DOCTORAL THESIS

Control of penguin polutions in the weak phase ϕ_s with $B_s^0 \to J/\psi \overline{K}^{*0}$ decays with the LHCb detector.

Author:
Vasilis SYROPOULOS

Supervisor:
prof. Dr. Gerhard RAVEN

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

LHCb
Faculy of Science

April 20, 2016

Contents

Chapter 1

Introduction

Write a nice introduction after the chapter is complete.

1.1   The Standard Model of Particle Physics

After a few decades of research in the subatomic scale it has been possible
to account for a large number of phenomena in nature, based on the exis-
tance of only a handfull of particles. Which is impressive given the vast
dimmensions that the universe spans over. Two distinct categories emerge
from the above collection of particles, namely the *gauge bossons* responsible
for mediating three out of the four know fundamental forces of nature, and
quarks plus leptons which are the constituents of matter (and antimatter)[1].
The resently discoverd higgs bosson [refff] is a special category by itself due
to each spacial role of explaining how particles acquire mass. Figure Fig-
ure **??** illustrates those matter particles and their interactions via the gauge
bossons (also known as force cariers). The mathematical "language" neces-
sary to describe those interactions is spelled out by the *Standard Model* of
Particle Physics which is a quantum field theory[2]. In this framework par-
ticles are described by quantum fields which is a dificult to grasp but very
usefull way to represent a particle[3].

$$\mathscr{L}_{SM} = \underbrace{i\bar{\psi}D^{\mu}\gamma_{\mu}\psi}_{\mathscr{L}_{\text{Kinetic}}} + \underbrace{(D^{\mu}\phi)^{\dagger}(D^{\mu}\phi) + V(\phi)}_{\mathscr{L}_{\text{Higgs}}} + \underbrace{Y_{ij}\bar{\psi}_{Li}\phi\psi_{Rj} \; + \; h.c.}_{\mathscr{L}_{\text{Yukawa}}}, \qquad (1.1)$$

The above mentioned mathematical language in its most compact form
i shown in Eq. 1.1. The last equation is the so called *lagrangian*[4] of the Stan-
dard Model. The first *Kinnetic*, term of Eq. 1.1 describes the possible inter-
actions between the quarcks or the leptons. The second, *Higgs*, term is the
one that includes the masses of the gauge and higgs bossons. Whereas the
last, *Yukawa*, attributes masses to the quarks and leptons[5] Quarcks or lep-
tons and the higgs fields are represented by ψ and ϕ in Eq. 1.1 respectivelly.

---

[1]Anitmatter is not as exotic as it sounds. It is a state of matter where the signs of all
quantum numbers (like the electric charge) of a particle are fliped.

[2]Quantum field theories are[ref], roughly speaking, the result of combining Einstein's
special theory of relativity[reference] and quantum mechanics[ref].

[3]For further reading on quantum filed theories [refff]

[4]Lgrangians are elegant equations that discribe the dynamics of a system. They are used
both in classical physical systems such as motions of planets and in the quantum world of
tiny distances.

[5]The *Kinnetic* and *Higgs* terms appear more natureally in the lagrangian wheras the last
one is added by hand, which is perfectly fine.

The γ_μ is a Dirac matrix that takes care of the additional structure of the fields due to their intrinsic spin. $V(\phi)$ is the potential term of the higgs field and it contains its mass. The L and R subscripts denote the left and right handed projections of the quark or lepton fields. Y_{ij} are called Yuakawa cuplings. Their meaning and relevance to the current thesis is shown in Section **??**. The symbol D_μ introduces the interactions of the gauge bossons with the quarks or leptons quantum fields and it is called covariant derivative[ref????].
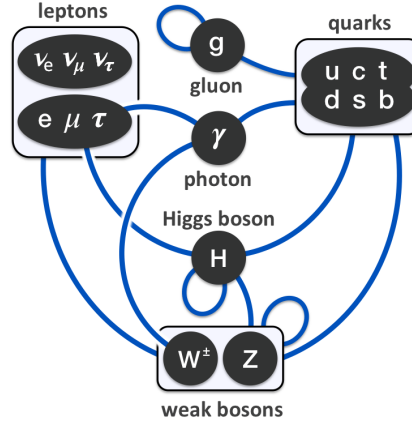

FIGURE 1.1: Standard Model matter particles. Combinations of three and two quarks build hadrons and messons respectivelly. Hadrons are for example protons and neutrons wheares a messon could be the B_s^0 particle which is relevant for the current thesis. Leptons can be charged (e,μ,τ) and uncharged (ν_e,ν_μ,ν_τ). The first of the charged ones is the familiar electron that exists in the neucleous of every atom and the rest are heavier "brothers" of it. Netrinos on the other hand do not take part in any known nucleous-like formation they are also massless within the Standard Model making them extremelly (and notoriously) difficult to detect. They live in the least well known "neighborhoud"of the Standard Model. Their presence is indirectly implied in the radioactive decay of atoms and. The gauge bossons γ, g, $W^\pm$ Z^0 are responsible for mediating the electromagnetic, strong, and weak interactions between quarcks or leptons repsectively. These possible interactions are indicated by the blue lines. Some of the gauge bossons can interact with themselves which is illustrated by the blue closed loop lines.

The core element that makes Eq. 1.1 elegant is the so called *local gauge invariance*[ref??]. Briefely speaking this means that Eq. 1.1 is invariant (or symmetric) under spacetime dependant phase[6] transformations of the quark or lepton fields. By construction, the Lagrangian obeys the symmetry group[7] $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. This means that there are three distinct transformations of the ψ fields that Eq. 1.1 is symetric to. Each one introduces

---

[6]the quark and lepton quantum fields are complex valued quantities. Thus phase here means complex phase.

[7]with the mathematical notion of a group implied

the electromagnetic, weak and strong interactions between the fields. The mathematics behind the above mentioned transoformations really gives the current paragreaph proper meaning but it would be completeley out of context here. Further reading on the symmetries of th Standard Model can be found in [?????????]

Within the accuracy of the current experiments the Standard Model has seen its predictions confirmed to a great extend. The most recent and perhaps one of the most crucial meaning the discovery of the mechanism that particles acquire mass, makes the Standard Model looks quite robust. However there are phenomena and observations that it can not account for. Perhaps the most stricking one is the absence of any discription about the most familiar and strong force of nature, meaning gravity[8]. or the peculiar value of the cosmological constant leading to dark energy interpretations[[]].But even at the heart of the Standard Model there obscure aspects for example the well established fact that neutrinos have non zero mass[[]] or the unexplained amount of the observed matter-antimatter assymetry in the universe [[]]. For all of the above reasons plus our curiosity driven nature scientists are compled to continue testing the Standard Model and look for deviations of its predictions.

1.2 Flavour Physics

$$\mathscr{L}_{\text{Yukawa}} = -Y_{ij}^d \bar{\psi}_{Li} \phi \psi_{Rj} \ + \ h.c. \tag{1.2}$$

- Yukawa couplings.

- Quark mixing

- CKM picture

- three generatios anthopic principle.

- The falvour physics observables, puzzle fasi, hits for Np, We need everything.

Lastly, the vast majority of the observed matter is built from u, d quarks and e which makes the explanation of the triplet like structure of quarcks and leptons very enigmatic and difficult to understand from first principles

1.3 The $B_s^0 - \bar{B}_s^0$ system and the ϕ_s parameter

- Effective hamiltonian for $B_s^0 - \bar{B}_s^0$ (mass and cp eigenstates)

- feynman diagrams and qft and lagrangian.s (maybe in the next sectyion)

- ϕ_s parameterization

- ϕ_s status

- NP in ϕ_s

---

[8]Gravity is not renormalisable[[]] theory and thus cannot be described as a quantum field theory

1.4   Higher order effects in ϕ_s and the $B_s^0 \to J/\psi \overline{K}^{*0}$ decay

- Decay topologies

- Feynman diagrams

- $B_s^0 \to J/\psi \overline{K}^{*0}$ as a control chanell for ϕ_s

- Estimate of the expected yield

- Define parameters of interest

- Flavour specific final state (no oscilations, the sign tags the flavour at production)

- Event display

$$A^{CP}(B_{(s)}^0 \to f_{(s)}) =$$

$$\frac{\int_0^\infty \left[\Gamma(B_{(s)}^0 \to \bar{f}_{(s)}) + \Gamma(B_{(s)}^{\bar{0}} \to \bar{f}_{(s)})\right]\mathrm{d}t \; - \; \int_0^\infty \left[\Gamma(B_{(s)}^0 \to f_{(s)}) + \Gamma(B_{(s)}^{\bar{0}} \to f_{(s)})\right]\mathrm{d}t}{\int_0^\infty \left[\Gamma(B_{(s)}^0 \to \bar{f}_{(s)}) + \Gamma(B_{(s)}^{\bar{0}} \to \bar{f}_{(s)})\right]\mathrm{d}t \; + \; \int_0^\infty \left[\Gamma(B_{(s)}^0 \to f_{(s)}) + \Gamma(B_{(s)}^{\bar{0}} \to f_{(s)})\right]\mathrm{d}t} \quad (1.3)$$

$$A^{CP} = \frac{\Gamma(B_{(s)}^{\bar{0}} \to \bar{f}_{(s)}) - \Gamma(B_{(s)}^0 \to f_{(s)})}{\Gamma(B_{(s)}^{\bar{0}} \to \bar{f}_{(s)}) + \Gamma(B_{(s)}^0 \to f_{(s)})} \quad (1.4)$$

Assuming negligible cpv in mixing, Mention this.

$$A_{CP}^{\mathrm{raw}}(B_{(s)}^0 \to f_{(s)}) = \frac{N^{\mathrm{obs}}(\bar{f}_{(s)}) - N^{\mathrm{obs}}(f_{(s)})}{N^{\mathrm{obs}}(\bar{f}_{(s)}) + N^{\mathrm{obs}}(f_{(s)})} \quad (1.5)$$

Chapter 2

Data Analysis

The parameters of interest as described in Section 1.4 are extracted from the analysis of the Run I 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 a weighted maximum likelihood fit to the angular distributions of $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ is performed, where the weights have been computed with the *sPlot* technique [21] described in Section 2.1.3. Prior to any analysis the data obtained from the detector need to be filtered to reduce the presence of background events. Thus the current chapter addresses first the event selection and then the angular analysis.

Section 2.1 describes all the steps involved to reduce background in the data. For this purpose it is necessary to model the $J/\psi K \pi$ invariant mass distribution in the region of the B_s^0 meson, which is addressed in Section 2.1.3. In Section 2.2 the angular analysis is described.

2.1 Event Selection

There are effectively four stages of offline event selection taking place after the data has been recorded. First a set of selection criteria is applied (hereafter *Stripping*). Each of those criteria[1] requires from a certain observable to satisfy some condition. For example the mass of the mother particle B_s^0 must be within a certain range[2]. A detailed table of the Stripping selection is given in Table A.1. The purpose of stripping is to perform a loose rejection of the combinatorial background and prepare the data for the next stages of selection. Stripping is also a way to save disk space for offline event storage, since it is impossible to save anything that the detector writes out. The remaining three selection steps after Stripping are described in each of the subsequent subsections.

2.1.1 Multivariate Based Selection

The $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ signal yield out of the full $3\,\text{fb}^{-1}$ data-set is expected to be low, see Section 1.4. Thus, one would like to reject as much background as possible while keeping all the signal. One way to do that would be to tighten the Srtripping selection criteria one by one[3]. Alternatively a multivariate approach (here after *MVA*) is adopted. In that case a set of variables

---

[1] also referred to as a *cut*

[2] Or another type of cut could be the probability that a final state particle, like the pion, is misidentified as a muon.

[3] also known as *cut-based analysis*

are combined by the MVA algorithm to produce a single variable, the *classifying variable*. The MVA approach tries exploits the correlations between input variables to get the maximum discriminating power out of their combination. The classifying variable ranges from −1 to +1. Signal candidates tend to peak closer to +1, whereas background candidates closer to -1.
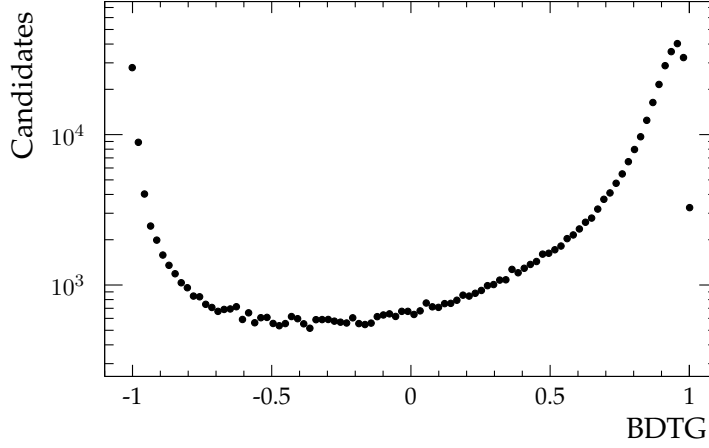

FIGURE 2.1: BDTG distribution on real data. y-axis in logarithmic scale. Signal (Background) like events peak to the right (left).

For the current analysis the TMVA toolkit [10] is used for the MVA selection. In order to construct the classifying variable two sets, one signal-like and one background-like data sets, are needed. This pair of data-sets is fed 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. Next another independent pair of sets is used to asses how successful the training step was, this step is called *testing*. For the signal-like samples, $B_s^0 \to J/\psi \overline{K}^{*0}$ Monte-Carlo simulated data (here after simulated data) are used. The B_s^0 mass window for that sample is $\pm 25 \, \mathrm{MeV}/c^2$ around the B_s^0 peak. As for the background-like sample, candidates from the high mass side-band with invariant masses between $5401.3 \, \mathrm{MeV}/c^2$ and $5700 \, \mathrm{MeV}/c^2$ are used. Note that the simulated data are treated in the same way as the real data do by all the processing steps after the L0 trigger level. A boosted decision tree (here after BDTG) is used as the classifying variable.

The BDTG shows a good discrimination power over signal and background distributions Figure 2.1. Note also that the the BDTG was checked for possible over-training. The last is a situation that the BDTG becomes sensitive to statistical fluctuations between statistically compatible distributions, such as the testing and training samples. The result is worst signal from background separation. Once the training and testing steps are complete a cutoff value on the BDTG is applied so that it maximizes the following figure of merit (hereafter *FoM*) [26]:

$$F(w_i) = \frac{\left(\sum w_i\right)^2}{\sum w_i^2},$$

(2.1)

where w_i are weights associated to each event (here after *sWeights*), and

calculated with the *sPlot* technique[4]. This FoM can be understood as an effective signal yield, which is inversely proportional to the square root of the total number of events. For a range of BDTG values the *sPlot* technique provides a different set of *sWeights* based on which a different FoM value can be obtained. The optimum BDTG value is chosen as the one that maximizes $F(w_i)$.
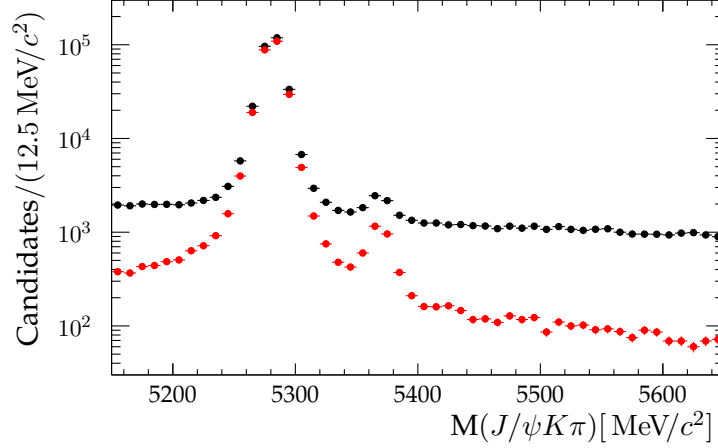

FIGURE 2.2: Mass distribution before (black) and after (red) BDTG selection. y-axis in logarithmic scale

2.1.2 Peaking Backgrounds

After applying the BTDG cut there is still some combinatorial background remaining which is removed further using the *sPlot* technique described in Section 2.1.3. However, there is one more crucial step that is necessary to address prior to the *sPlot* technique, namely that of the peaking backgrounds. Studies of simulated samples show contributions from several specific backgrounds, such as $B_s^0 \rightarrow J/\psi K^+ K^-$, $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ and $B_d^0 \rightarrow J/\psi \pi^+ \pi^-$. Those backgrounds are the result of misidentifying a final state particle by the event building algorithms, for example a pion can be confused for a kaon. The invariant mass distribution of the above mentioned misidentified backgrounds peak near the $B_{d(s)}^0 \rightarrow J/\psi K \pi$ signal peaks. Furthermore given the low number of expected peaking background events, see Table 2.1, their distribution is rather discontinuous with sharp spikes. This type of distributions have to be removed before the next selection step described in Section 2.1.3 is applied. The last technique is more effective when the distributions involved are broader and continuous. Therefore, specific simulated events with negative weights are appended to the data sample such that they cancel out the contribution of the peaking background events[5]. The current section briefly addresses the treatment of peaking backgrounds with negative weights (this is a two step procedure) as well as the special treatment of the $\Lambda_b^0 \rightarrow J/\psi p \pi^-$ peaking background.

---

[4]The *sPlot* technique is a statistical tool to unfold data distributions [21]. For example disentangle signal from background. Further details in Section 2.1.3.

[5]The term "cancel out" obtains more meaning in the context of the likelihood of each event which is built in the next selection step Section 2.1.3 based on a given mass PDF.

As a first step towards removing the peaking background contribution it is necessary to have an estimate on the expected peaking backgrounds yields on data. This is done using simulated data and based on the expression:

$$N_{\mathrm{exp}} = 2 \times \sigma_{b\bar{b}} \times f_q \times BR \times \varepsilon \times \mathcal{L}, \tag{2.2}$$

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 forms a meson with another quark type q), ε is the total efficiency (reconstruction, selection and trigger) and $\mathcal{L}$ is the luminosity of the data. Finally BR stands for the branching fraction of the particular peaking background. Since simulated data are used it is necessary to estimate the effective luminosity of that simulated sample and scale it to the luminosity of the data. After that the number of peaking background events from the simulated sample is a valid estimate of the one in real data[6].

The second and last step of the peaking background removal is to apply an angular correction factor to account for the fact that simulated events are distributed uniformly in phase space[7] and hence do not contain the proper polarization amplitudes, mentioned in Section 1.4. This can cause the peaking background yield estimations to be wrong because the simulated events are distributed in the $(\Omega, m_{K\pi})$ space in a different way than the actual peaking background in the data. The amplitude analysis of $B_d^0 \to J/\psi\pi^+\pi^-$, $B_s^0 \to J/\psi\pi^+\pi^-$, $B_s^0 \to J/\psi K^+K^-$ and $\Lambda_b^0 \to J/\psi p K^-$ which has been performed in [8], [7], [1] and [15] respectively provides enough information for the amplitude structure of those modes. Therefore the simulated events are weighted with

$$w_{\mathrm{MC}} = \frac{P_{\mathrm{DATA}}(\Omega, m_{hh}|A_i)}{P_{\mathrm{MC}}(\Omega, m_{hh})}, \tag{2.3}$$

where P_{DATA} and P_{MC} are normalized PDFs[8] and A_i stands for the particular amplitude structure of the certain peaking background mode. The final peaking background yields estimation after the above mentioned steps is shown in Table 2.1

As it was previously mentioned, the $\Lambda_b^0 \to J/\psi p\pi^-$ peaking background is treated specially. Instead of following the above procedure, the $\Lambda_b^0 \to J/\psi p\pi^-$ invariant mass line shape is modeled and statistically subtracted in the same way as the combinatorial background in the next section. The reasons for this different treatment with respect to other peaking backgrounds are twofold:

- The full amplitude structure of $\Lambda_b^0 \to J/\psi p\pi^-$ decays is not yet known, and thus the weighting of the simulated samples is not possible.

