

---

---

VRIJE UNIVERSITEIT

Controlling Penguins

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan  
de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. V. Subramaniam,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de Faculteit der Exacte Wetenschappen  
op dinsdag 3 oktober 2017 om 11.45 uur  
in het auditorium van de universiteit,  
De Boelelaan 1105

door

Vasileios Syropoulos

geboren te Thessaloniki, Griekenland

---

---

promotor: prof.dr. H.G. Raven

copromotor: dr. P. Koppenburg

---

---

Cover designed by Antwnis Karanautis

Printed by Gildeprint - The Netherlands  
ISBN: 978-94-6233-686-5



This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). It was carried out at the National Institute for Subatomic Physics (Nikhef) in Amsterdam.

---

---



---

---

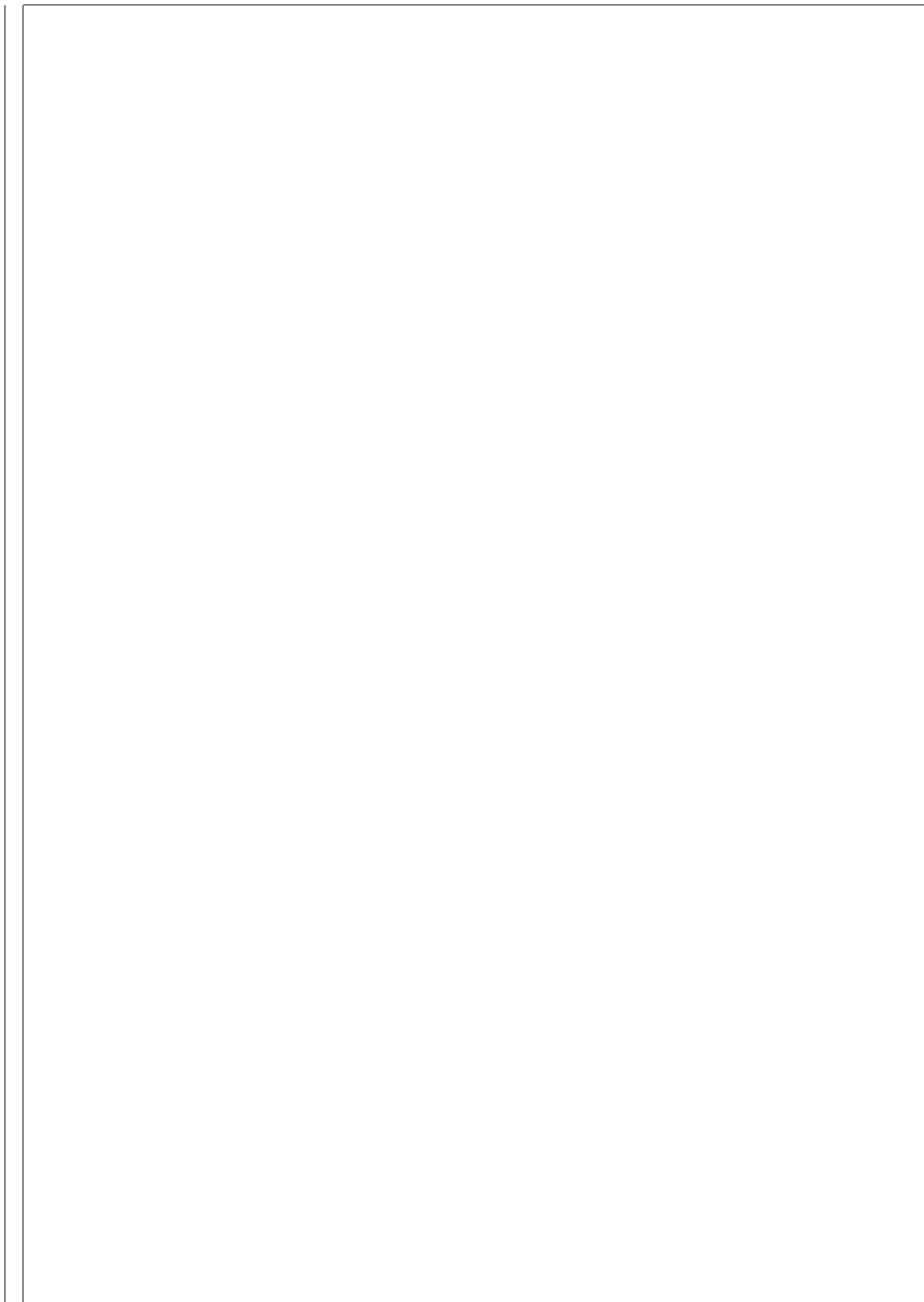
---

*To my parents*

---

---

---



---

---

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Standard Model and the Weak Interaction . . . . .	2
1.2	Flavor Physics . . . . .	6
1.3	The $B_s^0$ Meson and $CP$ violation . . . . .	12
1.4	The Weak Phase $\phi_s$ . . . . .	15
1.4.1	Measuring $\phi_s$ . . . . .	18
1.4.2	Probing New Physics . . . . .	20
1.4.3	Higher Order Effects in $\phi_s$ . . . . .	21
<b>2</b>	<b>The LHCb Detector</b>	<b>23</b>
2.1	Tracking Systems . . . . .	25
2.2	Particle Identification . . . . .	28
2.3	Calorimetry . . . . .	29
2.4	The trigger system . . . . .	30
2.5	$B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ Decay in LHCb . . . . .	33
<b>3</b>	<b>High Level Trigger Muon Identification in Run 2</b>	<b>37</b>
3.1	HLT1 Muon-ID in the LHC Run 2 . . . . .	37
3.1.1	HLT1 muon algorithm . . . . .	38
3.2	Matching VELO tracks to Muon Hits . . . . .	41
3.2.1	The MatchVeloMuon algorithm . . . . .	41
3.2.2	Upgrade to MatchVeloTTMuon . . . . .	43
<b>4</b>	<b>Data Analysis</b>	<b>51</b>
4.1	Candidate Selection . . . . .	51
4.1.1	Multivariate Based Selection . . . . .	52
4.1.2	Reflection Backgrounds . . . . .	55
4.1.3	sWeighting and Invariant Mass Distribution . . . . .	58
4.2	Angular Analysis . . . . .	62

---

4.2.1	Angular Dependence . . . . .	63
4.2.2	Acceptance . . . . .	65
4.2.3	Acceptance Corrections . . . . .	76
4.2.4	$K\pi$ Invariant Mass . . . . .	79
4.2.5	Production and Detection Asymmetries . . . . .	82
4.2.6	Likelihood fit and Total Decay Rate . . . . .	84
4.3	Normalization of $B_s^0 \rightarrow J/\psi K^- \pi^+$ . . . . .	90
4.4	Results . . . . .	93
4.4.1	Parameters of Interest . . . . .	94
4.4.2	Systematic Uncertainties . . . . .	94
4.4.3	Likelihood Scans . . . . .	100
4.4.4	Pseudo-experiments Study . . . . .	105
<b>5</b>	<b>Controlling Penguins in <math>\phi_s</math></b>	<b>109</b>
5.1	Amplitude Structure of $B_s^0 \rightarrow J/\psi \phi$ . . . . .	109
5.2	Hadronic Factors and SU(3) Symmetry . . . . .	111
5.3	Formalism . . . . .	112
5.4	Estimating Penguin Parameters . . . . .	115
5.4.1	The $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ Channel . . . . .	116
5.4.2	The $B^0 \rightarrow J/\psi \rho^0$ Channel . . . . .	117
5.4.3	Fitting Strategy . . . . .	118
5.5	Results . . . . .	120
5.5.1	Further Crosschecks . . . . .	122
<b>A</b>	<b>Additional Tables</b>	<b>127</b>
<b>B</b>	<b>Horizontal Weighting</b>	<b>135</b>
<b>References</b>		<b>141</b>
<b>Summary - Samenvatting</b>		<b>147</b>
<b>Acknowledgments</b>		<b>163</b>

---

# Chapter 1

## Introduction

The main focus of the current thesis is to provide the required ingredients for a high precision measurement of the weak phase  $\phi_s$ . The latter is an important parameter related to  $CP$  violation, which, as explained throughout the current chapter, is related to the matter-antimatter asymmetry in the universe. Measuring  $\phi_s$  enables one to look for deviations from the established theory of elementary particles. Given the state of the art measurement of  $\phi_s$  [1], it is clear that in order to probe potential deviations,  $\phi_s$  needs to be measured with increased precision. However, entering this high precision regime one finds out that there are sub-leading contributions that need to be controlled first. Unless this is achieved, a high precision measurement of  $\phi_s$  can not provide insight on possible deviations. While measuring  $\phi_s$  is based on the analysis of  $B_s^0 \rightarrow J/\psi \phi$  decays, controlling these sub-leading contributions requires a different decay channel, like  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ , which plays the role of a control channel. The reason for this becomes apparent in Section 5.1. Essentially, estimating these contributions is more precise when using  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  instead of  $B_s^0 \rightarrow J/\psi \phi$  decays, as their effect is enhanced in the former.

The current chapter places  $\phi_s$  into perspective with respect to the established theory of particle physics. Chapter 2 introduces the LHCb detector. The analysis of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decays that are required to control the above mentioned higher order effects is presented in Chapter 4. The interpretation and impact of the analysis can be found in Chapter 5. Lastly, Chapter 3 is dedicated to a particular part of the LHCb data acquisition system, which is responsible for selecting a certain type of elementary particles called *muons*, see Table 1.1. Note that

the latter particles are a necessary ingredient for the measurements of  $\phi_s$  and the analysis of  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$  decays.

Finally, the work described in Chapter 4 resulted in the analysis of the  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$  decays [2], and contributed to the  $\phi_s$  measurement [1]. The work performed in Section 3.2 is an essential part of the muon identification software update [3] of the LHCb detector in preparation for the Run 2 data taking period.

## 1.1 Standard Model and the Weak Interaction

After a century of research at the subatomic level, a large number of phenomena have been accounted for, based on the existence of only a handful of elementary particles. Given the vast dimensions that the universe spans, this is an impressive achievement. The known collection of particles can be classified in distinct categories, namely *gauge bosons*, responsible for mediating all the known fundamental interactions of nature (with the exception of gravity) and *fermions* which are the constituents of matter and antimatter. Antimatter fermions carry quantum numbers, such as electric charge, of opposite sign compared to matter fermions. Fermions can be divided further into *quarks* and *leptons*. The latter do not interact through the *strong* interaction, which is one of the above-mentioned fundamental interactions.

	1st generation	2nd generation	3rd generation
quarks	up (u)	charm (c)	top (t)
	down (d)	strange (s)	beauty (b)
leptons	e - neutrino ( $\nu_e$ )	$\mu$ - neutrino ( $\nu_\mu$ )	$\tau$ - neutrino ( $\nu_\tau$ )
	electron ( $e^-$ )	muon ( $\mu^-$ )	tau ( $\tau^-$ )

TABLE 1.1: Standard Model fermions.

The recently discovered Higgs boson [4, 5] plays a special role in explaining how particles acquire mass, and does not function as a gauge boson since it does not mediate any interaction. Fermions, listed in Table 1.1, are the fundamental building blocks of matter and antimatter.

The mathematical framework necessary to describe the interactions between the fermions is called the *Standard Model* of Particle Physics [6–8] which is a quantum field theory. In this framework particles are treated as excited states of their underlying field.

The exact details of the elementary particle interactions are incorporated in the *Lagrangian* of the Standard Model:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Kinetic}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (1.1)$$

Note the three main terms. The first, or *Kinetic*, term describes the possible interactions between the fermions. The second, or *Higgs*, term is the one that contains the Higgs potential, responsible for generating masses for the gauge and Higgs bosons. The last, or *Yukawa*, term is introduced to couple the Higgs field with the fermions and generate masses for them. It is interesting to point out that the Kinetic and Higgs terms originate in a natural way. Specifically, they are the result of exploiting the symmetries that Eq. 1.1 exhibits. By construction, the Standard Model Lagrangian obeys the symmetry group  $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ . This means that there are three distinct types of local transformations that leave Eq. 1.1 invariant. Each type of transformation is associated respectively the strong, *weak* and *electromagnetic* interactions between the fermion following the concept of *local gauge invariance* [9, 10].

As previously mentioned, gauge bosons mediate fundamental interactions. These are the electromagnetic plus the weak nuclear interactions, which are unified under a common description, and the strong nuclear interaction. The first one is mediated by the photon,  $\gamma$ , which interacts with, or more appropriately, *couples to*, any particle that carries an electric charge quantum number. The strong interaction is mediated by gluons,  $g$ , and it couples to a different quantum number, the *color*. Only quarks have a non-zero color quantum number which is typically labeled as either red, green, or blue. Lastly, the weak interaction is mediated by three gauge bosons,  $W^\pm$  and  $Z^0$ , which couple to all fermions via the *weak isospin* quantum number, explained in the next paragraph.

Standard Model particles come in three *generations*, illustrated in Table 1.1. The weak interaction dominantly couples to fermions in the same generation, effectively changing an up quark (u) to a down (d) for

example. Although quarks and leptons are both fermions, a quark can never become a lepton via the weak interaction or *vice versa* within the Standard Model. The above paradigm of up and down quarks extends to the rest of the generations as well, resulting in the up-type (u,c,t) and the down-type (d,s,b) quarks. Similarly for the up-type (e,  $\mu$ ,  $\tau$ ) and the down-type ( $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ ) leptons. The above-mentioned up-down type transitions are commonly called *flavor transitions* and are based on the *weak isospin*. The latter is a quantum number that each fermion has. The value of this quantum number depends on whether it is of up or down type. The charged weak bosons  $W^\pm$  have the ability to change the flavor of quarks whereas the neutral  $Z^0$  does not, see Figure 1.1. An important subtlety is related to the right handed anti-neutrinos, for which there is no up-down type structure, *i.e.* they form a *singlet representation*. Handedness or *chirality* is the quantum number associated to the spin-momentum dot product. Chirality can be +1, right-handed, or -1, left-handed.

It should be noted that the weak interaction violates both the parity,  $P$ , and the charge,  $C$ , symmetries. The first one is associated with left-right symmetry in three dimensional space. The second is associated with inverting the signs of all quantum numbers of a particle. In both cases Eq. 1.1 is not invariant under these parity and charge transformations. What is interesting experimentally is the observation [11, 12] that the weak interaction interacts only with left handed neutrinos or right-handed anti-neutrinos. Given that neutrinos interact only with the weak interaction, and assuming zero mass for the neutrinos (which is a good approximation in view of the measured neutrino masses [13–16]), it follows that there is no need for right-handed neutrinos (and left-handed anti-neutrinos), according to the Standard Model. The implication of the above is that  $P$  and  $C$  are maximally violated, indicating the peculiar structure of the weak interaction. It is reminded that the electromagnetic and strong interactions respect these symmetries, although the  $CP$  conservation of the strong interaction is still to be fully understood due to the unsolved *strong CP problem* [17]. In view of these observations the combined  $CP$  transformation appears to be a symmetry of nature which seems to imply a matter-antimatter symmetry in the universe [18]. However the latter is not consistent with observations

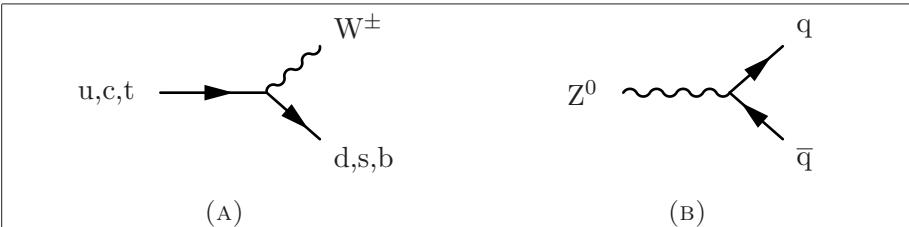


FIGURE 1.1: Weak interactions of quarks in the Standard Model. Left: A charged weak boson changes the flavor of an up-type quark to a down-type within the same generation. Right: A neutral weak boson decays into a quark anti-quark pair. Time flows from left to right.

and nature does distinguish between matter and antimatter through the weak interaction. The origin of  $CP$  violation in the Standard Model is addressed in Section 1.2.

Within the accuracy of the current accelerator experiments the Standard Model has seen its predictions confirmed. The recent discovery of the Higgs boson, which is an essential part of the mechanism through which particles acquire mass, makes the Standard Model seem complete. However, there are established phenomena and observations that it can not account for. Perhaps the most striking one is the absence of any description of the most familiar, yet the weakest, interaction of nature, meaning gravity, which has not been incorporated in a quantum field theory so far. But even at the heart of the Standard Model there are unexplained aspects, for example the well established fact that neutrinos have non-zero mass [13–16].

Furthermore, the current thesis is about matter anti-matter asymmetry. Thus it is crucial to mention the observed matter-antimatter asymmetry in the universe [19–21]. Specifically, ordinary matter, *i.e.* matter that consists of Standard Model particles, accounts for only a small fraction of the total matter density inferred to be present in the universe [22]. As a result, hints of alternative models that could explain the observed matter anti-matter asymmetry are searched in particle decays, such as  $B_s^0 \rightarrow \mu^+ \mu^-$ ,  $B^0 \rightarrow \mu^+ \mu^-$  [23] and  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  [24] or

from the branching ratios,  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau)$  and  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)$  ratio measurement [25]. Lastly, nearly all of the ordinary matter in the universe consists of particles from the first generation of fermions. Interestingly enough there is no fundamental reason embedded in the Standard Model as to why there are exactly three generations. For all of the above reasons, the scientific method compels us to continue testing Standard Model predictions and look for ways to improve it.

## 1.2 Flavor Physics

As it was mentioned in Section 1.1, quarks and leptons acquire mass through the Yukawa term of the Standard Model Lagrangian. An important aspect of the weak interaction emerges from that term, namely the *quark flavor mixing*. A brief description of the latter and its relevance to *CP* violation is given in what follows. The section concludes with introducing the *CKM-mixing matrix*, which is the core of *flavor physics* in the quark sector of the Standard Model.

### Quark flavor mixing

By applying the Higgs mechanism [26, 27] to the Standard Model Lagrangian the Higgs field obtains a *vacuum expectation value*, which corresponds to the lowest energy value of the Higgs potential. After this step the Yukawa term for the quark fields, ignoring quark-Higgs field interaction terms, is:

$$-\mathcal{L}_{\text{Yukawa}} = \left[ y_{ij}^d \bar{d}_{Li} d_{Rj} + y_{ij}^u \bar{u}_{Li} u_{Rj} \right] \frac{v}{\sqrt{2}} + h.c. + \dots \quad (1.2a)$$

$$= \left[ m_{ij}^d \bar{d}_{Li} d_{Rj} + m_{ij}^u \bar{u}_{Li} u_{Rj} \right] + h.c. + \dots, \quad (1.2b)$$

$$\text{with } m_{i,j}^{u,d} = \frac{v}{\sqrt{2}} y_{ij}^{u,d},$$

where  $v$  is the Higgs vacuum expectation value and  $y_{ij}^{u,d}$  are complex valued numbers called *Yukawa couplings*. The latter are free parameters that represent the coupling strength between Higgs and quark fields.

Note that for equations 1.2 to 1.5 it is implied that  $u$  and  $d$  indicate any up or a down type quark. The generation of the up(down) type quark is specified by the indices  $i, j$ , whereas, indices  $L, R$  indicate the left or right handedness of the quark field. Finally the matrix  $m^{u,d}$  expresses the desired quark masses.

The quark fields,  $u$  and  $d$ , in Eq. 1.2 have a definite quantum number that labels the generation to which they belong and also whether they are of up or down type. This quantum number is commonly called *flavor* and thus the quark fields are flavor eigenstates. By construction the mass matrix is not diagonal which means that a quark with a well defined flavor does not have a well defined mass. Or, in more formal phrasing, the flavor and mass eigenstates of the quark fields do not coincide. In order to obtain proper quark masses the matrix  $m^{u,d}$  has to be diagonalized, as:

$$m_{\text{diag}}^{d,u} = V_L^{d,u} m^{d,u} \left( V_R^{d,u} \right)^\dagger, \quad (1.3)$$

where the matrices  $V$  are required to be unitary. Since Eq. 1.2 has to stay intact after  $m^{d,u}$  is replaced with  $m_{\text{diag}}^{d,u}$ , quark fields need to be rotated as well as shown in Eq. 1.4, such that they cancel the additional  $V$  matrices of Eq. 1.3.

$$(d_i^m)_{L,R} = \left( V_{ij}^d d_j \right)_{L,R}, \quad (u_i^m)_{L,R} = \left( V_{ij}^u u_j \right)_{L,R}. \quad (1.4)$$

At this point the quark fields in the rest of the Standard Model Lagrangian are still flavor eigenstates. The field rotations of Eq. 1.4 need to be applied in these terms as well. Specifically, it must be applied to the kinetic term involving quark interactions with the charged weak bosons  $W^\pm$ , also known as *charged current* interaction, shown in Eq. 1.5. The charged current interaction in the same equation is expressed in two ways. One with quark fields expressed as flavor eigenstates, Eq. 1.5a and two as mass eigenstates, Eq. 1.5b.

$$\mathcal{L}_{\text{Kinetic}}^{CC} \propto \bar{u}_{Li} \gamma_\mu W^{-\mu} d_{Ri} + \bar{d}_{Li} \gamma_\mu W^{+\mu} u_{Ri} \quad (1.5a)$$

$$\propto \bar{u}_{Li}^m V_{\text{CKM}} \gamma_\mu W^{-\mu} d_{Ri}^m + \bar{d}_{Li}^m V_{\text{CKM}} W^{+\mu} \gamma_\mu u_{Ri}^m, \quad (1.5b)$$

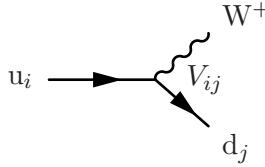


FIGURE 1.2: Feynman diagram where an up-type quark couples to any of the three, (d,s,b), down type quarks, via a  $W^+$  boson. Time flows from left to right.

with  $V_{CKM} \equiv V_L^u V_L^{d\dagger}$  and  $W^\pm$  are the charged weak boson fields, whereas  $\gamma_\mu$  are Dirac matrices. The resulting matrix is the *CKM mixing matrix*:

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{CKM}} \begin{pmatrix} d^m \\ s^m \\ b^m \end{pmatrix}. \quad (1.6)$$

The charged current term, Eq. 1.5b, incorporates an important aspect of the weak interaction, namely the *quark flavor mixing*. The latter is due to the fact that quark mass eigenstates are superpositions of the flavor eigenstates. By definition the  $V_{CKM}$  rotates only the down type quarks, implying that the mass eigenstates of the up-type quarks are identical to the flavor eigenstates. Thus an up type quark can change its flavor to any of the down type quarks, see Figure 1.2, with a certain probability. The probability of such a transition is given by the corresponding element of the CKM mixing matrix.

### CKM mixing matrix

The CKM matrix, or simply CKM matrix, is a unitary matrix which, as previously mentioned, describes the strength of quark couplings, or in other words, the probability of a certain quark flavor transitions. The elements of the CKM matrix have been measured, see *e.g.* Chapter 12 of [28], showing the following structure for their magnitudes:

$$|V_{\text{CKM}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \simeq \begin{pmatrix} 1 & 0.2 & 0.008 \\ 0.2 & 1 & 0.04 \\ 0.008 & 0.04 & 1 \end{pmatrix}. \quad (1.7)$$

Essentially the structure implies that transitions between generations are suppressed with respect to transitions within the same generation in a hierarchical way. The most suppressed transitions are between the first and third generations followed by the ones between the second and third and the least suppressed are between first and second.

The unitarity of the CKM matrix is at the center of flavor physics. In order to achieve such tests a parametrization of the CKM matrix is useful. After exploiting the unitarity of  $V_{\text{CKM}}$  and all the redundant quark field phases, the CKM matrix has, by construction, 3 real parameters and one complex phase. The choice of the CKM matrix parametrization is *a priori* arbitrary. However due to the observed structure the *Wolfenstein* [29, 30] parametrization is a standard choice. The Wolfenstein parametrization utilizes three real parameters,  $\lambda, A, \rho$  and an imaginary one  $i\eta$ , as follows:

$$|V_{\text{CKM}}| = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.8)$$

As previously mentioned, the CKM matrix is unitary, meaning that  $V_{\text{CKM}}V_{\text{CKM}}^\dagger = I_{3 \times 3}$ . This leads to the orthogonality relations. The latter are sums of complex numbers that are equal to zero, and thus can be represented as *unitarity triangles* in the complex plane. There are six orthogonality relations, two of which, shown in Eq. 1.9, are relevant for this thesis since the CKM elements present in these relations govern the dynamics in the  $B_s^0$  and  $B^0$  meson systems.

$$B^0 : V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (1.9a)$$

$$B_s^0 : V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0. \quad (1.9b)$$

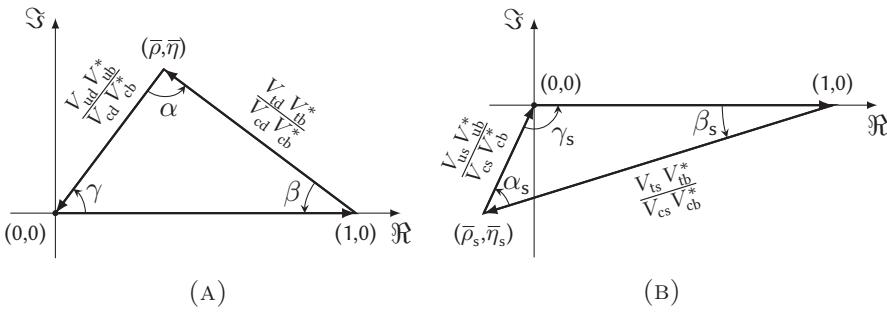


FIGURE 1.3: CKM-unitarity triangles.  $B^0$  triangle (left), corresponding to Eq. 1.9a.  $B_s^0$  triangle (right), corresponding to Eq. 1.9b. Triangle sides have been normalized, see text. This way one of the sides is real with unit length. Note that triangles are not drawn to scale. Figures from [31].

These two relations, after dividing by  $V_{cd}V_{cb}^*$ ,  $V_{cs}V_{cb}^*$  respectively for  $B^0$  and  $B_s^0$ , are illustrated in Figure 1.3. Note that the CKM element  $V_{ts}$  has a complex part at higher order in  $\lambda$ , see section 13.3 of [28]. By inspecting the triangles one can define some of the angles as follows:

$$\begin{aligned} \alpha &\equiv \arg \left( -\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right), \quad \beta \equiv \arg \left( -\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right), \quad \gamma \equiv \arg \left( -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right), \\ \beta_s &\equiv \arg \left( -\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} \right). \end{aligned} \quad (1.10)$$

Note that the definitions of Eq. 1.10 are independent of the quark field phases. Thus the above angles are useful observables, regardless of the chosen CKM matrix parametrization.

There is no fundamental reason known in the Standard model for the observed hierarchy of the CKM elements. One of the main goals of flavor physics is to verify the consistency of the CKM picture. Overlaying measurements of flavor physics observables in the complex  $\bar{\rho} - \bar{\eta}$  plane of Figure 1.4 should show a compatible picture regarding the position of the apex of the unitarity triangle. The apexes of the  $B^0$  and  $B_s^0$  triangles

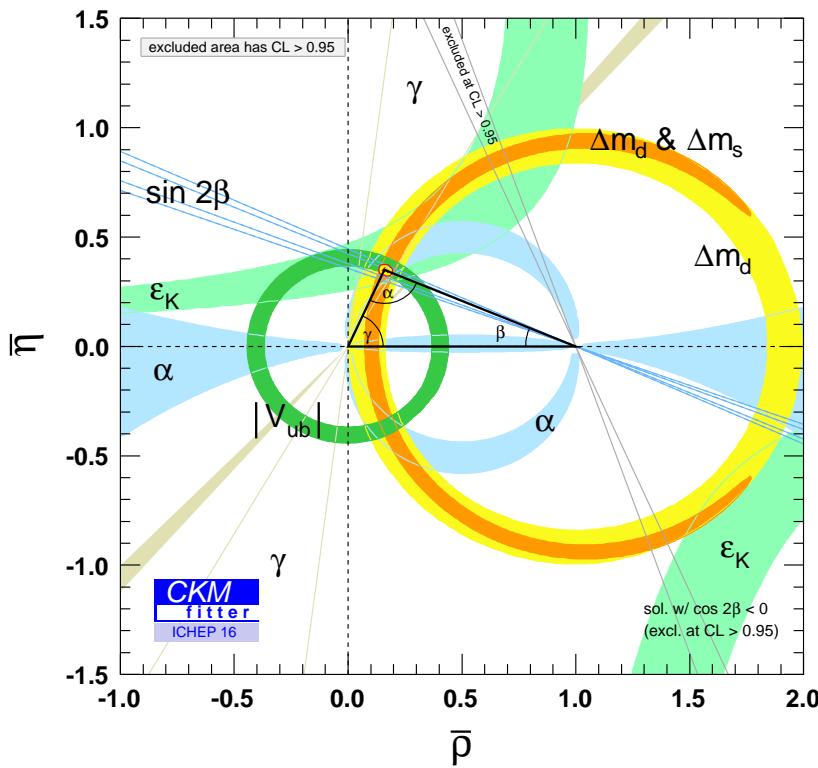


FIGURE 1.4: Global fit of the  $B^0$  triangle [32]. Inputs from measurements of flavor physics observables are indicated by colored bands.

are defined in a convention independent way as:

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \quad \bar{\rho}_s + i\bar{\eta}_s = -\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*}. \quad (1.11)$$

For completeness it is interesting to mention that in the lepton sector a similar matrix is present. The observed hierarchy there is completely different which is yet another intriguing feature of nature.

### CP violation and fermion masses

Having introduced both the fermion masses and  $CP$  violation it is interesting to point out their common origin in the Standard Model [33]. Elaborating more, from one hand the CKM matrix was introduced to couple the Higgs to the fermions and thus provide them with a mass. On the other hand the fact that  $V_{CKM}$  is a complex valued matrix allows for  $CP$  violation in the Standard Model via the weak interaction. This is because the charged current interaction of Eq. 1.5b is invariant under  $CP$  operations only if  $V_{CKM} = V_{CKM}^*$ , which is evidently not the case.

## 1.3 The $B_s^0$ Meson and $CP$ violation

The  $B_s^0$  meson is an electrically neutral particle that consists of a quark anti-quark pair, specifically  $(\bar{b} s)$ . The two constituents of the mesons are bound by the strong interaction. An important feature of the  $B_s^0$  meson, and all heavy neutral mesons, is that it can change into its antiparticle, the  $\bar{B}_s^0$  meson and *vice versa*. This feature is called *meson-antimeson oscillations* and it is possible in the Standard Model via *box diagram* of Figure 1.5. Meson oscillations play a central role in flavor physics. Particularly because they are sensitive to the existence of new particles as will be explained in Section 1.4.2.  $CP$  violating effects can appear in  $B_s^0 - \bar{B}_s^0$  oscillations, in the subsequent decay as well as in the interference between oscillation and decay. The current section as well as Section 1.4 qualitatively addresses the way that these effects manifest themselves.

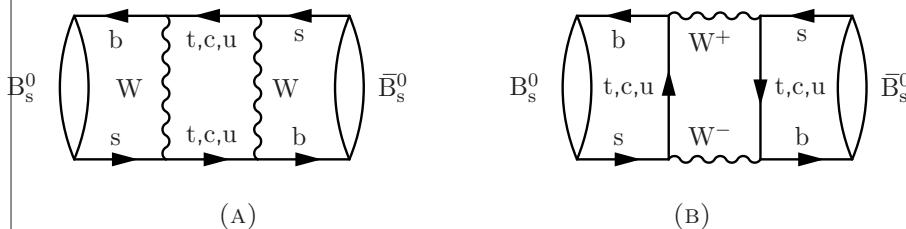


FIGURE 1.5: Leading order diagrams for  $B_s^0 - \bar{B}_s^0$  oscillations. Figures from [31].

### The $B_s^0 - \bar{B}_s^0$ System

The behavior of the  $B_s^0 - \bar{B}_s^0$  is described phenomenologically, where the two mesons are treated as a coupled quantum mechanical system. This approach [34, 35] uses an *effective Hamiltonian* [36, 37] to describe the time evolution of the system. The wave-functions of the mesons are:

$$|B_s^0\rangle \equiv |\bar{b}s\rangle, \quad |\bar{B}_s^0\rangle \equiv |b\bar{s}\rangle, \quad (1.12)$$

implying that the states  $B_s^0$  and  $\bar{B}_s^0$  have a definite quark content, or in other words, they are flavor eigenstates. The effective Hamiltonian,  $\mathcal{H}_{\text{eff}}$ , is a 2x2 non-diagonal non-Hermitian matrix. The non-diagonal feature is due to the presence of  $B_s^0 - \bar{B}_s^0$  transition amplitudes, whereas the non Hermitian description takes into account the probability that the system will eventually decay. By diagonalizing the effective Hamiltonian one can obtain the mass eigenstates of the system:

$$\begin{aligned} |B_{s,H}^0\rangle &= p|B_s^0\rangle + q|\bar{B}_s^0\rangle, \\ |B_{s,L}^0\rangle &= p|B_s^0\rangle - q|\bar{B}_s^0\rangle. \end{aligned} \quad (1.13)$$

Note the difference between the masses,  $m_{B_{s,H}^0}$  and  $m_{B_{s,L}^0}$ , of the previous mass eigenstates and the average mass,  $m_{B_s^0}$ , of the  $B_s^0 - \bar{B}_s^0$  system, which is used in Section 4.1. Similarly the lifetimes  $\tau_H$  and  $\tau_L$ , refer to the lifetime of the heavy and light mass eigenstates, whereas the  $\tau_{B_s^0}$  is the average lifetime of the previous system. More details on the exact calculations can be found both in section 13.1 of [28] and in [31, 37]. The amount of mixing between the mass and flavor eigenstates is governed by the complex parameters  $p, q$ . Theoretical calculations [38] and experimental measurement [39] point to the ratio  $q/p$  being compatible with one. In addition the difference in mass and decay width between the two mass eigenstates are important observables of the  $B_s^0 - \bar{B}_s^0$  system and are denoted as  $\Delta m_s$  and  $\Delta\Gamma_s$  respectively. Note that decay widths,  $\Gamma_x$  are the inverse of the corresponding lifetime  $\tau_x$ , where  $x : \{B_s^0, H, L\}$ .

### Classification of $CP$ Violation effects

The ratio  $|q/p|$  is related to a type of  $CP$  violation. Specifically,  $CP$  violation *in the mixing* (of the  $B_s^0 - \bar{B}_s^0$  system). This type of asymmetry vanishes if  $|q/p| = 1$  which implies that the mass eigenstates of Eq. 1.13 consist of equal amounts of the flavor eigenstates in Eq. 1.12.

Another type of  $CP$  violation, *in the decay* or *direct*  $CP$  violation, is based on the amplitude,  $A_f$ , of a certain meson decay to some final state  $f$ . Given the  $CP$  conjugated  $\bar{A}_{\bar{f}}$  amplitude, this type of asymmetry vanishes when  $|\bar{A}_{\bar{f}}/A_f| = 1$ , which implies that the probabilities for the  $B \rightarrow f$  and  $\bar{B} \rightarrow \bar{f}$  processes are equal. Note that unlike  $CP$  violation in mixing which is restricted to neutral mesons,  $CP$  violation in the decay can manifest itself in the decays of charged mesons.

There is one more type of  $CP$  violation and that originates from the *interference* between a neutral meson decaying directly to a final state  $f$  or oscillating to its antiparticle and then decaying to the same final state  $f$ . Due to its relevance to the current thesis, this type of  $CP$  violation is explained in more detail in the next section. All the above classifications are based on section 13.1.4 of [28].

Finally, it is useful at that stage to introduce the decay rate formulas of the  $B_s^0$  and  $\bar{B}_s^0$  mesons in the next set of equations. For a full derivation refer to [28, 31, 37].

$$\Gamma(\bar{B}_s^0 \rightarrow f) = |q/p|^2 |A_f|^2 (1 + |\lambda_f|^2) \frac{e^{-\Gamma_s t}}{2} \left[ \cosh \frac{\Delta\Gamma_s t}{2} + \mathcal{A}_{CP}^{\Delta\Gamma} \sinh \frac{\Delta\Gamma_s t}{2} - \mathcal{A}_{CP}^{\text{dir}} \cos \frac{\Delta m_s t}{2} + \mathcal{A}_{CP}^{\text{mix}} \sin \frac{\Delta m_s t}{2} \right] \quad (1.14a)$$

$$\Gamma(B_s^0 \rightarrow f) = |A_f|^2 (1 + |\lambda_f|^2) \frac{e^{-\Gamma_s t}}{2} \left[ \cosh \frac{\Delta\Gamma_s t}{2} + \mathcal{A}_{CP}^{\Delta\Gamma} \sinh \frac{\Delta\Gamma_s t}{2} + \mathcal{A}_{CP}^{\text{dir}} \cos \frac{\Delta m_s t}{2} - \mathcal{A}_{CP}^{\text{mix}} \sin \frac{\Delta m_s t}{2} \right]. \quad (1.14b)$$

The letter  $t$  in the same equation represents time and  $\lambda_f$  is a parameter commonly used to parameterize  $CP$  violation in meson decays which is defined as:

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f}. \quad (1.15)$$

The parameter  $\lambda_f$  above is not to be confused with the  $\lambda$  of the Wolfenstein parametrization. Note also the coefficients of the trigonometric functions in Eq. 1.14. These parameters become relevant in Chapter 4 and Chapter 5 and are thus defined in the following equation:

$$\mathcal{A}_{CP}^{\text{dir}} \equiv \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \quad \mathcal{A}_{CP}^{\text{mix}} \equiv \frac{2\Im\lambda_f}{1 + |\lambda_f|^2}, \quad \mathcal{A}_{CP}^{\Delta\Gamma} \equiv \frac{2\Re\lambda_f}{1 + |\lambda_f|^2}. \quad (1.16)$$

The first asymmetry,  $\mathcal{A}_{CP}^{\text{dir}}$ , quantifies  $CP$  violation in the decay while the second asymmetry,  $\mathcal{A}_{CP}^{\text{mix}}$ , expresses the mixing induced  $CP$  asymmetry. The last one,  $\mathcal{A}_{CP}^{\Delta\Gamma}$  is associated with the difference in the decay rate of the mass eigenstates of Eq. 1.13.

## 1.4 The Weak Phase $\phi_s$

The weak phase  $\phi_s$  is an important parameter of the  $B_s^0 - \bar{B}_s^0$  system. It is related to the third type of  $CP$  violation mentioned in Section 1.3, namely in the interference between the decay and mixing amplitudes, see Figure 1.6. In the current section  $\phi_s$  and its relevance to the search for new particles is discussed.

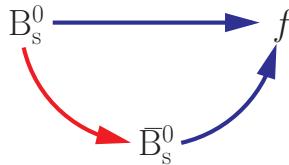


FIGURE 1.6: The two interfering decay paths leading to the same final state.

The parameter  $\phi_s$  manifests itself in the time dependence of the  $b \rightarrow c\bar{c}s$  transitions. An example of the latter is shown in Figure 1.7.

In addition, the final state of the  $B_s^0$  meson has to be a  $CP$  eigenstate, which implies that both  $B_s^0$  and  $\bar{B}_s^0$  can decay into it, such that the two decay paths illustrated in Figure 1.6 can interfere with each other. The asymmetry associated to the interference between mixing and decay is:

$$a_{CP}^{\text{interf}}(t) \equiv \frac{\Gamma(B_s^0 \rightarrow f) - \Gamma(\bar{B}_s^0 \rightarrow f)}{\Gamma(B_s^0 \rightarrow f) + \Gamma(\bar{B}_s^0 \rightarrow f)}. \quad (1.17)$$

Note the common final state  $f$  that both  $B_s^0$  and  $\bar{B}_s^0$  can decay to, which is due to the fact that  $f$  is a  $CP$  eigenstate. Assuming no  $CP$  violation in both  $B_s^0 - \bar{B}_s^0$  mixing and in the decay to the final state  $f$  and using the  $B_s^0 - \bar{B}_s^0$  decay rate formulas in Eq. 1.14, the asymmetry of Eq. 1.17 can be expressed as:

$$a_{CP}^{\text{interf}}(t) = \frac{-\Im(\lambda_f) \sin(\Delta m_s t)}{\cosh(\frac{1}{2}\Delta\Gamma_s t) - \Re(\lambda_f) \sinh(\frac{1}{2}\Delta\Gamma_s t)}. \quad (1.18)$$

It can be seen that the previous asymmetry shows an oscillatory behavior versus time, due to the presence of the  $\sin(\Delta m_s t)$  term. The oscillation frequency is driven by  $\Delta m_s$ , whereas the amplitude by  $\Im\lambda_f$ .

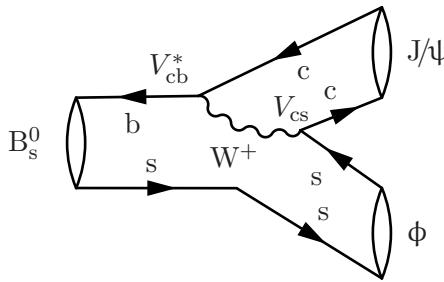


FIGURE 1.7: Leading order tree diagram of the  $B_s^0 \rightarrow J/\psi \phi$  decay.

It is interesting to point out that according to Eq. 1.18 and Eq. 1.16  $CP$  asymmetry can still be observed in the interference between decay and mixing even if no  $CP$  violation occurs in either the mixing or decay. This is possible if the imaginary part,  $\Im\lambda_f$ , in Eq. 1.15 is not zero. The

asymmetry of Eq. 1.18 can vanish at a particular point of time or for a given time interval. Because of that Eq. 1.18 should not be integrated over time implying that the previous asymmetry has to be measured as a function of the *decay time* of the  $B_s^0$  ( $\bar{B}_s^0$ ) meson. Where decay time is the time interval between the production and decay of a particle, for example the  $B_s^0$  particle.

The imaginary part of Eq. 1.15 can be expressed as a combination of CKM elements, given a certain final state  $f$ , like  $J/\psi\phi$ . The latter final state is the channel that yields the most precise  $\phi_s$  measurement with respect to other  $b \rightarrow c\bar{s}$  modes. Assuming that for this final state only the leading order diagrams of Figure 1.5 and Figure 1.7 contribute to the  $B_s^0 - \bar{B}_s^0$  mixing and decay respectively; Then the leading order Standard Model estimate of  $\Im(\lambda_f)$  is written as:

$$\Im(\lambda_f) = \Im \left( \frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}} V_{cs}^*V_{cb} \frac{1}{V_{cs}V_{cb}^*} \right), \quad (1.19)$$

where the first fraction is related to the mixing part  $q/p$  whereas the second and third come from the  $\bar{B}_s^0$  and  $B_s^0$  decay amplitudes. Note that the t quark dominates the loop in Figure 1.5 over the c and u quarks. Re-writing the imaginary part of Eq. 1.19 considering Eq. 1.10 leads to:

$$\Im \left( \left[ \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} \right]^2 \right) = \sin \left( 2 \arg \left( \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} \right) \right) = \sin(-2\beta_s). \quad (1.20)$$

Assuming tree level  $b \rightarrow c\bar{s}$  transitions, like the one in Figure 1.7, the parameter  $\phi_s$  is finally defined as:

$$\phi_s^{\text{SM,tree}} \equiv -2\beta_s. \quad (1.21)$$

Lastly, an implication arises from the choice of the intermediate resonances,  $J/\psi$  and  $\phi$ , which both have a non-zero spin quantum number. Specifically, the sum of the  $J/\psi$  and  $\phi$  spin vectors has to be compensated by orbital angular momentum,  $l$ , such that the  $J/\psi\phi$  state has a zero total angular momentum, matching the total angular momentum of the initial  $B_s^0$  meson (angular momentum conservation is implied). As a

result, the presence of orbital angular momentum makes it such that the  $J/\psi\phi$  is not a pure  $CP$ -eigenstate, but rather an admixture of a  $CP$ -odd and  $CP$ -even components. Thus, for a particular  $J/\psi\phi$  decay the value of the total angular momentum,  $l$ , dictates whether the  $CP$  eigenvalue of the previous decay is  $CP$ -odd or  $CP$ -even:

$$CP|J/\psi\phi\rangle = (-1)^l|J/\psi\phi\rangle. \quad (1.22)$$

As a consequence Eq. 1.20 needs to be modified as:

$$\sin(-2\beta_s) \rightarrow (-1)^l \sin(-2\beta_s) \quad (1.23)$$

It turns out that in order to properly measure  $\phi_s$  it is necessary to disentangle the  $CP$ -odd and  $CP$ -even components by means of an angular analysis of  $B_s^0 \rightarrow J/\psi\phi$  decays. The angular analysis is based on the angular distributions of  $\phi \rightarrow K^+K^-$  and  $J/\psi \rightarrow \mu^+\mu^-$ , see Section 4.2.1 for the angular analysis of  $B_s^0 \rightarrow J/\psi\bar{K}^{*0}$  decays. For completeness it is mentioned that the equivalent weak phase and decay channel in the  $B^0 - \bar{B}^0$  is  $\phi_d = -\sin 2\beta$  and  $B^0 \rightarrow J/\psi K_s^0$  respectively. The only difference is that no angular analysis is necessary since the final state  $J/\psi K_s^0$  is a pure  $CP$  eigenstate.

### 1.4.1 Measuring $\phi_s$

The parameter  $\phi_s$  has been measured at the LHCb experiment by studying the decays of  $B_s^0 \rightarrow J/\psi\phi$  [1]. The particular choice of the final state, where  $\phi \rightarrow K^+K^-$  and  $J/\psi \rightarrow \mu^+\mu^-$ , is preferred for several reasons. First the large number of selected  $B_s^0 \rightarrow J/\psi\phi$  decays available at LHCb, roughly 90k at the end of the Run 2 is an advantage. Second, it is relatively easy to select  $\phi$  mesons over other mesons that decay to a  $K^+K^-$  pair. This is due to the fact that the  $\phi$  meson resonance dominates the  $K^+K^-$  mass distribution within a window of about  $\pm 30 \text{ MeV}/c^2$  around its mass. This allows for the analysis to be almost independent of the  $m_{KK}$  mass variable simplifying the procedure. However an angular analysis is necessary to disentangle the odd and even  $CP$  components of the final state  $J/\psi\phi$ . In addition, as it was already implied  $\phi_s$  can be correctly probed only when the decay time dependence of the  $B_s^0 \rightarrow J/\psi\phi$

decays is taken into account. Furthermore knowledge on whether the final state  $J/\psi\phi$  originated from a  $B_s^0$  or a  $\bar{B}_s^0$  is necessary since it significantly improves the sensitivity of the  $\phi_s$  measurement. The technique used to deduce this information is called *flavor tagging* and it complicates the analysis further. On top of that experimental effects such as angular and time efficiencies plus modeling the detector's time resolution need to be taken into account.

