UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXV CICLO DEL DOTTORATO DI RICERCA IN FISICA

Search for New Physics in the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\,\phi$ Decays at CDF

Settore scientifico-disciplinare
FIS/04 Fisica Nucleare e Subnucleare

Dottorando
Mirco Dorigo

Direttore della Scuola
Prof. Paolo Camerini

Supervisori
Dr. Anna Maria Zanetti
Dr. Marco Rescigno

Tutore Universitario
Prof. Lorenzo Vitale

Anno Accademico 2011/2012

Abstract

We present a search for physics beyond the standard model (SM) through a measurement of the violation of the charge-parity (*CP*) symmetry in two decays of the B_s^0 meson using data collected by the Collider Detector at Fermilab (CDF) in proton-antiproton collisions at the center-of-mass energy of 1.96 TeV. We exploit the decays $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ and $B_s^0 \to \phi(\to K^+K^-)\phi(\to K^+K^-)$, for which the SM accurately predicts very small or vanishing *CP* violation; both decay modes are very sensitive to new sources of *CP* violation expected in a broad class of SM extensions. We analyze the time-dependent *CP* asymmetry of the $B_s^0 \to J/\psi\,\phi$ decays collected in the full CDF Run II dataset for providing the final measurement of the B_s^0-$\bar{B}_s^0$ mixing phase, $2\beta_s$, and we present the first measurement of *CP* violation in $B_s^0 \to \phi\phi$ decays, through the determination of two time-integrated *CP* asymmetries, $\mathcal{A}_v$ and $\mathcal{A}_u$, using an original method developed in this work. We find: $-0.06 < \beta_s < 0.30$ at the 68% confidence level; $\mathcal{A}_v = (-12.0 \pm 6.4(\text{stat}) \pm 1.6(\text{syst}))\%$; and $\mathcal{A}_u = (-0.7 \pm 6.4(\text{stat}) \pm 1.8(\text{syst}))\%$. In addition, we provide measurements of the decay width difference between the light and heavy mass eigenstates of the B_s^0 meson, $\Delta\Gamma_s = 0.068 \pm 0.026(\text{stat}) \pm 0.009(\text{syst})\,\text{ps}^{-1}$; and of their mean lifetime, $\tau_s = 1.528 \pm 0.019(\text{stat}) \pm 0.009(\text{syst})\,\text{ps}$. All results are among the most precise determinations from a single experiment and exhibit an excellent agreement with the SM predictions. They have been published in two letters in *Physical Review* [1, 2].

Contents

Introduction **vii**

1 Search for new physics in the B_s^0-$\bar{B}_s^0$ mixing **1**

1.1 Probing the new physics scales with flavor physics . . . 1

1.2 Flavor constraints of the SM . . . 5

1.3 B_s^0-$\bar{B}_s^0$ mixing bounds on new physics . . . 7

1.3.1 The B_s^0-$\bar{B}_s^0$ mixing phase . . . 11

1.4 CKM phases from time-dependent CP asymmetries . . . 13

1.5 B_s^0-$\bar{B}_s^0$ mixing phase in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays . . . 16

2 Phenomenology of $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays **19**

2.1 Angular amplitudes of $B \to V_1 V_2$ decays . . . 19

2.1.1 Angles bases . . . 21

2.2 The B_s^0-$\bar{B}_s^0$ mixing phase in the amplitudes time-evolution . . . 26

2.2.1 Time evolution of the $(S+P)$-waves system . . . 29

2.2.2 Discrete symmetries . . . 33

2.3 Time-integrated triple products asymmetries . . . 35

2.4 Measurements' strategies . . . 39

2.4.1 $B_s^0 \to J/\psi\phi$ analysis . . . 39

2.4.2 $B_s^0 \to \phi\phi$ analysis . . . 41

3 Experimental Apparatus **43**

3.1 The Fermilab's Tevatron collider . . . 43

3.1.1 The proton beam . . . 44

3.1.2 The antiproton beam . . . 45

3.1.3 The collisions . . . 45

3.1.4 Run II performances and achievements . . . 46

3.2 The CDF II detector . . . 47

3.2.1 Tracking system . . . 49

3.2.2 Particle identification . . . 56

3.2.3 Muons system . . . 61

3.2.4 Other detectors . . . 63

3.2.5 Trigger and data acquisition systems . . . 65

3.2.6 Operations and data quality . . . 67

3.2.7 Offline data processing . . . 69
3.2.8 Monte Carlo simulation . . . 70

4 Flavor tagging **73**
4.1 Flavor tagging principles . . . 73
4.2 Opposite Side Tagger . . . 75
4.2.1 Calibration of OST in the full Run II sample . . . 76
4.3 Same Side Kaon Tagger . . . 83
4.4 Effect of the tagging in the β_s measurements . . . 86

5 $B_s^0 \to J/\psi\phi$ Data Set **89**
5.1 The dimuon trigger . . . 89
5.2 Reconstruction and kinematic variables . . . 91
5.3 Offline selection . . . 92
5.3.1 Neural Network Selection . . . 93
5.4 Final sample . . . 98
5.5 Monte Carlo samples . . . 100
5.5.1 Phase-space model . . . 100
5.5.2 Simulated $B_s^0 \to J/\psi K^+K^-$ data . . . 103
5.5.3 Simulated $B^0 \to J/\psi K^+\pi^-$ data . . . 106

6 Maximum Likelihood Fit of the $B_s^0 \to J/\psi\phi$ analysis **109**
6.1 Maximum Likelihood method . . . 109
6.2 The Probability Distribution Function . . . 110
6.2.1 PDF of the B mass . . . 112
6.2.2 PDF of transversity angles and ct for the signal . . . 113
6.2.3 PDF of the decay-length of the background . . . 117
6.2.4 PDF of transversity angles for the background . . . 118
6.2.5 PDF of the decay-length uncertainty . . . 119
6.2.6 PDF of flavor tagging variables . . . 121
6.2.7 Summary of the PDF parameters . . . 121
6.3 Investigation of the ML fit's features . . . 123
6.3.1 Pull studies . . . 126
6.3.2 Fit of realistic simulated data samples . . . 130

7 Measurements of B_s^0 lifetime, $\Delta\Gamma_s$, and $B_s^0 \to J/\psi K^+K^-$ amplitudes. **135**
7.1 Fit results for β_s fixed at the SM expectation . . . 135
7.2 Systematic Uncertainties . . . 141

7.3 Final results . . . 145
7.4 Estimation of the S-wave fraction . . . 146

8 Mixing phase bounds **151**
8.1 Measurement method . . . 151
8.2 Results of the β_s-floating fit . . . 153
8.3 Confidence regions . . . 155
8.4 Confidence intervals . . . 161

9 $B_s^0 \to \phi\phi$ Data Set **163**
9.1 Trigger on displaced tracks . . . 163
9.2 Offline selection . . . 169
9.3 Final sample . . . 173
9.4 Monte Carlo sample . . . 175

10 Measurement of triple products asymmetries **177**
10.1 Maximum likelihood fit . . . 177
10.2 Validation of the ML fit . . . 181
10.2.1 Pulls distribution . . . 181
10.2.2 Fit to Monte Carlo data samples . . . 182
10.3 Data sample fit . . . 185
10.4 Systematic uncertainties . . . 187

11 Conclusions **193**

A Derivation of the time-dependent CP asymmetry **199**

B Resonances in the K^+K^- mass spectrum **203**

C Time-evolution of the $B_s^0 \to J/\psi\phi$ amplitudes **207**

D Validation plots of phase space $B_s^0 \to J/\psi\phi$ Monte Carlo **213**

E Pulls distribution of the $B_s^0 \to J/\psi\phi$ fit **217**

F Variables correlation in the fit to the $J/\psi K^+K^-$ and K^+K^- masses. **221**

Bibliography **225**

Introduction

The recent observation of a Higgs-like boson confirms the standard model (SM) as an accurate and consistent description of the fundamental constituents of matter and their interactions at the electroweak energy scale and beyond [3, 4]. However, it is well established that the SM is necessarily invalidated at some energy scale not probed so far, hence, the search for SM-extensions is a chief goal in today's particle physics.[1] Quark-flavor physics offers a rich opportunity for this purpose and can probe energy scales higher than directly accessible with current and foreseen high-energy colliders [6, 7]. A plethora of precise measurements in the decays of strange and bottom hadrons pursued in the past two decades [8, 9], along with a mature phenomenology that provides accurate predictions, allows the redundant determination of quarks-mixing parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [10, 11] that can be compared for precision tests of the overall picture. All measurements performed so far are consistently described by few CKM parameters [8], nevertheless, sub-10% effects from new physics in the quarks dynamic are not ruled out by the current experimental precision [12].

In this dissertation we analyze two decay modes of the bottom-strange mesons, $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ and $B_s^0 \to \phi(\to K^+K^-)\phi(\to K^+K^-)$. The experimental effort to investigate the B_s^0 phenomenology has started only recently, and still limited information on the violation of the charge-parity (CP) symmetry in the B_s^0-meson physics is available [9].[2] In addition, theoretically-solid predictions are provided for some B_s^0-decays now experimentally-accessible [13]. The expectation of the size of CP violation in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decay is one of them; the CKM hierarchy provides vanishing expected-values of CP asymmetries in these decays, that can be significantly increased in a broad class of SM extensions. These features makes $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ golden channels in the indirect search for new physics [14, 15]. Specifically, the time-dependent CP asymmetry of the decays allows us to probe the phase of the B_s^0-$\bar{B}_s^0$ *mixing* amplitude, *i. e.* the phase of the amplitude governing flavor oscillations between B_s^0 and $\bar{B}_s^0$ meson. The mixing phase is approximated by $2\beta_s$, where β_s is expressed in terms of the CKM matrix-elements as $\arg(V_{ts}V_{tb}^\star/(V_{cs}V_{cb}^\star))$; the overall constraints of the CKM-matrix gives $\beta_s \approx 0.02$ [13], which is negligible with respect to the current experimental sensitivity.

First determinations of β_s by the CDF and D0 experiments, suggested an interesting but mildly significant deviation from the SM expectation [16, 17]. Even if not conclusive, later measurements have pointed towards a better consistency with the CKM expectation [18, 19, 20, 21]. This makes the theoretical understanding of some anomalies observed in the B-physics sector even more puzzling [13]; for instance, the 4σ departure from the SM expectation of the dimuon asymmetry observed by D0, which is tightly correlated with β_s, if generated in the

[1] An extensive literature is presents. For very recent developments see Ref. [5] and references therein.

[2] The violation of the CP symmetry in a physics process is its non-invariance for the transformations that inverts all spatial coordinates and replaces particles with by their antiparticles.

B_s^0-$\bar{B}_s^0$ mixing [22]. Providing more precise, even redundant, measurements in the B_s^0-$\bar{B}_s^0$ mixing phenomenology is therefore mandatory. Measurements of *CP* violation in $B_s^0 \to \phi\phi$ decays could add helpful complementary information. However, the small datasets available for this decay mode, which suffers a low decay-rate with respect to the $B_s^0 \to J/\psi\phi$ decay, do not allow yet for a measurement of the time-dependent *CP* asymmetry for probing directly the mixing phase.

This document presents the final measurement of β_s using $B_s^0 \to J/\psi\phi$ decays collected in the full CDF Run II dataset, and the world's first measurement of *CP* violation in $B_s^0 \to \phi\phi$ decays, through the determination of two time-integrated *CP*-violating asymmetries, the triple-products asymmetries, originally designed for this work. In addition, it presents measurements of the decay width difference between the light and heavy mass eigenstates of the B_s^0 meson, $\Delta\Gamma_s$; of their mean lifetime, τ_s; and of the angular-momentum composition of $B_s^0 \to J/\psi\phi$ decays. The *CP* violation in $B_s^0 \to \phi\phi$ decays is found to be less than 10%; the measurement of β_s is consistent with zero within about 10% accuracy; and $\Delta\Gamma_s$ is measured with a resolution comparable with the SM prediction uncertainty. The results presented in this thesis are among the most precise determinations available to date from a single experiment and agree with measurements from other experiments and with the SM expectations [19, 20, 21, 13, 23]. They contribute to establish that large non-SM contributions to the B_s^0 mixing phase are unlikely, and provide important constraining information for phenomenological model building. All results have been published in two letters to *Physical Review* [1, 2].

The final measurement of β_s supersedes the previously-published CDF result using half of the current dataset [18], and, in addition to the tuning of the analysis-tools for the larger dataset, presents a significantly improved understanding of the sample composition, in terms of the $J/\psi K^+K^-$ signal and the $J/\psi K^+\pi^-$ background (due to the misidentification of the pion as a kaon). We demonstrate that the $J/\psi K^+\pi^-$ contamination gives a background larger than typically assumed, which mimics a greater S-wave signal than present, if neglected. The measured S-wave component of the signal has indeed shown discrepancies among experiments. Along the same line, the development of a new simulation of the $B_s^0 \to J/\psi K^+K^-$ decays, completed with the whole resonances structure of the K^+K^- mass spectrum, is used to build confidence in the determinations by our analysis of the S-wave fraction, of $\Delta\Gamma_s$ and of β_s.

The $B_s^0 \to \phi\phi$ analysis is built upon the first measurement of the polarization amplitudes of this decay mode [2, 24], and introduces for the first time in literature the two *CP*-violating asymmetries in the $B_s^0 \to \phi\phi$ decay. The asymmetries have been devised in collaboration with phenomenologists, who have consolidated the implications of a possible nonzero asymmetry of this decay mode [25, 26]. These asymmetries are particularly useful for *CP*-violation studies of rare decays at hadron-collider experiments, because the developed method for their measurement does not require flavor-tagging information.

In the following, the organization of the thesis is outlined. The first part of the dissertation is dedicated to aspects common to both analyses. Chapter 1 briefly summarizes the motivation behind the measurements presented in the thesis in the general contest of the indirect searches of new physics. The angular and decay-time distribution of the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays which allows the measurement of β_s and the triple-product asymmetries are discussed in Chap. 2 along with a short review of the current experimental situation and an outline of the analyses. Chapter 3 describes the experimental apparatus, with a brief description of the Tevatron accelerator, the CDF II detector, and the data-taking operations. Chapter 4 describes the algorithms used in CDF to determine the flavor of the B_s^0 meson at production and the updated calibration of the opposite-tagging algorithm with the whole dataset.

Then, we move to the specific description of each measurements, starting with the analysis of the $B_s^0 \to J/\psi\phi$ decays. The reconstruction and selection of the data set are described in Chap. 5; the simulated samples used in several part of the analysis are also presented here. The maximum likelihood estimator for measuring the observables of interest is reported in Chap. 6 along with the tests aimed at evidencing its features. The results of the measurement of the B_s^0 lifetime, the widths difference $\Delta\Gamma_s$, and the angular-momentum composition of $B_s^0 \to J/\psi\phi$ decays are presented in Chap. 7. Finally, Chap. 8 is entirely devoted to the derivation of the mixing phase results.

The last part of the dissertation deals with the analysis of the $B_s^0 \to \phi\phi$ decays. Chapter 9 outlines the selection of the data and the final sample used in the measurement of the asymmetries; the simulated samples are also described here. Chapter 10 presents the distributions of the triple product variables and the maximum likelihood fit employed for the measurement of their asymmetries; the results are reported along with the estimation of their systematic uncertainties.

In the end, Chap. 11 is devoted to draw conclusions from this work with a few remarks on future prospects.

Chapter 1

Search for new physics in the B_s^0-$\bar{B}_s^0$ mixing

In this chapter we introduce the physics motivation behind the measurement of the B_s^0-$\bar{B}_s^0$ mixing phase. We briefly summarize the basics of the indirect searches for new physics and the current knowledge of the quarks sector of the SM. In particular, we focus on the role of the B_s^0-$\bar{B}_s^0$ phenomenology. We finally explain the features that make $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays promising to detect sensible deviations from the expected values of the B_s^0-$\bar{B}_s^0$ mixing phase.

1.1 Probing the new physics scales with flavor physics

The standard model (SM) of particle physics provides a mathematical description of three fundamental interactions, namely the strong, weak, and electromagnetic interactions, that act among the elementary spin-half particles, the quarks and the leptons.[1] The structure of the model stems from symmetries under transformations of a gauge group and implies that the various interactions are mediated by spin-one force carriers: eight massless gluons for the strong interaction; two charged massive bosons, $W^{\pm}$, and a single neutral massive boson, Z^0, for the weak interaction; and a massless photon, γ, for the electromagnetic interaction. The quarks and the leptons must reside in representations of the gauge group. Finally, the SM includes a spin-zero particle, the Higgs boson, to allow for the generation of particles masses from the spontaneous symmetry breaking of the gauge group of the electroweak interaction. Quarks and leptons can interact via Higgs-mediated interactions that, unlike gauge interactions, are not ruled by symmetry principles. These are called *Yukawa* interactions and are responsible for *flavor* physics. The term *flavor* is used to describe several copies of the quarks and leptons of the model that have same quantum charges. There are four of them in the SM, each coming in three flavors: up-type quarks (u, c, t), down-type quarks (d, s, b), charged leptons (e, μ, τ), and neutrinos (ν_e, ν_μ, ν_τ).

The Higgs boson has eluded all experimental efforts to discover it, until July 2012 when the CMS and ATLAS Collaborations announced the discovery of a new resonance produced in proton-proton collisions of the Large Hadron Collider (LHC) [3, 4]. The new particle has a mass of approximately $125\,\mathrm{GeV}/c^2$, and it shows properties compatible with the SM Higgs

[1] For a general introduction on the SM and high energy physics see Ref. [27].

interpretation. The discovery of the Higgs boson would complete the validation of the SM, which has explained very successfully most of the experimental phenomena probed so far [5]. However, it is well known that the SM has to be extended. Classical gravity, well described by general relativity, should break down at energy scales close to 10^{19} GeV, the Planck scale, where quantum effects of gravity may arise [28]. The SM is necessarily invalidated at such energy, and, above the Planck energy, a theory of quantum gravity must replace the SM.

The calculation of the Higgs mass in the current model is affected by divergences introduced by radiative corrections that invalidates the SM at an energy scale which depends on the measured mass of the Higgs particle. This energy scale represents the *cut-off* of the effective model, *i. e.*, the energy above which the model must be replaced by a more fundamental theory. Either a fine-tuning of the model parameters that can push the cut-off till the Planck scale, or new particles generated from a new symmetry principle present below the Planck scale till the electroweak scale, were argued [29]. Assuming the newly-observed particle at LHC is the SM Higgs boson, the measured value of its mass consolidates the SM at the electroweak scale, and allows a cut-off of the theory at larger energy scale[5]. The questions if new particles are present in the energy range from $\mathcal{O}(1)$ TeV till the Plank scale is therefore still open. It is also motivated from cosmological arguments for overcoming some experimental observations yet unexplained by currents models [8] (for instance, the observation that the 84% mass of the Universe is made of non-baryonic and non-luminous matters).

Heavy new particles can be produced *directly* in high-energy collisions and observed through their decay products. The key ingredient is the available center-of-mass energy: the higher the energy is, the heavier the particles that can be produced, the higher the new physics scale probed. Clearly, others key parameters are the rate of production of the new particles, that depend on the kind of process employed in the high-energy collisions, and the standard-physics background that can overwhelm the new signal. Ultimately, such ingredients can be expressed as functions of the new particles mass. Thus the success of this *direct* search depends on the unknown energy scale of the new physics particles. Pushing forward the energy frontier will require building colliders with ever-higher center-of-mass energies. However, the new physics scale can be at an energy not reachable with current and foreseen technology.

A complementary way is to infer the presence of new particles *indirectly* in processes where they could be virtually exchanged among SM particles, by detecting deviations from expectations precisely calculated in the SM [6, 7]. For indirect searches the production threshold is not an issue. Because quantum effects become as smaller as heavier are the virtual particles at play, higher new physics scales are explored by increasing the precision of the measurements while controlling the SM contributions with sufficient accuracy to disentangle genuine new effects from SM uncertainties. Flavor physics is the best candidate for indirect searches. Experimental access to a plethora of precisely measurable processes, along with a mature phenomenology that provides accurate predictions, allows the redundant determination of several SM free parameters that can be compared for precision tests of the overall picture [12]. The indirect approach

(a) (b)

Figure 1.1: Two examples of flavor transition of the b quark in terms of Feynman diagrams. In (a), the *tree*-level transition $b \to c$, where the b quark changes its flavor and charge $(-1/3)$ becoming a c quark (with charge $+2/3$) through a W^+ boson; in (b), the *loop*-mediated transition $b \to s$, where the b quark changes its flavor by exchanging an up-type quark (either u, or c, or t) and a W boson with the s quark. In this case, the charge of the initial quark and the charge of the final quark are the same. Tree-level transition involves quarks of different type (up-down and down-up), while loop-mediated transitions can change the flavor of two quarks of the same type (up-up and down-down).

has demonstrated itself very successful in building the current model of particles physics. For instance, the theoretical explanation [30] of the rare decay rate of quark transitions that change the strangeness flavor by two unit (such as $K^0 \to \mu^+\mu^-$ decays) was used to put upper limits on the c-quark mass before its experimental discovery [31, 32]. Another more recent example is the prediction of the large value of the top quark mass before its observation at the Tevatron experiments [33, 34], from the constraints imposed by the measurement of the B^0-$\bar{B}^0$ mesons oscillations [35].

The flavor structure of the SM is impressively predictive and peculiar. Flavor violation is allowed only in the quarks sector. Weak interactions mediated by $W^\pm$ bosons that changes the flavors of the quarks (flavor-changing charged-currents, FCCC) are universal; flavor transitions mediated by neutral currents (flavor-changing neutral-currents, FCNC) are highly suppressed. The latter cannot occur at *tree-level*, *i. e.*, through the mediation of a weak boson only, but they do require the intermediate exchange of a quark and a W boson (*loop* transition). In Fig. 1.1 we show this features with two examples of Feyman diagrams, *i. e.*, representations of the two flavor-changing transitions in terms of the elementary particles involved. The FCNC are further suppressed in the SM by the Glashow-Illiopoulos-Maiani (GIM) mechanism [36], namely the smallness of the mass differences between second- and first-generation quarks, and by the hierarchical structure of *quark-mixing* angles which gives the rotation of the quark flavor basis with respect to the weak-interaction basis. These special features are crucial in order to explain the observed pattern of weak decays, and any extension of the SM must account for them.

To describe new physics effects in flavor physics, we can use a generic effective-theory approach [6]. This strategy is very general and model-independent. Assuming the new physics scale to be heavier than the electroweak energy scale, we can describe new physics effects by means of a generalization of the Fermi theory. The SM Lagrangian represents a part of a more

Bounds on Λ (TeV)	Bounds on c_i ($\Lambda = 1$ TeV)	Mesons mixing
10^2–10^5	10^{-11}–10^{-7}	K^0-$\bar{K}^0$
10^3–10^4	10^{-7}–10^{-8}	D^0-$\bar{D}^0$
10^2–10^3	10^{-7}–10^{-6}	B^0-$\bar{B}^0$
10^2–10^3	10^{-5}	B_s^0-$\bar{B}_s^0$

Table 1.1: Bounds from experimental constraints on meson-mixing [6]. Mesons-mixing will be discussed in Sect. 1.3, in particular the case B_s^0-$\bar{B}_s^0$.

general local Lagrangian, which includes a series of operators with dimension $d > 4$, $O_i^{(d)}$, constructed in terms of SM fields, with couplings $c_i^{(d)}$ suppressed by inverse powers of an effective scale Λ, which represents the cut-off of the effective theory:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(d)}}{\Lambda^{(d-4)}} O_i^{(d)} \tag{1.1}$$

Bounds on Λ can be derived assuming an effective coupling $c_i \approx 1$; alternatively, the bounds on the respective couplings can be given assuming that $\Lambda \sim \mathcal{O}(1)$ TeV. This general approach allows us to analyze realistic extensions of the SM in terms of few parameters (the coefficients of the higher-dimensional operators). Consider a generic new physics model, where the suppression of FCNC processes is due only to the large masses of the particles that mediate them, *i. e.* the couplings are of order one. We can find the bounds on the scale Λ in order to satisfy the measurements of FCNC decay rates. Depending on the flavor process under analysis, this model-independent approach gives bounds of the order:

$$\Lambda \gtrsim 10^2 \text{ TeV}. \tag{1.2}$$

This implies that any new physics at TeV scale must have a highly-tuned flavor structure, *i. e.* the coupling c_i should have very suppressed values. As an example of the constraint that can be set following this method, Tab. 1.1 lists the bounds derived from measurements related to the mixing of neutral mesons for some operators that change the flavor of the decaying quark by two unit in Eq. (1.1).[2]

This gives an overview of the basics of the indirect method in searching for new physics with flavor constraints. Experimental measurements allow us to test and establish general features of the new theory (either its energy scale or its flavor structure), which hold independently of the dynamical details of the model. The phenomenology of the B, D, and K mesons is extremely useful for this purpose [12]. This thesis presents the analysis of two decay modes of the B_s^0 meson, $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$, which allows a measurement of the B_s^0-$\bar{B}_s^0$ mixing observables. The experimental effort to investigate the B_s^0 phenomenology has been started only recently, and still limited information in the B_s^0-meson sector of the CKM matrix is available [9],

[2] We will discuss the phenomenology of neutral-mesons mixing in Sect. 1.3.

although theoretically-solid predictions for some B_s^0-decays that are experimentally accessible are provided [13]. Moreover, recent measurements point to anomalies in the B-physics sector which could be explained by new physics in the B_s^0-$\bar{B}_s^0$ mixing amplitude [37, 13]; for instance, the 4σ departure from the SM expectation of the dimuon asymmetry observed by D0, could be tightly correlated with new physics amplitudes in the B_s^0-$\bar{B}_s^0$ mixing [22]. Providing further, more precise, even redundant, experimental information is therefore mandatory.

1.2 Flavor constraints of the SM

The starting point of the indirect approach to probe the new physics scale is that, in several realistic new physics models, we can neglect non-standard effects in all cases in which the corresponding effective operator is generated at tree level within the SM [6]. This general assumption implies that the experimental determination of the quark-mixing couplings via tree-level processes is free from the contamination of new physics contributions. We now briefly summarize the current knowledge on the quark flavor structure of the SM model to emphasize the need to look at loop-mediated process for searching for non-SM effects.

Within the SM the only source of flavor-changing interactions stems from the Yukawa sector of the quarks and it is originated from a rotation of the quarks flavor basis with respect to the weak-interaction basis. Such rotation is given by a unitary 3×3 complex matrix,

$$V_{\rm CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix},$$

that is the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [10, 11]. As a result of the fact that $V_{\rm CKM}$ is not diagonal, the $W^\pm$ gauge bosons (only them) couple to quark (mass eigenstates) of different generations. Thus FCCC occurs at tree-level, while FCNC are mediated only by loops (Fig. 1.1). The current knowledge of the CKM matrix elements moduli, as obtained from Ref. [12], is the following:

$$|V_{\rm CKM}| = \begin{pmatrix} 0.97426^{+0.00022}_{-0.00014} & 0.22539^{+0.00062}_{-0.00095} & 0.003501^{+0.000196}_{-0.000087} \\ 0.22526^{+0.00062}_{-0.00095} & 0.97345^{+0.00022}_{-0.00018} & 0.04070^{+0.00116}_{-0.00059} \\ 0.00846^{+0.00043}_{-0.00015} & 0.03996^{+0.00114}_{-0.00062} & 0.999165^{+0.000024}_{-0.000048} \end{pmatrix}. \tag{1.3}$$

Using just the observed hierarchy $|V_{ub}| \ll |V_{cb}| \ll |V_{us}|$, and $|V_{cd}| \ll 1$, one can expand $V_{\rm CKM}$ in powers of $\lambda = |V_{us}|$, the sine of the Cabibbo angle [38]. The expansion up to and including terms $\mathcal{O}(\lambda^5)$ is given by [39]:

$$V_{\rm CKM} = \begin{pmatrix} 1 - \lambda^2/2 - \lambda^4/8 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + A^2\lambda^5[1 - 2(\rho + i\eta)]/2 & 1 - \lambda^2/2 - \lambda^4(1 + 4A^2)/8 & A\lambda^2 \\ A\lambda^3[1 - (1 - \lambda^2/2)(\rho + i\eta)] & -A\lambda^2 + A\lambda^4[1 - 2(\rho + i\eta)]/2 & 1 - A^2\lambda^4/2 \end{pmatrix} + \mathcal{O}(\lambda^6), \tag{1.4}$$

where A, ρ and η are the remaining real parameters, all of order unity ($\lambda \approx 0.23$, $A \approx 0.80$, $\rho \approx 0.14$, $\eta \approx 0.34$ [12]).

The unitarity of the CKM matrix, $V_{\rm CKM}V_{\rm CKM}^\dagger = 1$, leads to a set of 9 equations,

$$\sum_{k\in\{u,c,t\}} V_{ki}V_{kj}^\star = \delta_{ij} \qquad (i,j \in \{d,s,b\}).$$

Six of them require the sum of three complex quantities to vanish. They are termed *triangular* because they define triangles in the complex plane. Remarkably, the area of all these triangles is a constant equals to $J_{CP}/2$. The symbol J_{CP} is the Jarlskog invariant, a combination of the CKM elements that quantifies the violation of the charge-parity (CP) symmetry in the SM, *i. e.*, the non-invariance of physics processes when all spatial coordinates are inverted and particles are replaced by their antiparticles. The CP symmetry is violated only if $J_{CP} \neq 0$, as confirmed by current measurements [12]: $J_{CP} = (2.884^{+0.253}_{-0.053}) \cdot 10^{-5}$. Any CP-violating quantity in the SM must be proportional to J_{CP}, reflecting the fact that a single complex phase appears in the 3×3 CKM matrix. This feature makes the SM implementation of CP violation (in principle) very predictive and all possible CP asymmetry measurements are correlated by their common origin from a single parameter of the theory.

One triangular equation of particular phenomenological interest is referred to as the unitarity triangle (UT):

$$V_{ud}V_{ub}^\star + V_{cd}V_{cb}^\star + V_{td}V_{tb}^\star = 0; \tag{1.5}$$

The UT can be rewritten in the normalized form:

$$R_t e^{-i\beta} + R_u e^{i\gamma} = 1 \tag{1.6}$$

where

$$R_t = \left|\frac{V_{td}V_{tb}^\star}{V_{cd}V_{cb}^\star}\right|,\ R_u = \left|\frac{V_{ud}V_{ub}^\star}{V_{cd}V_{cb}^\star}\right|,\ \beta = \arg\left(-\frac{V_{cd}V_{cb}^\star}{V_{td}V_{tb}^\star}\right),\ \gamma = \arg\left(-\frac{V_{ud}V_{ub}^\star}{V_{cd}V_{cb}^\star}\right), \tag{1.7}$$

are, respectively, two sides and two angles of the UT. The third side is the unity vector, and the third angle is $\alpha = \pi - \beta - \gamma = \arg(-V_{td}V_{tb}^\star/(V_{ud}V_{ub}^\star))$. Given the definition in Eq. (1.6), all the information related to the UT is encoded in one complex number:

$$\bar{\rho} + i\bar{\eta} = R_u e^{i\gamma} \tag{1.8}$$

which corresponds to the coordinates $(\bar{\rho},\bar{\eta})$ in the complex plane of the only nontrivial apex of the UT. One can assume that flavor-changing processes are fully described by the SM and can check the consistency of the various measurements with this assumption. The values of λ and A are known accurately from, respectively, $K \to \pi l\nu$ and $b \to cl\nu$ decays [8]:

$$\lambda = 0.2257 \pm 0.0010, \quad A = 0.814 \pm 0.022. \tag{1.9}$$

Then, one can express all the relevant observables as a function of the two remaining parameters, ρ and η (or equivalently, $\bar{\rho}$ and $\bar{\eta}$), and check whether there is a range in the (ρ,η) plane that is consistent with all measurements. The resulting constraints in the $(\bar{\rho},\bar{\eta})$ plane are shown in

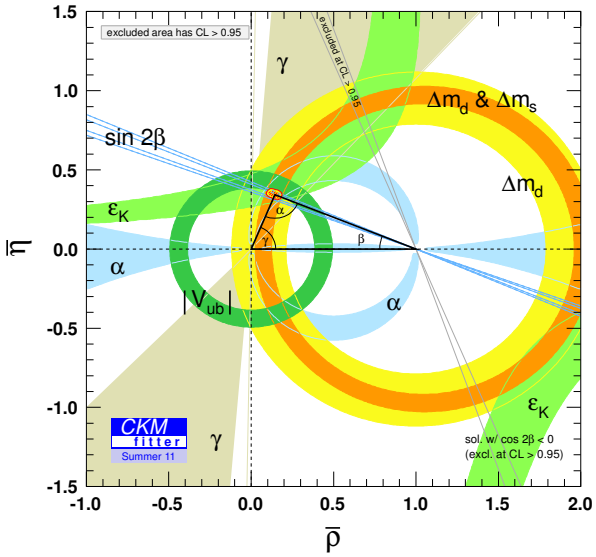


Figure 1.2: Constraints in the $(\bar{\rho}, \bar{\eta})$ plane. The red hashed region of the global combination corresponds to 68% CL.

Fig. 1.2. The consistency of the various constraints is impressive. Specifically, the following ranges for ρ and η can account for all the measurements [8]:

$$\rho = +0.135^{+0.031}_{-0.016}, \quad \eta = +0.349 \pm 0.017. \tag{1.10}$$

Given the very good overall consistency of the global fit, it is very likely that flavor violation and *CP* violation in flavor-changing processes are dominated by the CKM mechanism. Such remarkable success of the SM suggests that arbitrary new physics contributions in flavor-changing processes that occurs at tree-level are highly suppressed with respect to SM contributions [6]. A greater chance for discovering new physics effects might reside in the study of loop-mediated FCNC transitions, such as the ones that mediates neutral-mesons oscillations described in the next section.

1.3 B_s^0-$\bar{B}_s^0$ mixing bounds on new physics

Neutral mesons, with the exception of the pion, are subject to the *mixing*, *i. e.*, oscillations from particle to antiparticles through weak transitions that change the meson flavor by two units, $\Delta F = 2$. For example, the FCNC quark transition that underlies the B_s^0-$\bar{B}_s^0$ oscillations, $b \to \bar{b}s\bar{s}$, changes the beauty flavor, $\Delta B = 2$; Feynman diagrams describing these transitions are shown in Fig. 1.3. We now briefly recall the mixing phenomenology and then discuss the constraints imposed by measurements of B_s^0-$\bar{B}_s^0$ mixing observables on new physics using the effective Lagrangian, Eq. (1.1).

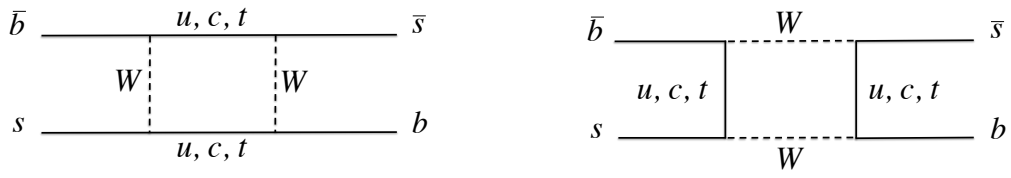


Figure 1.3: Feynman diagrams of $b \to \bar{b}\bar{s}s$ transitions that underlie the B_s^0-$\bar{B}_s^0$ mixing. These two diagrams are called *box* diagrams.

The system involves the meson states M^0 and $\bar{M}^0$, and all the states they can decay into.[3] Before the meson decays, the state is a coherent superposition of the two meson states, $|M^0(0)\rangle + |\bar{M}^0(0)\rangle$. The time evolution of the state is determined by a 2×2 hamiltonian $\mathbf{H}$, that can be written in terms of hermitian matrices, $\mathbf{M}$ and $\mathbf{\Gamma}$, as $\mathbf{H} = \mathbf{M} - \frac{\mathbf{i}}{\mathbf{2}}\mathbf{\Gamma}$. The matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ are associated with transitions via off-shell (dispersive) and on-shell (absorptive) intermediate states, respectively. Off-diagonal elements M_{12} and Γ_{12} (of $\mathbf{M}$ and $\mathbf{\Gamma}$, respectively), are associated with flavor-changing transitions $M^0 \rightleftarrows \bar{M}^0$. Since $\mathbf{H}$ is not diagonal, the meson states, M^0 and $\bar{M}^0$ are not mass eigenstates, and thus do not have well defined masses and widths. We denote the eigenvectors of $\mathbf{H}$ as the light and heavy eigenstates, M_L and M_H, respectively, with masses $m_L < m_H$. Such eigenstates are given by

$$|M_{L,H}\rangle = p|M^0\rangle \pm q|\bar{M}^0\rangle, \tag{1.11}$$

with the normalization $|p|^2 + |q|^2 = 1$. The time evolution of the mass eigenstates is governed by the two eigenvalues, $m_H - \frac{i}{2}\Gamma_H$ and $m_L - \frac{i}{2}\Gamma_L$, of which the real and imaginary parts represent their masses and decay-widths, respectively. Due to the *CPT* theorem [41], the mass of M^0 and $\bar{M}^0$ and their lifetime must be the same, thus, $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. The average mass and width are given by

$$\begin{aligned} m &= \frac{m_H + m_L}{2} = M_{11} = M_{22}, \\ \Gamma &= \frac{\Gamma_H + \Gamma_L}{2} = \Gamma_{11} = \Gamma_{22}. \end{aligned} \tag{1.12}$$

The mass difference Δm and the width difference $\Delta\Gamma$ are defined as follows:

$$\begin{aligned} \Delta m &= m_H - m_L, \\ \Delta\Gamma &= \Gamma_L - \Gamma_H. \end{aligned} \tag{1.13}$$

Note that Δm is positive by definition, while the sign of $\Delta\Gamma$ is to be determined experimentally. We choose the definition that corresponds to a positive decay width difference in the B_s^0 system, $\Delta\Gamma_s > 0$, in the SM. Its sign has been determined very recently by the LHCb Collaboration to be positive [42].

[3] The formalism of neutral-mesons mixing described here and in App. A is discussed in details in Ref. [40].

We define the phases $\phi_M = \arg(M_{12})$ and $\phi_\Gamma = \arg(-\Gamma_{12})$, and we introduce the relative phase between M_{12} and Γ_{12}:

$$\phi = \phi_M - \phi_\Gamma, \tag{1.14}$$

which is the phase of the interference between the mixing amplitude and the amplitude of all decays of M^0 and $\bar{M}^0$ into common final states. Then, one can write Δm and $\Delta\Gamma$ in terms of $|M_{12}|$, $|\Gamma_{12}|$ and ϕ. A simplified solution may be derived when $\Delta m \gg \Delta\Gamma$ and $|M_{12}| \gg |\Gamma_{12}|$, which is the case that empirically holds for B_s^0 mesons:

$$\Delta m = 2|M_{12}| + \mathcal{O}(o^2), \tag{1.15}$$

$$\Delta\Gamma = 2|\Gamma_{12}|\cos\phi + \mathcal{O}(o^2), \tag{1.16}$$

where $o = |\Gamma_{12}/M_{12}| \sim 10^{-3}$. We also need an approximate expression for q/p. It is convenient to define a small parameter:

$$a = \mathfrak{Im}(\Gamma_{12}M_{12}^\star) = \left|\frac{\Gamma_{12}}{M_{12}}\right|\sin\phi, \tag{1.17}$$

and then q/p becomes:

$$\frac{q}{p} = -e^{-i\phi_M}\left[1 - \frac{a}{2}\right] + \mathcal{O}(o^2). \tag{1.18}$$

We can predict meson-antimeson mixing and FCNC amplitudes within the SM and compare them with the available experimental information on Δm_q, $\Delta\Gamma_q$, and ϕ_q, thereby constraining the couplings of the $\Delta F = 2$ operators in Eq. (1.1) and derive bound, as for instance the ones in Tab. 1.1. The subscript q takes the value d to indicate an observable related to B^0-$\bar{B}^0$ mixing, and the value s for B_s^0-$\bar{B}_s^0$ mixing. We will focus on the B_s^0 mesons from now. Deviations from the SM $\Delta B = 2$ amplitude induced by new physics effects might be conveniently parametrized in terms of the shift induced in the modulo, $|\Delta_s|$, and in the *CP*-violating phase, $\phi_s^{\rm NP}$ [6]:

$$\Delta_s \equiv \frac{\langle B_s^0|\mathcal{L}_{\rm eff}|\bar{B}_s^0\rangle}{\langle B_s^0|\mathcal{L}_{\rm SM}|\bar{B}_s^0\rangle} = |\Delta_s|e^{2i\phi_s^{\rm NP}}, \tag{1.19}$$

such that $M_{12}^s = M_{12}^{s,\rm SM}\Delta_s$. Equation (1.19) is the basic equation for the search of new physics in the phenomenology of the mesons oscillations in the model independent approach. In B_s^0-$\bar{B}_s^0$ mesons oscillations, the magnitude of the SM amplitude is suppressed both by the GIM mechanism and by the hierarchical structure of the CKM matrix. The top quark give the dominant contribution in the box diagrams (Fig. 1.3):

$$A_{\rm SM}^{\Delta B=2} \propto \frac{G_F^2 m_t^2}{16\pi^2}(V_{ti}^\star V_{tj})^2\langle\bar{B}_s^0|O|B_s^0\rangle F\left(\frac{M_W^2}{m_t^2}\right), \tag{1.20}$$

where G_F is the Fermi constant of the weak interaction; m_t and M_W are the mass of the top and W boson, respectively; $\langle\bar{B}_s^0|O|B_s^0\rangle$ is the matrix element in terms of the $\Delta B = 2$ operator of the SM; F is a function of $\mathcal{O}(1)$ that gives the dependence on the masses of the top quark and the W boson, which are the virtual intermediate particles in this amplitude [6].

Calculation of Δm_s are proportional to $|A_{\rm SM}^{\Delta B=2}|$. Comparison of experimental data with SM expectation of Δm_s tells us that the magnitude of the new physics amplitude cannot be larger than the SM amplitude, since:

$$\Delta m_s = 2|M_{12}^{s,\rm SM}||\Delta_s|. \tag{1.21}$$

The first measurement of Δm_s was published by CDF in 2006: $\Delta m_s = 17.77 \pm 0.11(\text{stat}) \pm 0.07(\text{syst})\,\text{ps}^{-1}$ [43]. The current world average is $\Delta m_s = 17.69 \pm 0.08\,\text{ps}^{-1}$ [9]. Comparison of the experimental result with theoretical calculation, $\Delta m_s = 17.3 \pm 2.6\,\text{ps}^{-1}$ [13], does not allow large deviations in modulo with respect to the SM for the B_s^0-$\bar{B}_s^0$ mixing amplitude.[4] The constraint is more stringent if we consider the ratio $\Delta m_d/\Delta m_s$ in which the hadronic uncertainties cancel out to a large extent [13].

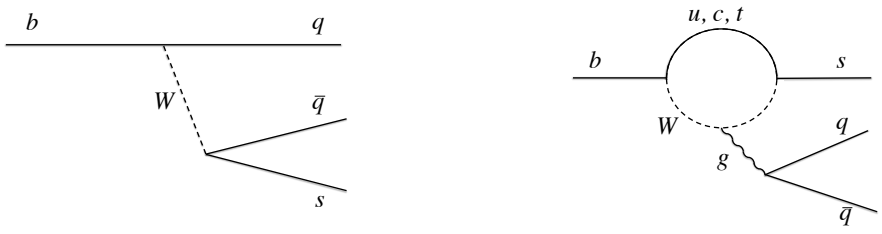
In a similar way, we can compare the prediction of the decay width difference $\Delta\Gamma_s$ with measurements. The determination of $\Delta\Gamma_s$ stems from the dispersive part of the B_s^0-$\bar{B}_s^0$ amplitude, Γ_{12}^s, (Eq. (1.16)) and is much more complicated than the calculation of Δm_s, because one can not integrate out at once all the particles inside the loop of the box diagram. A recent theoretical determination of $\Delta\Gamma_s$ yields to $0.087 \pm 0.021\,\text{ps}^{-1}$ in the SM [13]. First attempts to measure $\Delta\Gamma_s$ with the $B_s^0 \to J/\psi\,\phi$ decay were pursued at Tevatron by CDF and D0 since 2007, obtaining some hints of a nonzero value [44, 45]. The first nonzero measurement with a significance greater than 5σ was presented by LHCb in March 2012 [46], with $\Delta\Gamma_s = 0.116 \pm 0.019\,\text{ps}^{-1}$. The most recent result from D0 reads $\Delta\Gamma_s = 0.163 \pm 0.065\,\text{ps}^{-1}$ [19]; the latest result from CDF is presented in this thesis, while the previous CDF measurement is $\Delta\Gamma_s = 0.075 \pm 0.036\,\text{ps}^{-1}$ [18]. The current average, including also constrains from other decays, is $0.091 \pm 0.011\,\text{ps}^{-1}$ [9], in agreement with the SM expectation. We will discuss the new physics bound that might be obtained from these measurement.

Consider the quark transitions underlying decays common to B_s^0 and $\bar{B}_s^0$. We can schematically write each contribution of these decays in term of a_i amplitudes and CKM-matrix elements:

$$\underbrace{V_{cb}V_{cs}^\star}_{\mathcal{O}(\lambda^2)} a^T_{c\bar{c}s} + \underbrace{V_{ub}V_{us}^\star}_{\mathcal{O}(\lambda^4)} a^T_{u\bar{u}s} + \underbrace{V_{tb}^\star V_{ts}}_{\mathcal{O}(\lambda^2)} a^P_{s\bar{u}u} + \underbrace{V_{tb}^\star V_{ts}}_{\mathcal{O}(\lambda^2)} a^P_{s\bar{c}c} + \underbrace{V_{tb}^\star V_{ts}}_{\mathcal{O}(\lambda^2)} a^P_{s\bar{d}d} + \underbrace{V_{tb}^\star V_{ts}}_{\mathcal{O}(\lambda^2)} a^P_{s\bar{s}s} + \ldots \tag{1.22}$$

The apex T stands for a tree-mediated amplitude; the apex P stands for a loop-mediated transition named *penguin*; they are shown as Feynman diagrams in Fig. 1.4(a) and Fig. 1.4(b), respectively. The penguin transition is dominated by the t-quark contribution in the loop in the same way as the top dominates in the box diagram. Loop-mediated contributions with the exchange of the top quark can enter only in the off-shell part of the mixing amplitude. Being Γ_{12} the dispersive part of the mixing amplitude, the leading contribution is given by the two first term in Eq. (1.22), representing $b \to c\bar{c}s$ and $b \to u\bar{u}s$ tree transition, respectively. The former is Cabibbo-suppressed, $V_{cb}V_{cs}^\star \sim \mathcal{O}(\lambda^2)$, the latter doubly Cabibbo-suppressed, $V_{ub}V_{us}^\star \sim \mathcal{O}(\lambda^4)$.

[4] The theoretical uncertainty is currently dominated by uncertainties in the hadronic matrix calculation due to non perturbative-QCD contribution.



(a) (b)

Figure 1.4: Tree-level $b \to q\bar{q}s$ transition in (a) and penguin $b \to s\bar{q}q$ transition in (b). The symbol g in the penguin diagram stand for a gluon boson.

Then $b \to c\bar{c}s$ tree transitions dominate Γ_{12}^s, which are highly CKM constrained and disfavored for searching for new physics effects. Thus, we reasonably assume no NP effects in Γ_{12}^s. Some authors [47] have challenged this assumption by postulating new B_s^0 decay channels with large branching fractions, however, any competitive new decay mode will increase the total B_s^0 width which is severely constrained by precise measurements of the B_s^0 lifetime [9, 8, 13]. The only way for NP to affect $\Delta\Gamma_s$ is the phase ϕ_s, which can only lower $\Delta\Gamma_s$ compared to its SM value [13, 14]:

$$\Delta\Gamma_s = 2|\Gamma_{12}^{s,\mathrm{SM}}| \cos\left(\phi_s^{\mathrm{SM}} + \phi_s^{\mathrm{NP}}\right). \tag{1.23}$$

1.3.1 The B_s^0-$\bar{B}_s^0$ mixing phase

The phase ϕ_s is the difference of the phase of $\phi_M^s = \arg(M_{12}^s)$ and $\phi_\Gamma^s = \arg(-\Gamma_{12}^s)$ of Eq. (1.14):

$$\phi_s = \phi_M^s - \phi_\Gamma^s. \tag{1.24}$$

The two phases ϕ_M^s and ϕ_Γ^s depend on the convention of the CKM parameterization. However, the difference ϕ_s is phase-independent. In the SM its value is extremely small [13]:

$$\phi_s^{\mathrm{SM}} = -0.0038 \pm 0.0010. \tag{1.25}$$

This is caused by two effects [14]. The first stems from the fact discussed above that Γ_{12}^s is dominated by the decay $b \to c\bar{c}s$, and $(V_{cb}V_{cs})$ is close to mixing phase, $\phi_M^s = 2\arg(V_{tb}V_{ts}^\star)$. Second, the small correction to ϕ_Γ^s involving $V_{ub}V_{us}$ is further suppressed by a a factor m_u^2/m_b^2. In the search for a sizable new physics contributions to ϕ_s these doubly-Cabibbo suppressed transitions can safely be neglected, as we will do in the following. For this reason, it is customary to refer to new physics contribution to ϕ_s as new contribution to the mixing phase ϕ_M^s. Despite the strict bounds on $|A_{\mathrm{NP}}^{\Delta B=2}|$, new physics operator can participate in the B_s^0-$\bar{B}_s^0$ transitions, and enhance the phase ϕ_s with respect to the value in Eq. (1.25) [14]:

$$\phi_s = \phi_s^{\mathrm{SM}} + \phi_s^{\mathrm{NP}} \tag{1.26}$$

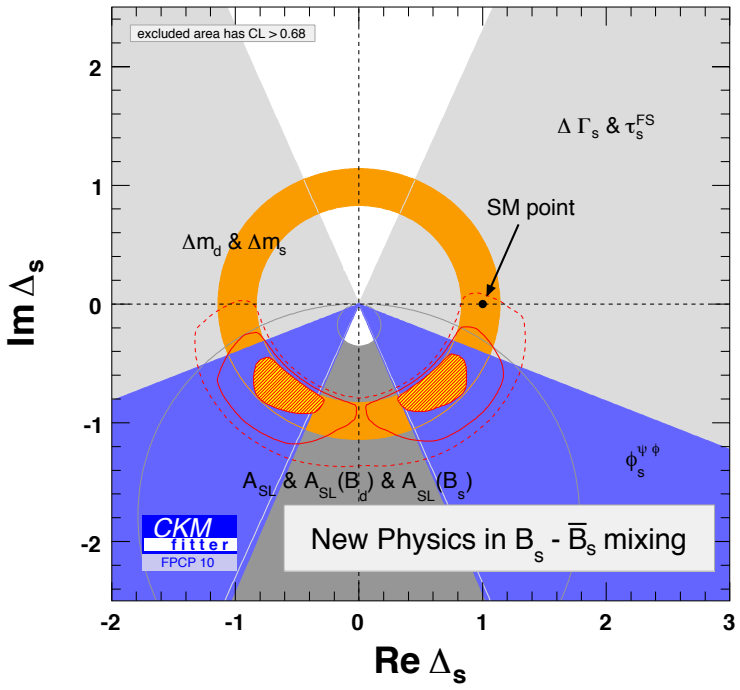


Figure 1.5: Bounds in the $(\mathfrak{Re}\Delta_s, \mathfrak{Im}\Delta_s)$ plane from measurements as of fall 2010 [12]. The orange circle represents the bounds on $|\Delta_s|$ given by Δm_s and Δm_d measurements; the blue bands are the constraints from ϕ_s; the light-grey band are derived from measurements of $\Delta\Gamma_s$ and of the B_s^0 lifetime, τ_s; the dark-grey band is given by measurements of $\mathcal{A}_{sl}^b$. The SM point is $(1, 0)$.

When we started the work described in this thesis, the phase ϕ_s was very poorly constrained by experiments [9]. A summary of the situation in fall 2010 is summarized by the plot in Fig. 1.5, which show the experimental bounds in the $(\mathfrak{Re}\Delta_s, \mathfrak{Im}\Delta_s)$ plane [12]. At that time, new physics in B_s^0-$\bar{B}_s^0$ mixing could accommodate the different deviations from the SM expectations, such as the measurement by the D0 Collaboration of the dimuon asymmetry. The dimuon asymmetry $\mathcal{A}_{\rm sl}^b$ is a linear combination of the semileptonic asymmetries in the B^0 and B_s^0 systems:

$$\mathcal{A}_{\rm sl}^b = C^d a_{\rm sl}^d + C^s a_{\rm sl}^s. \tag{1.27}$$

where the coefficients C^d and C^s depend on the mean mixing probabilities and the production fractions of B^0 and B_s^0 mesons, respectively. The other two parameters in Eq. (1.27) are the *CP* asymmetries in charged-current semileptonic decays of neutral B mesons, $B_{(s)}^0, \bar{B}_{(s)}^0 \to \ell^\pm X$:

$$a_{\rm sl}^q = \frac{\Gamma(\bar{B}_{(s)}^0 \to \ell^+ X) - \Gamma(B_{(s)}^0 \to \ell^- X)}{\Gamma(\bar{B}_{(s)}^0 \to \ell^+ X) + \Gamma(B_{(s)}^0 \to \ell^- X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} = a^q + \mathcal{O}((a^q)^2). \tag{1.28}$$

where a^q is the parameter of Eq. (1.17). If a^q is zero, there is no *CP* violation in mixing. In the SM [13]:

$$\mathcal{A}_{\rm sl}^{b,\rm SM} = (-0.0228^{+0.005}_{-0.006})\%, \tag{1.29}$$

which is negligible compared to present experimental sensitivity. The D0 measurement is largely deviating from the SM expectation with a significance of about 4σ [22]: $\mathcal{A}_{\rm sl}^b = (-0.787 \pm 0.172(\text{stat}) \pm 0.093(\text{syst}))\%$. Single measurements of $a_{\rm sl}^d$ [48] and $a_{\rm sl}^s$ [49, 50] are also provided with different methodologies, but the large uncertainties prevent a definitive conclusion. If the discrepancy is originated from new physics in the B_s^0 sector, it should be tightly correlated with ϕ_s:

$$a_{\rm sl}^s = \frac{|\Gamma_{12}^{s,\rm SM}|}{|M_{12}^{s,\rm SM}|}\frac{\sin(\phi_s^{\rm SM}+\phi_s^{\rm NP})}{|\Delta_s|} \tag{1.30}$$

The model independent fit of all flavor data in Fig. 1.5 is consistent with no new physics in B_s^0-$\bar{B}_s^0$ mixing within about 2 standard deviation, but there is still some room for sizable deviations from the SM expectations, being ϕ_s very poorly constrained. Adding experimental information on this phase is therefore mandatory. The mixing phase is accessible via the measurement of a time-dependent *CP* asymmetry as we are going to explain in the next section.

1.4 CKM phases from time-dependent *CP* asymmetries

As far as decays of B mesons are concerned, *CP* violation appears as a difference between the rates of a decay and its *CP* conjugate and is accounted for by the direct *CP* asymmetry:[5]

$$\mathcal{A}_{CP} = \frac{\Gamma(\bar{B}\to\bar{f}) - \Gamma(B\to f)}{\Gamma(\bar{B}\to\bar{f}) - \Gamma(B\to f)} = \frac{|\bar{A}_{\bar{f}}|^2 - |A_f|^2}{|\bar{A}_{\bar{f}}|^2 + |A_f|^2} \tag{1.31}$$

where f is the final state and $\bar{f}$ is its *CP*-conjugate states, and A_f and $\bar{A}_{\bar{f}}$ are the two decay amplitudes. Charged mesons can violate *CP* only in the decay. In the case of neutral mesons, $B_{(s)}^0$, *CP* violation can also occur in the mixing itself and in the interference between decay occurring with and without mixing. Those occurrences allow for additional opportunities to observe *CP* violation in neutral mesons.

The deviation of $|q/p|$ from one, *i. e.*, $a \neq 0$, is a measure of *CP* violation in the mixing, as shown by the expression of the semileptonic *CP* asymmetry of Eq. (1.28). In the case of the B_s^0 meson, the SM predict that *CP* violation in mixing can be neglected. Indeed, the SM value is $a_{\rm sl}^s = (0.0019 \pm 0.0003)\%$ [13]. We will assume $|q/p| = 1$ from now on, which means that q/p is determined by a pure phase, and focus on the case of the B_s^0 meson.

A third manifestation of *CP* violation in neutral meson decays is through the interference between mixing and decay. The key observable in this case is the time-dependent *CP* asymmetry

[5] We point to Ref. [40] for a complete description of *CP* violation in mesons decays and mixing.

of the decays into a final state f accessible to both B_s^0 and $\bar{B}_s^0$. We consider final states that are *CP* eigenstates:

$$CP|f\rangle = \eta|f\rangle, \tag{1.32}$$

where η can be ± 1 depending on the intrinsic *CP* parity of f. We write the decay amplitude as A_f and $\bar{A}_f$, respectively for the decay $B_s^0 \to f$ and $\bar{B}_s^0 \to f$. Neglecting *CP* violation in mixing, the time-dependent *CP* asymmetry is:

$$\begin{aligned} \mathcal{A}_{CP}(t) &= \frac{\Gamma(B_s^0 \to \bar{f}) - \Gamma(\bar{B}_s^0 \to f)}{\Gamma(B_s^0 \to \bar{f}) + \Gamma(\bar{B}_s^0 \to f)} \\ &\simeq \frac{\mathcal{C}_f \cos(\Delta m_s t) - \mathcal{S}_f \sin(\Delta m_s t)}{\cosh(\Delta\Gamma_s t/2) + \mathcal{S}'_f \sinh(\Delta\Gamma_s t/2)}. \end{aligned} \tag{1.33}$$

where:

$$\mathcal{C}_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \quad \mathcal{S}_f = \frac{2\mathfrak{Im}(\lambda_f)}{1 + |\lambda_f|^2}, \quad \mathcal{S}'_f = \frac{2\mathfrak{Re}(\lambda_f)}{1 + |\lambda_f|^2}, \tag{1.34}$$

and the key complex parameter that encodes *CP* violation is:

$$\lambda_f \equiv \eta \frac{q}{p} \frac{\bar{A}_f}{A_f}. \tag{1.35}$$

Hence, *CP* violation happens when λ_f either carries a nonzero phase or has a non-unitary modulus and Eq. (1.33) exhibits two different sources of *CP* violation: *direct CP* violation probed by $(1 - |\lambda_f|^2)$ (with $\mathcal{C}_f \simeq -\mathcal{A}_{CP}$); and *interference CP* violation, probed by $\mathfrak{Im}(\lambda_f) \neq 0$. The derivation of the time-dependent *CP* asymmetry is presented in App. A.

There are two types of phases that may appear in the amplitudes A_f and $\bar{A}_f$.[6] Complex parameters in any Lagrangian term that contributes to the amplitude will appear in complex conjugate form in the *CP* conjugate amplitude. Thus, their phases appear in A_f and $\bar{A}_f$ with *opposite signs* and these phases are *CP-odd*. In the SM, this phases are relative phases of the CKM matrix elements which enter in the quarks flavor-transitions, hence *CP*-odd phases are also called *weak phases*.

A second type of phases can appear in decay amplitudes even when the Lagrangian is real. They stem from possible contributions of intermediate on-shell states in the decay process. These phases are the same in A_f and $\bar{A}_{\bar{f}}$ and are therefore *CP-even*. One type of such phases are given by rescattering due to the strong interactions in the final state (*final-state-interaction*, FSI). For this reason these phases are called *strong phases*. The FSI allows the various final states of the weak decay to scatter elastically or inelastically via non-weak interaction. For instance, the total amplitude can include contributions from processes $B_s^0 \to f' \to f$, where the decay $B_s^0 \to f'$ is weak, and the state f' subsequently scatters into f via the strong (or electromagnetic) interaction. Strong phases are hard to calculate since they usually originate from non-perturbative QCD dynamics.

[6] There is a third kind of phases: spurious phases due to an arbitrary choice of phase convention, which do not originate from any dynamics. For simplicity, we set these unphysical phases to zero.

If there are two (or more) contributions to the decay amplitude A_f (with modulo a_i, weak phases ϕ_i, and strong phases δ_i), for instance, from two different Feynman diagrams contributing to the process:

$$\begin{aligned} A_f &= |a_1|e^{i(\phi_1+\delta_1)} + |a_2|e^{i(\phi_2+\delta_2)}, \\ \bar{A}_f &= \eta(|a_1|e^{i(-\phi_1+\delta_1)} + |a_2|e^{i(-\phi_2+\delta_2)}), \end{aligned} \tag{1.36}$$

then:

$$\lambda_f = -\eta e^{-2i\Phi_1} \frac{1 + re^{i(\Phi_1-\Phi_2)}e^{i\Delta}}{1 + re^{-i(\Phi_1-\Phi_2)}e^{i\Delta}} \tag{1.37}$$

where $r = a_2/a_1$, $\Phi_1 = \phi_1 - \phi_M^s$, and $\Phi_2 = \phi_1 - \phi_M^s$. If a_2 is much smaller than a_1, we obtain:

$$\mathcal{C}_f \simeq 2r \sin(\Phi_1 - \Phi_2) \sin\Delta, \tag{1.38}$$

$$\mathcal{S}_f \simeq \eta[\sin 2\Phi_1 - 2r \cos 2\Phi_1 \sin(\Phi_1 - \Phi_2) \cos\Delta] \tag{1.39}$$

$$\mathcal{S}'_f \simeq \eta[\cos 2\Phi_1 + 2r \cos 2\Phi_1 \sin(\Phi_1 - \Phi_2) \cos\Delta]. \tag{1.40}$$

The theoretical estimate of *CP* violation is then plagued by hadronic uncertainties of the strong phases and of r and the interference-*CP*violation observable does not provide a clean measurement of a single weak phase. For this reason, finding peculiar situations where these problems can be overcome is crucial. This occurs in decays that are dominated by a single weak phase, *i. e.*, when there is only one amplitude contributing to A_f,

$$\begin{aligned} A_f &= |a|e^{i\phi_A}e^{i\delta}, \\ \bar{A}_f &= \eta|a|e^{-i\phi_A}e^{i\delta}, \end{aligned} \tag{1.41}$$

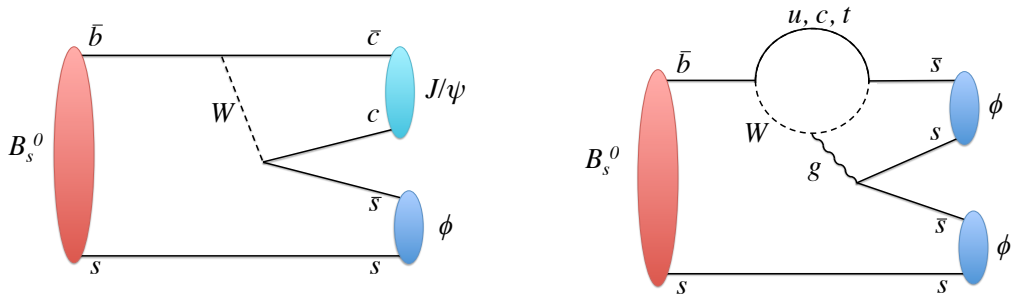
In that case, the calculation of λ_f is also free of hadronic uncertainties:

$$\lambda_f = -\eta e^{2i(\phi_M^s - \phi_A)}. \tag{1.42}$$

The crucial point is that both the moduli of the decay amplitudes and the FSI effects that those amplitudes contain cancel out in λ_f. This effectively eliminates the hadronic uncertainties from the computation of the *CP*-violating parameter λ_f. Note that $|\lambda_f|^2 = 1$, so that direct *CP* violation in Eq. (1.33) is canceled out. This is expected, since the decay amplitude has only one contribution. What really matter is only the phase:

$$\phi^f \equiv \arg(\lambda_f) = 2(\phi_M^s - \phi_A). \tag{1.43}$$

The phase ϕ^f can be expressed in term of the CKM matrix element and depending on the meson decays under analysis, the time-dependent *CP* asymmetries allow for measuring a specific angle of one of the unitary triangles. Note that the phase ϕ^f of Eq. (1.43) is different from the phase $2\phi_s$ of Eq. (1.24), because in general $\phi_A \neq \phi_\Gamma^s$. However, for particular decays the measurement of ϕ^f allow a clean measurement of ϕ_s and ϕ_M^s. For this reason, we analyze the form assumed by ϕ^f in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays in the following section.



(a) (b)

Figure 1.6: Leading amplitudes of the $B_s^0 \to J/\psi\phi$ decay (a) and of the $B_s^0 \to \phi\phi$ decay (b).

1.5 B_s^0-$\bar{B}_s^0$ mixing phase in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays

Neglecting *CP* violation in B_s^0-$\bar{B}_s^0$ mixing, we write the pure phase of q/p as follow:

$$\phi_M^s = \arg\left(\frac{q}{p}\right) = \arg\left(\frac{V_{ts}V_{tb}^\star}{V_{ts}^\star V_{tb}}\right). \tag{1.44}$$

The underlying quark transitions of $B_s^0 \to J/\psi\phi$ decays is the flavor-changing decay $b \to c\bar{c}s$, where the s quark of the B_s^0 meson doesn't take part in the weak transition, and acts as *spectator* (see Fig. 1.6(a)). Such amplitude can be parametrized as the sum of a leading tree contribution (Fig. 1.4(a) with $q\bar{q} = c\bar{c}$) and a second term from the $b \to sc\bar{c}$ penguin contribution, $a_p e^{i\theta} e^{\pm i\gamma}$, (Fig. 1.4(b) with $q\bar{q} = c\bar{c}$):

$$A_{c\bar{c}s}, \bar{A}_{c\bar{c}s} \propto \left[1 + \left(\frac{\lambda^2}{1-\lambda^2}\right) a_p e^{i\theta} e^{\pm i\gamma}\right]. \tag{1.45}$$

Naively, the parameter $a_p e^{i\theta}$ measures the ratio of penguin- to tree-diagram-like amplitude and is loop-suppressed [14]. The weak phase factor $e^{i\gamma}$ associated to a_p is strongly Cabibbo suppressed by λ^2. Hence, we take only the leading tree contribution to the $b \to c\bar{c}s$. Since only one amplitude contribute to the decay, there is no direct *CP* violation. Thus, we set $|A_{c\bar{c}s}| = |\bar{A}_{c\bar{c}s}|$ and we have that

$$\frac{\bar{A}_{c\bar{c}s}}{A_{c\bar{c}s}} = \frac{V_{cs}^\star V_{cb}}{V_{cs} V_{cb}^\star} \tag{1.46}$$

is given by a pure phase. Putting together the expression of $\bar{A}_{c\bar{c}s}/A_{c\bar{c}s}$ and q/p, we have

$$\phi^{c\bar{c}s} \equiv \arg(\lambda_{c\bar{c}s}) = \arg\left(\frac{V_{ts}V_{tb}^\star}{V_{ts}^\star V_{tb}} \frac{V_{cs}^\star V_{cb}}{V_{cs} V_{cb}^\star}\right) = -2\beta_s \tag{1.47}$$

The phase β_s is the small angle stemming from the following unitarity equation

$$V_{us}V_{ub}^\star + V_{cs}V_{cb}^\star + V_{ts}V_{tb}^\star = 0, \tag{1.48}$$

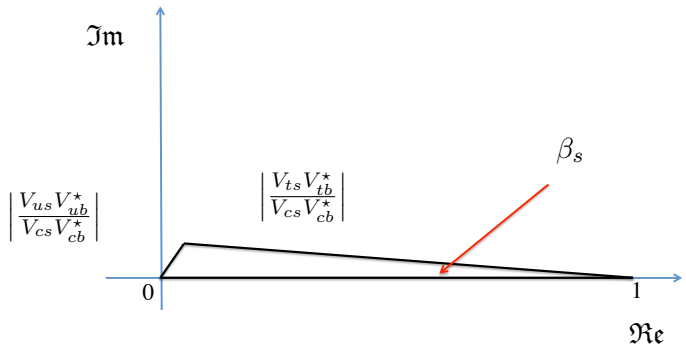


Figure 1.7: Graphical representation of the unitary triangle relevant for *CP* violation in the B_s^0 sector given by Eq. (1.48) normalized to $V_{cs}V_{cb}^\star$; the squashed angle is β_s. The angles are not to scale.

which represents the squashed triangle of Fig. 1.7, once it is normalized to $V_{cs}V_{cb}^\star$. In analogy of the angle β of the UT (Eq. (1.5)), β_s is defined as

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^\star}{V_{cs}V_{cb}^\star}\right) \tag{1.49}$$

and, from the experimental constraints on the CKM-matrix elements, takes the value [13]:

$$2\beta_s^{\rm SM} = 0.0363^{+0.0016}_{-0.0015}. \tag{1.50}$$

These definitions can be related to ϕ_s of Eq. (1.26). Both ϕ_s and β_s are predicted to be small in the SM. If new physics is present in B_s^0-$\bar{B}_s^0$ mixing, it will contribute to both phases, and they can be written as combinations of the SM and NP contributions. According to the chosen phase conventions [14]:

$$\phi_s = \phi_s^{\rm SM} + \phi_s^{\rm NP} \quad \text{and} \quad 2\beta_s = 2\beta_s^{\rm SM} - \phi_s^{\rm NP}. \tag{1.51}$$

With the current experimental sensitivity, the SM contributions $\phi_s^{\rm SM}$ and $2\beta_s^{\rm SM}$ can be treated as negligibly small, and the new physics phase would be expected to dominate. In this case, the measured phases can be related as:

$$2\beta_s \simeq -\phi_s. \tag{1.52}$$

Therefore, we will refer to the measurement of the phase β_s as the measurement of new physics contribution to the mixing phase. We will describe in the next chapter how we can access to β_s through the time-evolution of $B_s^0 \to J/\psi\phi$ decays.

Additional information on the origin of the new physics contribution to ϕ_s can be inferred in principle by analyzing the time-dependent *CP* asymmetry of $B_s^0 \to \phi\phi$ decays. In the SM the amplitude of the $B_s^0 \to \phi\phi$ decay is dominated by the penguin amplitude shown in Fig. 1.6(b).[7]

[7] Again, the s quark of the B_s^0 is a spectator with respect to the weak transition.

The decay amplitude can be decomposed in three contributions, a_q, one for each quark participating in the loop ($q = u, c, t$), as follow [15]:

$$\begin{aligned} A_{s\bar{s}s} &\propto V_{tb}^{\star}V_{ts}a_t + V_{cb}^{\star}V_{cs}a_c + V_{ub}^{\star}V_{us}a_u \\ &= \underbrace{V_{tb}^{\star}V_{ts}}_{\mathcal{O}(\lambda^2)}a_{tc} + \underbrace{V_{ub}^{\star}V_{us}}_{\mathcal{O}(\lambda^4)}a_{uc}, \end{aligned} \tag{1.53}$$

where in the second line, the unitarity of the CKM matrix is used to eliminate the c-quark contribution, and we have defined $a_{tc} \equiv a_t - a_c$ and $a_{uc} \equiv a_u - a_c$. The a_{ac} amplitude is doubly-Cabbibo suppressed, $V_{ub}^{\star}V_{us} \sim \mathcal{O}(\lambda^4)$, and we will neglect it. Considering only the t-quark amplitude in $\mathcal{A}_{s\bar{s}s}$, there cannot be *CP* violation in the decay and $\bar{A}_{s\bar{s}s}/A_{s\bar{s}s}$ is a pure phase, being $|A_{s\bar{s}s}| = |\bar{A}_{s\bar{s}s}|$. We have:

$$\frac{\bar{A}_{s\bar{s}s}}{A_{s\bar{s}s}} = \frac{V_{ts}^{\star}V_{tb}}{V_{ts}V_{tb}^{\star}} \tag{1.54}$$

Putting together this expression with q/p in Eq. (1.44):

$$\phi^{s\bar{s}s} \equiv \arg(\lambda_{s\bar{s}s}) = \underbrace{\arg\left(\frac{V_{ts}V_{tb}^{\star}}{V_{ts}^{\star}V_{tb}}\right)}_{\phi_M^{s,\mathrm{SM}}} + \underbrace{\arg\left(\frac{V_{ts}^{\star}V_{tb}}{V_{ts}V_{tb}^{\star}}\right)}_{-\phi_M^{s,\mathrm{SM}}} = 0 \tag{1.55}$$

Thus, there is no *CP*-violation in the $B_s^0 \to \phi\phi$ decays. The presence of subleading transitions in the decay amplitude with the same CKM coupling but different QCD-topology with respect to the $b \to s\bar{s}s$ penguin-amplitude doesn't spoil the conclusion of no *CP*-violation [51].[8]

Now, we can have the following scenarios. New physics is present in $\Delta B = 2$ transition but not in $\Delta B = 1$. In this case, the phase ϕ_M^s will be altered, $\phi_M^s = \phi_M^{s,\mathrm{SM}} + \phi_M^{s,\mathrm{NP}}$, while the phase of the decay remains $\phi_f = \arg(\bar{A}_{s\bar{s}s}/A_{s\bar{s}s}) = -\phi_M^{s,\mathrm{SM}}$, then:

$$\phi^{s\bar{s}s} = \phi_M^{s,\mathrm{NP}}. \tag{1.56}$$

If there is new physics for the $\Delta B = 1$ penguin decay, then the interfering amplitudes are the SM decay amplitude, $A_{s\bar{s}s}^{\mathrm{SM}}$, and the new physics decay amplitude, $A_{s\bar{s}s}^{\mathrm{NP}}$: we can have an alteration of $\lambda_{s\bar{s}s}$ through Eq. (1.37), resulting in $\phi^{s\bar{s}s} \neq 0$ (and additional strong phases), independently on what happens to the mixing phase. In any case, the crucial point is the following: we have *CP* violation in $B_s^0 \to \phi\phi$ only if a new physics amplitude contributing either in the decay or in the mixing (or both) is present with a different weak phase with respect to the SM amplitudes.

The small data samples available for this decay mode, which suffer a lower decay rate with respect to the $B_s^0 \to J/\psi\phi$ decays, do not allow for a suited analysis for measuring $\phi^{s\bar{s}s}$. However, we can infer the presence of new physics through a measurement of *CP*-violating asymmetries that are expected to vanish in the SM, given the above arguments. Such *CP*-violating asymmetry stems from the rich dynamic of $B_s^0 \to \phi\phi$ decays, which is described in the next chapter.

[8] Some phenomenological models would demand for sub-leading amplitudes such as the *penguin-annihilation* [52, 51] to explain the hierarchy of the polarization amplitudes measured in $B_s^0 \to \phi\phi$ [2] decays and other $b \to s$ penguin-dominated decays of the B^0 and B_s^0 mesons. Polarization amplitudes are defined in details in Sect. 2.1.

Chapter 2

Phenomenology of $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays

In this chapter we discuss the phenomenology of the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays. We present the angular distributions of the decays and the time-development of the angular amplitudes which provides the access to the B_s^0-$\bar{B}_s^0$ mixing phase, and we introduce the two time-integrated asymmetries of $B_s^0 \to \phi\phi$ decays that allow to measure the *CP* violation in this decay mode. Finally, we present the experimental strategy of the measurements presented in this dissertation.

2.1 Angular amplitudes of $B \to V_1 V_2$ decays

The nomenclature $B \to V_1 V_2$ stands for the class of B-meson decays into two vector mesons, which are spin-1 particles. Both the $B_s^0 \to J/\psi\phi$ and the $B_s^0 \to \phi\phi$ decays belong to this category. The conservation of the angular momentum implies that there are three allowed states of relative angular momentum for the V_1V_2 pair, $L = 0, 1, 2$. Thus, the overall decay amplitude is decomposed in three independent amplitudes, which can be expressed in terms of three equivalent bases: the basis of *partial waves* (S, P, D), the *transversity* basis $(A_0, A_\parallel, A_\perp)$, and the *helicity* basis (H_0, H_+, H_-). Partial waves are directly related to the states of relative angular momentum: the S-wave is the $L = 0$ state, the P-wave is the $L = 1$ state, and the the D-wave is the $L = 2$ state. The transversity basis is convenient for identifying the *CP* parity of the final state. The transversity is the projection of the spin in the plane orthogonal to the particle's momentum; for a massive meson, it corresponds to a linear polarization: A_0 is the longitudinal polarization, while $A_\parallel$ and $A_\perp$ are the transverse polarization with the linear polarization vectors parallel and perpendicular to each others, respectively. The helicity basis is useful to express the amplitude in terms of the helicity states, $\lambda = 0, \pm 1$. The helicity is the projection of the spin along the momentum vector of a particle. For a massive meson, it corresponds to a right/left circular polarization ($\lambda = \pm 1$) and one longitudinal polarization ($\lambda = 0$).

Consider the $B \to V_1 V_2$ decay followed by the decays of the intermediate vector states, $V_1 \to P_1^+ P_1^-$ and $V_2 \to P_2^+ P_2^-$. In the rest frame of $V_{1(2)}$ we identify two decay angles, $\vartheta_{1(2)}$ and $\varphi_{1(2)}$, as the polar and azimuthal angles of the $P_{1(2)}^+$ momentum vector (see Fig. 2.1).

The angular distribution of the decay is written as [53]:

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\vartheta_1 \mathrm{d}\cos\vartheta_2 \mathrm{d}\varphi_1 \mathrm{d}\varphi_2} \propto \left| \sum_{\lambda=0,\pm1} H_\lambda D^{1\star}_{\lambda,0}(\varphi_1, \vartheta_1, 0) D^{1\star}_{\lambda,0}(\varphi_2, \vartheta_2, 0) \right|^2 \tag{2.1}$$

where λ is the helicity of the V_i particle (the helicity of V_1 and V_2 is the same because of angular momentum conservation); the D-function is defined in terms of the Wigner d-function, $D^j_{mn}(\alpha, \beta, \gamma) = e^{-im\alpha} d^j_{mn}(\beta) e^{-in\gamma}$ [8]. In Eq. (2.1), H_λ are the *helicity amplitudes*. The helicity states are not CP eigenstates, being the helicity a pseudoscalar which changes sign under the parity transformation. Therefore, the $\lambda = \pm 1$ amplitudes do not have definite CP quantum number, but their combination, such as:

$$\begin{aligned} A_0 &\equiv H_0 \\ A_\parallel &\equiv \frac{1}{\sqrt{2}} (H_+ + H_-) , \\ A_\perp &\equiv \frac{1}{\sqrt{2}} (H_+ - H_-) , \end{aligned} \tag{2.2}$$

have definite CP parity [54]. The amplitudes defined by Eq. (2.2) are the *transversity* amplitudes. In the case of J/ψ and ϕ vector mesons, both V_1 and V_2 have intrinsic quantum numbers $C = -1$ and $P = -1$; then, the A_0 and the $A_\parallel$ amplitudes are CP-even states, while the $A_\perp$ is a CP-odd state. It can be proved by relating the transversity amplitudes to the partial wave basis:

$$\begin{aligned} A_0 &= -\sqrt{\frac{1}{3}} S + \sqrt{\frac{2}{3}} D, \\ A_\parallel &= \sqrt{\frac{2}{3}} S + \sqrt{\frac{1}{3}} D. \\ A_\perp &= P. \end{aligned} \tag{2.3}$$

Indeed, the S and D states are CP-even, and the P state is CP-odd, because of the relation:

$$CP|V_1 V_2\rangle = (-1)^L |V_1 V_2\rangle, \tag{2.4}$$

In particular:

$$\begin{aligned} (CP) A_0 &= \bar{A}_0, \\ (CP) A_\parallel &= \bar{A}_\parallel, \\ (CP) A_\perp &= -\bar{A}_\perp. \end{aligned} \tag{2.5}$$

and $|A_0| = |\bar{A}_0|$, $|A_\parallel| = |\bar{A}_\parallel|$, $|A_\perp| = |\bar{A}_\perp|$ if CP is conserved in the decay.

Whatever basis is chosen, the disentanglement of the amplitudes along with their interference are inferred through the analysis of angular distributions of the decay's product of the two vector mesons. In the following sections the angular distributions in the different basis are presented. The angular distributions of a $B \to V_1 V_2$ decay can be written by using two basis for the

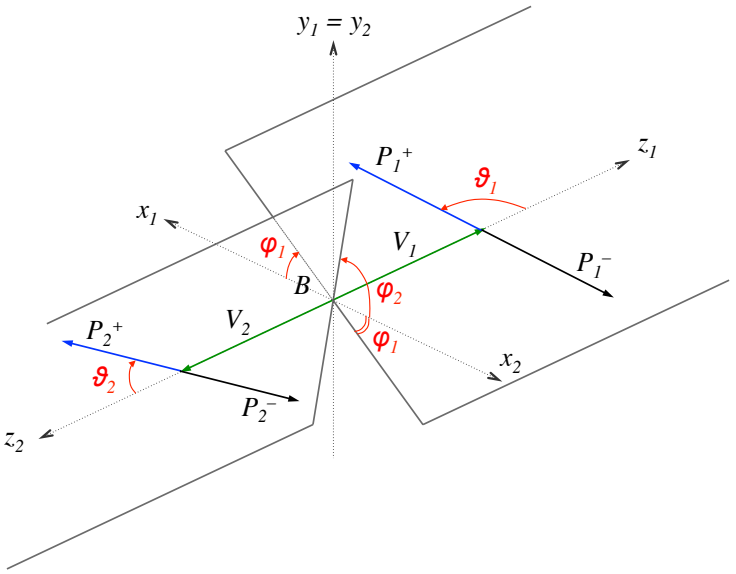


Figure 2.1: Graphical representation of the helicity angles ϑ_1, ϑ_2 and $\varphi = \varphi_1 + \varphi_2$.

angles' definition: the helicity angles basis and the transversity angles basis. Despite the same nomenclature of the amplitudes bases, the angles convention is independent of the choice of the amplitudes, *i. e.* one can use the transverse amplitudes to express the polarization of the decay and the helicity angles to give angular distributions (and viceversa), since there exist trivial transformations that relate the different bases, such as the one in Eq. (2.3). Since it is customary to express the measurement of the polarization of a decay by using the transversity amplitudes, in the following we relate the angular distributions to such amplitudes both in the case of the helicity and the transversity angles. In particular, it is more convenient to describe the $B_s^0 \to \phi\phi$ angular distributions by use of the helicity angles, because such decay involves two identical vector mesons, and the helicity basis is symmetric under the exchange $V_1 \leftrightarrow V_2$. On the other hand, the $B_s^0 \to J/\psi\,\phi$ angular distributions are traditionally described with transversity angles.

2.1.1 Angles bases

The helicity angles are defined as [2]:

$$\omega \equiv (\vartheta_1, \vartheta_2, \varphi), \tag{2.6}$$

where ϑ_1 and ϑ_2 were introduced in the previous section, entering Eq. (2.1), as the polar angles of the P_i^+ momentum in the $V_i \to P_i^+ P_i^-$ decays. The momenta of the two $P_i^+ P_i^-$ pairs form two decay planes intersecting at a straight line given by the momentum vector of V_i. The angle φ is the angle between the decay plane of the $P_i^+ P_i^-$ pairs, and it is defined as the sum of the azimuthal angles $\varphi \equiv \varphi_1 + \varphi_2$ (Fig. 2.1).

i	$\overline{(K)}_i$	$h_i^{J/\psi}(\cos\vartheta_1, \cos\vartheta_2, \varphi)$
1	$\lvert \overline{(A)}_0 \rvert^2$	$\frac{9}{32\pi} 4 \sin^2\vartheta_1 \cos^2\vartheta_2$
2	$\lvert \overline{(A)}_\parallel \rvert^2$	$\frac{9}{32\pi} \sin^2\vartheta_2 (\sin^2\varphi + \cos^2\vartheta_1 \cos^2\varphi)$
3	$\lvert \overline{(A)}_\perp \rvert^2$	$\frac{9}{32\pi} \sin^2\vartheta_2 (\cos^2\varphi + \cos^2\vartheta_1 \sin^2\varphi)$
4	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}_\parallel^\star)$	$-\frac{9}{32\pi} 2 \sin^2\vartheta_1 \sin^2\vartheta_2 \sin 2\varphi$
5	$\mathfrak{Re}(\overline{(A)}_\parallel \overline{(A)}_0^\star)$	$\frac{9}{32\pi} \sqrt{2} \sin 2\vartheta_1 \sin 2\vartheta_2 \cos\varphi$
6	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}_0^\star)$	$\frac{9}{32\pi} \sin 2\vartheta_1 \sin 2\vartheta_2 \sin\varphi$

Table 2.1: Angular functions in terms of helicity angles and corresponding transversity amplitudes for the $B_s^0 \to J/\psi\phi$ decay entering Eq. (2.7).

From Eq. (2.1), we can derive the angular dependence with helicity angles and helicity amplitudes. Then, using Eq. (2.3) we express the decay rate in terms of the transversity amplitudes. The angular distributions depend on the spins of the final state particles $P_i^+ P_i^-$. Considering the case of the $B_s^0 \to J/\psi\phi$ decay with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ decays, we obtain [26]:[1]

$$\frac{1}{\overline{(\Gamma)}} \frac{\mathrm{d}^3\,\overline{(\Gamma)}(\overline{(B)}_s^0 \to J/\psi\phi)}{\mathrm{d}\cos\vartheta_1 \mathrm{d}\cos\vartheta_2 \mathrm{d}\varphi} = \frac{\sum_{i=1}^6 \left[\overline{(K)}_i \, h_i^{J/\psi}(\cos\vartheta_1, \cos\vartheta_2, \varphi) \right]}{\lvert \overline{(A)}_0 \rvert^2 + \lvert \overline{(A)}_\parallel \rvert^2 + \lvert \overline{(A)}_\perp \rvert^2}, \tag{2.7}$$

and the $\overline{(K)}_i$ and $h_i^{J/\psi}(\cos\vartheta_1, \cos\vartheta_2, \varphi)$ terms are given in Tab. 2.1. We have considered as ϑ_1 the angle corresponding to the μ^+, and ϑ_2 the angle corresponding to the K^+. If there are only pseudoscalar mesons in the final states particles, as the case of the $\phi \to K^+K^-$ decay in the $B_s^0 \to \phi\phi$ decay, we obtain [26]:

$$\frac{1}{\overline{(\Gamma)}} \frac{\mathrm{d}^3\,\overline{(\Gamma)}(\overline{(B)}_s^0 \to \phi\phi)}{\mathrm{d}\cos\vartheta_1 \mathrm{d}\cos\vartheta_2 \mathrm{d}\varphi} = \frac{\sum_{i=1}^6 \left[\overline{(K)}_i \, h_i(\cos\vartheta_1, \cos\vartheta_2, \varphi) \right]}{\lvert \overline{(A)}_0 \rvert^2 + \lvert \overline{(A)}_\parallel \rvert^2 + \lvert \overline{(A)}_\perp \rvert^2}, \tag{2.8}$$

with $\overline{(K)}_i$ and $h_i(\cos\vartheta_1, \cos\vartheta_2, \varphi)$ terms given in Tab. 2.2.

[1] From now on, we use $\overline{(\text{—})}$ over a symbol x as a compact notation for indicating both x and its *CP*-conjugate $\bar{x}$.

i	$\overline{(K)}_i$	$h_i(\cos\vartheta_1, \cos\vartheta_2, \varphi)$
1	$\lvert \overline{(A)}_0 \rvert^2$	$\frac{9}{32\pi} 4 \cos^2\vartheta_1 \cos^2\vartheta_2$
2	$\lvert \overline{(A)}_\parallel \rvert^2$	$\frac{9}{32\pi} \sin^2\vartheta_1 \sin^2\vartheta_2 (1 + \cos 2\varphi)$
3	$\lvert \overline{(A)}_\perp \rvert^2$	$\frac{9}{32\pi} \sin^2\vartheta_1 \sin^2\vartheta_2 (1 - \cos 2\varphi)$
4	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}^\star_\parallel)$	$-\frac{9}{32\pi} 2 \sin^2\vartheta_1 \sin^2\vartheta_2 \sin 2\varphi$
5	$\mathfrak{Re}(\overline{(A)}_\parallel \overline{(A)}^\star_0)$	$\frac{9}{32\pi} \sqrt{2} \sin 2\vartheta_1 \sin 2\vartheta_2 \cos\varphi$
6	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}^\star_0)$	$\frac{9}{32\pi} \sin 2\vartheta_1 \sin 2\vartheta_2 \sin\varphi$

Table 2.2: Angular functions in terms of helicity angles and corresponding transversity amplitudes for the $B_s^0 \to \phi\phi$ entering Eq. (2.8).

In the transversity basis, illustrated in Fig. 2.2, the angles

$$\Omega \equiv (\Theta, \Phi, \Psi) \tag{2.9}$$

are defined in two different frames as follows [18]. Consider the $B_s^0 \to J/\psi\phi$ decay as example. The first two angles are calculated in the rest frame of the J/ψ, and the third in the rest frame of the ϕ. In the rest frame of the J/ψ, the ϕ meson direction defines the x-axis. The plane of the J/ψ and the K^+K^- pair defines the xy-plane with a positive projection of the K^+ momentum along the y-axis, $p_y(K^+) > 0$. The z-axis, defined in order to have a right-handed frame, is the transverse axis which gives the name "transversity" to the basis; indeed, the transversity is the particle's spin projection along such axis. Then, we define Θ and Φ, respectively, the azimuthal and polar angles of the μ^+ momentum vector in the J/ψ rest frame; the third angle, Ψ, is the angle between the K^+ momentum vector and the vector opposite to the J/ψ momentum vector in the ϕ meson rest frame.

It is possible to relate the transversity angles to the helicity angles. Considering again the $B_s^0 \to J/\psi\phi$ as example, we have $\Psi = \vartheta_2$, and [54]:

$$\begin{cases} \sin\vartheta_1 \sin\varphi = -\cos\Theta, \\ \sin\vartheta_1 \cos\varphi = -\sin\Theta \sin\Phi, \\ \cos\vartheta_1 = \sin\Theta \cos\Phi. \end{cases} \tag{2.10}$$

Then, from Equation (2.7) we can derive the angular distributions of the $B_s^0 \to J/\psi\phi$ decay in terms of transversity angles and transversity amplitudes, which reads:

$$\frac{1}{\overline{(\Gamma)}} \frac{\mathrm{d}^3 \overline{(\Gamma)}(\overline{(B)}^0_s \to J/\psi\phi)}{\mathrm{d}\cos\Theta \, \mathrm{d}\Phi \, \mathrm{d}\cos\Psi} = \frac{\sum_{i=1}^6 \left[\overline{(K)}_i \, f_i(\cos\Theta, \Phi, \cos\Psi) \right]}{\lvert \overline{(A)}_0 \rvert^2 + \lvert \overline{(A)}_\parallel \rvert^2 + \lvert \overline{(A)}_\perp \rvert^2}. \tag{2.11}$$

with the K_i and $f_i(\cos\Theta, \Phi, \cos\Psi)$ terms reported in Tab. 2.3 [44]. In Fig. 2.3 we show some examples of transversity angles distribution in the $B_s^0 \to J/\psi\phi$ decay for three sets of polarization amplitudes, illustrating how the distribution of these observables strongly depend on the underlying physics parameters.