- The peak of the misidentified $\Lambda_b^0 \to J/\psi p\pi^-$ decays in the $J/\psi\ K\pi$ mass spectrum is broader than those of the other peaking backgrounds see Figure 2.3, as a result the *sPlot* technique still effective.

---

[6]Assuming blah. It will be fixed in the next iteration.

[7]Phase space generated events use only kinematic variables for the generation thus there is no angular dependence. On the other hand real data have intrinsic angular dependence

[8]PDF is an acronym for Probability Density Function

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

TABLE 2.1: Approximated expected yields in $\pm 70\,\mathrm{MeV}/c^2$ $K\pi$ mass window of each background after re-weighting of the phase space (the $\Lambda_b^0 \to J/\psi p \pi^-$ decay is not weighted since no amplitude analysis for that decay is published).

2.1.3 sWeighting and Invariant Mass Distribution

After successfully appending simulated events with negative weights to remove the peaking background contributions, the data sample is left only with $B^0 \to J/\psi K\pi$, $B_s^0 \to J/\psi K\pi$, $\Lambda_b^0 \to J/\psi p\pi^-$, and combinatorial background. These four modes (also refereed to as species) are statistically disentangled using the *sPlot* technique [21]. The technique requires as an input a PDF that estimates the yield of each of the above specie. The current analysis uses the PDF of the $J/\psi K\pi$ invariant mass, hereafter just mass, where the yield of each specie is estimated with an extended maximum likelihood fit to the mass distribution on data, hereafter mass fit.

The mass PDF

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

- Combinatorial background: Exponential decay.

- $\Lambda_b^0 \to J/\psi p\pi^-$ decays: Amoroso distribution [11]

- B^0 and B_s^0 signals: Hypatia distribution [20].

An exponential decay distribution, e^{-km}, is found to be appropriate description of the combinatorial background. It accounts for events where random combinations of final state particles are combined to a signal candidate. Those kind of events are not expected to exhibit any structure in the mass distribution, they simply follow a exponential distribution with a negative slope in the mass spectrum. The slope is negative because with increasing mass the frequency of those random combinations drops.

As for the $\Lambda_b^0 \to J/\psi p\pi^-$ 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. Amoroso is actually

a family of distributions. Depending on the values of the shape parameters the resulting distribution can vary from an exponential to Gaussian like. This is a powerful feature that is exploited in order to describe the $\Lambda_b^0 \to J/\psi p \pi^-$ peaking background shape.

The Hypatia distribution is chosen mainly because it nicely describes the tails of the B_s^0 and B^0 invariant mass peaks which is of crucial importance. These two peaks are sufficiently close, see picture Figure 2.3, that, in case of bad modeling of the tails, event leakage between the B_s^0 and B^0 peaks takes place. 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/ψ mass constrain or badly reconstructed events caused by decays of the final state hadrons, see [20]. The tail parameters (α, n) are four in total (two for each tail) and are taken from a fit to MC simulated events with a known resolution. The last makes sure that the tail parameters do not rely on detector simulation imperfections. For the core of invariant mass distribution, Hypatia uses five shape parameters, namely ζ, β, λ, σ, μ. The first two are set to zero[9], the third one is taken from the previous MC simulated sample along with the tail parameters [10] whereas the last two which are the width and the mean of the core are allowed to float in the mass fit.

The mass fit

From MC studies, 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 bins of the $m_{K\pi}$ invariant mass. In addition, due to correlations between the mass and one of the variables used in the subsequent angular fit (meaning the $\cos\theta_\mu$ angle), the requirements of the *sPlot* technique are not satisfied and thus it cannot be applied. 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 mass final mass fit to the data is shown in Figure 2.3, whereas the overall B_s^0 and B^0 yields are:

$$N_{B^0} = 208656 \pm 462^{+78}_{-76} \qquad (2.4)$$

$$N_{B_s^0} = \quad 1808 \pm 51^{+38}_{-33}, \qquad (2.5)$$

where the first uncertainties are statistical and obtained from the quadratic sum of the ones in each fitting category, and the second uncertainties correspond to systematical. The correlations between the B^0 and B_s^0 yields in each fitting category are found to be small (smaller than 4%). Note that the mass fitting model has been validated with a procedure similar to the one described in Section 2.4.4.

---

[9]ζ is empirically found to be very small whereas $\beta = 0$ implies that the core is symmetric left and right with respect to the mean.

[10]In the limit of $\zeta = 0$ λ does not depend on detector effects but only on particle kinematics, the same way as the tail parameters do.
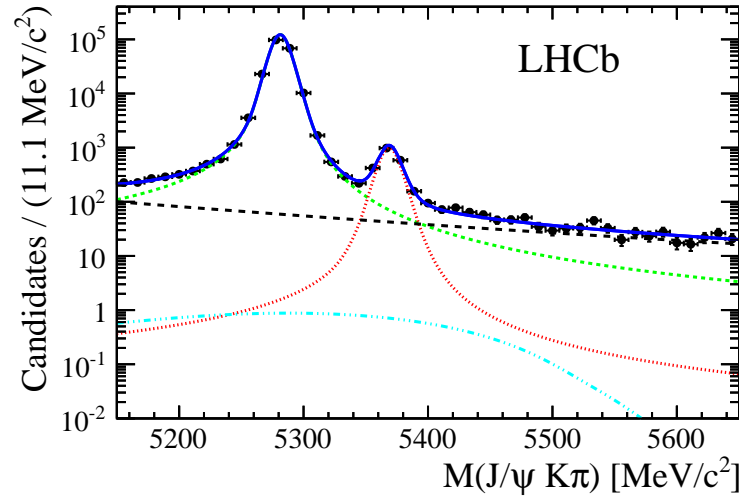
FIGURE 2.3: Invariant mass fit to the data. y-axis is in logarithmic scale.

sWeighting

Having performed the mass fit 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 (also called sWeight), to each event, 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 only that particular specie on the data while eliminating all the others. In the context of the current analysis events that are likely to be background such as combinatorial and $\Lambda_b^0 \to J/\psi p K^-$ are eliminated while $B_{(s)}^0 \to J/\psi \overline{K}^{*0}$ events are projected with the above mentioned technique (also referred to as sWeighting).

The advantage of the *sWeights* approach becomes apparent when performing the angular analysis, described in the subsequent section. Specifically since the resulting $B_s^0 \to J/\psi \overline{K}^{*0}$ weighted events can be described by the signal PDF, meaning there is no need to model the background shape of the angular distributions. Also the fit itself is faster and simpler to implement. On the other hand there are a few disadvantages such as the evaluation of some systematic uncertainties on the fitted angular parameters due to the mass PDF modeling. Any variation on those models implies recomputing the *sWeights*. Also special care needs to be taken such that the uncertainties of the fitted angular parameters are correct [21]. Particularly The uncertainties tend to be diluted in a fit with weighted events. This is because the per event weights in a weighted likelihood fit change the shape of the likelihood function. The last is where the statistical uncertainty estimation comes from. Thus to correct for this dilution the scale factor

$$\alpha = \left(\frac{\sum_i w_i}{\sum_i w_i^2} \right)^{1/2} \tag{2.6}$$

has to be applied to each weight before performing the angular fit. The scale factor essentially forces the sum of weights to be the same as the number of

events.

2.2   Angular Analysis

The core formalism and all the necessary steps towards building a PDF to describe the amplitude properties and angular structure of $B^0_{(s)} \to J/\psi \overline{K}^{*0}$ decays are presented in the current section. A brief description of the angular PDF derivation is given in Section 2.2.1. Treatment of the angular acceptance are addressed in Section 2.2.2 and Section 2.2.3. In Section 2.2.4 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 2.2.5. Some aspects of the maximum likelihood fit are presented in Section 2.2.6, along with the $\mathcal{A}^{CP}$ parameters of interest, which are introduced in a special way in angular fit when building the full decay rate PDF.

2.2.1   Angular Dependence

The decay of interest, $B^0_{(s)} \to J/\psi \overline{K}^{*0}$, is a P2VV process[11] with 4 particles in the final state. There are at least two ways to describe the angular dependence of that decay, namely the transversity framework [17, 25] and the helicity formalism [16, 18]. In both cases the goal is to come up with an angular dependent description of the total decay amplitude. In the transversity framework the decay amplitude is decomposed to the three possible polarization states of the intermediate particles J/ψ and K^{*0} with respect to the B^0_s rest frame using polarization vectors of the decay amplitude. For the current analysis the helicity formalism is adopted where the angular dependence is introduced by summing all possible spin configurations of the intermediate vector particles relative to their momentum direction in the B^0_s 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 of both approaches within the scope of a P2VV decay can be found in [14] and [19] respectively for the transvercity and helicity formalism.

The current paragraph aims at very briefly explaining the steps needed to arrive to an expression for the angular decay rate for $B^0_s \to J/\psi \overline{K}^{*0}$ decays, which is the sum of ten terms. As shown in Eq. 2.7

$$\frac{d\Gamma(B^0_{(s)} \to J/\psi \overline{K}^{*0})}{d\Omega \, dm_{K\pi}} \propto |\mathcal{A}(B^0_{(s)} \to J/\psi \overline{K}^{*0})|^2 =$$

$$= \left| \sum_i^{0,\|,\perp,S} A_i(\Omega, m_{K\pi}) \right|^2 \propto \sum_n a_n h_n M_n, \quad (2.7)$$

where the term $\mathcal{A}(B^0_{(s)} \to J/\psi \overline{K}^{*0})$ is the amplitude of the decay which is decomposed in its P-wave (0, $\|$, $\perp$) and S-wave (S) polarizations. The

---

[11]P2VV is an acronym referring to the spin of the particles involved in the decay. The B^0_s (and B^0) is spin-0 parity minus (Pseudo-scalar) particle, whereas the intermediate resonances J/ψ and ϕ are spin-1 (Vector) particles. Hence the acronym P2VV.

TABLE 2.2: Angular functions corresponding to each term in Eq. 2.7 for the $B^0_{(s)} \to J/\psi \overline{K}^{*0}$ decay. 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 are expressed in the helicity basis. The angles $\cos\theta_K, \cos\theta_\mu, \varphi_h$ are called *helicity angles* and in Figure 2.4 are put into perspective. The middle column express the angular dependence in an orthogonal basis and it is equivalent to the last column. The P and Y symbols denote associated Legendre polynomials and real valued spherical harmonics respectively.

a_n	$h_n(\Omega) \times 16\sqrt{\pi}$	$h_n(\Omega) \times \frac{32\pi}{9}$
$\|A_0\|^2$	$4\left(P_0^0 + 2\,P_2^0\right)\left(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0}\right)$	$2\,\cos^2\theta_K\,\sin^2\theta_\mu$
$\|A_\parallel\|^2$	$P_2^2\left(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} - \sqrt{\frac{3}{5}}\,Y_{2,+2}\right)$	$\sin^2\theta_K\left(1 - \sin^2\theta_\mu\,\cos^2\varphi_h\right)$
$\|A_\perp\|^2$	$P_2^2\left(2\,Y_{0,0} + \frac{1}{\sqrt{5}}\,Y_{2,0} + \sqrt{\frac{3}{5}}\,Y_{2,+2}\right)$	$\sin^2\theta_K\left(1 - \sin^2\theta_\mu\,\sin^2\varphi_h\right)$
$\Re(A_0^* A_\parallel)$	$+2\sqrt{2}\sqrt{\frac{3}{5}}\,P_2^1\,Y_{2,+1}$	$+\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(A_0^* A_\perp)$	$-2\sqrt{2}\sqrt{\frac{3}{5}}\,P_2^1\,Y_{2,-1}$	$-\frac{1}{\sqrt{2}}\,\sin 2\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$
$\Im(A_\parallel^* A_\perp)$	$+2\sqrt{\frac{3}{5}}\,P_2^2\,Y_{2,-2}$	$+\sin^2\theta_K\,\sin^2\theta_\mu\,\sin 2\varphi_h$
$\|A_S\|^2$	$4\,P_0^0\left(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0}\right)$	$\frac{2}{3}\,\sin^2\theta_\mu$
$\Re(A_0^* A_S)$	$8\sqrt{3}\,P_1^0\left(Y_{0,0} - \frac{1}{\sqrt{5}}\,Y_{2,0}\right)$	$\frac{4}{3}\sqrt{3}\,\cos\theta_K\,\sin^2\theta_\mu$
$\Re(A_\parallel^* A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}\,P_1^1\,Y_{2,+1}$	$+\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\cos\varphi_h$
$\Im(A_\perp^* A_S)$	$+6\sqrt{2}\frac{1}{\sqrt{5}}\,P_1^1\,Y_{2,-1}$	$+\frac{1}{3}\sqrt{6}\,\sin\theta_K\,\sin 2\theta_\mu\,\sin\varphi_h$

terms a_n come from squaring the amplitude and their exact expressions are shown in the first column of Table 2.2[12]. Note the appearance of P-wave self interference and SP-wave interference terms. The functions h represent the angular dependence of each term, whereas M_n stands for the $m_{K\pi}$ dependence of the amplitude and its special treatment is postponed for the next section. Table 2.2 lists the angular part of the decay rate (meaning the h functions) equation which is the result of applying the *helicity formalism* in the $B^0_{(s)} \to J/\psi \overline{K}^{*0}$ decay. The compact notation Ω stands for the so called *heliciy angles*. Three angles are required in a 4 body decay which along with the $m_{K\pi}$ total the degrees of freedom necessary to describe such a system. Those angles are $\theta_K, \theta_\mu, \varphi_h$ and are defined in Figure 2.4. In that figure it can be seen that θ_μ is the angle of the positive μ with respect to the momentum of the J/ψ in the B_s^0 rest frame. Similarly for θ_K, the K is used to define the angle with respect to the momentum of the intermediate $K\pi$ resonance. Finally φ_h is the relative angle between the $K\pi$ and dimuon decay planes, where each plane is defined in the rest frame of the respective intermediate resonance.

Lastly, the mathematically elegant orthogonal angular basis of P and Y [13] is adopted in the current analysis, since the integrals of type $\int_{-1}^{1} P\,d\cos\theta_K$

---

[12]It is interesting to see in [19] how the $\Re$ and $\Im$ parts of the interference terms show up.

[13]The decomposition of angular functions in an orthogonal basis follows naturally from the fact that spherical harmonics can be expressed as Wigner-D matrices. More details on this mapping can be found in [19]

and $\int_{\cos\theta=-1,\varphi=-\pi}^{\cos\theta=1,\varphi=\pi} Y\, d\cos\theta\, d\varphi$ are known analytically thus significantly reducing the fitting time and simplifying the implementation of the angular functions.
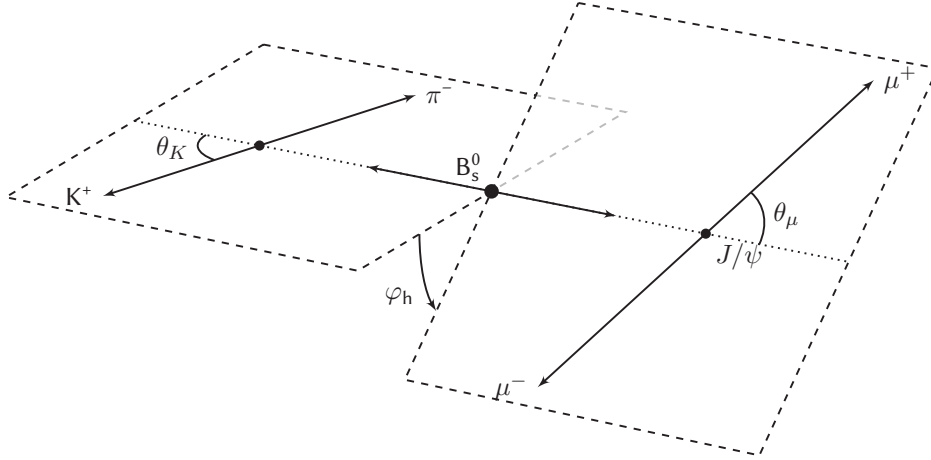

FIGURE 2.4: Definition of the decay angles in the helicity formalism.

2.2.2 Acceptance

While the effects of angular resolution are known to be negligible for this analysis, angular acceptance on the other hand can not be neglected. The current section introduces the acceptance shape. After that the parametrization of the acceptance is discussed and finally an important issue related to the choice of parametrization is dealt with.

The shape of the acceptance can be seen in Figure 2.5a to Figure 2.5c. The most striking feature is the drop of efficiency close to $\cos\theta_K = 1$, which implies that the pion is more likely to fly out of the geometrical acceptance of the detector, compared to a kaon, and thus not be detected. This feature can be understood given the definition of $\cos\theta_K$ in Figure 2.4. The scenario where $\cos\theta_K \to 1$ and hence $\theta_K \to 0$ corresponds to a soft pion after boosting to the lab frame. Since the pion is lighter than a kaon it is more likely that it flies out of the acceptance more often than a kaon does, creating the observed asymmetric efficiency loss in the $\cos\theta_K$ projection. With that said, it is intuitive that this type of efficiency loss is driven by the mass difference between the two particles. In the case of the $\cos\theta_\mu$ the inefficiency introduced as described before for the $\cos\theta_K$ projection, is symmetric with respect to the positive and negative muons since they have the same mass. As for the acceptance in the φ_h plane it is very close to being flat. By staring again Figure 2.4 one does not expect any reason for a strong acceptance effect in that projection. In addition to this type of acceptance effects any of the selection cuts applied could also introduce some inefficiency.

