 0.04
B_s Down-Down	30	29.34 ± 0.48	7.09	-0.66 ± 0.48	-0.02 ± 0.07	1.02 ± 0.06
B^0 Long-Long	5	5.35 ± 0.23	4.71	0.35 ± 0.23	0.10 ± 0.05	0.89 ± 0.05
B^0 Down-Down	10	12.20 ± 0.23	3.91	2.20 ± 0.23	0.36 ± 0.06	1.01 ± 0.07
B_s Long-Long	20	20.03 ± 0.35	6.48	0.03 ± 0.35	0.06 ± 0.05	0.90 ± 0.05
B_s Down-Down	25	25.06 ± 0.38	7.65	0.06 ± 0.38	0.10 ± 0.05	0.95 ± 0.05

5.3 $B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$ MODEL

The fit model for the $K_s^0 \pi \pi$ final state was extensively studied for the inclusive $B_{d,s} \rightarrow K_s^0 h h'$ analysis and for the measurement presented in this thesis only few changes are made.

5.3.1 Signal model

A sum of two Crystal Ball PDFs sharing the mean and the width is again used to parametrise the signal. All the considerations for the relations among the parameters in the B spectrum of the $K_s^0 K \pi$ fit also hold in this case. The only difference is that since the 1D fit suffers less in stability than the 2D one, here no constraint on the $B_s^0 - B_d^0$ mass difference is applied. The parameters extracted from simulation after the whole selection are summarised in Tab. 32 and the mass plots are shown in Fig. 51.

Table 32.: Fitted parameter values after the fit to $K_s^0 \pi \pi$ simulated events.

Parameter	Value
$m(B_s^0)$	5366.42 ± 0.18
$m(B^0)$	5279.64 ± 0.18
σ_1	15.31 ± 0.13
$\sigma(\text{LL})/\sigma(\text{DD})$	0.976 ± 0.013
α_1	1.42 ± 0.21
α_2/α_1	-1.56 ± 0.33
n_1	1.71 ± 0.13
n_2/n_1	1.07 ± 0.13
f_2	0.51 ± 0.14

5.3.2 Misidentified signal model

Contrary to what happens in the $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ case, the misidentification of a kaon as a pion from the inclusive $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$ decay is not negligible for the $B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$ decays. Hence this component will be present in the final fit to the data. On the other hand the double misidentification from the $B_{d,s}^0 \rightarrow K_s^0 K^+ K^-$

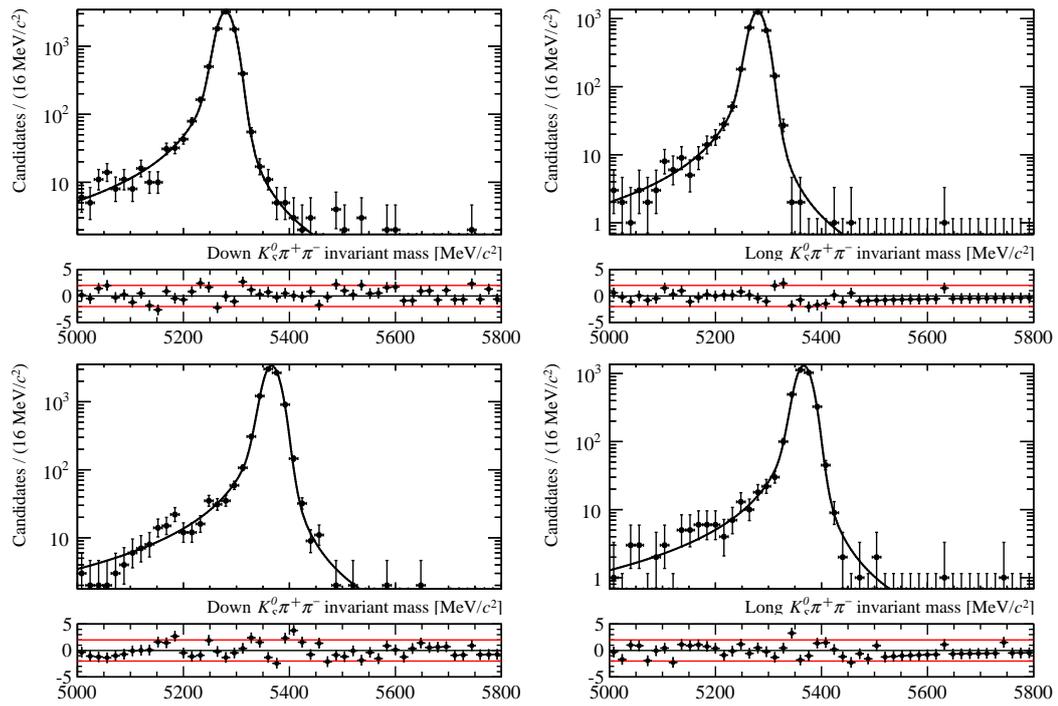


Figure 51.: $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ signal fit model fitted to simulated events. The pull distributions are shown underneath. Top: B^0 . Bottom: B_s^0 . Left: Down-Down K_S^0 sample. Right: Long-Long K_S^0 sample.

decays has been neglected. The misidentified signals are modeled with single CB functions plus a small flat combinatorial background. The following conditions were applied:

- The parameters for the B^0 and B_s^0 decays are the same, except the mean value;
- The Down-Down and Long-Long parameters are the same;
- Since the fit is very sensitive to the α and n parameters, they were fixed already in the fit to simulation.

The fit on the simulated data samples, after the PID, trigger and truth matched selection, is performed simultaneously on B^0 , B_s^0 , Down-Down and Long-Long and the results are shown in Fig. 52. The results of the fit are shown in Tab. 33.

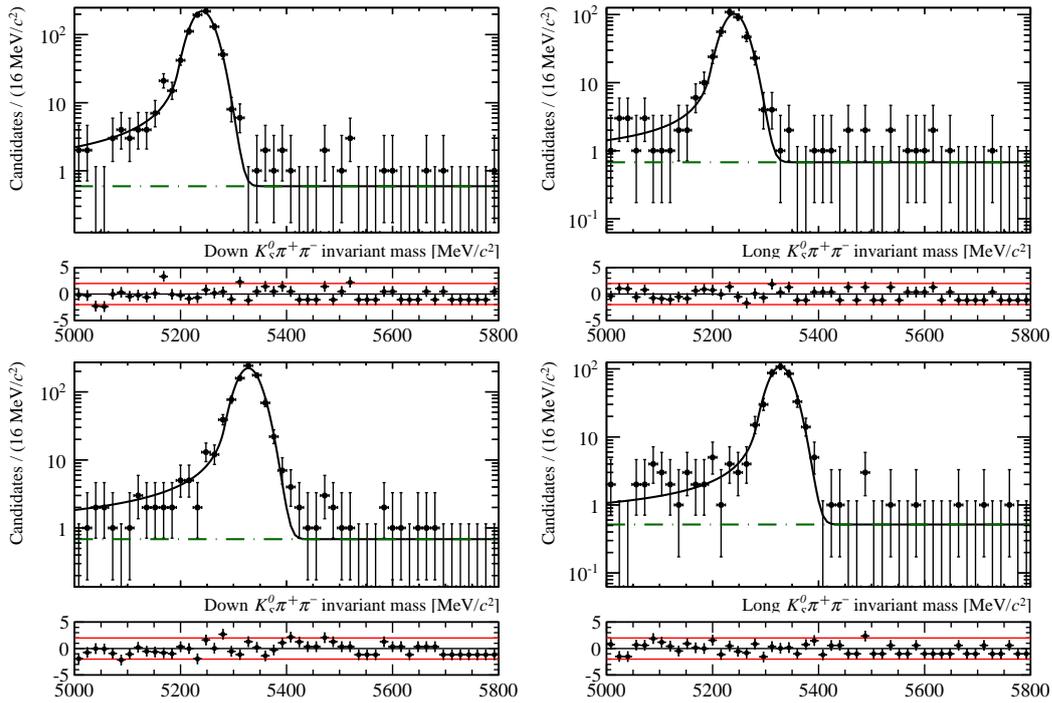


Figure 52.: $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ misidentified fit model fitted to simulated events. Top: B^0 . Bottom: B_s^0 . Left: Down-Down K_S^0 sample. Right: Long-Long K_S^0 sample.

Table 33.: Misidentified signal fit model parameters extracted from $B_{d,s}^0 \rightarrow K_s^0 K^\pm \pi^\mp$ simulated samples reconstructed as $B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$.

Parameter	Value
α	2 ± 2
n	1 ± 100
$\mu(B_d^0)$	5240.61 ± 0.73
$\mu(B_s^0)$	5327.94 ± 0.71
σ	22.47 ± 0.44

5.3.3 Partially reconstructed background model

The partially reconstructed backgrounds which may be present in the left sideband of the B spectrum were already listed in Tab. 8. Like for the signal, among all the simulation samples only a single decay for each category has been used to extract the parameters to fix on data. The shapes are again modeled with ARGUS functions convolved with a gaussian. The results of the fit for truth matched and selected simulated events are reported in Tab. 34 while the plots are shown in Fig. 53.

Table 34.: The fit parameters of the partially reconstructed decays.

Parameter	Value
$p(B^0 \rightarrow \eta'(\rho^0 \gamma) K_s^0)$	$1.12^{+0.16}_{-0.15}$
$s(B^0 \rightarrow \eta'(\rho^0 \gamma) K_s^0)$	-30^{+4}_{-4}
$p(B^0 \rightarrow K_s^0 \pi^+ \pi^- \gamma)$	$1.34^{+0.64}_{-0.54}$
$s(B^0 \rightarrow K_s^0 \pi^+ \pi^- \gamma)$	-13^{+11}_{-13}
$p(B \rightarrow Dh \rightarrow K_s^0 K \pi X)$	$0.74 \pm^{+0.41}_{-0.35}$
$s(B \rightarrow Dh \rightarrow K_s^0 K \pi X)$	-22^{+16}_{-18}
$p(B \rightarrow K_s^0 \pi \pi X)$	$0.79^{+0.50}_{-0.41}$
$s(B \rightarrow K_s^0 \pi \pi X)$	-34^{+19}_{-22}

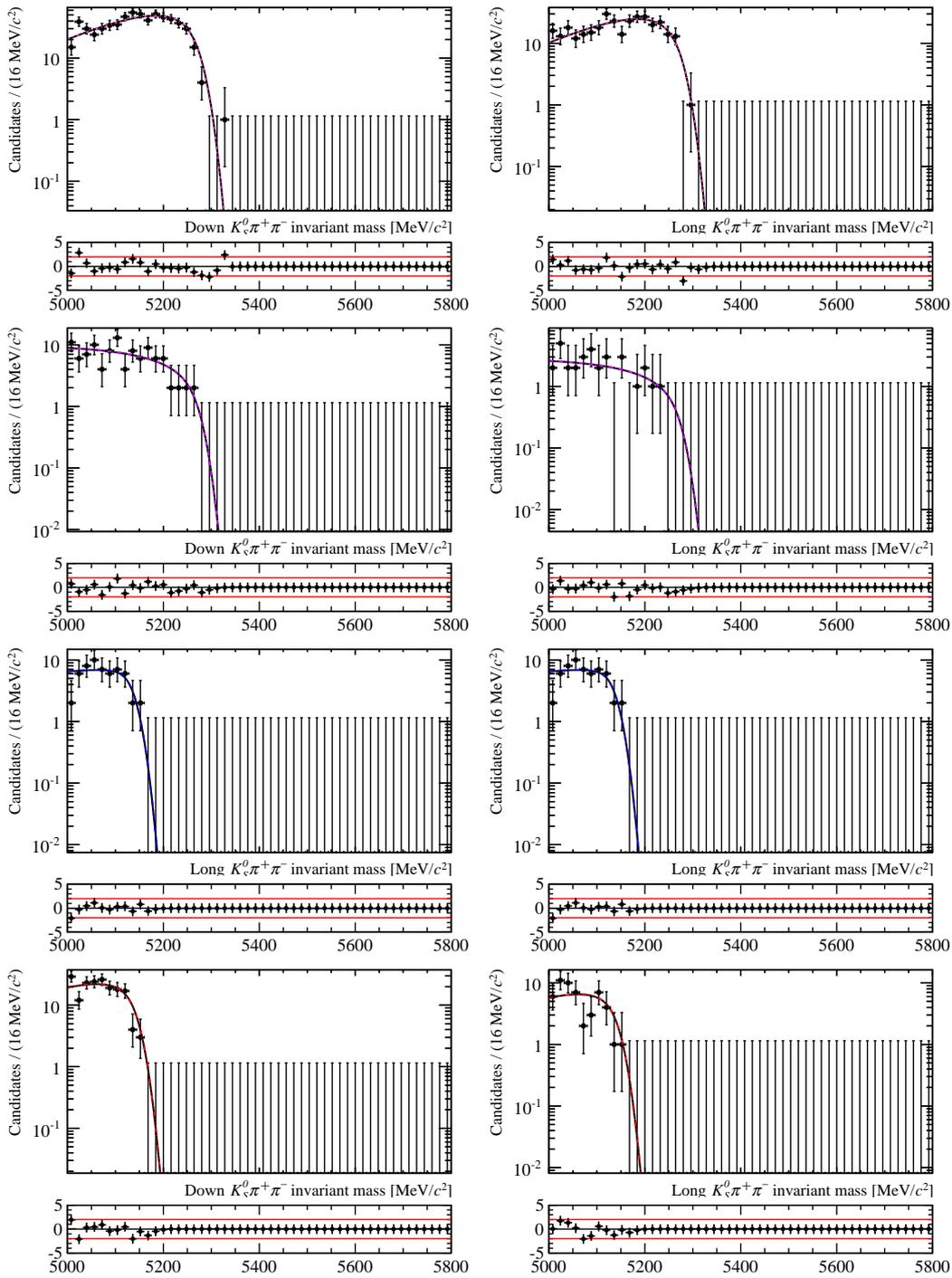


Figure 53.: $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ partially reconstructed backgrounds fitted to simulated events. From the Top to the Bottom: $B^0 \rightarrow \eta'(\eta \pi^+ \pi^-) K_S^0$, $B^0 \rightarrow K_S^0 \pi^+ \pi^- \gamma$, $B^+ \rightarrow D^0(K_S^0 \pi^+ \pi^-) \pi^+$, $B^0 \rightarrow K^{*0}(K_S^0 \pi^0) \rho^0(\pi^+ \pi^-)$. Left: Down-Down K_S^0 sample. Right: Long-Long K_S^0 sample.

5.3.4 Combinatorial background model

The combinatorial background is modeled with an exponential function. The slope is left free in data and only a scale factor to account for possible differences between Long-Long and Down-Down slopes is introduced.

5.3.5 Summary of the fit model

A summary of the PDF shapes used for the signal and the background fit model is given in Tab. 35. The free parameters in the final fit model for the $K_s^0 \pi \pi$ final state are given in Tab. 36. Any parameter without a K_s^0 reconstruction category label (DD or LL) is shared between the models during the simultaneous fit. The fit model for the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ decay was already well tested in the inclusive analysis so no additional studies were performed.