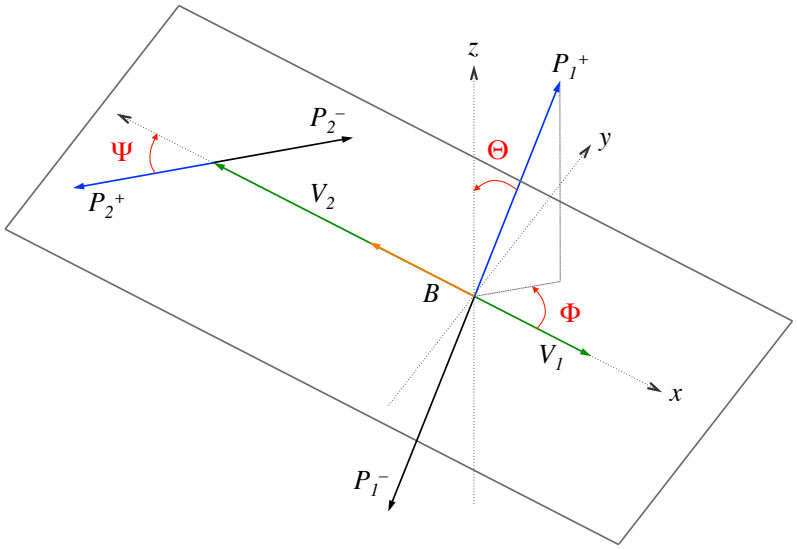
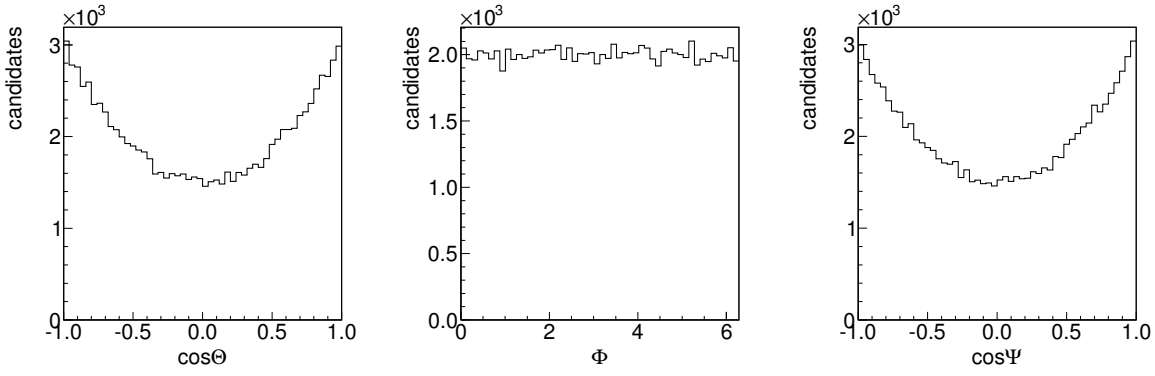


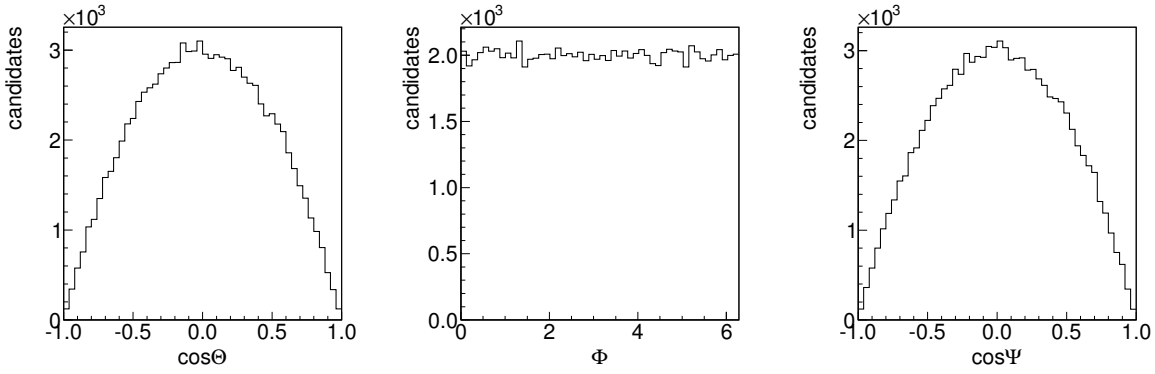
Figure 2.2: Graphical representation of the transversity angles Θ, Φ and Ψ.

i	$\overline{(K)}_i$	$f_i(\cos\Theta, \Phi, \cos\Psi)$
1	$\lvert\overline{(A)}_0\rvert^2$	$\frac{9}{32\pi} 2\cos^2\Psi(1-\sin^2\Theta\cos^2\Phi)$
2	$\lvert\overline{(A)}_\parallel\rvert^2$	$\frac{9}{32\pi}\sin^2\Psi(1-\sin^2\Theta\sin^2\Phi)$
3	$\lvert\overline{(A)}_\perp\rvert^2$	$\frac{9}{32\pi}\sin^2\Psi\sin^2\Theta$
4	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}^\star_\parallel)$	$-\frac{9}{32\pi}\sin^2\Psi\sin 2\Theta\sin\Phi$
5	$\mathfrak{Re}(\overline{(A)}_\parallel \overline{(A)}^\star_0)$	$\frac{9}{32\pi}\frac{\sqrt{2}}{2}\sin 2\Psi\sin^2\Theta\sin 2\Phi$
6	$\mathfrak{Im}(\overline{(A)}_\perp \overline{(A)}^\star_0)$	$\frac{9}{32\pi}\frac{\sqrt{2}}{2}\sin 2\Psi\sin 2\Theta\cos\Phi$

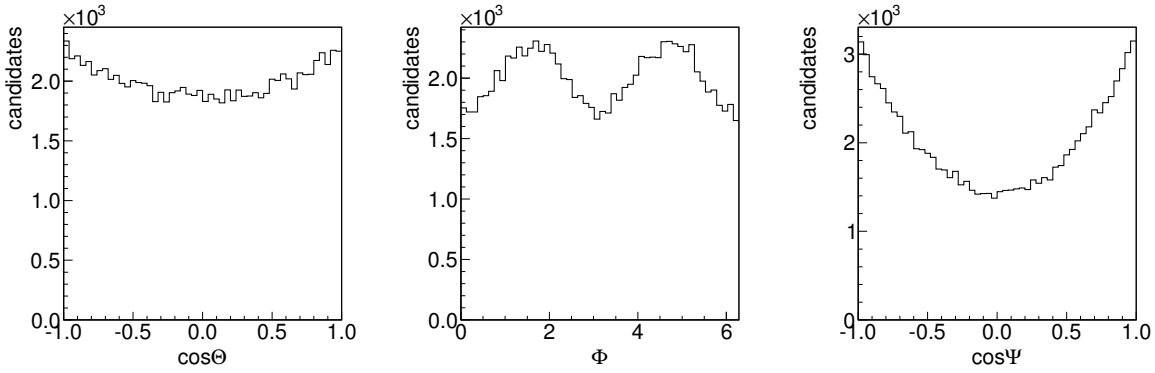
Table 2.3: Angular functions in terms of transversity angles and corresponding transversity amplitudes for the $B_s^0 \to J/\psi\phi$ decay entering Eq. (2.11).



(a)



(b)



(c)

Figure 2.3: Example of transversity angles distributions. In (a) the *CP*-odd amplitude $A_\perp$ is set zero, while both $|A_0|^2$ and $|A_\parallel|^2$ are set to 50%. In (b) the *CP*-even amplitudes A_0 and $A_\perp$ are set to zero. In (c) the amplitudes and strong phases are set to the value measured in Ref. [18]: $|A_0|^2 = 0.524$, $|A_\parallel|^2 = 0.231$, $\delta_\perp = 2.95$, and we take $\delta_\parallel = \pi$. Note that the angle Φ has a non-uniform distribution if there is an interference between the *CP*-even and *CP*-odd amplitudes.

2.2 The B_s^0-$\bar{B}_s^0$ mixing phase in the amplitudes time-evolution

We now consider the time evolution of the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays. We have to account for the evolution of the B_s^0 meson that exhibits the B_s^0-$\bar{B}_s^0$ mixing along with the decay transition to the final state.

The time evolution is independent of the angular distributions, and it is expressed by a time dependence of the polarization amplitudes, *i. e.* the K_i terms of Equations 2.7, 2.8 and 2.11:

$$K_i \to K_i(t),$$

where t is the decay-time. The transversity basis is particularly useful for expressing the time evolution, since it has definite *CP* properties. By using Eq. (A.3) and (A.5) in App. A and considering the *CP* parity of each transversity amplitudes, one can derive the time development for each $K_i(t)$ term. We can write the decay rate of an initially produced $\overset{(-)}{B}{}^0_s$ meson as a function of the decay time and transversity angles as derived in Ref. [15]:

$$\frac{1}{\overset{(-)}{\Gamma}} \frac{\mathrm{d}^4 \overset{(-)}{\Gamma}(\overset{(-)}{B}{}^0_s \to V_1V_2)}{\mathrm{d}t\, \mathrm{d}\cos\Theta\, \mathrm{d}\Phi\, \mathrm{d}\cos\Psi} = \frac{\sum_{i=1}^{6} \left[\overset{(-)}{K}{}_i(t)\, f_i(\cos\Theta, \Phi, \cos\Psi) \right]}{|\overset{(-)}{A}{}_0|^2 + |\overset{(-)}{A}{}_\parallel|^2 + |\overset{(-)}{A}{}_\perp|^2}. \tag{2.12}$$

An analogous equation holds in the helicity angles basis, where the $f_i(\cos\Theta, \Phi, \cos\Psi)$ functions are replaced by the $h_i(\cos\vartheta_1, \cos\vartheta_2, \varphi)$ functions. In Tab. 2.4, we list the $\overset{(-)}{K}{}_i(t)$, where we have defined the polarization amplitudes at $t = 0$ for an initially produced $\overset{(-)}{B}{}^0_s$ meson as:

$$\begin{aligned}
A_0(t=0) = A_0 &= \langle V_1V_2, 0 | B_s^0 \rangle, & \bar{A}_0(t=0) = \bar{A}_0 &= \langle V_1V_2, 0 | \bar{B}_s^0 \rangle, \\
A_\parallel(t=0) = A_\parallel &= \langle V_1V_2, \parallel | B_s^0 \rangle, & \bar{A}_\parallel(t=0) = \bar{A}_\parallel &= \langle V_1V_2, \parallel | \bar{B}_s^0 \rangle, \\
A_\perp(t=0) = A_\perp &= \langle V_1V_2, \perp | B_s^0 \rangle, & \bar{A}_\perp(t=0) = \bar{A}_\perp &= \langle V_1V_2, \perp | \bar{B}_s^0 \rangle.
\end{aligned} \tag{2.13}$$

The expressions in Tab. 2.4 includes the most general case for the time evolution of a B_s^0-decay into a vector-vector and self-conjugate final state. By computing:

$$\frac{K_i(t) - \bar{K}_i(t)}{K_i(t) + \bar{K}_i(t)}. \tag{2.14}$$

with $i = (1, 2, 3, 5)$, one can obtain the time-dependent *CP* asymmetry of Eq. (1.33), which includes both *direct* and *interference CP* violation.

i	$\overset{(-)}{K}_i$
1	$\frac{1}{2}e^{-\Gamma t}\Big[\Big(\lvert A_0\rvert^2+\lvert\bar{A}_0\rvert^2\Big)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\Big(\lvert A_0\rvert^2-\lvert\bar{A}_0\rvert^2\Big)\cos\Delta m_s t$ $-2\mathfrak{Re}(A_0^\star\bar{A}_0)\Big(\cos 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\overset{(+)}{-}\sin 2\phi_M^s\sin\Delta m_s t\Big)$ $-2\mathfrak{Im}(A_0^\star\bar{A}_0)\Big(\overset{(-)}{+}\cos 2\phi_M^s\sin\Delta m_s t+\sin 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\Big)\Big]$
2	$\frac{1}{2}e^{-\Gamma t}\Big[\Big(\lvert A_\parallel\rvert^2+\lvert\bar{A}_\parallel\rvert^2\Big)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\Big(\lvert A_\parallel\rvert^2-\lvert\bar{A}_\parallel\rvert^2\Big)\cos\Delta m_s t$ $-2\mathfrak{Re}(A_\parallel^\star\bar{A}_\parallel)\Big(\cos 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\overset{(+)}{-}\sin 2\phi_M^s\sin\Delta m_s t\Big)$ $-2\mathfrak{Im}(A_\parallel^\star\bar{A}_\parallel)\Big(\overset{(-)}{+}\cos 2\phi_M^s\sin\Delta m_s t+\sin 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\Big)\Big]$
3	$\frac{1}{2}e^{-\Gamma t}\Big[\Big(\lvert A_\perp\rvert^2+\lvert\bar{A}_\perp\rvert^2\Big)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\Big(\lvert A_\perp\rvert^2-\lvert\bar{A}_\perp\rvert^2\Big)\cos\Delta m_s t$ $+2\mathfrak{Re}(A_\perp^\star\bar{A}_\perp)\Big(\cos 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\overset{(+)}{-}\sin 2\phi_M^s\sin\Delta m_s t\Big)$ $+2\mathfrak{Im}(A_\perp^\star\bar{A}_\perp)\Big(\overset{(-)}{+}\cos 2\phi_M^s\sin\Delta m_s t+\sin 2\phi_M^s\sinh\left(\Delta\Gamma_s t/2\right)\Big)\Big]$
4	$\frac{1}{2}e^{-\Gamma t}\Big[\mathfrak{Im}(A_\perp A_\parallel^\star-\bar{A}_\perp\bar{A}_\parallel^\star)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\mathfrak{Im}(A_\perp A_\parallel^\star+\bar{A}_\perp\bar{A}_\parallel^\star)\cos\Delta m_s t$ $+\mathfrak{Im}(A_\perp\bar{A}_\parallel^\star-\bar{A}_\perp A_\parallel^\star)\Big(-\sinh\left(\Delta\Gamma_s t/2\right)\cos 2\phi_M^s\overset{(-)}{+}\sin\Delta m_s t\sin 2\phi_M^s\Big)$ $+\mathfrak{Re}(A_\perp\bar{A}_\parallel^\star+\bar{A}_\perp A_\parallel^\star)\Big(-\sinh\left(\Delta\Gamma_s t/2\right)\sin 2\phi_M^s\overset{(+)}{-}\sin\Delta m_s t\cos 2\phi_M^s\Big)\Big]$
5	$\frac{1}{2}e^{-\Gamma t}\Big[\mathfrak{Re}(A_\parallel A_0^\star+\bar{A}_\parallel\bar{A}_0^\star)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\mathfrak{Re}(A_\parallel A_0^\star-\bar{A}_\parallel\bar{A}_0^\star)\cos\Delta m_s t$ $+\mathfrak{Re}(A_\parallel\bar{A}_0^\star+\bar{A}_\parallel A_0^\star)\Big(-\sinh\left(\Delta\Gamma_s t/2\right)\cos 2\phi_M^s\overset{(-)}{+}\sin\Delta m_s t\sin 2\phi_M^s\Big)$ $+\mathfrak{Im}(A_\parallel\bar{A}_0^\star-\bar{A}_\parallel A_0^\star)\Big(\sinh\left(\Delta\Gamma_s t/2\right)\sin 2\phi_M^s\overset{(-)}{+}\sin\Delta m_s t\cos 2\phi_M^s\Big)\Big]$
6	$\frac{1}{2}e^{-\Gamma t}\Big[\mathfrak{Im}(A_\perp A_0^\star-\bar{A}_\perp\bar{A}_0^\star)\cosh\left(\Delta\Gamma_s t/2\right)\overset{(-)}{+}\mathfrak{Im}(A_\perp A_0^\star+\bar{A}_\perp\bar{A}_0^\star)\cos\Delta m_s t$ $+\mathfrak{Im}(A_\perp\bar{A}_0^\star-\bar{A}_\perp A_0^\star)\Big(-\sinh\left(\Delta\Gamma_s t/2\right)\cos 2\phi_M^s\overset{(-)}{+}\sin\Delta m_s t\sin 2\phi_M^s\Big)$ $+\mathfrak{Re}(A_\perp\bar{A}_0^\star+\bar{A}_\perp A_0^\star)\Big(-\sinh\left(\Delta\Gamma_s t/2\right)\sin 2\phi_M^s\overset{(+)}{-}\sin\Delta m_s t\cos 2\phi_M^s\Big)\Big]$

Table 2.4: General expressions of $K_i(t)$ ($\bar{K}_i(t)$) terms for $B_s^0\to V_1V_2$ decays, where both direct and interference *CP* violation are allowed.

In particular, direct *CP* violation is present when the moduli of the decay amplitudes and of their *CP* conjugate are different:

$$|A_i| \neq |\bar{A}_i|.$$

and the polarization amplitudes are written as follows:

$$\begin{aligned} A_0 &= |A_0| e^{i\phi_{A_0}} e^{i\delta_0}, \quad & \bar{A}_0 &= |\bar{A}_0| e^{-i\phi_{A_0}} e^{i\delta_0}, \\ A_\parallel &= |A_\parallel| e^{i\phi_{A_\parallel}} e^{i\delta_\parallel}, \quad & \bar{A}_\parallel &= |\bar{A}_\parallel| e^{-i\phi_{A_\parallel}} e^{i\delta_\parallel}, \\ A_\perp &= |A_\perp| e^{i\phi_{A_\perp}} e^{i\delta_\perp}, \quad & \bar{A}_\perp &= -|\bar{A}_\perp| e^{-i\phi_{A_\perp}} e^{i\delta_\perp}, \end{aligned} \tag{2.15}$$

with ϕ_{A_i} and δ_i respectively the weak and strong phases of the amplitudes. Thus, Tab. 2.4 represents a compact writings to emphasize the two distinct source of *CP* violation. The terms proportional to $\mathfrak{Re}(\overset{(-)}{A}_i \overset{(-)}{A}_j{}^\star)$ and to $\mathfrak{Im}(\overset{(-)}{A}_i \overset{(-)}{A}_j{}^\star)$ in the $\overset{(-)}{K}_i(t)$ encode the dependence on both ϕ_{A_i} and δ_i. The dependence on the mixing phase, ϕ_M^s, is factorized in the expressions, but only the phase differences $\phi_i^f = 2(\phi_M^s - \phi_{A_i})$ (Eq. (1.43)) have physical meaning and are observable. To see the explicit dependence on the phase ϕ_i^f, one should expand the terms containing ϕ_A^i. We can derive the expressions as functions of ϕ_i^f for the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays, considering the SM case of a single weak phase for each amplitude, $\phi_{A_j} = \phi_A$, and no direct *CP* violation, $|A_j| = |\bar{A}_j|$, as follows. First, the amplitudes become:

$$A_j = |A_j| e^{i\phi_A} e^{i\delta_j}, \quad \bar{A}_j = \pm |A_j| e^{-i\phi_A} e^{i\delta_j},$$

with the "+" sign for $j = (0, \parallel)$ and the "−" for $j = \perp$. We expand as an example the term $K_1(t)$ in Tab. 2.4. We have:

$$\begin{aligned} &|A_0|^2 + |\bar{A}_0|^2 = 2|A_0|^2, \\ &|A_0|^2 - |\bar{A}_0|^2 = 0, \\ &\mathfrak{Re}\left(A_0^\star \bar{A}_0\right) = |A_0|^2 \cos(2\phi_A), \\ &\mathfrak{Im}\left(A_0^\star \bar{A}_0\right) = -|A_0|^2 \sin(2\phi_A), \end{aligned}$$

then:

$$\begin{aligned} K_1(t) = |A_0|^2 e^{-\Gamma t} [&\cosh(\Delta\Gamma_s t/2) \\ &- \underbrace{(\cos 2\phi_A \cos 2\phi_M^s + \sin 2\phi_A \sin 2\phi_M^s)}_{\cos\phi^f} \sinh(\Delta\Gamma_s t/2) \\ &- \underbrace{(\cos 2\phi_A \sin 2\phi_M^s - \sin 2\phi_A \cos 2\phi_M^s)}_{-\sin\phi^f} \sin(\Delta m_s t)]. \end{aligned}$$

In the case of $B_s^0 \to J/\psi\phi$ decays, $\phi^f = \phi^{c\bar{c}s}$ of Eq. (1.47); in the case of $B_s^0 \to \phi\phi$ decays, $\phi^f = \phi^{s\bar{s}s}$ of Eq. (1.55). We focus now on the former decay: the time-evolution functions $K_i(t)$ are reported in Tab. 2.5, where we use a compact form, which highlights the presence

i	$\overline{K}_i(t)$	*CP* parity
1	$\lvert\overline{A}_0\rvert^2\,\overline{\mathcal{O}}^+(t)$	even
2	$\lvert\overline{A}_\parallel\rvert^2\,\overline{\mathcal{O}}^+(t)$	even
3	$\lvert\overline{A}_\perp\rvert^2\,\overline{\mathcal{O}}^-(t)$	odd
4	$\lvert\overline{A}_\parallel\rvert\lvert\overline{A}_\perp\rvert\,\overline{\mathcal{E}}_\mathfrak{I}(t,\delta_\perp-\delta_\parallel)$	mix
5	$\lvert\overline{A}_\parallel\rvert\lvert\overline{A}_0\rvert\cos\delta_\parallel\,\overline{\mathcal{O}}^+(t)$	even
6	$\lvert\overline{A}_\perp\rvert\lvert\overline{A}_0\rvert\,\overline{\mathcal{E}}_\mathfrak{I}(t,\delta_\perp)$	mix

Table 2.5: Expressions of $K_i(t)$ terms of the $B_s^0\to J/\psi\phi$ decay rate, where only interference *CP* violation is allowed. The third column reports the *CP* parity of each term. The formulae of $\mathcal{O}^\pm$ and $\mathcal{E}_\mathfrak{I}$ are given by Eq. (2.16), while the phases $\delta_\parallel$ and $\delta_\perp$ are defined by Eq. (2.19)–2.18.

of the same time dependence for final states with the same *CP*-parity. We have used the "β_s" convention as per Eq. (1.47) in the expressions, and defined:

$$\overline{\mathcal{O}}^+(t) = e^{-\Gamma t}\left(\cosh\frac{\Delta\Gamma_s t}{2} - \cos 2\beta_s \sinh\frac{\Delta\Gamma_s t}{2} \overset{(-)}{+} \sin 2\beta_s \sin\Delta m_s t\right), \tag{2.16}$$

$$\overline{\mathcal{O}}^-(t) = e^{-\Gamma t}\left(\cosh\frac{\Delta\Gamma_s t}{2} + \cos 2\beta_s \sinh\frac{\Delta\Gamma_s t}{2} \overset{(+)}{-} \sin 2\beta_s \sin\Delta m_s t\right), \tag{2.17}$$

$$\begin{aligned}\overline{\mathcal{E}}_\mathfrak{I}(t,\alpha) = e^{-\Gamma t}\Bigg(&\overset{(-)}{+}\sin\alpha\cos\Delta m_s t \overset{(+)}{-}\cos\alpha\cos 2\beta_s\sin\Delta m_s t \\ &- \cos\alpha\sin 2\beta_s\sinh\frac{\Delta\Gamma_s t}{2}\Bigg).\end{aligned} \tag{2.18}$$

The phase α in such equations represents the *CP*-conserving phase associate to the polarization amplitudes. Since only phase differences matter, a customary convention is to choose A_0 real and define the strong phases of the transverse amplitudes as:

$$\begin{aligned}\delta_\parallel &= \arg\left(\frac{A_\parallel}{A_0}\right) = \arg\left(\frac{\bar{A}_\parallel}{\bar{A}_0}\right), \\ \delta_\perp &= \arg\left(\frac{A_\perp}{A_0}\right) = \arg\left(\frac{\bar{A}_\perp}{\bar{A}_0}\right).\end{aligned} \tag{2.19}$$

2.2.1 Time evolution of the $(S+P)$-waves system

So far we have considered K^+K^- pairs originated from the decay of $\phi(1020)$ mesons only, both in the $B_s^0\to J/\psi\phi$ and in the $B_s^0\to\phi\phi$ decays. However, the reconstructed K^+K^- spectrum could be a mixture of two resonances, their interference, and also of a non-resonant production. We now focus on the decay $B_s^0\to J/\psi X$, with $X\to K^+K^-$, which is the case of the $B_s^0\to J/\psi\phi$ analysis.[2]

[2]Similar arguments can hold for the $B_s^0\to\phi\phi$ analysis. We can consider the general decay $B_s^0\to X(\to K^+K^-)X(\to K^+K^-)$, where X is one of the resonance in the K^+K^- spectrum described in the text.

Specifically, it was argued that neglecting a large contamination of a $s\bar{s}$-quark resonance of zero spin with mass close to the $\phi(1020)$ pole, the $f_0(980)$ [8], biases the estimation of the *CP*-odd fraction of the signal and alters the measurement of β_s [55]. Therefore, in order to properly disentangle the decay amplitudes in $B_s^0 \to J/\psi\phi$ decays, one should take into account the contributions of these different sources.

A nomenclature from partial wave formalism is usually assigned to the two resonances, according to their spin: the $f_0(980)$ is a spin-0 particle and it is called *S-wave*, while the $\phi(1020)$ is a spin-1 particle and it is denoted as *P-wave*.[3] The partial-wave classification of the K^+K^- spectrum, which is based on the spin of the resonance, should not be confused with the partial-wave basis of the polarization amplitudes, which is based on the value of the relative angular momentum between two resonances. In what follows, we will use the partial-waves nomenclature only for referring to the K^+K^- resonances spectrum, while for polarization amplitudes we will use only the transversity (or helicity) basis.

The differential decay rates considered so far – *e. g.* Eq. (2.12) – are parametrized as a function of decay (helicity or transversity) angles and decay time. However, they also depend on the invariant mass m of the K^+K^- resonances, and the amplitudes should be considered as a function of m as well [54, 56]. In cases where the amplitudes are dominated by a particular resonance in a given energy range, one can factorize the mass dependence, usually given by a relativistic Breit-Wigner distribution, $B_J(m)$, and the polarization dependence, $A_i(t)$ $(i = 0, \parallel, \perp)$; hence, the transversity amplitudes become [54]:

$$A_i(t) \to A_i'(m,t) = B_J(m) A_i(t), \tag{2.20}$$

where the index J represents the dependence on the spin of the resonance. We describe the dependence on $B_J(m)$ by analyzing the K^+K^- spectrum in App. B, and here we present the $B_s^0 \to J/\psi(K^+K^-)_{(S+P)}$ decay rate as a function of the transversity angles, the decay time, and the K^+K^- mass. To account for the total $S+P$ contribution in the decay rate we have to sum the P_{wave} amplitude and the S_{wave} amplitude; then the total decay rate can be decomposed as follows [56, 54]:

$$\begin{aligned} \frac{\mathrm{d}^5\,\overline{\Gamma}^{(-)}(\,\overline{B}^{(-)0}_s \to J/\psi K^+K^-)}{\mathrm{d}m\,\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi} &= |\,\overline{P}^{(-)}_{\text{wave}} + \overline{S}^{(-)}_{\text{wave}}|^2 \\ &= |\,\overline{P}^{(-)}_{\text{wave}}|^2 + |\,\overline{S}^{(-)}_{\text{wave}}|^2 + 2\,\mathfrak{Re}\left(\overline{P}^{(-)}_{\text{wave}}\,\overline{S}^{(-)\star}_{\text{wave}}\right), \end{aligned} \tag{2.21}$$

The P-wave component is given by:

$$|\,\overline{P}^{(-)}_{\text{wave}}|^2 = |B_\phi(m)|^2 \frac{\mathrm{d}^4\,\overline{\Gamma}^{(-)}(\,\overline{B}^{(-)0}_s \to J/\psi\phi)}{\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi}. \tag{2.22}$$

[3] There can be also a fraction of $B_s^0 \to J/\psi K^+K^-$ decays with a non-resonant K^+K^- pair that has a relative angular momentum $L = 1$ with respect the J/ψ state. However, it should be much smaller than the resonant fraction, because its amplitude involves the production of an extra $u\bar{u}$-quark pair with respect to the amplitude for the resonances' production. Hence, the non-resonant component will not be consider in the following.

where the term $\frac{\mathrm{d}^4\Gamma(B_s^0\to J/\psi\phi)}{\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi}$ is the $B_s^0\to J/\psi\phi$ decay rate as a function of the B decay-time and the transversity angles, Eq. (2.12); the term $|B_\phi(m)|^2$ is the relativistic Breit-Wigner distribution of the $\phi(1020)$ resonance, Eq. (B.7).

The $B_s^0\to J/\psi f_0$ decay is a pseudoscalar to vector-scalar decay which has only one allowed state of relative angular momentum ($L=1$); hence, the transition is given by one amplitude, $A_S(t)=|A_S(t)|e^{i\delta_S}$, where δ_S is its *CP*-conserving phase:

$$|\overset{(-)}{S}{}_{\text{wave}}|^2 = |B_{f_0}(m)|^2|A_S(t)|^2 f_7(\cos\Theta,\Phi,\cos\Psi). \tag{2.23}$$

The term $|B_{f_0}(m)|^2$ describes the $f_0(980)$ distribution in the K^+K^- spectrum and its is given by the asymmetric Flatté distribution of Eq. (B.8).

The interference term $\mathfrak{Re}(\overset{(-)}{P}{}_{\text{wave}}\,\overset{(-)}{S}{}^{\star}_{\text{wave}})$ is decomposed in three part, since the S-wave interferes with each polarization state of the $B_s^0\to J/\psi\phi$ decay:

$$\begin{aligned}\mathfrak{Re}\left(\overset{(-)}{P}{}_{\text{wave}}\,\overset{(-)}{S}{}^{\star}_{\text{wave}}\right) = &\,\mathfrak{Re}\left((B_\phi B^\star_{f_0})(m)\,(\overset{(-)}{A}{}_{\parallel}\,\overset{(-)}{A}{}^{\star}_{S})(t)\right) f_8(\cos\Theta,\Phi,\cos\Psi)\\ &+\mathfrak{Im}\left((B_\phi B^\star_{f_0})(m)\,(\overset{(-)}{A}{}_{\perp}\,\overset{(-)}{A}{}^{\star}_{S})(t)\right) f_9(\cos\Theta,\Phi,\cos\Psi)\\ &+\mathfrak{Re}\left((B_\phi B^\star_{f_0})(m)\,(\overset{(-)}{A}{}_{0}\,\overset{(-)}{A}{}^{\star}_{S})(t)\right) f_{10}(\cos\Theta,\Phi,\cos\Psi).\end{aligned} \tag{2.24}$$

The angular functions $f_i(\cos\Theta,\Phi,\cos\Psi)$ $(i=7,8,9,10)$ are presented in Tab. 2.6 [56, 20]. The $J/\psi f_0$ state is a pure *CP*-odd eigenstate; then the time-evolution of the amplitude A_S follows the same function of the $A_\perp(t)$ amplitude in the $B_s^0\to J/\psi\phi$ decay. Table 2.6 displays the time-evolution of the new terms involving the S-wave state: $|A_S(t)|^2$; $\mathfrak{Re}(\overset{(-)}{A}{}_{\parallel}(t)\,\overset{(-)}{A}{}^{\star}_{S}(t))$; $\mathfrak{Im}(\overset{(-)}{A}{}_{\perp}(t)\,\overset{(-)}{A}{}^{\star}_{S})(t))$; and $\mathfrak{Re}(\overset{(-)}{A}{}_{0}(t)\,\overset{(-)}{A}{}^{\star}_{S}(t))$.

Usually, in analyses of $B_s^0\to J/\psi\phi$ decays, the dependence on the mass m is integrated out and the disentanglement of the polarization amplitude are obtained with the analysis of the angular distributions only. With

$$\int |B_J(m)|^2\mathrm{d}m = 1, \tag{2.25}$$

$$\int (B_\phi B^\star_{f_0})(m)\mathrm{d}m = I_m, \tag{2.26}$$

we obtain:

$$\frac{1}{\overset{(-)}{\Gamma}}\frac{\mathrm{d}^4\,\overset{(-)}{\Gamma}(\overset{(-)}{B}{}^0_s\to J/\psi K^+K^-)}{\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi} = \frac{\sum_{i=1}^{10}\left[\overset{(-)}{K}{}_i(t)\,f_i(\cos\Theta,\Phi,\cos\Psi)\right]}{|\overset{(-)}{A}{}_0|^2+|\overset{(-)}{A}{}_\parallel|^2+|\overset{(-)}{A}{}_\perp|^2+|\overset{(-)}{A}{}_S|^2}, \tag{2.27}$$

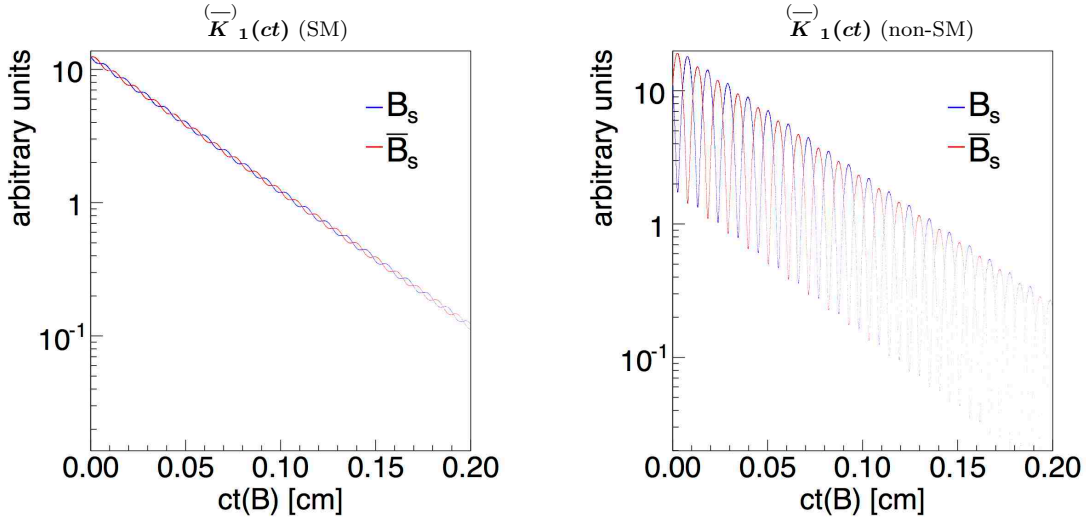
i	$\overline{K}_i(t)$	$f_i(\cos\Theta, \Phi, \cos\Psi)$	*CP* parity
1	$\lvert\overline{A}_0\rvert^2 \overline{\mathcal{O}}^+(t)$	$\frac{9}{32\pi} 2\cos^2\Psi(1-\sin^2\Theta\cos^2\Phi)$	even
2	$\lvert\overline{A}_\parallel\rvert^2 \overline{\mathcal{O}}^+(t)$	$\frac{9}{32\pi} \sin^2\Psi(1-\sin^2\Theta\sin^2\Phi)$	even
3	$\lvert\overline{A}_\perp\rvert^2 \overline{\mathcal{O}}^-(t)$	$\frac{9}{32\pi} \sin^2\Psi\sin^2\Theta$	odd
4	$\lvert\overline{A}_\parallel\rvert\lvert\overline{A}_\perp\rvert \overline{\mathcal{E}}_\mathfrak{I}(t,\delta_\perp-\delta_\parallel)$	$-\frac{9}{32\pi} \sin^2\Psi\sin 2\Theta\sin\Phi$	mix
5	$\lvert\overline{A}_\parallel\rvert\lvert\overline{A}_0\rvert\cos\delta_\parallel \overline{\mathcal{O}}^+(t)$	$\frac{9}{32\pi}\frac{\sqrt{2}}{2} \sin 2\Psi\sin^2\Theta\sin 2\Phi$	even
6	$\lvert\overline{A}_\perp\rvert\lvert\overline{A}_0\rvert \overline{\mathcal{E}}_\mathfrak{I}(t,\delta_\perp)$	$\frac{9}{32\pi}\frac{\sqrt{2}}{2} \sin 2\Psi\sin 2\Theta\cos\Phi$	mix
7	$\lvert\overline{A}_S\rvert^2 \overline{\mathcal{O}}^-(t)$	$\frac{3}{32\pi} 2(1-\sin^2\Theta\cos^2\Phi)$	odd
8	$I_m\lvert\overline{A}_\parallel\rvert\lvert\overline{A}_S\rvert \overline{\mathcal{E}}_\mathfrak{R}(t,\delta_\parallel-\delta_S)$	$\frac{3}{32\pi} 2\cos\Psi(1-\sin^2\Theta\cos^2\Phi)$	mix
9	$I_m\lvert\overline{A}_\perp\rvert\lvert\overline{A}_S\rvert\sin(\delta_\perp-\delta_S) \overline{\mathcal{O}}^-(t)$	$\frac{3}{32\pi}\frac{1}{\sqrt{2}} 2\sin\Psi\sin^2\Theta\sin 2\Phi$	odd
10	$I_m\lvert\overline{A}_0\rvert\lvert\overline{A}_S\rvert \overline{\mathcal{E}}_\mathfrak{R}(t,-\delta_S)$	$\frac{3}{32\pi}\frac{1}{\sqrt{2}} 2\sin\Psi\sin 2\Theta\cos\Phi$	mix

Table 2.6: Expressions of $K_i(t)$ ($\bar{K}_i(t)$) and $f_i(\cos\Theta, \Phi, \cos\Psi)$ terms of the $B_s^0 \to J/\psi K^+K^-$ decay rate, where the dependence on the mass m is integrated out. The last column reports the *CP* parity of each term. The formulae of $\mathcal{O}^\pm$ and $\mathcal{E}_\mathfrak{I}$ are given by Eq. (2.16) and Eq. (2.28). The coefficient I_m is the integral in Eq. (2.26).

with the expressions of $\overline{K}_i(t)$ and $f_i(\cos\Theta, \Phi, \cos\Psi)$ given in Tab. 2.6, where we have introduced the following shorthand in the time-evolution of the interference term between P- and S-wave with mixed *CP*-parity:

$$\overline{\mathcal{E}}_\mathfrak{R}(t,\alpha) = e^{-\Gamma t}\Big(\overset{(-)}{+}\cos\alpha\cos\Delta m_s t \overset{(+)}{-}\sin\alpha\cos 2\beta_s\sin\Delta m_s t - \sin\alpha\sin 2\beta_s\sinh\frac{\Delta\Gamma_s t}{2}\Big). \tag{2.28}$$

Figure 2.4 sketches the evolution of the amplitude K_1 ($|A_0(t)|^2$) of the $B_s^0 \to J/\psi K^+K^-$ decay rate in Eq. (2.27), as a function of the decay-length, separately for the B_s^0 and the $\bar{B}_s^0$ mesons, at the SM values of $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.90\,\text{ps}^{-1}$ in (a), and for $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.09\cos(2\beta_s) = 0.49\,\text{ps}^{-1}$ in (b). In both cases, the value of the oscillation frequency is fixed to $\Delta m_s = 17.77\,\text{ps}^{-1}$ [**?**], and polarization amplitudes and strong phases as measured in Ref. [18]. The plots shows how the time-evolution of the decay-amplitude changes for different values of β_s and $\Delta\Gamma_s$. Specifically, the squared moduli of the polarization amplitudes depend on the terms $\cos 2\beta_s\sinh(\Delta\Gamma_s t/2)$ and $\sin 2\beta_s\sin(\Delta m_s t)$; the former provides sensitivity to β_s even without distinction of the flavor of the B_s^0-meson at production, if $\Delta\Gamma_s$ is different from zero. The latter explains the reason of the different size of the oscillations amplitude. In App. C, we show the time-evolution of each term of the decay rate in Eq. (2.27) for the two values of $(\beta_s, \Delta\Gamma_s)$ reported above.



(a) (b)

Figure 2.4: Evolution of the amplitude K_1 of the $B_s^0 \to J/\psi K^+K^-$ decay rate in Eq. (2.27), as a function of ct. Blue line is for B_s^0 meson and red line is for $\bar{B}_s^0$. In (a) $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.09\,\text{ps}^{-1}$ (SM point), while in (b) $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.09\cos(2\beta_s) = 0.049\,\text{ps}^{-1}$. In both cases, $\Delta m_s = 17.77\,\text{ps}^{-1}$. Plots are obtained by means of the simulation described in Sect. 5.5.2.

2.2.2 Discrete symmetries

The decay rate in Eq. (2.27) features some discrete symmetries, *i. e.*, transformations of some of the observables of interest that leave the equations of the decay rate invariant.

We first consider the simpler case, where only the P-wave is present. Suppose we cannot distinguish the B_s^0 and the $\bar{B}_s^0$ meson at their production and suppose they are produced in the same amount (*untagged* sample). Thus, we sum the $B_s^0 \to J/\psi K^+K^-$ and the $\bar{B}_s^0 \to J/\psi K^+K^-$ decay rates. In such case, each oscillation term proportional to $\sin\Delta m_s t$ or $\cos\Delta m_s t$ is canceled out, since they appear with opposite sign in the $K_i(t)$ and $\bar{K}_i(t)$ terms, but we are still sensitive to β_s if $\Delta\Gamma_s \neq 0$. Then, the untagged decay rate is invariant under [44]:

$$\begin{cases} \beta_s \to \pi/2 - \beta_s \\ \Delta\Gamma_s \to -\Delta\Gamma_s \\ \delta_\parallel \to 2\pi - \delta_\parallel \\ \delta_\perp \to \pi - \delta_\perp \end{cases} \tag{2.29}$$

together with the reflection of this transformation with respect to $\beta_s = 0$:

$$\begin{cases} \beta_s \to -\beta_s \\ \Delta\Gamma_s \to \Delta\Gamma_s \\ \delta_\parallel \to \delta_\parallel \\ \delta_\perp \to \delta_\perp \end{cases} \quad \text{and} \quad \begin{cases} \pi/2 - \beta_s \to -\pi/2 + \beta_s \\ -\Delta\Gamma_s \to -\Delta\Gamma_s \\ 2\pi - \delta_\parallel \to 2\pi - \delta_\parallel \\ \pi - \delta_\perp \to \pi - \delta_\perp \end{cases} \tag{2.30}$$

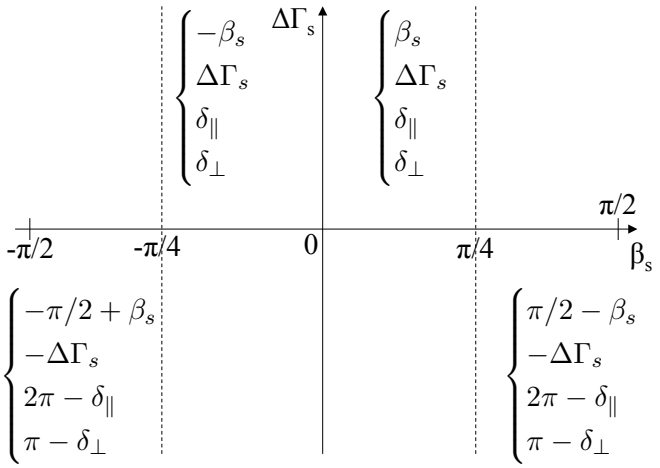


Figure 2.5: Example in the $(\beta_s, \Delta\Gamma_s)$ plane of the equivalent values of β_s, $\Delta\Gamma_s$, and strong phases, that leave the decay rate invariant.

Therefore a four-fold ambiguity is present for the values of the observables in the transformations. The four equivalent solutions are sketched in the $(\beta_s, \Delta\Gamma_s)$ plane in Fig. 2.5. When the differences of the B_s^0 and $\bar{B}_s^0$ meson decay-rates are taken into account, the transformations in Eq. (2.30) are not longer symmetries of the decay rate, and only the transformation of Eq. (2.29) leaves the decay rate invariant. We have then a two-fold ambiguity (with the cancellation of the solutions for $\beta_s < 0$ in Fig. 2.5).

Considering also the contribution of the S-wave state makes the argument more complicated. We first start by analyzing the decay rate in Eq. (2.21), but considering a simplify model, a nonrelativistic form of the $\phi(1020)$, such as:

$$|B_\phi(m)|^2 \propto \frac{m_\phi \Gamma_\phi}{(m_\phi^2 - m^2)^2 + m_\phi^2 \Gamma_\phi} \tag{2.31}$$

with a mass-independent width Γ_ϕ symmetric around the pole m_ϕ. For the $f_0(980)$ resonance, given its large and slowly-decreasing width, when we restrict to a narrow window around m_ϕ, we can consider the $f_0(980)$ shape as a flat distribution (see Fig. B.1 in App. B). Then, when considering the $(S+P)$ decay rate in Eq. (2.27) integrated over a finite interval $[m_1, m_2]$ in the K^+K^- mass spectrum, along with Eq. (2.29), the transformation [56]:

$$\delta_S \to 2\delta_{\rm BW} - \delta_S \tag{2.32}$$

leaves the decay rate of the $(S+P)$-wave system invariant. The phase $\delta_{\rm BW}$ is:

$$\delta_{\rm BW} = \arg\left[\log\left(\frac{m_2 - m_\phi + i\Gamma/2}{m_1 - m_\phi + i\Gamma/2}\right)\right], \tag{2.33}$$

and reduces to 0 in the limit of an infinitesimal interval of integration, and to $\pi/2$ in the limit of a finite interval symmetric around m_ϕ.

Then, considering a window around m_ϕ of about $20\,\text{MeV}/c^2$, the transformation:

$$\begin{cases} \beta_s \to \pi/2 - \beta_s \\ \Delta\Gamma_s \to -\Delta\Gamma_s \\ \delta_\parallel \to 2\pi - \delta_\parallel \\ \delta_\perp \to \pi - \delta_\perp \\ \delta_S \to \pi - \delta_S \end{cases} \tag{2.34}$$

is again a symmetry of the integrated decay rate. We have to remember, however, that the invariance under tranformation 2.34 requires the symmetry of the K^+K^- resonances. In the more general case, when considering the asymmetric shapes of the K^+K^- resonances described in App. B, the transformation 2.34 leads to an approximate symmetry, which is as spoiled as larger is the S-wave fraction in the sample, since the asymmetric shape of the integral I_m in Eq. (2.26) becomes more effective for larger fractions. When considering a S-wave fraction as the one measured in our data ($\approx 1\%$), we can treat the decay rate as symmetric under the transformation 2.34.

Since each solution of $(\beta_s, \Delta\Gamma_s)$ corresponds to a different set of strong phases, one may attempt to experimentally resolve the ambiguity by using a measurement of the variation of the strong phases as follows. The phase of the P-wave amplitude rises rapidly through the $\phi(1020)$ mass pole. On the other hand, the phase of the S-wave amplitude should vary relatively slowly. This is shown in Fig. B.1 of App. B. As a result, the phase difference between the S-wave and P-wave amplitudes, for instance $\delta_S - \delta_\perp$, falls rapidly with increasing m. By measuring this phase difference as a function of m and taking the solution with a decreasing trend around m_ϕ as the physical solution, the sign of $\Delta\Gamma_s$ is determined and the ambiguity in β_s is resolved. This analysis was performed by the LHCb Collaboration and led to the determination $\Delta\Gamma_s > 0$ [42].

2.3 Time-integrated triple products asymmetries

The current statistics of the $B_s^0 \to \phi\phi$ data sample are not sufficient for a suitable time-dependent analysis to measure the $\phi^{s\bar{s}s}$ phase of the time-dependent CP-asymmetry. However, an investigation of time-integrated observables that encode the same information of $\phi^{s\bar{s}s}$ about CP violation in this decay mode is accessible. Such asymmetries are defined in terms of angular distributions asymmetries of the functions of the helicity angles that we are going to describe in this section.

Integrating over ϑ_1 and ϑ_2 the decay rate of Eq. (2.8) one obtains the following distribution in φ:

$$\frac{\mathrm{d}\,\overline{\Gamma}}{\mathrm{d}\varphi} = \frac{1}{2\pi} \frac{\overline{K}_1 + 2\,\overline{K}_2 \sin^2\varphi + 2\,\overline{K}_3 \cos^2\varphi - 2\,\overline{K}_4 \sin 2\varphi}{|\overline{A}_0|^2 + |\overline{A}_\parallel|^2 + |\overline{A}_\perp|^2}. \tag{2.35}$$

The last term in this angular distribution provides a *triple products* (TP):

$$\sin 2\varphi = 2(\hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2)(\hat{\boldsymbol{n}}_1 \times \hat{\boldsymbol{n}}_2) \cdot \hat{\boldsymbol{z}}_1. \tag{2.36}$$

Considering Fig. 2.1, we have defined $\boldsymbol{z}_i$ as the direction of the V_i momentum in the B rest frame and denote by $\hat{\boldsymbol{z}}_i$ a unit vector in this direction. Unit vectors normal to the two decay planes and to their line of intersection defined by $\hat{\boldsymbol{z}}_{\boldsymbol{i}}$ are denoted as $\hat{\boldsymbol{n}}_i$; and they are defined in order that ϕ is the angle from $\hat{\boldsymbol{n}}_1$ to $\hat{\boldsymbol{n}}_2$. A TP asymmetry is defined as an asymmetry between the number of decays ($\overline{N}$) involving positive and negative values of $\sin 2\varphi$:

$$\begin{aligned} \overline{\mathcal{A}}^{(2)}_{\rm TP} &= \frac{\overline{N}(\sin 2\varphi \geq 0) - \overline{N}(\sin 2\varphi < 0)}{\overline{N}(\sin 2\varphi \geq 0) + \overline{N}(\sin 2\varphi < 0)} \\ &= \frac{\left[\int_0^{\pi/2} + \int_{\pi}^{3\pi/2}\right] \frac{\mathrm{d}\,\overline{\Gamma}}{\mathrm{d}\varphi}\mathrm{d}\varphi - \left[\int_{\pi/2}^{\pi} + \int_{3\pi/2}^{2\pi}\right] \frac{\mathrm{d}\,\overline{\Gamma}}{\mathrm{d}\varphi}\mathrm{d}\varphi}{\int_0^{2\pi} \frac{\mathrm{d}\,\overline{\Gamma}}{\mathrm{d}\varphi}\mathrm{d}\varphi}. \end{aligned} \tag{2.37}$$

Using Eq. (2.35):

$$\overline{\mathcal{A}}^{(2)}_{\rm TP} = \overline{+}\frac{4}{\pi}\frac{\overline{K}_4}{|\overline{A}_0|^2 + |\overline{A}_\parallel|^2 + |\overline{A}_\perp|^2}. \tag{2.38}$$

The dependence of the angular distribution Eq. (2.8) on ϑ_1 and ϑ_2 allow us to introduce a second TP:

$$\sin\varphi = (\hat{\boldsymbol{n}}_1 \times \hat{\boldsymbol{n}}_2) \cdot \hat{\boldsymbol{z}}_{\boldsymbol{1}}, \tag{2.39}$$

The corresponding asymmetry, $\overline{\mathcal{A}}^{(1)}_{\rm TP}$ involves $\overline{K}_6$. One defines an asymmetry with respect to values of $\sin\varphi$, the TP in Eq. (2.39), assigning it the sign of $\cos\vartheta_1\cos\vartheta_2$ (a T-even quantity) and integrating over all angles,

$$\overline{\mathcal{A}}^{(1)}_{\rm TP} = \frac{\overline{N}(\mathrm{sign}(\cos\vartheta_1\cos\vartheta_2)\sin\varphi > 0) - \overline{N}(\mathrm{sign}(\cos\vartheta_1\cos\vartheta_2)\sin\varphi < 0)}{\overline{N}(\mathrm{sign}(\cos\vartheta_1\cos\vartheta_2)\sin\varphi > 0) + \overline{N}(\mathrm{sign}(\cos\vartheta_1\cos\vartheta_2)\sin\varphi < 0)}. \tag{2.40}$$

By using Eq. (2.8):

$$\begin{aligned} \overline{\mathcal{A}}^{(1)}_{\rm TP} &= \frac{(\overline{I}_1 - \overline{I}_2) - (\overline{I}_3 - \overline{I}_4)}{\overline{I}_1 + \overline{I}_2 + \overline{I}_3 + \overline{I}_4} \\ &= \overline{+}\frac{2\sqrt{2}}{\pi}\frac{\overline{K}_6}{|\overline{A}_0|^2 + |\overline{A}_\parallel|^2 + |\overline{A}_\perp|^2}, \end{aligned} \tag{2.41}$$

where (with $x = \cos\vartheta_1$ and $y = \cos\vartheta_2$)

$$\begin{aligned} \overline{I}_1 &= \left(\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \int_0^{\pi} \mathrm{d}\varphi + \int_{-1}^0 \mathrm{d}x \int_{-1}^0 \mathrm{d}y \int_0^{\pi} \mathrm{d}\varphi\right) \frac{\mathrm{d}^3\,\overline{\Gamma}}{\mathrm{d}x\mathrm{d}y\mathrm{d}\varphi}, \\ \overline{I}_2 &= \left(\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \int_{\pi}^{2\pi} \mathrm{d}\varphi + \int_{-1}^0 \mathrm{d}x \int_{-1}^0 \mathrm{d}y \int_{\pi}^{2\pi} \mathrm{d}\varphi\right) \frac{\mathrm{d}^3\,\overline{\Gamma}}{\mathrm{d}x\mathrm{d}y\mathrm{d}\varphi}, \\ \overline{I}_3 &= \left(\int_0^1 \mathrm{d}x \int_{-1}^0 \mathrm{d}y \int_0^{\pi} \mathrm{d}\varphi + \int_{-1}^0 \mathrm{d}x \int_0^1 \mathrm{d}y \int_0^{\pi} \mathrm{d}\varphi\right) \frac{\mathrm{d}^3\,\overline{\Gamma}}{\mathrm{d}x\mathrm{d}y\mathrm{d}\varphi}, \\ \overline{I}_4 &= \left(\int_0^1 \mathrm{d}x \int_{-1}^0 \mathrm{d}y \int_{\pi}^{2\pi} \mathrm{d}\varphi + \int_{-1}^0 \mathrm{d}x \int_0^1 \mathrm{d}y \int_{\pi}^{2\pi} \mathrm{d}\varphi\right) \frac{\mathrm{d}^3\,\overline{\Gamma}}{\mathrm{d}x\mathrm{d}y\mathrm{d}\varphi}. \end{aligned}$$

Triple products are interesting because they are odd under time-reversal, T. However, in particle decays a TP can also stem from FSI, where all interactions conserved T. A genuine T-violating TP is given by comparing the TP value of a pair of CP-conjugate processes: then, assuming CPT invariance, a powerful tool for displaying CP violation is the investigation of TP asymmetries in a process and its CP-conjugate [57, 26, 15]. The TP asymmetries are similar to the direct CP asymmetries in this respect: they are both obtained by comparing a signal in a given decay with the corresponding signal in the CP-transformed process, and both are nonzero only if there are two interfering decay amplitudes. There is an important difference between the two. Denoting ϕ and δ as the relative weak and strong phases, respectively, between the two interfering amplitudes, the signal for direct CP violation given by Eq. (1.31) is proportional to $\sin\phi\sin\delta$ [40], while that for the CP-violating TP asymmetry is

$$\mathcal{A}_{\rm TP}^{CP} \propto \sin\phi\cos\delta \tag{2.42}$$

and TP asymmetries are maximal when the strong-phase difference vanishes [57].

The two TP asymmetries given in Eq. (2.38) and Eq. (2.41) in terms of transversity amplitudes are odd under time-reversal; however, to have a genuine CP-violating asymmetry we need to sum the TP asymmetry of the $B_s^0\to\phi\phi$ decay with the one of its CP conjugate, $\bar{B}_s^0\to\phi\phi$ [26, 15]:[4]

$$\begin{aligned}\mathcal{A}_{\rm TP,1}^{CP} &\equiv \mathcal{A}_{\rm TP}^{(1)} + \bar{\mathcal{A}}_{\rm TP}^{(1)} \\ &= -\frac{2\sqrt{2}}{\pi}\frac{K_6+\bar{K}_6}{|A_0|^2+|A_\parallel|^2+|A_\perp|^2+|\bar{A}_0|^2+|\bar{A}_\parallel|^2+|\bar{A}_\perp|^2},\end{aligned} \tag{2.43}$$

$$\begin{aligned}\mathcal{A}_{\rm TP,2}^{CP} &\equiv \mathcal{A}_{\rm TP}^{(2)} + \bar{\mathcal{A}}_{\rm TP}^{(2)} \\ &= -\frac{4}{\pi}\frac{K_4+\bar{K}_4}{|A_0|^2+|A_\parallel|^2+|A_\perp|^2+|\bar{A}_0|^2+|\bar{A}_\parallel|^2+|\bar{A}_\perp|^2}.\end{aligned} \tag{2.44}$$

We can infer information about CP-violating phases through these two asymmetries. In Equation (2.43) and Eq. (2.44) the two time-dependent asymmetries are proportional to the *untagged* sums $K_j+\bar{K}_j$ of the time-dependent functions in Tab. 2.4 with $j=4,6$; the term "untagged" means that we do not *tag* (*i. e.*, distinguish) the flavor of the B_s^0 meson at its production. We sum over the decay rates of the B_s^0 and the $\bar{B}_s^0$ mesons, assuming that they are equally produced (assumption that holds at Tevatron). We find [15]:

$$\begin{aligned}\mathcal{A}_{\rm TP,N}^{CP}(t) &\propto K_j(t)+\bar{K}_j(t) \\ &= e^{-\Gamma t}\Big[\mathfrak{Im}(A_\perp A_k^\star - \bar{A}_\perp\bar{A}_k^\star)\cosh(\Delta\Gamma_s t/2) \\ &- \big(\mathfrak{Im}(A_\perp\bar{A}_k^\star - \bar{A}_\perp A_k^\star)\cos 2\phi_M^s + \mathfrak{Re}(A_\perp\bar{A}_k^\star + \bar{A}_\perp A_k^\star)\sin 2\phi_M^s\big)\sinh(\Delta\Gamma_s t/2)\Big],\end{aligned} \tag{2.45}$$

[4] We sum the TP of $B_s^0\to\phi\phi$ and $\bar{B}_s^0\to\phi\phi$ instead of doing their difference, because the $K_4(t)$ and $K_6(t)$ involve the amplitude $A_\perp$, which changes sign under CP transformation; then $K_i(t)+\bar{K}_i(t)\propto\mathfrak{Im}(A_\perp A_j)(t) - \mathfrak{Im}(\bar{A}_\perp\bar{A}_j)(t)$, with $i=4,6$ and $j=\parallel,\perp$, respectively.

We denote with ϕ_k and δ_k respectively the weak and the strong phases of A_k as in Eq. (2.15). The term proportional to $\cos(\Delta\Gamma_s t/2)$ involves the decay amplitudes as follows:

$$\begin{aligned}\mathfrak{a}_j &\equiv \mathfrak{Im}(A_\perp A_k^\star - \bar{A}_\perp \bar{A}_k^\star) \\ &= \Big(|A_\perp||A_k| + |\bar{A}_\perp||\bar{A}_k|\Big)\cos(\delta_\perp - \delta_k)\sin(\phi_\perp - \phi_k) \\ &+ \Big(|A_\perp||A_k| - |\bar{A}_\perp||\bar{A}_k|\Big)\sin(\delta_\perp - \delta_k)\cos(\phi_\perp - \phi_k).\end{aligned} \tag{2.46}$$

and the term proportional to $\sin(\Delta\Gamma_s t/2)$ involves the mixing phase as well:

$$\begin{aligned}\mathfrak{b}_j &\equiv \mathfrak{Im}(A_\perp \bar{A}_k^\star - \bar{A}_\perp A_k^\star)\cos 2\phi_M^s + \mathfrak{Re}(A_\perp \bar{A}_k^\star + \bar{A}_\perp A_k^\star)\sin 2\phi_M^s \\ &= \Big(|A_\perp||\bar{A}_k| + |\bar{A}_\perp||A_k|\Big)\cos(\delta_\perp - \delta_k)\sin(2\phi_M^s - (\phi_\perp + \phi_k)) \\ &+ \Big(|A_\perp||\bar{A}_k| - |\bar{A}_\perp||A_k|\Big)\sin(\delta_\perp - \delta_k)\cos(2\phi_M^s - (\phi_\perp + \phi_k)).\end{aligned} \tag{2.47}$$

The measurements of TP asymmetries discussed in this thesis are time-integrated:

$$\int_0^\infty (K_j(t) + \bar{K}_j(t))\mathrm{d}t \tag{2.48}$$

then [15]:

$$\mathcal{A}_{\mathrm{TP,N}}^{CP} \propto \mathfrak{a}_j - \mathfrak{b}_j y_s, \tag{2.49}$$

where $y_s = \Delta\Gamma_s/\Gamma_s$. It is interesting to note that time-integrated asymmetries for untagged B_s^0 decays are not suppressed due to fast B_s^0-$\bar{B}_s^0$ oscillations by $\Gamma/\Delta m_s \approx 0.0005$, as they would be for time-dependent terms behaving like $\cos(\Delta m_s t)$ and $\sin(\Delta m_s t)$.

In the SM, for $B_s^0 \to \phi\phi$ decays we expect neither direct *CP* violation ($|A_i| = |\bar{A}_i|$ for $i = 0, \|, \perp$) nor interference *CP* violation, and a vanishing mixing phase ($\phi_\perp - \phi_k = 0$ and $2\phi_M^s - (\phi_\perp + \phi_k) = 0$). Suppose first that there is new physics only in the B_s^0-$\bar{B}_s^0$ mixing amplitude, *i. e.*, $2\phi_M^s - (\phi_\perp + \phi_k) \neq 0$. In this case, the first term $\mathfrak{a}_j$ in Eq. (2.49) is zero, but the second is nonzero. However, the asymmetry is suppressed by $y_s \sim 0.1$. The second possibility is that there is new physics in the decay, then both terms $\mathfrak{a}_j$ and $\mathfrak{b}_j$ in Eq. (2.49) are nonzero. And of course one can have new physics both in mixing and in decay, resulting in a nonzero asymmetry. However, the asymmetry is more sensitive to $\mathfrak{a}_j$ in both latter cases.

A particular situation reported in Ref. [26] occurs when $|A_i| = |\bar{A}_i|$, for $i = (0, \|, \perp)$, but $\phi_\perp - \phi_k \neq 0$. Then, the second line in Eq. (2.46) and Eq. (2.47) vanishes, and we have:

$$\mathcal{A}_{\mathrm{TP,N}}^{CP} \propto 2|A_\perp||A_k|\cos(\delta_\perp - \delta_k)\Big[\sin(\phi_\perp - \phi_k) + y_s \sin(2\phi_M^s - (\phi_\perp + \phi_k))\Big], \tag{2.50}$$

Consider the measured polarization amplitudes, $|A_i| \approx 0.33$ for each polarization state [2, 23], and the assumption $|\cos(\delta_\perp - \delta_0)| \approx |\cos(\delta_\perp - \delta_\|)| \approx 1$, which should hold to a good approximation based on $U(3)$ symmetry and measured value in $B^0 \to \phi K^\star(890)^0$ decays; taking into account the normalization factors, the size of the two asymmetries are about:

$$|(30 \div 40)\sin(\phi_\perp - \phi_0) - (3 \div 4)\sin(2\phi_M^s - (\phi_\perp + \phi_k))|\% \tag{2.51}$$

and in the SM, we expect $\mathcal{A}^{CP}_{\mathrm{TP},1} \approx \mathcal{A}^{CP}_{\mathrm{TP},2} = 0$ within 1% [51]. Then, a measurement of asymmetries larger than 1% is a clear indication of new physics. An analysis of these time-integrated triple-product asymmetries was never performed before, and it is presented here for the first time for the $B_s^0 \to \phi\phi$ decay.

2.4 Measurements' strategies

In the following we present the strategies to measure the *CP* violation in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\phi$ decays. For each analyses, we also briefly outline the current experimental situation as of the writing of this thesis.

2.4.1 $B_s^0 \to J/\psi\phi$ analysis

The first measurements of the *CP*-violating phase in $B_s^0 \to J/\psi\phi$ decays was finalized in 2008 by the CDF experiment [16]. It showed a mild, 1.5σ discrepancy from the SM. It was intriguing that the D0 experiment, few months later, found a similar, and consistent effect [17]. Indeed, the combination yielded a 2.2σ deviation from the SM [58]. This attracted some interest, further enhanced by the recent dimuon asymmetry results from the D0 collaboration [22], which probe the same dynamics as $B_s^0 \to J/\psi\phi$, and report suggestive, even more significant anomalies (see Sect. 1.3.1). In 2010, both the CDF and D0 collaborations updated their measurements of $B_s^0 \to J/\psi\phi$ time-evolution. CDF used an event sample based on $5.2\,\mathrm{fb}^{-1}$ of integrated luminosity [18], D0 on $8\,\mathrm{fb}^{-1}$ of integrated luminosity [19]. The results from both experiments, although consistent with the previous ones, showed an improved agreement with the SM. Also LHCb began recently to contribute, with a published measurement on only $340\,\mathrm{pb}^{-1}$ of data [20], which appears already very competititve. A more precise result with $1\,\mathrm{fb}^{-1}$ of data has been presented in February 2012 [59], but it is still preliminary. LHCb provides also the first experimental determination of the sign of $\Delta\Gamma_s$, by analyzing the *P*-*S* wave interference in bins of the K^+K^- mass to distinguish among the equivalent for β_s and $\Delta\Gamma_s$ due to the symmetries discussed in Sect. 2.2.2. The $\Delta\Gamma_s$ sign turns to be consistent with $\Gamma_L > \Gamma_H$ [42]. Table 2.7 reports a summary of the current experimental status along with a comparison of key experimental parameters.

We present the $B_s^0 \to J/\psi\phi$ analysis updated to the full sample collected by CDF during its operations, which corresponds to an integrated luminosity of $9.6\,\mathrm{fb}^{-1}$, almost doubling the statistic of the latest CDF measurement. The measurement of the phase β_s relies on an analysis of the time-evolution of the $B_s^0 \to J/\psi\phi$ decay in which decays from mesons produced as B_s^0 or $\bar{B}_s^0$ are studied independently, and the *CP*-parity of the final state is statistically determined using angular distributions. The analysis can be dissected in four main steps: selection and reconstruction of the signal event sample; preparation of the analysis tools; fit to the time-evolution; statistical procedure to extract results and uncertainties.

Parameter	LHCb ($340\,\text{pb}^{-1}$) [20]	D0 ($8\,\text{fb}^{-1}$) [19]	CDF ($5.2\,\text{fb}^{-1}$) [18]
β_s	-0.07 ± 0.10	$0.28^{+0.18}_{-0.19}$	$\approx 0.27 \pm 0.25$
$\Delta\Gamma_s$ [ps^{-1}]	0.123 ± 0.031	$0.163^{+0.065}_{-0.064}$	0.075 ± 0.036
$\sigma_t(B_s^0)$ [fs]	≈ 50	≈ 100	≈ 90
$\sigma_m(B_s^0)$ [MeV/c^2]	≈ 7	≈ 30	≈ 10
Effective tagging power	$\approx 2.1\%$	$\approx 2\%$	$\approx 4.7\%$
Signal yield	8 300 ($t > 0.3$ ps)	5 600	6 500
S/B at peak	33/1 ($t > 0.3$ ps)	1/3	2/1

Table 2.7: Summary of current experimental status and comparison of key experimental parameters. The D0 analysis uses an additional constraint from a theory assumption in the fit to the phase. The parameters $\sigma_t(B_s^0)$ and $\sigma_m(B_s^0)$ are the mean resolutions on the measurement of the B_s^0 decay time and mass, respectively. The effective tagging power is the measurement of the capability to distinguish the production of a B_s^0 from a $\bar{B}_s^0$ meson. The symbol S/B stands for the signal to background ratio.

For this update we follow the general analysis strategy used for earlier CDF publications [18, 16]. In particular the interesting physical parameters are extracted from an unbinned likelihood fit to the B_s^0 candidate mass, the angular variables in the transversity basis, the proper decay time, and flavor-tagging information. We adopt the same fitting code employed in the latest CDF measurement [18] with minimal simplifications and updated acceptance maps and other needed inputs. Signal contributions in the $B_s^0 \to J/\psi K^+K^-$ final state other than $B_s^0 \to J/\psi\phi$ signal itself are taken into account assuming an S-wave state for the K^+K^- system. Since the the K^+K^- mass m is not used as a input to the fit, its contribution is integrated as described in Sect. 2.2.1, assuming a relativistic Breit-Wigner shape with mass dependent width for the P-wave $\phi(1020)$ contribution and a flat shape for the S-wave contribution, given the small window used in the K^+K^- spectrum, $[1.009, 1.028]\,\text{GeV}/c^2$. Several improvements in the analysis are introduced with respect to the previous CDF measurements, such as a new calibration of a tagging algorithm (Sect. 4.2.1); an original, accurate determination of previously-neglected physics background that mimic signal (Sect. 7.4); and a test of the fit's reliability through a new simulation of the full $(S+P)$-wave decay (Sect. 5.5.2).

The $B_s^0 \to J/\psi\phi$ analysis exhibits the well known problem of discrete symmetries in the decay rate (Sect. 2.2.2) that corresponds to four ambiguities in the extracted physical parameters β_s and $\Delta\Gamma_s$, which are only marginally lifted by the $(S+P)$-wave interference. Half of the solutions are eliminated using the difference of B_s^0's and $\bar{B}_s^0$'s time evolution. Flavor tagging, furthermore, improves the statistical behaviour of the likelihood for $B_s^0 \to J/\psi\phi$ decays in the presence of limited event samples. Two algorithms are used for flavor tagging. In the present analysis, one of them, the opposite-side-tagging algorithm (Sect. 4.2), has been recalibrated using data corresponding to the final dataset. On the other hand, it was not possible to reliable calibrate the second algorithm, the same-side-kaon-tagging (Sect. 4.3), because the available

statistics of B_s^0 flavor specific decays from the hadronic trigger for the latest part of the data is limited. We employ this tagger only for the first part of the data, corresponding to $5.2\,\text{fb}^{-1}$ of integrated luminosity, where a reliable calibration is availble [18].

The complexity of the fit and the irreducible symmetries of the likelihood make the extraction of proper confidence intervals challenging from the simple fit results. A thorough work of simulation is needed to construct correct confidence regions and finally extract the results.

The $B_s^0 \to J/\psi\phi$ analysis is presented in the second part of this dissertation after the description of the CDF detector in Chap. 3.

2.4.2 $B_s^0 \to \phi\phi$ analysis

First evidence of the $B_s^0 \to \phi\phi$ decay was obtained in 2005 by CDF [60]. In 2010, CDF provides also the first measurement of its polarization amplitudes [2]. The latter is of particular interest because it gives insight in the so-called *polarization puzzle*, related to the amplitudes hierarchy in $b \to s$ penguin processes: the V–A structure of charged weak currents leads to the expectation of a dominant longitudinal polarization [61, 62, 63], while approximately equal longitudinal and transverse polarizations have been measured in $b \to s$ penguin-dominated B^0 and B^+ decay modes [64, 65, 66, 67]. This is explained in the SM by including either nonfactorizable penguin-annihilation effects [52] or FSI [68]. Recent theoretical predictions [61, 62, 63] indicate a longitudinal fraction in the 40%–70% range, when phenomenological parameters are adjusted to accommodate present experimental data. Explanations involving new physics in the $b \to s$ penguin process have also been proposed [69, 70]. Additional experimental information in B_s^0 penguin-dominated decays help distinguishing the various solutions.

Using $3\,\text{fb}^{-1}$ of integrated luminosity, CDF measures the polarization amplitudes reported in Tab. 2.8 with about 300 $B_s^0 \to \phi\phi$ candidates [2], which are confirmed by a recent measurement by LHCb [23]; the polarization measurement of another $b \to s$ decay of the B_s^0 meson, the $B_s^0 \to K^{0\star}\bar{K}^{0\star}$ decay, exhibits a similar result [71]. Although the pattern agrees with observations in the B^0 sector, longitudinal polarization is even more suppressed in B_s^0 penguin decay. Such amplitudes' hierarchy motivate the search for new physics through the measurement of *CP*-violating asymmetries.

Amplitude	$B_s^0 \to \phi\phi$ (CDF) [2]	$B_s^0 \to \phi\phi$ (LHCb) [23]	$B_s^0 \to K^{0\star}\bar{K}^{0\star}$ (LHCb) [71]
$\lvert A_0\rvert^2$	$0.348 \pm 0.041(\text{stat}) \pm 0.021(\text{syst})$	$0.365 \pm 0.022(\text{stat}) \pm 0.012(\text{syst})$	$0.31 \pm 0.12(\text{stat}) \pm 0.04(\text{syst})$
$\lvert A_\parallel\rvert^2$	$0.287 \pm 0.043(\text{stat}) \pm 0.011(\text{syst})$	$0.344 \pm 0.33(\text{stat}) \pm 0.016(\text{syst})$	$0.31 \pm 0.16(\text{stat}) \pm 0.06(\text{syst})$
$\lvert A_\perp\rvert^2$	$0.365 \pm 0.044(\text{stat}) \pm 0.027(\text{syst})$	$0.291 \pm 0.024(\text{stat}) \pm 0.010(\text{syst})$	$0.38 \pm 0.11(\text{stat}) \pm 0.04(\text{syst})$

Table 2.8: Summary of current experimental status of polarization amplitude measurements in the $b \to s$ decays of the B_s^0 meson.

The CDF statistic of the $B_s^0 \to \phi\phi$ data sample are not sufficient for measuring the time-dependent *CP* asymmetry as the case of the $B_s^0 \to J/\psi\phi$ decay. However, an investigation of the time-integrated triple product asymmetries is feasible. We provide the world's first measurement ot such asymmetries in this decay mode. We employ the same dataset collected by the CDF hadronic trigger used for the polarization measurement (Chap. 9). The measurement of the asymmetry is implemented in a fit to the mass distribution of the sample (Chap. 10). We split the dataset in two samples according to the sign of the triple products of each candidate. We fit simultaneously the two subsamples to extract the signal yields and compute the TP asymmetry. In this way, without distinguish the B_s^0 flavor at production, we measure the true *CP*-violating asymmetry, since B_s^0 and $\bar{B}_s^0$ mesons are produced in the same amount at Tevatron.

The $B_s^0 \to \phi\phi$ analysis is presented after the description of the $B_s^0 \to J/\psi\phi$ analysis.

Chapter 3

Experimental Apparatus

This chapter briefly describes the Tevatron collider and the CDF II detector. We focus on the subsystems most important for the analyses presented in this thesis the tracking and the particles identification systems, and the muon's detectors.

3.1 The Fermilab's Tevatron collider

The Tevatron is a circular synchrotron of 1 km in radius, located at the Fermi National Accelerator Laboratory (FNAL or Fermilab), about 50 km West from Chicago (IL), USA, that has ended its operations on September 30$^{\text{th}}$ 2011, after 28 years of activity. It has produced about 10^{13} collisions between bunches of protons against bunches of antiprotons accelerated in opposite direction, both at energies of 980 GeV. The $\sqrt{s} = 1.96$ TeV energy available in the center-of-mass after the collision was the world's highest before the beginning of LHC operations.

An important feature of a collider is the instantaneous luminosity, $\mathscr{L}$, that is the coefficient of proportionality between the rate of a given process and its cross-section σ:

$$\frac{dN}{dt}\ [\text{events/s}] = \mathscr{L}\left[\text{cm}^{-2}\text{s}^{-1}\right] \cdot \sigma\left[\text{cm}^{2}\right].$$

The time-integral of the luminosity (integrated luminosity) is therefore a measure of the expected number of events, N, produced in a finite time T:

$$N(T) = \int_0^T \mathscr{L}\ \sigma\ dt.$$

It is customary to report the amount of data delivered by the collider or collected by detectors in terms of integrated luminosity, fixing $\sigma = 1$. Assuming an ideal head-on $p\bar{p}$ collision with no crossing angle between the beams, the instantaneous luminosity at the Tevatron is defined as

$$\mathscr{L} = 10^{-5}\frac{N_p N_{\bar{p}} B f \beta\gamma}{2\pi\beta^{\star}\sqrt{(\epsilon_p + \epsilon_{\bar{p}})_x(\epsilon_p + \epsilon_{\bar{p}})_y}}\ F\left(\sigma_z/\beta^{\star}\right) \qquad \left[10^{30}\text{cm}^{-2}\text{s}^{-1}\right]$$

where N_p ($N_{\bar{p}}$) is the average number of protons (antiprotons) in each bunch ($N_p \approx 8.1 \cdot 10^{11}$ and $N_{\bar{p}} \approx 2.9 \cdot 10^{11}$), B (36) is the number of bunches per beam circulating into the ring, f (47.713 kHz) is the revolution frequency, $\beta\gamma$ is the relativistic factor of the Lorentz boost (1045.8 at 980 GeV), F is an empiric form factor which depends on the ratio between the longitudinal width of the bunch ($\sigma_z \approx 60$ cm) and the "betatron function" calculated at the interaction point

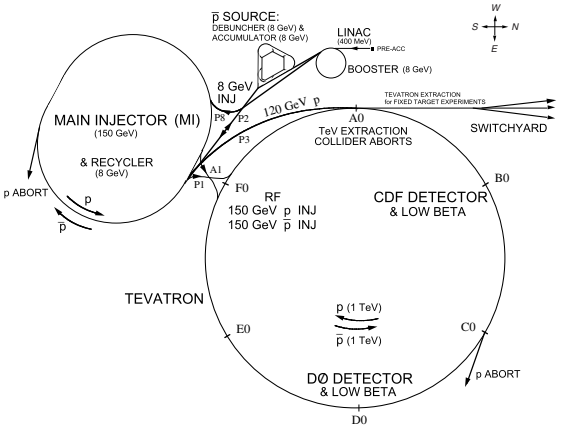


Figure 3.1: Sketch of the Fermilab's accelerators system.

($\beta^\star \approx 31\,\text{cm}$), and finally ϵ_p ($\epsilon_{\bar{p}}$) is the 95% normalized emittance of the proton (antiproton) beam ($\epsilon_p \approx 18\pi$ mm mrad and $\epsilon_{\bar{p}} \approx 13\pi$ mm mrad after injection). The luminosity is determined only through quantities that depend on the acceleration performance: at the Tevatron the most limiting factor of the luminosity is the availability of antiprotons ($N_{\bar{p}}$) since it is difficult to produce, to compact into bunches and to transfer them efficiently through all the acceleration stages.

The Tevatron is the final sector of a more complex accelerators system, entirely represented in Fig. 3.1, which provides beam to different typologies of experiments ($p\bar{p}$ collisions, fixed-target, test beam facilities). In the following sections we describe the procedure for obtaining a continuous period of collider operation, called a *store*, using the same collection of protons and antiprotons. Further details can be found in Refs. [72, 73].

3.1.1 The proton beam

The protons production starts with turning hydrogen gas into H^- ions. The gas, in the molecular state H_2, is placed in a container lined with molybdenum electrodes: a matchbox-sized, oval-shaped cathode and a surrounding anode, separated by 1 mm and held in place by glass ceramic insulators. A magnetron generates a plasma that forms H^- ions near the metal surface. These are then extracted out of the container by a 750 keV electrostatic field, applied by a commercial Cockcroft-–Walton generator, and injected into a 150 m long linear accelerator (*Linac*) which increases their energy up to 400 MeV. A carbon foil is used to strip the electrons from the H^- before the resulting protons are injected into the *Booster*. The Booster is a rapid cycling synchrotron (radius of 75 m) which accelerates the protons up to 8 GeV and compacts them into bunches of about 9×10^{12} particles each. The bunches are then transfered into the *Main Injector*, a synchrotron which brings their energy up to 150 GeV, and finally into the Tevatron where superconducting magnets keep them on an approximately circular orbit waiting for the antiproton beam to be injected.

3.1.2 The antiproton beam

While the energy of the protons bunches circulating in the Main Injector reaches 120 GeV, they are slammed to a rotating 7 cm thick nickel target. Spatially wide-spread antiprotons are produced and focused into a beam via an appropiate magnetic lens which separates them from other charged interaction products. The emerging antiprotons have a bunch structure similar to that of the incident protons and are stored in a *Debuncher*. This is a storage ring where the momentum spread of the antiprotons is reduced while maintaining a constant energy of 8 GeV, via stochastic cooling stations. Many cycles of Debuncher cause the destruction of the bunch structure which results in a continuous beam of antiprotons. At the end of the process the monochromatic antiprotons are transferred into the *Accumulator* which is a triangle-shaped storage ring where they are further cooled and stored until the cycles of the Debuncher are completed. When a current sufficient to create 36 bunches with the required density is available ($\approx 3 \times 10^{12}$ particles per bunch), the antiprotons are injected into the Main Injector, here accelerated up to 150 GeV and finally transferred into the Tevatron where 36 bunches of protons are already circulating in opposite direction.

3.1.3 The collisions

When 36 bunches of both protons and antiprotons are circulating in the Tevatron, the energy of the machine is increased in about 10 s from 150 to 980 GeV and the beams collide every 396 ns at the two interaction points, DØ and BØ, where the DØ and the CDF II detectors are respectively located. Special quadrupole magnets (*low-β squeezers*) located at both extremities of the detectors along the beam pipe "squeeze" the beam in the longitudinal direction to maximize the luminosity inside the detectors. Then the beam transverse profile is shaped to its optimized configuration by mean of iron plates which act as collimators and sweep away the transverse beam halo. The interaction region thus achieves a roughly Gaussian distribution in both transverse ($\sigma_T \approx 30\,\mu$m) and longitudinal ($\sigma_z \approx 28$ cm) planes with its center in the nominal interaction point. When the beam profile is narrow enough and the conditions are safely stable the detectors are powered and the data taking starts.

The number of overlapping inelastic interactions N for each bunch crossing is a Poisson-distributed variable that depends on the instantaneous luminosity. The observed distribution of the multiplicity of interaction vertexes yields $\bar{N} \approx 0.2$, 1.0, 2.0, and 6.0 for respectively, $\mathscr{L} \approx 10^{31}$, 5×10^{31}, 10^{32}, and 3×10^{32} luminosities. In the last years of operations, when collisions start the peak luminosity is almost regularly above $3.2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. While collisions are taking place the luminosity decreases exponentially because of the beam-gas and beam-halo interactions.[1] In the meantime, antiproton production and storage continues. When the antiproton stack is sufficiently large and the circulating beams are degraded the detector high-voltages are switched off and the store is dumped. The beam is extracted via a switch-yard and

[1] The decrease is about a factor of 3 (5) for a store of ≈ 10 (20) hrs.

Date		$\sqrt{s}$ [TeV]	$\mathscr{L}$ [$\mathrm{cm^{-2}s^{-1}}$]	$\int \mathscr{L}\, dt$ [$\mathrm{pb^{-1}}$]
Mar 1983	End of the construction	–	–	–
Jul 1983	Proton energy: 512 GeV	–	–	–
Oct 1983	Fixed-target program	–	–	–
Feb 1984	Proton energy: 800 GeV	–	–	–
Oct 1985	First $p\bar{p}$ collisions	1.6	10^{24}	–
Oct 1986	Proton energy 900 GeV	–	–	–
Jun 1988–May 1989	Run 0	1.8	2×10^{30}	4.5
Aug 1992–Feb 1996	Run I	1.8/0.63	28×10^{30}	180
Aug 2000	Beam energy: 980 GeV	–	–	–
Mar 2001–Sept 2011	Run II	1.96	$(5 - 440) \times 10^{30}$	12000
Jun 2011	Best performances	1.96	440×10^{30}	11350

Table 3.1: Chronological overview of the Tevatron operation and performance. The fourth column reports the peak luminosity. The fifth column reports the delivered integrated luminosity. The last row shows the best performances achieved in therm of peak luminosity and the corresponding integrated luminosity at that time.

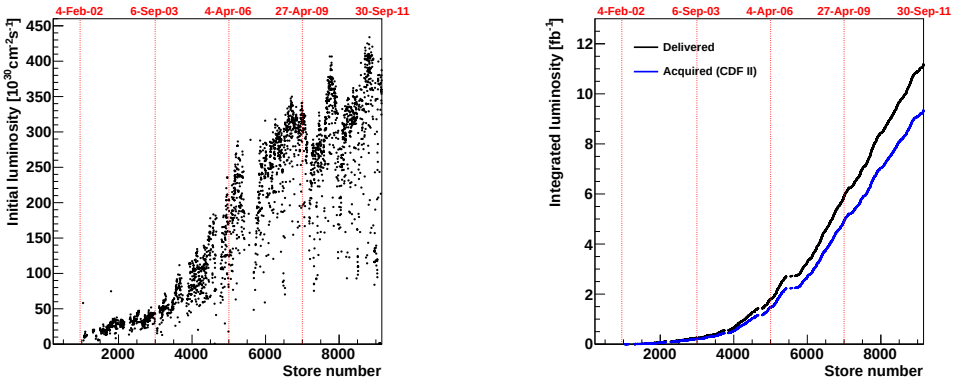
sent to an absorption zone. Beam abortion can occur also accidentally when the temperature of a superconducting magnet shift above the critical value and a magnet quenches destroying the orbit of the beams. The time between the end of a store and the beginning of collisions of the next one is typically ≈ 1 hr; during this time calibrations of the sub-detectors and test runs with cosmics are performed.

3.1.4 Run II performances and achievements

The Tevatron was commissioned in 1983 as the first large-scale superconducting synchrotron in the world and, since then, various periods of operations occurred. Each period of Tevatron collider operations is conventionally identified as a Run. Table 3.1 contains a summary of the Tevatron operations and performance since its construction.

The Tevatron $p\bar{p}$ collider has been the centerpiece of the world's high energy physics program for almost a quarter of century, until it was overtaken by the LHC first in the colliding beam energy, in November 2009, and then in terms of colliding beam luminosity, in April 2011. In this section we briefly describe the achievements of such a long history. We focus on the last period of Tevatron operations, since March 2001 till September 2011, which is commonly referred to as *Run II*. Indeed, the analyses reported in this thesis use data collected in Run II.

Figure 3.2 (a) shows the collider performance history in terms of instantaneous luminosity. At the end of the Run II, typical Tevatron luminosities were well constantly above $3.2 \times 10^{32}\,\mathrm{cm^{-2}s^{-1}}$, with record peak above $4.4 \times 10^{32}\,\mathrm{cm^{-2}s^{-1}}$, thus ultimately exceeding its initial luminosity design by a factor of 400. The plot also shows the steadily increase occurred during the years after numerous improvements, some were implemented during operation and oth-



(a) (b)

Figure 3.2: Initial (a) and integrated (b) luminosity as a function of store number/time.

ers were introduced during regular shutdown periods. They took place in all accelerators and addressed all parameters affecting luminosity, proton and antiproton intensities, emittances, optics functions, bunch length, losses, reliability and availability, etc. The Tevatron integrated luminosity has greatly progressed over the years (see Fig. 3.2 (b)) and, at the end of Run II operations, a total of $12\,\mathrm{fb}^{-1}$ of $p\bar{p}$ collisions has been delivered to both CDF and DØ.

3.2 The CDF II detector

The CDF II detector is an azimutally and forward-backward symmetric apparatus designed to study $p\bar{p}$ collisions at the Tevatron, installed at the BØ interaction point (see Fig. 3.1). It was designed, built, and operated by a team of physicists, technicians, and engineers that spans 60 institutions of 13 countries. Several upgrades modified the design of the original facility commissioned in 1985.[2] The most extensive upgrade started in 1995 and led to the current detector whose operation is generally referred to as CDF II.

The CDF II detector was in operation between 2001 and 2011. It is a multi-purpose cylindrical-shaped detector, about 15 m in length and 15 m in diameter, composed of several specialized subsystems, each one designed to perform a different task, as shown in Fig. 3.3. High resolution three-dimensional charged particle tracking is achieved through an integrated system consisting of three silicon inner subdetectors and a large outer drift chamber, all contained in a superconducting solenoid, 1.5 m in radius and 4.8 m in length, which generates a 1.4 T magnetic field parallel to the beam axis. Outside the magnet a calorimeter system provides electrons and photons identification and finely segmented sampling of energy flow coming from final state particles. A system of muon chambers plus scintillators, is instead used to track and identify muons which pass through the calorimeters interacting as minimum-ionizing-particles.

[2]Originally, the CDF acronym was meant for Collider Detector Facility.

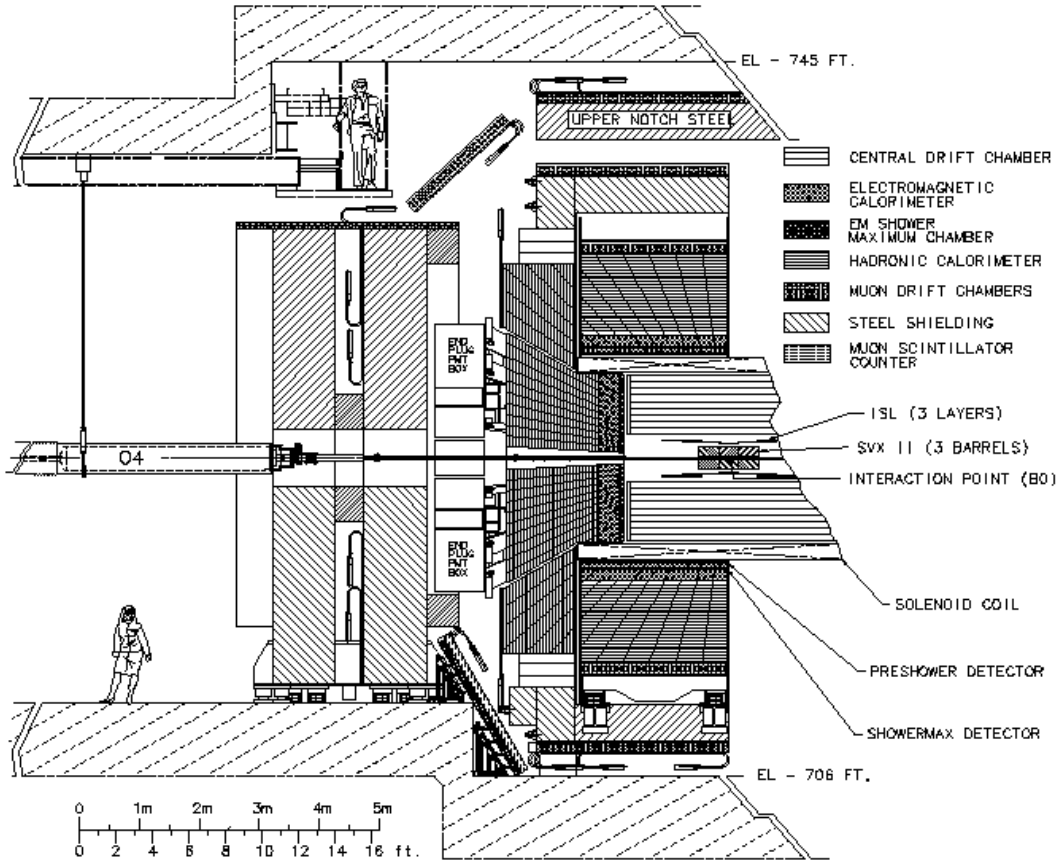


Figure 3.3: Elevation view of one half of the CDF II detector.

In the following sections we describe the general features of the most important subsystems for the analyses reported in this thesis, the tracking system (Sect. 3.2.1), the particle identification (Sect. 3.2.2) and the muon's detectors (Sect. 3.2.3). A comprehensive description of the entire apparatus is given in Ref. [74]. Before, we need to define few conventions.

CDF II employs a right-handed Cartesian coordinates system with the origin in the BØ interaction point, assumed coincident with the center of the drift chamber. The positive z axis lies along the nominal beam-line pointing toward the proton direction. The (x, y) plane is therefore perpendicular to either beams, with positive y axis pointing vertically upward and positive x axis in the horizontal plane of the Tevatron, pointing radially outward with respect to the center of the ring. A cylindrical (r, ϕ, z) coordinates system is particularly convenient to describe the detector geometry. Throughout this thesis, *longitudinal* (or *axial*) means parallel to the proton beam direction (*i. e.* to the z axis) and *transverse* means perpendicular to the proton direction (*i. e.* in the (x, y) or (r, ϕ) plane).