Acceptance parametrization

As for the parametrization of the angular acceptance there are at least two ways to proceed, namely the *efficiency moments* [19] and the *normalization weights* [19, 23]. The current analysis implements the first one, where the
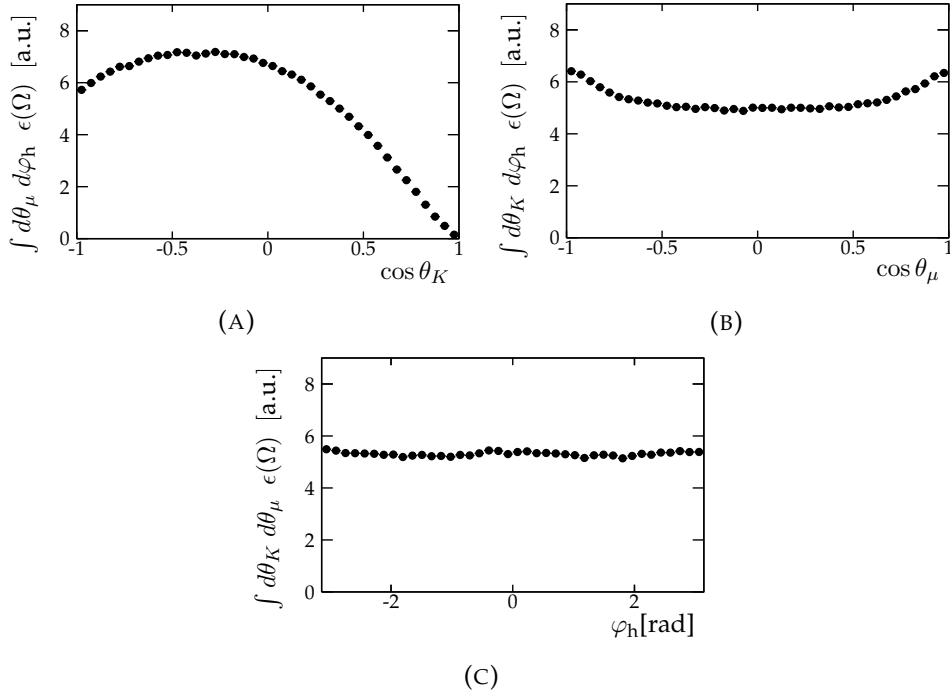
(A)

(B)

(C)

FIGURE 2.5: Angular acceptance shape from simulated data. Each event in the simulated sample has been weighted with the inverse integral of the PDF that was used to generate them with, resulting to those three distributions that give a feeling of the acceptance shape.

efficiency function is parametrized using orthogonal functions (nicely combined with the angular description formalism in the previous subsection). Such a parameterization can be seen in the next equation Eq. 2.8. The symbol $\alpha_i^{jk}(\Omega)$ is introduced as a shorter notation.

$$\epsilon(\Omega) = \sum_{ijk} c_{jk}^i \, P_i^0(\cos\theta_K) \, Y_{jk}(\cos\theta_\mu, \varphi_\mathrm{h}) \equiv c_{jk}^i \, \alpha_i^{jk}(\Omega), \qquad (2.8)$$

where the familiar from the previous section spherical harmonics (Y) and associated Legendre polynomials (P) show up as the basis functions on which the angular acceptance is decomposed to the PY basis by means of the coefficients c_k^i, hereafter efficiency moments (or simply moments). The reason of choosing the particular parameterizations is obviously for the known analytical integrals of product of two P or Y. Following the above parameterizations one just needs to choose a set of c_{jk}^i and find a way compute them. In principle there are infinite efficiency moments since the sum in Eq. 2.8 runs through all possible values of the ijk indices. However in practice only a small number of efficiency moments contribute significantly to the proper description of the angular acceptance shape. The efficiency moments used to describe the acceptance shape are shown in the first column of Table 2.3. The choice of those particular 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 then is defined as the ratio of two angular dependent 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 later is the angular PDF which the simulated data have been generated with. It follows that the efficiency is defined as the ratio of the first over the later and it is shown in equation Eq. 2.9.

$$\epsilon(\Omega) \equiv \frac{P^{\text{obs}}(\Omega)}{P^{\text{gen}}(\Omega)}. \tag{2.9}$$

Computing the efficiency moments is done by projecting the efficiency as defined in Eq. 2.9 on the orthogonal PY basis. This projection is shown in Eq. 2.10. Where the factor $j + \frac{1}{2}$ is used to account for the fact that the associated Legendre polynomials form an orthogonal basis but in fact the orthonormal behavior is required to properly compute the efficiency moments.

$$
\begin{aligned}
c_{jk}^i &\equiv (j + \frac{1}{2}) \int d\Omega \ \alpha_i^{jk}(\Omega) \ \epsilon(\Omega) \\
&= (j + \frac{1}{2}) \int d\Omega \ P^{\text{obs}}(\Omega) \ \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)}.
\end{aligned} \tag{2.10}
$$

The last step in the calculation of the efficiency moments is the computation of the integral in Eq. 2.10. This is done by means of simulated data and the integral is computed following the concept of Monte Carlo integration. This implies that the integral in Eq. 2.10 can be estimated by a sum over the simulated data (which they have passed through the exact same steps selection as the real data) as shown in Eq. 2.11.

$$E\left[\int d\Omega \ \alpha_i^{jk}(\Omega) \frac{P^{\text{obs}}(\Omega)}{P^{\text{gen}}(\Omega)}\right] = \frac{1}{N^{\text{obs}}} \sum_e^{N^{\text{obs}}} \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)}. \tag{2.11}$$

Combining Eq. 2.10 and Eq. 2.11 the final expression for computing the efficiency moments is shown in Eq. 2.12 and the result of the computations can be found it table Table 2.3/

$$c_{jk}^i \equiv (j + \frac{1}{2}) \frac{1}{N^{\text{obs}}} \sum_e^{N^{\text{obs}}} \frac{\alpha_i^{jk}(\Omega)}{P^{\text{gen}}(\Omega)} \tag{2.12}$$

moment	central value	standard deviation
c_{00}^0	constrained	-
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 2.3: Efficiency moments of $B_s^0 \to J/\psi \overline{K}^{*0}$. The bigger the value of a moment the more the corresponding function contributes to the overall efficiency function. The efficiency functions corresponding to the moments c_{00}^1 and c_{00}^2 account for most of the hard acceptance drop in the $\cos\theta_K$ projection. The acceptance in the $\cos\theta_\mu$ projection is much softer compared to the one $\cos\theta_K$. This can be deduced from Figure 2.5b as well as from the values of moments like c_{jk}^0 compared to the values of c_{00}^1 and c_{00}^2. The value of the c_{00}^0 is constrained for reasons that have to do with the value of the acceptance close to $\cos\theta_K = 1$. Details on the constrain are explained in the current subsection. The efficiency moments are correlated the corresponding correlation matrix can be found in Table B.6

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 that is adopted in this analysis it can happen that the acceptance function Eq. 2.8 can take negative values depending on the choice of certain efficiency moments. In other words the approach of efficiency moments does not guarantee that the acceptance shape described by the same equation is a positive definite quantity. Note that this is an artifact of the parametrization and by no means is there something wrong in the data or the detector. The problem is mathematical in nature and there is no standard solution (known to the author when those lines were written). The approach followed to solve this problem has two steps. First a constrain on the acceptance function, Eq. 2.8, is introduced to force its value positive and second a constrained fit to the efficiency moments is performed to optimize the acceptance shape description. The result is a hybrid acceptance function that is positive definite.

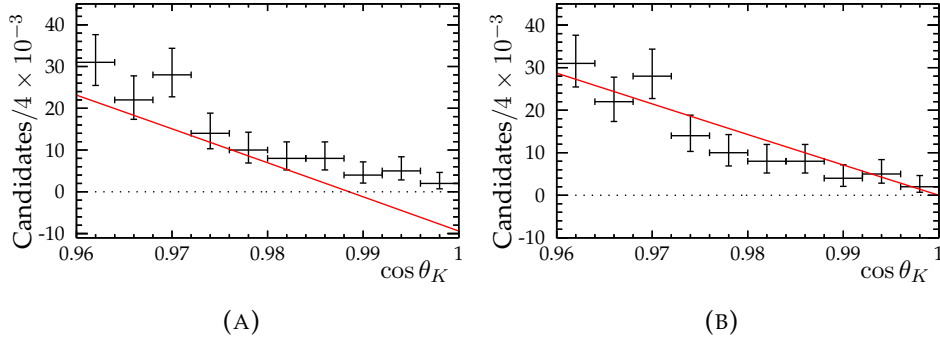Before laying out the formulas behind the above mentioned steps it is

(A)   (B)

FIGURE 2.6: Angular $\cos\theta_K$ PDF projection corrected with
the angular efficiency function before (A) and after (B) con-
straining its value at $\cos\theta_K = 1$, shown as the red curve.
The black points show the $\cos\theta_K$ distribution in the simu-
lated sample. The x-axis scale does not correspond to the
full $\cos\theta_K$ range but it is zoomed in to clearly show the
problematic region. It is clear in the right hand side of the
figure that the problem is lifted by constraining the accep-
tance.

useful to literally see the problem caused by the parametrization, see Fig-
ure 2.6a and the corresponding caption. Having done that the most natural
thing to do before making any effort to force the value of Eq. 2.8, here-
after efficiency function, positive at $\cos\theta_K = 1$ is to make the best choice
of efficiency moments. The moments are chosen such that the value of the
efficiency function is as "less negative" as possible[14].

Having chosen the best possible set of efficiency moments the value of
the efficiency function is constrained by solving it for one of the efficiency
moments, which can be easily done for c_{00}^0. The constrain that is imposed
is shown in Eq. 2.13

$$\epsilon(\cos\theta_K = 1, \cos\theta_\mu, \varphi_\mathrm{h}) = 0. \tag{2.13}$$

Starting from the efficiency function of Eq. 2.8 (where the sum now runs
over the indices of the first column in Table 2.3):

---

[14]It is actually not trivial to find a good choice of efficiency moments. Two things are rel-
evant when trying to describe a distribution with orthogonal functions like P and Y. 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 effi-
ciency 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 2.3) 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 of its three
dimensions at the same time.

$$\epsilon(\Omega) = \sum_{ijk} c^i_{jk} P^0_i(\cos\theta_K) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h})$$

$$= c^0_{00} P^0_0(\cos\theta_K) Y_{00}(\cos\theta_\mu, \varphi_{\rm h}) + \sum_{ijk}^{ijk \neq 000} c^i_{jk} P^0_i(\cos\theta_K) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}).$$

$$(2.14)$$

Given Eq. 2.13 and also $P^0_0 = 1$, $Y_{00} = \frac{1}{2\sqrt{\pi}}$ Eq. 2.14 can be easily solved for c^0_{00} resulting in Eq. 2.15.

$$c^0_{00} = -2\sqrt{\pi} \sum_{ijk}^{ijk \neq 000} c^i_{jk} P^0_i(1) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) \qquad (2.15)$$

Eq. 2.14 is used along with Eq. 2.15 resulting finally in Eq. 2.16. The last is an updated efficiency function which is guaranteed to be positive definite at $\cos\theta_K = 1$.

$$\epsilon(\Omega) = \left[-2\sqrt{\pi} \sum_{ijk}^{ijk \neq 000} c^i_{jk} P^0_i(1) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) \right] \overbrace{P^0_0(\cos\theta_K)}^{1} \overbrace{Y_{00}(\cos\theta_\mu, \varphi_{\rm h})}^{\frac{1}{2\sqrt{\pi}}}$$

$$+ \sum_{ijk}^{ijk \neq 000} c^i_{jk} P^0_i(\cos\theta_K) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) =$$

$$= \left[-\sum_{ijk}^{ijk \neq 000} c^i_{jk} Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) \right] + \sum_{ijk}^{ijk \neq 000} c^i_{jk} P^0_i(\cos\theta_K) Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) =$$

$$= \sum_{ijk}^{ijk \neq 000} c^i_{jk} \left[P^0_i(\cos\theta_K) - 1 \right] Y_{jk}(\cos\theta_\mu, \varphi_{\rm h}) \qquad (2.16)$$

The result of the above constrain is shown in figure Figure 2.6b, where it can be seen that the acceptance projection on $\cos\theta_K = 1$ is not zero anymore. In addition it has also been checked whether the acceptance is negative in 2D and 3D, since it is possible for a 2D projection of the acceptance to take negative values as well while the 1D projections are positive. Two dimentional projection of the constrained acceptance can be seen in Figure 2.7.

Constraining the acceptance in this manner described above comes with a small penalty, meaning that the shape of the 1D acceptance projections is distorted a bit. In order to undo this distortion and describe better the acceptance shape an additional step is implemented. As it was previously mentioned the acceptance function of Eq. 2.16 is further fitted to the simulated data where the efficiency moments are allowed to float with Gaussian constraints of one standard deviation and properly taking into account their correlations by means of a multivariate Gaussian constrain. The fitting PDF is shown in Eq. 2.17 and it consists of the PDF that the simulated data was

generated with multiplied with the efficiency function (2.16) times the multivariate Gaussian constrain. The generated parameters are fixed so that only the efficiency moments are allowed to float. The multivariate Gaussian function $G(\vec{c})$, used to constrain the efficiency moments within one standard deviation is shown in Eq. 2.17.

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

$$\text{with } G(\vec{c}) = \frac{1}{\sqrt{(2\pi)^k |C|}} \, e^{-\frac{1}{2}(\vec{c} - <\mathbf{c}>)^{\text{T}} C^{-1} (\vec{c} - <\mathbf{c}>)}, \quad (2.17)$$

where $\vec{c}$ represents the floating efficiency moments, whereas $< \vec{c} >$ are the efficiency moments computed with Eq. 2.15. 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. 2.16.
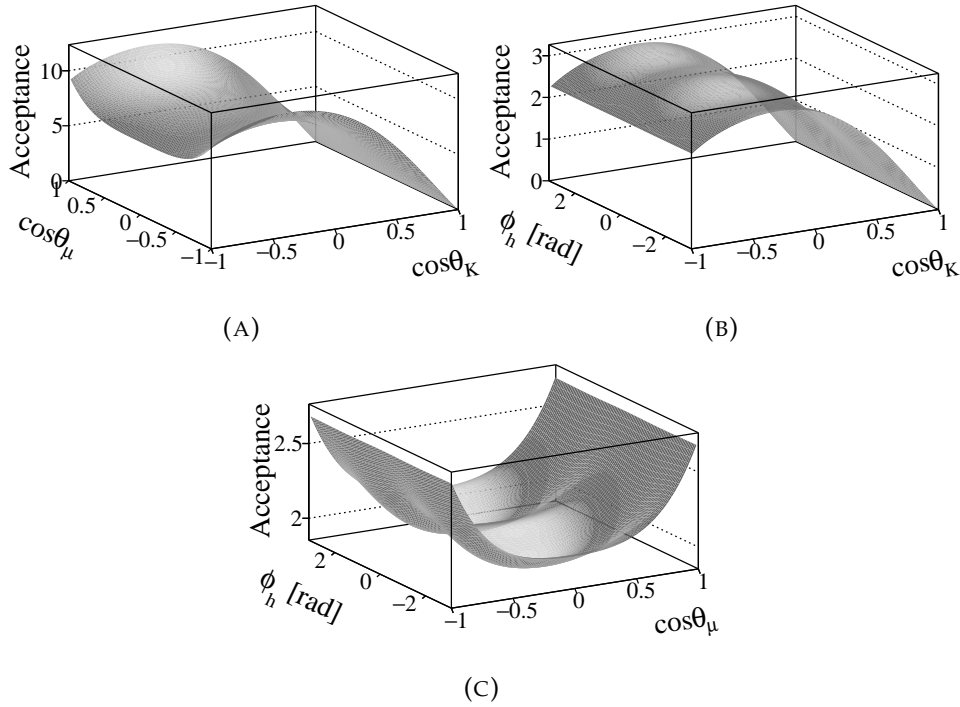
FIGURE 2.7: Two dimensional efficiency projections of the $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ simulation sample, after the constrain of Eq. 2.13. The equivalent plots for $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$ are indistinguishable by eye. The scale of the y-axis is arbitrary.

Alternative parameterization

As it was 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 2.2 as a basis to project the efficiency of Eq. 2.9, instead of the orthogonal basis a_i^{jk} of Eq. 2.8. Looking at Eq. 2.18 it is helpful to realize how the normalization weights, ξ_n, are introduced into the fitting PDF. One can see that they

effectively do not multiply the angular PDF directly but modify the relative contributions between the angular functions f_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.

$$\mathcal{P} \propto \frac{\sum_n f_k(\Omega)}{\sum_n \int d\Omega \; \epsilon(\Omega) \; f_k(\Omega)}, \quad \text{where} \int d\Omega \; \epsilon(\Omega) \; f_k(\Omega) \equiv \xi_n \qquad (2.18)$$

The advantages of this method is that it is immune from any issue with negative efficiency values. After all the acceptance shape is not parametrized at all. On the other hand the fitting time increases due to the fact that the PDF is not expressed in an orthogonal basis any more. Also other implementation and plotting issues arise when using normalization weights. Lastly the two methods can give 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 2.2 and the efficiency function in Eq. 2.8, then the mapping of $c^i_{jk} \to \xi_n$ follows naturally by exploiting the orthogonality of P and Y. More details on the normalization weights approach can be found in [19].

2.2.3 Acceptance Corrections

As it was mentioned in section Section 2.2.2 the angular acceptance is determined with simulated $B^0_s \to J/\psi \overline{K}^{*0}$ events. Thus the acceptance determination can be trusted only if the simulation accurately describes the detector as well as the physical amplitudes of the $B^0_s \to J/\psi \overline{K}^{*0}$ decay. Specifically real and simulated data differences[15] in the momentum distributions of the final state particles may reflect a different acceptance in data compared to MC. On the other hand, non perfect acceptance simulation could be also attributed to differences in the underlying physical parameter values generating the simulated data and those present on real data. For example, the presence of S-wave in the real 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 defined from the momenta of the final state particles.

In order to over-come the problem of simulation imperfections, the simulated sample is weighted using an iterative procedure. At each step of the weighting procedure the simulated data are corrected for the current best estimate of the physics parameters and the two dimensional $p_{K^\pm}, p_{\pi^\mp}$ momentum distribution as well. The efficiency moments are re-evaluated using the weighted simulated sample and the fit to the real data is repeated. This changes the best estimate of the physical parameters of interest. This procedure is repeated until the parameters of interest converge. The full procedure by steps can be summarized as:

1. Calculate an initial set of efficiency moments using uncorrected $B^0_s \to J/\psi \overline{K}^{*0}$ simulated data.

---

[15] Final state momenta distributions were found to differ at a few percent level between real and simulated data.

2. Perform a fit on the $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ data using the initial efficiency moments, and obtain the first estimate of the physical parameters.

3. Weight each event in the simulated sample for the difference between the obtained from the previous fit angular PDF and the angular PDF used to generate the simulated data sample with. This essentially means that the underlying physics of the re-weighted simulated sample corresponds to the physics obtained in the previous step.

4. Compare the $p_{K\pm}, p_{\pi\mp}$ momentum distribution between the physics weighted simulated sample of the previous step with the real data, and weight again each simulated event for the difference.

5. Re-estimate the efficiency moments using the physics plus momentum corrected $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ simulated sample and repeat the fit to $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ real data.

6. Go back to step 4 and repeat until change in the parameters of interest are negligible ($< 0.01\sigma$).

Iteration	1	2	3	4	5	6
$\mathcal{A}_0^{\mathrm{CP}}$	-0.0	-0.03	+0.03	+0.0	-0.0	-0.0
$\mathcal{A}_S^{\mathrm{CP}}$	-0.01	-0.17	+0.09	+0.06	+0.03	+0.01
$\mathcal{A}_\parallel^{\mathrm{CP}}$	+0.13	+0.04	-0.09	-0.04	-0.02	-0.01
$\mathcal{A}_\perp^{\mathrm{CP}}$	-0.19	+0.15	+0.02	+0.01	+0.0	+0.0
A_0	-1.15	+0.96	+0.19	+0.0	-0.02	-0.01
$A_\parallel$	-0.05	-0.06	+0.03	+0.04	+0.02	+0.01
$\delta_\parallel - \delta_0$	-0.37	+0.23	+0.09	+0.03	+0.01	+0.0
$\delta_\perp - \delta_0$	-0.57	+0.28	+0.16	+0.07	+0.03	+0.01
f_S^1	+0.86	-0.66	-0.32	-0.12	-0.05	-0.02
f_S^2	+0.84	-0.27	-0.18	-0.09	-0.05	-0.02
f_S^3	+0.0	-0.02	-0.0	+0.01	+0.02	+0.01
f_S^4	+0.47	-0.21	-0.14	-0.06	-0.04	-0.02
δ_S^1	-2.08	+0.9	+0.66	+0.32	+0.13	+0.04
δ_S^2	+0.13	-0.05	-0.02	-0.03	-0.01	-0.01
δ_S^3	-0.9	+0.48	+0.19	+0.08	+0.04	+0.03
δ_S^4	-2.43	+1.1	+0.53	+0.24	+0.13	+0.07

TABLE 2.4: Parameters of interest central values differences after each iteration. The differences are shown in units of one standard deviation. Larger changes are observed in the first two iterations. Subsequently the differences get progressively smaller, which implies that the parameters of interest converge.

variables	range	#bins
$p_{K\pm}$	$[0, 140]$ GeV/c^2	10
$p_{\pi\pm}$	$[0, 60]$ GeV/c^2	10

TABLE 2.5: Binning scheme for each of the re weighted variables. Bins have equal width.

