THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Constituants Elémentaires Systèmes Complexes

présentée par

Gian Piero DI GIOVANNI

pour obtenir le grade de

Docteur de l'Université "Pierre Et Marie Curie"

Mixing and CP Violation in the B_s^0 Meson System at CDF

Mélange et Violation de CP dans le Systéme des Mésons B_s^0 *à CDF*

Soutenue le 27 Juin 2008 devant le jury composé de:

M	Philip Patrick ALLPORT	
Mme	Paula EEROLA	
M	Ivan Krešimir FURIĆ	
M	Yosef NIR	
M	Marco RESCIGNO	
M	Alberto RUIZ JIMENO	
Mme	Aurore SAVOY-NAVARRO	Directeur de thèse
M	Pascal VINCENT	

Contents

List of Figures

List of Tables

Introduction

The Standard Model is the theoretical framework developed to describe the known elementary particles and their interactions. Over the past decades, it has been stringently tested experimentally, proving itself very successful, so far. Its predictions involving the fundamental building blocks of matter, the quarks, the leptons and the vector bosons have been confirmed. In this picture, the physics of the b quark represents one of the most active research areas in high energy physics to challenge the Standard Model predictions and limitations. B_s^0 mesons are very rare particles and the Tevatron accelerator, which collides protons and anti-protons at a center-of-mass energy of 1.96 TeV, provides a simultaneous access to large samples of strange and non-strange b-mesons necessary for precision measurements, offering a great opportunity to study the B_s^0 flavor sector, before the start-up of CERN Large Hadronic Collider (LHC).

The B_s^0 ($\overline{B}_s^0$) mesons, formed by a quark $\bar{b}(b)$ and a quark $s(\bar{s})$, are not pure mass eigenstates. The mesons can then oscillate into its antiparticles via weak interacting processes, whose quark coupling is governed by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. A precise measurement of the B_s^0 mixing oscillation frequency would indeed offer a powerful constraint to the Standard Model: a value different from the predicted one would be a clear indication of New Physics which contributes with additional processes to modify the magnitude of the mixing amplitude.

The meson flavor oscillation, or mixing, is a very well established phenomenon in particle physics since it was first observed in the kaon system [1]. For the B mesons case, the rate at which the $B - \overline{B}$ transformations occur are characterized by the mass difference of the two mass eigenstates, denoted as Δm. While the B^0 oscillation was precisely determined at the B factories [2, 3], the B_s^0 oscillation was recently observed at CDF [4]. Part of this thesis is, indeed, devoted to the description of the analysis which, ending a 20-years-long effort, leaded to the definitive Δm_s measurement. The Standard Model expects an high value for the B_s^0 oscillation frequency, compared to the analogous B^0 mixing, which makes its determination extremely demanding. Large samples and excellent detector performances, particularly challenging at a $p\bar{p}$ collider, are a requirement to accomplish such measurement. The data are recorded by the CDF II detector (Collider Detector at Fermilab Run II) installed in one of the two collision points of the Tevatron accelerator. CDF II, exploiting several novel aspects of its multi-purpose detector, is designed to pursue a large spectrum of physics phenomena studies. In detail, the highly sophisticated trigger development, based on the excellent performances of the tracking system (a combination of silicon layers and

drift chambers) made possible to collect, at CDF, the world largest fully reconstructed B_s^0 hadronic decays data sample. Thanks to its good precision in the determination of the proper-decay-length and resolution, the latter provide most part of the statistical power to the B_s^0 mixing analysis and played a key role in the competition for the final Δm_s measurement [1].

In addition to the measurement of the oscillation frequency, the lifetime difference between the two eigenstates and the phase β_s, responsible for mixing-induced CP asymmetries in the B_s^0 decay modes, are of great interest in the determination of the B_s^0 system properties. The measurement of the CP violation effect is, in fact, a prerequisites for understanding the imbalance between matter and anti-matter: the Big Bang theory predicts to have produced an equal amount of matter and anti-matter, but our Universe is matter-dominated. So far, the phenomenology of CP violation in B^0 and B^+ decays has been explained utilizing the single source of CP violation arising from the CKM mechanism within the Standard Model and it has been proven to be an extremely successful description. However, a comparable experimental knowledge in B_s^0 decays has been lacking. In this dissertation, I will present the first time-dependent angular analysis of the $B_s^0 \to J/\psi\phi$ decay mode, when the initial state of the B_s^0 meson is identified exploiting the flavor tagging information. Such information allows to separate the time evolution of mesons originally produced as B_s^0 or $\bar{B}_s^0$, to extract separate lifetimes for the mass eigenstates (B_s^L and B_s^H, where the superscripts L and H stay for "light" and "heavy"), then to measure the width difference, $\Delta\Gamma_s \equiv \Gamma_s^L - \Gamma_s^H = 1/\tau_{B_s^L} - 1/\tau_{B_s^H}$, and to measure the CP violating phase β_s. The latter, responsible for CP violating effects in $B_s^0 \to J/\psi\phi$ decays, is defined as $-2\beta_s = -2\arg(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*})$ where V_{ij} are the elements of the already mentioned CKM quark mixing matrix. Since it is expected to be tiny in the Standard Model predictions, $2\beta_s^{SM} \sim \mathcal{O}(0.04)$, the measurement of sizable value of $2\beta_s$ inconsistent with zero would indicate New Physics. These quantities are then extremely useful for over-constraining the Unitarity Triangle and thereby infer if CP violation in the quark sector is entirely explained within the Standard Model.

The two analyses presented in the thesis, the B_s^0 mixing analysis and the $B_s^0 \to J/\psi\phi$ angular analysis, share most of the technical implementations and features. Thus, my choice was to pursue in parallel the common aspects of the analyses, avoiding, whenever possible, repetitions. Each Chapter is split in two part, the first one dedicated to the B_s^0 mixing analysis and the second one describing the angular analysis on the $B_s^0 \to J/\psi\phi$ decay mode. They are organized as follows.

In Chapter 1 we present the theoretical framework of the B_s^0 neutral mesons system. After a general introduction on the Standard Model, we focus on the quantities which are relevant to the Δm_s measurement and the CP violation phenomena, un-

[1] Before the definitive Δm_s observation reported in the dissertation DØ collaboration released an analysis which set a two-sided bound on B_s^0 oscillation frequency at 90% C.L. [5]

derlying the details concerning the study of pseudo-scalar to vector vector decays, $P \to VV$, which allow to carry out an angular analysis. A discussion on the implication of the performed measurements in the search of physics beyond the Standard Model is presented.

The accelerator facilities and the CDF II detector are reported in Chapter 2. While describing the detector, more emphasis is given to the components fundamental to perform B physics analyses at CDF.

The Chapter 3 is focused on the reconstruction and selection of the data samples. The Chapter starts with a description of the on-line trigger requirements, according to the B_s^0 sample considered, followed by the offline selection criteria implemented to reconstruct B_s^0 semileptonic and hadronic decays, fully and partially reconstructed, for the B_s^0 mixing analysis, as well as the $\mathrm{B}_s^0 \to J/\psi\phi$ decay mode for the angular analysis.