Since the protons and antiprotons are composite particles, the actual interaction occurs between their individuals partons (valence or sea quarks and gluons). Each parton c arries a varying fraction of the (anti)proton momentum, not known on a event-by-event basis.

As a consequence of the possible imbalance in the longitudinal components of the momenta of interacting partons, possible large velocities along z for the center-of-mass of the parton-level interaction may occur. In the hadron collisions environment, it is customary to use a variable invariant under z boosts as an unit of relativistic phase-space, instead of the polar angle θ. This variable is the *rapidity*, defined as

$$Y = \frac{1}{2} \ln\left(\frac{E + p\cos\theta}{E - p\cos\theta}\right),$$

where $(E, \boldsymbol{p})$ is the energy-momentum four-vector of the particle. Under a $\hat{\boldsymbol{z}}$ boost to an inertial frame with velocity β, the rapidity of a particle transforms linearly, according to $Y \to Y' = Y + \tanh^{-1}\beta$, therefore Y is invariant since $dY \equiv dY'$. However, a measurement of rapidity still requires a detector with accurate identification capabilities because of the mass term entering E. For practical reasons, it is often preferred to replace Y with its approximate expression η in the ultra-relativistic limit (usually valid for products of high-energy collisions):

$$Y \overset{p \gg m}{\longrightarrow} \eta + \mathcal{O}(m^2/p^2),$$

where the *pseudo-rapidity*,

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right),$$

is only function of the momenta. As the event-by-event longitudinal position of the actual interaction is distributed around the nominal interaction point with 28 cm r.m.s width, it is useful to distinguish the *detector pseudo-rapidity*, η_{det}, measured with respect to the $(0, 0, 0)$ nominal interaction point, from the *particle pseudo-rapidity*, η, measured with respect to the z position of the real vertex where the particle originated.

3.2.1 Tracking system

Tracking refers to the measurement of charged particles trajectories within the detector volume. This allows the determination of the charge and the momentum of a particle, which are essential for the analyses presented in this thesis, since both vector mesons from the B_s^0 meson decay are reconstructed through particle pairs of opposite charge. In particular, the excellent CDF II tracking performance results in a mass resolution of $\approx 9\,\text{MeV}/c^2$ for a B meson and a mass resolution of $\approx 1\,\text{MeV}/c^2$ for the $\phi(1020)$ resonance.

Within an uniform axial magnetic field, the trajectory of a charged particle produced with non-zero initial velocity in the bending plane of the magnet is described by an helix, which can be uniquely parameterized by the following set of equations:

$$\begin{cases} x = R\sin(2Cs + \phi_0) - (R + d_0)\sin\phi_0 \\ y = -R\cos(2Cs + \phi_0) + (R + d_0)\cos\phi_0 \\ z = z_0 + s\lambda \end{cases},$$

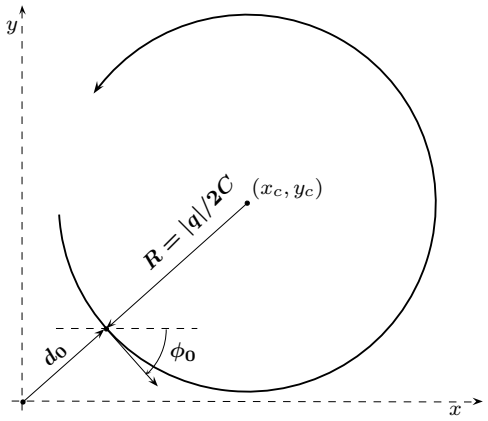


Figure 3.4: Schematic view of a positively charged track in the plane transverse to an axial magnetic field $\boldsymbol{B} = (0, 0, -B)$.

where, given the projected length along the track, s, one can find the corresponding (x, y, z) coordinates of the trajectory by means of five parameters (see Fig. 3.4):

C – signed half-curvature of the helix, defined as $C = q/2R$, where R is the radius of the helix and q is the charge of the track. This is directly related to the transverse momentum: $p_T = cB/(2|C|)$, where c is the speed of light and B is the magnetic field of the solenoid;

ϕ_0 – ϕ angle of the particle at the point of closest approach to the z axis;

d_0 – signed impact parameter, *i. e.* the radial distance of closest approach to the z axis, defined as $d_0 = q(\sqrt{x_c^2 + y_c^2} - R)$, where (x_c, y_c) are the coordinates of the center;

λ – the helix pitch, *i. e.* $\cot\theta$, where θ is the polar angle of the helix at the point of its closest approach to the z axis. This is directly related to the longitudinal component of the momentum: $p_z = p_T \cot\theta$;

z_0 – the z coordinate of the point of closest approach.

Charged particles leave small charge depositions as they pass through the alternative layers of the tracking system. Using a set of spatial measurements of these depositions ("hits"), pattern recognition algorithms can reconstruct the particle original trajectory measuring the five parameters of the helix that best match to the observed path in the tracking detector. At CDF II this is an integrated system consisting of three silicon inner sub-detectors and a large outer drift chamber, all contained in a 1.4 T magnetic field of a solenoid parallel to the beams and pointing in the negative z direction. The silicon detectors provide excellent impact parameter, azimuthal angle and z_0 resolution while the drift chamber provides excellent resolution of the curvature and ϕ_0. Together they provide a very accurate measurements of the helical paths of charged particles.

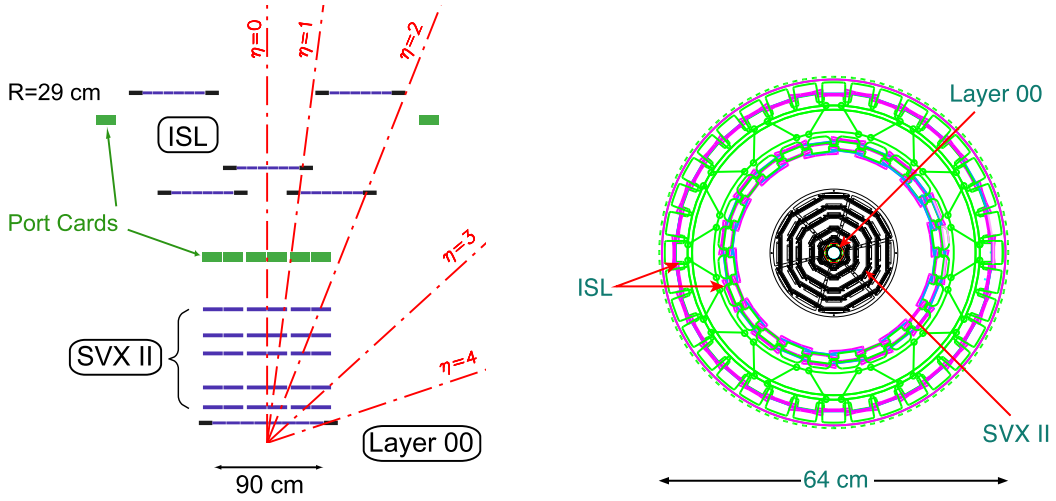


Figure 3.5: View of the CDF II silicon system, including the SVX II cooling bulkheads and ISL support structure, in the (r, z) (a) and (r, ϕ) (b) planes. The z scale is highly compressed.

3.2.1.1 The inner silicon tracker

The CDF II silicon tracking system, as shown in Fig. 3.5 in both (r, ϕ) and (r, z) projections, is composed of three approximately cilindrical coaxial subsystems: the *Layer00* (L00), the *Silicon VerteX detector* (SVX II) and the *Intermediate Silicon Layer* (ISL).

L00 [75] is the innermost subsystem and consists of a one layer of single-sided, AC-coupled, microstrip silicon sensors installed at radii of 1.35 and 1.62 cm on a mechanical structure in direct contact with the beam pipe. It provides full azimuthal and $|z| \lesssim 47$ cm longitudinal coverage. Longitudinally adjacent sensors ($0.84 - 1.46\,\text{cm} \times 7.84\,\text{cm}$) are ganged in modules of 15.7 cm active-length arranged into twelve partially-overlapping ϕ sectors, and six longitudinal barrels. These radiation-tolerant sensors are biased to $\mathcal{O}(500\,\text{V})$, which allows full depletion after $\mathcal{O}(5\,\text{MRad})$ integrated radiation doses. The strips are parallel to the beam axis allowing sampling of tracks in the (r, ϕ) plane. The inter-strip implant pitch of 25 μm with floating alternate strips results in 50 μm read-out pitch. The analog signals of the 13 824 channels are fed via fine-pitch cables, up $\sim$ 50 cm long, to the front-end electronics outside the tracking volume.

The SVX II [76] is a fine resolution silicon micro-strip vertex detector which provides five three-dimensional samplings of tracks at 2.45, 4.1, 6.5, 8.2 and 10.1 cm (or, depending on the ϕ sector, at 2.5, 4.6, 7.1, 8.7 and 10.6 cm) of radial distance from the beam with full pseudo-rapidity coverage in the $|\eta_{\text{det}}| \lesssim 2$ region. This corresponds to a length of $|z| \lesssim 96$ cm along the beam-line, sufficient to cover the $\sigma_z \approx 28$ cm longitudinal spread of the luminous region. The SVX II has a cylindrical geometry coaxial with the beam, and its mechanical layout is segmented in three 32 cm axial sections ("barrels") times twelve 30° azimuthal sectors ("wedges") times five

equally-spaced radial layers. A small overlap between the edges of adjacent azimuthal sectors helps wedge-to-wedge alignment (see Fig. 3.5 (b)). Sensors in a single layer are arranged into independent longitudinal read-out units, called "ladders". Each ladder comprises two, double-sided sensors and a multi-layer electronic board, all glued on a carbon-fiber support. Front-end electronics, biasing circuits, and fan-out are located on the board that serves the pair of sensors whose strips are wire-bonded together resulting in a 15 cm active length. At a given radial layer and azimuth, each barrel contains pairs of ladders stacked length-wise head-to-head to keep the read-out electronic at the two outside extremities of the barrel. The active surface consists of double-sided, AC-coupled, $7.5\,\text{cm} \times 1.5 - 5.8\,\text{cm}$ silicon sensors with microstrips implanted on a $300\,\mu$m thick, high resistivity bulk. Bias is applied through integrated poly-silicon resistors. On one side, all sensors have axial strips spaced by approximately 60–65 μm, for a precise reconstruction of the ϕ coordinate. On the reverse side, the following combination of read-out pitch (strip orientations with respect to the beam) is used: 141 μm (90°), 125.5 μm (90°), 60 μm($-1.2°$), 141 μm (90°), 65 μm (1.2°), from the innermost to the outermost layer for reconstructing the z coordinate. A total of 405 504 electronics channels are used for SVX II.

The ISL [77] detector is placed at intermediate radial distance between the SVX II and the drift chamber and has polar coverage up to $|\eta_{\text{det}}| < 2$ and a total length of 174 cm along z. At $|\eta_{\text{det}}| \lesssim 1$ a single layer of silicon sensors is mounted on a cylindrical barrel at radius of 22.6 cm (or 23.1 cm). At $1 \lesssim |\eta_{\text{det}}| \lesssim 2$ two layers of silicon sensors are arranged into two pairs of concentric barrels (inner and outer). In the inner (outer) barrel, staggered ladders alternate at radii of 19.7 and 20.2 cm (28.6 and 29 cm). One pair of barrels is installed in the forward region, the other one in the backward region. Each barrel is azimuthally divided into a 30° structure matching the SVX II segmentation. The basic read-out unit consists of an electronic board and three sensors ganged together resulting in a total active length of 25 cm. ISL employs 888 $5.7\,\text{cm} \times 7.5(6.7)\,\text{cm}$ double-sided, $300\,\mu$m thick sensors. Each sensor has axial strips spaced by $112\,\mu$m on one side, and 1.2° angled strips spaced 112–146 μm on the reverse, for a total of 303 104 channels.

All 722 432 channels from the $\sim 7.0\,\text{m}^2$ silicon active-surface employ 5 644 radiation-tolerant, custom integrated read-out chips of the same type. This chip allows independent cycles of digitization of data and analog processing of subsequent data. The discriminated differential pulse from each channel is preamplified, digitized and propagated to the downstream data-acquisition. The ISL and the SVX II, whose mass is approximately 128 kg, share the carbon-fiber supporting structure.

The total amount of material in the silicon system, averaged over ϕ and z, varies roughly as $0.1X_0/\sin\theta$ in the $|\eta_{\text{det}}| \lesssim 1$ region, and roughly doubles in $1 \lesssim |\eta_{\text{det}}| \lesssim 2$ because of the presence of cables, cooling bulk-heads, and portions of the support frame.[3] The average amount of energy loss for a charged particle is roughly 9 MeV. The total heat load of the silicon system is approximately 4 kW. To prevent thermal expansion, relative detector motion,

[3]The symbol X_0 indicates the radiation length.

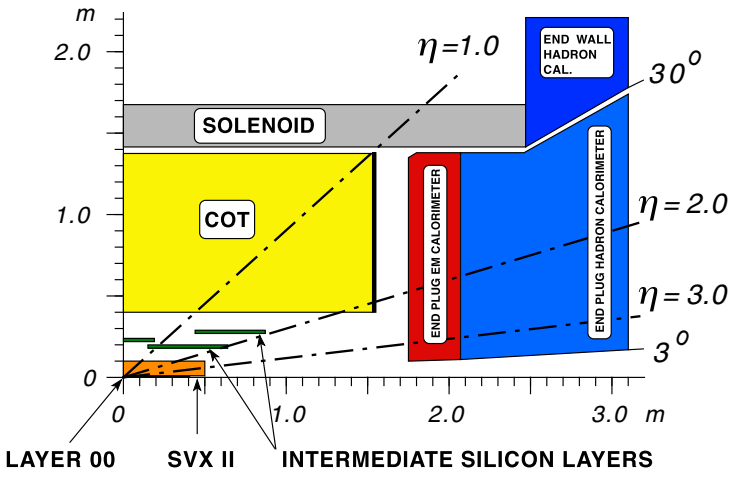


Figure 3.6: Elevation view of one quadrant of the inner portion of the CDF II detector showing the tracking volume surrounded by the solenoid and the forward calorimeters.

increased leakage-current, and chip failure due to thermal heating, the silicon detectors and the associated front-end electronics are held at roughly constant temperature ranging from -6°C to -10°C for L00 and SVX II, and around 10°C for ISL, by an under-pressurized water and ethylene-glycol coolant flowing in aluminum pipes integrated in the supporting structures.[4]

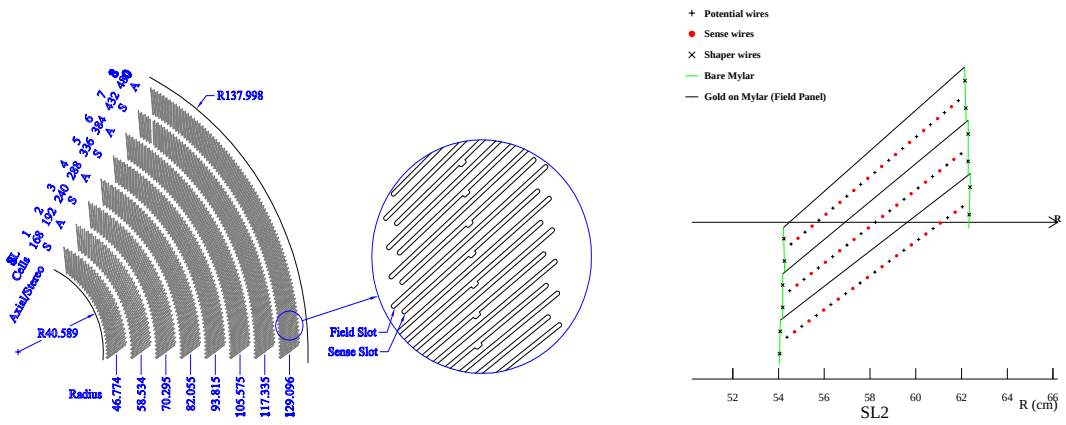
The resolution on the hit position for all silicon sensors is about 11 μm in the (r, ϕ) plane, thus allowing to reach about 20 μm resolution on the impact parameter of high-p_T tracks which degrades to about 35 μm at 2 GeV/c. This precision provides a powerful help to identify long-lived hadrons containing heavy-flavored quarks already at trigger level.

3.2.1.2 The central outer tracker

A large multi-wire, open-cell drift chamber called the *Central Outer Tracker* (COT) [78] extends, in the central pseudo-rapidity region $|\eta_{\text{det}}| \lesssim 1$, the silicon tracking system (see Fig. 3.6).

The COT has an hollow-cylindrical geometry, its active volume spans from 43.4 to 132.3 cm in radius and $|z| \lesssim 155$ cm in the axial direction. Arranged radially into eight "super-layers", it contains 96 planes of wires that run the length of the chamber between two end-plates (see Fig. 3.7 (a)). Each super-layer is divided into ϕ cells; within a cell, the trajectory of a charged particle is sampled at 12 radii (spaced 0.583 cm apart) where sense wires (anodes) are strung. Four super-layers employ sense-wires parallel to the beam axis, for the measurement of the hit coordinates in the (r, ϕ) plane. These are radially interleaved with four stereo super-layers whose wires are alternately canted at angles of 2° and -2° with respect to the beam-line. Combined read-out of stereo and axial super-layers allows the measurement of the (r, z) hit coordinates. Each super-layer is azimuthally segmented into open drift cells. Figure 3.7 (b) shows the

[4] The pressure of the cooling fluid is maintained under the atmospheric pressure to prevent leaks in case of damaged cooling pipes.



(a) (b)

Figure 3.7: A 1/6 section of the COT end-plate (a); for each super-layer the total number of cells, the wire orientation (axial or stereo), and the average radius in cm are given; the enlargement shows in details the slot where the wire planes (sense and field) are installed. Sketch of an axial cross-section of three cells in the super-layer 2 (b); the arrow points into the radial direction.

drift cell layout, which consists of a wire plane closed azimuthally by cathode sheets spaced approximately 2 cm apart. The wire plane contains sense wires alternating with field-shaping wires, which control the gain on the sense wires optimizing the electric field intensity. The cathode is a 6.35 μm thick Mylar sheet with vapor-deposited gold shared with the neighboring cell.[5] Innermost and outermost radial extremities of a cell (*i. e.*, the boundaries between super-layers) are closed both mechanically and electrostatically by Mylar strips with an additional field-shaping wire attached, the shaper wire.

Both the field sheet and the wire plane have a center ($z \approx 0$) support rod that limits motion due to electrostatic forces. Each wire plane contains 12 sense, 13 field-shaping and 4 shaper wires, all made of 40 μm diameter gold-plated tungsten. Wire planes are not aligned with the chamber radius: a $\zeta = 35°$ azimuthal tilt partially compensates for the Lorentz angle of the drifting electrons in the magnetic field.[6] The tilted-cell geometry helps in the drift velocity calibration, since every high-p_T (radial) track samples the full range of drift distances within each super-layer. Further benefit of the tilt is that the left-right ambiguity is resolved for particles coming from the z axis since the ghost track in each super-layer appears azimuthally rotated by $\arctan[2\tan(\zeta)] \approx 54°$, simplifying the pattern recognition problem. On the other

[5] Gold, used also for the wires, was chosen because of its good conductivity, high work function, resistance to etching by positive ions, and low chemical reactivity.

[6] In the presence of crossed electric ($\boldsymbol{E}$) and magnetic ($\boldsymbol{B}$) fields, electrons drifting in a gas move at an angle ζ with respect to the electric field direction, given by $\zeta \approx \arctan\left(\frac{v(E,B=0)B}{kE}\right)$, where $v(E, B = 0)$ is the drift velocity without a magnetic field, and k is a $\mathcal{O}(1)$ empirical parameter that depends on the gas and on the electric field. A common solution for this problem consists in using tilted cells (*i. e.* tilted drift electric field) that compensate the Lorentz angle linearizing the time-to-distance relation.

hand this tilt angle causes an unavoidable difference in the reconstruction efficiencies between positively and negatively charged tracks which cross the sense wires with different incidence angle. Such difference is particularly enhanced for low-p_T tracks and introduces unwanted spurious asymmetries of the order of few percents.

A 50 : 50 gas admixture of argon and ethane bubbled through isopropyl alcohol (1.7%) flows in the active volume of the chamber with its pressure being continuously monitored by four probes. High voltage is applied to the sense and field-shaping wires to generate a 1.9 kV/cm drift electric field. This value, combined with the drift gas, results in a maximum drift-time of about 177 ns along a maximum drift-distance of 0.88 cm, allowing for read-out and processing of the COT data between two consecutive bunch-crossings. The average 180 kV/cm field present at the surface of the sense wire produces typical gains of 2×10^4. The sense wires are read-out by the front-end chip, which provides input protection, amplification, shaping, baseline restoration, discrimination and charge measurement. The input-charge information is encoded (logarithmically) in the signal width for dE/dx sampling and is fed to a time-to-digital converter that records leading and trailing-edge times of signal in 1 ns bins.

The COT single-hit resolution is 140 μm, including a 75 μm contribution from the ~ 0.5 ns uncertainty on the measurement of the $p\bar{p}$ interaction time. Internal alignments of the COT cells are maintained within 10 μm using cosmic rays. Curvatures effects from gravitational and electrostatic sagging are under control within 0.5% by equalizing the difference of E/p between electrons and positrons as a function of $\cot\theta$.

3.2.1.3 Tracking algorithms and performances

The reconstruction of a charged particle trajectory consists in determining the helix parameters of Pag. 49 trough a fit of the reconstructed hits in the tracking sub-detectors with two basic steps: clustering multiple close measurements coming from the same track and pattern-recognition algorithm to joint the hits along the whole track arc. CDF employs several algorithms for tracks reconstruction depending on which component of the detector a particle travels through. The principal one, used to track the particles in the central region ($|\eta_{\text{det}}| \lesssim 1$), is the *Outside-In* (OI). In this algorithm tracks are first reconstructed in the COT and then extrapolated inward to the silicon. This approach guarantees fast and efficient tracking with high purity. The greater radial distance of the COT with respect to the silicon tracker results in a lower track density and consequent fewer accidental combination of hits in the track reconstruction. Due to the limited COT coverage and the strict hits requirement (at least 4 out of 8 super-layers must contain a valid hit), tracking in the forward region requires different algorithms that are not described here because not used in these analyses. A concise overview of all the algorithms used at CDF is given in Refs. [79, 80, 81], in the following we briefly summarize how the OI works and which performances are achieved.

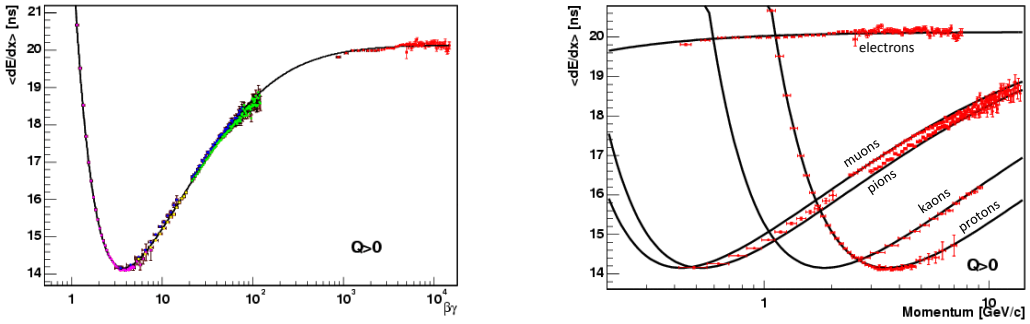
In the first step of pattern recognition, cells in the axial super-layers are searched for sets of 4 or more hits that can be fit to a straight line. Once these "segments" of hits are found, there are two approaches that can be followed to reconstruct a track. One approach is to link together the segments which are consistent with lying tangent to a common circular path[7]. The other approach is to constrain its circular fit to the beam-line. Once a circular path is found in the (r, ϕ) plane, segments and hits in the stereo super-layers are added depending on their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. Once a track is reconstructed in the COT, it is extrapolated inward to the silicon system. Based on the estimated errors on the track parameters, a three-dimensional "road" is formed around the extrapolated track. Starting from the outermost layer, and working inwards, silicon hits found inside the road are added to the track. As hits get added, the road gets narrowed, according to the knowledge of the updated track parameters and their covariance matrix. Reducing the width of the road reduces the chance of adding a wrong hit to the track, and also reduces the computation time. In the first pass of this algorithm, only axial hits are considered; while in a second pass, hits with stereo information are also added to the track. At the end, the track combination with the highest number of hits and lowest χ^2/ndf for the five parameters helix fit is kept.

COT efficiency for tracks with p_T larger (smaller) than 1 GeV/c is typically 98-99% (95%) depending on the isolation. The typical resolutions on track parameters are: $\sigma_{p_T}/p_T^2 \approx 0.15\%\,(\text{GeV}/c)^{-1}$, $\sigma_{\phi_0} \approx 0.035^\circ$, $\sigma_{d_0} \approx 250\,\mu$m, $\sigma_\theta \approx 0.17^\circ$ and $\sigma_{z_0} \approx 0.3\,$cm for tracks with no silicon information nor beam constrained. The silicon information improves the impact parameter resolution which, depending on the number (and radial distance) of the silicon hits, may reach $\sigma_{d_0} \approx 20\,\mu$m (not including the transverse beam size). This value, combined with the $\sim 30\,\mu$m transverse beam size, is sufficiently small with respect to the typical transverse decay-lengths of heavy flavors (a few hundred microns) to allow separation of their decay-vertices from production vertices. The silicon tracker improves also the stereo resolutions to $\sigma_\theta \approx 0.06^\circ$ and $\sigma_{z_0} \approx 70\,\mu$m, while the transverse momentum and the azimuthal resolutions remain approximately the same of COT-only tracks. Transverse momentum resolution can be further improved to about $\sigma_{p_T}/p_T^2 \approx 0.05\%\,(\text{GeV}/c)^{-1}$ when tracks are beam constrained.

3.2.2 Particle identification

Particle identification (PID) plays an important role in two aspects of the $B_s^0 \to J/\psi\phi$ analysis, as a component in the discriminating variables of the offline selection (Chap. 5), and in selecting kaon tracks for flavor tagging (Chap. 4). The two quantities used as PID to distinguish pions from kaons at CDF are the ionization energy loss in the COT and the Time of Flight.

[7] The helical track, when projected onto the (r, ϕ) plane, is a circle.



(a) (b)

Figure 3.8: Universal curve showing combination of pions, kaons, protons and muons (a). Momentum dependence for muons, pions, kaons, protons and electrons at CDF (b).

3.2.2.1 PID with the COT

The ionization energy loss (dE/dx) of a charged particle as it moves through matter is given by the Bethe-Bloch formula [27]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi N e^4}{m_e c^2 \beta^2} q^2 \left[\log \frac{2 m_e c^2 \beta^2 \gamma^2}{I^2} - \beta^2 \right],$$

where N is the number density of electrons in the material of interest, e is the electron charge, m_e the electron mass, q the particle's charge, βc the particle's speed, $\gamma = 1/\sqrt{1-\beta^2}$, and I is the mean excitation energy for atoms in the material. For a material with known properties, such as the gas in the COT, this equation can yield a measurement of the particle's mass when combined with a momentum measurement, which can then be compared to known particle masses in order to estimate the particle type. The Bethe-Bloch formula is empirically modified to better model the CDF detector, as [82]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2} \left[c_1 \log \left(\frac{\beta\gamma}{b + \beta\gamma} \right) + c_0 \right] + a_1(\beta - 1) + a_2(\beta - 1)^2 + C, \tag{3.1}$$

where a_i, b, c_j and C are free parameters which float when fitting the data. From Eq. (3.1), a *universal curve* can be plotted against $\beta\gamma$. An example of the CDF universal curve for several particle types is shown in Fig. 3.8 (a), while Fig. 3.8 (b) plots the momentum dependence of measured dE/dx for different particles, which demonstrates the ability to separate particle types using dE/dx and momentum. This separation when dE/dx is plotted against momentum occurs because the dependence on $\beta\gamma$ is an implicit function of mass and momentum.

The amount of ionization charge produced by a charged particle near a COT sense wires affects the signal strength in the wire. The dE/dx is measured as the amount of charge, above a threshold value, which is proportional to the width (Δt) of the pulse from the readout chip.

Thus *dE/dx* values are given in nanoseconds rather than a unit of energy. While the measured *dE/dx* should only depend on the boost, $\beta\gamma$, of a particle the measurement capability of the COT for *dE/dx* is not perfect and introduces effects due to both environmental and kinematic variables. These dependences reduce the power of the *dE/dx* variable to separate between particle types, so in order to optimise PID at CDF the measured *dE/dx* must be calibrated to remove or minimise these dependences. The calibration taken to correct the detector and kinematic effects on measured *dE/dx* in the COT are described in Ref. [83]. To calibrate the *dE/dx* pure samples of pions and of kaons are used. These are obtained from D^0 decays, from the decay chain $D^\star \to D^0\pi^+$, $D^0 \to K^-\pi^+$ and its charge conjugate.

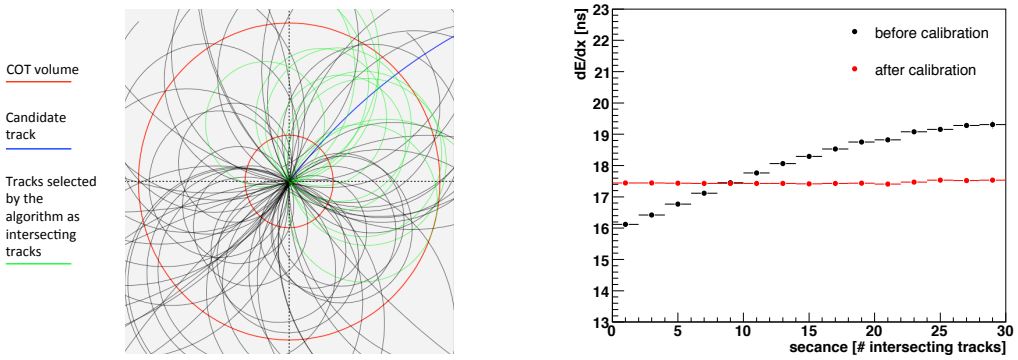
The following effects are corrected for:

- *time.*
  The dependence of *dE/dx* on time is understood to be due to several properties of the drift chamber which have varied since the start of Run II. One example of this is that in 2006, the two inner layers of the COT had *dE/dx* read out switched off in order to maintain tracking capabilities at high luminosity. The fact that the delivered luminosity from the Tevatron has increased significantly over the run period also affects measured *dE/dx*. Additionally, aging of the COT can affect the amount of charge collected for dE/dx measurement. This effect exhibits correlations with luminosity, number of COT hits, and track density; variations in these parameters are calibrated for in a 4-dimensional simultaneous correction.

- *Luminosity.*
  This interdependence is to be expected given the time dependence of luminosity itself, and the increase in occupancy of the COT in higher luminosity running will affect the number of hits, and the density of tracks. Higher occupancy in the COT means a larger number of tracks which could be in the region of the track of interest, and could contribute to the measured charge deposit for that track. This causes an artificially raised *dE/dx* measurement for higher luminosity events.

- *COT hits.*
  There is a significant dependence on the number of associated COT hits for a track. This is itself correlated with other occupancy related variables.

- *Secance* (track density).
  A variation dependent on the track density in the vicinity of the candidate, measured as the number of r–ϕ intersections of the candidate track with other tracks within the COT (see Fig. 3.9 (a)). There is no longitudinal segmentation in the wires of the COT, so the charge deposited by any hit axially in the region of the candidate track can be counted towards the *dE/dx* of that track. Secance is a variable which has been constructed to count the number of tracks intersecting the candidate track, which was found to be a more effective



(a) (b)

Figure 3.9: Sketch of the secance: the blue track is the candidate, the green tracks are those that the algorithm selects as intersecting tracks in the COT volume (a); variations in measured dE/dx as a function of the secance before and after calibration (b).

quantity than a more traditional approach such as track isolation. An assumption has to be made in correcting this parameter, that the number of reconstructed tracks for an event is proportional to the total number of tracks, as only the reconstructed tracks are accessible in the dataset.

- η_{det}.
  The pseudorapidity is a function of transverse momentum, and therefore there will be some intrinsic variation in measured dE/dx with η_{det} as the momentum dependence is a physical effect, η_{det} is the only variable which exhibits a significant momentum dependence. This parameter demonstrated independence from the other parameters, so can be corrected for separately.

- ϕ_0.
  This parameter can be corrected for independently as the effect has been shown to factorise with the other studied variables, modulo a global shift due to the change in mean dE/dx measured. Variations in the measured dE/dx for particles traversing different sectors of the COT can be explained by the temperature gradient and the flow of gas within the chamber. The variable chosen to measure this effect is the track ϕ_0, which gives the initial azimuthal direction within the COT and therefore is expected to show clearly any variations which occur due to the effects mentioned.

Figure 3.9 (a) shows the effect of the correction of the dE/dx for the secance; similar effects are given for the other variables.

The curves for the four studied particle types, positive and negative kaons and pions, can be compared in Fig. 3.10, showing the separation between the particle species after the calibration. As is clear from this plot, there is a systematic difference in the average dE/dx response of the

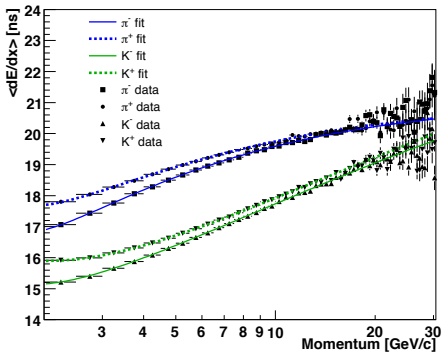
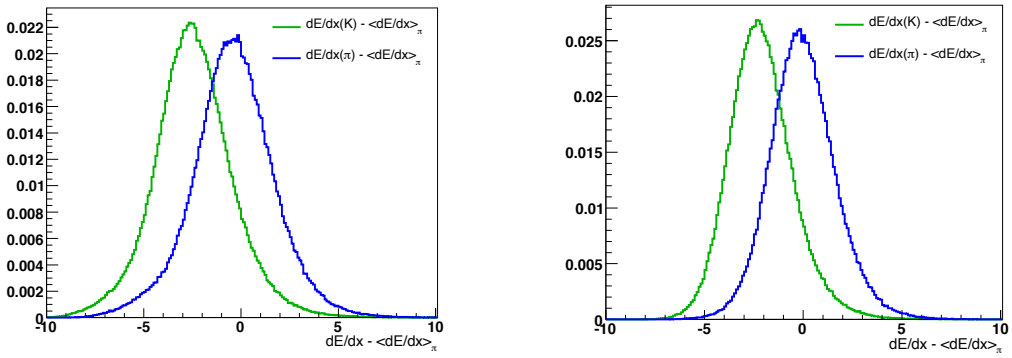


Figure 3.10: *dE/dx* dependence on momentum for different particle types after calibration.



(a) (b)

Figure 3.11: Data distributions for *dE/dx* residuals before (a) and after (b) calibration.

COT for positively and negatively charged particles, and this difference is momentum dependent. The change in separation of the curves shown causes a drop off in performance with momentum for *dE/dx* as a PID tool, but for the momentum ranges used in the $B_s^0 \to J/\psi\phi$ analysis described in this thesis.

Using the predicted *dE/dx* curves, it is possible to construct *dE/dx* residuals, $(\langle dE/dx \rangle_{\text{measured}} - \langle dE/dx \rangle_{\text{predict}})$. To assess the effect of calibrating the *dE/dx* on the PID performance of this variable, the main figure of merit used is the pion-kaon separation power. This is studied by looking at the separation between residual *dE/dx* distributions for pions and kaons. Separation power ranges from 0% (no separation) to 100% (perfect separation), where perfect separation would mean no overlap in the distributions. The value of separation power between K and π for uncalibrated *dE/dx* is 48.8%, which increases after calibration to 56.4%, a 22% improvement (see Fig. 3.11).

3.2.2.2 Time of flight detector

The *Time Of Flight* (TOF) is a complementary quantity to *dE/dx*, as it is most powerful at separating pions from kaons at low momentum, $p_T < 1.5\,\text{GeV}/c$. Particle identification with TOF is based on the relation

$$m = \frac{p}{c}\sqrt{\frac{(ct_\text{f})^2 - 1}{L}}$$

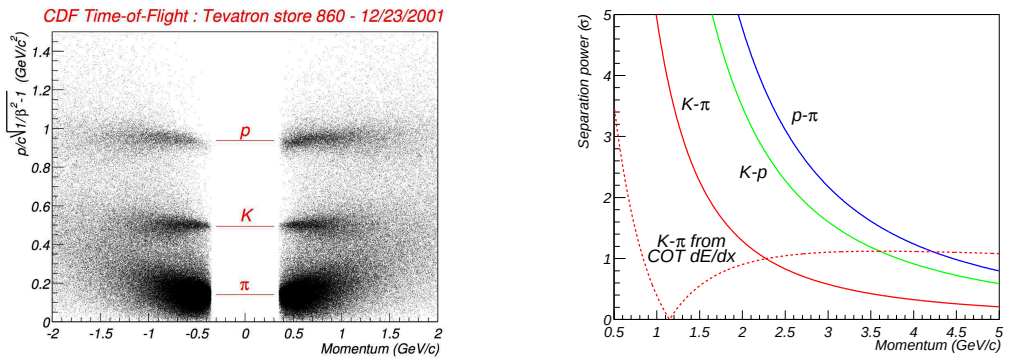
where m is the predicted mass (which can be used to identify a particle by comparison with known particle masses), c is the speed of light, L is the track length, p is the particle momentum, and t_f is the time of flight.

The TOF detector [84] is a cylindrical array made of 216 scintillating bars and it is located between the external surface of the COT and the cryostat containing the superconducting solenoid. Bars are 280 cm long and oriented along the beam axis all around the inner cryostat surface at an average radial distance of 138 cm. Both longitudinal sides of the bars collect the light pulse into photomultiplier and measure accurately the timing of the two pulses. The time between the bunch crossing and the scintillation signal in these bars defines the β of the charged particle while the momentum is provided by the tracking system. The measured mean time resolution is now 110 ps. This guarantees a separation between charged pions and kaons with $p_T \lesssim 1.6\,\text{GeV}/c$ equivalent to 2σ, assuming Gaussian distributions. Unfortunately, also in the usual luminosity conditions ($\mathscr{L} \gtrsim 5\times10^{31}\ \text{cm}^{-2}\text{s}^{-1}$) the occupancy of the single bars determines a degradation in efficiency, which is about 60% per track.

Like *dE/dx*, TOF should be dependent only on the mass and momentum of a particle, however the detector and event kinematics can also affect this quantity. The measured TOF therefore must be calibrated in order to obtain optimal separation between particle species. This is done using pure samples of each particle type separately; the full calibration method for this variable is described in [85]. Figure 3.12 shows an example of the TOF distribution for different particles, demonstrating the separation according to mass, and the separation power for several particle types compared to the K–π separation power of *dE/dx*.

3.2.3 Muons system

Muon identification is crucial for the $B_s^0 \to J/\psi\,\phi$ analysis; the online selection based on the reconstruction of a muons pair is used to select $J/\psi \to \mu^+\mu^-$ events, which make up half of the final state particles of the $B_s^0 \to J/\psi\,\phi$ decay. The system devoted to muon identification is the furthest detector component from the beam pipe. Muons pass through the rest of the detector losing little energy, and if they have sufficient momentum they reach the muon drift chambers. A particle entering the muon chambers leaves a track which is registered as a muon *stub*. As they are charged, muons leave tracks in the COT; if a COT track is matched to a muon stub these can be combined to make up a muon candidate.



(a) (b)

Figure 3.12: TOF distribution of different particles (a) and separation power for TOF compared to that of dE/dx (b).

The design of the detector is such that other types of particle than muons should be absorbed by the material between the beam pipe and the first of the muon detectors. The CDF muon system [86, 87] consists of several subcomponents, the main part is the *Central MUon* detector (CMU), which was the initial muon system of the Run I detector. During Run I, the muon system was upgraded by adding the *Central Muon eXtension* (CMX) and *Central Muon Upgrade* (CMU) components. These components were improved and finalized for Run II. An additional section, the *Intermediate MUon* detector (IMU), extends the coverage in η_{det} to the forward region. The η_{det}–ϕ coverage of the different muon detector components is shown in Fig. 3.13. Features and properties of the muon detectors are summarised in Tab. 3.2. The CMU, CMP, CMX and IMU are drift chambers, there are also three scintillators, the CSP, CSX and BSU, which are close to each drift chamber. These serve two main purposes, they are used for triggering and deliver timing information to reduce background by identifying which beam crossing produced a specific muon.

	CMU	CMP	CMX	IMU
η_{det} range	$\lvert\eta_{\text{det}}\rvert < 0.6$	$\lvert\eta_{\text{det}}\rvert < 0.6$	$0.6 < \lvert\eta_{\text{det}}\rvert < 1.0$	$1.0 < \lvert\eta_{\text{det}}\rvert < 1.5$
Drift tube cross section [cm^2]	2.68×6.35	2.5×15	2.5×15	2.5×8.4
Drift tube length [cm]	226	640	180	363
Minimum muon p_T [GeV/c]	1.4	2.2	1.4	1.4–2.0
pion interaction lengths	5.5	7.8	6.2	6.2–20

Table 3.2: Some important design parameters of the CDF II muon detectors.

The CMU is made up of 144 modules each with 16 cells filled with the same mixture of gases as the COT. In the centre of each cell there is a 50 μm stainless steel sense hambers, these are referred to as *punch-throughs* which result in *fake* muons. The first upgraded muon detector, the CMP, is behind an additional 60 cm of steel, further reducing the chance of fake

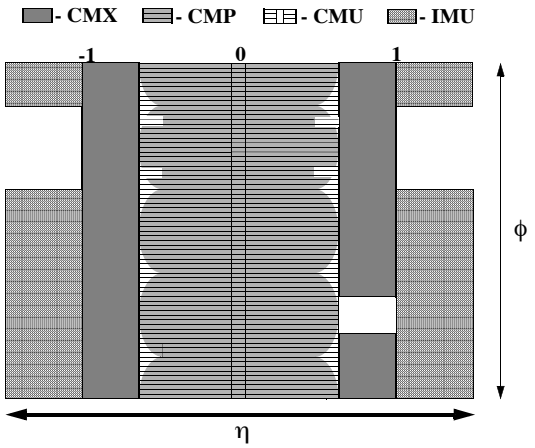


Figure 3.13: Layout of the CDF II muon detector components in the $\eta_{\rm det}$–ϕ plane.

muon events being recorded. CMP measurements therefore improve the muon identification, and a muon candidate with hits in both the CMU and CMP (called a CMUP muon) has higher precision than one with CMU information only. The CMP is rectangular in design, therefore has variable $\eta_{\rm det}$ coverage with respect to ϕ, as shown in Fig. 3.13. Sense wires in the CMP are read out individually, and groups of hits in nearby wires can be combined with CMU information for triggering. The CMX is made up of conical sections with four layers of twelve drift tubes, and is associated with the CSX set of scintillation counters. These detectors extend the $\eta_{\rm det}$ coverage of the muon system, lying at either end of the central detector components. There is no additional shielding for the CMX, but the large angle from the interaction point means that to leave a stub in the CMX, particles travel through significantly more material than those which reach the CMU. The furthest reaching muon detector in $\eta_{\rm det}$ is the IMU, which has trigger capability for muons with $|\eta_{\rm det}| < 1.5$ and can be used in offline reconstruction of up to $|\eta_{\rm det}| < 2.0$ muons. The drift chambers of the IMU of the same type as in the CMP, and are arranged in barrel structures. The readout from the IMU is linked with scintillator timing information for the trigger and muon identification.

3.2.4 Other detectors

In this section we briefly describe the CDF II sub-detectors not used or less important for our analyses.

3.2.4.1 Calorimeters

Outside the solenoid, scintillator-based calorimeters cover the region $|\eta_{\rm det}| \leqslant 3.6$, and are devoted to the measurement of the energy deposition of photons, electrons and hadrons using the shower sampling technique. Muons, which are minimally ionising and therefore deposit only a small fraction of energy passing through the calorimeter material, leave little trace in the

calorimeters and pass through to the muon detectors which are described in the next section. Calorimeters are particularly important in identifying neutral particles, which do not leave tracks in the inner detectors. They are also essential in deducing the likelihood of the presence of a neutrino in an event, as neutrinos do not interact with the detector but can be observed by the absence of energy that should be present due to energy-momentum conservation, called *missing energy*.

The basic structure of the calorimeters consists of alternating layers of passive absorber and plastic scintillator. Neutral particles and charged particles with $p_T \gtrsim 350\,\mathrm{MeV}/c$ are likely to escape the solenoid's magnetic field and penetrate into the calorimeters. These are finely segmented in solid angle around the nominal collision point and coarsely segmented radially outward from the collision point (in-depth segmentation). Angular segmentation is organized in projective towers. Each tower has a truncated-pyramidal architecture having the imaginary vertex pointing to the nominal interaction point and the base is a rectangular cell in the $(\eta_{\mathrm{det}}, \phi)$ space. Radial segmentation of each tower instead consists of two compartments, the inner (closer to the beam) devoted to the measure of the electromagnetic (EM) component of the shower, and the outer devoted to the measure of the hadronic (HA) fraction of energy. These two compartments are read independently through separated electronics channels. A different fraction of energy release in the two compartments distinguishes photons and electrons from hadronic particles.

CDF calorimeters are divided in several independent subsystems: the *Central* EM and HA calorimeters, the CEM and CHA, surround the tracking system radially, and cover the pseudorapidity range $|\eta_{\mathrm{det}}| < 1.1$ [88, 89]; coverage in η_{det} is extended by the *Plug* calorimeters, PEM and PHA, to $|\eta_{\mathrm{det}}| < 3.6$, and the region between the central and plug areas is bridged by the *end-Wall* HA (WHA) calorimeters [90, 91, 92]. The EM calorimeters are enhanced by showermax detectors, which are gas filled wire and strip chambers. These give position measurements which can be matched to tracks, and a transverse profile of the shower, to separate photons from neutral pions. Additionally, preshower scintillator tile chambers are positioned on the front of the central calorimeter wedges and the first layer of the PEM. These improve soft (low momentum) photon and electron identification.

3.2.4.2 Cherenkov luminosity counters

CDF has two dedicated luminosity detectors, the *Cherenkov Luminosity Counters* (CLC), which are innovative devices for making precision measurements in the high luminosity regime. Prior to CDF II, luminosity measurement at hadron machines was usually carried out with scintillating counters which recorded the number of bunch crossings with no interactions. For high luminosities this technique is not practical as the number of bunch crossings with no interactions is minimal, so it is necessary to directly measure the number of interactions. High precision luminosity measurements are essential for analyses, such as cross-section measurements, which

require knowledge of the total integrated luminosity in a dataset. From the CLCs, the average number of particles per bunch crossing is measured by the amount of Cherenkov light collected, and this can be used to estimate the number of inelastic $p\bar{p}$ interactions in each bunch crossing. The CLCs are placed at either end of the CDF detector, in the end plug calorimeters, covering the pseudo-rapidity range $3.7 < |\eta_{\text{det}}| < 4.7$ [93, 94]. They are each made up of 48 long conical Cherenkov counters, filled with isobutane gas, arranged in 3 concentric layers about the beampipe. Small, fast PMT are used to collect the Cherenkov light, and backgrounds such as secondary particles are excluded by setting suitable light thresholds.

3.2.5 Trigger and data acquisition systems

The CDF II trigger system is a key element that makes the presented measurements possible. Identification of decays of heavy-flavored mesons is challenging in the Tevatron collider environment due to the large inelastic $p\bar{p}$ cross section and high particle multiplicities at 1.96 TeV. In order to collect these events the trigger system must reject more than 99.99% of the collisions while retaining good efficiency for signal. In this section, we describe the CDF II trigger structure and the algorithms used in collecting samples enriched in $J/\psi \to \mu^+\mu^-$ for the $B_s^0 \to J/\psi\,\phi$ analysis and samples enriched in hadronic B decays for the $B_s^0 \to \phi\phi$ analysis.

At the typical Tevatron instantaneous luminosity, approximately 2.6×10^6 inelastic collisions per second occur, corresponding to one interaction per bunch-crossing on average. Since the read-out of the entire detector needs about 2 ms on average, after the acquisition of one event, another approximately 5 000 interactions would remain unrecorded. When an event recording is prevented because the system is busy with a different event or a different task, this is called *deadtime*. Expressing the same concept in terms of information units, the average size of information associated to each event is 140 kB. Even in case of deadtime-less read-out of the detector, in order to record all events, an approximate throughput and storage rate of 350 GB/s would be needed, largely beyond the possibility of currently available technology. The read-out system has to reduce the 2.3 MHz interaction-rate to the 100 Hz storage rate attainable at CDF II. The challenge for the whole system is to be smart enough to cut-off events that do not have the minimal requirements to be reconstructed or seem to contain well-known processes, that do not need further study, focusing the acquisition system on the interesting processes. Figure 3.14 shows a scheme to explain how the information flows trough the different parts. To suppress unwanted events, the CDF II *Data AcQuisition system* (DAQ) is segmented in three levels, each level receiving the accepted event from the previous one, and, provided with detector information with increasing complexity and with more time for processing, determines if one of a set of existing criteria is verified by the event.

Prior to any trigger level, the bunched structure of the beam is exploited to reject cosmic-ray events by gating the front-end electronics of all sub-detectors in correspondence of the bunch-crossing. The front-end electronics of each sub-detector, packaged in *Vesa Module Eurocard*

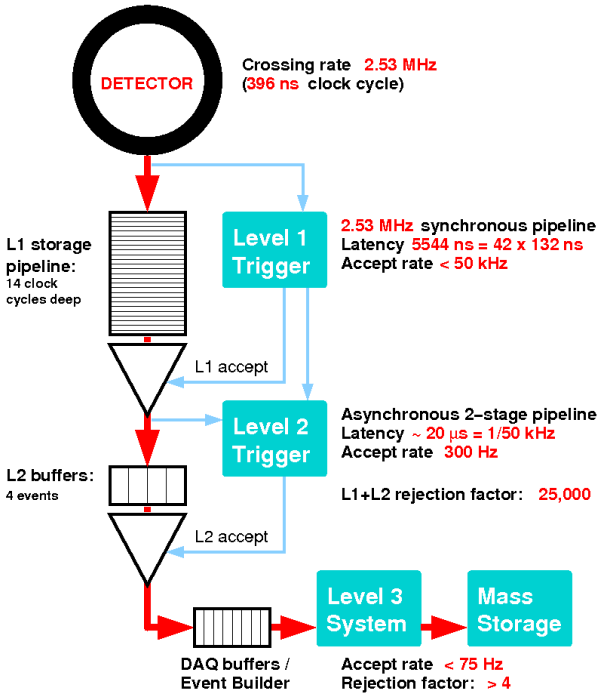


Figure 3.14: Functional block diagram of the CDF II trigger and data acquisition systems.

(VME) modules hosted in about 120 crates, has a 42-cells deep pipeline synchronized with the Tevatron clock-cycle set to 132 ns. The Tevatron clock picks up a timing marker from the synchrotron RF and forwards this bunch-crossing signal to the trigger and to the front-end electronics. Since the inter-bunch time is 396 ns, three times the Tevatron clock-cycle, the pipeline can collect data corresponding to a maximum of 14 bunch-crossings. The pipeline depth gives the amount of time that *Level 1* (L1) trigger has to decide to accept or reject an event otherwise the buffer content is overwritten: $14 \times 396\,\text{ns} = 5.5\,\mu\text{s}$. An event accepted by the L1 is passed to the *Level 2* (L2) buffer, where the number of buffers in the pipeline is 4, that gives $4 \times 5.5\,\mu\text{s} = 22\,\mu\text{s}$. This means that if an event is accepted by the L1 and the L2 does not have a free buffer, deadtime will incur. L2 output rate is low enough to avoid in general deadtime problem in the connection between L2 and *Level 3* (L3).

At L1, a synchronous system of custom-designed hardware processes a simplified subset of data in three parallel streams to reconstruct coarse information from the calorimeters (total energy and presence of single towers over threshold), the COT (two-dimensional tracks in the transverse plane), and the muon system (muon stubs in the CMU, CMX and CMP chambers). A decision stage combines the information from these low-resolution physics objects, called *primitives*, into more sophisticated objects, *e. g.* track primitives are matched with muon stubs, or tower primitives, to form muon, electron, or jet objects, which then undergo some basic selections.[8]

[8] A jet is a flow of observable secondary particles produced in a spatially collimated form.

At L2, an asynchronous system of custom-designed hardware processes the time-ordered events accepted by the L1. Additional information from the shower-maximum strip chambers in the central calorimeter and from the axial layers of the SVX II detector is combined with L1 primitives to produce L2 primitives. A crude energy-clustering is done in the calorimeters by merging the energies in adjacent towers to the energy of a seed tower above threshold. L1 track primitives matched with consistent shower-maximum clusters provide refined electron candidates whose azimuthal position is known with 2° accuracy. Information from the (r, ϕ) sides of the SVX II is combined with L1 tracks primitives to form two-dimensional tracks with resolution similar to the offline one. Finally, an array of programmable processors makes the trigger decision, while the L2 objects relative to the following event accepted at L1 are already being reconstructed. Samples enriched with heavy-flavor particles are selected at L2 by the displaced-track trigger, based on the *Silicon Vertex Trigger* (SVT) [95, 96]. It provides a precise measurement of the track impact parameter, defined as the distance of closest approach to the beam axis in the transverse plane. Decays of heavy-flavor particles, such as the $B_s^0 \to \phi\phi$, are identified by requiring two tracks with $120\,\mu\text{m} < \text{d0} < 1.0\,\text{mm}$ and applying a requirement on the two-dimensional decay length, $L_{xy} > 200\,\mu\text{m}$. The SVT trigger is used to select on-line the $B_s^0 \to \phi\phi$ sample and it will be described in details in Sect. 9.1.

The digitized output relative to the L2-accepted event reaches L3 via optical fibers and it is fragmented in all sub-detectors. It is collected by a custom hardware switch that arranges it in the proper order and transfers it to commercial computers, organized in a modular and parallelized structure of 16 subsystems. The ordered fragments are assembled in the *event record*, a block of data that univocally corresponds to a bunch-crossing and is ready for the analysis of the L3 software. The event reconstruction benefits from full detector information and improved resolution with respect to the preceding trigger levels, including three-dimensional track reconstruction, tight matching between tracks and calorimeter or muon information, and calibration information. If an event satisfies the L3 requirements, the corresponding event record is transferred to mass storage at a maximum rate of 20 MB/s. A fraction of the output is monitored in real time to search for detector malfunctions, to derive calibrations constants and to graphically display events. The L3 decision is made after the full reconstruction of the event is completed and the integrity of its data is checked, a process that takes a few milliseconds.

3.2.6 Operations and data quality

The data-taking efficiency is plotted in Fig. 3.15 as a function of time. The average over the whole Run II is about 85%. The inefficiency is approximately equally shared in a 5% arising at the beginning of the store, when the detector is not powered while waiting for stable beam conditions, a 5% due to trigger deadtime, and a 5% due to unexpected detector or DAQ problems.

Each time that at least one of the trigger paths fires, an *event* is labeled with a progressive

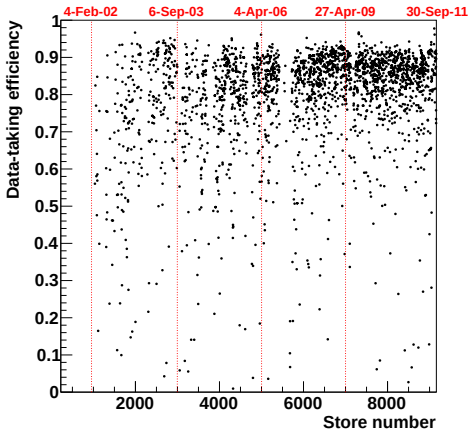


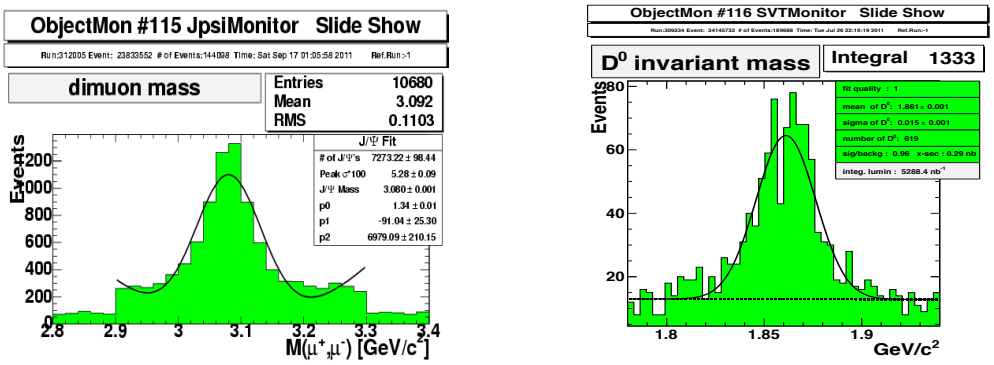
Figure 3.15: Data-taking efficiency as a function of store number/time.

number. Events are grouped into runs, *i. e.* periods of continuous data-taking in constant configurations of trigger table, set of active sub-detectors and so forth.[9] Several parameters of the operations (*e. g.* beam-line position and slope, set of calibrations, etc.) are stored in the database on a run-averaged format.

All data manipulations occurring some time after the data are written to permanent memories are referred to as offline processes, as opposed to the online operations that take place in real time, during the data-taking. The most important offline operation is the processing with a centralized *production* analysis that generates collections of high-level physics objects suitable for analysis, such as tracks, vertexes, muons, electrons, jets, etc. from low-level information such as hits in the tracking sub-detectors, muon stubs, fired calorimeter towers, etc. During the production, more precise information about the detector conditions (*e. g.* calibrations, beam-line positions, alignment constants, masks of malfunctioning detector-channels, etc.) and more sophisticated algorithms are used than those ones available at the L3 of the trigger. The production may be repeated when improved detector information or reconstruction algorithms become available.

To ensure homogeneous data-taking conditions, each run undergoes a quality inspection. online shift operators, offline production operators, and sub-detector experts certify in what fraction of data the running conditions for all relevant sub-detectors are compliant to physics-quality standards. When detectable problems of the detector occur, the data-taking is quickly stopped, so very short runs are likely to contain corrupted data. Runs with fewer than 10^8 live Tevatron clock-cycles, or fewer than 10^4 (10^3) L1 (L2) accepts, or containing data corresponding to an integrated luminosity $\int \mathscr{L}\, dt < 1\,\text{nb}$ are excluded from physics analysis. Online shift

[9] The data acquisition might need to be interrupted and recovered for several motivations, including the need for enabling or disabling a sub-detector, the need for a change in the trigger table, a problem in the DAQ chain and so forth.



(a) (b)

Figure 3.16: Examples of online plots to monitor the data-taking, the $J/\psi \to \mu^+\mu^-$ (a) and $D^0 \to K^-\pi^+$ (b) yields.

operators further exclude the runs in which temporary or test trigger tables were used.[10] Runs whose data underwent problems or software crashes during the production are excluded offline.

Accurate integrated luminosity measurements are ensured in physics-quality data by requiring the CLC to be operative during the data-taking and by verifying that a set of luminosity and beam-monitor probe quantities are within the expected ranges. Shift operators ensure that L1 and L2 triggers operate correctly and that the rate of SVX II data corruption errors is smaller than 1%.[11] SVT experts verify that the online fit and subtraction of the beam position is done correctly and that the SVT occupancy is within the expected limits. In addition, higher level quantities, such as event yields of $J/\psi \to \mu^+\mu^-$ and $D^0 \to K^-\pi^+$ decays are monitored online and are required to be within the expected ranges (see Fig. 3.16). For analyses using COT information, the minimum integrated luminosity required is 10 nb and the fraction of noisy COT channels is required to be smaller than 1%.

3.2.7 Offline data processing

After data has been selected by triggers and written out, it is subjected to an offline production process. This includes refitting tracks, fitting vertices, reconstructing decay chains, and packaging the information necessary to allow for physics measurements in a user-friendly format.

3.2.7.1 Tracks refitting

Tracks are refit offline with the most accurate available detector information and calibration. Energy loss corrections are applied depending on a track's particle type. If L00 hit information

[10] It is sometimes necessary to test new configurations of the trigger selections in a real data-taking condition to monitor trigger rates, performance and so on.

[11] The read-out of the silicon detector and the proper integration of the information in the online infrastructure is a complex operation which, occasionally, leads to a certain fraction of data to be improperly processed.

is available, it is added to the track refitting. Requirements on matching between muon stubs and COT tracks are made more stringent. Trigger confirmation is required for muons: the muons coming from the B decay of interest must pass the dimuon trigger requirement. In other words, the event must have been selected by the dimuon trigger based on the final state muons from a B decay, not from random muons in the event that happened to meet the trigger requirements.

3.2.7.2 Vertex fitting

Vertex fitting is central to reconstructing a decay chain. Tracks observed in the detector are matched together as coming from a common origin, where the parent particle decayed. The parent particle may itself be a decay product of a heavier particle. The decay chain is traced backwards from the stable tracks as a series of intermediate decay vertices, leading to the primary vertex at the interaction point. Momenta can be assigned to the decay particles, based on the measured momenta of the final state tracks.

The decays used in this thesis were reconstructed using the CTVMFT vertex fitting program [97]. Vertices can be fit either in 3D or in 2D, providing the user with a handle on the goodness of fit for the vertices in the form of a χ^2 probability. Knowledge of the primary vertex is critical to the measurement of the B meson lifetime. Multiple methods of determining the primary vertex exist. For our analyses, the primary vertex location is recalculated for each event, using tracks in the event that were not related to the B meson decay of interest. This is referred to as the *event-by-event primary vertex method*, and it provides an accurate assessment of the primary vertex position, with a resolution of $\approx 20\,\mu$m.

3.2.7.3 Final data format

The data used in this thesis was stored in the BSTNTUPLE framework [98]. The framework is based on the standard STNTUPLE framework [99]. It stores information about stable and decaying particle objects, together with information for flavor tagging and particle identification. It allows the user the flexibility to access different decay modes for B- and D-physics analyses, without requiring excessive CPU or storage space.

3.2.8 Monte Carlo simulation

Estimation of the fraction of events of a certain type that escape the detector acceptance, or detailed studies of the expected response of the detector to the passage of particles is a common need in many analyses. Usually, complex detector geometries and the numerous effects that need to be accounted for in predicting their response make it the analytical derivation of the relevant distributions impractical or impossible. The alternative method is to use numerical simulation to carry out these tasks. The algorithms used all involve some type of random sampling to simulate processes, and are collectively called *Monte Carlo* (MC) simulation.

The simulation is divided into several steps, which reproduce in order the main physical processes and processing steps involved in collecting data from real $p\bar{p}$ interactions. We provide here a short overview of the standard CDF II simulation used in B-physics analyses. Further details can be found in Ref. [100].

The first step in simulation is the treatment of the $p\bar{p}$ hard scattering, and the outgoing quark and gluon collision products, followed by simulation of the fragmentation and hadronization processes which yield hadrons and associated jets. We used the BGenerator package [101]: it concentrates on producing only one B meson per event, which yields a great advantage in computational speed. On the other hand, by design it does not mimic the full collision environment, as the Pythia package could perform. For our purposes, BGenerator is sufficient since we wish to model single B decay samples.

The second step is the simulation of the full decay chain of the B mesons under study. For this task, we use the EvtGen package [102]. EvtGen is specialized for heavy flavor decays and accounts correctly for quantum mechanical interference effects.

The third step in simulation incorporates the interaction of the decay products with the detector material. For this task we use the CdfSim package [100], which is a CDF-specific full detector simulation. In the standard CDF II simulation, the detector geometry and material are modelled using the version 3 of the Geant package [103] tuned to test-beam and collision data. Geant receives in input the positions, the four-momenta, and the identities of all particles produced by the simulated collisions that have long enough lifetimes to exit the beam pipe. It simulates their paths in the detector, modelling their interactions (*bremsstrahlung*, multiple scattering, nuclear interactions, photon conversions, etc.) and the consequent generation of signals on a single channel basis. Specific packages substitute Geant for some sub-detectors: the calorimeter response is simulated with Gflash, a faster parametric shower-simulator [104] tuned for single-particle response and shower-shape using test-beam data (8–230 GeV electrons and charged pions) and collision data (0.5–40 GeV/c single isolated tracks); the drift-time within the COT is simulated using the Garfield standard package [105] further tuned on data; the charge-deposition model in the silicon uses a parametric model, tuned on data, which accounts for restricted Landau distributions, production of δ-rays, capacitive charge-sharing between neighboring strips, and noise [106].[12] Furthermore, the actual trigger logic is simulated. The output of the simulated data mimics the structure of collision data, allowing their analysis with the same reconstruction programs used for collision data.

The detector and trigger configuration undergo variations during data-taking. Minor variations may occur between runs, while larger variations occur, for instance, after major hardware improvements, or Tevatron shut-down periods. For a more detailed simulation of the actual experimental conditions, the simulation has been interfaced with the online database that reports, on a run-by-run basis, all known changes in configuration (position and slope of the beam line,

[12]The δ-rays are knock-on electrons emitted from atoms when the passage of charged particles through matter results in transmitted energies of more than a few keV in a single collision.

relative mis-alignments between sub-detectors, trigger-table used, set of SVT parameters) and local or temporary inefficiencies in the silicon tracker (active coverage, noisy channels, etc.). This allows us to simulate the detailed configuration of any set of real runs and to use it, after proper luminosity reweighing, for modeling the realistic detector response in any given subset of data. Finally, the output of the MC simulation is then processed with the software package which reconstructs B decays and writes the output events in the BSTNTUPLE format.

Chapter 4

Flavor tagging

This chapter outlines the methods to infer the flavor of the B_s^0 meson at production, known as *flavor tagging*. We briefly describes the algorithms used in CDF on such purpose along with the procedure to calibrate their performances. Finally,we discuss the impact of the tagging in the β_s measurement.

4.1 Flavor tagging principles

The knowledge of the B_s meson flavor at the production is an important ingredient in the measurement of the B_s^0-$\bar{B}_s^0$ mixing phase. Indeed, the distinction of the B_s^0-meson flavor at production eliminates two of the four symmetries of the decay rate in Eq. (2.27) (see Sect. 2.2.2) and enhances the sensitivity to β_s through the access to the terms proportional to the oscillation frequency. Measuring if a B meson was produced as a B_s^0 or $\bar{B}_s^0$ is called *flavor tagging*. The properties of the $p\bar{p}\to b\bar{b}$ production process (see Fig. 4.1), and of the b quark hadronization and fragmentation (see Fig. 4.2), are used in two CDF flavor tagging algorithms, the Same Side Kaon Tagger (SSKT) and the Opposite Side Tagger (OST). They have been developed for the CDF B_s^0-$\bar{B}_s^0$ mixing measurement [43], and many references exist for both the OST[107, 108, 109] and the SSKT [110, 111, 112]. A complete description of the algorithm is beyond the scope of this thesis; only a brief overview of the methods is given.

Before proceeding with the description of the two algorithms, it is useful to introduce some quantities commonly used to characterize flavor taggers:

- the *tag decision* ξ. It is a discrete variable, that can take the value -1, 0, or 1. The tag decision id $\xi = -1$ if that the meson at the production has been tagged as $\bar{B}_s^0$; $\xi = 1$ if the initial meson has been tagged as $\bar{B}_s^0$; and $\xi = 0$ if the tagger could not make a decision.

- the *tagging efficiency* ε. It is the fraction of events for which a tag decision can be made. It is defined as follows:
$$\varepsilon = \frac{N_{\text{tag}}}{N_{\text{unt}} + N_{\text{tag}}} \tag{4.1}$$
where N_{tag} is the number of tagged events and N_{unt} is the number of events for which a tagging decision has not been taken.

- the *dilution* D, a quantity defined in order to characterize the rate of mis-tagging for a particular algorithm. It is defined as:
$$D = 1 - 2P_W = \frac{N_R - N_W}{N_R + N_W} \tag{4.2}$$

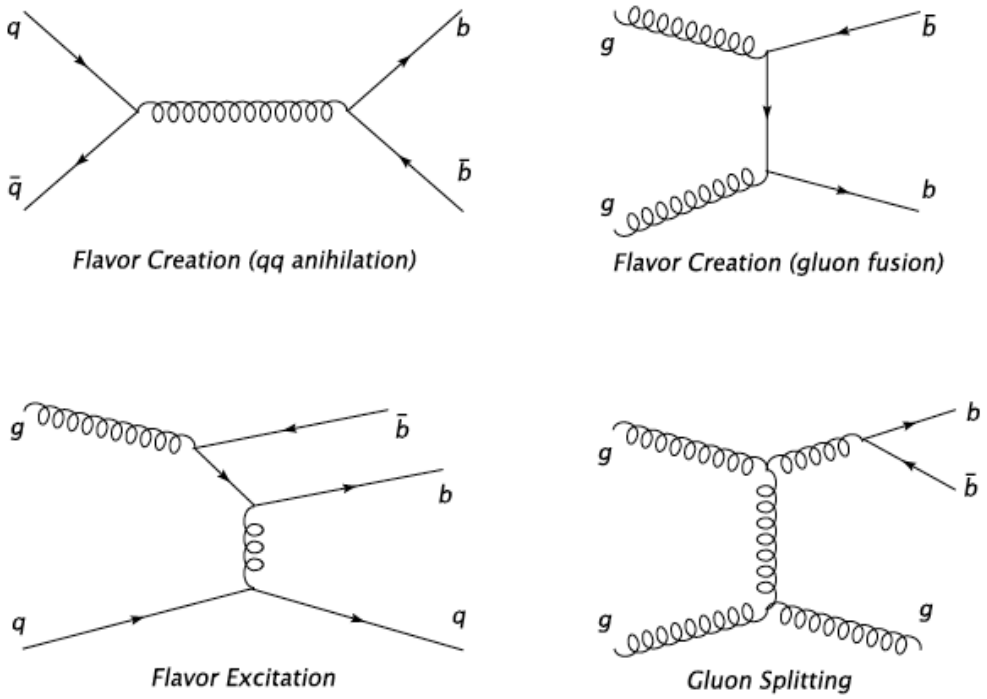


Figure 4.1: Feynman diagrams for $b\bar{b}$ pair production at lowest order. The top graphs are the lowest order contributions of the so called *flavor creation*, where only a quarks and an antiquark or two gluons interact. Bottom-left is the *flavor excitation*; bottom-right is the *gluon splitting*. In the flavor excitation a b from the sea quark of one of the interacting particles is scattered out in the strong interaction with a parton of the other hadron. The gluon splitting takes place when a gluon produced in the hadron collision results in a $b\bar{b}$ pair in the fragmentation process.

where N_R is the number of right-tagged events, N_W is the number of wrong-tagged events, and P_W is the probability for an event to be wrongly tagged. With this definition, when the tagger does not work properly, and it assigns randomly the tag decision, $P_W = 0.5$, so $D = 0$. In the case of an algorithm that makes no mistakes in the tag decision assignment, the dilution is $D = 1$. Moreover P_W is parametrized as a function of the calibration parameters, so that the dilution ca be predicted event by event.

- the *effective tagging efficiency* defined as εD^2, where D in this case represents the average dilution over the whole sample, is usually used as the figure of merit to measure the performance of a flavor tagger.

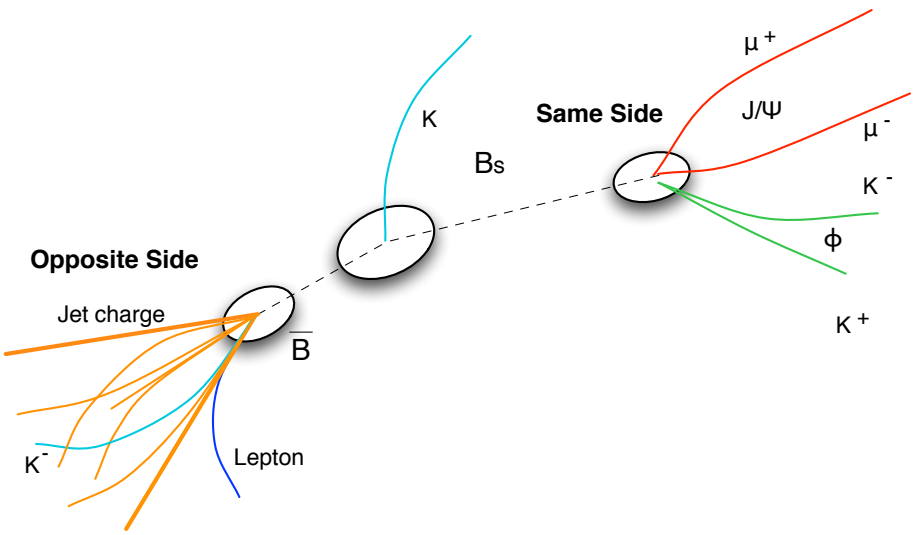


Figure 4.2: Sketch of the production and decay of a B_s^0 meson. The event's features exploited by the SSKT and the OST are shown.

4.2 Opposite Side Tagger

The OST exploits the incoherent production of $b\bar{b}$ at Tevatron, Fig. 4.1, and uses the information from the hadronization and decay of the other b-quark produced in the $p\bar{p}$ collision along with the b-quark belonging to the B meson of interest. The b or $\bar{b}$ quark on the opposite side of the B candidate hadronizes into a b-hadron. The b-hadron can be tagged by its decay products, although the situation is somewhat complicated if it is a B^0 meson, which can mix before decaying. If the opposite side quark is tagged as a b, the B candidate must contain a $\bar{b}$, and if the opposite side quark is tagged as a $\bar{b}$, the B candidate must contain a b. The opposite side flavor tagger is a combination of several algorithms: the soft muon tagger (SMT), the soft electron tagger (SET), and the jet charge tagger (JQT). The names refer to the charged object used to determine the flavor of the opposite side b quark.

The lepton taggers SMT and SET utilize the charge of a muon or of an electron, respectively, to determine the production flavor of their parent B meson. As they are the products of semi-leptonic decays $b \to c\,\ell^-\bar{\nu}_\ell X$ and $\bar{b} \to \bar{c}\,\ell^+\nu_\ell X$, one can infer that a positive muon or positron comes from a $\bar{b}$ at production, and a negative muon or electron from a b. The efficiency of these taggers is rather low, of order 20%, and is defined by the branching fraction of B to semileptonic decays [8]. Furthermore, detector and tracking inefficiencies lower the lepton-tagging algorithms efficiency. The mis-tag rate is increased by the possibility of mixing, if the lepton came from a B^0 decay. Mis-tags are also caused by the presence of $b \to c \to \ell X$, which can be misidentified as $b \to \ell X$ decays. In this case, the lepton will have the opposite sign as if it came directly from a b transition, and the tag decision will be the opposite of the correct decision.

The predicted tagging dilution for the lepton taggers is a function of p_T and of the lepton likelihood (the confidence in the lepton identification derived from calorimeter and dE/dx data). Well identified leptons with high p_T lead to a good dilution. The improvement of tagging performances with higher-p_T leptons is due to the fact that leptons from b decays are likely to have higher momentum in the transverse plane than those from lighter quarks, due to a larger available phase space.

The JQT infers the flavour of the candidate B_s^0 from the charge of the opposite-side b-jet. The jet charge is calculated as the momentum weighted sum of all the jet particles charge. Tracks are requested to be isolated from the candidate B meson, as it is important to look at jets only in the opposite side. Neural network algorithms are then used to find the jet most likely to come from a b quark. If the net charge is approximately $-1/3$, the jet is presumed to have come from a b quark, if it is $+1/3$, the jet is presumed to have come from a $\bar{b}$ quark. The dilution for this tagger is parametrized as a linear function of the jet charge and the probability that the jet contains a b (or $\bar{b}$ quark).

These three OST are not independent, since they can share tracks, and this effect needs to be accounted for when using the three taggers together. They are combined to give a single opposite side tagging decision by means of a neural network procedure. All OST processes are independent from the B candidate hadronization, so it is possible to use the same opposite side tagging algorithm for sample of different B meson types. This means that the algorithms can be developed or calibrated with the high statistics of B-light mesons samples, and then applied in the B_s^0-decays analysis.

4.2.1 Calibration of OST in the full Run II sample

The combined OST algorithm is calibrated on the dimuon samples,[1] for determining a global *scale factor* S_D, which will be applied to the event by event dilution to account for the mismatch between the predicted dilution and the actual dilution of the data. This is achieved by applying the OST algorithms to the full Run II data sample of fully reconstructed $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ decays (charge conjugates implied everywhere). The tagging decision and associated dilution of the algorithm is compared to the actual b quark content of the meson at decay time, which is known from the charge of the kaon. This is also the flavor at production since charged B do not oscillate. Two scale factors are estimates, for B^+ and B^- separately, to account for any charge related asymmetry in the tagging algorithms.

[1] the dimuon sample is a sample enriched of $J/\psi \to \mu^+\mu^-$ decays, collected online by means of the dimuon trigger that will be discussed in Sect. 5.1.

4.2.1.1 Data sample

We reconstruct exclusive $B^+ \to J/\psi K^+$ using the dimuon trigger. We use data collected throughout the whole Run II and corresponding to an integrated luminosity of $\mathcal{L} \approx 10\,\text{fb}^{-1}$. The offline selection is applied in two stages. After event reconstruction the events are subject to a loose preselection, aimed at speeding up the downstream processing, which comprises the following requirements:

- $5.16 < M(J/\psi K^+) < 5.40\,\text{GeV}/c^2$;
- ≥ 3 axial hits per track in the silicon detector for muons;
- ≥ 3 axial hits per track in the silicon detector for the kaon;
- successful XFT-muons match;
- $0.0 \leq \sigma_{ct} \leq 0.1\,\mu\text{m}$.

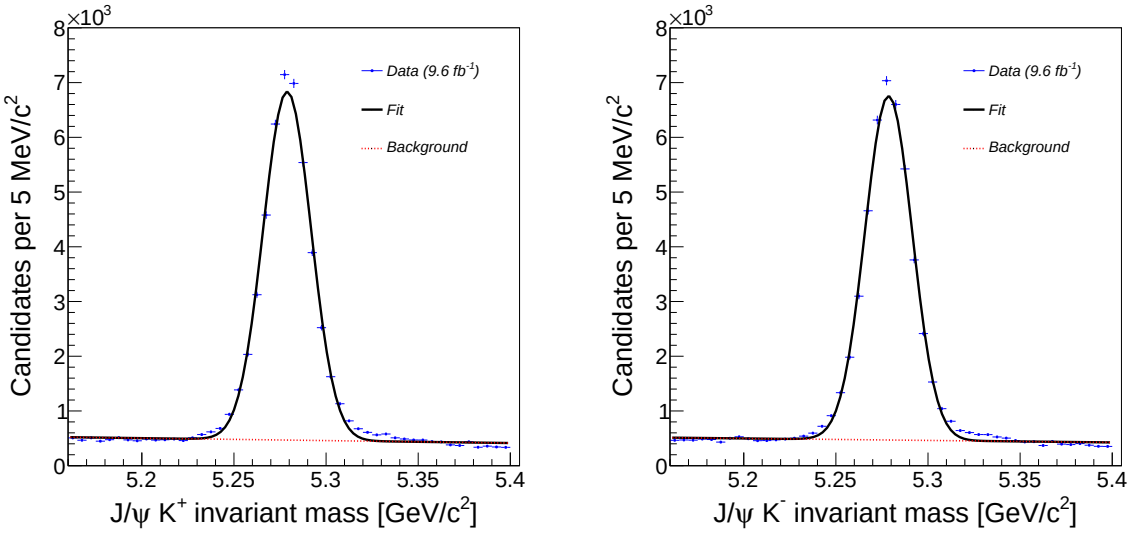
Then the preselected data are fed into an artificial Neural Network (NN) [113], an artificial classifier that combines the event variables into a single output variable and classifies whether an event is signal-like or background-like, to achieve an improved background suppression.[2] We reused the NN training and optimization discussed in [18], selecting events with a NN output ≥ 0.8 which had been shown to maximize the $S/\sqrt{S+B}$ figure of merit, where S is the number of signal events, and B is the number of background events. In addition to the NN requirement, we impose also a threshold on the decay length at $60\,\mu\text{m}$. This rejects a large part of the combinatorial background while preserving about 85% of the signal. The scale factors that will be obtained with this additional requirement exhibit an increased consistency between B^+ and B^- sample, allowing the use of a single, common scale factor.

The resulting $B^+ \to J/\psi K^+$ sample is shown in Fig. 4.3. A simple gaussian fit over a linear background finds approximately 40 000 B^- decays and 41 000 B^+ decays. The signal yield is consistent with what we expected from the previous iteration (see Fig. 4.4). Central mass values and widths are consistent as well.

4.2.1.2 Calibration procedure and results

For each $B^\pm \to J/\psi K^\pm$ decay, we compare the true flavor (as indicated by the kaon charge) with the flavor identified by the OST algorithm. Indeed, for each event the tagging algorithm provides a tag (b or $\bar{b}$) and a predicted dilution that quantifies the reliability of the tag. The algorithms to compute the dilution were designed and developed at the beginning of Run II using semileptonic B decays, where the charge of the lepton identified the flavor of the parent B hadron. However the lepton could come from a mixed B^0 meson or from a sequential $b \to c \to \ell$ decay. These and other subtler effects yield mistags. Hence, the performance of the algorithm in the current

[2] The basics of the NN will be described in Sect. 5.3.1.



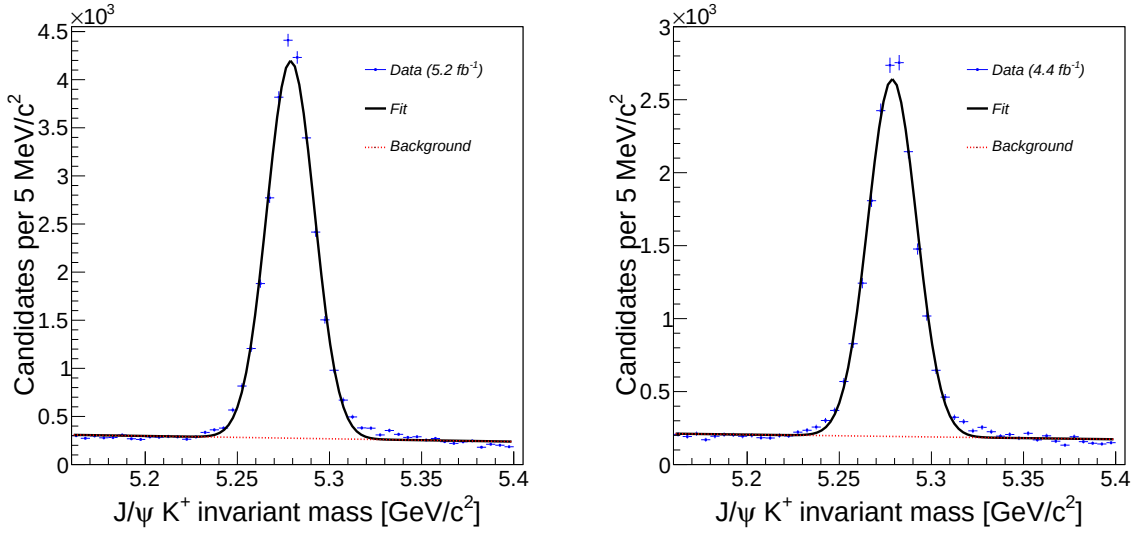
(a) (b)

Figure 4.3: The $J/\psi K^+$ mass distribution (a) and the $J/\psi K^-$ mass distribution (b) for the full Run II data sample.

B^+ sample could slightly deviate from the performance as predicted using semileptonic decays. Such deviations are modeled through a scale factor correction, which should be extracted from data. This is done by comparing the known dilution (since we know the B^+ flavor exactly) with the dilution predicted by the algorithm.

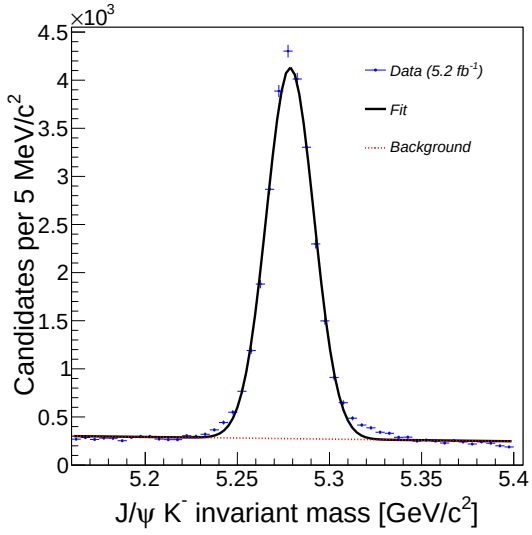
We divide the sample in independent subsamples according to their predicted dilution. For each bin of predicted dilution we count the number of right (wrong) tags to extract the actual dilution. Then we graph the actual dilution as a function of the predicted one (Fig. 4.5) to determine the scale factor. All dilutions distributions are background-subtracted. The scale factor is determined as the slope of the straight line fits of Fig. 4.5. For the entire dataset, we use two scale factors for the opposite side tagger, one for the B^+ and one for the B^-, in order to allow for any asymmetry in the tagging algorithms. We find $S_D^+ = 1.09 \pm 0.05$ and $S_D^- = 1.08 \pm 0.05$ respectively with a total average dilution of $\overline{D} = 6.88 \pm 0.03\%$. Since the calibrated values of the scale factor for the B^+ and B^- are approximately equal we use the average of the two in the fit and the error to compute the measurement as a systematic uncertainty. The measured scale factor in ≈ 10 fb^{-1} of data is $S_D = 1.085 \pm 0.05$, with a tagging efficiency of $92.8 \pm 0.1\%$ and a mean predicted dilution of $11.30 \pm 0.06\%$. The total effective tagging power is $\epsilon(S_D D)^2 = 1.39 \pm 0.05\%$.

As a check, we determine separately the scale factors for the first 5.2 fb^{-1} of data, and for last 4.8 fb^{-1} of data, in Fig. 4.6. No major drift in performance are identified In addition, we determine OST tagging efficiency and dilution for different periods of data and summarize

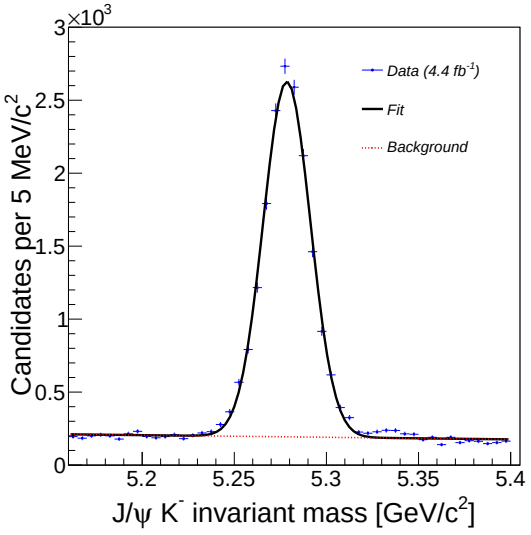


(a)

(b)



(c)

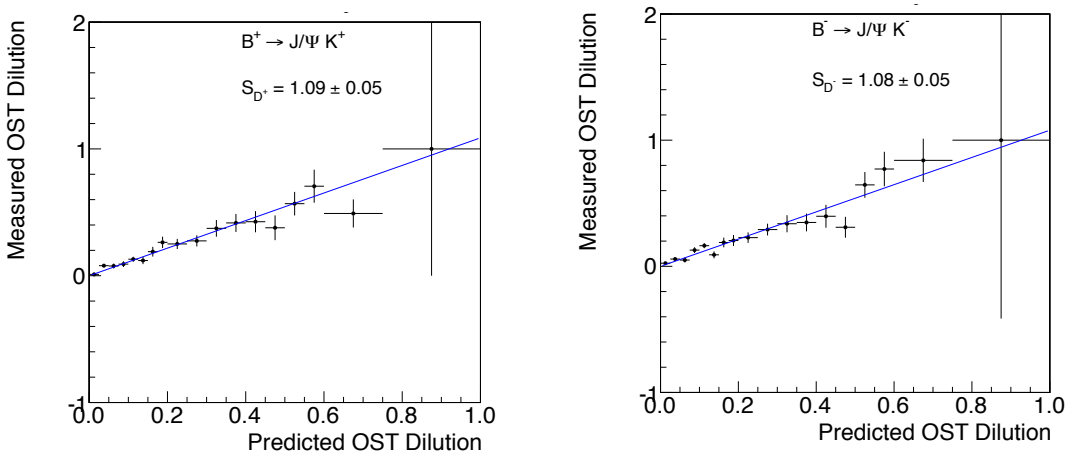


(d)

Figure 4.4: Mass distributions of $J/\psi K^+$ for the first $5.2\,\mathrm{fb}^{-1}$ (a) and for the second $4.8\,\mathrm{fb}^{-1}$ (b) of data. Mass distributions of $J/\psi K^-$ for the first $5.2\,\mathrm{fb}^{-1}$ (c) and for the second $4.8\,\mathrm{fb}^{-1}$ (d) of data.

Scale Factor	previous calibration [18]	0–5.2 fb^{-1}	5.2–10 fb^{-1}	0–10 fb^{-1}
${S_D}^+$	0.93 ± 0.09	1.09 ± 0.06	1.08 ± 0.08	1.09 ± 0.05
${S_D}^-$	1.12 ± 0.10	1.06 ± 0.07	1.10 ± 0.08	1.08 ± 0.05
ϵ	$94.3 \pm 0.3\%$	$93.8 \pm 0.1\%$	$91.2 \pm 0.2\%$	$92.8 \pm 0.1\%$
$\sqrt{\overline{D^2}}$	$11.5 \pm 0.02\%$	$11.26 \pm 0.08\%$	$11.36 \pm 0.10\%$	$11.30 \pm 0.06\%$

Table 4.1: OST performance for B^+ and B^- in different parts of the data, compared with the 5.2 fb^{-1} analysis.



(a) (b)

Figure 4.5: Measured dilution as a function of predicted dilution for B^+ (a) and B^- (b) for the final Run II sample.

the results in Tab. 4.1. We derive the scale factors and efficiencies in periods of data with approximately similar statistics ($\sim 1.7\,\text{fb}^{-1}$ each), to ensure stability and consistence throughout all parts of the data. Figure 4.7 shows as the scale factors are fairly stable through data but a tendency towards a decreased tagging efficiency as a function of time (data-taking periods) is observed. The decrease in average efficiency does not exceed 3% by comparing earlier data with later data. This is consistent with a luminosity-dependent effect, as shown in Fig. 4.7, where data are shown as a function of time (integrated luminosity) and as a function of instantaneous luminosity. Such a reduction in efficiency impacts only (very marginally) the resolution on the extracted physical parameters but is not expected to bias them.