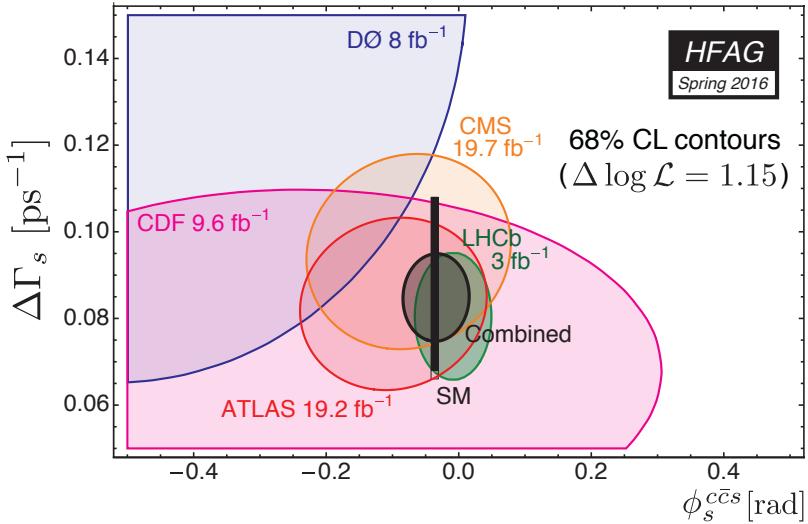


FIGURE 1.8: Likelihood contours of  $\Delta\Gamma_s - \phi_s$  (here  $\phi_s \equiv \phi_s^{c\bar{s}}$ ). Many individual measurements are combined to the black ellipse. Standard Model prediction is represented by the black band. The LHCb measurement mentioned in Eq. 1.24a is illustrated by the green ellipse. The combined measurements agree with the Standard Model predictions. Plot from [40].

The status of the  $\phi_s$  measurements is shown in Figure 1.8, where it can be seen that LHCb has the most precise measurement to date. The combined  $\phi_s$  value from LHCb [1], shown in Eq. 1.24a, includes two channels  $B_s^0 \rightarrow J/\psi\phi$  and  $B_s^0 \rightarrow J/\psi\pi^+\pi^-$ :

$$\phi_s^{\text{LHCb}} = -0.010 \pm 0.039(\text{total}) \text{ rad}, \quad (1.24a)$$

$$\phi_s^{\text{SM,tree}} = -0.03761^{+0.00073}_{-0.00082} \text{ rad}. \quad (1.24b)$$

The previous result is consistent with global fits to the Standard Model parameters Eq. 1.24b [32] (updated with summer the 2015 result), within the current experimental uncertainty.

### 1.4.2 Probing New Physics

As explained at the end of Section 1.1 Standard Model needs to be extended to describe effects and phenomena beyond our current theoretical framework. In many of the Standard Model extensions some new particle (field) is introduced. For example *Supersymmetry* [41–43] increases the number of elementary particles by predicting a supersymmetric partner for each one of them. Supersymmetry is appealing since it could solve important problems such as the *hierarchy problem* related to the value of the Higgs boson mass.

There are typically two ways that these particles are sought for, namely *on-shell* and *off-shell*. The first implies that the particle is created in high energy collisions and its presence is inferred by detecting the products of its decay. A typical example is the recent discovery of the Higgs boson at LHC [4, 5]. In the second scenario the presence of a new particle modifies some observable quantity, like  $\phi_s$ , in such a way that Standard Model cannot explain. In that case it is not necessary for this particle to be created in the first place. One of the most significant examples of this type is the existence of the t quark, whose large mass drastically affects the  $B^0 - \bar{B}^0$  oscillation frequency [44]. The second approach is the one that is followed in flavor physics.

Particularly, in the case of  $B_s^0 \rightarrow J/\psi \phi$  decays, the predicted Standard Model value of Eq. 1.24b is precise enough such that the presence of new particles could shift its value [45–47] significantly. Any deviation (small or large) is a direct evidence for physics beyond the Standard Model, making  $\phi_s$  an excellent observable for probing New Physics [45, 46]. This situation is shown in Eq. 1.25 where the measured value  $\phi_s^{\text{eff}}$

is interpreted as the sum of the Standard Model prediction  $\phi_s^{\text{SM,tree}}$  and the introduced from New Physics shift  $\Delta\phi_s^{\text{NP}}$ .

$$\phi_s^{\text{eff}} = \phi_s^{\text{SM,tree}} + \Delta\phi_s^{\text{NP}}. \quad (1.25)$$

In view of the combined  $\phi_s$  result illustrated in Figure 1.8 New Physics effects, if there in the first place, must be small [37, 48]. This is supported by other interesting measurements of flavor physics observables where deviations from Standard Model are not significant enough to claim the presence of New Physics effects. Such hints come from the rare decays  $B_s^0 \rightarrow \mu^+ \mu^-$ ,  $B^0 \rightarrow \mu^+ \mu^-$  [23] and  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  [24] or from the branching ratios,  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau)$  and  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)$  ratio measurement [25]. An interesting fact that could be helpful in the future is that all these deviations and tensions are correlated with each other. This implies that under the assumption of a certain New Physics model all the flavor physics observables need to form a coherent picture such that the assumed model can be identified [45]. The latter points out the power of flavor physics.

Lastly, the smallness of New Physics hints compels the scientific community to continue measuring the flavor physics observables with increased precision, both from the theory and the experimental side. However, in this precision era higher order effects become important and need to be taken into account, as explained in the following subsection. Given that,  $\phi_s$  will most likely continue to play an important role in the pursuit for New Physics.

### 1.4.3 Higher Order Effects in $\phi_s$

As mentioned in Section 1.4  $B_s^0 \rightarrow J/\psi \phi$  decays are dominated by the tree diagram of Figure 1.7. However the decay also receives contributions from higher order suppressed process, for example the *penguin* diagrams (or topologies) shown in Figure 1.9. The  $\phi_s$  measurement of Eq. 1.24a ignores the contribution of these penguin diagrams [1]. This has been a reasonable assumption before the LHCb measurement which showed that the measured  $\phi_s^{\text{eff}}$  value is consistent with the Standard Model prediction  $\phi_s^{\text{SM,tree}}$ , within the experimental accuracy.

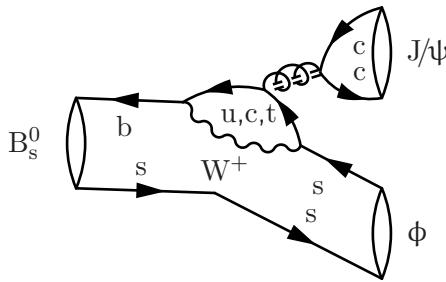


FIGURE 1.9: Higher order-penguin diagram. This diagram is suppressed as explained in Section 5.

However, with the increasing accuracy at the LHCb experiment the previous assumption needs to be revisited. Specifically, because of the fact that penguin topology contributions shifts the leading tree level Standard Model prediction. This shift could be misinterpreted as  $\Delta\phi_s^{\text{NP}}$  if these penguin contributions are not properly estimated. Thus in order to correctly probe New Physics in future measurements the penguin topology contributions  $\Delta\phi_s^{\text{SM,peng}}$  have to be estimated. The above situation is spelled out as:

$$\phi_s^{\text{eff}} = \phi_s^{\text{SM,tree}} + \Delta\phi_s^{\text{SM,peng}} + \Delta\phi_s^{\text{NP}}. \quad (1.26)$$

Calculations for the penguin contributions to  $\Delta\phi_s^{\text{SM,peng}}$  are available [49, 50]. However, these calculations are not precise, since they involve non-perturbative long-distance QCD effects. Thus an alternative approach, according to [37, 48–51], is followed in order to determine  $\Delta\phi_s^{\text{SM,peng}}$  by relying on experimental data rather than theoretical calculations. Furthermore, the SU(3) flavor symmetry, is invoked since it increases the precision on the penguin contributions estimation. The full description of the necessary formalism to determine the penguin shift,  $\Delta\phi_s^{\text{SM,peng}}$ , is given in Chapter 5.

## Chapter 2

# The LHCb Detector

The LHCb experiment is an international collaboration and is located at the *European Organization for Nuclear Research*, CERN. There are four major collider experiments at that location; Two are general purpose experiments, ATLAS and CMS, whereas ALICE and LHCb, are dedicated to *heavy ion* and *flavor* physics respectively. Heavy ion physics studies the very energetic state of matter where quarks behave almost as free particles; flavor physics is introduced in Section 1.2. All of the above experiments aim at detecting, recording and analyzing particle collisions, produced by the *Large Hadron Collider*, LHC. The latter is the world's most powerful particle accelerating machine. The LHC machine accelerates, stores and collides two beams of protons. It is also capable of handling heavy ion beams, *e.g.* lead, thus serving the heavy ion physics program of the experiments. In both cases particle storage is achieved by magnetically forcing the beams to follow a closed, almost circular, trajectory. The two beam trajectories collide with each other at four specific *interaction points* where the experiments are located. The LHC accelerator, as well as the four previously mentioned experiments, are about 100 m underground. Given the complexity and technical difficulties, building, operating and maintaining all these machines is a remarkable achievement.

As mentioned in the previous paragraph, the LHCb experiment is dedicated to studying flavor physics, and particularly B mesons, introduced in Section 1.3. The geometry of the LHCb detector, shown in Figure 2.1, is such that it exploits a particular feature of the b quark production (which is one of the constituents of the B mesons). Specifically, b quarks produced in such a collision cluster mainly along the

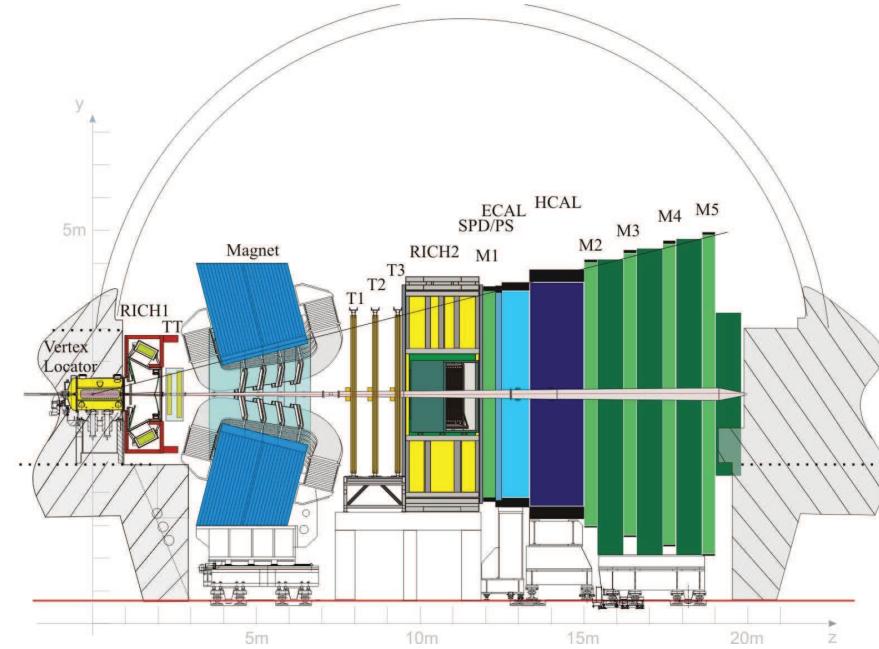


FIGURE 2.1: Schematic side view (vertical cross section) of the LHCb detector.

proton beam axis. In other words approximately 50% of the b quarks *fly* inside a cone with an opening angle of  $90^\circ$  centered around either directions of the beam axis. A large fraction of the previous quarks fly inside an even smaller solid angle around the same axis, as suggested by Figure 2.2. The previous region of solid angle is also called *forward region*. As a result, a  $4\pi$  detector geometry like the other LHC experiments, where the center of collision is surrounded by layers of detectors, is not necessary. Instead the LHCb detector covers one of the two forward regions along the beam axis, and thus approximately 25% of the produced b hadrons<sup>1</sup> are detectable by LHCb. The full LHCb detector specifications, design and performance are described in [53]. The approach of the current chapter is to briefly address the function and

<sup>1</sup> *Hadrons* are bound, by the strong interaction, states of two, three and recently [52] of four or five quarks.

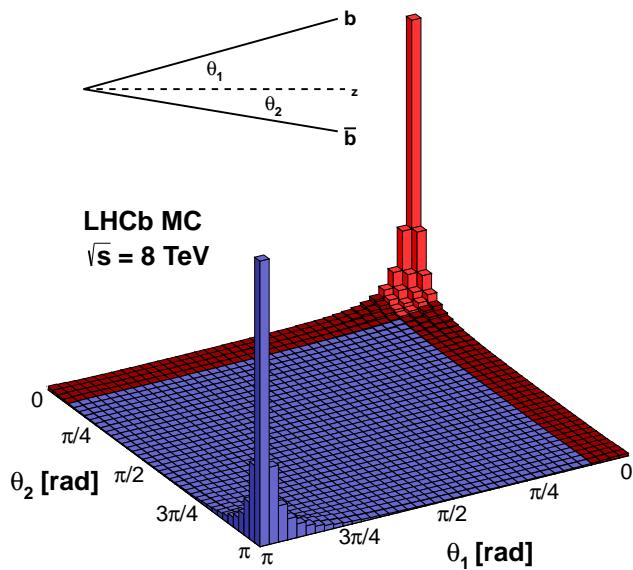


FIGURE 2.2: Decay angles distribution of b quark pair produced in LHCb. Angles are defined with respect to the beam axis  $z$ .

performance of the detector components throughout Section 2.1 to Section 2.3. Emphasis is given to the sub-detectors that are relevant to the analysis described in Chapter 4 as well as to the trigger system which is of crucial importance to the LHCb experiment.

## 2.1 Tracking Systems

The tracking sub-detectors are responsible for registering spacial coordinates of charged particles. The coordinates from multiple layers of tracking sub-systems are processed by the tracking algorithms to produce *tracks*, which are objects representing the trajectory of a charged particle. There are three distinct tracking sub-detectors in LHCb shown in Figure 2.1; Closest to the interaction point is the *VErtex LOfator*, VELO, then just before the LHCb magnet is the *Tracker Turicensis*, TT,

and lastly the T stations. The latter consists of the *Inner Tracker*, IT, and the *Outer Tracker*, OT, covering the area after the LHCb magnet. There are several track types that are *reconstructed* by the tracking algorithms, depending on which of the sub-detectors provided information for the track reconstruction.

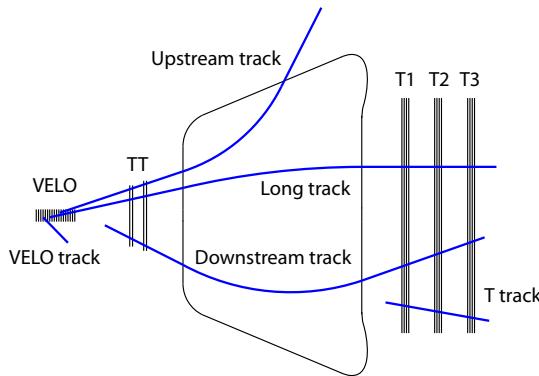


FIGURE 2.3: Various track types in LHCb.

The various track types are shown in Figure 2.3. The most common tracks used are *Long* which have the best momentum resolution. Long tracks include information from the VELO and optionally from the TT tracking systems, see Chapter 3.

A charged particle that traverses the LHCb magnet is deflected. The amount of deflection is inversely proportional to its momentum, which is exploited by tracking algorithms to estimate the momentum of a track. The relative momentum resolution of the tracking system, shown in Figure 2.4a, is  $\Delta p/p \sim 0.5\%$  for low momentum tracks and up to 1% for 200 GeV/c momentum tracks.

Measuring the mass of a particle, like the  $B_s^0$  meson, is achieved by reconstructing particles in order to build a complete cascade of particle decays (also referred to as *decay channel*). The mass measurement precision varies depending on the specific decay channel. Two body

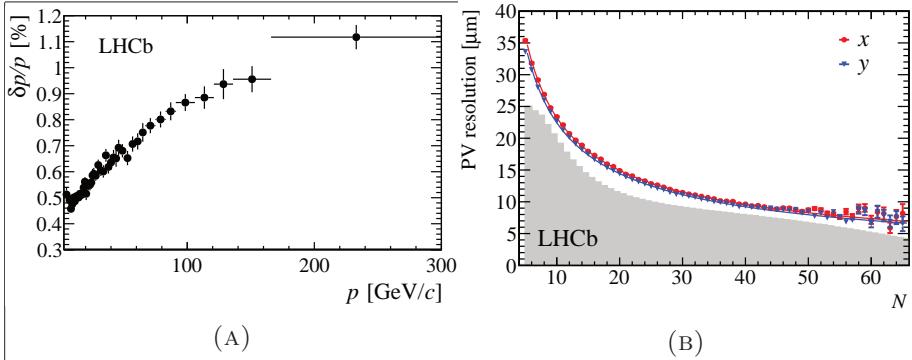


FIGURE 2.4: Tracking performance plots. (A): Relative Long track momentum resolution as a function of momentum. (B): Primary vertex resolutions versus the number of tracks present in the vertex,  $N$ . The gray area indicates the distribution of  $N$ .

$B$  decays including a  $J/\psi$  have a very precise mass resolution of about  $8 \text{ MeV}/c^2$ . On the other hand decay channels including neutral particles, *i.e.*  $B_s^0 \rightarrow \phi\gamma$ , have a worse resolution of about  $100 \text{ MeV}/c^2$ . Note that neutral particles are not detected by the tracking system. Instead the calorimetry system, introduced in Section 2.3, is responsible for identifying and reconstructing these type of particles.

Many important analyses performed with the LHCb detector rely on measuring the *decay time* of a particle, like the  $B_s^0$  meson, which is the time interval between the production and decay of a particle. This particular measurement requires knowledge of the *flight distance*, which is the distance between the point where the two protons collided, called *Primary Vertex* (PV) and the point where the  $B_s^0$ , in that case, decayed. This latter point is called *Secondary Vertex* (SV). The decay time is the actual quantity that a typical *lifetime dependent* analysis, *e.g.*  $B_s^0 \rightarrow J/\psi\phi$ , requires. The VELO sub-detector was designed for optimum spacial resolution by placing the VELO sensors as close to the beam as possible,  $\sim 8 \text{ mm}$ . The VELO sensors are the closest detector element along the entire LHC beam. The PV position resolution depends on the number of tracks used to build the vertex and can be seen in Figure 2.4b.

The average decay time resolution measured with  $B_s^0 \rightarrow J/\psi \phi$  data is  $\sim 45$  fs, which enables the  $\phi_s$  measurement discussed in Section 1.4.1.

Lastly, a big part of the LHCb physics program includes muons. The muon detector system consists of *Multi Wire Proportional Chambers* (MWPC). It is positioned after the calorimeters, with the exception of the first muon station, M1. Note that muons can penetrate through a large volume of lead, like the one present in the calorimeter system, without decaying or loosing much energy. There are five muon stations in total; Each station is divided in four concentric regions. The granularity<sup>2</sup> of each station becomes progressively smaller as one moves towards the inner regions which are closer to the beam. This is done to account for the increasing particle density. This way the special resolution is higher exactly where it is needed. The horizontal coordinate hit resolution in all regions is superior compared to the vertical coordinate. Given that the LHCb magnet bends charged particles in the horizontal plane, the position resolution in this plane is crucial for the momentum measurement. As a result the (horizontal,vertical) resolution varies from (4, 10) mm to (150, 180) mm respectively for the inner M1 region and the outer M5 region. The overall muon system efficiency is covered in [54].

## 2.2 Particle Identification

The kaons and pions are indistinguishable in the absence of dedicated *Particle IDentification*, PID, detectors. Furthermore, a particle that is not stopped by the calorimeters is most likely identified as a muon, since muons can penetrate the entire LHCb detector without decaying. In addition, properly identifying electrons, photons and pions is a challenging task that require sophisticated setup and algorithms. These experimental difficulties are tackled by exploiting information from several sub-detector systems; thus increasing the purity of the analyzed data.

The two *Cherenkov* based PID sub-detectors, RICH1 and RICH2, aim at identifying particles in the low and high momentum regions respectively. The way the RICH systems provide particle identification

---

<sup>2</sup>In this context granularity is the distance between the wires of the MWPC.

information is based on the velocity measurement implied by the angle of emission of Cherenkov radiation. The muon system also plays a role in PID. The presence of hits in the muon stations that could be associated to a certain track increase the probability that this track corresponds to a muon, see Chapter 3. Lastly, the calorimeter systems aim at increasing the chances of correctly identifying electrons, photons and pions, as mentioned in the following subsection. Further details regarding the PID performance of LHCb can be found in the overall detector [53], RICH system [55], Muon-ID [56] performance studies respectively.

## 2.3 Calorimetry

Calorimeters in LHCb measure the energy of particles provided that all of the subsequent particle decays and interactions are contained within the calorimeter. Furthermore in the case of neutral particles, like photons, calorimeters are the only means with which these neutral particles can be detected; the tracking system only detects charged particles. There are four sub-detectors that enable calorimetry at LHCb. First the combination of the *Scintillating Pad Detector* (SPD) and the *Pre Shower* (PS) detectors with a lead layer in between them enables to distinguish  $\pi^0$  from photons, and photons from electrons. Immediately after, the electromagnetic (ECAL) and hadronic (HCAL) calorimeters measure the energy of photons and neutral hadrons respectively. Note that the  $\pi^0$  decays into two photons as a result the two photons might be interpreted as a single by the ECAL. The calorimeter system is not explicitly used in the analysis of Chapter 4. However, they are implicitly used since information from the SPD and PS sub-detectors are used by the L0 trigger system to estimate the overall *multiplicity* of an event. Multiplicity essentially provide a quick estimate, to the L0 trigger system, on whether a certain event is interesting from physics point of view. Furthermore, the lead material that is inter-layered in the ECAL and HCAL active detector material, stops most of the hadrons and thus preventing them from being mis-identified as muons by the subsequent muon stations. Lastly, information from the ECAL and HCAL is used by PID algorithms utilising information from the rest of the PID systems.

## 2.4 The trigger system

The trigger system is responsible for keeping the amount of data that the detector writes to offline storage at a manageable level, since it is not feasible to save every collision-*event*. The three levels of the LHCb trigger system decide whether a certain event is interesting from physics point of view and should thus be written out to storage. The criteria used to filter out the interesting events get progressively more strict at each trigger level. The whole trigger chain is optimized for maximum signal purity given the amount of computing and storage infrastructure. In more detail the first trigger level, L0, reads part of the detector at a rate of 40 MHz, while the subsequent *Higher Level Trigger*, HLT1 and HLT2, writes out to storage at 5 kHz during the 2012 run ( and at 12.5 kHz in Run 2). The organization of the trigger is done in *trigger lines*. Essentially, each trigger line is a collection of selection criteria similar to the ones applied during the offline data analysis, but looser.

Furthermore, the HLT trigger lines are software implementations and can thus be quite general as to what they trigger on, or they can be very specific. For example a trigger line could require the presence of two muons forming a vertex with an invariant mass window around the nominal  $J/\psi$  mass. Or in case of particular decay channel a trigger line could reconstruct the entire  $n$  – body particle cascade and perhaps require particular quantities to satisfy some criteria, *i.e.* the distance between the primary and secondary vertex. The first trigger level is a pure hardware implementation whilst the other two are software only. Complicated selection criteria are not affordable due to limited computing power, given the 40 MHz input rate of L0. After the rate is reduced down to 1 MHz the data is read out and passed to the subsequent higher level trigger, where tracking and PID information become available.

The overall efficiency of the LHCb trigger system is  $\sim 90\%$  for dimuon trigger lines and  $\sim 30\%$  for multi-body lines with hadronic final states. The efficiency is estimated with the *tis-tos* method [57]. In the context of this method a given reconstructed object, *i.e.* a muon track, might cause a trigger line to fire, given that it fulfills the requirements of that particular line. In that case the muon track is characterized as *Trigger On Signal*. On the other hand the same trigger line might fire

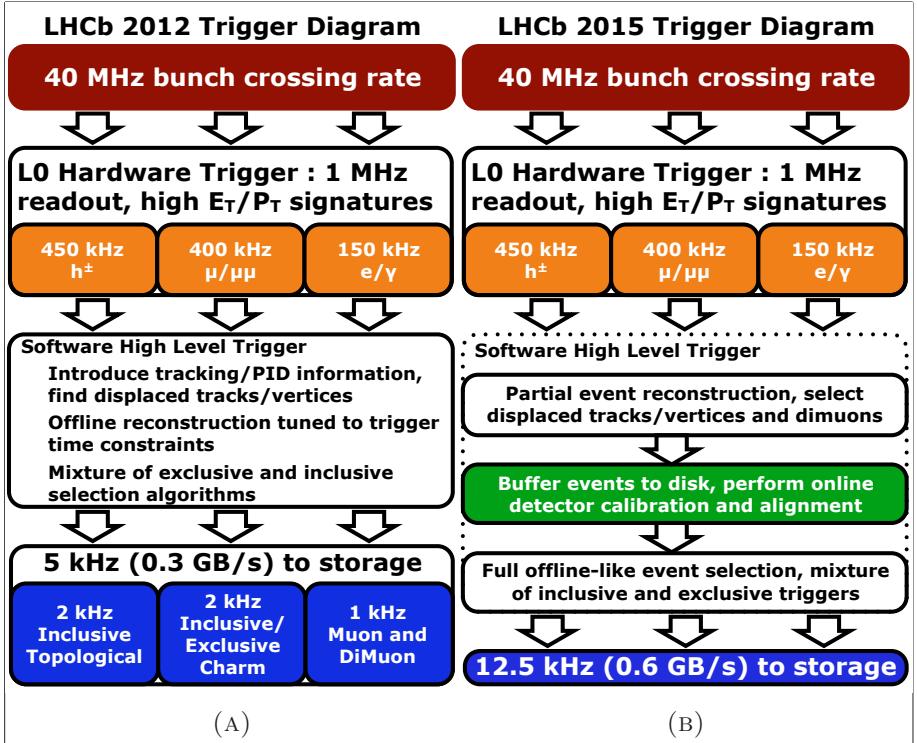


FIGURE 2.5: Run 1 (left) and Run 2 (right) trigger system schemes.

regardless of the muon track, for example because something else in the event might cause the line to fire. In that case the muon track is characterized as *Trigger Independent of Signal*. The *tos efficiency* of a given trigger line is estimated as the ratio of *tis&tos* over *tis*. A complete performance study of the trigger system can be found in chapter 5 of [53]. Muons play a key role in the current thesis; The efficiencies of the muon trigger lines of the LHCb trigger system are shown in Figure 2.6.

There are two more features of the LHCb trigger system worth mentioning. First is the *deferred* trigger scheme, according to which, 20% of the data triggered by L0 are temporarily stored and processed by the HLT during the periods where the LHC does not provide colliding

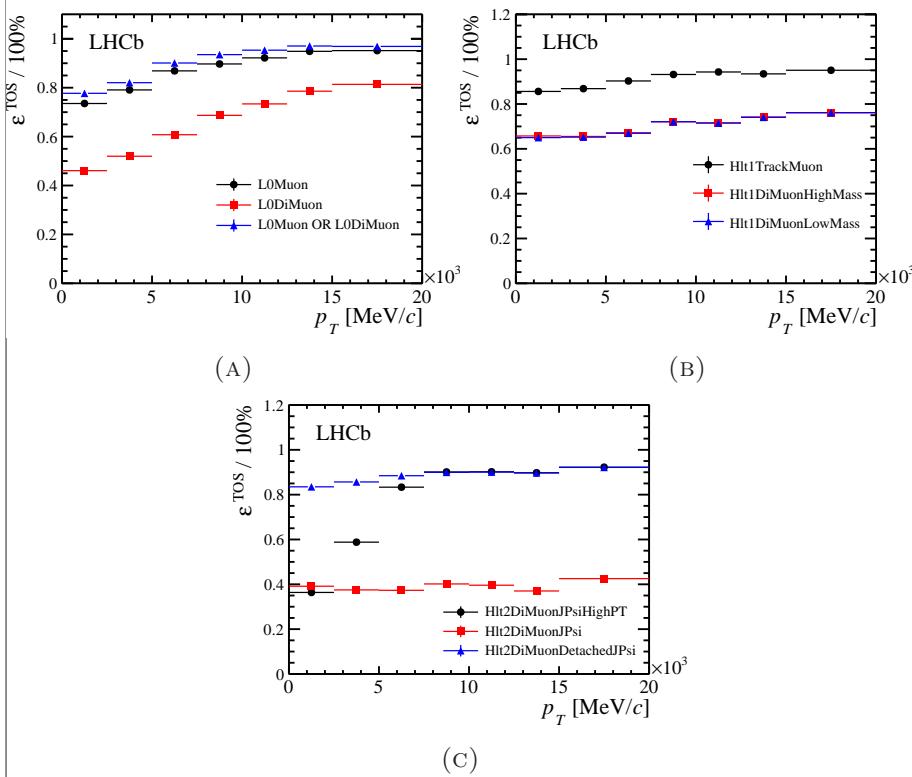


FIGURE 2.6: Muon trigger line tos efficiencies, see text, during Run 1 respectively for the L0 (A), HLT1 (B) and HLT2 (C) trigger stages.

beams and thus the LHCb trigger infrastructure would otherwise idle. This increases the amount of available computing power which can be used to improve other features of the trigger system. The second interesting feature is the novel technique of *online detector alignment and calibration* that was enabled in Run 2, see [58]. The technique makes it possible that the output data quality of the trigger becomes identical to the offline data quality. This means that offline data analysis can be performed straight from the output of the trigger, which saves computing resources and accelerates the analysis.

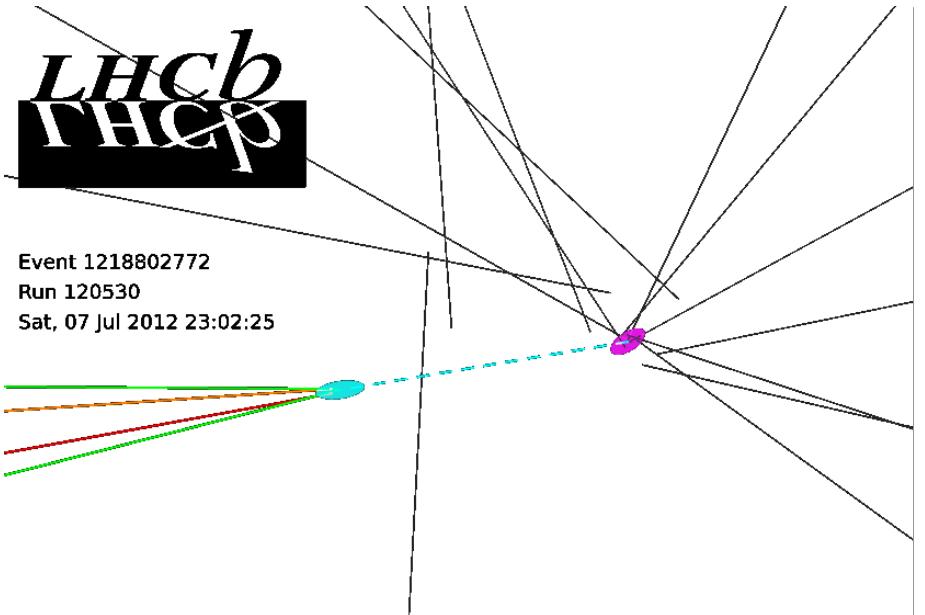


FIGURE 2.7: A  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  candidate. Purple and cyan circles correspond to the primary and secondary vertices respectively. The dashed cyan line illustrates the  $B_s^0$  flight distance. See text for more details.

## 2.5 The $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ Decay in LHCb

The  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  process, with  $J/\psi \rightarrow \mu^+ \mu^-$  and  $\bar{K}^{*0} \rightarrow K^- \pi^+$ , is a typical channel for which LHCb has a high detection efficiency. The presence of muon in the final state yields a high trigger efficiency, while the intermediate  $K^{*0}$  resonance decays within the VELO into a  $K^+$  and a  $\pi^-$  thus resulting in two Long tracks. Thanks to PID information and the optimized selection described in Section 2.2 most of the background is filtered out. The only difficult point is the low branching fraction of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay compared to  $B^0 \rightarrow J/\psi K^{*0}$  whose invariant mass peak is very close to the one of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ , as it can be seen in Figure 4.2 and Figure 4.3. The total signal yield accumulated during Run 1 is about 1800  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  candidates.

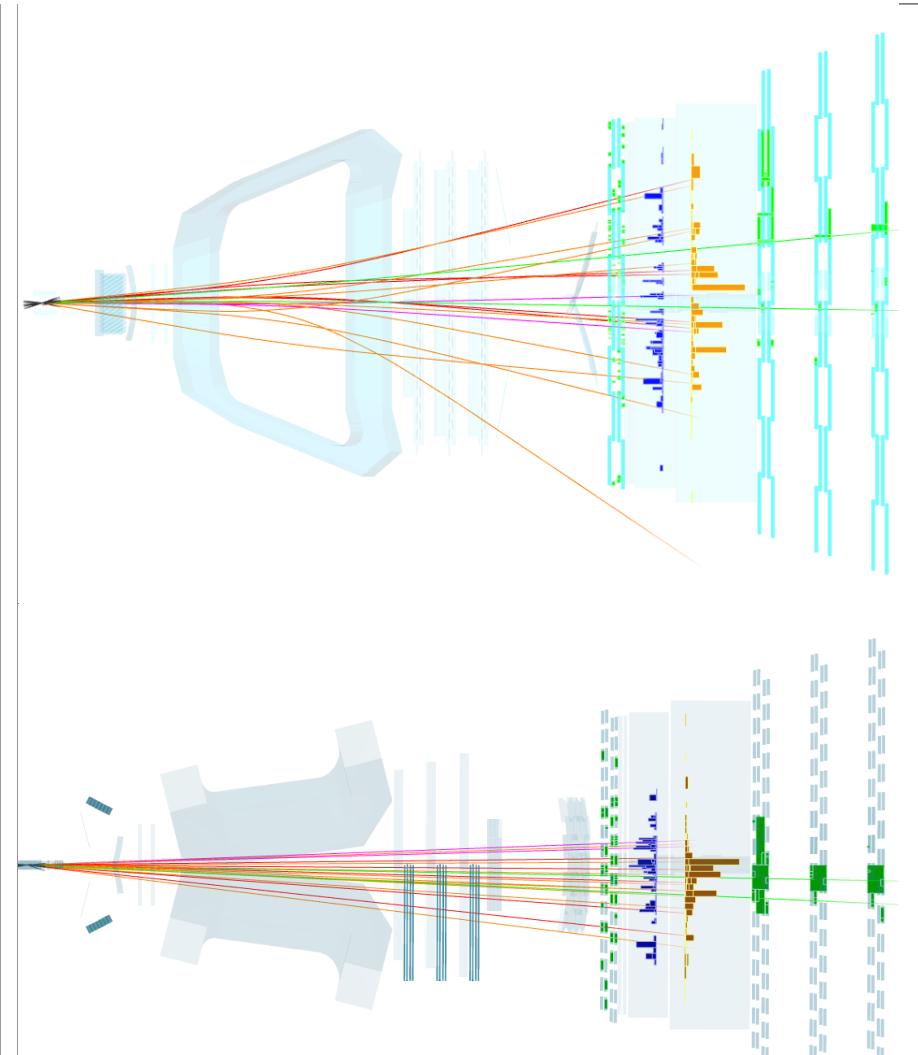
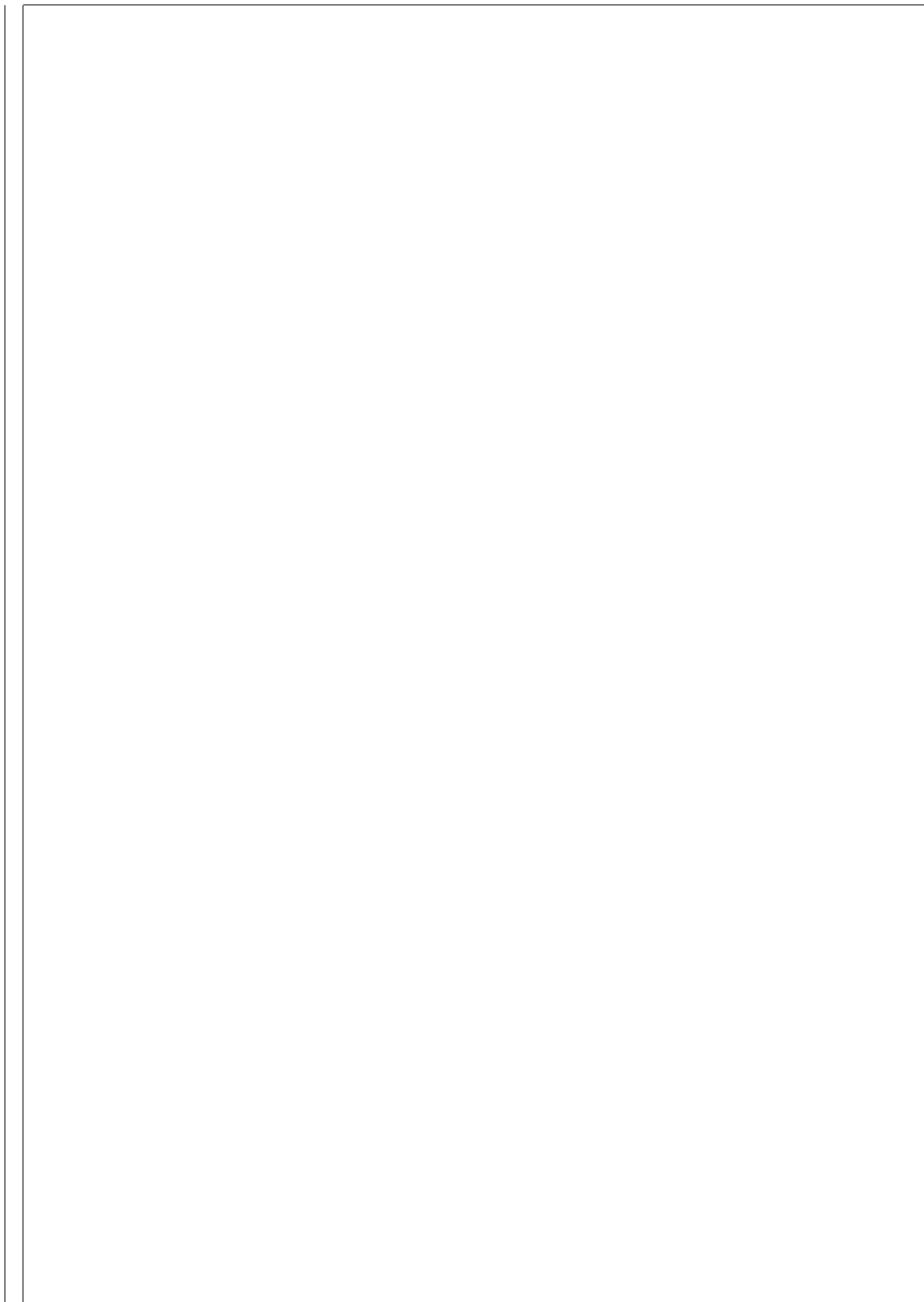


FIGURE 2.8: Horizontal (top) and vertical (bottom) cross section of an LHCb event. See text for details.

An LHCb event is shown in Figure 2.8, where ECAL and HCAL energy depositions are shown as blue and orange bars respectively; the length of each bar is proportional to the energy deposited. Muon hits are illustrated with green blocks. Long tracks are indicated by red, orange, green, magenta and blue colored lines corresponding to kaons, pions, muons, protons and electrons respectively. A reconstructed  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay candidate is shown in Figure 2.7. Where VELO tracks are indicated by black lines. Note how some of the VELO tracks appear to originate from the primary vertex and how  $B_s^0$  final state tracks point at the secondary vertex.

---

---



---

## Chapter 3

# High Level Trigger Muon Identification in Run 2

Muons are important for the LHCb experiment and are easy to identify since they penetrate matter deeply. Therefore they provide a robust way to identify  $J/\psi$  particles. Given that a large fraction of the LHCb physics program is based on the identification of muons in the final state: for example the measurement of  $\phi_s$  through  $B_s^0 \rightarrow J/\psi \phi$  decays, the study of  $CP$  violation through semileptonic decays, and several rare decays such as  $B^0 \rightarrow K^{*0} \mu^+ \mu^-$  and  $B_s^0 \rightarrow \mu^+ \mu^-$ . It is therefore of major importance to maintain and improve the efficiency and purity of identified muons in the Run 2 data-taking of LHCb. During the shutdown period between Run 1 and 2, effort has been put to optimize the muon identification software, Muon-ID, which is used in the LHCb trigger system, introduced in Section 2.4. A brief quantitative description of the Muon-ID improvements is given in Section 3.1. Furthermore, special care is taken to improve the detection efficiency of low momentum muons, this issue is addressed in Section 3.2. Note that the latter improvement is particularly useful for the part of the LHCb physics program involving newly introduced rare processes, *e.g.*  $\Sigma^+ \rightarrow p \mu^+ \mu^-$  [59] and  $K_s^0 \rightarrow \mu^+ \mu^-$  [60].

### 3.1 HLT1 Muon-ID in the LHC Run 2

The Muon-ID [61, 62] that is used both in the High Level Trigger (HLT) [63] and in the offline event processing, or simply *offline*, has been revisited in view of the LHC Run 2. Note that during Run 1 the HLT

version of the Muon-ID was different from the one used offline in order to conform with the limitations imposed by the available computing resources. The main reasons for revisiting the software are; first to unify the Muon-ID used in HLT and offline and second to optimize the trigger efficiency of low transverse momentum muons. In particular, the Muon-ID has undergone a significant re-factory resulting in a modularized common code base between HLT and offline [64]. Because of that Muon-ID in the HLT1 and offline are identical <sup>1</sup>, which, combined with the novel online detector alignment and calibration endeavor [58], see also Figure 2.5b, allows the trigger to produce offline quality data. After the modularization, readability and maintainability of the code is significantly improved. An increase in performance of the Muon-ID code both in CPU and memory usage is also achieved.

The algorithm sequence of the HLT1 muon trigger lines, defined in Section 2.4, during Run 1 was tuned in order to comply with the output rate limitations given the existing computing infrastructure, named *Event Filter Farm* (EFF). Since then the EFF has undergone a significant upgrade resulting in an increased processing power along with increased storage capabilities. In view of this infrastructure upgrade as well as the Run 2 LHC running conditions the total output rate of the entire trigger system was increased from 5 kHz to 12.5 kHz, see Figure 2.5. The impact on HLT1 is that its output rate is roughly doubled and additional processing time is available. This boost in computing power allowed for changes in the Muon-ID algorithm sequence described in the following subsection.

### 3.1.1 HLT1 muon algorithm

The role of the trigger system and the decisions that it provides are presented in Section 2.4. The HLT1 muon trigger lines are mainly organized in single muon and dimuon ones; both exploit the same muon identification procedure which is described in the current subsection. The efficiency of the muon lines during Run 1 is shown in Figure 2.6. It can be observed that a low  $p_T$  turn on and some efficiency loss in the dimuon lines are present. In order to understand the origin of this

---

<sup>1</sup>The Muon-ID in HLT2 was the same as offline already in Run 1.

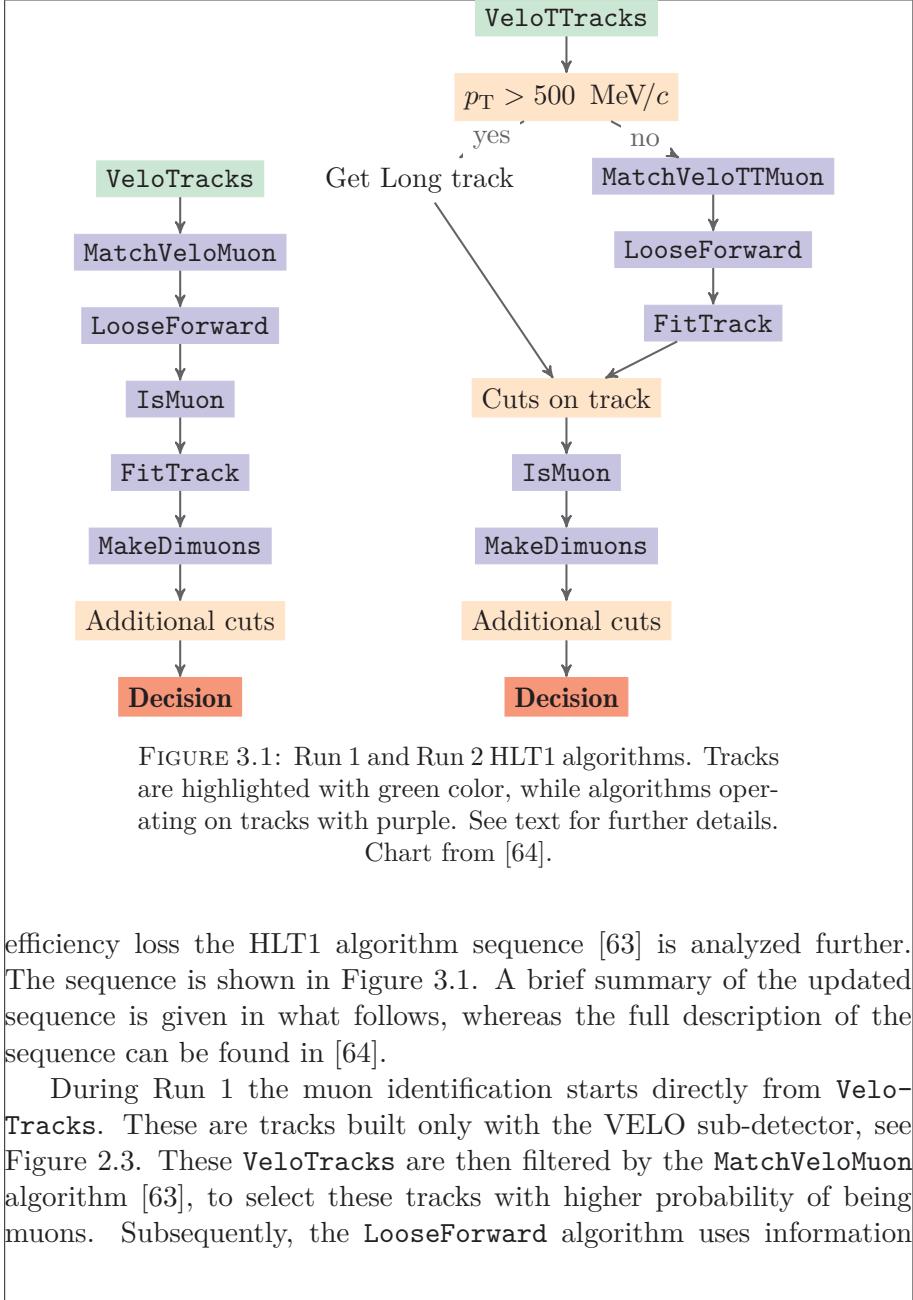


FIGURE 3.1: Run 1 and Run 2 HLT1 algorithms. Tracks are highlighted with green color, while algorithms operating on tracks with purple. See text for further details.

Chart from [64].

efficiency loss the HLT1 algorithm sequence [63] is analyzed further. The sequence is shown in Figure 3.1. A brief summary of the updated sequence is given in what follows, whereas the full description of the sequence can be found in [64].