The subsequent Chapter 4 is dedicated to the revision of the technical ingredients needed in the final analyses. The B_s^0 mixing elements are firstly described. The methodology historically used in the oscillation searches, the "amplitude scan", is here introduced together with the calibration of the proper-decay-time resolution and the flavor tagging algorithms, in particular a closer examination of the same-side tagger performances is given. The $\mathrm{B}_s^0 \to J/\psi\phi$ angular analysis elements description then follows, focusing on the performances and the eventual differences with respect to the B_s^0 oscillation search.

The final results of the analyses are obtained with the use of an unbinned likelihood fitting framework: Chapter 5 presents the general principles behind this methodology and a description of both the maximum likelihood fitters employed.

Chapter 6 contains the conclusive results on the B_s^0 analyses. They are presented in an historical fashion: the measurement of the B_s^0 oscillation frequency is followed by the first flavor tagged $\Delta\Gamma_s$ and β_s measurements. The impact and the constraints on the parameters of the flavor model is part of the discussion in the Chapter.

As cross-check of the B_s^0 angular analysis, the $\mathrm{B}^0 \to J/\psi K^{*0}$ decay mode has been additionally studied. Its angular analysis shows a competitive sensitivity with the B factories in measuring the parameters which define the decay. Not only this contributes to enforce the reliability of the entire framework, but it constitutes an excellent result by itself. Thus, we devote the entire Chapter 7 to the sole discussion of the angular analysis of the $\mathrm{B}^0 \to J/\psi K^{*0}$ decay mode.

Chapter 1

Theoretical Overview of B_s^0 Neutral Mesons System

In this chapter an introduction to the various quantities entering the time evolution of B_s^0 mesons and their decay amplitudes is given. The effects of the measurement of these parameters in the current model of particle physics are presented and their implications in selected scenarios of new physics are reviewed.

1.1 Matter in the Standard Model

The Standard Model of Particle Physics (SM) provides, at present, the best description of the properties of elementary particles and their interactions. It is defined by a gauge group, $\mathrm{SU}_{321} = \mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$, which describes the symmetries of the theory. The group is directly factorisable and the (local) symmetries which correspond to the three factors explicitly written above are color, weak isospin, and hypercharge. The transformations of the fields which describe fundamental particles are governed by the representations of the groups which are assigned to them. Matter is classified in three families of quarks:

$$\begin{aligned} &Q = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \\ &u_R\,, d_R\,, c_R\,, s_R\,, t_R\,, b_R\,, \end{aligned} \tag{1.1.1}$$

usually referred to as "up", "down", "charm", "strange", "top", and "bottom"-type quarks, and leptons:

$$\begin{aligned} &L = \begin{pmatrix} \nu_L^e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_L^\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_L^\tau \\ \tau_L \end{pmatrix}, \\ &e_R\,, \mu_R\,, \tau_R\,, \end{aligned} \tag{1.1.2}$$

where the subscripts L and R indicate left- and right-handed fields, doublets and singlets, respectively, with respect to transformations of the SU(2) component of SU_{321}. Table 1.1 summarizes the SU_{321} quantum numbers of the fields which experience

Field	SU(3)	SU(2)	U(1)
Q	3	2	$+1/6$
u_R, c_R, t_R	$\overline{3}$	1	$-2/3$
d_R, s_R, b_R	$\overline{3}$	1	$+1/3$
L	1	2	$+1/2$
e_R, μ_R, τ_R	1	1	-1

TABLE 1.1: Quantum numbers of matter. Right-handed neutrinos would have quantum numbers equal to (1,1,0), corresponding to the representation provided by the identity.

		Mass [GeV/c^2]	Charge
Quarks	u	1.5 to $3.0 \cdot 10^{-3}$	$\frac{2}{3}$
	d	3 to $7 \cdot 10^{-3}$	$-\frac{1}{3}$
	c	1.25 ± 0.09	$\frac{2}{3}$
	s	$95 \pm 25 \cdot 10^{-3}$	$-\frac{1}{3}$
	t	$174.2 \pm 3.3^{\dagger}$	$\frac{2}{3}$
	b	4.70 ± 0.07	$-\frac{1}{3}$
Leptons	ν_e	$< 225 \cdot 10^{-9}$ CL 95%	0
	e	$0.51099092 \pm 0.00000004 \cdot 10^{-3}$	-1
	ν_μ	$< 0.19 \cdot 10^{-3}$ CL 90%	0
	μ	$105.658369 \pm 0.000009 \cdot 10^{-3}$	-1
	ν_τ	$< 18.2 \cdot 10^{-3}$ CL 95%	0
	τ	$1776.99\ ^{+0.29}_{-0.26} \cdot 10^{-3}$	-1

TABLE 1.2: The families of matter in the SM. The latest measurements and fits are reported from Reference [6].
† Direct observation of top events.

gauge interactions in the SM. The right-handed counterpart of neutrinos ν is not included because it would transform trivially with respect to the entire group and thus have no gauge interactions. The properties of the fundamental components of matter are described in Table 1.2.

The gauge structure of particle interactions in the SM has been verified by many experiments, while the exploration of the flavor sector has not been as comprehensive. Interactions which couple quarks belonging to different families are mediated by W bosons. In the formalism of the SM, it is possible to describe the phenomenon by replacing the lower terms of the three quark doublets of SU(2) in Equation 1.1.1 with linear combinations of them, obtaining:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.1.3}$$

The matrix V contains the parameters that govern quark mixing, and relates the

physical quarks, the mass eigenstates d, s, and b, to the flavor eigenstates, indicated by the primed notation, which represent the states participating in charged-current weak interactions. This matrix is usually referred to as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7, 8].

The conservation of probability requires the CKM matrix to be unitary. This constraint, in the case of three families of quarks, leaves 3^2 free parameters in a 3×3 matrix, only four of which constitute physical degrees of freedom. The freedom to define arbitrary phases for the quark fields allows for the elimination of other $2 \times 3 - 1$ parameters, which are unphysical phases. The four physical parameters can be chosen to be three real angles and one complex phase, which is responsible of CP-violating effects in the SM. Another common representation of the CKM matrix uses the Wolfenstein parameters λ (the sine of the Cabibbo angle), A, ρ, and η [9]. The CKM matrix is traditionally expressed as a power series in terms of λ:

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{1.1.4}$$

Because λ is about 0.2, the power series converges rapidly. The expression above shows that the CKM matrix is almost diagonal, and that off-diagonal terms decrease with powers of λ the further they are from the diagonal.