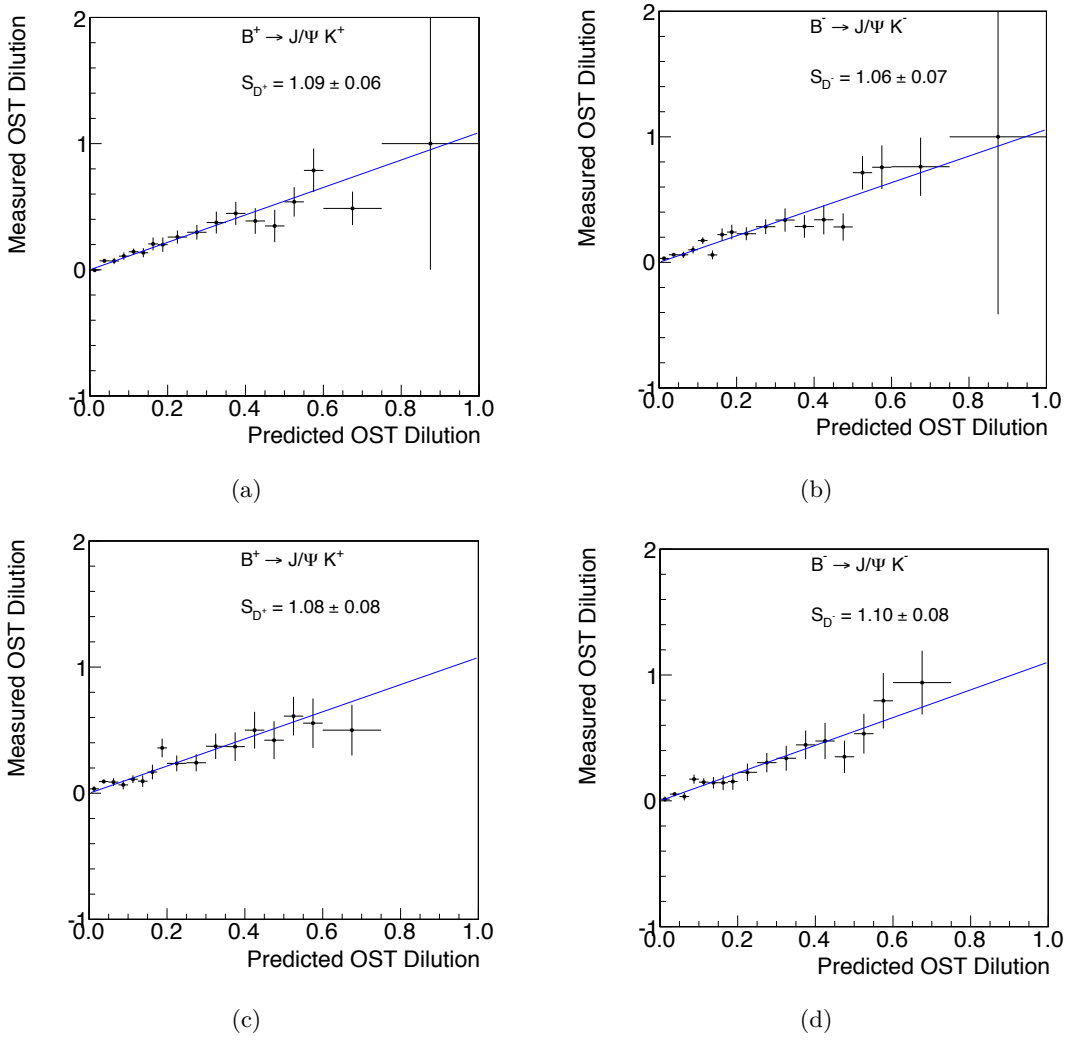


Figure 4.6: Measured dilution as a function of predicted dilution for B^+ (a) and B^- (b) for the first $5.2\,\mathrm{fb}^{-1}$ of data. Measured dilution as a function of predicted dilution for B^+ (c) and B^- (b) for the last $4.8\,\mathrm{fb}^{-1}$ of data.

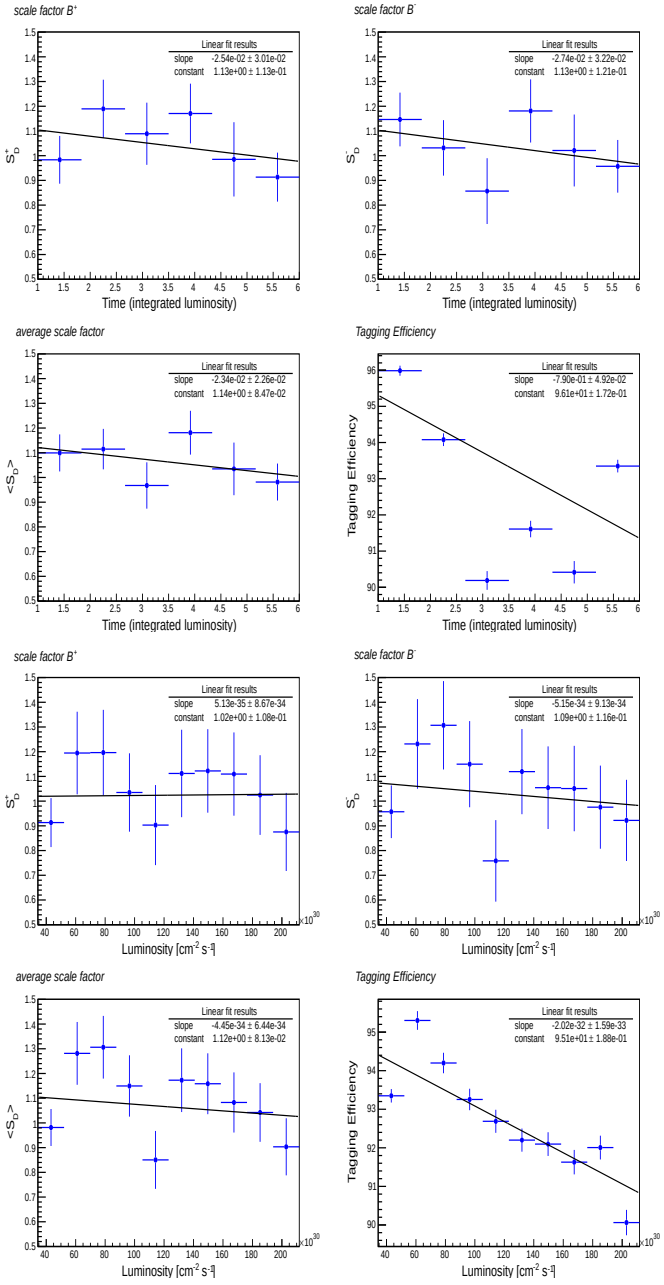


Figure 4.7: Top four plots: OST dilution scale factors for B^+ (top, left), B^- (top, right) and an average of the two (bottom, left) and efficiency (bottom, right) as a function of time (integrated luminosity). Each point in the x-axis correspond to approximately $1.7\,\mathrm{fb}^{-1}$ of data. Bottom four plots: OST dilution scale factors for B^+ (top, left), B^- (top, right), average them (bottom, left) and tagging efficiency (bottom, right) as a function of instantaneous luminosity.

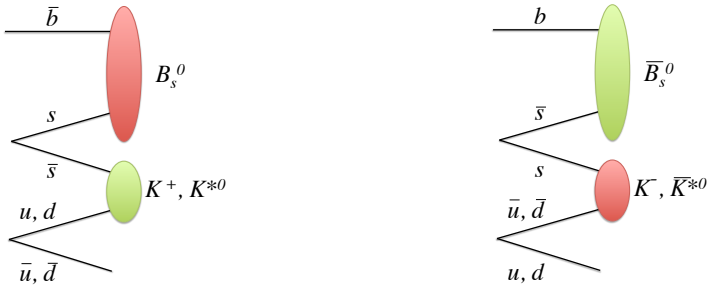


Figure 4.8: Sketch of the fragmentation of a B_s^0 and of a $\bar{B}_s^0$ meson.

4.3 Same Side Kaon Tagger

The SSKT uses the fragmentation tracks of the B mesons candidate (meaning the B meson of interest) to determine its flavor. The tagger is supposed to identify the flavor of the s-quark (*i. e.* s or $\bar{s}$) of the B_s^0 meson candidate in a kaon produced alongside the B_s^0 (or $\bar{B}_s^0$). In the case of a B_s^0 meson ($\bar{b}s$) at the production, the *strange* quark of the kaon, that is the produced partner of the s of the B_s^0, should be a $\bar{s}$. If we have a $\bar{B}_s^0$ at the production, the kaon should be formed by an s quark. In other words if a K^+ is identified, the meson at the production was a B_s^0, instead in the case of a K^- identification the initial meson was a $\bar{B}_s$. The principle is sketched in Fig. 4.8. The SSKT power benefits from the use of PID to identify the associated track as a kaon. This decreases the mistag rate, which occurs if a pion is mistakenly identified as an associated kaon. Particle identification, as well as information about track momentum, is used to decide which track to choose as the tagging track, if there are multiple possible tagging tracks in an event.

Obviously, the SSKT is specific to B_s^0 decays, eliminating the possibility of a calibration in a high statistics samples of B-light meson decays, as in the case of OST. Its last calibration was done using $5.2\,\text{fb}^{-1}$ of data by repeating the measurement of the B_s^0-$\bar{B}_s^0$ mixing frequency and extracting the dilution [18]. The probability for observing the B_s^0 in a flavor eigenstate as a function of time is:

$$P(t)_{B_s^0(\bar{B}_s^0)} \propto |1 \pm \cos \Delta m_s t| \tag{4.3}$$

Adding in the effect of the measured dilution, our measurement probability becomes:

$$P(t)_{B_s^0(\bar{B}_s^0)} \propto |1 \pm D \cos \Delta m_s t| \tag{4.4}$$

This can be re-expressed introducing an amplitude, A, such that:

$$P(t)_{B_s^0(\bar{B}_s^0)} \propto |1 \pm A D_p \cos \Delta m_s t| \tag{4.5}$$

where D_p is the predicted dilution. The likelihood that uses this probability is normalized such that at the correct value of the mixing frequency Δm_s the amplitude A is one. The mixing frequency can be measured by what is called an *amplitude scan*: the amplitude is measured

for each value of Δm_s in a given range. For the true value of Δm_s, the amplitude should be one, while it is zero when Δm_s is far from the true value. In the calibration, the amplitude range values are interpreted in the following way: an amplitude consistent with one means that the tagger assesses its decisions and thus its performance correctly. A value smaller than one indicates that it overestimates itself. According to that a value greater than one implies an underestimation of the decision power. After the mixing amplitude is determined, it serves as a scale factor for the dilution. The mixing frequency is determined simultaneously together with the amplitude; the main reason for doing this is providing a cross-checks, by comparing the mixing frequency with previous measurement [43], and thus to inspire confidence into the calibration.

The calibration of the SSKT with the Δm_s measurement was made on $5.2\,\text{fb}^{-1}$of data [114]. Four channels collected using the *displaced vertex trigger* were employed:[3]

- $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-, \phi \to K^+K^-$;
- $B_s^0 \to D_s^- \pi^+, D_s^- \to K^*K^-, K^* \to K^+\pi^-$;
- $B_s^0 \to D_s^- \pi^+, D_s^- \to \pi^+\pi^+\pi^-$;
- $B_s^0 \to D_s^- \pi^+\pi^+\pi^-, D_s^- \to \phi\pi^-, \phi \to K^+K^-$.

where about half of the signal events come from the first channel. The Δm_s amplitude scan on $5.2\,\text{fb}^{-1}$ of data is shown in Fig. 4.9. The maximum amplitude in Δm_s occurs at $\Delta m_s = 17.79 \pm 0.07\,\text{ps}^{-1}$, a value consistent with the CDF measurement [43]. The size of the amplitude at maximum and the measured dilution scale factor for the SSKT is $S_D = 0.94 \pm 0.2$. We measure a tagging efficiency of 52.2 ± 0.7, and an average predicted dilution on B_s^0 signal of $25.9 \pm 5.4\%$. The SSKT tagging power is $3.5 \pm 1.4\%$.

We have not yet extended the calibration of the SSKT to the full $10\,\text{fb}^{-1}$ data sample. This is due to the marginal increase in calibration sample statistics with respect to the $5.2\,\text{fb}^{-1}$ analysis. Figure 4.10(a) compares the mass distribution of the $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi\pi^-, \phi \to K^+K^-$ signal for samples corresponding to $5.2\,\text{fb}^{-1}$ and $7.2\,\text{fb}^{-1}$ datasets. The 10% increase in sample statistics against a 40% increase in nominal integrated luminosity is due to the displaced-tracks trigger being severely suppressed in the latest period of the data-taking operations (see Sect. 9.1). We therefore use the calibration obtained with $5.2\,\text{fb}^{-1}$ of data [18] and, accordingly, use this tagger only in the first half of our sample. This conservative choice prevents potential problems arising from variations of the tagger performances as a function of time. The impact of the missing SSKT in the latest data is limited. We check the impact on the resolution of β_s by means of pseudo-experiments. We generate a large set of simulated $B_s^0 \to J/\psi\,\phi$ samples with the same statistic as in data (signal + background) and we perform the analysis to measure β_s for each sample, in two configurations; in the first one, the SSKT is used in all the data sample; in

[3]The displaced vertex trigger will be described in Sect. 9.1.

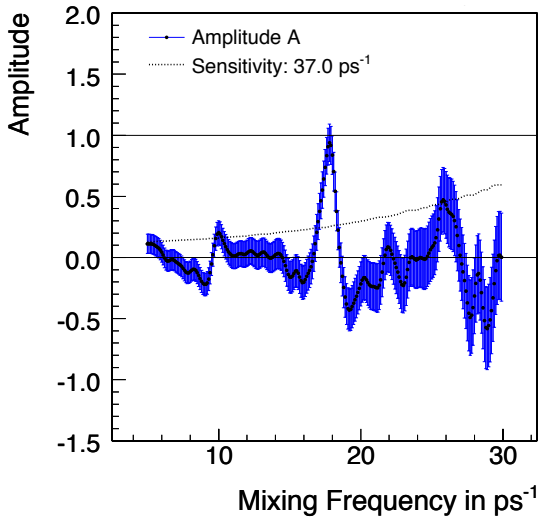
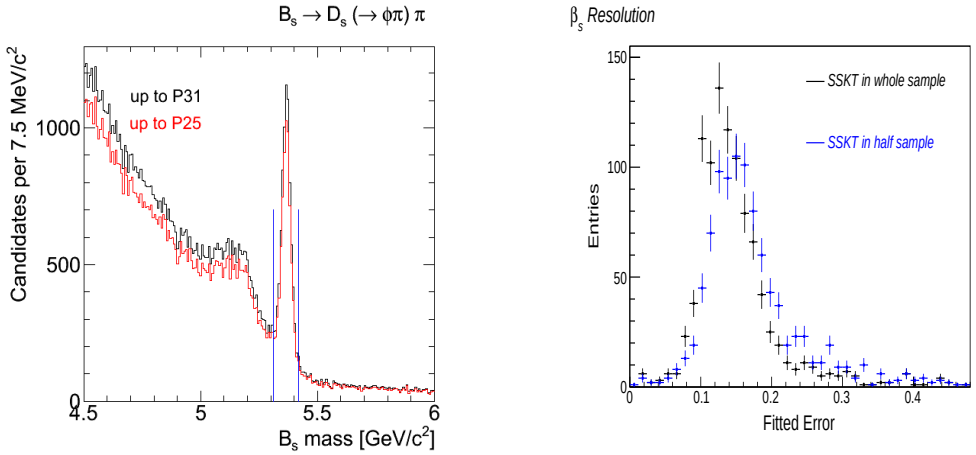


Figure 4.9: Amplitude scan as function of Δm_s using $5.2\,\text{fb}^{-1}$ of data to determine the SSKT dilution scale [114].

the second one, the SSKT is exploited in a fraction of data only, corresponding to the size of the first $5.2\,\text{fb}^{-1}$ of data. In both cases, in the analysis and in the generation of the samples we employ the SSKT performances measured in the calibration. This investigation shows that the average worsening in the β_s resolution when restricting the SSKT to half of the sample is only 12%, as shown in Fig. 4.10(b). Considering that the new data will suffer from a considerably larger uncertainty on SSKT dilution that gain will probably be even smaller.



(a) (b)

Figure 4.10: In (a), $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-, \phi \to K^+K^-$ mass distribution in $5.2\,\text{fb}^{-1}$ (red) and $7.2\,\text{fb}^{-1}$(black) of data. In (a), comparison between distributions of the β_s uncertainty from the analysis of simulated samples with SSKT used in the whole data sample (black points) and up to 54% of the sample only (blue points).

experiment	OST (%)	SSKT (%)
CDF	1.39 ± 0.05	3.5 ± 1.4
D0	2.48 ± 0.21	–
LHCb	2.1 ± 0.1	1.3 ± 0.4

Table 4.2: Summary of the tagging power at different experiment, for the OST and SSKT, in the $B_s^0 \to J/\psi\phi$ analysis.

4.4 Effect of the tagging in the β_s measurements

The total tagging power is quite low at hadron colliders, $\mathcal{O}(4\%)$, compared to the tagging power at the B factories, $\mathcal{O}(30\%)$. We summarize the tagging power performances for different experiments in Tab. 4.2. The main effect of the tagging in this analysis is to break the $\beta_s \to -\beta_s$ symmetry discussed in Sect. 2.2.2, that exists in the untagged decay rate, removing half of the allowed region in the $(\beta_s, \Delta\Gamma_s)$ space.

However, the tagging power is not large enough to substantially reduce the errors on the remaining solutions, and each of the four untagged solutions has comparable errors to those on the tagged solutions, as shown in Fig. 4.11, where we compare the results in the $(\beta_s, \Delta\Gamma_s)$ plane for tagged and untagged analysis of one simulated samples of $B_s^0 \to J/\psi\phi$ decays. If the tagging power were greater, we would expect our sensitivity to β_s to be substantially better in the tagged case, and the errors on β_s to be smaller; the tagging allows to access to the following terms of the decay rate:

$$\sin 2\beta_s \sin(\Delta m_s t) \;\text{ and }\; \cos 2\beta_s \sin(\Delta m_s t), \tag{4.6}$$

that are not present in the untagged rate. Greater the tagging power ϵD^2, greater is the equivalent fraction of the total sample $\epsilon D^2 N$ (where N is the total number of events), that provides the two additional terms for measuring β_s. Nevertheless, the sensitivity provided by these terms depends on the experimental resolution on the B_s^0 decay-time; since Δm_s is quite large ($\approx 18\,\text{ps}^{-1}$), better the sensitivity, better the resolution of the fast B_s^0-$\bar{B}_s^0$ oscillation. Then, effectively, the different sensitivity between the tagged and untagged analysis is due by the combined effect of the tagging power and the decay-time resolution. The latter will be discussed in Sect. 6.2.2.

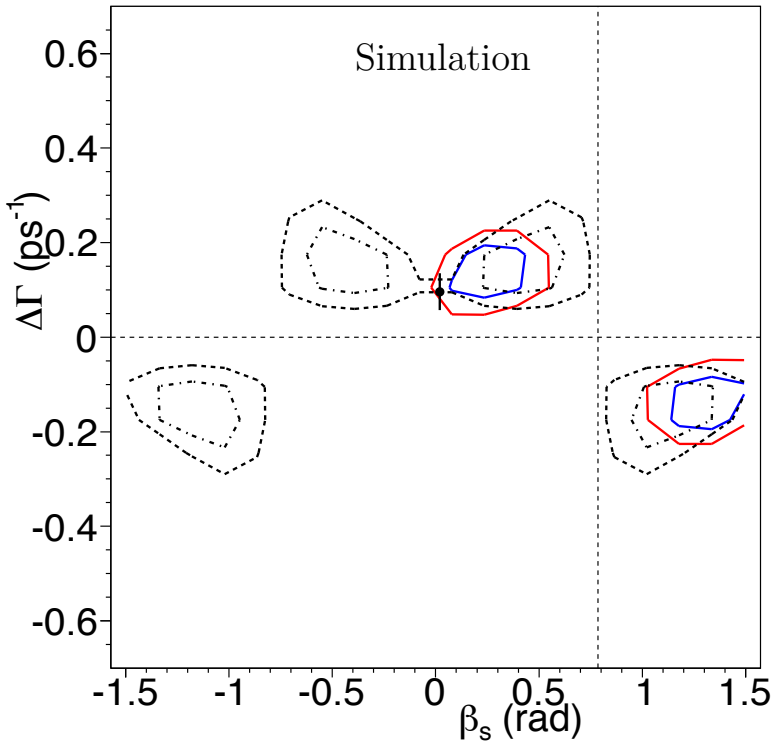


Figure 4.11: Confidence regions at 68% C.L. (blue) and 95% C.L. (red) for the analysis of a pseudo-experiment with (bold lines) and without (light lines) using the tagging information.

Chapter 5

$B_s^0 \to J/\psi\phi$ Data Set

This chapter describes the data samples for the analysis of $B_s^0 \to J/\psi\phi$ decays. We present the selection of the data. We outline the system used to trigger on events with two muons for collecting the candidates of interest. An offline selection, based on a neural network discriminator, is then applied to optimize the measurement of B_s^0 mixing phase. The final optimized sample for the measurement is presented. Finally, the MC simulated data samples used in several parts of the analysis are described.

5.1 The dimuon trigger

The data sample used in the analysis of $B_s^0 \to J/\psi\phi$ decays was collected by CDF over the entire period of the Tevatron operations in the Run II, since March 2001 till September 2011. It corresponds to an integrated luminosity of approximately $9.6\,\text{fb}^{-1}$, after discarding all data taking periods during which a detector subsystem critical for the analysis was malfunctioning or turned off. A sample enriched of $B_s^0 \to J/\psi\phi$ decay's candidates is selected online by means of a trigger system based on the reconstruction of two muon candidates with opposite charge originating by the same vertex; this is the dimuon trigger [115].

The dimuon trigger relies on a clear signature of two muons coming from $J/\psi \to \mu^+\mu^-$ decays. In order to make trigger decisions, it uses the XFT tracking and muon system information available at L1. The L2 and L3 play a small role in the event selection decision. L2 is used to tighten any existing requirements of L1, *e. g.*, on the transverse momentum, and L3 uses more precise determination of several event variables, such as the transverse momentum of tracks, better track-stub matching, dimuon mass, etc. The dimuon trigger data is an ideal sample to develop time-dependent analysis, because the trigger does make any requirement on the impact parameter of the XFT tracks. An impact parameter cut would distort the B_s^0 proper decay time distribution and complicate the time-dependent portion of the measurement.

Although we refer to it as a single entity, the dimuon trigger is in fact a combination of two triggers: CMU-CMU, where both muons are found in the most central muon chambers, and CMU-CMX, where one muons is found in the CMU and one in CMX. We describe the CMU-CMU trigger, and then comment on the differences in the CMU-CMX one. The following terminology is specific to triggering on CMU muons. A *stack* is a set of four drift cells stacked on top of each other. The CMU has 288 stacks in each of the East and West sides of the detector. A L1 stub is a track segment in a stack such that cells 1 and 3 or cells 2 and 4 have hits separated by no more than 396 ns. A *tower* is a set of two neighboring stacks. A tower has

Level 1	Level 2	Level 3
2 XFT tracks with opposite charge 2 muon stubs for each track $p_T(\mu) > 1.5(2.2)\,\text{GeV}/c$ for CMU (CMX) $\Delta\phi_6(\text{CMU},\text{CMU}) < 135°$	$\lvert z_0(\mu_1) - z_0(\mu_2)\rvert < 5\,\text{cm}.$	$2.7 < M(\mu^+\mu^-) < 4\,\text{GeV}/c^2$

Table 5.1: Basic dimuon trigger requirements.

fired when one or both stacks have a L1 stub, and is empty otherwise. A muon tower is a fired tower matched with an XFT track. In order to keep the L1 decision time within the L1 latency, only the information about which towers have fired is used in triggering, rather than detailed hit positions and direction. The XFT reports the p_T and ϕ_6, as well as the charge of the track to the XTRP. The XTRP extrapolates this track to the CMU radius and creates a footprint, a 3σ window in ϕ (wide enough to account for multiple Coulomb scattering). If a tower is found within that footprint, it is a muon tower. The CMU-CMU trigger requires that at least two muon towers are found such that they are either on opposite sides of the detector or are separated by at least two other towers. The CMU-CMX trigger uses a very similar algorithm. The changes to the decision algorithm arise from the differences between the CMU and CMX detectors. In the CMU-CMX case, only XFT tracks with $p_T > 2.2\,\text{GeV}/c$ are used as the extra material that muons pass through to reach the CMX limits further the momentum requirements on the muon, and no azimuthal separation is required because the muons are by definition in different subdetector volumes.

Trigger algorithms are among the few elements of the experimental apparatus which are continuously improved and optimized, as this process does not require performing expensive and time-consuming hardware upgrades to the detector. The dimuon trigger has undergone constant revision in order to carry out such optimization. While the core logic outlined above is more or less constant, other parameters have been changed often to improve the trigger. Such parameters include requirements on the p_T of the XFT tracks, the difference in ϕ between the two muons ($\Delta\phi$), and their transverse mass (M_T). In addition, some of the triggers are prescaled, which means that only one every N events are kept. This is done in order to deal with periods of high luminosity when triggering on every event that passes the nominal requirements would overwhelm the DAQ system. An improved implementation of prescaling is the dynamic prescaling, where the trigger automatically adjusts its prescale over the data taking period as luminosity decreases, in order to accept a larger fraction of the fewer number of events that satisfy the nominal trigger conditions as the luminosity drops. The various combinations of these requirements result in slightly different trigger requirements. The basic dimuon trigger requirements are summarized in Tab. 5.1. The data collected by the trigger go through the offline data process described in Sect. 3.2.7. The final BSTNTUPLE output is then used for the reconstruction of the decays candidates and the offline selection.

5.2 Reconstruction and kinematic variables

The decay modes are fully reconstructed offline, meaning that all final state muon and kaon tracks have been identified to reconstruct the B_s^0 meson. The muon pairs are analysed and fitted to a common vertex to obtain an estimated vertex position and vertex-constrained tracks. Important quantities such as the J/ψ invariant mass and transverse momentum are then estimated from the refitted tracks. Then to find ϕ candidates, oppositely charged pairs of non-muon tracks coming from a displaced vertex determined by a kinematic fitting algorithm are examined. The two tracks are initially assumed to be kaons. At a later stage, a probability to actually be kaons is assigned based on *dE/dx* and TOF information. The kaons and muons tracks are used to locate the *secondary vertex*, *i. e.* the decay point, in a kinematic fit to a common vertex. After the 4-track vertex fit is performed, the best fit values of the B_s^0 momenta and its daughter particles are obtained. A dimuon mass constraint to the known J/ψ mass [8] improves the B_s^0 mass resolution. A $B_s^0 \to J/\psi\phi$ decay candidate is then formed.

The most important variable in the study of the time-evolution of the B_s^0 decay is the B-meson transverse decay-length (L_{xy}): the displacement of the secondary vertex with respect to the primary one, projected onto the transverse momentum vector of the decaying particle:

$$L_{xy} = \frac{\boldsymbol{p_T} \cdot \boldsymbol{V}}{p_T}, \tag{5.1}$$

where $\boldsymbol{V}$ is the 2-dimensional vector from the primary to the secondary vertex in the transverse plane, and $\boldsymbol{p}_T$ is the B transverse momentum vector. From this, the proper decay length, *i. e.* the time to decay in the hadron's rest frame times the speed of light, is:

$$ct = \frac{ML_{xy}}{p_T}, \tag{5.2}$$

where M is the world average B_s^0 meson mass [8]. Associated with the proper decay length is its event by event uncertainty σ_{ct} which is obtained from the error on L_{xy}, the components from the other parameters being treated as negligible.

The momenta of the daughter particles are used to calculate the transversity angles defined in Sect. 2.1.1. They can be calculated by first boosting the four-momenta of the decay particles into the B_s^0 meson rest frame. The world average B_s^0 mass [8] and the reconstructed momentum of the B_s^0 are used to calculate the boost vector. To then boost into the J/ψ frame the world average J/ψ mass [8] is used, together with the reconstructed J/ψ momentum, and to calculate the boost vector into the ϕ meson rest frame the reconstructed K^+K^- mass and momentum are used. The use of the world average or the reconstructed mass for each particle type is motivated by the natural width of the particle. For ϕ meson, its natural width is greater than the resolution of the CDF detector, so the reconstructed mass is used, but the B_s^0 and J/ψ have widths much smaller than the resolution so it is more accurate to use the world average mass.

The transversity angle Ψ of the K^+ is defined as:

$$\cos\Psi = -\frac{\boldsymbol{p}^{\phi}_{K^+} \cdot \boldsymbol{p}^{\phi}_{J/\psi}}{p^{\phi}_{K^+} p^{\phi}_{J/\psi}}, \tag{5.3}$$

where $\boldsymbol{p}^B_A$ is the three momentum of particle A in the rest frame of particle B. A coordinate system with versors $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ is defined in order to calculate the other two angles:

$$\begin{cases} \hat{\boldsymbol{x}} = \frac{\boldsymbol{p}^{J/\psi}_{\phi}}{p^{J/\psi}_{\phi}}, \\ \hat{\boldsymbol{y}} = \frac{\boldsymbol{p}^{J/\psi}_{K^+} - \left(\boldsymbol{p}^{J/\psi}_{K^+} \cdot \hat{\boldsymbol{x}}\right)\hat{\boldsymbol{x}}}{|\boldsymbol{p}^{J/\psi}_{K^+} - \left(\boldsymbol{p}^{J/\psi}_{K^+} \cdot \hat{\boldsymbol{x}}\right)\hat{\boldsymbol{x}}|}, \\ \hat{\boldsymbol{z}} = \hat{\boldsymbol{x}} \times \hat{\boldsymbol{y}}. \end{cases} \tag{5.4}$$

Then, the angles Φ and $\cos\Theta$ are:

$$\Phi = \arctan\left(\left(\frac{\boldsymbol{p}^{J/\psi}_{\mu^+}}{p^{J/\psi}_{\mu^+}} \cdot \hat{\boldsymbol{y}}\right) \Big/ \left(\frac{\boldsymbol{p}^{J/\psi}_{\mu^+}}{p^{J/\psi}_{\mu^+}} \cdot \hat{\boldsymbol{x}}\right)\right), \tag{5.5}$$

$$\cos\Theta = \frac{\boldsymbol{p}^{J/\psi}_{\mu^+}}{p^{J/\psi}_{\mu^+}} \cdot \hat{\boldsymbol{z}}. \tag{5.6}$$

The signs of $\boldsymbol{p}^{J/\psi}_{\mu^+} \cdot \hat{\boldsymbol{x}}$ and $\boldsymbol{p}^{J/\psi}_{\mu^+} \cdot \hat{\boldsymbol{y}}$ are used to resolve the ambiguity of the angle Φ.

5.3 Offline selection

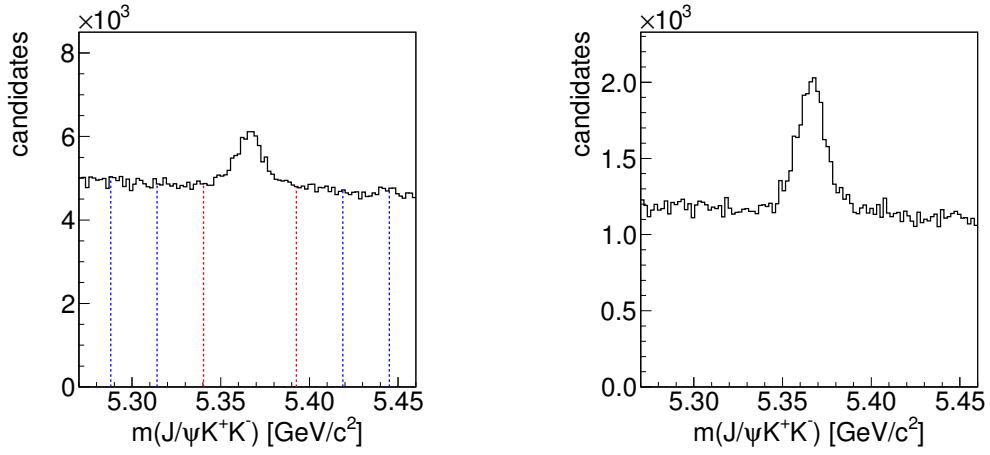
To get the maximal useful information from the data, we need to reduce the number of the uninteresting events (referred to as *background*) which degrade the sensitivity on the observables of interest. An optimized selection with respect to the measurement of the β_s is applied offline for that purpose. The offline selection is implemented in two steps. A set of basic and loose selection requirements (*rectangular cuts*) reduces the volume of downstream processing. Rectangular cuts veto any event which does not fall into an accepted range for any given variable. We then apply the final selection using a multivariate classifier, implemented through an artificial *Neural Network* (NN), which takes information from all input variables and combines it into a single decision variable. It is described in the next section.

The requirements applied in the first stage of the offline selection are:

- > 9 axial and stereo hits in the COT and > 3 hits in the silicon detector, for each track;
- $5.1 < m(J/\psi K^+K^-) < 5.6\,\mathrm{GeV}/c^2$;
- $3.04 < m(\mu^+\mu^-) < 3.14\,\mathrm{GeV}/c^2$;
- $1.009 < m(KK) < 1.028\,\mathrm{GeV}/c^2$;

- $p_T(B) > 4\,\text{GeV}/c$;
- $p_T(\phi) > 1\,\text{GeV}/c$;
- $p_T(K) > 0.4\,\text{GeV}/c$;
- $\chi^2_{r\varphi} < 50$,

where $\chi^2_{r\varphi}$ is the χ^2 of the 2-dimensional vertex fit in the $(r\varphi)$ plane. The resulting $J/\psi K^+K^-$ mass distribution is shown in Fig. 5.1 (a). A distinct signal enhancement is visible atop a significant background. In the following, we will use the *sidebands* of the reconstructed B_s^0 mass (m_B) distribution to study and model the background features. Sidebands are defined as: $(5.291 < m_B < 5.315) \cup (5.417 < m_B < 5.442)\,\text{GeV}/c^2$. The signal region is defined as the range $5.34 < m_B < 5.39\,\text{GeV}/c^2$, corresponding roughly to a $\pm 2.5\sigma$ window around the known B_s^0 mass, with mass resolution of $\sigma \approx 10\,\text{MeV}/c^2$.

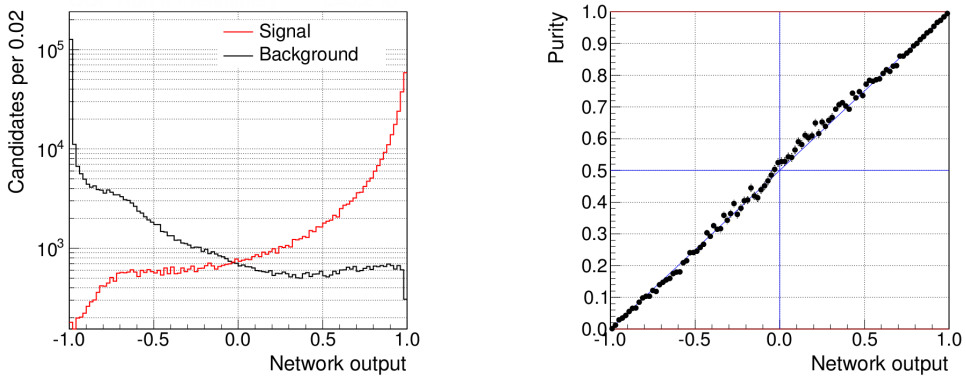


(a) (b)

Figure 5.1: Mass distribution of B_s^0 candidates after event preselection only (a), and after the rectangular cuts described in Sect. 5.3.1 for the data validation of variables in input to the NN (b). The blue dotted lines represent the bounds of the sidebands regions, while the red dotted lines represent the bounds of the signal region.

5.3.1 Neural Network Selection

The artificial NN classifier adopted in the optimization of the data sample is the `NeuroBayes` tool applied in the previous CDF analysis [113, 18]. The network was trained to recognize the signal properties using 0.35 million simulated signal events, and for the background using 0.3 million data events from the B_s^0 mass sidebands. The simulated signal data is described in detail in Sect. 5.5.1. The NN combines the information from all the inputs into a single output variable, that classifies whether an event is signal-like or background-like on a continuous scale



(a) (b)

Figure 5.2: NN output distribution for training sample (a), red line is signal, black is background. Signal purity as a function of NN output (b).

between -1 (background) and 1 (signal) (see Fig. 5.2 (a)). The following variables are used as input to the NN:

- $\chi^2_{r\phi}$, χ^2 of the 2-dimensional (r, φ) vertex fit;
- $P(\chi^2, p)$, χ^2 probability for the 3-dimensional vertex fit;
- $p_T(p)$, momentum component transverse to the beam direction for particle p;
- $LL_\mu(p)$, value for a likelihood based quantity used for muon identification;
- $LL_K(p)$, value of a likelihood based discriminant for kaon identification.

The muon and kaon likelihoods are quantities used for particle identification. The algorithm determining the muon likelihood is described in Ref. [116]. The kaon likelihood [117] is a combined discriminant constructed from the kaon track specific ionization energy loss, dE/dx, and its TOF information. The variables are ranked in order of decreasing discriminating power as follows: the transverse momentum p_T of the ϕ meson, the kaon likelihood, the muon likelihood for the J/ψ muon daughters, $\chi^2_{r\phi}$ for the B_s^0 decay vertex reconstruction, the transverse momentum p_T of the B_s^0 meson, and the vertex probabilities associated to the B_s^0, ϕ, and J/ψ candidates. The output of the NN for signal and background using the training samples is shown in Fig. 5.2 (a). In Fig. 5.2 (b) the signal purity as a function of NN output is shown, where the gradient shows the expected correlation between the network output and the signal purity.

The NN was retrained in the last iteration of this analysis as described in Ref. [18]. Because the data quality is known to be very similar between data added in this update and previous data, we expect the performance of the NN discriminator in our sample to be to a good approximation comparable with the performance on the $5.2\,\text{fb}^{-1}$ sample of the latest analysis. Hence, we do not explore the possibility to include new variables nor retraining the NN. Indeed,

it is very unlikely that any selection bias would be induced in the results, although the final selection could be a little suboptimal. However, we compared the distributions of all network input variables in the first $5.2\,\mathrm{fb}^{-1}$ of data and the second $4.4\,\mathrm{fb}^{-1}$. We expect this comparison to be sensitive to effects such like large changes in the trigger-path composition of our sample, since kinematic thresholds may differ in each path. The data were required to satisfy the following requirements $p_T(\mu) > 1.5\,\mathrm{GeV}/c$, $p_t(K) > 0.6\,\mathrm{GeV}/c$, four-track vertex probability > 0.001, and $p_T(B) > 5\,\mathrm{GeV}/c$. The B_s^0 mass distribution after such selection is shown in Fig. 5.1 (b). Sideband-subtracted distributions have been compared at this stage, where the nomenclature *sideband subtraction* stands for an operation performed on real data to extract the shape of the signal distributions. It is done by subtracting to a given distribution of the events in the signal region of the B_s^0 mass distribution the distribution of those events which are located in the sidebands, properly normalized to the background fraction expected in the signal region. Figure 5.3 and Fig. 5.4 show the comparison. No large discrepancy is observed, as shown by the Kolmogorov probability [118] displayed on top of each plot. We also compared the distribution of the NN output between old and new data (see Fig. 5.4 (c) and (d)), independently for the sideband subtracted sample, and for the sidebands.

It is necessary to select the best NN output value to cut on in order to achieve a high degree of signal purity as well as a good signal yield. Selecting a high NN output value as a cutoff will produce a very pure signal, but will restrict the number of signal events available therefore reducing the statistical power of the sample. A threshold which is too low will increase the statistics, but the separation between signal and background will be inefficient. As in Ref. [18], we choose the threshold on the NN output that minimizes the expected average variance on the mixing phase as determined in large samples of statistical trials generated with different choices of true values for β_s and $\Delta\Gamma_s$. The procedure is briefly outlined. Pseudo experiments are generated corresponding to different NN cut values by choosing the associated signal and background numbers for each cut value. The fast Monte Carlo simulation used to generate these pseudo experiments is described in Sect. 5.5. Studies are carried out for three potential true values of β_s; 0.3, 0.5, and the SM value 0.02. For each β_s value, the decay width difference $\Delta\Gamma_s$ is calculated according to Eq. (2.27), and all other variables are taken according to their best fit values from [18]. About 700 pseudo experiments are generated and fit for each case, and the β_s statistical uncertainty is checked at each NN cut value. The error distribution is fitted with a Landau function to determine the most probable value of the statistical error for β_s at each NN cut value for the three input values of β_s, these fitted values are plotted against the NN cut level and shown in Fig. 5.5 (a) for $\beta_s = 0.02$. From this study, it was determined that the optimal NN cut value is 0.2, which minimizes β_s errors by increasing as much as possible the signal yield. It is a looser cut than would be selected by a commonly used optimization procedure that maximize the figure of merit $S/\sqrt{S+B}$ (where S and B are the number of signal and background events, respectively) as represented in Fig. 5.5 (b).

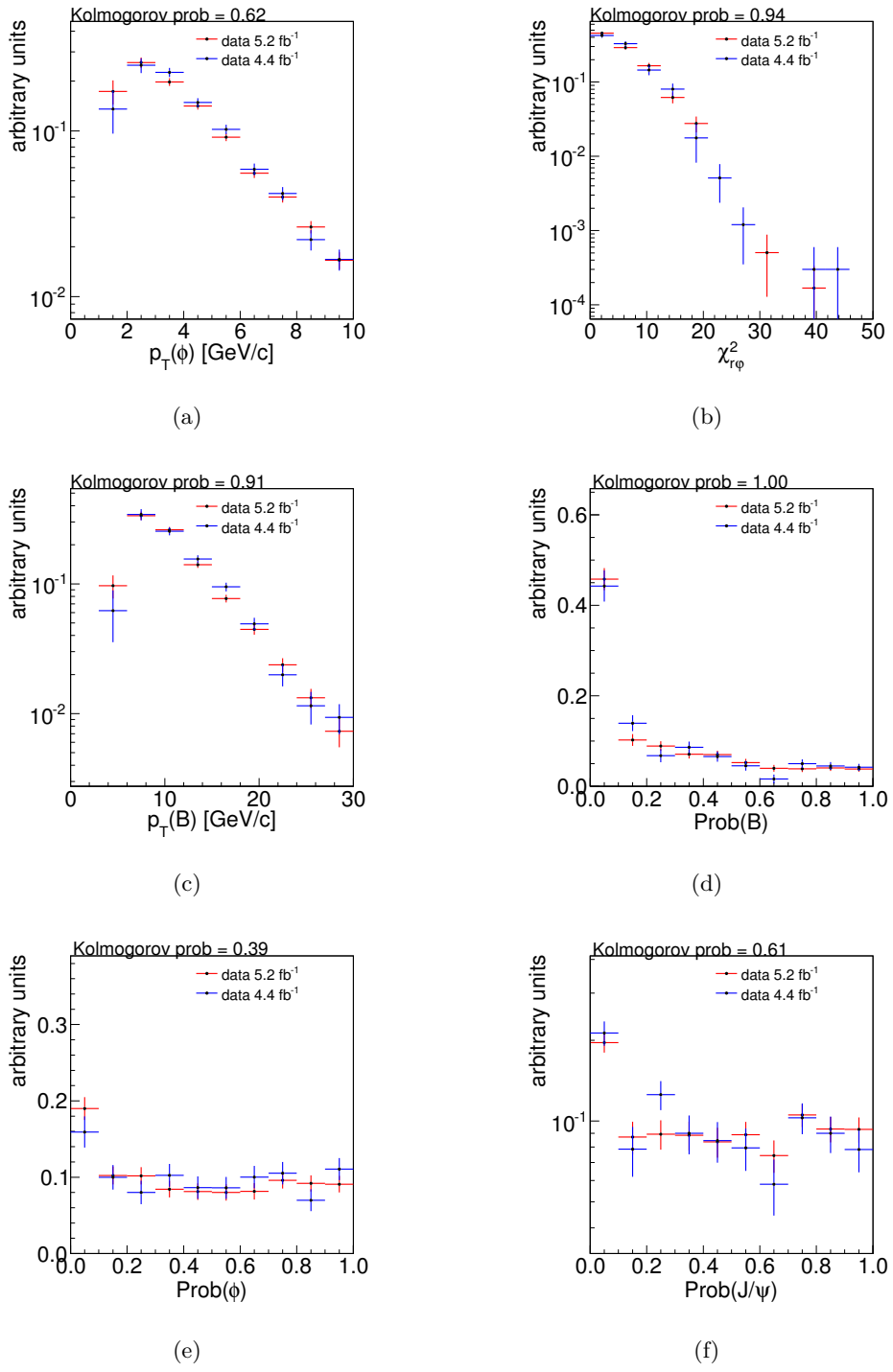


Figure 5.3: Comparison between the distributions of NN input variables between the first 5.2 fb^{-1} (red) and the last 4.4 fb^{-1} (blue) of data. Top to bottom, left to right $p_T(\phi)$ (a), $\chi^2_{r\varphi}$ (b), $p_T(B)$ (c), χ^2 probability of the B_s^0 vertex fit (d), χ^2 probability of the ϕ vertex fit (e), χ^2 probability of the J/ψ vertex fit (f).

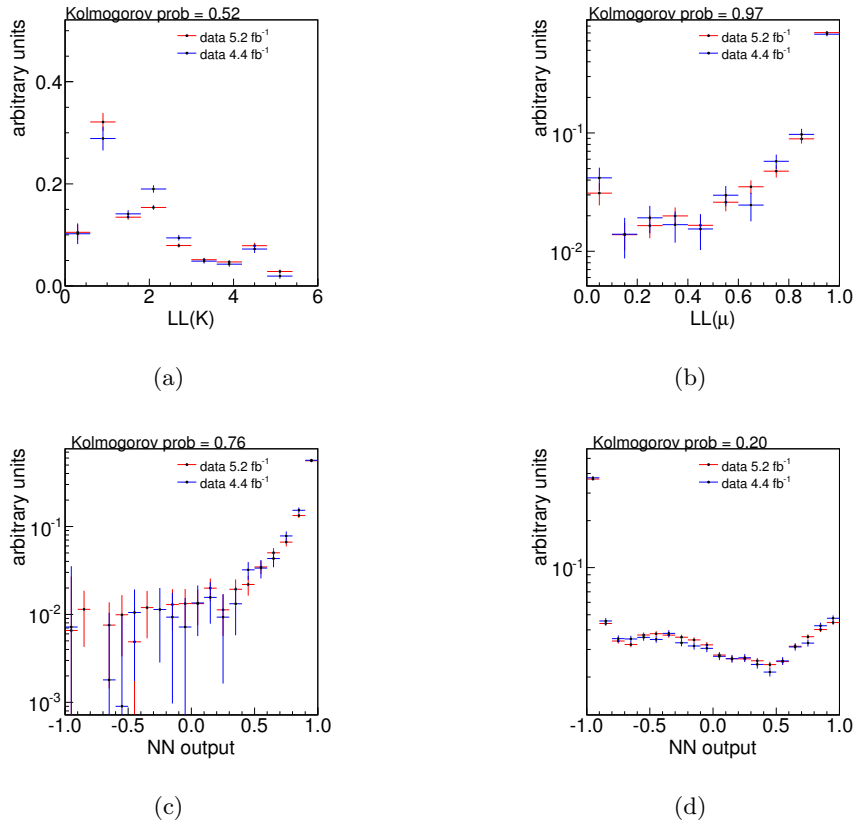


Figure 5.4: Comparison between the distributions of NN input variables between the the first 5.2 fb^{-1} (red) and the last 4.4 fb^{-1} (blue) of data: $LL_K(K)$ (a), $LL_\mu(\mu)$ (b). Comparison between the NN output distributions in the first 5.2 fb^{-1} of data (blue) and the full data set (red). The NN output variable for (sideband-subtracted) signal events only (c) and for sidebands events only (d).

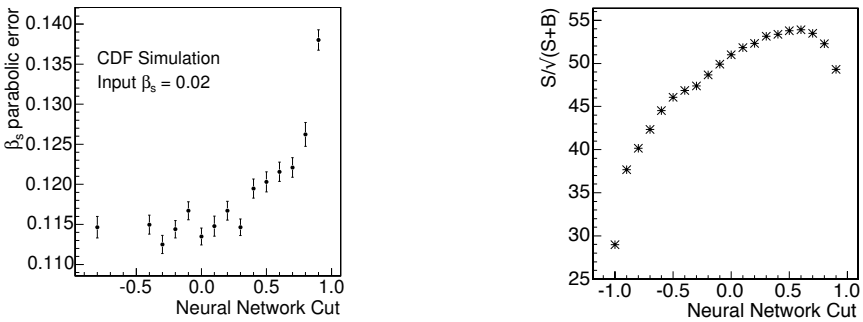
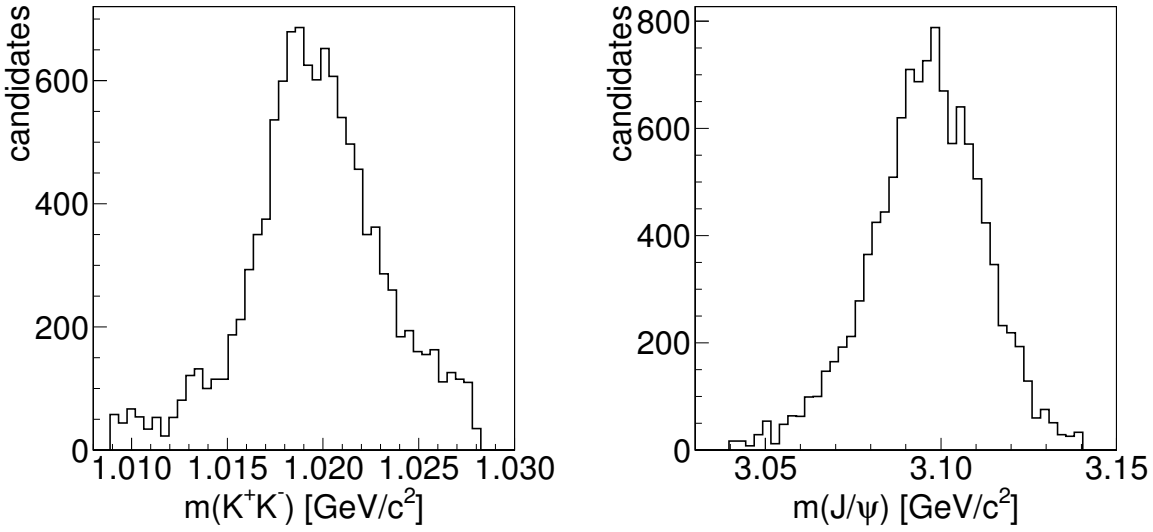


Figure 5.5: Fitted β_s statistical errors versus NN cut value for a true value of $\beta_s = 0.02$ in pseudo experiments, from which a value of 0.2 is chosen for the final sample selection (a). Figure of merit $S/\sqrt{S+B}$ versus NN cut value (b).

5.4 Final sample

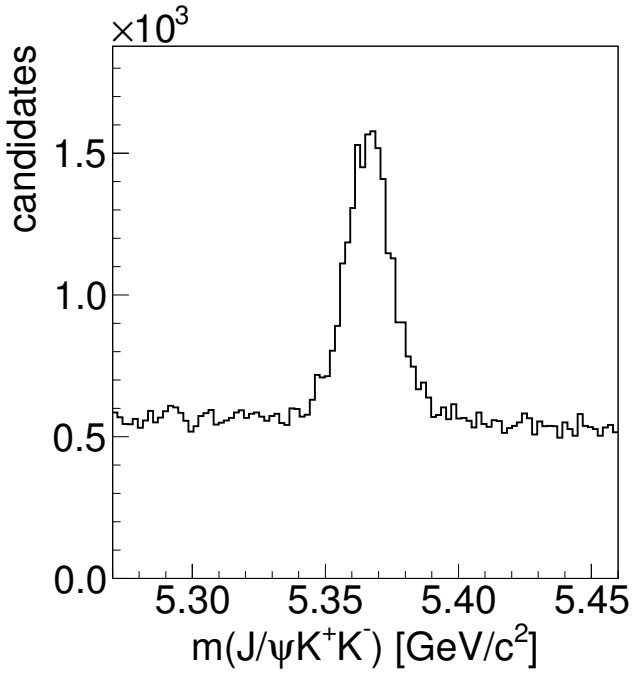
The resulting $J/\psi K^+K^-$ mass distribution is shown in Fig. 5.6 (a), along with sidebands-subtracted distributions of $m(K^+K^-)$ (a) and of $m(J/\psi)$ (b). A prominent signal structure emerges from a smooth background approximately constant in mass in the $m(J/\psi K^+K^-)$ spectrum. This may include contributions from $B_s^0 \to J/\psi\phi$ as well as $B_s^0 \to J/\psi f_0$ decays. The signal can be satisfactorily approximated by a single Gaussian distribution centered at the nominal B_s^0 mass, 8.9 MeV/c^2–wide, containing $10\,950 \pm 110$ events. The background is nearly saturated by the combinatorial component. This is mainly promptly-produced and dominated by accidental combinations of two charged tracks with a real J/ψ decay. Indeed, Fig. 5.7 (b) shows the large suppression of the background when the requirement $ct > 60\,\mu$m is applied, preserving more than 85% of the signal.

Given the limited PID of the CDF detector, a contribution from $B^0 \to J/\psi K^+\pi^-$ decays, where the pion is mis-identified as a kaon (B^0 *cross-feed*) is known to contribute predominantly in the higher-mass side of the signal. In particular, in the limited window of the K^+K^- mass used in the preselection of the data, the main contribution is due to a $K^\star(890)^0 \to K^+\pi^-$ decay mis-recontructed as a $\phi \to K^+K^-$ decay. This contribution is extensively studied in a separate simultaneous fit of the $J/\psi K^+K^-$ and K^+K^- mass distributions described in Sect. 7.4 and found to be approximately 8% of the signal candidates. A systematic uncertainty on the final measurement will be associated to such physics background (see Sect. 7.2).

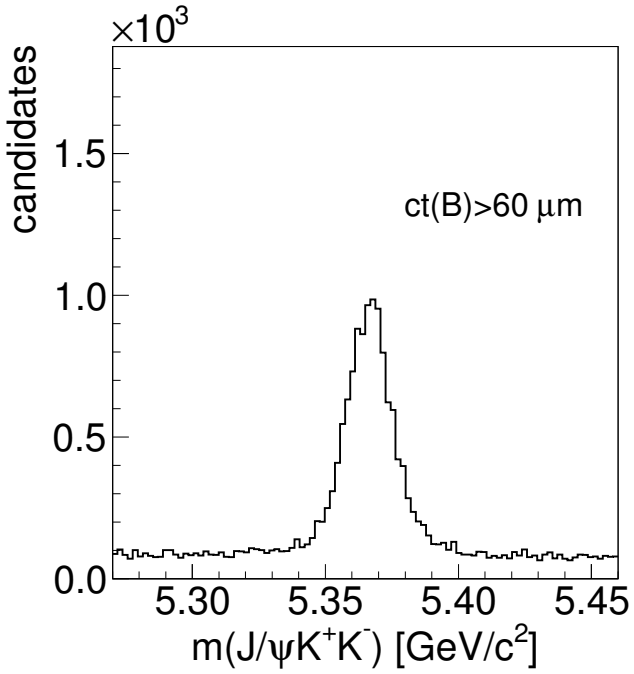


(a) (b)

Figure 5.6: Sidebands-subtracted distributions of $m(K^+K^-)$ (a) and of $m(J/\psi)$ (b) and distribution of $m(J/\psi K^+K^-)$ (a) for the final $B_s^0 \to J/\psi\phi$ data sample after the NN selection. Distribution of $m(J/\psi K^+K^-)$ adding the cut $ct(B) > 60\,\mu$m (b).



(a)



(b)

Figure 5.7: Sidebands-subtracted distributions of $m(K^+K^-)$ (a) and of $m(J/\psi)$ (b) and distribution of $m(J/\psi K^+K^-)$ (a) for the final $B_s^0 \to J/\psi\phi$ data sample after the NN selection. Distribution of $m(J/\psi K^+K^-)$ adding the cut $ct(B) > 60\,\mu$m (b).

5.5 Monte Carlo samples

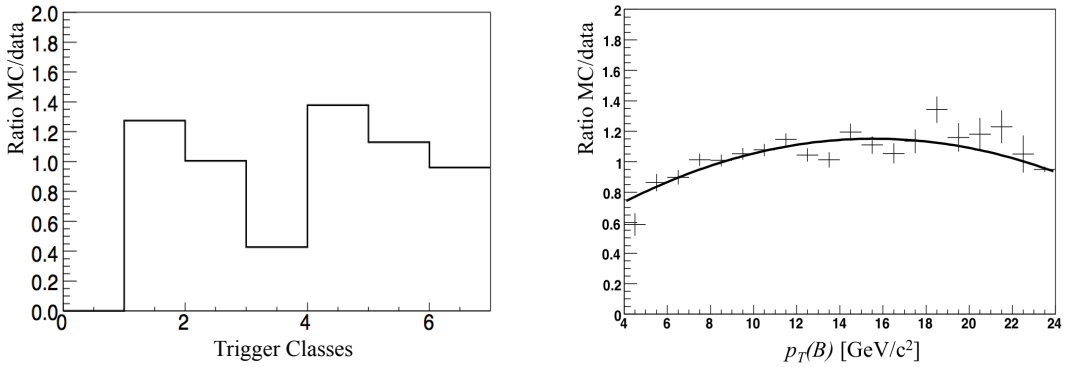
Simulation of the $B_s^0 \to J/\psi\phi$ decay and other relevant decays is employed at several points of the analysis, when an analytic or data-based approach is not feasible. We can distinguish three main techniques, described in the following sections, that provide us with several samples used for different purposes:

- full realistic MC simulation of the B_s^0 production and decay, and of the subsequent detector response. It is used to determine the detector efficiency in the 3-dimensional space of the transversity angles due to the non-hermeticity of the detector. This is described in Sect. 5.5.1.
- Simulation of $B_s^0 \to J/\psi K^+K^-$ samples that includes the full $(P+S)$-waves decay rate and its dependence on the K^+K^- mass, as described Sect. 2.2.1. Such samples are employed to built confidence in the determination of the S-wave fraction as computed by the fit to the transversity angles distributions. It is described in Sect. 5.5.2.
- Generation of samples of $B^0 \to J/\psi K^+\pi^-$ decays that simulate $(P+S)$-wave $K^+\pi^-$ resonances as they are reconstructed in the K^+K^- spectrum with the mis-identification of the pion as a kaon. This simulation are used to extract the templates for the auxiliary fit of the $J/\psi K^+K^-$ and K^+K^- mass distributions for the precise determination of the sample composition. It is described in Sect. 5.5.3.

There is a fourth MC technique used: simplified generation of the $B_s^0 \to J/\psi\phi$ decay's variables called *toys* or *pseudo-experiments*. They are particularly used for studies where a full realistic MC simulation is not necessary, and a simplified simulation of the data is all that is required to test the fit behaviors, such as biases in the estimation of the parameters. The pseudo-experiment events are created by randomly sampling the probability density functions of the maximum likelihood (see Sect. 6.2). The events are generated according to a set of inputs defined by the user for each parameter in the likelihood. The sample of pseudo-experiment events can then be fit in the same way as a sample of data events. When generating pseudo-experiments, we can include background events, parameterized in the same manner as the data. Also, simplified versions of the likelihood can be used to generate pseudo-experiments. For instance, signal-only toys may be used, or toys that do not account for detector efficiencies for some studies of systematic uncertainties.

5.5.1 Phase-space model

In order to study the detector sculpting of the transversity angles distributions, a simulated event sample of $B_s^0 \to J/\psi\phi$ decays obtained through a phase-space model from EVTGEN has been used [102]. Because all spins of the final-state particles are averaged in this model, the angular distributions at generator level are flat and allow one to account for the deviations



(a) (b)

Figure 5.8: Simulation weights. Discrete weights associated to the trigger classes (a). Continuous weight function, as resulting from a fit to the ratio of the $p_T(B)$ distributions of MC and data (b).

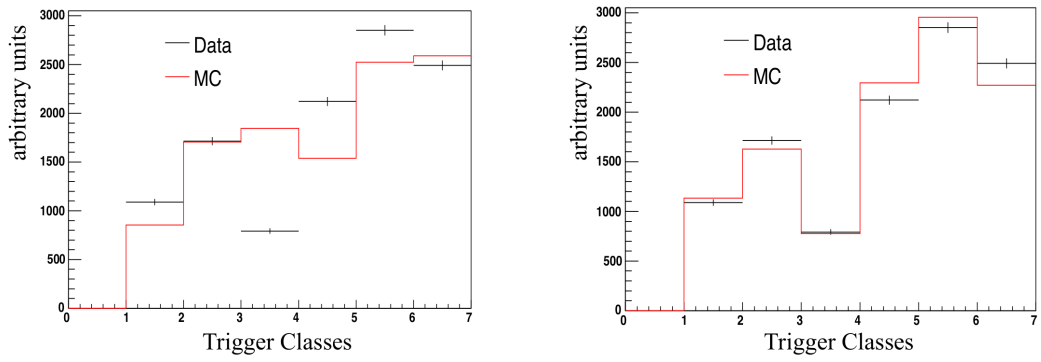
from a uniform acceptance due to instrumental effects only. The simulation follows the three steps described in Sect. 3.2.8, then the events are selected using the same preselection and NN selection as data.

The angular acceptance model depends on the agreement between data and the generated MC in variables that affect the angular decay features of the J/ψ and the ϕ. The MC sample of 100 million events is modelled using input parameters from the first 1 fb^{-1} of CDF data, and result in some inconsistencies with the current dataset which need to be corrected for. A known mismodeling [18] in the $p_T(B)$ spectrum between simulation and data is observed and shown in Fig. 5.9 (a). Because this may affect the distributions of the transversity angles, data-versus-MC compatibility has been investigated in several other variables. Since trigger prescales modify the trigger composition of the sample, and different trigger paths have different p_T thresholds, a prescale-dependent reweighing has been applied. In the following, we compare sideband-subtracted data distributions with simulation distributions, which are reweighed to match data whenever necessary.

The procedure involves three steps: the first takes into account the different triggers mix of our dataset; the second enforces the agreement in the $p_T(B)$ spectrum; and the third step accounts for the combination of both effects.

Trigger path mixture: the candidates are split into two groups, depending on whether a CMU-CMU or a CMU-CMX muon pair has fired the trigger. Then, each of the two classes is split in the following three classes:

- both muons have $p_T > 3\,\text{GeV}/c$;
- both muons have $p_T > 2\,\text{GeV}/c$ and at least one muon has $p_T > 3$ GeV/c;
- all events left.



(a) (b)

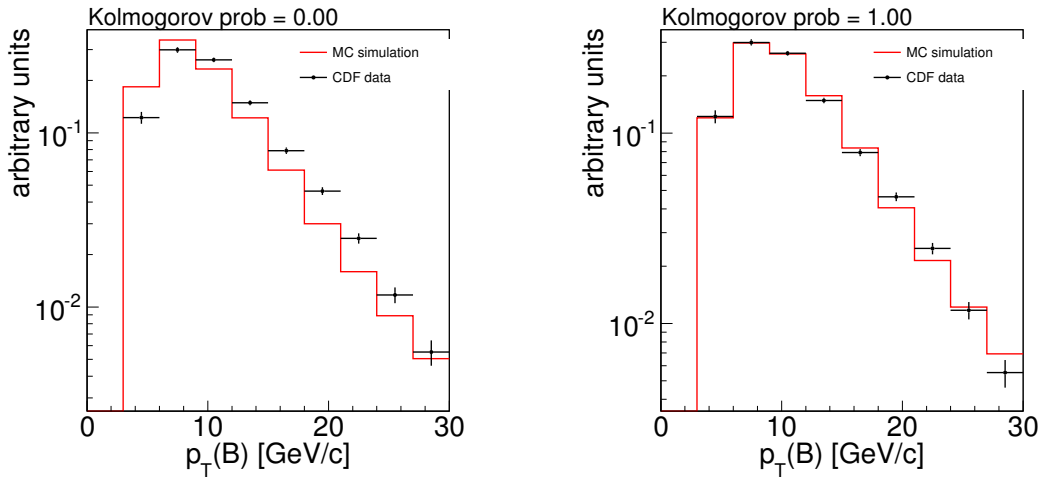
Figure 5.9: Comparison between the data and phase space MC distributions of the trigger classes. The red histograms are MC data, the black points sideband-subtracted data. The MC distribution on the left is not reweighed; the plot on the rights is obtained after the MC reweighing.

We reweigh the simulation to reproduce the fraction of candidates belonging to each of these six classes observed in data. The classes are, as a first approximation, mutually exclusive and such that their union comprises the whole data sample. Figure 5.8 (a) shows the distribution of the resulting weights. The heights of the first three bins represent the weights for the three CMU-CMU trigger classes, the other bins refer to the CMU-CMX triggers.

The $p_T(B)$ distributions in data and simulation is compared after the previous step in the 4–24 GeV/c range. The reweighing function is extracted from the results of a second order polynomial fit to the ratio of the distributions (see Figure 5.8 (b)).

Combined effect of trigger path admixture and $p_T(B)$ distribution: the weight associated to each simulated event is the product of the weight associated to the trigger class and the weight associated to the $p_T(B)$ distribution (see Fig. 5.9 and 5.10).

A third weight is used for the comparison of simulated events to data and takes into account the appropriate angular dependence of the decay, which purposely was not simulated by EVTGEN. To obtain the weight for the angular decay distributions, we take the untagged decay rate of Eq. (2.27) and integrate it over decay time. We use the SM expectation of β_s and $\Delta\Gamma_s$, and the value measured in Ref. [18] for the polarization amplitudes, strong phases and mean B_s^0 lifetime. In App. D, Fig. D.1 – Fig. D.3 show the good agreement between the data and the simulation for other relevant variables.



(a) (b)

Figure 5.10: Comparison between the data and phase space MC distributions of $p_T(B)$. The red histograms are MC data, the black points sideband-subtracted data. The MC distribution on the left is not reweighed; the plot on the rights is obtained after the MC reweighing.

5.5.2 Simulated $B_s^0 \to J/\psi K^+K^-$ data

We validate the robustness of the analysis by using high-statistics samples generated by a realistic MC that fully simulates the $(P+S)$-waves interference independently of the probability density functions used in the fit. This is a new consistency test aimed at further understanding potential biases in the estimation of the physical parameters of interest and at shedding some light on the inconsistency between the size of the S-wave contributions observed between experiments. Based on angular-information only (no K^+K^- mass is ever used in fits) D0 claims a 17% fraction [19], in contrast to the findings from CDF, LHCb, and ATLAS in the 1–2% range [18, 20, 21]. An improperly-modeled component of misidentified S-wave $B^0 \to J/\psi K^+\pi^-$ decays could mimic a KK S-wave and potentially explain the discrepancies, as will be discussd in Sect. 7.4

Since EvtGen does not provide a model that simulates $B_s^0 \to J/\psi K^+K^-$ decays with all the resonances substructure, we wrote an independent MC generator that fully simulates the $B_s^0 \to J/\psi K^+K^-$ decay amplitude as a function of the transversity angles, the B_s^0 proper decay–time and the K^+K^- mass. The generation proceeds in two steps:

1. we simultaneously draw samplings of the transversity angles, the B_s^0 proper decay–time, and the K^+K^- mass distributions, from the theoretical decay rate of $B_s^0 \to J/\psi K^+K^-$ decays of Eq. (2.21), which comprises both $\phi(1020) \to K^+K^-$ and $f_0(980) \to K^+K^-$ components, along with their interference. We generate an equal amount of initially produced B_s^0 and $\bar{B}_s^0$ mesons.

2. We introduce detector effects by simulating the smearing of ct, using as a model the sum of two gaussians whose parameter are taken from the fit to data (see Sect. 6.2.3). The effect of the detector angular acceptance is simulated using the sampling-rejection method [118] with a 3-dimensional histogram of the angular efficiency derived with the phase-space MC described above. The experimental tagging is simulated at this stage with realistic performances as measured in data (see Chap. 4).

We use the relativistic Breit-Wigner distribution for the $\phi(1020)$ and the Flatté shape for the $f_0(980)$, as described in App. B. The Breit-Wigner parameters are fixed to the world averaged values [8]. The parameters of the Flatté, from a BES measurement [119], are the f_0 mass (965 ± 10 MeV/c^2) and the coupling constants to π and to K decays ($g_\pi = 165 \pm 18$ and $g_K/g_\pi = 4.21 \pm 0.33$). Figure 5.5.2 shows the generated K^+K^- spectra for several S-wave fractions from threshold up to 1.1 GeV/c^2. Examples of the generated transversity angles, $ct(B)$ and K^+K^- distributions are shown in Fig. 5.12. They compare well to sidebands-subtracted data when the generation parameters match the values found in data.

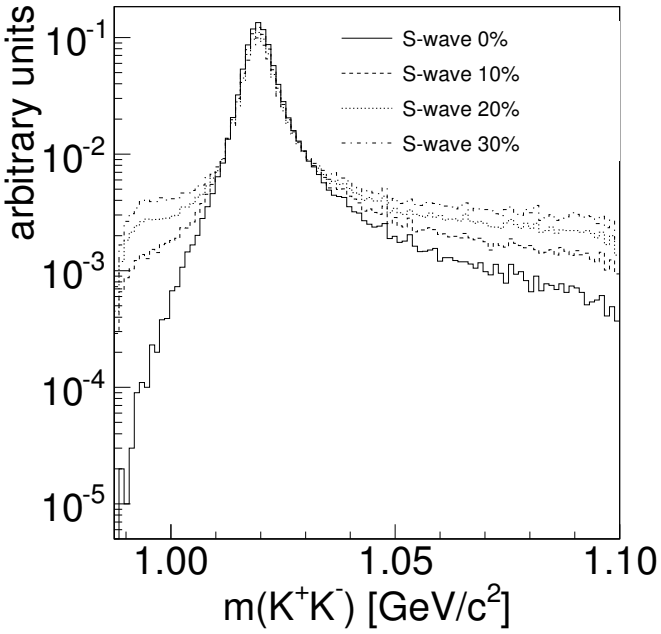


Figure 5.11: Example of K^+K^- spectrum from $B_s^0 \to J/\psi K^+K^-$ simulation for different fractions of S-wave in [0.988,1.100] GeV/c^2.

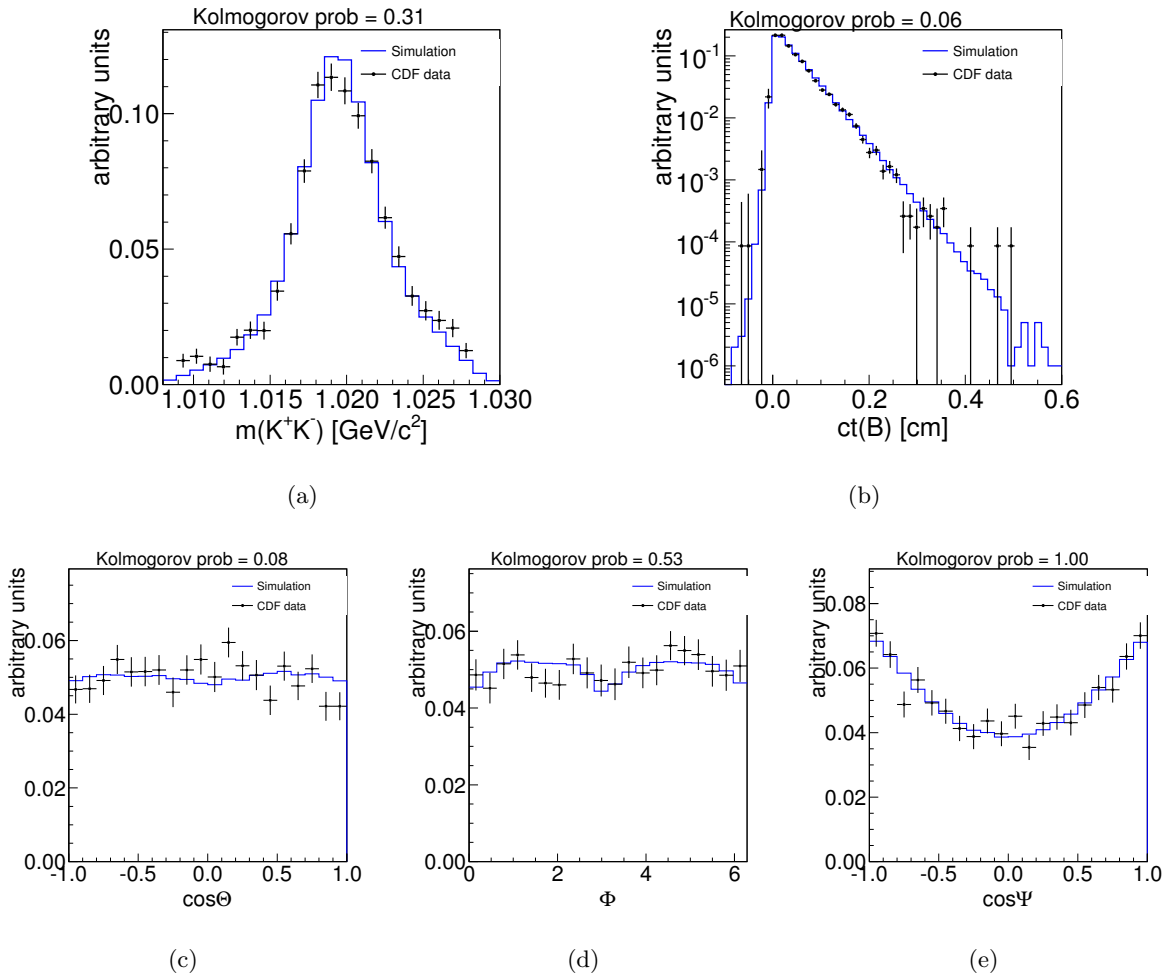
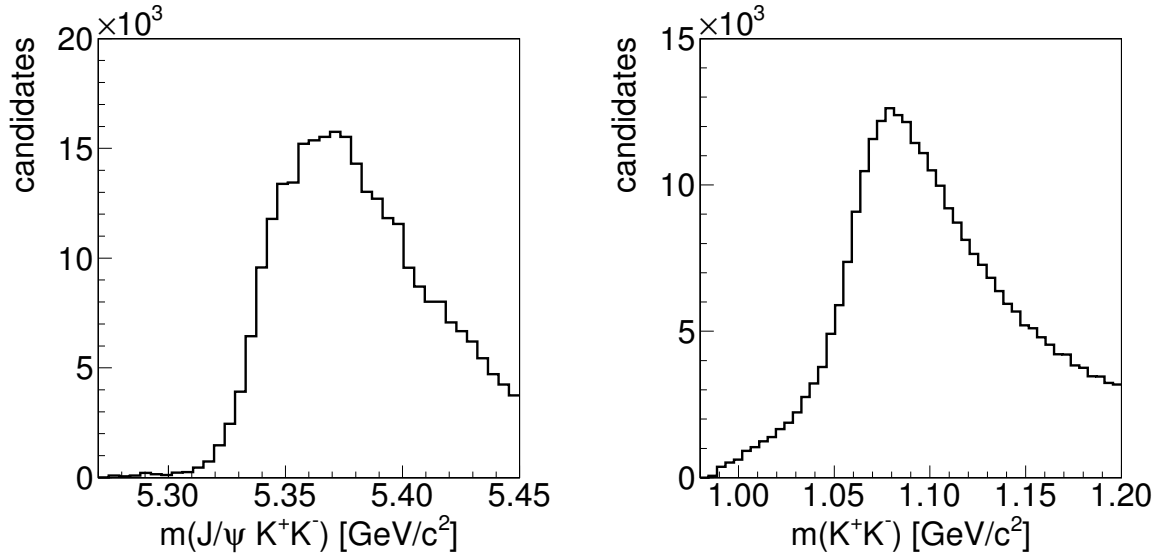


Figure 5.12: Comparison between the data and the $B_s^0 \to J/\psi K^+K^-$ decays simulation. The blue histograms are simulated data, the black points sideband-subtracted data. The simulated data are generated with value of polarization amplitudes, strong phases, $\Delta\Gamma_s$, β_s and Δm_s as measured in Ref. [18]; Top to bottom, left to right: the K^+K^- spectrum generated with 1% of S-wave fraction in $[1.009, 1.028]\,\mathrm{GeV}/c^2$ (a); the $ct(B)$ (b); the transversity angles $\cos\Theta$ (c), Φ (d), and $\cos\Psi$ (e). Example of K^+K^- spectrum from $B_s^0 \to J/\psi K^+K^-$ simulation for different fractions of S-wave in $[0.988,1.100]\,\mathrm{GeV}/c^2$.

5.5.3 Simulated $B^0 \to J/\psi K^+\pi^-$ data

A fraction of $B^0 \to J/\psi K^+\pi^-$ is entering our sample because of decays candidates where the pion is misreconstructed as a kaon. We need a simulation of $B^0 \to J/\psi K^+\pi^-$ decays reconstructed as $B_s^0 \to J/\psi K^+K^-$ decays for extracting the templates to use in the fit for a precise estimation of the sample composition.

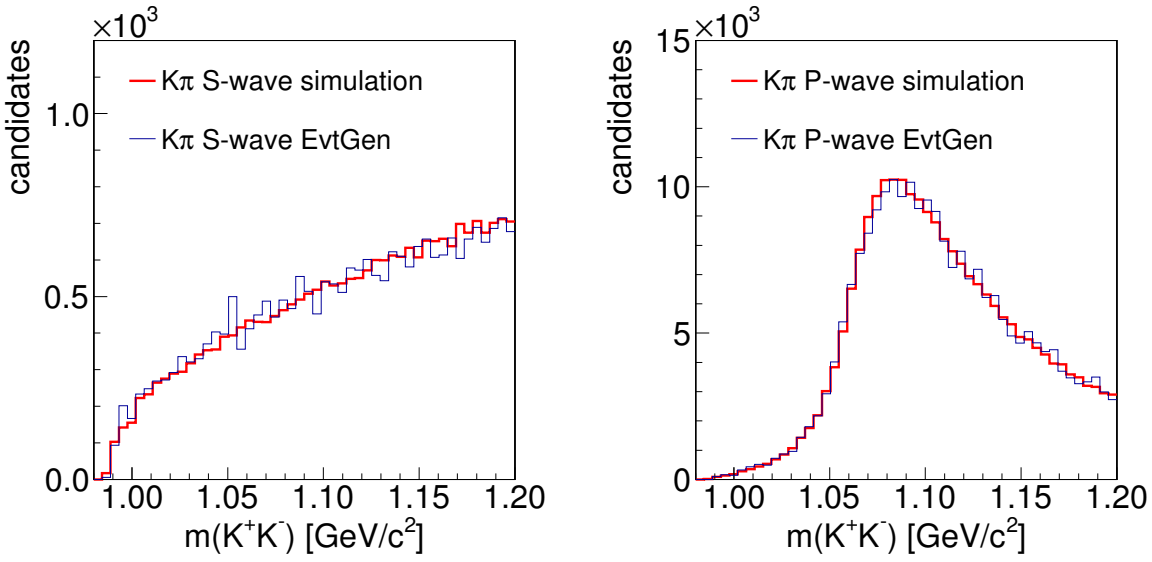
The EVTGEN package [102] provides a model for the generation of the $K^+\pi^-$ P-wave resonance, the $B^0 \to J/\psi K^\star(890)^0$ decay, and the $K^+\pi^-$ phase space model of the $B^0 \to J/\psi K^+\pi^-$ decay, which reasonable reproduces the S-wave shape in the low $K^+\pi^-$ spectrum, but it cannot simulate the full resonance structures with their interference. Therefore, the $K\pi$ component is modeled from a custom simulation, where the S- and P-wave are simulated along with their interference as measured by *BABAR* [120]. Using the decay rate reported in Ref. [120], we generate a sample of $B^0 \to J/\psi K^+\pi^-$ decays with a $K^+\pi^-$ fraction fixed to 7.3% in $[0.8, 1.0]\,\text{GeV}/c^2$ Ref. [120]. We then reassign momenta and masses of the generated $K^+\pi^-$ particles as they were reconstructed like a K^+K^- pair. We show the resulting distributions of $m(J/\psi K^+K^-)$ and $m(K^+K^-)$ in Fig. 5.13. The P-wave and S-wave simulations have been compared separately



(a) (b)

Figure 5.13: Distributions of $m(J/\psi K^+K^-)$ (a) and $m(K^+K^-)$ (b) for simulated $B^0 \to J/\psi K^+\pi^-$ decay reconstructed as $B_s^0 \to J/\psi K^+K^-$.

with the CDF simulation, which is based on EVTGEN generator of the $B^0 \to J/\psi K^\star(890)^0$ and $B^0 \to J/\psi K^+\pi^-$ phase-space decays, respectively, reconstructed as $B_s^0 \to J/\psi K^+K^-$ decays. The EVTGEN-based simulation comprises the simulation of the CDF detector, reconstruction and NN selection. The comparison is reported in Fig. 5.14 and shows good agreement.



(a) (b)

Figure 5.14: Comparison of the custom simulation of $B^0 \to J/\psi K^+\pi^-$ decays reconstructed as $B_s^0 \to J/\psi K^+K^-$ (red) and the EvtGen-based simulation (blue) separately for the S-wave (a) and the P-wave (b).

The S- and P- interference cancels out when integrating in the transversity angles for a symmetric detector acceptance over the transversity angles space, because the angular functions of the interference terms of the decay rate are anti-symmetric with respect to the origin of the transversity angles space. Given the form of the acceptance for $B^0 \to J/\psi K^+\pi^-$ decays, the sculpting of the angular distributions is asymmetric as displayed in Fig. 5.15, the mass spectrum is altered, giving an enhancement in the low-mass region and a suppression in the high-mass range, as shown in Fig. 5.16 (a). The $B^0 \to J/\psi K^+\pi^-$ simulation of the full $S + P$ decay rate has been validated with CDF data by comparing the efficiency-unfolded angular distributions of simulated events and data and finding good agreement (Fig. 5.16 (b)). The only external constraint introduced is the ratio between S- and P-wave $K\pi$ contributions.

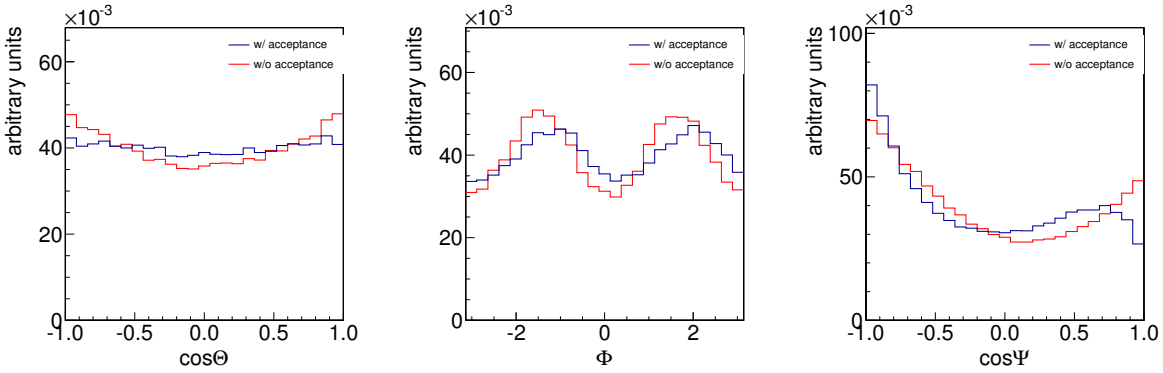
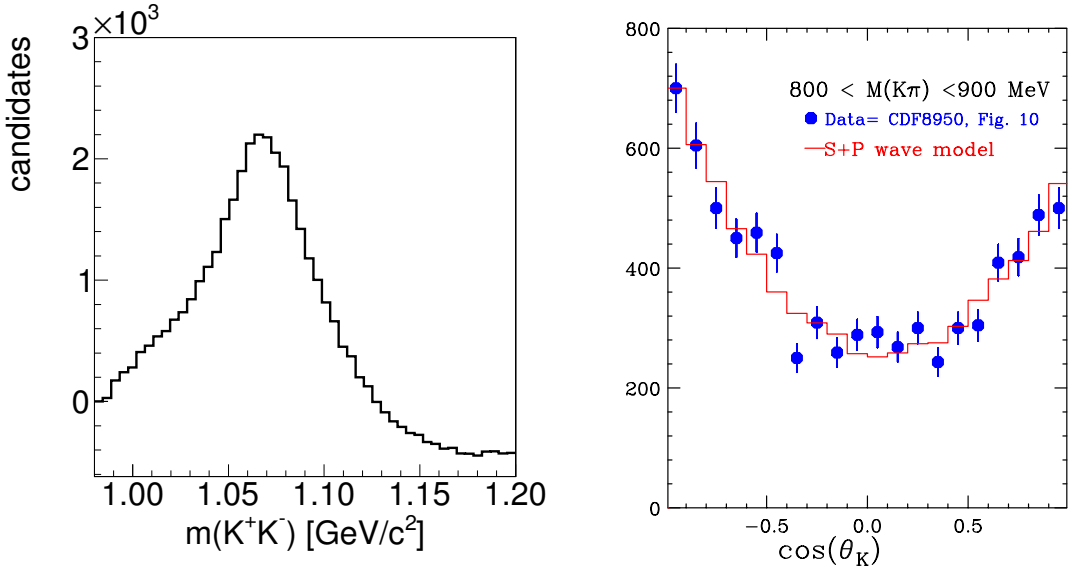


Figure 5.15: Transversity angles distribution for simulated $B^0 \to J/\psi K^+\pi^-$ decays. The red histograms represent distributions where the data are simulated without the effect of angular acceptance. The blue histograms includes the effect of the angular acceptance.



(a) (b)

Figure 5.16: Interference between $K^+\pi^-$ S- and P-wave reconstructed in the K^+K^- spectrum (a). Comparison between the $\cos\Theta$ distribution between MC and CDF data for $B^0 \to J/\psi K^+\pi^-$ decays (b); data are background-subtracted and efficiency-unfolded

Chapter 6

Maximum Likelihood Fit of the $B_s^0 \to J/\psi\phi$ analysis

This chapter describes the multivariate likelihood function used to extract the physical parameters of interest, in particular β_s and $\Delta\Gamma_s$. Properly normalized probability density functions for signal and background are constructed in each of the observed variables. These are combined to make one multivariate probability density function from which a likelihood is constructed. The likelihood has degenerate minima corresponding to symmetries in its parameterization which must be interpreted appropriately; the investigation of its principal features along with its validation is reported.

6.1 Maximum Likelihood method

An unbinned maximum likelihood (ML) fit is used to extract the values of the interesting physics parameters describing $B_s^0 \to J/\psi\phi$ decays. In what follows we briefly remind the ML method. The parameters estimation is obtained through the maximization of the likelihood function with respect to a set of parameters which are assumed to describe the data. Suppose we have a set of n measured quantities $\bar{\boldsymbol{x}} = (\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_n)$ of the event-variables' vector $\boldsymbol{x} = (x^1, \ldots, x^N)$ described by a joint probability density function (PDF) $P(\boldsymbol{x}|\boldsymbol{\zeta})$, where $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_k)$ is a set of k parameters whose values are unknown. The likelihood function is given by the PDF evaluated with $\bar{\boldsymbol{x}}$, but expressed as a function of the k parameters, *i. e.* $\mathcal{L}(\boldsymbol{\zeta}) = P(\bar{\boldsymbol{x}}|\boldsymbol{\zeta})$. If the measurements $\bar{\boldsymbol{x}}_i$ are statistically independent and each follow the PDF $P(\boldsymbol{x}|\boldsymbol{\zeta})$, then the joint PDF for $\boldsymbol{x}$ factorizes and the likelihood function is

$$\mathcal{L}(\boldsymbol{\zeta}) = \prod_{i=1}^{n} P(\bar{\boldsymbol{x}}_i|\boldsymbol{\zeta}). \tag{6.1}$$

The ML method takes the estimators $\hat{\boldsymbol{\zeta}}$ to be those values of $\boldsymbol{\zeta}$ that maximize $\mathcal{L}(\boldsymbol{\zeta})$. It is usually easier to work with $\ln\mathcal{L}$; both $\mathcal{L}$ and $\ln\mathcal{L}$ are maximized for the same parameter values $\hat{\boldsymbol{\zeta}}$. The ML estimators can be found by solving the likelihood equations:

$$\frac{\partial \ln\mathcal{L}}{\partial\zeta_i} = 0, \;\; i = 1, \ldots, k. \tag{6.2}$$

ML estimators are approximately unbiased and efficient for large data samples [8], under quite general conditions, and the method has a wide range of applicability. The inverse V^{-1} of the

covariance matrix $V_{ij} = \mathrm{cov}[\hat{\zeta}_\mathrm{i}, \hat{\zeta}_\mathrm{j}]$ for a set of ML estimators can be estimated by using

$$\hat{V}_{ij}^{-1} = -\frac{\partial^2 \ln \mathcal{L}}{\partial\zeta_i \partial\zeta_j}\bigg|_{\hat{\xi}}. \tag{6.3}$$

For most of the fits the solution of Eq. (6.2) is analytically impossible to find. Thus, a numerical method must be used. The software package MINUIT is used on that purpose, through the minimization of the function $-2\log\mathcal{L}$ [121].

Once a fit is performed, *i. e.* once the estimators $\hat{\boldsymbol{\zeta}}$ are found, we would quantify the goodness of the results obtained. There is no direct method for testing the goodness-of-fit of an unbinned ML fit. Different approaches have been proposed, none of them rigorously correct. A practical method is comparing the distributions of data with the joint PDF evaluated with the set of parameters $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}}$. We can define the *fit projection* onto the variable x^i of the variables vector $\boldsymbol{x}$ as the following one-dimensional function:

$$\mathscr{P}(x^i|\hat{\boldsymbol{\zeta}}) = \int P(\boldsymbol{x}|\hat{\boldsymbol{\zeta}})\,\mathrm{d}x^1 \ldots \mathrm{d}x^{i-1}\mathrm{d}x^{i+1} \ldots \mathrm{d}x^N \tag{6.4}$$

which is the predicted distribution for x^i under the assumed values for the fit parameters, and it can be overlaid to the experimental data. This allows us to detect possible discrepancies between the observed distributions and our model. Then, we perform the comparison of the fit projection onto x^i with its data distribution making a χ^2 test [118]. The test returns a rough evaluation of the goodness-of-fit, since it doesn't take into account the correlations among the variables: it has to be considered only as a qualitative indicator.

6.2 The Probability Distribution Function

In Sect. 5.4 we have distinguished two main components of the $B_s^0 \to J/\psi\phi$ data sample: the events which come from the decay under study, the signal, and the events which are not related to it, the background. Accordingly, the PDF of an event is the sum of two components, P_s and P_b, which describes the signal and the background distributions respectively:

$$P(\boldsymbol{x}|\boldsymbol{\zeta}) = f_s P_s(\boldsymbol{x}|\boldsymbol{\zeta}_s) + (1 - f_s) P_b(\boldsymbol{x}|\boldsymbol{\zeta}_b), \tag{6.5}$$

where f_s is the fraction of signal events ($0 \leq f_s \leq 1$), and we have distinguished the parameters of the signal, $\boldsymbol{\zeta}_s$, and the ones of the background, $\boldsymbol{\zeta}_b$. Each PDF is normalized to unit integral and it is decomposed in the products of PDFs when it is appropriate to treat event variables as independent.