Table 35.: Summary of PDF shapes used for the $B_{d,s}^0 \rightarrow K_s^0 \pi^+ \pi^-$ fit model.

Component	m_B PDF
$B^0 \rightarrow K_s^0 \pi^+ \pi^-$	Double Crystal Ball
Misid. $B^0 \rightarrow K_s^0 \pi^+ \pi^-$	Single Crystal Ball
$B \rightarrow Dh \rightarrow K_s^0 K \pi X$	ARGUS
$B \rightarrow K_s^0 \pi \pi X$	ARGUS
$B^0 \rightarrow \eta'(\rho^0 \gamma) K_s^0$	ARGUS
$B^0 \rightarrow K_s^0 \pi^+ \pi^- \gamma$	ARGUS
Combinatorial	Exponential

Table 36.: Free fit model parameters for the $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ decays.

Free parameter symbol	Description
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{LL})$	Yield of Long-Long $B^0 \rightarrow K_S^0 \pi^+ \pi^-$
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{DD})$	Yield of Down-Down $B^0 \rightarrow K_S^0 \pi^+ \pi^-$
$N_{B_s^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{LL})$	Yield of Long-Long $B_s^0 \rightarrow K_S^0 \pi^+ \pi^-$
$N_{B_s^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{DD})$	Yield of Down-Down $B_s^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{\text{Misid.} B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{LL})$	Ratio of yield of Long-Long Misid $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{\text{Misid.} B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-}(\text{DD})$	Ratio of yield of Down-Down Misid $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{B \rightarrow Dh \rightarrow K_S^0 K \pi X}$	Ratio of yield of yield of $B \rightarrow Dh \rightarrow K_S^0 K \pi X$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{B \rightarrow K_S^0 \pi \pi X}$	Ratio of yield of $B \rightarrow K_S^0 \pi \pi X$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{B^0 \rightarrow \eta'(\rho^0 \gamma) K_S^0}$	Ratio of yield of $B^0 \rightarrow \eta'(\rho^0 \gamma) K_S^0$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$R_{B^0 \rightarrow K_S^0 \pi^+ \pi^- \gamma}$	Ratio of yield of $B^0 \rightarrow K_S^0 \pi^+ \pi^- \gamma$ and $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$
$N_{\text{Comb}}(\text{DD})$	Yield of Down-Down Combinatorial.
$N_{\text{Comb}}(\text{LL})$	Yield of Long-Long Combinatorial.
\bar{m}_{B^0}	Mean of the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ B^0 mass peak.
$\bar{m}_{B_s^0}$	Mean of the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ B_s^0 mass peak.
σ_B	Width of the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ B mass peak.
r_{σ_B}	Ratio of Down-Down and Long-Long widths of m_B .
s	Exponential slope for Down-Down Combinatorial of m_B
r	Ratio of the LL and DD exponential slopes for Combinatorial of m_B

5.4 FIT RESULTS ON DATA

5.4.1 $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ fit results

The results of the fit to the data sample for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays are shown in Fig. 54 and the fit parameters results are given in Tab. 37. Data and fit agree within 2σ in almost the entire mass range. The values of the fitted signal yields are compatible with zero for the unfavored $B^0 \rightarrow K^{*0} K_S^0$ decay, while non-zero yields result for the $B_S^0 \rightarrow K^{*0} K_S^0$ mode. More detailed considerations on the number of signal candidates and on the relative significance are reported in Sec. 8.1.

Table 37.: Fitted parameter values after the fit to data for the $K_S^0 K \pi$ final state. Only statistical uncertainties from the fit are given.

Fit parameter	Fitted value
$\bar{m}_{B_S^0}$	5372 ± 2
\bar{m}_{K^*}	882 ± 5
σ_B	14.43 ± 0.26
$\Gamma(K^*)$	49 ± 1
s	-0.00327 ± 0.00059
r	0.80 ± 0.37
$t_{K^*}(DD)$	272 ± 281
$t_{K^*}(LL)$	427 ± 340
$N_{B^0 \rightarrow K^{*0} K_S^0}(DD)$	2 ± 3
$N_{B^0 \rightarrow K^{*0} K_S^0}(LL)$	1 ± 2
$N_{B_S^0 \rightarrow K^{*0} K_S^0}(DD)$	21 ± 6
$N_{B_S^0 \rightarrow K^{*0} K_S^0}(LL)$	25 ± 6
$R_{B^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	5 ± 8
$R_{B_S^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	0.94 ± 0.41
$N_{Comb}(DD)$	96 ± 15
$N_{Comb}(LL)$	35 ± 15
$N(B \rightarrow Dh)$	11 ± 9
$N(B \rightarrow K^* K^*)$	6 ± 4

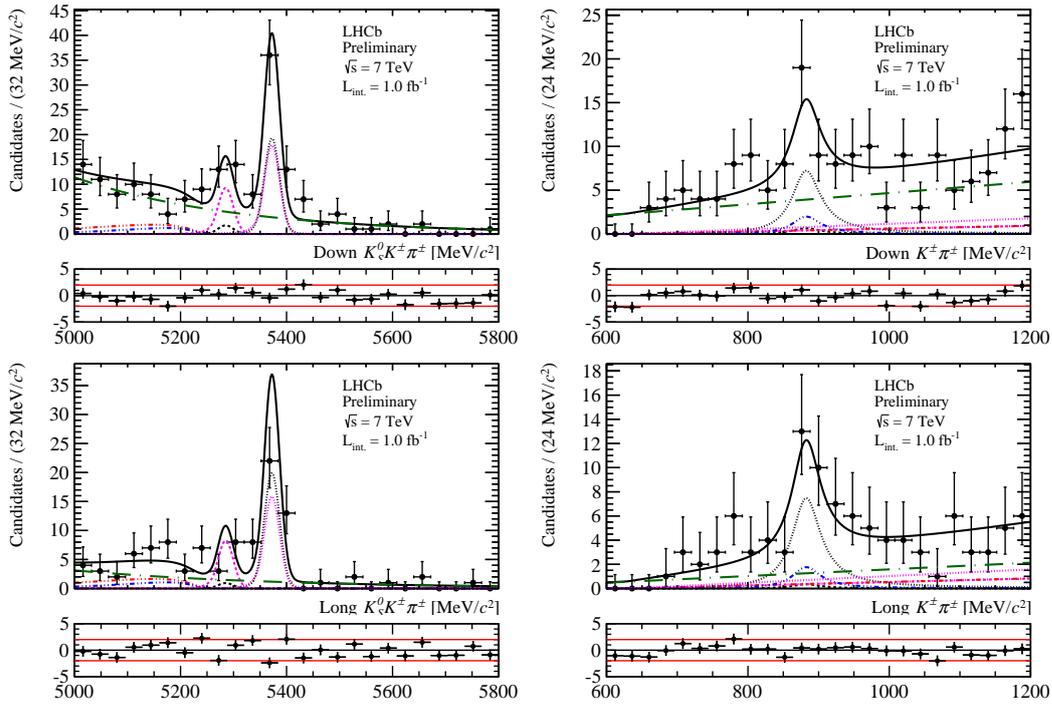


Figure 54.: $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ fit results. Top: Down-Down category. Bottom: Long-Long category Left: $m_{K_S^0 K \pi}$. Right: $m_{K \pi}$. The 2011 data are displayed by the points. Each component of the fit model is shown on the plot. The signals are in dashed black, the non resonant background is in magenta, the combinatorics is the green dashed line, partially reconstructed charmed transitions in red, partially reconstructed charmless transition in blue. The overall fit is represented by the black line.

5.4.2 $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ fit results

The results of the fit to the data sample for the $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ decay are shown in Fig. 55 and the fit parameters results are given in Tab. 38. Also in this case the residuals show a good agreement in the entire mass range.

Table 38.: Fitted parameter values after the fit to data for the $K_S^0 \pi \pi$ final state. Only statistical uncertainties from the fit are given.

Fit parameter	Fitted value
\bar{m}_{B^0}	5283.94 ± 0.78
$\bar{m}_{B_S^0}$	5374 ± 5
σ_B	17.94 ± 0.88
r_{σ_B}	1.181 ± 0.093
s	-0.00198 ± 0.00014
r	2.07 ± 0.24
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-} (DD)$	828 ± 41
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-} (LL)$	341 ± 23
$N_{B_S^0 \rightarrow K_S^0 \pi^+ \pi^-} (DD)$	71 ± 19
$N_{B_S^0 \rightarrow K_S^0 \pi^+ \pi^-} (LL)$	38 ± 9
$N_{Comb} (DD)$	2554 ± 97
$N_{Comb} (LL)$	292 ± 41
$R_{Misid B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-} (DD)$	0.0505 ± 0.0054
$R_{Misid B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-} (LL)$	0.0480 ± 0.0059
$R_{B \rightarrow Dh \rightarrow K_S^0 K \pi X}$	0.472 ± 0.056
$R_{B^0 \rightarrow \eta' (\rho^0 \gamma) K_S^0}$	0.0399 ± 0.0040
$R_{B^0 \rightarrow K_S^0 \pi^+ \pi^- \gamma}$	0.0404 ± 0.0080
$R_{B \rightarrow K_S^0 \pi \pi X}$	0.054 ± 0.037

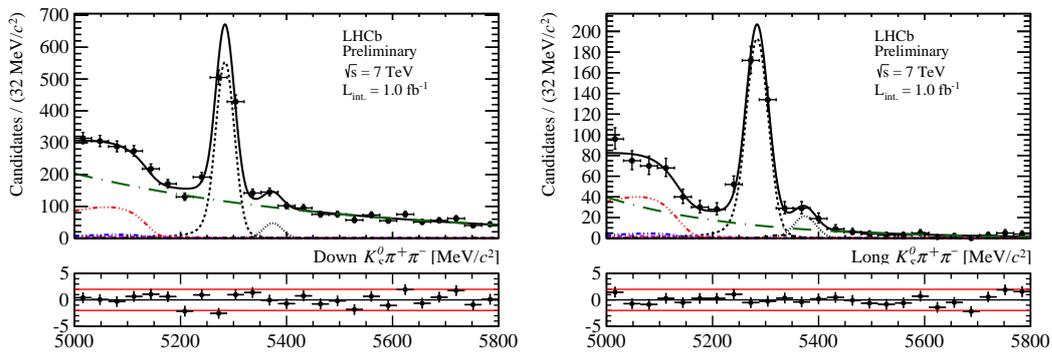


Figure 55.: $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ fit results. Left: Down-Down category. Right: Long-Long category. The 2011 data are described by the points. Each component of the fit model is displayed on the plot. The signal and the crossfeed are in dashed black, the combinatorics is the green dashed line, partially reconstructed charmed transitions in red, partially reconstructed charmless transition in blue, partially reconstructed radiative in violet and partially reconstructed radiative resonant in magenta. The overall fit is represented by the black line.

6 | EFFICIENCIES

A key ingredient for the measurement of a branching ratio is the determination of the efficiencies due to reconstruction and selection effects. In this chapter the determination of the efficiencies is reported first for the signal and then for the normalization channel.

6.1 GENERAL CONSIDERATIONS

In the definition of the branching ratio (Eq. 40), different efficiencies appear: the acceptance efficiency ϵ^{Acc} , the combined trigger, reconstruction, stripping and of-line selection efficiencies (ϵ^{Sel}) and the PID efficiency (ϵ^{PID}). The use of Monte Carlo simulated data to estimate ϵ^{Sel} is considered reliable, while the particle identification efficiency is taken instead from data. To estimate the reconstruction and stripping efficiencies only the generator level cuts are applied, while the PID efficiencies are estimated for events that pass all the other cuts. The factorisation of the efficiency is useful to understand the different systematic effects that can affect the measurement. However, the final estimate is given by an overall combination of the intermediate steps. The efficiencies for the signal and the normalization channel are obtained with two different approaches:

- The efficiencies for the signal can be easily determined counting the number of events which pass a certain selection criteria in simulated samples;
- For the normalization channel a more sophisticated method is necessary. The dynamical structure of three body decays can be described in the Dalitz plane. This can be in general rather complex due to the presence of two body decays, non resonant contributions and interference among them. Generally the efficiencies are not flat over the Dalitz plane and it is thus a bad approximation to consider that the efficiency does not vary on it. The method developed for the $B_{d,s} \rightarrow K_s^0 h h'$ analysis consists of calculating the efficiencies in bins of the Dalitz plane and weighting them based on the position in phase space determined from the data.

6.2 $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ EFFICIENCIES

6.2.1 Detector acceptance

In order for a B meson to be reconstructed it is necessary that the daughter tracks are within the LHCb acceptance, corresponding to a polar angle between 10 and 400 mrad. To avoid the full simulation of all events, already at the generator level an acceptance cut is applied. The efficiencies for this cut are taken from the table of generator and simulation statistics provided to the collaboration. The values were obtained generating ten thousand signal decays for each mode and each of the magnet polarities. The efficiencies were then calculated from the number of passing and failing decays. The results for the MC signals respectively in Magnet Down and Up configurations are shown in Tab. 39 and 40.

Table 39.: Generator level efficiency in the Magnet Up configuration.

Decay mode	particle (%)	antiparticle (%)	average (%)
$B^0 \rightarrow K^{*0} K_s^0$	23.72 ± 0.10	23.58 ± 0.10	23.65 ± 0.08
$B_s^0 \rightarrow K^{*0} K_s^0$	23.35 ± 0.10	23.64 ± 0.10	23.50 ± 0.08

Table 40.: Generator level efficiency in the Magnet Down configuration.

Decay mode	particle (%)	antiparticle (%)	average (%)
$B^0 \rightarrow K^{*0} K_s^0$	23.55 ± 0.08	23.71 ± 0.08	23.63 ± 0.06
$B_s^0 \rightarrow K^{*0} K_s^0$	23.76 ± 0.08	23.50 ± 0.08	23.63 ± 0.06

The efficiencies are in agreement for the particle (B^0 and B_s^0) and the antiparticle (\bar{B}_d^0 and \bar{B}_s^0) and the average is computed as the unweighted mean of the previous two. Also, the generator cut acceptance is independent of the initial flavor of the B meson (B^0 or B_s^0). No systematic difference is observed between Magnet Up and Down configurations.

6.2.2 Reconstruction and stripping efficiencies

The reconstruction and stripping efficiencies are combined for simplicity. Table 41 lists the computed efficiency values. The efficiencies are evaluated from

simulated signal events which have already passed the acceptance requirements. The two K_s^0 categories have different reconstruction efficiency since they are selected using different cuts, i.e. specific efficiencies for each category are reported. The calculation of the statistical error is done using the binomial formula

$$\sigma_\epsilon = \sqrt{\epsilon(1 - \epsilon)/N}, \quad (58)$$

where ϵ is the efficiency of a certain requirement and N is the number of events before the selection cut is applied.

Table 41.: Stripping and reconstruction efficiencies from Monte Carlo simulations. The errors collected in this table are only due to the statistical uncertainties of the Monte Carlo sample.