The condition of unitarity is expressed as follows:

$$\begin{aligned} \sum_k V_{ki} V_{kj}^* &= \delta_{ij} \quad k \in u,c,t \quad i,j \in d,s,b\,, \\ \sum_i V_{ki} V_{li}^* &= \delta_{kl} \quad k,l \in u,c,t \quad i \in d,s,b\,. \end{aligned} \tag{1.1.5}$$

These equations produce a set of six independent expressions which equate the sum of three complex numbers to zero or unity, and are geometrically equivalent to triangles in the complex plane. The expression obtained above with $i = d$ and $j = b$ is of particular interest because the three terms which appear in it are of the same order in λ. It thus represents a triangle the sides of which are of about the same size, due to the structure of the CKM matrix. The equation is explicitly:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,. \tag{1.1.6}$$

The expression which is obtained by dividing the equation above by its second term defines the *Unitarity Triangle*. A sketch of the Unitarity Triangle is shown in Figure 1.1. The three angles are commonly called α, β, and γ (or ϕ_2, ϕ_1, and ϕ_3) and are related to the CKM matrix elements as follows:

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right)\,, \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)\,, \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)\,. \tag{1.1.7}$$

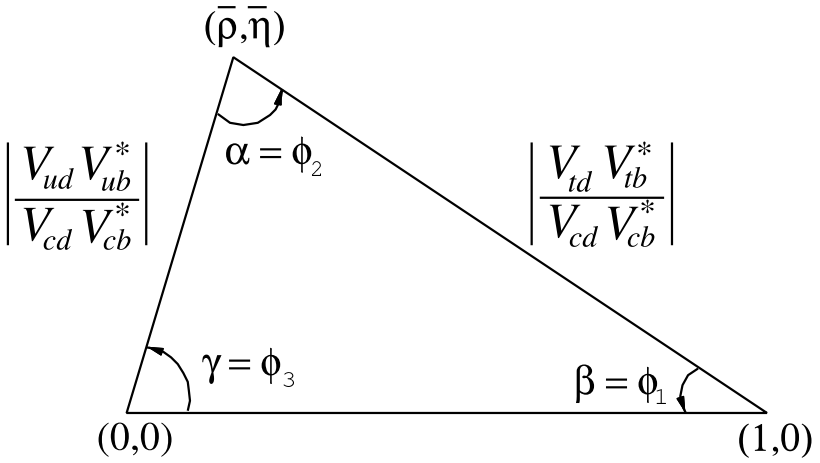

FIGURE 1.1: A sketch of the Unitarity Triangle.

It is convenient to define the rescaled Wolfenstein parameters $\bar{\rho}$ and $\bar{\eta}$ as follows:

$$\bar{\rho} + i\bar{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,. \tag{1.1.8}$$

This definition is phase-convention independent, and ensures that the matrix V written in terms of A, λ, $\bar{\rho}$, and $\bar{\eta}$ is unitary to all orders in λ. In terms of $\bar{\rho}$ and $\bar{\eta}$, the following relations hold:

$$\tan\alpha = \frac{\bar{\eta}}{\bar{\eta}^2 - \bar{\rho}(1-\bar{\rho})}\,, \quad \tan\beta = \frac{\bar{\eta}}{1-\bar{\rho}}\,, \quad \tan\gamma = \frac{\bar{\eta}}{\bar{\rho}}\,. \tag{1.1.9}$$

The length of the sides of the normalized Unitarity Triangle are given by:

$$R_c \equiv 1 \tag{1.1.10}$$

$$R_u \equiv \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\bar{\rho}^2+\bar{\eta}^2} = 1 - \frac{\lambda^2}{2}\frac{1}{\lambda}\left|\frac{V_{ud}}{V_{cb}}\right| \tag{1.1.11}$$

$$R_t \equiv \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\bar{\rho})^2+\bar{\eta}^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{ts}}\right| \tag{1.1.12}$$

The presence of CP violating effects in the SM is indicated by any of the three angles being different from zero or π. In fact CP violation is guaranteed by the presence of the phase η: in SM any quantities describing CP violation is expressed as a function of η. The measurements of the parameters A, λ, $\bar{\rho}$, and $\bar{\eta}$ reported by the latest analyses are collected in Table 1.3. It is worth noting that λ and A are known with a considerably higher precision than $\bar{\rho}$ and $\bar{\eta}$. Table 1.4 summarizes the current measurements of the angles of the Unitarity Triangle.

Parameter	Value [6]
λ	0.2272 ± 0.0010
A	$0.818\ ^{+0.007}_{-0.017}$
$\bar{\rho}$	$0.221\ ^{+0.064}_{-0.028}$
$\bar{\eta}$	$0.340\ ^{+0.017}_{-0.045}$

TABLE 1.3: Results of the latest fits for the CKM parameters in the Wolfenstein representation.

Parameter	Value [6]
α or ϕ_2	$(99\ ^{+13}_{-8})^\circ$
$\sin 2\beta$ or $\sin 2\phi_1$	0.0687 ± 0.032
γ or ϕ_3	$(63\ ^{+15}_{-12})^\circ$

TABLE 1.4: Results of the latest fits for the angles of the Unitarity Triangle.

1.2 The Neutral B Meson System

In this section the principle of the neutral $\mathrm{B}-\overline{\mathrm{B}}$ system time evolution are introduced. The details concerning the specific B_s^0 system will follow the general discussion.

Neutral B^0 mesons are composed of an $\bar{b}$ anti-quark and d quark. Similarly B_S^0 mesons are composed of an $\bar{b}$ anti-quark and an s quark:

$$|B^0\rangle = |\bar{b}d\rangle \qquad |\bar{B}^0\rangle = |b\bar{d}\rangle$$
$$|B_s^0\rangle = |\bar{b}s\rangle \qquad |\bar{B}_s^0\rangle = |b\bar{s}\rangle$$

Given the description we present applies equally to B_d and B_s mesons, we refer, for convenience, to B and $\overline{\mathrm{B}}$ to indicate the eigenstates of the strong interaction, i.e., the pure $\bar{b}q$ and $b\bar{q}$ states (here $q = d, s$), while B_H and B_L will represent the mass eigenstates.

In absence of weak flavor-changing interaction the $|\mathrm{B}\rangle$ and $|\overline{\mathrm{B}}\rangle$ would have been eigeinstates of the hamiltonian. With the inclusion of weak terms, the Hamiltonian is no longer diagonal in the $|B\rangle$, $|\overline{\mathrm{B}}\rangle$ basis, and transitions between the states are allowed. The phenomenon is known as B meson mixing.

The time evolution of the $\mathrm{B}-\overline{\mathrm{B}}$ system is governed by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} |B(t)\rangle \\ |\overline{\mathrm{B}}(t)\rangle \end{pmatrix} = \mathcal{H}\begin{pmatrix} |B(t)\rangle \\ |\overline{\mathrm{B}}(t)\rangle \end{pmatrix} \tag{1.2.1}$$

$$\text{where} \quad \mathcal{H} = \begin{pmatrix} m & M_{12} \\ M_{12}^* & m \end{pmatrix} + \frac{i}{2}\begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix} \tag{1.2.2}$$

with the 2×2 hermitian mass and decay matrices. The diagonal elements of the Hamiltonian describe the mass and decay width of the flavor eigenstates. CPT in-

variance guarantees that the two eigenstates have the same mass m and decay width Γ. The off-diagonal terms represent virtual (M_{12}) and real (Γ_{12}) particle-antiparticle transitions and, when different from zero, imply that mass and flavor eigenstates are not the same. The Hamiltonian $\mathcal{H}$ is diagonalized, by definition, in the basis of its eigenstates, B_H and B_L, which have definite mass and width ($\Gamma = 1/\tau$, where τ indicates the lifetime).