The $p_{K\pm}, p_{\pi\mp}$ weighting of the MC sample is done using 2D histograms, with a binning scheme that is shown in Table 2.5. At least 6 iterations are necessary to achieve convergence. Detailed evolution of the parameters of

interest after each iteration are shown in Table 2.4. The central values of the parameters of interest does not change a lot impling that the overall effect of the acceptance corrections is small. Figure 2.8 show (in arbitrary units) the shape of the acceptance beforea nd after corrections. Lastly, the acceptance can be floated in the angular fit to the data in order to assign a systematic uncertainty, see Section 2.4.2. The last makes, in hindsight, the acceptance correction procedure (with the current number of observed evetns) not necessary.

(A)

(B)

(C)
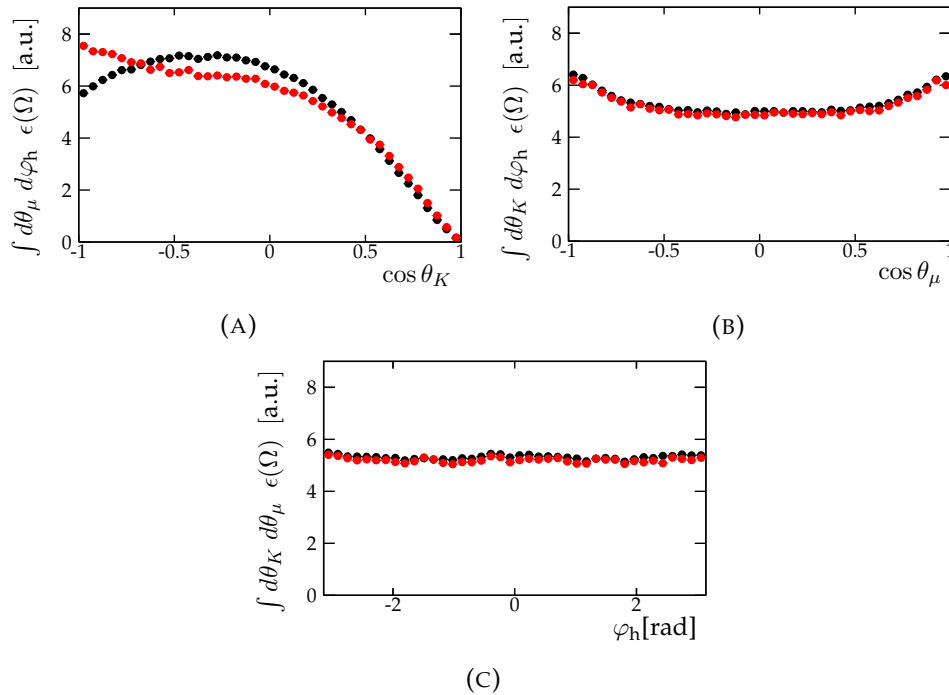
FIGURE 2.8: Angular acceptance shape from simulated data before(black points) and after(red points) acceptance corrections. Changes due to the last are hardly noticable for two of the decay angles. The $\cos\theta_K$ angle is affected more. Since the y-axis scale is aribitrary it is not posible to quantify the amount of change. The plots show cahnges in the acceptance shape only. The plots are made in the same way as in Figure 2.5.

2.2.4 $K\pi$ **Invariant mass**

Dependence on $m_{K\pi}$ and CSP factors

The dependence of the $B^0_{d(s)} \to J/\psi K\pi$ decay amplitude on $m_{K\pi}$ as it was introduced in Eq. 2.7 is treated in a special way. The M_n functions are not implemented but they are integrated over $m_{K\pi}$ instead the fitting real data sample is binned in four parts. where in each part the $m_{K\pi}$ line-shape is assumed to be flat. 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 well known P-wave. In addition the angular structure of the S-wave is not known from first principles. Thus, developing a full model for the both the P-wave and S-wave is not the first priority in this analysis.

Despite the $m_{K\pi}$ integration there is still some residual dependence left. It originates from integrals like the one in equation Eq. 2.19

$$\int dm_{K\pi} \; \Phi \; M_n^* M_m, \qquad (2.19)$$

where $M_{n,m}$ run in the same way as in Eq. 2.7. Φ is a phase space factor[16] and it eventually manifests only in the SP-wave interference terms of Table 2.3 as shown in Eq. 2.20.

$$\frac{\int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \; \Phi \; s^* \times p}{\int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \; \Phi \; |s|^2 \; \int_{m_{K\pi}^-}^{m_{K\pi}^+} dm_{K\pi} \; \Phi |p|^2} = C_{\mathrm{SP}} \; e^{-i\theta_{\mathrm{SP}}} \;, \qquad (2.20)$$

where s and p stand for the S-wave and P-wave line-shapes respectively. Evaluating those complex valued integrals yield the so called C_{SP} factors which are inserted in the angular functions of Table 2.2. Particularly the real part is inserted as a multiplicative factor in front of all SP-wave interference terms[17]. Whereas the imaginary part, is absorbed by the phase difference of the same interference terms[18]. Floating the C_{SP} factors in the fit is not preferred due to high correlations with 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. 2.7. The above mentioned S-wave line-shapes that are necessary to compute the C_{SP} factors are a Isobar combination of a $K^*(892)^0$ and $K_1^*(1410)^0$ for p and a LASS parametrization of a $K_1^*(1430)^0$ with a non resonant term. LASS is a common S-wave parametrization in amplitude analysis, it is a combination of a relativistic Breit-Wigner with angular momentum barrier factors. Whereas Isobar is a superposition of resonances. I was not able to find a public reference for these. I will leave it like this and make sure I can explain what I am saying in this paragraph

sWeighted $m_{K\pi}$ distribution

While the B_s^0 and B^0 muon sWeighted distributions have very similar shape, the $m_{K\pi}$ spectrum exhibit different shapes. Indeed, the B_s^0 $m_{K\pi}$ sPlot seems to be slightly distorted, Figure 2.9a. Since there is no reason why the K^{*0} P-wave should have different shape between $B_s^0 \to J/\psi \overline{K}^{*0}$ and $B^0 \to J/\psi K^{*0}$. This could be due to 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 peaking 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 $m_{K\pi}$ distribution is corrected for angular efficiency effects. This way the interference between the $K\pi$ S-wave and the K^{*0} P-wave vanishes.

---

[16]$\Phi = \left[\lambda(m_\mu, m_K, m_\pi) \lambda(m_{B_s^0}, m_{J/\psi}, m_\mu) \right] / 4m_\mu^2$ and
$\lambda(m_1, m_2, m_3) = \left[m_1^2 - 2m_1^2(m_2^2 + m_3^2) + m_2^2 + m_3^2 - 2m_2^2 m_3^2 \right] / m_1^2$
[17]for example $\Re(A_0^* A_S) \to C_{\mathrm{SP}} * \Re(A_0^* A_S)$
[18] $\arg(A_0^* A_S) \to \arg(A_0^* A_S) - \theta_{\mathrm{SP}}$

This happens because for a perfect acceptance the interference terms in Table 2.7 integrated over all angles evaluate to zero. Which is exactly what is shown in Figure 2.9b. The B_s^0 $m_{K\pi}$ distribution is closer to the one of the B^0 after applying the efficiency correction. This is a clear indication of the presence of stronger interference in the B_s^0 case compared to the B^0 one.
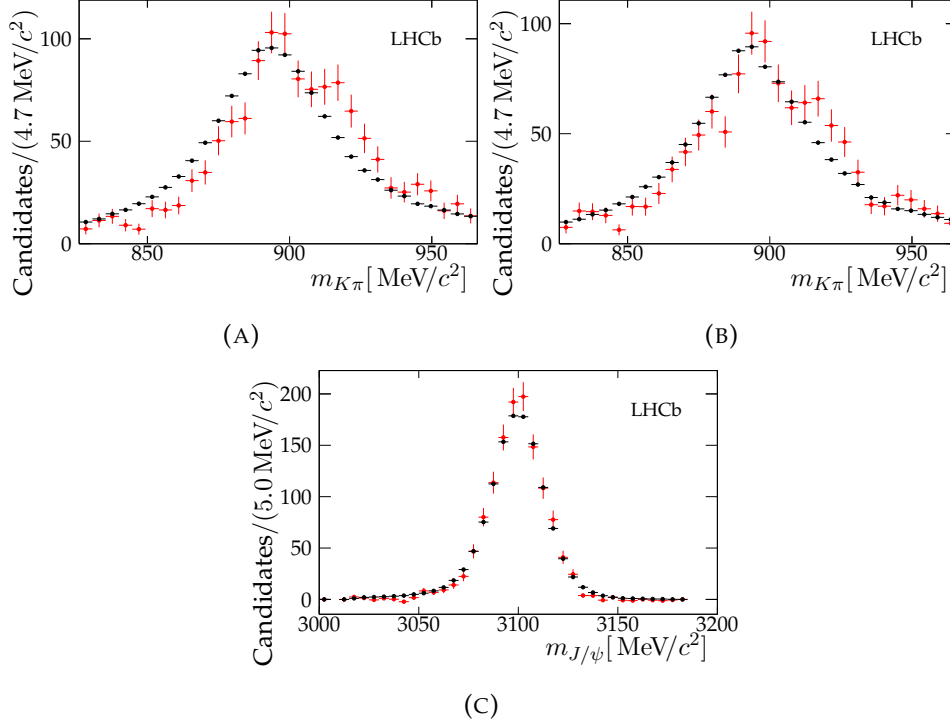

(A)

(B)

(C)

FIGURE 2.9: Comparison of the intermediate state particles mass distributions between B_s^0 and B^0. The plots are background subtracted (sWeighted). Top row shows the $m_{K\pi}$ distribution before (left) and after (right) correcting for angular acceptance effects. The correction is applied by weighting each event with the inverse of the acceptance function as obtained in Eq. 2.16. The J/ψ mass background subtracted distribution is shown at the bottom. It was checked that the acceptance correction has negligible effects on the shapes of the J/ψ mass distributions.

2.2.5 Production and Detection Asymmetries

The nature of the current analysis is such that it is affected by asymmetries related to experimental effects. This is intuitive considering the $\mathcal{A}^{\mathrm{CP}}$ parameter of interest. As explained in Section 1.4 the last quantifies if the decay $B_s^0 \to J/\psi \overline{K}^{*0}$ takes place more often that its charge conjugate. Now in case the collisions at LHC favor one over the other or the detector has different efficiency for positive and negative kaons (or pions) then the $\mathcal{A}^{\mathrm{CP}}$ asymmetries will be diluted from those experimental effects. It happens so that both of those effects are in action. Particularly the LHC is a $p\,p$ colider and not a $p\,\bar{p}$. This implies that at the region of collision there are more b quarks $\bar{b}$ quarks. Thus in such an environment the b quarks are enhanced and hence the $\bar{B}_s^0 \to J/\psi K^{*0}$ decays as well. This effect is quantified by the

production asymmetry $\mathcal{A}^{\mathrm{prod}}$. The other effect mentioned before is called detection asymmetry, $\mathcal{A}^{\mathrm{det}}$, and it becomes relevant at the LHCb detector as described earlier in the current paragraph. Both of those effects need to be estimated so that the observed $\mathcal{A}^{\mathrm{CP}}_{\mathrm{raw}}$ is corrected for as shown in Eq. 2.21.

$$\mathcal{A}^{\mathrm{CP}}(B_s^0 \to J/\psi\overline{K}^{*0}) = \mathcal{A}^{\mathrm{CP}}_{\mathrm{raw}} - \zeta\mathcal{A}^{\mathrm{det}} - \kappa\mathcal{A}^{\mathrm{prod}}, \qquad (2.21)$$

where $\zeta = -1$, and κ account for the dilution due to $B_s^0 - \bar{B}_s^0$ oscillations [3]. The factor κ is shown in Eq. 2.22 and it evaluates to 0.06%, which implies that the $B_s^0 - B_s^0$ oscillations are so fast that nearly all the production asymmetry is washed out by them.

$$\kappa = \frac{\int_0^\infty e^{-\Gamma_s t}\cos(\Delta m_s t)\,\varepsilon(t)\mathrm{d}t}{\int_0^\infty e^{-\Gamma_s t}\cosh\left(\frac{\Delta\Gamma_s}{2}t\right)\varepsilon(t)\mathrm{d}t}, \qquad (2.22)$$

where the quantities Γ_s, $\Delta\Gamma_s$ and Δm_s have been described in Section 1.3 and the decay time efficiency function $\varepsilon(t)$ is estimated on sWeighted $B_s^0 \to J/\psi\overline{K}^{*0}$ data[19]

Using the result from the dedicated LHCb measurement of the production asymmetry [5] and replacing the yield fractions in bins of B_s^0 momentum of Table 3 in the same reference with the sWeighted $B_s^0 \to J/\psi\overline{K}^{*0}$ ones of the current analysis, one gets the production asymmetry estimate shown in Eq. 2.23.

$$\mathcal{A}^{\mathrm{prod}} = (-1.64 \pm 2.28(\mathrm{stat}) \pm 0.55(\mathrm{syst}))\% \qquad (2.23)$$

$$\mathcal{A}^{\mathrm{det}} = (-1.086 \pm 0.531(\mathrm{stat}))\% \qquad (2.24)$$

Coming now to the detection asymmetry it is computed starting from the LHCb study of the detection asymmetry in [4] and weighting the kaon momentum distribution found there to much the one of the $B_s^0 \to J/\psi\overline{K}^{*0}$ decay of the current analysis. The result is shown in Eq. 2.24. Why do we quote only statistical uncertainty

2.2.6　Likelihood fit and Total Decay Rate

Nearly all the necessary ingredients to fit for the parameters of interest are already covered in the previous subsections. However there are a few important details still pending for description. The current subsection addresses the implementations arising from the necessary splitting of the data in parts and the impact of this in the adopted fitting strategy which is a simultaneous maximum likelihood fit. Lastly but most importantly the issue of defining the $\mathcal{A}^{CP}$ asymmetries as free parameters in the angular fit and how this becomes relevant to the normalization of the PDFs involved is discussed in the last part of the current subsection.

---

[19]and parametrized with a the function $\varepsilon(t) = \frac{[1+\beta(t-t_0)][a(t-t0)]^n}{1+[a(t-t_0)]^n}$ found in [6]

Simultaneous Likelihood fit

There are several categories that the data are split to for particular reasons. For example the special treatment of the $m_{K\pi}$ dependence, as described in Section 2.2.4 requires that the data are split in four $m_{K\pi}$ bins. This results in defining a corresponding PDF for each data category[20]. Some parameters of interest differ between the categories while some others are shared. However in order to avoid doing four independent fits and 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[21] such the corresponding component PDF is invoked when the total PDF needs to be evaluated. The current section summarizes all the categories that the data are split to and all the implications arising from that. Lastly a brief insight in some aspects of the likelihood parameter estimation method is included.

Starting from the angular PDF terms as shown in Table 2.7. The last cannot describe both $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$. There is a relative minus sign difference in the imaginary terms originating from complex conjugating the decay amplitude of $B_s^0 \to J/\psi \overline{K}^{*0}$. As it was explained in Section 1.4 those decays are flavor specific which means that a B_s^0 meson has always a negative kaons in the final state particles. The situations is inverted in the case of a $\bar{B}_s^0$ meson. Thus the data can easily be split into two parts based on the sign of one of the hadrons[22]. Furthermore, as it can be seen in Table B.3 the agreement between the $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$ samples is not too bad and could be averaged into one set of efficiency moments. However, the angular acceptance as described in Section 2.2.2 is calculated separately for the those two samples. This is done to avoid involving the $\mathcal{A}_i^{\text{CP}}$ parameters of interest in the efficiency moments computation[23] and thus significantly simplifying the above computation.

	Bin 0	Bin 1	Bin 2	Bin 3
C_{SP}	0.9681	0.9312	0.9519	0.9880

TABLE 2.6: C_{SP} values for each $m_{K\pi}$ bin.

The next category that the data are split to was already mentioned in Section 2.2.4 that dealt with the $m_{K\pi}$ dependence. There are four $m_{K\pi}$ bins defining four more categories in the data. In addition and contrary to the above mentioned data category, there are parameters of interest that are different in each $m_{K\pi}$ bin. Namely the S-wave fraction f_S and phase δ_S. In addition the C_{SP} factors of Section 2.2.4 need to be calculated and inserted separately for each component PDF resulting in Table 2.6, where the C_{SP} factors are shown for each $m_{K\pi}$ bin. The acceptance is also split in the $m_{K\pi}$ category since it was found to vary significantly different in each $m_{K\pi}$ bin, see Table B.4 and Table B.5. The chosen $m_{K\pi}$ binning is shown in Table 2.7.

---

[20]also referred to as component PDF

[21]also referred to as slice

[22]In the current analysis the sign of the K is used to distinguish between the $B_s^0 \to J/\psi \overline{K}^{*0}$ (negative K) and $\bar{B}_s^0 \to J/\psi K^{*0}$ (positive K)

[23]The $\mathcal{A}_i^{\text{CP}}$ parameters are used to average the $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$ decay amplitudes in one parameter, more details in Section 2.2.6.

Bin 0	Bin 1	Bin 2	Bin 3
$[826, 861)$	$[861, 896)$	$[896, 931)$	$[931, 966)$

TABLE 2.7: Definitions of $m_{K\pi}$ bins. Units are in MeV/c^2.

The data categories as described in the above paragraphs result an a simultaneous fit of 8 categories. A single angular acceptance is computed form the combined 2011 and 2012 data samples on grounds of tables Table B.1 and Table B.1.

Having addressed the simultaneous feature of the fit it is interesting to have 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 shown in Eq. 2.25. By definition the likelihood function is the product of probabilities of every event according to the given PDF. Note that opposite to the PDF the likelihood is a function of the free parameters given all the events 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.

$$L(\vec{p}; \vec{x}) = \prod_{e}^{n} \mathcal{P}(\vec{x}; \vec{p}). \tag{2.25}$$

Whereas the inverse of the covariance matrix is given by the negative second derivative of the likelihood at the maximum as shown in Eq. 2.26.

$$\left(V^{-1}\right)_{ij} = -\left.\frac{\partial^2 \ln L}{\partial p_1 \partial p_2}\right|_{\vec{p}_{\mathbf{max}}}. \tag{2.26}$$

Another typical way to estimate and better visualize the standard deviations is shown in Section 2.4.3.