We consider the following event variables: the mass of B candidates and its uncertainty (m and σ_m); the B decay-length and its uncertainty (ct and σ_{ct}); the flavor tag with its predicted dilution (ξ_i and D_i) for each tagging algorithm ($i = \mathrm{OST},\ \mathrm{SSKT}$); and finally the three transversity angles $\Omega = (\cos\Theta, \Phi, \cos\Psi)$. The resolution of the transversity angles have been

proven to be negligible in previous analyses and will not be considered [18]. Since definitions of the decay-length and transversity angles as well as the tagging responses do not depend on the $J/\psi K^+K^-$ mass m and resolution σ_m, the signal PDF for an event factorizes as:[1]

$$P_s(\boldsymbol{x}) = P_s(m, \sigma_m)P(ct, \Omega, \sigma_{ct}, \xi, D), \tag{6.6}$$

where we have defined the PDF of the $J/\psi K^+K^-$ mass and its uncertainty, $P_s(m, \sigma_m)$, that will be written as $P_s(m|\sigma_m)P_s(\sigma_m)$. The PDF for the other variables is decomposed using the condition-probability definition as follows:

$$P(ct, \Omega, \sigma_{ct}, \xi, D) = P(ct, \Omega|\sigma_{ct}, \xi, D)P_s(\sigma_{ct})P_s(\xi)P_s(D), \tag{6.7}$$

given the fact that the decay-length error is independent from the tagging responses. We have introduced: the PDF of the B decay-length uncertainty $P_s(\sigma_{ct})$; the PDF of the flavor tagging variables $P_s(\xi)$ and $P_s(D)$, which combines the OST and SSKT responses; and the condition probability $P(ct, \Omega|\sigma_{ct}, \xi, D)$. The latter is given by the differential decay rate of Eq. (2.27) for $B_s^0 \to J/\psi K^+K^-$ decays, which differs for an initially produced B_s^0 or $\bar{B}_s^0$ meson, hence, we need the dependence on the tagging decisions (ξ, D) in the PDF; the experimental resolution on the decay-length spoils the fast B_s^0-$\bar{B}_s^0$ oscillation, and must be taken into account for considering the actual sensitivity on the mixing terms of the decay rate.

Since the combinatorial background is given by random tracks accidentally satisfying the selection criteria, its PDF fully factorizes as follows:

$$P_b(\boldsymbol{x}) = P_b(m|\sigma_m)P_s(\sigma_m)P_b(ct|\sigma_{ct})P_b(\sigma_{ct})P_b(\Omega)P_b(\xi)P_b(D), \tag{6.8}$$

with obvious notations for the PDF of each variables. The other source of background, the $B^0 \to J/\psi K^+\pi^-$ decays described in Sect. 5.4, is not modeled in the PDF and it will be accounted for in the systematic uncertainties and studied in greater details in Sect. 7.4.

When PDFs have the same dependence on a given variable for the signal and the background, and there is no interest in the determination of its parameters, they can be considered as an overall factor of the total (signal plus background) PDF that can be neglected in the maximization of the likelihood. This is assumed for the mass-resolution PDFs, that are treated as equivalent, $P_s(\sigma_m) = P_b(\sigma_m)$ (we will consider the effects of this assumption in Sect. 7.2). Therefore, the total PDF does not contain them and it finally reads as follows:

$$\begin{aligned} P(\boldsymbol{x}) =& f_s P_s(m|\sigma_m)P(ct, \Omega|\sigma_{ct}, \xi, D)P_s(\sigma_{ct})P_s(\xi)P_s(D) + \\ & (1 - f_s)P_b(m|\sigma_m)P_b(ct|\sigma_{ct})P_b(\sigma_{ct})P_b(\Omega)P_b(\xi)P_b(D). \end{aligned} \tag{6.9}$$

In the following sections we describes the PDF of each variables.

[1] We will omit in the notations the dependence on the parameters $\boldsymbol{\zeta}$ for simplicity from now on

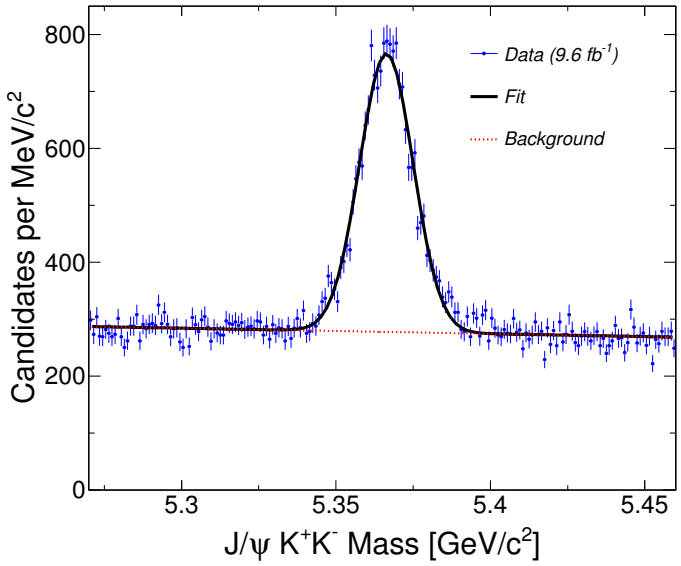


Figure 6.1: Distribution of the $J/\psi K^+K^-$mass with the fit projection overlaid.

6.2.1 PDF of the B mass

Since the B_s^0 peak is well defined in the $J/\psi K^+K^-$ spectrum with a narrow width of about $9\,\text{MeV}/c^2$ atop a flat background distribution, the $J/\psi K^+K^-$ mass of the events is primary useful in order to statistically separate signal candidates from background. The signal mass distribution $P_s(m|\sigma_m)$ is modeled by a single Gaussian function with central value M, fixed to the B_s^0 world average value [8], and a event-by-event width given by the mass resolution (σ_m), scaled using a global scale factor (s_m) to account for a general mis-estimation on the mass uncertainties. The PDF is then constructed by normalizing the Gaussian over the mass window determined in the selection, $m_{\text{min}} = 5.27\,\text{GeV}/c^2$ and $m_{\text{max}} = 5.46\,\text{GeV}/c^2$; it reads:

$$P_s(m) = \frac{1}{N} e^{-\frac{1}{2}\left(\frac{m-M}{s_m\sigma}\right)^2}, \tag{6.10}$$

where N is the normalization factor.

The background mass model is a straight-line function that doesn't depend on the mass resolution:

$$P_b(m) = p_1 m + \frac{1 - \frac{p_1}{2}(m_{\text{max}}^2 - m_{\text{min}}^2)}{m_{\text{max}} - m_{\text{min}}} \tag{6.11}$$

where p_1 is the slope of the line. Figure 6.1 shows the $J/\psi K^+K^-$ spectrum with the fit projection overlaid.

6.2.2 PDF of transversity angles and ct for the signal

The PDF of the signal describing the multidimensional distributions of transversity angles and decay-length of B candidates is not separable, since it is modeled by the differential decay rate in Eq. (2.27),

$$\frac{\mathrm{d}^4\,\overset{(-)}{\Gamma}(\,\overset{(-)}{B}{}^0_s \to J/\psi K^+K^-)}{\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi}, \tag{6.12}$$

properly normalized. The decay rate is different for the decay of a B_s^0 or a $\bar{B}_s^0$ meson and we have to rely on the tagging response to assign the correct PDF. Sculpting of the angular distributions caused by non-hermeticity of the detector and selection criteria must be taken into account as well as the resolution on the measured ct of the event and the limited tagging capability. In Fig. 6.2, we show how the combined effect of the decay-length resolution, of the tagging efficiencies and of the tagging dilutions, spoils the tracing of the theoretical time-evolution of one of the decay amplitude of Eq. (2.27).

The angular acceptance is assumed independent of ct and is modeled by a multiplicative term $A(\Omega)$ representing the angular acceptance, which is parametrized in three dimensions using a set of real spherical harmonics, $Y_{lm}(\Theta,\Phi)$, and Legendre polynomials, $P_k(\cos\Psi)$, as basis functions with ranges $0<\Theta<\pi$, $0<\Psi<\pi$, and $0<\Phi<2\pi$:[2]

$$a^k_{lm} P_k(\cos\Psi) Y_{lm}(\Theta,\Phi). \tag{6.13}$$

The real spherical harmonics are expanded according to the Laplace series:

$$Y_{lm}(\Theta,\Phi) = \sum_{l,m} \left[C_{lm}\cos(m\Phi) + S_{lm}\sin(m\Phi)\right] P_l^m(\cos\Theta), \tag{6.14}$$

where each term is expanded as a function of the Legendre polynomial used to fit $\cos\Psi$:

$$\begin{aligned} S_{lm} &= \sum_k S^k_{lm}\sqrt{\frac{(2k+1)}{2}}, \\ C_{lm} &= \sum_k C^k_{lm}\sqrt{\frac{(2k+1)}{2}}. \end{aligned} \tag{6.15}$$

The parameters a^k_{lm} are obtained from a fit of 100 million simulated events of $B_s^0 \to J/\psi\phi$ decays of the MC described in Sect. 5.5.1, where all transversity angles are generated flat. These MC events, which have been passed through the full CDF detector simulation, enable us to examine how the initially flat distributions are sculpted by the detector acceptance, thus allowing us to determine the angle-dependent efficiencies of the reconstructed particle candidates. The MC distributions of $\cos\Theta$, Φ, and $\cos\Psi$ are filled into a 3-dimensional histogram with 20 bins in each variable; these distributions are shown in Fig. 6.3. The largest effect is an approximate 15% (peak-to-valley) excursion in the angle Φ. Figure 6.4 shows the 2-dimensional projection in (Θ,Φ) of the fit to the 3-dimensional acceptance distribution.

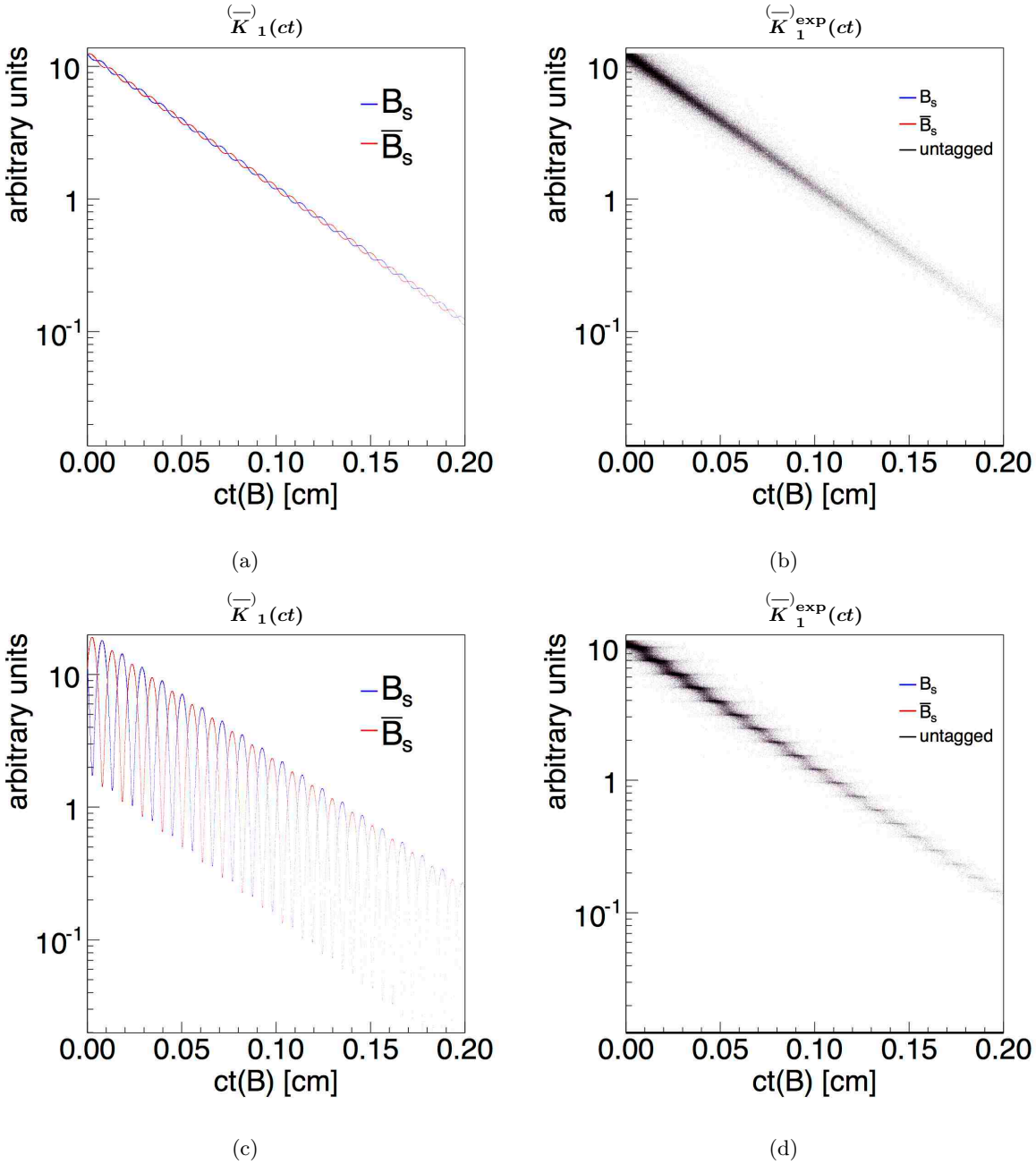
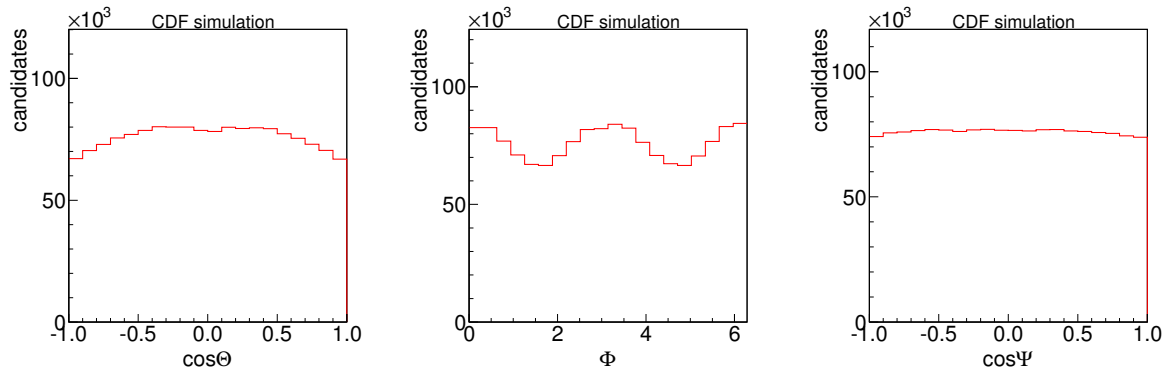


Figure 6.2: Evolution of the amplitude $K_1(ct)$ of the decay rate Eq. (2.27) as a function of the B decay-length: comparison between the ideal case for a perfect tagging and decay-length measurement (left column), and the experimental case when efficiencies, dilutions, and resolutions are introduced as measured in data (right column). In (a) and (b), $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.09\,\mathrm{ps}^{-1}$ (SM point). In (c) and (d), $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.049\,\mathrm{ps}^{-1}$. We can distinguish three main effects: the smearing of the oscillatory curves due to the decay-length resolution; the loss of the distinction between the B_s^0 and $\bar{B}_s^0$ curves due to the limited tagging power; the reduction of the amplitude of the oscillations due to the dilution D (the terms $\sin(\Delta m_s t)$ and $\cos(\Delta m_s t)$ become respectively $D\sin(\Delta m_s t)$ and $D\cos(\Delta m_s t)$).



(a) (b) (c)

Figure 6.3: Observed distribution of the transversity angles $\cos\Theta$ (a), Φ (b), and $\cos\Psi$ (c) for simulated data generated flat in the angular variable space.

The ct resolution caused a smearing of the time-evolution functions in Eq. (2.27). Therefore, the exponential functions describing the decay, $e^{-\Gamma t}\sinh(\Delta\Gamma_s t/2)$ and $e^{-\Gamma t}\cosh(\Delta\Gamma_s t/2)$, as well as the oscillating functions for the B_s^0-$\bar{B}_s^0$ mixing probability, $e^{-\Gamma t}\cos(\Delta m_s t)$ and $e^{-\Gamma t}\sin(\Delta m_s t)$, must be convoluted with the experimental resolution function, $\mathcal{R}(ct|\sigma_{ct})$. The latter is empirically parameterized with a sum of two Gaussians:

$$\mathcal{R}(ct|\sigma_{ct}) = f_{\mathcal{R}}\frac{e^{-\frac{1}{2}\left(\frac{ct}{s_{ct_1}\sigma_{ct}}\right)^2}}{\sqrt{2\pi}s_{ct_1}\sigma_{ct}} + (1-f_{\mathcal{R}})\frac{e^{-\frac{1}{2}\left(\frac{ct}{s_{ct_2}\sigma_{ct}}\right)^2}}{\sqrt{2\pi}s_{ct_2}\sigma_{ct}}. \tag{6.16}$$

whose parameters are floating in the fit and are determined mainly by the prompt ct-background described in Sect. 6.2.3. They are the relative fraction among the Gaussians, $f_{\mathcal{R}}$, and two scale factors, s_{ct_1} and s_{ct_2}, to account for general mis-estimation of the decay-length error. From the distribution of decay-length uncertainties, we find the average of the decay-length resolution function at $\sigma_{ct} \approx 30\,\mu$m, with a root-mean-square deviation of about $12\,\mu$m. Once the analytical form of the resolution function is given, the smeared terms properly normalized replace the time evolution functions in the decay rate. Therefore, the PDF becomes:

$$\overset{(-)}{P}(ct,\Omega|\sigma_{ct}) = A(\Omega)\frac{\mathrm{d}^4\,\overset{(-)}{\Gamma}(\overset{(-)}{B}{}^0_s \to J/\psi K^+K^-)}{\mathrm{d}t\,\mathrm{d}\cos\Theta\,\mathrm{d}\Phi\,\mathrm{d}\cos\Psi} \otimes \mathcal{R}(ct|\sigma_{ct}). \tag{6.17}$$

A compact formalism for the implementation of Eq. (6.17), specifically useful to calculate and incorporate the normalization of the PDF including the experimental effects, was developed in Ref. [56] and it is still adopted here. In such formalism, the polarization amplitudes at $t = 0$ are derived from the estimation of the time-integrated rate, $|a_i|^2$, to each of the polarization

[2]The number of basis functions needed to model the acceptance is determined by the size of the data sample; in our case we take the first 72 coefficients of the expansion.

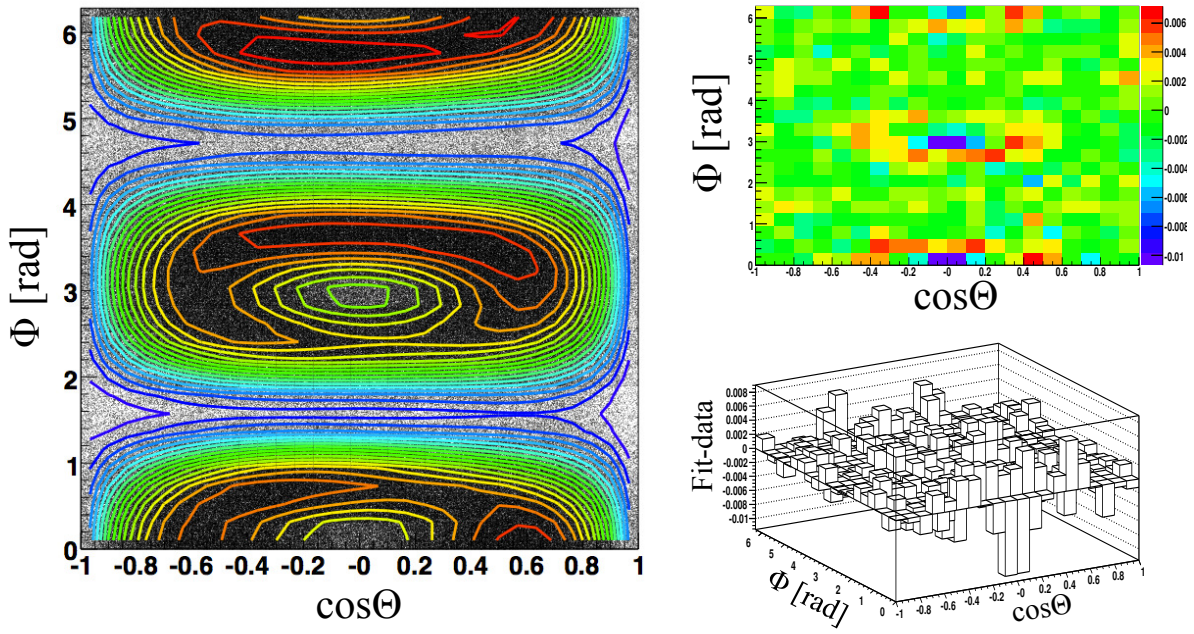


Figure 6.4: Representation of the iso-likelihood lines in the two-dimensional angular space $(\cos\Theta, \Phi)$ as resulting from the fit to simulated events to parametrize the angular acceptance (a). Residuals of the fitted parametrization with respect to the original histogram (b).

states. To translate between the $|a_i|^2$ and the $|A_i|^2$ one can use the following transformations:

$$\begin{aligned} |A_0|^2 &= \frac{|a_0|^2}{1+(y-1)|a_\perp|^2}, \\ |A_\parallel|^2 &= \frac{|a_\parallel|^2}{1+(y-1)|a_\perp|^2}, \\ |A_\perp|^2 &= \frac{y|a_\perp|^2}{1+(y-1)|a_\perp|^2}. \end{aligned} \tag{6.18}$$

where $y \equiv (1-z)/(1+z)$ and $z \equiv (\Delta\Gamma_s/(2\Gamma))\cos 2\beta_s$. Given the unitarity constraint $|a_0|^2 + |a_\parallel|^2 + |a_\perp|^2 = 1$, we can fit the polarization rates by using the fraction:

$$\begin{aligned} \alpha_\perp &\equiv |a_\perp|^2, \\ \alpha_\parallel &\equiv \frac{|a_\parallel|^2}{1-|a_\perp|^2}. \end{aligned} \tag{6.19}$$

and then we automatically have $|a_0|^2 = (1-\alpha_\perp)(1-\alpha_\parallel)$. The inclusion of the potential S-wave contamination from decays of the scalar $f_0(980)$ resonance was also worked out and included in Ref. [56]. We retain the choice of using the likelihood incorporating this S-wave component as the central analysis fit, although the fraction of S-wave is found to be about 1% only (see [18]). The mixing frequency Δm_s entering the decay rate has been Gaussian-constrained in the fit to the measured value [43] and its uncertainty is taken as standard deviation of the Gaussian.

The decay rate for a initially produced $\bar{B}_s^0$ meson is different because of the sign-flip in front of the $\cos(\Delta m_s t)$ and $\sin(\Delta m_s t)$ functions in the decay rate. Therefore, we introduce the flavor tag decision ξ to chose between the two probabilities $P(ct, \Omega|\sigma_{ct})$ and $\bar{P}(ct, \Omega|\sigma_{ct})$. We need to include the responses of the two tagging algorithms, indicated with index 1 for the OST and with index 2 for the SSKT. Each dilution D_i of the tagging decision is multiplied by a scale factor s_i to account for some mis-estimation of the algorithm. These scale factors are separately extracted from the dedicated calibration of both taggers described in Chap. 4. Then, in the likelihood each scale factor is free to float within a Gaussian-constraint which has the results of the calibration and its uncertainty as central value and standard deviation respectively. We use a single scale factor of OST dilutions for both tagging decisions (B or $\bar{B}$) instead of using two separate scale factors. This choice is motived by the result of the updated calibration of the OST which gives the same scale factor for the $B^+ \to J/\psi K^+$ and $B^- \to J/\psi K^-$ decays. We used the OST in the whole dataset, while the SSKT is used for the first $5.2\,\text{fb}^{-1}$ only because of the lack of its calibration in the second part of data (see Sect. 4.3). Finally the PDF that includes all the terms described in this section is:

$$\begin{aligned} P(ct, \Omega|\sigma_{ct}, D, \xi) = &\frac{1+\xi_1 s_1 D_1}{1+|\xi_1|}\frac{1+\xi_2 s_2 D_2}{1+|\xi_2|} P(ct, \Omega|\sigma_{ct}) \\ &+ \frac{1-\xi_1 s_1 D_1}{1+|\xi_1|}\frac{1-\xi_2 s_2 D_2}{1+|\xi_2|} \bar{P}(ct, \Omega|\sigma_{ct}). \end{aligned} \tag{6.20}$$

Since the two taggers search for tagging information (tracks, jets) in complementary regions in space, we treat them as independent tags. It was verified that the tagging decisions and predicted dilutions are indeed independent during their development.

6.2.3 PDF of the decay-length of the background

The background decay-length model is inherited from several CDF measurements of B hadron lifetimes in similar final states [122]. It consists of three main components:

- a prompt peak which accounts for most of the combinatorial background events that are expected to have no significant lifetime. It is modeled by a δ function at $ct = 0$, convoluted with the same resolution function as the signal decay time dependence;

- two exponential functions with positive lifetime (defined for $ct > 0$) used to describe events with misreconstructed vertex and events of the longer lived background, such as other b-hadron decays;

- an exponential with negative lifetime (defined for $ct < 0$) is needed to account for those background events that present a negative decay-length because of a mis-reconstructed secondary vertex.

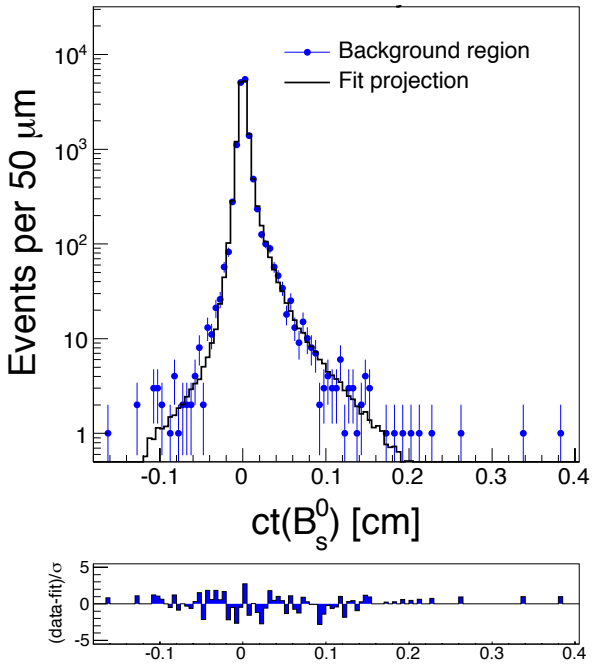


Figure 6.5: Decay-length distribution of sidebands events along with the fit projection overlaid. The χ^2/d.o.f. of the fit is 91/63.

The prompt peak component has a relevant role since allows us to determine the resolution function in ct, Eq. (6.16). The total background decay-length PDF reads:

$$\begin{aligned} P_b(ct|\sigma_{ct}) =& (1-f_{\rm pr})\left[f_{\Lambda_1}\frac{e^{-\Lambda_1(ct)}}{\Lambda_1} + (1-f_{\Lambda_1})\left(f_{\Lambda_n}\frac{e^{\Lambda_n(ct)}}{\Lambda_n} + (1-f_{\Lambda_n})\frac{e^{-\Lambda_2(ct)}}{\Lambda_2}\right)\right] \\ &+ f_{\rm pr}\left[\delta(ct)\otimes\mathcal{R}(ct|\sigma_{ct})\right], \end{aligned} \tag{6.21}$$

where $f_{\rm pr}$ is the fraction of the prompt background; Λ_1, Λ_2, and Λ_n are the inverse of the effective lifetimes of the background events distributed according to the long and short-lived positive exponential as well as the negative exponential, respectively, while f_{Λ_1} and f_{Λ_n} are their corresponding fractions. In Fig. 6.5 we report the ct distribution of background along with fit projection overlaid.

6.2.4 PDF of transversity angles for the background

The background PDF of the transversity angles is parametrized empirically from data of the sidebands of $J/\psi K^+K^-$ mass distribution. Each transversity angle distribution is verified to be largely uncorrelated to the other two angles (see Fig. 6.6). For this reason each angle distribution is modeled separately, and $P(\Omega) = P(\cos\Theta)P(\Phi)P(\cos\Psi)$. Moreover, the background angular distributions are assumed independent from ct. We check that it is a fairly good approximation as shown in Fig. 6.6. We consider $f(\cos\Theta) \propto 1 - a\cos^2(\Theta)$ and $f(\Phi) \propto 1 + b\cos(2\Phi)$ (where a

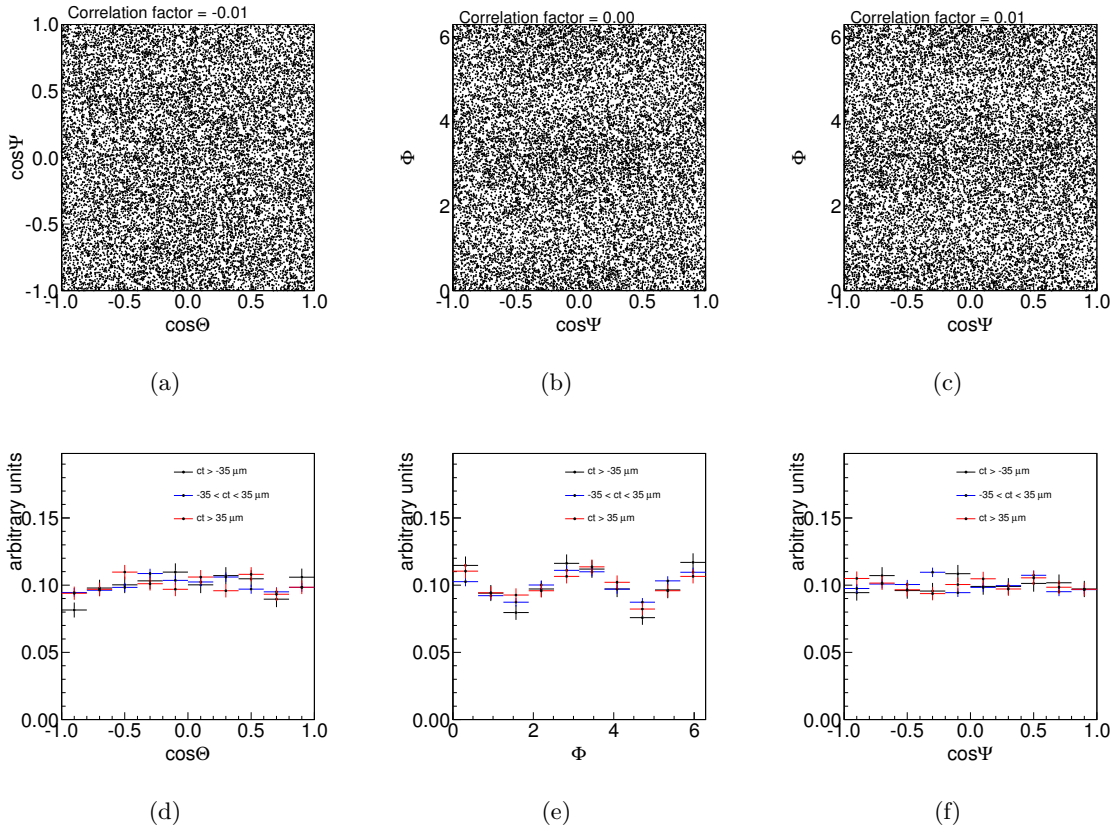


Figure 6.6: Scatter plots of $\cos\Theta$ versus $\cos\Psi$ (a), $\cos\Theta$ versus Φ (b), and $\cos\Psi$ versus Φ (c). Angular distribution of $\cos\Theta$ (d), Φ (e), and $\cos\Psi$ (f), in slice of ct.

and b are fit parameters), while we adopt a flat distribution for $\cos(\Psi)$. Such functions follow closely the shapes of the angular efficiencies, which suggests that the underlying transversity angle distributions of the background events are flat. In Fig. 6.7 the projection of fit to the sidebands events is shown.

6.2.5 PDF of the decay-length uncertainty

As we are using candidate-by-candidate expected decay-time uncertainties, which are not distributed identically for the signal and background events, it is necessary to include a PDF for the separate uncertainty distributions. The PDF has been built using Gamma functions as follows:

$$P(\sigma_{ct}) = f_P \frac{(\sigma_{ct})^{a_1} e^{\frac{\sigma_{ct}}{b_1}}}{b_1^{a_1+1}\Gamma(a_1+1)} + (1-f_P)\frac{(\sigma_{ct})^{a_2} e^{\frac{\sigma_{ct}}{b_2}}}{b_2^{a_2+1}\Gamma(a_2+1)}, \tag{6.22}$$

where a_1, b_1, a_2, b_2 define the mean and the width of respectively the first and the second distribution, and f_P define the fraction of the first distribution. Both the background and the signal PDF has the same form of eq. 6.22 with two different set of parameters. These parameters are found with a preliminary fit to mass and lifetime distributions on the data, and they are then

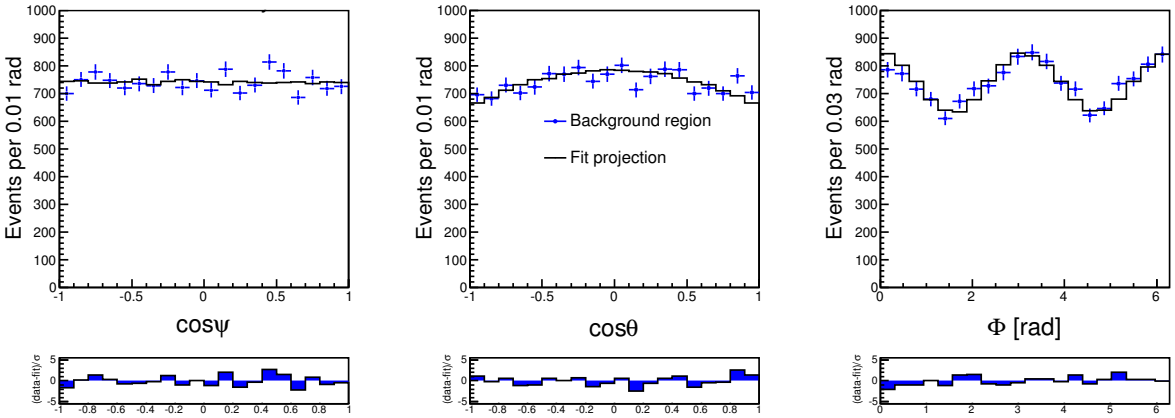
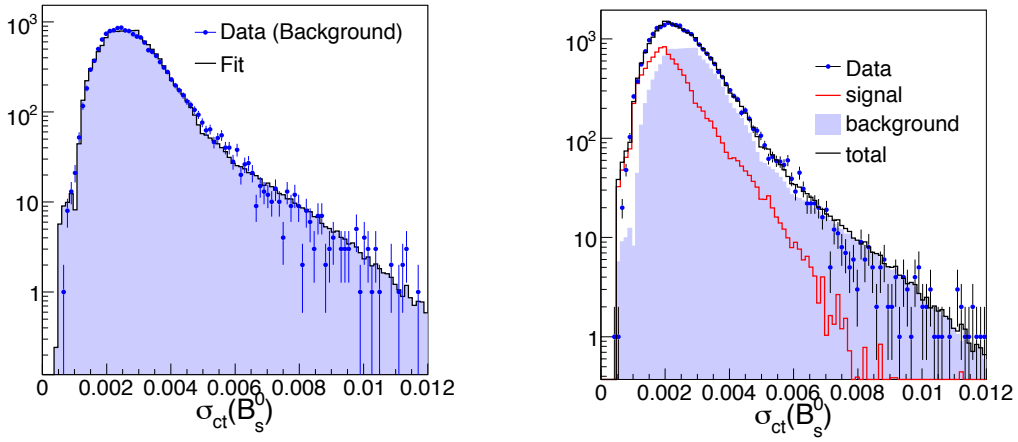


Figure 6.7: Angular distributions of background events with fit projections overlaid. The χ^2/d.o.f. of the fits are 27/19 ($\cos\Psi$), 28/19 ($\cos\Theta$), and 21/19 (Φ).



(a) (b)

Figure 6.8: Distribution of σ_{ct} in background (a) and signal (b) regions with the fit projections overlaid.

fixed in the full likelihood used for the complete analysis. The choice of fixing these parameter in the full fit, based mainly on the technical reason of simplifying the likelihood and reducing the minimization time, has been checked to have limited impact in the extraction of the observables of interest, as it will be discussed in Sect. 7.2. Distributions of decay time uncertainties with fit projections overlaid are shown in fig. 6.8.

6.2.6 PDF of flavor tagging variables

The probability of a particular combined tag decision for the signal is dependent on the efficiency, ϵ_i, of each of the taggers:

$$P_s(\xi) = \begin{cases} (1-\epsilon_1)(1-\epsilon_2), & \text{for } \xi_1 = 0 \text{ and } \xi_2 = 0, \\ \epsilon_1(1-\epsilon_2), & \text{for } \xi_1 = \pm 1 \text{ and } \xi_2 = 0, \\ (1-\epsilon_1)\epsilon_2, & \text{for } \xi_1 = 0 \text{ and } \xi_2 = \pm 1, \\ \epsilon_1\epsilon_2, & \text{for } \xi_1 = \pm 1 \text{ and } \xi_2 = \pm 1, \end{cases} \tag{6.23}$$

Tagging information for background events is essentially without physical meaning, as they are mostly combinations of unrelated tracks which pass the signal selection criteria. Knowledge of the fractions of positively and negatively charged events is sufficient to describe any tagging asymmetry present in the background. Therefore, the PDF $P_b(\xi)$ contains a term to correct for any charge asymmetry in the taggers, a disparity in the fraction of background events tagged as B_s^0 over $\bar{B}_s^0$.

The PDFs of the dilution, $P_i(D_j)$, is modeled with a template that consists of an histogram, taken from the data itself. As the predicted dilution distributions are different for signal and background events, separate histograms are produced for the signal and the background. Different histograms are produced for different taggers. The signal histograms are produced by using background subtracted data; the background dilution histograms are complementary produced with sidebands data. These distributions are reported in Fig. 6.9 and for OST (a) and SSKT (b).

Because the SSKT is used only in the first $5.2\,\text{fb}^{-1}$ of data, we factorize the likelihood as follows:

$$\mathcal{L} = \mathcal{L}_{\text{OST+SSKT}} \times \mathcal{L}_{\text{OST}} \tag{6.24}$$

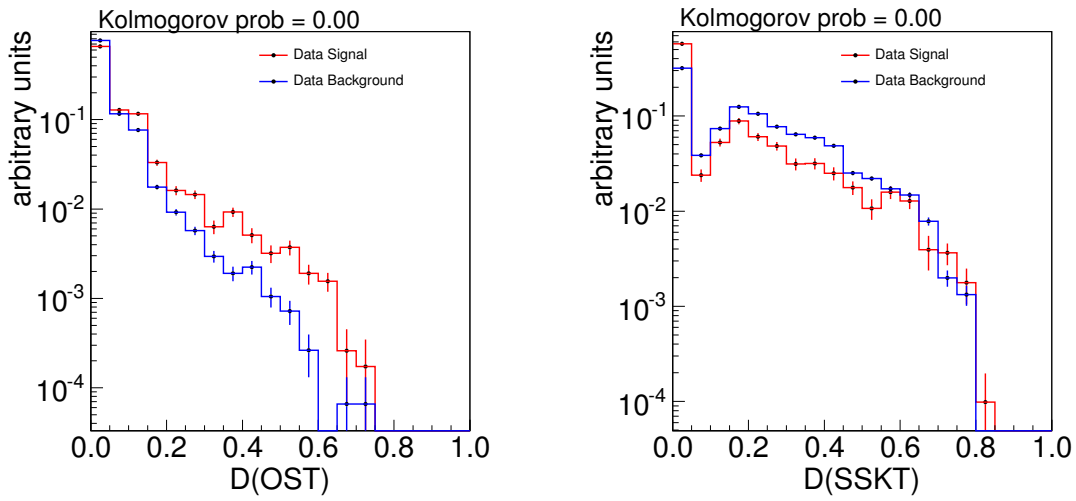
where $\mathcal{L}_{\text{OST+SSKT}}$ refers to the likelihood of the first part of the sample and its PDF's contains both OST and SSKT observables and parameters, while $\mathcal{L}_{\text{OST}}$ refers to the second part of the sample where only the OST information are exploited. In particular, $\mathcal{L}_{\text{OST}}$ has four less fitting parameters with respect to $\mathcal{L}_{\text{OST+SSKT}}$: the SSKT tagging efficiency for signal and background (ϵ_s(SSKT) and ϵ_b(SSKT)), the SSKT background tag asymmetry (A(SSKT)) and the SSKT dilution scale factor S_D(SSKT). Moreover, the SSKT dilution PDFs are not present in $\mathcal{L}_{\text{OST}}$, while the tag decision PDF Eq. (6.23) and the tagged PDF Eq. (6.20) have only the index for the OST case.

6.2.7 Summary of the PDF parameters

Table 6.1 lists all of the fit parameters along with a very short description. The parameters of the lifetime-uncertainty PDF are excluded since they are not floating in the fit, but determined with a preliminary mass and lifetime fit.

Parameter	Description
β_s	*CP*-violating phase of B_s^0-$\bar{B}_s^0$ mixing amplitude
$\Delta\Gamma_s$	lifetime difference $\Gamma_L - \Gamma_H$
$\alpha_\perp$	fraction $\lvert a_\perp\rvert^2$
$\alpha_\parallel$	fraction $\lvert a_\parallel\rvert^2/(1-\lvert a_\perp\rvert^2)$
$\delta_\perp$	$\arg(A_\perp A_0^\star)$
$\delta_\parallel$	$\arg(A_\parallel A_0^\star)$
$c\tau$	B_s^0 mean lifetime
$\lvert A_S\rvert^2$	fraction of S-wave K^+K^- component in the signal
δ_S	strong phase of S-wave amplitude
Δm_s	B_s^0-$\bar{B}_s^0$ mixing frequency
f_s	Signal fraction
s_m	Mass uncertainty scale factor
p_1	mass background slope
s_{ct_1}	lifetime uncertainty scale factor 1
s_{ct_2}	lifetime uncertainty scale factor 2
$f_{\mathcal{R}}$	relative fraction of Gaussians of the decay-length resolution
$f_{\rm pr}$	fraction of prompt background
f_{Λ_n}	fraction of background with Λ_n lifetime
f_{Λ_1}	fraction of background with Λ_1 lifetime
Λ_n	Inverse of lifetime of a background component
Λ_1	Inverse of lifetime of a background component
Λ_2	Inverse of lifetime of a background component
a	parameter in background fit to Φ
b	parameter in background fit to $\cos\Theta$
S_D(OST)	OST dilution scale factor
S_D(SSKT)	SSKT dilution scale factor
ϵ_b(OST)	OST tagging efficiency for background
ϵ_b(SSKT)	SSKT tagging efficiency for background
A(OST)	OST background tag asymmetry
A(SSKT)	SSKT background tag asymmetry
ϵ_s(OST)	OST tagging efficiency for signal
ϵ_s(SSKT)	SSKT tagging efficiency for signal

Table 6.1: Summary of the fitting parameters for: the decay-length and transversity angles PDF of the signal (green); the mass PDF (red); decay-length PDF of the signal (blue); decay-length PDF of the background (orange); angular PDF of the background (yellow); tagging parameters (gray).



(a) (b)

Figure 6.9: Dilution histograms for OST (a) and for SSKT (b). Comparison between sidebands and sidebands-subtracted data.

6.3 Investigation of the ML fit's features

The ML estimators show significant biases that depend on the true values of the parameters, challenging any attempt at bias corrections. This has been tracked to originate from a combination of the complications due to likelihood symmetries, which introduce multiple, equivalent solutions; to the sensitivity to some parameters depending on the estimated values of others; and to the current data size being still insufficient to approximate the asymptotic regime.

A complication is originated by the symmetry of the untagged decay rate discussed in Sect. 2.2.2, when there is no distinction on the flavor of the B meson at production. In that case, transformations of strong phases are independent on the transformations on β_s and $\Delta\Gamma_s$, as they do not need to be simultaneous for the symmetry to hold, with the result of a four-fold ambiguity in the $(\beta_s, \Delta\Gamma_s)$ space. Moreover, an additional complication is due to the fact that the strong phase $\delta_\perp$ appears always and only in a product with $\sin 2\beta_s$ in the untagged likelihood. As a result, in case of small CP violation there is no sensitivity to $\delta_\perp$, and the fit tends to bias the result as by increasing the value of β_s to gain sensitivity on $\delta_\perp$ as an additional parameter available to describe the statistical fluctuations. For this reason, we choose to neither give results by using the untagged likelihood nor to rely on it for a control check of the fit with flavor tagging. In Figure 6.10 we report some comparisons of the tagged and untagged likelihood for the estimation of the mixing phase in simulated data samples as confidence regions in the $(\beta_s, \Delta\Gamma_s)$ plane. The systematic bias towards larger value of the mixing phase when its generation value is close to zero is visible for the untagged likelihood (along with the four fold ambiguity). The figures show also the irregular shapes of both the tagged and untagged

iso-likelihood contours which largely fluctuate sample by sample.

Given the limited experimental tagging power, a sort of residual four-fold ambiguity still affects the likelihood when using the likelihood with the tagging. Because of such approximate symmetries, MINUIT can occasionally find a local minimum in the minimization process. The local minimum is a point of the multidimensional space of $(\beta_s, \Delta\Gamma_s, \delta_\parallel, \delta_\perp, \delta_S)$ with a value of the likelihood function being larger than the value of the likelihood estimated in the quasi-symmetric point $(\beta_s, \Delta\Gamma_s, 2\pi - \delta_\parallel, \pi - \delta_\perp, \pi - \delta_S)$. Furthermore, the PDF in eq. 6.9 is invariant under the simultaneous transformations of all the four parameter $(\beta_s, \Delta\Gamma_s, \delta_\parallel, \delta_\perp, \delta_S)$ in $(\pi - \beta_s, -\Delta\Gamma_s, 2\pi - \delta_\parallel, \pi - \delta_\perp, \pi - \delta_S)$, as described in Sect. 2.2.2. Such a symmetry can be lifted for large values of the fraction of S-wave in the sample, due to the asymmetric shape of the integral I of the relativistic Breit Wigner function in Eq. (2.27) around the ϕ pole. We have numerically checked the discrete symmetries of our likelihood by sampling some points of the space parameters. When dealing with an S-wave fraction like the one expected in data ($\approx 1\%$), with the current event sample size, we can consider the likelihood completely symmetric under the above transformations, since the difference in the likelihood value is not significant when it is evaluated in exactly symmetric points of the parameters space.

The exact symmetry is broken when a parameter affected by the transformation is fixed to a given value. In that case, estimation on remaining parameters are more reliable since the biases are considerably reduced. Therefore, we choose to perform two type of analysis. The first analysis aims at measuring precisely the width difference $\Delta\Gamma_s$, the mean lifetime and the polarization amplitude of $B_s^0 \to J/\psi\phi$ decay, when the mixing phase is set at its SM expectation. Such measurements are done with the maximum likelihood fit described above in which the mixing phase value is fixed at the SM prediction and it is no longer a fit parameter. We refer to this analysis as the *β_s-fixed fit*. The approximate symmetry still may give problem in the minimization process because of the strong phases. We actually found a minimum of the likelihood at the boundary of the symmetry point for $\delta_\parallel$ which prevent us to give a point estimate for such parameter. We check numerically in the fit to data that the minimum found in the likelihood minimization is the global minimum by comparing this value with the likelihood value calculated in its quasi-symmetric point with respect to the strong phases (constraining the value of $\Delta\Gamma_s$ to be positive, as expected in the SM solution of β_s), as per the approximate-symmetry transformation.

The second analysis aims at the estimation of the mixing phase. We call it the *β_s-floating fit*, in which we use the likelihood employed in the β_s-fixed fit and we treat the mixing phase as a fit parameters to be determined in the minimization process. Given the problems presented above, the estimated uncertainties are unlikely to represent actual confidence regions with the desired level of confidence. Thus, we do no use the point estimates (*i. e.*, a central value with uncertainties) to quote results of the parameters of interest. We instead provide confidence intervals for the mixing phase and confidence regions in the $(\beta_s, \Delta\Gamma_s)$ subspace, using the Neyman construction of confidence region, which will be described in Sect. 8.1.

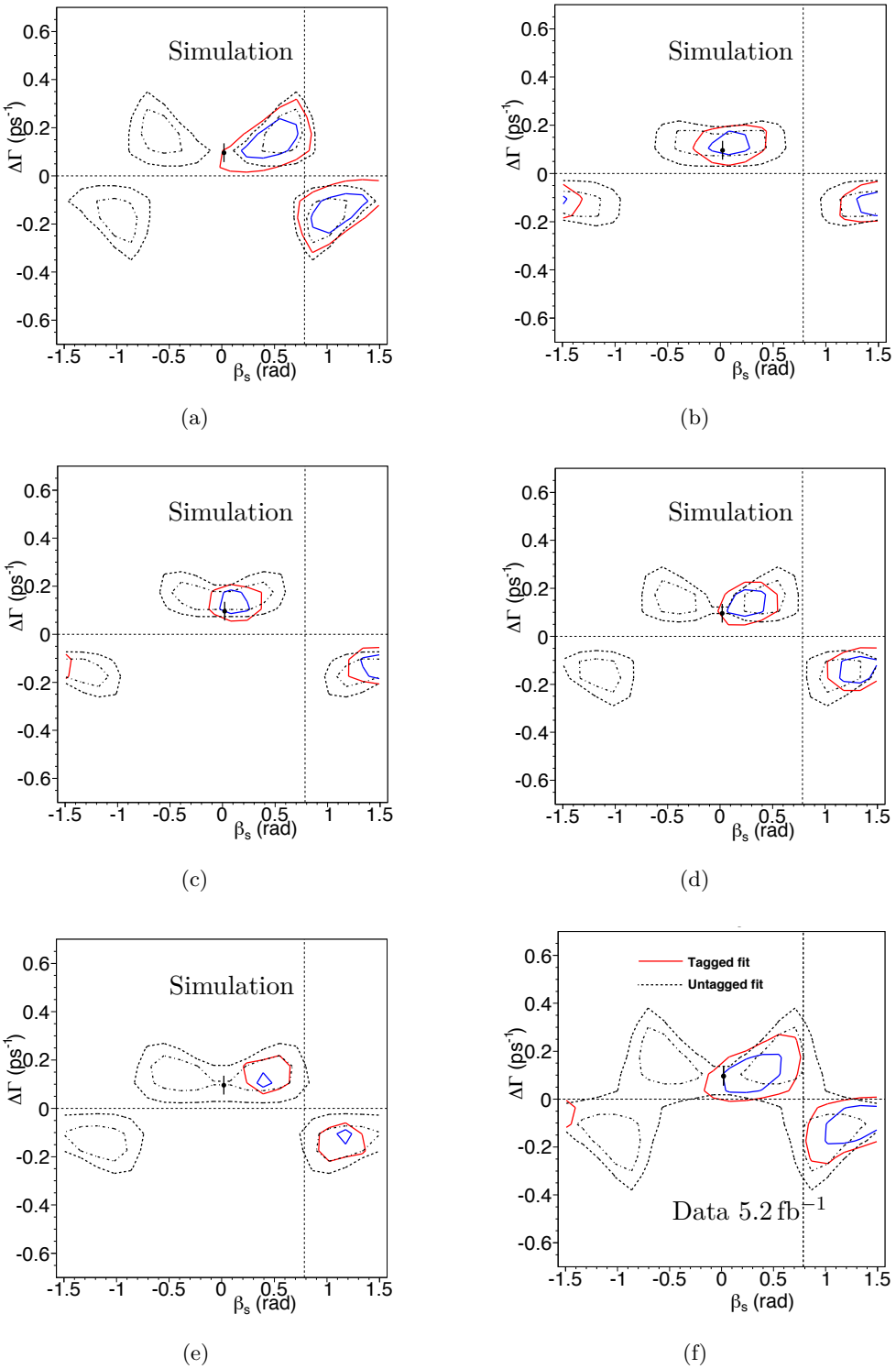


Figure 6.10: Confidence regions for the 68% C.L.(blue, internal curves) and 95% C.L. (red, external curves) in the $(\beta_s, \Delta\Gamma_s)$ plane for 5 pseudo-experiments (a)–(e) [123]. The last plot (f) is done on real data with $5.2\,\text{fb}^{-1}$ as published in Ref. [18]. Solid curves are iso-likelihood contours for analyses with flavor-tagging; dashed curves are for analysis without flavor-tagging. In the generation of pseudo experiments, $\beta_s = 0.22$ and $\Delta\Gamma_s = 0.1\,\text{ps}^{-1}$.

In the following sections, we investigate the features and biases in the parameters estimation given the above problems, both for the β_s-fixed and the β_s-floating fits.

6.3.1 Pull studies

We check the behavior of the likelihood estimator in pseudo-experiments by looking at the distribution of normalized deviations (*pulls*) from the generated value of each parameter. For a variable x which has a Gaussian distribution with mean μ and width σ, the pull

$$g = \frac{x - \mu}{\sigma} \tag{6.25}$$

is expected to be Gaussian by definition. The central limit theorem shows that this principle can in fact be extended to non-Gaussian parameters, such as a lifetime, which would be expected to have a Gaussian distribution of measured values if suitably large datasets are analyzed. In this case, the pull

$$g = \frac{\zeta_m - \zeta_g}{\sigma_\zeta} \tag{6.26}$$

is expected to be a Gaussian of zero mean and unit standard deviation, where ζ_m and ζ_g are the measured and generated (true) values of the parameter being studied, and σ_ζ is the error on the measurement. A deviation from the Gaussian distribution in this case can indicate either a bias in the fitting technique, or a mis-calculation of the errors which enter in the denominator.

We investigate the pulls of the fitted parameters by studying 1000 pseudo-experiments each with the same statistics as in data, *i. e.*, ∼56 000 events (signal plus background) generated by sampling the PDF function as described in Sect. 5.5. In the generation of the samples, we use the value estimated in the fit to real data for all parameters of the PDF.

6.3.1.1 β_s-fixed fit

We first inspect the pull distribution of the parameters of physics interest $\Delta\Gamma_s$, $c\tau$, $\alpha_\parallel$, $\alpha_\perp$, and the strong phases $\delta_\perp$ and $\delta_\parallel$ in the β_s-fixed fit (see Fig. E.1 in App. E). The pulls are regular for all quantities but $\delta_\parallel$ and $|A_S|^2$. The pull of $\alpha_\perp$ presents a bias of the order of ≈ 0.2 standard deviations; being Gaussian-distributed, we will add this bias in the systematic uncertainty of this parameter, as well as for $\alpha_\parallel$, which has a smaller bias, ≈ 0.1 standard deviations. The $\delta_\parallel$ pull shows a non-Gaussian behavior, which makes challenging to quote a reliable point estimate. For estimated values close to the symmetric point π, the minimization process cannot always unambiguously tell apart the two symmetric and close minima, and tends to return the boundary value, π, as estimate (Fig. 6.11).

It is useful to look also at the pull distribution for the S-wave amplitude $|A_S|^2$. The S-wave amplitude parameter $|A_S|^2$ is floating in the minimization within the interval [0,1] to prevent it from assuming negative, unphysical values. The value obtained for the $|A_S|^2$ by the fit to data is previous CDF measurement is $(1.8 \pm 2.3)\%$ [18], *i. e.* very close to the boundary. Being also

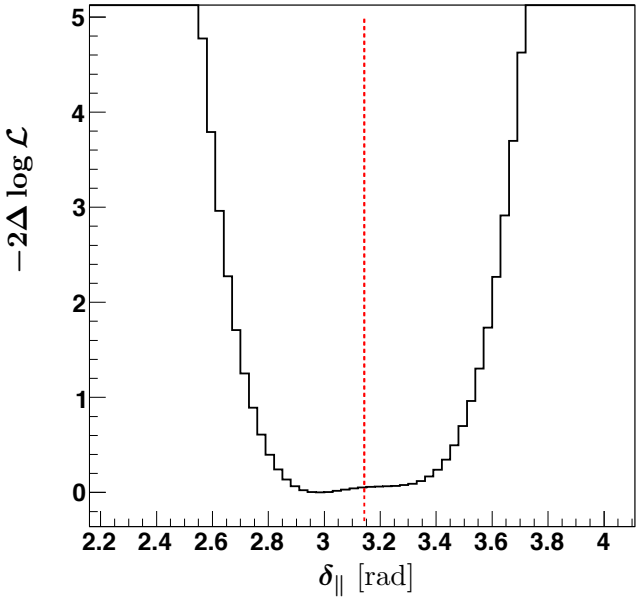


Figure 6.11: Likelihood-ratio profile of $\delta_{\parallel}$ in one pseudo-experiments. The two likelihood minima are close being in proximity of the symmetry point $\delta_{\parallel} = \pi$ (red dotted line). The fit can converge either in a minimum or in the other, and this spreads the error on the parameter.

the value used in the generation of the pseudo-experiments, this explains the odd behavior of the pull distribution for $|A_S|^2$ (Fig. E.1 (h)). Figure 6.12 (a) shows that more than half of the times the fit finds a value for $|A_S|^2$ that is either zero or in the interval between 0.0 and 0.02. The residual plot in Figure 6.12 (b) shows how the boundary at $|A_S|^2 = 0$ is responsible of the values found for $|A_S|^2$ and therefore the pull distribution behavior (Figure 6.12 (c)). In addition, sometimes the fit shows convergence problems, because the MINUIT minimizer gets stuck at the limit for $|A_S|^2$ (see Ref. [121]). We fix those cases by restarting the minimization from the local minimum of the likelihood, moving the starting point of $|A_S|^2$. To support the hypothesis that the boundary was causing the bad behavior of the $|A_S|^2$ pull distribution, another set of pseudo-experiments has been generated, with the generation value of $|A_S|^2 = 25\%$ far away from the boundary. We expect a gaussian distribution centered on 0.25 for the values of $|A_S|^2$ fitted on the pseudo-experiment, and a Normal distribution for the pulls. Figure 6.12 (d)–(f) show the obtained distribution which is in agreement with the expectations.

We verified whether the mean fit uncertainty for physics parameters in the toy studies is comparable with the uncertainty observed in data. This comparison is reported in Tab. 6.2, together with the parameters describing the pull distributions (mean and width). The maximum likelihood values are reliable, and the observed uncertainties agree with what expected from pseudo-experiments with the exception of $\delta_{\parallel}$, where the r.m.s. of the fit value is much larger than the one estimated in data. We observed that out of a large number of pseudo-experiments, it happen that the fit may converge to a minimum different than the generation value for $\delta_{\parallel}$, due to the proximity of the two minima (Fig. 6.11).

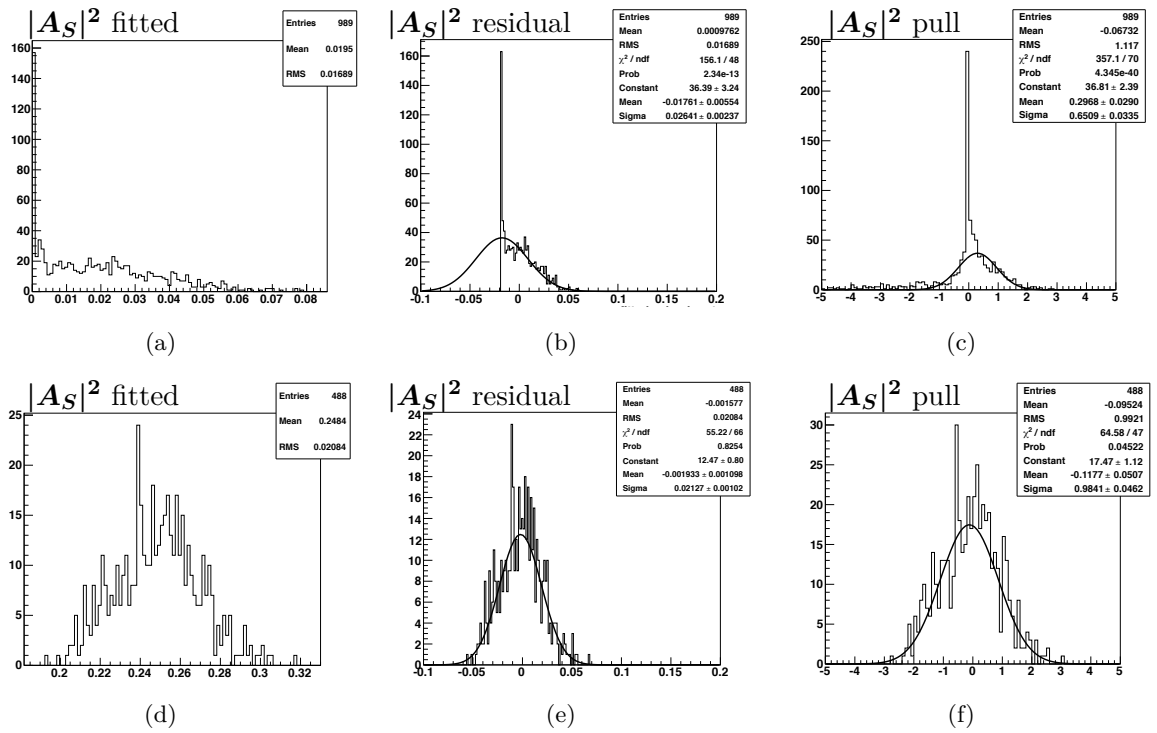


Figure 6.12: Value, residual and pull distributions for the S-wave amplitude $|A_S|^2$; pseudo-experiments generated with the $|A_S|^2$ value of 1.8% correspond to plots (a)–(c), pseudo-experiments generated with the $|A_S|^2$ value of 25% correspond to plots (d)–(f)

6.3.1.2 β_s-floating fit

We performed studies with pseudo-experiments generated at $\beta_s = 0.11$ (the central value found in our data fit on the whole data sample). Figure E.2 in App. E shows the pull distribution for the main physics parameters. When β_s is allowed to float a bias is present on $\Delta\Gamma_s$, and those of $\alpha_\perp$ and $\alpha_\parallel$, although still limited, become larger. Figure E.2 (c) shows that no significant bias is affecting β_s, probably owing to the larger data sample size compared to previous iterations. For $\delta_\parallel$, $|A_S|^2$ and δ_S, similar considerations as for the β_s-fixed case hold. In general, the mean error of the parameters in pseudo-experiments are considerably larger than the errors found by the fit in data. Following the same scheme adopted in describing the β_s-fixed fit, in table Tab. 6.3 we report a comparison between the mean error on a given parameter in the toy and the error obtained by fitting the data. In Fig. 6.13 we show mean and width of the pull distributions for all the parameters appearing in the full likelihood function.

6.3.1.3 Pulls studies with random input values of parameters

We repeat the pulls study of the β_s-fixed fit by generating about 2000 pseudo-experiments with input values of the PDF parameters randomly taken from a flat distributions. Those distributions are centered at the values of the parameters found in the fit to data, and have a

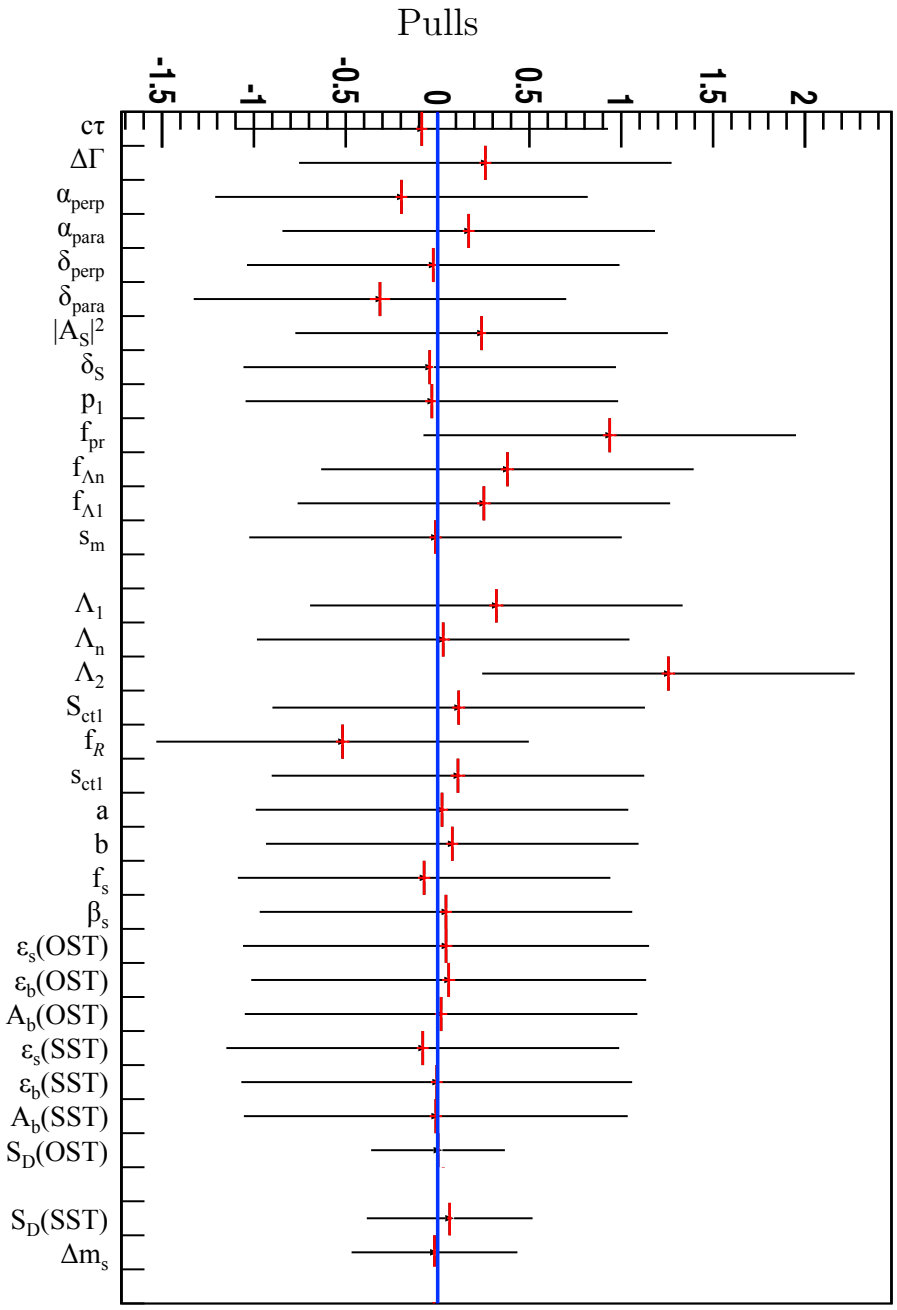


Figure 6.13: Summary of the pull distributions for all parameters of the likelihood function for the β_s-floating fit. The red markers and the black lines are the means and the widths of the pulls, respectively. The blue line is drawn as reference. The scale factors of the tagging dilutions along with Δm_s present a smaller pull-width because they are gaussian-constrained in the fit.

Parameter	Pull mean	Pull σ	Mean Error	Fit Error
$c\tau$	-0.043 ± 0.036	1.049 ± 0.028	0.00061	0.00062
$\Delta\Gamma_s$	0.016 ± 0.034	0.998 ± 0.028	0.028	0.029
$\alpha_\perp$	-0.180 ± 0.032	0.942 ± 0.028	0.012	0.012
$\alpha_\parallel$	0.072 ± 0.032	0.957 ± 0.027	0.013	0.012
$\delta_\perp$	-0.049 ± 0.027	0.802 ± 0.030	0.58	0.61

Table 6.2: Mean and σ of the pull distribution of the β_s-fixed fit; variable mean error and in the last column the fitted parameter error.

Parameter	Pull mean	Pull σ	Mean Error	Fit Error
β_s	0.046 ± 0.033	0.984 ± 0.032	0.197	0.123
$c\tau$	-0.086 ± 0.030	0.914 ± 0.025	0.00059	0.00058
$\Delta\Gamma_s$	0.260 ± 0.032	0.965 ± 0.028	0.038	0.028
$\alpha_\perp$	-0.197 ± 0.032	0.949 ± 0.028	0.012	0.011
$\alpha_\parallel$	0.170 ± 0.032	0.963 ± 0.027	0.013	0.011
$\delta_\perp$	-0.020 ± 0.019	0.537 ± 0.026	0.67	0.74

Table 6.3: Mean and σ of the β_s-floating fit pull distribution; variable mean error and in the last column the fitted parameter error.

width 5 times larger than the estimated errors, avoiding unphysical regions of the parameters space. The parameters are randomized all together for each pseudo-experiments. The test has the purpose to check possible biases against different true values of the fitting parameters. If the likelihood fit present biases for particular true values of a given parameter, the corresponding pull would show a deformation of the Normal distribution. In such a case, a dedicated study can be performed for that parameter.

We already know and discussed the critical points in the parameter space for strong phases (the symmetry points) and the S-wave fraction (the boundary at zero). We want to span all other regions of the parameters space. The results of the test is presented Fig. E.3 in App. E, and a summary in Fig. 6.14, while we report the result of the test in Tab. 6.4 for the main physics parameters. The pulls are regular for all quantities. In particular, we see that the pulls for $\delta_\parallel$ and $|A_S|^2$ are now regular, since their generated values are mostly populating regions far from the problematic points (respectively, $\delta_\parallel = \pi$ and $|A_S|^2 = 0$).

6.3.2 Fit of realistic simulated data samples

We validate the robustness of the analysis by using high-statistics samples generated by the MC simulation of $B_s^0 \to J/\psi K^+K^-$ decays described in Sect. 5.5.2, that fully simulates the P- and S-waves interference. This is a new consistency test with respect to the previous analyses aimed at further understanding potential biases in the estimation of physical parameters of interest and

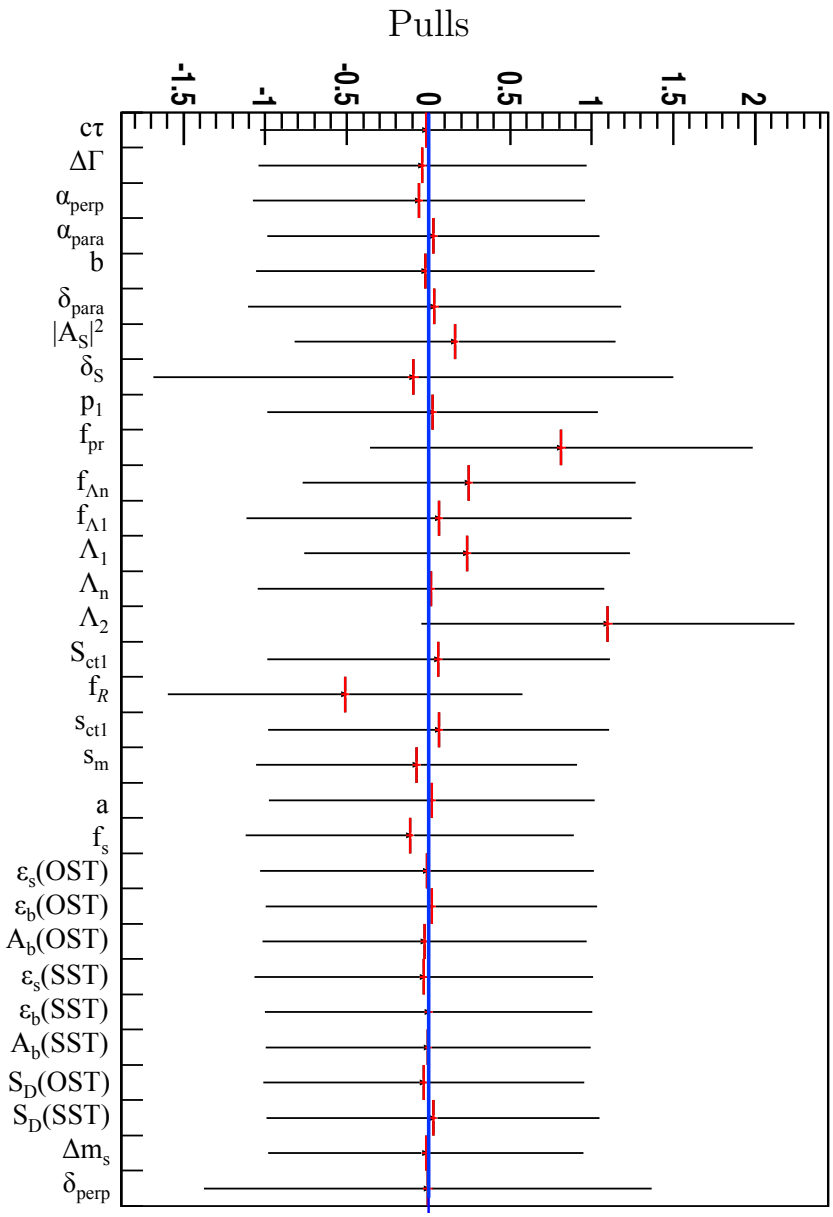


Figure 6.14: Summary of the pull distributions for all the variables of the likelihood function for the β_s-fixed fit with random values of PDF parameters in the generation of the pseudo-experiments. The red markers and the black lines are the means and the widths of the pulls, respectively. The blue line is drawn as reference.

Parameter	Pull mean	Pull σ
$c\tau$	-0.013 ± 0.024	0.983 ± 0.018
$\Delta\Gamma_s$	-0.038 ± 0.024	0.976 ± 0.019
$\alpha_\perp$	-0.058 ± 0.023	0.9621 ± 0.018
$\alpha_\parallel$	0.2837 ± 0.023	0.9838 ± 0.019
$\delta_\perp$	-0.051 ± 0.031	1.211 ± 0.030
$\delta_\parallel$	-0.037 ± 0.025	0.994 ± 0.024
$\lvert A_S\rvert^2$	-0.161 ± 0.019	0.950 ± 0.020

Table 6.4: Mean and σ of the β_s-fixed fit pull distribution with random values of PDF parameters in the generation of the pseudo-experiments.

testing the implementation of the complicated decay-rate expressions of the likelihood, since the simulation is written with a different formalism. Specifically, we want to prove the reliability of the estimation of the S-wave fraction by the angular analysis. Indeed, based on angular-information only (the KK-mass variable is never used in fits) D0 claims a $(17.3 \pm 3.6)\%$ fraction in contrast to findings from CDF, LHCb, and ATLAS in the 0–4% range. An improperly-modeled component of misidentified S-wave from $B^0 \to J/\psi K\pi$ decays could mimic the KK S-wave and potentially explain the D0 discrepancy, as we discuss in Sect. 7.4.

We generate 8 samples of 50 000 signal events only (no background is simulated), which corresponds to 5 times the statistics of signal events in real data. Table 6.5–6.8 summarize the results of the interesting parameters for the fit of those samples. Point estimates are generally very close to the generated values except for the strong phase δ_S, even when the sample is generated with a large S-wave fraction. However, we consider the results satisfactory, since we do not quote δ_S in our measurement given the small fraction of S-wave seen in data which spoils the sensitivity on that parameter, and the simplified model used in the PDF parametrization of the S-wave component that can alter the estimation of the S-wave phase (see Sect. 2.2.2). The test shows that the fit can actually estimates the proper S-wave fraction and there are no macroscopic biases in the $\Delta\Gamma_s$ estimates. Estimates of β_s seems to be systematically smaller than the generated value, but the effect is generally less than a standard deviation. Figure 6.15 summaries the pull for $|A_S|^2$, $\Delta\Gamma_s$, and β_s, observed in the fits to the 8 simulated samples.

	Input	Fitted	Pull
$c\tau$	0.049	0.04870 ± 0.00023	-1.3
$\|A_0\|^2$	0.1610	0.1589 ± 0.0048	-0.4
$\|A_\perp\|^2$	0.4633	0.4572 ± 0.0061	-1.0
$\delta_\parallel$	3.9	3.84 ± 0.048	-1.4
$\delta_\perp$	0.8	-0.46 ± 0.41	-3.0
δ_S	1.22	3 ± 10	0.2
$\|A_S\|^2$	0	0 ± 0.00025	0.0
$\Delta\Gamma_s$	0.09	0.087 ± 0.012	-0.3
β_s	0.02	-0.068 ± 0.054	-1.6

	Input	Fitted	Pull
$c\tau$	0.0458	0.04570 ± 0.00025	-0.4
$\|A_0\|^2$	0.5167	0.5123 ± 0.0053	-0.8
$\|A_\perp\|^2$	0.2521	0.2484 ± 0.0069	-0.5
$\delta_\parallel$	3.04	3.06 ± 0.23	0.1
$\delta_\perp$	2.73	2.34 ± 0.66	-0.6
δ_S	1.22	2.64 ± 0.90	1.6
$\|A_S\|^2$	0.01	0.0100 ± 0.0045	0.0
$\Delta\Gamma_s$	0.09	0.0914 ± 0.011	0.1
β_s	0.02	-0.022 ± 0.064	-0.7

Table 6.5: Fit to $B_s^0 \to J/\psi K^+K^-$ simulated data: sample 1 (left) and 2 (right).

	Input	Fitted	Pull
$c\tau$	0.0458	0.04579 ± 0.00021	0.0
$\|A_0\|^2$	0.4982	0.4958 ± 0.0053	-0.5
$\|A_\perp\|^2$	0.2790	0.2740 ± 0.0068	-0.7
$\delta_\parallel$	3.04	3.00 ± 0.14	-0.3
$\delta_\perp$	2.73	2.28 ± 0.48	-0.9
δ_S	1.22	0.62 ± 0.23	-2.7
$\|A_S\|^2$	0	0 ± 0.0023	0.0
$\Delta\Gamma_s$	0	0.0075 ± 0.0092	0.8
β_s	0.02	-0.086 ± 0.090	-1.2

	Input	Fitted	Pull
$c\tau$	0.0458	0.04600 ± 0.00029	0.7
$\|A_0\|^2$	0.4977	0.4937 ± 0.0051	-0.8
$\|A_\perp\|^2$	0.2798	0.2754 ± 0.0067	-0.7
$\delta_\parallel$	3.04	3.19 ± 0.18	0.9
$\delta_\perp$	2.73	2.77 ± 0.33	0.1
δ_S	1.22	3.38 ± 0.71	3.0
$\|A_S\|^2$	0.01	0.0099 ± 0.0080	0.0
$\Delta\Gamma_s$	0.09	0.076 ± 0.015	-1.0
β_s	-0.8	-0.807 ± 0.069	-0.1

Table 6.6: Fit to $B_s^0 \to J/\psi K^+K^-$ simulated data: sample 3 (left) and 4 (right).

	Input	Fitted	Pull
$c\tau$	0.0458	0.04574 ± 0.00024	-0.2
$\|A_0\|^2$	0.5230	0.5170 ± 0.0052	-1.2
$\|A_\perp\|^2$	0.2431	0.2443 ± 0.0068	0.2
$\delta_\parallel$	3.25	3.15 ± 0.14	-0.7
$\delta_\perp$	2.95	2.27 ± 0.44	-1.6
δ_S	2.98	0.17 ± 0.43	-6.5
$\|A_S\|^2$	0.042	0.034 ± 0.0086	-0.9
$\Delta\Gamma_s$	0.123	0.122 ± 0.011	-0.1
β_s	-0.075	-0.090 ± 0.052	-0.3

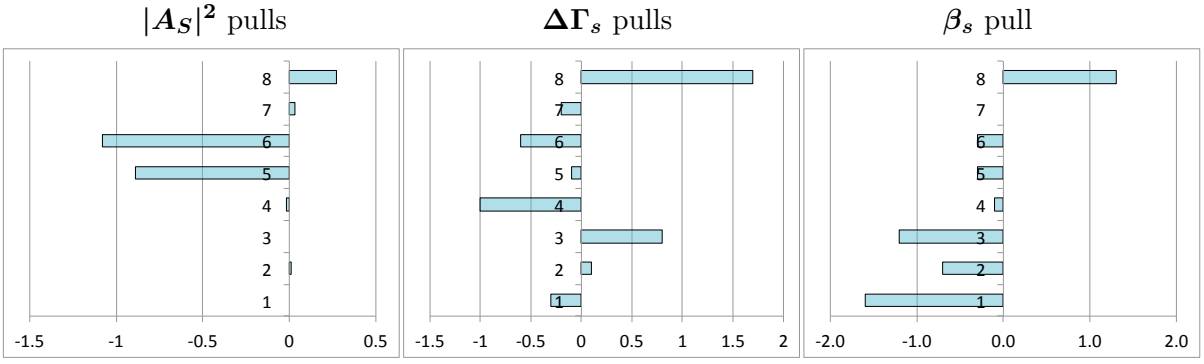
	Input	Fitted	Pull
$c\tau$	0.0458	0.04577 ± 0.00025	-0.1
$\|A_0\|^2$	0.5280	0.5170 ± 0.0053	-0.8
$\|A_\perp\|^2$	0.2358	0.2349 ± 0.0066	-0.1
$\delta_\parallel$	3.15	3.11 ± 0.24	-0.2
$\delta_\perp$	2.73	2.71 ± 0.32	-0.1
δ_S	2.98	2.20 ± 0.33	-2.4
$\|A_S\|^2$	0.17	0.1611 ± 0.0082	-1.1
$\Delta\Gamma_s$	0.163	0.155 ± 0.013	-0.6
β_s	0.22	0.206 ± 0.052	-0.3

Table 6.7: Fit to $B_s^0 \to J/\psi K^+K^-$ simulated data: sample 5 (left) and 6 (right).

	Input	Fitted	Pull
$c\tau$	0.0458	0.045813 ± 0.00031	0.0
$\|A_0\|^2$	0.5424	0.5363 ± 0.0052	-1.2
$\|A_\perp\|^2$	0.2251	0.2301 ± 0.0068	0.7
$\delta_\parallel$	3.15	3.02 ± 0.32	-0.4
$\delta_\perp$	2.8	2.31 ± 0.36	-1.4
δ_S	3.5	-0.81 ± 0.30	-14.4
$\|A_S\|^2$	0.5	0.5003 ± 0.0081	0.0
$\Delta\Gamma_s$	0.09	0.087 ± 0.017	-0.2
β_s	0.02	0.023 ± 0.070	0.0

	Input	Fitted	Pull
$c\tau$	0.0458	0.04539 ± 0.0011	-0.4
$\|A_0\|^2$	0.5424	0.5550 ± 0.011	1.1
$\|A_\perp\|^2$	0.2251	0.2107 ± 0.012	-1.2
$\delta_\parallel$	3.15	3.22 ± 0.48	0.1
$\delta_\perp$	2.8	3.65 ± 0.69	1.2
δ_S	0.5	3.67 ± 0.69	4.6
$\|A_S\|^2$	0.8	0.8020 ± 0.0075	0.3
$\Delta\Gamma_s$	0.09	0.155 ± 0.039	1.7
β_s	0.02	0.109 ± 0.069	1.3

Table 6.8: Fit to $B_s^0 \to J/\psi K^+K^-$ simulated data: sample 7 (left) and 8 (right).



(a) (b) (c)

Figure 6.15: Summary of the pulls of $|A_S|^2$ (a), $\Delta\Gamma_s$ (b) and β_s (c) measured in the 8 fits to $B_s^0 \to J/\psi K^+K^-$ simulated data.

Chapter 7

Measurements of B_s^0 lifetime, $\Delta\Gamma_s$, and $B_s^0 \to J/\psi K^+K^-$ amplitudes.

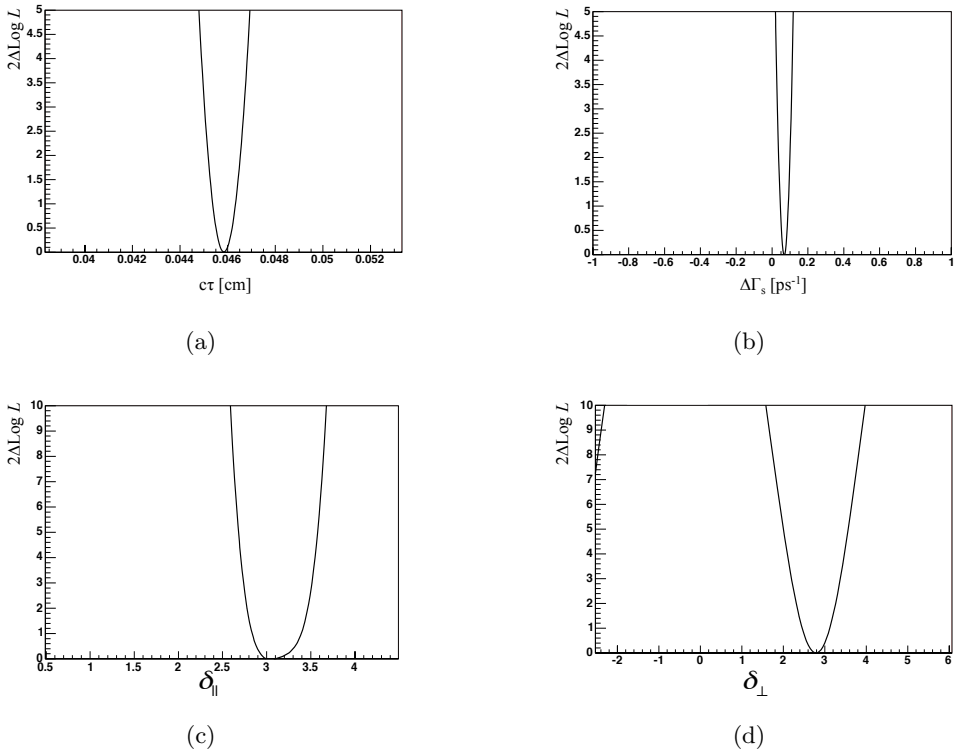
In this chapter the results of the measurement of the B_s^0 lifetime, the width difference $\Delta\Gamma_s$, the polarization amplitudes $|A_0(0)|^2$ and $|A_\parallel(0)|^2$, and the strong phase $\delta_\perp$, when β_s is fixed to its SM expectation are presented. A full study of the effects of systematic uncertainties on each parameter is presented. Projections of the fit results onto individual event variables in data are shown. Finally, we present a dedicated study for estimating the fraction of the S-wave component along with the $B^0 \to J/\psi K^+\pi^-$ background.

7.1 Fit results for β_s fixed at the SM expectation

As was demonstrated in Sect. 6.3, the ML fit produces no or minimally biased measurements of key physics parameters with β_s fixed to its expected value in the SM (0.02). Aside from constraining β_s, the likelihood fit used to produce these measurements is identical to that used in the final measurement of β_s. Thus, in addition to giving the precise measurements of $c\tau$, $\Delta\Gamma_s$ and the angular amplitudes, comparing these values to theoretical predictions and other measurements provides a good cross check of the full analysis technique.

The fit parameters for which point estimates can be quoted are the proper decay length, $c\tau$ (from which the B_s^0 lifetime, τ_s, can be calculated), the decay width difference, $\Delta\Gamma_s$, the polarization fractions $\alpha_\perp$ and $\alpha_\parallel$, and the strong phase $\delta_\perp$. The fit results for the physics observables in the case where the phase β_s is fixed to its SM value are reported in the rightmost column of Tab. 7.1 along with the parameters correlations in Tab. 7.2. We also report the results of the fit performed only to the new data added in this final measurement with respect to the previous CDF measurement [18]. They show good consistency with past results. In Fig. 7.1 we shows the scan of the likelihood for $\Delta\Gamma_s$, $\delta_\perp$ and $\delta_\parallel$, which present a regular parabolic shape for the former two parameters, but a non-parabolic one for the latter.

We also cross-check the results with two alternative fits, of which results are reported in Tab. 7.3. The first check uses the untagged likelihood in the fit. Indeed, the untagged likelihood provide reliable estimates when β_s is set to the SM value for all physical parameters [123], except for $\delta_\perp$ which does not enter the untagged likelihood, and $\delta_\parallel$, which is affected by the problems due to the symmetry transformation. The second check exploits the fit in which the likelihood



(a) (b) (c) (d)

Figure 7.1: Scan of the likelihood for *ct* (a), $\Delta\Gamma_s$ (b), $\delta_\parallel$ (c), and $\delta_\perp$ (d). The scan is obtained by calculating the value of the likelihood in different points of a given parameter with all other parameters fixed at the values obtained in the global minimum of the likelihood. The scan is not a profile-likelihood (re-minimization of all parameters for each point), and gives only a qualitative test of the shape of the likelihood around the minimum of the parameter.

contains only the P-wave parameterization in the PDF of the signal; since the default fit found a negligible value of the S-wave, $(0.8 \pm 1.8)\%$, the alternative fit is expected to give results of the others physical parameters very close to the default fit's ones. Indeed, the cross checks are very satisfactory.

The strong phase, $\delta_\parallel$, was shown to demonstrate non-Gaussian behavior even with β_s fixed to zero (Sect. 6.3.1), therefore we do not quote a point estimate for such parameter. The S-wave fraction measured in data doesn't allow us to quote reliable estimation of the phase δ_S, being the fraction measured compatible with zero. A dedicated, alternative fit is considered in Sect. 7.4 for a precise measurement of the S-wave fraction $|A_S|^2$ and cross-check the obtained results. The mixing frequency Δm_s is not reported, since this parameter is Gaussian-constrained in the fit to the value measured by CDF [43]. As a control check, we perform a fit in which we release the Gaussian-constraint and we treat Δm_s as a floating fit's parameter. The likelihood of the β_s-fixed fit is sensitive to Δm_s, mainly though the time-evolution of the interference between the polarization amplitudes, which contains terms as:

$$\cos\alpha \cos 2\beta_s \sin(\Delta m_s t) \tag{7.1}$$

Parameter	Previous result ($5.2\,\mathrm{fb}^{-1}$)	New data only ($4.4\,\mathrm{fb}^{-1}$)	All sample
$\Delta\Gamma_s$ [ps^{-1}]	0.075 ± 0.035	0.029 ± 0.033	0.068 ± 0.026
$\alpha_\perp$	0.266 ± 0.014	0.296 ± 0.017	0.279 ± 0.010
$\alpha_\parallel$	0.306 ± 0.015	0.309 ± 0.019	0.309 ± 0.012
$\delta_\perp$	2.95 ± 0.64	3.45 ± 0.84	2.79 ± 0.53
$\delta_\parallel$	3.08 ± 0.63	3.11 ± 0.40	3.09 ± 0.36
$c\tau$ [μm]	458.6 ± 7.5	457.2 ± 9.3	458.2 ± 5.8
$\lvert A_S\rvert^2$	0.019 ± 0.027	0.00 ± 0.03	0.008 ± 0.018
δ_S	1.37 ± 0.77	2 ± 2	1.26 ± 0.75

Table 7.1: Summary of the results of the main physical parameters in the fit to data with β_s fixed to its SM value (rightmost column). We report also the results of previous CDF measurement (second column) compared with the results of the fit to new data only (third column).

	$\Delta\Gamma_s$	$\alpha_\perp$	$\alpha_\parallel$	$\delta_\perp$
$c\tau$	0.52	-0.16	0.07	0.03
$\Delta\Gamma_s$		-0.17	0.06	-0.01
$\alpha_\perp$			-0.53	-0.01
$\alpha_\parallel$				0.05

Table 7.2: Matrix of correlation coefficients among main parameters of the β_s-fixed fit.

where α is the strong phases between the amplitudes. In the amplitudes interference, there are also terms that evolves with $\sin\alpha\cos(\Delta m_s t)$, which are suppressed by the fact that α takes value close to π. The oscillatory terms in the time-evolution of the squared moduli of the amplitudes, $\sin 2\beta_s \sin(\Delta m_s t)$, are instead suppressed by $\beta_s = 0.02$. The results of such test is reported in Tab. 7.3 and show good agreement with respect to the default fit. The found value of Δm_s is very close to the CDF measured value, and build confidence on the sensitivity of the oscillatory terms in the likelihood parametrization and on the decay-length resolution model. However, the sensitivity on Δm_s is limited, as the likelihood-ratio profile in Fig. 7.2 shows. Hence, we retain the choice to constrain the parameter in the fit.