The mass eigeinstates at time $t = 0$ are linear combination of $|B\rangle$ and $|\bar{B}\rangle$:

$$|\mathrm{B}_L\rangle = p|\mathrm{B}\rangle + q|\overline{\mathrm{B}}\rangle\,, \tag{1.2.3}$$
$$|\mathrm{B}_H\rangle = p|\mathrm{B}\rangle - q|\overline{\mathrm{B}}\rangle\,, \tag{1.2.4}$$

where:

$$|p|^2 + |q|^2 = 1 \quad \text{and} \quad \frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}\,. \tag{1.2.5}$$

Omitting the calculation details, the time evolution of flavor state may be written as:

$$|\mathrm{B}(t)\rangle = g_+(t)|\mathrm{B}(0)\rangle + \frac{q}{p}g_-(t)|\overline{\mathrm{B}}(0)\rangle\,, \tag{1.2.6}$$
$$|\overline{\mathrm{B}}(t)\rangle = \frac{p}{q}g_-(t)|\mathrm{B}(0)\rangle + g_+(t)|\overline{\mathrm{B}}(0)\rangle\,, \tag{1.2.7}$$

where:

$$g_\pm(t) = \frac{1}{2}\left[e^{-\left(im_L + \frac{1}{2}\Gamma_L\right)t} \pm e^{-\left(im_H + \frac{1}{2}\Gamma_H\right)t}\right]\,. \tag{1.2.8}$$

Particular interest is given to the determination of the probability densities $\mathcal{P}_{\mathrm{B}\to\overline{\mathrm{B}}}(t)$ and $\mathcal{P}_{\mathrm{B}\to\mathrm{B}}(t)$ to observe flavor eigenstates produced at $t = 0$ which decay with the opposite or the same flavor, respectively, at time t. In the limit of $|q/p| = 1$ and $(\Gamma_L - \Gamma_H)/\Gamma$ negligibly small, both of which are good approximations for the B^0 and B_s^0 mesons, the probability densities are given by:

$$\mathcal{P}_{\mathrm{B}\to\overline{\mathrm{B}}}(t) = \mathcal{P}_{\overline{\mathrm{B}}\to\mathrm{B}}(t) = \frac{\Gamma}{2}e^{-\Gamma t}\left[1 - \cos\left(\Delta m t\right)\right]\,, \tag{1.2.9}$$
$$\mathcal{P}_{\mathrm{B}\to\mathrm{B}}(t) = \mathcal{P}_{\overline{\mathrm{B}}\to\overline{\mathrm{B}}}(t) = \frac{\Gamma}{2}e^{-\Gamma t}\left[1 + \cos\left(\Delta m t\right)\right]\,, \tag{1.2.10}$$

with

$$\Gamma = \frac{\Gamma_L + \Gamma_H}{2} = \frac{1}{\tau_{B_S}} \quad \text{and} \quad \Delta m = m_H - m_L. \tag{1.2.11}$$

The frequency of flavor transitions corresponds, as explicitily shown, to the mass difference between the two mass eigeinvalues of the system, therefore Δm constitues the target observable of a time dependent flavor oscillation measurement. With the convention $\hbar = c = 1$, Δm is conveniently described as the oscillation frequency in units of inverse time, typically ps^{-1}.

Another important parameter of the B system is the lifetime difference defined as:

$$\Delta\Gamma = \Gamma_L - \Gamma_H \tag{1.2.12}$$

Note that both Δm and $\Delta\Gamma$ are defined to be positive quantitities. Thus, in analogy with the kaon system, the heavy state is the long lived while the light one is the short lived. The lifetime difference is expected to be larger in the B_s^0 system than the B^0 one. The experimental information $\Delta m_s \gg \Gamma_s$ model independently implies $|\Gamma_{12}| \ll |M_{12}|$. By expanding in terms of $\Gamma_{12}/|M_{12}$, we obtain:

$$\frac{q}{p} = -e^{-i\phi_M}\left[1 - \frac{1}{2}\left|\frac{\Gamma_{12}}{M_{12}}\right| sin\phi_s\right] + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right) \tag{1.2.13}$$

$$\Delta m = 2|M_{12}|\left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \simeq 2|M_{12}| \tag{1.2.14}$$

$$\Delta\Gamma = 2|\Gamma_{12}|cos\phi_s\left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \simeq 2|\Gamma_{12}|cos\phi_s \tag{1.2.15}$$

$$\tag{1.2.16}$$

where

$$\phi_M \equiv arg M_{12}, \qquad \phi_s \equiv arg(-\frac{\Gamma_{12}}{M_{12}}) \tag{1.2.17}$$

Introducing B_{even} and B_{odd}, eigenstates of the CP operator, the following relation is obtained:

$$|B_L\rangle = \frac{1 + e^{i\phi_s}}{2}|B_{even}\rangle - \frac{1 - e^{i\phi_s}}{2}|B_{odd}\rangle\,, \tag{1.2.18}$$

$$|B_H\rangle = -\frac{1 - e^{i\phi_s}}{2}|B_{even}\rangle + \frac{1 + e^{i\phi_s}}{2}|B_{odd}\rangle\,. \tag{1.2.19}$$

The SM thus predicts that B_L is almost completely CP-even and B_H CP-odd.

1.3 B_s^0 Mixing in the Standard Model

In the SM the leading-order diagrams to describe the neutral B meson mixing are represented in the two diagrams in Figure 1.2.

The contribution to the loops in Figure 1.2 is calculated to be proportional to the mass of the quark which appear in the loop [10]. The mass of the top quark is $\mathcal{O}(10^2)$ times greater than the mass of the charm and up quarks, as seen in Table 1.2, and thus the top quark contribution to the loop dominates. With this assumption, the oscillation frequency is proportional to elements of the quark mixing matrix V:

$$\Delta m_q \propto f_B^2 \hat{B} m_B |V_{tq} V_{tb}^*|^2\,, \tag{1.3.1}$$

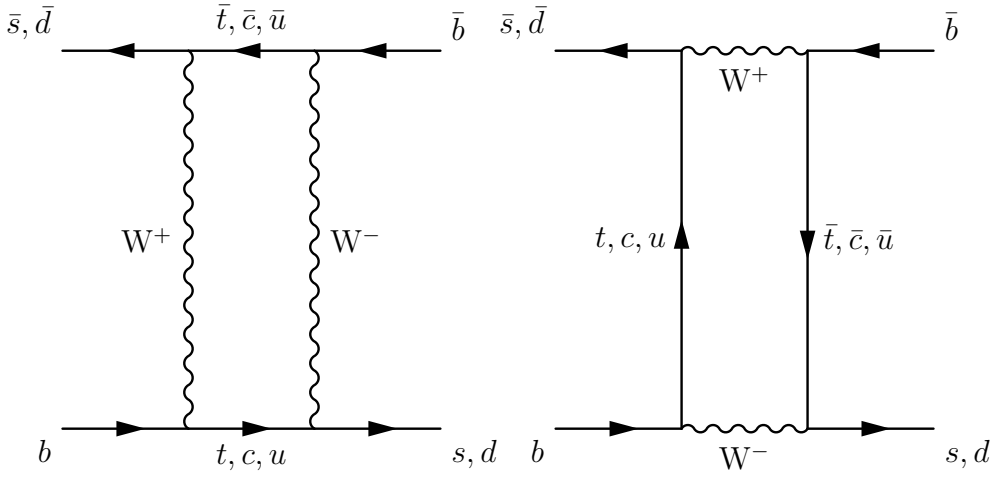

FIGURE 1.2: Lowest order diagrams for B mixing.