Coming now to the concept of the extended likelihood. In Eq. 2.25 the number of events is fixed. However it is possible to extend the last equation such that the expected number of events shows up as a parameter in the likelihood function which can be floated. This is useful in cases where the PDF consists of two species, such as signal and background, or if the expected number of events depends on the parameters of interest $\vec{p}$. For the current analysis the later is exploited in order to float the the $\mathcal{A}^{\mathrm{CP}}$ asymmetries in the fit, see end of current subsection. Extending the PDF is achieved by multiplying the likelihood with a poison distribution as shown in Eq. 2.27.

$$L(\vec{p}; \vec{x}) = P(n; \mu(\vec{p})) \prod_{e}^{n} \mathcal{P}(\vec{x}; \vec{p}), \tag{2.27}$$

where $\mu(\vec{p})$ is the expected number of events. As it can be seen it depends on the parameters of interest which in principle reduces the uncertainty on them since additional information is added to the likelihood function. For purely practical reasons one would typically choose to minimize the negative logarithm of the likelihood, which after some simple algebra and using the well known expression of the Poisson distribution results in Eq. 2.28,

$$-\ln L(\vec{p};\vec{x}) = -\ln\left(\frac{\mu(\vec{p})}{n\,!}e^{-\mu(\vec{p})}\right) - \ln\left(\prod_e^n \mathcal{P}(\vec{x};\vec{p})\right)$$

$$= -n\ln(\mu(\vec{p})) + \mu(\vec{p}) - \sum_e^n \ln(\mathcal{P}(\vec{x};\vec{p})). \qquad (2.28)$$

Note that constant terms in the likelihood do not affect either the shape of the likelihood or the position of the minimum and thus neglected.

CP **Asymmetries as Free Parameters**

The *CP* asymmetries discussed in Section 1.4 can be measured from the raw asymmetry observed in data. This is done by counting $\bar{B}_s^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \bar{K}^{*0}$ events and use them as shown Eq. 2.29 defines.

$$\mathcal{A}_{\text{raw}}^{\text{CP}} \equiv \frac{N^+ - N^-}{N^+ + N^-}, \qquad (2.29)$$

where N^+ and N^- stands for the number of observed $\bar{B}_s^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \bar{K}^{*0}$ respectively. Going one step further $\mathcal{A}_{\text{raw}}^{\text{CP}}$ can be implemented as a free parameter in the fit. The advantage is that the entire statistical uncertainty propagation for those parameters is built in the likelihood thus avoiding unnecessary calculations. In addition the above quantities can be measured in a polarization dependent way. In order to vary $\mathcal{A}_{\text{raw}}^{\text{CP}}$ in the fit it is necessary to re-parameterize them in a different way shown in Eq. 2.30.

$$\mathcal{A}_i^{\text{CP}} = \frac{N^+|A_i^+|^2 - N^-|A_i^-|^2}{N^+|A_i^+|^2 + N^-|A_i^-|^2} \text{ and } |A_i|^2 = \frac{N^+|A_i^+|^2 - N^-|A_i^-|^2}{N^+ + N^-}, \quad (2.30)$$

where $|A_i^+|^2$ and $|A_i^-|^2$ stand for the polarized amplitude present in the $\bar{B}_s^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \bar{K}^{*0}$ samples respectively. The above equations can be solved for $|A_i^+|^2$ $|A_i^-|^2$. The idea is that the last parameters will be used as amplitudes for the $\bar{B}_s^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \bar{K}^{*0}$ PDFs respectively. This way both samples are fitted simultaneously in one fit. Solving Eq. 2.30 for $|A_\pm|^2$ yields:

$$|A_i^+|^2 = \xi\,|A_i|^2(1 + \mathcal{A}_i^{\text{CP}}) \text{ and } |A_i^-|^2 = \frac{\xi}{2\xi - 1}|A_i|^2(1 - \mathcal{A}_i^{\text{CP}}). \qquad (2.31)$$

The $|A_i^+|^2$ and $|A_i^-|^2$ are the amplitude expressions that are used along with the angular functions of Table 2.7. Whereas the normalization factors in front of them can be re-expressed as function of $\mathcal{A}_{\text{raw}}^{\text{CP}}$ using Eq. 2.29.

$$\xi = \frac{N^+ + N^-}{2N^+} = \frac{1}{2}\left(1 + \frac{N^-}{N^+}\right) =$$

$$= \frac{1}{2}\left(1 + \frac{1 - \mathcal{A}_{\text{raw}}^{\text{CP}}}{1 + \mathcal{A}_{\text{raw}}^{\text{CP}}}\right) = \frac{1}{1 + \mathcal{A}_{\text{raw}}^{\text{CP}}} \qquad (2.32)$$

$$\frac{\xi}{2\xi - 1} = \frac{1}{2 - 1/\xi} = \frac{1}{1 - \sum_i \mathcal{A}_{\text{raw}}^{\text{CP}}} \qquad (2.33)$$

The quantities that are actually fitted are the $|A_i|^2$ with $i : \{0, \|, \perp, S\}$. Having said that, and by looking at Eq. 2.31 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 $i : 0, \|, \perp, S$) 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 which is shown in the following equation.

$$|A_S|^2 = \mathsf{f_S} \quad \text{and} \quad |A_k|^2 = (1 - \mathsf{f_S})\mathsf{f}_k. \qquad (2.34)$$

Note that the P-wave fractions are also constrained so that they add up to one by imposing $\sum \mathsf{f}_k = 1$ with $k : \{0, \|, \perp\}$. The fractions f_k and $\mathsf{f_S}$ are the actual parameters that one wish to measure and correspond to those presented in Section 1.4. The final expressions of the $|A_i^\pm|^2$ after all the above is given in Eq. 2.35

$$|A_k^\pm|^2 = \frac{1 \pm \mathcal{A}_k^{\text{CP}}}{1 \pm \sum_i \mathcal{A}_i^{\text{CP}}}(1 - \mathsf{f_S})\mathsf{f}_k \quad \text{and} \quad |A_S^\pm|^2 = \frac{1 \pm \mathcal{A}_S^{\text{CP}}}{1 \pm \sum_i \mathcal{A}_i^{\text{CP}}}\mathsf{f_S}, \qquad (2.35)$$

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

There is one more step required to properly build the full 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 events it would be wrong since the observed number of events are diluted by production and detection asymmetry. Given that the parameters of interest, $\mathcal{A}^{CP}$, quantify the difference between the observed $\bar{B}_s^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \overline{K}^{*0}$ decays, it follows that the observed number of those decays need to be corrected for. With input from Section 2.2.5 where the production and detection asymmetries were estimated, the observed yields are made dependent on the $\mathcal{A}^{CP}$ asymmetry in Eq. 2.36

$$N_{CP}^+ = \frac{N_{\text{raw}}^{\text{total}}}{2}\left(\mathsf{f_S}\,\mathcal{A}_S^{\text{CP}} + \sum_k^{0, \|, \perp}(1 - \mathsf{f_S})\mathsf{f}_k\mathcal{A}_k^{\text{CP}}\right)$$

$$N_{CP}^- = N_{\text{raw}}^{\text{total}} - N_{CP}^+ \qquad (2.36)$$

and then subsequently corrected for the production and detection asymmetry in Eq. 2.37.

$$N^\pm = N_{CP}^\pm\left(1 \pm \mathcal{A}^{\text{det}} \mp \mathcal{A}^{\text{prod}}\right). \qquad (2.37)$$

The above parametric yields are the ones that the PDFs will be normalized to and enter as the expected number of events in the extended likelihood term of the PDFs as discussed earlier in the current subsection.

2.3 Normalization of $B_s^0 \rightarrow J/\psi K\pi$

The current section addresses the issue of the $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ branching fraction estimation, hereafter $\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})$, which is one of the necessary ingredients for estimating the penguin contributions to ϕ_s, addressed in Section 3.1. Absolute branching fraction measurements are more complicated to perform compared to relative ones. The $\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})$ 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 extracted.

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

Equation 2.38 shows how $\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})$ is normalized with respect to $\text{Br}(B_s^0 \rightarrow J/\psi \phi)$, hereafter $\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})_\phi$. It requires as input the observed number of $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ events. The first one is taken from Eq. 2.5 as for the later a bit more effort is needed. 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 $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ in Table A.1.

$$\frac{\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})}{\text{Br}(B_s^0 \rightarrow J/\psi \phi)} = \frac{N_{B_s^0 \rightarrow J/\psi K\pi}}{N_{B_s^0 \rightarrow J/\psi K^+ K^-}} \times \frac{\varepsilon^{MC}_{B_s^0 \rightarrow J/\psi \phi}}{\varepsilon^{MC}_{B_s^0 \rightarrow J/\psi \overline{K}^{*0}}} \times$$
$$\times \frac{\omega_{B_s^0 \rightarrow J/\psi \phi}}{\omega_{B_s^0 \rightarrow J/\psi \overline{K}^{*0}}} \times \frac{\text{Br}(\phi \rightarrow K^+ K^-)}{\text{Br}(K^* \rightarrow K\pi)}.$$
$$(2.38)$$

Also the BDTG classifier is applied to the reference channel with the same cut values as in the $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$ channel. The ratio of observed number of events is represented by the first fraction in Eq. 2.38, Whereas the ratio of total efficiencies between the two channels by the second. The last is corrected for differences in efficiencies between real and simulated data. Additionally there is a correction factor to account for the presence of S-wave in the two channels[24].

$$\text{Br}(B_s^0 \rightarrow J/\psi \overline{K}^{*0})_\phi = (4.25 \pm 0.20 \pm 0.13 \pm 0.36) \times 10^{-5}, \qquad (2.39)$$

---

[24]The penguin estimation assumes no S-wave contribution to $B_s^0 \rightarrow J/\psi \overline{K}^{*0}$. Each of the factors ω in Eq. 2.38 correspond to the ratio of two angular PDFs. One with the S-wave parameters set to zero and P-wave to those obtained in the angular fit, whereas the other with both P-wave and S-wave parameters as obtained from the same fit. The ω factors can be understood as an overall S-wave fraction which corrects the observed number of events.

The branching fractions of the intermediate resonance decaying to a di-hadron system are obtained from [22] [25]. The $\mathrm{Br}(B_s^0 \to J/\psi\phi)$ [1] is updated with the latest b quark hadronization fraction[26] [24]. Whereas the efficiency and the angular correction ratios are about 2 and 0.9 respectively. All the above information is used to extract $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})_\phi$ shown in Eq. 2.39, where the uncertainties are statistical systematic and due to $\mathrm{Br}(B_s^0 \to J/\psi\phi)$ respectively.

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

Equation 2.40 shows how $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})$ is normalized with respect to $\mathrm{Br}(B^0 \to J/\psi K^{*0})$, hereafter $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})_d$.

$$\frac{\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})}{\mathrm{Br}(B^0 \to J/\psi K^{*0})} = \frac{N_{B_s^0 \to J/\psi K\pi}}{N_{B^0 \to J/\psi K^{*0}}} \times \frac{f_d}{f_s} \times \frac{\varepsilon_{B^0 \to J/\psi K^{*0}}^{MC}}{\varepsilon_{B_s^0 \to J/\psi\overline{K}^{*0}}^{MC}} \times \frac{\omega_{B^0 \to J/\psi K^{*0}}}{\omega_{B_s^0 \to J/\psi\overline{K}^{*0}}}, \tag{2.40}$$

The last is similar to Eq. 2.38 as far as the way efficiency and angular correlation ratios are obtained. The observed number of events are both obtained from Eq. 2.5 and the latest measurement of the hadronization fraction from [24]. The $\mathrm{Br}(B^0 \to J/\psi K^{*0})$ value is taken from [9]. The final result for $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})_d$ is shown in Eq. 2.41

$$\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})_d = (3.85 \pm 0.18 \pm 0.15 \pm 0.22 \pm 0.42) \times 10^{-5}, \tag{2.41}$$

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

Averaged $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})$

Both estimates of Eq. 2.39 and Eq. 2.41 are compatible within uncorrelated systematics. A weighted average of these two estimates is performed where correlations are taken into account. Correlations between $\omega_{B^0 \to J/\psi K^{*0}}$ and $\omega_{B_s^0 \to J/\psi\overline{K}^{*0}}$ originate from the expected number of events estimated with the mass fit in Section 2.1. In addition the efficiency ratios cannot be treated separately and they have a common factor, namely $\varepsilon_{B_s^0 \to J/\psi\overline{K}^{*0}}^{MC}$. The averaged $\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})$ is shown in Eq. 2.42

$$\mathrm{Br}(B_s^0 \to J/\psi\overline{K}^{*0})_d = (4.13 \pm 0.16 \pm 0.25 \pm 0.24 \pm 0.42) \times 10^{-5}, \tag{2.42}$$

The last is in agreement with previous measurements [2]

---

[25] $\mathrm{Br}(K^* \to K\pi) = 2/3$ whereas $\mathrm{Br}(\phi \to K^+K^-) \simeq 1/2$

[26]The quantity $\frac{f_d}{f_s}$ expresses the ratio of probabilities that a b quark will form a meson with a d quark over an s quark

2.4 Results

2.4.1 Parameters of Interest

After selecting the data as described in Section 2.1 the angular model of Section 2.2 is fitted to the data. The results of this fit are presented in the current section. Table 2.8 summarizes the parameters of interest best fit values whereas Figure 2.10 shows the a plot of the fitted PDF to the $B_s^0 \to J/\psi \overline{K}^{*0}$ data. The correlation matrix can be found in Table B.7.

(A)

(B)

(C)
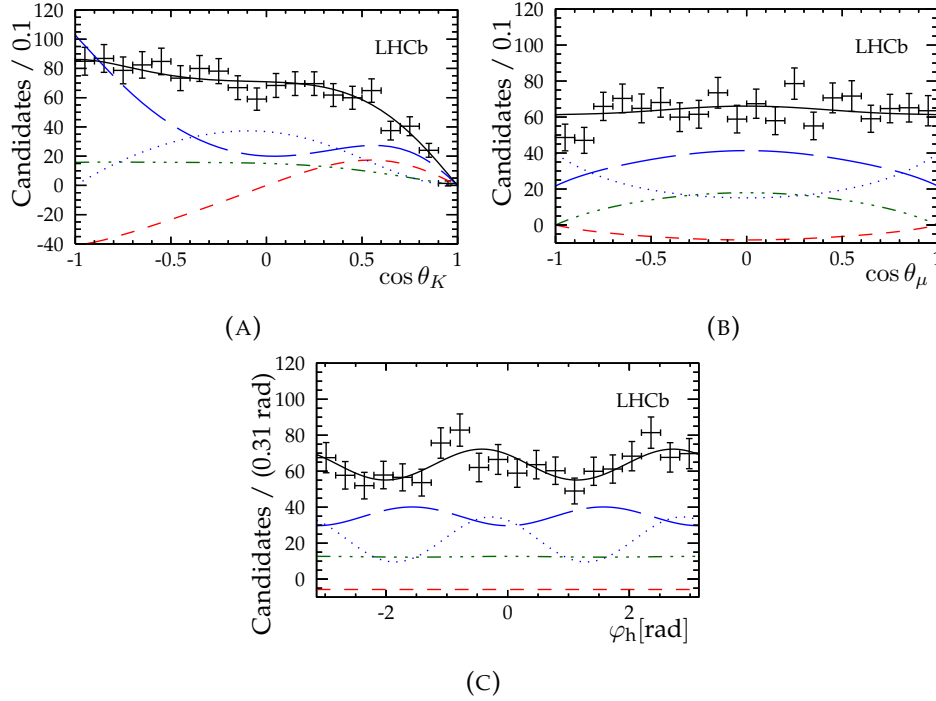
FIGURE 2.10: Fitted angular PDF projections on top of the $B_s^0 \to J/\psi \overline{K}^{*0}$ sWeighted data. Black solid: total, blue dashed P-wave even, blue dotted P-wave odd plus P-wave self interference, green dotted: S-wave, red dotted dashed S-P interference. Notice the negative values of the SP-wave interference. Pull distributions are shown as well.

2.4.2 Systematic Uncertainties

The different source of systematic uncertainties on the B_s^0 angular parameters, as well as on the branching fraction are presented in the current section.

Angular acceptance

The efficiency moments as described in Section 2.2.2 are computed with a limited size simulation sample and are 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 with a multivariate Gaussian constrain within 5 standard deviations from their nominal

parameter	value	statistical uncertainty
$\mathcal{A}_0^{\mathrm{CP}}$	-0.063	$(+0.063, -0.063)$
$\mathcal{A}_\parallel^{\mathrm{CP}}$	$+0.155$	$(+0.159, -0.157)$
$\mathcal{A}_\perp^{\mathrm{CP}}$	-0.087	$(+0.097, -0.097)$
$\mathcal{A}_S^{\mathrm{CP}}$	$+0.129$	$(+0.098, -0.097)$
f_0	$+0.481$	$(+0.024, -0.025)$
$f_\parallel$	$+0.183$	$(+0.028, -0.027)$
$\delta_\parallel - \delta_0$	-2.766	$(+0.163, -0.166)$
$\delta_\perp - \delta_0$	-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	$+0.578$	$(+0.111, -0.120)$
δ_S^1	$+0.265$	$(+0.136, -0.131)$
δ_S^2	-0.532	$(+0.229, -0.195)$
δ_S^3	-1.682	$(+0.189, -0.240)$
δ_S^4	-2.070	$(+0.142, -0.160)$

TABLE 2.8: Best fit values of the parameters of interest.

value. The multivariate Gaussian takes properly into account their correlations. In addition this systematic covers any potential difference in the acceptance between data and simulated sample that has not been corrected by the iterative procedure described in Section 2.2.3. Furthermore, the constrain of Eq. 2.13 used to force the acceptance positive assumes that the acceptance is zero at $\cos\theta_K = 1$ for any value of $(\cos\theta_\mu, \varphi_{\mathrm{h}})$. This assumption might not be exactly true. It could be that the value of the acceptance at that point is very small but not zero. This systematic effect is accounted for by adding a term[27] to Eq. 2.16 and allowing it to float in the above mentioned fit to the data. The differences in the parameters of interest between the best fit of Table 2.8 and the fit with the floating acceptance is taken as a systematic uncertainty.

	$\mathcal{A}_0^{\mathrm{CP}}$	$\mathcal{A}_\parallel^{\mathrm{CP}}$	$\mathcal{A}_\perp^{\mathrm{CP}}$	$\mathcal{A}_S^{\mathrm{CP}}$
Mass parameters and B^0 contamination	0.001	0.001	–	0.001
D-wave contribution	0.002	0.015	0.002	–
Mass—$\cos\theta_\mu$ correlations	0.014	$+0.009 \atop -0.012$	0.016	$+0.023 \atop -0.029$
Angular acceptance	0.003	0.003	0.002	0.001
Peaking background	$+0.003 \atop -0.004$	$+0.012 \atop -0.004$	0.002	0.001
C_{SP} factors	0.001	0.002	0.002	–
Fit bias	0.0139	0.003	0.012	0.038
Total uncertainties	0.31	$+0.14 \atop -0.13$	0.21	$+0.45 \atop -0.49$

TABLE 2.9: Systematic uncertainties for the $\mathcal{A}^{\mathrm{CP}}$ parameters of interest. The total uncertainties is the quadratic sum of all the individual systematic uncertainties expressed in percentage of the statistical uncertainty.

---

[27] with the appropriate $\frac{1}{2\sqrt{\pi}}$ normalization

$m(J/\psi K^+ \pi^-)$ **mass model**

Not all the PDF parameters are free to vary in the mass fit to the data as described in Section 2.1. These parameters are fixed to values taken from simulated events. In order to account for systematic uncertainties due to these fixed parameters, the mass fit is repeated 1000 times. In each fit to the data the values of those fixed parameters are taken from a Gaussian distribution. The mean and width of the Gaussian are taken from a mass fit to the simulated data (this last fit is the one that those fixed parameters were estimated in the first place.) Correlations among those 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 like distribution for each of the angular parameters of interest. The one standard deviation interval of this distribution is assigned as a systematic to that parameter.