Decay mode	MagUp (%)	MagDown (%)
$B^0 \rightarrow K^{*0}K_s^0$ Down-Down	2.12 ± 0.01	2.18 ± 0.01
$B_s^0 \rightarrow K^{*0}K_s^0$ Down-Down	2.11 ± 0.01	2.14 ± 0.01
$B^0 \rightarrow K^{*0}K_s^0$ Long-Long	0.92 ± 0.03	0.93 ± 0.03
$B_s^0 \rightarrow K^{*0}K_s^0$ Long-Long	0.91 ± 0.03	0.92 ± 0.03

6.2.3 Trigger efficiencies

The trigger efficiency is normally quoted considering all the levels combined. Simply to have a more accurate perception of the effects of these criteria, the efficiency of each individual decision presented in Sec.4.4.3 is reported in Tabs.42-43 for both magnet polarities. The efficiencies are calculated relative to the previous step.

One can notice that at Level-0, the Down-Down efficiency is systematically higher than that for Long-Long. The origin of this effect may come from the larger boost, on average, of the Down-Down candidates, which could affect the L0Hadron T0S decision. Regarding the HLT1 decision, the Long-Long efficiency is higher than that for Down-Down. A more striking difference is introduced by the HLT2 decision. This is expected, since the Downstream tracks were not included in the HLT2 reconstruction in 2011.

Table 42.: Trigger, stripping and reconstruction efficiencies and relative efficiency for each step from Monte Carlo simulations for the MagDown polarity. The errors collected in this table are only due to the statistical uncertainties of the Monte Carlo sample.

Decay mode	L0/Strip (%)	HLT1/L0 (%)	HLT2/HLT1 (%)	Trig/Strip (%)
$B^0 \rightarrow K^{*0} K_S^0$ Down-Down	46.61 ± 0.33	64.37 ± 0.47	26.92 ± 0.54	8.08 ± 0.18
$B_s^0 \rightarrow K^{*0} K_S^0$ Down-Down	47.05 ± 0.34	65.12 ± 0.47	26.84 ± 0.55	8.22 ± 0.19
$B^0 \rightarrow K^{*0} K_S^0$ Long-Long	42.07 ± 0.51	69.69 ± 0.73	60.73 ± 0.93	17.81 ± 0.39
$B_s^0 \rightarrow K^{*0} K_S^0$ Long-Long	43.04 ± 0.51	71.84 ± 0.71	64.60 ± 0.89	19.98 ± 0.41

Table 43.: Trigger, stripping and reconstruction efficiencies and relative efficiency for each step from Monte Carlo simulations for the MagUp polarity. The errors collected in this table are only due to the statistical uncertainties of the Monte Carlo sample.

Decay mode	L0/Strip (%)	HLT1/L0 (%)	HLT2/HLT1 (%)	Trig/Strip (%)
$B^0 \rightarrow K^{*0} K_S^0$ Down-Down	46.41 ± 0.34	64.86 ± 0.48	27.53 ± 0.56	8.29 ± 0.19
$B_s^0 \rightarrow K^{*0} K_S^0$ Down-Down	48.06 ± 0.34	64.46 ± 0.47	25.93 ± 0.53	8.03 ± 0.18
$B^0 \rightarrow K^{*0} K_S^0$ Long-Long	43.46 ± 0.52	70.43 ± 0.72	62.05 ± 0.91	18.99 ± 0.41
$B_s^0 \rightarrow K^{*0} K_S^0$ Long-Long	44.54 ± 0.51	70.39 ± 0.70	62.26 ± 0.89	19.52 ± 0.41

6.2.4 Combined offline selection

In this section the combined efficiencies of the previous sections plus the efficiency due to the other offline cuts are reported altogether. The PID efficiencies have been determined separately in Section 6.2.5. The combined efficiencies are given in Tab. 44.

Table 44.: Absolute reconstruction and selection efficiencies.

Decay mode	MagUp (%)	MagDown (%)	Average (%)
$B^0 \rightarrow K^{*0}K_s^0$ Down-Down	0.096 ± 0.003	0.095 ± 0.003	0.096 ± 0.003
$B_s^0 \rightarrow K^{*0}K_s^0$ Down-Down	0.095 ± 0.003	0.098 ± 0.003	0.097 ± 0.003
$B^0 \rightarrow K^{*0}K_s^0$ Long-Long	0.068 ± 0.003	0.065 ± 0.003	0.067 ± 0.003
$B_s^0 \rightarrow K^{*0}K_s^0$ Long-Long	0.069 ± 0.003	0.069 ± 0.003	0.069 ± 0.003

The efficiency for the different B flavors within the same K_s^0 category are in agreement. While the difference between the efficiencies for the two K_s^0 categories reflects the different nature of the tracks.

6.2.5 PID efficiencies

The PID efficiencies corresponding to the cuts shown in Tab. 17 are evaluated with a data driven method which takes into account the possible correlations of the track kinematics. The method is split into two parts:

1. Use calibration tracks of known ID to make PID performance histograms, i.e. the efficiency for a given PID requirement as a function of different kinematic variables. The histograms can be produced for a specified magnet polarity and particle type. The performances of the RICH sub-detectors may vary during the data taking period. Therefore also the efficiency is not constant. To account for this, different efficiency histograms are made with subsets of the calibration tracks, corresponding to the running subperiods. The efficiency histograms for each run range are weighted by the integrated luminosity of the run range they correspond to. Finally a weighted average of the efficiency histograms is performed.
2. Take as input the performance histograms produced in the previous step and reweight them with a reference sample with the same kinematics of the signal decay. Then determine both the individual track, as well as the overall

efficiencies for each event in the reference sample. Finally the average PID efficiency over all events in the reference sample is determined.

In this analysis the chosen binning scheme corresponds to 16 bins in p and 4 bins in p_T , where the bin boundaries are tuned using $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ simulated events to give approximately the same number of tracks in each bin. For each event the p and p_T distributions for the kaon and pion tracks are found. These values are then used with the corresponding efficiency histograms to extract the two track efficiencies. These efficiencies are multiplied together to form an event efficiency for every event in the sample. The efficiencies are then averaged to give the final value. The estimated PID efficiencies for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ modes are reported in Tab. 45.

Table 45.: PID efficiencies for the $B^0 \rightarrow K^{*0} K_S^0$ and $B_s^0 \rightarrow K^{*0} K_S^0$ modes.

Decay mode	MagUp (%)	MagDown (%)	Average (%)
$B^0 \rightarrow K^{*0} K_S^0$ Down-Down	75.768 ± 0.005	75.988 ± 0.004	75.878 ± 0.005
$B_s^0 \rightarrow K^{*0} K_S^0$ Down-Down	75.854 ± 0.005	76.227 ± 0.004	76.041 ± 0.005
$B^0 \rightarrow K^{*0} K_S^0$ Long-Long	73.175 ± 0.005	75.241 ± 0.004	74.208 ± 0.005
$B_s^0 \rightarrow K^{*0} K_S^0$ Long-Long	73.975 ± 0.005	74.303 ± 0.004	74.139 ± 0.005

6.2.6 Summary of the efficiencies

In Tab. 46 a summary of the efficiencies averaged over magnet polarity is given for the different decay modes.

Table 46.: Summary of the efficiencies for the $B^0 \rightarrow K^{*0} K_S^0$ and $B_s^0 \rightarrow K^{*0} K_S^0$ modes.

Decay mode	Acceptance (%)	Selection (%)	PID (%)	Total (%)
$B^0 \rightarrow K^{*0} K_S^0$ Down-Down	25.14 ± 0.07	0.096 ± 0.003	75.878 ± 0.005	0.0183 ± 0.0013
$B_s^0 \rightarrow K^{*0} K_S^0$ Down-Down	23.57 ± 0.07	0.097 ± 0.003	76.041 ± 0.005	0.0174 ± 0.0012
$B^0 \rightarrow K^{*0} K_S^0$ Long-Long	25.14 ± 0.07	0.067 ± 0.003	74.208 ± 0.005	0.0125 ± 0.0009
$B_s^0 \rightarrow K^{*0} K_S^0$ Long-Long	23.57 ± 0.07	0.069 ± 0.003	74.139 ± 0.005	0.0121 ± 0.0008

6.3 $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ EFFICIENCIES

The normalization channel is considered inclusively in the final state, this means that it can be composed of several intermediate quasi-two-body states and also a non resonant contribution. Therefore, the population of events throughout the phase space is driven by the dynamics of the decay process. If the efficiency varies significantly, the integrated phase space efficiency may be a poor representation of the effective efficiency. However a priori the phase space distribution of the events is unknown and has to be determined from the data.

This can be achieved using the so called *sPlot* technique (explained in details in [76]), which allows one to unfold the contribution of signal and background sources to the distribution of a data sample in a given variable. Using a fit to a discriminating observable it is possible to extract the signal and background yields. The method uses these yields, the PDFs and the correlation matrix of the fit to calculate an *sWeight* for each species, on a per event basis. The sum of the *sWeights* over all events in the sample is equal to the fitted yields. The calculated signal *sWeights* can be used to reweight the distribution of a variable, not used in the likelihood fit, and to recover its signal distribution.

Then a common method to deal with the determination of the efficiency for three body decays is given by weighting the efficiency, using the signal *sWeights*, based on the position in phase space. To describe the phase space the square Dalitz plot variables (θ' , m') are used (see App. A). The method consists schematically of the following steps:

- Determine the total efficiency in bins of the square Dalitz plane [θ' , m'];
- Use the *sPlot* technique to extract the signal *sWeights* ω for each event. The weights are extracted by performing a fit to the m_B distribution, used as discriminating observable (Fig. 56);
- Determine the efficiency corrected yield:

$$N_{\text{corr}} = \sum_i \frac{\omega_i}{\epsilon_i}$$

where ϵ_i is the efficiency in the i^{th} bin

- The average efficiency is then given by

$$\bar{\epsilon} = \frac{N}{N_{\text{corr}}}$$

where $N = \sum_i \omega_i$

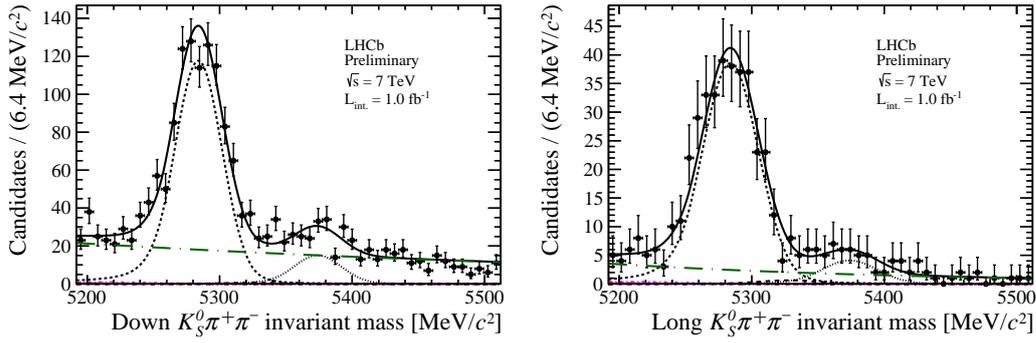


Figure 56.: $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ spectra: results of the $sWeight$ extraction. Left: Down-Down category. Right: Long-Long category.

A 10×10 binning is chosen, with the additional requirement of having a similar amount of observed events in each bin. The results are summarized in Tab. 47 and the binned efficiencies together with the associated uncertainties are given in Fig. 57. The results are averaged over magnet polarity and the error is statistical only. Just for completeness in Tab. 48 a summary of the efficiencies averaged over magnet polarity without doing the reweighting procedure is given. The total efficiency is in agreement with the one coming from the reweighting procedure, showing the fact that the efficiencies for the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ decay are not varying so much in the Dalitz plane.

Table 47.: Total efficiencies for Down-Down and Long-Long $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ [9].

Decay mode	Total (%)
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Down-Down	0.0336 ± 0.0010
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Long-Long	0.0117 ± 0.0009

Table 48.: Summary of the efficiencies for the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ mode without reweighting the Dalitz plane.

Decay mode	Acceptance (%)	Selection (%)	PID (%)	Total (%)
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Down-Down	18.76 ± 0.06	0.242 ± 0.005	78.977 ± 0.005	0.0359 ± 0.0022
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Long-Long	18.76 ± 0.06	0.096 ± 0.003	77.527 ± 0.007	0.0140 ± 0.0008

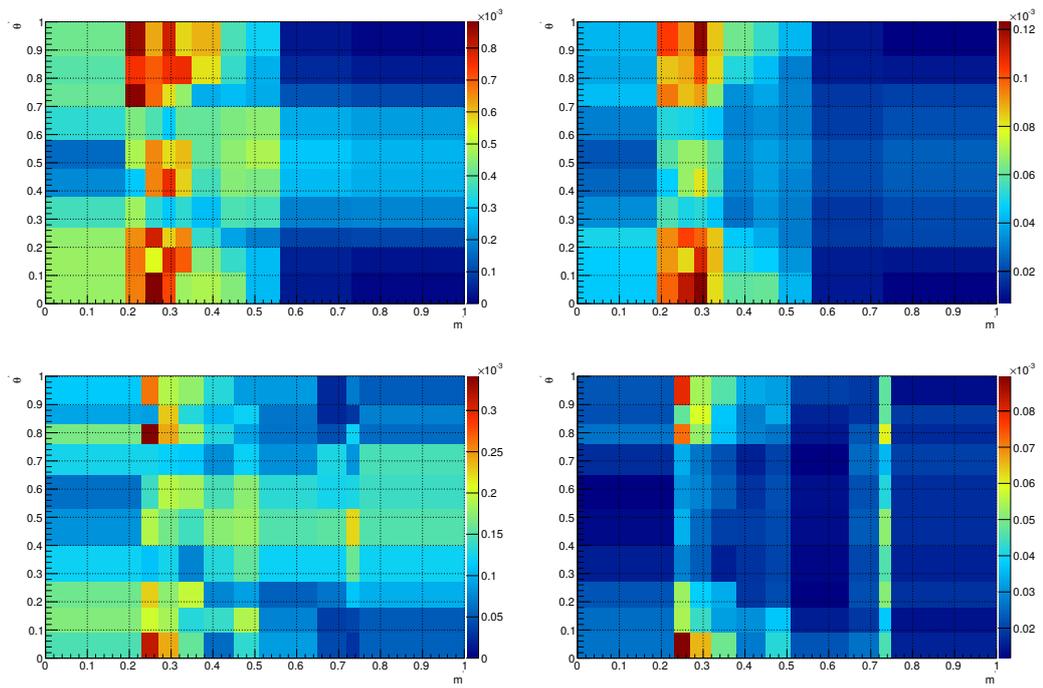


Figure 57.: Efficiencies (left) and their errors (right) for the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ decay. Top: DD category. Bottom: LL category [9].