Parameter	Value [11]
f_{B^0}	$216 \pm 9 \pm 19 \pm 7$ MeV
$f_{\mathrm{B}^0_s}$	$260 \pm 7 \pm 26 \pm 9$ MeV
$\hat{B}_{\mathrm{B}^0}$	$0.836 \pm 0.027\ ^{+0.056}_{-0.062}$
$\hat{B}_{\mathrm{B}^0_s}/\hat{B}_{\mathrm{B}^0}$	$1.017 \pm 0.016\ ^{+0.056}_{-0.017}$

TABLE 1.5: Latest Lattice QCD estimates of form factors and bag factors of B mesons.

where $q = d, s$. Lattice QCD provides estimates of the form factor f_{B} and the bag factor $\hat{B}$ for B^0 and B^0_s mesons. The current best estimates for these parameters are reported in Table 1.5. The parameters are known with a precision of about 10%, which is thus the best level at which V_{tq} can be measured using Equation 1.3.1.

Nevertheless, most of the hadronic uncertainties that separately affect f_{B} and $\hat{B}$ cancel out, when considering the ratio between Δm_d and Δm_s, and yield to a direct relation with elements of the CKM matrix:

$$\frac{\Delta m_s}{\Delta m_d} = \xi^2 \frac{m_{\mathrm{B}^0_s}}{m_{\mathrm{B}^0}} \frac{|V_{ts}|^2}{|V_{td}|^2}\,, \tag{1.3.2}$$

where:

$$\xi = \frac{f_{\mathrm{B}^0_s}}{f_{\mathrm{B}^0}} \sqrt{\frac{\hat{B}_{\mathrm{B}^0_s}}{\hat{B}_{\mathrm{B}^0}}} = 1.210\ ^{+0.047}_{-0.039}\ [11]\,. \tag{1.3.3}$$

Thus, the measurement of the ratio $\Delta m_d/\Delta m_s$ allows a precise estimation of $|V_{td}|/|V_{ts}|$. The CKM matrix elements are not predicted by the SM and must be extracted from experimental data, therefore the measurement of B^0_s mixing, combined with the well-measured B^0 frequency ($0.507 \pm 0.005\ ps^{-1}$), provide a constrain of the

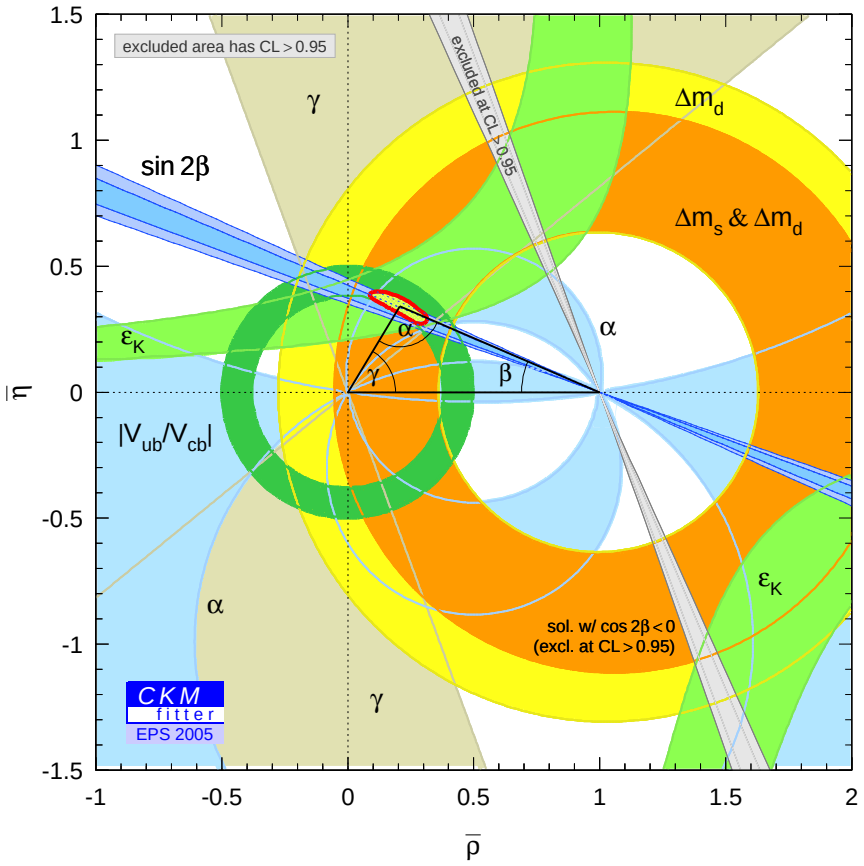

FIGURE 1.3: CKM fit [13] of the unitarity triangle presented at EPS2005, before the first measurement of Δm_s [14].

lenght of upper-right side of the Unitarity Triangle shown in Figure 1.1:

$$R_t \equiv \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\bar{\rho})^2+\bar{\eta}^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{ts}}\right| \tag{1.3.4}$$

In Figure 1.3 the status of the constraints in the $\bar{\rho}$–$\bar{\eta}$ as of the EPS 2005 [12] conference is depicted. At that time, the combination used, as experimental input for Δm_s, the 95% CL limit $\Delta m_s > 14.4$ ps^{-1}. The UT apex position is poorly constrained along the side R_t, thus a definitive measurement of Δm_s will then result in a stringent test of the unitarity of the CKM matrix. Assuming the SM to be fundamental theory and utilizing the currently available experimental results on CKM parameters, the CKM Fitter group [13] fit predicts a value of $\Delta m_s = 18.3\ ^{+6.5}_{-1.5}$ ps^{-1}.

1.4 Lifetime Difference and CP-Violating Phase in the B_s^0 System

In this section we describe the angular distributions of the decay $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^{*0}$, their time dependence and appropriate weighting functions. Both the decays considered here are decays of a pseudo-scalar to a vector-vector intermediate state ($P \to VV$). In such decays one can statistically distinguish their parity P by looking at the angular correlation among the final state particles. For the $B_s^0 \to J/\psi\phi$ both J/ψ and ϕ are C-odd eigeinstates, so the properties of $J/\psi\phi$ state under parity P are the same as those under CP. In principle it gives an handle to separate the light and heavy mass eigeinstates (equations 1.2.18 and 1.2.19). The $B_s^0 \to J/\psi\phi$ decays to an admixture of CP eigeinstates therefore these distributions will be used in an angular analysis to disentagle the light and heavy mass eigeinstates contribution to extract information about the lifetime difference $\Delta\Gamma_s$ and the CP-violating phase β_s. The study of the kinematically similar decay $B^0 \to J/\psi K^{*0}$ is used as calibration sample for the main B_s^0 analysis. The obtained results will be cross-checked with the most precise B factory ones [references], to test the robustness of the entire framework.