In addition, another systematic to account for event leakage from the B^0 to the B_s^0 peak is assigned. As mentioned in Section 2.1 α 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 α of the B^0 peak is further scrutinized. Particularly the mass fit to the data is repeated with α set to infinity and α floating. The first case, according to the Hypatia function, implies that any contribution to the tails other than resolution is ignored. Those 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 those two case is assigned as a systematic.red Confirm this with Simon. ANA note is not clear enough.

	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
Peaking background	$+0.004$ -0.003	0.002	0.020	0.010
C_{SP} factors	—	0.001	—	—
Fit bias	0.001	—	0.007	0.016
Total uncertainties	$+0.40$ -0.38	$+0.26$ -0.27	$+0.45$ -0.44	$+0.24$ -0.38

TABLE 2.10: Systematic uncertainties for the P-wave parameters of interest. The total uncertainties is the quadratic sum of all the individual systematic uncertainties expressed in percentage of the statistical uncertainty.

Peaking backgrounds

Peaking background estimates have an impact on the mass fit and eventually on the computed set of *sWeights*. Those estimates are summarized in Table 2.1, where one can see that they come with a statistical uncertainty as well. To account for the last uncertainty the estimates of all peaking background species ($B_s^0 \to J/\psi K^+ K^-$, $B_s^0 \to J/\psi \pi^+ \pi^-$, $B_d^0 \to J/\psi \pi^+ \pi^-$, and

$\Lambda_b^0 \to J/\psi pK^-$) are shifted one at a time by 1 standard deviation. The result is a different set of *sWeights* for each specie. Subsequently the angular fit is performed with that set of *sWeights*. The difference in the angular parameters between the last fit and the nominal angular fit is taken as a systematic uncertainty for the particular peaking background specie. Individual contributions from all species to a certain angular parameter are added in quadrature and assigned as a systematic.

Correlations between mass and angles in the *sPlot* context

Correlations between the B_s^0 invariant mass and the $\cos\theta_\mu$ angular distribution is taken into account by performing the mass fit in 5 bins of $\cos\theta_\mu$. Since the choice of 5 bins is arbitrary the mass fit is repeated in 3 and 6 $\cos\theta_\mu$ bins 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 last quadratic sum is assigned as systematic.

Fit bias

Potential fit biases of the fitting procedure are addressed in Section 2.4.4 and the corresponding results are reported in Table 2.14. The reported biases on the parameters of interest are assigned as a systematic uncertainty.

	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	0.007	0.015
Angular acceptance	0.000	0.000	0.000	0.003
Peaking 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
Total uncertainties	$+0.66$ -0.43	$+0.18$ -0.23	$+0.27$ -0.41	$+0.14$ -0.13

TABLE 2.11: Systematic uncertainties for the S-wave fractions. The total uncertainties is the quadratic sum of all the individual systematic uncertainties expressed in percentage of the statistical uncertainty.

C_{SP} factors

The choice of S-wave and P-wave models affect the values of the C_{SP} factors, which they finally enter the angular fit via the SP-wave interference terms of the angular PDF, see Section 2.2.4. The choice of those models has a small but measurable impact on the parameters of interest and it was made such that the chosen models optimally describe the $m_{K\pi}$ distribution on data. A conservative systematic to the parameters of interest is assigned by choosing the worst possible model and repeating the angular fit. The difference in the fitted parameters from the nominal angular fit corresponds to the assigned systematic uncertainty.

Nuisance CP asymmetries

Asymmetries induced in the B_s^0 meson production and from detector imperfections as described in Section 2.2.5, directly affect the $\mathcal{A}_i^{\mathrm{CP}}$ parameters of interest, see Section 2.2.6. The estimated production and detection asymmetries are allowed to float 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. Thus no separate systematic is assigned.

	δ_S^1	δ_S^2	δ_S^3	δ_S^4
Mass parameters and B^0 contamination	$^{+0.000}_{-0.010}$	$-$	$-$	$-$
D-wave contribution	0.020	0.030	0.080	0.040
Mass—$\cos\theta_\mu$ correlations	0.050	0.020	0.032	0.010
Angular acceptance	0.003	0.004	0.036	0.005
Peaking background	0.010	0.010	$^{+0.030}_{-0.020}$	$^{+0.070}_{-0.040}$
C_{SP} factors	0.010	$-$	$-$	0.001
Fit bias	0.063	0.009	0.024	0.013
Total uncertainties	$^{+0.62}_{-0.65}$	$^{+0.17}_{-0.20}$	$^{+0.53}_{-0.41}$	$^{+0.58}_{-0.37}$

TABLE 2.12: Systematic uncertainties for the S-wave phases. The total uncertainties is the quadratic sum of all the individual systematic uncertainties expressed in percentage of the statistical uncertainty.

D-wave contribution

Any D-wave contribution to the considered $m_{K\pi}$ window is neglected in the current analysis. However, in the official LHCb result[] fix the stupid citation entry of the same analysis D-wave contributions are taken into account. Thus the systematic uncertainty quoted there is also assigned to the current analysis.

Uncertainties from $\frac{f_d}{f_s}$, $\mathrm{Br}(B^0 \to J/\psi K^{*0})$, and $\mathrm{Br}(B_s^0 \to J/\psi\phi)$

The uncertainties from the external parameters $\frac{f_d}{f_s}$, $\mathrm{Br}(B^0 \to J/\psi K^{*0})$, and $\mathrm{Br}(B_s^0 \to J/\psi\phi)$ are taken as systematic for this analysis. They are added in

Source	$\frac{\mathrm{Br}(B_s^0 \to J/\psi \bar{K}^{*0})}{\mathrm{Br}(B^0 \to J/\psi K^{*0})}$	$\frac{\mathrm{Br}(B_s^0 \to J/\psi \bar{K}^{*0})}{\mathrm{Br}(B_s^0 \to J/\psi\phi)}$
Statistical	0.14	0.20
$\frac{f_d}{f_s}$	0.17	0
Efficiency ratio	0.039	0.054
Peaking backgrounds effect on the yield	0.018	0.025
Angular corrections (syst)	0.092	0.070
Mass fit model	0.057	0.078
Fit bias	0.0050	0.0068
Quadratic sum	0.18 (0.25 w/ $\frac{f_d}{f_s}$)	0.23

TABLE 2.13: Systematic uncertainties for the branching fraction measurements. Numbers are in units of 10^{-2}.

quadrature to the estimation of $\mathrm{Br}(B_s^0 \to J/\psi \overline{K}^{*0})$.

Summary

All the above systematic uncertainties related to the angular parameters are summarized from Table 2.9 to Table 2.12. Two sources of systematics are found to be dominant. First the size of the Monte Carlo sample used to estimate the normalization weights of the angular acceptance; the correlation between the B_s^0 invariant mass and the cosine of the helicity angle θ_μ. 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 2.13.
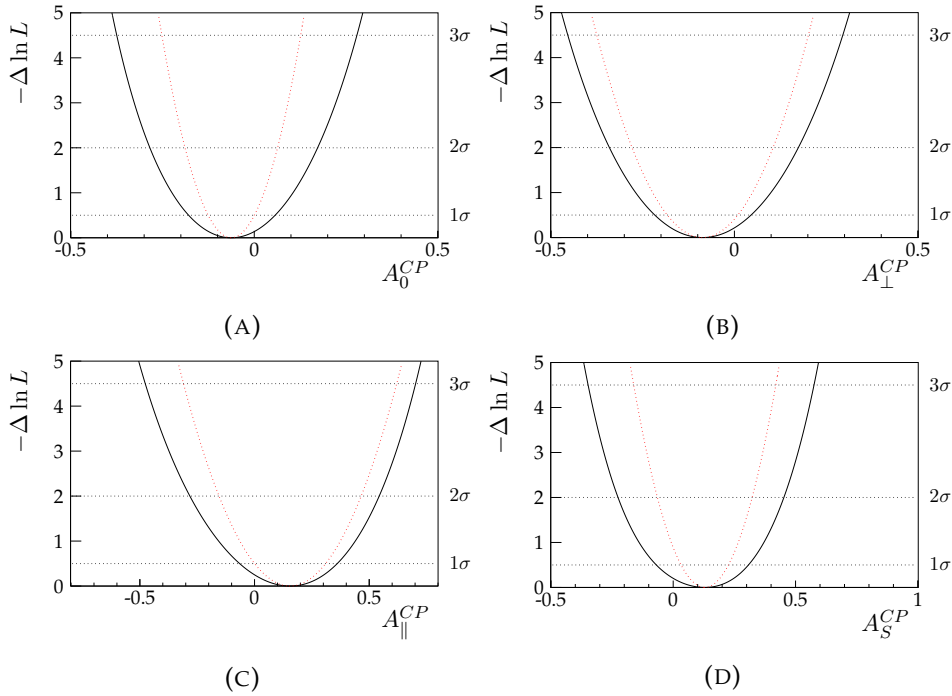
2.4.3 Likelihood scans

FIGURE 2.11: Profile likelihood scans of $\mathcal{A}^{\mathrm{CP}}$ parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a parabola, with a symmetric Gaussian error. No peculiar structure is observed apart from the non Gaussian nature of the parameters

The likelihood function presented in Section 2.2.6 is revisited in the current subsection. This time it is visualized in order to gain more confidence that the fit result presented in Section 2.4.1 comes indeed from a valid minimum of the likelihood function. One can easily get a good estimate on the shape of the function by expanding it as a Taylor series around the minimum [28]. Assuming that the data sample $\vec{x}$ in Eq. 2.25 is large enough and

---

[28]The constant term of the expansion is irrelevant to the position of the minimum as it moves the value of the likelihood function vertically. Whereas the first derivative is zero by definition.

also that the variables $\vec{p}$ in the same equation are Gaussian in nature, then one can easily find out that the Taylor expansion of the negative logarithm of Eq. 2.25 projected over one parameter μ out of $\vec{p}$ has a parabolic shape as shown in Eq. 2.43

$$\ln L(\mu; \vec{x}) \simeq \frac{1}{2} s^2, \quad \text{with} \quad s = \left(\frac{\mu - \hat{\mu}}{\sigma_\mu} \right), \tag{2.43}$$

where $\hat{\mu}$ is the value of μ at the minimum and σ_μ is the statistical uncertainty associated to μ. The last equation is called *profile likelihood*.
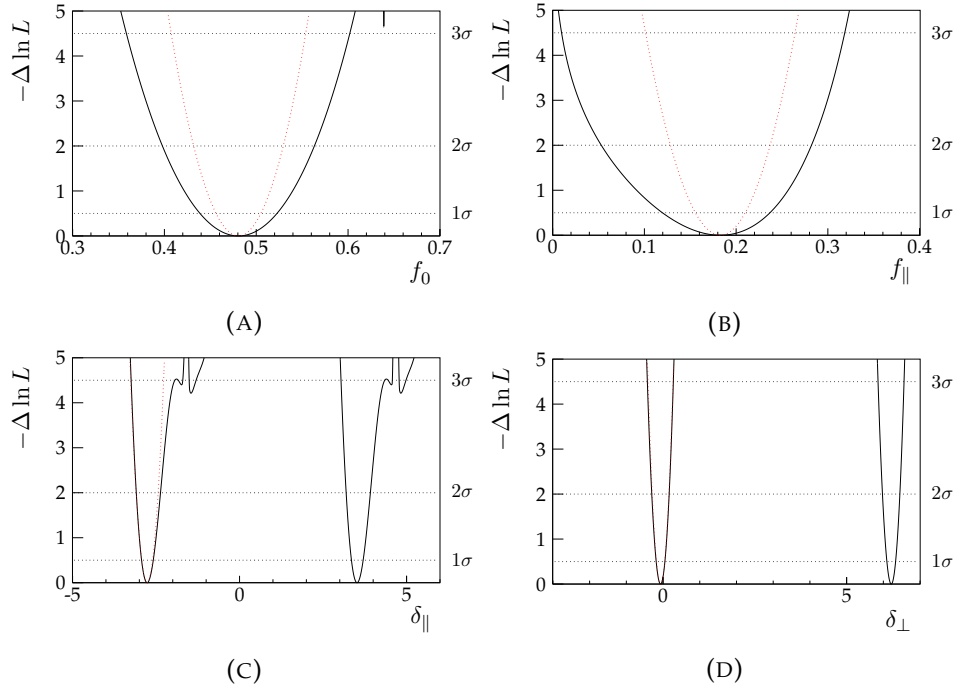
(A)

(B)

(C)

(D)

FIGURE 2.12: 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 parabola, with a symmetric Gaussian error. No peculiar structure is observed for the P-wave fractions apart from the non Gaussian nature of the parameters. Whereas the phases clearly show the existence of the double minimum related to the symmetry of the fitted PDF explained in Section 1.4 Do not forget to mention this. Also $\delta_\parallel$ appears to have other local minima at above 3σ and as such are neglected, since only global minima are relevant to the likelihood parameter estimation.

Based on the above mentioned assumptions two important conclusions can be drawn given a profile likelihood plot. First the non Gaussian nature of a parameter of interest can be visualized by how much the corresponding profile likelihood deviates from a perfect parabola. Non perfect parabola is mainly an indication of low sample size[29]. Second, Eq. 2.43 serves as an alternative method of estimating the statistical uncertainty of a parameter

---

[29]Also a bad looking profile likelihood with spikes and discontinuities is a very good indication that the PDF used to build the likelihood is problematic. For example a bad mistake in the implementation can cause discontinuities in the profile likelihood plot.

of interest given a profile likelihood plot. Specifically s number of standard deviations are defined from the two points where Eq. 2.43 intersects the profile likelihood function (for the desired value of s) and projecting to the horizontal axis[30].

It follows that the errors defined in the the above mentioned way can be asymmetric left and right of the minimum. The sample size for some of the data categories in the current analysis is quite small. This is depicted in the corresponding profile likelihood plots of the parameters of interest relevant to those data categories. This is also the reason why asymmetric errors are used.

Going one step further on the statistical error interpretation, it is useful to realize what kind of statement is usually attributed to the statistical uncertainty of one standard deviation. Basically the claim is that one standard deviation guaranties that the probability that the true value of the parameter of interest μ^{true} will be inside the interval $\hat{\mu}^{+\sigma_\mu}_{-\sigma_\mu}$ is 68%[31]. The last interval is usually quoted as a 1σ *confidence interval* [32]

2.4.4 Toy Experiments Study

The current subsection addresses the robustness of the fitting angular PDF. This is done by means of a so called *toy experiments* study. This procedure can point out any potential biases that are introduced by the angular fitting model. Those biases are finally assigned as a systematic to the parameters of interest.

In the current toy experiments study the angular PDF is used to generate data-sets of $\Omega = (\cos\theta_K, \cos\theta_\mu, \varphi_{\text{h}})$ 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 *pull* variable of Eq. 2.44.

$$P^i = \frac{p^i - p^{\text{gen}}}{\sigma^i}, \quad \begin{array}{l} \sigma^i = \sigma^i_- \ \text{ if } \ p^i - p^{\text{gen}} > 0 \\ \sigma^i = \sigma^i_+ \ \text{ if } \ p^i - p^{\text{gen}} < 0 \end{array} , \tag{2.44}$$

where, p^{gen} and p^i are the generated and fitted (from each toy experiment) 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 value difference in units of one standard deviation.

In an unbiasing fitting model the pull distribution of a certain parameter follows a Gaussian distribution of zero mean and width one[33]. This can be translated to a useful statement about the parameter, $\hat{p}$, and upper (lower) error, $\hat{\sigma}_+(\hat{\sigma}_-)$ estimates of Section 2.4.1. Particularly it means that the probability that the interval $[\hat{p} - \hat{\sigma}_-, \hat{p} + \hat{\sigma}_+]$ contains p^{true} is 68%[34]. Or

---

[30]Complex as it might seem to be the above become apparent in the profile likelihood plots that follow in the current subsection

[31]The last percentage comes from the fraction of the total area of a normal distribution corresponding to one standard deviation.

[32]and it is in the core of the so called *frequentist statistics* interpretation of the probability.

[33]Provided that the Central Limit Theorem (CLT) of statistics is satisfied.

[34]In this context the term "true" stands for the value of p that nature "chose" and it is unknown. The value of p^{gen} on the other hand is known and for the purposes of the toy experiment is a suitable replacement of the p^{true}.
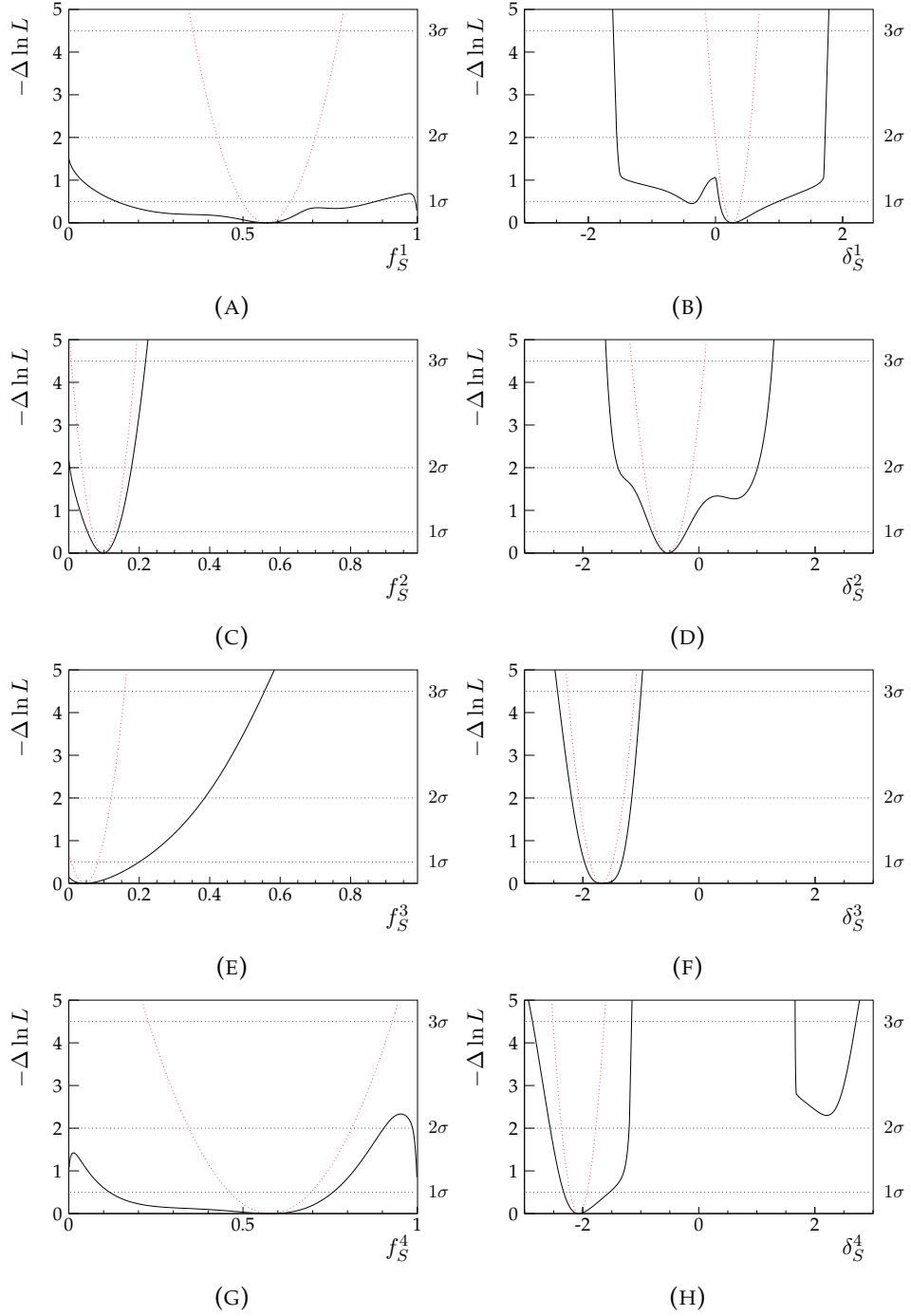
FIGURE 2.13: Profile likelihood scans of S-wave related parameters of interest. The black line corresponds to the profile likelihood function projected over the corresponding parameter. The red curve is a parabola, with a symmetric Gaussian error. There are several interesting features in those plots. First of all f_S^1 and f_S^4 minima are very broad due to the low number of events in that $m_{K\pi}$ bin. Second the f_S^2 and f_S^3 minima are close to a physical boundary.

in other words if the real experiment would be repeated n times, then 68% of these times the p^{true} would be contained in the previous interval.