During Run 1 the muon identification starts directly from **VeloTracks**. These are tracks built only with the VELO sub-detector, see Figure 2.3. These **VeloTracks** are then filtered by the **MatchVeloMuon** algorithm [63], to select these tracks with higher probability of being muons. Subsequently, the **LooseForward** algorithm uses information

from the tracking stations, behind the magnet and upgrades the `Velo-Tracks` to Long tracks, mentioned in Section 2.1. After this, the `IsMuon`, which implements the offline Muon-ID, algorithm is applied. In addition there are some quality cuts between each algorithm, such as the number of hits in the VELO sub-detector assigned to a track or the fitted track  $\chi^2/\text{nDoF}$  as well as momentum cuts applied after the `LooseForward` tracking algorithm. The minimum momentum and transverse momentum is  $6\text{ GeV}/c$  and  $0.5\text{ GeV}/c$  respectively.

The efficiency breakdown, performed in [64], of the Run 1 HLT1 muon lines showed that the main source of efficiency loss is due to quality and momentum cuts. Also the `MatchVeloMuon` algorithm reduces the efficiency by roughly 4%. Given these diagnostics plus the Run 2 capabilities of the EFF, the updated HLT1 muon lines algorithm sequence is described in the following paragraph.

For the majority of the tracks, which have transverse momentum larger than  $500\text{ MeV}/c$ , the sequence is different with respect to Run 1: full tracking is performed for all these tracks, so that the start for muon identification is already with Long tracks. After soft quality cuts these get passed through the standard `IsMuon` algorithm in order to be identified. Positively identified muons are then combined into dimuons as in Run 1. The combination of the use of Long tracks and standard muon identification is much more efficient than the procedure in Run 1, and is exactly the same as the one used by the offline reconstruction. Tracks with transverse momentum smaller than  $500\text{ MeV}/c$  cannot be immediately upgraded to Long tracks for timing reasons. Their lower momentum requires larger search regions in the tracking stations behind the magnet, thus increasing the number of possible hits that need to be tested for compatibility. However, some of them are muons, and the efficiency for these could be recovered using a different strategy, explained in Section 3.2.

## 3.2 Matching VELO tracks to Muon Hits

In Run 1 all `VeloTracks` in HLT1 were filtered by the `MatchVeloMuon` algorithm, as mentioned in Section 3.1.1. Given the available information from the muon stations, the latter algorithm performs a quick check on whether a VELO track is compatible with being identified as a muon by the subsequent tracking algorithms, before these algorithms are actually run. If the VELO track passes certain requirements it is accepted, otherwise it is discarded permanently and not processed further by the HLT muon lines. As a result valuable computing time is saved, since the `LooseForward` tracking and fitting algorithms are time costly.

### 3.2.1 The MatchVeloMuon algorithm

The `MatchVeloMuon` algorithm takes a VELO track as input and checks whether it is compatible with being a muon by searching for hits inside an area in the M3 muon station. This area is called *Field of Interest* (FoI). The hits found inside the M3 FoI are called *seed hits* and they are used as starting space points to look for more hits in the rest of the muon stations. For each M3 seed hit, additional hits are searched for by extrapolating from each M3 seed hit further to the rest of the muon stations. Since there is no magnetic field between the muon stations, the above-mentioned extrapolations are simple straight lines. The corresponding FoIs for the rest of the muon stations are smaller, compared to the M3 FoI, since there is no additional deflection involved. These smaller FoIs do account for the fact that the presence of heavy material between the muon stations affects the trajectory of a muon passing through. This effect is called *multiple scattering*, and it is more pronounced for low momentum tracks.

M3 is the first station to search for hits that could be assigned to a certain VELO track: This is because the number of hits in M2 is large compared to M3, see Figure 3.2 in [65]; Thus a VELO track candidate would result in a large number of seed hits, which is not computationally feasible. Note that the muon reconstruction algorithm (`IsMuon`) starts from the M3 station as well, for the same reason.

Regarding the issue of FoI size computation: As it can be seen in Figure 2.1, the LHCb magnet is between the VELO tracker and the muon system. Thus a track traversing the area from the VELO to the muon stations is deflected by the magnetic field based on its momentum. The lower the momentum of a track the more it is deflected by the magnetic field. No momentum or charge information is available for the VELO track before the `LooseForward` algorithm is run. Thus, the size of the M3 FoI is defined assuming a minimum track momentum of 6 GeV/c. This particular momentum value corresponds to the threshold value of the subsequent `IsMuon` algorithm.

At least three hits in the muon stations M2-M5 are required by the `MatchVeloMuon` algorithm in order to accept the initial VELO track. At the end of this procedure all *muon candidate tracks* that could result from an initial VELO track are considered. Within the scope of the current chapter a muon candidate is defined as the track formed by muon hits only. Note that at this stage of the algorithm an estimate of the momentum of the initial VELO track becomes available for each candidate muon track using the kick method [66]. The kick method provides an estimate of the track momentum by comparing the slope of the track before and after passing through the LHCb magnetic field. The relation between the change in slope and the track momentum is empirically estimated in [65].

As a last step the candidate muon tracks originating from the above procedure are filtered by a  $\chi^2/n\text{DoF}$  requirement. The  $\chi^2$  is simple to compute since there is no magnetic field in the region of the muon stations and all tracks are almost straight lines. Deviations from the straight line approach are only due to multiple scattering, mentioned above. The  $\chi^2$  requirement reduces the chances that random combinations of muon hits can fake a muon track. The requirement is particularly useful for low momentum tracks for which the number of M3 seed hits is larger. This is because the corresponding M3 FoI are larger, thus the number of possible random combinations of muon tracks increases.

An important detail that increases the discriminating power of the above-mentioned  $\chi^2$  is based on the assumption that the bending induced by the LHCb magnetic field can be approximated by a single instantaneous change in the slope of the charged particle trajectory. This

instantaneous change happens at the *focal plane*. The special coordinates of the focal plane form a *magnet hit*, which is included in the  $\chi^2$  computation, as if it was an actual hit. Note that this instantaneous change in slope is an approximation and in reality the track follows a smooth trajectory inside the whole area of the magnetic field, see Figure 2.8. However estimating this trajectory consumes valuable computation time during HLT1 run-time, while `MatchVeloMuon` can perform efficiently enough without a smooth track trajectory using the focal plane approximation. The position of the  $z$  focal plane as a function of the track parameters has been studied [66] and is used in the `MatchVeloMuon` algorithm. As mentioned, including information on the estimated focal plane helps rejecting fake muon tracks, because the magnet hit essentially requires from muon hits to align in a straight line that passes through the magnet hit. The latter is true for real muon tracks only.

### 3.2.2 Upgrade to `MatchVeloTTMuon`

As mentioned in Section 3.1, due to the increased HLT1 rate in Run 2 it becomes possible to remove the `MatchVeloMuon` algorithm from the sequence of the HLT1 muon lines and instead push all `VeloTracks` with  $p_T > 500 \text{ MeV}/c$  further in the algorithm sequence. Additionally, another improvement is the use of information from the TT tracker when reconstructing `VeloTracks` [67]. Note that the TT is positioned close enough to the LHCb magnet to allow for an initial track momentum estimate. The TT momentum resolution is inferior,  $\sim 15\%$ , compared to the one of a Long track. However, adding information from the TT the purity of the of the `VeloTracks` increases [68]. This is achieved by reducing the number of fake tracks also known as *ghost tracks*. The resulting set of tracks is called `VeloTTTracks`.

The above-mentioned improvements offer the opportunity to reconstruct muons with  $p_T < 500 \text{ MeV}/c$ , hereafter low momentum muons, by improving the old `MatchVeloMuon` algorithm and using it in newly introduced HLT1 muon lines which aim at reconstructing low  $p_T$  muons.

The upgraded `MatchVeloTTMuon` algorithm improves with respect to the old `MatchVeloMuon` algorithm in several points. First the available momentum and charge information from the TT is used to open

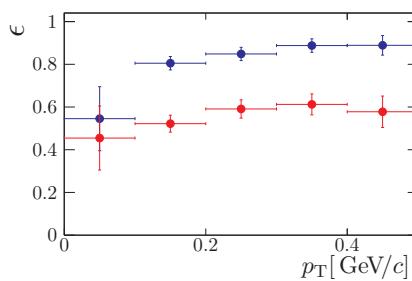


FIGURE 3.2: Efficiency comparison between the old (red) and new (blue) matching algorithms in bins of transverse momentum.

smaller M3 FoIs. This reduces the number of M3 seed hits and hence increases the purity of the `MatchVeloTTMuon` algorithm output for reasons explained earlier in Section 3.2.1. Second, the parametrization of the magnet  $z$  focal plane as well as the uncertainties on the focal plane coordinates are improved using simulated data. Furthermore, the vertical, non-bending, plane is included in the  $\chi^2$  computation which was not the case in the old `MatchVeloMuon` algorithm. Note that the vertical component of the track trajectory is almost a straight line. The two previous improvements increase the discriminating power of the  $\chi^2$ . Finally, the required number of hits in the muon stations mentioned before, is increased to four in order to further suppress fake muon tracks.

An important caveat arises from the fact some of the `VeloTTTracks` tracks do not have any information from the TT. These tracks are mainly normal high momentum tracks that pass through the TT station's inner acceptance hole where there is no active detector. Since there is no reason to discard these tracks, the upgraded `MatchVeloTTMuon` treats them as the old `MatchVeloMuon` algorithm did, meaning that the M3 FoI sizes are adjusted such that they cover the possibility for both positive and negative charged tracks.

After the improvements mentioned in the above paragraph the efficiency of the upgraded `MatchVeloTTMuon` algorithm is estimated using simulated  $\Sigma^+ \rightarrow p\mu^+\mu^-$  events, where the mean transverse momentum is  $\sim 300$  MeV/ $c$ . The muons tracks from this decay, which passed the

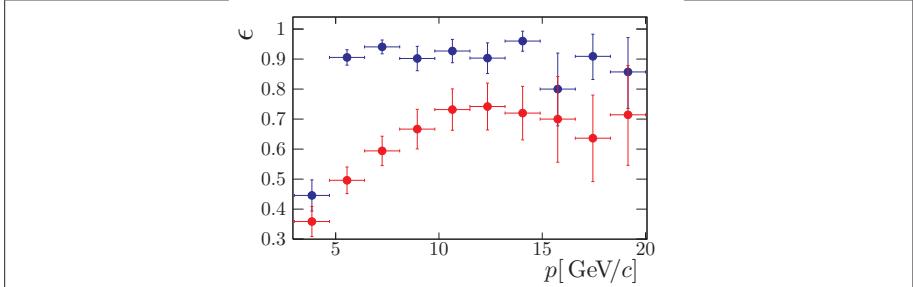


FIGURE 3.3: Efficiency comparison between the old (red) and new (blue) matching algorithms in bins of total momentum.

full track reconstruction resulting in Long tracks, are further filtered to ensure that they indeed correspond to a muon particle. This is possible by exploiting the available information in the simulated event and thus defining *true muons*. An important concept that is used for the `MatchVeloTTMuon` efficiency definition is the *overlap*. Specifically, a certain VELO track overlaps with a true muon (which is a Long track) when the VELO track has more than 70% VELO hits in common with the VELO segment of the true muon track. The number of all `VeloTTracks` that overlap with a true muon define the denominator in the `MatchVeloTTMuon` efficiency, while these `VeloTTracks` that are in addition accepted by `MatchVeloTTMuon` make the number in the numerator.

The efficiency is compared against the old `MatchVeloMuon` algorithm, where a common cutoff value for the  $\chi^2/\text{nDoF}$  of 2 is chosen. The efficiency versus the total momentum, can be seen in Figure 3.3, where the improvement of the upgraded `MatchVeloTTMuon` algorithm is shown to be significant. Given that the aim of the `MatchVeloTTMuon` upgrade is to optimize for high efficiency of low transverse momenta muons the efficiency is also projected versus  $p_T$  and shown in Figure 3.2. As mentioned in Section 3.1, the high  $p_T$  muons will not be processed by the `MatchVeloTTMuon` algorithm and thus the previous figure focuses on the region where  $p_T \in [0, 0.5] \text{ GeV}/c$ . The relative efficiency increase with respect to the old `MatchVeloMuon` algorithm in the above region is  $\sim 50\%$ . Finally, the improvement of the  $\chi^2/\text{nDoF}$  discriminating power between

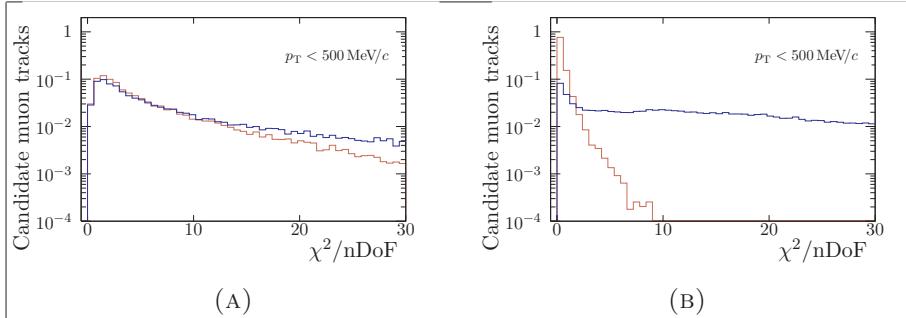


FIGURE 3.4: Distribution of the  $\chi^2$  between the old (left) and new (right) matching algorithms. Only low momentum muon tracks are considered. Matched (not matched) muon tracks are shown with orange (blue).

the old and upgraded algorithm is shown in Figure 3.4.

### **z focal plane parametrization**

The empirical parametrization of the magnet focal plane  $z$  coordinate already used in the old `MatchVeloMuon` [65], is updated. Furthermore the uncertainties on the focal plane intersection estimate are empirically estimated using simulated data. Specifically, since the true muon track is known it is possible to tune these uncertainties such that they yield the same momentum estimation as the full computation of the muon trajectory would do. Note that a momentum estimate is available by the `MatchVeloTTMuon` algorithm as soon as a seed hit in M3 is found, using the kick method.

### **M3 Field of Interest**

As explained in Section 3.2.1 the M3 FoI  $x$  coordinates are computed assuming a minimum track momentum, of  $6 \text{ GeV}/c$ , and thus maximum deflection. For a  $6 \text{ GeV}/c$  muon the deflection caused by the LHCb magnet is computed using the empirical parametrization mentioned in Section 3.2.2 of [65]. Using this technique makes it possible to estimate the region in M3 that a muon track will arrive at without performing the

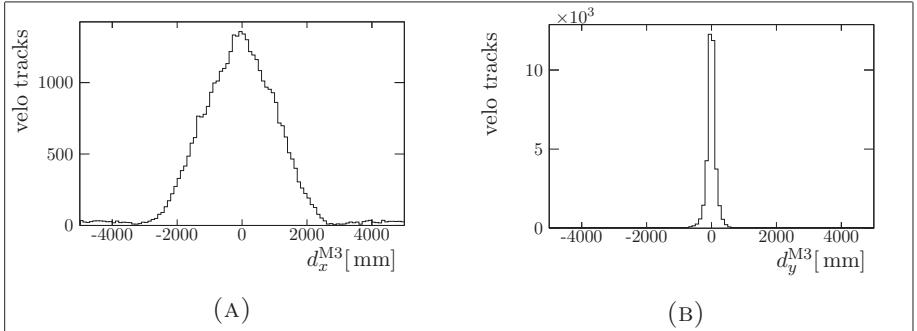


FIGURE 3.5: Hit pseudo-residuals in M3. The distributions correspond to the distance between the center of the FoI and the true M3 hit in the  $x$  (left) and  $y$  (right) planes respectively.

computation expensive calculation of its trajectory through the LHCb magnet, also referred to track *extrapolation*. In order to visualize the FoI sizes, simulated events are used to compute the distance,  $d^{\text{M3}}$ , of the true hit in M3 from the middle of the FoI, respectively in the  $x$  and  $y$  planes. The results are shown in Figure 3.5. The average FoI half-width is 2.7 m and 0.4 m respectively for the  $x$  and  $y$  planes. This implies that most of the times the true hit is inside the FoI. The outliers in the Figure 3.5a in with  $|d^{\text{M3}}| > 3$  m correspond to cases were the track charge estimate by the TT is wrong. Remember that charge information is used to look for hits only in the one side of the M3 station. Lastly the  $y$  plane is the non-bending plane hence the FoI sizes are smaller, compared to the ones for the  $x$  plane, since only multiple scattering is present in that plane.

### HLT1 timing and rate considerations

After all the improvements the `MatchVeloTTMuon` algorithm needs to be tuned to comply with the HLT1 output rate of 12 kHz mentioned in Section 3.2.1. Given the cutoff  $\chi^2/\text{nDoF}$  value of the upgraded `MatchVeloTTMuon` is chosen to be 2. Based on the signal and background distributions of Figure 3.4a and Figure 3.6a, the previous cutoff implies that; A VELO track which is accepted by the `MatchVeloTTMuon`

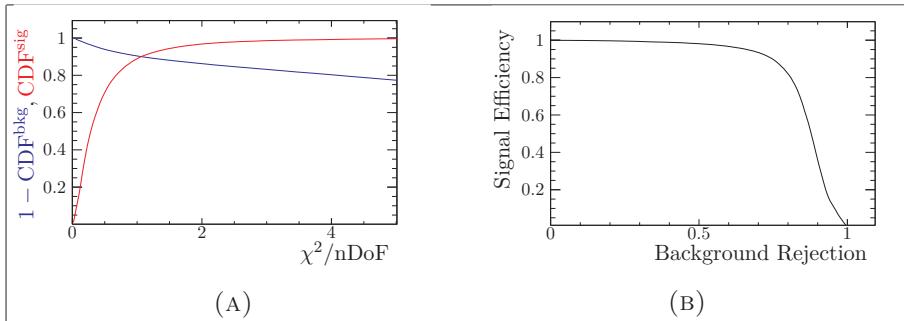


FIGURE 3.6: Tuning of the `MatchVeloTTMuon` algorithm. Left:  $\chi^2$  Cumulative Distribution Function, CDF, of matched (red) and not matched (blue) muon tracks (the blue curve is actually  $1 - \text{CDF}$ ). Right: ROC curve made from the previous  $\chi^2$  CDF. Only soft muon tracks are considered.

algorithm has 97% probability of being a real muon against originating from random combinations in the muon stations. Furthermore the ROC curve of Figure 3.6b is made from the CDF distributions of Figure 3.6a, illustrating the discriminating power of the  $\chi^2$  computed from the `MatchVeloTTMuon` algorithm. The overall `MatchVeloTTMuon` efficiency in the  $p_T < 0.5 \text{ GeV}/c$  region for the above cutoff value is  $\epsilon = 55 \pm 1.5\%$ . In addition it was checked that the changes introduced in the upgraded `MatchVeloTTMuon` algorithm do not increase the algorithm's running time.

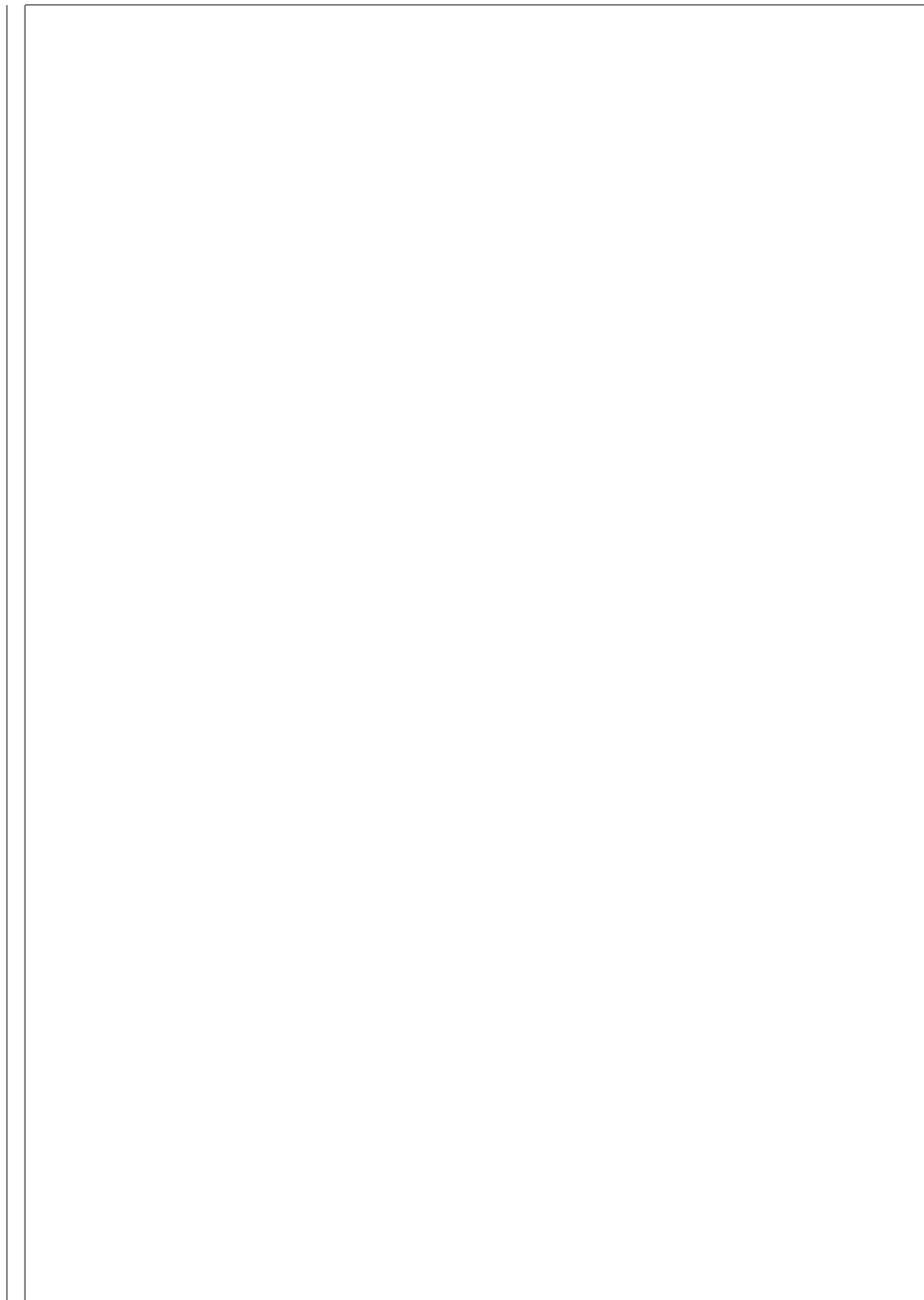
### Possible improvements

The most straightforward way to improve the matching algorithm in terms of efficiency is by relaxing the value of the  $\chi^2/\text{nDoF}$  cut-off value. On the other hand, doing so will increase the number of output tracks to the point that the HLT1 rate requirement of 12 kHz is not fulfilled. Alternatively, one might attempt to improve the  $\chi^2/\text{nDoF}$  discriminating variable itself, by finding a better parametrization for the magnet hit position and uncertainty. However the current parametrization seems to be quite good with respect to the one present in the old `MatchVeloMuon`

algorithm. Lastly, in case the momentum estimate of TT improves, for example the UT tracker that will replace the TT in the LHCb upgrade, then it is worth trying the FoI used in the Muon-ID code mentioned in Section 3.1. The previous FoIs are momentum dependent search windows, implying that smaller FoIs are more suitable for high momentum muons and *vice versa* for low momentum ones. This effectively reduces the number of M3 seed hits and thus increasing the purity of output tracks which might allow to relax the  $\chi^2/\text{nDoF}$  requirement.

---

---



---

## Chapter 4

# Data Analysis

In the current chapter the necessary ingredients to estimate penguin topology contributions to the  $B_s^0 \rightarrow J/\psi \phi$  decay, mentioned in Section 1.4.3, are obtained. This is achieved by analyzing Run 1 LHCb data, accumulated over the years 2011 and 2012, which correspond to a total integrated luminosity of  $3\text{ fb}^{-1}$ . To fit these data and extract the parameters of interest, a maximum likelihood fit is performed to the angular distributions of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decays. Prior to any analysis, the data obtained from the detector are filtered to reduce the presence of background. Thus the current chapter addresses the  $B_s^0$  candidate selection first and then the angular analysis. Section 4.1 describes the steps employed to reduce various sources of background in the data. In Section 4.2 the angular analysis is described. Results, treatment of systematic uncertainties as well as a validation of the fitting procedure are presented in Section 4.4.

### 4.1 Candidate Selection

Following the trigger system, addressed in Section 2.4, there are effectively four stages of offline event selection which take place after the data has been recorded. First a set of selection criteria applied is called *Stripping*. The purpose of stripping is to roughly select the triggered data according to their physics content. Furthermore, a loose rejection of the *combinatorial* background is performed such that the data are prepared for the next selection stages. Combinatorial background is the type of background typical in particle collision where a large number of

particles is produced. It is caused by random combinations of non-signal particles that build a signal *candidate* during the event reconstruction, mentioned in Section 2.1. Stripping is also a way to save disk space when permanently storing events offline. Each of the stripping selection criteria, also referred to as a *cut*, requires an observable to satisfy some condition. For example, the mass of the  $B_s^0$  candidate must be within a certain range. Or another type of cut could be the probability that a final state particle, like the pion, is misidentified as a muon. A detailed table of the Stripping selection is given in Table A.1. The remaining three selection steps after Stripping involve a multivariate based selection, Section 4.1.1, further reduction of the combinatorial background, Section 4.1.3 and specific background removal, Section 4.1.2.

Prior to addressing the various selection steps involved in detail, it is useful to introduce the *sPlot* technique, which is used to statistically subtract the combinatorial background further after the first candidate selection step of Section 4.1.1. The latter technique is a statistical tool to unfold data distributions [69], for example signal and background. A *Probability Density Function*, PDF, that is able to estimate the contribution-yield of signal and background, in that case, is required as an input. The *sPlot* technique uses information from the likelihood, built from the previous PDF, to assign a weight to each candidate. The weights are also called *sWeights* and are computed such that the contribution of the background component in data is effectively canceled. Thus a large *sWeight* value implies higher signal probability, for the particular candidate. In Section 4.1.3 the above mentioned PDF is described and a bit more details on the *sPlot* technique are provided.

### 4.1.1 Multivariate Based Selection

The  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$  signal yield out of the  $3 \text{ fb}^{-1}$  data-set is expected to be small, see Section 2.5. Thus, one would like to reject as much background as possible while keeping most of the signal. One way to do that would be to tighten the Stripping selection criteria one by one. Instead, a multivariate approach, hereafter MVA, is adopted. In that case a set of observables is combined by an MVA algorithm to produce a single variable, the *classifier variable*. The MVA approach can exploit the

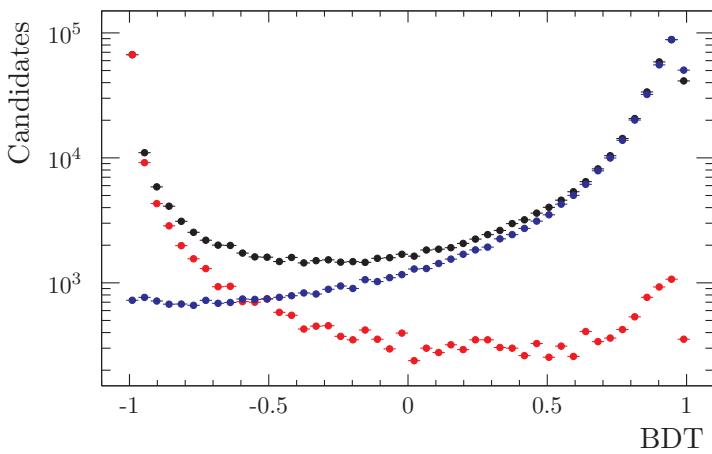


FIGURE 4.1: Data BDT distribution (black points). Signal (blue points) and background (red points) distributions come from simulated data and the  $B_s^0$  mass side-band control region, see text for details.

correlations between input variables to get the maximum discriminating power out of their combination. The classifying variable is constructed such that it ranges from  $-1$  to  $+1$ . Signal (background) candidates tend to accumulate near  $+1$  ( $-1$ ), see Figure 4.1.

In the current analysis the MVA algorithm is a *Boosted Decision Tree*, BDT in short, which is a class of machine learning algorithm. Further information on the BDT algorithm used can be found in Section 7.2 of [70]. For the current analysis the TMVA toolkit [70] is used for the MVA selection. In order to construct the classifying variable, a pair of data sets is needed. One data set must be a representative sample of the signal and one of the background. This pair of data sets is provided as input to the MVA algorithm to construct the classifying variable. In that step the MVA algorithm is *trained* to distinguish between the signal and background input data sets. After the training step is completed a second, independent, pair of sets is used to asses how successful the training step was. This step is called *testing*. For the signal representative samples,  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  Monte-Carlo (MC) simulated data

are used, hereafter simulated data. For the background representative sample data candidates are used. These candidates are taken from a control region away from the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  mass peak, usually referred to as *mass side-band*. This way the background representative sample is not affected by non-perfect simulation and it is also independent of the signal representative sample. Note that the simulated data are treated by all the processing steps, after the L0 trigger, in the same way as data.

The two well-separated peaks of the data BDT distribution, indicated with black points in Figure 4.1, suggest that the BDT variable achieves a good discrimination between signal and background candidates. The BDT was also checked for possible over-training. The latter is a situation where the BDT variable becomes sensitive to statistical fluctuations of the training samples. This means that statistically compatible distributions, such as the testing and training samples, appear to have significantly different BDT distributions. This is not a desirable situation which leads to bad BDT performance. Once the training and testing steps are complete, a cutoff value on the BDT is applied so that it maximizes the following figure of merit:

$$F(w_i) = \frac{(\sum w_i)^2}{\sum w_i^2}, \quad (4.1)$$

where  $w_i$  are weights associated to each candidate and computed with the *sPlot* technique, mentioned earlier in the introduction of the current section. The above-mentioned figure of merit, hereafter FoM, increases in value with the sum of *sWeights*, in the numerator. Large values of *sWeights* translate to higher effective signal yield. This is because candidates that are more likely to be signal receive a larger weight  $w_i$ . It follows that the optimum set of weights will have the higher FoM value compared to any other set of weights. Thus the FoM of Eq. 4.1 essentially expresses the signal yield. For a range of BDT values the *sPlot* technique provides a different set of *sWeights* based on which a different FoM value can be obtained, thus allowing for optimizing the BDT cutoff value.

### 4.1.2 Reflection Backgrounds

After applying the BDT cut there is still some combinatorial background remaining, see Figure 4.2, which is statistically subtracted further by means of the *sWeights* in Section 4.1.3. However, there is one more crucial step that is necessary to address prior to the application of the *sPlot* technique, which is the treatment of *reflection* backgrounds. This type of background is due to particle misidentification. Specifically a final state particle is assigned the wrong mass hypothesis. As a result the invariant mass of the  $B_s^0$  candidate, is shifted with respect to  $B_s^0$  candidates where all the final state particles were correctly identified. Studies of simulated samples show contributions from several specific background sources, such as  $B_s^0 \rightarrow J/\psi K^+ K^-$ ,  $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$  and  $B^0 \rightarrow J/\psi \pi^+ \pi^-$ . These backgrounds are the result of misidentifying a final state particle, for example a pion can be confused for a kaon. The invariant mass distribution of the above-mentioned misidentified backgrounds accumulate near the  $B_{(s)}^0 \rightarrow J/\psi K\pi$  signal peaks. Furthermore given the low number of expected reflection background, see Table 4.1, their distribution is quite not smooth and continuous. This type of background is removed before the next selection step is applied as it is more effective when the distributions involved are smooth. Therefore, specific simulated decays with negative weights are injected to the data sample such that they cancel out the contribution of the above-mentioned reflection background. The term "cancel out" obtains its exact meaning in the context of the likelihood which is built in Section 4.1.3 based on the mass probability density function. The current section addresses the treatment of the reflection background with negative weights as well as the special treatment of the  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  background.

As a first step towards removing the above-mentioned type of background contribution it is necessary to have an estimate on the expected reflection background yield on data. This is done using simulated data and based on the expression:

$$N_{\text{exp}} = 2 \times \sigma_{b\bar{b}} \times f_q \times \mathcal{B} \times \varepsilon \times \mathcal{L}, \quad (4.2)$$

where  $\sigma_{b\bar{b}}$  is the cross section for the production of a pair of bottom quarks,  $f_q$  is the hadronization fraction (probability that the  $b$  quark

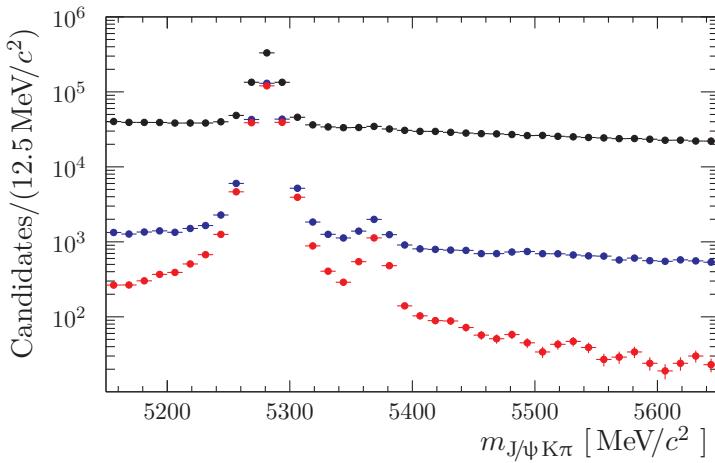


FIGURE 4.2: Mass distribution after stripping selection (black points), before (blue points) and after (red points) BDT selection. The peaks around 5300 and 5360  $\text{MeV}/c^2$  correspond to the  $B^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay channels respectively.

forms a  $B_q$  meson, where  $q = d, s$ ),  $\varepsilon$  is the total efficiency (including reconstruction, selection and trigger) and  $\mathcal{L}$  is the integrated luminosity of the data. Finally  $\mathcal{B}$  stands for the branching fraction of the particular background of Table 4.1. Since simulated data are used it is necessary to determine the effective integrated luminosity in the simulated sample and adjust it such that it matches the estimated integrated luminosity of the data. After that, the fraction of reflection background from the simulated sample is a valid estimate of the reflection background fraction in data (assuming same total efficiency between the samples).

The second and last step of the reflection background removal is to apply an angular correction factor to account for the fact that simulated decays are distributed uniformly in the  $(\Omega, m_{K\pi})$  phase space and hence do not represent a proper decay amplitude structure. This can cause the reflection background yield estimations to be imperfect because the simulated events are distributed in the previous phase space in a different way than the actual reflection background in the data. The symbol

Sample	$\pm 70 \text{ MeV}/c^2$ window	
$B^0 \rightarrow J/\psi \pi^+ \pi^-$	2011	$51 \pm 10$
	2012	$115 \pm 23$
$B_s^0 \rightarrow J/\psi \pi^+ \pi^-$	2011	$9.3 \pm 2.1$
	2012	$25.0 \pm 5.4$
$B_s^0 \rightarrow J/\psi K^+ K^-$	2011	$10.1 \pm 2.3$
	2012	$19.2 \pm 4.0$
$\Lambda_b^0 \rightarrow J/\psi p K^-$	2011	$36 \pm 17$
	2012	$90 \pm 43$
$\Lambda_b^0 \rightarrow J/\psi p \pi^-$	2011	$13.8 \pm 5.3$
	2012	$27.3 \pm 9.0$

TABLE 4.1: Estimated reflection yields in a  $\pm 70 \text{ MeV}/c^2$   $m_{K\pi}$  window of each background specie after weighting for the angular distributions, with the exception of the  $\Lambda_b^0 \rightarrow J/\psi p \pi^-$  as explained at the end of Section 4.1.2.

$\Omega$  represents the angles that describe the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay, see Section 4.2.1 for further details. The amplitude analyses of  $B^0 \rightarrow J/\psi \pi^+ \pi^-$  [71],  $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$  [72],  $B_s^0 \rightarrow J/\psi K^+ K^-$  [73] and  $\Lambda_b^0 \rightarrow J/\psi p K^-$  [74] provide the necessary knowledge to compute the correction factors:

$$w_{\text{MC}}^i = \frac{P_{\text{DATA}}^i(\Omega, m_{\text{hh}} | A_k^i)}{P_{\text{MC}}^i(\Omega, m_{\text{hh}})}, \quad (4.3)$$

where  $P_{\text{DATA}}$  and  $P_{\text{MC}}$  are normalized PDFs. The index  $i$  labels the species of the particular reflection background, *i.e.*  $B^0 \rightarrow J/\psi \pi^+ \pi^-$ ,  $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ ,  $B_s^0 \rightarrow J/\psi K^+ K^-$ ,  $\Lambda_b^0 \rightarrow J/\psi p K^-$ . The polarization states of each of the above species is labeled by the  $k = (0, \parallel, \perp, S, D)$  index. Where the first three indices represent the P-wave and the last two the S-wave and D-wave respectively, see Section 4.2.1 for more details. The estimated yields of reflection background after the above-mentioned steps are shown in Table 4.1

As it was previously mentioned, the  $\Lambda_b^0 \rightarrow J/\psi p \pi^-$  background is treated differently. Instead of following the above procedure, the  $\Lambda_b^0 \rightarrow J/\psi p \pi^-$  invariant mass line shape is modeled and statistically subtracted in the same way as the combinatorial background in the next

section. The reason for this different treatment with respect to other reflection backgrounds is due the  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  amplitude structure [75], which was not sufficiently known when the analysis was concluded. Thus weighting of the simulated samples was not possible. Since the shape of the misidentified  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  decays in the  $J/\psi K\pi$  mass spectrum is broad enough, see Figure 4.3, the *sPlot* technique can be used to subtract this specific reflection background.

### 4.1.3 sWeighting and Invariant Mass Distribution

After injecting candidates with negative weights to cancel reflection background contributions, the data sample is effectively composed of,  $B^0 \rightarrow J/\psi K^+\pi^-$ ,  $B_s^0 \rightarrow J/\psi K^-\pi^+$ ,  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$ , and combinatorial background not removed by the multivariate based selection. These four modes are statistically disentangled using the *sPlot* technique [69]. The technique requires as an input a PDF for each of the above species. The PDF has to be *extended*, which means that it is possible to estimate the yield of each species by means of an extended maximum likelihood fit. The current analysis uses the  $J/\psi K\pi$  invariant mass PDF, hereafter just mass, to perform the extended maximum likelihood fit to the mass distribution of data, hereafter mass fit. It is important to point out that the *sPlot* technique assumes no correlation between the mass and other variables. A systematic uncertainty corresponding to the validity of this assumption will be assigned later.

#### The mass PDF

Before discussing the mass fit description it is useful to briefly describe the PDFs used for each of the following individual species:

- Combinatorial background: Exponential distribution.
- $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  decays: Amoroso distribution [76]
- $B^0$  and  $B_s^0$  signals: Hypatia distribution [77].

An exponential distribution,  $e^{-km}$ , is found to be an appropriate description of the combinatorial background. It accounts for cases where

random combinations of final state particles are combined to a signal candidate. Such candidates are not expected to exhibit any structure in the mass distribution, they simply follow an exponential distribution with a negative slope. The negative slope can be understood when considering two main effects. First, low invariant mass implies a low momentum of the particles ( $K, \pi, \mu$ ) used to reconstruct the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  candidate. Second, there are more low momentum particles produced in a  $p - p$  collision compared to high momentum ones. This is because elastic-soft strong interactions between colliding *partons*, quarks and gluons contained inside the proton, are more likely to happen compared to central, head-on, ones.

As for the  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  background, the Amoroso distribution was found to provide good description of the data. The parameters of the distribution are obtained from simulated data and then fixed in the mass fit to the data. The first two parameters are the mean and width of the distribution, whereas the other two are related to the shape. The Amoroso distribution describes an entire family of distributions: Depending on the values of the shape parameters the resulting distribution can vary from exponential to a Gaussian distribution. This is a powerful feature that is exploited in order to describe the  $\Lambda_b^0 \rightarrow J/\psi p\pi^-$  reflection background shape.

The Hypatia distribution has two main features which makes it a suitable description of  $B_s^0$  and  $B^0$  signal invariant mass distributions. First, a hyperbolic core is found to describe the peak of the previous distributions adequately. Second, invariant mass resolution effects are taken into account by marginalizing over an approximate  $m_{J/\psi K\pi}$  mass uncertainty distribution. Overall the Hypatia function adequately models the tails of the  $B_s^0$  and  $B^0$  invariant mass distribution which is of crucial importance. This is because the previous invariant mass distributions are sufficiently close, see Figure 4.3, such that bad modeling of the tails implies that a certain number of  $B^0$  decays are associated to  $B_s^0$  and *vice versa*. Several effects contribute to the structure of the tails apart from mass resolution of the detector. These effects can be radiative tails (a charged final state particle radiates a photon), interplay of radiative tail with  $J/\psi$  mass constraint (applied when building the  $B_s^0$  candidate) or badly reconstructed candidates caused by decays of the

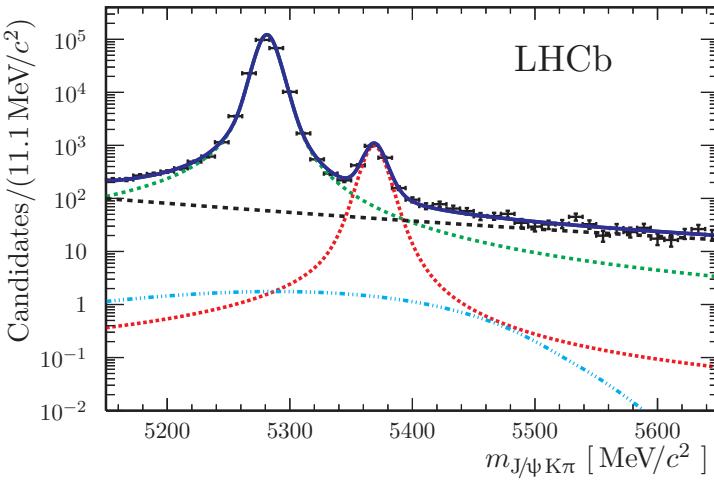


FIGURE 4.3: Invariant mass fit to the data, shown with black points. The green and red curves correspond to the  $B^0 \rightarrow J/\psi K^+ \pi^-$  and  $B_s^0 \rightarrow J/\psi K^- \pi^+$  mass PDF components. The combinatorial and  $\Lambda_b^0 \rightarrow J/\psi p \pi^-$  background components are shown with black and cyan colors respectively. The total mass PDF, blue curve, is overlaid on top of the data. Figure from [2].

final state hadrons, see [77]. The tail parameters  $(\alpha, n)$  are four in total (two for each tail) and are taken from a fit to MC simulated events with a known resolution. This makes sure that the tail parameters do not rely on detector simulation imperfections. For the core of invariant mass distribution, Hypatia requires five shape parameters, namely  $\zeta, \beta, \lambda, \sigma, \mu$ . The first two are set to zero since  $\zeta$  is empirically found to be very small whereas  $\beta = 0$  implies that the core is symmetric left and right with respect to the mean. The third parameter is taken from the previous simulated sample along with the tail parameters. The latter, in the limit of  $\zeta = 0$ ,  $\lambda$  does not depend on detector effects but only on particle kinematics, the same way as the tail parameters do. Lastly, the remaining two parameters, the width and the mean of the core, are determined by the mass fit to the data.

### The mass fit

From studies with simulated data, some of the  $B_s^0$  and  $B^0$  Hypatia parameters appear to be significantly correlated with the  $m_{K\pi}$  invariant mass. Since these parameters need to be fixed in the mass fit, the latter is performed in intervals of the  $m_{K\pi}$  invariant mass. In addition, due to correlations between the mass and one of the variables,  $\cos\theta_\mu$ , used in the subsequent angular fit, the requirements of the *sPlot* technique are not satisfied, see Section 4.1.3, and thus it cannot be applied directly. Therefore, each  $m_{K\pi}$  invariant mass sub-sample is divided further in intervals of  $\cos\theta_\mu$  where the *sPlot* technique can be applied. The final fit to the data is shown in Figure 4.3, and the overall  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B^0 \rightarrow J/\psi K^{*0}$  yields are:

$$N_{B^0} = 208656 \pm 462^{+78}_{-76}, \quad (4.4)$$

$$N_{B_s^0} = 1808 \pm 51^{+38}_{-33}, \quad (4.5)$$

where the first uncertainties are statistical, obtained from the quadratic sum of the ones in each fitting category, and the second are systematical uncertainties. The correlations between the  $B^0$  and  $B_s^0$  yields in each fitting category are found to be less than 4%. Note that the mass fitting model has been validated with a procedure commonly called *pseudo-experiments*. This procedure checks if the fitting model introduces a bias to the parameters of interest. Further details can be found in Section 4.4.4 where a pseudo-experiment study is performed for the angular fitting model of Section 4.2.6.

### sWeighting

After the mass fit is performed, all the necessary ingredients to remove the remaining background are in place. The fitted mass PDF is now given to the *sPlot* algorithm. The algorithm assigns a weight to each candidate based on the likelihood function built from the input PDF. The main idea of *sPlot* is that for each of the species of the total PDF a set of weights can be computed such that they project out only that particular species while effectively subtracting the others. Based on this

technique,  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  candidates are selected out of the full data sample, since this is the decay of interest.

The advantage of the *sWeights* approach becomes apparent when performing the angular analysis, described in the subsequent section. Specifically, the resulting weighted candidates can be described by the signal PDF only, meaning there is no need to explicitly model the background angular distributions. As a result the fit itself is faster and simpler to implement. On the other hand there are a few disadvantages such as the evaluation of systematic uncertainties on the estimated angular parameters due to the mass PDF modeling: Any variation of these models implies re-computing the *sWeights*. Also, care must be taken to ensure that the uncertainties of the fitted angular parameters are properly estimated [69]. The uncertainties tend to be underestimated in a fit with weighted candidates. This is because the per candidate weighs in a weighted likelihood fit change the shape of the likelihood function, which is where the statistical uncertainty estimation comes from. Thus to correct for this distortion the scale factor:

$$\alpha = \left( \frac{\sum_i w_i}{\sum_i w_i^2} \right)^{1/2}, \quad (4.6)$$

is applied to each weight before performing the angular fit. The scale factor essentially ensures that the sum of weights equals the number of candidates. A detailed derivation of  $\alpha$  can be found in [31].

## 4.2 Angular Analysis

The core formalism and all the necessary steps towards building a PDF to describe the amplitude properties and angular structure of  $B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0}$  decays are presented in the current section. A brief description of the angular PDF derivation is given in Section 4.2.1. Treatment of the angular acceptance is addressed in Section 4.2.2 and Section 4.2.3. In Section 4.2.4 the implications arising from the  $m_{K\pi}$  dependence are discussed. The effect of asymmetries introduced by detector imperfections such as production and detection asymmetries are dealt with in Section 4.2.5. Some aspects of the maximum likelihood fit are presented

in Section 4.2.6, along with the  $\mathcal{A}_{CP}^k$  parameters of interest, which are introduced in a special way in the angular fit.