1.4.1 Time Dependent Angular Distributions in the $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^{*0}$

One of the most suitable coordinate bases, from a theoretical and experimental point of view, to investigate the angular correlations among the final state particles in the $P \to VV$ decays is the transversity basis, expressed in terms of three angles θ_T, ϕ_T, ψ_T. They are defined as follows. Consider the decay $B_s^0 \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$. In the J/ψ rest of frame, the ϕ flight direction defines the x axis, while the plane of K^+K^- system defines the $x-y$ plane with the y axis oriented such that $p_y(K^+) > 0$. By adopting a right-handed coordinate system, the ambiguity in the choice of the z axis is solved. The angle θ_T is defined as the angle between μ^+ flight direction and the positive direction of the z axis. The angle ϕ_T is the angle between the x axis and the projection of the μ^+ onto the $x-y$ plane. Finally, ψ_T is the angle of the K^+ in the ϕ rest frame relative to the negative direction of J/ψ in that frame. Throughout this dissertation we will denote as $\vec{\omega} = \{\cos\theta_T, \phi_T, \cos\psi_T\}$ this set of three angular variables.

The genereric expression of the differential rate at time t for the $P \to VV$ decays is []:

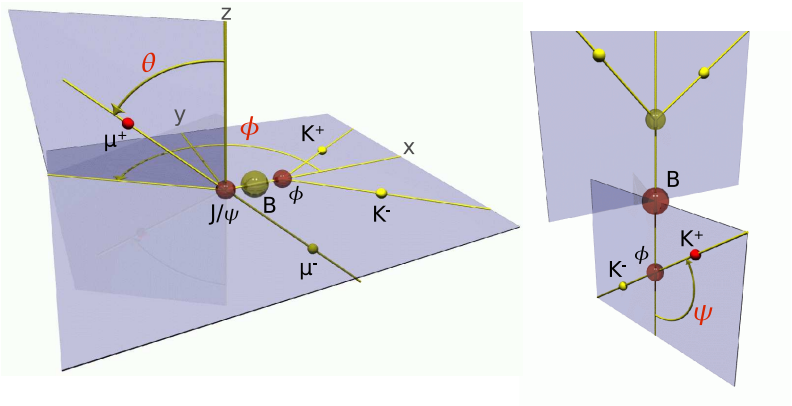

FIGURE 1.4: Transversity basis and angle definition in the case of $B_s^0 \to J/\psi\phi$ decay.

$$\begin{aligned}\frac{d^4P(\vec{\omega},t)}{d\vec{\omega}dt} &\propto \left|A_0(t)\right|^2 f_1(\vec{\omega}) + \left|A_\parallel(t)\right|^2 f_2(\vec{\omega}) \\ &+ \left|A_\perp(t)\right|^2 f_3(\vec{\omega}) + Im\{A_\parallel^*(t)A_\perp(t)\} f_4(\vec{\omega}) \\ &+ Re\{A_0^*(t)A_\parallel(t)\} f_5(\vec{\omega}) + Im\{A_0^*(t)A_\perp(t)\} f_6(\vec{\omega}) \\ &\equiv \sum_k \mathcal{O}^{(k)}(t) f^{(k)}(\theta_T, \phi_T, \psi_T)\end{aligned} \tag{1.4.1}$$

where the appropriate weighting function $f_i(\vec{\omega})$ (index i runs from 1 to 6) are:

$$\begin{aligned}f_1(\vec{\omega}) &\equiv \frac{9}{32\pi} 2\cos^2(\psi_T)[1-\sin^2(\theta_T)\cos^2(\phi_T)] \\ f_2(\vec{\omega}) &\equiv \frac{9}{32\pi} \sin^2(\psi_T)[1-\sin^2(\theta_T)\sin^2(\phi_T)] \\ f_3(\vec{\omega}) &\equiv \frac{9}{32\pi} \sin^2(\psi_T)\sin^2(\theta_T) \\ f_4(\vec{\omega}) &\equiv \frac{9}{32\pi} \sin^2(\psi_T)\sin(2\theta_T)\sin(\phi_T) \\ f_5(\vec{\omega}) &\equiv -\frac{9}{32\pi}\frac{1}{\sqrt{2}} \sin(2\psi_T)\sin^2(\theta_T)\sin(2\phi_T) \\ f_6(\vec{\omega}) &\equiv \frac{9}{32\pi}\frac{1}{\sqrt{2}} \sin(2\psi_T)\sin(2\theta_T)\cos(\phi_T)\end{aligned} \tag{1.4.2}$$

The differential rate is expressed as a sum over the time evolution of possible bilinear combinations of decay amplitudes $A_\alpha(t)$, with $\alpha = \{0, \parallel, \perp\}$. The three decay amplitudes corresponds to linear polarization of the vector mesons (J/ψ, K^{*0}/ϕ), which are either longitudinal (A_0), or transverse to their flight direction and parallel ($A_\parallel$) or perpendicular ($A_\perp$) to one another. The amplitude $A_\parallel$ results in decays where the two vector mesons are emitted with relative orbital angular momentum $L = 1$,

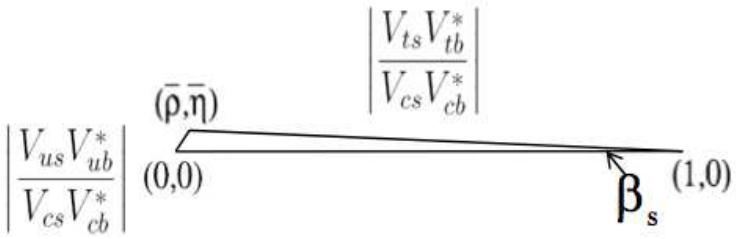

FIGURE 1.5: A sketch of the B_s^0 Unitarity Triangle referred to equation 1.4.5.

thus it is associated with the P-odd decays. The amplitudes A_0 and $A_\parallel$ are associated with (mixtures of) the $L = 0$ and $L = 2$ decays, therefore P-even[reference]. Only relative phases of the amplitudes can enter physics observables, so we are free to fix the phase of one of them. Our convention is to chose $arg(A_0) = 0$ and define $\delta_\parallel \equiv arg(A_\parallel)$ and $\delta_\perp \equiv arg(A_\perp)$. The amplitudes satistify the normalization condition $|A_0(t)|^2 + |A_\parallel(t)|^2 + |A_\perp(t)|^2 = 1$. When not differently specified, we define $|A_\alpha(t = 0)| \equiv |A_\alpha|$ with $\alpha = \{0, \parallel, \perp\}$. The equation 5.7.2 is valid for initially produced B ($\overline{\mathrm{B}}$) mesons when the upper (lower) sign is taken for the intereference between the (C)Peven and odd amplitudes. The time evolution of the amplitudes depends on whetever we have a B or $\overline{\mathrm{B}}$ at the production time and on the decay analyzed. For the case of $B_s^0 \to J/\psi\phi$, we get the following formulas [15].