7 | SYSTEMATICS

In this chapter the different sources of systematic uncertainties are described. Since the branching fraction of the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decay is normalised to the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ channel, and the selection cuts of the two modes are similar, few sources of systematics are expected to contribute. These are due to:

- Event selection:
 - Trigger selection
 - PID selection
 - Binning scheme
- Fit model:
 - Fixed parameters
 - Shape models
 - S-wave interference
 - Fit bias
- f_s / f_d

7.1 EVENT SELECTION

7.1.1 Trigger requirements

The trigger contributes to the systematic uncertainty due to the discrepancy between simulated samples and data. To estimate this systematic the efficiency calibration curves, produced exploiting the large data sample of $D^{*\pm} \rightarrow D^0 \pi_s^\pm$, where $D^0 \rightarrow K\pi$, recorded by the experiment are used [77]. The efficiency was obtained with the so called *TISTOS* method [78]

$$\epsilon = \frac{N(\text{TIS and TOS})}{N(\text{TIS})} \quad (59)$$

where $N(\text{TIS and TOS})$ is the number of tracks independent of the trigger (TIS) and which pass the L0 hadron trigger condition (TOS), and $N(\text{TIS})$ is the total number of tracks independent of the trigger (see Sec.3.2.4). A candidate is classified as TOS if the 3×3 cell cluster, built around the HCAL cell hit by the projection of the track measured by the tracking system, shares at least one cell in the HCAL, with a 2×2 cluster of a L0Calo hadron candidate with a E_T above the trigger threshold. A candidate is classified as TIS if the event was triggered by the L0 Muon system, in order not to bias the efficiency computation. The plots, averaged for Magnet Up and Magnet Down configurations, as a function of the E_T^{real} of each track, for the inner and the outer part of the HCAL, are reported in Fig. 58. The

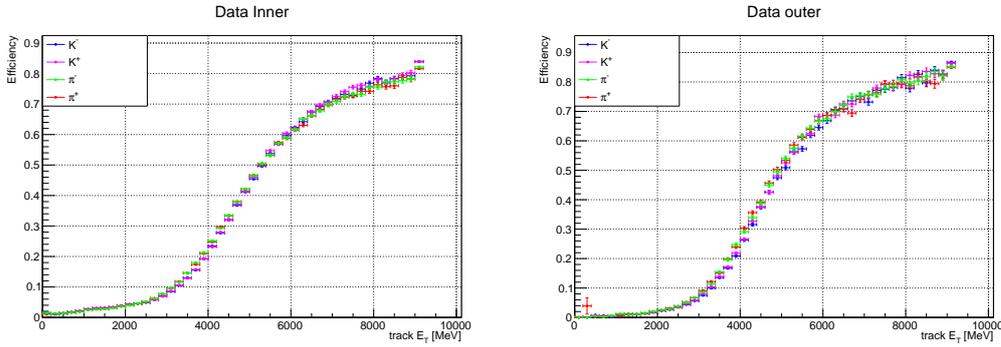


Figure 58.: The trigger efficiency in bins of E_T^{true} for each type of tracks considered in the analysis, extracted from data.

trigger efficiencies show a general good agreement for the different types of tracks and no significant difference is observed comparing positive and negative tracks.

These data-driven calibration curves can be used to reweight MC events allowing to extract the hadron trigger efficiency for each track. The efficiency for each channel is then computed from the individual track efficiencies, assuming that they are uncorrelated. The results are reported in the Tab. 49 and compared to the full Monte Carlo prediction.

The relative difference of these estimated efficiencies and those predicted by the MC for two different modes is assigned as a systematic uncertainty related to the trigger efficiency. The values are shown in Tab. 50.

7.1.2 Binning scheme for the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ efficiencies

The systematic uncertainty associated to the choice of the binning in the Dalitz plane was obtained by varying the binning and redoing the extraction of the effi-

Table 49.: L0TOS trigger efficiencies from simulated events and from a weighted average of calibration efficiencies.

Channel	L0TOS (MC)(%)	L0TOS (weighted) (%)
$B^0 \rightarrow K^{*0} K_S^0$ Down-Down	31.77 ± 0.22	32.38 ± 0.17
$B_S^0 \rightarrow K^{*0} K_S^0$ Down-Down	32.12 ± 0.22	34.03 ± 0.18
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Down-Down	30.25 ± 0.22	31.57 ± 0.16
$B^0 \rightarrow K^{*0} K_S^0$ Long-Long	26.83 ± 0.32	27.76 ± 0.13
$B_S^0 \rightarrow K^{*0} K_S^0$ Long-Long	27.86 ± 0.33	30.58 ± 0.16
$B^0 \rightarrow K_S^0 \pi^+ \pi^-$ Long-Long	27.96 ± 0.35	28.68 ± 0.14

Table 50.: Relative uncertainties on the ratio of branching fractions due to the trigger modeling uncertainty in the simulated events.

Relative B.F.	Down-Down K_S^0	Long-Long K_S^0
$B^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$	2.4%	0.9%
$B_S^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$	6.6%	6.5%

ciencies with the binning optimized to have approximately equivalent numbers of weights in each bin. Weighted average and weighted rms of the efficiencies were obtained knowing the corresponding statistical uncertainties. Tab. 51 reports the central value of the efficiency and its uncertainty (first column), the weighted average and the weighted rms of the efficiencies when varying the binning scheme (second column) and the systematic uncertainty finally assigned (last column) computed by adding in quadrature the relative statistical uncertainty and the weighted rms divided by the weighted average.

Table 51.: Systematic uncertainties from binning scheme. The central value of the efficiency with its statistical uncertainty, the weighted average and the weighted rms extracted by varying the binning scheme and the final systematic uncertainty [62].

Channel	Efficiency	Average and rms	Uncertainty
$B^0 \rightarrow K_s^0 \pi^+ \pi^-$ Down-Down	0.0336 ± 0.0010	0.0350 ± 0.0009	3.9%
$B^0 \rightarrow K_s^0 \pi^+ \pi^-$ Long-Long	0.0117 ± 0.0009	0.0116 ± 0.0002	7.9%

7.1.3 PID requirements

The main source of systematic uncertainty associated with the extraction of the particle identification efficiencies comes from the limited size of the reference sample used to re-weight the calibration efficiencies. The method described in Sec. 6.2.5 only accounts for the statistical uncertainty due to the size of the calibration samples which gives the uncertainties on the efficiency histogram bins. However this is generally significantly smaller than the other sources of uncertainty associated with the calibration procedure. The software package used to determine the efficiencies offers the possibility to recalculate the statistical uncertainty on the average PID efficiency propagating the uncertainties due to the reweighting and limited reference sample size. The estimated PID efficiencies for the $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ modes taking into account the discussed uncertainties are reported in Tab. 52.

Table 52.: PID efficiencies for the $B^0 \rightarrow K^{*0}K_S^0$ and $B_s^0 \rightarrow K^{*0}K_S^0$ modes. The errors include the statistical uncertainty due to the size of the reference samples.

Decay mode	MagUp efficiency (%)	MagDown efficiency (%)
$B^0 \rightarrow K^{*0}K_S^0$ Down-Down	75.768 ± 0.628	75.988 ± 0.603
$B_s^0 \rightarrow K^{*0}K_S^0$ Down-Down	75.854 ± 0.593	76.227 ± 0.606
$B^0 \rightarrow K^{*0}K_S^0$ Long-Long	73.175 ± 0.711	75.241 ± 0.704
$B_s^0 \rightarrow K^{*0}K_S^0$ Long-Long	73.975 ± 0.691	74.303 ± 0.684

7.2 FIT MODEL

Possible systematic uncertainties are introduced because of the model used to fit the PDF to the data. We consider three possible sources.

7.2.1 Fixed shape parameters

Some parameters in the fit model are fixed to values determined from simulated events, which are subject to an uncertainty from the statistics of the sample. In order to estimate these uncertainties, the fixed parameter values for the PDFs are varied according to their covariance matrices, which are taken from the fits that determined the values of the parameters. Then the new values were used to fit again the data and produce new values for the yields. The difference between the new values and the nominal one is then fitted with a gaussian function and the absolute mean value plus the uncertainty is considered as the uncertainty. The systematic uncertainties are associated to the following fixed parameters:

- α , n and f_2 of the two Crystal Ball PDF for the signal in the $K_S^0 K \pi$ and $K_S^0 \pi \pi$ spectrum;
- α and n of the two Crystal Ball PDF for the non resonant component in the $K_S^0 K \pi$ spectrum;
- Threshold t of the linear function for the non resonant component in the $K \pi$ spectrum;
- Parameters of the ARGUS function used to describe the partially reconstructed background in the $K_S^0 K \pi$ and $K_S^0 \pi \pi$ spectrum;

- Threshold x of the linear function used to describe the partially reconstructed background in the $K\pi$ spectrum;
- Parameters of the cross-feed in the $K_S^0\pi\pi$ spectrum;

Tables 53-54 show the value of the different components of the systematic uncertainty due to the fixing of the shape parameters. The total systematic error is defined as the sum in quadrature of all the components.

Table 53.: Systematic uncertainties on the $B^0 \rightarrow K^{*0}K_S^0$ yields due to the fixing of shape parameters. Uncertainties are divided into the contributions from the signal PDF parameters, partially reconstructed background parameters, and parameters of the non resonant background. Uncertainties are given in terms of the absolute number of events, which are added in quadrature for the full systematic uncertainty.

Decay	sig.	non reso.	part.reco.	Tot
$N_{B^0 \rightarrow K^{*0}K_S^0}^{DD}$	0.1	0.2	0.3	0.4
$N_{B^0 \rightarrow K^{*0}K_S^0}^{LL}$	0.1	0.1	0.2	0.2
$N_{B_S^0 \rightarrow K^{*0}K_S^0}^{DD}$	0.2	0.1	0.4	0.5
$N_{B_S^0 \rightarrow K^{*0}K_S^0}^{LL}$	0.1	0.1	0.2	0.2

Table 54.: Systematic uncertainties on the $B^0 \rightarrow K_S^0\pi^+\pi^-$ yields due to the fixing of shape parameters. Uncertainties are divided into the contributions from the signal PDF parameters, cross-feed background parameters and partially reconstructed background parameters. Uncertainties are given in terms of the absolute number of events, which are added in quadrature for the full systematic uncertainty.

Decay	sig.	cross-feed	part.reco.	Tot
$N_{B^0 \rightarrow K_S^0\pi^+\pi^-}^{DD}$	1.9	1.7	3.1	4.0
$N_{B^0 \rightarrow K_S^0\pi^+\pi^-}^{LL}$	1.5	3.4	1.7	4.1

7.2.2 Shape model

The systematic uncertainties related to the choice of the signal and background shape are evaluated considering various changes to the nominal model. The

dataset is then fitted using the alternative model and the variation of the new yields with respect to the nominal one are taken as the uncertainties. The alternative shapes are considered for:

- *Signal model*: a Crystal Ball plus a gaussian function instead of a double Crystal Ball function;
- *Non resonant model*: the linear function used to model this component is an approximation valid only in a small range. As alternative shape the so-called "Lass" function [74] is used: this is composed of a linear function plus a relativistic spin-0 Breit Wigner, which should model more correctly the high mass part of the spectrum;
- *Combinatorial model*: instead of the exponential (B) and the linear function (K^{*0}) a second order Chebyshev polynomial and an exponential are used;

The partially reconstructed background and cross-feed shapes suffer from a large statistical uncertainty due to limited MC statistics, so the previous systematic should cover also any variation of shape.

Tables 55 and 56 show the value of the different components of the systematic uncertainty due to the shape model. The total systematic error is defined as the sum in quadrature of all the components.

Table 55.: Systematic uncertainties on the signal yields due to the shape model. Uncertainties are divided into the contributions from the signal, non resonant and combinatorial background models. Given in terms of the absolute number of events, which are added in quadrature for the full systematic uncertainty.

Yield	sig.	non reso.	combo	Tot
$N_{B^0 \rightarrow K^{*0} K_S^0}^{DD}$	0.02	0.6	0.2	0.6
$N_{B^0 \rightarrow K^{*0} K_S^0}^{LL}$	0.2	0.4	0.1	0.5
$N_{B_s^0 \rightarrow K^{*0} K_S^0}^{DD}$	0.6	1.9	0.4	2.0
$N_{B_s^0 \rightarrow K^{*0} K_S^0}^{LL}$	0.0	1.3	0.4	1.4

7.2.3 Fit model: S-wave interference

For the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ decays there is an additional source of systematic uncertainty. In fact the fit model does not account for the interference between the

Table 56.: Systematic uncertainties on the signal yields due to the shape model. Uncertainties are divided into the contributions from the signal and combinatorial background models. Given in terms of the absolute number of events, which are added in quadrature for the full systematic uncertainty.

Yield	sig.	combo	Tot
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}^{\text{DD}}$	10	23	25
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}^{\text{LL}}$	5	31	31

$K^*(892)^0$ resonance and the other intermediate states like the non resonant component or the $K^*(1430)^0$ resonance (referred to as S-wave). Following the same approach used in [10] a useful variable to estimate the associated systematic is the decay angle θ , defined as the angle between the flight direction of the K^+ in the K^* rest frame, with respect to the K^* flight direction in the B rest frame

$$\cos\theta = \frac{(p_B \cdot p_{K^*})(p_B \cdot p_{K^+}) - p_B^2(p_{K^*} \cdot p_{K^+})}{\sqrt{[(p_B \cdot p_{K^*})^2 - p_B^2 p_{K^*}^2][(p_B \cdot p_{K^+})^2 - p_B^2 p_{K^+}^2]}}. \quad (60)$$

Using a simplified toy model, it is possible to assert that the angular distribution can be modeled by a second order polynomial

$$y = p_0 + p_1 x + p_2 x^2 \quad (61)$$

where the constant part describes the non resonant component, the linear part is related to the interference and finally the quadratic part is associated to the signal.

From the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ signal MC shown in Fig. 59, it is possible to observe that after the entire selection the angular distribution is very difficult to parametrise for values of $\cos\theta$ less than zero, where the acceptance effects related to tracks with low p_T are more evident. For values between zero and one instead, it increases as a function of $\cos^2\theta$.