### 4.2.1 Angular Dependence

The decay of interest,  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ , is a  $P \rightarrow VV$  process with 4 charged particles in the final state. The symbols in  $P \rightarrow VV$  refers to the spin of the particles involved in the decay. Specifically, the  $B_s^0$  (and  $B^0$ ) is a spin zero parity minus particle (or pseudo-scalar), whereas the intermediate resonances  $J/\psi$  and  $\bar{K}^{*0}$  are spin one (or vector) particles. The final state particles are two muons,  $\mu$ , and a kaon-pion pair,  $K\pi$ , coming from the decay of the  $J/\psi$  and  $\bar{K}^{*0}$  respectively. The angular dependence of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay can be described using either the transversity framework [78, 79] or the helicity formalism [80, 81]. For the current analysis the helicity formalism is adopted. This implies that the angular dependence is introduced by summing all possible spin configurations of the intermediate vector particles relative to their momentum direction in the  $B_s^0$  rest mass frame and squaring the sum. Or, in more compact wording, from summing up all possible helicity configurations of the intermediate vector particles. A detailed derivation can be found in [82] and [31] respectively for the transversity and helicity formalism.

Here, the steps to arrive at an expression for the angular decay rate of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  are briefly described. The total decay amplitude  $\mathcal{A}(B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0})$  is decomposed in its P-wave ( $0, \parallel, \perp$ ) and S-wave ( $S$ ) polarizations as shown in the following equation:

$$\begin{aligned} \frac{d\Gamma(B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0})}{d\Omega dm_{K\pi}} &\propto |\mathcal{A}(B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0})|^2 = \\ &= \left| \sum_i^{0, \parallel, \perp, S} A_i(\Omega, m_{K\pi}) \right|^2 \propto \sum_n a_n h_n M_n. \end{aligned} \quad (4.7)$$

Note that the S-wave is not a  $P \rightarrow VV$  process since no intermediate  $K^{*0}$  resonance is produced. The terms  $a_n$  originate from squaring the amplitude and their exact expressions are shown in Table 4.2. Note the appearance of the  $\Re$  and  $\Im$  interference terms as well. The full

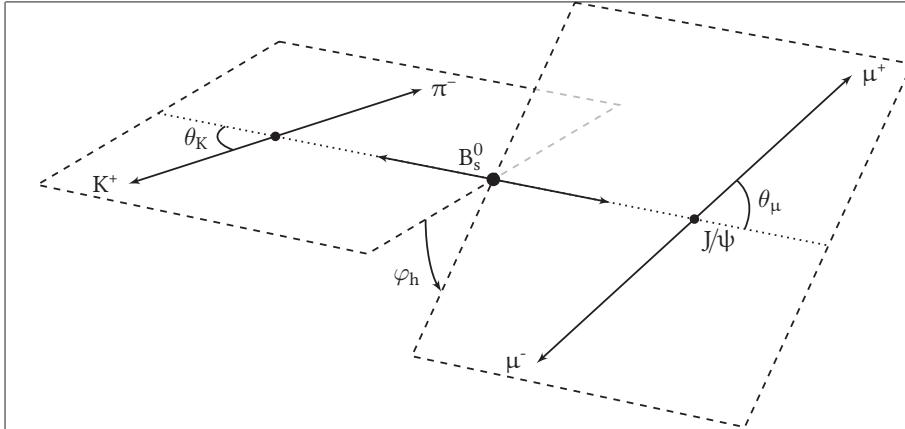


FIGURE 4.4: Definition of helicity angles.

derivation of the above table can be found in [31]. The functions  $h_n$  and  $M_n$  represent the angular dependence of each  $a_n$  term and the  $m_{K\pi}$  dependence of the amplitude respectively. The description of the  $m_{K\pi}$  dependence is postponed for the next section. Table 4.2 lists the angular part of the decay rate equation, *i.e.* the  $h_n$  functions, which are the result of applying the helicity formalism in the  $B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0}$  decay. In order to describe the 4-body decay  $B_{(s)}^0 \rightarrow J/\psi \bar{K}^{*0}$ , three angles, denoted by  $\Omega$ , plus  $m_{K\pi}$  are required. As shown in Figure 4.4, the three *helicity angles* angles  $\theta_K$ ,  $\theta_\mu$  and  $\varphi_h$ , are defined as follows:  $\theta_\mu$  is the angle of the positive  $\mu$  with respect to the momentum of the  $J/\psi$  in the  $B_s^0$  rest frame; similarly for  $\theta_K$ , the  $K$  is used to define the angle with respect to the momentum of the intermediate  $K\pi$  resonance; finally  $\varphi_h$  is the relative angle between the  $K\pi$  and dimuon decay planes, where each plane is defined in the rest frame of the corresponding intermediate resonance.

Lastly, the orthogonal angular basis of associated Legendre polynomials,  $P_n^m$ , and real-valued spherical harmonics,  $Y_{ij}$ , is adopted in the current analysis. This choice is favored by the fact that the integrals of the product of either two  $P_n^m$  or two  $Y_{ij}$  are known analytically. The latter significantly improves the description and implementation of the angular acceptance, see next section. Note that the decomposition of angular functions in an orthogonal basis follows naturally since spherical

$a_n$	$h_n(\Omega) \times 16\sqrt{\pi}$
$ A_0 ^2$	$4(P_0^0 + 2P_2^0)(Y_{0,0} - \frac{1}{\sqrt{5}}Y_{2,0})$
$ A_{\parallel} ^2$	$P_2^2(2Y_{0,0} + \frac{1}{\sqrt{5}}Y_{2,0} - \sqrt{\frac{3}{5}}Y_{2,+2})$
$ A_{\perp} ^2$	$P_2^2(2Y_{0,0} + \frac{1}{\sqrt{5}}Y_{2,0} + \sqrt{\frac{3}{5}}Y_{2,+2})$
$\Re(A_0^*A_{\parallel})$	$+2\sqrt{2}\sqrt{\frac{3}{5}}P_2^1Y_{2,+1}$
$\Im(A_0^*A_{\perp})$	$-2\sqrt{2}\sqrt{\frac{3}{5}}P_2^1Y_{2,-1}$
$\Im(A_{\parallel}^*A_{\perp})$	$+2\sqrt{\frac{3}{5}}P_2^2Y_{2,-2}$
$ A_S ^2$	$4P_0^0(Y_{0,0} - \frac{1}{\sqrt{5}}Y_{2,0})$
$\Re(A_0^*A_S)$	$8\sqrt{3}P_1^0(Y_{0,0} - \frac{1}{\sqrt{5}}Y_{2,0})$
$\Re(A_{\parallel}^*A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}P_1^1Y_{2,+1}$
$\Im(A_{\perp}^*A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}P_1^1Y_{2,-1}$

TABLE 4.2: Angular functions corresponding to each term in Eq. 4.7. Pure and interference P-wave terms are shown in the upper part, whereas the S-wave plus SP-wave interference in the lower. The angular functions,  $h_n$ , are expressed in the helicity basis. The  $P$  and  $Y$  denote associated Legendre polynomials and real valued spherical harmonics respectively.

harmonics can be expressed as Wigner- $D$  matrices.

### 4.2.2 Acceptance

The effects of angular resolution are known to be small for this analysis [83], however, sources introducing angular acceptance need to be explicitly accounted for. Deviations from a perfect angular acceptance can be introduced by selection criteria mentioned in Section 4.1 or by the detector itself. As explained in Section 2 the LHCb detector does not cover the full  $4\pi$  solid angle. The current section introduces the treatment of the angular acceptance function. After that the parametrization of the acceptance is discussed and finally an important issue related to the choice of parametrization is dealt with.

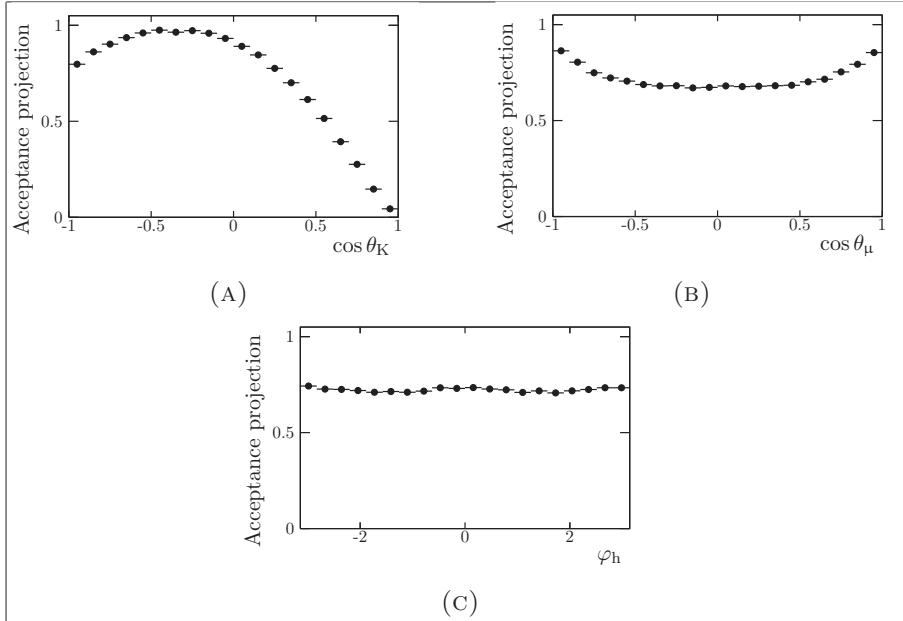


FIGURE 4.5: Angular acceptance shape from simulated data. Each candidate is assigned a weight. Each weight is equal to one over the integral of the PDF used to generate the angular distribution of the simulated data.

The shape of the acceptance can be seen in Figure 4.5. The most striking feature is the drop of efficiency close to  $\cos \theta_K = 1$ . The latter is mainly caused by one of the candidate selection requirements of Table A.1, specifically the requirement on the K and  $\pi$  transverse momentum,  $p_T$ , to be larger than 500 MeV/c. Before attempting to explain the acceptance shape it is useful to illustrate how the decay angles are distributed according to the PDF of Eq. 4.7, *i.e.* without any acceptance effect, see Figure 4.6.

Understanding the shape of the acceptance in the  $\cos \theta_K$  projection begins with the observation that the phase space region of  $(p_{T,\pi} - \cos \theta_K)$  is not uniformly populated. This is related to kinematic constraints, explained in the next paragraph. Qualitatively speaking, it turns out that

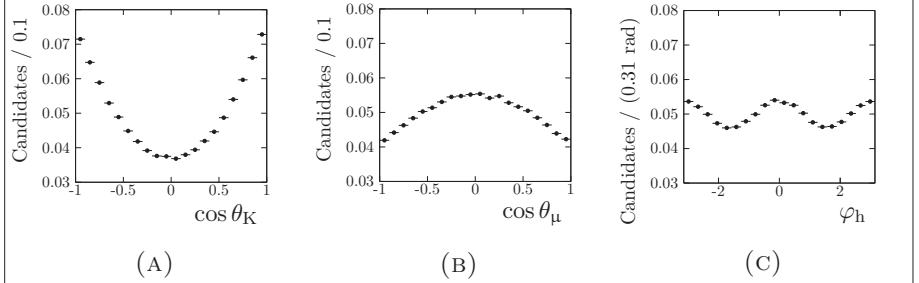


FIGURE 4.6: Decay angles distributions from simulation data. No detector effects and no selection are included. Histograms are normalized to unity area.

$B_s^0$  decays tend to cluster close to the region of  $(p_{T,\pi} - \cos \theta_K|_{-1})$  compared to the  $(p_{T,\pi} - \cos \theta_K|_{+1})$  region. This picture is less pronounced in the case of  $(p_{T,K} - \cos \theta_K)$  phase space region, where the  $p_{T,K}$  spectrum is harder and the  $(p_{T,K} - \cos \theta_K)$  is more symmetric with respect to  $\cos \theta_K$ , see Figure 4.7. Given these correlations it follows that applying the  $p_T$  requirement on both K and  $\pi$  makes the  $\cos \theta_K$  distribution asymmetric around  $\cos \theta_K = 0$ .

The above is related to kinematic constrains in two body decays particularly close to phase space boundaries. The latter constraint manifests itself when computing the angle of the pion with respect to the boost direction in the lab frame, as shown in the next equation:

$$\tan \theta_\pi^{\text{lab}} = \frac{p_{T,\pi}^{\text{lab}}}{p_{L,\pi}^{\text{lab}}} = \frac{p_{T,\pi}}{\gamma (p_{L,\pi} + \beta E_\pi)} = \frac{p_\pi \sin \theta_\pi}{\gamma (p_\pi \cos \theta_\pi + \beta E_\pi)}, \quad (4.8)$$

where,  $\theta_\pi^{\text{lab}}$  is the angle of the pion momentum vector with respect to the boost direction in the lab frame. Absence of the lab superscript refers to quantities computed in the  $K^{*0}$  rest frame. The subscripts T,L label the transverse and longitudinal components of the pion momentum,  $p_\pi$ . The symbol  $E$  stands for the energy of the pions. The Lorentz boost factor and the relativistic speed are labeled by  $\gamma$  and  $\beta$  respectively. The angle  $\theta_\pi$  is the angle of the pion momentum, in the  $K^{*0}$  rest frame,

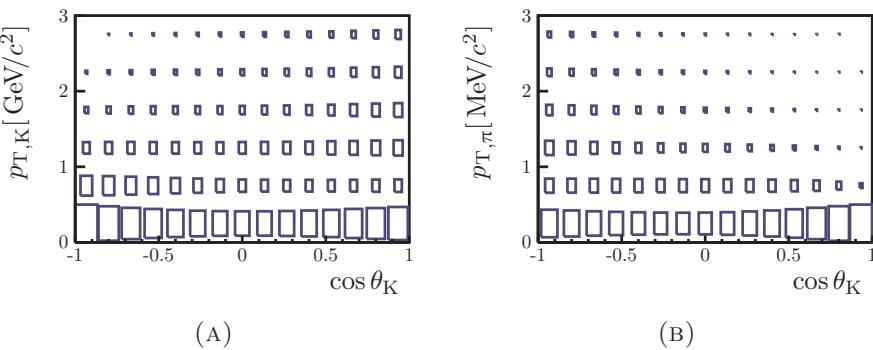


FIGURE 4.7: Two dimensional histograms of kaon (left) and pion (right)  $p_T$  in the lab frame versus  $\cos \theta_K$ . No detector effects and no selection are included. Box area reflects number of entries.

which is related to  $\theta_K$  in Figure 4.4 as:  $\theta_\pi = \theta_K + \pi$ . Now, the origin of the above-mentioned asymmetry is understood better in the limit of a massless pion, in which case Eq. 4.8 reduces to:

$$\tan \theta_\pi^{\text{lab}} = \frac{\sin \theta_\pi}{\gamma (\cos \theta_\pi + 1)}. \quad (4.9)$$

Note that the energy of a massless particle is equal to its total momentum, which implies that the Lorentz factor  $\gamma$  is infinite. Also  $\beta \simeq 1$  for pions coming from  $B_{(s)}^0 \rightarrow J/\psi K\pi$  decays in LHCb. Thus, Eq. 4.9 implies that decays where  $\cos \theta_\pi = -1$  are not allowed at all in the massless pion limit since the denominator in the same equation becomes undefined. The situation corresponds to a low transverse momentum pion for which  $\cos \theta_K = -\cos \theta_\pi = +1$ . This explains why the shape of the  $\cos \theta_K$  distribution becomes asymmetric when excluding low momenta pions. The effect is general and it is driven by the mass difference between the two particles in a two body decay.

Coming now to the case of the  $\cos \theta_\mu$  acceptance shape, the two decay products of the  $J/\psi$  have the same mass and thus get the same momentum fraction out of the mother  $J/\psi$  particle. A consequence of the latter is that the transverse momentum spectrum of the two muons

is identical, excluding detector resolution, and thus the two muons populate the  $(p_T - \cos \theta_\mu)$  plane similarly. Based on the latter it follows that excluding low transverse momentum muons does not introduce any asymmetry in the  $\cos \theta_\mu$  distribution.

As for the acceptance in the  $\varphi_h$  plane it is very close to being uniform. No strong acceptance effect in that projection is expected.

### Acceptance parametrization

Regarding the angular acceptance parametrization, there are at least two ways to proceed, namely the *efficiency moments* [31] and the *normalization weights* [31, 83]. The current analysis implements the first one. The efficiency function is thus parameterized using orthogonal polynomials, which, when combined with the angular description formalism in the previous subsection, allows for fast analytic integration.

$$\epsilon(\Omega) = \sum_{ijk} c_{jk}^i P_i^0(\cos \theta_K) Y_{jk}(\cos \theta_\mu, \varphi_h) \equiv c_{jk}^i \alpha_i^{jk}(\Omega), \quad (4.10)$$

Such a parametrization is shown in the previous equation where the real-valued spherical harmonics ( $Y$ ) and associated Legendre polynomials ( $P$ ) are used as the basis functions to which the angular acceptance is decomposed by means of the coefficients  $c_{jk}^i$ , hereafter efficiency moments or simply moments. Following the parametrization of Eq. 4.10 one just needs to obtain a set of  $c_{jk}^i$ . Note that there are *a priori* infinite efficiency moments since the sum in Eq. 4.10 runs through all allowed values of the  $ijk$  indices. However, in practice only a small number of efficiency moments contribute significantly to the angular acceptance shape. This is due to the fact that higher moments introduce variations on the acceptance shape that are effectively washed out by the finite angular resolution.

The efficiency moments used to describe the acceptance shape are shown in the first column of Table 4.3. The larger the value of a moment, the more the corresponding function contributes to the overall efficiency function. The efficiency functions  $P_1^0 Y_{00}$  and  $P_2^0 Y_{00}$  account for most of the hard acceptance drop in the  $\cos \theta_K$  projection. This is supported by

moment	central value	standard deviation
$c_{00}^0$	+3.7454	0.0023
$c_{00}^1$	-2.0982	0.0084
$c_{00}^2$	-1.9116	0.0126
$c_{00}^3$	-0.0317	0.0151
$c_{00}^4$	+0.1146	0.0175
$c_{00}^5$	+0.1384	0.0193
$c_{20}^0$	+0.2820	0.0066
$c_{22}^0$	+0.0612	0.0058
$c_{40}^0$	+0.0864	0.0066
$c_{20}^1$	-0.1900	0.0115
$c_{40}^1$	-0.0720	0.0118
$c_{20}^2$	-0.1315	0.0161
$c_{22}^2$	-0.0674	0.0117

TABLE 4.3: Efficiency moments of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ .

the large value of the corresponding moments  $c_{00}^1$  and  $c_{00}^2$ . The choice of the particular efficiency moments, is connected to an important issue with the acceptance structure close to  $\cos \theta_K = 1$  and it is addressed at the end of the current section.

In order for the efficiency moments  $c_{jk}^i$  to be computed a definition of the efficiency is necessary. Note that this computation relies on simulated data where all the generating conditions are known. The efficiency is defined as the ratio of two angular PDFs, namely  $P^{\text{obs}}(\Omega)$  and  $P^{\text{gen}}(\Omega)$ . The first is the angular PDF that the simulated data follow after all stages of selection, whereas the latter is the angular PDF with which the simulated data has been generated. It follows that the efficiency is defined as:

$$P^{\text{obs}}(\Omega) = \epsilon(\Omega) P^{\text{gen}}(\Omega) \quad \Rightarrow \quad \epsilon(\Omega) \equiv \frac{P^{\text{obs}}(\Omega)}{P^{\text{gen}}(\Omega)}. \quad (4.11)$$

Computing the efficiency moments is done by projecting the efficiency

as defined in Eq. 4.11 on the orthogonal  $PY$  basis. This projection is shown in the following equation:

$$\begin{aligned} c_{jk}^i &\equiv \left( j + \frac{1}{2} \right) \int d\Omega \alpha_i^{jk}(\Omega) \epsilon(\Omega) \\ &= \left( j + \frac{1}{2} \right) \int d\Omega P^{\text{obs}}(\Omega) \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)}. \end{aligned} \quad (4.12)$$

The factor  $j + \frac{1}{2}$  is used to make it such that the associated Legendre polynomials form an orthonormal basis; implying that the integral of the product of two associated Legendre polynomials is equal to the *Kronecker delta* symbol. The latter is required to properly normalize the efficiency moments. The final step in the calculation of the efficiency moments is the computation of the integral in Eq. 4.12. This is done by means of simulated data based on the concept of Monte Carlo integration. This implies that the integral in Eq. 4.12 can be estimated by a sum over the simulated data (which they have passed through the exact same selection steps as data) as:

$$E \left[ \int d\Omega \alpha_i^{jk}(\Omega) \frac{P^{\text{obs}}(\Omega)}{P^{\text{gen}}(\Omega)} \right] = \frac{1}{N^{\text{obs}}} \sum_e^{N^{\text{obs}}} \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)}. \quad (4.13)$$

Combining Eq. 4.12 and Eq. 4.13 the final expression for computing the efficiency moments is shown in Eq. 4.14 and the result of the computations can be found in table Table 4.3.

$$c_{jk}^i \equiv (j + \frac{1}{2}) \frac{1}{N^{\text{obs}}} \sum_e^{N^{\text{obs}}} \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)}. \quad (4.14)$$

Acceptance effects in the  $\cos \theta_\mu$  projection are less pronounced compared to the one  $\cos \theta_K$ . This can be deduced from Figure 4.5b as well as from the values of moments like  $c_{jk}^0$  compared to the values of  $c_{00}^1$  and  $c_{00}^2$ . Lastly, the efficiency moments are correlated and the corresponding correlation matrix can be found in Table A.7.

### Choice of efficiency moments

The choice of efficiency moments has to do with the structure of the acceptance at  $\cos \theta_K = 1$ . At that point the acceptance goes to zero which is not a problem by itself. However, within the parametrization of the acceptance function, Eq. 4.10 may become negative. In other words the method of efficiency moments does not guarantee that the acceptance function is a positive definite quantity. Note that this is an artifact of the parametrization. The problem is mathematical in nature and there is no standard solution (known to the author when these lines were written). The approach followed to solve this problem has two steps: First a constraint on the efficiency function, Eq. 4.10, is introduced to force it to be positive definite and second, a fit to the efficiency moments is performed to optimize the acceptance shape description. The result is a hybrid acceptance function that is by construction positive definite.

Before laying out the formulas behind the above-mentioned steps it is useful to visualize the problem caused by the parametrization, see Figure 4.8a and the corresponding caption. Subsequently, the most natural thing to do before making any effort to force the value of Eq. 4.10 positive at  $\cos \theta_K = 1$  is to make the best possible choice of efficiency moments. The moments are chosen such that the efficiency function becomes negative at the largest possible value of  $\cos \theta_K$ . In other words the crossing point of the red curve with the horizontal axis in Figure 4.8a, has to be as close to  $\cos \theta_K = 1$  as possible.

It is actually not trivial to find a good choice of efficiency moments. Two things are relevant when trying to describe a distribution with orthogonal functions like  $P_n^m$  and  $Y_{ij}$ . One is the shape of the orthogonal function itself and two the sign and size of the calculated moment. For example the efficiency term  $c_{00}^1 P_1 Y_{00}$  is necessary to describe the drop of efficiency in  $\cos \theta_K$  but the size of  $c_{00}^1$  is too large making the total efficiency function negative. Thus one could add the term  $P_3 Y_{00} c_{00}^3$  where  $c_{00}^3$  is positive (see Table 4.3 ) to counter the large negative value of  $c_{00}^1$  but unfortunately it is not enough. Things get more complicated when trying to choose efficiency moments that affect the acceptance shape in all three dimensions at the same time.

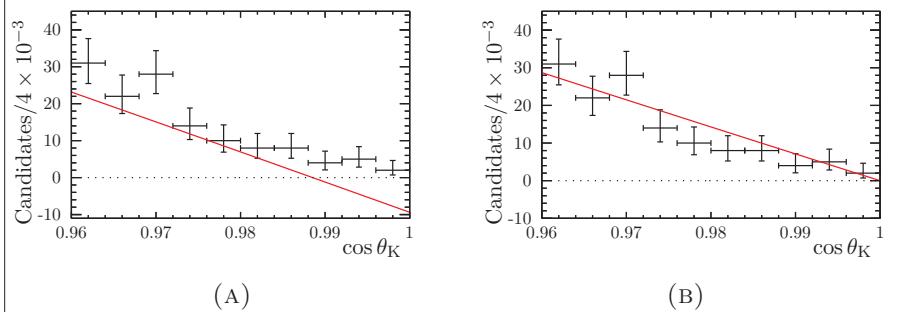


FIGURE 4.8: PDF projection of  $\cos \theta_K$  corrected with the angular efficiency function before (left) and after (right) constraining its value at  $\cos \theta_K = 1$ , shown as the red curve. The black points show the  $\cos \theta_K$  distribution in the simulated sample. Zooming close to  $\cos \theta_K = 1$  clearly shows the problematic region.

Having chosen the best possible set of moments the efficiency function has to be constrained such that the condition,

$$\epsilon(\cos \theta_K = 1, \cos \theta_\mu, \varphi_h) = 0, \quad (4.15)$$

is satisfied. It turns out that this condition can be easily satisfied if the efficiency basis of Eq. 4.10, is modified as follows:

$$\alpha_i^{jk}(\Omega) \rightarrow \left( \alpha_i^{jk}(\Omega) \right)' = \left[ P_i^0(\cos \theta_K) - P_i^0(1) \right] Y_{jk}(\cos \theta_\mu, \varphi_h), \quad (4.16)$$

resulting in the next efficiency function:

$$\epsilon(\Omega) = \sum_{ijk} c_{jk}^i \left[ P_i^0(\cos \theta_K) - P_i^0(1) \right] Y_{jk}(\cos \theta_\mu, \varphi_h). \quad (4.17)$$

Note that the subscript  $i$  in  $P_i^0(1)$  is irrelevant since  $P_i^0(1) = 1$  for any value of  $i$ . The result of the above constraint is shown in Figure 4.8b, where it can be seen that the acceptance projection on  $\cos \theta_K = 1$  is not negative anymore. In addition it has also been checked whether the

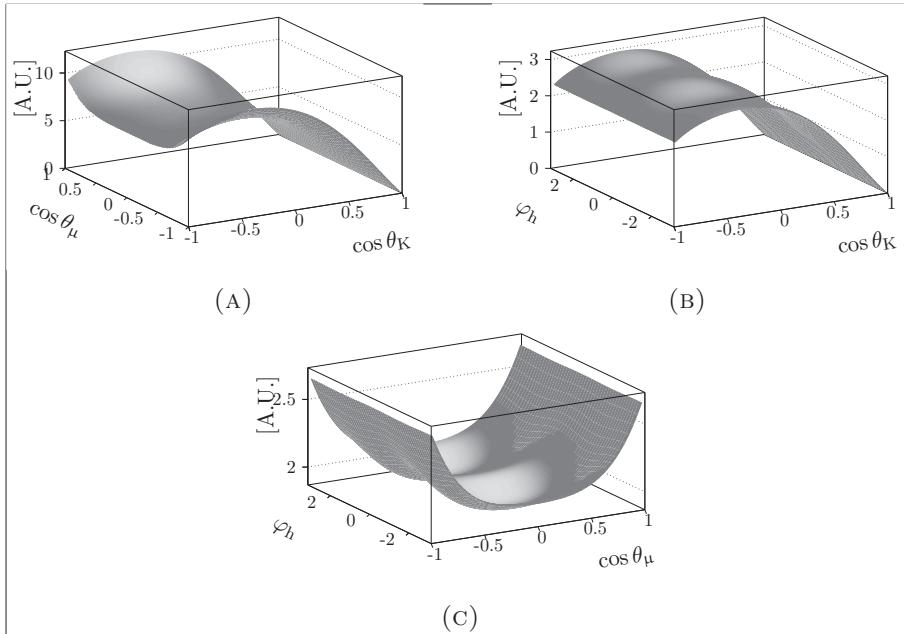


FIGURE 4.9: 2D efficiency projections of the constrained efficiency function.  $y$  axis in arbitrary units.

acceptance is negative in 2D and 3D, since it is possible for a 2D projection of the acceptance to take negative values whilst the 1D projections are positive. Two dimensional projection of the constrained acceptance can be seen in Figure 4.9.

Constraining the acceptance in the manner described above comes with a small penalty: the shape of the 1D acceptance projections is potentially slightly distorted. In order to reduce this distortion and describe better the acceptance shape an additional step is implemented. As previously mentioned, the acceptance function of Eq. 4.17 is fitted to the simulated data where the efficiency moments are allowed to vary from their measured values, of Table 4.3. This feature is implemented by a multivariate Gaussian constraint where correlations between efficiency moments are taken into account. The fitting PDF is shown in Eq. 4.18 and it consists of the PDF with which the simulated data were

generated, multiplied with the efficiency function (4.17) times the multivariate Gaussian constraint. The physics parameters used to generate the simulation sample are fixed and only the efficiency moments are allowed to vary. The multivariate Gaussian,  $G(\vec{c})$ , used to constrain the efficiency moments is shown in Eq. 4.19.

$$\mathcal{P} = P^{\text{gen}}(\Omega) \times \epsilon(\Omega; \vec{c}) \times G(\vec{c}), \quad (4.18)$$

with

$$G(\vec{c}) = \frac{1}{\sqrt{(2\pi)^k |C|}} e^{-\frac{1}{2}(\vec{c} - \langle \vec{c} \rangle)^T C^{-1}(\vec{c} - \langle \vec{c} \rangle)}, \quad (4.19)$$

where  $\vec{c}$  represents the estimated efficiency moments and  $\langle \vec{c} \rangle$  are the efficiency moments of Table 4.3.  $C$  and  $|C|$  are the covariance matrix of the efficiency moments and its determinant respectively. Lastly,  $k$  is the number of efficiency moments present in Eq. 4.17.

### Alternative parametrization

As mentioned before there is at least one more way to parameterize the angular acceptance, namely the *normalization weights*. The main difference between this approach and the method adopted in this analysis is that essentially the normalization weights use the angular functions of Table 4.2 as a basis to project the efficiency of Eq. 4.11, instead of the orthogonal basis  $a_i^{jk}$  of Eq. 4.10. Looking at Eq. 4.20 it is helpful to realize how the normalization weights,  $\xi_a$ , are introduced into the fitting PDF. One can see that they effectively do not multiply the angular PDF directly. Instead the normalization weights modify the relative contributions between the angular functions  $h_n$ , in the normalization integral of the fitting PDF. The normalization weights are also calculated following the concept of Monte Carlo integration similarly to the case of efficiency moments.

$$\mathcal{P} \propto \frac{\sum_n h_n(\Omega)}{\sum_n \int d\Omega \epsilon(\Omega) h_n(\Omega)}, \quad \text{where } \int d\Omega \epsilon(\Omega) h_n(\Omega) \equiv \xi_a. \quad (4.20)$$

The advantages of this method is that it is by construction immune from any issue with localized negative efficiency values: the acceptance shape is not parameterized at all. The disadvantage on the other hand is that directly plotting the fitting PDF becomes impossible when using normalization weights. This is because the normalization integral of the fitting PDF is modified in a non-trivial way and other methods need to be implemented in order to visualize the PDF. Lastly the two methods can yield identical values on the fitted parameters by mapping the normalization weights to efficiency moments. This can be done if one solves the normalization integral,  $\int d\Omega P^{\text{gen}}(\Omega) \times \epsilon(\Omega)$  of the PDF of Table 4.2 and the efficiency function in Eq. 4.10, then the mapping of  $c_{jk}^i \rightarrow \xi_a$  follows naturally by exploiting the orthogonality of  $P$  and  $Y$ . However, the above-mentioned mapping is not unique and not always invertible. It depends on the choice of the particular efficiency moments. More details on the normalization weights can be found in [31] and [83].

### 4.2.3 Acceptance Corrections

The angular acceptance described in section Section 4.2.2 is determined with simulated  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  events. Thus the acceptance determination can be trusted only if the simulation accurately describes the detector as well as the physical amplitudes of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay. Specifically real and simulated data differences in the momentum distributions of the final state particles may reflect a different acceptance in real compared to simulated data. Such differences are at a few percent level between real and simulated data. On the other hand, differences between data and simulation could be also attributed to differences in the underlying physical parameter values generating the simulated data compared to these present in data. For example, the presence of S-wave in the data (and its absence on simulated data) can cause a difference in the observed kaon momentum spectrum even though the angular acceptance may be perfectly simulated. Both of the above-mentioned effects have an impact on the decay angles distribution since the helicity angles are functions of the momenta of the final state particles.

In order to overcome the problem of simulation imperfections, the simulated sample is weighted using an iterative procedure. At each

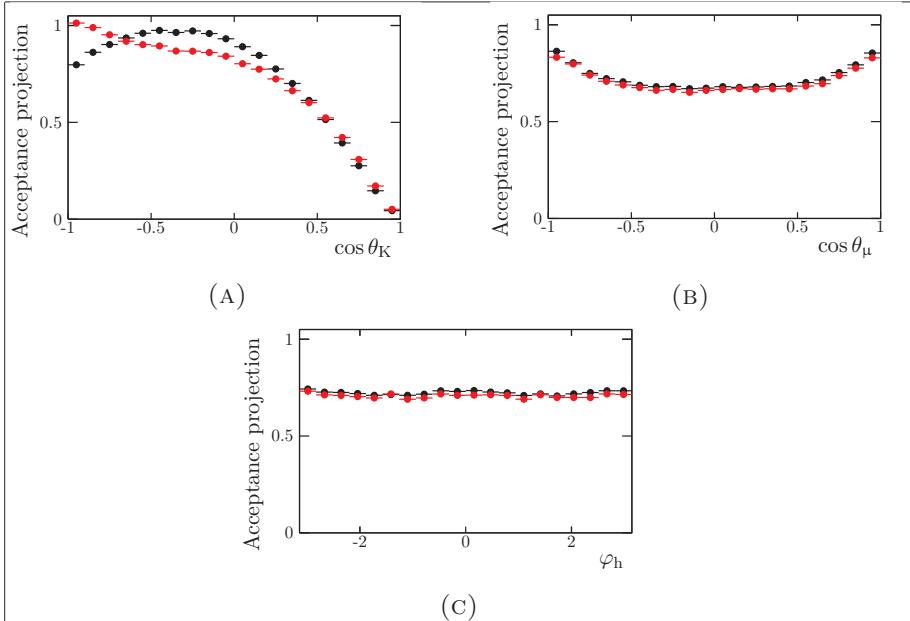


FIGURE 4.10: Angular acceptance shape from simulated data before (black points) and after (red points) acceptance corrections.

step of the weighting procedure the simulated data are corrected for the current best estimate of the physics parameters and the two dimensional ( $p_K, p_\pi$ ) momentum distribution. The efficiency moments are re-evaluated using the weighted simulated sample and the fit to the data is repeated. This changes the best estimate of the parameters of interest. This procedure is repeated until these parameters converge. The full procedure by steps can be summarized as:

1. Compute an initial set of efficiency moments using uncorrected  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  simulated data.
2. Fit the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data using the initial efficiency moments, and obtain the first estimate of the physics parameters.

Iteration	1	2	3	4	5	6	total
$\mathcal{A}_{CP}^0$	-0.00	-0.03	+0.03	+0.00	-0.00	-0.00	+0.00
$\mathcal{A}_{CP}^{\parallel}$	+0.13	+0.04	-0.09	-0.04	-0.02	-0.01	+0.01
$\mathcal{A}_{CP}^{\perp}$	-0.19	+0.15	+0.02	+0.01	+0.00	+0.00	-0.01
$\mathcal{A}_{CP}^S$	-0.01	-0.17	+0.09	+0.06	+0.03	+0.01	+0.01
$f_0$	-1.15	+0.96	+0.19	+0.00	-0.02	-0.01	-0.03
$f_{\parallel}$	-0.05	-0.06	+0.03	+0.04	+0.02	+0.01	-0.01
$\delta_{\parallel}$	-0.37	+0.23	+0.09	+0.03	+0.01	+0.0	-0.01
$\delta_{\perp}$	-0.57	+0.28	+0.16	+0.07	+0.03	+0.01	-0.02
$f_S^1$	+0.86	-0.66	-0.32	-0.12	-0.05	-0.02	-0.31
$f_S^2$	+0.84	-0.27	-0.18	-0.09	-0.05	-0.02	+0.23
$f_S^3$	+0.00	-0.02	-0.00	+0.01	+0.02	+0.01	+0.02
$f_S^4$	+0.47	-0.21	-0.14	-0.06	-0.04	-0.02	-0.00
$\delta_S^1$	-2.08	+0.90	+0.66	+0.32	+0.13	+0.04	-0.03
$\delta_S^2$	+0.13	-0.05	-0.02	-0.03	-0.01	-0.01	+0.01
$\delta_S^3$	-0.90	+0.48	+0.19	+0.08	+0.04	+0.03	-0.08
$\delta_S^4$	-2.43	+1.10	+0.53	+0.24	+0.13	+0.07	-0.36

TABLE 4.4: Parameters of interest after each iteration. The amount of change is expressed as a fraction of one standard deviation. Differences get progressively smaller, implying convergence.

3. Weight each candidate in the simulated data for the difference between the angular PDF obtained from the previous fit and the angular PDF used to generated the simulated data. Implying that the underlying physics of the simulated data corresponds to the estimated physics parameters obtained previously.
4. Compare the 2D  $(p_K, p_\pi)$  momentum distribution between the weighted simulated data of the previous step with the data, and then weigh again each simulated candidate for the difference.
5. Re-estimate the efficiency moments using the physics and momentum corrected  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  simulated data and repeat the fit to the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data.
6. Go back to step three and repeat until the change in the parameters of interest is negligible.

The  $(p_K, p_\pi)$  weighting of the simulated data sample is done using 2D histograms, following the intervals defined in Table 4.5. At least 6 iterations are necessary to achieve convergence. Detailed evolution of the parameters of interest after each iteration are shown in Table 4.4. The central values of the parameters of interest do not change significantly implying that the overall effect of the acceptance corrections is small. Figure 4.10 shows the shape of the acceptance before and after corrections. This is also supported by the effect of these corrections on the parameters of interest which are shown in Table 4.4.

variables	range	#intervals
$p_K$	$[0, 140]$ GeV/c	10
$p_\pi$	$[0, 60]$ GeV/c	10

TABLE 4.5: Weighting variable intervals.

Lastly, the acceptance can be floated in the angular fit to the data in order to assign a systematic uncertainty, see Section 4.4.2. This feature makes, in hindsight, the acceptance correction procedure unnecessary.

#### 4.2.4 $K\pi$ Invariant Mass

The dependence of the  $B_{(s)}^0 \rightarrow J/\Psi K\pi$  decay amplitude on  $m_{K\pi}$  as introduced in Eq. 4.7 is treated in a special way. The  $M_n$  functions are not implemented, but instead they are integrated over in four intervals. The choice of this strategy is good enough given the fact that the  $m_{K\pi}$  region relevant for this analysis is dominated by the  $K^*$  P-wave, while the S-wave component is small. As a result the development of a full model for both the P-wave and S-wave, which is not the first priority in the current analysis, can be avoided. Despite the  $m_{K\pi}$  integration there is still some residual dependence left. It originates from integrals like the one in the following equation:

$$\int dm_{K\pi} \Phi M_n^* M_m, \quad (4.21)$$

where  $M_{n,m}$  run the same as in Eq. 4.7 and the phase space factor:

$$\Phi = \frac{1}{4} \frac{\lambda(m_{B_s^0}, m_{K\pi}, m_{J/\psi}) \lambda(m_{K\pi}, m_K, m_\pi)}{m_{B_s^0} m_{K\pi}}, \quad (4.22)$$

with,

$$\lambda(m_1, m_2, m_3) = [m_1^4 + m_2^4 + m_3^4 - 2m_1^2 m_2^2 - 2m_1^2 m_3^2 - 2m_2^2 m_3^2]^{1/2}.$$

Integrals like Eq. 4.21 eventually manifest themselves only in the SP-wave interference terms of Table 4.3, since the  $m = n$  integrals are by construction equal to one. These integrals are written as:

$$\frac{\int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \Phi s^* \times p}{\int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \Phi |s|^2 \int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \Phi |p|^2} = C_{SP} e^{-i\theta_{SP}}, \quad (4.23)$$

where  $s$  and  $p$  stand for the S-wave and P-wave line-shapes respectively. Evaluating these complex valued integrals yield the  $C_{SP}$  factors which are inserted in the angular functions of Table 4.2. Particularly, the real part is inserted as a multiplicative factor in all of the SP-wave interference terms, for example  $\Re(A_0^* A_S) \rightarrow \Re(A_0^* A_S) \cdot C_{SP} \cdot \cos(\delta_0 + \delta_S + \theta_{SP})$ . The imaginary part,  $\theta_{SP}$ , is absorbed in the overall phase of the SP-wave interference term. Varying either  $C_{SP}$  or  $\theta_{SP}$  in the fit is not preferred due to high correlations between the decay amplitudes. Instead it would be better to extend the angular fit so that the  $m_{K\pi}$  dependence is modeled directly via the  $M_n$  terms of Eq. 4.7. The above-mentioned line-shapes,  $s$  and  $p$  that are necessary to compute the  $C_{SP}$  factors are a combination of a  $K^*(892)^0$  and  $K_1^*(1410)^0$  for the P-wave, as in [84]. For the S-wave a LASS parametrization [85] is used, consisting of a linear combination of the  $K_1^*(1430)^0$  resonance with a non-resonant term, coming from elastic scattering.

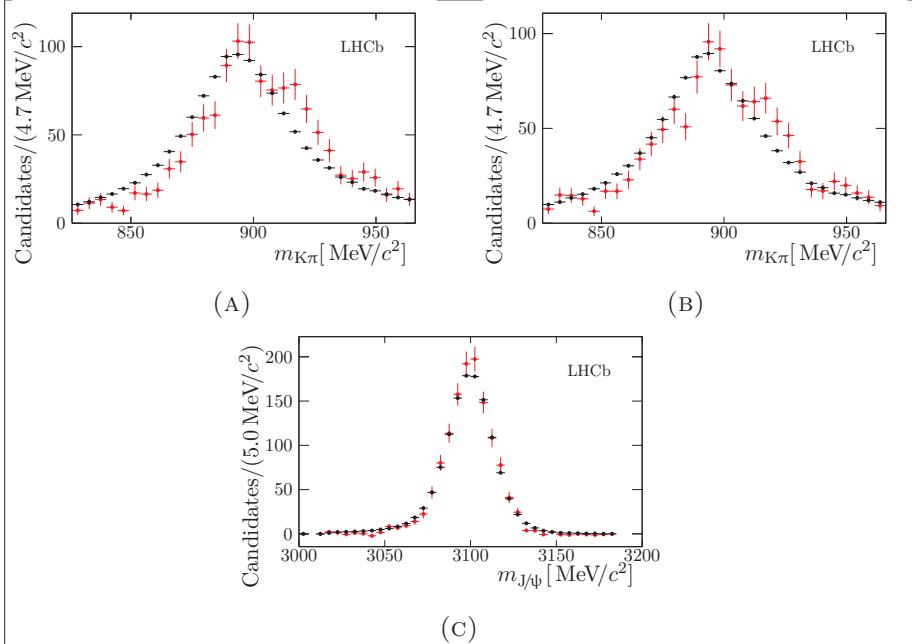


FIGURE 4.11: Background subtracted mass distributions of the intermediate state particles. Black (red) points correspond to  $B^0$  ( $B_s^0$ ) decays. Top row shows the  $m_{K\pi}$  distribution before (A) and after (B) angular acceptance corrections. The  $J/\psi$  mass is shown at the bottom.

### Background subtracted $m_{K\pi}$ distribution

While the  $B_s^0$  and  $B^0$  background subtracted muon pair mass,  $m_{\mu^+\mu^-}$ , have very similar shapes, the  $m_{K\pi}$  spectra exhibit different shapes. Indeed, as suggested by Figure 4.11a the  $B_s^0$   $m_{K\pi}$  *sPlot* seems to be slightly distorted compared to the one of  $B^0$   $m_{K\pi}$ . Since there is no *a priori* reason why the  $K^{*0}$  P-wave should have different shape between  $B_s^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow J/\psi K^{*0}$ ; This could indicate the presence of interference between the  $K\pi$  S-wave and the  $K^{*0}$ , which may be stronger in the  $B_s^0$  decays compared to the  $B^0$  ones. In order to check the validity of this hypothesis, two additional studies are performed. First it was

checked that the non-resonant background treatment propagated to the *sWeights* was not responsible for this behavior. Indeed, no significant difference between the  $B_s^0 m_{K\pi}$  spectrum using *sWeights* computed with and without simulated data injection was found. Second, the  $m_{K\pi}$  is corrected for angular acceptance effects, by weighting each candidate with the inverse of the acceptance function of Eq. 4.17. This way the interference between the  $K\pi$  S-wave and the  $K^{*0}$  P-wave vanishes, since for a perfect acceptance the interference terms in Table 4.7, integrated over all angles, evaluate to zero. This is demonstrated in Figure 4.11b. The  $B_s^0 m_{K\pi}$  distribution is closer to the one of the  $B^0$  after applying the efficiency correction. The effect is quantified by a Kolmogorov-Smirnov test, where its value before and after angular acceptance corrections is 0.470 and 0.833 respectively. This indicates the presence of stronger interference in the  $B_s^0$  case compared to the  $B^0$ . Lastly, It was checked that the acceptance correction has negligible effects on the shapes of the  $J/\psi$  mass distributions

#### 4.2.5 Production and Detection Asymmetries

The nature of the current analysis is such that it is affected by asymmetries related to experimental effects. Particularly, the difference between the number of observed  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  decays is related to the main parameters of interest, introduced in Section 4.2.6. These parameters essentially quantify the direct  $CP$  asymmetry between the previous decays. Considering now the case where collisions at the LHC would favor the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay over its  $CP$  conjugate,  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$ ; or that the detector might have different efficiency for positive and negative kaons (or pions), then the observed  $\mathcal{A}_{CP}$  will be biased by these experimental effects. In fact it happens so that both of these effects are present. The first one, where the creation of a  $B_s^0$  or  $\bar{B}_s^0$  is favored is labeled as production asymmetry,  $\mathcal{A}^{\text{prod}}$ , whereas the second effect is labeled as detection asymmetry,  $\mathcal{A}^{\text{det}}$ . The production and detection asymmetries are defined in the following equation:

$$\mathcal{A}^{\text{prod}} = \frac{\sigma(\bar{B}_s^0) - \sigma(B_s^0)}{\sigma(\bar{B}_s^0) + \sigma(B_s^0)}, \quad \mathcal{A}^{\text{det}} = \frac{\epsilon^{\text{det}}(K^{*0}) - \epsilon^{\text{det}}(\bar{K}^{*0})}{\epsilon^{\text{det}}(K^{*0}) + \epsilon^{\text{det}}(\bar{K}^{*0})}, \quad (4.24)$$

where  $\sigma$  is the production cross section of either  $B_s^0$  or  $\bar{B}$ , whereas  $\epsilon_{\text{det}}$  is the efficiency of detecting  $B_s^0$  ( $\bar{B}_s^0$ ) decays into the final state which is a  $J/\psi \bar{K}^{*0}$  ( $J/\psi K^{*0}$ ), see also Section 4.2.6. Both of the asymmetries have to be estimated so that the experimentally measured asymmetry,  $\mathcal{A}_{CP}^{\text{raw}}$ , defined in Eq. 4.34, can be corrected for as follows:

$$\mathcal{A}_{CP}(B_s^0 \rightarrow J/\psi \bar{K}^{*0}) = \mathcal{A}_{CP}^{\text{raw}} + \mathcal{A}^{\text{det}} - \kappa \mathcal{A}^{\text{prod}}, \quad (4.25)$$

where the corrected asymmetry  $\mathcal{A}_{CP}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$  is related to the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay amplitudes, defined in Eq. 4.7, as:

$$\begin{aligned} \mathcal{A}_{CP}(B_s^0 \rightarrow J/\psi \bar{K}^{*0}) &= \\ &= \frac{|\mathcal{A}(\bar{B}_s^0 \rightarrow J/\psi K^{*0})|^2 - |\mathcal{A}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})|^2}{|\mathcal{A}(\bar{B}_s^0 \rightarrow J/\psi K^{*0})|^2 + |\mathcal{A}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})|^2}, \end{aligned} \quad (4.26)$$

The factor  $\kappa$  of Eq. 4.25 accounts for the dilution of the production asymmetry due to  $B_s^0 - \bar{B}_s^0$  oscillations [86] and is:

$$\kappa = \frac{\int_0^\infty e^{-\Gamma_s t} \cos(\Delta m_s t) \varepsilon(t) dt}{\int_0^\infty e^{-\Gamma_s t} \cosh\left(\frac{\Delta\Gamma_s}{2} t\right) \varepsilon(t) dt}. \quad (4.27)$$

It evaluates to 0.06%, which implies that the  $B_s^0 - \bar{B}_s^0$  oscillation frequency is large compared to the average lifetime,  $\Gamma_s$ . As a result the production asymmetry is nearly completely washed out. The quantities  $\Delta\Gamma_s$  and  $\Delta m_s$  have been introduced in Section 1.3, while the decay time efficiency function  $\varepsilon(t)$  is estimated on background subtracted  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data and parameterized as in [87].