Furthermore, potential biases to a certain parameter of interest arising

FIGURE 2.14: Pull distributions of the parameters of the $\mathcal{A}_i^{\text{CP}}$ interest. The distributions imply that the corresponding parameters are Gaussian distributed.

from the fitting model can be deduced from the mean of the pull and error distributions as shown in Eq. 2.45.

$$\text{b}_p = |\langle P \rangle|\ \langle \sigma^i \rangle, \quad \begin{matrix} \sigma^i = \sigma^i_+ & \text{if} & \langle P \rangle < 0 \\ \sigma^i = \sigma^i_- & \text{if} & \langle P \rangle > 0 \end{matrix} \quad . \tag{2.45}$$

The pull distributions for the parameters of interest are shown from Figure 2.14 to Figure 2.15. The corresponding biases are summarized in Table 2.14.
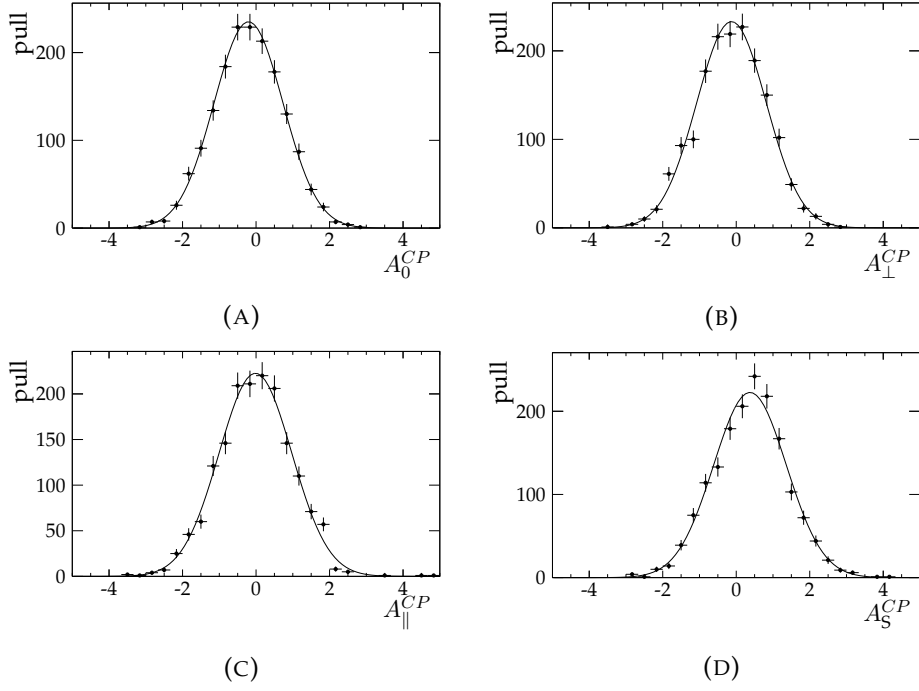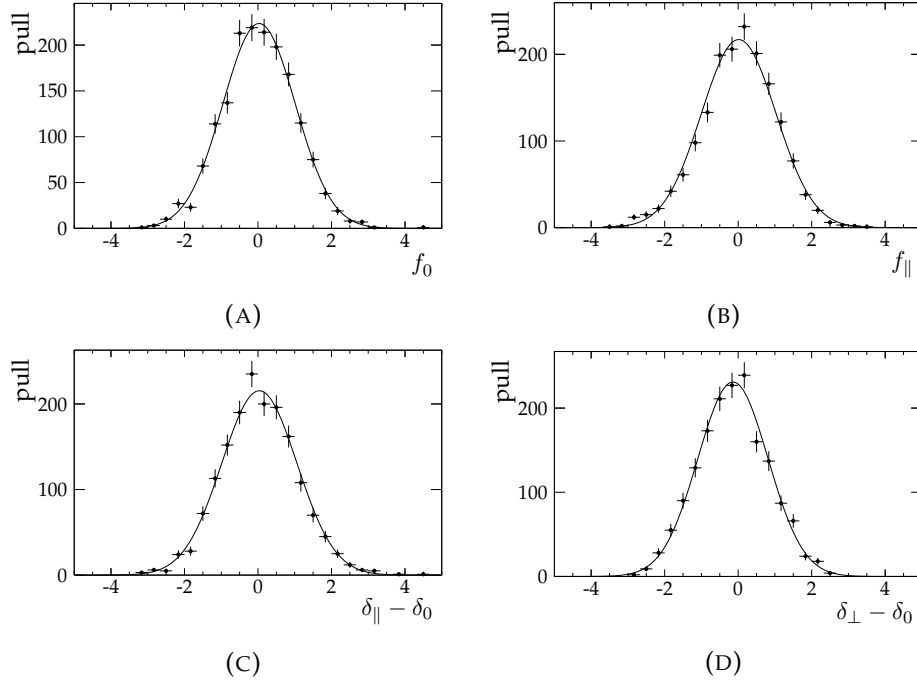
FIGURE 2.15: Pull distributions of the parameters of the P-wave interest. The distributions imply that the corresponding parameters are Gaussian distributed.

parameters	mean		rms		
	fit	pull	fit	pull	bias
$\mathcal{A}_0^{\mathrm{CP}}$	-0.205	-0.201	0.938	0.937	0.013
$\mathcal{A}_\perp^{\mathrm{CP}}$	-0.129	-0.126	0.946	0.943	0.012
$\mathcal{A}_\parallel^{\mathrm{CP}}$	-0.016	-0.020	0.990	0.986	0.003
$\mathcal{A}_S^{\mathrm{CP}}$	$+0.374$	$+0.375$	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
$\delta_\perp - \delta_0$	-0.142	-0.145	0.954	0.951	0.016
$\delta_\parallel - \delta_0$	$+0.037$	$+0.041$	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
δ_S^1	-0.394	-0.391	1.010	1.000	0.063
δ_S^2	-0.037	-0.037	1.086	1.079	0.009
δ_S^3	$+0.138$	$+0.135$	0.921	0.916	0.024
δ_S^4	-0.090	-0.092	0.975	0.965	0.013

TABLE 2.14: Summary of toy study. Mean and rms of each pull distribution is shown for two case. The first labeled as pull implies that the values are computed directly from the distribution, whereas the second labeled as fit corresponds to the fitted Gaussian curve on to of each pull distribution.
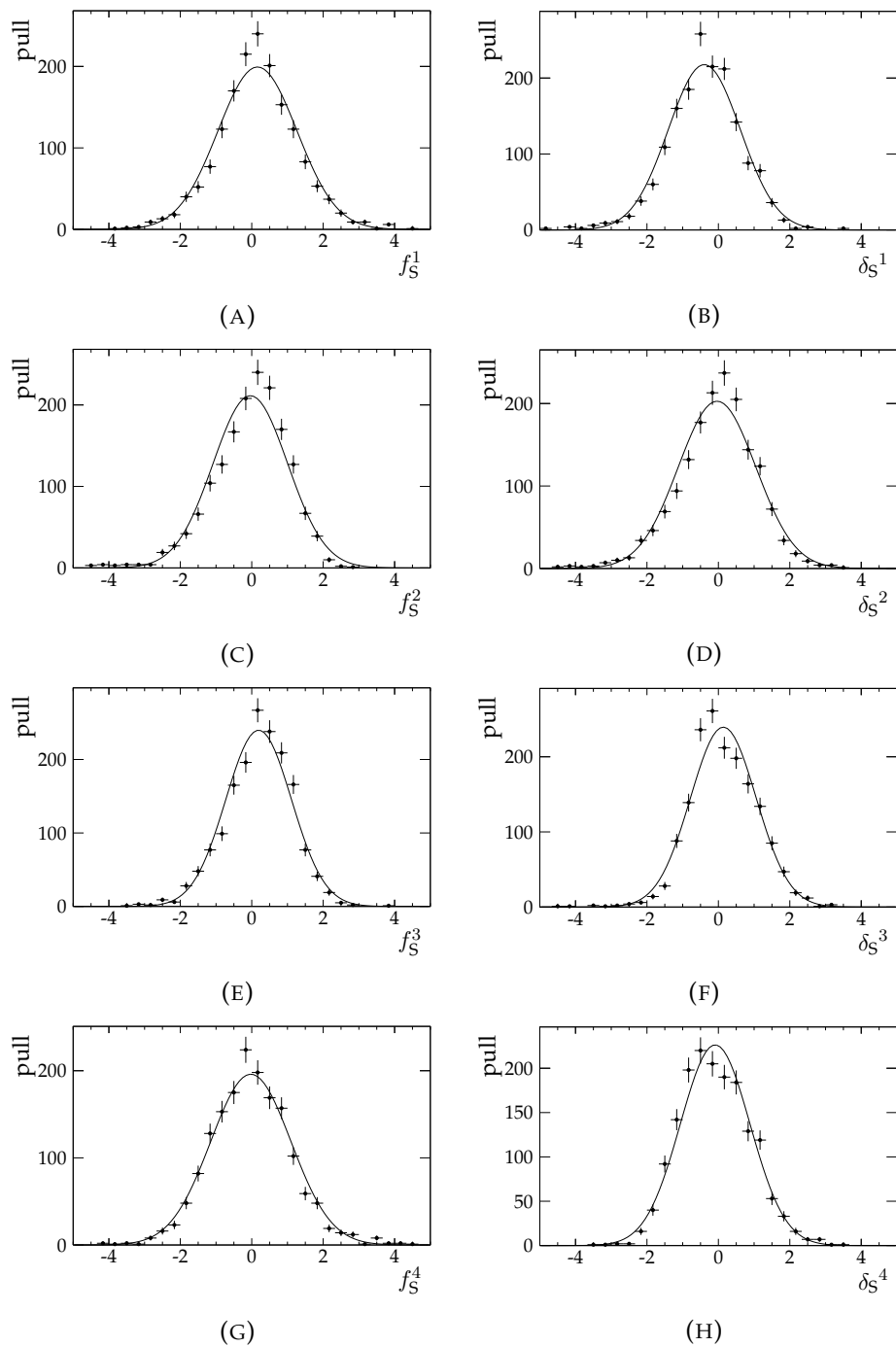
(A)

(B)

(C)

(D)

(E)

(F)

(G)

(H)

FIGURE 2.16: Pull distributions of the parameters of the S-wave interest. The distributions imply that the corresponding parameters are Gaussian distributed. Comment on the deltaS3 and fS3 after you accumulate all toys.

Chapter 3

Effects of Penguins on ϕ_S

3.1 Formalism

- H_i, related to the branching ratios and polarisation fractions,

$$
\begin{aligned}
H_i & \equiv \frac{1}{\epsilon}\left|\frac{\mathcal{A}'_i}{\mathcal{A}_i}\right|^2 \frac{\mathrm{PhSp}\left(B_s^0 \to J/\psi\phi\right)}{\mathrm{PhSp}(B_s^0 \to J/\psi\overline{K}^{*0})} \frac{\mathcal{B}B_s^0 \to J/\psi\overline{K}^{*0}_{\mathrm{theo}}}{\mathcal{B}B_s^0 \to J/\psi\phi_{\mathrm{theo}}} \frac{f_i}{f'_i}, (3.1)\\
& = \frac{1 - 2a_i\cos(\theta_i)\cos\gamma + a_i^2}{1 + 2\epsilon a'_i\cos\theta'_i\cos\gamma + \epsilon^2 a_i'^2},
\end{aligned}
$$

- A_i^{CP}, the direct CP violation asymmetries[1].

$$
A_i^{CP} = -\frac{2a_i\sin\theta_i\sin\gamma}{1 - 2a_i\cos\theta_i\cos\gamma + a_i^2}. \tag{3.2}
$$

$$
\mathrm{PhSp}(B \to V_1 V_2) \equiv \frac{1}{16\pi m_B}\Phi\left(\frac{m_{V_1}}{m_B}, \frac{m_{V_2}}{m_B}\right), \tag{3.3}
$$

$$
a_i = a'_i, \qquad \theta_i = \theta'_i, \tag{3.4}
$$

$$
\tan(\Delta\phi_{s,i}) = \frac{2\epsilon a'_i\cos\theta'_i\sin\gamma + \epsilon^2 a_i'^2\sin 2\gamma}{1 + 2\epsilon a'_i\cos\theta'_i\cos\gamma + \epsilon^2 a_i'^2\cos 2\gamma}. \tag{3.5}
$$

$$
\left|\frac{\mathcal{A}'_0(B_s^0 \to J/\psi\phi)}{\mathcal{A}_0(B_s^0 \to J/\psi\overline{K}^{*0})}\right| = 1.23 \pm 0.16, \tag{3.6}
$$

$$
\left|\frac{\mathcal{A}'_\parallel(B_s^0 \to J/\psi\phi)}{\mathcal{A}_\parallel(B_s^0 \to J/\psi\overline{K}^{*0})}\right| = 1.28 \pm 0.15, \tag{3.7}
$$

$$
\left|\frac{\mathcal{A}'_\perp(B_s^0 \to J/\psi\phi)}{\mathcal{A}_\perp(B_s^0 \to J/\psi\overline{K}^{*0})}\right| = 1.20 \pm 0.12, \tag{3.8}
$$

$$
H_0 = 0.99 \pm 0.07\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst}) \pm 0.27\,(|\mathcal{A}'_i/\mathcal{A}_i|) = 0.99 \pm 0.28, \quad (3.9)
$$
$$
H_\parallel = 0.91 \pm 0.14\,(\mathrm{stat}) \pm 0.08\,(\mathrm{syst}) \pm 0.21\,(|\mathcal{A}'_i/\mathcal{A}_i|) = 0.91 \pm 0.27, \;(3.10)
$$
$$
H_\perp = 1.47 \pm 0.14\,(\mathrm{stat}) \pm 0.11\,(\mathrm{syst}) \pm 0.28\,(|\mathcal{A}'_i/\mathcal{A}_i|) = 1.47 \pm 0.33. \;(3.11)
$$

$$
\gamma = \left(73.2^{+6.3}_{-7.0}\right)^\circ \quad (\text{CKMfitter [12]}), \tag{3.12}
$$

---

[1]Conventions: $A_i^{CP} = -\mathcal{A}_{\mathrm{dir}}^{CP}$ used in Ref. [13]

$$\Re[a_0] = \quad 0.01^{+0.97}_{-0.32}\,, \quad \Im[a_0] = \quad 0.025^{+0.035}_{-0.031}\,, \quad \chi^2_{\min} = 1.1 \times 10^{-7}\,, \quad (3.13)$$

$$\Re[a_\parallel] = \quad 0.31^{+0.58}_{-0.51}\,, \quad \Im[a_\parallel] = -0.082^{+0.074}_{-0.087}\,, \quad \chi^2_{\min} = 1.2 \times 10^{-3}\,, \quad (3.14)$$

$$\Re[a_\perp] = -0.44^{+0.27}_{-0.21}\,, \quad \Im[a_\perp] = \quad 0.037^{+0.079}_{-0.076}\,, \quad \chi^2_{\min} = 1.5 \times 10^{-6}\,, \quad (3.15)$$

$$a_0 = 0.03^{+0.97}_{-0.03}\,, \qquad\qquad \theta_0 = \quad \left(64^{+116}_{-244}\right)^\circ\,, \qquad (3.16)$$

$$a_\parallel = 0.32^{+0.58}_{-0.32}\,, \qquad\qquad \theta_\parallel = - \left(15^{+150}_{-14}\right)^\circ\,, \qquad (3.17)$$

$$a_\perp = 0.45^{+0.21}_{-0.27}\,, \qquad\qquad \theta_\perp = \quad (175 \pm 10)^\circ\,. \qquad (3.18)$$

$$a_0 = 0.03^{+0.97}_{-0.03}\,, \qquad\qquad \theta_0 = \quad \left(64^{+116}_{-244}\right)^\circ\,, \qquad (3.19)$$

$$a_\parallel = 0.32^{+0.58}_{-0.32}\,, \qquad\qquad \theta_\parallel = - \left(15^{+150}_{-14}\right)^\circ\,, \qquad (3.20)$$

$$a_\perp = 0.45^{+0.21}_{-0.27}\,, \qquad\qquad \theta_\perp = \quad (175 \pm 10)^\circ\,. \qquad (3.21)$$

$$\Delta\phi^{J/\psi\phi}_{s,0} = \quad 0.001^{+0.087}_{-0.011}\,(\text{stat})^{+0.013}_{-0.008}\,(\text{syst})^{+0.048}_{-0.030}\,(|\mathcal{A}'_i/\mathcal{A}_i|) = \quad 0.001^{+0.100}_{-0.033}\,,$$
$$(3.22)$$

$$\Delta\phi^{J/\psi\phi}_{s,\parallel} = \quad 0.031^{+0.049}_{-0.038}\,(\text{stat})^{+0.013}_{-0.013}\,(\text{syst})^{+0.031}_{-0.033}\,(|\mathcal{A}'_i/\mathcal{A}_i|) = \quad 0.031^{+0.059}_{-0.052}\,,$$
$$(3.23)$$

$$\Delta\phi^{J/\psi\phi}_{s,\perp} = -0.046^{+0.012}_{-0.012}\,(\text{stat})^{+0.007}_{-0.008}\,(\text{syst})^{+0.017}_{-0.024}\,(|\mathcal{A}'_i/\mathcal{A}_i|) = -0.046^{+0.022}_{-0.028}\,,$$
$$(3.24)$$

$$\Delta\phi^{J/\psi\phi}_{s,0} = \quad \left(0.1^{+5.0}_{-0.6}\,(\text{stat})^{+0.7}_{-0.5}\,(\text{syst})^{+2.7}_{-1.7}\,(|\mathcal{A}'_i/\mathcal{A}_i|)\right)^\circ = \quad \left(0.1^{+5.7}_{-1.9}\right)^\circ\,,$$
$$(3.25)$$

$$\Delta\phi^{J/\psi\phi}_{s,\parallel} = \quad \left(1.8^{+2.8}_{-2.2}\,(\text{stat})^{+0.7}_{-0.7}\,(\text{syst})^{+1.8}_{-1.9}\,(|\mathcal{A}'_i/\mathcal{A}_i|)\right)^\circ = \quad \left(1.8^{+3.4}_{-3.0}\right)^\circ\,,$$
$$(3.26)$$

$$\Delta\phi^{J/\psi\phi}_{s,\perp} = - \left(2.6^{+0.7}_{-0.7}\,(\text{stat})^{+0.5}_{-0.4}\,(\text{syst})^{+1.4}_{-1.0}\,(|\mathcal{A}'_i/\mathcal{A}_i|)\right)^\circ = - \left(2.6^{+1.6}_{-1.3}\right)^\circ\,.$$
$$(3.27)$$

3.2   External Inputs

3.3   Fit results

3.4 $\mathrm{SU}(3)$ breaking

Appendix A

Datasets and Stripping

This appendix is meant to give more details to the familiar with the LHCb jargon reader.