The angular distribution can be extracted from data using the *sPlot* technique and unfolding only the signal component. However due to very low number of signal yields it would be really difficult to fit the distribution. For this reason both the signal and the non resonant components are considered together. To extract the *sWeights*, the mass window for the fit has been chosen in order to reduce the presence of the partially reconstructed backgrounds to a negligible amount ($m_B > 5200 \text{ MeV}/c^2$). All the shape parameters are then fixed to the values obtained from the nominal fit and a one dimensional fit is performed in the $K_S^0 K \pi$ invariant

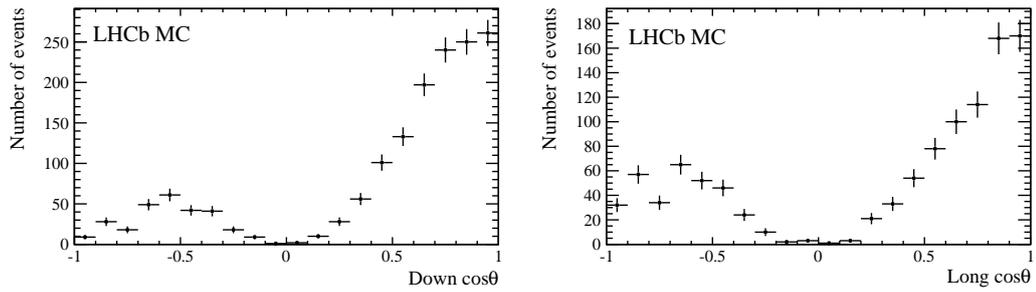


Figure 59.: Distribution of $\cos\theta$ for $B_s^0 \rightarrow K^{*0}K_s^0$ MC events. Left: Down-Down. Right: Long-Long.

mass. The angular distributions obtained by the *sPlot* technique (Fig. 60) are then fitted with two alternative models:

- Signal + S-wave + interference: all the parameters of the polynomial function are free to float;
- Signal + S-wave: the interference term coefficient (p_1) is fixed to zero

For each of the models the integral of the fit function is determined and the relative difference between the values is taken as a systematic uncertainty. The results are given in Tab. 57.

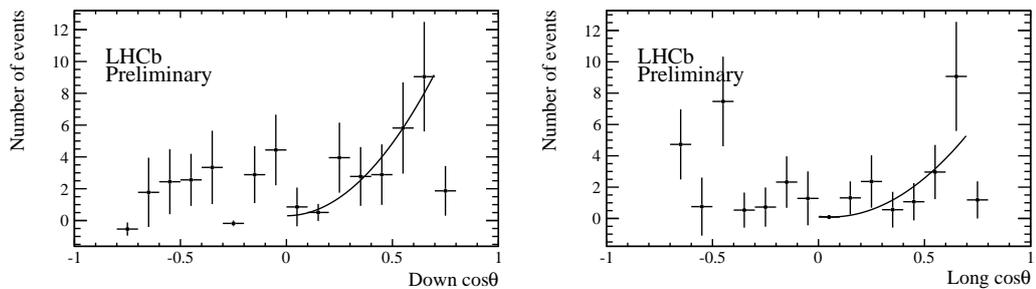


Figure 60.: Distribution of $\cos\theta$ for $B_s^0 \rightarrow K^{*0}K_s^0$ data events. Left: Down-Down. Right: Long-Long.

Due to the very low statistics of the $B^0 \rightarrow K^{*0}K_s^0$ decay, even considering together the signal and the S-wave, it is not possible to perform a fit to the angular distribution. Hence the same systematics as for the $B_s^0 \rightarrow K^{*0}K_s^0$ is taken.

Table 57.: Measured systematic uncertainties on measured yields from the interference between the signal and the non resonant components.

Yield	K_S^0 Down-Down	K_S^0 Long-Long
$N_{B^0 \rightarrow K^{*0} K_S^0}$	2.9%	3.1%
$N_{B_S^0 \rightarrow K^{*0} K_S^0}$	2.9%	3.1%

7.2.4 Fit bias

As discussed in Sec. 5.1 the maximum likelihood estimate is unbiased for a large number of events. This is not the case for the size of the data sample available in this analysis. To determine the bias on the signal yield 1000 toy studies are performed using the yields and PDF parameters from the fit to the data. The residuals of the yields are then fitted with gaussian functions and the final systematic uncertainty on each yield is taken to be the absolute value of the mean plus its uncertainty. The plots on Fig. 61 show the residuals and the gaussian function used to fit the distribution, together with the results on the μ parameter. The yield parameters of the B_S^0 do not show any particular bias. On the other hand the distributions for the yield parameters of the B^0 show some bias, reflecting again the problem of the likelihood fit with low statistics. The value of the bias for each decay is then given in Tab. 58.

Table 58.: Absolute value of the fitted mean plus uncertainty of residuals from fits to 1000 toy data samples.

Decay	K_S^0 Down-Down	K_S^0 Long-Long
$N_{B^0 \rightarrow K^{*0} K_S^0}$	0.4	0.1
$N_{B_S^0 \rightarrow K^{*0} K_S^0}$	0.5	0.4
$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}$	2.0	1.8

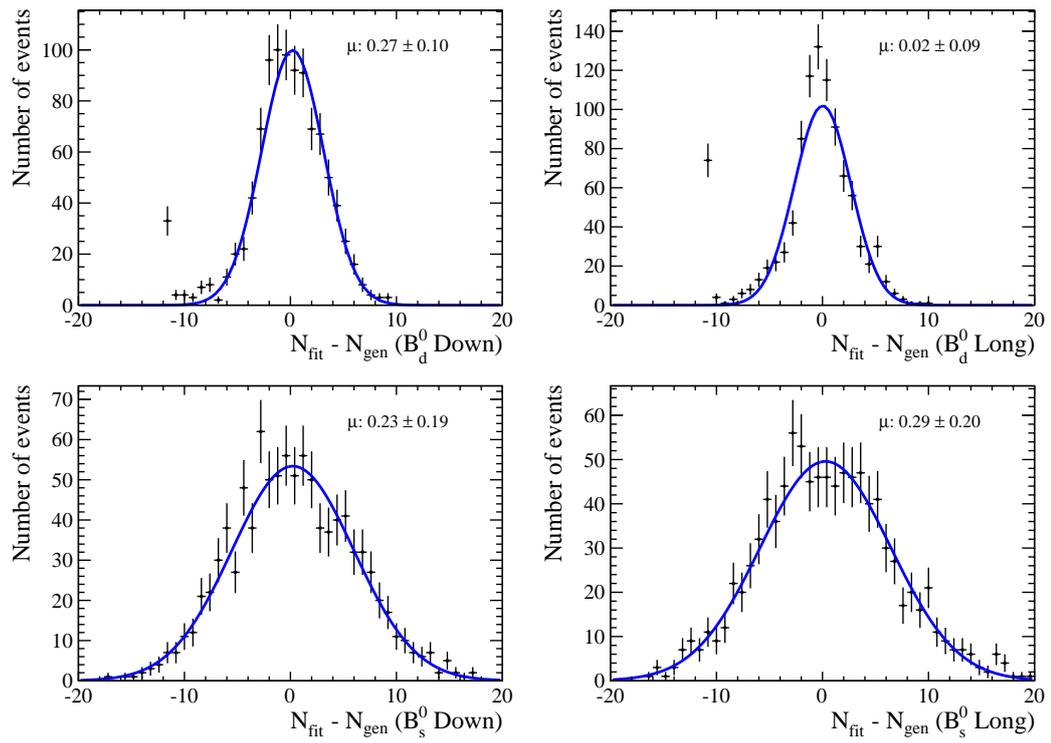


Figure 61.: Residual distributions from fits to toy events. Top: $B^0 \rightarrow K^{*0} K^0$. Bottom: $B_s^0 \rightarrow K^{*0} K^0$. Left: Down-Down. Right: Long-Long

7.3 SUMMARY

In this section the systematic uncertainties are summarised. The systematics contributing to the observation significance are given in Tab. 59-60. These are given in terms of the absolute yield of the relevant decay. The systematics contributing to the branching fraction measurements and limits, are given in Tab. 61-62 in terms of the relative branching fractions. The uncertainties on the efficiency measurements are quoted as additional sources of systematic uncertainty on the branching fractions.

Table 59.: Summary of systematic uncertainties on the Down-Down K_S^0 yields, given in terms of absolute number of events measured.

Contribution	$N_{B^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	$N_{B_s^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}$
Fit par.	0.4	0.5	4.0
Fit model	0.6	2.0	3.0
Fit bias	0.4	0.5	2.0
S-wave interference	0.06	0.6	–
Total	0.8	2.2	5.4

Table 60.: Summary of systematic uncertainties on the Long-Long K_S^0 yields, given in terms of absolute number of events measured.

Contribution	$N_{B^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	$N_{B_s^0 \rightarrow K_S^0 K^\pm \pi^\mp}$	$N_{B^0 \rightarrow K_S^0 \pi^+ \pi^-}$
Fit par.	0.2	0.2	4.0
Fit model	0.5	1.4	5.0
Fit bias	0.1	0.4	1.8
S-wave interference	0.03	0.8	–
Total	0.5	1.7	6.7

Table 61.: Systematic uncertainties on the Down-Down K_S^0 branching fractions, relative to the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ branching fraction. The fit systematic is from the combined systematic contributions detailed in Tab. 59.

Contribution	$B^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$	$B_s^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$
Trigger	0.02	0.07
Selection	0.08	0.08
PID	0.01	0.01
Fit	0.36	0.11
Total	0.37	0.15
f_s/f_d	-	0.06

Table 62.: Systematic uncertainties on the Long-Long K_S^0 branching fractions, relative to the $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ branching fraction. The fit systematic is from the combined systematic contributions detailed in Tab. 60.

Contribution	$B^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$	$B_s^0 \rightarrow K^{*0} K_S^0 / B^0 \rightarrow K_S^0 \pi^+ \pi^-$
Trigger	0.09	0.07
Selection	0.11	0.10
PID	0.01	0.01
Fit	0.30	0.07
Total	0.33	0.14
f_s/f_d	-	0.06

8 | RESULTS

In this chapter the final results about the significance of the signal and the measurement of the branching ratio or the upper limit are given.

8.1 OBSERVATION SIGNIFICANCE

To determine the signal yields an unbinned extended maximum likelihood fit was used. Under this condition a commonly used method to determine the significance of the signal is given by the likelihood ratio test [79]. The method consists in building two likelihoods: the first (\mathcal{L}) is the nominal one, with the default values of the fit, the second (\mathcal{L}_0) has the number of signal yields fixed to zero. Given the ratio of the two likelihoods

$$\lambda = \frac{\mathcal{L}_0}{\mathcal{L}}, \quad (62)$$

Wilks' theorem [79] states that the likelihood and the χ^2 distribution are related by the formula

$$-2 \ln \lambda \approx \chi^2, \quad (63)$$

where $\ln \lambda$ is also known as *Delta Log-Likelihood* $\Delta \ln \mathcal{L}$. The significance of the measurement expressed in standard deviations σ is given by

$$\sigma = \sqrt{\chi^2}. \quad (64)$$

From the technical point of view the following procedure has been implemented:

- Perform many fits, fixing the signal yield to values in a sufficiently wide range, including the nominal result and the null result. The scans are performed separately for the two K_s^0 categories;
- Determine the $\Delta \ln \mathcal{L} = \ln(\mathcal{L}(N_{\text{sig}})/\mathcal{L}(N_{\text{fix}}))$, where $\mathcal{L}(N_{\text{sig}})$ and $\mathcal{L}(N_{\text{fix}})$ are the likelihoods from the nominal fit and from the fit where the signal yield is fixed to a given value;

- Convert the negative log likelihood to the likelihood;
- Convolve the likelihood with a gaussian of width equal to the total systematic uncertainty on the signal yield, that is the systematic uncertainty related only to the fit model;
- Convert again the likelihood into the negative log likelihood;
- Determine the significance as $\sqrt{-2\Delta \ln \mathcal{L}}$;
- The combination for the two K_S^0 categories is performed summing the value of the χ^2 for each K_S^0 sample at the null hypothesis. The corresponding χ^2 with two degrees of freedom is used to give a p-value. Then the equivalent χ^2 value with one degree of freedom and the corresponding significance are found.

The profiles of the negative log likelihood for the $B_s^0 \rightarrow K^{*0}K_S^0$ are shown in Fig. 62. The dashed red line shows the $-\Delta \ln \mathcal{L}$ profile, which takes into account only the statistical error. The solid blue curve shows the $-\Delta \ln \mathcal{L}$ distribution, convolved with a Gaussian of width equal to the total fit systematic uncertainties on the measured yields. The significance of each of the signals are reported in Tab. 63.

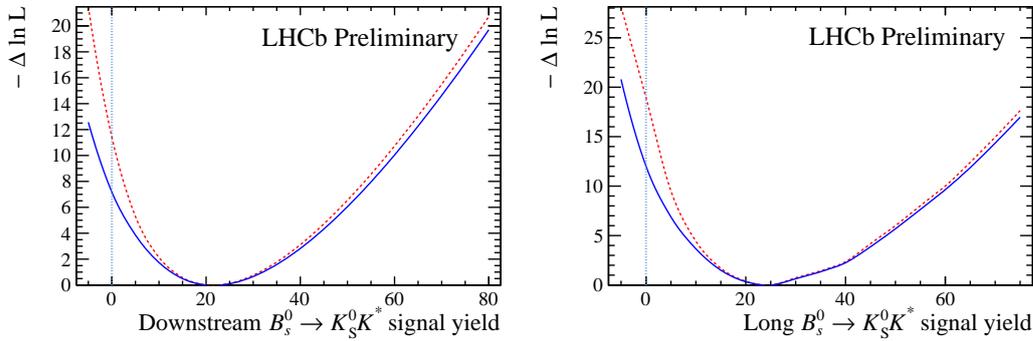


Figure 62.: Difference in the negative log-likelihood as a function of the signal yield for (left) Down-Down and (right) Long-Long $B_s^0 \rightarrow K^{*0}K_S^0$. The red line corresponds to the statistical scan whereas the blue line includes also the fit model systematic uncertainties.

Since the $B^0 \rightarrow K^{*0}K_S^0$ is not significant only a limit is placed on the relative branching fraction (Sec. 8.2). For the $B_s^0 \rightarrow K^{*0}K_S^0$ the combined significance is bigger than 5σ , this corresponds to a first observation of this decay and a branching ratio measurement is reported in Sec. 8.3.

Table 63.: Signal significance (statistical only and including fit systematics) for $B_s^0 \rightarrow K^{*0}K_s^0$ decays.

	$B_s^0 \rightarrow K^{*0}K_s^0$		
	DD	LL	Combined
Stat. significance	4.8	6.2	7.2
Tot. significance	3.8	4.9	6.1

8.2 UPPER LIMIT

Since there is no evidence for the $B^0 \rightarrow K^{*0}K_s^0$ decay, a limit was calculated on the relative branching fraction. A similar method to that used to determine the significance of the signals is adapted in the following way.

- The master formula for the relative branching fraction is used to calculate the yields of $B^0 \rightarrow K^{*0}K_s^0$ giving specific branching ratios values. The yields in the fit were then fixed to these values so that the likelihood scans for the Down-Down and Long-Long categories are presented as a function of the branching ratio;
- Each likelihood is convolved with a gaussian considering all sources of systematic uncertainties that are not correlated between Down-Down and Long-Long;
- The Down-Down and Long-Long $-\Delta \ln \mathcal{L}$ scans are summed;
- The combined likelihood is convolved with a Gaussian considering the correlated systematics. Only the PID systematic is considered to be fully correlated;
- The likelihood is integrated from 0 to infinity;
- The value of the branching ratio is the one such that the integral from 0 to this value corresponds to 90% (or 95%) of the integral from 0 to infinity.