In Tab. 7.4, we report the result of all parameter from the fit. Few comments are in order. The large fraction of the prompt background (with null lifetime), about 88% of the total background, allows for a good modeling of the resolution function in ct, with a precise determination of the scale factors for the decay-length uncertainty. The ct resolution is dominated ($\approx 85\%$) by the Gaussian with the smaller scale factor (1.308 ± 0.012). The background at $ct < 0$, which is due to badly reconstructed events, have similar lifetime as the background at $ct > 0$ with short lifetime, in agreement with the expectation that such mis-reconstruction effect

Parameter	Untagged fit	Fit w/o S-wave	Δm_s floating	Default fit
$\Delta\Gamma_s$ [ps^{-1}]	0.068 ± 0.027	0.069 ± 0.026	0.069 ± 0.026	0.068 ± 0.026
$\alpha_\perp$	0.279 ± 0.011	0.277 ± 0.010	0.279 ± 0.011	0.279 ± 0.010
$\alpha_\parallel$	0.310 ± 0.012	0.310 ± 0.012	0.309 ± 0.012	0.309 ± 0.012
$\delta_\perp$	–	2.66 ± 0.49	2.64 ± 0.67	2.79 ± 0.53
$\delta_\parallel$	–	–	3.09 ± 0.36	3.09 ± 0.36
$c\tau$ [μm]	458.5 ± 5.8	458.3 ± 5.8	458.3 ± 5.8	458.2 ± 5.8
$\lvert A_S \rvert^2$	0.008 ± 0.022	–	0.008 ± 0.018	0.008 ± 0.018
δ_S	1.22 ± 0.80	–	1.26 ± 0.80	1.26 ± 0.80
Δm_s [ps^{-1}]	–	17.74 ± 0.11	17.65 ± 0.24	17.74 ± 0.11

Table 7.3: Comparison of the results of the the physics parameters for the default fit (rightmost column) with alternative fits (central columns). In the second column, the alternative fit exploit the untagged likelihood. In the third column, the likelihood of the alternative fit doesn't have the S-wave component in its parametrization. In the fourth column, the results of the fit with Δm_s treated as a floating parameter.

is symmetric with respect to $ct = 0$. The efficiency of both OST and SSKT are in agreement with expectations from the dedicated calibrations (Chap. 4). The dilution scale factors were Gaussian-constrained in the fit, therefore, their estimations are naturally consistent with the values determined in the taggers calibrations. Fit projections are in Fig. 7.3, showing overall data model agreement. Specifically, we find good agreement between fit projections for sideband subtracted-data and the signal model.

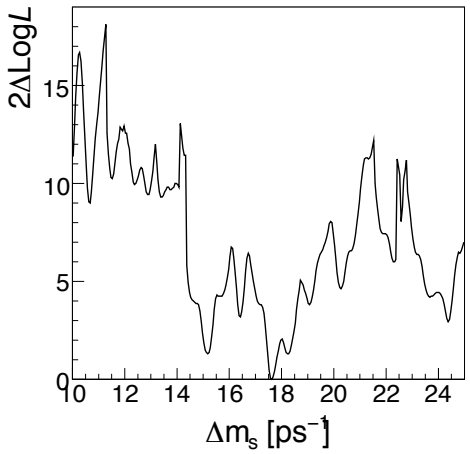
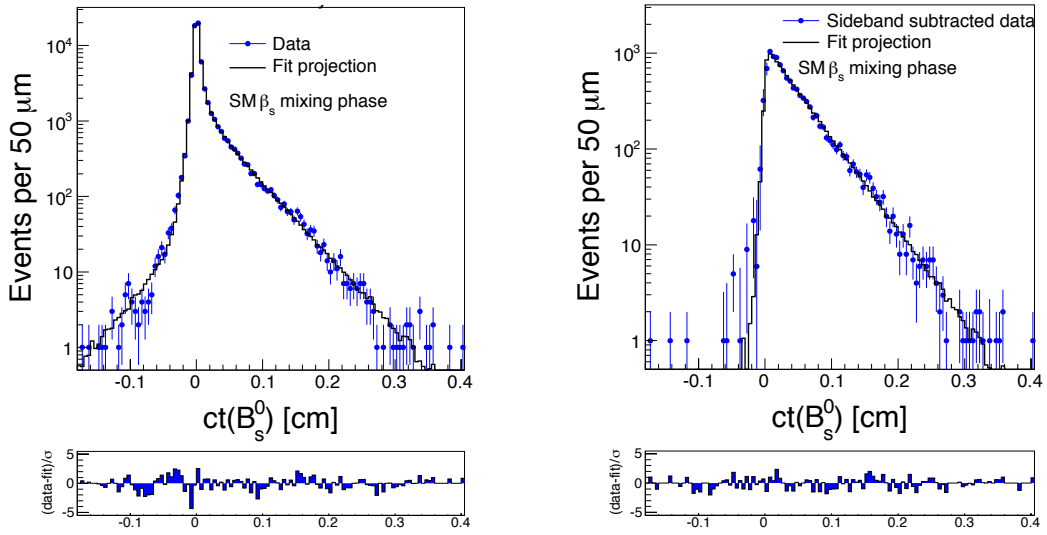


Figure 7.2: Likelihood-ratio profile for Δm_s. The profile is obtained by minimizing the likelihood at a fixed point of Δm_s with respect to all others parameters. The minimum of the likelihood is at $\Delta m_s = 17.65\,\text{ps}^{-1}$, in agreement with the measured value $\Delta m_s = (17.77 \pm 0.11)\,\text{ps}^{-1}$ [43], but the irregular shape and the presence of local minima spoils the sensitivity on Δm_s.

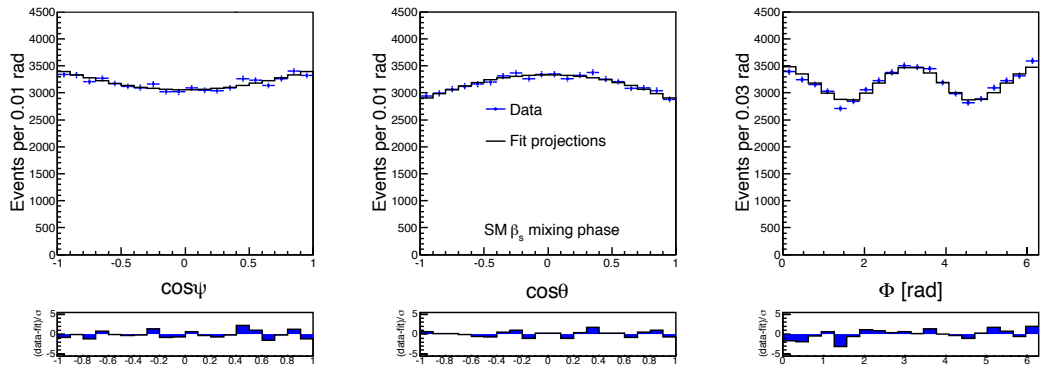
Parameter	Description	
β_s	CP-violating phase of B_s^0-$\bar{B}_s^0$ mixing amplitude	fixed to 0.02
$\Delta\Gamma_s$	lifetime difference $\Gamma_L - \Gamma_H$	$0.068 \pm 0.026\,\mathrm{ps}^{-1}$
$\alpha_\perp$	fraction $\lvert a_\perp\rvert^2$	0.279 ± 0.010
$\alpha_\parallel$	fraction $\lvert a_\parallel\rvert^2/(1-\lvert a_\perp\rvert^2)$	0.309 ± 0.012
$\delta_\perp$	$\arg(A_\perp A_0^\star)$	2.79 ± 0.53
$\delta_\parallel$	$\arg(A_\parallel A_0^\star)$	3.09 ± 0.36
$c\tau$	B_s^0 mean lifetime	$458.2 \pm 5.8\,\mu\mathrm{m}$
$\lvert A_S\rvert^2$	fraction of S-wave K^+K^- component in the signal	0.008 ± 0.018
δ_S	strong phase of S-wave amplitude	1.26 ± 0.75
Δm_s	B_s^0-$\bar{B}_s^0$ mixing frequency	$17.74 \pm 0.11\,\mathrm{ps}^{-1}$
f_s	Signal fraction	0.1721 ± 0.0018
s_m	Mass uncertainty scale factor	1.727 ± 0.017
p_1	mass background slope	$-1.89 \pm 0.42\,(\mathrm{GeV}/c^2)^{-1}$
s_{ct_1}	lifetime uncertainty scale factor 1	1.308 ± 0.012
s_{ct_2}	lifetime uncertainty scale factor 2	3.34 ± 0.13
$f_{\mathcal{R}}$	relative fraction of Gaussians of the decay-length resolution	0.852 ± 0.010
f_{pr}	fraction of prompt background	0.8839 ± 0.0039
f_{Λ_n}	fraction of background with Λ_n lifetime	0.210 ± 0.038
f_{Λ_1}	fraction of background with Λ_1 lifetime	0.718 ± 0.039
Λ_n	Inverse of lifetime of a background component	$0.040 \pm 0.032\,\mu\mathrm{m}^{-1}$
Λ_1	Inverse of lifetime of a background component	$0.0439 \pm 0.0038\,\mu\mathrm{m}^{-1}$
Λ_2	Inverse of lifetime of a background component	$0.0134 \pm 0.0010\,\mu\mathrm{m}^{-1}$
a	parameter in background fit to Φ	0.1441 ± 0.0062
b	parameter in background fit to $\cos\Theta$	0.169 ± 0.013
S_D(OST)	OST dilution scale factor	1.089 ± 0.049
S_D(SSKT)	SSKT dilution scale factor	0.85 ± 0.18
ϵ_b(OST)	OST tagging efficiency for background	0.7592 ± 0.0019
ϵ_b(SSKT)	SSKT tagging efficiency for background	0.7266 ± 0.0026
A(OST)	OST background tag asymmetry	0.4975 ± 0.0025
A(SSKT)	SSKT background tag asymmetry	0.4956 ± 0.0033
ϵ_s(OST)	OST tagging efficiency for signal	0.9280 ± 0.0030
ϵ_s(SSKT)	SSKT tagging efficiency for signal	0.5222 ± 0.0067

Table 7.4: Results of the fitting parameters for the β_s-fixed fit to the entire data set.

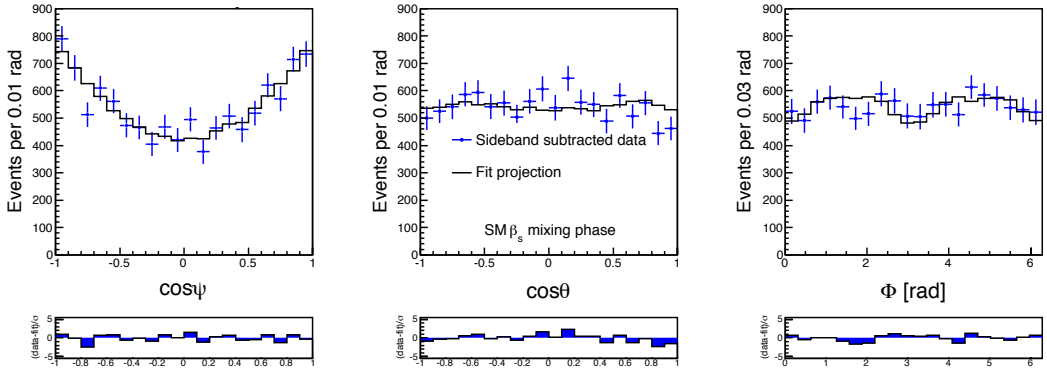


(a)

(b)



(c)



(d)

Figure 7.3: Projections onto *ct* distribution for the β_s-fixed fit in the complete fit range (χ^2/d.o.f. = 138/108) (a) and for sideband subtracted data (χ^2/d.o.f. = 97/97) (b); projections onto transversity angles distributions for the β_s-fixed fit in the complete fit range (χ^2/d.o.f. = 18/19 for $\cos\Psi$, χ^2/d.o.f. = 10/19 for $\cos\Theta$, and χ^2/d.o.f. = 29/19 for Φ) (c) and for sideband subtracted data (χ^2/d.o.f. = 17/19 for $\cos\Psi$, χ^2/d.o.f. = 23/19 for $\cos\Theta$, and χ^2/d.o.f. = 13/19 for Φ) (d).

7.2 Systematic Uncertainties

Systematics uncertainties are assigned to account for potential mis-modelings of the fit. For each systematic effect, we estimate the uncertainty using two ensembles of a few hundred pseudo-experiments each, one generated according to the default fit model, the other using a model modified in order to introduce the individual effect under study. The same randomization seeds are used to minimize the statistical fluctuations. For each observable, the systematic uncertainty is the mean of the distribution of the difference between the results of the fit on the default pseudo-experiments and the results on the modified pseudo-experiments:

$$\sigma_{\text{syst}} = |\langle \Delta_{\text{syst}} - \Delta_{\text{ref}} \rangle|, \tag{7.2}$$

with $\Delta_i = \zeta_i^{\text{fit}} - \zeta_i^{\text{gen}}$ and ζ_i^{fit} is the value of the parameter fitted in the pseudo-experiment, ζ_i^{gen} is the value used in the generation of the pseudo-experiments, and the index i stands for the pseudo-experiment generated according to the modified (syst) or default (ref) models. Being the value ζ^{gen} the same for the modified and default pseudo-experiments, the systematic uncertainty is simply:

$$\sigma_{\text{syst}} = |\langle \zeta_{\text{syst}}^{\text{fit}} - \zeta_{\text{ref}}^{\text{fit}} \rangle|. \tag{7.3}$$

When the statistical uncertainty on the mean value of the systematic shift is larger than the systematic shift itself, we use the statistical uncertainty rather than the central value for the systematic uncertainty. Results are reported in Tab. 7.5.

In the following we list the considered systematic effects.

Signal Angular Efficiency

The angular acceptance of the detector (Sect. 6.2.2) is modeled with a linear combination of Legendre polynomials and spherical harmonics with expansion coefficients determined from fitting a three-dimensional histogram in reweighed simulated events (Sect. 5.5.1). Inaccurate modeling or inaccuracies in the reweighing may introduce a systematic uncertainty. Because the reweighing procedure has the dominant impact in modifying the acceptance, an alternative fit on data has been performed by using non-reweighed simulation for the acceptance (Fig. 7.4). The difference between the values obtained for the observables of interest with this fit and the values obtained with the default fit has been taken as systematics uncertainty for the signal angular efficiency modeling effect. In addition, pseudo-experiments are generated with a modified model using an histogram for the angular distributions of the non-reweighed MC sample, such that both the effect of the mis-parametrization and the effect of the MC re-weighting are accounted for. This effect dominates the systematic uncertainty on the polarization amplitudes (Tab. 7.5).

Signal mass model

Mismodelings of the signal mass can impact the final results. The default model for the B_s^0 mass is a single Gaussian. An alternative model using two Gaussian, with common

Source of systematic effect	$c\tau$ [μm]	$\Delta\Gamma_s$ [ps^{-1}]	$\lvert A_\parallel(0)\rvert^2$	$\lvert A_0(0)\rvert^2$	$\delta_\perp$
Signal Angular Efficiency	0.29	0.0014	0.0134	0.0162	0.076
Mass Signal Model	0.17	0.0007	0.0006	0.0020	0.018
Mass Bkg Model	0.14	0.0006	0.0003	0.0002	0.034
ct Resolution	0.52	0.0010	0.0004	0.0002	0.066
ct Bkg	1.31	0.0057	0.0006	0.0012	0.064
Angular Bkg	0.46	0.0037	0.0011	0.0022	0.009
Sigma mass	0.85	0.0006	0.0003	0.0002	0.036
Sigma ct	0.63	0.0006	0.0003	0.0002	0.038
$B_d \to J/\psi K^*$ cross-feed	0.42	0.0055	0.0009	0.0058	0.039
SVX alignment	2.0	0.0004	0.0002	0.0001	0.034
Pull bias	0.2	0.0012	0.0021	0.0008	0.02
TOT	2.8	0.009	0.014	0.018	0.15

Table 7.5: Systematic uncertainties of the SM fit.

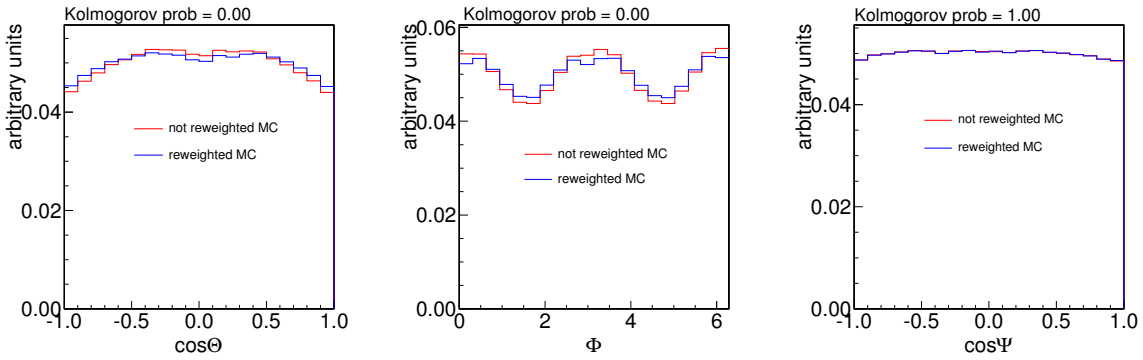


Figure 7.4: Comparison between the reweighed and non-reweighed MC simulation used for the angular acceptance.

means and independent and floating widths and relative fractions has been fit to data to extract parameters used to generate modified pseudo-experiments. The impact on the systematic uncertainties is moderate.

Background mass model

Similarly, if the chosen linear background mass model is inadequate it could contribute to the systematic uncertainty. An alternative model consisting in a decreasing exponential is applied on data to extract parameters to be used in the generation of pseudo-experiments. The impact of this source on systematic uncertainties is limited.

Lifetime resolution model

Mismodelings of the lifetime resolution can strongly impact the lifetime-related observables. The default model assumes a two-Gaussian resolution function. An alternative three-Gaussian resolution model, of which parameters have been extracted from a lifetime-fit to sideband data, has been used to generate the pseudo-experiments that are then fit with the default model.

Background lifetime model

Inadequate modeling of the lifetimes of various background components can systematically affect the B_s^0 lifetime measurement. We used an alternative model using the histogram of the proper time of the events populating the signal mass sidebands (with ct resolution parameters fixed at the results of the default fit) to generate pseudo-experiments to be fit with the default model. This has a relevant impact on several observables.

Angular background model and correlations

Effects from mismodeling of the angular background distribution are assessed with an alternative fit on pseudo-experiments generated using the histograms of the three transversity angles of the events populating the mass sidebands. Also, two-dimensional histograms of the angles versus σ_{ct} are used to probe potential effects from correlations among these quantities.

Mass uncertainty distribution

The mass resolution is assumed to be the same for signal and background events. We tested any residual effects through an alternative fit where mass uncertainty distributions are modeled by histograms of B_s^0 sideband data for background events and sideband subtracted signal region data for signal events separately (Fig. 7.5). To generate the pseudo-experiments, we have sampled the background uncertainties from separate upper and lower sideband histograms according to the generated B_s^0 mass thus accounting also for effects caused by neglecting small correlations between σ_m and mass.

ct uncertainty distribution

To account for a possible mis-parametrization of the ct error distributions effect, we per-

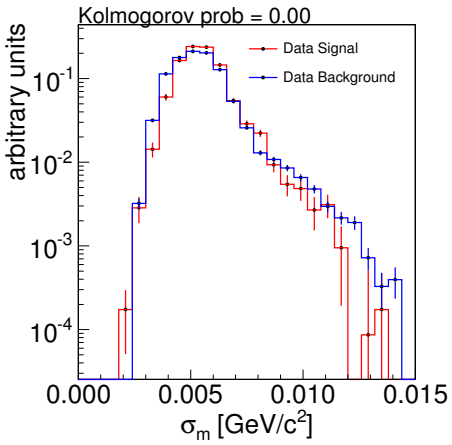


Figure 7.5: Comparison of distributions of σ_m for sideband and sideband-subtracted data.

form an alternative fit exploiting the ct uncertainty distributions taken from data histograms rather than from the model described in Sect. 6.2.5. The PDFs used in the alternative fits are simply the σ_{ct} histograms of B_s^0 sideband data for background events and sideband subtracted data for signal events, while to generate the pseudo-experiments, we have sampled the background uncertainties from separate upper and lower sideband histograms according to the generated B_s mass. This has been done in order to account for any effect caused by small correlations between σ_{ct} and the B_s^0 invariant mass.

$B^0 \to J/\psi K^+\pi^-$ contribution

The $B^0 \to J/\psi K^+\pi^-$ events misreconstructed as $B_s^0 \to J/\psi K^+K^-$ decays are not included in the default fit, though a fraction of these leaks into the B_s^0 signal region. We estimate the size of this contribution from data to be $(7.99 \pm 0.20)\%$ of the B_s^0 signal (see Sect. 7.4). We generate pseudo-experiments according to this fraction and fit with the default model which does not account for this component. The B_s^0 lifetime, width and transversity amplitudes of the generated events are taken from a previous CDF measurement [124]. The effect from this contribution has some limited impact on the final results.

SVX alignment

A systematic uncertainty is associated to the assumption that the silicon detector is perfectly aligned. The effects of plausible misalignments in lifetime-related observables have been addressed in detail in early Run II analyses, estimating a systematic uncertainty on the decay-length $c\tau$ in lifetime measurements of $2\,\mu$m. This value is assumed for our lifetime determination, and also to assess secondary effects propagating through correlations to the other parameters of interest. Pseudo-experiments in which the decay time in each event is randomly shifted $\pm 2\,\mu$m have been generated and fitted with the default fit and the mean of the differences between fit results used as systematic uncertainty.

Pull bias

We assign a systematic uncertainty to the parameters that present a bias, as studied in Sect. 6.3.1, which consists in the the mean shift $\Delta_i = |\zeta_i^{\rm fit} - \zeta_i^{\rm gen}|$ where $\zeta_i^{\rm fit}$ is the value of the parameter fitted in the pseudo-experiments, $\zeta_i^{\rm gen}$ is the value used in the generation of the pseudo-experiments, which is fixed to the values found in the fit to the data (Tab. 7.4).

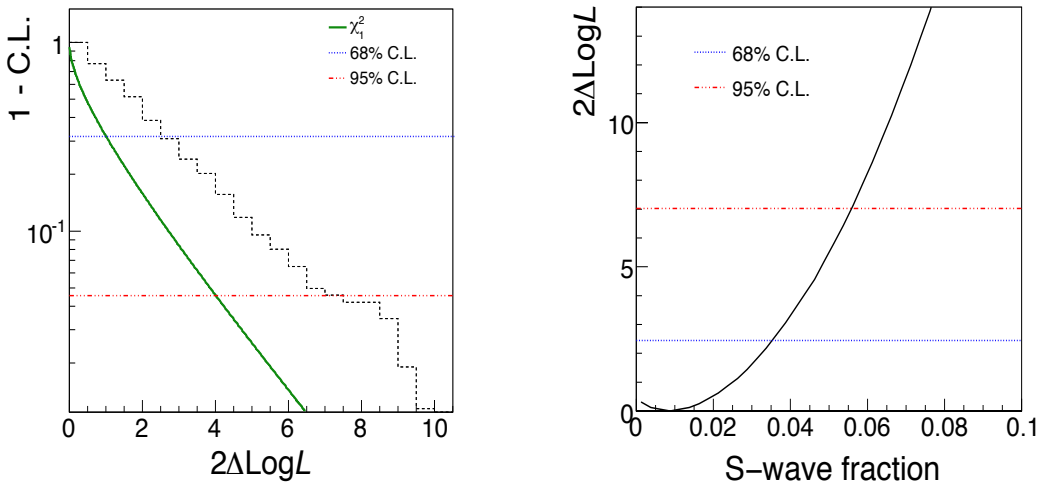
7.3 Final results

The relevant results for the SM fit using the full data set are

$$\begin{aligned} c\tau &= 458.2 \pm 5.8(\text{stat}) \pm 2.8(\text{syst})\ \mu\text{m}, \\ \Delta\Gamma_s &= 0.068 \pm 0.026(\text{stat}) \pm 0.009(\text{syst})\ \text{ps}^{-1}, \\ |A_\parallel|^2 &= 0.229 \pm 0.010(\text{stat}) \pm 0.014(\text{syst}), \\ |A_0|^2 &= 0.512 \pm 0.012(\text{stat}) \pm 0.018(\text{syst}), \\ \delta_\perp &= 2.79 \pm 0.53(\text{stat}) \pm 0.15(\text{syst}). \end{aligned}$$

The results are in good agreement with results from other experiments and with theoretical predictions. From the measurements of lifetime and decay width difference and their correlations we derive $\Delta\Gamma_s/\Gamma_s = 0.1045 \pm 0.048(\text{stat}) \pm 0.027(\text{syst})$, while from the world average value of the B^0 lifetime we derive $\tau_s/\tau_d = 1.006 \pm 0.015$ (stat+syst).

The fraction of S-wave found by the fit in the KK mass range $(1.009, 1.028)\,\text{GeV}/c^2$ is $(0.8 \pm 1.8)\%$, consistent with zero. Hence, we want to set an upper limit on this fraction. We use the profile-likelihood ratio statistic as a function of the S-wave fraction (Fig. 7.6). The method of the profile-likelihood ratio will be explained in details in Sect. 8.1, since it is used to extract the main results of this analysis, the bounds on the mixing phase. The extraction of the confidence interval must take into account the bound imposed in the likelihood to prevent negative S-wave fraction. We generate a large number of pseudo-experiments according to the likelihood. The pseudo-experiment are generated using the values determined by the fit on data for all the nuisance parameters and for the S-wave fraction. We perform two fits for each pseudo-experiments to calculate the likelihood ratio: one with the S-wave fraction fixed to the generation value, and the other with the S-wave fraction floating in the fit together with all nuisance parameters. We then plot the cumulative histogram of the likelihood ratios from pseudo-experiments. In Fig. 7.6 (a) we report the obtained map that provide us with the likelihood ratio needed to obtain the desired p-value. We then perform the likelihood-ratio profile in the $|A_S|^2$ space in Fig. 7.6 (b). Using the likelihood-ratio's map, we set an upper limit on the fraction of S-wave in our data of 5.6% at 95% C.L., and 3.5% at 68% C.L., in agreement with previous CDF results, and LHCb and ATLAS determination, but in contrast to findings by the D0 Collaboration in a similar mass window, $(17.3 \pm 3.6)\%$.



(a) (b)

Figure 7.6: Mapping of p-value $(1 - \text{C.L.})$ as a function of $2\Delta \log \mathcal{L}$ as evaluated in pseudo-experiments (a). The green curve represents the cumulative distribution for a χ^2 statistic, to show the deviation from the ideal case of the distribution obtained in pseudo-experiments. The map is used to draw the intervals for the fraction of S-wave. In (b), the likelihood-ratio profile for $|A_S|^2$. The profile is obtained by minimizing the likelihood at a fixed point of $|A_S|^2$ with respect to all others parameters. Setting the obtained likelihood ratios in the map (a), we derive the upper limit on the fraction of S-wave in our data of 5.6% at 95% C.L. (red dashed line), and 3.5% at 68% C.L (blue dotted line).

7.4 Estimation of the S-wave fraction

A simpler, alternative fit has been used to check the determination of the size of the S-wave component and of the $B^0 \to J/\psi K^+\pi^-$ background. A simultaneous likelihood fit of the unbinned $J/\psi K^+K^-$ and K^+K^- mass distributions is used to extract the fractions of interest, which are compared with the corresponding values obtained in the default fit from the angular analysis. Some key modifications of the analysis are introduced in order to better describe the details of the sample composition. We refine the templates of the $K\pi$ background by including its full resonance structure, both the P- and S- $K^+\pi^-$ waves along with their interference as described in Sect. 5.5.3. The relative size of the $K^+\pi^-$ background is determined by the fit, we only fix its S-wave to P-wave ratio to the value measured by *BABAR* [120].

The sample is enlarged by accepting events populating an extended K^+K^- mass window, which offers a longer lever arm to fit the details of the threshold region of the K^+K^- mass. A slightly different selection with respect to the default analysis provides a greater sensitivity to the small S-wave component; we introduced the following changes:

- the K^+K^- mass window ranges from threshold to 1.2 GeV/c^2 (was 1.009–1.028 GeV/c^2) providing access to the S-wave-enriched low-mass region and high-mass data to properly model the $K^\star(890)^0 \to K^+\pi^-$ background.

- the $ct(B) > 60\,\mu$m requirement suppresses combinatorial background by preserving roughly 85% of the signal.

The usual threshold on the NN output (NN>0.2) is applied.

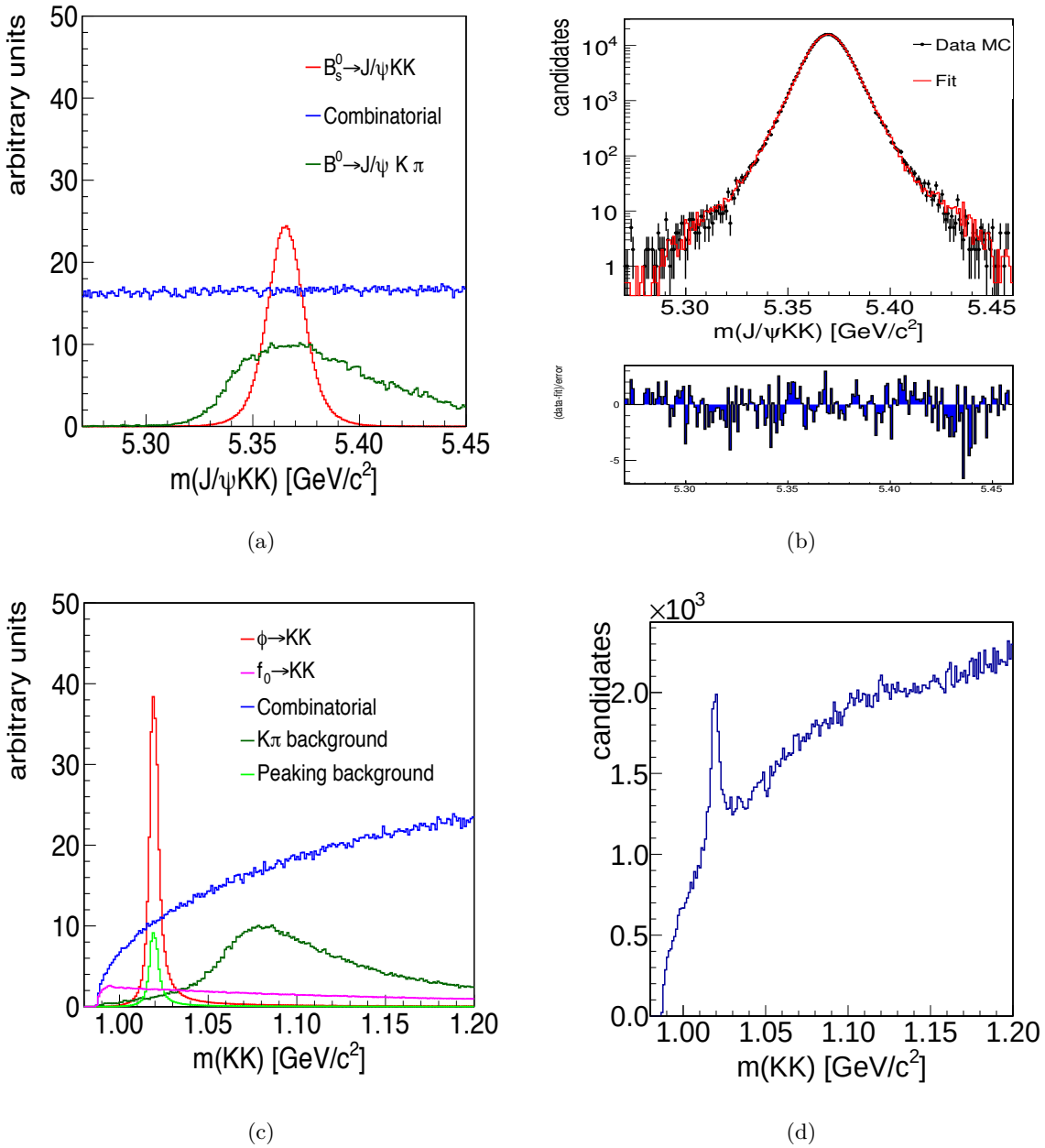
The $J/\psi K^+K^-$ mass distribution features three components (Fig. 7.7 (a)): a smooth, nearly-constant combinatorial background, which is modeled with a linear function, $B_{\rm c}(m_B)$; the B_s^0 signal component, which comprises both the P-wave ($J/\psi\phi$) and the S-wave ($J/\psi K^+K^-$); the B^0 broad peak for $B^0 \to J/\psi K^+\pi^-$ decays where the pion is misreconstructed as a kaon. The signal is modeled with a sum of three Gaussian functions, $G(m_B)$, whose parameters are extracted from a fit of a large sample of simulated decays and fixed in all subsequent steps of this check (Fig. 7.7 (b)). The B^0 line shape, $B_{B^0}(m_B)$, is determined from the simulation described in Sect. 5.5.3.

The K^+K^- mass distribution features five distinct components (Fig. 7.7 (c)): the narrow signal from $\phi(1020)$ mesons associated to a B_s^0 decay; the S-wave component; the $K^+\pi^-$ background due to misidentification of the pion as a kaon; the combinatorial background; a peaking background presumably due to a real $\phi(1020)$ associated to a pair of random tracks or muons. The $\phi(1020)$ resonance is modeled with a relativistic Breit-Wigner distribution, $W(m_{KK})$, whose parameters are fixed to those reported by the Particle Data Group [8] (Sect. 2.2.1), convoluted with a Gaussian resolution function whose width, σ_{KK}, is a free fit parameter. The S-wave component is modeled using a Flatté distribution, $F(m_{KK})$, whose parameters are those measured by BES [119] (Sect. 2.2.1). The combinatorial background is empirically modeled from an histogram, $B_{\rm c}(m_{KK})$, of the mass distribution of events populating the low-mass sideband of the B_s^0 peak, depleted in actual B^0 decays; this sample is selected by requiring $ct(B) < 60\,\mu$m, an orthogonal sample with respect to the fitted one (Fig. 7.7 (d)). To adjust the normalization of the peaking background present in $B_{\rm c}(m_{KK})$, an additional Breit Wigner component is assigned in the background PDF, $W_{\rm p}(m_{KK})$, and the fraction of such background, $f_{\rm p}$ is free to float, to take into account the possible dependence of the size of this component on the $J/\psi K^+K^-$ mass. The $K\pi$ component, $B_{K\pi}(m_{KK})$, is modeled from the dedicated simulation described in Sect. 5.5.3, where the S- and P-wave are simulated along with their interference as measured by *BABAR* (Fig. 7.7 (c)). The only external constraint introduced for this component is the ratio between the S- and the P-wave $K\pi$ contributions. It is $(7.3 \pm 1.8)\%$ of a $K^\star(890)^0$ signal in the range $[0.8, 1.0]\,\text{GeV}/c^2$ of the $K^+\pi^-$ spectrum [120].

The total PDF reads:

$$\begin{aligned} P(m_B, m_{KK}) =& f_s G(m_B)\Big[(1 - f_{S\rm w})W(m_{KK}) + f_{S\rm w}F(m_{KK})\Big] + \\ &+ f_d B_{B^0}(m_B) B_{K\pi}(m_{KK}) + \\ &+ (1 - f_s - f_d) B_{\rm c}(m_B)\Big[(1 - f_{\rm p})B_{\rm c}(m_{KK}) + f_{\rm p}W_{\rm p}(m_{KK})\Big]. \end{aligned} \tag{7.4}$$

The floating parameters of the fit are the overall fractions of B_s^0 and B^0 events, respectively f_s and f_d; the fraction of S-wave events relative to the total B_s^0 events, $f_{S\rm w}$; the B_s^0 mass and



(a) (b)

(c) (d)

Figure 7.7: Templates used to model the $J/\psi K^+K^-$ spectrum (a). Distribution of $J/\psi K^+K^-$ of simulated events with fit-projection overlaid to extract the B_s^0 signal template (b). Templates used to model the K^+K^- spectrum (c). Distribution of K^+K^- for events in the low-mass sideband of the B_s^0 peak with $ct(B) < 60\,\mu$m.

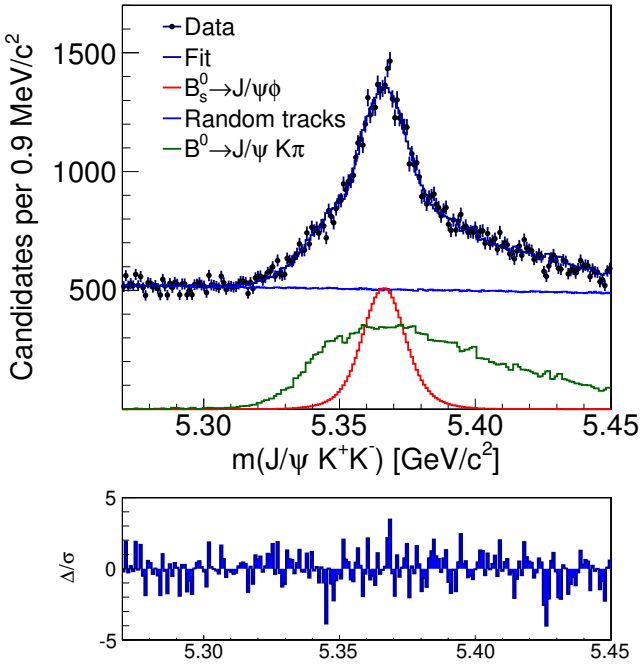
width, respectively M_B and σ_B; the K^+K^- mass resolution, σ_{KK}; the slope of the combinatorial background in the $J/\psi K^+K^-$ mass, λ; the fraction of peaking K^+K^- background, $f_{\rm p}$. No fraction is constrained to be in $[0,1]$. We have assumed that the PDF is separable for the $J/\psi K^+K^-$ and the K^+K^- variables by neglecting correlations among the masses, and checked that the biases on the estimated fractions introduced by this assumption are small with respect to the statistical errors, by using pseudo-experiments as explained in App. F.

The data distributions with fit projections are shown in Fig. 7.8, showing good agreement between data and model. The determinations of the S-wave fraction and the B^0 fractions show some sensitivity to the arbitrary choice made for the modeling of the combinatorial background at the level of one statistical standard deviation. The results of the fit are reported in Tab. 7.6.

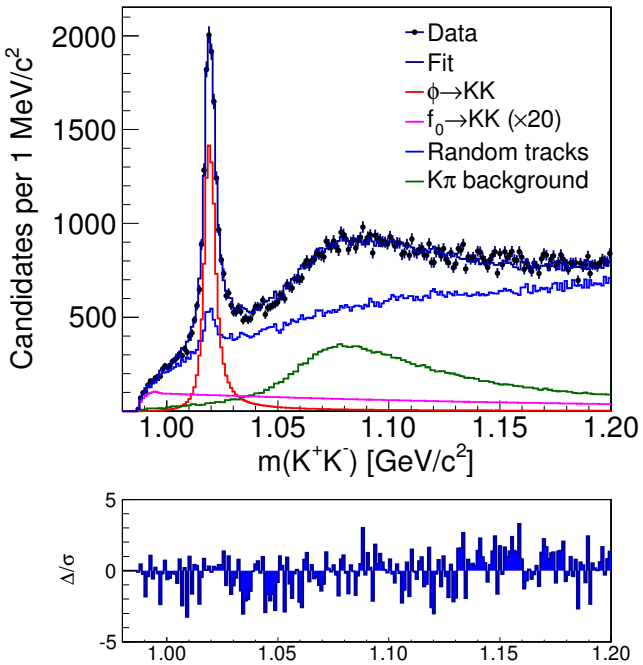
Parameter	Fitted value
f_s	0.0828 ± 0.0013
$M_B[\mathrm{GeV}/c^2]$	5.36626 ± 0.00011
$\sigma_B[\mathrm{GeV}/c^2]$	0.007033 ± 0.000088
$\lambda[(\mathrm{GeV}/c^2)^{-1}]$	-1.96 ± 0.40
f_d	0.2148 ± 0.0031
$f_{S\mathrm{w}}$	0.050 ± 0.014
σ_{KK}	0.001074 ± 0.000047
$f_{\rm p}$	-0.0056 ± 0.0011

Table 7.6: Results of the simultaneous fit to the $J/\psi K^+K^-$ and K^+K^- mass distributions.

We scale the fraction of K^+K^- S-wave and $K^+\pi^-$ background to the selection window used in the main angular fit. With the current model, the S-wave contribution to the B_s^0 signal in the K^+K^- window $(1.009, 1.028)\,\mathrm{GeV}/c^2$, is determined to be $(0.79 \pm 0.21)\%$, confirming the default fit result from the angular analysis and the previous CDF, LHCb and ATLAS determinations. The contribution of misreconstructed B^0 decays in the same K^+K^- window is found to be $(7.99 \pm 0.20)\%$ of the B_s^0 signal. This last determination is significantly larger than the 2% values typically derived with simulation by assuming only P-wave B^0 decays [18]. If neglected, this additional B^0 component could mimic a larger K^+K^- S-wave than present, specially if limited PID performance are available to discriminate kaon tracks from pion tracks. We check the assumption by fitting the K^+K^- spectrum removing the $K^+\pi^-$ S-wave model from the PDF. In this case we found a double fraction of K^+K^- S-wave, 1.5%, and half of the $K^+\pi^-$ background, 3.9%.



(a)



(b)

Figure 7.8: Distribution of $J/\psi K^+K^-$ (a) and K^+K^- (b) masses with fit projections overlaid. The χ^2/d.o.f. are 251/200 and 336/194, respectively.

Chapter 8

Mixing phase bounds

This chapeter present the main results of the analysis of the time-evolution of $B_s^0 \to J/\psi\phi$ decays. We provide confidence intervals that bound the mixing phase β_s and confidence regions in the $(\beta_s, \Delta\Gamma_s)$ space. We use a profile-likelihood ratio ordering with frequentist inclusion of systematic uncertainties to treat the uncertainties in a rigorous way that ensures the desired level of confidence and the robustness of our intervals. The experimental precision on the mixing phase is only second to the best published determination by LHCb, and in conjunction with it, establishes that large non-SM contributions to the B_s^0-$\bar{B}_s^0$ mixing phase are unlikely.

8.1 Measurement method

We chose not to use the results of the β_s-floating fit to quote point estimates (i.e. central value $\pm$ uncertainty) for the parameters of interest, specifically the mixing phase and $\Delta\Gamma_s$. The maximum likelihood estimators show significant biases that depend on the true values of the parameters as detailed in Sect. 6.3. In addition, the estimated uncertainties are unlikely to represent actual confidence regions with the desired level of confidence, as we will prove in the following. We abandon the point estimates by resorting to the Neyman construction of a fully frequentist confidence region [125].

Constructing correct and informative confidence regions from highly multi-dimensional likelihoods is challenging. In our case, determining the full 32-dimensional confidence space is computationally prohibitive. More importantly, the choice of the *ordering algorithm*, *i. e.*, the procedure chosen to cover regions of the parameters space until the desired confidence level is attained, is non-trivial. An arbitrariness is associated to the choice of the ordering algorithm, but one needs to avoid that the projection of the region onto the $(\beta_s, \Delta\Gamma_s)$ subspace of interest includes most, if not all, of the allowed values, thus yielding a useless result.

We choose to replace the likelihood, $\mathcal{L}(\beta_s, \Delta\Gamma_s, \boldsymbol{\zeta})$ with the profile likelihood, $\mathcal{L}_p(\beta_s, \Delta\Gamma_s, \hat{\boldsymbol{\zeta}})$. For every point in the $(\beta_s, \Delta\Gamma_s)$ plane, $\hat{\boldsymbol{\zeta}}$ are the values of nuisance parameters that maximize the likelihood in that point. The profile-likelihood ratio (PLR),

$$-2\Delta \ln \mathcal{L}_p = -2 \ln \mathcal{L}_p(\beta_s, \Delta\Gamma_s, \hat{\boldsymbol{\zeta}}) - 2 \ln \mathcal{L}(\hat{\beta}_s, \Delta\hat{\Gamma}_s, \hat{\boldsymbol{\zeta}}), \tag{8.1}$$

where $\mathcal{L}(\hat{\beta}_s, \Delta\hat{\Gamma}_s, \hat{\boldsymbol{\zeta}})$ is the value at the global minimum floating β_s and $\Delta\Gamma_s$ along with the nuisance parameters, can be used in principle as a χ^2 statistic with two degree of freedom (χ^2_2)

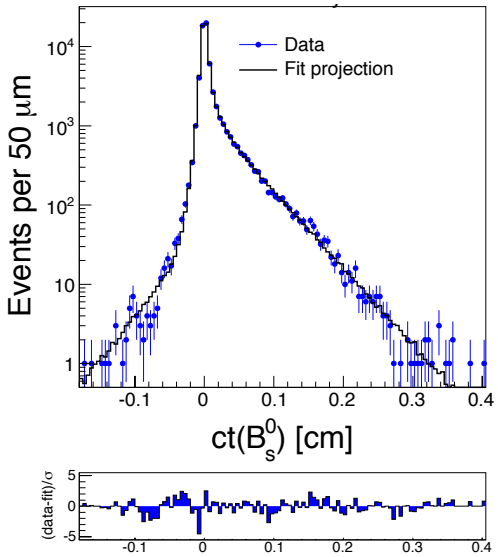
to derive confidence regions in the two-dimensional space $(\beta_s, \Delta\Gamma_s)$. The procedure to calculate the PLR is the following. We first fit the data with all parameters floating. Then, we define a 32×48 grid on the $(\beta_s, \Delta\Gamma_s)$ plane in the range $-\pi/2 < \beta_s < \pi/2$ and $-0.3 < \Delta\Gamma_s < 0.3\,\mathrm{ps}^{-1}$, and for each grid element we fit the data by floating all parameters but β_s and $\Delta\Gamma_s$, which are fixed to the values corresponding to the point to probe. The value of the likelihood in each of the two steps provide a PLR value for each point in the $(\beta_s, \Delta\Gamma_s)$ plane.

Simulations show that the observed distribution of PLR deviates from the χ^2 one. Specifically, the resulting confidence regions contain true values of the parameters with lower probability than the nominal confidence level (CL) because the observed PLR distribution has longer tails than a χ^2 (see for instance, Fig. 8.7(a)). In addition, the PLR distribution appears to depend on the true values of the nuisance parameters, which are unknown. We use therefore the simulation of a large number of pseudo-experiments to derive the actual distribution of PLR. The effect of systematic uncertainties is accounted for by randomly sampling a limited number of points in the space of all nuisance parameters and using the most conservative of the resulting PLR distributions to calculate the final confidence level. We obtain these distributions by generating 16 ensembles of 1000 pseudo-experiments each. In each ensemble, the true values of β_s and $\Delta\Gamma_s$ correspond to the probed point, while the true values of the nuisance parameters are a random sampling from an hypercube centered at their best fit values in data, with side corresponding to 10 standard deviations. We ensure coverage over a wide range of possible values, but always within their physically allowed range, in order to guarantee the desired CL of the final regions against systematic fluctuations of the parameters. Profile-likelihood ratios are determined for each of these pseudoexperiment exactly as for data. The ensemble giving the broadest PLR distribution is chosen to construct the final regions. For each point in the $(\beta_s, \Delta\Gamma_s)$ grid, we calculate the p-value as the fraction of pseudo-experiments from this ensemble in which a PLR value as large or larger than in data is observed. The $(\beta_s, \Delta\Gamma_s)$ region where the p-value is larger than $1 - x$ forms the $x\%$ CL region.

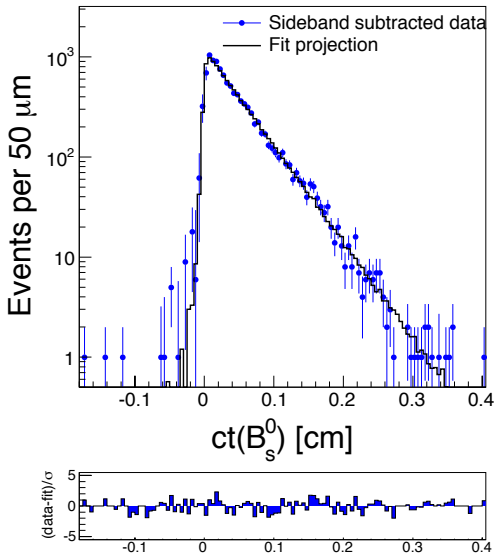
In practice we observe that the PLR distributions is fairly independent of the probed value of $(\beta_s, \Delta\Gamma_s)$, so we don't need to generate pseudo-experiments for each $(\beta_s, \Delta\Gamma_s)$ point. It suffices to compare the PLR observed in data for each point to the PLR distributions generated in a single point. Because the main goal of this analysis is to quantify compatibility of our data with the SM expectations, we choose the SM value $(\beta_s = 0.02, \Delta\Gamma_s = 0.090\,\mathrm{ps}^{-1})$ to generate the reference PLR distribution. An idea of the deviation of the observed PLR distribution from the expected χ^2_2 distribution is shown in Fig. 8.7 (a). Including the coverage adjustment and the effect of systematic uncertainties we need to change the value of PLR by approximately 2 units in order that projections on the $(\beta_s, \Delta\Gamma_s)$ plan contains the true values with 95% CL, compared with the nominal value of 5.99 for the ideal χ^2_2 statistic.

8.2 Results of the β_s-floating fit

Before proceeding with the extraction of the confidence regions and intervals, we present the results of the β_s-floating fit, which are reported in Tab. 8.1. They are perfectly consistent with the results of the β_s-fixed fit (Tab. 7.4), as expected. The minimum of the likelihood is found to be at $\beta_s = (0.11 \pm 0.12)$, in agreement with the SM expectation. The decay difference, $\Delta\Gamma_s = (0.068 \pm 0.027)\,\mathrm{ps}^{-1}$, is also found consistent with the SM prediction. The projections of the β_s-floating fit are shown in Fig. 8.1 and Fig. 8.2.



(a)

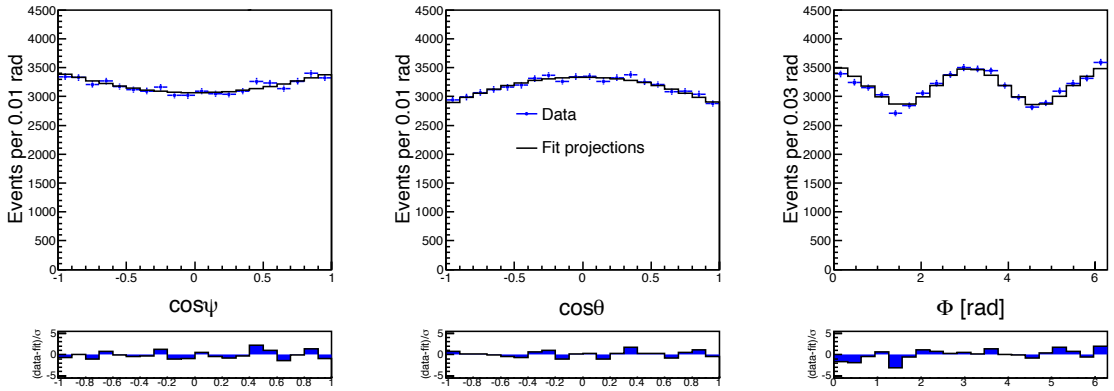


(b)

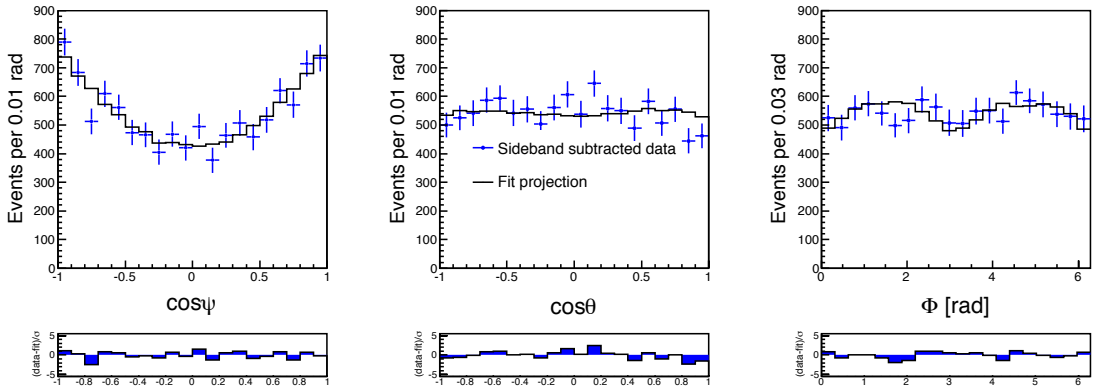
Figure 8.1: Projections onto ct distribution for the β_s-floating fit in the complete fit range (χ^2/d.o.f.=141/107) (a) and for sideband subtracted data (χ^2/d.o.f.=98/97) (b).

Parameter	Description	
β_s	CP-violating phase of B_s^0-$\bar{B}_s^0$ mixing amplitude	0.11 ± 0.12
$\Delta\Gamma_s$	lifetime difference $\Gamma_L - \Gamma_H$	$0.068 \pm 0.027\,\mathrm{ps}^{-1}$
$\alpha_\perp$	fraction $\lvert a_\perp\rvert^2$	0.279 ± 0.011
$\alpha_\parallel$	fraction $\lvert a_\parallel\rvert^2/(1-\lvert a_\perp\rvert^2)$	0.309 ± 0.012
$\delta_\perp$	$\arg(A_\perp A_0^\star)$	2.73 ± 0.53
$\delta_\parallel$	$\arg(A_\parallel A_0^\star)$	3.04 ± 0.36
$c\tau$	B_s^0 mean lifetime	$457.9 \pm 5.8\,\mu\mathrm{m}$
$\lvert A_S\rvert^2$	fraction of S-wave K^+K^- component in the signal	0.008 ± 0.016
δ_S	strong phase of S-wave amplitude	1.21 ± 0.65
Δm_s	B_s^0-$\bar{B}_s^0$ mixing frequency	$17.72 \pm 0.11\,\mathrm{ps}^{-1}$
f_s	Signal fraction	0.1721 ± 0.0018
s_m	Mass uncertainty scale factor	1.727 ± 0.017
p_1	mass background slope	$-1.89 \pm 0.42\,(\mathrm{GeV}/c^2)^{-1}$
s_{ct_1}	lifetime uncertainty scale factor 1	1.308 ± 0.012
s_{ct_2}	lifetime uncertainty scale factor 2	3.34 ± 0.13
$f_{\mathcal{R}}$	relative fraction of Gaussians of the decay-length resolution	0.852 ± 0.010
f_{pr}	fraction of prompt background	0.8839 ± 0.0039
f_{Λ_n}	fraction of background with Λ_n lifetime	0.210 ± 0.038
f_{Λ_1}	fraction of background with Λ_1 lifetime	0.718 ± 0.039
Λ_n	Inverse of lifetime of a background component	$0.040 \pm 0.032\,\mu\mathrm{m}^{-1}$
Λ_1	Inverse of lifetime of a background component	$0.0439 \pm 0.0038\,\mu\mathrm{m}^{-1}$
Λ_2	Inverse of lifetime of a background component	$0.0134 \pm 0.0010\,\mu\mathrm{m}^{-1}$
a	parameter in background fit to Φ	0.1441 ± 0.0062
b	parameter in background fit to $\cos\Theta$	0.169 ± 0.013
$S_D(\mathrm{OST})$	OST dilution scale factor	1.085 ± 0.049
$S_D(\mathrm{SSKT})$	SSKT dilution scale factor	0.87 ± 0.17
$\epsilon_b(\mathrm{OST})$	OST tagging efficiency for background	0.7592 ± 0.0019
$\epsilon_b(\mathrm{SSKT})$	SSKT tagging efficiency for background	0.7266 ± 0.0026
$A(\mathrm{OST})$	OST background tag asymmetry	0.4975 ± 0.0025
$A(\mathrm{SSKT})$	SSKT background tag asymmetry	0.4956 ± 0.0033
$\epsilon_s(\mathrm{OST})$	OST tagging efficiency for signal	0.9280 ± 0.0030
$\epsilon_s(\mathrm{SSKT})$	SSKT tagging efficiency for signal	0.5222 ± 0.0067

Table 8.1: Results of the fitting parameters for the β_s-floating fit to the entire data set.



(a)



(b)

Figure 8.2: Projections onto transversity angles distributions for the β_s-floating fit in the complete fit range (χ^2/d.o.f. = 18/19 for $\cos\Psi$, χ^2/d.o.f. = 10/19 for $\cos\Theta$, and χ^2/d.o.f. = 29/19 for Φ) (a) and for sideband subtracted data (χ^2/d.o.f. = 18/19 for $\cos\Psi$, χ^2/d.o.f. = 23/19 for $\cos\Theta$, and χ^2/d.o.f. = 14/19 for Φ) (b)

8.3 Confidence regions

We now present the confidence regions as profile-likelihood ratio in the $(\beta_s, \Delta\Gamma_s)$ plane. Because of the approximate symmetries of the likelihood, MINUIT can occasionally find a local minimum (see Sect. 6.3). Therefore, we compute two PLR values for each point of the plane. One is computed starting the minimization in the $\Delta\Gamma > 0$ and $\delta_{\|} < \pi$ (0 and π are symmetry points for those parameters) subplane, the other is computed starting the minimization in $\Delta\Gamma < 0$ and $\delta_{\|} > \pi$. We proved in simulation that this procedure allows the identification of the global minimum by comparing the value of the likelihood from the two minimizations, without imposing any constraint on the domain of fit parameters. Thus, for each point of the plane, the minimum of the values found in the two minimization procedures is chosen to construct the PLR distribution. The latter is shown in Fig. 8.3. If the PLR is considered as an ideal χ^2

statistic with 2 degrees of freedom, we set the confidence regions displayed in Fig. 8.4, which must be corrected to ensure the proper coverage. Before doing the correction, we perform some checks.

We first analyze only the first $5.2\,\text{fb}^{-1}$ of data and check the compatibility with the previous published CDF results to validate the fitting technique. Figure 8.4 (b) shows the comparison of the confidence regions obtained compared with the curves of the published analysis, before any coverage-corrections. The iso-PLR curves in the $(\beta_s, \Delta\Gamma_s)$ plan overlap and the fit gives the same minimum for $\beta_s = 0.23$ and $\Delta\Gamma_s = 0.97\,\text{ps}^{-1}$, as expected.

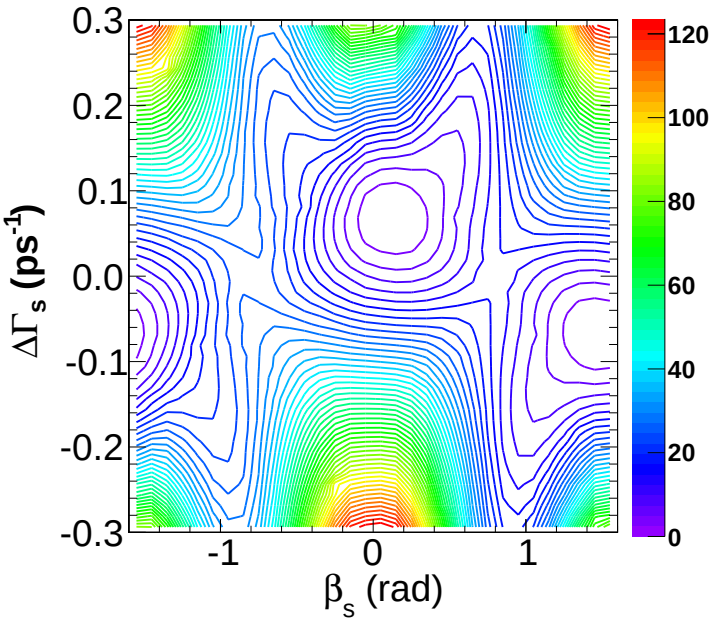
Figure 8.5 (a) shows the confidence region for the fit to only the newly added data since the published analysis using only OST, while in Fig. 8.5 (b) also the SSKT is applied (assuming same performances as in the first $5.2\,\text{fb}^{-1}$). The regions in Fig. 8.5 (a) are wider than the ones in Fig. 8.5 (b) because of the limited tagging power of the OST with respect to the SSKT, which doesn't suffice to completely brake the symmetry of the untagged likelihood, yielding the 4 equivalent minima to enlarge and spread the confidence regions. The results of the β_s-floating fit are in agreement with the estimations previously found in the first $5.2\,\text{fb}^{-1}$ of data [18]. The main physical parameter are reported in Tab. 8.2 for comparison. We see that the mixing phase in later part of data is consistent with the SM expectation and the minimum of the likelihood in the β_s subspace is very close to zero.

Parameter	Previous result ($5.2\,\text{fb}^{-1}$)	New data only ($4.4\,\text{fb}^{-1}$)
β_s	0.24 ± 0.13	0.09 ± 0.25
$\Delta\Gamma_s$ [ps^{-1}]	0.097 ± 0.035	0.029 ± 0.033
$\alpha_\perp$	0.264 ± 0.014	0.297 ± 0.013
$\alpha_\parallel$	0.307 ± 0.015	0.309 ± 0.014
$\delta_\perp$	3.03 ± 0.52	3.50 ± 0.84
$\delta_\parallel$	3.02 ± 0.47	3.11 ± 0.40
$c\tau$ [μm]	459.0 ± 7.3	456.4 ± 6.7

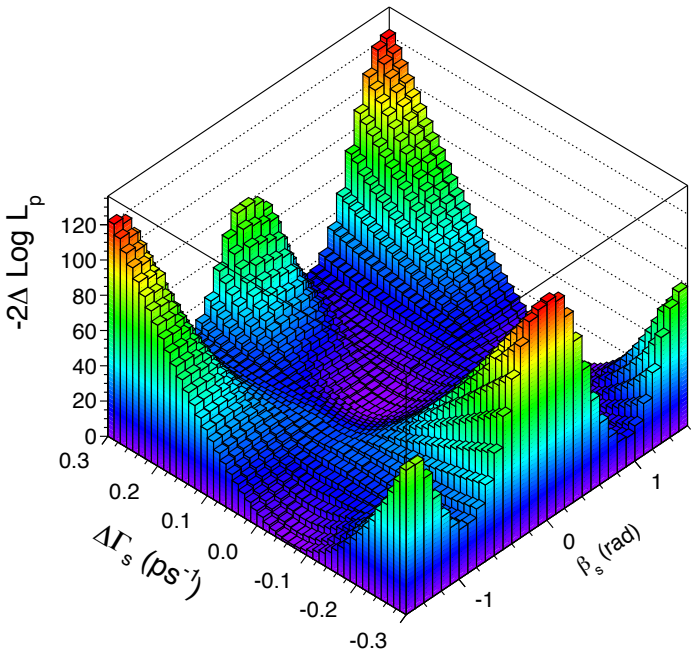
Table 8.2: Summary of the results of the main physical parameters in the β_s-floating fit to newly data added (rightmost column). We report also results of previous CDF measurement (center column).

Figure 8.6(a) shows the comparison between the PLR distribution for the default fit and the fit where the SSKT is used in all the data, assuming that the performances of the tagger do not change in the second part of the data. The resulting confidence regions are very similar, confirming that our choice of not using an uncalibrated version of SSKT in the second half of data (Sect. 4.3) does not significantly compromise the quality of the result.

As a further control check, we compute the PLR using two different OST scale factors, one for B_s^0 and the other for $\bar{B}_s^0$, to allow for a tagging asymmetry. This was the used procedure in the previous iteration of the analysis, since in the past calibration of the OST, the scale

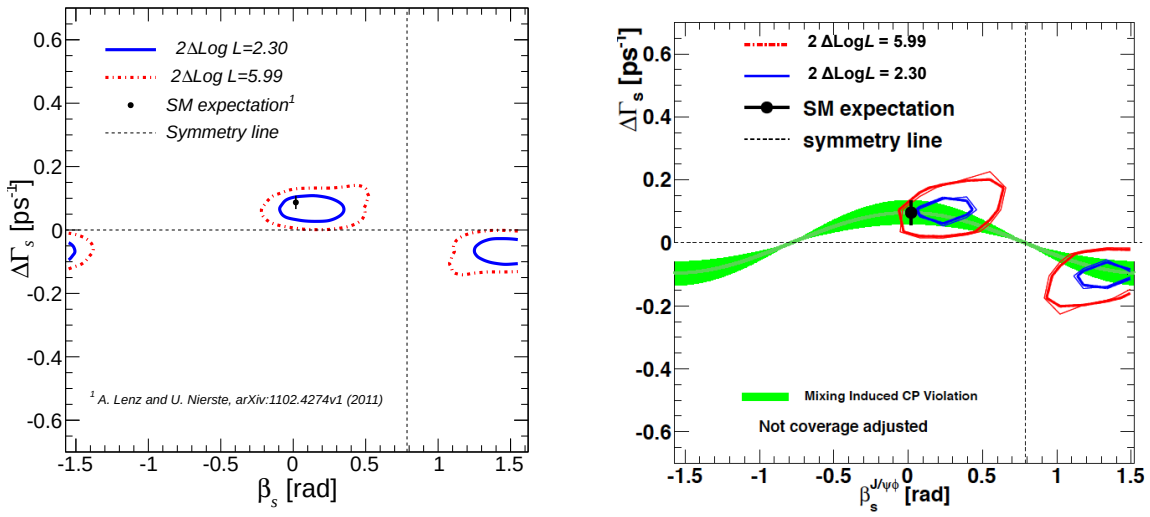


(a)



(b)

Figure 8.3: Contours of iso-PLR (a) and two-dimensional distribution (b) in the $(\beta_s, \Delta\Gamma_s)$ plane.

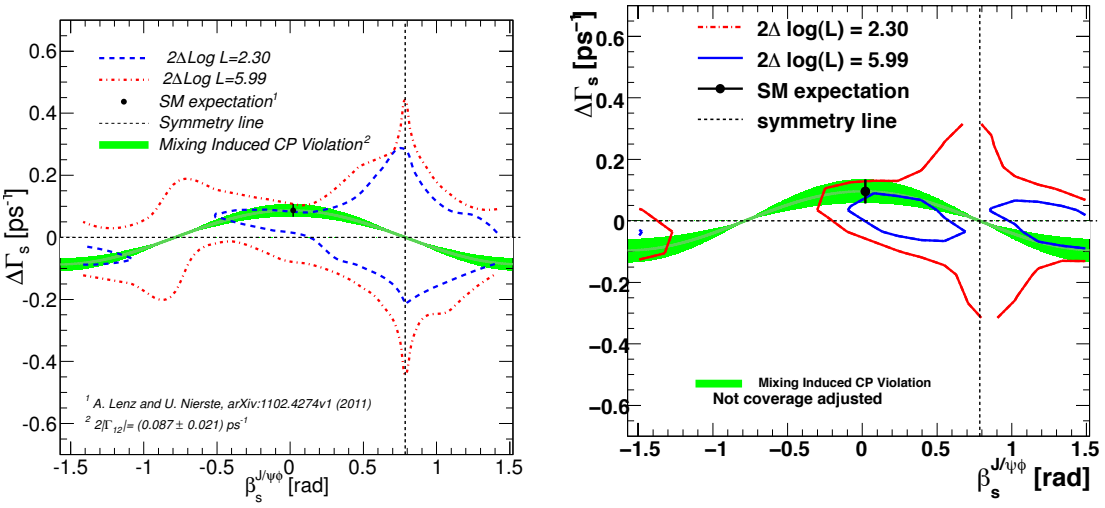


(a) (b)

Figure 8.4: Confidence regions in the $(\beta_s, \Delta\Gamma_s)$ plane before any coverage adjustement for the analysis of the full data set (a). Comparison of the iso-PLR curves before corrections between the previous CDF analysis (light contours) and the new analysis run in the first $5.2\,\text{fb}^{-1}$ of the data (bold contours) (b).

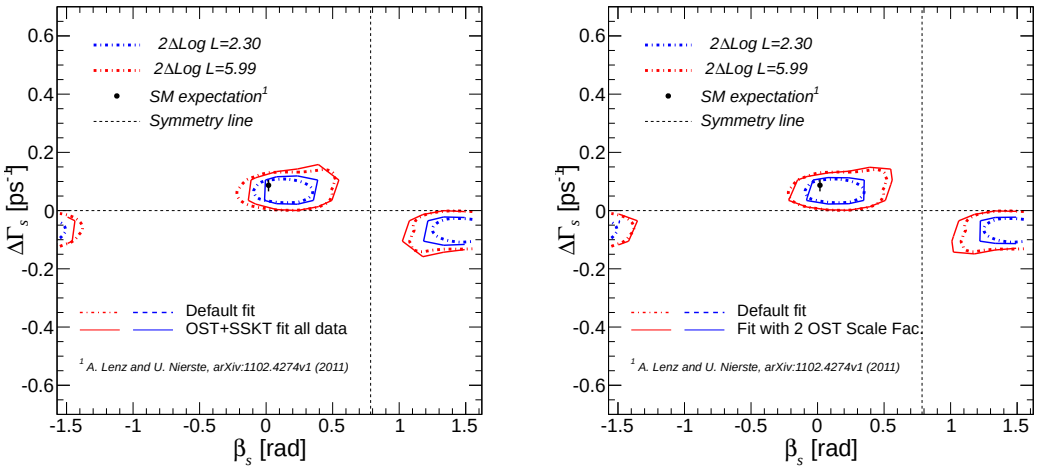
factors for the B^+ and the B^- mesons were found to be different by 2 standard deviations, in contrast to the more precise calibration with the whole data set reported in Sect. 4.2.1. The two scale factors are Gaussian-constrained in the fit by the values measured in the calibration: $S^+ = 1.09 \pm 0.05$ and $S^- = 1.08 \pm 0.05$. The results of the fits are in perfect agreement with respect to the fit with a single scale factor. Moreover, the two scale factors are found to be consistent, giving an *a posteriori* check of our choice. They are: $S^+ = 1.086 \pm 0.048$ and $S^- = 1.083 \pm 0.048$. The PLR obtained is shown in Fig. 8.6(b) in comparison with the likelihood profile of the default fit. The confidence regions present a perfect overlap.

The map between the PLR and the p-value obtained with simulation is shown in Fig. 8.7(a). The solid black line represents PLR for pseudo-experiment generated with the nuisance parameters set to the values measured in data. The colored dotted lines are PLR distributions for pseudo-experiments introducing variation of nuisance parameters for systematics. The worst case is chosen to obtain the final confidence regions. A PLR of 3.56 must be set to guarantee the correct coverage for a 68% CL, while a PLR of 8.13 corresponds to 95% CL. The adjusted confidence regions are reported in Fig. 8.7(b). The SM point has a p-value of 0.59 indicating full agreement of CDF data with the SM hypothesis.



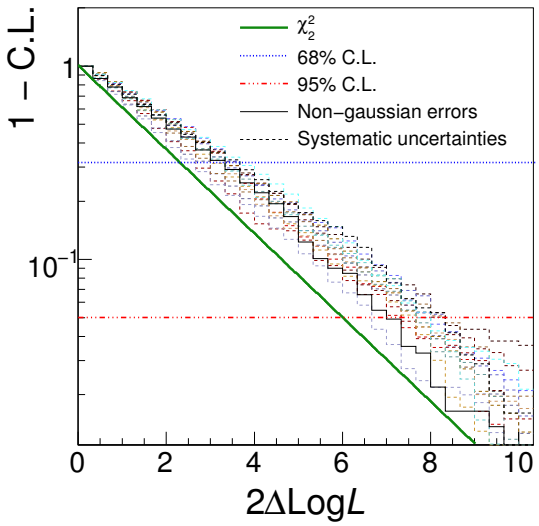
(a) (b)

Figure 8.5: Curves of iso-PLR in the $(\beta_s, \Delta\Gamma_s)$ plane (no coverage adjustment) from a fit to the newly added data only (last $4.4\,\text{fb}^{-1}$). Only OST is applied (a), OST and SSKT are applied (b).

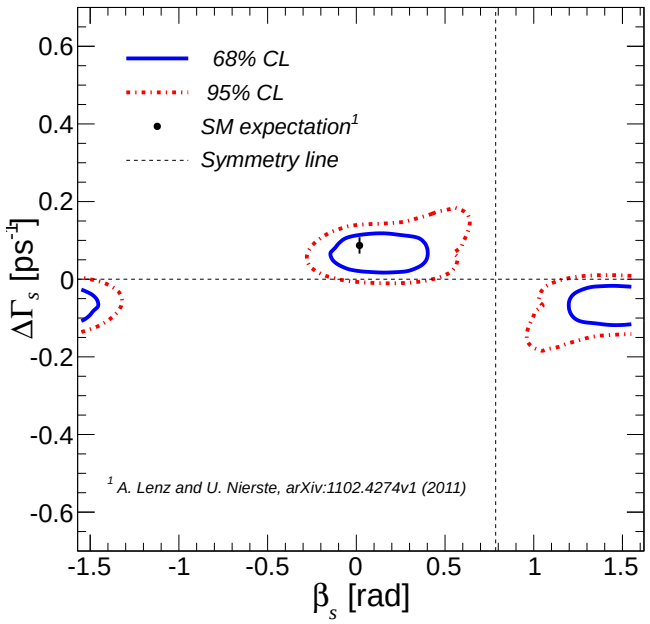


(a) (b)

Figure 8.6: Comparison of curves of iso-PLR in the $(\beta_s, \Delta\Gamma_s)$ plane (no coverage adjustment) for the central fit and a fit in which SSKT is used in all data, assuming constant performances of the tagger in time (a). Comparison between the iso-PLR done for the dafault fit and the fit where two different OST scale factors are used for B_s^0 and for $\bar{B}_s^0$ (b).



(a)



(b)

Figure 8.7: The p-value (1−CL) as a function of PLR for the coverage adjustment of $(\beta_s, \Delta\Gamma_s)$ confidence regions (a). The green curve is the cumulative of the χ^2 statistic corresponding to the ideal case. The solid black line represents PLR for pseudo-experiment generated with the nuisance parameters set to the values measured in data. The colored dotted lines are PLR distributions for pseudo-experiments introducing variation of nuisance parameters for systematics. The worst case is chosen to draw coverage-adjusted confidence regions in (b) for the 68% CL and 95% CL.

8.4 Confidence intervals

The PLR distributions is evaluated also as a function of β_s only by the same method previously described and treating $\Delta\Gamma_s$ as a nuisance parameter. We divide the interval $-\pi/2 < \beta_s < \pi/2$ in 100 equally-spaced bins, and we fit the data by floating all parameters but β_s, which are fixed to the values corresponding to the point to probe. In order to calculate the PLR, we subtract to the value of the $-2\ln\mathcal{L}(\beta_s, \hat{\boldsymbol{\zeta}})$ found in each point of the interval the value $-2\ln\mathcal{L}(\hat{\beta}_s, \hat{\boldsymbol{\zeta}})$ at the global minimum with all parameters floating.

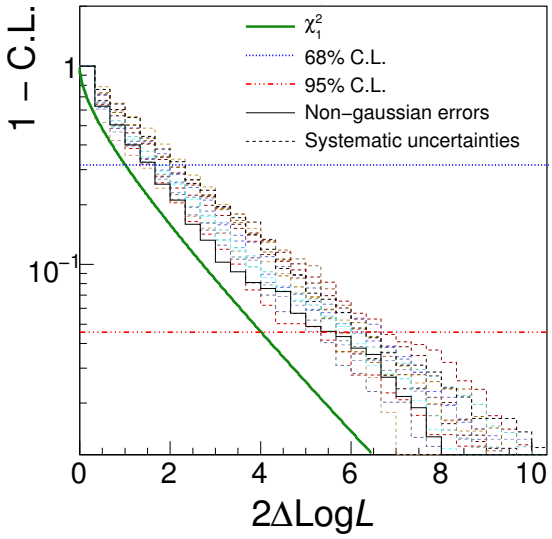
Following a procedure analogous to the two-dimensional case, we extracted the correspondence between the PLR cumulative distribution and p-values, and used it to ensure right coverage properties to our intervals (see Fig. 8.8(a)). After the adjustment, the PLR corresponding to a 68% CL interval and to a 95% CL interval are respectively 2.12 and 6.62, instead of 1 and 4 of the ideal χ^2_1 statistic. This broadens the intervals of about 40%. They are reported in Fig. 8.8(b). We found:

$$\begin{aligned} \beta_s &\in [-\pi/2, -1.51] \cup [-0.06, 0.30] \cup [1.26, \pi/2] \text{ at } 68\% \text{ CL}, \\ \beta_s &\in [-\pi/2, -1.36] \cup [-0.21, 0.53] \cup [1.04, \pi/2] \text{ at } 95\% \text{ CL}. \end{aligned} \tag{8.2}$$

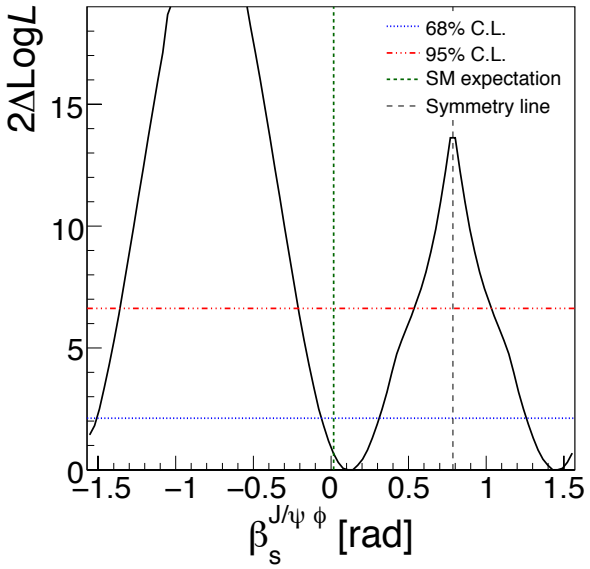
Using the recent determination of the sign of $\Delta\Gamma_s$ [42] which resolve the ambiguity of the solutions, we choose the intervals which include the SM expectation: $\beta_s \in [-0.06, 0.30]$ at 68% CL, and $\beta_s \in [-0.21, 0.53]$ at 95% CL. The SM point has a p-value of 0.54 indicating full agreement of CDF data with SM hypothesis. Table 8.3 compare the new and the previous intervals: the updated analysis has a 35% improvement of the 68% CL interval.

CL	$5.2\,\text{fb}^{-1}$	$9.6\,\text{fb}^{-1}$
68%	$[0.02, 0.52]$	$[-0.06, 0.30]$
95%	$[-0.11, 0.65]$	$[-0.21, 0.53]$
SM p-value	0.30	0.54

Table 8.3: Comparison of the new and previous intervals of β_s for the solution closest to the SM expectation.



(a)



(b)

Figure 8.8: The p-value $(1 - \mathrm{CL})$ as a function of PLR for the coverage adjustment of β_s intervals (a). The green curve is the cumulative of the χ^2 statistic corresponding to the ideal case. The solid black line represents PLR for pseudo-experiment generated with the nuisance parameters set to the values measured in data. The colored dotted lines are PLR distributions for pseudo-experiments introducing variation of nuisance parameters for systematics. The worst case is chosen to set coverage-adjusted confidence intervals displayed in (b) as the intersection of the PLR of data with the blue line for the 68% CL, and with the red line for the 95% CL.

Chapter 9

$B_s^0 \to \phi\phi$ Data Set

This chapter describes the data samples for the analysis of $B_s^0 \to \phi\phi$ decays. We first present the selection of the CDF data sample, outlining the system used to trigger on events with a displaced vertex from the $p\bar{p}$ interaction point. An offline selection is applied to reduce the background in the final sample for measuring the *CP*-violating asymmetries. Finally, we describe the MC simulated data samples used in the analysis.

9.1 Trigger on displaced tracks

In this section we describe that part of the CDF II trigger system which provides the ability to recognize tracks consistent with decay of long-lived hadrons, decaying in displaced vertexes with respect to the primary $p\bar{p}$ interaction point.[1] Such trigger is crucial to collect samples of hadronic B decays, such as $B_s^0 \to \phi(\to K^+K^-)\phi(\to K^+K^-)$ decays.

When a long-lived particle decays after traveling some distance from the point where it was created, the trajectories of the decay products generally do not point back to the collision point. The distance of closest approach of the extrapolated trajectory to the collision point, the impact parameter d, is the key quantity for the detection of secondary vertices. A typical situation is sketched in Fig. 9.1. When a long-lived particle is created in the primary collision, it travels some distance L before decaying, and will generate secondary tracks originating from a secondary vertex. Then, the impact parameter is $d = L \sin\theta$, where θ is the angle between the direction of the secondary track and the direction of the parent particle, and $L = \beta\gamma ct$, where β and γ are the usual parameters of the Lorentz transformation from the rest frame of the parent particle to the laboratory, t is the proper decay time, and c is the speed of light. If the particles involved in the process are all relativistic, then on average, θ is of the order of $1/\gamma$, and for $\gamma \gg 1$, the impact parameter is

$$\langle d \rangle \simeq \beta\gamma c\tau(1/\gamma) = \beta c\tau \simeq c\tau, \tag{9.1}$$

where τ is the lifetime of the parent particle. For particles containing the b quark, $c\tau \approx 450\,\mu$m, and for particles containing the c quark, $c\tau$ can be as small as $125\,\mu$m. These values set the scale for the precision of the impact-parameter measurement needed to detect secondary vertices from heavy-flavor decay to the order of a few tens of micrometers. We describe now how this is achieve in the CDF II detector.

[1] For a detailed description we point to Ref. [96].

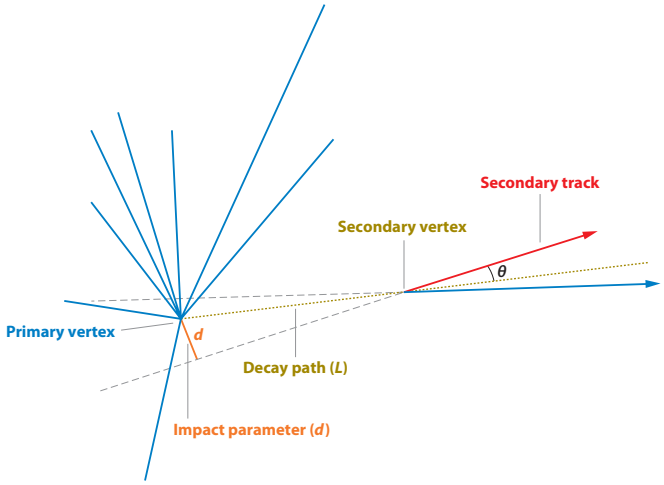


Figure 9.1: When a long-lived particle decays after traveling some distance, the trajectories of the decay products do not point back to the collision point. The distance of closest approach of the extrapolated trajectory to the collision point is the impact parameter.

Using information from the COT, at L1, the *eXtremely Fast Tracker* (XFT) [126, 127] reconstructs trajectories of charged particles in the (r, ϕ) plane for each proton-antiproton bunch crossing. The XFT is a custom processor that uses pattern matching to first identify short segments of tracks and then to link them into full-length tracks. After classifying the hits of the four axial super-layers in *prompt* (0–66 ns) or *delayed* hits (67–220 ns), depending upon the observed drift-time within the cell, track segments are reconstructed in each axial super-layer. A pattern-matching algorithm searches for coincidences between the observed combinations of hits in each super-layer (a minimum of 11 out of 12 hits is required) and a set of predetermined patterns. If a coincidence between segments crossing four super-layers is found, two-dimensional XFT-tracks are reconstructed by linking the segments. The segments are compared with a set of about 2 400 predetermined patterns corresponding to all tracks with $p_T \gtrsim 1.5\,\text{GeV}/c$ originating from the beam-line. The comparison proceeds in parallel in each of the 288 azimuthal 1.25° sectors in which XFT logically divides the chamber. If no track is found using all four super-layers, then the best track found in the innermost three super-layers is output. The track-finding efficiency and the fake-rate with respect to the offline tracks depend on the instantaneous luminosity and were measured to be about 96% and 3%, respectively, at $\mathscr{L} \simeq 10^{31}\,\text{cm}^{-2}\text{s}^{-1}$. The observed momentum resolution is $\sigma_{p_T}/p_T^2 \approx 1.7\%\,(\text{GeV}/c)^{-1}$, and the azimuthal resolution is $\sigma_{\phi_6} \approx 0.3°$, where ϕ_6 is the azimuthal angle of the track measured at the sixth COT super-layer, located at 106 cm radius from the beam-line. Events are selected for further processing when two XFT-tracks satisfying trigger criteria on basic variables are found. The variables are the product of any combination of two particles' charges (opposite or same sign), the opening angle of the two tracks in the transverse plane ($\Delta\phi_6$), the two particles' transverse momenta and their scalar sum.

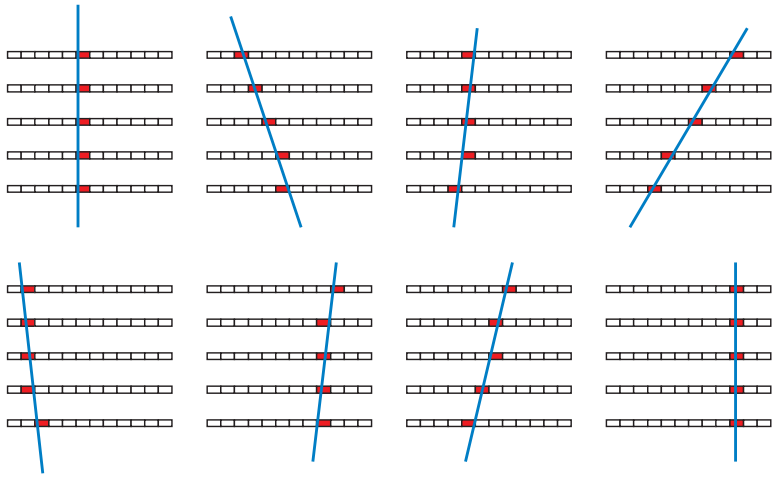


Figure 9.2: Schematic illustration of combinations of super-bins (in the transverse plane) corresponding to the passage of charged particles in five radial silicon layers.

At L2 the information from the SVX II detector is incorporated into the trigger track reconstruction by the *Silicon Vertex Trigger* (SVT) [95, 96]. Charge clusters in the silicon, which SVT finds by converting a list of channel numbers and pulse heights into charge-weighed hit centroids, are used by a pattern recognition algorithm, which is formed of two subsequent stages. First, a low-resolution stage is implemented by grouping together adjacent detector channels into *super-bins*. Their width in the azimuthal direction is programmable, with 250 μm typical values. A set containing about 95% of all super-bin combinations compatible with the trajectory of a charged particle with $p_T \gtrsim 2\,\mathrm{GeV}/c$ originated from the beam-line (*patterns*) is calculated in advance from simulation and stored in a special design memories called *Associative Memories* (AMs). For each of $12(\phi) \times 6(z) = 72$ wedges in SVX II, the $\approx 5\,500$ most probable patterns are stored. Online, an algorithm detects low-resolution candidate tracks, called *roads*, by matching super-bins containing hits with the stored patterns. A road is a combination of five super-bins in different SVX II layers plus the XFT track parameters, which are logically treated as additional hits (see Fig. 9.2).

In the AMs system, maximum parallelism is exploited to speed-up the processing, using a working principle similar to the one of the bingo game: while the silicon hits are being read out, each player marks the matching super-bins on his score-card; each bingo corresponds to a road and is retained for further processing. A maximum of 64 roads per event, each one having a maximum of 8 hits per super-bin, is output. At this stage, pattern recognition is done during detector read-out with no additional processing time. The resolution is coarse enough to reduce the fraction of accidental combinations, but fine enough to separate most tracks. Once a track is confined to a road, most of the pattern recognition is done, leaving the remaining ambiguities, as multiple hits in the same super-bin, to the stage of track fitting.

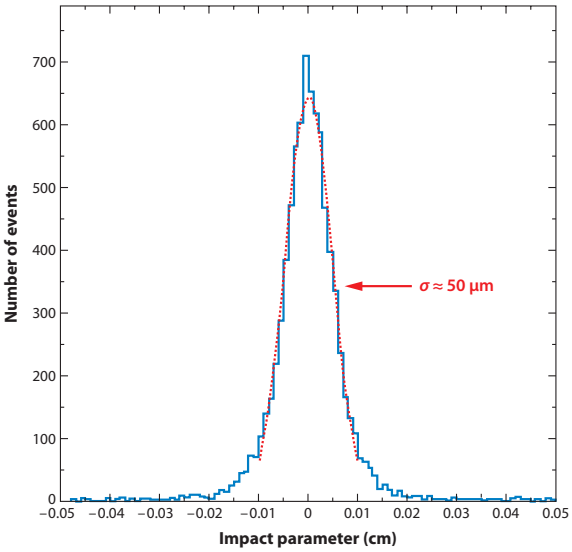


Figure 9.3: Impact parameter distribution as measured by the SVT. The width of the distribution is interpreted as the convolution of the actual transverse size of the beam spot ($\approx 25\,\mu$m) with the impact-parameter resolution of the SVT ($\approx 43\,\mu$m).

In principle, no exact linear relation exists between the transverse parameters of a track in a solenoidal field, and the coordinates at which the track intersects a radial set of flat detector planes. But for $p_T \gtrsim 2\,\text{GeV}/c$ and $|d| \lesssim 1\,$mm a linear fit biases the reconstructed d by at most a few percent. The track-fitting process exploits this feature by expanding the non-linear constraints and the parameters of the real track to first order with respect to the reference track associated to each road. A linear expansion in the hit positions of both the track parameters and the χ^2 is used. The fit process is thus reduced to computing a few scalar products, which is done within 250 ns or less than 10 clock cycles. The needed constants, which depend on detector geometry and alignments, are evaluated in advance and stored in an internal memory. The output of the SVT are the reconstructed parameters of the two-dimensional track in the transverse plane: ϕ_0, p_T and d. The list of parameters for all found tracks is sent to L2 for trigger decision.