Using the result from the dedicated LHCb measurement of the production asymmetry [88] and replacing the yield fractions in intervals of

$B_s^0$  momentum of Table 3 in the same reference with the background subtracted  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  ones of the current analysis, one gets the following value for the production asymmetry:

$$\mathcal{A}^{\text{prod}} = -1.64 \pm 2.28(\text{stat}) \pm 0.55(\text{syst}) \%. \quad (4.28)$$

The detection asymmetry is computed starting from [89] and weighting the kaon momentum distribution to much the observed distribution in  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data. The result is:

$$\mathcal{A}^{\text{det}} = -1.086 \pm 0.531(\text{stat}) \%. \quad (4.29)$$

#### 4.2.6 Likelihood fit and Total Decay Rate

Nearly all the necessary ingredients to fit for the parameters of interest are covered in the previous subsections. However there are a few important details remaining. The current subsection addresses the implications arising from the necessary splitting of the data in slices and the impact of this in the adopted fitting strategy. Lastly but most importantly the issue of defining the polarization dependent  $CP$  asymmetries,  $\mathcal{A}_{CP}^k$ , as free parameters in the angular fit and how this becomes relevant in the normalization of the PDFs involved is discussed in the last part of the current subsection.

##### Simultaneous likelihood fit

There are several reasons as to why the data are divided into sub-samples. For example, the special treatment of the  $m_{K\pi}$  dependence, as described in Section 4.2.4 requires that the data is split in four  $m_{K\pi}$  intervals. This implies that for each data category, a corresponding component PDF must be defined. Some parameters of interest differ between the categories while some others are shared. In order to avoid performing four independent fits and subsequently averaging the common parameters of interest, a simultaneous fit is implemented. In a simultaneous fit each component PDF is assigned to a certain part of the data such that the corresponding component PDF is invoked when the total PDF is evaluated. The current section summarizes all the

categories that the data is split into and the implications arising from this. Lastly a brief insight in some aspects of the likelihood parameter estimation method are included.

The first data category is related to the fact that the angular PDF terms of Table 4.2 cannot describe both  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$ . There is a relative minus sign difference in the imaginary terms originating from complex conjugating the decay amplitude of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ . For reasons explained in Section 5.3 the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decays always have a negative kaon in the final state and *vice versa* in the case of  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  decays. Thus the data must be split according to the charge of the kaon. Furthermore, as can be seen in Table A.4 the agreement between the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  samples is sufficient for them to be averaged into one set of efficiency moments. However, the angular acceptance as described in Section 4.2.2 is calculated separately for these two samples. This is done to avoid involving the  $\mathcal{A}_{CP}^i$  parameters of interest in the estimation of the efficiency moments, which significantly simplifies the necessary computation. Note that the  $\mathcal{A}_{CP}^i$  parameters are used to average the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  decay amplitudes into one parameter, see Section 4.2.6.

	interval 1	interval 2	interval 3	interval 4
$m_{K\pi}$ range	[826, 861)	[861, 896)	[896, 931)	[931, 966]
$C_{SP}$ factor	0.9681	0.9312	0.9519	0.9880

TABLE 4.6: Definitions of  $m_{K\pi}$  intervals and the corresponding  $C_{SP}$  factor values.

The next observable used to split the data is already mentioned in Section 4.2.4 that dealt with the  $m_{K\pi}$  dependence. There are four  $m_{K\pi}$  intervals defining four more categories in the data. Furthermore, and contrary to the above-mentioned data category, there are parameters of interest that are different in each  $m_{K\pi}$  interval. Namely the S-wave fraction  $f_S$  and phase  $\delta_S$ . In addition, the  $C_{SP}$  factors of Section 4.2.4 need to be calculated and inserted separately for each component PDF resulting in Table 4.6, where the  $C_{SP}$  factors are shown for each  $m_{K\pi}$  interval. The acceptance is also split in the  $m_{K\pi}$  category since it was

found to vary significantly between  $m_{K\pi}$  intervals, see Table A.5 and Table A.6. The  $m_{K\pi}$  intervals chosen are shown in Table 4.6.

The data categories described above result in a simultaneous fit of eight categories. A single angular acceptance is computed from the combined Run 1 data on grounds of tables Table A.2 and Table A.3.

Having addressed the simultaneous nature of the fit it is interesting to provide an insight on the connection between a PDF and the likelihood estimation of the parameters of interest. Thus, starting from any PDF  $\mathcal{P}$  a likelihood function can be built as follows:

$$L(\vec{p}; \vec{x}) = \prod_e^n \mathcal{P}(\vec{x}; \vec{p}). \quad (4.30)$$

By definition the likelihood function is the product of probabilities of every candidate according to the given PDF. Note that, contrary to the PDF, the likelihood is a function of the parameters given all the candidates in the data. It is useful to realize that at the maximum of the likelihood function one gets the best estimate for the parameters of interest,  $\vec{p}$ , given the data,  $\vec{x}$ . The inverse of the covariance matrix for the estimated parameters is given by the negative second derivative of the log likelihood at the maximum,

$$\left( V^{-1} \right)_{ij} = - \left. \frac{\partial^2 \ln L}{\partial p_1 \partial p_2} \right|_{\vec{p}_{\max}}. \quad (4.31)$$

Regarding the concept of the extended likelihood: In Eq. 4.30 the number of candidates is fixed. However, it is possible to extend the previous equation such that the expected number of candidates shows up as a parameter in the likelihood function and can thus be estimated. This is useful in cases where the PDF consists of a sum over two species, such as signal and background, or if the expected number of candidates depends on the parameters of interest  $\vec{p}$ . For the current analysis, the latter is exploited in order to estimate the  $\mathcal{A}_{CP}^i$  asymmetries in the fit. The PDF is extended by multiplying it with the Poisson probability to observe  $n$  events when expecting  $\mu(\vec{p})$  candidates, see Eq. 4.32. The expected number of candidates  $\mu(\vec{p})$  depends on the parameters of interest

which increases the statistical uncertainties. This is because the likelihood has more freedom as the number of parameters increases thus the shape of the likelihood becomes less steep leading to larger uncertainty.

$$L(\vec{p}; \vec{x}) = P(n; \mu(\vec{p})) \prod_e^n \mathcal{P}(\vec{x}; \vec{p}), \quad (4.32)$$

For purely practical reasons one would typically choose to minimize the negative logarithm of the likelihood, which after some simple algebra and using the known expression of the Poisson distribution results in:

$$\begin{aligned} -\ln L(\vec{p}; \vec{x}) &= -\ln \left( \frac{\mu(\vec{p})^n}{n!} e^{-\mu(\vec{p})} \right) - \ln \left( \prod_e^n \mathcal{P}(\vec{x}; \vec{p}) \right) \\ &= -n \ln(\mu(\vec{p})) + \mu(\vec{p}) + \ln n! - \sum_e^n \ln(\mathcal{P}(\vec{x}; \vec{p})) \\ &= \mu(\vec{p}) - \sum_e^n \ln(\mu(\vec{p}) \mathcal{P}(\vec{x}; \vec{p})) + \text{const.} \end{aligned} \quad (4.33)$$

Constant terms in the likelihood, like  $\ln n!$ , affect neither the shape of the likelihood nor the position of the minimum and can thus be omitted.

### **$CP$ asymmetries as free parameters**

The raw asymmetry,  $\mathcal{A}_{CP}^{\text{raw}}$ , discussed in Section 4.2.5, can be measured by counting  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decays and use them to determine the following asymmetry:

$$\mathcal{A}_{CP}^{\text{raw}} \equiv \frac{N^+ - N^-}{N^+ + N^-}, \quad (4.34)$$

where  $N^+$  and  $N^-$  are the number of observed  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  respectively. Going one step further,  $\mathcal{A}_{CP}^{\text{raw}}$  can be implemented as a parameter to be estimated in the fit. The advantage is that the statistical uncertainty propagation for these parameters is implicit in the likelihood thus avoiding explicit calculations. In addition, the above

quantities can then be estimated separately for each polarization. In order to vary  $\mathcal{A}_{CP}^{\text{raw}}$  in the fit it is necessary to re-parameterize them as:

$$\mathcal{A}_{CP}^i = \frac{N^+|A_i^+|^2 - N^-|A_i^-|^2}{N^+|A_i^+|^2 + N^-|A_i^-|^2} \quad (4.35)$$

$$\text{and } |A_i|^2 = \frac{N^+|A_i^+|^2 - N^-|A_i^-|^2}{N^+ + N^-}, \quad (4.36)$$

where  $|A_i^+|^2$  and  $|A_i^-|^2$  are the polarized amplitude in the  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  samples respectively. The above equations can be solved for  $|A_i^+|^2$  and  $|A_i^-|^2$ . The idea is that the latter parameters will be used as amplitudes for the  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  PDFs respectively. This way both samples are fitted simultaneously in one fit. Solving Eq. 4.35 for  $|A_{\pm}|^2$  yields:

$$|A_i^+|^2 = \xi |A_i|^2 (1 + \mathcal{A}_{CP}^i) \text{ and } |A_i^-|^2 = \frac{\xi}{2\xi - 1} |A_i|^2 (1 - \mathcal{A}_{CP}^i). \quad (4.37)$$

The  $|A_i^+|^2$  and  $|A_i^-|^2$  are the amplitude expressions that are used along with the angular functions of Table 4.7. The normalization factors in front of them can be re-expressed as function of  $\mathcal{A}_{CP}^{\text{raw}}$  using Eq. 4.34.

$$\begin{aligned} \xi &= \frac{N^+ + N^-}{2N^+} = \frac{1}{2} \left( 1 + \frac{N^-}{N^+} \right) = \\ &= \frac{1}{2} \left( 1 + \frac{1 - \mathcal{A}_{CP}^{\text{raw}}}{1 + \mathcal{A}_{CP}^{\text{raw}}} \right) = \frac{1}{1 + \mathcal{A}_{CP}^{\text{raw}}}, \end{aligned} \quad (4.38)$$

$$\frac{\xi}{2\xi - 1} = \frac{1}{2 - 1/\xi} = \frac{1}{1 - \sum_i \mathcal{A}_{CP}^{\text{raw}}}. \quad (4.39)$$

Now, the quantities that are actually estimated in the likelihood fit are the  $|A_i|^2$  with  $i : \{0, \parallel, \perp, S\}$ . Having said that, and by looking at Eq. 4.37, one understands that in order for  $|A_i^+|^2$  and  $|A_i^-|^2$  to be meaningful and consistent, the sum of all amplitudes  $\sum |A_i|^2$  (with  $i :$

$0, \parallel, \perp, S$  ) must add up to one. Thus imposing the condition  $\sum |A_i|^2 = 1$  one can re-express the  $|A_i|^2$  as relative fractions:

$$|A_S|^2 = f_S \quad \text{and} \quad |A_k|^2 = (1 - f_S)f_k. \quad (4.40)$$

Note that the P-wave fractions are also constrained so that they add up to one by imposing  $\sum f_k = 1$  with  $k : \{0, \parallel, \perp\}$ . The fractions  $f_k$  and  $f_S$  are the actual parameters that one wishes to estimate and correspond to these used in the subsequent analysis of Chapter 5. The final expressions of the  $|A_i^\pm|^2$  after the above is given by:

$$|A_k^\pm|^2 = \frac{1 \pm \mathcal{A}_{CP}^k}{1 \pm \sum_i \mathcal{A}_{CP}^i} (1 - f_S)f_k \quad \text{and} \quad |A_S^\pm|^2 = \frac{1 \pm \mathcal{A}_{CP}^S}{1 \pm \sum_i \mathcal{A}_{CP}^i} f_S, \quad (4.41)$$

where the  $i$  index runs through all amplitudes whereas  $k$  only runs through the P-wave amplitudes.

There is one final step required to build the complete extended PDF. That is the normalization of each component PDF. Tempting as it might be to normalize the component PDFs to the number of observed candidates it would be wrong since the observed number of candidates are diluted by production and detection asymmetry. Given that the parameters of interest,  $\mathcal{A}_{CP}^i$ , quantify the difference between the observed  $B_s^0 \rightarrow J/\psi K^{*0}$  and  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decays, it follows that the observed number of these decays need to be corrected for. With input from Section 4.2.5 where the production and detection asymmetries were estimated, the observed yields are made dependent on  $\mathcal{A}_{CP}^i$  as:

$$N_{CP}^+ = \frac{N_{\text{raw}}^{\text{total}}}{2} \left( f_S \mathcal{A}_{CP}^S + \sum_k^{0, \parallel, \perp} (1 - f_S)f_k \mathcal{A}_{CP}^k \right), \quad (4.42a)$$

$$N_{CP}^- = N_{\text{raw}}^{\text{total}} - N_{CP}^+ \quad (4.42b)$$

and then subsequently corrected for the production and detection asymmetry resulting in:

$$N^\pm = N_{CP}^\pm (1 \pm \mathcal{A}^{\text{det}} \mp \mathcal{A}^{\text{prod}}). \quad (4.43)$$

The above yields are the expected number of candidates in the extended likelihood term of the PDFs as discussed earlier.

### 4.3 Normalization of $B_s^0 \rightarrow J/\psi K^- \pi^+$

The current section addresses the determination of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  branching fraction, hereafter  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$ . This is one of the necessary ingredients for estimating the penguin contributions to  $\phi_s$ , addressed in Chapter 5. Absolute branching fraction measurements are more complicated to perform compared to relative ones. Because a dedicated analysis must be built up and optimize for the channel of interest. Also in a relative branching fraction measurement some systematic uncertainties drop out, *i.e.* the systematic related to the luminosity estimate. Thus, the  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$  is normalized with respect to two different channels, namely  $B_s^0 \rightarrow J/\psi \phi$  and  $B^0 \rightarrow J/\psi K^{*0}$  and a single average is computed.

#### Normalization with respect to $B_s^0 \rightarrow J/\psi \phi$

Equation 4.44 shows how  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$  is normalized with respect to the  $B_s^0 \rightarrow J/\psi \phi$  channel. The branching fraction defined in the previous equation will hereafter be denoted as  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_\phi$ . It requires as input the observed number of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B_s^0 \rightarrow J/\psi \phi$  decays. The first one is taken from Eq. 4.5 whereas the latter requires more effort. Particularly, the selection steps applied to  $B_s^0 \rightarrow J/\psi \phi$  are chosen such that they are as similar as possible to that of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ , see Table A.1.

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)} = \frac{N_{B_s^0 \rightarrow J/\psi K^- \pi^+}}{N_{B_s^0 \rightarrow J/\psi K^+ K^-}} \times \frac{\varepsilon_{B_s^0 \rightarrow J/\psi \phi}^{MC}}{\varepsilon_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}^{MC}} \times \frac{\omega_{B_s^0 \rightarrow J/\psi \phi}}{\omega_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}} \times \frac{\mathcal{B}(\phi \rightarrow K^+ K^-)}{\mathcal{B}(\bar{K}^{*0} \rightarrow K^- \pi^+)}. \quad (4.44)$$

Also the BDT classifier is applied to the reference channel with the same cut values as in the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  channel. The ratio of observed number of candidates is represented by the first fraction in Eq. 4.44, whereas the ratio of total efficiency between the two channels is denoted by the second. The latter fraction is corrected for known differences in efficiency between data and simulated data. Additionally, there is a correction factor  $\omega$  to account for the presence of S-wave in the two channels, since the penguin estimation assumes no S-wave contribution to  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ . Each of the factors  $\omega$  in Eq. 4.44 correspond to the ratio of two angular PDFs. One with the S-wave parameters set to zero and P-wave to these obtained in the angular fit, whereas the other with both P-wave and S-wave parameters as obtained from the same fit. The  $\omega$  factors can thus be understood as an overall S-wave fraction.

The branching fractions of the intermediate resonance decaying to a dihadron system are obtained from [28]. The  $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$  in [73] is updated with the latest b quark hadronization fraction,  $f_d/f_s$  [90]. This quantity expresses the ratio of probabilities that a b quark will form a meson with a d quark over an s quark. The efficiency and the angular correction ratios are about 2 and 0.9 respectively. The above information is used to determine the following branching fraction:

$$\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_\phi = (4.20 \pm 0.20 \pm 0.13 \pm 0.36) \times 10^{-5}, \quad (4.45)$$

where the uncertainties are statistical, systematic and due to  $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$  respectively.

**Normalization with respect to  $B^0 \rightarrow J/\psi K^{*0}$** 

Alternatively  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$  can be normalized with respect to the  $B^0 \rightarrow J/\psi K^{*0}$  channel to obtain  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_d$ :

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})}{\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})} = \frac{N_{B_s^0 \rightarrow J/\psi K^- \pi^+}}{N_{B^0 \rightarrow J/\psi K^{*0}}} \times \times \frac{f_d}{f_s} \times \frac{\varepsilon_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}^{MC}}{\varepsilon_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}^{MC}} \times \frac{\omega_{B^0 \rightarrow J/\psi K^{*0}}}{\omega_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}}, \quad (4.46)$$

The previous equation is similar to Eq. 4.44 as far as the way efficiency and angular correlation ratios are obtained. The observed number of candidates are both obtained from Eq. 4.5 and the latest measurement of the hadronization fraction from [90]. The  $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$  value is taken from [91], where the S-wave component is explicitly modeled and does contribute in the  $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$  estimation. The final result is:

$$\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_d = (3.95 \pm 0.18 \pm 0.16 \pm 0.23 \pm 0.43) \times 10^{-5}, \quad (4.47)$$

where the first two uncertainties are statistical and systematic and the rest are coming from  $f_d/f_s$  and  $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$  respectively.

**Averaged  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$** 

Both estimates of Eq. 4.45 and Eq. 4.47 are compatible within uncorrelated systematic uncertainties. A weighted average of these two estimates is performed where correlations are taken into account. Correlations are introduced via the common parameters  $N_{B_s^0 \rightarrow J/\psi K^- \pi^+}$  and  $\omega_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}$ . In addition the efficiency ratios cannot be treated separately and they have a common factor, namely  $\varepsilon_{B_s^0 \rightarrow J/\psi \bar{K}^{*0}}^{MC}$ . The averaged  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$  is given by:

$$\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0}) = (4.14 \pm 0.18 \pm 0.26 \pm 0.24) \times 10^{-5}. \quad (4.48)$$

The first two uncertainties are statistical and systematic and the last is due to  $f_d/f_s$ . The result of Eq. 4.48 is in good agreement with previous measurements [92].

## 4.4 Results

After selecting the data as described in Section 4.1 the angular model of Section 4.2 is fitted to the data. The results of the fit are presented in

parameter	value	statistical uncertainty
$A_{CP}^0$	-0.063	(+0.063, -0.063)
$A_{CP}^{\parallel}$	+0.155	(+0.159, -0.157)
$A_{CP}^{\perp}$	-0.087	(+0.097, -0.097)
$A_{CP}^S$	+0.129	(+0.098, -0.097)
$f_0$	+0.481	(+0.024, -0.025)
$f_{\parallel}$	+0.183	(+0.028, -0.027)
$\delta_{\parallel}$	-2.766	(+0.163, -0.166)
$\delta_{\perp}$	-0.067	(+0.113, -0.116)
$f_S^1$	+0.567	(+0.066, -0.077)
$f_S^2$	+0.098	(+0.034, -0.027)
$f_S^3$	+0.045	(+0.045, -0.030)
$f_S^4$	+0.578	(+0.111, -0.120)
$\delta_S^1$	+0.265	(+0.136, -0.131)
$\delta_S^2$	-0.532	(+0.229, -0.195)
$\delta_S^3$	-1.682	(+0.189, -0.240)
$\delta_S^4$	-2.070	(+0.142, -0.160)

TABLE 4.7: Best fit of the parameters of interest.

the current section. Section 4.4.2 addresses systematic uncertainties. The likelihood function is profiled against the parameters of interest in Section 4.4.3 to gain confidence on the validity of the minimum. Finally a pseudo-experiment study is performed in Section 4.4.4 to validate the fitting procedure and determine fit bias.

#### 4.4.1 Parameters of Interest

Table 4.7 summarizes the best fit values of the angular fit whereas Figure 4.12 show the projections of the fitted angular PDF on the data. Notice the destructive interference of the SP-wave. The correlation matrix of the angular fit can be found in Table A.8.

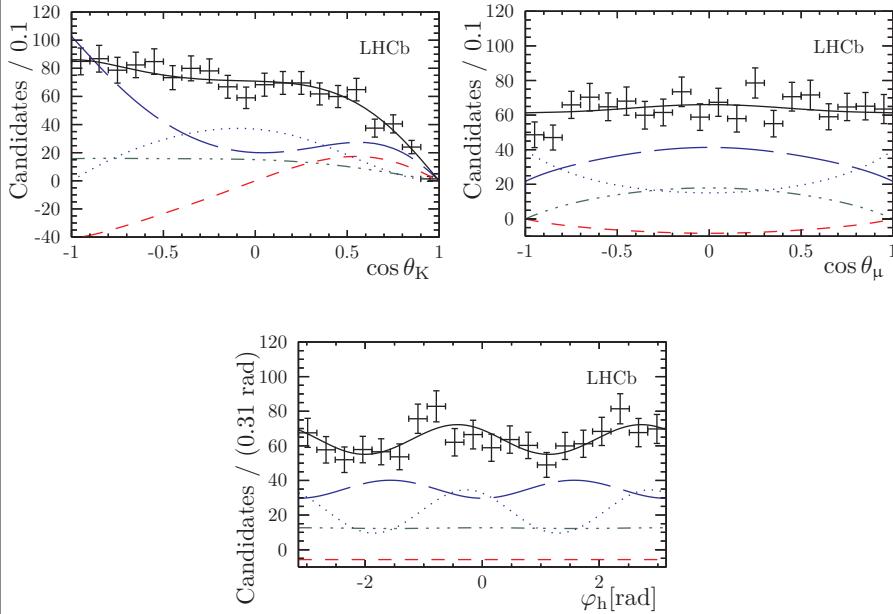


FIGURE 4.12: Fitted angular PDF projections (black solid) on top of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  background subtracted data (crosses). Blue dashed: P-wave even, blue dotted: P-wave odd plus interference, green dotted: S-wave, red dotted dashed S-P interference.

#### 4.4.2 Systematic Uncertainties

The different source of systematic uncertainties of the  $B_s^0$  angular parameters, as well as on the branching fraction are presented in the current section and are summarized from Table 4.8 to Table 4.12. The total

uncertainties in the above tables is the quadratic sum of all the individual systematic uncertainties and they are expressed as fractions of the standard deviation.

	$\mathcal{A}_{CP}^0$	$\mathcal{A}_{CP}^{\parallel}$	$\mathcal{A}_{CP}^{\perp}$	$\mathcal{A}_{CP}^S$
Mass parameters and $B^0$ contamination	0.001	0.001	—	0.001
D-wave contribution	0.002	0.015	0.002	0.008
Mass— $\cos \theta_\mu$ correlations	0.014	$^{+0.009}_{-0.012}$	0.016	$^{+0.023}_{-0.029}$
Angular acceptance	0.003	0.003	0.002	0.001
Reflection background	0.004	$^{+0.012}_{-0.004}$	0.002	0.001
$C_{SP}$ factors	0.001	0.002	0.002	—
Fit bias	0.013	0.003	0.012	0.038
Statistical uncertainty ( $\sigma$ )	0.063	$^{+0.159}_{-0.157}$	0.097	$^{+0.098}_{-0.097}$
Systematic uncertainty (% of $\sigma$ )	0.310	$^{+0.140}_{-0.130}$	0.210	$^{+0.460}_{-0.500}$

TABLE 4.8: Systematic uncertainties of  $\mathcal{A}_{CP}^i$ .

### Angular acceptance

The efficiency moments as described in Section 4.2.2 are computed with a limited simulation sample and thus their central values are subjected to statistical fluctuations. To account for this limitation the efficiency moments are allowed to float in the angular fit to the data. A multivariate Gaussian, similar to Eq. 4.19 is used to constrain the efficiency moments around their nominal values of Table 4.3. Note that the multivariate Gaussian includes their correlations. In addition this systematic covers potential difference in the acceptance between data and simulated sample that has not been corrected by the iterative procedure described in Section 4.2.3. Furthermore, the constraint of Eq. 4.15 used to force the acceptance positive assumes that the acceptance is zero at  $\cos \theta_K = 1$  for any value of  $(\cos \theta_\mu, \varphi_h)$ . This assumption might not hold exactly. It could be that the value of the acceptance at that point is very small but non-zero. This systematic effect is accounted for by adding a term to Eq. 4.17 and allowing it to float in the above-mentioned fit to the data. The size of the latter effect was found to be small and the above

assumption is almost true. The differences in the parameters of interest between the best fit of Table 4.7 and the fit with this additional degree of freedom is taken as a systematic uncertainty.

	$f_0$	$f_{\parallel}$	$\delta_{\parallel}$	$\delta_{\perp}$
Mass parameters and $B^0$ contamination	—	—	—	—
D-wave contribution	0.004	0.003	—	—
Mass— $\cos \theta_{\mu}$ correlations	0.007	0.006	0.070	$^{+0.020}_{-0.040}$
Angular acceptance	0.003	0.001	0.002	0.001
Reflection background	$^{+0.004}_{-0.003}$	0.002	0.020	0.010
$C_{SP}$ factors	—	0.001	—	—
Fit bias	0.001	—	0.007	0.016
Statistical uncertainty ( $\sigma$ )	$^{+0.024}_{-0.025}$	$^{+0.028}_{-0.027}$	$^{+0.163}_{-0.166}$	$^{+0.113}_{-0.116}$
Systematic uncertainty (% of $\sigma$ )	$^{+0.400}_{-0.380}$	$^{+0.260}_{-0.270}$	$^{+0.450}_{-0.440}$	$^{+0.240}_{-0.380}$

TABLE 4.9: P-wave systematic uncertainties.

### $m_{J/\psi K\pi}$ mass model

Not all the PDF parameters are free to vary in the mass fit to the data as described in Section 4.1. These parameters are fixed to values taken from simulated events. In order to account for systematic uncertainties due to these parameters, the mass fit is repeated 1000 times. In each fit to the data the values of these fixed parameters are drawn from a Gaussian distribution. The mean and width of the Gaussian are taken from a mass fit to the simulated data, which was used to determine these parameters. Correlations among these parameters are taken into account in this procedure. Lastly, for each one of the 1000 different mass fits a new sets of *sWeights* is computed and subsequently the angular fit is performed. The result is one Gaussian distribution for each of the angular parameters of interest. The one standard deviation interval of the previous distribution is assigned as a systematic to that parameter.

In addition, another systematic to account for associating  $B^0$  candidates to the  $B_s^0$  peak is assigned. As mentioned in Section 4.1  $\alpha$  is the parameter controlling the tails of the mass peak and it is fixed during

the mass fit to the data. Complementary to the treatment of the fixed parameters in the above paragraph, the parameter  $\alpha$  of the  $B^0$  peak is further scrutinized. Particularly the mass fit to the data is repeated with  $\alpha$  set to infinity and  $\alpha$  floating. The first case, according to the Hypatia function, implies that any contribution to the tails other than resolution is ignored. These two mass fits yield two more set of *sWeights* and two more angular fits. The difference in the value of the fitted angular parameters between these two case is assigned as a systematic.

	$f_S^1$	$f_S^2$	$f_S^3$	$f_S^4$
Mass parameters and $B^0$ contamination	0.001	—	—	—
D-wave contribution	0.010	0.005	0.008	0.002
Mass— $\cos \theta_\mu$ correlations	$+0.040$ $-0.028$	0.003	$+0.006$ $-0.016$	0.015
Angular acceptance	0.000	0.000	0.000	0.003
Reflection background	0.002	$+0.000$ $-0.001$	—	$+0.002$ $-0.000$
$C_{SP}$ factors	0.002	0.001	0.002	0.001
Fit bias	0.013	0.002	0.006	0.003
Statistical uncertainty ( $\sigma$ )	$+0.066$ $-0.077$	$+0.034$ $-0.027$	$+0.045$ $-0.030$	$+0.111$ $-0.120$
Systematic uncertainty (% of $\sigma$ )	$+0.660$ $-0.430$	$+0.180$ $-0.230$	$+0.260$ $-0.630$	$+0.140$ $-0.130$

TABLE 4.10: S-wave systematic uncertainties.

## Reflection backgrounds

Reflection background estimates have an impact on the mass fit and eventually on the computed set of *sWeights*. These estimates are summarized in Table 4.1, where one can see that they come with a statistical uncertainty as well. To account for the previous uncertainty the estimates of all reflection background species ( $B_s^0 \rightarrow J/\psi K^+ K^-$ ,  $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ ,  $B^0 \rightarrow J/\psi \pi^+ \pi^-$ , and  $\Lambda_b^0 \rightarrow J/\psi p K^-$ ) are shifted, one at a time by one standard deviation. The result is a different set of *sWeights* for each species. Subsequently the angular fit is performed with that set of *sWeights*. The difference in the angular parameters between the previous fit and the nominal angular fit is taken as a systematic uncertainty for the particular reflection background species. Individual

contributions from all species to a certain angular parameter are added in quadrature and assigned as a systematic uncertainty.

### Correlations between mass and angles in the *sPlot* context

Correlations between the  $B_s^0$  invariant mass distribution and the  $\cos \theta_\mu$  angular distribution are taken into account by performing the mass fit in five intervals of  $\cos \theta_\mu$ . Since the choice of five intervals is arbitrary the mass fit is repeated in three and six  $\cos \theta_\mu$  intervals as well. The result is two more sets of *sWeights*. For each set the angular fit is performed and the differences in each of the parameters of interest are added in quadrature. The quadratic sum is assigned as systematic uncertainty.

	$\delta_S^1$	$\delta_S^2$	$\delta_S^3$	$\delta_S^4$
Mass parameters and $B^0$ contamination	0.010	—	—	—
D-wave contribution	0.020	0.030	0.080	0.040
Mass— $\cos \theta_\mu$ correlations	0.050	0.040	0.020	0.010
Angular acceptance	0.003	0.004	0.036	0.005
Reflection background	0.010	0.010	0.030	$^{+0.070}_{-0.040}$
$C_{SP}$ factors	0.010	—	—	0.001
Fit bias	0.063	0.009	0.024	0.013
Statistical uncertainty ( $\sigma$ )	$+0.136$ $-0.131$	$+0.229$ $-0.195$	$+0.189$ $-0.240$	$+0.142$ $-0.160$
Systematic uncertainty (% of $\sigma$ )	$+0.620$ $-0.650$	$+0.230$ $-0.270$	$+0.520$ $-0.410$	$+0.580$ $-0.370$

TABLE 4.11: S-wave systematic uncertainties.

### Fit bias

Potential fit biases of the fitting procedure are addressed in Section 4.4.4 and the corresponding results are reported in Table 4.13. The reported biases on the parameters of interest are assigned as a systematic.

### $C_{SP}$ factors

The choice of S-wave and P-wave models affect the values of the  $C_{SP}$  factors, which enter the angular fit via the SP-wave interference terms

of the angular PDF, see Section 4.2.4. The choice of model has a small but measurable impact on the parameters of interest and it was made such that the chosen model describes the  $m_{K\pi}$  distribution best on the observed data. A conservative systematic to the parameters of interest is assigned by choosing the worst possible model and repeating the angular fit. The change in the fitted parameters from the nominal angular fit is assigned as systematic uncertainty.

### Nuisance $CP$ asymmetries

Asymmetries induced in the  $B_s^0$  meson production and from detector imperfections as described in Section 4.2.5, directly affect the  $\mathcal{A}_{CP}^i$  parameters of interest, see Section 4.2.6. The estimated production and detection asymmetries are allowed to vary in the final angular fit to the data with a Gaussian constrain. This way the impact of their uncertainties on the parameters of interest is built in the likelihood function and reflected in the statistical uncertainty. As a result, no explicit systematic uncertainty is assigned.

Source	$\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_d$	$\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})_\phi$
Efficiency ratio	0.04	0.05
Angular corrections ( $\omega$ )	0.09	0.07
Mass model (effect on the yield)	0.06	0.08
$f_d/f_s$	0.17	—
$\mathcal{B}(\phi \rightarrow K^+ K^-)$	—	0.04
Statistical uncertainty ( $\sigma$ )	0.14	0.19
Systematic uncertainty (% of $\sigma$ )	1.70	0.95

TABLE 4.12: Branching fraction systematic uncertainties. Numbers expressed in units of  $10^{-2}$ .

### D-wave contribution

Any D-wave contribution to the considered  $m_{K\pi}$  window is neglected in the current analysis. However, the impact of D-wave has been studied

in [2]. The systematic uncertainty quoted there is also assigned to the current analysis.

### Uncertainties from $f_d/f_s$ , $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$ , and $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$

The uncertainties from the external parameters  $f_d/f_s$ ,  $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$ , and  $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$  are propagated and taken as systematic uncertainties to the estimation of  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$ .

### Summary

All the above systematic uncertainties related to the angular parameters are summarized from Table 4.8 to Table 4.11. Two sources of systematic uncertainties are found to be dominant. First the size of the simulated sample used to estimate the efficiency moments of the angular acceptance; second the correlation between the  $B_s^0$  invariant mass resolution and the cosine of the helicity angle  $\theta_\mu$ . The total uncertainties on the angular parameters remain largely dominated by the statistical uncertainties. Furthermore, systematic uncertainties related to the branching fraction measurement are shown in Table 4.12.

### 4.4.3 Likelihood Scans

The likelihood function presented in Section 4.2.6 is revisited in the current subsection. This time it is visualized in order to gain more confidence that the fit result presented in Section 4.4.1 comes indeed from a valid minimum of the likelihood function. One can easily get an initial estimate on the shape of the function by expanding it as a Taylor series around the minimum. Note that the constant term of the expansion is irrelevant to the position of the minimum as it moves the value of the likelihood function vertically. Assuming that the data sample  $\vec{x}$  in Eq. 4.30 is large enough and also that the variables  $\vec{p}$  in the same equation are uncorrelated and Gaussian in nature, then the Taylor expansion of Eq. 4.30 has the following parabolic shape:

$$\ln L(\mu; \vec{x}) \simeq \frac{1}{2} s^2, \quad \text{with} \quad s = \left( \frac{\mu - \hat{\mu}}{\sigma_\mu} \right), \quad (4.49)$$

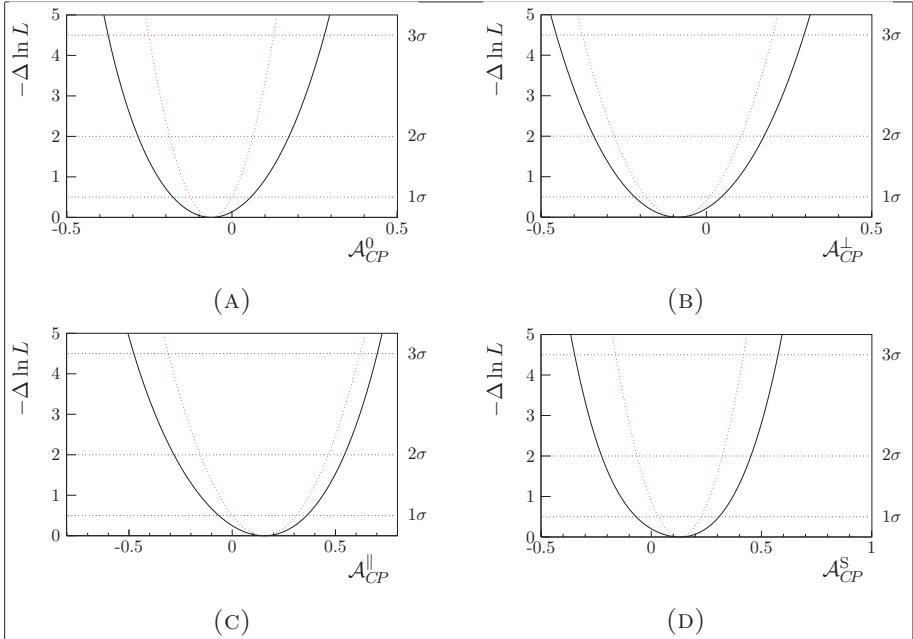


FIGURE 4.13: Profile likelihood scans of  $\mathcal{A}_{CP}$  parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a (symmetric) parabola.

where  $\hat{\mu}$  denotes the values of the parameters of interest,  $\mu$ , at the minimum and  $\sigma_\mu$  is the corresponding statistical uncertainty (note that  $\mu$  and  $\vec{p}$  in the context of Eq. 4.49 and Eq. 4.33 denote the same thing). Since the above assumptions regarding the parameters,  $\mu$  might not be always true, it is useful to define the *profile likelihood* function. The values of the latter function are obtained, given the likelihood of Eq. 4.33, by moving a certain parameter  $\mu_i$  away from its value at the minimum and fixing it to that value. Subsequently, the likelihood function is minimized with respect to the rest of the parameters  $\mu$  obtaining the value of the profile likelihood  $L(\mu_i; \vec{x})$  at the above value of  $\mu_i$ . The procedure is repeated until the desired range of  $\mu_i$  is covered and the values of the profile likelihood,  $L(\mu_i; \vec{x})$ , are obtained.

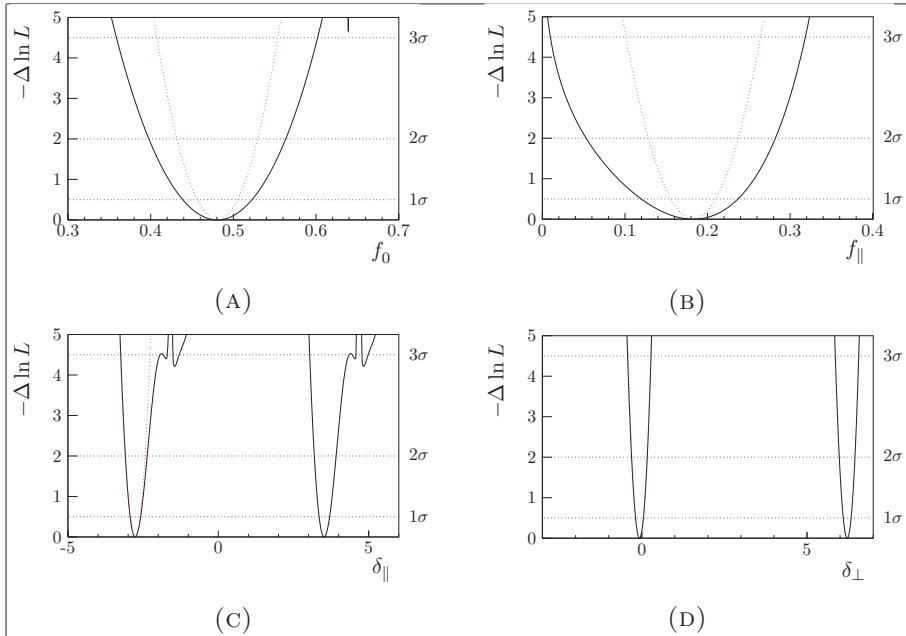


FIGURE 4.14: Profile likelihood scans of P-wave related parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a (symmetric) parabola. The phases appear to have a double minimum. This is actually expected given the symmetries that the fitted PDF of Eq. 4.7 is known to have.

Based on the definition of the profile likelihood two important conclusions can be drawn. First the non Gaussian nature of a parameter of interest can be visualized by how much the corresponding profile likelihood plot deviates from a perfect parabola. A non parabolic shape is an indication of low sample size or high correlations between the parameters. Also a non-smooth profile likelihood might point to mistakes in the implementation of the PDF used to build the likelihood or the presence of almost degenerate ambiguities.

Second, Eq. 4.49 serves as an alternative method of statistical uncertainty estimation for a parameter of interest given a profile likelihood

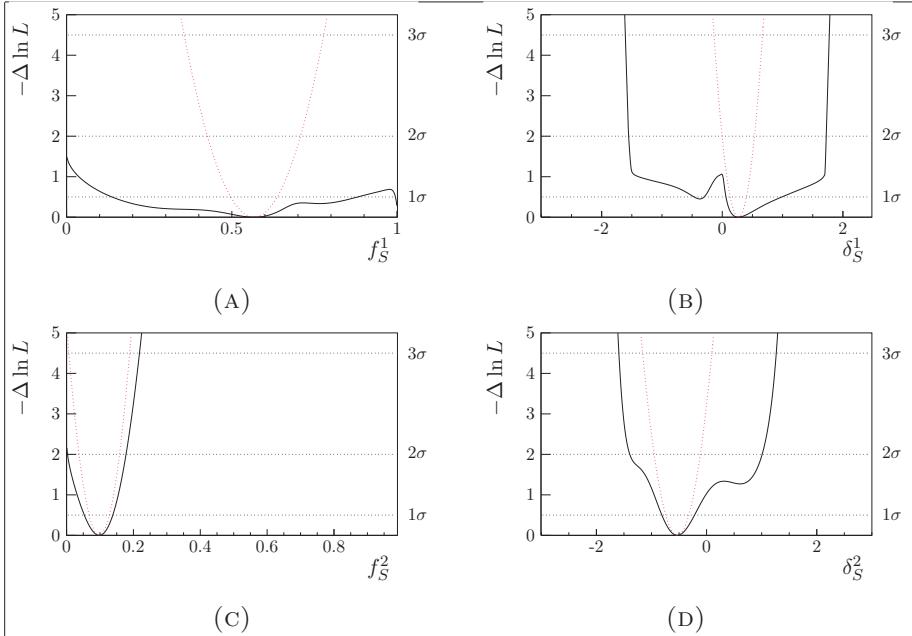


FIGURE 4.15: Profile likelihood scans of S-wave related parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a (symmetric) parabola. There are two interesting features in these plots. First  $f_S^1$  minimum is very broad due to the low number of candidates in that  $m_{K\pi}$  interval. Second the  $f_S^2$  minimum is close to a physical boundary.

plot. Specifically,  $s$  number of standard deviations are defined from the two points where the horizontal line,  $-\Delta \ln L = s$ , intersects the profile likelihood function and projecting down to the horizontal axis. It follows that the uncertainties defined in the the above-mentioned way can be asymmetric left and right of the minimum. The small sample size of the data categories, is depicted in the corresponding profile likelihood plots of the parameters of interest relevant to these data categories. This is also the reason why asymmetric uncertainties are quoted.

Going one step further on the statistical uncertainty interpretation,

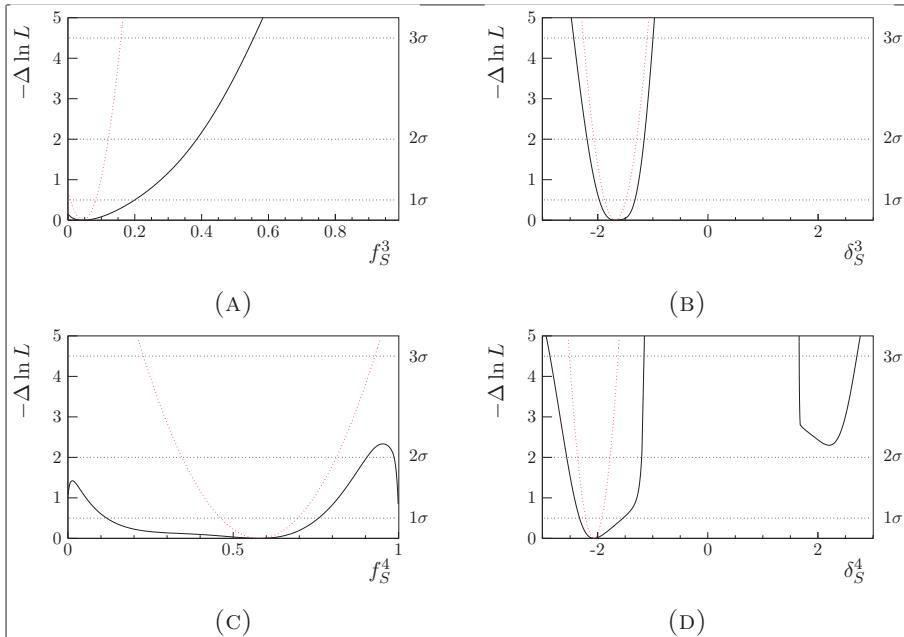
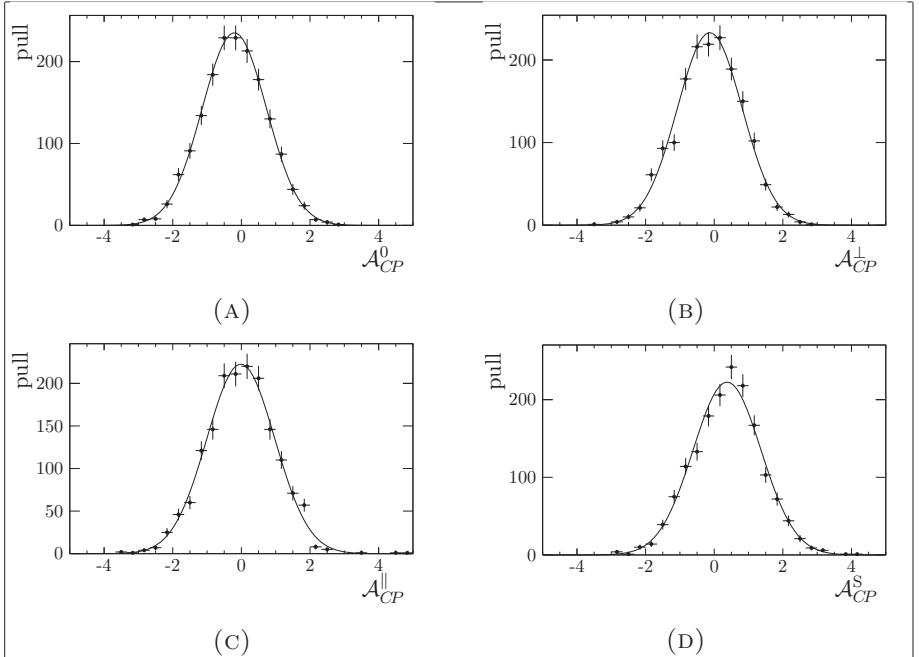


FIGURE 4.16: Profile likelihood scans of S-wave related parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a (symmetric) parabola. There are two interesting features in these plots. First  $f_S^4$  minimum is very broad due to the low number of candidates in that  $m_{K\pi}$  interval. Second the  $f_S^3$  minimum is close to a physical boundary.

it is useful to realize what kind of statement is usually attributed to the statistical uncertainty of one standard deviation, within the frequentist interpretation of probability. Basically the claim is that one standard deviation guarantees that the probability that the true value of the parameter of interest  $\mu^{\text{true}}$  will be inside the interval  $\hat{\mu}_{-\sigma_\mu}^{+\sigma_\mu}$  is 68%. The previous interval is usually quoted as a  $1\sigma$  *confidence interval*. Note that all of the above are true only in case the likelihood of a certain parameters of interest has a parabolic shape or in other words that the CLT theorem of statistics is satisfied.