The final likelihood curve, plotted as a function of the relative branching fraction is shown in Fig. 63. Also shown are the 90% and 95% confidence limits for the branching ratio which read

$$\frac{\mathcal{B}(B^0 \rightarrow K^{*0}K_s^0)}{\mathcal{B}(B^0 \rightarrow K_s^0 \pi^+ \pi^-)} < 0.012(0.015) \text{ at } 90\% (95\%) \text{ CL.} \quad (65)$$

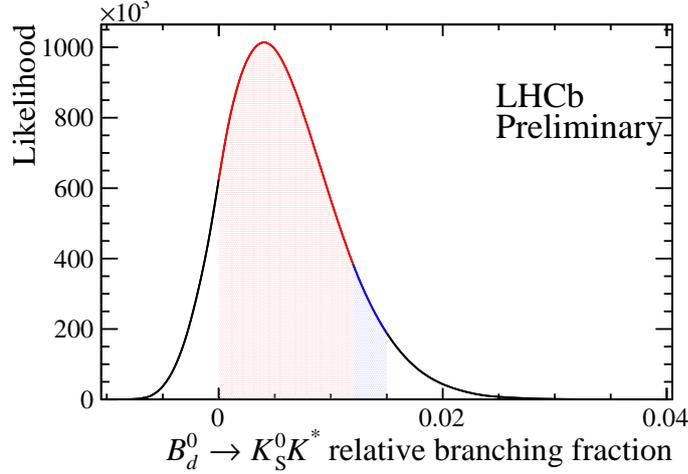


Figure 63.: Integration of the combined Long-Long and Down-Down likelihood curves (including systematic uncertainties) in the positive region to give 90% (red) and 95% (red + blue) confidence level upper limits for $B^0 \rightarrow K^{*0} K_s^0$.

8.3 BRANCHING RATIO MEASUREMENT

As the statistical evidence for the discovery of the $B_s^0 \rightarrow K^{*0} K_s^0$ decay modes is reached, a measurement of the branching fraction relative to the branching fraction of the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ is given. The master formula for the branching ratio is

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_s^0)}{\mathcal{B}(B^0 \rightarrow K_s^0 \pi^+ \pi^-)} = \frac{f_d}{f_s} \cdot \frac{\epsilon_{B_s^0 \rightarrow K^{*0} K_s^0}}{\epsilon_{B^0 \rightarrow K_s^0 \pi^+ \pi^-}} \frac{N_{B_s^0 \rightarrow K^{*0} K_s^0}}{N_{B^0 \rightarrow K_s^0 \pi^+ \pi^-}}, \quad (66)$$

where the efficiencies include all the efficiencies determined in Sec. 6. The value used for the f_s/f_d is that one coming from the combination of the hadronic and semileptonic measurements [61]

$$f_s/f_d = 0.259 \pm 0.015.$$

A summary of the total selection efficiency and yield inputs entering the branching fraction calculations are given in Tab. 64 for each reconstruction category.

The results, given separately for Down-Down and Long-Long categories, read

$$\begin{aligned} \frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_s^0)}{\mathcal{B}(B^0 \rightarrow K_s^0 \pi^+ \pi^-)_{\text{Downstream}}} &= 0.19 \pm 0.06(\text{stat}) \pm 0.03(\text{syst}) \pm 0.01(f_s/f_d), \\ \frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_s^0)}{\mathcal{B}(B^0 \rightarrow K_s^0 \pi^+ \pi^-)_{\text{Long}}} &= 0.28 \pm 0.07(\text{stat}) \pm 0.04(\text{syst}) \pm 0.02(f_s/f_d). \end{aligned}$$

The results are compatible within 1 standard deviation of the statistical and systematic uncertainties in quadrature. Then the obtained relative branching fractions are

Table 64.: Summary of the efficiencies and yields used in the calculation of the branching fractions. Uncertainties are only given for the statistical source of uncertainty from the measured yields.

Input	Down-Down K_s^0	Long-Long K_s^0
$\epsilon_{B_s^0 \rightarrow K^{*0} K_s^0}$	0.000174	0.000121
$\epsilon_{B^0 \rightarrow K_s^0 \pi^+ \pi^-}$	0.000336	0.000117
$N_{B_s^0 \rightarrow K^{*0} K_s^0}$	21 ± 6	25 ± 6
$N_{B^0 \rightarrow K_s^0 \pi^+ \pi^-}$	828 ± 41	341 ± 23

averaged by weighting each K_s^0 reconstruction mode by its total uncertainty (except the hadronisation fractions uncertainty, which is fully correlated for Down-Down and Long-Long).

$$\frac{\mathcal{B}(B_s^0 \rightarrow K^{*0} K_s^0)}{\mathcal{B}(B^0 \rightarrow K_s^0 \pi^+ \pi^-)} = 0.22 \pm 0.06(\text{stat}) \pm 0.02(\text{syst}) \pm 0.01(f_s/f_d).$$

The absolute branching fraction is calculated using the measured branching fraction of the normalisation channel $\mathcal{B}(B^0 \rightarrow K^0 \pi^+ \pi^-) = (4.96 \pm 0.20) \times 10^{-5}$ [17]

$$\mathcal{B}(B_s^0 \rightarrow K^{*0} K^0) = 10.9 \pm 2.9(\text{stat}) \pm 1.0(\text{syst}) \pm 0.5(f_s/f_d) \pm 0.4(BF).$$

9

SUMMARY AND CONCLUSIONS

The $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decays have been studied using pp collision data at a centre of mass energy of 7 TeV, recorded by the LHCb experiment during 2011 and corresponding to an integrated luminosity of 1 fb^{-1} .

The decay $B_S^0 \rightarrow K^{*0}K_S^0$ is observed for the first time with a total significance of 6.1σ , which takes into account the statistical and the systematic uncertainties. The $B^0 \rightarrow K^{*0}K_S^0$ signal instead is not found to be significant.

The absolute branching fraction or the limit measurements in terms of the well measured $B^0 \rightarrow K^0\pi^+\pi^-$ branching fraction is determined for the two modes and can be compared to the theoretical predictions. In Tab. 65 the theoretical values of the branching fractions, calculated using the QCD factorization, the perturbative QCD and a six quark effective Hamiltonian models, and the measured values are reported. The naive combination of the theoretical predictions is obtained summing the two values and considering for each source of uncertainty the sum in quadrature.

Table 65.: Theoretical predictions, measurement and limit of $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decay branching fractions given in units of 10^{-6} . For the $B_S^0 \rightarrow K^{*0}K_S^0$ measured value, the first uncertainty is statistical and the second systematic. The limit on the $B^0 \rightarrow K^{*0}K_S^0$ corresponds to the 95% CL limit.

Branching fraction	QCDF [32]	pQCD [33]	Eff. Hamiltonian [30]	Measured value
$\mathcal{B}(B_d \rightarrow \bar{K}^*K^0)$	$0.70^{+0.18+0.28}_{-0.15-0.25}$	-	-	-
$\mathcal{B}(B_d \rightarrow K^*\bar{K}^0)$	$0.47^{+0.36+0.43}_{-0.17-0.27}$	-	-	-
$\mathcal{B}(B_d \rightarrow K^*K^0)$	$1.17^{+0.40+0.51}_{-0.23-0.37}$	-	-	< 0.74
$\mathcal{B}(B_S^0 \rightarrow \bar{K}^*K^0)$	$10.5^{+3.3+5.1}_{-2.8-4.5}$	$7.3^{+2.5+2.1+0.0}_{-1.7-1.3-0.0}$	$8.5^{+1.8+1.5}_{-2.1-1.6}$	
$\mathcal{B}(B_S^0 \rightarrow K^*\bar{K}^0)$	$10.1^{+7.5+7.7}_{-3.6-4.8}$	$4.3^{+0.7+2.2+0.0}_{-0.7-1.4-0.0}$	$7.1^{+0.2+1.3}_{-0.4-1.1}$	
$\mathcal{B}(B_S^0 \rightarrow K^*K^0)$	$20.6^{+8.2+9.2}_{-4.6-6.6}$	$11.6^{+2.6+3.0}_{-1.8-1.9}$	$15.6^{+1.8+2.0}_{-2.1-1.9}$	$10.9 \pm 2.9 \pm 1.0 \pm 0.5 \pm 0.4$

The branching fraction of the $\mathcal{B}(B_d \rightarrow K^*\bar{K}^0)$ decay is calculated only in the QCD Factorization approach and given the large theoretical uncertainties the value is compatible with the limit found in this thesis. The central value of the measured $\mathcal{B}(B_S^0 \rightarrow K^*K^0)$ is very close to the prediction given by the perturbative QCD model but also for this decay, due to the big uncertainties, it is in agreement with

the other calculations. These results leads to the conclusion that it is not feasible to look for deviations from the theoretical predictions until the calculations will be more accurate. However the measured values serve also as input to the theory and to further constrain possible New Physics sources.

Different questions arise about the future: how can the uncertainties be reduced? Is it possible to use these kind of decays to have more information about CP violating phases? This analysis used the full 2011 LHCb data set, however in 2012 other 2 fb^{-1} of data were recorded at a centre of mass energy of 8 TeV, with a corresponding increase of the bb cross-section. In addition in 2012 the trigger efficiency for decays containing a K_S^0 was improved. Both these conditions can help to reduce the statistical uncertainties.

Moreover CP violation studies are very appealing as crucial test of the SM and a method to determine time-dependent CP asymmetries with non CP eigenstate decays has been proposed in Ref. [80] and [81]. Taking into account the effective flavour tagging efficiency of the order of 5% at LHCb, at least a few hundreds of these decays would be needed to perform this kind of measurement.

Another possibility to study these decays is to exploit a full angular analysis, which allows for the reduction of both statistical and systematic uncertainties. In addition such analysis would increase the sensitivity to relative phases, giving the possibility to search for CP violation beyond the Standard Model. This analysis is currently under way in LHCb, using the entire sample of 2011 and 2012 data.

A

DALITZ PLOT

The Dalitz plot [17] is a useful method to parametrize the phase space of a spin-zero particle (B) decaying to three spin-zero particles (1,2,3). The daughter particles can be paired in three possible ways, and the invariant masses of the pairs are labelled m_{12} , m_{23} and m_{13} . Any two of these are enough to remove the remaining degrees of freedom and to completely determine the decay kinematics in the parent particle rest frame. The scatter plot formed by plotting m_{ij}^2 against m_{ik}^2 is referred to as a Dalitz plot, as shown in Fig.64. From the conservation of energy and momentum all of the points in this plot should lie within a well-defined kinematic limit.

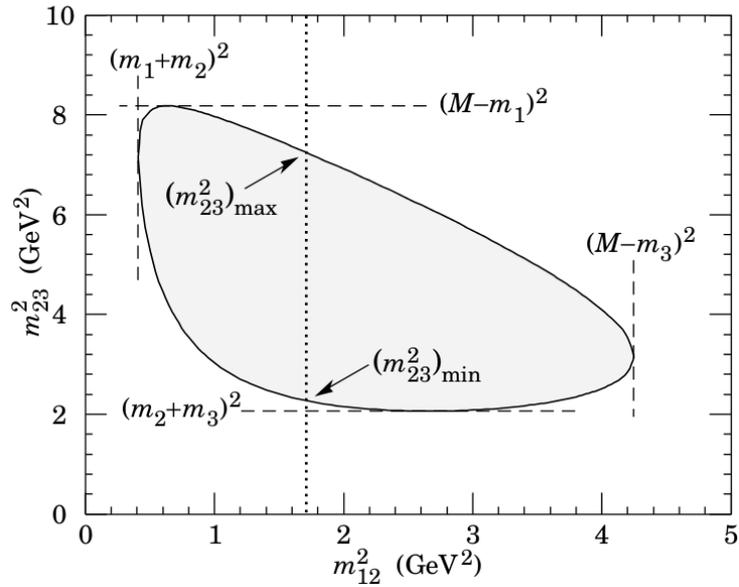


Figure 64.: Dalitz plot for a typical three-body final state decay [17].

The decay rate $d\Gamma$ of a particle with mass M , can be written in terms of the m_{ij}^2 variables and the spin-averaged matrix element for the decay \mathcal{M} as

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{8M} |\mathcal{M}|^2 dm_{ij}^2 dm_{jk}^2 \quad (67)$$

Non-uniformities in a Dalitz plot provide invaluable insights into the nature of \mathcal{M} . They suggest the presence of intermediate two-body resonances, whose typical structure in the Dalitz plot depends on the angular momentum. For charmless B meson decays the signal events populate regions near the edges and corners of the Dalitz plot so that the conventional Dalitz plot is not the best parametrization to describe efficiency variation. However it is possible to apply the concept of *square Dalitz plot*, performing a transformation into new variables with range between 0 and 1:

$$m_{ij}^2 m_{jk}^2 \longrightarrow |\det J| dm' d\theta' \quad (68)$$

with

$$m' \equiv \frac{1}{\pi} \arccos \left(2 \frac{m_{jk} - m_{jk}^{\min}}{m_{jk}^{\max} - m_{jk}^{\min}} - 1 \right) \quad (69)$$

$$\theta' \equiv \frac{1}{\pi} \theta_{jk} \quad (70)$$

where $m_{jk}^{\max} = m_B - m_i$ and $m_{jk}^{\min} = m_j - m_k$ are the boundaries of m_{jk} . θ_{jk} is the angle between i and the bachelor with opposite charge from the B flavour in the jk rest frame. J is the Jacobian of the transformation.

B | MC DATA COMPARISON

Data and Monte Carlo are compared for $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ for events after the BDT selection within $\pm 3\sigma$ around the nominal B^0 mass. Figs. 65 and 66 show the comparison for variables used in the analysis.