Events for this analysis are selected from:

- 1 fb^{-1}, collected in 2011 at $\sqrt{s} = 7\,\text{TeV}$, called `Reco14-Stripping20r1`.

- 2 fb^{-1}, collected in 2012 at $\sqrt{s} = 8\,\text{TeV}$, called `Reco14-Stripping20r0p1`.

The total luminosity of the processed files is 3.0 fb^{-1}.

The simulated samples used are:

- 5 M $B_s^0 \to J/\psi \overline{K}^{*0}$ signal events, from `Sim08g-Reco14a-Stripping20` processing, with TCK 0x409f0045 in flagging mode, which is representative of the data taken during 2012.

- 5 M $B_s^0 \to J/\psi \overline{K}^{*0}$ signal events, from `Sim08g-Reco14a-Stripping20r1` processing, with TCK 0x40760037 in flagging mode, which is representative of the data taken during 2011.

- 1 M $B^0 \to J/\psi K^{*0}$ signal events, from `Sim08a-Reco14a-Stripping20` processing, with TCK 0x409f0045 in flagging mode, which is representative of the data taken during 2012.

- 1 M $B^0 \to J/\psi K^{*0}$ signal events, from `Sim08b-Reco14a-Stripping20r1` processing, with TCK 0x40760037 in flagging mode, which is representative of the data taken during 2011.

- 5 M $B_s^0 \to J/\psi \phi$ signal events, from `Sim08a-Reco14a-Stripping20` processing, with TCK 0x409f0045 in flagging mode, which is representative of the data taken during 2012.

- 5 M $B_s^0 \to J/\psi \phi$ signal events, from `Sim08a-Reco14a-Stripping20r1` processing, with TCK 0x40760037 in flagging mode, which is representative of the data taken during 2011.

All these samples were generated using `Pythia8`. Each one contains simulated data from both magnet polarities.

The `BetaSBs2JpsiKstarWideLine` stripping line used for this analysis. Table A.1 lists the selection criteria in that stripping line and the offline final selection cuts used, which applies a mass constraint of $150\,\text{MeV}/c^2$ around the reconstructed J/ψ peak. The analysis is not restricted to any particular trigger line, i.e. an event should just pass at least one of the LHCb trigger lines.

Decay mode	Cut parameter	Stripping line	Final selection
$J/\psi \to \mu^+\mu^-$	$\Delta LL \mu\pi(\mu)$	> 0	–
	$\chi^2_{\text{vtx}}/\text{nDoF}(J/\psi)$	< 16	–
	$\chi^2_{\text{DOCA}}/\text{nDoF}(J/\psi)$	< 20	–
	$\chi^2_{\text{IP}}(\mu)$	–	> 16
	$p_{\text{T}}(\mu)$	> 0.5 GeV/c	–
	IsMuon(μ)	> 0	–
	$|M(\mu^+\mu^-) - M(J/\psi)|$	< 150 MeV/c^2	–
$\overline{K}^{*0} \to K^-\pi^+$	$\Delta LL K\pi(K)$	> 0	–
	$p_{\text{T}}(K)$	> 0.5 GeV/c	> 0.5 GeV/c
	Prob$_{\text{ghost}}$(track)(K)	< 0.8	–
	$\chi^2_{\text{IP}}(K)$	–	> 2
	ProbNNK(K)	–	> 0.21
	ProbNNK(K)/ProbNNp(K) (*)	–	> 0.99
	$\Delta LL K\pi(\pi)$	< 0	–
	$p_{\text{T}}(\pi)$	> 0.5 GeV/c	> 0.5 GeV/c
	Prob$_{\text{ghost}}$(track)(π)	< 0.8	–
	$\chi^2_{\text{IP}}(\pi)$	–	> 2
	ProbNNK(π)	–	< 0.01
	ProbNN$\pi(\pi)$/ProbNNp(π) (*)	–	> 21.9
	$\chi^2_{\text{vtx}}/\text{nDoF}(\overline{K}^{*0})$	< 25	–
	$\chi^2_{\text{DOCA}}/\text{nDoF}(\overline{K}^{*0})$	< 30	–
	$M(K^-\pi^+)$	$\in [750, 1900]$ MeV/c^2	–
	$|M(K^-\pi^+) - 896|$	–	< 70 MeV/c^2
$B_s^0 \to J/\psi\overline{K}^{*0}$	$M(B_s^0)$	$\in [5100, 5700]$ MeV/c^2	$\in [5150, 5650]$ MeV/c^2
	$\chi^2_{\text{vtx}}/\text{nDoF}(B_s^0)$	< 10	–
	DIRA(B_s^0)	> 0.999	–
	VS	> 1.5 mm	–
$B^+ \to J/\psi K^+$ veto	$|M(J/\psi, K) - 5279|$	–	> 60 MeV/c^2
BDTG (MVA)	2011 data/MC (**)	–	> 0.2
	2012 data/MC (**)	–	> 0.12

TABLE A.1: Selection criteria for $B_s^0 \to J/\psi\overline{K}^{*0}$ candidates in our stripping line and offline final selection. Also, (*) cuts are introduced for $\Lambda_b^0 \to J/\psi p K^-$ and $\Lambda_b^0 \to J/\psi p \pi^-$ peaking backgrounds suppression M(J/ψ) is the PDG nominal mass for J/ψ. VS stands for the separation between a vertex w.r.t. its associated PV. Explain some of the variables above maybe...

Appendix B

Additional Tables

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

TABLE B.1: Comparison between efficiency moments values obtained from the 2011 and 2012 $\bar{B}_s^0 \rightarrow J/\psi K^{*0}$ simulation samples. The overall χ^2 over the 12 degrees of freedom is 1.5, corresponding to a χ^2 probability of 11%.

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

TABLE B.2: Comparison between efficiency moments values obtained from the 2011 and 2012 $B_s^0 \to J/\psi \overline{K}^{*0}$ simulation samples. The overall χ^2 over the 12 degrees of freedom is 2.4, corresponding to a χ^2 probability of 0.37%.

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

TABLE B.3: Comparison between efficiency moments values obtained from the $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$ simulation samples. Production periods 2011 and 2012 are merged. The overall χ^2 over the 12 degrees of freedom is 1.7, corresponding to a χ^2 probability of 5.7%.

moment	$m_{K\pi}$ bin 1	$m_{K\pi}$ bin 2	$m_{K\pi}$ bin 3	$m_{K\pi}$ bin 4
c_{00}^1	-3.1421 ± 0.0536	-2.2392 ± 0.0194	-1.8757 ± 0.0172	-1.3891 ± 0.0449
c_{00}^2	+0.6198 ± 0.0745	-1.5482 ± 0.0305	-1.7036 ± 0.0246	-1.5303 ± 0.0479
c_{00}^3	-1.2819 ± 0.0861	-0.5077 ± 0.0364	-0.3875 ± 0.0289	-0.3059 ± 0.0622
c_{00}^4	+0.1335 ± 0.1025	+0.4200 ± 0.0425	+0.0522 ± 0.0334	-0.7166 ± 0.0691
c_{00}^5	+0.1604 ± 0.1146	+0.1465 ± 0.0464	+0.1763 ± 0.0375	-0.0305 ± 0.0770
c_{20}^0	+0.3156 ± 0.0352	+0.3023 ± 0.0141	+0.2957 ± 0.0127	+0.2540 ± 0.0258
c_{22}^0	+0.0069 ± 0.0267	+0.0414 ± 0.0103	+0.0725 ± 0.0113	+0.0930 ± 0.0257
c_{40}^0	+0.0984 ± 0.0326	+0.0766 ± 0.0145	+0.0887 ± 0.0127	+0.0410 ± 0.0262
c_{20}^1	-0.2174 ± 0.0759	-0.2362 ± 0.0317	-0.1925 ± 0.0214	-0.1725 ± 0.0445
c_{40}^1	-0.1033 ± 0.0676	-0.0682 ± 0.0333	-0.0638 ± 0.0215	-0.0503 ± 0.0456
c_{20}^2	-0.0738 ± 0.0962	-0.0742 ± 0.0477	-0.1120 ± 0.0306	-0.1287 ± 0.0526
c_{22}^2	+0.0039 ± 0.0609	-0.0487 ± 0.0220	-0.0745 ± 0.0233	-0.0434 ± 0.0451

TABLE B.4: Final efficiency moments values from the $B_s^0 \to J/\psi \overline{K}^{*0}$ simulation sample after applying the corrections described in Section 2.2.3. It is interesting to notice, even without any χ^2 test check that the efficiency moments vary more between the $m_{K\pi}$ bins than it does between production periods (see Table B.1, Table B.2) or between $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$ simulation samples (see Table B.3). The situation is similar in case of no corrections applied as well.

moment	$m_{K\pi}$ bin 1	$m_{K\pi}$ bin 2	$m_{K\pi}$ bin 3	$m_{K\pi}$ bin 4
c_{00}^1	-4.0320 ± 0.0562	-2.8430 ± 0.0150	-1.8788 ± 0.0173	-0.5687 ± 0.0365
c_{00}^2	+1.2914 ± 0.1030	-1.2462 ± 0.0227	-1.7805 ± 0.0246	-3.0352 ± 0.0458
c_{00}^3	-1.0834 ± 0.1225	+0.4463 ± 0.0274	-0.5942 ± 0.0289	-0.8462 ± 0.0543
c_{00}^4	+0.5572 ± 0.1397	-0.1800 ± 0.0311	+0.2157 ± 0.0325	+0.1747 ± 0.0652
c_{00}^5	-0.1778 ± 0.1538	-0.0238 ± 0.0345	+0.2763 ± 0.0365	+0.3265 ± 0.0707
c_{20}^0	+0.2204 ± 0.0446	+0.2853 ± 0.0117	+0.2893 ± 0.0124	+0.3296 ± 0.0252
c_{22}^0	+0.0052 ± 0.0330	+0.0321 ± 0.0101	+0.0413 ± 0.0108	+0.0599 ± 0.0240
c_{40}^0	+0.0400 ± 0.0423	+0.0863 ± 0.0118	+0.0714 ± 0.0123	+0.0708 ± 0.0248
c_{20}^1	-0.2249 ± 0.1083	-0.2692 ± 0.0215	-0.2004 ± 0.0219	-0.1287 ± 0.0389
c_{40}^1	-0.0109 ± 0.1007	-0.0732 ± 0.0222	-0.0649 ± 0.0217	-0.0526 ± 0.0386
c_{20}^2	-0.0013 ± 0.1450	-0.0534 ± 0.0292	-0.1268 ± 0.0307	-0.2497 ± 0.0491
c_{22}^2	-0.0262 ± 0.0882	+0.0252 ± 0.0204	-0.0496 ± 0.0222	-0.0652 ± 0.0464

TABLE B.5: Final efficiency moments values from the $\bar{B}_s^0 \to J/\psi K^{*0}$ simulation sample after applying the corrections described in Section 2.2.3. It is interesting to notice, even without any χ^2 test check that the efficiency moments vary more between the $m_{K\pi}$ bins than it does between production periods (see Table B.1, Table B.2) or between $B_s^0 \to J/\psi \overline{K}^{*0}$ and $\bar{B}_s^0 \to J/\psi K^{*0}$ simulation samples (see Table B.3). 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_{20}^0	c_{40}^0	c_{20}^1	c_{40}^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_{20}^1	0.734	-0.546	0.463	-0.328	0.186	-0.919		-0.402	1.0			
c_{40}^1	0.331	-0.247	0.227	-0.186	0.133	-0.41		-0.914	0.465	1.0		
c_{20}^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 B.6: Efficiency moments correlation matrix for $B_s^0 \to J/\psi \overline{K}^{*0}$. 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 2.2.2.

	$\mathcal{A}_0^{CP}$	$\mathcal{A}_\parallel^{CP}$	$\mathcal{A}_\perp^{CP}$	$\mathcal{A}_S^{CP}$	f_0	$f_\parallel$	$\delta_\parallel$	$\delta_\perp$	f_S^1	f_S^2	f_S^3	f_S^4	δ_S^1	δ_S^2	δ_S^3	δ_S^4
$\mathcal{A}_0^{CP}$	+1.00															
$\mathcal{A}_\parallel^{CP}$	-	+1.00														
$\mathcal{A}_\perp^{CP}$	-	-0.42	+1.00													
$\mathcal{A}_S^{CP}$	+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	+1.00							
f_S^2	-	-	-	-	+0.10	-	-	-0.11	-	+1.00						
f_S^3	-	-	-	-	-	-	-	-	-	-	+1.00					
f_S^4	+0.13	-	-0.15	-	-0.11	-	+0.10	-	-	-	-	+1.00				
δ_S^1	-	-	-	-	-	-	-	+0.18	-0.34	-	-	-	+1.00			
δ_S^2	-	-	-	-	+0.14	-0.13	+0.20	+0.17	-	-0.55	-	-	-	+1.00		
δ_S^3	-	-	-	-	+0.25	-	-	-	-	-	+0.23	-	-	-	+1.00	
δ_S^4	-	-0.12	-	-	+0.13	-	-	-	-	-	-	-0.38	-	-	-	+1.00

TABLE B.7: Correlation matrix of the parameters of interest. Correlations below 10% are not shown. Comment a bit on the values and move it to the result section

Bibliography

[1] R Aaij et al. "Amplitude analysis and the branching fraction measurement of $\bar{B}_s^0 \to J/\psi K^+ K^-$". In: *Phys.Rev.* D87.7 (2013), p. 072004. arXiv: 1302.1213 [hep-ex].

[2] R. Aaij et al. "Amplitude analysis of $B^0 \to (\pi^+\pi^-)(\pi^+\pi^-)$ decays and first observation of $B^0 \to \rho^0\rho^0$". In: (2014). in preparation.

[3] R. Aaij et al. "First observation of CP violation in the decays of B_s^0 mesons". In: *Phys. Rev. Lett.* 110 (2013), p. 221601. DOI: 10.1103/PhysRevLett.110.221601. arXiv: 1304.6173 [hep-ex].

[4] R. Aaij et al. "Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays". In: *JHEP* 07 (2014), p. 041. DOI: 10.1007/JHEP07(2014)041. arXiv: 1405.2797 [hep-ex].

[5] R. Aaij et al. "Measurement of the $\bar{B}^0$–B^0 and $\bar{B}_s^0$–B_s^0 production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV". In: *Phys. Lett.* B739 (2014), p. 218. DOI: 10.1016/j.physletb.2014.10.005. arXiv: 1408.0275 [hep-ex].

[6] R. Aaij et al. "Measurement of the semileptonic CP asymmetry in B^0–$\bar{B}^0$ mixing". In: (2014). to appear in Phys. Rev. Lett. arXiv: 1409.8586 [hep-ex].

[7] Roel Aaij et al. "Measurement of resonant and CP components in $\bar{B}_s^0 \to J/\psi\pi^+\pi^-$ decays". In: *Phys.Rev.* D89.9 (2014), p. 092006. arXiv: 1402.6248 [hep-ex].

[8] Roel Aaij et al. "Measurement of the resonant and CP components in $\overline{B}^0 \to J/\psi\pi^+\pi^-$ decays". In: *Phys.Rev.* D90.1 (2014), p. 012003. arXiv: 1404.5673 [hep-ex].

[9] K. Abe et al. "Measurements of branching fractions and decay amplitudes in $B \to J/\psi K^*$ decays". In: *Phys.Lett.* B538 (2002), pp. 11–20. DOI: 10.1016/S0370-2693(02)01969-X. arXiv: hep-ex/0205021 [hep-ex].

[10] A. Hoecker P et al. "TMVA Toolkit for Multivariate Data Analysis with ROOT". In: *PoS ACAT:040* 040 (2007). arXiv: 0703039 [physics.data-an].

[11] L. Amoroso. "Richerche intorno alla curve die redditi". In: *Ann. Mat. Pura. Appl.* 21 (1925), pp. 123–159.

[12] J. Charles et al. "Current Status of the Standard Model CKM Fit and Constraints on $\Delta F = 2$ New Physics". In: *Phys.Rev.* D91.7 (2015). Online update: CKM 2014, p. 073007. DOI: 10.1103/PhysRevD.91.073007. arXiv: 1501.05013 [hep-ph].

[13] Kristof De Bruyn and Robert Fleischer. "A Roadmap to Control Penguin Effects in $B_d^0 \to J/\psi K_S^0$ and $B_s^0 \to J/\psi\phi$". In: *JHEP* 1503 (2015), p. 145. DOI: 10.1007/JHEP03(2015)145. arXiv: 1412.6834 [hep-ph].

[14] Daan van Eijk and L M Bouter. "Ageing and the Decay of Beauty: Radiation Hardness of the LHCb Outer Tracker and Time-Dependent CP Violation using $B_s^0 \to J/\psi\phi$". Presented 2012. PhD thesis. Vrije U., Amsterdam, 2012. URL: https://cds.cern.ch/record/1484405.

[15] Yuanning Gao, Zhenwei Yang, and Xuhao Yuan. "Measurement of Λ_b differential production cross section in $\Lambda_b \to J/\psi pK^-$ decay". In: (2014).

[16] M. Jacob and G.C. Wick. "On the general theory of collisions for particles with spin". In: *Annals Phys.* 7 (1959), pp. 404–428. DOI: 10.1016/0003-4916(59)90051-X.

[17] G Krämer, T Mannel, and W F Palmer. "Angular correlations in the decays $B \to VV$ using heavy quark symmetry". In: *Z. Phys. C* 55.DESY-92-042. IKDA-92-8 (1992), pp. 497–502.

[18] Robert K. Kutschke. *An Angular Distribution Cookbook*. internal CLEO note, 1996.

[19] Jeroen van Leerdam. "Measurement of CP Violation in Mixing and Decay of Strange Beauty Mesons". Presented 2016. PhD thesis. Vrije U., Amsterdam, 2012. URL: BLAH.

[20] Diego Martinez Santos and Frederic Dupertuis. "Mass distributions marginalized over per-event errors". In: *Nucl.Instrum.Meth.* A764 (2014), pp. 150–155. DOI: 10.1016/j.nima.2014.06.081. arXiv: 1312.5000 [hep-ex].

[21] Francois R. Le Diberder Muriel Pivk. "sPlot: a statistical tool to unfold data distributions". In: *Nucl.Instrum.Meth.* A555 (2005), pp. 356–369. arXiv: 0402083 [physics.data-an].

[22] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[23] Tristan du Pree, M H M Merk, and H G Raven. "Search for a Strange Phase in Beautiful Oscillations". Presented on 22 Oct 2010. PhD thesis. Amsterdam: Vrije U., Amsterdam, 2010. URL: https://cds.cern.ch/record/1299931.

[24] "Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions". In: (2013). URL: https://cds.cern.ch/record/1559262.

[25] G Valencia. "Angular correlations in the decay $B \to VV$ and CP violation". In: *Phys. Rev. D* 39.BNL-42229 (1988), pp. 3339–3345.

[26] Y. Xie. "Figure of merit for "$B_s^0 \to KK$ event selection optimization". In: *LHCb internal WG presentation* (Feb. 2013). URL: https://indico.cern.ch/conferenceDisplay.py?confId=236950.