FIGURE 4.17:  $\mathcal{A}_{CP}^i$  pull distributions.

#### 4.4.4 Pseudo-experiments Study

The current subsection addresses the robustness of the fitting angular PDF. This is done by means of a *pseudo-experiments* study. This procedure can indicate biases introduced to the parameters of interest by the fitting model. These biases are finally assigned as a systematic.

In the current pseudo-experiments study the angular PDF is used to generate data-sets of  $\Omega = (\cos \theta_K, \cos \theta_\mu, \varphi_h)$  distributions. The generated data-set is fitted again with the same PDF and a new set of fitted parameters is obtained. Potential biases on the parameters of interest are estimated based on the following *pull* variable:

$$P^i = \frac{p^i - p^{\text{gen}}}{\sigma^i}, \quad \begin{cases} \sigma^i = \sigma_-^i & \text{if } p^i - p^{\text{gen}} > 0 \\ \sigma^i = \sigma_+^i & \text{if } p^i - p^{\text{gen}} < 0 \end{cases}, \quad (4.50)$$

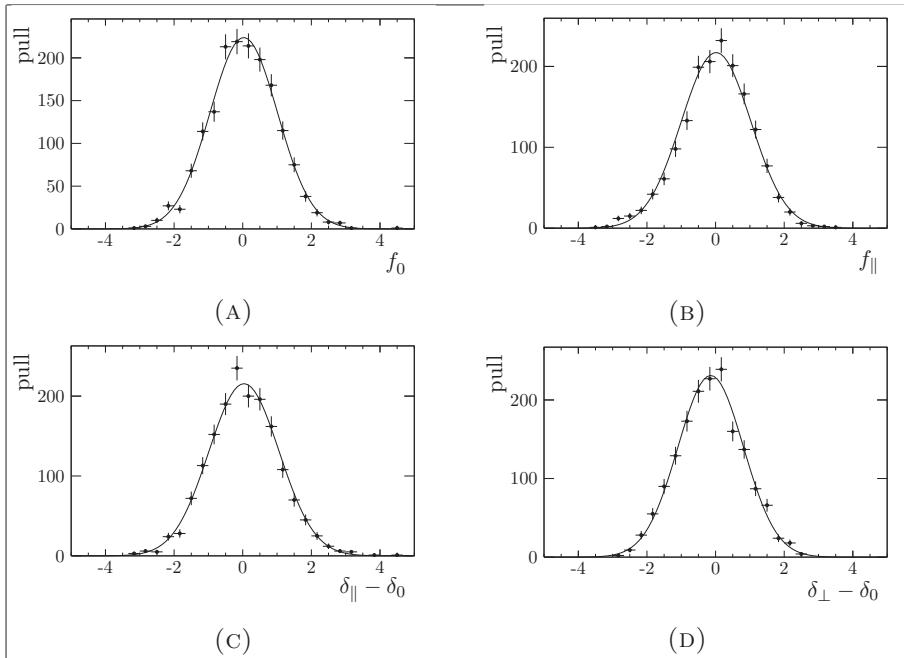


FIGURE 4.18: P-wave pull distributions.

where,  $p^{\text{gen}}$  and  $p^i$  are the generated and fitted values of a given parameter of interest respectively. The upper (lower) uncertainties  $\sigma_+^i$  ( $\sigma_-^i$ ) are used in case  $p^i$  is on the left (right) side of the minimum. Essentially the pull variable expresses the generated-fitted parameter difference in units of one standard deviation.

In a non-biasing fitting model the pull distribution of a certain parameter follows a Gaussian distribution of zero mean, provided that the Central Limit Theorem (CLT) of statistics is satisfied. This can be translated to a useful statement about the parameter,  $\hat{p}$ , and upper (lower) uncertainty,  $\hat{\sigma}_+(\hat{\sigma}_-)$  estimates of Section 4.4.1. Particularly it means that the probability that the interval  $[\hat{p} - \hat{\sigma}_-, \hat{p} + \hat{\sigma}_+]$  contains  $p^{\text{true}}$  is 68%. Or in other words if the real experiment would be repeated  $n$  times, then 68% of these times the  $p^{\text{true}}$  would be contained in the previous interval. In this context the term "true" stands for the value of  $p$  that nature "chose" and it is unknown. The value of  $p^{\text{gen}}$  on the

other hand is known and for the purposes of the pseudo-experiment is a suitable replacement of  $p^{\text{true}}$ .

parameters	mean		rms		bias
	fit	pull	fit	pull	
$\mathcal{A}_{CP}^0$	-0.205	-0.201	0.938	0.937	0.013
$\mathcal{A}_{CP}^{\perp}$	-0.129	-0.126	0.946	0.943	0.012
$\mathcal{A}_{CP}^{\parallel}$	-0.016	-0.020	0.990	0.986	0.003
$\mathcal{A}_{CP}^S$	+0.374	+0.375	0.991	0.986	0.038
$f_0$	+0.030	+0.031	0.986	0.984	0.001
$f_{\parallel}$	+0.015	+0.016	1.016	1.010	0.000
$\delta_{\perp}$	-0.142	-0.145	0.954	0.951	0.016
$\delta_{\parallel}$	+0.037	+0.041	1.023	1.025	0.007
$f_S^1$	+0.161	+0.162	1.107	1.105	0.013
$f_S^2$	-0.039	-0.041	1.044	1.041	0.002
$f_S^3$	+0.193	+0.188	0.920	0.915	0.006
$f_S^4$	-0.025	-0.026	1.123	1.117	0.003
$\delta_S^1$	-0.394	-0.391	1.010	1.000	0.063
$\delta_S^2$	-0.037	-0.037	1.086	1.079	0.009
$\delta_S^3$	+0.138	+0.135	0.921	0.916	0.024
$\delta_S^4$	-0.090	-0.092	0.975	0.965	0.013

TABLE 4.13: Pseudo-experiment study summary.

Furthermore, potential biases to a certain parameter of interest arising from the fitting model can be deduced from the mean of the pull and uncertainty distributions as:

$$b_p = |\langle P \rangle| \langle \sigma^i \rangle, \quad \begin{cases} \sigma^i = \sigma_+^i & \text{if } \langle P \rangle < 0 \\ \sigma^i = \sigma_-^i & \text{if } \langle P \rangle > 0 \end{cases} . \quad (4.51)$$

The pull distributions are shown from Figure 4.17 to Figure 4.19. The corresponding biases are summarized in Table 4.13. Mean and rms of each pull distribution are shown for two case. The "pull" column is computed directly from the pull distributions, whereas the "fit" indicates the fitted Gaussian curve on to of each pull distribution.

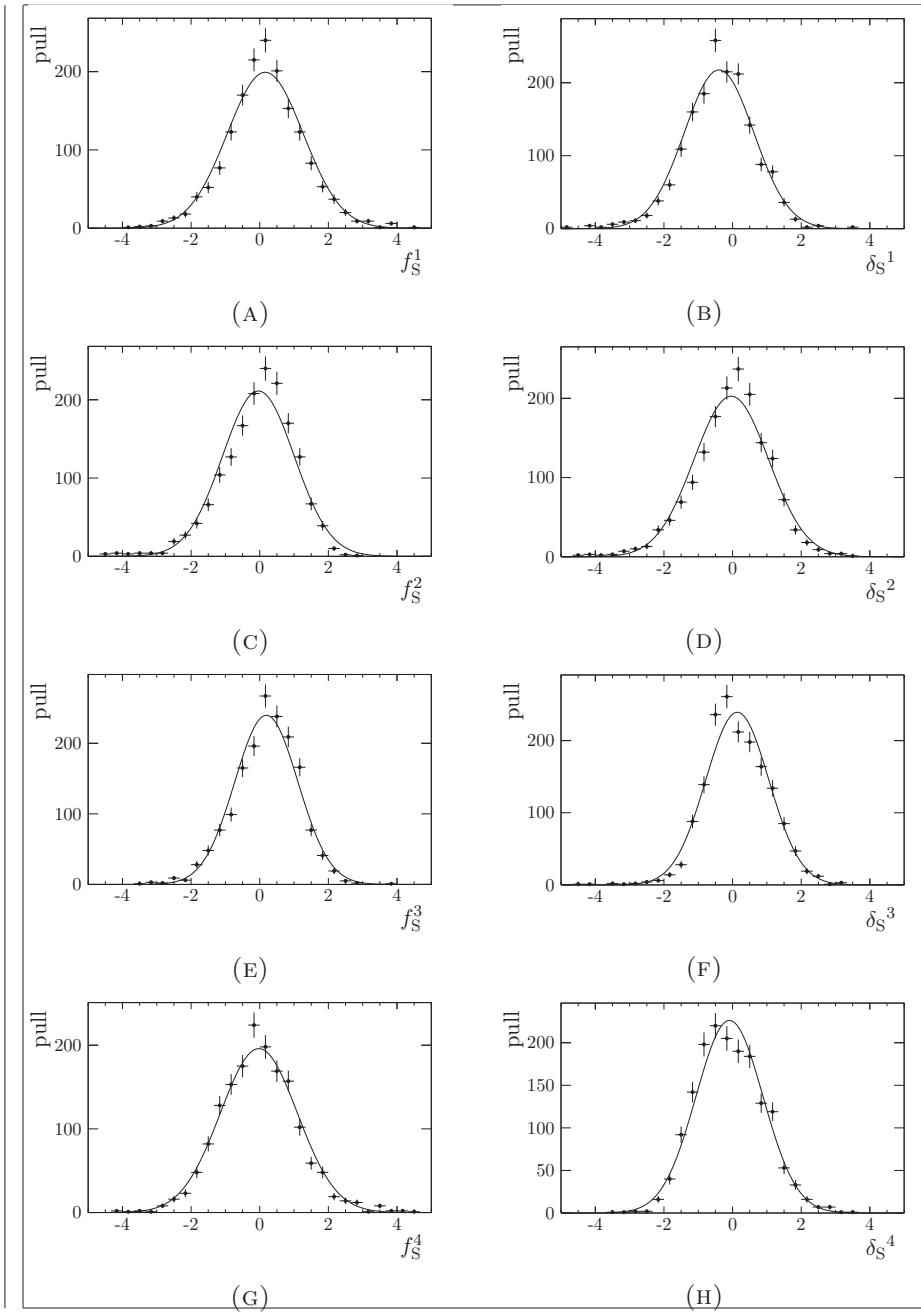


FIGURE 4.19: S-wave pull distributions.

## Chapter 5

# Controlling Penguins in $\phi_S$

As explained in Section 1.4.3 suppressed contributions from penguin topology become increasingly relevant in the pursuit of New Physics from flavor physics experiments. The issue of estimating the contributions from penguin topologies to the total  $B_s^0 \rightarrow J/\psi \phi$  amplitude is addressed in the current chapter. The adopted formalism focuses on involving as many channels similar to  $B_s^0 \rightarrow J/\psi \phi$  as possible. This way the sensitivity to the parameters that quantify penguin topologies increases. From Section 5.1 to Section 5.4 the fundamentals of the necessary formalism are introduced. Details of the penguin parameter estimation are explained in Section 5.4. Results are given in Section 5.5. Note that the strategy followed in this chapter along with the quoted equations are suggested by [48, 51, 93].

### 5.1 Amplitude Structure of $B_s^0 \rightarrow J/\psi \phi$

The  $B_s^0 \rightarrow J/\psi \phi$  decay amplitude can take place via four different topologies. Two of them were already introduced in Figure 1.7 and Figure 1.9; the *color suppressed (C)* and *penguin (P)* topologies respectively. According to [37, 48] the other two types, *penguin-annihilation* and *exchange* illustrated in Figure 5.1, can be neglected given the current experimental precision. It is also assumed that the  $\phi$  meson is a pure ( $s\bar{s}$ ) state, based on [51]. Given these assumptions the  $B_s^0 \rightarrow J/\psi \phi$  amplitude is decomposed in Eq. 5.1, taking into account the relevant CKM elements involved in each topology.

$$\mathcal{A} \left( B_s^0 \rightarrow (J/\psi \phi)_k \right) = V_{us} V_{ub}^* P_u^k + V_{cs} V_{cb}^* \left[ C^k + P_c^k \right] + V_{ts} V_{tb}^* P_t^k, \quad (5.1)$$

where the subscripts in the penguin topologies,  $P$ , denote the flavor of the quark present inside the loop of Figure 1.9. The superscript  $k$  is identical to the one defined in Eq. 4.40 of Chapter 4 and denotes the  $B_s^0 \rightarrow J/\psi \phi$  amplitude polarization. The above expression has to be parameterized in such a way that it is possible to probe the penguin contributions to the  $B_s^0 \rightarrow J/\psi \phi$  decay amplitude. Given the unitarity nature of the CKM matrix and using the Wolfenstein parametrization of Eq. 1.8, the decay amplitude of Eq. 5.1 can be rewritten as:

$$\mathcal{A} \left( B_s^0 \rightarrow (J/\psi \phi)_k \right) = \eta_k \left( 1 - \frac{\lambda^2}{2} \right) \mathcal{A}_k \left[ 1 - \epsilon a_k e^{i\theta_k} e^{i\gamma} \right], \quad (5.2)$$

where the following definitions are used:

$$\mathcal{A}_k \equiv |V_{cb}| \left[ C^k + P_c^k - P_t^k \right], \quad a_k e^{i\theta_k} \equiv R_b \left[ \frac{P_c^k - P_t^k}{C^k + P_c^k - P_t^k} \right], \quad (5.3)$$

with,

$$\epsilon = \frac{\lambda^2}{1 - \lambda^2} \text{ and } R_b = \left( 1 - \frac{\lambda^2}{2} \right) \frac{1}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|. \quad (5.4)$$

The quantities  $a_k$  and  $\theta_k$  in Eq. 5.2 parameterize the penguin contribution to the overall  $B_s^0 \rightarrow J/\psi \phi$  decay amplitude. The hadronic parameters  $\mathcal{A}_k$  are combinations of  $C$  and  $P_q$  decay amplitudes and are discussed in more detail, in Section 5.2. The CKM angle  $\gamma$  is introduced via the  $V_{ub}$  matrix element. The parameter  $\lambda$  here is the one of the CKM matrix parametrization of Section 1.8. The eigenvalue,  $\eta_k$ , of the final state  $f$  is identified with the one in Eq. 1.15. Note the suppression factor  $\epsilon \sim 0.05$  in Eq. 5.2 which implies that penguin contributions to the  $B_s^0 \rightarrow J/\psi \phi$  amplitude are smaller by approximately two orders of magnitude with respect to tree level contributions.

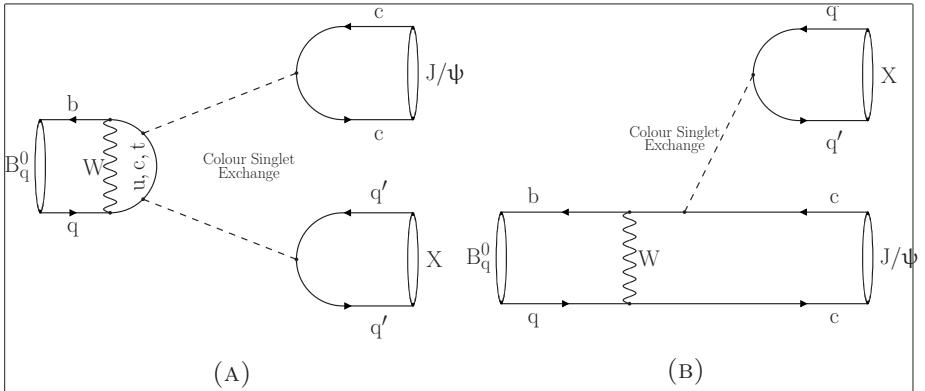


FIGURE 5.1: Higher order, penguin annihilation (A) and exchange (B),  $B_q \rightarrow J/\psi X$  decay topologies. Figures from [37].

## 5.2 Hadronic Factors and $SU(3)$ Symmetry

Hadronic factors  $\mathcal{A}_k$  are introduced by the  $H_k$  observable of Eq. 5.7. Despite the fact that these factors can be calculated theoretically, as for example in [37]. The calculations involve the computation of  $B_q \rightarrow J/\psi X$  decay amplitudes which is difficult. The difficulty comes from the *hadronization* process, which is the process of quarks forming bound states, hadrons, such as the  $J/\psi$  and  $\phi$ . During this process quarks might interact with each other via the strong interaction. This class of strong interactions, also referred to as *long distance QCD effects* cannot be computed within the context of *perturbation theory* [37, 48]. The computation of the  $B_q \rightarrow J/\psi X$  decay amplitude easier under the *factorization* assumption [94–97]. However, according to [37, 48], the power of this approach is limited in the case of  $B_q \rightarrow J/\psi X$  decays and corrections are required to account for *non-factorizable* effects during the process of hadronization. Hence, the hadronic factors  $\mathcal{A}_k$  computed from theory suffer from large uncertainties.

The strategy described in the current chapter avoids, as much as possible, the use of hadronic factors  $\mathcal{A}_k$  from theory. However, the  $H_k$  observable, does require the explicit calculation of ratios of hadronic

factors. Thus on the one hand introducing the previous observable provides additional information for estimating the penguin parameters, on the other hand it introduces further uncertainty from non-factorizable QCD effects. As will be explained in Section 5.4 introducing the previous the  $H_k$  observable can also be avoided.

An important feature that the current chapter makes use of is the  $SU(3)_F$  flavor symmetry [98, 99]. In a naive interpretation,  $SU(3)_F$  flavor symmetry implies that the three lightest quarks (u,d,s) are indistinguishable by the strong interaction. The last symmetry is broken by the fact that the above-mentioned quarks have indeed different mass. However, these differences are small with respect to the hadronization scale  $\Lambda_{\text{QCD}}$  of the strong interaction, and as a result  $SU(3)_F$  remains a useful approximate symmetry. Particularly, in the case of  $B_q$  meson decays, hadronic parameters from different decay modes can be related to one another and thus their calculation can be avoided. The penguin parameter estimation of the current chapter is an example of  $SU(3)_F$ .

According to [100, 101] the estimated amount of  $SU(3)_F$  breaking is about 20%. The impact of the symmetry breaking has implications in the penguin parameter estimation of the current chapter. These implications are addressed in Section 5.5.1. Other  $SU(3)_F$  breaking estimates can be found in [28, 102].

### 5.3 Formalism

Following [51] the penguin parameters  $\alpha_f$  and  $\beta_f$  of Eq. 5.2 can be related to the  $CP$  asymmetries of Eq. 1.16 as follows:

$$\mathcal{A}_{CP}^{\text{dir}}(B_q \rightarrow (f)_k) = \frac{a_k \sin \theta_k \sin \gamma}{1 - 2a_k \cos \theta_k \cos \gamma + a_k^2}, \quad (5.5a)$$

$$\begin{aligned} \mathcal{A}_{CP}^{\text{mix}}(B_q \rightarrow (f)_k) &= \\ &= \eta_k \left[ \frac{\sin \phi_q - 2a_k \cos \theta_k \sin(\phi_q + \gamma) + a_k^2 \sin(\phi_q + 2\gamma)}{1 - 2a_k \cos \theta_k \cos \gamma + a_k^2} \right], \end{aligned} \quad (5.5b)$$

where  $\phi_q$  represents the Standard Model prediction for either the weak phase  $\phi_s$  of Eq. 1.24b, or the equivalent weak phase  $\phi_d$  in the  $B^0$  meson [28]. Note the notation conversion regarding the direct  $CP$  asymmetry,  $\mathcal{A}_{CP}^{\text{dir}}(B_q \rightarrow (f)^k) = \mathcal{A}_{CP}^k$ , as quoted in the current chapter and in Section 4.2.6 respectively. Equations 5.5 form a system in terms of  $(a_k, \theta_k)$  that can in principle be solved. After an estimation of  $(a_k, \theta_k)$  is available the penguin shift to  $\phi_s$  could be determined as:

$$\tan(\Delta\phi_s^k) = \frac{2a_k \epsilon \cos \theta_k \sin \gamma + \epsilon^2 a_k^2 \sin 2\gamma}{1 + 2\epsilon a_k \cos \theta_k \cos \gamma + \epsilon^2 a_k^2 \cos 2\gamma}, \quad (5.6)$$

where  $\Delta\phi_s^k$  denotes the polarization dependent version of  $\Delta\phi_s^{\text{SM, peng}}$  of Eq. 1.26. Note that in the case where Eq. 5.5 refers to the  $B_s^0 \rightarrow J/\psi \phi$  channel, the penguin suppression factor  $\epsilon$  needs to be applied. Thus the transformation  $(a_k \rightarrow -\epsilon a_k)$  has to be performed.

As already mentioned in the introduction, the aim is to increase the precision on  $\Delta\phi_s$  by involving channels similar to  $B_s^0 \rightarrow J/\psi \phi$  in the penguin parameter estimation based on the  $SU(3)$  quark symmetry. Thus, the estimation of  $(a_k, \theta_k)$  will come from a  $\chi^2$  fit to the measured values of the observables of Eq. 5.5 where the channels  $B_s^0 \rightarrow J/\psi \bar{K}^0$  and  $B^0 \rightarrow J/\psi \rho^0$  enter the fit. Details in Section 5.4.3.

### Information from Branching ratios

Additional information from branching ratios can be exploited via the  $H$  observable [93], defined by the following equation:

$$H_k \equiv \frac{1}{\epsilon} \left| \frac{\mathcal{A}_k}{\mathcal{A}'_k} \right|^2 \frac{\text{PhSp}(B_s^0 \rightarrow J/\psi \phi)}{\text{PhSp}(B_q \rightarrow f)} \frac{\tau_{B_s^0}}{\tau_{B_q}} \frac{\mathcal{B}(B_q \rightarrow f)_{\text{theo}}}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)_{\text{theo}}} \frac{f'_k}{f_k}, \quad (5.7)$$

where the superscript prime(') labels quantities related to  $B_q \rightarrow f$ . The subscript "theo" is there to distinguish between definitions of branching fraction, as shown in Eq. 5.11. The polarization fractions  $f_k$  are identical to the ones defined in Eq. 4.40. The average decay time of the  $B_s^0$  and  $B_q$  is denoted by  $\tau_{B_s^0}$  and  $\tau_{B_q}$  respectively. The phase space factor  $\text{PhSp}(X)$  is defined as:

$$\text{PhSp} (B_q \rightarrow J/\Psi X) = [m_{B_q} \Phi (m_{J/\Psi}/m_{B_q}, m_X/m_{B_q})]^3, \quad (5.8)$$

where

$$\Phi (x, y) = \sqrt{(1 - (x + y)^2)(1 - (x - y)^2)}, \quad (5.9)$$

is the standard two body decay phase space function.

The  $H_k$  observable is related to the penguin parameters [93] as:

$$H_k = \frac{1 - 2a_k \cos \theta_k \cos \gamma + a_k^2}{1 + 2\epsilon a'_k \cos \theta_k \cos \gamma + \epsilon^2 a_k^2}. \quad (5.10)$$

The above observable is useful for decay channels, such as  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$ , where  $\mathcal{A}_{CP}^{\text{mix}}$  vanishes, as it provides the second equation necessary to solve for  $(a_k, \theta_k)$ . Such channels where  $B_s^0 - \bar{B}_s^0$  mixing is not active, are called *flavor specific*. In that case only the decays  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\Psi K^{*0}$  are possible and thus neither  $B_s^0$  nor  $\bar{B}_s^0$  can first oscillate and then decay to the above channels respectively. Despite providing the second equation the  $H_k$  observable relies on external input, specifically the hadronic quantities  $\mathcal{A}_k$ . The latter come from theoretical calculations and introduce uncertainty when evaluating the  $H_k$ , as explained in Section 5.2. For this reason the  $H_k$  observable is not a clean quantity and thus not always preferred.

Lastly, note that the  $H_k$  observable is constructed in terms of the theoretical branching fractions defined at zero decay time, which differ from the experimentally measured time-integrated branching fractions [103] due to the non-zero decay-width difference,  $\Delta\Gamma_s$ , of the  $B_s^0$  meson [40]. The necessary conversion factor is:

$$\frac{\mathcal{B}(B \rightarrow f)_{\text{theo}}}{\mathcal{B}(B \rightarrow f)_{\text{exp}}} = \frac{1 - y_s^2}{1 - y_s \cos(\phi_s^{\text{SM, tree}})}, \quad (5.11)$$

with  $y_s = \Delta\Gamma_s/2\Gamma_s$ . The parameter  $\Gamma_s$  is introduced in Section 1.3. The exact numbers for the  $CP$ -even,  $CP$ -odd, and flavor specific final states are reported in Section 9.1 of [2].

### Invoking SU(3) Symmetry

The  $SU(3)_F$  symmetry is used to relate channels like  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B_s^0 \rightarrow J/\psi \rho^0$  to  $B_s^0 \rightarrow J/\psi \phi$ . In practice this relation is implemented by assuming that the penguin parameters  $(a_k, \theta_k)$  are the same for all the above channels, as shown in Eq. 5.18. In addition  $SU(3)_F$  is also exploited in hadronic parameter ratios between different channels, as explained in Section 5.4.3.

## 5.4 Estimating Penguin Parameters

In the recent LHCb publication [1] of the  $\phi_s$  measurement one can find the best fit value for the  $\lambda_{J/\psi \phi}$  parameter, mentioned in Eq. 1.15. These two parameters are summarized as in the following equation:

$$\phi_s^{J/\psi \phi} = -0.058 \pm 0.049(\text{stat}) \text{ rad}, \quad (5.12a)$$

$$\lambda^{J/\psi \phi} = +0.964 \pm 0.019(\text{stat}), \quad (5.12b)$$

and can be used to estimate the penguin parameters  $(a, \theta)$ . Based on Eq. 1.16 and according to the formalism in Section 5.3, the  $\Delta\phi_s$  estimation due to penguin topologies based on  $B_s^0 \rightarrow J/\psi \phi$  decays comes with an uncertainty of about 0.05 rad, see Eq. 5.125 of [37], which is not precise enough given the uncertainty of  $\phi_s^{J/\psi \phi}$ . This is related to the suppression factor  $\epsilon$ , mentioned in Section 5.1, which the penguin topologies receive in the  $B_s^0 \rightarrow J/\psi \phi$  decay.

In order to increase the precision on the penguin parameters the  $SU(3)$  flavor symmetry is invoked such that more channels similar to  $B_s^0 \rightarrow J/\psi \phi$  are involved in the computation of  $(a_k - \theta_k)$ . Note that the information from these channels entering through the observables in Eq. 5.5 and Eq. 5.7 has to be polarization dependent. This is due to the fact that new physics dynamics might enter in a different way in each polarization, as mentioned in [37]. This section introduces the additional channels,  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B^0 \rightarrow J/\psi \rho^0$  and provides details of the fitting strategy to estimate the penguin parameters.

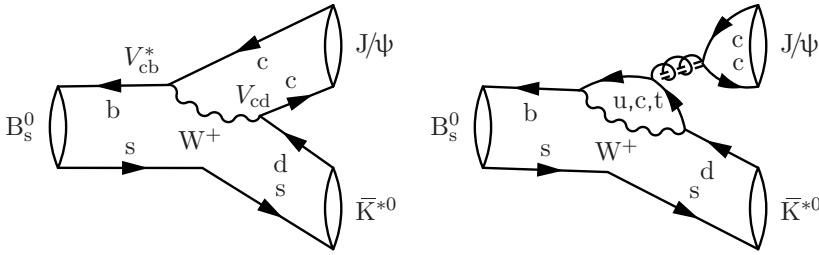


FIGURE 5.2: Leading order diagrams of the decay  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ . Left: Color-suppressed tree topology. Right: Penguin topology.

### 5.4.1 The $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ Channel

The  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  channel is a flavor specific decay with the same topology as  $B_s^0 \rightarrow J/\psi \phi$ . The corresponding diagrams are illustrated in Figure 5.2. The  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  amplitude is parameterized following the same concept as in the case of  $B_s^0 \rightarrow J/\psi \phi$  resulting in [48]:

$$\mathcal{A}(B_s^0 \rightarrow (J/\psi \bar{K}^{*0})_k) = -\lambda \mathcal{A}'_k [1 - a'_k e^{i\theta'_k} e^{i\gamma}], \quad (5.13)$$

where primed ' quantities from here are associated with the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay only. Note the absence of the suppression factor  $\epsilon$  in Eq. 5.13, which implies that the penguin diagram contributes as much as the color suppressed tree diagram does to the total amplitude, contrary to the case of the  $B_s^0 \rightarrow J/\psi \phi$  decay.

As mentioned in Section 5.3 the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  channel provides access to  $\mathcal{A}_{CP}^{\text{dir}}$  only. Thus, additional information is required, via the  $H'_k$  observable, in order to probe  $a_k$  and  $\theta_k$ . Both observables are based on measurements that are described in Chapter 4 of the current thesis. The first observable is reported in Table 4.7, with the associated systematic quoted in Table 4.8. The  $H'_k$  observable, constructed from  $\mathcal{B}(B_s^0 \rightarrow J/\psi \bar{K}^{*0})$ , is reported in Eq. 4.48 and is determined to be:

$$H'_0 = 0.99 \pm 0.07 \text{ (stat)} \quad 0.06 \text{ (syst)} \quad 0.27 \text{ (hadr)}, \quad (5.14a)$$

$$H'_{\parallel} = 0.91 \pm 0.14 \text{ (stat)} \quad 0.08 \text{ (syst)} \quad 0.21 \text{ (hadr)}, \quad (5.14b)$$

$$H'_{\perp} = 1.47 \pm 0.14 \text{ (stat)} \quad 0.11 \text{ (syst)} \quad 0.28 \text{ (hadr)}. \quad (5.14c)$$

The last uncertainty is due to the hadronization factor  $|\mathcal{A}_k/\mathcal{A}'_k|$  necessary for the construction of  $H'_k$ . The hadronization factors are calculated theoretically based on [36]. The exact numbers used can be found in [37]. Note that the uncertainty on  $H'_k$  is dominated by these factors.

$$\Delta\phi_{s,0}^{J/\psi\phi} = +0.001^{+0.100}_{-0.033} \text{ rad}, \quad (5.15a)$$

$$\Delta\phi_{s,\parallel}^{J/\psi\phi} = +0.031^{+0.059}_{-0.052} \text{ rad}, \quad (5.15b)$$

$$\Delta\phi_{s,\perp}^{J/\psi\phi} = -0.046^{+0.022}_{-0.028} \text{ rad}, \quad (5.15c)$$

Prior to any combination the penguin parameters have been estimated based on the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  channel only [2], resulting in Eq. 5.15, where most of the uncertainty is statistical in nature. The  $\chi^2$  fit performed here is similar to the one described in Section 5.4.3, hence details of the fitting strategy are postponed until later.

### 5.4.2 The $B^0 \rightarrow J/\psi \rho^0$ Channel

The topology of the  $B^0 \rightarrow J/\psi \rho^0$  decay is shown in Figure 5.3. The amplitude structure is identical to that of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  mode [93]:

$$\mathcal{A}\left(B^0 \rightarrow \left(J/\psi \rho^0\right)_k\right) = -\lambda \mathcal{A}_k'' \left[1 - a_k'' e^{i\theta_k''} e^{i\gamma}\right], \quad (5.16)$$

where the double primed symbol  $''$  from here on labels parameters related to the  $B^0 \rightarrow J/\psi \rho^0$  decay. Note the absence of the suppression factor  $\epsilon$ .

The  $B^0 \rightarrow J/\psi \rho^0$  channel provides access to both  $\mathcal{A}_{CP}^{\text{dir}}$  and  $\mathcal{A}_{CP}^{\text{mix}}$  since the final state  $J/\psi \rho$ , just like  $J/\psi \phi$  (see Section 1.4), is an admixture of  $CP$ -odd and  $CP$ -even eigenstates and thus  $B^0 - \bar{B}^0$  oscillations are active. Both of the above observables are measured in the time dependent angular analysis of  $B^0 \rightarrow J/\psi \pi^+ \pi^-$  decays from LHCb [104]. The

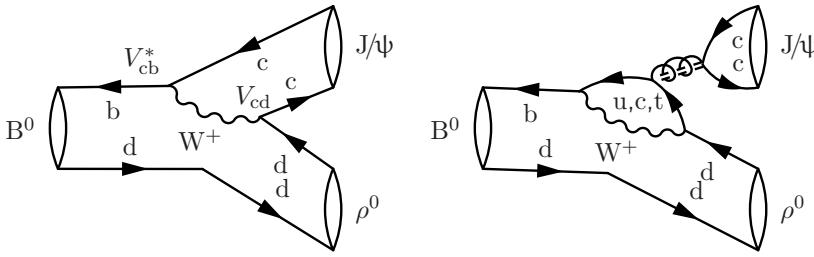


FIGURE 5.3: Leading order tree diagram of the decay  $B^0 \rightarrow J/\psi \rho^0$ .

penguin parameters as determined in Section 5.5.3. of [37] are shown in the following equations:

$$\Delta\phi_{s,0}^{J/\psi\phi} = -0.000^{+0.011}_{-0.014} \text{ rad}, \quad (5.17a)$$

$$\Delta\phi_{s,\parallel}^{J/\psi\phi} = +0.001^{+0.012}_{-0.017} \text{ rad}, \quad (5.17b)$$

$$\Delta\phi_{s,\perp}^{J/\psi\phi} = +0.003^{+0.012}_{-0.016} \text{ rad}, \quad (5.17c)$$

where most of the uncertainty is statistical in nature. Note the small uncertainty on  $\Delta\phi_{s,k}^{J/\psi\phi}$  compared to Eq. 5.15. According to [37, 48] this is attributed to the particular value of the penguin parameter  $\theta_k''$ , which is  $\sim 90^\circ$  (while in the case of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ ,  $\theta_k' \sim 10^\circ$ ). Specifically, as it can be seen in Eq. 5.6, the algebraic structure of the same equation is such that,  $\tan(\Delta\phi_s^k)$  becomes minimal for values of  $\theta_k$  around odd multiples of  $\pi/2$ .

### 5.4.3 Fitting Strategy

The parameters  $(a_k, \theta_k)$  of Eq. 5.2, which quantify the penguin topology contributions to the  $B_s^0 \rightarrow J/\psi \phi$  decay, are estimated by means of a  $\chi^2$  fit. The  $\chi^2$  is defined after some assumptions have been made. First, perfect  $SU(3)_F$  symmetry between the  $B_s^0 \rightarrow J/\psi \phi$  and the additional channels  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B^0 \rightarrow J/\psi \rho^0$  is assumed, implying:

$$a_k = a'_k = a''_k \quad \theta_k = \theta'_k = \theta''_k. \quad (5.18)$$

Second, the hadronic ratios necessary for building  $H_k$  are assumed to be the same between the two additional channels  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B^0 \rightarrow J/\psi \rho^0$ , which translates in:

$$|\mathcal{A}_k/\mathcal{A}'_k| = |\mathcal{A}_k/\mathcal{A}''_k|. \quad (5.19)$$

Doing so makes it possible to avoid using theoretical calculations of the hadronic factors. Instead the single hadronic ratio, based on the assumption of Eq. 5.19, is directly estimated from the fit, since it is treated as a free parameter. This way the uncertainty from the theoretical calculations on the hadronic parameters does not enter the fit. This choice is supported by [37, 48] where it can be seen that the hadronic parameters estimated from experimental data are more precise than the ones coming from theoretical calculations.

Under the above assumptions the observables related to the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $B^0 \rightarrow J/\psi \rho^0$  channels along with their corresponding experimental measurements, as described in Section 5.4.2 and Section 5.4.1, are used to define a  $\chi^2$  in the following equation:

$$\chi_k^2 = \sum \left( \frac{O_k^{\text{theo}} - O_k^{\text{exp}}}{\sigma(O_k^{\text{exp}})} \right)^2, \quad (5.20)$$

where,

$$O_k \in \left\{ \left( \mathcal{A}_{CP}^{\text{dir}} \right)'_k, H'_k, \left( \mathcal{A}_{CP}^{\text{dir}} \right)''_k, \left( \mathcal{A}_{CP}^{\text{mix}} \right)''_k, H''_k \right\}. \quad (5.21)$$

Note also that the CKM angle  $\gamma$  and the weak phase  $\phi_d$  are necessary for computing  $\mathcal{A}_{CPk}^{\text{dir}'}$ ,  $\mathcal{A}_{CPk}^{\text{dir}''}$  and  $\mathcal{A}_{CPk}^{\text{mix}''}$ . The previous parameters are allowed to vary in the  $\chi^2$  fit with a Gaussian constrain around their measured values which are shown in Table 5.1.

Minimizing the above  $\chi^2$  yields the penguin parameters  $(a_k, \theta_k)$ . The results of the fit presented in Section 5.5.

parameter	value	source
$\gamma$	$(73.2^{+6.3}_{-7.0})^\circ$	CKMfitter [102]
$\phi_d$	$0.767 \pm 0.029$ rad	De Bruyn K.[37]

TABLE 5.1: Constrains entering the  $\chi^2$  fit.

## 5.5 Results

The results of the  $\chi^2$  fit to the penguin parameters, described in Section 5.4 are the following:

$$a_0 = 0.01^{+0.10}_{-0.01} \quad \theta_0 = -\left(82^{+262}_{-97}\right)^\circ \quad |\mathcal{A}_0/\mathcal{A}'_0| = 1.212^{+0.075}_{-0.057}, \quad (5.22a)$$

$$a_{\parallel} = 0.07^{+0.11}_{-0.05} \quad \theta_{\parallel} = -\left(84^{+93}_{-129}\right)^\circ \quad |\mathcal{A}_{\parallel}/\mathcal{A}'_{\parallel}| = 1.230^{+0.102}_{-0.078}, \quad (5.22b)$$

$$a_{\perp} = 0.05^{+0.12}_{-0.05} \quad \theta_{\perp} = +\left(53^{+127}_{-232}\right)^\circ \quad |\mathcal{A}_{\perp}/\mathcal{A}'_{\perp}| = 1.038^{+0.083}_{-0.065}. \quad (5.22c)$$

The quoted uncertainties are the quadratic sum of statistical and systematic uncertainties. A 20%  $SU(3)_F$  breaking is assumed for the systematic, according to what was mentioned at the end of Section 5.2, by performing a special fit explained in Section 5.5.1. Two dimensional contours are shown in Figures 5.4 to 5.6, where the parameters of interest,  $(a_k, \theta_k)$ , are transformed to rectangular coordinates for better visibility. There it can be seen that the  $H_k$  observables have a smaller impact on the best fit result due to their associated large uncertainty. Also note the tension between the  $H_k$  observables and the best fit result, which suggests that the assumption Eq. 5.19 is probably not fully satisfied. The best fit estimates of Eq. 5.22 are subsequently translated, using Eq. 5.6, to the phase shifts  $\Delta\phi_{s,k}^{J/\psi\phi}$  as:

$$\Delta\phi_{s,0}^{J/\psi\phi} = 0.000^{+0.009}_{-0.011} \text{ (stat)}^{+0.006}_{-0.007} \text{ (syst)} = 0.000^{+0.010}_{-0.014}, \quad (5.23a)$$

$$\Delta\phi_{s,\parallel}^{J/\psi\phi} = 0.001^{+0.010}_{-0.014} \text{ (stat)}^{+0.006}_{-0.007} \text{ (syst)} = 0.001^{+0.012}_{-0.016}, \quad (5.23b)$$

$$\Delta\phi_{s,\perp}^{J/\psi\phi} = 0.003^{+0.010}_{-0.014} \text{ (stat)}^{+0.006}_{-0.007} \text{ (syst)} = 0.003^{+0.012}_{-0.016}, \quad (5.23c)$$

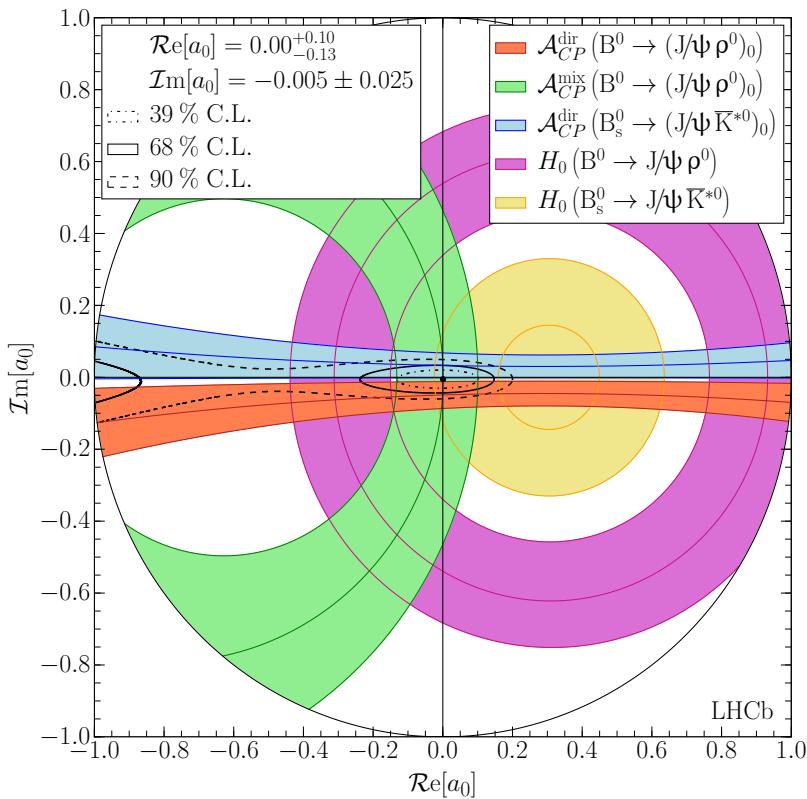
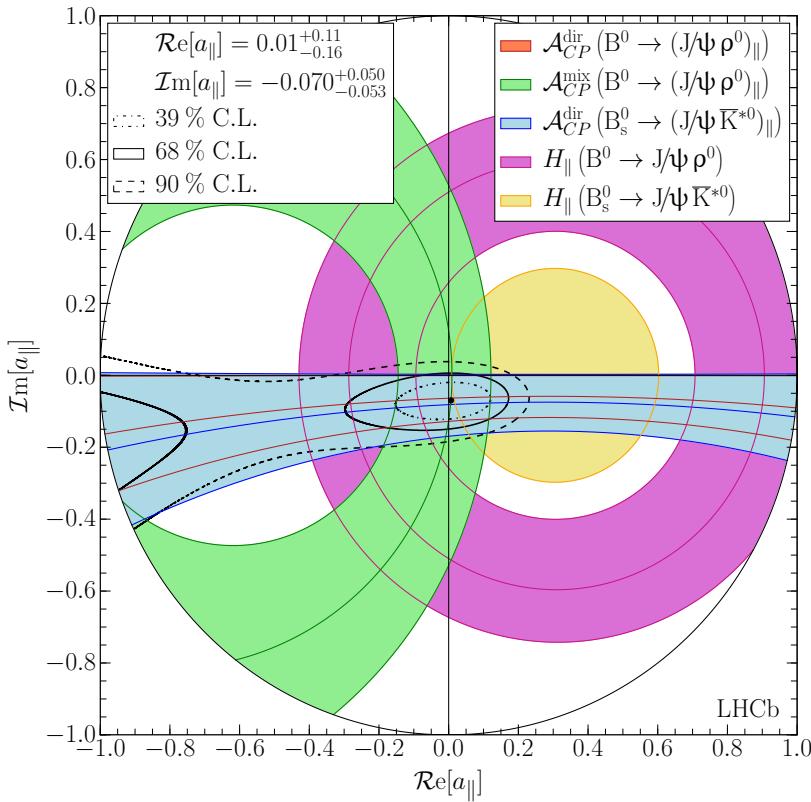


FIGURE 5.4: Confidence levels  $a_0$ ,  $\theta_0$ . Measurements of Eq. 5.21 are illustrated with colored bands, which correspond to one standard deviation. Figure is based on [37] and updated with the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  polarization fractions and  $\mathcal{A}_{CP}^{dir}$  from Table 4.7.

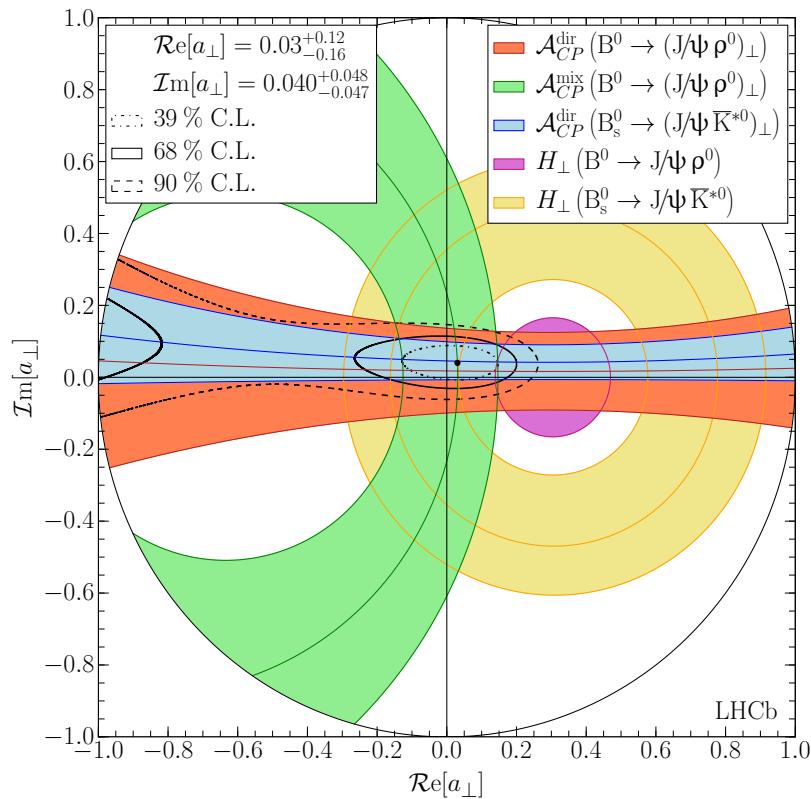
where the uncertainties are statistical and systematic due to SU(3) breaking respectively. Statistical uncertainty dominates the  $\Delta\phi_{s,k}^{J/\psi\phi}$  measurement. Note that the penguin parameters  $(a_k, \theta_k)$  are ratios of hadronic amplitudes. Given that, it is interesting to point out that any factorizable SU(3)<sub>F</sub> breaking effects entering through the assumption of Eq. 5.18 cancel out [93]. Thus the results reported here suffer from non-factorizable SU(3)<sub>F</sub> breaking which are small, according to [37, 48].