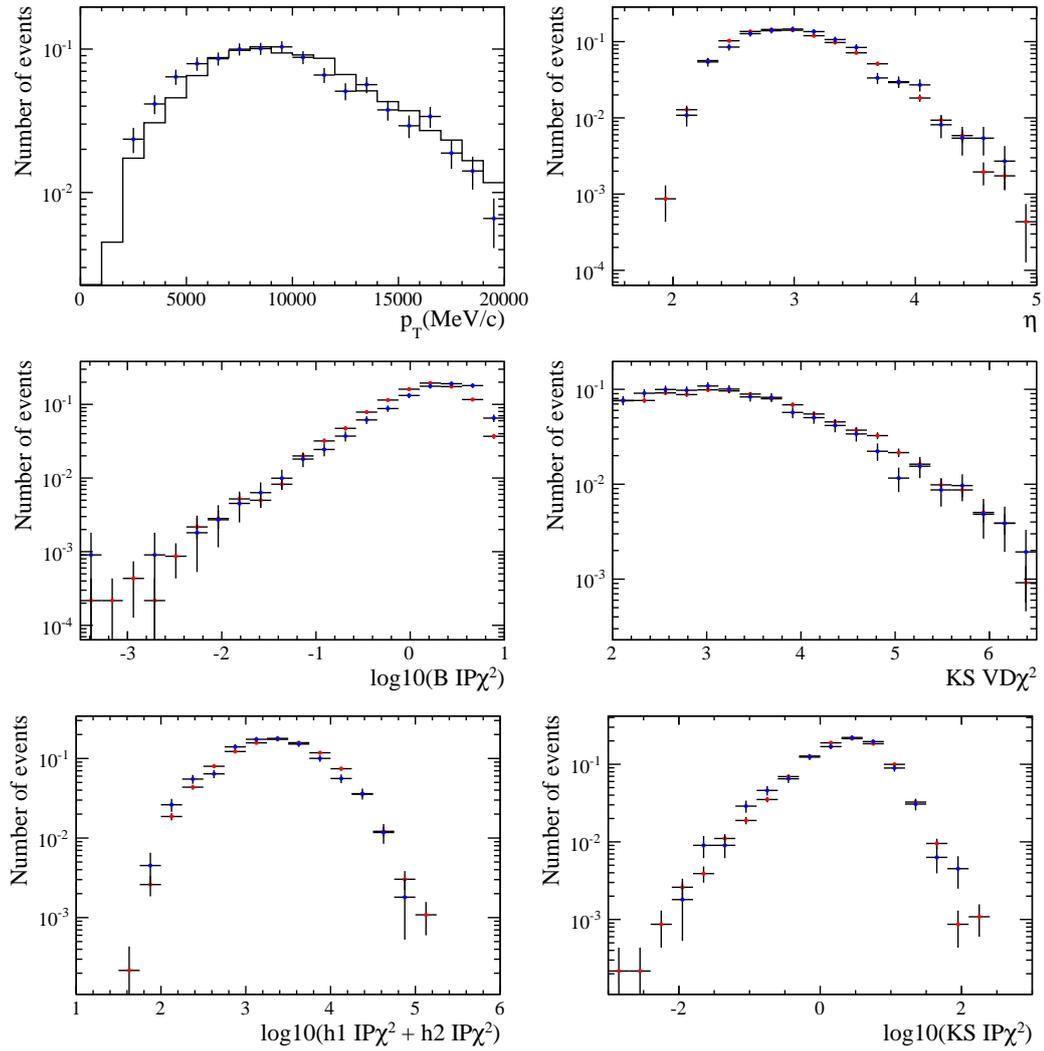


Figure 65.: Comparison of variables used in the analysis for the Down-Down K_S^0 .

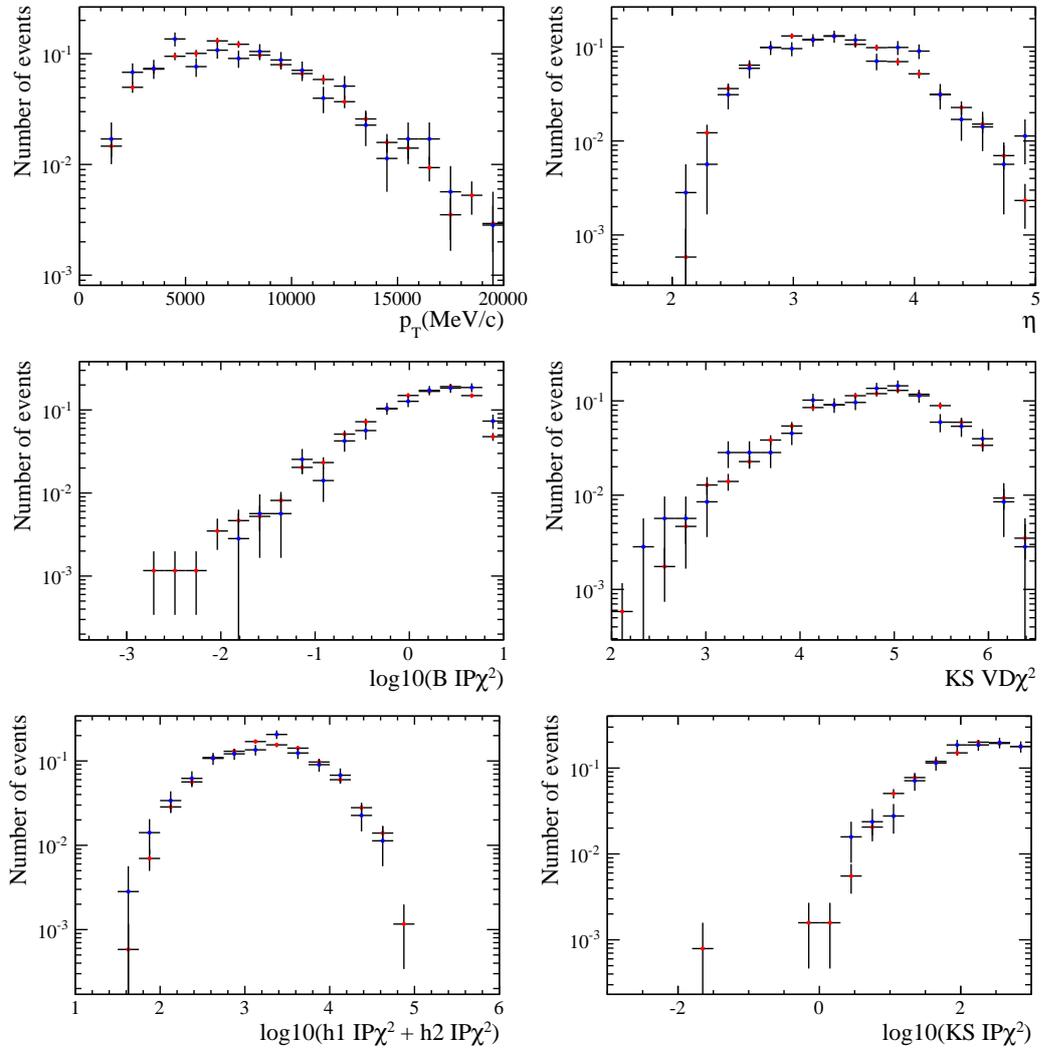


Figure 66.: Comparison of variables used in the analysis for the Long-Long K_S^0 .

BIBLIOGRAPHY

- [1] G. Altarelli, *Status of the Standard Model at the LHC*, *Nuovo Cimento B Serie* **123** (2008) 257–269. [1]
- [2] N. A. Bahcall et al., *The Cosmic Triangle: Revealing the State of the Universe*, *Science* **284** (1999) 1481. [1]
- [3] Super-Kamiokande Collaboration, *Evidence for an oscillatory signature in atmospheric neutrino oscillation*, *Phys. Rev. Lett.* **93** (2004) 101801. [1, 5]
- [4] IceCube Collaboration, *Measurement of atmospheric neutrino oscillations with IceCube*, *Phys. Rev. Lett.* **111** (2013) 081801. [1, 5]
- [5] A. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, *Sov. Phys. Usp.* **34** (1991) 392. [1]
- [6] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *Adv. Study Inst. Ser. B Phys.* **59** (1980) 135. [1]
- [7] L. Silvestrini, *Searching for New Physics in $b \rightarrow s$ Hadronic Penguin Decays*, *Ann. Rev. Nuc. Par. Sc.* **57** (2007) 405-440. [2, 16, 17]
- [8] J. Iliopoulos S. L. Glashow in L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D2* **D2** (1970) 12851292. [2, 7]
- [9] LHCb collaboration, *Study of $B_{(s)}^0 \rightarrow K_S^0 h^+ h'^-$ decays with first observation of $B_s^0 \rightarrow K_S^0 K^\pm \pi^\mp$ and $B_s^0 \rightarrow K_S^0 \pi^+ \pi^-$* , *J. High Energy Phys.* **10** (2013) 143. [2, 19, 61, 72, 114, 115, 150, 151, 155]
- [10] LHCb Collaboration, *Observation of $B_s^0 \rightarrow K^{*\pm} K^\mp$ and evidence for $B_s^0 \rightarrow K^{*-} \pi^+$ decays*, *arXiv:1407.7704*. [2, 124]
- [11] S. Glashow, *Partial-symmetries of weak interactions*, *Nucl. Phys.* **22** (1961) 579. [5]
- [12] J. Ward A. Salam, *Ectromagnetic and weak interactions*, *Phys. Lett.* **13** (1964) 168. [5]
- [13] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264. [5]

- [14] D. Perkins, *Introduction to High Energy Physics*, (Cambridge University Press 4th Edition 2000). [5]
- [15] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1. [6]
- [16] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30. [6]
- [17] J. Beringer *et al.*, *Particle Data Group*, *Phys. Rev.* **D86** (2012) 010001. [6, 8, 10, 11, 19, 53, 67, 135, 139, 149, 153]
- [18] G. Altarelli, *The Standard Model of particle physics*, (Encyclopedia of Mathematical Physics 2005). [6]
- [19] A. Djouadi, *The anatomy of electro-weak symmetry breaking I*, *Physics Reports* **457** (2005) 1. [7, 8]
- [20] M. Kobayashi in T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, *Prog. Theor. Phys.* **49** (1973) 652. [8]
- [21] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945-1947. [9]
- [22] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, *Phys. Rev. Lett.* **55** (1985) 1039. [9]
- [23] UTfit Collaboration, <http://www.utfit.org/UTfit>. [11, 149]
- [24] LHCb Collaboration, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \rightarrow J/\psi K^+ K^-$ and $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ decays*, *Phys. Rev.* **D87** (2013) 112010. [14]
- [25] R. Fleischer, *B Physics and CP Violation*, *Lect. Notes Phys.* **647** (2004) 42-77. [14, 15, 149]
- [26] G. Buchalla *et al.*, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125-1144. [15]
- [27] J. D. Björken, *Flavor Mixing and CP Violation*, *Nucl. Phys. (Proc. Suppl.)* **B11**, 325 (1989). [17]

- [28] M. Dugan in B. Grinstein, *QCD basis for factorization in decays of heavy mesons*, *Phys. Lett.* **B255** (1991) 583. [17]
- [29] M. Beneke et al., *QCD Factorization in $B \rightarrow \pi K, \pi\pi$ Decays and Extraction of Wolfenstein Parameters*, *Nucl. Phys.* **B606** (2001) 245. [17]
- [30] C. Zhuang F. Su, Y.-L. Wu in Y.-B. Yang, *Charmless $B_s \rightarrow PP, PV, VV$ Decays Based on the Six-Quark Effective Hamiltonian with Strong Phase Effects II*, *Eur. Phys. J.* **C72** (2012) 1914. [17, 19, 137]
- [31] *Feynman Diagram Library (FDL)*, <http://www.physik.uzh.ch/~che/FeynDiag/index.php>. [18]
- [32] H.-Y. Cheng in C.-K. Chua, *QCD Factorization for Charmless Hadronic B_s Decays Revisited*, *Phys. Rev.* **D80** (2009) 114026. [19, 137]
- [33] A. Ali et al., *Charmless non-leptonic B_s decays to PP, PV and VV final states in the pQCD approach*, *Phys. Rev.* **D76** (2007) 074018. [19, 137]
- [34] L. Evans in P. Bryant, *LHC machine*, *JINST* **3** (2008) S08001. [21]
- [35] CERN document server, <https://cds.cern.ch/record/40525>. [21, 149]
- [36] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. [22]
- [37] CMS Collaboration, *The CMS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08004. [22]
- [38] ALICE Collaboration, *The ALICE Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08002. [22]
- [39] LHCb Collaboration, *The LHCb Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08005. [22, 25, 26, 28, 32, 33, 34, 35, 36, 37, 39, 41, 42, 149, 150]
- [40] C. Lefevre, *The cern accelerator complex*, *CERN-DI-0606052* (2006). [22, 149]
- [41] LHCb Collaboration, *LHCb Material for presentations*, <http://lhcb-operationsplots.web.cern.ch/lhcb-operationsplots/2011Plots.htm>. [23, 24, 30, 31, 48, 149, 150]
- [42] J. W. Huston J. M. Campbell in W. J. Stirling, *Hard Interactions of Quarks and Gluons: a Primer for LHC Physics*, *Rept. Prog. Phys.* **70** (2007) 89. [24, 149]

- [43] LHCb Collaboration, *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, *Phys. Lett. B* **694** (2010) 209. [24]
- [44] LHCb Collaboration, *LHCb velo (vertex locator): TDR* (CERN, 2000). [27]
- [45] LHCb Collaboration, *LHCb inner tracker: TDR* (CERN, 2000). [29]
- [46] LHCb Collaboration, *LHCb outer tracker: TDR* (CERN, 2000). [31]
- [47] LHCb Collaboration, *LHCb magnet: TDR* (CERN, 2000). [31]
- [48] LHCb Collaboration, *LHCb rich: TDR* (CERN, 2000). [35]
- [49] LHCb Collaboration, *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J. C* **73** (2013) 2431. [38, 150]
- [50] LHCb Collaboration, *LHCb calorimeters: TDR* (CERN, 2000). [39]
- [51] LHCb Collaboration, *LHCb muon system: TDR* (CERN, 2000). [40]
- [52] LHCb Collaboration, *LHCb trigger system: TDR* (CERN, 2000). [41]
- [53] I. Belyaev et al., *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, *J. Phys. Conf. Ser.* **331** (2011) 032047. [44, 48]
- [54] S.Marena T.Sjostrand in P.Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 0, 26. [44, 48]
- [55] The EvtGen particle decay simulation package, *D. Lange*, *Nucl. Instrum. Meth.* **A462** (2001) 152. [44]
- [56] Geant4 Collaboration, *Geant: A simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250. [44]
- [57] LHCb Collaboration, *The LHCb digitization program*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/>. [44]
- [58] LHCb Collaboration, *The LHCb High Level Trigger program*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/>. [44]
- [59] *World wide LHC computing grid*, [http://wlcg-public.web.cern.ch/http://wlcg-public.web.cern.ch/](http://wlcg-public.web.cern.ch/http://wlcg-public.web.cern.ch/http://wlcg-public.web.cern.ch/). [45]
- [60] LHCb Collaboration, *The LHCb reconstruction program*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/>. [45]

- [61] LHCb Collaboration, *Updated average f_s/f_d b -hadron production fraction ratio for 7 TeV pp collisions*, *EPS-HEP 2013* (2013). [47, 134]
- [62] LHCb Collaboration, *First observation of the decay $B_s^0 \rightarrow K^{*0} \bar{K}^{*0}$* , *Phys. Lett. B* **709** (2012) 50–58. [52, 120]
- [63] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*, *PoS A CAT 040* (2007). [57, 58, 67, 150]
- [64] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, *The Annals of Statistics* **29** (2001) 5. [58]
- [65] Y. Freund in R.E. Schapire, *A decision-theoretic generalization of on-line learning and an application to Boosting*, *J. of Computer and System Science* **55** (1997) 119. [58]
- [66] LHCb Collaboration, *The LHCb Trigger and its Performance in 2011*, *JINST* **8** (2013) P04022. [59]
- [67] V. V. Gligorov, *A single track HLT1 trigger*, *LHCb-PUB-2011-003*. [59]
- [68] LHCb Collaboration, V. Gligorov, C. Thomas, and M. Williams, *LHCb-PUB-2011-016*. [59]
- [69] M. Needham, *Clone track identification using the kullback-liebler distance*, Tech. Rep. LHCb-2008-002 (2008). [62]
- [70] G. Punzi, *Sensitivity of Searches for New Signals and its Optimization*, <http://arxiv.org/abs/physics/0308063>. [73]
- [71] R. J. Barlow, *Statistics - A Guide to the Use of Statistical Methods in the Physical Sciences*, (WILEY-VCH Verlag 1989). [77]
- [72] *Minuit Manual*, <http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html>. [77]
- [73] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, (PhD thesis, Institute of Nuclear Physics, Krakow 1986). [81]
- [74] BaBar Collaboration, *Time-dependent and time-integrated angular analysis of $B \rightarrow \phi K_S^0 \pi^0$ and $\phi K^\pm \pi^\mp$* , *Phys. Rev. D* **78** (2008) 092008. [81, 123]
- [75] R. D. Cousins, *Why isn't every physicist a Bayesian?* [82]