$$
\begin{aligned}
&\sum_k \mathcal{O}^{(k)}(t) f^{(k)}(\theta_T, \phi_T, \psi_T) = \\
&\quad \left|A_0\right|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma|t}{2} - \sin(2\beta_s)\sin(\Delta m t)\right] f_1(\vec{\omega}) + \\
&\quad \left|A_\parallel\right|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma|t}{2} - \sin(2\beta_s)\sin(\Delta m t)\right] f_2(\vec{\omega}) + \\
&\quad \left|A_\perp\right|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} + |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma|t}{2} + \sin(2\beta_s)\sin(\Delta m t)\right] f_3(\vec{\omega}) + \\
&\quad \left|A_0\right|\left|A_\parallel\right| e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma|t}{2} - \sin(2\beta_s)\sin(\Delta m t)\right] \cos(\delta_2 - \delta_1) f_5(\vec{\omega}) + \\
&\quad \left|A_\parallel\right|\left|A_\perp\right| e^{-\Gamma t}\left[\sin\delta_1\cos(\Delta m t) - \cos\delta_1\cos(2\beta_s)\sin(\Delta m t) + \cos\delta_1\sin(2\beta_s)\sinh\frac{\Delta\Gamma t}{2}\right] f_4(\vec{\omega}) + \\
&\quad \left|A_0\right|\left|A_\perp\right| e^{-\Gamma t}\left[\sin\delta_2\cos(\Delta m t) - \cos\delta_2\cos(2\beta_s)\sin(\Delta m t) + \cos\delta_2\sin(2\beta_s)\sinh\frac{\Delta\Gamma t}{2}\right] f_6(\vec{\omega}).
\end{aligned}
\tag{1.4.3}
$$

while for the P-conjougate decays $\overline{\mathrm{B}}_s \to J/\psi\phi$ takes the form:

$$\begin{aligned}
&\sum_k \bar{\mathcal{O}}^{(k)}(t) f^{(k)}(\theta_T, \phi_T, \psi_T) = \\
&\quad \big|A_0\big|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma| t}{2} + \sin(2\beta_s)\sin(\Delta m t)\right] f_1(\vec{\omega}) + \\
&\quad \big|A_\parallel\big|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma| t}{2} + \sin(2\beta_s)\sin(\Delta m t)\right] f_2(\vec{\omega}) + \\
&\quad \big|A_\perp\big|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} + |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma| t}{2} - \sin(2\beta_s)\sin(\Delta m t)\right] f_3(\vec{\omega}) + \\
&\quad \big|A_0\big|\big|A_\parallel\big| e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos(2\beta_s)|\sinh\frac{|\Delta\Gamma| t}{2} + \sin(2\beta_s)\sin(\Delta m t)\right] \cos(\delta_2 - \delta_1) f_5(\vec{\omega}) + \\
&\quad \big|A_\parallel\big|\big|A_\perp\big| e^{-\Gamma t}\left[-\sin\delta_1\cos(\Delta m t) + \cos\delta_1\cos(2\beta_s)\sin(\Delta m t) + \cos\delta_1\sin(2\beta_s)\sinh\frac{\Delta\Gamma t}{2}\right] f_4(\vec{\omega}) + \\
&\quad \big|A_0\big|\big|A_\perp\big| e^{-\Gamma t}\left[-\sin\delta_2\cos(\Delta m t) + \cos\delta_2\cos(2\beta_s)\sin(\Delta m t) + \cos\delta_2\sin(2\beta_s)\sinh\frac{\Delta\Gamma t}{2}\right] f_6(\vec{\omega}).
\end{aligned} \tag{1.4.4}$$

The magnitudes of the corresponding decay amplitudes are assumed to be equal for particle or anti-particle decay. Thus we have to deal only with mixing-induced CP violation and there is no direct CP violation, i.e. $|A_0(0)| = |\bar{A}_0(0)|$, $|A_\parallel(0)| = |\bar{A}_\parallel(0)|$, $|A_\perp(0)| = |\bar{A}_\perp(0)|$. It is important to notice that the information on the Δm could in principle be extracted from a time dependent analyses of tagged $\mathrm{B}_s^0 \to J/\psi\phi$ samples, but for the analysis reported in this dissertation we will use as input for Δm_s the value obtained from a dedicated measurement of the B_s^0 mixing oscillation frequency described in chapter [put reference to chapter I will write and to PRL].
The angle β_s, shown in Figure 1.5, is defined as $\beta_s = arg[-(V_{ts}V_{tb}^*)/(V_{cs}V_{cb}^*)]$. In the B_s^0 system it is the equivalent to the angle β for the B^0 system reported in Figure 1.1. It belongs to a different Unitarity Triangle coming from the following equation:

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0\,. \tag{1.4.5}$$

Re-writing explicitily the Equations 1.4.3 and 1.4.4, we get the following expression for the $\mathrm{B}_s^0 \to J/\psi\phi$ decay amplitudes:

$$
\begin{aligned}
&\frac{d^4P}{dtd\vec{\omega}}(t,\vec{\omega}) \propto \\
&\quad \big|A_0\big|^2\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}-\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_1+ \\
&\quad \big|A_\parallel\big|^2\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}-\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_2+ \\
&\quad \big|A_\perp\big|^2\left[\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_H t}+\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_3+ \\
&\quad \big|A_0\big|\big|A_\parallel\big|\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}-\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]\cos(\delta_2-\delta_1)f_5+ \\
&\quad \big|A_\parallel\big|\big|A_\perp\big|\left[-\cos\delta_1\sin(2\beta_s)\frac{e^{-\Gamma_L t}-e^{-\Gamma_H t}}{2}+\sin\delta_1\cos(\Delta m t)e^{-\Gamma t}-\cos\delta_1\cos(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_4+ \\
&\quad \big|A_0\big|\big|A_\perp\big|\left[-\cos\delta_2\sin(2\beta_s)\frac{e^{-\Gamma_L t}-e^{-\Gamma_H t}}{2}+\sin\delta_2\cos(\Delta m t)e^{-\Gamma t}-\cos\delta_2\cos(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_6.
\end{aligned}
\tag{1.4.6}
$$

$$
\begin{aligned}
&\frac{d^4\bar{P}}{dtd\vec{\omega}}(t,\vec{\omega}) \propto \\
&\quad \big|A_0\big|^2\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}+\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_1+ \\
&\quad \big|A_\parallel\big|^2\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}+\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_2+ \\
&\quad \big|A_\perp\big|^2\left[\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_H t}-\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_3+ \\
&\quad \big|A_0\big|\big|A_\parallel\big|\left[\frac{1+\cos(2\beta_s)}{2}e^{-\Gamma_L t}+\frac{1-\cos(2\beta_s)}{2}e^{-\Gamma_H t}+\sin(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]\cos(\delta_2-\delta_1)f_5+ \\
&\quad \big|A_\parallel\big|\big|A_\perp\big|\left[-\cos\delta_1\sin(2\beta_s)\frac{e^{-\Gamma_L t}-e^{-\Gamma_H t}}{2}-\sin\delta_1\cos(\Delta m t)e^{-\Gamma t}+\cos\delta_1\cos(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_4+ \\
&\quad \big|A_0\big|\big|A_\perp\big|\left[-\cos\delta_2\sin(2\beta_s)\frac{e^{-\Gamma_L t}-e^{-\Gamma_H t}}{2}-\sin\delta_2\cos(\Delta m t)e^{-\Gamma t}+\cos\delta_2\cos(2\beta_s)\sin(\Delta m t)e^{-\Gamma t}\right]f_6.
\end{aligned}
\tag{1.4.7}
$$