FIGURE 5.5: Confidence levels of  $a_{\parallel}$ ,  $\theta_{\parallel}$ .

The previous statement is supported by the values of the fitted hadronic parameter ratios which are close to the ones calculated assuming factorization holds. As mentioned in [37, 48], either the non factorization  $SU(3)_F$  effects or their ratio with respect to the factorizable  $SU(3)_F$  ones are small.

### 5.5.1 Further Crosschecks

There are two important issues that the fitting strategy, as described in Section 5.4.3, does not take into account. Namely,  $SU(3)_F$  breaking

FIGURE 5.6: Confidence levels of  $a_\perp$  and  $\theta_\perp$ .

effects and correlations from the experimental measurement between the observables Eq. 5.21. Both issues are addressed in what follows.

### SU(3) breaking

Potential  $SU(3)_F$  effects manifest themselves in the calculations of the hadronic parameters ratios when computing the  $H_k$  observables and in the assumption of Eq. 5.18. However, due to the strategy followed the hadronic parameter ratios are not affected by  $SU(3)_F$  breaking since they are determined by the fit. As far as the assumption of Eq. 5.18 is concerned, a special fit is performed using only observables related to

$B^0 \rightarrow J/\psi \rho^0$  channel, to investigate the effect of  $SU(3)_F$ . For that, the parameters  $a_k$  and  $\theta_k$  are re-expressed as:

$$a_k \rightarrow \xi_k a''_k, \quad \theta_k \rightarrow \delta_k + \theta''_k. \quad (5.24)$$

The quantities  $\xi_k$  and  $\delta_k$  are allowed to vary in the fit and can thus absorb potential  $SU(3)_F$  effects that will break the assumption of Eq. 5.18. The central values of  $\xi_k$  and  $\delta_k$  assume perfect  $SU(3)_F$  symmetry, meaning  $\xi_k = 1$  and  $\delta_k = 0$ . However, a Gaussian constraint allows them to vary within a range of up to 50% breaking  $SU(3)_F$ . Specifically, the uncertainties within which  $\xi_k$  and  $\delta_k$  are allowed to vary are  $[0, 0.5]$  and  $[0, 40]^\circ$  respectively. Repeating the above-mentioned special fit shows that the value of  $\Delta\phi_s^k$  does not depend on the amount of allowed  $SU(3)_F$  symmetry breaking. This is most likely due to the particular structure of Eq. 5.6, *i.e.* the value of  $a_k$ , or  $\xi_k$  in that case becomes less pronounced for values of  $\theta_k$  which are close to  $90^\circ$ .

### Correlations of polarization states

Correlations between the  $k$  polarizations of the observables entering the  $\chi^2$  fit are ignored in the current analysis. These correlations enter mainly via the  $\mathcal{A}_{CP}^k$  and  $f_k$ . For example, see the correlation matrix Table A.8 of the angular analysis performed in Chapter 4. In order to check the effect of these correlations on the estimation of the penguin parameters  $(a_k, \theta_k)$ , an additional  $\chi^2$  fit is performed. In that fit the correlations are indeed taken into account. The results show a small increase on the uncertainty of the penguin parameter  $\theta_0$ . In addition, the correlated fit result is compatible with the nominal fit of Section 5.4.3, thus correlations are finally ignored. However, with future and more precise measurements correlations might be important and maybe have to be taken into account.

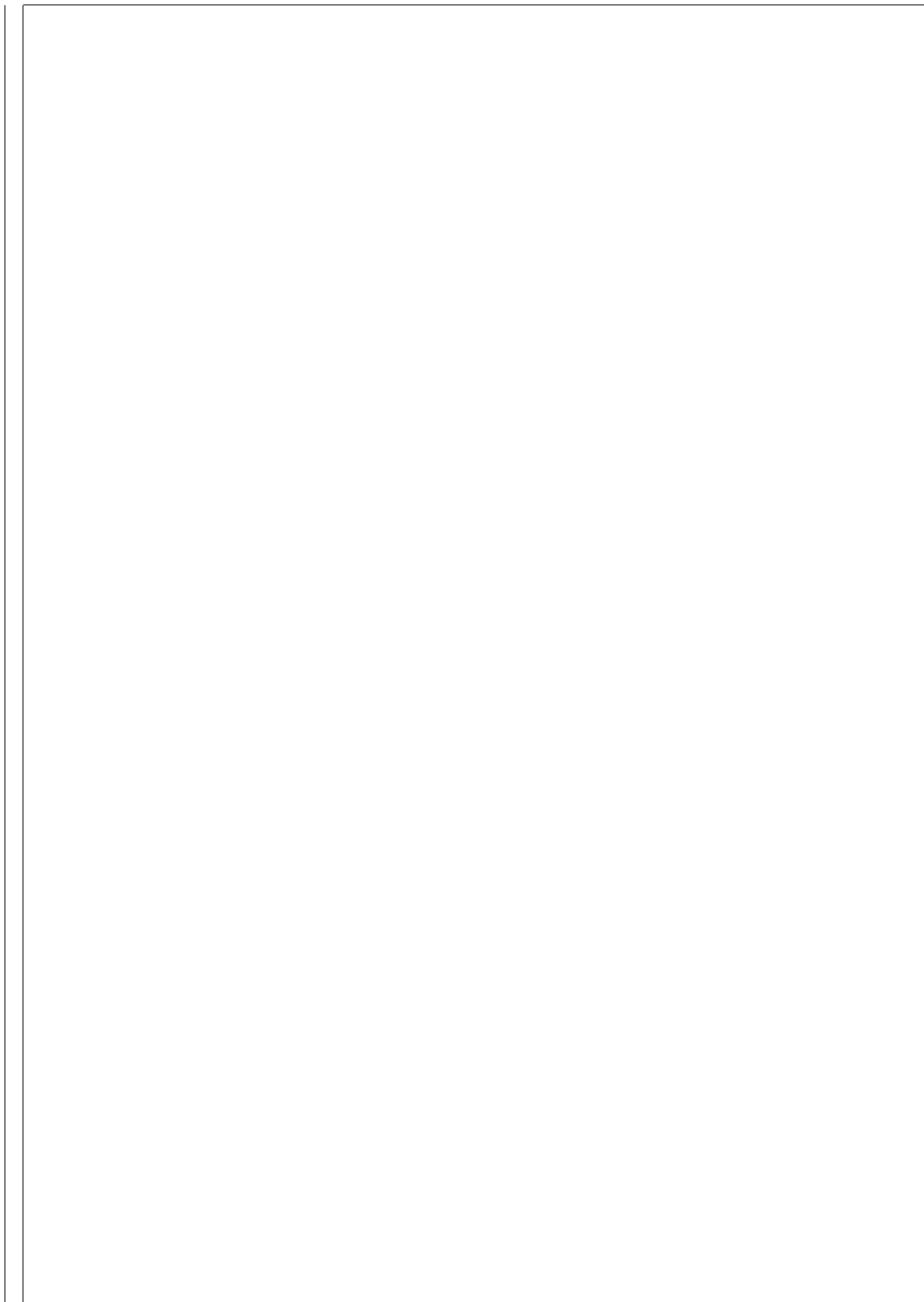
### Conclusions

The shifts  $\Delta\phi_{s,k}^{J/\psi\phi}$  quoted in Eq. 5.23 suggest that contributions of penguin topologies to the  $B_s^0 \rightarrow J/\psi \phi$  decay amplitude are small,  $< 1^\circ$ . Given the also small  $\phi_s$  measured value from LHCb shown in Eq. 1.24a

it becomes mandatory to control penguin contributions in future  $\phi_s$  measurements. Furthermore, the sensitivity from the experimental side is interesting with respect to the Standard Model prediction, see Eq. 1.24. Thus, potential deviations from this prediction would probably play a central role in future and more precise  $\phi_s$  measurements. Increasing the amount of data in the LHC Run 2 might not be enough to yield a significant claim on the presence of physics beyond the Standard Model and hence the upgraded LHCb detector becomes important in the pursuit for New Physics with  $\phi_s$  in the future.

---

---



---

# Appendix A

# Additional Tables

The current appendix contains tables related to the selection described in Section 4.1. Also tables relevant to the angular acceptance of Section 4.2.2 are included. The correlation matrix between the parameters of the angular fit result of Table 4.7 can be found [here](#).

### Candidate selection variables

The symbols used in Table A.1 are explained in the current paragraph. The variables  $\Delta LL$  and  $PNN$  are particle identification variables. The first is the likelihood ratio of two particle ID hypotheses, whereas the particle identification information that the second provides is based on a neural network approach. The  $IsMuon$  variable is another particle identification variable dedicated to muons. The quality of the fitted vertices of two reconstructed tracks is expressed by  $\chi^2_{\text{vtx}}/\text{nDoF}(\text{J}/\psi)$ , with  $\text{nDoF} = 1$ , where  $\text{nDoF}$  is the number of degrees of freedom of that fit. The impact parameter,  $IP$ , is the distance in 3D space between a track and a vertex, *e.g.* the primary vertex mentioned in Section 2.1. A certain type of fake tracks called *ghost tracks* is suppressed by the  $P_{\text{ghost}}(\text{track})$  variable. The variable  $\text{DIRA}(B_s^0)$  is the cosine of the angle between the momentum of the  $B_s^0$  particle and the direction of flight from the best PV to the  $B_s^0$  decay vertex. The  $VS$  variable stands for the separation between a vertex with respect to its associated PV. The nominal values of particle masses are taken from *Particle Data Group* [28], wherever required. Masses are in units of  $\text{MeV}/c^2$ . Cuts labeled with  $(*)$  are introduced for  $\Lambda_b^0 \rightarrow \text{J}/\psi p K^-$  and  $\Lambda_b^0 \rightarrow \text{J}/\psi p \pi^-$  reflection backgrounds suppression.

Decay mode	Cut parameter	Stripping line	Final selection
$J/\psi \rightarrow \mu\mu$	$\Delta LL\mu\pi(\mu)$	$> 0$	–
	$\chi^2_{\text{vtx}}/\text{nDoF}(J/\psi)$	$< 16$	–
	$\chi^2_{\text{IP}}(\mu)$	–	$> 16$
	$p_T(\mu)$	$> 0.5 \text{ GeV}/c$	–
	$\text{IsMuon}(\mu)$	$> 0$	–
	$ m_{\mu^+\mu^-} - m_{J/\psi} $	$< 150$	–
$\bar{K}^{*0} \rightarrow K\pi$	$\Delta LLK\pi(K)$	$> 0$	–
	$p_T(K)$	$> 0.5 \text{ GeV}/c$	$> 0.5 \text{ GeV}/c$
	$P_{\text{ghost}}(\text{track})(K)$	$< 0.8$	–
	$\chi^2_{\text{IP}}(K)$	–	$> 2$
	$\text{PNNK}(K)$	–	$> 0.21$
	$\text{PNNK}(K)/\text{PNNp}(K)$ (*)	–	$> 0.99$
	$\Delta LLK\pi(\pi)$	$< 0$	–
	$p_T(\pi)$	$> 0.5 \text{ GeV}/c$	$> 0.5 \text{ GeV}/c$
	$P_{\text{ghost}}(\text{track})(\pi)$	$< 0.8$	–
	$\chi^2_{\text{IP}}(\pi)$	–	$> 2$
	$\text{PNNK}(\pi)$	–	$< 0.01$
	$\text{PNN}\pi(\pi)/\text{PNNp}(\pi)$ (*)	–	$> 21.9$
	$\chi^2_{\text{vtx}}/\text{nDoF}(\bar{K}^{*0})$	$< 25$	–
	$\chi^2_{\text{DOCA}}/\text{nDoF}(\bar{K}^{*0})$	$< 30$	–
	$m_{K\pi}$	[750, 1900]	–
	$ m_{K\pi} - m_{K^{*0}} $	–	$< 70$
$B_s^0$	$m_{B_s^0}$	[5100, 5700]	[5150, 5650]
	$\chi^2_{\text{vtx}}/\text{nDoF}(B_s^0)$	$< 10$	–
	$\text{DIRA}(B_s^0)$	$> 0.999$	–
	VS	$> 1.5 \text{ mm}$	–
$B^+$ veto	$ m_{(J/\psi, K)} - 5279 $	–	$> 60$
BDT	2011 data/MC (**)	–	$> 0.2$
	2012 data/MC (**)	–	$> 0.12$

TABLE A.1:  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  selection criteria.

moment	2012	2011	difference
$c_{00}^0$	$+3.7522 \pm 0.0033$	$+3.7442 \pm 0.0031$	$-0.0080 \pm 0.0046 (-1.8\sigma)$
$c_{00}^1$	$-2.1273 \pm 0.0119$	$-2.1387 \pm 0.0114$	$-0.0114 \pm 0.0165 (-0.7\sigma)$
$c_{00}^2$	$-1.8893 \pm 0.0179$	$-1.9145 \pm 0.0171$	$-0.0252 \pm 0.0247 (-1.0\sigma)$
$c_{00}^3$	$-0.0218 \pm 0.0213$	$-0.0131 \pm 0.0204$	$+0.0087 \pm 0.0295 (+0.3\sigma)$
$c_{00}^4$	$+0.1397 \pm 0.0244$	$+0.1259 \pm 0.0233$	$-0.0139 \pm 0.0337 (-0.4\sigma)$
$c_{00}^5$	$+0.1191 \pm 0.0268$	$+0.1489 \pm 0.0257$	$+0.0298 \pm 0.0371 (+0.8\sigma)$
$c_{20}^0$	$+0.2950 \pm 0.0093$	$+0.3007 \pm 0.0089$	$+0.0056 \pm 0.0129 (+0.4\sigma)$
$c_{22}^0$	$+0.0397 \pm 0.0082$	$+0.0611 \pm 0.0079$	$+0.0213 \pm 0.0114 (+1.9\sigma)$
$c_{40}^0$	$+0.0829 \pm 0.0093$	$+0.0844 \pm 0.0089$	$+0.0014 \pm 0.0129 (+0.1\sigma)$
$c_{20}^1$	$-0.2259 \pm 0.0162$	$-0.2096 \pm 0.0152$	$+0.0162 \pm 0.0222 (+0.7\sigma)$
$c_{40}^1$	$-0.0670 \pm 0.0164$	$-0.0746 \pm 0.0151$	$-0.0076 \pm 0.0223 (-0.3\sigma)$
$c_{20}^2$	$-0.1037 \pm 0.0224$	$-0.1667 \pm 0.0209$	$-0.0630 \pm 0.0307 (-2.1\sigma)$
$c_{22}^2$	$-0.0351 \pm 0.0166$	$-0.0660 \pm 0.0160$	$-0.0309 \pm 0.0231 (-1.3\sigma)$

TABLE A.2: Comparison between efficiency moments values obtained from the 2011 and 2012  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  simulation samples. The overall  $\chi^2$  over the 13 degrees of freedom is 1.6, corresponding to a  $\chi^2$  probability of 6.5%.

moment	2012	2011	difference
$c_{00}^0$	$+3.7432 \pm 0.0033$	$+3.7476 \pm 0.0033$	$+0.0044 \pm 0.0047 (+0.9\sigma)$
$c_{00}^1$	$-2.0708 \pm 0.0120$	$-2.1206 \pm 0.0119$	$-0.0497 \pm 0.0169 (-2.9\sigma)$
$c_{00}^2$	$-1.8925 \pm 0.0178$	$-1.9287 \pm 0.0179$	$-0.0362 \pm 0.0252 (-1.4\sigma)$
$c_{00}^3$	$-0.0556 \pm 0.0212$	$-0.0179 \pm 0.0216$	$+0.0377 \pm 0.0303 (+1.2\sigma)$
$c_{00}^4$	$+0.1004 \pm 0.0243$	$+0.1249 \pm 0.0250$	$+0.0245 \pm 0.0349 (+0.7\sigma)$
$c_{00}^5$	$+0.1455 \pm 0.0267$	$+0.1471 \pm 0.0279$	$+0.0016 \pm 0.0386 (+0.0\sigma)$
$c_{20}^0$	$+0.2690 \pm 0.0093$	$+0.2908 \pm 0.0093$	$+0.0218 \pm 0.0131 (+1.7\sigma)$
$c_{22}^0$	$+0.0566 \pm 0.0083$	$+0.0651 \pm 0.0080$	$+0.0085 \pm 0.0115 (+0.7\sigma)$
$c_{40}^0$	$+0.0832 \pm 0.0093$	$+0.0898 \pm 0.0094$	$+0.0066 \pm 0.0132 (+0.5\sigma)$
$c_{20}^1$	$-0.1722 \pm 0.0159$	$-0.2050 \pm 0.0165$	$-0.0328 \pm 0.0229 (-1.4\sigma)$
$c_{40}^1$	$-0.0759 \pm 0.0159$	$-0.0674 \pm 0.0174$	$+0.0085 \pm 0.0236 (+0.4\sigma)$
$c_{20}^2$	$-0.1273 \pm 0.0219$	$-0.1227 \pm 0.0235$	$+0.0046 \pm 0.0321 (+0.1\sigma)$
$c_{22}^2$	$-0.0613 \pm 0.0168$	$-0.0724 \pm 0.0162$	$-0.0111 \pm 0.0234 (-0.5\sigma)$

TABLE A.3: Comparison between efficiency moments values obtained from the 2011 and 2012  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  simulation samples. The overall  $\chi^2$  over the 13 degrees of freedom is 2.6, corresponding to a  $\chi^2$  probability of 0.15%.

moment	$\bar{B}_s^0 \rightarrow J/\psi K^{*0}$	$B_s^0 \rightarrow J/\psi \bar{K}^{*0}$	difference
$c_{00}^0$	$+3.7481 \pm 0.0023$	$+3.7455 \pm 0.0023$	$-0.0026 \pm 0.0033 (-0.8\sigma)$
$c_{00}^1$	$-2.1366 \pm 0.0083$	$-2.0982 \pm 0.0084$	$+0.0384 \pm 0.0118 (+3.3\sigma)$
$c_{00}^2$	$-1.9024 \pm 0.0124$	$-1.9116 \pm 0.0126$	$-0.0092 \pm 0.0177 (-0.5\sigma)$
$c_{00}^3$	$-0.0112 \pm 0.0147$	$-0.0317 \pm 0.0151$	$-0.0205 \pm 0.0211 (-1.0\sigma)$
$c_{00}^4$	$+0.1331 \pm 0.0168$	$+0.1146 \pm 0.0175$	$-0.0185 \pm 0.0243 (-0.8\sigma)$
$c_{00}^5$	$+0.1271 \pm 0.0185$	$+0.1384 \pm 0.0193$	$+0.0112 \pm 0.0268 (+0.4\sigma)$
$c_{20}^0$	$+0.3006 \pm 0.0064$	$+0.2820 \pm 0.0066$	$-0.0186 \pm 0.0092 (-2.0\sigma)$
$c_{22}^0$	$+0.0518 \pm 0.0057$	$+0.0612 \pm 0.0058$	$+0.0094 \pm 0.0081 (+1.2\sigma)$
$c_{40}^0$	$+0.0832 \pm 0.0064$	$+0.0864 \pm 0.0066$	$+0.0031 \pm 0.0092 (+0.3\sigma)$
$c_{20}^1$	$-0.2221 \pm 0.0111$	$-0.1900 \pm 0.0115$	$+0.0321 \pm 0.0160 (+2.0\sigma)$
$c_{40}^1$	$-0.0711 \pm 0.0111$	$-0.0720 \pm 0.0118$	$-0.0009 \pm 0.0163 (-0.1\sigma)$
$c_{20}^2$	$-0.1425 \pm 0.0153$	$-0.1315 \pm 0.0161$	$+0.0110 \pm 0.0222 (+0.5\sigma)$
$c_{22}^2$	$-0.0547 \pm 0.0115$	$-0.0674 \pm 0.0117$	$-0.0127 \pm 0.0164 (-0.8\sigma)$

TABLE A.4: Comparison between efficiency moments values obtained from the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  simulation samples. Production periods 2011 and 2012 are merged. The overall  $\chi^2$  over the 13 degrees of freedom is 1.6, corresponding to a  $\chi^2$  probability of 6.7%.

moment	$m_{K\pi}$	interval 1	$m_{K\pi}$	interval 2
$c_{00}^0$	$-3.4978 \pm 0.0134$	$+3.7214 \pm 0.0056$		
$c_{00}^1$	$-3.1421 \pm 0.0536$	$-2.2392 \pm 0.0194$		
$c_{00}^2$	$+0.6198 \pm 0.0745$	$-1.5482 \pm 0.0305$		
$c_{00}^3$	$-1.2819 \pm 0.0861$	$-0.5077 \pm 0.0364$		
$c_{00}^4$	$+0.1335 \pm 0.1025$	$+0.4200 \pm 0.0425$		
$c_{00}^5$	$+0.1604 \pm 0.1146$	$+0.1465 \pm 0.0464$		
$c_{20}^0$	$+0.3156 \pm 0.0352$	$+0.3023 \pm 0.0141$		
$c_{22}^0$	$+0.0069 \pm 0.0267$	$+0.0414 \pm 0.0103$		
$c_{40}^0$	$+0.0984 \pm 0.0326$	$+0.0766 \pm 0.0145$		
$c_{20}^1$	$-0.2174 \pm 0.0759$	$-0.2362 \pm 0.0317$		
$c_{40}^1$	$-0.1033 \pm 0.0676$	$-0.0682 \pm 0.0333$		
$c_{20}^2$	$-0.0738 \pm 0.0962$	$-0.0742 \pm 0.0477$		
$c_{22}^2$	$+0.0039 \pm 0.0609$	$-0.0487 \pm 0.0220$		

moment	$m_{K\pi}$	interval 3	$m_{K\pi}$	interval 4
$c_{00}^0$	$+3.7287 \pm 0.0045$	$+3.6867 \pm 0.0094$		
$c_{00}^1$	$-1.8757 \pm 0.0172$	$-1.3891 \pm 0.0449$		
$c_{00}^2$	$-1.7036 \pm 0.0246$	$-1.5303 \pm 0.0479$		
$c_{00}^3$	$-0.3875 \pm 0.0289$	$-0.3059 \pm 0.0622$		
$c_{00}^4$	$+0.0522 \pm 0.0334$	$-0.7166 \pm 0.0691$		
$c_{00}^5$	$+0.1763 \pm 0.0375$	$-0.0305 \pm 0.0770$		
$c_{20}^0$	$+0.2957 \pm 0.0127$	$+0.2540 \pm 0.0258$		
$c_{22}^0$	$+0.0725 \pm 0.0113$	$+0.0930 \pm 0.0257$		
$c_{40}^0$	$+0.0887 \pm 0.0127$	$+0.0410 \pm 0.0262$		
$c_{20}^1$	$-0.1925 \pm 0.0214$	$-0.1725 \pm 0.0445$		
$c_{40}^1$	$-0.0638 \pm 0.0215$	$-0.0503 \pm 0.0456$		
$c_{20}^2$	$-0.1120 \pm 0.0306$	$-0.1287 \pm 0.0526$		
$c_{22}^2$	$-0.0745 \pm 0.0233$	$-0.0434 \pm 0.0451$		

TABLE A.5: Final efficiency moments values from the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  simulation sample after applying the corrections described in Section 4.2.3. It is interesting to notice, even without any  $\chi^2$  test check that the efficiency moments vary more between the  $m_{K\pi}$  intervals than it does between production periods ( see Table A.2, Table A.3) or between  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  simulation samples (see Table A.4). The situation is similar in case of no corrections applied as well.

moment	$m_{K\pi}$	interval 1	$m_{K\pi}$	interval 2
$c_{00}^0$	$+3.4164 \pm 0.0188$	$+3.6963 \pm 0.0042$		
$c_{00}^1$	$-4.0320 \pm 0.0562$	$-2.8430 \pm 0.0150$		
$c_{00}^2$	$+1.2914 \pm 0.1030$	$-1.2462 \pm 0.0227$		
$c_{00}^3$	$-1.0834 \pm 0.1225$	$+0.4463 \pm 0.0274$		
$c_{00}^4$	$+0.5572 \pm 0.1397$	$-0.1800 \pm 0.0311$		
$c_{00}^5$	$-0.1778 \pm 0.1538$	$-0.0238 \pm 0.0345$		
$c_{20}^0$	$+0.2204 \pm 0.0446$	$+0.2853 \pm 0.0117$		
$c_{22}^0$	$+0.0052 \pm 0.0330$	$+0.0321 \pm 0.0101$		
$c_{40}^0$	$+0.0400 \pm 0.0423$	$+0.0863 \pm 0.0118$		
$c_{20}^1$	$-0.2249 \pm 0.1083$	$-0.2692 \pm 0.0215$		
$c_{40}^1$	$-0.0109 \pm 0.1007$	$-0.0732 \pm 0.0222$		
$c_{20}^2$	$-0.0013 \pm 0.1450$	$-0.0534 \pm 0.0292$		
$c_{22}^2$	$-0.0262 \pm 0.0882$	$+0.0252 \pm 0.0204$		

moment	$m_{K\pi}$	interval 3	$m_{K\pi}$	interval 4
$c_{00}^0$	$+3.7321 \pm 0.0044$	$+3.8477 \pm 0.0087$		
$c_{00}^1$	$-1.8788 \pm 0.0173$	$-0.5687 \pm 0.0365$		
$c_{00}^2$	$-1.7805 \pm 0.0246$	$-3.0352 \pm 0.0458$		
$c_{00}^3$	$-0.5942 \pm 0.0289$	$-0.8462 \pm 0.0543$		
$c_{00}^4$	$+0.2157 \pm 0.0325$	$+0.1747 \pm 0.0652$		
$c_{00}^5$	$+0.2763 \pm 0.0365$	$+0.3265 \pm 0.0707$		
$c_{20}^0$	$+0.2893 \pm 0.0124$	$+0.3296 \pm 0.0252$		
$c_{22}^0$	$+0.0413 \pm 0.0108$	$+0.0599 \pm 0.0240$		
$c_{40}^0$	$+0.0714 \pm 0.0123$	$+0.0708 \pm 0.0248$		
$c_{20}^1$	$-0.2004 \pm 0.0219$	$-0.1287 \pm 0.0389$		
$c_{40}^1$	$-0.0649 \pm 0.0217$	$-0.0526 \pm 0.0386$		
$c_{20}^2$	$-0.1268 \pm 0.0307$	$-0.2497 \pm 0.0491$		
$c_{22}^2$	$-0.0496 \pm 0.0222$	$-0.0652 \pm 0.0464$		

TABLE A.6: Final efficiency moments values from the  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  simulation sample after applying the corrections described in Section 4.2.3. It is interesting to notice, even without any  $\chi^2$  test check that the efficiency moments vary more between the  $m_{K\pi}$  intervals than it does between production periods ( see Table A.2, Table A.3) or between  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  and  $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$  simulation samples (see Table A.4). The situation is similar in case of no corrections applied as well.

	$c_{00}^1$	$c_{00}^2$	$c_{00}^3$	$c_{00}^4$	$c_{00}^5$	$c_{22}^0$	$c_{22}^1$	$c_{40}^0$	$c_{40}^1$	$c_{20}^1$	$c_{20}^2$	$c_{22}^2$
$c_{00}^1$	1.0											
$c_{00}^2$	-0.828	1.0										
$c_{00}^3$	0.571	-0.836	1.0									
$c_{00}^4$	-0.359	0.521	-0.826	1.0								
$c_{00}^5$	0.184	-0.289	0.525	-0.828	1.0							
$c_{22}^0$	-0.635	0.478	-0.385	0.24	-0.132	1.0						
$c_{20}^0$							1.0					
$c_{40}^0$	-0.273	0.199	-0.184	0.152	-0.108	0.366		1.0				
$c_{40}^1$	0.734	-0.546	0.463	-0.328	0.186	-0.919		-0.402	1.0			
$c_{20}^1$	0.331	-0.247	0.227	-0.186	0.133	-0.410		-0.914	0.465	1.0		
$c_{40}^2$	-0.742	0.579	-0.531	0.425	-0.286	0.742		0.408	-0.9	-0.464	1.0	
$c_{22}^2$							0.439					1.0

TABLE A.7: Efficiency moments correlation matrix for  $B_s^0 \rightarrow J/\Psi \bar{K}^{*0}$  (Similar correlations are observed for the  $\bar{B}_s^0 \rightarrow J/\Psi K^{*0}$  sample). Some of the correlations between the moments are quite large. The choice of efficiency moments was not done based on minimum correlations as it is explained in Section 4.2.2.

	$\mathcal{A}_{CP}^0$	$\mathcal{A}_{CP}^{\parallel}$	$\mathcal{A}_{CP}^{\perp}$	$\mathcal{A}_{CP}^S$	$f_0$	$f_{\parallel}$	$\delta_{\parallel}$	$\delta_{\perp}$
$\mathcal{A}_{CP}^0$	+1.00							
$\mathcal{A}_{CP}^{\parallel}$	-	+1.00						
$\mathcal{A}_{CP}^{\perp}$	-	-0.42	+1.00					
$\mathcal{A}_{CP}^S$	+0.10	-	-	+1.00				
$f_0$	-	-	-	-	+1.00			
$f_{\parallel}$	-	-0.15	-	-	-0.35	+1.00		
$\delta_{\parallel}$	-	+0.13	-	-	-	-0.11	+1.00	
$\delta_{\perp}$	-	-	-	-	-	-	+0.66	+1.00
$f_S^1$	-	-	-	+0.16	+0.11	-	-	-0.18
$f_S^2$	-	-	-	-	+0.10	-	-	-0.11
$f_S^3$	-	-	-	-	-	-	-	-
$f_S^4$	-	+0.13	-	-0.15	-0.11	-	+0.10	-
$\delta_S^1$	-	-	-	-	-	-	-	+0.18
$\delta_S^2$	-	-	-	-	+0.14	-0.13	+0.20	+0.17
$\delta_S^3$	-	-	-	-	+0.25	-	-	-
$\delta_S^4$	-	-0.12	-	-	+0.13	-	-	-

	$f_S^1$	$f_S^2$	$f_S^3$	$f_S^4$	$\delta_S^1$	$\delta_S^2$	$\delta_S^3$	$\delta_S^4$
$f_S^1$	+1.00							
$f_S^2$	-	+1.00						
$f_S^3$	-	-	+1.00					
$f_S^4$	-	-	-	+1.00				
$\delta_S^1$	-0.34	-	-	-	+1.00			
$\delta_S^2$	-	-0.55	-	-	-	+1.00		
$\delta_S^3$	-	-	+0.23	-	-	-	+1.00	
$\delta_S^4$	-	-	-	-0.38	-	-	-	+1.00

TABLE A.8: Correlation matrix of the parameters of interest. Correlations below 10% are not shown.

## Appendix B

# Horizontal Weighting

The current appendix addresses the issue of manipulating some input distribution, *source*, such that it becomes statistically compatible with a given, *target*, distribution. A common method to do that involves binning the two distributions with non-unit weights. This method is sometimes called *vertical weighting* and it is described in the following section. The main motivation behind the technique advertised in the current appendix is to circumvent problems arising from binning distributions, especially with increasing number of bins or in case of small sample sizes. Note that the bigger the number of bins the more precise the matching of the two distributions will be. The advertised technique is described in the current appendix and a choice for its name could be *horizontal weighting*. The technique is inspired by discussions with Gerhard Raven and Diego Martinez Santos.

Lastly a matching example is given at the of the current appendix. This example is a typical problem in high energy physics originating from non perfect simulation. Specifically, certain kinematic distributions might differ significantly between simulated events and data. Given that simulated events are commonly used to control acceptance, resolution or other detector effects; what is typically done is to correct the simulated data sample by matching its kinematic distributions to the ones observed in data. An example of a situation where this problem arises can be found in Section 4.2.3.

## Vertical Weighting

The vertical weighting is straightforward approach to matching distributions. The first step is to bin the source (S) and target (T) distributions with the same binning scheme. After that a weight,  $w_i$ , is assigned to each entry in a given bin,  $i$ , such that the bin contents of the source distributions matches these of the target, see Eq. B.1. Essentially this technique moves vertically each bin,  $i$ , of the source distribution in order to match the  $i^{th}$  bin contents of the target.

$$w_i = \frac{T_i}{S_i}. \quad (\text{B.1})$$

The advantages of the vertical weighting approach is that it is easy to understand and implement. However, there are some disadvantages that result from the binning itself. For example it can happen that a given source or target bin has zero entries for a given binning scheme. This situation becomes more pronounced in the case of large number of bins, which as already mentioned improves the precision of the matching. In addition it can also happen that any of the distributions is weighted and the sum of weights in a given bin is negative. Both of the above situation require some justification as to how these problematic cases can be handled. In addition, it can also happen that the source to target matching needs to be done in many dimensions, as mentioned in the introduction of the current appendix. In that case the number of bins increase rapidly, and thus the number of problematic bins as well, to the point that it is no longer possible to match the source distributions to the target one. Note that variables corresponding to these dimensions are in general correlated with each other. Hence, doing several one dimensional weighting steps will simply ignore these correlations.

## Horizontal Weighting

The horizontal approach to matching distributions is meant to bypass the problem of binning, especially in many dimensions, and thus make it possible to match an arbitrary number of variables between source and target. The basic idea of the approach is to apply as chain of transformations to both source and target distributions, such that they become

uncorrelated, see Eq. B.2a and Eq. B.2b respectively. Subsequently the transformation chain that has been applied to the target distributions are inverted and then applied to the source ones, Eq. B.2c. The three transformations involved in the current approach are: Transformation (*A*) converts the input distribution to flat. Transformation (*B*) converts a flat distribution to a normal distribution while (*C*) removes the correlation between two distributions. The necessary mathematical tools to perform the above transformations are presented in the next paragraph.

$$S \times A_S \rightarrow S_{\text{flat}} \times B_S \rightarrow S_{\text{gaus}} \times C_S \rightarrow S_{\text{uncor gaus}}, \quad (\text{B.2a})$$

$$T \times A_T \rightarrow T_{\text{flat}} \times B_T \rightarrow T_{\text{gaus}} \times C_T \rightarrow T_{\text{uncor gaus}}, \quad (\text{B.2b})$$

$$S_{\text{matched}} = S_{\text{uncorgaus}} \times C_T^{-1} \times B_T^{-1} \times A_T^{-1}. \quad (\text{B.2c})$$

The transformation steps, *A* and *B*, that appear in the above logical steps make use of well known mathematical theorems, namely *Inverse Transformation Sampling* and *Probability Integral Transform*. No proof of the above theorems is included since the intention of the current appendix is to quickly demonstrate the advertised technique. However, there is plethora of examples online.

**Theorem 1 (Inverse Transformation Sampling)** *Let  $u$  be a variable uniformly distributed. Consider another random variable  $x$  that has a Cumulative Distribution Function (CDF), call it  $F_x$ . Then the variable  $x' = F_x^{-1}(u)$  is distributed the same way as  $x$  does.*

**Theorem 2 (Probability Integral Transform)** *Let  $x$  be a random variable distributed according to a PDF,  $P(x)$ . Let  $F_x$  be the CDF of  $P(x)$ . Then the variable  $u = F_x(x)$  is uniform. This theorem is the inverse of Theorem 1.*

For completion the CDF,  $F_x(y)$ , of a certain PDF,  $P(x)$ , where  $y$  takes its values from the same domain as  $x$ , is by definition:

$$F_x(y) = \int_{-\infty}^y P(t)dt = \text{Probability}(x \leq y). \quad (\text{B.3})$$

Both of the above theorems can only be directly applied to continuous distributions. However, exact analytic shape of the PDFs or CDFs involved is not required. This is because the CDF of any variable  $x$  can be built by essentially computing the integral-sum of Eq. B.3. In order to do that  $x$  is binned so that the cumulative sum of each bin can be computed. This is a straightforward step, where the number of entries in each bin are added to the number of entries of the next bin until the range of  $x$  is exhausted. It is important to point out that this binning can be arbitrarily fine without having the problems explained in the introduction, since it is not used to much any distribution. However, from implementation point of view the larger the number of bins the slower the algorithm performs. (Here is where the built-in function `numpy.digitise()` of `python` proves to be useful, as the timing scales nicely with the number of bins.)

The transformations  $A$  and  $B$  mentioned in Eq. B.2, are direct implementations of Theorem 2. Similarly the inverted versions,  $A^{-1}$  and  $B^{-1}$ , of the previous transformations are direct implementations of Theorem 1. The only difference between the case of the transformations  $B$  and  $B^{-1}$  the shapes of the PDF and CDF involved are known analytically. The previous transformations involve the *inverse error function* and the *complementary error function* respectively, which are both well known distributions.

Coming now to the last transformation step,  $C$ , of Eq. B.2. This step essentially performs a linear transformation to two correlated Gaussian distributed variables such that they become uncorrelated. The method followed is essentially a standard *Matrix diagonalization*; The relevant steps required are described in Method 1

**Method 1** *Let  $\vec{x}$  be a set of correlated variables, with a corresponding covariance matrix  $C$ . Let  $P^{-1}$  be the matrix with the eigen-vectors of  $C$  as columns. The set of values  $\vec{x}' = P\vec{x}$  is an uncorrelated set of  $\vec{x}$ . The set of values  $\vec{x} = P^{-1}\vec{x}'$  is the corresponding correlated set of  $\vec{x}'$ .*

The covariance is computed using the following standard formula:

$$c_{ij} = \frac{1}{N} \sum (x_i - \hat{x}_i)(x_j - \hat{x}_j), \text{ with } \hat{x} = \frac{1}{N} \sum x. \quad (\text{B.4})$$

Finding the eigen-vectors of  $C$  is a bit more lengthy to quote in the current appendix. Nevertheless, it is a straightforward well known problem for which there are many implemented algorithms as well.

Lastly an important implementation issues is clarified. Specifically, the CDF is built by binning a given distribution,  $x$ , and associating the a value of the integral of Eq. B.3, name it  $c_i$ , to each bin  $x_i$ . In that case the CDF is just an  $x_i \rightarrow c_i$  map structure, implying that the  $x \rightarrow c$  function is not continuous by definition. Of the binning can be increased but one could argue that this is more brute force that solving the problem. The way to make the  $x \rightarrow c$  mapping continues is using the well known technique of *linear interpolation*, shown in Figure B.5, thus solving the first of the two implementation issues.

$$c(x) = v_i + \frac{c_{i+1} - c_i}{x_{i+1} - x_i} * (x - x_i). \quad (\text{B.5})$$

The ingredients necessary to apply the horizontal weighting technique presented have been covered. An implementation of the above advertised technique as well as an example can be found in the following link on GitHub: <https://github.com/vsyropou/multi-dim-matching>.

### Example-Discussion

To demonstrate the advertised technique a typical problem in high energy physics is addressed. Specifically, the  $K^{*0}$  particle from the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  mode decays into a  $K$  and a  $\pi$ . The momenta distributions of these two particles ( $p(K) - p(\pi)$ ) is found to defer between the simulation sample (source) and background subtracted data (target). Furthermore, the previous distributions are correlated and combined with a fine binning will yield problems as explained earlier in the vertical weighting subsection.

After applying the advertised technique the ( $p(K) - p(\pi)$ ) distributions become statistically compatible, avoiding the risks associated to a multidimensional fine binning scheme. The matched distributions can be seen at the bottom of Figure B.1. In addition, a Kolmogorov–Smirnov test is performed to quantify the matching of each source-target distribution. The results are summarized in Table B.1.

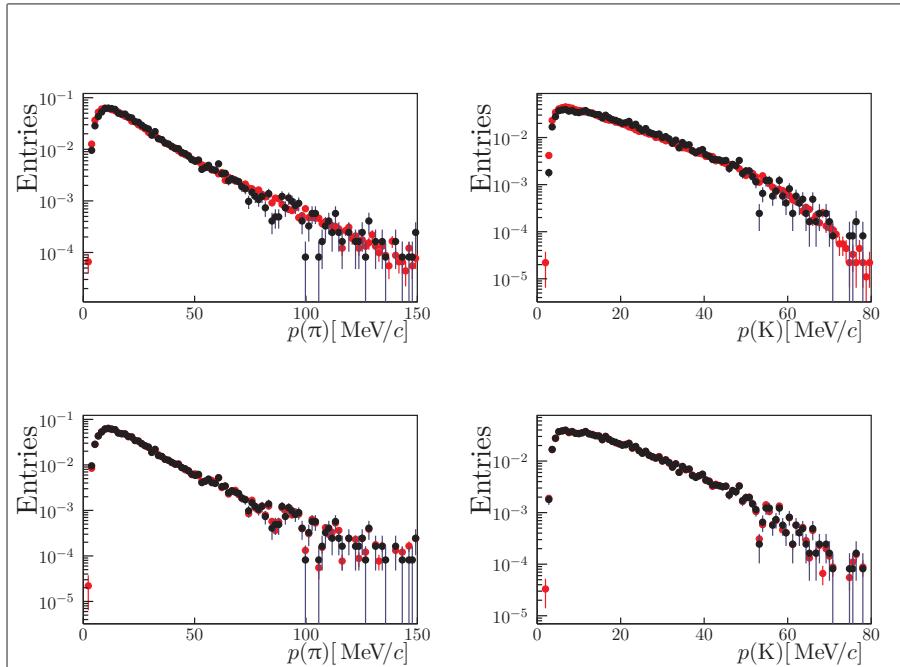


FIGURE B.1: Comparison before and after matching. Source (Target) distributions are shown in red (blue) color. The upper (lower) two distributions are the original (matched) distributions.

distribution	KS before matching	KS after matching
$p(K)$	$10^{-9}$	0.998
$p(\pi)$	$10^{-25}$	1.000

TABLE B.1: KS test between source and target for each of the two distributions ( $p(K) - p(\pi)$ ). Better agreement is achieved after matching.

# References

- [1] LHCb collaboration, *Phys. Rev. Lett.* 114.4 (2015), p. 041801. arXiv: 1411.3104 [hep-ex].
- [2] LHCb collaboration, *JHEP* 11 (2015), p. 082. arXiv: 1509.00400 [hep-ex].
- [3] Johannes Albrecht et al. Tech. rep. LHCb-PUB-2017-007. CERN, 2017. URL: <http://cds.cern.ch/record/2253050>.
- [4] CMS collaboration, *Phys. Lett.* B716 (2012), pp. 30–61. arXiv: 1207.7235 [hep-ex].
- [5] ATLAS collaboration, *Phys. Lett.* B716 (2012), pp. 1–29. arXiv: 1207.7214 [hep-ex].
- [6] S. L. Glashow, *Nuclear Physics* 22.4 (1961), pp. 579 –588.
- [7] S. Weinberg, *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
- [8] A. Salam, *Conf. Proc.* C680519 (1968), pp. 367–377.
- [9] I. Aitchison et al. v. 1. Taylor & Francis, 2012.
- [10] F. Halzen et al. Wiley, 1984.
- [11] C. S. Wu et al., *Phys. Rev.* 105 (4 1957), pp. 1413–1415.
- [12] R. L. Garwin et al., *Phys. Rev.* 105 (4 1957), pp. 1415–1417.
- [13] Super-Kamiokande collaboration, *Phys.Rev.Lett.* 81 (1998), pp. 1562–1567. arXiv: hep-ex/9807003 [hep-ex].
- [14] KamLAND collaboration, *Phys.Rev.Lett.* 90 (2003), p. 021802. arXiv: hep-ex/0212021 [hep-ex].
- [15] SNO collaboration, *Phys.Rev.Lett.* 89 (2002), p. 011301. arXiv: nucl-ex/0204008 [nucl-ex].
- [16] DAYA-BAY Collaboration, *Phys.Rev.Lett.* 108 (2012), p. 171803. arXiv: 1203.1669 [hep-ex].

- [17] R. D. Peccei, *Lect. Notes Phys.* 741 (2008), pp. 3–17. arXiv: [hep-ph/0607268 \[hep-ph\]](#).
- [18] A. Sakharov, *Pisma Zh.Eksp.Teor.Fiz.* 5 (1967), pp. 32–35.
- [19] P. Huet et al., *Phys.Rev.* D51 (1995), pp. 379–394. arXiv: [hep-ph/9404302 \[hep-ph\]](#).
- [20] M. Gavela et al., *Nucl.Phys.* B430 (1994), pp. 382–426. arXiv: [hep-ph/9406289 \[hep-ph\]](#).
- [21] M. Gavela et al., *Mod.Phys.Lett.* A9 (1994), pp. 795–810. arXiv: [hep-ph/9312215 \[hep-ph\]](#).
- [22] WMAP collaboration, *Astrophys.J.Suppl.* 208 (2013), p. 19. arXiv: [1212.5226 \[astro-ph.CO\]](#).
- [23] CMS et al., *Nature* 522 (2015), pp. 68–72. arXiv: [1411.4413 \[hep-ex\]](#).
- [24] LHCb collaboration, *JHEP* 02 (2016), p. 104. arXiv: [1512.04442 \[hep-ex\]](#).
- [25] LHCb collaboration, *Phys. Rev. Lett.* 115.11 (2015), p. 111803. arXiv: [1506.08614 \[hep-ex\]](#).
- [26] F. Englert et al., *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323.
- [27] P. Higgs, *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509.
- [28] Particle Data, *Chin. Phys.* C38 (2014), p. 090001.
- [29] L. Wolfenstein, *Phys.Rev.Lett.* 51 (1983), p. 1945.
- [30] A. Buras et al., *Phys. Rev.* D50 (1994), pp. 3433–3446. arXiv: [hep-ph/9403384 \[hep-ph\]](#).
- [31] J. Leerdam. PhD thesis. Amsterdam U., 2016. URL: <https://cds.cern.ch/record/2156941>.
- [32] J. Charles et al., *Phys. Rev.* D84 (2011). Online update: CKM 2015, p. 033005. arXiv: [1106.4041 \[hep-ph\]](#).
- [33] M. Kobayashi et al., *Prog.Theor.Phys.* 49 (1973), pp. 652–657.
- [34] V. Weisskopf et al., *Z.Phys.* 63 (1930), pp. 54–73.
- [35] V. Weisskopf et al., *Z.Phys.* 65 (1930), pp. 18–29.