- [76] M. Pivk in F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, *Nucl. Instr. and Meth.* **A555**, 356 (2005). [[113](#)]
- [77] A. Martin Sanchez et al., *Performances of the LHCb Lo Calorimeter Trigger?*, *LHCb-PUB-2011-026*. [[117](#)]
- [78] S. Tolk et al., *Data-driven measurement of trigger efficiencies with the TisTos method*, *LHCb-INT-2013-038*. [[117](#)]
- [79] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (1938) 60. [[131](#)]
- [80] M. Pierini M. Ciuchini in L. Silvestrini, *$B_s \rightarrow K^{*0} \bar{K}^{*0}$ decays: the golden channels for new physics searches*, *Phys. Rev. Lett.* **100** (2008) 031802. [[138](#)]
- [81] R. Aleksan et al., *CP violation using non-CP eigenstate decays of neutral B mesons*, *Nucl. Phys. B* **361** (1991) 141. [[138](#)]

LIST OF FIGURES

Figure 1	A sketch of the unitarity triangle [17].	10
Figure 2	Current experimental status of the unitarity triangles relevant for the B^0 (left) and B_s^0 (right) systems [23].	11
Figure 3	Tree-level Feynman diagram governing the transition of a non-leptonic B hadron decay [25].	15
Figure 4	QCD penguin Feynman diagram governing the transition of a non-leptonic B hadron decay [25].	15
Figure 5	EW Feynman diagrams governing the transition of a non-leptonic B hadron decay [25].	15
Figure 6	Feynman diagram of the $B_s^0 \rightarrow K^{*0} K_s^0$ decay.	18
Figure 7	A schematic view of the LHC collider and the position of the four experiments ATLAS, CMS, ALICE, LHCb [35].	21
Figure 8	Schematic view of the complex of CERN accelerators [40].	22
Figure 9	Integrated luminosity delivered and recorded by LHCb [41].	23
Figure 10	Leading order Feynman diagrams of b -hadron production [42].	24
Figure 11	Azimuthal angle distribution of the $b\bar{b}$ quark pairs [41].	24
Figure 12	Layout of the LHCb detector in the y - z section [39].	26
Figure 13	Top view of the VELO silicon sensors [39].	28
Figure 14	Schematic view of R and ϕ sensors [39].	28
Figure 15	Resolutions in the reconstruction of primary vertex coordinates as a function of the number of tracks in the event [41].	30
Figure 16	Impact parameter resolution of tracks in the x and y directions [41].	31
Figure 17	Layout of the four TT layers [39].	32
Figure 18	Layout of one of the IT layers [39].	33
Figure 19	Overall layout of the Outer Tracker [39].	34
Figure 20	Layout of the LHCb dipole magnet [39].	34
Figure 21	Scheme of the different LHCb track types [39].	35
Figure 22	Schematic layout of the RICH1 and RICH2 detectors [39].	36

Figure 23	Cherenkov angle θ_C as a function of particle momentum p [39].	37
Figure 24	Efficiency for the identification and misidentification of kaons and pions as a function of particle momentum [49].	38
Figure 25	Lateral segmentation of the SPD/PS and ECAL (left) and HCAL (right) [39].	39
Figure 26	Side view of the five stations of the muon detector [39].	41
Figure 27	LHCb trigger scheme [39].	42
Figure 28	LHCb integrated luminosities for 2011 [41].	48
Figure 29	Left: data mass of $K_s^0 \pi^- p$ combinations with Down-Down K_s^0 candidates. Right: $p K_s^0$ combinations with Down-Down K_s^0 candidates for $K_s^0 \pi^+ p$ combinations consistent with Λ_b mass.	51
Figure 30	Diagram for the definition of the transverse momentum.	55
Figure 31	Diagram for the definition of the impact parameter.	56
Figure 32	Diagram for the definition of the flight distance.	56
Figure 33	Diagram for the definition of the DIRA.	56
Figure 34	Illustration of multivariate selections.	57
Figure 35	Schematic view of a decision tree [63].	58
Figure 36	Scheme explaining the procedure adopted to train and test the BDTs.	68
Figure 37	Distribution of the variables used to build the BDT for the Long-Long K_s^0	70
Figure 38	Distribution of the variables used to build the BDT for the Down-Down K_s^0	71
Figure 39	Output from the training and test sample [9].	72
Figure 40	$B \rightarrow K_s^0 K^\pm \pi^\mp$ collision data, upper sideband background modeled by an exponential PDF.	74
Figure 41	Punzi figure of merit and signal efficiency for different BDT requirements.	75
Figure 42	Correlation between the fit variables for the $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decays.	80
Figure 43	$B^0 \rightarrow K^* K$ signal fit model fitted to simulated events.	83
Figure 44	$B_s^0 \rightarrow K^* K$ signal fit model fitted to simulated events.	84
Figure 45	$B^0 \rightarrow K_s^0 K^\pm \pi^\mp$ non resonant fit model fitted to simulated events.	86
Figure 46	$B_s^0 \rightarrow K_s^0 K^\pm \pi^\mp$ non resonant fit model fitted to simulated events.	87

Figure 47	$B^+ \rightarrow D^0(K_s^0\pi^-\pi^+)K^+$ background fit model fitted to simulated events. 89
Figure 48	$B_s^0 \rightarrow K^*(892)^0(K^\pm\pi^\mp)K^*(892)^0(K_s^0\pi^0)$ background fit model fitted to simulated events 90
Figure 49	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ combinatorial background fit model for the Down-Down category. 91
Figure 50	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ blind fit results. 94
Figure 51	$B_{d,s}^0 \rightarrow K_s^0\pi^+\pi^-$ signal fit model fitted to simulated events. 97
Figure 52	$B^0 \rightarrow K_s^0\pi^+\pi^-$ misidentified fit model fitted to simulated events. 98
Figure 53	$B^0 \rightarrow K_s^0\pi^+\pi^-$ partially reconstructed backgrounds fitted to simulated events. 100
Figure 54	$B_{d,s}^0 \rightarrow K^{*0}K_s^0$ fit results. 104
Figure 55	$B_{d,s}^0 \rightarrow K_s^0\pi^+\pi^-$ fit results. 106
Figure 56	$B_{d,s}^0 \rightarrow K_s^0\pi^+\pi^-$ spectra: results of the <i>sWeight</i> extraction 114
Figure 57	Efficiencies for the $B^0 \rightarrow K_s^0\pi^+\pi^-$ decay [9]. 115
Figure 58	The trigger efficiency in bins of E_T^{true} extracted from data. 118
Figure 59	Distribution of $\cos\theta$ for $B_s^0 \rightarrow K^{*0}K_s^0$ MC events 125
Figure 60	Distribution of $\cos\theta$ for $B_s^0 \rightarrow K^{*0}K_s^0$ data events 125
Figure 61	Residual distributions from fits to toy events. 127
Figure 62	Difference in the negative log-likelihood as a function of the signal yield. 132
Figure 63	Integration of the combined Long-Long and Down-Down likelihood curves (including systematic uncertainties) in the positive region $B^0 \rightarrow K^{*0}K_s^0$. 134
Figure 64	Dalitz plot for a typical three-body final state decay. 139
Figure 65	Comparison of variables used in the analysis for the Down-Down K_s^0 . 141
Figure 66	Comparison of variables used in the analysis for the Long-Long K_s^0 . 142

LIST OF TABLES

Table 1	Properties of the fundamental fermions in the Standard Model [17].	6
Table 2	Gauge bosons and Higgs particle in the Standard Model [17].	6
Table 3	Theoretical predictions of $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decays branching fractions.	19
Table 4	Experimental measurements of charmless B decay branching fractions	19
Table 5	Number of generated signal and background Monte Carlo events for each decay mode for each magnet polarity.	49
Table 6	Charmonium and charm background contributions.	52
Table 7	Partially reconstructed background for the $B_{d,s}^0 \rightarrow K^{*0}K_S^0$ decays.	53
Table 8	Partially reconstructed background for the $B^0 \rightarrow K_S^0\pi^+\pi^-$ decay.	53
Table 9	Trigger selection requirements after reconstruction of decays.	59
Table 10	Hlt1 trigger selection requirements.	60
Table 11	Hlt2 trigger selection requirements on the candidates.	60
Table 12	Hlt2 trigger selection requirements on the K_S^0 candidates.	61
Table 13	Requirements on the charged pion and kaon tracks.	62
Table 14	Requirements for the Long-Long and Down-Down K_S^0 candidates.	63
Table 15	Requirements on the B candidates from four-momentum combination.	65
Table 16	Preselection cuts applied to the stripped events.	65
Table 17	PID cuts applied to the particles reconstructed as pions or kaons.	66
Table 18	Charm vetoes.	66
Table 19	Fitting ranges of the m_B and $m_{K_S^0\pi}$ variables.	67
Table 20	Number of events used to train the BDT.	69

Table 21	List of variables used to build the BDT discriminant for the two K_S^0 categories. 69
Table 22	BDT variable importance ranking (0,1). Variables are presented in the hierarchical order as obtained for Down-Down and Long-Long K_S^0 . 73
Table 23	Definition of the signal and background mass intervals used to optimize the final offline selection cuts. 74
Table 24	BDT optimal cut and efficiencies for signal and background. 75
Table 25	Correlation coefficients between the $m_{K_S^0 K \pi}$ and $m_{K \pi}$ fit variables for simulated signal events. 80
Table 26	Fitted parameter values for the fit to $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ simulated samples. 85
Table 27	Fitted parameter values for the fit to $B_{d,s}^0 \rightarrow K_S^0 K^\pm \pi^\mp$ simulated samples. 86
Table 28	The fit parameters of the partially reconstructed decays. 89
Table 29	Summary of PDF shapes used for the $B_{d,s}^0 \rightarrow K^{*0} K_S^0$ fit model. 91
Table 30	Free fit model parameters for the $K_S^0 K \pi$ final state. 92
Table 31	Signal yields obtained from fits to toy samples with gaussian constraints applied. 95
Table 32	Fitted parameter values after the fit to $K_S^0 \pi \pi$ simulated events. 96
Table 33	Misidentified signal fit model parameters extracted from $B_{d,s}^0 \rightarrow K_S^0 K^\pm \pi^\mp$ simulated samples reconstructed as $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$. 99
Table 34	The fit parameters of the partially reconstructed decays. 99
Table 35	Summary of PDF shapes used for the $B_{d,s}^0 \rightarrow K_S^0 \pi^+ \pi^-$ fit model. 101
Table 36	Free fit model parameters for the $K_S^0 K \pi$ final state. 102
Table 37	Fitted parameter values after the fit to data for the $K_S^0 K \pi$ final state. 103
Table 38	Fitted parameter values after the fit to data for the $K_S^0 \pi \pi$ final state. 105
Table 39	Generator level efficiency in the Magnet Up configuration. 108
Table 40	Generator level efficiency in the Magnet Down configuration. 108
Table 41	Stripping and reconstruction efficiencies from Monte Carlo simulations. 109

Table 42	Trigger, stripping and reconstruction efficiencies and relative efficiency for each step from Monte Carlo simulations for the MagDown polarity. 110
Table 43	Trigger, stripping and reconstruction efficiencies and relative efficiency for each step from Monte Carlo simulations for the MagUp polarity. 110
Table 44	Absolute reconstruction and selection efficiencies. 111
Table 45	PID efficiencies for the $B^0 \rightarrow K^{*0}K_S^0$ and $B_S^0 \rightarrow K^{*0}K_S^0$ modes. 112
Table 46	Summary of the efficiencies for the $B^0 \rightarrow K^{*0}K_S^0$ and $B_S^0 \rightarrow K^{*0}K_S^0$ modes. 112
Table 47	Total efficiencies for Down-Down and Long-Long $B^0 \rightarrow K_S^0\pi^+\pi^-$ [9]. 114
Table 48	Summary of the efficiencies for the $B^0 \rightarrow K_S^0\pi^+\pi^-$ mode without reweighting the Dalitz plane. 114
Table 49	LOTOS trigger efficiencies from simulated events and from a weighted average of calibration efficiencies. 119
Table 50	Relative uncertainties on the ratio of branching fractions due to the trigger modeling uncertainty in the simulated events. 119
Table 51	Systematic uncertainties from binning scheme. 120
Table 52	PID efficiencies for the $B^0 \rightarrow K^{*0}K_S^0$ and $B_S^0 \rightarrow K^{*0}K_S^0$ modes. 121
Table 53	Systematic uncertainties on the $B^0 \rightarrow K^{*0}K_S^0$ yields due to the fixing of shape parameters. 122
Table 54	Systematic uncertainties on the $B^0 \rightarrow K_S^0\pi^+\pi^-$ yields due to the fixing of shape parameters. 122
Table 55	Systematic uncertainties on the signal yields due to shape model. 123
Table 56	Systematic uncertainties on the signal yields due to shape model. 124
Table 57	Measured systematic uncertainties on measured yields from the interference between the signal and the non resonant components. 126
Table 58	Absolute value of the fitted mean plus uncertainty of residuals from fits to 1000 toy data samples. 126
Table 59	Summary of systematic uncertainties on the Down-Down K_S^0 yields. 128
Table 60	Summary of systematic uncertainties on the Long-Long K_S^0 yields. 128

Table 61	Systematic uncertainties on the Down-Down K_s^0 branching fractions, relative to the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ branching fraction. 129
Table 62	Systematic uncertainties on the Long-Long K_s^0 branching fractions, relative to the $B^0 \rightarrow K_s^0 \pi^+ \pi^-$ branching fraction. 129
Table 63	Signal significance (statistical only and including fit systematics) for $B_s^0 \rightarrow K^{*0} K_s^0$ decays. 133
Table 64	Summary of the efficiencies and yields used in the calculation of the branching fractions. 135
Table 65	Theoretical predictions, measurement and limit of $B_{d,s}^0 \rightarrow K^{*0} K_s^0$ decay branching fractions. 137

COLOPHON

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". classicthesis is available for both LaTeX and mLyX:

<http://code.google.com/p/classicthesis/>

Final Version as of October 30, 2014 (classicthesis).