The SVT measures the impact parameter with a r.m.s. width $\sigma_{d^{\text{SVT}}} \approx 43\,\mu$m, with an average latency of $24\,\mu$s, $9\,\mu$s of which being spent waiting for the start of the read-out of silicon data. This resolution is comparable with the offline one, for tracks not using L00 hits, and yields a distribution of impact parameter of prompt tracks with respect to the z axis with $\sigma_{d^{\text{SVT}}} \approx 50\,\mu$m when combined with the transverse beam-spot size, as shown in Fig. 9.3. The SVT efficiency is higher than 85%.[2] The impact parameter is a quantity measured with respect to the beam. If the actual beam position in the transverse plane is shifted by an amount

[2]This efficiency is defined as the ratio between the number of tracks reconstructed by SVT and all XFT-matched offline silicon tracks that are of physics analysis quality.

d_{beam} with respect to the origin of the SVT reference frame, all prompt tracks appear to SVT as having $\mathcal{O}(d_{\text{beam}})$ impact parameters. This is relevant since the beam is usually displaced from its nominal $(0, 0, z)$ position. Between Tevatron stores, $\mathcal{O}(500\,\mu\text{m})$ displacements in the transverse plane and $\mathcal{O}(100\,\mu\text{rad})$ slopes with respect to the detector axis may occur. In addition, the beam can drift by $\mathcal{O}(30\,\mu\text{m})$ in the transverse plane even during a single store. However, a simple geometric relation prescribes that the impact parameter of a track, calculated with respect to a point displaced from its production vertex, is a sinusoidal function of its azimuthal coordinate:

$$d = y_0 \cos\phi_0 - x_0 \sin\phi_0, \tag{9.2}$$

where (x_0, y_0) are the coordinates of the production vertex. Using Eq. (9.2), SVT measures the actual coordinates of the beam position with respect to the detector system and subtracts them from the measured impact parameters, in order to provide physical impact parameters. Using about 10^5 tracks every 30 seconds, six transverse beam positions (one for each SVX II semi-barrel) are determined online. The six samplings (one for each SVX II barrel) along the z direction provide a measurement of the slope of the beam with respect to the nominal z axis. For the proper measurement of impact parameters, the beam slope is more harmful than the transverse drift, because it breaks the cylindrical symmetry of the system. The SVT does not have access to the z_0 coordinate of tracks. For each track, only the longitudinal coordinate of the SVX II half-barrel that detected the track is known. But half-barrels are too long (16 cm) to allow for a reliable correction of the beam slope. When significant slopes are observed, the Tevatron beam division is alerted and they apply a corrective action on the magnets.

The L2 trigger selections used in this analysis typically requires two SVT-tracks with impact parameter greater than $100\,\mu$m and smaller than 1 mm. In addition, the L2 trigger requires the transverse decay length, L_{xy}, to exceed $200\,\mu$m, where L_{xy} is calculated as the projection of the vector from the primary vertex to the two track vertex in the transverse plane along the vectorial sum of the transverse momenta of the tracks. The trigger based on SVT collects large quantities of long-lived B hadrons, rejecting most part of the prompt background. However, through its impact-parameter-based selection, the SVT trigger also biases the observed proper decay time distribution. This has important consequences on time-dependent analysis of hadronic B decays.

The L3 trigger uses a full reconstruction of the event with all detector information, but uses a simpler tracking algorithm and preliminary calibrations relative to the ones used offline and retests the criteria imposed by L2. In addition, the difference in z of the two tracks at the point of minimum distance from the primary vertex, Δz_0, is required not to exceed 5 cm, removing events where the pair of tracks originate from different collisions within the same crossing of p and $\bar{p}$ bunches.

Over the course of a Tevatron store, the available trigger bandwidth varies because trigger rates fall as instantaneous luminosity falls. Higher trigger rates at high luminosity arise from both a larger rate for real physics processes as well as a larger fake trigger rate due to multiple $p\bar{p}$ interactions. To fully exploit the available trigger bandwidth, we employ three main variants

Version	Level 1	Level 2	Level 3
High-p_T	$p_T > 2.5\,\text{GeV}/c$	$p_T > 2.5\,\text{GeV}/c$	$p_T > 2.5\,\text{GeV}/c$
	Opposite charge	Opposite charge	Opposite charge
	$\Delta\phi_6 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$
	$\sum p_T > 6.5\,\text{GeV}/c$	$\sum p_T > 6.5\,\text{GeV}/c$	$\sum p_T > 6.5\,\text{GeV}/c$
		$0.1 < d < 1\,\text{mm}$	$0.1 < d < 1\,\text{mm}$
		$L_{xy} > 200\,\mu\text{m}$	$L_{xy} > 200\,\mu\text{m}$
			$\lvert\Delta z_0\rvert < 5\,\text{cm}$
			$\lvert\eta\rvert < 1.2$
Medium-p_T	$p_T > 2\,\text{GeV}/c$	$p_T > 2\,\text{GeV}/c$	$p_T > 2\,\text{GeV}/c$
(or scenario A)	Opposite charge	Opposite charge	Opposite charge
	$\Delta\phi_6 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$
	$\sum p_T > 5.5\,\text{GeV}/c$	$\sum p_T > 5.5\,\text{GeV}/c$	$\sum p_T > 5.5\,\text{GeV}/c$
		$0.1 < d < 1\,\text{mm}$	$0.1 < d < 1\,\text{mm}$
		$L_{xy} > 200\,\mu\text{m}$	$L_{xy} > 200\,\mu\text{m}$
			$\lvert\Delta z_0\rvert < 5\,\text{cm}$
			$\lvert\eta\rvert < 1.2$
Low-p_T	$p_T > 2\,\text{GeV}/c$	$p_T > 2\,\text{GeV}/c$	$p_T > 2\,\text{GeV}/c$
	$\Delta\phi_6 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$	$2^\circ < \Delta\phi_0 < 90^\circ$
	$\sum p_T > 4\,\text{GeV}/c$	$\sum p_T > 4\,\text{GeV}/c$	$\sum p_T > 4\,\text{GeV}/c$
		$0.1 < d < 1\,\text{mm}$	$0.1 < d < 1\,\text{mm}$
		$L_{xy} > 200\,\mu\text{m}$	$L_{xy} > 200\,\mu\text{m}$
			$\lvert\Delta z_0\rvert < 5\,\text{cm}$
			$\lvert\eta\rvert < 1.2$

Table 9.1: Selections for the three versions of the displaced-tracks trigger used in this analysis. The criteria refer to track-pairs. The p_T, d and η requirements are applied to both tracks. The $\sum p_T$ refers to the scalar sum of the p_T of the two tracks.

of the displaced-tracks trigger. The three selections are summarized in Tab. 9.1 and are referred to as the low-p_T, medium-p_T (called also *scenario A* in CDF jargon) and high-p_T selections, according to their requirements on minimum transverse momentum. At high luminosity, the higher purity, but less efficient, high-p_T selection is employed. As the luminosity decreases over the course of a store, trigger bandwidth becomes available and the other selections are utilized to fill the available trigger bandwidth and maximize the B yield. The rates are controlled by the application of prescaling, which rejects a predefined fraction of events accepted by each trigger selection dependent on the instantaneous luminosity.

9.2 Offline selection

We use the data collected by the displaced-tracks trigger in a period starting from March 2001 till April 2008, which corresponds to an integrated luminosity of $2.9\,\text{fb}^{-1}$. Such data sample was selected for the update of the measurement of the branching ratio of the $B_s^0 \to \phi\phi$ decay in 2009 and for the first measurement of its polarization amplitudes in 2010 [2]. Given the marginal increase in the sample size due to the severe suppression of the displaced-tracks trigger in the latest part of the data taking, we choose to not add new data, of which handling and validation would have slowed the release of the results.

We reconstruct B_s^0 mesons from BSTNTUPLE (Sect. 3.2.7) by first forming $\phi(1020) \to K^+K^-$ candidates decays from opposite-sign track pairs and with mass within $15\,\text{MeV}/c^2$ of the known $\phi(1020)$ mass. We form $B_s^0 \to \phi\phi$ candidates by fitting to a single vertex the $\phi\phi$ candidates pairs. At least one pair of tracks in the B_s^0 candidate must satisfy the trigger requirements. In the reconstruction of the decay we refine the calculation of the variables of the events, sketched in Fig. 9.4, such as the momenta of the particle, the decay-length in the transverse plane, L_{xy}, and the impact parameter corresponding to each particle. We require at this stage two basic loose cuts to reduce the downstream: $L_{xy} > 200\,\mu\text{m}$ and that the transverse momentum $p_T > 5\,\text{GeV}/c$, for each B_s^0 candidate.

The momenta of the daughter particles are used to calculate the helicity angles defined in Sect. 2.1.1. Starting from four-momenta of the kaons, we derive four-momenta of the ϕ particles and boost all particles in the B_s^0 rest frame, where the direction of the momenta of $\phi_{1(2)}$ defines the unit vector $\hat{\boldsymbol{z}}_{1(2)}$. We boost each kaon pairs in the corresponding ϕ rest frame, where we calculate the polar angle of the $K^+_{1(2)}$ particle with respect to the $\hat{\boldsymbol{z}}_{1(2)}$ vector of its mother particle, $\cos\vartheta_1$ and $\cos\vartheta_2$. The vector product of the momentum of the charged kaon, $\boldsymbol{p}(K^+_{1(2)})$, with $\hat{\boldsymbol{z}}_{1(2)}$ gives the unit vector $\hat{\boldsymbol{n}}_{1(2)}$ orthogonal to each ϕ decay plane:

$$\hat{\boldsymbol{n}}_{1(2)} = \frac{\boldsymbol{p}(K^+_{1(2)}) \times \hat{\boldsymbol{z}}_{1(2)}}{|\boldsymbol{p}(K^+_{1(2)}) \times \hat{\boldsymbol{z}}_{1(2)}|}. \tag{9.3}$$

The angle between $\hat{\boldsymbol{n}}_1$ and $\hat{\boldsymbol{n}}_2$ is the angle between the two decay planes, *i. e.*, the third helicity angle φ.

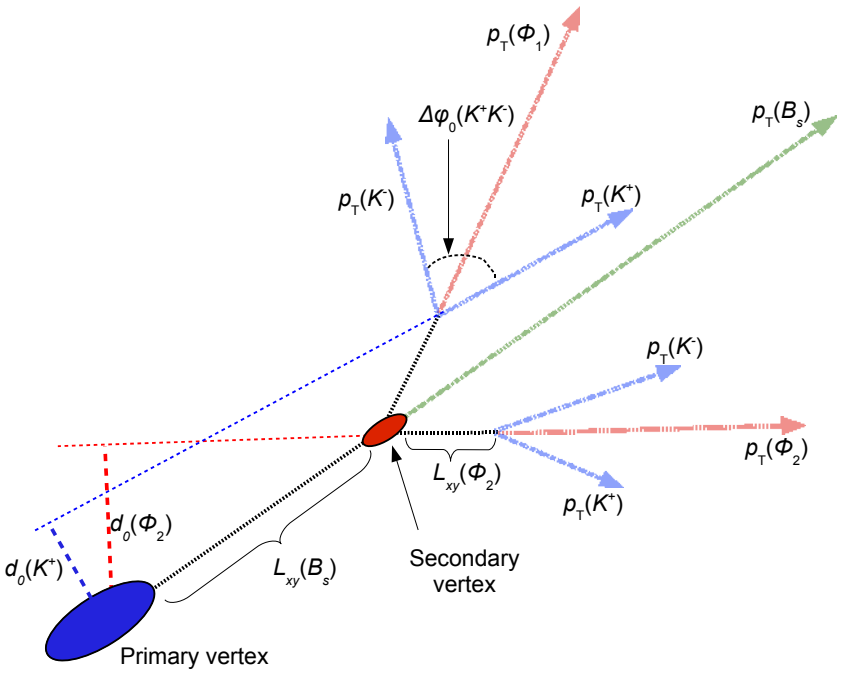


Figure 9.4: Sketch of a $B_s^0 \to \phi\phi$ decay in the transverse plane. Ellipses indicate vertexes, arrows indicate the transverse momenta (*i. e.* the direction) of charged particles. Nothing is to scale.

Figure 9.5 shows the distribution of the $K^+K^-K^+K^-$ invariant mass after the trigger selection only. No evident signal peak is recognizable in the B mass window and a dedicated offline selection to suppress the background is in order. A fast and reliable method to evaluate the resolution expected from a measurement is provided by the Minimum Variance Bound (MVB) [128, 129]. Given the data, the MVB provides an upper bound to the precision that can be achieved on a parameter, whatever the estimation procedure used. In the simplified case of a counting experiment to determine the number S of signal events within a total number of $S+B$ events (where B is the number of background events), the expected statistical resolution σ on the signal yield, estimated with the MVB, is given by:

$$\frac{1}{\sigma} \propto \frac{S}{\sqrt{S+B}} \tag{9.4}$$

The optimal selection would be one that maximizes Eq. (9.4): this expression, which is rigorously valid for a counting experiment, is still sufficiently accurate in the case of a likelihood fit of a continuous distribution, as we will do to extract the signal yields for computing the asymmetries in our analysis. The maximization depends on the selection requirements for a set of variables which are the most representative of a signal candidate. In our case the event variables exploited are:

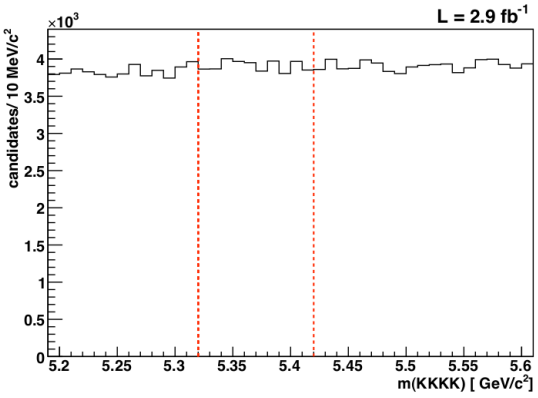


Figure 9.5: The $K^+K^-K^+K^-$ spectrum after the trigger selection with the requirements $L_{xy}(B_s^0) > 200\,\mu$m and $p_T(B_s^0) > 5\,\text{GeV}/c$ only. The red dotted lines represent bounds of the region around the B_s^0 mass.

- $L_{xy}(B)$, the transverse decay-length of the reconstructed B_s^0;
- $d(B)$, the impact parameter of the reconstructed B_s^0;
- $d(\phi_{\text{max}})$, the impact parameter of the ϕ with higher momentum;
- $p_T^{\text{min}}(K)$, the transverse momentum of the softer kaon;
- χ^2_{xy}, the χ^2 of the fit used in the reconstruction of the secondary vertex.

The distributions of these variables are shown in Fig. 9.6. The optimization of the selection requirements was developed in the context of our previous work on the measurement of the $B_s^0 \to \phi\phi$ decays branching ration [130]. An optimized selection has been studied using signal MC events (described later in Sect. 9.4) and a model for background obtained from mass sidebands, 5.02–5.22 GeV/c^2 and 5.52–5.72 GeV/c^2, in order not to introduce bias from using events in the signal region. The optimized requirements for the extraction of the signal that maximize Eq. (9.4) are reported in Tab. 9.2.

variable	requirement
$L_{xy}(B)$	$> 330\,\mu$m
$p_T^{\text{min}}(K)$	$> 0.7\,\text{GeV}/c$
χ^2_{xy}	< 17
$d(B)$	$< 65\,\mu$m
$d(\phi_{\text{max}})$	$> 85\,\mu$m

Table 9.2: Requirements used in the optimized selection of the $B_s^0 \to \phi\phi$ dataset.

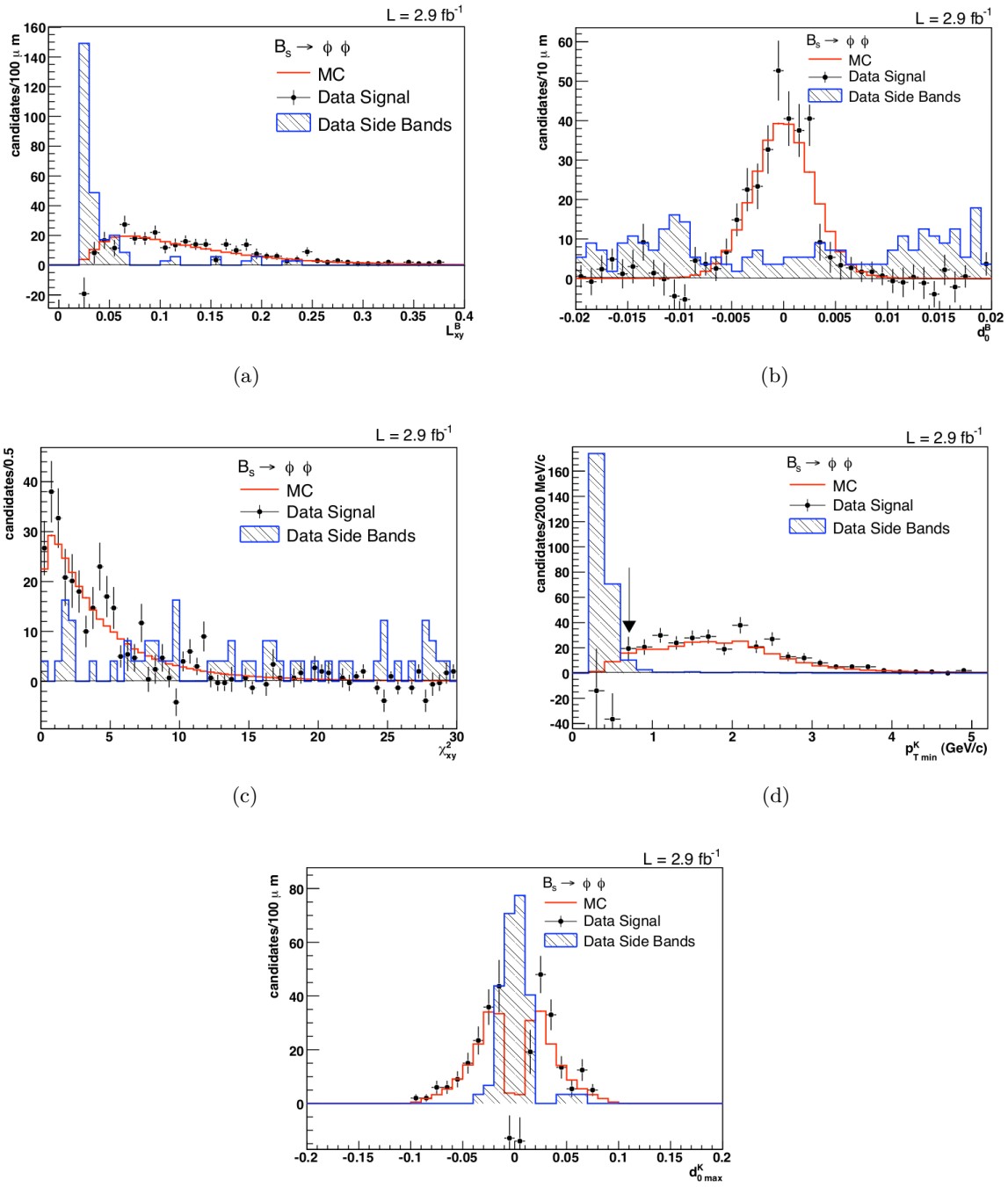


Figure 9.6: Variables used in the $B_s^0 \to \phi\phi$ selection: $L_{xy}(B)$ (a), $d(B)$ (b), χ^2_{xy} (c), $p_T^{\min}(K)$ (d), and $d(\phi_{\max})$ (e). The black points are side-bands subtracted data; the red histograms are simulated events; the blue histograms are sidebands data distribution.

9.3 Final sample

After applying the optimized cuts, we obtain the $K^+K^-K^+K^-$ invariant mass spectrum in Fig. 9.7. We distinguish three main component of the candidate mass distribution: a Gaussian-like signal peak with a width of about 20 MeV/c^2, which comprises about 300 events; a smooth constant distribution given by random combinations of charged tracks accidentally satisfying the selection requirements, which is the dominant background in the analysis; other B decays due to incorrect mass assignment of their decay-products. The latter is expected to peak under the signal, and we have considered two B^0-meson decays, the $B^0 \to \phi K^\star(890)^0$ and the $B^0 \to K^\star(890)^0 K^\star(890)^0$, when respectively one or two pions are reconstructed as kaons. We have estimated with simulation the number of these decay modes that pass the selection and enter the sample; we expect 3% of the signal yield for $B^0 \to \phi K^\star(890)^0$ decays, and 0.01% of the signal yield for $B^0 \to K^\star(890)^0 \bar{K}^\star(890)^0$ decays. In the following, we consider the sidebands regions as the intervals 5.244–5.294 GeV/c^2 and 5.444–5.494 GeV/c^2, and the signal region is in 5.32–5.42 GeV/c^2.

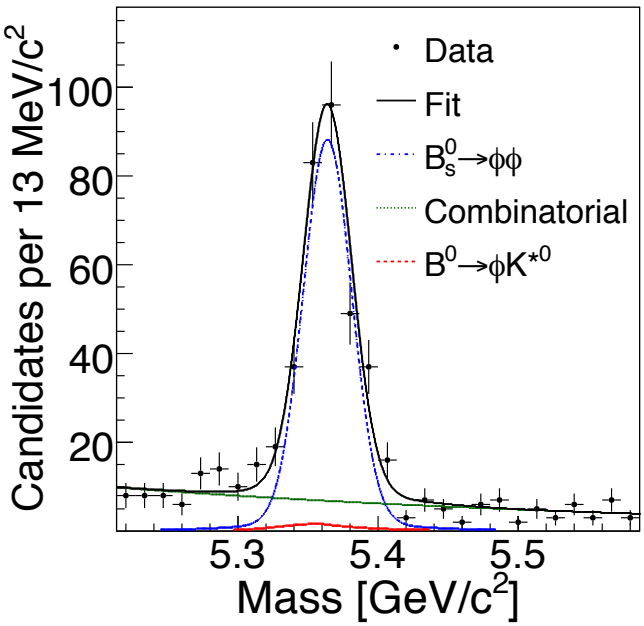
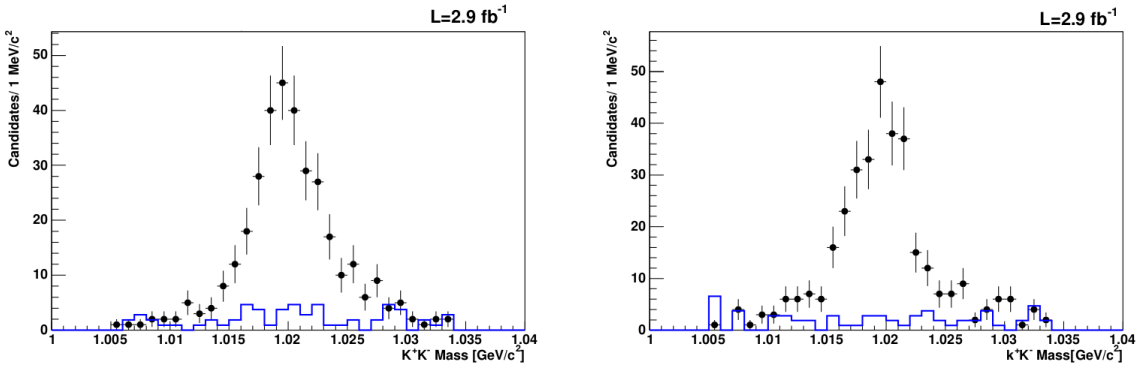


Figure 9.7: Distribution of the $K^+K^-K^+K^-$ mass after the optimized selection with the fit projections for the sample components overlaid. The signal is parametrized with the sum of two Gaussian distributions; the combinatorial background is modeled with a decreasing exponential; the shape of the $B^0 \to \phi K^\star(890)^0$ peak is extracted from simulations and its fraction is fixed in the fit.



(a) (b)

Figure 9.8: Distributions of K^+K^- mass of the ϕ with higher (a) and lower (b) momentum. Black points are sidebands-subtracted data; blue histograms are sidebands data.

Figure 9.8 shows the K^+K^- mass distributions after the selection, for the ϕ mesons with higher and lower momentum. In Fig. 9.9 the two distributions are summed and a fit is performed using the convolution of a Breit-Wigner distribution and a Gaussian function for the experimental resolution. The width of the Breit-Wigner function is fixed to the world average value [8], while the mass resolution is about 1 MeV/c^2.

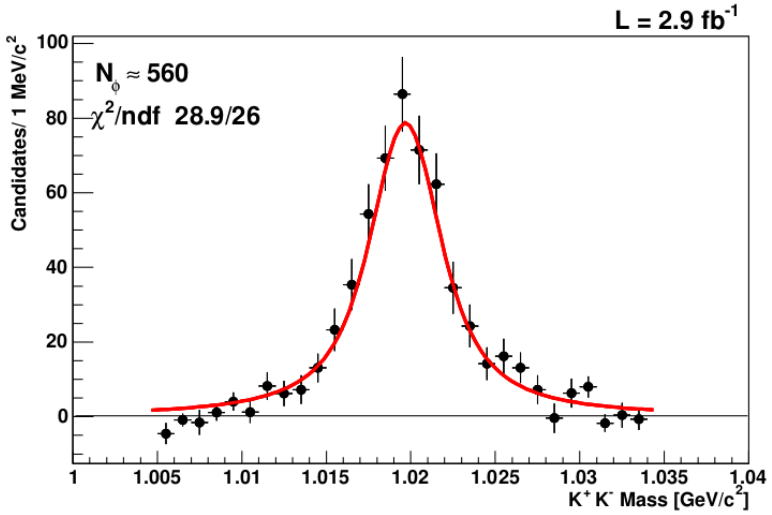


Figure 9.9: Sum of the K^+K^- spectum of the two ϕ candidates. The black points are sidebands-subtracted data; the red line is the fit performed with the convolution of a Breit-Wigner function (resonance) and a Gaussian function (experimental resolution).

9.4 Monte Carlo sample

The simulation of $B_s^0 \to \phi\phi$ decays follows the techniques described in Sect. 3.2.8. Simulations are exploited in the analysis to model the detector and selection acceptance, and to test the measurement methods and tools.

In order to study the detector sculpting on the triple products distributions of which we want to measure the asymmetries, a simulated sample of $B_s^0 \to \phi\phi$ decays obtained through a phase-space model from EvtGen has been used [102]. As for the phase-space model adopted in the $B_s^0 \to J/\psi\phi$ analysis, all spins of the final-state particles are averaged and the angular distributions at generator level are flat. Then, we fully simulated the detector response and the reconstruction of the events. We apply to the simulated sample the same selection as for data and store the information in the same format.

The reliability of the model depends on the agreement between data and the generated MC in variables that affect the angular distributions. Some inconsistencies are expected due to the inability of fully simulating the experimental conditions of the data acquisition, such as the instantaneous luminosity variations and the evolution of the trigger paths with their prescales. Similarly to the $B_s^0 \to J/\psi\phi$ analysis, a mismodeling in the $p_T(B)$ spectrum between simulation and data is observed. We extract the weights for reweighing the simulated $B_s^0 \to \phi\phi$ distribution from the results of a linear fit to the ratio of the distributions of the $p_T(B)$ spectrum of simulated events and data, in Fig. 9.10. The data distribution is obtained from a control sample of about 1 800 $B_s^0 \to J/\psi\phi$ decays collected with the displaced-tracks trigger in the same condition of the $B_s^0 \to \phi\phi$ decays, and the MC distribution is obtained from simulated $B_s^0 \to J/\psi\phi$ decays passing the same trigger conditions as in data and the same offline selection. Figure 9.6 shows the comparison between sideband-subtracted data distributions and simulated distributions, which are reweighed to match data. Further distributions are shown in Fig. 9.11.

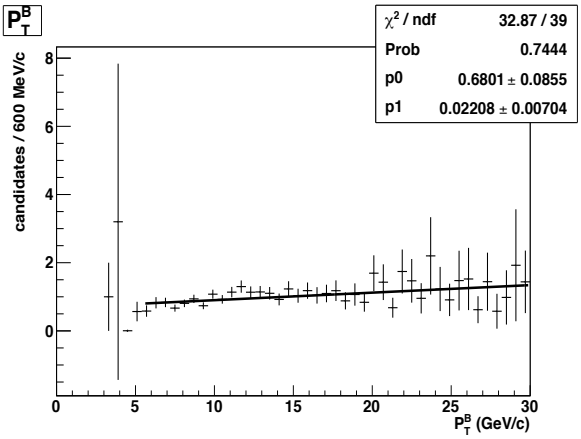
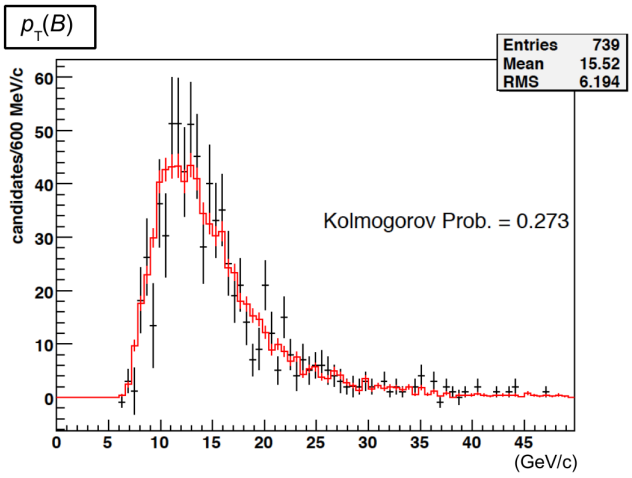
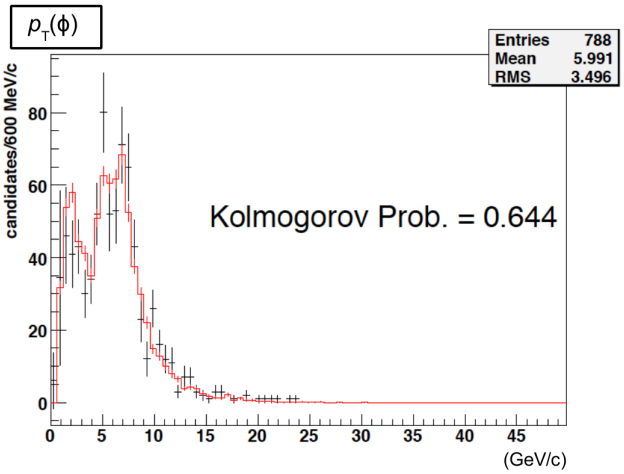


Figure 9.10: Continuous weights, as resulting from a fit to the ratio of the distributions of the $p_T(B)$ of simulated and collected $B_s^0 \to J/\psi\phi$ decays by the displaced-tracks trigger.



(a)



(b)

Figure 9.11: Comparison of data and MC after the reweighing: p_T spectra of the B_s^0 (a) and the ϕ (b) candidates.

Chapter 10

Measurement of triple products asymmetries

In this chapter we present the measurement of the triple products asymmetries of $B_s^0 \to \phi\phi$ decays. We first investigate the triple products distributions and we develop a maximum likelihood fit to measure the asymmetries. Some checks of the fitting technique are presented along with the results and their systematic uncertainties.

10.1 Maximum likelihood fit

In Sect. 2.3 we have introduced two functions of the helicity angles that define the triple products in the $B_s^0 \to \phi\phi$ decays, here reported again for convenience:

$$v \equiv (\hat{\boldsymbol{n}}_1 \times \hat{\boldsymbol{n}}_2) \cdot \hat{\boldsymbol{z}}_{\mathbf{1}} = \sin\varphi, \tag{10.1}$$

$$u \equiv 2(\hat{\boldsymbol{n}}_1 \cdot \hat{\boldsymbol{n}}_2)(\hat{\boldsymbol{n}}_1 \times \hat{\boldsymbol{n}}_2) \cdot \hat{\boldsymbol{z}}_{\mathbf{1}} = \sin 2\varphi. \tag{10.2}$$

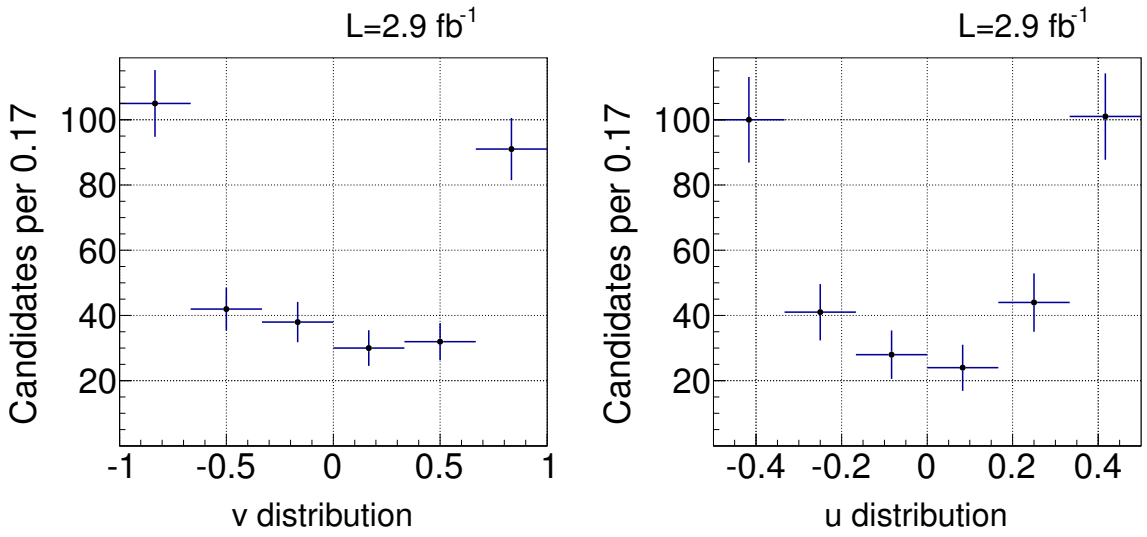
For sake of simplicity in notations, we have called v the first TP and u the second TP. Given the CP symmetric initial state, a Tevatron collision can yield a b quark only in association with a $\bar{b}$ quark (see Fig. 4.1 in Chap. 4). This ensures an equal amount of B_s^0 and $\bar{B}_s^0$ mesons at production, because the probability to produce an s and an $\bar{s}$ quark is the same and is independent from the flavor of the b quark. Thus, the signal sample collected is an equal mixture of $B_s^0 \to \phi\phi$ and $\bar{B}_s^0 \to \phi\phi$ decays. Such condition guarantees that the measurement of triple products asymmetries, $\mathcal{A}_v$ and $\mathcal{A}_u$, results in actual CP-violating asymmetries, Eq. (2.43) and Eq. (2.44). In fact, without the distinction of the production flavor of the B_s^0 meson, what we actually see is the sum of the triple products asymmetries in $B_s^0 \to \phi\phi$ and $\bar{B}_s^0 \to \phi\phi$ decays:

$$\mathcal{A}_v \equiv \mathcal{A}_{\mathrm{TP}}^{(1)} + \bar{\mathcal{A}}_{\mathrm{TP}}^{(1)} = \mathcal{A}_{\mathrm{TP},1}^{CP}, \tag{10.3}$$

$$\mathcal{A}_u \equiv \mathcal{A}_{\mathrm{TP}}^{(2)} + \bar{\mathcal{A}}_{\mathrm{TP}}^{(2)} = \mathcal{A}_{\mathrm{TP},2}^{CP}. \tag{10.4}$$

The v and the u distributions of $B_s^0 \to \phi\phi$ candidates in the signal region are shown in Fig. 10.1. We can have a first raw estimate of the asymmetries in these distributions by calculating the difference of the number of events with $v > 0$ (M^+) and $v < 0$ (M^-), normalized to the entire sample size:

$$\mathcal{A}_v^{\mathrm{raw}} = \frac{M^+ - M^-}{M^+ + M^-}, \tag{10.5}$$



(a) (b)

Figure 10.1: Distribution of v (a) and u (b) for $B_s^0 \to \phi\phi$ candidates in the signal region.

and with similar notation,

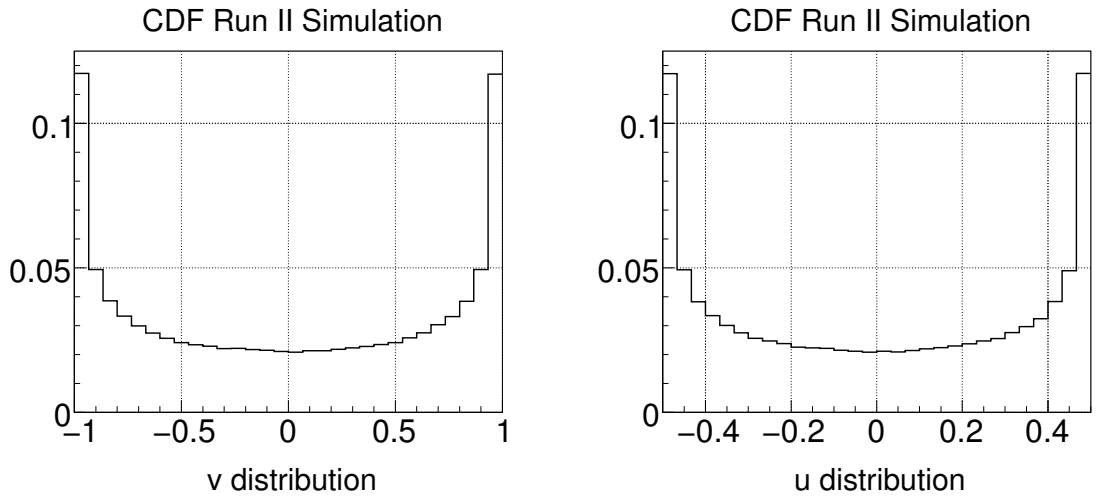
$$\mathcal{A}_u^{\text{raw}} = \frac{N^+ - N^-}{N^+ + N^-}. \tag{10.6}$$

With $M^+ = 153$ and $M^- = 185$; $N^+ = 169$ and $N^- = 169$; the raw asymmetries are:

$$\begin{aligned} \mathcal{A}_v^{\text{raw}} &= (-9.5 \pm 5.4)\% \\ \mathcal{A}_u^{\text{raw}} &= (0.0 \pm 5.4)\%. \end{aligned} \tag{10.7}$$

These asymmetries must be corrected for the asymmetries introduced by the background, and for spurious asymmetries due to experimental effects (detector acceptance and selection requirements). We now develop a maximum likelihood (ML) fit to the unbinned $K^+K^-K^+K^-$ mass distribution to refine the calculation of the asymmetries by accounting for the signal events only. In the following, we describe the fit referring to the triple product u for the measurement of $\mathcal{A}_u$, but exactly the same technique is applied for measuring the asymmetry $\mathcal{A}_v$ of the triple product v.

We consider all candidates in the range 5.2–5.6 GeV/c^2, and we split this sample into two subsets, made of N^+ and N^- events, according to the sign of u. The splitting depends on the asymmetry of the signal, *i. e.*, the TP asymmetry; on the asymmetry of the background; and on instrumental spurious asymmetries. The latter are estimated in a sample of about 400 thousand $B_s^0 \to \phi\phi$ events, simulated with the phase space MC presented in Sect. 9.4. The v and u distributions of the simulated decays are shown in Fig. 10.2, and the calculation of their



(a) (b)

Figure 10.2: Distributions of v (a) and u (b) for $B_s^0 \to \phi\phi$ events of phase space MC.

asymmetries gives:

$$\begin{aligned} \mathcal{A}_u^{\rm acc} &= (-0.14 \pm 0.15)\%, \\ \mathcal{A}_v^{\rm acc} &= (-0.08 \pm 0.15)\% \end{aligned} \tag{10.8}$$

There is no statistically significant spurious asymmetry at the per mill level and thus this effect can be safely neglected. Additional detector charge-asymmetries due to different detection efficiency of K^+ and K^- in the COT, or different interaction cross section of the positive and negative kaon, are automatically canceled since the $K^+K^-K^+K^-$ final state is completely symmetric and thus common to B_s^0 and $\bar{B}_s^0$.

We now perform a simultaneous ML fit to the mass distribution of the subsamples, N^+ and N^-, to extract the number of signal events, N_s^+ and N_s^-, which gives the asymmetry of the signal, $\mathcal{A}_u$:

$$\mathcal{A}_u = \frac{N_s^+ - N_s^-}{N_s^+ + N_s^-} = \frac{N_s^+ - N_s^-}{N_s}, \tag{10.9}$$

where $N_s = N_s^+ + N_s^-$ is the number of all signal events in the entire sample. Considering the fractions of signal events in each subsample, f_s^i (where $i = +, -$), and the fraction of signal in the entire sample, f_s, we can write:

$$\mathcal{A}_u = \frac{f_s^+ N^+ - f_s^- N^-}{f_s N}, \tag{10.10}$$

being $f_s N = f_s^+ N^+ + f_s^- N^-$. We must consider fluctuations of N^+ and N^- with the constrain $N = N^+ + N^-$. Thus, we treat N^+ and N^- as binomially distributed, with a probability p to

have N^+ events given N total. The asymmetry takes the following form:

$$\mathcal{A}_u = \frac{f_s^+ p - f_s^-(1-p)}{f_s}, \tag{10.11}$$

and the dependence on the total number of events cancels out. Neglecting spurious instrumental asymmetries, the splitting probability p is a combination of the total signal fraction f_s, the signal asymmetry $\mathcal{A}_u$, and a background asymmetry $\mathcal{A}_b$:

$$p = \frac{1}{2}(1 + f_s\mathcal{A}_u + (1-f_s)\mathcal{A}_b). \tag{10.12}$$

We can write each f_s^i in terms of $\mathcal{A}_u$, $\mathcal{A}_b$, and f_s:

$$\begin{aligned} f_s^+ &= \frac{f_s}{p}\frac{1+\mathcal{A}_u}{2} \\ f_s^- &= \frac{f_s}{(1-p)}\frac{1-\mathcal{A}_u}{2} \end{aligned} \tag{10.13}$$

where p is given by Eq. (10.12).

Following the same parameterization that we have used in the measurement of the $B_s^0 \to \phi\phi$ branching ratio and polarization amplitudes [2], the PDF for each subsample is:

$$P^i(m^i|\boldsymbol{\zeta}^i) = f_s^i\mathcal{G}(m^i|M,\sigma) + (1-f_s^i)\mathcal{B}(m^i|\lambda), \tag{10.14}$$

where $\boldsymbol{\zeta}^i$ is the parameters vector, m^i is the $K^+K^-K^+K^-$ mass of a $B_s^0 \to \phi\phi$ candidate in the subsample of N^i events, and:

- $\mathcal{G}(m^i|M,\sigma)$ is the sum of two Gaussian distributions with same mean value M but different resolutions, σ and $k\sigma$, which describes the signal component:

$$\mathcal{G}(m^i|M,\sigma) = h\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(m^i-M)^2}{2\sigma^2}} + (1-h)\frac{1}{\sqrt{2\pi}k\sigma}e^{-\frac{(m^i-M)^2}{2k^2\sigma^2}}, \tag{10.15}$$

  where h is the fraction of a Gaussian distribution with respect to the other one. This choice is commonly used to describe mass distribution in the CDF tracker and takes into account detector effects introducing tails in the mass distributions. In the fit the multiplicative factor k and the fraction h are fixed from a fit to large MC data sample, M is fixed to the world average value of the B_s^0 mass [8], while σ is floating in the fit.

- $\mathcal{B}(m^i|\lambda)$ describes the combinatorial background:

$$\mathcal{B}(m^i|\lambda) = \frac{\lambda e^{-\lambda m^i}}{e^{-\lambda m_{\min}} - e^{-\lambda m_{\max}}}, \tag{10.16}$$

  where λ is the slope of the exponential function floating in the fit, $m_{\min} = 5.2\,\text{GeV}/c^2$ and $m_{\max} = 5.6\,\text{GeV}/c^2$.

We are neglecting the 3% contribution of $B^0 \to \phi K^\star(890)^0$ decays present in our sample (which is expected to not have triple-products asymmetries anyway [26]), and we will assign a systematic uncertainty instead.

The likelihood for each subsample is simply:

$$\mathcal{L}^i(\boldsymbol{m}^i|\boldsymbol{\zeta}^i) = \prod_{j=1}^{N^i} P^i(m^i_j|\boldsymbol{\zeta}^i), \tag{10.17}$$

where $\boldsymbol{m}^i = \{m^i_1, \ldots, m^i_{N^i}\}$ is the vector of the masses of the events in the subsample N^i. Since the two subsamples are independent, we can perform a simultaneous fit taking the product of the two likelihoods, and considering the binomial PDF for N^+ and N^-, $f(N^+, N^-|p)$. Thus, the total likelihood reads:

$$\begin{aligned} \mathcal{L}(\boldsymbol{m}|\boldsymbol{\zeta}) &= \mathcal{L}^+(\boldsymbol{m}^+|\boldsymbol{\zeta}^+)\mathcal{L}^-(\boldsymbol{m}^-|\boldsymbol{\zeta}^-)f(N^+, N^-|p) = \\ &= \prod_{j=1}^{N^+} P^+(m^+_j|\boldsymbol{\zeta}) \prod_{i=1}^{N^-} P^-(m^-_j|\boldsymbol{\zeta}) \Big(p^{N^+}(1-p)^{N^-} \Big) \end{aligned} \tag{10.18}$$

Using the same parameter σ and λ in the parametrization of the N^+ and N^- events, the total floating parameters in the final fit are the following five:

$$\boldsymbol{\xi} = \{\mathcal{A}_u, \mathcal{A}_b, f_s, \sigma, \lambda\}. \tag{10.19}$$

The same likelihood is used in the estimation of $\mathcal{A}_v$, with the following replacements:

$$\begin{aligned} u &\to v, \\ N^+ &\to M^+, \\ N^- &\to M^-, \\ \mathcal{A}_u &\to \mathcal{A}_v. \end{aligned} \tag{10.20}$$

10.2 Validation of the ML fit

We check the implementation of the ML fit and its performances by making the test of the pulls distributions and by fitting simulated samples.

10.2.1 Pulls distribution

We generate 1000 pseudo-experiments with the statistic of the data, where:

- N^+ is generated from a binomial distribution given the splitting probability p (Eq. (10.12)) and the total number of events N, while N^- is simply $N - N^+$;
- N^+ and N^- mass values are randomly polled from the corresponding PDF, *i. e.* from $P^+(m|\boldsymbol{\zeta})$ and $P^-(m|\boldsymbol{\zeta})$, respectively.

Parameter	generation value	Mean fitted value	Mean fitted error	Pull mean	Pull variance
$\mathcal{A}_u$	-0.007	-0.009	0.064	-0.02 ± 0.03	1.02 ± 0.02
$\mathcal{A}_b$	-0.042	0.044	0.086	-0.02 ± 0.03	1.05 ± 0.02
f_s	0.616	0.617	0.029	0.03 ± 0.03	0.99 ± 0.02
σ	0.0174	0.0175	0.0011	-0.03 ± 0.03	0.98 ± 0.02
λ	2.68	2.68	0.67	-0.01 ± 0.03	1.00 ± 0.02

Table 10.1: Pulls study results for 1000 pseudo-experiments: set A.

Parameter	generation value	Mean fitted value	Mean fitted error	Pull mean	Pull variance
$\mathcal{A}_v$	-0.120	-0.121	0.064	-0.02 ± 0.03	1.00 ± 0.02
$\mathcal{A}_b$	0.009	0.007	0.086	-0.04 ± 0.03	1.07 ± 0.02
f_s	0.615	0.616	0.029	0.04 ± 0.03	1.03 ± 0.02
σ	0.0174	0.0174	0.0011	-0.01 ± 0.03	1.03 ± 0.02
λ	2.68	2.68	0.67	-0.01 ± 0.03	1.01 ± 0.02

Table 10.2: Pulls study results for 1000 pseudo-experiments: set B.

The pseudo-experiments are fitted using the likelihood defined in Eq. (10.18). For each parameter ζ^i, the corresponding pull g^i is defined by Eq. (6.26). We generate the events using two different sets (A and B) of input parameters and apply the described procedure for each of them. The results of the test are very satisfactory and are summarized in Tab. 10.1 and Tab. 10.2. The pull distributions for the sets A e B are reported in Fig. 10.3 and Fig. 10.4, respectively, and shows normally distributed pulls for all parameters.

10.2.2 Fit to Monte Carlo data samples

We fit two MC samples that differ from each others for the decay model used in the generation of the events. One sample is a subset of the phase-space MC used to extract the spurious asymmetries induced by the detector and the selection. The second MC sample is generated with a decay model of EVTGEN that simulates the polarization of the decay. In the generation of the sample we set the polarization amplitudes measured in our precedent analysis of $B_s^0 \to \phi\phi$ decays [2]. We don't expect any TP asymmetry in both MC samples, since no *CP* violation is applied in the simulation. The results of the fit to the two samples are summarized in Tab. 10.3, and in both cases, we measure a zero asymmetry within at most 0.5% uncertainty given by the limited MC event sample statistics. Therefore, we do not need a systematic uncertainty for residual asymmetries that are not modeled in the signal component of the likelihood.

MC sample	$\mathcal{A}_v$	$\mathcal{A}_u$
phase-space MC	$(-0.08 \pm 0.16)\%$	$(-0.14 \pm 0.16)\%$
polarized MC	$(-0.30 \pm 0.54)\%$	$(0.31 \pm 0.54)\%$

Table 10.3: Fit results for $\mathcal{A}_u$ and $\mathcal{A}_v$ for two MC datasets. No asymmetry is expected.

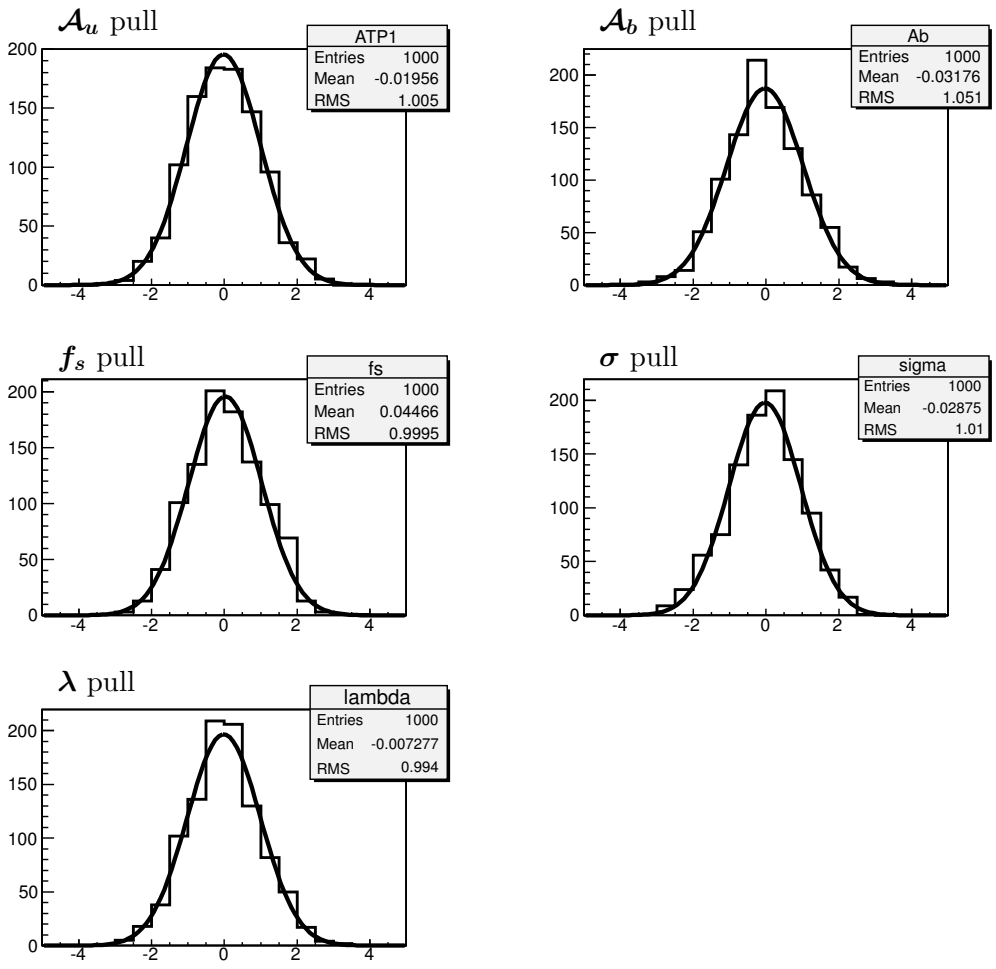


Figure 10.3: Pulls distributions for 1000 pseudo-experiments: set A.

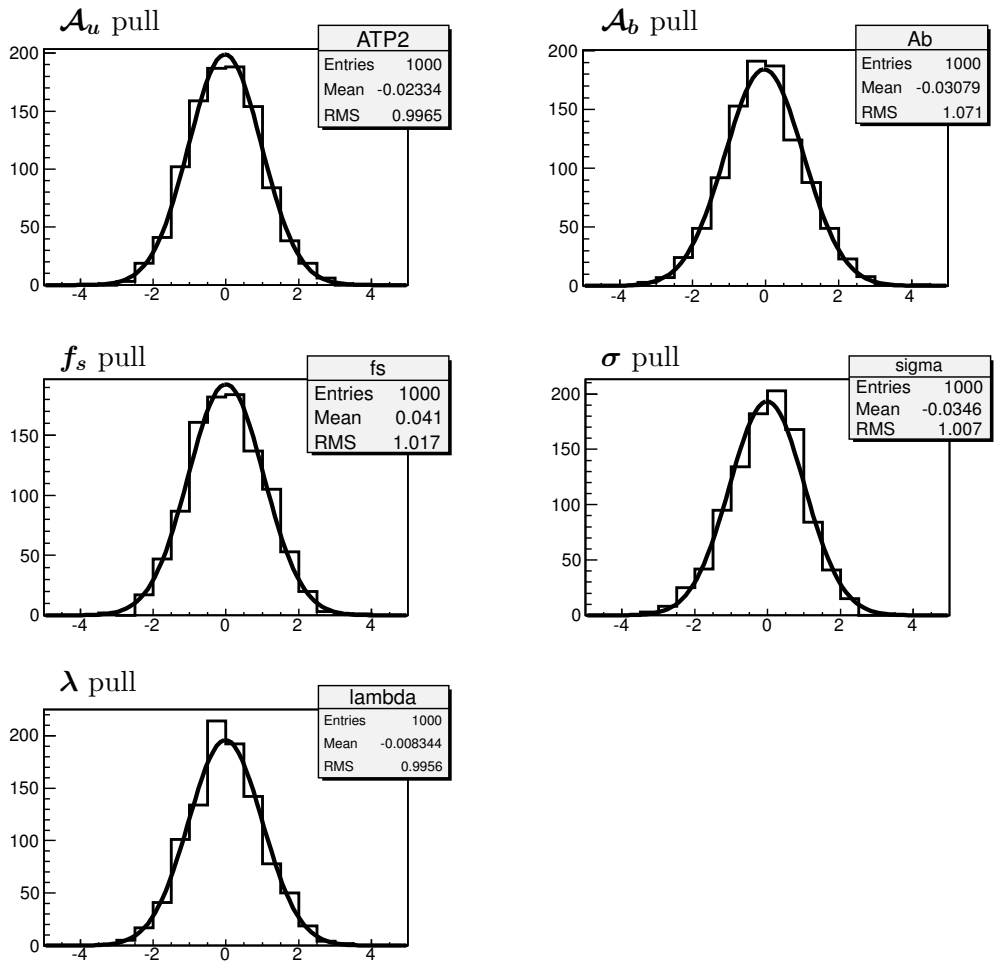


Figure 10.4: Pulls distributions for 1000 pseudo-experiments: set B.

10.3 Data sample fit

The results of the data fit are reported in Tab. 10.4 for both asymmetries, while in Tab. 10.5 the correlation coefficients of the parameters are shown. Both triple-product asymmetries are statistically consistent with zero as well as the background asymmetries. As expected, the two fit gives the same value of the signal fraction f_s, the width of the signal peak σ, and the slope of the background λ. The fit projection on the two subsamples are shown in Fig. 10.5 and Fig. 10.6 for the fit to measure $\mathcal{A}_u$ and $\mathcal{A}_v$, respectively, showing good agreement with data distributions.

Parameter	Fitted value
$\mathcal{A}_v$	-0.120 ± 0.064
$\mathcal{A}_b$	0.010 ± 0.086
f_s	0.615 ± 0.030
σ	0.0174 ± 0.0012
λ	2.68 ± 0.66

Parameter	Fitted value
$\mathcal{A}_u$	-0.007 ± 0.064
$\mathcal{A}_b$	-0.042 ± 0.087
f_s	0.616 ± 0.030
σ	0.0174 ± 0.0011
λ	2.68 ± 0.66

Table 10.4: Fit results of the fits to measure the $\mathcal{A}_v$ (left) and $\mathcal{A}_u$ (right) asymmetries.

	$\mathcal{A}_v$	$\mathcal{A}_b$	f_s	σ	λ
$\mathcal{A}_v$	1.00	-0.23	0.03	0.02	0.00
$\mathcal{A}_b$		1.00	0.02	0.02	0.00
f_s			1.00	0.35	-0.02
σ				1.00	-0.02
λ					1.00

	$\mathcal{A}_u$	$\mathcal{A}_b$	f_s	σ	λ
$\mathcal{A}_u$	1.00	-0.23	0.04	0.06	0.00
$\mathcal{A}_b$		1.00	-0.06	-0.09	-0.00
f_s			1.00	0.36	-0.02
σ				1.00	-0.02
λ					1.00

Table 10.5: Correlation coefficients of the parameters in the fit to measure the $\mathcal{A}_v$ (left) and $\mathcal{A}_u$ (right) asymmetries.

We check the results of the fit by computing the asymmetries of the distributions in Fig. 10.1, subtracting the background contribution. We estimate the symmetries of the background by looking at the v and u distributions for those background events populating the sidebands of the mass distribution, normalized to the fraction of combinatorial background that lies in the signal region (see Fig. 10.7). The asymmetries of the events in the sidebands are consistent with zero, respectively $\mathcal{A}_v^{\rm bkg} = (5 \pm 6)\%$ and $\mathcal{A}_u^{\rm bkg} = (1 \pm 6)\%$. Then, we subtract the triple-products distributions of the sidebands to the distributions in Fig. 10.1, and we found:

$$\begin{aligned} \mathcal{A}_v^{\rm sub} &= (-11.8 \pm 5.9)\%, \\ \mathcal{A}_u^{\rm sub} &= (-0.2 \pm 5.9)\%, \end{aligned} \tag{10.21}$$

in agreement we the estimations of the fits to the mass distributions.

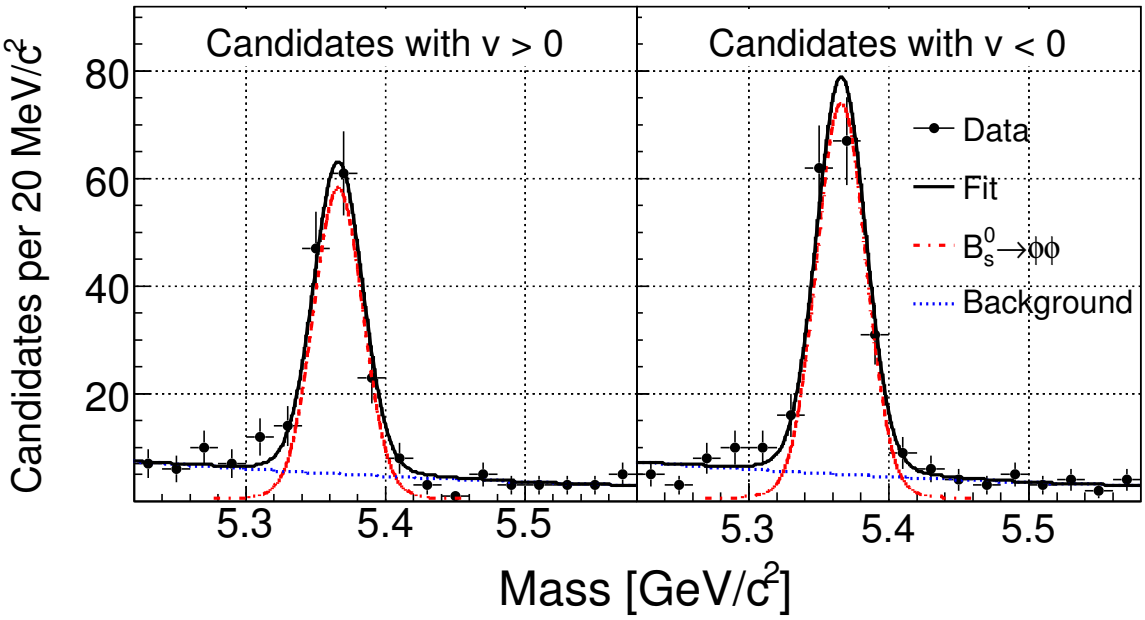


Figure 10.5: Fit projections for $\mathcal{A}_v$: M^+ (left) and M^- (right) samples. The curves are the fitting functions and represent: in red the $B_s^0 \to \phi\phi$ signal, in blue the combinatorial background, and in black the total projection.

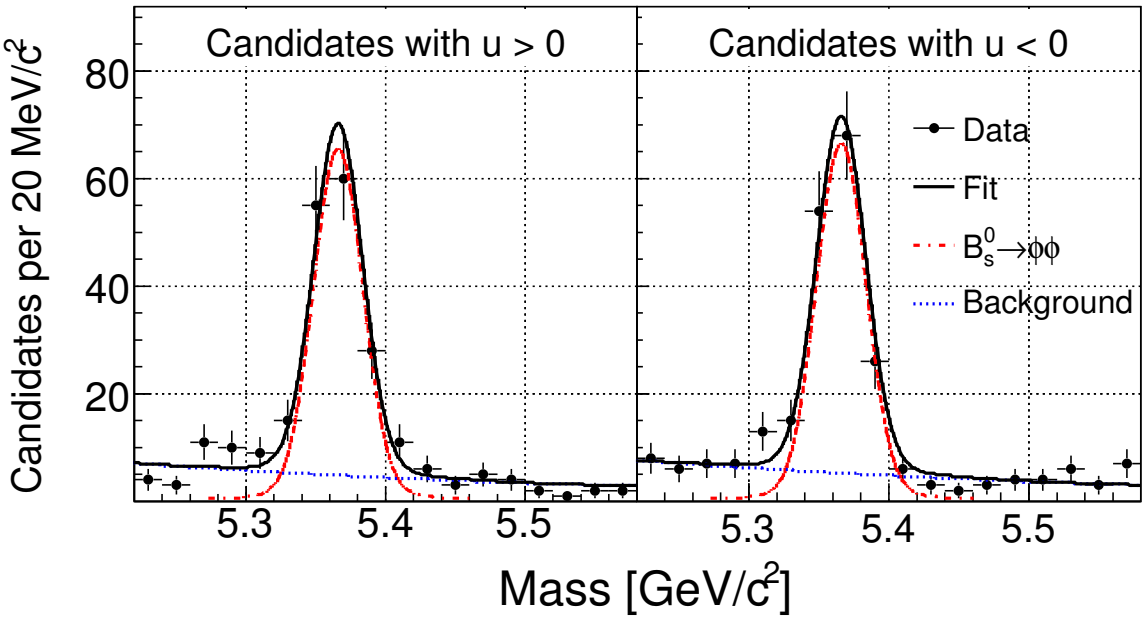
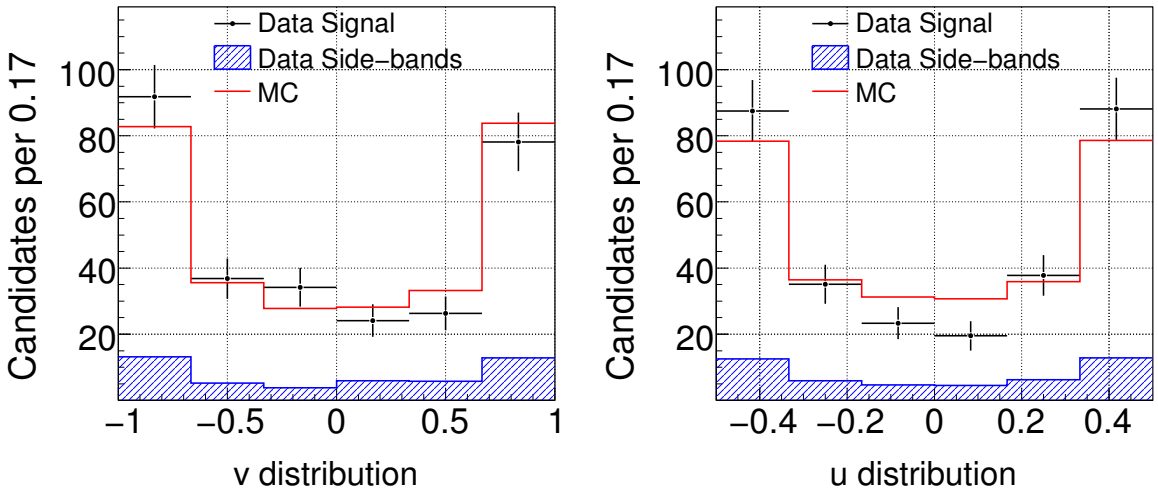


Figure 10.6: Fit projections for $\mathcal{A}_u$: N^+ (left) and N^- (right) samples. The curves are the fitting functions and represent: in red the $B_s^0 \to \phi\phi$ signal, in blue the combinatorial background, and in black the total projection.



(a) (b)

Figure 10.7: Comparison of distributions of v (a) and u (b) for sideband-subtracted data and MC events.

We compare the sidebands-subtracted distributions with distribution of the MC events simulated with the same polarization as measured in data [2] and without *CP* violation. The comparison is shown in Fig. 10.7 and reported in Tab. 10.6.

	$\mathcal{A}_u$	$\mathcal{A}_v$
data	(-0.7 ± 6.4)%	(-12.0 ± 6.4)%
MC	(-0.31 ± 0.54)%	(-0.30 ± 0.54)%

Table 10.6: Data-MC comparison for the $\mathcal{A}_u$ and $\mathcal{A}_v$ asymmetries.

10.4 Systematic uncertainties

The resolution on the measurement of the triple products asymmetries is limited by the statistic of the data sample. Thus, the estimation of the systematic uncertainties requires a reasonable assessment of the main sources of systematic uncertainties only. We have considered the following effects.

Physics background

Due to the small contribution of the $B^0 \to \phi K^\star(890)^0$ component (3%, see Sect. 9.3), it has been neglected in the fit and we compute its contribution as a systematic error. We generate a set of 1000 pseudo-experiments where the expected fraction of background events are all added

either to the N^+ (M^+) or to the N^- (M^-) sample. In this way we maximize their effect on the asymmetry measurement. The results are summarized in Tab. 10.7. The biggest shift between the input parameter and the mean fitted value is taken as the systematic uncertainty. It is about 1% for both $\mathcal{A}_u$ and $\mathcal{A}_v$.

	Mean $\mathcal{A}_u$	Mean $\mathcal{A}_v$
Input (fixed)	-0.007	-0.120
Fitted with $B^0 \to \phi K^\star(890)^0$ in N^+ (M^+)	-0.002	-0.110
Fitted with $B^0 \to \phi K^\star(890)^0$ in N^- (M^-)	-0.019	-0.130

Table 10.7: Results obtained adding the $B^0 \to \phi K^\star(890)^0$ component as explained in text.

$B^0 \to \phi\phi$ decays

In the measurement of the polarization amplitude of $B_s^0 \to \phi\phi$ decays, a systematic was assigned to account for the fluctuation in the $K^+K^-K^+K^-$ spectrum on the left of the B_s^0 peak (Fig. 9.7) which might be interpreted as a signal of $B^0 \to \phi\phi$ decays. Another component was added in the background model, parametrized as a double Gaussian function with the same width used for the B_s^0 peak and with a mean value fixed to the mass of the B^0 mesons [8]. The $B^0 \to \phi\phi$ decay is not yet observed, and only a limit of its branching ratio is given by *BABAR*, $< 2 \times 10^{-7}$ at 90% CL [8]. Given this limit, assuming the same detection efficiency of $B_s^0 \to \phi\phi$ decays, and considering that the fragmentation probability of the s quark $\approx 1/4$ of the probability of the d quark, we can expect at most ten signal events in our sample.[1]

We retain the choice of assigning a systematic, and we add a Gaussian-like component in the background model centered at the B^0 mass. We introduce two more floating parameters in the likelihood, $f_{B^0}^+$ and $f_{B^0}^-$, which are the fraction of this background in the N^+ (M^+) and N^- (M^-) subsets, respectively. In Fig. 10.8 and 10.9 we show the fit projections for the $\mathcal{A}_v$ and $\mathcal{A}_u$ estimations, respectively. The results of these fits are reported in Tab. 10.8. For $\mathcal{A}_u$ we obtain a shift of 0.9% and for $\mathcal{A}_v$ of 0.5% with respect to the default likelihood. These are the systematic uncertainties we assigned to this background effect.

Combinatorial background

We change the parameterization of the combinatorial background. We perform a linear fit on the $B_s^0 \to \phi\phi$ sidebands to estimate the parameter values that describe our background contribution. We then use these values as input to the generation of the pseudo-experiments where the linear background is simulated. The simulated data are then fitted with our standard model; a shift of 0.2% and 0.4% are seen respectively on $\mathcal{A}_u$ and $\mathcal{A}_v$.

[1] A very preliminary study done later shows that our data are compatible with the hypothesis of a background fluctuation.

Parameter	Fitted value
$\mathcal{A}_u$	0.001 ± 0.060
$\mathcal{A}_b$	-0.058 ± 0.083
f_s	0.633 ± 0.031
σ	0.0177 ± 0.0012
λ	1.71 ± 0.81
$f^+_{B^0}$	0.060 ± 0.027
$f^-_{B^0}$	0.027 ± 0.029

Parameter	Fitted value
$\mathcal{A}_v$	-0.114 ± 0.063
$\mathcal{A}_b$	0.006 ± 0.084
f_s	0.632 ± 0.030
σ	0.0177 ± 0.0012
λ	1.69 ± 0.80
$f^+_{B^0}$	0.054 ± 0.031
$f^-_{B^0}$	0.036 ± 0.022

Table 10.8: Fit results for the fit to measure the $\mathcal{A}_u$ (left) and $\mathcal{A}_v$ (right) asymmetries when $B^0 \to \phi\phi$ decays background contribution is added.

S-wave contributions

We considered the effects of a scalar contamination given by two decays potentially present in our sample: the $B^0_s \to \phi f_0(980)$ and the non-resonant $B^0_s \to \phi K^+ K^-$. We estimated with simulations a fraction of 4.6% and of 0.9% of the signal yield, respectively (for details, see Ref. [131]). We generate pseudo-experiments where these contributions are added in the above proportions to the standard signal+background pseudo-experiments and fitted with the default likelihood. The events of the background decays are randomly taken from mass histograms generated with MC simulation and we add all of them either to the N^+ (M^+) or to the N^- (M^-) sample, in order to maximize their effect on the asymmetry measurement. The results of the fit to pseudo-experiments are reported in Tab. 10.9 for the effect of $B^0_s \to \phi K^+ K^-$ decays and in Tab. 10.10 for the $B^0_s \to \phi f_0(980)$. We assign as uncertainty the biggest shifts in each case, and we add in quadrature them to quote the full systematic for the S-wave contributions: it is 1.1% for $\mathcal{A}_u$ and 1.0% for $\mathcal{A}_v$.

	Mean $\mathcal{A}_u$	Mean $\mathcal{A}_v$
Input (fixed)	-0.007	-0.120
Fitted with $B^0_s \to \phi K^+ K^-$ in N^+ (M^+)	-0.007	-0.120
Fitted with $B^0_s \to \phi K^+ K^-$ in N^- (M^-)	-0.010	-0.122

Table 10.9: Results obtained adding the $B^0_s \to \phi K^+ K^-$ component in pseudo-experiments.

Signal model

In the default likelihood we used a double gaussian to model the signal, and we fixed the parameters of one of them fitting a large sample of MC data. We account for the effect of this specific choice, by using pseudo-experiments, where a single gaussian PDF parameterizes the signal component as an alternative model. The found systematic uncertainty is 0.1% for both

	Mean $\mathcal{A}_u$	Mean $\mathcal{A}_v$
Input (fixed)	-0.007	-0.120
Fitted with $B_s^0 \to \phi f_0(980)$ in N^+ (M^+)	0.000	-0.110
Fitted with $B_s^0 \to \phi f_0(980)$ in N^- (M^-)	-0.018	-0.129

Table 10.10: Results obtained adding the $B_s^0 \to \phi f_0(980)$ component in pseudo-experiments.

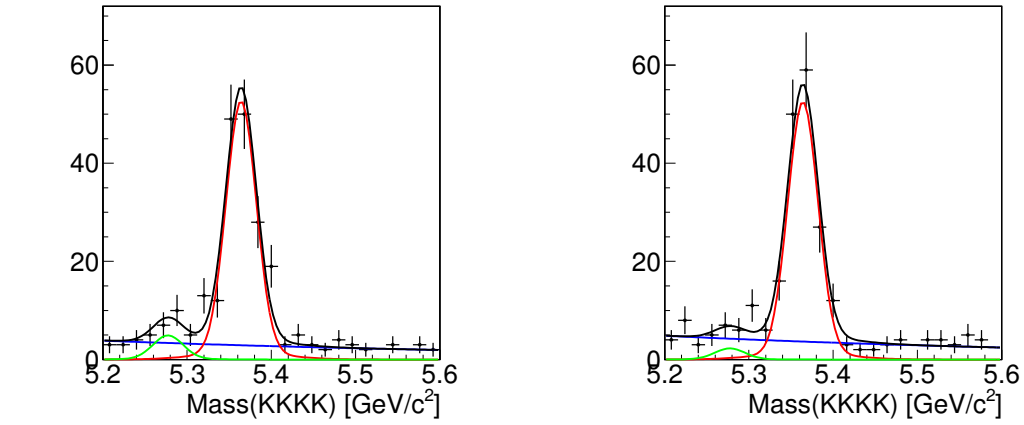
the asymmetries.

Adding in quadrature the different systematic uncertainties we obtain a total systematic uncertainty of 1.8% for $\mathcal{A}_u$ and of 1.6% for $\mathcal{A}_v$; both systematic uncertainties are less than 1/3 smaller with respect to the statistical errors.

Finally, the results of the measurement of the triple-product asymmetries, including statistical and systematic uncertainties, are:

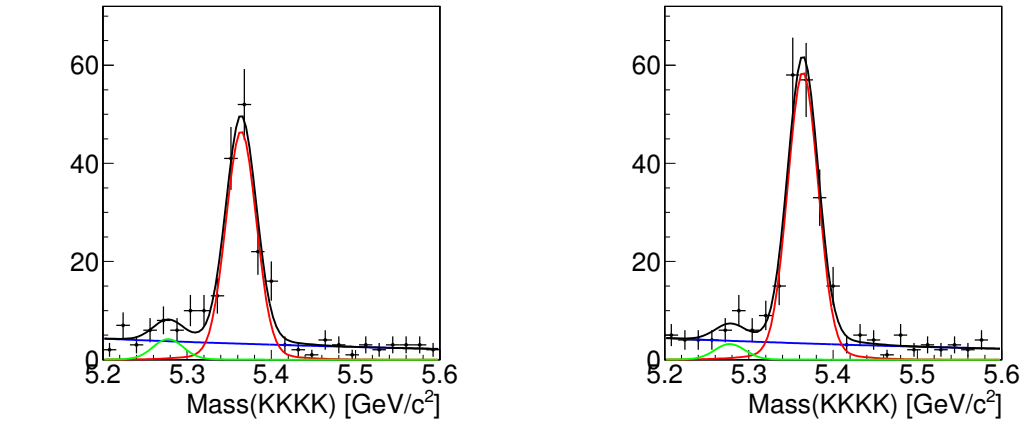
$$\begin{aligned} \mathcal{A}_v &= (-12.0 \pm 6.4(\text{stat}) \pm 1.6(\text{syst}))\%, \\ \mathcal{A}_u &= (-0.7 \pm 6.4(\text{stat}) \pm 1.8(\text{syst}))\%. \end{aligned}$$

The two measured asymmetries are statistically consistent with the no *CP* violation hypothesis, although $\mathcal{A}_v$ is 1.8 standard deviations different from zero, adding in quadrature the systematic and statistical uncertainties. This result represents the first search of *CP* violation performed in $B_s^0 \to \phi\phi$ decays. We discuss further on it in the next chapter, which is devoted to draw the conclusions of the dissertation.



(a) (b)

Figure 10.8: Fit to data for $\mathcal{A}_u$ with the $B^0 \to \phi\phi$ component added. Subsamples with $u > 0$ (a) and $u < 0$ (b). The curves represent: in red the $B_s^0 \to \phi\phi$ signal, in blue the combinatorial background, in green the $B^0 \to \phi\phi$ component, and in black the sum of the three.



(a) (b)

Figure 10.9: Fit to data for $\mathcal{A}_v$ with the $B^0 \to \phi\phi$ component added. Subsamples with $v > 0$ (a) and $v < 0$ (b). The curves represent: in red the $B_s^0 \to \phi\phi$ signal, in blue the combinatorial background, in green the $B^0 \to \phi\phi$ component, and in black the sum of the three.

Chapter 11

Conclusions

In this thesis we have reported on a search for non-SM physics in two decays modes of the B_s^0 meson through the measurement of the time-dependent *CP* asymmetry of the $B_s^0 \to J/\psi\phi$ decays, and of two time-integrated *CP* asymmetries of the $B_s^0 \to \phi\phi$ decays. The former allows to determine the value of the angles β_s of the unitarity triangle relevant for B_s^0 physics, which is expected to be ≈ 0.02 by the overall constraints of the CKM-matrix with $\mathcal{O}(0.1)\%$ accuracy. The angle $2\beta_s$ is the phase of the B_s^0-$\bar{B}_s^0$ mixing amplitude to a very good approximation, and a larger measured value than expected would indicate new physics in the $b \to \bar{b}s\bar{s}$ quark-transitions. The time-integrated *CP* asymmetries of $B_s^0 \to \phi\phi$ decays, $\mathcal{A}_v$ and $\mathcal{A}_u$, are given by asymmetries of helicity-angles functions, equivalent to the triple-product asymmetries discussed in literature. For $B_s^0 \to \phi\phi$ decays, they are predicted to vanish within 1% in the SM and a measurement of larger values with respect to expectations indicates new physics either in the $b \to s\bar{s}s$ or in the $b \to \bar{b}s\bar{s}$ quark-transitions (or both).

We have performed a new measurement of β_s using $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decays collected in the final CDF Run II dataset, and we have obtained:

$$-0.06 < \beta_s < 0.30 \text{ at the } 68\% \text{ confidence level.}$$

In addition, we have provided precise measurements of the decay width difference $\Delta\Gamma_s$; of the mean B_s^0 lifetime, τ_s; of the $B_s^0 \to J/\psi\phi$ polarization amplitudes, $|A_0|^2$ and $|A_\parallel|^2$, and of the *CP*-conserving phase $\delta_\perp$; as follows:

$$\begin{aligned}
\tau_s &= 1.528 \pm 0.019(\text{stat}) \pm 0.009(\text{syst}) \text{ ps},\\
\Delta\Gamma_s &= 0.068 \pm 0.026(\text{stat}) \pm 0.009(\text{syst}) \text{ ps}^{-1},\\
|A_\parallel|^2 &= 0.229 \pm 0.010(\text{stat}) \pm 0.014(\text{syst}),\\
|A_0|^2 &= 0.512 \pm 0.012(\text{stat}) \pm 0.018(\text{syst}),\\
\delta_\perp &= 2.79 \pm 0.53(\text{stat}) \pm 0.15(\text{syst}).
\end{aligned}$$

We have analyzed what was the largest sample of $B_s^0 \to \phi(\to K^+K^-)\phi(\to K^+K^-)$ decays to perform the world's first measurement of the asymmetries $\mathcal{A}_v$ and $\mathcal{A}_u$, and we have found:

$$\begin{aligned}
\mathcal{A}_v &= (-12.0 \pm 6.4(\text{stat}) \pm 1.6(\text{syst}))\%,\\
\mathcal{A}_u &= (-0.7 \pm 6.4(\text{stat}) \pm 1.8(\text{syst}))\%.
\end{aligned}$$

decay mode	observable	CDF	D0	LHCb	ATLAS	SM
$B_s^0 \to J/\psi\phi$	N events	11 000	6 600	21 200	22 700	
	β_s	$[-0.06, 0.30]$ @68% CL	$0.28^{+0.18}_{-0.19}$	0.001 ± 0.052	-0.11 ± 0.21	0.0184 ± 0.0009
	$\Delta\Gamma_s$ [ps^{-1}]	0.068 ± 0.027	$0.163^{+0.065}_{-0.064}$	0.116 ± 0.019	0.053 ± 0.023	0.087 ± 0.021
	τ_s [ps]	1.528 ± 0.021	$1.443^{+0.038}_{-0.035}$	1.521 ± 0.020	1.477 ± 0.017	1.489 ± 0.031
	$\lvert A_0\rvert^2$	0.512 ± 0.022	$0.558^{+0.017}_{-0.019}$	0.523 ± 0.025	0.528 ± 0.011	0.531 ± 0.022
	$\lvert A_\parallel\rvert^2$	0.229 ± 0.017	$0.231^{+0.024}_{-0.030}$	0.255 ± 0.020	0.220 ± 0.011	0.230 ± 0.028
	$\delta_\perp$	2.79 ± 0.55	$--$	2.90 ± 0.37	$--$	2.97 ± 0.18
$B_s^0 \to \phi\phi$	N events	300	$--$	800	$--$	
	$\mathcal{A}_v$[%]	-12.0 ± 6.6	$--$	1.0 ± 4.0	$--$	$< \pm 1$
	$\mathcal{A}_u$[%]	-0.7 ± 6.6	$--$	-5.5 ± 4.0	$--$	$< \pm 1$

Table 11.1: Comparison of the results obtained in this thesis (green column) with most recent results of other experiments (results of LHCb and ATLAS are preliminary) [19, 59, 23, 21], and the SM expectations in the rightmost column [13, 15]. The SM expectation of τ_s is given considering $\tau_s = (0.98 \pm 0.02)\tau_d$ [13], where τ_d is the world average value of the B^0 lifetime [8]. The theoretical predictions on the polarization amplitudes and $\delta_\perp$ are based on [132] and polarization measured in $B^0 \to J/\psi K^\star(890)^0$ decays [8]. In all experimental values, the uncertainties reported in the table includes in quadrature the statistical and the systematic uncertainty. The values of β_s from D0 and ATLAS are obtained with constraints on the strong phases; in addition ATLAS uses the constraint $\Delta\Gamma_s > 0$. The values of $\Delta\Gamma_s$, τ_s, $|A_0|^2$, and $|A_\parallel|^2$, from D0, LHCb and ATLAS are obtained in the analysis with β_s not constrained to its SM value.

All results are among the most precise determinations from a single experiment and exhibit an excellent agreement with the SM predictions and with measurements from other experiments (see Tab. 11.1). They have been published in two letters to *Physical Review* [1, 2].

Figure 11.1, Fig. 11.2(a) and Fig. 11.2(b), show the comparison between the most recent results from experiments in the $(\beta_s, \Delta\Gamma_s)$ plane, for $\Delta\Gamma_s$ alone, and for τ_s, respectively. The results in Tab. 11.1 are obtained with very different sample statistics, thus, in Tab. 11.2 we scale the statistical resolutions on β_s and $\Delta\Gamma_s$ of other experiments to the CDF signal yield for comparing the experimental sensitivity of the analyses.[1] The sensitivity to the mixing phase is mainly driven by two parameters: the B_s^0 decay-time resolution, σ_t, and the flavor-tagging power, ϵD^2. The higher boost of the B_s^0 mesons produced in the forward direction as in the LHCb, significantly helps in improving the decay-time resolution in this experiment, resulting in a better resolution on β_s with respect to other detectors. The lack of the flavor-tagging in the current ATLAS analysis explains the poorer sensitivity on β_s of this experiment. Conversely, the measurement of $\Delta\Gamma_s$ benefits from a better sample purity and it is almost independent from the tagging power and the decay-time resolution; indeed, the sensitivity on $\Delta\Gamma_s$ stems from the statistical separation through the angular analysis of the *CP*-even and *CP*-odd amplitudes, which evolves with different lifetimes, and requires neither flavor-tagging nor a highly-tuned decay-time resolution for the fit of the exponential decays. This comparison shows how the very good mass resolution, tagging

[1] For CDF we consider the gaussian-uncertainty on β_s as provided by the fit in Tab. 8.1 of Sect. 8.3, before the coverage-correction of the confidence interval.

experiment	σ_t [fs]	ϵD^2 [%]	$\sigma_m(B_s^0)$ [MeV/c^2]	S/B at peak	$\sigma(\beta_s)$	$\sigma(\Delta\Gamma_s)$ [ps^{-1}]
CDF	100	$\approx 4\%$	≈ 10	2/1	0.12	0.026
D0	100	$\approx 2\%$	≈ 30	1/3	0.15	0.050
LHCb	50	$\approx 2\%$	≈ 7	30/1	0.07	0.024
ATLAS	100	$--$	≈ 12	3/1	0.30	0.030

Table 11.2: Comparison of the experiments performances in the measurement of β_s and $\Delta\Gamma_s$. We consider the statistical uncertainty only, scaled to the same signal yield (11 000 events). The parameter in the table are the decay-time resolution, σ_t; the tagging power, ϵD^2; the B_s^0 mass resolution, $\sigma_m(B_s^0)$; and the signal-to-background ratio (S/B) in the $J/\psi K^+K^-$ mass distribution at the B_s^0 mass peak.

capability, and decay-time resolution of the CDF II detector, made it a competitive experiment in the flavor-physics sector, although it was not originally optimized for this kind of physics.

The measurements of β_s and $\Delta\Gamma_s$ are still limited by statistics. Larger datasets can be analyzed only at the LHC experiments in the next years: the Tevatron experiments have shut down in September 2011, and the super-B-factory at KEK in Japan, of which commissioning should start in 2015, will have limited samples of B_s^0-meson decays. The LHCb experiments has now approximately 3 fb^{-1} of data available, three times more data than currently used in the analysis; ATLAS has on tape about 27 fb^{-1} of data, more than 5 times the statistic analyzed so far for $B_s^0 \to J/\psi\phi$ decays (assuming constant data-taking configurations). Besides the better statistical resolution due to larger datasets, improvements in the analyses techniques are expected; LHCb has already provided an additional measurement of β_s with another decay mode, $B_s^0 \to J/\psi(\to \mu^+\mu^-) f_0(\to \pi^+\pi^-)$ [133], resulting in about 18% improvement on the resolution of β_s when combined with the $B_s^0 \to J/\psi\phi$ analysis.

In addition, a larger dataset of $B_s^0 \to \phi\phi$ decays at LHCb (about 2 500 in 3 fb^{-1}) should allow to perform the analysis of the time-dependent *CP* asymmetry with similar techniques of the $B_s^0 \to J/\psi\phi$ analysis, to probe directly the mixing phase and $\Delta\Gamma_s$ in this decay mode. Assuming similar sensitivity to β_s with $B_s^0 \to J/\psi\phi$ decays, we can expect ≈ 0.3 resolution on $\phi^{s\bar{s}s}$ and $\approx 0.05\,\text{ps}^{-1}$ resolution on $\Delta\Gamma_s$.

Figure 11.1 presents also the combination as computed by *HFAG* [9] of all measurements in the $(\beta_s, \Delta\Gamma_s)$ plane, which is consistent with the SM prediction. Combining all data till summer 2012 and allowing only for new physics in M_{12}^s, the bounds shown in Fig. 11.3(a) in the $(\mathfrak{Re}\Delta_s, \mathfrak{Im}\Delta_s)$ plane are derived by the *CKMfitter* group [12]. While in 2010 new physics in the B_s^0-$\bar{B}_s^0$ mixing could very well accommodate the different deviations from the SM expectations seen at that time (see Fig. 1.5 in Chap. 1), this is not the case anymore in 2012. The model independent fit of all flavor data is consistent with no new physics in B_s^0-$\bar{B}_s^0$ mixing so far, although there is still some room for sizable deviations from the SM expectations. The overall constraint to $\arg(\Delta_s)$ is consistent with zero within $\approx 5\%$, almost a factor two worse than the corresponding constraint in the B^0 sector, $\arg(\Delta_d)$, derived from the measurements of $\sin 2\beta$. Anyway, it's already evident a tension between the direct determination of β_s and the dimuon

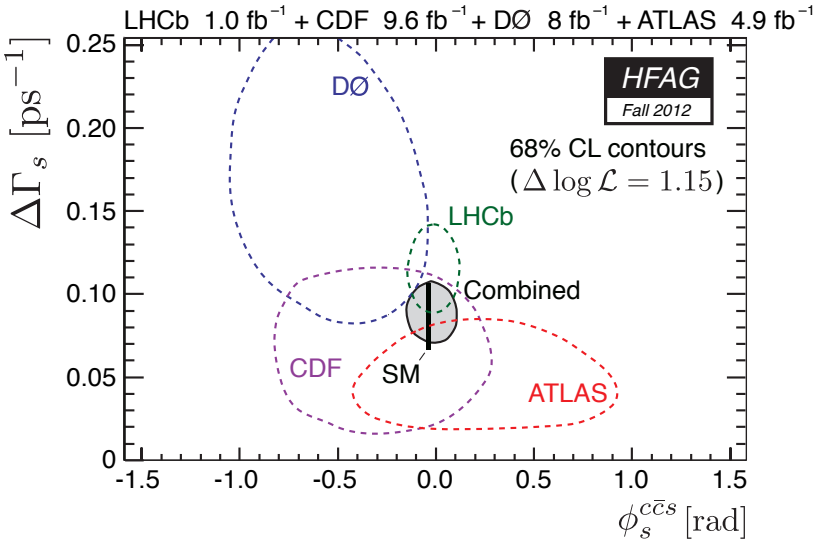


Figure 11.1: The 68% confidence level contours in the $(\phi^{c\bar{c}s}, \Delta\Gamma_s)$ plane (where $\phi^{c\bar{c}s} = -2\beta_s$) from CDF (this thesis), D0 [19] and LHCb [59], ATLAS [21], and their combined contour (solid line and shaded area), as well as the SM predictions (black marker) from global fit of the CKM matrix [13]. The combined result is consistent with the CKM predictions at the level of 0.14 standard deviations.

asymmetry, A_{sl}, measured by D0 [22], as shown by Fig. 11.3(b). Assuming that new physics in the B_s^0-$\bar{B}_s^0$ mixing is causing the anomaly, the large central value of the dimuon asymmetry is still an unsolved problem, because it can not be explained by new physics contributions to M_{12}^s alone; in addition, an enhancement of Γ_{12}^s compared to its SM value by a factor of 3 up to 34 would be required [13], which is not confirmed by latest measurements of $\Delta\Gamma_s$ and τ_s. To investigate these issues further, a better control over hadronic uncertainties from the theoretical side as well as more precise experimental inputs are mandatory.