- [36] A. Dighe et al., *Eur. Phys. J.* C6 (1999), pp. 647–662. arXiv: [hep-ph/9804253 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9804253).
- [37] K. Bruyn. PhD thesis. Vrije U., Amsterdam, 2015. URL: <https://cds.cern.ch/record/2048174>.
- [38] A. Lenz et al., *CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010.* 2011. arXiv: [1102.4274 \[hep-ph\]](https://arxiv.org/abs/1102.4274).
- [39] LHCb collaboration, (2016). arXiv: [1605.09768 \[hep-ex\]](https://arxiv.org/abs/1605.09768).
- [40] HFAG collaboration, (2014). and online updates at <http://www.slac.stanford.edu/xorg/hfag/>. arXiv: [1412.7515 \[hep-ex\]](https://arxiv.org/abs/1412.7515).
- [41] Y. Golfand et al., *JETP Lett.* 13 (1971), pp. 323–326.
- [42] D. Volkov et al., *Phys.Lett.* B46 (1973), pp. 109–110.
- [43] J. Wess et al., *Nucl.Phys.* B70 (1974), pp. 39–50.
- [44] ARGUS collaboration, *Phys. Lett.* B192 (1987). [,51(1987)], pp. 245–252.
- [45] A. Buras, *PoS EPS-HEP2009* (2009), p. 024. arXiv: [0910.1032 \[hep-ph\]](https://arxiv.org/abs/0910.1032).
- [46] C. Chiang, *JHEP* 1004 (2010), p. 031. arXiv: [0910.2929 \[hep-ph\]](https://arxiv.org/abs/0910.2929).
- [47] A. Datta et al., *Phys.Rev.* D80 (2009), p. 075006. arXiv: [0905.2105 \[hep-ph\]](https://arxiv.org/abs/0905.2105).
- [48] K. Bruyn et al., *JHEP* 1503 (2015), p. 145. arXiv: [1412.6834 \[hep-ph\]](https://arxiv.org/abs/1412.6834).
- [49] X. Liu et al., *Phys. Rev.* D89.9 (2014), p. 094010. arXiv: [1309.0313 \[hep-ph\]](https://arxiv.org/abs/1309.0313).
- [50] P. Frings et al., *Phys. Rev. Lett.* 115.6 (2015), p. 061802. arXiv: [1503.00859 \[hep-ph\]](https://arxiv.org/abs/1503.00859).
- [51] S. Faller et al., *Phys. Rev.* D79 (2009), p. 014005. arXiv: [0810.4248 \[hep-ph\]](https://arxiv.org/abs/0810.4248).
- [52] LHCb collaboration, *Phys. Rev.* D95 (2016). arXiv: [1606.07898 \[hep-ex\]](https://arxiv.org/abs/1606.07898).

- [53] LHCb collaboration, *Int. J. Mod. Phys. A*30.07 (2015), p. 1530022. arXiv: 1412.6352 [hep-ex].
- [54] A. Alves et al., *J. Instrum.* 8 (2012), P02022. 29 p. arXiv: 1211.1346 [hep-ex].
- [55] LHCb RICH group, *Eur. Phys. J. C* 73 (2012), 2431. 25 p. arXiv: 1211.6759 [hep-ex].
- [56] LHCb MuonID group, *J. Instrum.* 8 (2013), P10020. 17 p. arXiv: 1306.0249 [hep-ex].
- [57] H. Morata et al. Tech. rep. LHCb-2008-073. CERN, 2010. URL: <http://cds.cern.ch/record/1152284>.
- [58] LHCb collaboration, *Comput. Phys. Commun.* (2016). arXiv: 1604.05596 [physics.ins-det].
- [59] Tech. rep. LHCB-CONF-2016-013. CERN, 2016. URL: <https://cds.cern.ch/record/2214008>.
- [60] Tech. rep. LHCB-CONF-2016-012. CERN, 2016. URL: <https://cds.cern.ch/record/2224464>.
- [61] G. Lanfranchi et al. Tech. rep. CERN, 2009. URL: <https://cds.cern.ch/record/1202759>.
- [62] A. Sarti et al. Tech. rep. CERN, 2010. URL: <https://cds.cern.ch/record/1235824>.
- [63] R. Aaij et al. Tech. rep. CERN, 2011. URL: <https://cds.cern.ch/record/1384386>.
- [64] K. Dungs. PhD thesis. Dortmund U., 2015. URL: <https://cds.cern.ch/record/2063310>.
- [65] R. Aaij. PhD thesis. Vrije U., Amsterdam, 2015. URL: <https://cds.cern.ch/record/2037893>.
- [66] L. Hommels. PhD thesis. Vrije U., Amsterdam, 2006. URL: <https://cds.cern.ch/record/999327>.
- [67] C. Parkes et al. Tech. rep. CERN, 2015. URL: <https://cds.cern.ch/record/1994172>.
- [68] E. Bowen et al. Tech. rep. CERN, 2016. URL: <https://cds.cern.ch/record/2105078>.

- [69] M. Pivk et al., *Nucl.Instrum.Meth.* A555 (2005), pp. 356–369. arXiv: physics.data-an/0402083 [physics.data-an].
- [70] A. Hoecker et al., *PoS ACAT:040* 040 (2007). arXiv: physics.data-an/0703039 [physics.data-an].
- [71] LHCb collaboration, *Phys.Rev.* D90.1 (2014), p. 012003. arXiv: 1404.5673 [hep-ex].
- [72] LHCb collaboration, *Phys.Rev.* D89.9 (2014), p. 092006. arXiv: 1402.6248 [hep-ex].
- [73] LHCb collaboration, *Phys.Rev.* D87.7 (2013), p. 072004. arXiv: 1302.1213 [hep-ex].
- [74] Y. Gao et al., (2014). URL: <https://cds.cern.ch/record/1701984>.
- [75] LHCb collaboration, *JHEP* 07 (2014), p. 103. arXiv: 1406.0755 [hep-ex].
- [76] L. Amoroso, *Ann. Mat. Pura. Appl.* 21 (1925), pp. 123–159.
- [77] D. Santos et al., *Nucl.Instrum.Meth.* A764 (2014), pp. 150–155. arXiv: 1312.5000 [hep-ex].
- [78] G. Valencia, *Phys. Rev. D* 39.BNL-42229 (1988), pp. 3339–3345.
- [79] G. Krämer et al., *Z. Phys. C* 55.DESY-92-042. IKDA-92-8 (1992), pp. 497–502.
- [80] M. Jacob et al., *Annals Phys.* 7 (1959), pp. 404–428.
- [81] R. Kutschke. internal CLEO note, 1996.
- [82] D. Eijk. PhD thesis. Vrije U., Amsterdam, 2012. URL: <https://cds.cern.ch/record/1484405>.
- [83] T. Pree. PhD thesis. Amsterdam: Vrije U., Amsterdam, 2010. URL: <https://cds.cern.ch/record/1299931>.
- [84] D. Herndon et al., *Phys. Rev. D* 11 (11 1975), pp. 3165–3182.
- [85] D. Aston et al., *Nuclear Physics B* 296.3 (1988), pp. 493 –526.
- [86] LHCb collaboration, *Phys. Rev. Lett.* 110 (2013), p. 221601. arXiv: 1304.6173 [hep-ex].

- [87] LHCb collaboration, (2014). to appear in Phys. Rev. Lett. arXiv: 1409.8586 [hep-ex].
- [88] LHCb collaboration, *Phys. Lett.* B739 (2014), p. 218. arXiv: 1408.0275 [hep-ex].
- [89] LHCb collaboration, *JHEP* 07 (2014), p. 041. arXiv: 1405.2797 [hep-ex].
- [90] LHCb collaboration. Tech. rep. LHCb-CONF-2013-011. CERN, 2013. URL: <https://cds.cern.ch/record/1559262>.
- [91] Belle Collaboration, *Phys. Lett.* B538 (2002), pp. 11–20. arXiv: hep-ex/0205021 [hep-ph].
- [92] LHCb collaboration, *Phys. Rev.* D86 (2012), p. 071102. arXiv: 1208.0738 [hep-ex].
- [93] R. Fleischer, *Phys. Rev.* D60 (1999), p. 073008. arXiv: hep-ph/9903540 [hep-ph].
- [94] O. Haan et al., *Nuclear Physics B* 22.2 (1970), pp. 448–460.
- [95] M. Wirbel et al., *Zeitschrift für Physik C Particles and Fields* 29.4 (1985), pp. 637–642.
- [96] N. Cabibbo et al., *Physics Letters B* 73.4 (1978), pp. 418–422.
- [97] D. Fakirov et al., *Nuclear Physics B* 133.2 (1978), pp. 315–326.
- [98] M. Gell-Mann, *Physics Letters* 8.3 (1964), pp. 214–215.
- [99] Y. Neeman, *Nuclear Physics* 26.2 (1961), pp. 222–229.
- [100] M. Nagashima et al., *Mod. Phys. Lett.* A23 (2008), pp. 1175–1185. arXiv: hep-ph/0701199 [hep-ph].
- [101] M. Gronau, *Phys. Lett.* B727 (2013), pp. 136–140. arXiv: 1308.3448 [hep-ph].
- [102] CKMfitter group, *Phys. Rev.* D91.7 (2015). Online update: CKM 2015, p. 073007. arXiv: 1501.05013 [hep-ph].
- [103] K. Bruyn et al., *Phys. Rev.* D86 (2012), p. 014027. arXiv: 1204.1735 [hep-ph].
- [104] LHCb collaboration, *Phys. Lett.* B742 (2015), pp. 38–49. arXiv: 1411.1634 [hep-ex].

# Summary

Within the domain of modern physics there exists a field that attempts to answer a fundamental question, that puzzled philosophers since the ancient times. Namely, identify and understand the fundamental-indivisible building blocks that constitute the natural world. Remarkably enough more than 2 millennia after Democritus, who introduced the idea of atoms or  $\alpha\tauομα$  which in Greek means indivisible, contemporary scientists still have not found a definite answer. It seems that the observable universe consists of a handful of elementary particles and are classified in two distinct categories: Namely *gauge bosons*, responsible for mediating all the known particle interactions of nature (with the exception of gravity) and *fermions* which are the constituents of matter. The recently discovered Higgs particle [4, 5] is a special type of boson that plays a central role in explaining how particles acquire mass. Fermions can be divided further into *quarks*, which constitute particles like the proton and the neutron. In general *Leptons* do not build stable atoms. However, electrons constitute the nucleus of atoms and together with protons and neutrons do build atoms. Leptons, along with quarks, are present in crucial processes like the formation and evolution of stars (which has direct impact on the way life evolved on earth).

## Particle Physics and The Standard Model

The state of the art mathematical framework necessary to describe interactions between the fermions is called *Standard Model* [6–8]. Describing a quantum mechanical process like particle interactions means being able to compute the probability for a certain outcome in that particular particle interaction and it involves non-trivial calculations. Thus the established predictive power of the Standard Model is an important achievement. Furthermore, the recently discovered Higgs boson [4, 5],

which plays a special role in the explanation of how particles acquire mass, makes Standard Model complete.

Despite its success, there are established phenomena and observations that the Standard Model does not account for. Perhaps the most striking one is the absence of any description of the most familiar, yet the weakest, interaction of nature, gravity. Or the observed amount of matter-antimatter imbalance in the universe [19–21] (Note that antimatter is a well understood state of matter that has its quantum numbers signs, such as electric charge, flipped with respect to matter.) The above phenomena are a few examples that reveal the incompleteness of the model. Thus the scientific method compels to continued testing Standard Model predictions and look for ways to improve it. Significant deviation from these predictions is a hint for the presence of *New Physics* beyond the established model.

### ***CP* violation and New Physics**

According to the dominant theory, matter and antimatter were created in equal proportions during the initial phase of the big bang [19–21]. This idea invites the notion of a symmetry between matter and antimatter, known as Charge-Parity or *CP* symmetry. Perfect *CP* symmetry implies that nature interacts with matter in the same way as antimatter. However, at current time the observable universe appears to be almost entirely populated by matter. Thus the origin of *CP violation* follows as a natural question for which the Standard Model has to provide an explanation.

The established idea that nature indeed favors processes where the particles involved are matter and not antimatter particles is captured by the Standard Model. Despite this, it cannot account for the observed amount of matter-antimatter asymmetry [19–21]. Hence, other sources of *CP* violation beyond the Standard Model have to be active. A typical place to search for hints of non Standard Model *CP* violation is in parameters of the model where its predicted value of *CP* violation is very small. In that case any significant observation of *CP* violation is a direct hint for the presence of New Physics. Collecting many such hints helps particle physicists to identify the weaknesses in the Standard

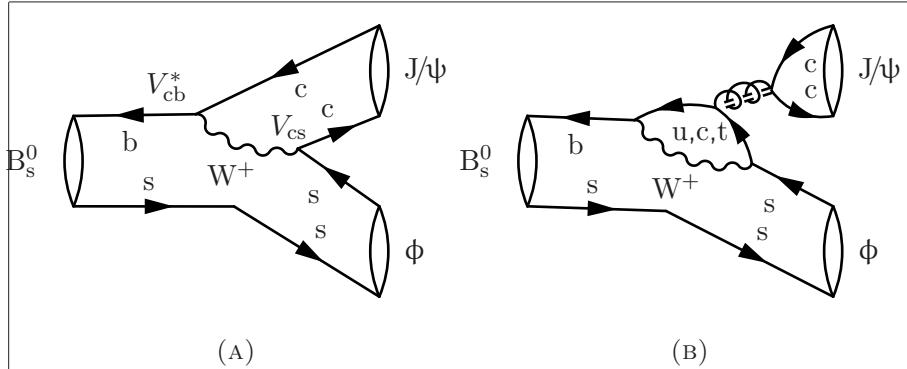


FIGURE S.2: Leading order *tree* (left) and sub-leading penguin (right)  $B_s^0 \rightarrow J/\psi \phi$  decay topologies.

Model as well as choose between the various alternative models, such as *Supersymmetry* e.g. [41–43].

### The Weak Phase $\phi_s$

An interesting parameter where the Standard Model prediction is nearly zero is the weak phase  $\phi_s$ . This parameter manifests itself in  $b \rightarrow c\bar{c}s$  interactions, where a  $b$  quark decays into three other quarks. The most promising decay channel to measure  $\phi_s$  is the  $B_s^0 \rightarrow J/\psi \phi$  decay, see Figure S.2a. The Standard Model  $\phi_s$  prediction as well as its most precise measurement by LHCb is:

$$\phi_s^{\text{LHCb}} = -0.010 \pm 0.039 \text{ rad}, \quad (\text{S.6a})$$

$$\phi_s^{\text{SM,tree}} = -0.03761 {}^{+0.00073}_{-0.00082} \text{ rad}. \quad (\text{S.6b})$$

Given the above measurement it follows that  $\phi_s$  is compatible with the prediction and any New Physics effects that might appear in  $\phi_s$  must be small. From an experimental point of view the situation is only now becoming interesting, where the statistical uncertainty of the experimental

measurement is approaching the Standard Model prediction. Thus future  $\phi_s$  measurements define a high precision era where the presence, or not, of New Physics will become apparent.

However, entering this promising era comes along with an important consideration that has to be taken into account in order to make a robust claim about New Physics hints in  $\phi_s$ . In particular there are certain processes within the Standard Model that  $\phi_s$  measurements did not take into account so far. The magnitude of these processes, which originate from *penguin topologies* and shown in Figure S.2, are known to be small. Despite that, the contribution of penguin topologies shift the Standard Model prediction,  $\phi_s^{\text{SM,tree}}$ , by a small amount,  $\Delta\phi_s^{\text{peng}}$ . Considering also that, as implied by Eq. S.6, potential New Physics in  $\phi_s$  are also small; contributions to  $\phi_s$  from Standard Model penguin topologies are crucial to disentangle from potential New Physics contributions, which also shift the Standard Model prediction by  $\Delta\phi_s^{\text{NP}}$ . The situation is depicted in following equation:

$$\phi_s^{\text{eff}} = \phi_s^{\text{SM,tree}} + \Delta\phi_s^{\text{peng}} + \Delta\phi_s^{\text{NP}}, \quad (\text{S.7})$$

where  $\phi_s^{\text{eff}}$  represents the  $\phi_s$  measurement. Thus, in order to overcome the above mentioned limitations it becomes mandatory to estimate contributions from penguin topology using different decay channels, also called control channels. These control channels, *e.g.*  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$ , have identical tree and penguin topologies with  $B_s^0 \rightarrow J/\psi \phi$  and by exploiting certain quark symmetries it is possible to properly estimate the size of  $\Delta\phi_s^{\text{NP}}$  and thus assess the presence of potential New Physics effects.

## Analyzing Particle Collisions

Having introduced the parameter  $\Delta\phi_s^{\text{peng}}$  and its role in the pursuit for New Physics it is interesting to point out some aspects relevant to the experimental measurement of the previous parameter. First and foremost by utilizing state of the art technologies scientists have built the most powerful accelerator, LHC, at CERN to improve and challenge the Standard Model. The machine is capable of accelerating two beams of protons up to almost the speed of light and collide them at a particular point in space called the interaction point, around which detectors are

located. To put these very energetic collisions into perspective; the energy density of a single proton-proton collision at LHC is approximately as high as when the universe was about a billionth of second old. This effectively allows particle physicists to look back in time and probe information about the state of matter, and antimatter, in the early universe.

The LHCb experiment, located at CERN, is a dedicated experiment to the study of  $CP$  violation. The design of the detector, both at the hardware and software level, is optimized for detecting special interactions of *heavy* quarks, like the  $b$  and the  $c$  quarks. Emphasis is also given to the detection of muons, particularly to the ones that fly inside a small cone along the beam direction. Note that muons can be found among the decay products of many decay channels that LHCb is interested in.

After recording particle collisions that are interesting for physics measurements, like  $\Delta\phi_s^{\text{peng}}$ , the stored data are further processed such that the presence of noise is suppressed as much as possible. In addition biases and finite resolution effects introduced by detector imperfections are taken into account such that the physics model, which estimates the parameters of interest, is corrected for these effects and thus provides valid and robust estimates when fitting the data. The fitting process is based on the principle of *maximum likelihood*. The likelihood is a function of the parameters of interest given the observed data. At its maximum the likelihood function provides the *best fit* estimate for the parameters of interest given the data that were used to build the previous likelihood function.

Visualizing the result of the fit to the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  control channel involves plotting the fitted model on top of the data, see Figure S.3. The data consist of three variables,  $(\cos\theta_K, \cos\theta_\mu, \varphi_h)$ , which are related to the direction of the  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  decay particles.

Having performed the likelihood fit to the data and following the strategy indicated by [37, 48, 51, 93] the final result regarding penguin topology contributions to  $\phi_s$  is:

$$\begin{aligned} \Delta\phi_{s,0}^{\text{peng}} &= 0.000 {}^{+0.010}_{-0.014}, & \Delta\phi_{s,\parallel}^{\text{peng}} &= 0.001 {}^{+0.012}_{-0.016}, \\ \Delta\phi_{s,\perp}^{\text{peng}} &= 0.003 {}^{+0.012}_{-0.016}. \end{aligned} \quad (\text{S.8})$$

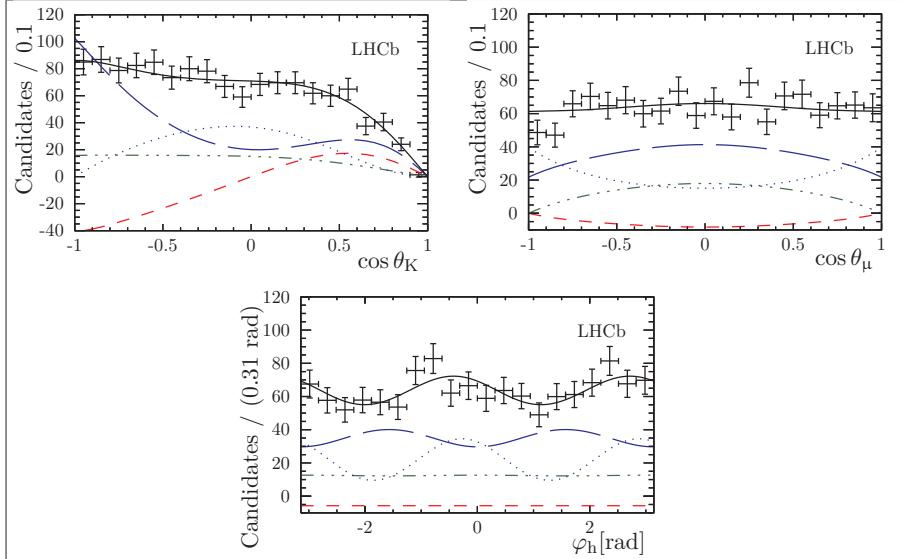


FIGURE S.3: Fitted model (black line) plotted on top of  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data (black crosses). Various components of the model relevant to its  $CP$  structure are shown with colored lines.

where quantities are expressed in rad ( $\phi_s$  is the argument of a complex number). Note that the parameter  $\Delta\phi_s^{\text{peng}}$  is split in its *polarization* components. Essentially the  $B_s^0 \rightarrow J/\psi \phi$  decay probability, as well as the  $\phi_s$  value are in principle different depending on the configuration of the particle spin quantum numbers in the  $B_s^0 \rightarrow J/\psi \phi$  decay resulting in three polarization components  $\parallel$ ,  $\perp$  and 0 respectively for parallel, perpendicular and longitudinal. Thus, the penguin parameter  $\Delta\phi_s^{\text{peng}}$  is expressed in a similar footing.

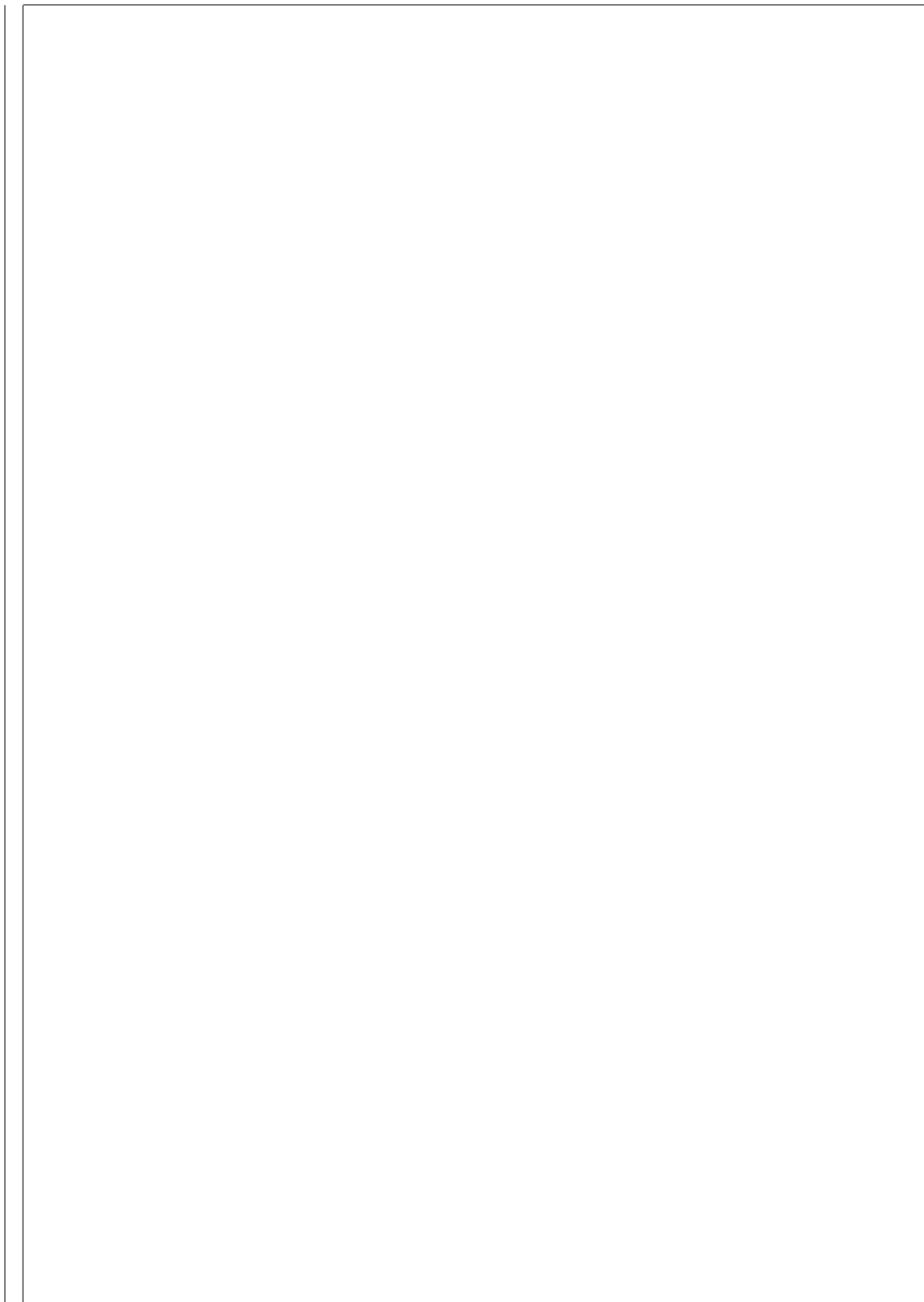
### Impact and Conclusions

The shifts  $\Delta\phi_s^{\text{peng}}$  quoted in Eq. S.8 suggest that contributions of penguin topologies to the  $B_s^0 \rightarrow J/\psi \phi$  decay amplitude are indeed small,  $< 0.017$  rad or  $< 1^\circ$ . Given the also small  $\phi_s$  measured value, shown in Eq. S.6a, it becomes mandatory to control penguin contributions in

future  $\phi_s$  measurements. Increasing the amount of data in the LHC Run 2 might not be enough to yield a significant claim on the presence of physics beyond the Standard Model and hence the upgraded LHCb detector becomes important in the pursuit for New Physics with  $\phi_s$  in the future.

---

---



---

# Samenvatting

Er is een vakgebied binnen de moderne fysica dat de fundamentele vragen der materie probeert te beantwoorden. Deze vragen, welke filosofen van alle tijden hebben beziggehouden, gaan over het identificeren en begrijpen van de fundamentele ondeelbaar bouwstenen waaruit de wereld is opgebouwd. Opmerkelijk is dat meer dan twee millennia nadat de filosoof Democitus de notie van atomen of  $\alpha\tauομα$  - Grieks voor ondeelbaar - heeft geïntroduceerd, hedendaagse wetenschappers nog steeds geen sluitende antwoorden hebben op deze fundamentele vragen. Het lijkt erop dat het zichtbare heelal bestaat een handvol elementaire deeltjes die zijn ondergebracht in twee verschillende categoriën. Namelijk *gauge bosons*, welke verantwoordelijk zijn voor het mediëren van alle bekende krachten tussen de deeltjes (met de uitzondering van de zwaartekracht), en *fermions* welke de bouwblokken zijn van materie. Het recentelijk ontdekte Higgs deeltje [4, 5] is een speciaal type boson dat een centrale rol speelt in het genereren van massa voor deeltjes. Fermions kunnen verder worden onderverdeeld in twee categoriën. *Quarks* vormen samengestelde deeltjes zoals het proton en het neutron. Het zwaarste quark echter, het t quark, zal vervallen voordat het een samengesteld deeltje kan vormen. Het bekendste *lepton* is het electron. Samen met de protonen en neutronen vormen elektronen de atomen. Leptonen en quarks spelen een cruciale rol in processen zoals de formatie van sterren en hebben zo een directe impact op de manier waarop leven op aarde is geëvolueerd.

## Deeltjes fysica en Het Standaard Model

Het Standaard Model is een state of the art mathematisch model dat de interacties tussen fermionen beschrijft [6–8]. Dit is een quantum mechanisch proces. Het beschrijven zo'n process vereist het berekenen

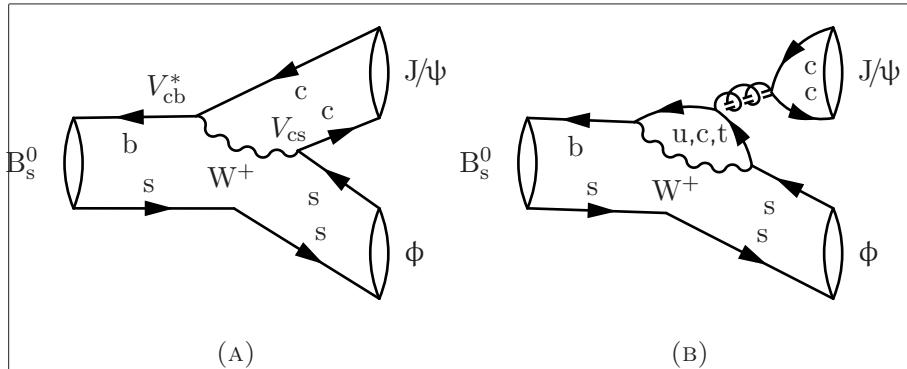
van de kans op een bepaalde uitkomst van het proces. Dit behelst niet-triviale berekeningen. Het voorspellend vermogen van het Standaard Model kan daarom worden beschouwd als een belangrijke prestatie. Het recentelijk ontdekte Higgs boson [4, 5], dat een belangrijke rol speelt in de verklaring hoe deeltjes massa verkrijgen, maakt het Standaard Model compleet.

Ondanks het succes van het Standaard Model, zijn er nog steeds fenomenen en observaties welke het model niet kan verklaren. Wellicht de meest opvallende is de afwezigheid van een beschrijving van de meest bekende, en toch de meest zwakke, natuurkracht: de zwaartekracht. Een ander voorbeeld betreft antimaterie; een toestand van materie, waarbij de quantum numbers voor elektrische lading en pariteit omgekeerd zijn aan die van materie. De waargenomen materie-antimaterie disbalans in het heelal [19–21] is niet voldoende beschreven in het Standaard Model. De bovenstaande fenomenen zijn slechts een paar voorbeelden die de onvolledigheid van het model blootleggen. Daarom is het des te meer van belang de voorspellingen van het Standaard Model wetenschappelijk te toetsen en zo het model te verbeteren. Significante verschillen tussen de metingen en de voorspellingen kunnen een aanwijzing zijn voor *Nieuwe Fysica*.

### ***CP* schending en Nieuwe Fysica**

Materie en antimaterie zijn gecreëerd in gelijke hoeveelheden tijdens de eerste fase van de oerknal, aldus de gevestigde theoriën [19–21]. Dit idee suggereert de notie van een symmetrie tussen materie en antimaterie: *CP* symmetrie. Perfecte *CP* symmetrie betekent dat interacties van materie op dezelfde manier verlopen als interacties van antimaterie. Echter, het huidige heelal bestaat bijna geheel uit materie. Het ontstaan van *CP* schending is een vraagstuk waar het Standaard Model geen sluitende verklaring voor geeft.

Het idee dat de natuur processen prefereert waar de betrokken deeltjes materie zijn boven processen waar de deeltjes antimaterie zijn wordt beschreven in het Standaard Model. Echter, het model kan de grootte van de asymmetrie tussen materie en antimaterie niet voldoende verklaren. Daarom wordt er gezocht naar andere bronnen van *CP* schending,



FIGUUR S.4: Eerste orde *tree* (links) en hogere order *pinguïn* (rechts)  $B_s^0 \rightarrow J/\psi \phi$  verval diagrammen.

buiten het Standard Model. Een manier om hier naar te zoeken is door die parameters te bestuderen waar de Standaard Model een minimale  $CP$  schending voorspelt. Een significante observatie van  $CP$  schending geeft dan een directe hint voor de aanwezigheid van Nieuwe Fysica. Het verzamelen van veel van zulke hints helpt fysici bij het identificeren van de zwakke plekken in het Standaard Model. Daarnaast geeft het richting aan de zoektocht naar alternatieve modellen, zoals *Supersymmetry* e.g. [41–43].

## De zwakke fase $\phi_s$

Een interessante parameter waar de Standaard Model een waarde van ongeveer nul voorspelt is de zwakke fase  $\phi_s$ . Deze parameter manifesteert zichzelf in de interactie  $b \rightarrow c\bar{c}s$ , waar een  $b$  vervalt naar drie andere quarks. Een veelbelovend vervalkanaal voor de meting van  $\phi_s$  is  $B_s^0 \rightarrow J/\psi \phi$ , zie Figure S.4a. De Standaard Model voorspelling van  $\phi_s$  en de meest precieze meting van deze parameter door LHCb zijn:

$$\phi_s^{\text{LHCb}} = -0.010 \pm 0.039 \quad \text{rad}, \quad (\text{S.9a})$$

$$\phi_s^{\text{SM,tree}} = -0.03761^{+0.00073}_{-0.00082} \text{ rad.} \quad (\text{S.9b})$$

Uit bovenstaande meting volgt dat de waarde van  $\phi_s$  is overeenstemming is met de Standaard Model voorspelling. Elke mogelijke bijdrage van Nieuwe Fysica aan  $\phi_s$  moet daarom heel klein zijn. Vanuit het experimentele oogpunt wordt deze meting nu pas echt interessant: de statistische onzekerheid van de meting is bijna klein genoeg om een verschil tussen de gemeten waarde en de SM voorspelling aan te kunnen tonen. De toekomstige metingen van  $\phi_s$  markeren dan ook een tijelperk van hoge precisie metingen, waar het bestaan van Nieuwe Fysica al dan niet zal worden aangetoond.

Echter, wanneer experimentele metingen met zulke hoge precisie kunnen worden uitgevoerd kunnen zelfs de kleinste effecten zichtbaar worden. Er is een belangrijke bijdrage aan de experimentele meting van  $\phi_s$  welke tot dusver niet in acht werd genomen. Het betreft de bijdragen van zogenaamde *pinguin topologien*, zie Figure S.4. Hoewel de bijdragen van deze pinguïn topologien klein zijn, zullen ze de Standaard Model voorspelling,  $\phi_s^{\text{SM,tree}}$ , veranderen met  $\Delta\phi_s^{\text{peng}}$ . Gegeven het feit dat potentiële Nieuwe Fysica bijdragen aan  $\phi_s$  klein zijn (Eq. S.9), is het van cruciaal belang de bijdragen van pinguïn topologien te onderscheiden van die door Nieuwe Fysica. Immers, Nieuwe Fysica verandert de Standaard Model voorspelling ook met een kleine waarde  $\Delta\phi_s^{\text{NP}}$ , zoals weergegeven in de volgende vergelijking:

$$\phi_s^{\text{eff}} = \phi_s^{\text{SM,tree}} + \Delta\phi_s^{\text{peng}} + \Delta\phi_s^{\text{NP}}, \quad (\text{S.10})$$

met  $\phi_s^{\text{eff}}$  de experimentele  $\phi_s$  meting. Het is dus van groot belang de bijdragen van pinguïn topologien correct af te schatten. Hiervoor worden andere  $B_s^0$  vervalkanalen gebruikt, zogenaamde controle kanalen. De gekozen controle kanalen, bijvoorbeeld  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  hebben dezelfde tree en pinguïn topologien als  $B_s^0 \rightarrow J/\psi \phi$ . Door gebruik te maken van bepaalde quark symmetriën is het mogelijk om nauwkeurig de grootte van  $\Delta\phi_s^{\text{peng}}$  te bepalen. Dit is noodzakelijk voor precieze meting van  $\Delta\phi_s^{\text{NP}}$  en de bepaling van de aanwezigheid van Nieuwe Fysica.

### De analyse van botsingen van deeltjes

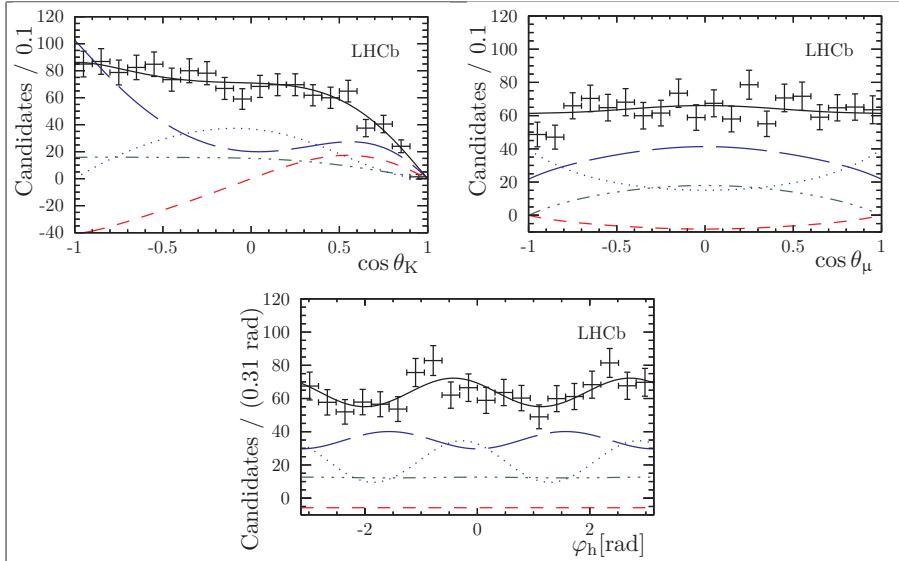
Zoals beschreven speelt de parameter  $\Delta\phi_s^{\text{peng}}$  een belangrijke rol in de zoektocht naar Nieuwe Fysica. Een aantal interessant aspecten van de

meting van deze parameter worden hieronder toegelicht. Ten eerste de LHC, op CERN, de meest krachtige deeltjes versneller op aarde. Deze is gebouwd door wetenschappers waarbij gebruik is gemaakt van state of the art technologien. De machine kan twee proton bundels versnellen tot ongeveer de snelheid van het licht. De bundels worden met elkaar in botsing gebracht op een bepaald punt in de ruimte, het zogenaamde interactie punt, waaromheen detectoren zijn geïnstalleerd. De energie dichtheid van een zo'n proton-proton botsing is net zo hoog als toen het heelal een miljardste van een seconde oud was. Om deze zeer energetische botsingen in perspectief te plaatsen: de energiedichtheid van een enkele proton-proton botsing is vergelijkbaar met de energie van het heelal een miljardste van een second na de oerknal. Door deze omstandigheden met de LHC te creëren kunnen fysici als het ware terug kijken in de tijd. Op deze manier kan informatie worden verkregen over de toestand van materie en antimaterie in het vroege heelal.

Het LHCb experiment, op CERN, is speciaal ontworpen voor de studie van *CP* schending. Het ontwerp van de detector, zowel de hardware als de software, is geoptimaliseerd voor het meten van interacties van zogenaamde *zware* quarks (b en c quarks). Muonen zijn een belangrijk vervalproduct in veel vervalkanalen waarin LHCb geïnteresseerd is. Daarom is bij het ontwerp van LHCb veel aandacht besteed aan het detecteren van muonen, met name van muonen die propageren in een smalle kegel in de richting van de bundel.

Data van proton-proton botsingen in de LHC zijn verzameld om hierop interessante analyses te kunnen verrichten, waaronder de meting van  $\Delta\phi_s^{\text{peng}}$ . De data worden opgeslagen en bewerkt zodanig dat de aanwezigheid van ruis zoveel mogelijk is onderdrukt. Imperfecties in de detector kunnen biases en resolutie effecten introduceren in de analyse. Hiervoor wordt gecorrigeerd zodanig dat een fit van het fysica model aan de data een valide en robuuste schatting geeft van de te meten fysica parameters. Het fit proces is gebaseerd op het principe van de *maximum likelihood*. De likelihood is een functie van de te meten parameters, gegeven de geobserveerde data. Wanneer deze functie maximaal is, geeft het de beste schatting van de te meten parameters gegeven de data waaraan de fit is gemaakt.

Om het resultaat van de fit aan het  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  controle kanaal



FIGUUR S.5: Het gefitte model (zwarte lijn) is op de  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  data (zwarte kruisen) afgebeeld. De verschillende componenten van het model die relevant zijn voor de  $CP$  structuur zijn afgebeeld met gekleurde lijnen.

te visualiseren wordt het model op de data afgebeeld, zie Figure S.5. Er worden drie variabelen bepaald uit de data:  $\cos \theta_K$ ,  $\cos \theta_\mu$  en  $\varphi_h$ . Deze variabelen zijn allen gerelateerd aan de bewegingsrichting van de vervalproducten van het  $B_s^0 \rightarrow J/\psi \bar{K}^{*0}$  verval.

Uit de likelihood fit aan de data, met in acht name van [37, 48, 51, 93], volgt voor de bijdragen van de pinguïn topologien aan  $\phi_s$ :

$$\begin{aligned} \Delta\phi_{s,0}^{\text{peng}} &= 0.000^{+0.010}_{-0.014}, & \Delta\phi_{s,||}^{\text{peng}} &= 0.001^{+0.012}_{-0.016}, \\ \Delta\phi_{s,\perp}^{\text{peng}} &= 0.003^{+0.012}_{-0.016}. \end{aligned} \quad (\text{S.11})$$

met de parameters uitgedrukt in rad ( $\phi_s$  is het argument van een complex getal). De parameter  $\Delta\phi_s^{\text{peng}}$  is gesplitst in drie *polarisaties*. Het

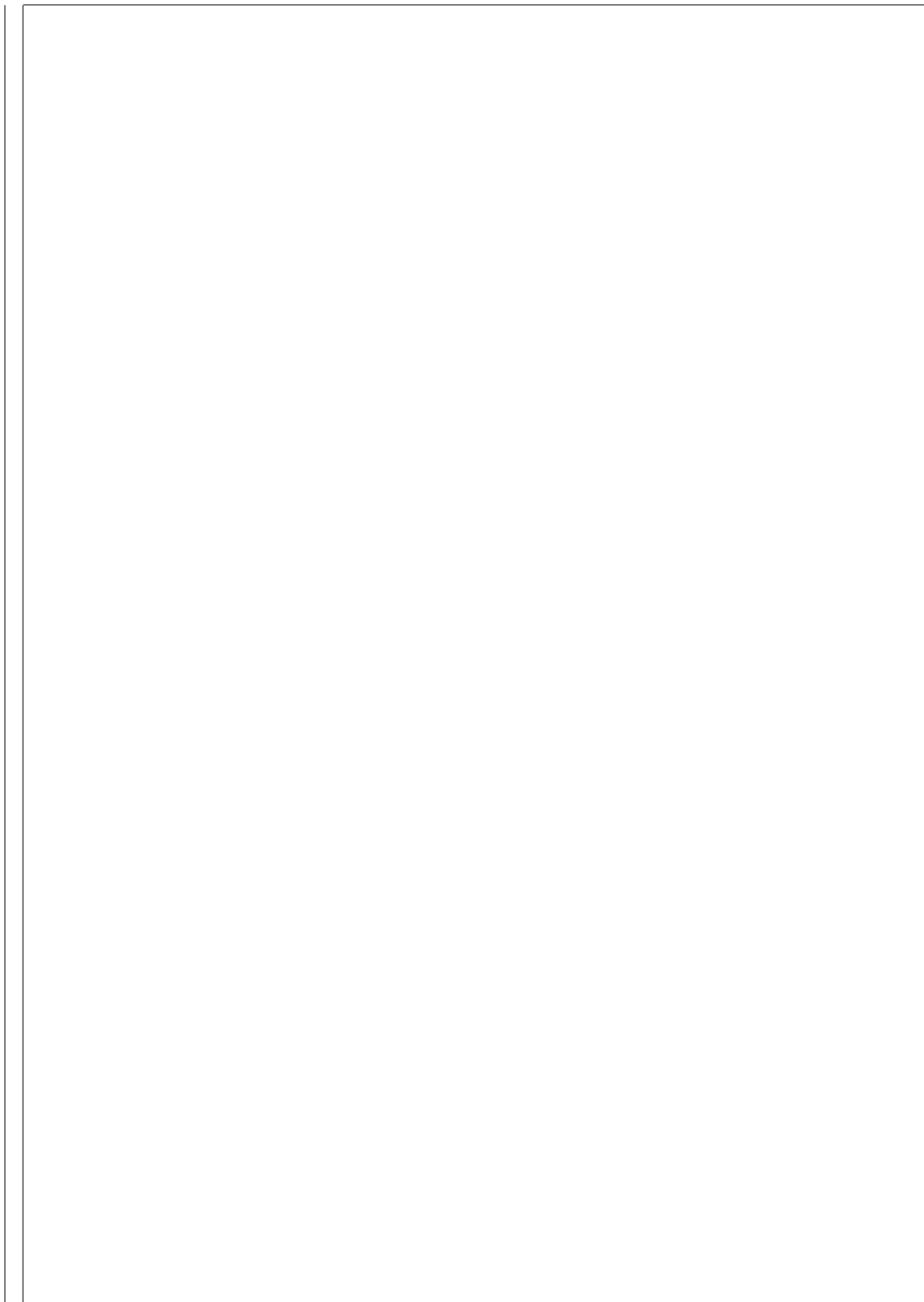
decay  $B_s^0 \rightarrow J/\psi \phi$  hangt af van de configuratie van de spin quantum getallen van de deeltjes. Dit resultert in drie polarisatie toestanden:  $\parallel$ ,  $\perp$  en 0 respectievelijk voor parallel, loodrecht, longitudinaal. Op eenzelfde manier hangt  $\phi_s$  af van de polarisatie en kan de waarde van  $\phi_s$  voor elke polarisatie worden gemeten.

### Impact en conclusies

Uit de gemeten waarden van  $\Delta\phi_s^{\text{peng}}$  in Eq. S.11 volgt dat de bijdrage van pinguïn topologien aan de  $B_s^0 \rightarrow J/\psi \phi$  vervalsamplitude inderdaad klein is, namelijk  $< 0.017$  rad of  $< 1^\circ$ . Gegeven dat de gemeten waarde van  $\phi_s$  óók klein is (zie Eq. S.9a), is het toch noodzakelijk rekening te houden met bijdrage van de pinguïn topologien wanneer  $\phi_s$  wordt gemeten. Om een uitspraak te kunnen doen over het bestaan van Nieuwe Fysica in  $\phi_s$  is het noodzakelijk de hoeveelheid beschikbare data te vergoten. Wellicht is de hoeveelheid data die zal worden verzameld tijdens LHCb Run 2 niet genoeg, en zal de data uit de upgrade fase pas uitsluitsel geven.

---

---



---

# Acknowledgments

By means of the current thesis I pay tribute to my parents, the two people that always supported me. Perhaps without even really realizing themselves, they provided the necessary conditions for me to develop both an endless curiosity about everything around me, and a thirst for the truth, which are essentially the reasons that I got involved in research. Regarding my nikhef colleagues, Roel Aaij and Jeroen van Leerdam deserve a big and warm thank you. These people always found time to supported me at the beginning of my PhD, answering all my questions and guiding me into programming in a clean and readable way. Special thanks goes to Diego Martinez Santos the most clever, modest and cool scientist I ever worked close with. Furthermore, another person that was always an endless tank of knowledge at my disposal is my supervisor, Gerhard Raven. I thank him a lot for his patience throughout my PhD as well as for his time that would always find for me to discuss with him. For the same reasons I also thank my co-promotor, Patrick Koppenburg, and especially for his detailed review of my thesis (and for coping with my spelling mistakes). Special thanks also go to the stuff members of the nikhef LHCb group, and particularly to Marcel Merk, for providing a healthy, full of knowledge and of course fun working environment. Furthermore, I particularly thank my colleagues and friends Panagiwi Tsopela and Jacco de Vries, who I both met at the beginning of my master in 2010. We have faced the same difficulties and helped each other many times. But above all I thank them for being kind and fun people, which are not so common features. Last but not least I thank my colleague and friend Rose Koopman for volunteering to translate my summary.