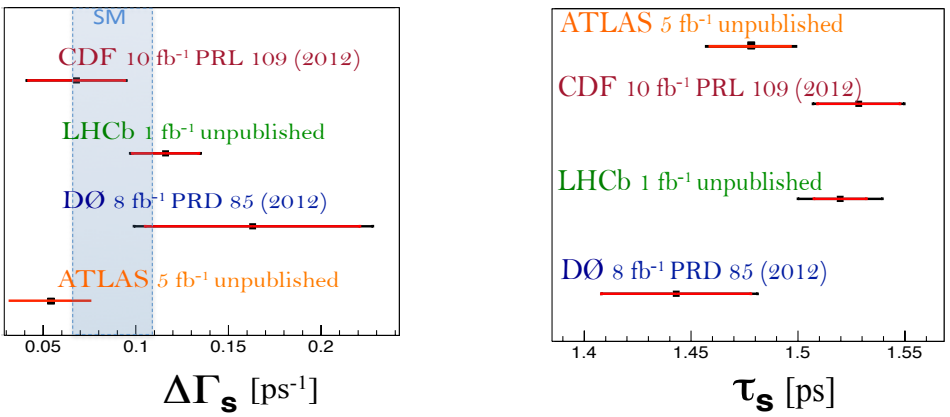
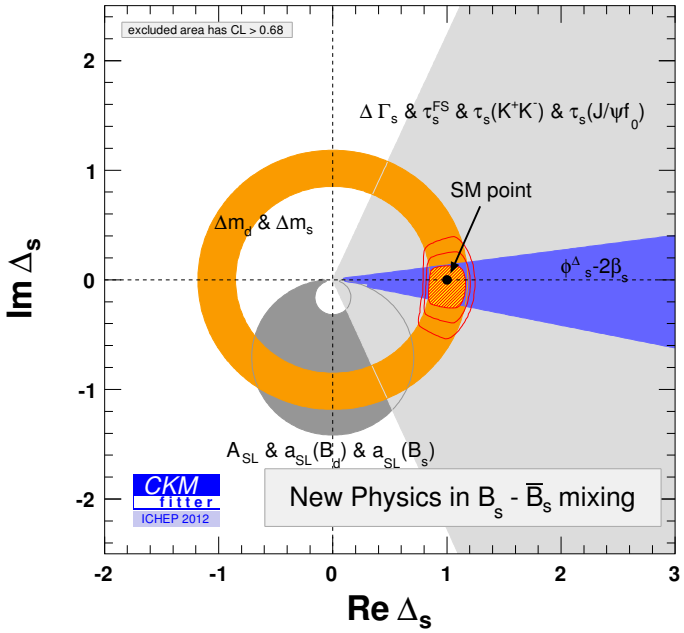
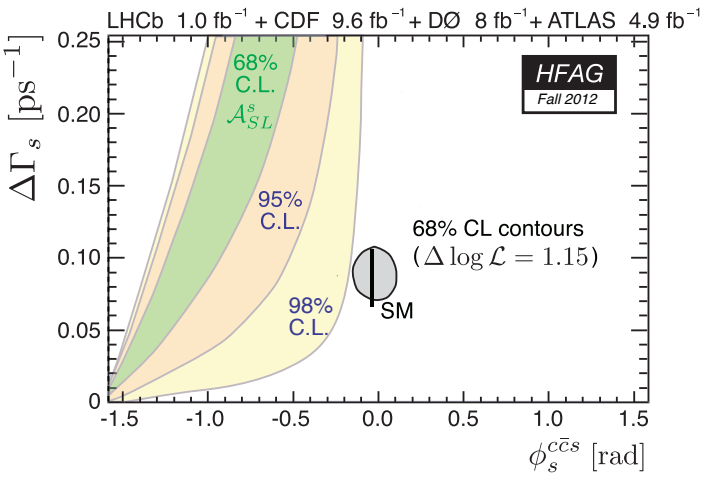


Figure 11.2: Comparisons of the most recent determinations of $\Delta\Gamma_s$ and of τ_s. Red bars represent the statistical uncertainties, while black bars include statistical and systematic uncertainties in quadrature.



(a)



(b)

Figure 11.3: In (a), bounds in the $(\mathfrak{Re}\Delta_s, \mathfrak{Im}\Delta_s)$ plane from measurements as of summer 2012 [12]. In (b), the $(\phi^{c\bar{c}s}, \Delta\Gamma_s)$ combined contour computed by HFAG [9] and the SM expectations together with the regions allowed at 68% and 95% confidence level by the average measurements of $A_{sl} = -0.0109 \pm 0.0040$ [9] and $\Delta m_s = 17.719 \pm 0.043\,\mathrm{ps}^{-1}$ [9], through the relation $\tan\phi_s = A_{sl}\Delta m_s/\Delta\Gamma_s$; these regions are drawn under the assumption that possible new physics will not affect the phase difference $\phi^{c\bar{c}s} - \phi_s$.

Appendix A

Derivation of the time-dependent CP asymmetry

In what follow we analyze the time-evolution of neutral mesons and derive the time-dependent CP asymmetry of Eq. (1.33).

After solving the eigenvalue problem of the mixing hamiltonian the mass eigenstates are

$$|P_{L,H}\rangle = p|P^0\rangle \pm q|\bar{P}^0\rangle, \tag{A.1}$$

with the normalization $|p|^2 + |q|^2 = 1$. The time evolution of the mass eigenstates is governed by the two eigenvalues, $m_H - \frac{i}{2}\Gamma_H$ and $m_L - \frac{i}{2}\Gamma_L$:

$$|P_{H,L}(t)\rangle = e^{-i(m_{H,L}+i\Gamma_{H,L}/2)t}|P_{H,L}(0)\rangle \tag{A.2}$$

Then, we can write the time evolution of P^0 and $\bar{P}^0$ from Eq. (A.2) and their definition in Eq. (A.1):

$$\begin{aligned} |P^0(t)\rangle &= g_+(t)|P^0\rangle + \frac{q}{p}g_-(t)|\bar{P}^0\rangle, \\ |\bar{P}^0(t)\rangle &= g_+(t)|\bar{P}^0\rangle + \frac{p}{q}g_-(t)|P^0\rangle, \end{aligned} \tag{A.3}$$

where

$$\begin{aligned} g_+(t) &= e^{-imt-\Gamma/2t}\left[\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right], \\ g_-(t) &= e^{-imt-\Gamma/2t}\left[-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right]. \end{aligned} \tag{A.4}$$

The following formulae are very useful:

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{e^{-\Gamma t}}{2}\left[\cosh\frac{\Delta\Gamma t}{2} \pm \cos(\Delta m t)\right], \\ g_+^\star(t)g_-(t) &= -\frac{e^{-\Gamma t}}{2}\left[\sinh\frac{\Delta\Gamma t}{2} + i\sin(\Delta m t)\right]. \end{aligned} \tag{A.5}$$

The probabilities of having a P^0 at any time t if it was either a P^0 or a $\bar{P}^0$ at $t = 0$ are:

$$\begin{aligned}
\mathcal{P}(P^0 \to P^0) &= |\langle P^0(t)|P^0(0)\rangle|^2 \\
&= |g_+(t)|^2 = \frac{e^{-\Gamma t}}{2}\left[\cosh\frac{\Delta\Gamma t}{2} + \cos(\Delta m t)\right], \\
\mathcal{P}(\bar{P}^0 \to P^0) &= |\langle P^0(t)|\bar{P}^0(0)\rangle|^2 \\
&= \left|\frac{p}{q}\right|^2 |g_-(t)|^2 = \left|\frac{p}{q}\right|^2 \frac{e^{-\Gamma t}}{2}\left[\cosh\frac{\Delta\Gamma t}{2} - \cos(\Delta m t)\right].
\end{aligned} \tag{A.6}$$

The mixing parameters Δm and $\Delta\Gamma$, or the corresponding dimensionless variables $x = \Delta m/\Gamma$ and $y = \Delta\Gamma/(2\Gamma)$, characterize the mixing, but are not related to CP violation. Just for completeness, we briefly comment on the values of x and y in neutral mesons. We can distinguish the following cases: $|y| \ll 1$ and $y \ll x$ for the B^0 meson; $y \approx x$ with $y \ll 1$ for the D^0 meson, with $y \approx 1$ and for the K^0 meson; $|y| \approx 1$ and $y \ll x$ for the B_s^0 meson, where $y \approx 0.1$. We show the unmixed and mixed intensities as a function of the dimensionless variable, Γt, for initially pure states of K^0, D^0, B^0, and B_s^0, in Fig. A.1.

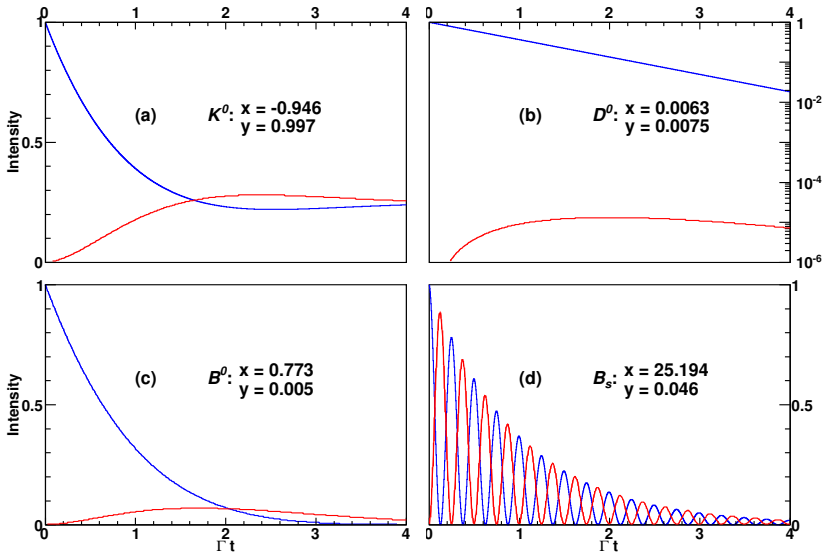


Figure A.1: The unmixed (blue) and mixed (red) intensities for an initially pure K^0 (a), D^0 (b), B^0 (c), and B_s^0 (d) state. The vertical scale in (b) is logarithmic, the others linear. The figures are taken from Ref. [134].

The time-dependent decay rates into the final state f for an initially produced P^0 and $\bar{P}^0$ are calculated by using Eq. (A.3) and Eq. (A.5); they reads:

$$\Gamma(P^0 \to f) = |A_f|^2 \frac{e^{-\Gamma t}}{2}(H + I), \tag{A.7}$$

$$\Gamma(\bar{P}^0 \to f) = |A_f|^2 \left|\frac{q}{p}\right|^2 \frac{e^{-\Gamma t}}{2}(H - I), \tag{A.8}$$

with:

$$
\begin{aligned}
H &\equiv \left(1 + |\lambda_f|^2\right) \cosh \frac{\Delta\Gamma t}{2} - 2\mathfrak{Re}(\lambda_f) \sinh \frac{\Delta\Gamma t}{2} \\
I &\equiv \left(1 - |\lambda_f|^2\right) \cos(\Delta m t) + 2\mathfrak{Im}(\lambda_f) \sin(\Delta m t).
\end{aligned}
\tag{A.9}
$$

where $A_f = \langle f|P^0\rangle$, and λ_f is given in Eq. (1.35). These expression gives us the probability, divided by dt, that the state which initially was P^0 (or $\bar{P}^0$) decays into the final state f during the time interval $[t, t + dt]$. The difference of Eq. (A.7) and Eq. (A.8) dived by their sum gives the time-dependent CP asymmetry of Eq. (1.33); it can be expressed in the compact form with the I and H functions as follows:

$$
\mathcal{A}_{CP}(t) = \frac{I}{H} \tag{A.10}
$$

Appendix B

Resonances in the K^+K^- mass spectrum

In what follows, we describe the relativistic Breit-Wigner distribution, $B_J(m)$, that enters Eq. (2.20) of Sect. 2.2.1 by analyzing the K^+K^- spectrum.

In general, a relativistic spin-J Breit-Wigner complex amplitude R_J describes a resonance masses with spin J [54]:

$$R_J(m) = \frac{m_J \Gamma_J(m)}{(m_J^2 - m^2) - i m_J \Gamma_J(m)} = e^{i\delta_J(m)} \sin\delta_J(m), \tag{B.1}$$

where we use the following convention:

$$\cot\delta_J(m) = \frac{m_J^2 - m^2}{m_J \Gamma_J(m)}. \tag{B.2}$$

with $\Gamma_J(m)$ and m_J the resonance width and mass, respectively. If the resonance can decay into several final states, the total width is the sum of the partial widths for any open channel:

$$\Gamma_J(m) = \sum_i \Gamma_J^{(i)}(m). \tag{B.3}$$

The form of the mass-dependent width is a function of the Blatt-Weisskopf factors [8], which weight the amplitudes to account for the spin-dependent effect. In the case of $J = 1$, we have:

$$\Gamma_J^{(i)}(m) = \Gamma_J^{(i)} \frac{m_J}{m} \frac{1 + r^2 q_J^2}{1 + r^2 q^2} \left(\frac{q}{q_J}\right)^3, \tag{B.4}$$

with q the momentum of the daughter particles in the resonance system after its two-body decay (q_J is evaluated at $m = m_J$). The mass-independent width $\Gamma_J^{(i)}$ depends on the final state i; r is the interaction radius, an empirical parameter of order of 1 fm (or few GeV^{-1}). For the $\phi(1020)$ resonance we consider $r = 3\,\mathrm{GeV}^{-1}$ and the partial widths of the decays: $\phi \to K^+K^-$ $(48.8 \pm 0.5)\%$, $\phi \to K_L^0 K_S^0$ $(34.2 \pm 0.4)\%$, and $\phi \to \rho\pi$ plus $\phi \to \pi^+\pi^-\pi^0$ $(15.32 \pm 0.32)\%$ [8].

The parametrization of the scalar $f_0(980)$ mass distribution requires more attention. Studies of K^+K^- spectrum were performed by the BES experiment [119]. It was found that the $f_0(980)$ resonance is described by an asymmetric shape called Flatté distribution [135]. The $f_0(980)$ can decay either into a pair of pions or into a pair of kaons. However, the threshold for the K^+K^- production is above the $f_0(980)$ pole, thus the decay width is asymmetric across the

pole. The mass dependence is described by means of a parametrization of a form similar to Eq. (B.2), with a decay width:

$$\Gamma_{f_0}(m) = \begin{cases} g_\pi \sqrt{m^2/4 - m_\pi^2} + g_K \sqrt{m^2/4 - m_K^2}, & \text{if } m^2 > m_K^2 \\ g_\pi \sqrt{m^2/4 - m_\pi^2} + i g_K \sqrt{m_K^2 - m^2/4}, & \text{if } m^2 < m_K^2, \end{cases} \tag{B.5}$$

where m_K and m_π are the mass of the charged kaon and the charged pion, respectively. The $g_{K(\pi)}$ parameter may be considered as the coupling-constant-squared for the resonance coupling to the $K^+K^-(\pi^+\pi^-)$ channel.

The invariant amplitude, $B_J(m)$, is proportional to $R_J(m)$:

$$B_J(m) \propto \frac{m}{q} R_J(m) \tag{B.6}$$

and its squared modulus gives the form of the mass distribution describing the resonance in the K^+K^- spectrum; the latter reads:

$$|B_\phi(m)|^2 = \frac{qm}{q_\phi m_\phi} \frac{m\Gamma_\phi^{(KK)}(m)}{(m_\phi^2 - m^2)^2 + m_\phi^2 \Gamma_\phi^2(m)} \tag{B.7}$$

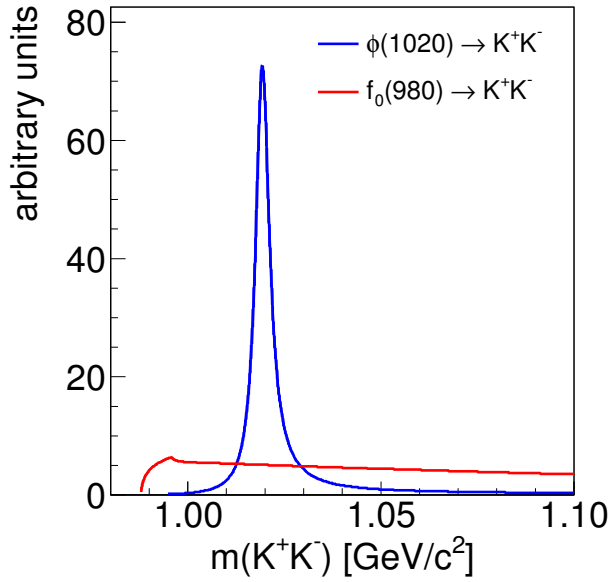
for the $\phi(1020)$ state, while it is:

$$|B_{f_0}(m)|^2 = \frac{m_{f_0}\Gamma_{f_0}^{(KK)}(m)}{(m_{f_0}^2 - m^2)^2 + m_{f_0}^2 \Gamma_{f_0}(m)} \tag{B.8}$$

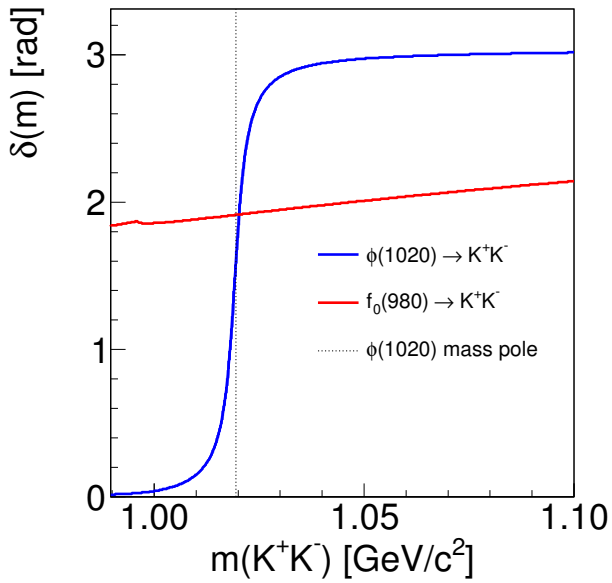
for the $f_0(980)$ resonance. Finally, given the above expressions Eq. (B.7), Eq. (B.8) and Eq. (B.2) the mass amplitudes are easily parametrized as:

$$B_J(m) = \sqrt{|B_J(m)|^2} e^{i\delta_J(m)}. \tag{B.9}$$

In Fig. B.1 (a) we show the distributions $|B_\phi(m)|^2$ and $|B_{f_0}(m)|^2$, and in Fig. B.1 (b) the phases $\delta_\phi(m)$ and $\delta_{f_0}(m)$.



(a)



(b)

Figure B.1: Shape of the $\phi(1020)$ and $f_0(980)$ resonances in the K^+K^- spectrum (a) and their phases $\delta_J(m)$ as a function of the K^+K^- mass (b).

Appendix C

Time-evolution of the $B_s^0 \to J/\psi\,\phi$ amplitudes

In the following pages, we report figures that sketch the time-evolution of the ten amplitudes of the $B_s^0 \to J/\psi K^+K^-$ decay rate described by Eq. (2.27) in Sect. 2.2.1, as a function of the B_s^0 decay-length, separately for the B_s^0 and the $\bar{B}_s^0$ mesons. Figure C.1 and Fig. C.2 are obtained with the SM values of $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.90\,\mathrm{ps}^{-1}$. Figure C.3 and Fig. C.4 are for $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.09\cos(2\beta_s) = 0.049\,\mathrm{ps}^{-1}$. In all plots, the value of the oscillation frequency is fixed to $\Delta m_s = 17.77\,\mathrm{ps}^{-1}$, as measured by CDF [43]; and polarization amplitudes and strong phases as measured in Ref. [18].

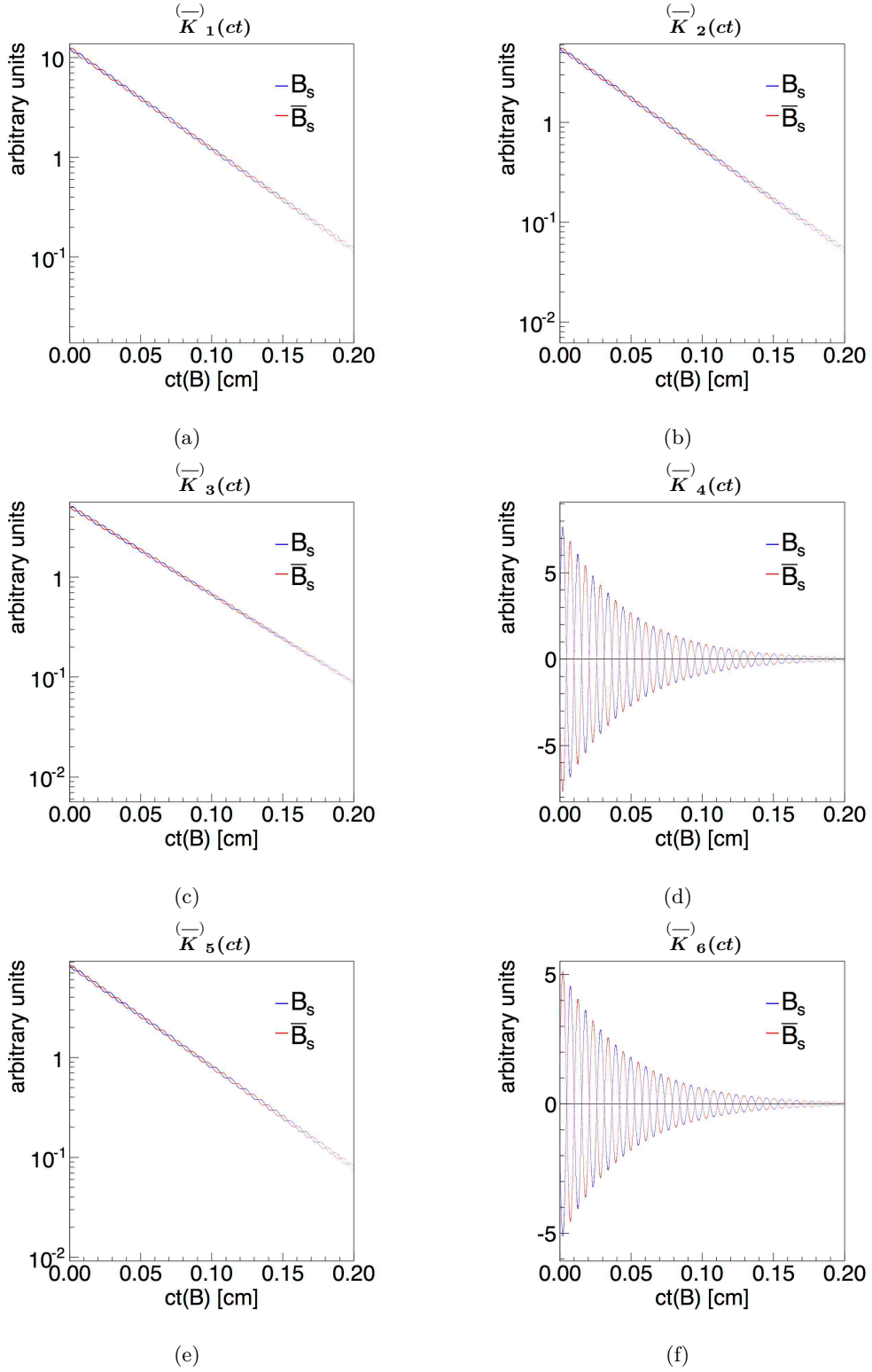


Figure C.1: Evolution of the six P-wave amplitudes of the $B_s^0 \to J/\psi K^+K^-$ decay rate, as a function of ct. Blue line is for B_s^0 meson and red line is for $\bar{B}_s^0$. From left to right, top to bottom: $K_1(ct)$ (a), $K_2(ct)$ (b), $K_3(ct)$ (c), $K_4(ct)$ (d), $K_5(ct)$ (e), $K_6(ct)$ (f). Here $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.09\,\text{ps}^{-1}$ (SM point), $\Delta m_s = 17.77\,\text{ps}^{-1}$. Plots are obtained by means of the simulation described in Sect. 5.5.2.

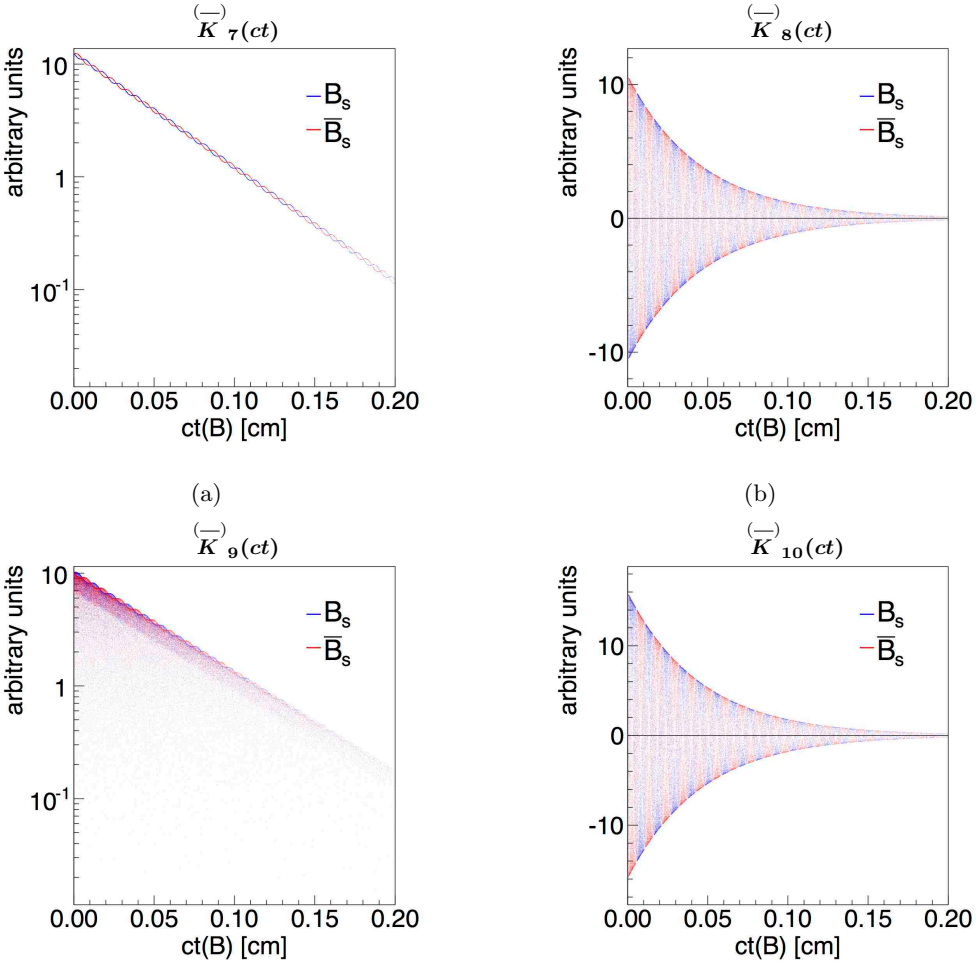


Figure C.2: Evolution of the S-wave and $(P+S)$-interference amplitudes of the $B_s^0 \to J/\psi K^+K^-$ decay rate, as a function of ct. Blue line is for B_s^0 meson and red line is for $\bar{B}_s^0$. From left to right, top to bottom: $K_7(ct)$ (a), $K_8(ct)$ (b), $K_9(ct)$ (c), $K_{10}(ct)$ (d). Here $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.09\,\mathrm{ps}^{-1}$ (SM point), $\Delta m_s = 17.77\,\mathrm{ps}^{-1}$. The interference terms depend on the resonances' phases $\delta_\phi(m) - \delta_{f_0}(m)$ that give the smearing of the points. Plots are obtained by means of the simulation described in Sect. 5.5.2

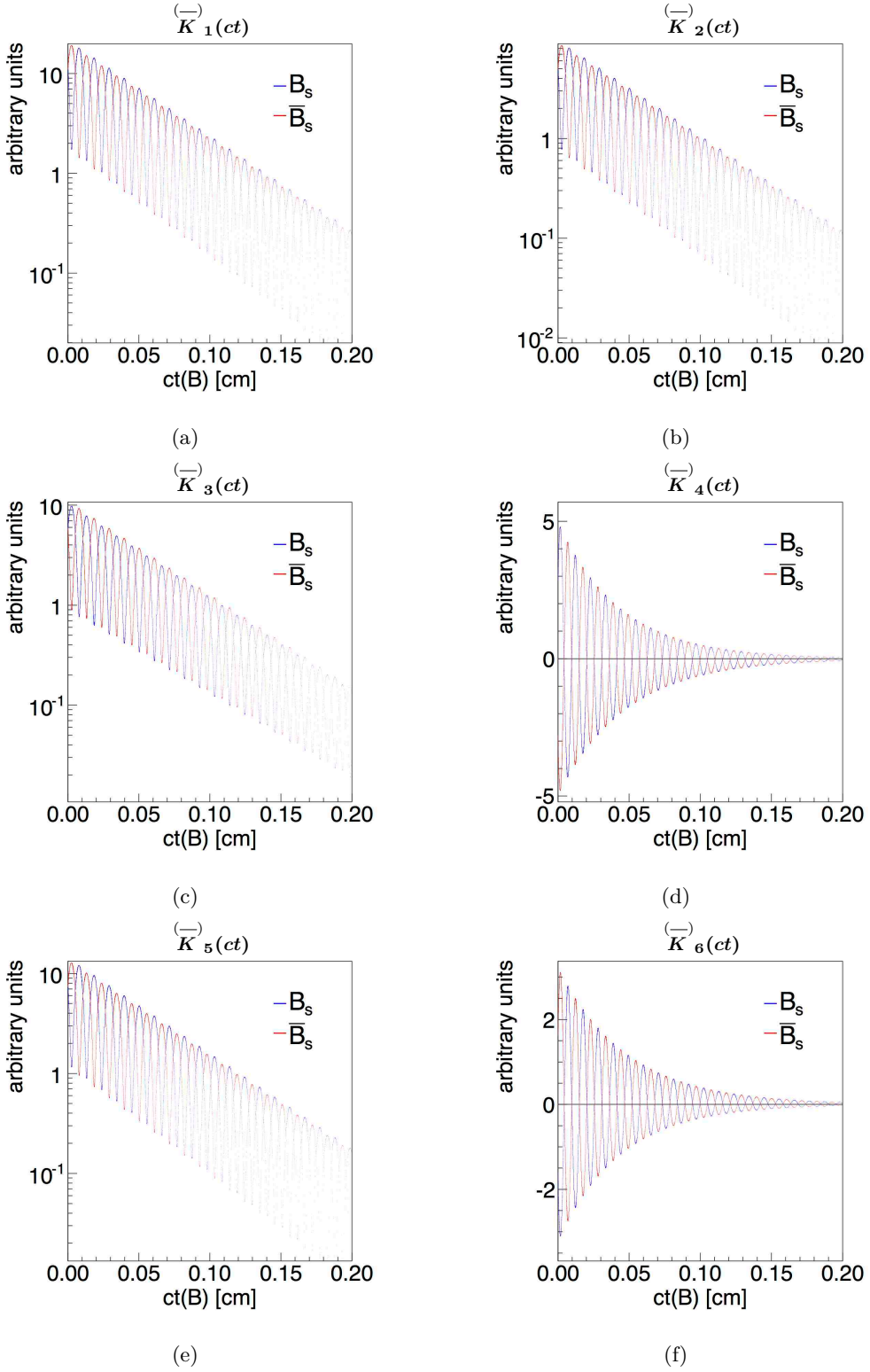


Figure C.3: Evolution of the six P-wave amplitudes of the $B_s^0 \to J/\psi K^+K^-$ decay rate, as a function of ct. Blue line is for B_s^0 meson and red line is for $\bar{B}_s^0$. From left to right, top to bottom: $K_1(ct)$ (a), $K_2(ct)$ (b), $K_3(ct)$ (c), $K_4(ct)$ (d), $K_5(ct)$ (e), $K_6(ct)$ (f). Here $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.09\cos(2\beta_s) = 0.049\,\text{ps}^{-1}$, $\Delta m_s = 17.77\,\text{ps}^{-1}$. Plots are obtained by means of the simulation described in Sect. 5.5.2

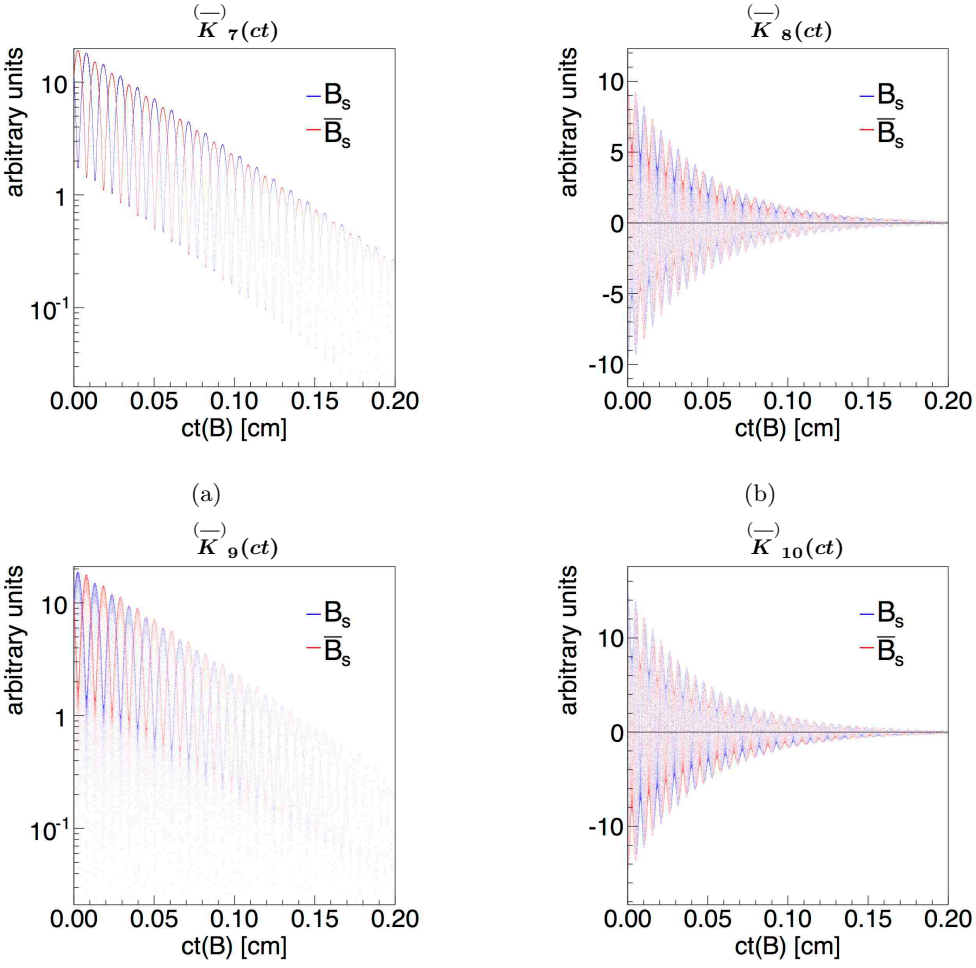
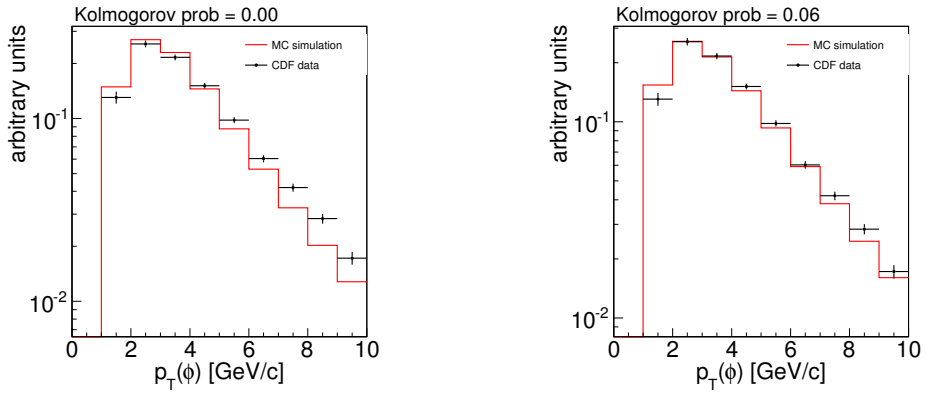


Figure C.4: Evolution of the S-wave and $(P+S)$-interference amplitudes of the $B_s^0 \to J/\psi K^+K^-$ decay rate, as a function of ct. Blue line is for B_s^0 meson and red line is for $\bar{B}_s^0$. From left to right, top to bottom: $K_7(ct)$ (a), $K_8(ct)$ (b), $K_9(ct)$ (c), $K_{10}(ct)$ (d). Here $\beta_s = 0.5$ and $\Delta\Gamma_s = 0.09\cos(2\beta_s) = 0.049\,\mathrm{ps}^{-1}$, $\Delta m_s = 17.77\,\mathrm{ps}^{-1}$. The interference terms depend on the resonances' phases $\delta_\phi(m) - \delta_{f_0}(m)$ that give the smearing of the points. Plots are obtained by means of the simulation described in Sect. 5.5.2

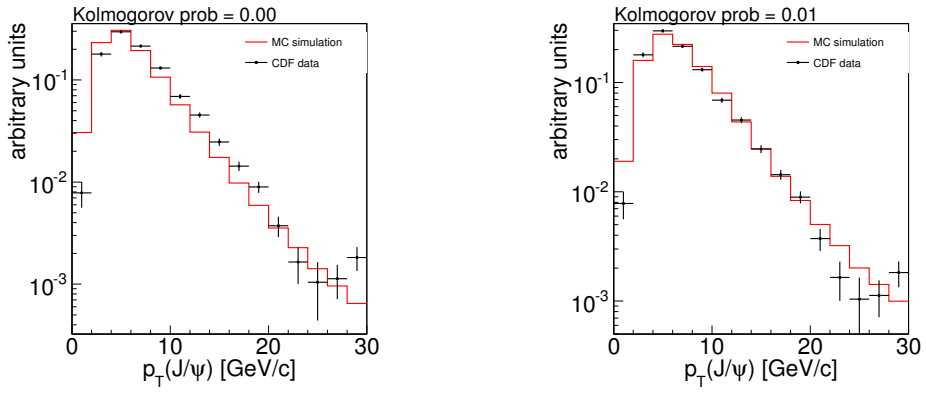
Appendix D

Validation plots of phase space $B_s^0 \to J/\psi\phi$ Monte Carlo

In this appendix we report the comparison plots between sidebands-subtracted data and phase space MC of the $B_s^0 \to J/\psi\phi$ decays, before and after the reweighing procedure describe in Sect. 5.5.1.



(a) (b)



(c) (d)

Figure D.1: Comparison between the data and phase space MC. The red histograms are MC data, the black points sideband-subtracted data. The MC distributions on the left column are not reweighed; the plot on rights are after the MC reweighing. Top to bottom, $p_T(B)$, $p_T(\phi)$, $p_T(J/\psi)$.

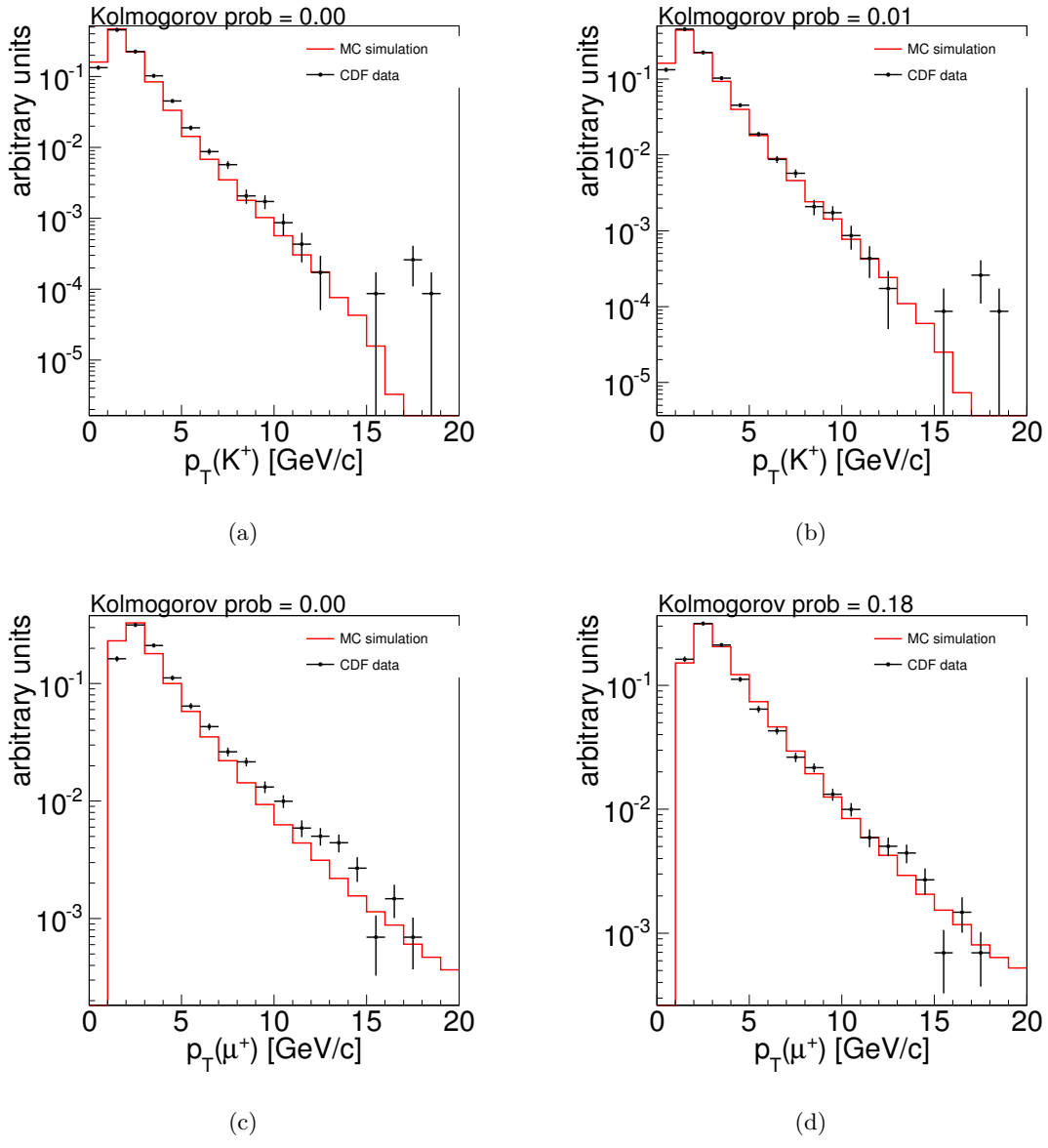


Figure D.2: Comparison between the data and phase space MC. The red histograms are MC data, the black points sideband-subtracted data. The MC distributions on the left column are not reweighed; the plot on rights are after the MC reweighing. Top to bottom, $p_T(K^+)$ and $p_T(\mu^+)$.

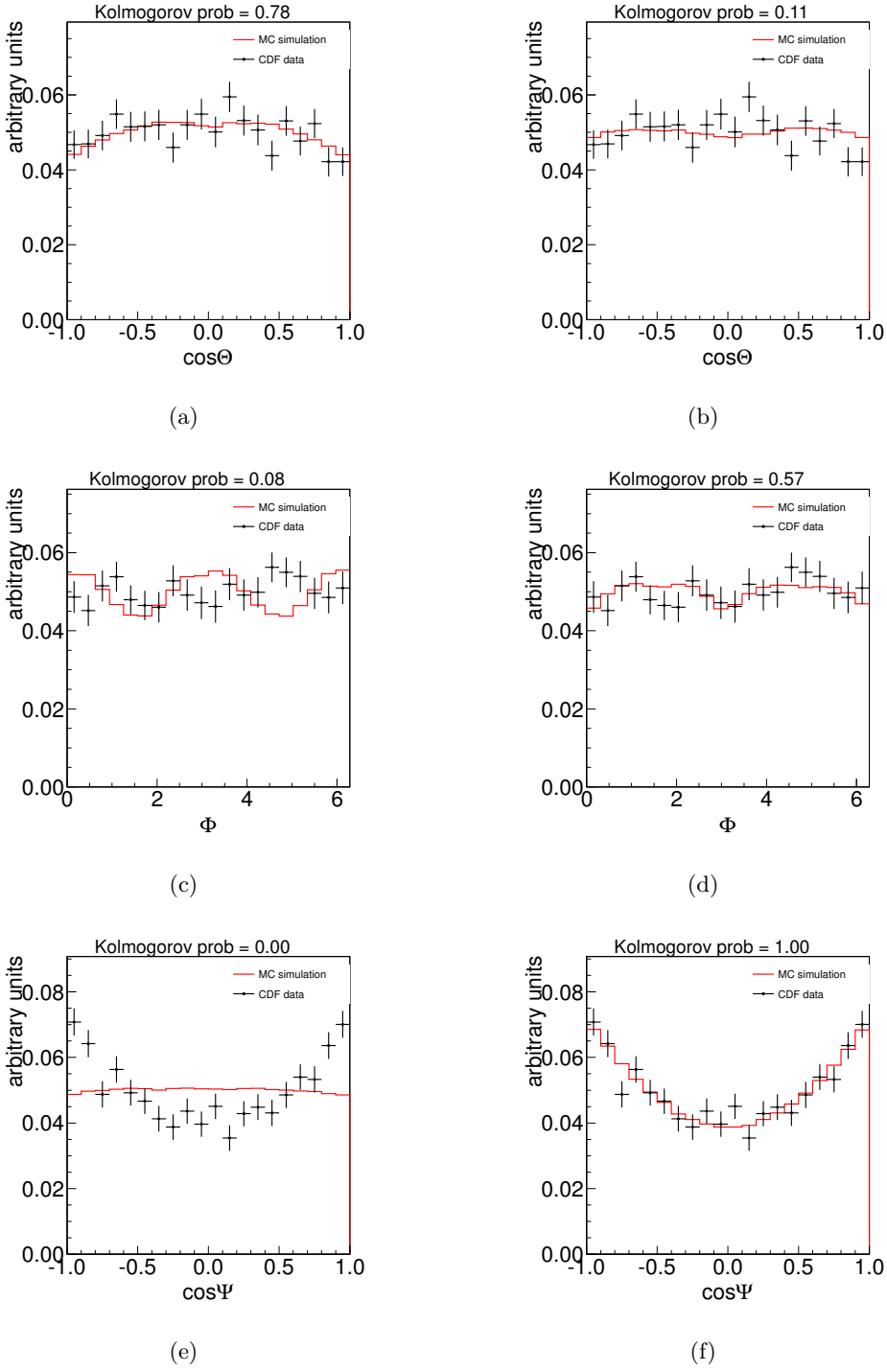


Figure D.3: Comparison between the data and phase space MC. The red histograms are MC data, the black points sideband-subtracted data. The MC distributions on the left column are not reweighed; the plot on rights are after the MC reweighing. Top to bottom, the transversity angles $\cos\Theta$, Φ and $\cos\Psi$.

Appendix E

Pulls distribution of the $B_s^0 \to J/\psi\phi$ fit

In the following, we report the plots of the pulls distribution of the ML fit for the $B_s^0 \to J/\psi\phi$ analysis discussed in Sect. 6.3.

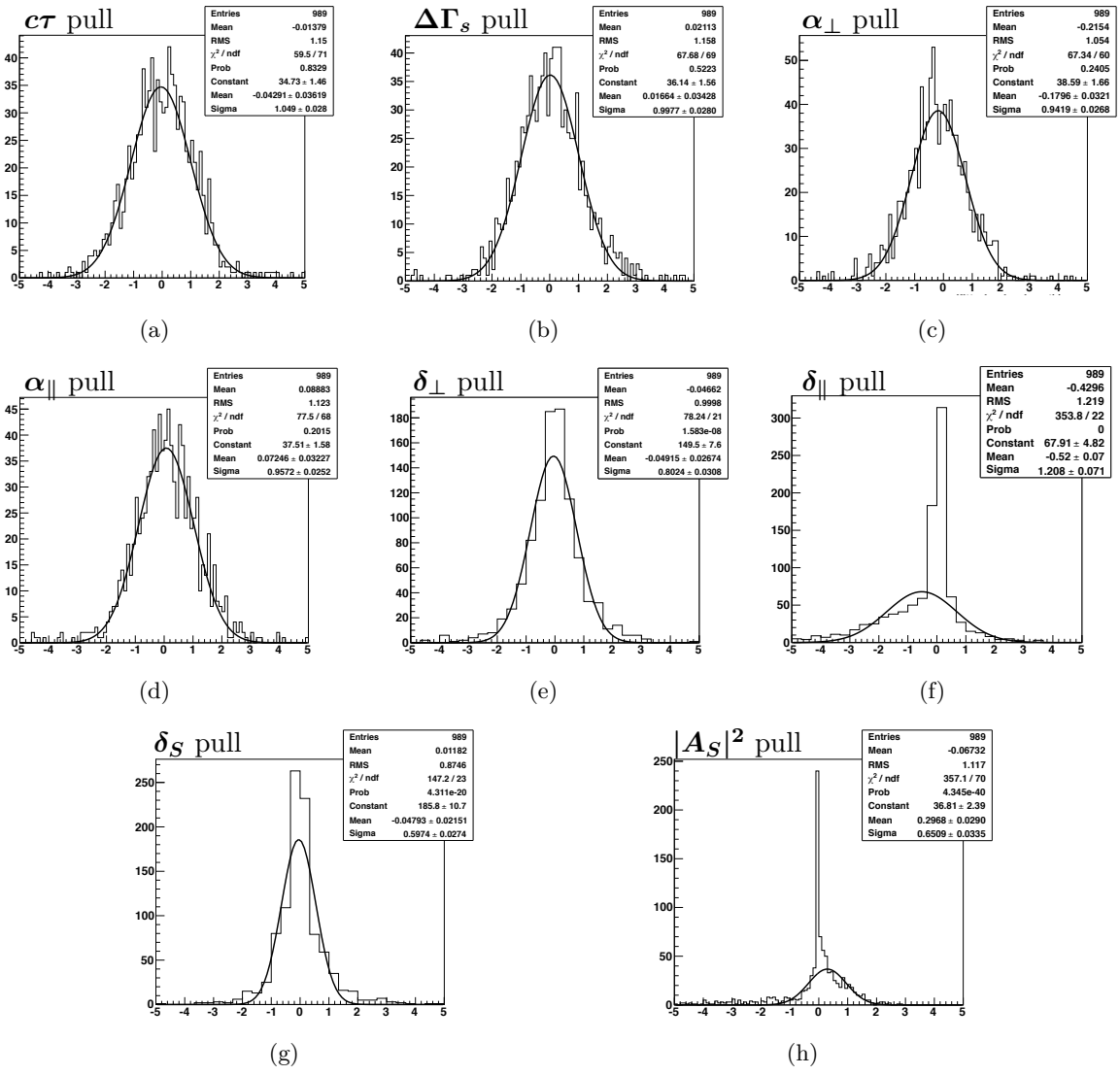


Figure E.1: Pull distributions for the main physics parameters of the β_s-fixed fit.

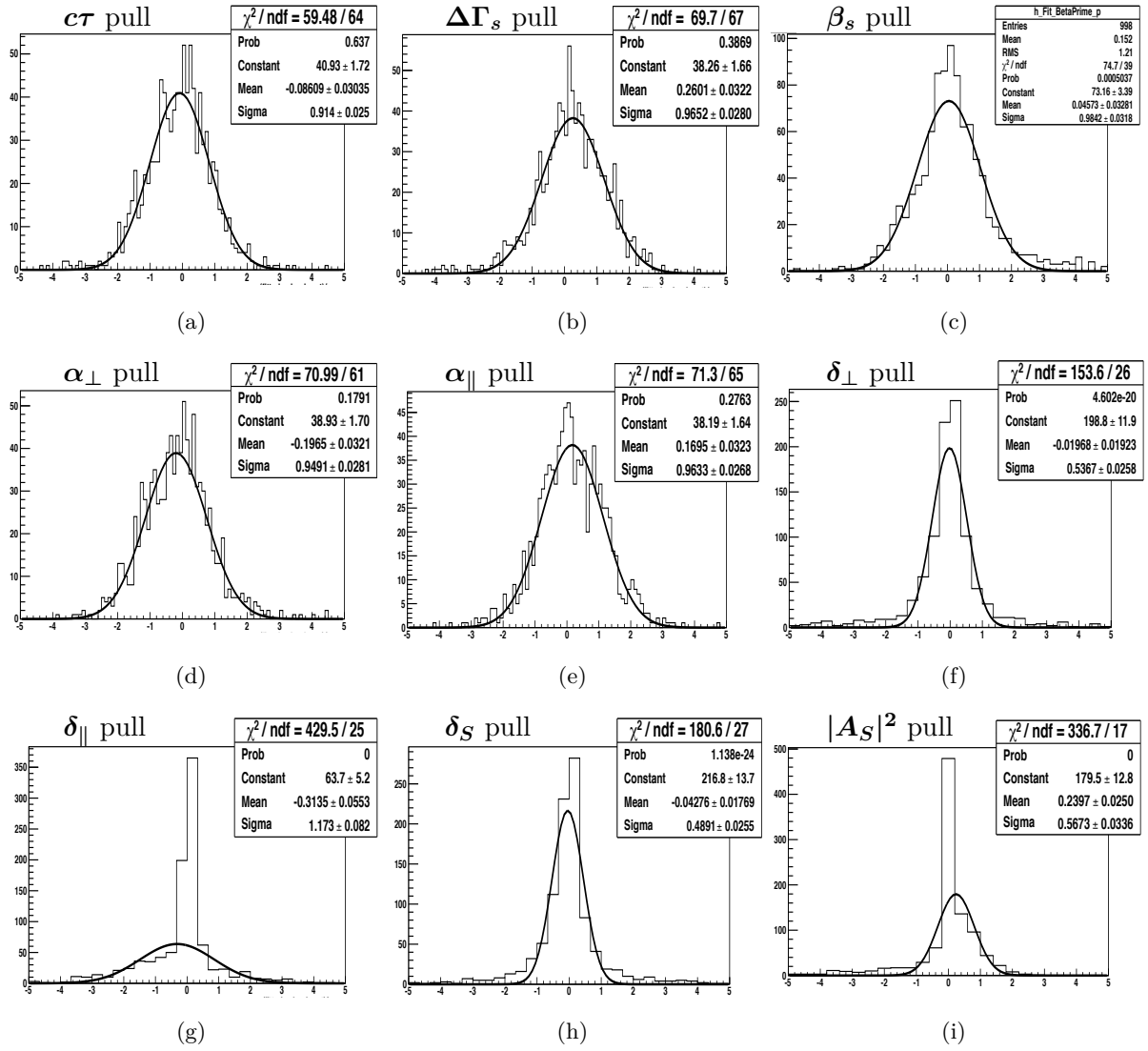


Figure E.2: Pull distributions for the main physics parameters of the β_s-floating fit.

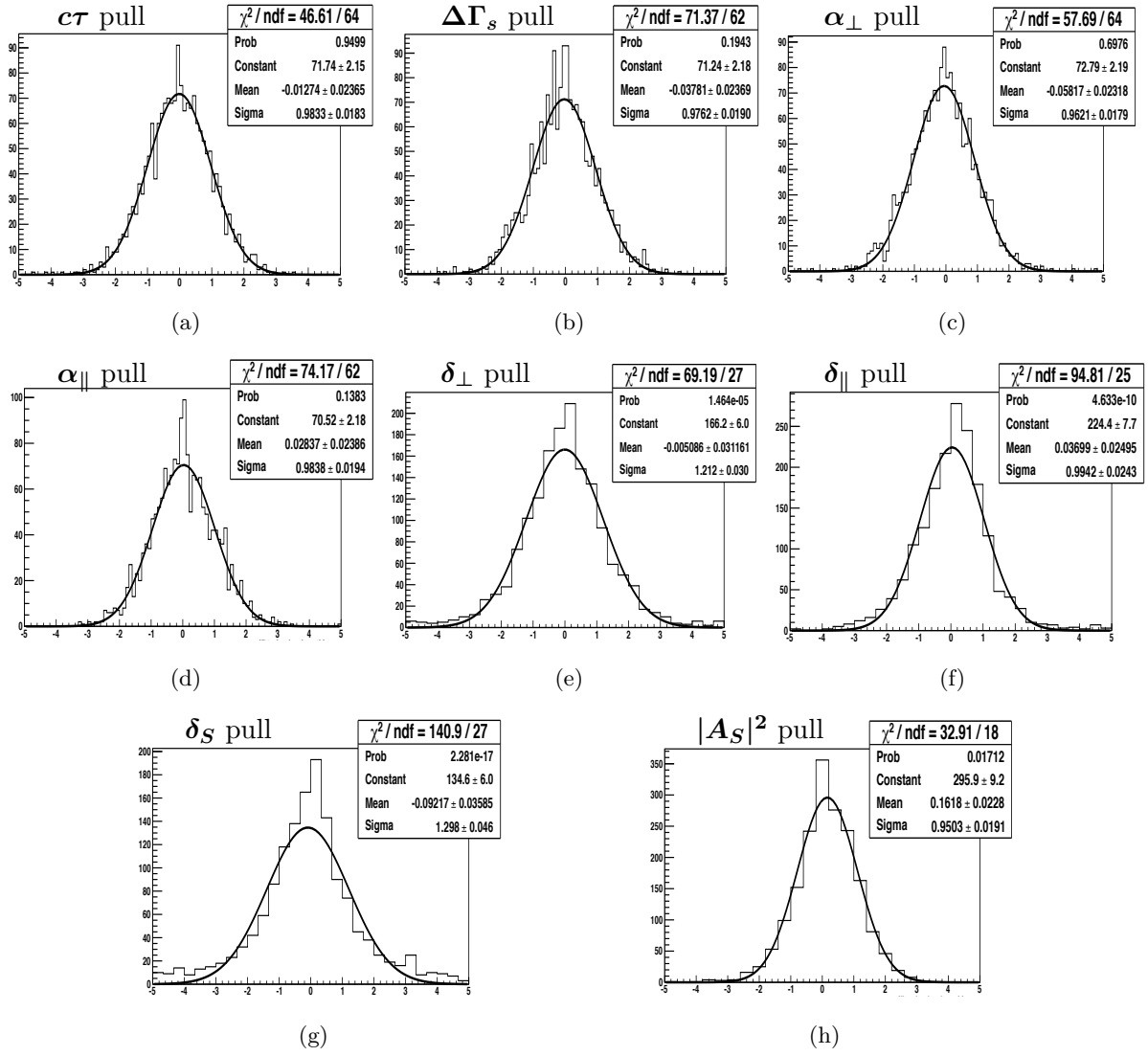


Figure E.3: Pull distributions for the main physics parameters of the β_s-fixed fit with random values of PDF parameters in the generation of the pseudo-experiments.

Appendix F

Variables correlation in the fit to the $J/\psi K^+K^-$ and K^+K^- masses.

In the simultaneous fit to the $J/\psi K^+K^-$ and K^+K^- mass distributions described in Sect. 7.4, we have assumed that the PDF in Eq. (7.4) is separable for the $J/\psi K^+K^-$ and the K^+K^- variables by neglecting correlations among the masses. We indeed expect a negligible correlation between the B_s^0, the $\phi(1020)$ and $f_0(980)$ signal, given the good resolution of the mass measurements (Fig. F.1). The combinatorial background have small correlations among the two variables (Fig. F.1), while a visible effect is expected for the $B^0 \to J/\psi K^+\pi^-$ component (Fig. F.1). Indeed, the shape of the B^0 spectrum reconstructed as the B_s^0 signal strongly depend on the momentum of pion in the mis-assignment of the mass hypothesis. Therefore, we check with simulation the impact of our approximation in the PDF.

We generate a set of pseudo-experiments according to the PDF, *i. e.*, with the factorization hypothesis. We fit the pseudo-experiments to have a reference value for each parameter of the residuals between generated and fitted values. Such references show unbiased estimates of the fit's parameters. Then, we generate a set of pseudo-experiments according to 2-dimensional histograms of the $J/\psi K^+K^-$ and K^+K^- masses, with the desired fraction for each component. Hence, the events are generated with the correlations among variables. We fit such pseudo-experiments with the usual PDF and compare the mean value of the residuals of each parameter with the corresponding residuals found in the reference (Fig. F.2). We observe deviations smaller than the fit uncertainties for the fitted fractions. Thus, given our cross-check purpose of this fit, we retain the simplified model of PDF.

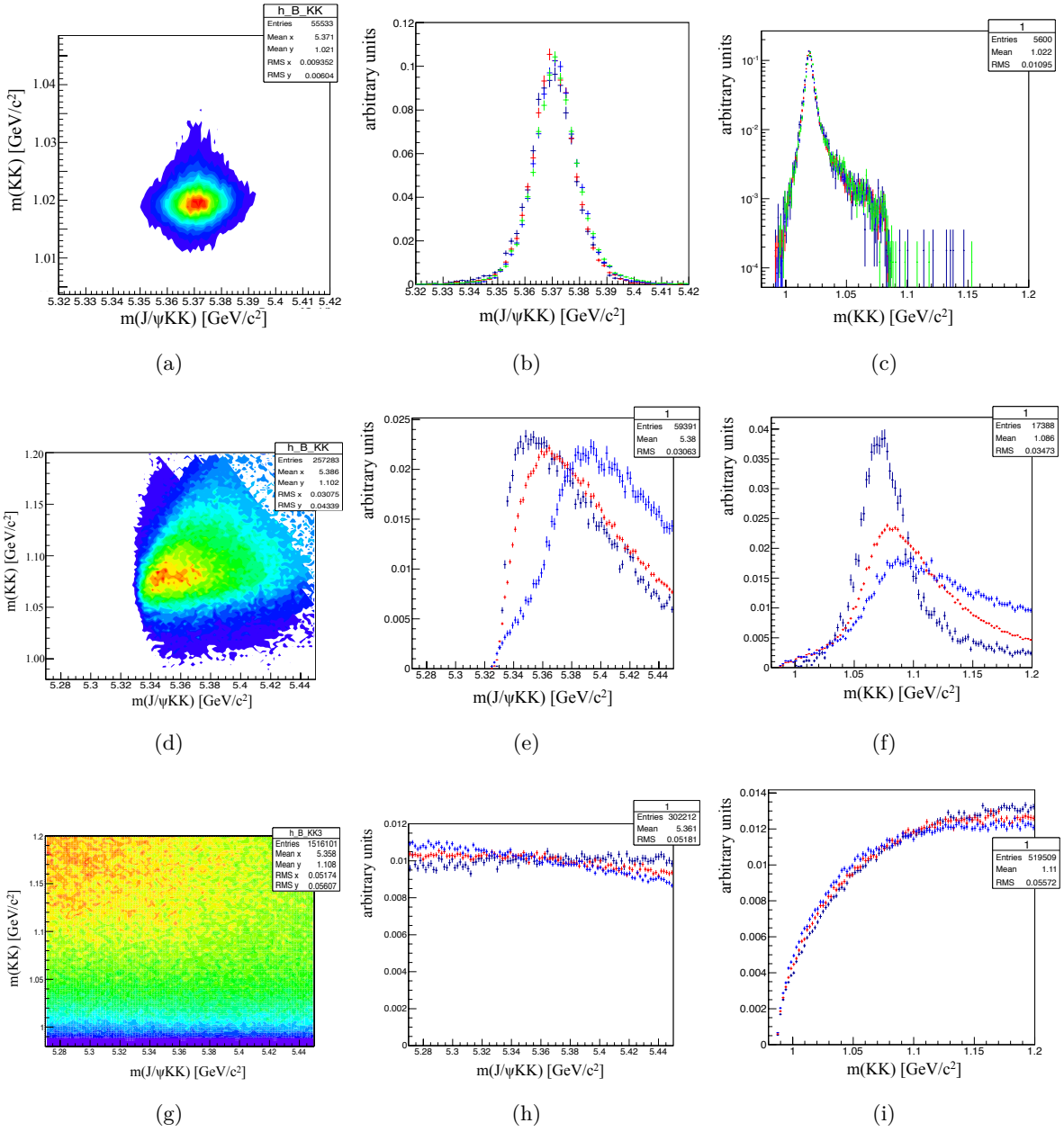
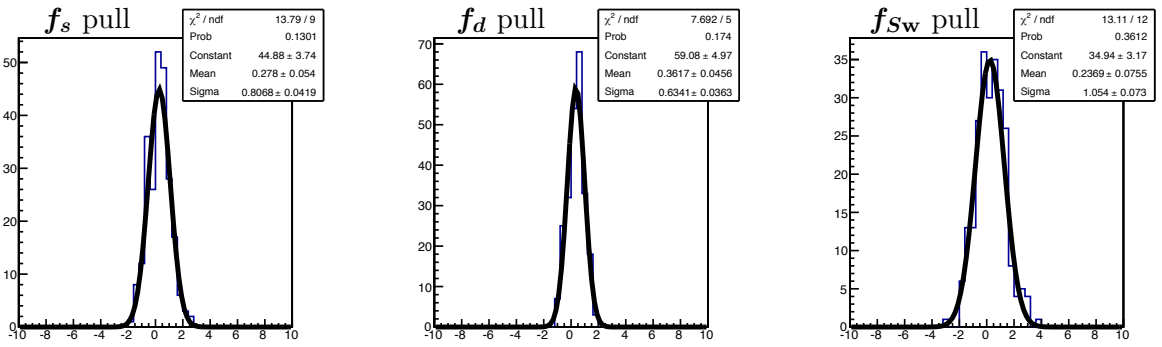
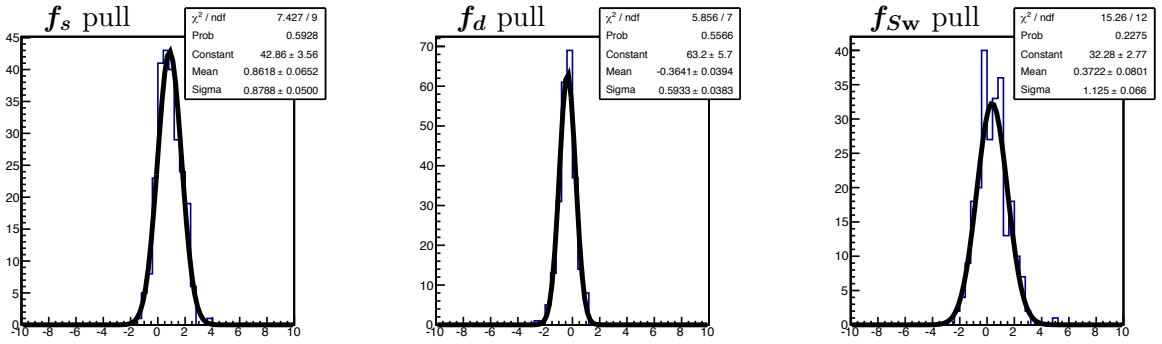


Figure F.1: Checks of the correlations between $J/\psi K^+K^-$ and K^+K^- masses for the different component of the data sample with simulated events. Scatter plot in $J/\psi K^+K^-$ and K^+K^- for the B_s^0–$\phi(1020)$signal (correlation factor 2%) (a), B^0–$K^+\pi^-$component (correlation 20%) (d), and the combinatorial background (correlation -3%) (g). Distribution of $J/\psi K^+K^-$ in slices of K^+K^- for B_s^0 (b), B^0 (e), and combinatorial background (h). Distribution of K^+K^- in slices of $J/\psi K^+K^-$ for $\phi(1020)$ (c), $K^+\pi^-$ (f), and combinatorial background (i).



(a)



(b)

Figure F.2: Pull distributions for 240 pseudo-experiments generated according to the PDF Eq. (7.4) (top row) and according to 2-dimensional histograms as described in the text (bottom row).

Bibliography

[1] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of the Bottom-Strange Meson Mixing Phase in the Full CDF Data Set*, Phys. Rev. Lett. **109**, 171802 (2012) [arXiv:1208.2967 (hep-ex)].

[2] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of Polarization and Search for CP Violation in $B_s^0 \to \phi\phi$ Decays*, Phys. Rev. Lett. **107**, 261802 (2011) [arXiv:1107.4999 (hep-ex)].

[3] G. Aad *et al.* (ATLAS Collaboration), *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716**, 1 (2012) [arXiv:1207.7214 (hep-ex)].

[4] S. Chatrchyan *et al.* (CMS Collaboration), *Observation of a new boson at a mass of 125* GeV *with the CMS experiment at the LHC*, Phys. Lett. **B716**, 30 (2012) [arXiv:1207.7235 (hep-ex)].

[5] R. Barbieri, *ICHEP2012 Physics Highlights*, (2012) [arXiv:1212.3440 (hep-ph)], contribution to the 36^{th} *International Conference on High energy Physics*, Melbourn, July 4–11, 2012.

[6] G. Isidori *et al.*, *Flavor Physics Constraints for Physics Beyond the Standard Model*, Ann. Rev. Nucl. Part. Sci. **60**, 355 (2010) [arXiv:1002.0900 (hep-ph)].

[7] M. Ciuchini and A. Stocchi, *Physics Opportunities at the Next Generation of Precision Flavor Physics*, Ann. Rev. Nucl. Part. Sci. **61**, 491 (2011) [arXiv:1110.3920 (hep-ph)].

[8] J. Beringer *et al.* (Particle Data Group), *Review of particle physics*, Phys. Rev. **D86**, 010001 (2012).

[9] Y. Amhis *et al.* (Heavy Flavor Averaging Group), *Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012*, (2012) [arXiv:1207.1158 (hep-ex)], and updates at http://www.slac.stanford.edu/xorg/hfag.

[10] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963).

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[12] J. Charles *et al.* (CKMfitter Group), *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41**, 1 (2005) [arXiv:hep-ph/0406184], updated results and plots available at http://ckmfitter.in2p3.fr.

[13] A. Lenz, *Theoretical update of B-Mixing and Lifetimes*, (2012) [arXiv:1205.1444 (hep-ph)].

[14] I. Dunietz *et al.*, *In pursuit of new physics with B_s decays*, Phys. Rev. **D63**, 114015 (2001) [arXiv:hep-ph/0012219 (hep-ph)].

[15] A. Datta *et al.*, *New Physics in $b \to s$ Transitions and the $B^0_{d,s} \to V_1 V_2$ Angular Analysis*, Phys. Rev. **D86**, 076011 (2012) [arXiv:1207.4495 (hep-ph)].

[16] T. Aaltonen *et al.* (CDF Collaboration), *First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in $B^0_s \to J/\psi\phi$ Decays*, Phys. Rev. Lett. **100**, 161802 (2008) [arXiv:0712.2397 (hep-ex)].

[17] V. Abazov *et al.* (D0 Collaboration), *Measurement of B^0_s mixing parameters from the flavor-tagged decay $B^0_s \to J/\psi\phi$*, Phys. Rev. Lett. **101**, 241801 (2008) [arXiv:0802.2255 (hep-ex)].

[18] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of the CP-Violating Phase $\beta_s^{J/\psi\phi}$ in $B^0_s \to J/\psi\phi$ Decays with the CDF II Detector*, Phys. Rev. **D85**, 072002 (2012) [arXiv:1112.1726 (hep-ex)].

[19] V. Abazov *et al.* (D0 Collaboration), *Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B^0_s \to J/\psi\phi$ in $8\,fb^{-1}$ of $p\bar{p}$ collisions*, Phys. Rev. **D85**, 032006 (2012) [arXiv:1109.3166 (hep-ex)].

[20] R. Aaij *et al.* (LHCb Collaboration), *Measurement of the CP-violating phase ϕ_s in the decay $B^0_s \to J/\psi\phi$*, Phys. Rev. Lett. **108**, 101803 (2012) [arXiv:1112.3183 (hep-ex)].

[21] G. Aad *et al.* (ATLAS Collaboration), *Time-dependent angular analysis of the decay $B^0_s \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase ϕ_s by ATLAS*, (2012) [arXiv:1208.0572 (hep-ex)].

[22] V. Abazov *et al.* (D0 Collaboration), *Measurement of the anomalous like-sign dimuon charge asymmetry with $9\,fb^{-1}$ of $p\bar{p}$ collisions*, Phys. Rev. **D84**, 052007 (2011) [arXiv:1106.6308 (hep-ex)].

[23] R. Aaij *et al.* (LHCb Collaboration), *Measurement of the polarization amplitudes and triple product asymmetries in the $B^0_s \to \phi\phi$ decay*, Phys. Lett. **B713**, 369 (2012) [arXiv:1204.2813 (hep-ex)].

[24] M. Dorigo, *Measurement of the polarization amplitudes of the $B^0_s \to \phi\phi$ decay at CDF II*, 2009, University of Trieste, master's thesis, available as FERMILAB-MASTERS-2009-02.

[25] A. Datta *et al.*, *Searching for New Physics with B-Decay Fake Triple Products*, Phys. Lett. **B701**, 357 (2011) [arXiv:1103.2442 (hep-ph)].

[26] M. Gronau and J. L. Rosner, *Triple product asymmetries in K, $D_{(s)}$ and $B_{(s)}$ decays*, Phys. Rev. **D84**, 096013 (2011) [arXiv:1107.1232 (hep-ph)].

[27] D. H. Perkins, *Introduction to High Energy Physics* (Cambridge University Press, Cambridge, 2000), 4^{th} edition.

[28] C. Quigg, *Unanswered Questions in the Electroweak Theory*, Ann. Rev. Nucl. Part. Sci. **59**, 505 (2009) [arXiv:0905.3187 (hep-ph)].

[29] M. S. Chanowitz, *Electroweak Symmetry Breaking: Unitarity, Dynamics, Experimental Prospects*, Ann. Rev. Nucl. Part. Sci. **38**, 323 (1988).

[30] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2**, 1285 (1970).

[31] J. Augustin *et al.* (SLAC-SP-017 Collaboration), *Discovery of a Narrow Resonance in e+ e- Annihilation*, Phys. Rev. Lett. **33**, 1406 (1974).

[32] J. Aubert *et al.* (E598 Collaboration), *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33**, 1404 (1974).

[33] F. Abe *et al.* (CDF Collaboration), *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74**, 2626 (1995) [arXiv:hep-ex/9503002].

[34] S. Abachi *et al.* (DØ Collaboration), *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74**, 2422 (1995) [arXiv:hep-ex/9411001].

[35] H. Albrecht *et al.* (ARGUS COLLABORATION), *Observation of B0 - anti-B0 Mixing*, Phys. Lett. **B192**, 245 (1987).

[36] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2**, 1285 (1970).

[37] M. Bona *et al.* (UTfit Collaboration), *First Evidence of New Physics in $b \longleftrightarrow s$ Transitions*, PMC Phys. **A3**, 6 (2009) [arXiv:0803.0659 (hep-ph)].

[38] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[39] A. Buras *et al.*, *Waiting for the top quark mass, $K^+ \to \pi^+ \nu\overline{\nu}$, $B^0_{(s)} - \overline{B}^0_{(s)}$ mixing and CP asymmetries in B decays*, Phys. Rev. **D50**, 3433 (1994) [arXiv:hep-ph/9403384].

[40] G. C. Branco *et al.*, *CP Violation* (Oxford Science Publications, 1999).

[41] G. Luders, *Proof of the TCP theorem*, Annals Phys. **2**, 1 (1957).

[42] R. Aaij *et al.* (LHCb Collaboration), *Determination of the sign of the decay width difference in the B_s^0 system*, Phys. Rev. Lett. **108**, 241801 (2012) [arXiv:1202.4717 (hep-ex)].

[43] A. Abulencia *et al.* (CDF Collaboration), *Observation of B_s^0-$\bar{B}_s^0$ Oscillations*, Phys.Rev.Lett. **97**, 242003 (2006) [arXiv:hep-ex/0609040 (hep-ex)].

[44] T. Aaltonen *et al.* (CDF collaboration), *Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays*, Phys. Rev. Lett. **100**, 121803 (2008) [arXiv:0712.2348 (hep-ex)].

[45] V. Abazov *et al.* (D0 Collaboration), *Measurement of the lifetime difference in the B_s^0 system*, Phys. Rev. Lett. **95**, 171801 (2005) [arXiv:hep-ex/0507084 (hep-ex)].

[46] R. Aaij *et al.* (LHCb Collaboration), *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays at LHCb*, LHCb-CONF-2012-002 (2012).

[47] C.-W. Chiang *et al.*, *New Physics in $B_s^0 \to J/\psi\phi$: A General Analysis*, JHEP **1004**, 031 (2010) [arXiv:0910.2929 (hep-ph)].

[48] V. M. Abazov *et al.* (D0 Collaboration), *Measurement of the semileptonic charge asymmetry in B^0 meson mixing with the D0 detector*, Phys. Rev. **D86**, 072009 (2012) [arXiv:1208.5813 (hep-ex)].

[49] R. Aaij *et al.* (LHCb Collaboration), *Measurement of the flavour-specific CP violating asymmetry $a_s l^s$ in B_s^0 decays*, (2012).

[50] V. Abazov *et al.* (D0 Collaboration), *Measurement of the semileptonic charge asymmetry using $B_s^0 \to D_s \mu X$ decays*, Phys. Rev. Lett. 110, **011801**, [7 pages] (2013) [arXiv:1207.1769 (hep-ex)].

[51] M. Beneke *et al.*, *Branching fractions, polarisation and asymmetries of $B \to VV$ decays*, Nucl. Phys. **B774**, 64 (2007) [arXiv:hep-ph/0612290 (hep-ph)].

[52] A. L. Kagan, *Polarization in$B \to VV$ decays*, Phys. Lett. **B601**, 151 (2004) [arXiv:hep-ph/0405134 (hep-ph)].

[53] S. Amato *et al.*, *LHCb's Sensitivity to New CP-Violating Phases in the Decay $B_s^0 \to \phi\phi$*, LHCb 2007-047 (2007).

[54] S. T'Jampens, *Etude de la violation de la symetrie CP dans les canaux charmonium-$K^\star(892)$ par une analyse angulaire complete dependante du temps*, Ph.D. thesis, University of Paris VI and VII, 2003, available as *BABAR* thesis THESIS-03/016.

[55] S. Stone and L. Zhang, *S-waves and the Measurement of CP Violating Phases in B_s^0 Decays*, Phys. Rev. **D79**, 074024 (2009) [arXiv:0812.2832 (hep-ph)].

[56] F. Azfar *et al.*, *Formulae for the Analysis of the Flavor-Tagged Decay $B_s^0 \to J/\psi\phi$*, JHEP **1011**, 158 (2010) [arXiv:1008.4283 (hep-ph)].

[57] A. Datta and D. London, *Triple-product correlations in $B \to V_1 V_2$ decays and new physics*, Int. J. Mod. Phys. **A19**, 2505 (2004) [arXiv:hep-ph/0303159 (hep-ph)].

[58] The CDF and D0 Collaborations, *Combination of D0 and CDF Results on $\Delta\Gamma_s$ and the CP-Violating Phase β_s*, CDF Note 9787, D0 Note 5928-CONF (2009).

[59] The LHCb Collaboration, *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays at LHCb*, LHCb-CONF-2012-002 (2012).

[60] D. Acosta *et al.* (CDF Collaboration), *First evidence for $B_s^0 \to \phi\phi$ decay and measurements of branching ratio and A_{CP} for $B^+ \to \phi K^+$*, Phys. Rev. Lett. **95**, 031801 (2005) [arXiv:hep-ex/0502044 (hep-ex)].

[61] M. Beneke *et al.*, *Branching fractions, polarisation and asymmetries of $B \to VV$ decays*, Nucl. Phys. **B774**, 64 (2007) [arXiv:hep-ph/0612290 (hep-ph)].

[62] H. Cheng and C. Chua, *QCD Factorization for Charmless Hadronic B_s^0 Decays Revisited*, Phys. Rev. **D80**, 114026 (2009) [arXiv:0910.5237 (hep-ph)].

[63] A. Ali *et al.*, *Charmless non-leptonic B_s^0 decays to PP, PV and VV final states in the pQCD approach*, Phys. Rev. **D76**, 074018 (2007) [arXiv:hep-ph/0703162 (HEP-PH)].

[64] B. Aubert *et al.* (*BABAR* Collaboration), *Rates, polarizations, and asymmetries in charmless vector-vector B meson decays*, Phys. Rev. Lett. **91**, 171802 (2003) [arXiv:hep-ex/0307026 (hep-ex)].

[65] B. Aubert *et al.* (*BABAR* Collaboration), *Time-Dependent and Time-Integrated Angular Analysis of $B \to \phi K_S \pi^0$ and $B \to \phi K^+ \pi^-$*, Phys. Rev. **D78**, 092008 (2008) [arXiv:0808.3586 (hep-ex)].

[66] K. Chen *et al.* (Belle Collaboration), *Measurement of branching fractions and polarization in $B \to \phi K^{(\star)}$ decays*, Phys. Rev. Lett. **91**, 201801 (2003) [arXiv:hep-ex/0307014 (hep-ex)].

[67] K. Chen *et al.* (BELLE Collaboration), *Measurement of polarization and triple-product correlations in $B \to \phi K^\star$ decays*, Phys.Rev.Lett. **94**, 221804 (2005) [arXiv:hep-ex/0503013 (hep-ex)].

[68] A. Datta *et al.*, *Testing Explanations of the $B \to \phi K^\star$ Polarization Puzzle*, Phys. Rev. **D76**, 034015 (2007) [arXiv:0705.3915 (hep-ph)].

[69] C. Huang *et al.*, *MSSM anatomy of the polarization puzzle in $B \to \phi K^\star$ decays*, Phys. Rev. **D73**, 034026 (2006) [arXiv:hep-ph/0511129 (hep-ph)].

[70] C. Chen and C. Geng, *Scalar interactions to the polarizations of* $B \to \phi K^\star$, Phys. Rev. **D71**, 115004 (2005) [arXiv:hep-ph/0504145 (hep-ph)].

[71] R. Aaij *et al.* (LHCb Collaboration), *First observation of the decay* $B_s^0 \to K^{\star 0}\bar{K}^{\star 0}$, Phys. Lett. **B709**, 50 (2012) [arXiv:1111.4183 (hep-ex)].

[72] S. D. Holmes (ed.), *Tevatron Run II Handbook*, FERMILAB-TM-2484 (1998).

[73] S. Holmes, R. S. Moore, and V. Shiltsev, *Overview of the Tevatron Collider Complex: Goals, Operations and Performance*, (2011) [arXiv:1106.0909 (physics.acc-ph)].

[74] R. Blair *et al.* (CDF Collaboration), *The CDF II detector: Technical design report*, FERMILAB-DESIGN-1996-01 (1996).

[75] C. S. Hill (On behalf of the CDF Collaboration), *Operational experience and performance of the CDF II silicon detector*, Nucl. Instrum. Meth. **A530**, 1 (2004).

[76] A. Sill (CDF Collaboration), *CDF Run II silicon tracking projects*, Nucl. Instrum. Meth. **A447**, 1 (2000).

[77] A. A. Affolder *et al.* (CDF Collaboration), *Intermediate silicon layers detector for the CDF experiment*, Nucl. Instrum. Meth. **A453**, 84 (2000).

[78] A. A. Affolder *et al.* (CDF Collaboration), *CDF central outer tracker*, Nucl. Instrum. Meth. **A526**, 249 (2004).

[79] P. Gatti, *Performance of the new tracking system at CDF II*, Ph.D. thesis, University of Pisa, 2001, available as FERMILAB-THESIS-2001-23.

[80] S. Menzemer, *Track reconstruction in the silicon vertex detector of the CDF II experiment*, Ph.D. thesis, Karlsruhe University, 2003, available as FERMILAB-THESIS-2003-52, IEKP-KA-2003-04.

[81] C. P. Hays *et al.*, *Inside-out tracking at CDF*, Nucl. Instrum. Meth. **A538**, 249 (2005).

[82] S. S. Yu *et al.*, *COT dE/dx Measurement and corrections*, CDF Note 6361 (2004).

[83] L. Oakes *et al.*, *Extended and improved dE/dx calibration for* $3\,fb^{-1}$ *analyses*, CDF Note 9592 (2009).

[84] D. Acosta *et al.* (CDF Collaboration), *A time-of-flight detector in CDF II*, Nucl. Instrum. Meth. **A518**, 605 (2004).

[85] M. Feindt *et al.*, *Particle Dependent Time-of-Flight Calibration*, CDF Note 9618 (2008).

[86] G. Ascoli *et al.*, *CDF Central Muon Detector*, Nucl. Instrum. Meth. **A268**, 33 (1988).

[87] C. Ginsburg (on behalf of the CDF Collaboration), *CDF Run 2 muon system*, Eur. Phys. J. **C33**, S1002 (2004).

[88] L. Balka *et al.* (CDF Collaboration), *The CDF Central Electromagnetic Calorimeter*, Nucl. Instrum. Meth. **A267**, 272 (1988).

[89] S. Bertolucci *et al.* (CDF Collaboration), *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Instrum. Meth. **A267**, 301 (1988).

[90] M. Albrow *et al.* (CDF Collaboration), *The CDF plug upgrade electromagnetic calorimeter: Test beam results*, Nucl. Instrum. Meth. **A480**, 524 (2002).

[91] M. Albrow *et al.*, *A preshower detector for the CDF plug upgrade: Test beam results*, Nucl. Instrum. Meth. **A431**, 104 (1999).

[92] G. Apollinari *et al.*, *Shower maximum detector for the CDF plug upgrade calorimeter*, Nucl. Instrum. Meth. **A412**, 515 (1998).

[93] F. Abe *et al.* (CDF Collaboration), *Measurement of the $\bar{p}p$ total cross-section at $\sqrt{s} = 546$ GeV and 1800 GeV*, Phys. Rev. **D50**, 5550 (1994).

[94] C. Avila *et al.* (E811 Collaboration), *A Measurement of the proton-antiproton total cross-section at $\sqrt{s} = 1.8$ TeV*, Phys. Lett. **B445**, 419 (1999).

[95] B. Ashmanskas *et al.* (CDF-II Collaboration), *The CDF silicon vertex trigger*, Nucl. Instrum. Meth. **A518**, 532 (2004) [arXiv:physics/0306169 (physics)].

[96] L. Ristori and G. Punzi, *Triggering on Heavy Flavors at Hadron Colliders*, Ann. Rev. Nucl. Part. Sci. **60**, 595 (2010).

[97] J. Marriner, *Secondary vertex fit with mass and pointing constraints (CTVMFT)*, CDF Note 1996 (1993).

[98] A. Di Canto, *Bstntuple's User Guide*, CDF Note 10711 (2012).

[99] P. Murat *et al.*, Stntuple, cdfcodebrowser.fnal.gov/CdfCode/source/Stntuple/.

[100] E. Gerchtein and M. Paulini, in *eConf. C0303241 TUMT005* (PUBLISHER, 2003), [physics/0306031].

[101] K. Anikeev *et al.*, *Description of BGenerator II*, (1999), cDF Note 5092.

[102] D. J. Lange, *The EvtGen Particle Decay Simulation Package*, Nucl. Instrum. Meth. (2001).

[103] R. Braun *et al.*, *GEANT: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual*, 1978, CERN-DD-78-2-REV.

[104] P. A. Movilla-Fernandez, in *Proceedings of the XII International Conference on Calorimetry in High Energy Physics (CALOR 06)*, edited by R. Yoshida S. Magill (PUBLISHER, 2006), [physics/0608081].

[105] R. Veenhof, *GARFIELD, recent developments.*, Nucl. Instrum. Methods **A419**, (1998).

[106] M. Gold, *Description of the Parameterized Charge Deposition Model*, CDF Note 5871 (2002).

[107] B_s^0-Mixing-Group, *Combined opposite side flavor tagger*, CDF Note 8314 (2006).

[108] G. Giurgiu *et al.*, *Muon B Flavor Tagging - A Likelihood Approach*, CDF Note 7043 (2004).

[109] A. Schmidt, *Combined B Flavour Tagging with Artificial Neural Networks*, Ph.D. thesis, Karlsruhe University, 2006, available at http://www-ekp.physik.uni-karlsruhe.de/pub/web/thesis/iekp-ka2006-11.pdf.

[110] B_s^0-Mixing-Group, *Same Side Kaon Tagging Studies in Fully Reconstructed Decays*, CDF Note 7979 (2006).

[111] B_s^0-Mixing-Group, *Same side Kaon Tagging Studies in Fully Reconstructed Decays on the New Data*, CDF Note 8182 (2006).

[112] B_s^0-Mixing-Group, *Optimization of the Same Side Kaon tagging algorithm combining PID and kinematic variables*, CDF Note 8344 (2006).

[113] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, ArXiv Physics e-prints (2004) [:arXiv:physics/0402093].

[114] M. Feindt *et al.*, *Measurement of Mixing Frequency and Amplitude of Neutral B-Mesons*, CDF Note 10025 (2010).

[115] M. Bishai *et al.*, *Run-II Dimuon Trigger Optimization and Efficiency Measurement*, CDF Note 6004 (2002).

[116] G. Giurgiu, *B Flavor Tagging Calibration and Search for B(s) Oscillations in Semileptonic Decays with the CDF Detector at Fermilab*, Ph.D. thesis, Carnegie Mellon University, 2005, available as FERMILAB-THESIS-2005-41.

[117] M. Milnik, *Measurement of the lifetime difference and CP-violating phase in $B_s^0 \to J/\psi\phi$ decays*, Ph.D. thesis, Karlsruhe University, 2007, available as FERMILAB-THESIS-2007-38.

[118] F. James, *Statistical Method in Experimental Physics* (World Scientific Publishing Co. Pte. Ltd., 2006).

[119] M. Ablikim *et al.* (BES Collaboration), *Resonances in $J/\psi \to \phi\pi^+\pi^-$ and ϕK^+K^-*, Phys. Lett. **B607**, 243 (2005) [arXiv:hep-ex/0411001 (hep-ex)].

[120] B. Aubert *et al.* (BABAR Collaboration), *Ambiguity-free measurement of* $\cos(2\beta)$*: Time-integrated and time-dependent angular analyses of $B \to J/\psi K\pi$*, Phys. Rev. **D71**, 032005 (2005) [arXiv:hep-ex/0411016 (hep-ex)].

[121] F. James *et al.*, *Minuit, Function Minimization and Error Analysis*, Minuit Reference Manual (2007).

[122] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of b Hadron Lifetimes in Exclusive Decays Containing a J/ψ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,\text{TeV}$*, Phys. Rev. Lett. **106**, 121804 (2011).

[123] L. Oakes, *Measurement of the CP violating phase β_s in $B_s^0 \to J/\psi\phi$ decays*, Ph.D. thesis, University of Oxford, available as fermilab-thesis-2010-60.

[124] K. Anikeevs, *Measurement of the lifetimes of B_s^0 meson mass eigenstates.*, Ph.D. thesis, Massachusetts Institute of Technology, available as fermilab-thesis-2004-12.

[125] R. L. Berger and D. Boos, *P values maximized over a confidence set for the nuisance parameter*, Journal of the American Statistical Association **89**, 1012 (1994).

[126] E. J. Thomson *et al.*, *Online track processor for the CDF upgrade*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).

[127] A. Abulencia *et al.* (CDF Collaboration), *The CDF II eXtremely fast tracker upgrade*, Nucl. Instrum. Meth. **A572**, 358 (2007).

[128] C. R. Rao, *Information and accuracy attainable in the estimation of statistical parameters*, Bull. Calcutta Math. Soc. **37**, (1945).

[129] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, 1946).

[130] B. Di Ruzza *et al.*, *Study of the $B_s^0 \to \phi\phi$ decay and measurement of its Branching Ratio*, CDF Internal Note **9743**, (2009).

[131] M. Dorigo *et al.*, *Measurement of the Polarization Amplitudes for the $B_s^0 \to \phi\phi$ Decays*, CDF Internal Note **10073**, (2010).

[132] M. Gronau and J. L. Rosner, *Flavor symmetry for strong phases and determination of β_s, $\Delta\Gamma_s$ in $B_s^0 \to J/\psi\phi$*, Phys. Lett. **B669**, 321 (2008) [arXiv:0808.3761 (hep-ph)].

[133] R. Aaij *et al.* (LHCb Collaboration), *Measurement of the CP-violating phase ϕ_s in $B_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Lett. **B713**, 378 (2012) [arXiv:1204.5675 (hep-ex)].

[134] C. A. Chavez *et al.*, *Charm meson mixing: An experimental review*, Int. J. Mod. Phys. **A27**, 1230019 (2012) [arXiv:1209.5806 (hep-ex)].

[135] S. M. Flatte, *Coupled - Channel Analysis of the $\pi\eta$ and K anti-K Systems Near K anti-K Threshold*, Phys. Lett. **B63**, 224 (1976).