In the B^0-$\overline{B}^0$ system the situation is different: given $|\beta| \gg 0$, impling $\frac{q}{p} \neq 1$, the mass eigeinstates are no CP eigeinstates. Taking into account $|A_\alpha| = |\bar{A}_\alpha|$ and using the transversity base, where the particle ϕ is replaced by the K^{*0} the angular distributions and time dependences for the *flavor-specific* $B^0 \to J/\psi(\to l^+l^-)K^{*0}(\to K^\pm\pi^\mp)$ modes are[hep-ph/9804253]:

$$\frac{d^3\Gamma[B^0(t) \to J/\psi(\to l^+l^-)K^{*0}(\to K^+\pi^-)]}{d^3\vec{\omega}} = \frac{9}{32\pi}\cos^2(\frac{\Delta mt}{2})e^{-\Gamma_d t}$$
$$\times \left[\left|A_0\right|^2 f_1 + \left|A_\parallel\right|^2 f_2 + \left|A_\perp\right|^2 f_3 + Im(A_\parallel^* A_\perp) f_4 + Re(A_0^* A_\parallel) f_5 + Im(A_0^* A_\perp) f_6 \right]$$

$$\frac{d^3\Gamma[\overline{B}^0(t) \to J/\psi(\to l^+l^-)\overline{K}^{*0}(\to K^-\pi^+)]}{d^3\vec{\omega}} = \frac{9}{32\pi}\cos^2(\frac{\Delta mt}{2})e^{-\Gamma_d t}$$
$$\times \left[\left|A_0\right|^2 f_1 + \left|A_\parallel\right|^2 f_2 + \left|A_\perp\right|^2 f_3 - Im(A_\parallel^* A_\perp) f_4 + Re(A_0^* A_\parallel) f_5 - Im(A_0^* A_\perp) f_6 \right]$$

$$\frac{d^3\Gamma[B^0(t) \to J/\psi(\to l^+l^-)\overline{K}^{*0}(\to K^-\pi^+)]}{d^3\vec{\omega}} = \frac{9}{32\pi}\sin^2(\frac{\Delta mt}{2})e^{-\Gamma_d t}$$
$$\times \left[\left|A_0\right|^2 f_1 + \left|A_\parallel\right|^2 f_2 + \left|A_\perp\right|^2 f_3 - Im(A_\parallel^* A_\perp) f_4 + Re(A_0^* A_\parallel) f_5 - Im(A_0^* A_\perp) f_6 \right]$$

$$\frac{d^3\Gamma[\overline{B}^0(t) \to J/\psi(\to l^+l^-)K^{*0}(\to K^+\pi^-)]}{d^3\vec{\omega}} = \frac{9}{32\pi}\cos^2(\frac{\Delta mt}{2})e^{-\Gamma_d t}$$
$$\times \left[\left|A_0\right|^2 f_1 + \left|A_\parallel\right|^2 f_2 + \left|A_\perp\right|^2 f_3 + Im(A_\parallel^* A_\perp) f_4 + Re(A_0^* A_\parallel) f_5 + Im(A_0^* A_\perp) f_6 \right] \tag{1.4.8}$$

Summing over the initially produced B^0 and $\overline{B}^0$, without taking into account the flavor identication at the production time, the following decay distribution is obtained:

$$\frac{d^4P(\vec{\omega},t)}{d\vec{\omega}dt} \propto \Bigg[\left|A_0\right|^2 f_1(\vec{\omega}) + \left|A_\parallel\right|^2 f_2(\vec{\omega})$$
$$+ \left|A_\perp\right|^2 f_3(\vec{\omega}) \pm Im\{A_\parallel^* A_\perp\} f_4(\vec{\omega}) \tag{1.4.9}$$
$$+ Re\{A_0^* A_\parallel\} f_5(\vec{\omega}) \pm Im\{A_0^* A_\perp\} f_6(\vec{\omega}) \Bigg] e^{-\Gamma_d t}$$

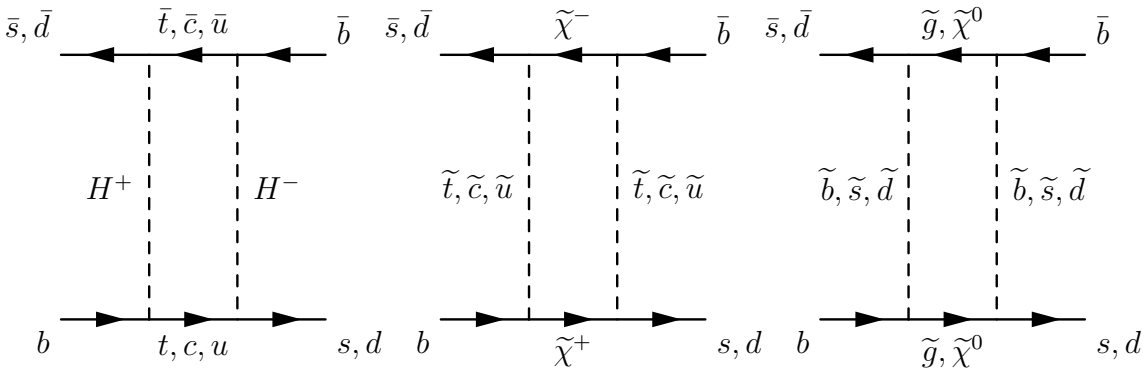

FIGURE 1.6: Contributions to B_s^0 mixing in MSSM with MFV.

where the upper(lower) sign is used for the flavor specific final state of $B^0(\overline{B}^0)$ with $K^+\pi^-$ ($K^-\pi^+$). In this decays mode all the linear term in $\Delta\Gamma_d$ disappear, so sensitity to the lifetime difference is lost. Nevertheless, the equation 1.4.9 can be used to extract the $|A_\alpha|$ and the phases $\delta_\parallel$ and $\delta_\perp$. These are valuable information because they make possible to test the SU(3) flavor simmetry in which the decay amplitudes and phases for the B^0 and the B_s^0 are expected to agree.

1.5 Beyond the Standard Model

The measurement of the B_s^0 system parameters such as the oscillation frequency, the lifetime difference or the CP-violation phase allows to test stringently the SM expectations, providing probes for New Physics (NP).
As shown in Figure 1.2 the $B_s^0 - \overline{B}_s^0$ mixing amplitude is a CKM-suppressed loop-induced fourth order weak interaction process, therefore very sensitive to NP. In fact, simple extensions of the SM could introduce additional fields which, mediating the $B_s^0 - \overline{B}_s^0$ transitions via box-diagrams, may provide large modifications to Δm_s expectation. An example of possible additional contributions to the B_s^0 mixing are shown in Figure 1.6. They refer to the case of Minimal Supersymmetric extension with flavor conservation. In this model, for instance, the contributions come from the masses of the charginos $\widetilde{\chi}_2^\pm$, stop $\widetilde{t}$ and charged Higgs bosons $H^\pm$.

Concerning the CP-violation aspect, the mixing-induced CP asymmetries in the B_s^0 decays practically vanish, because governed by the tiny phase β_s. It is clear that NP does not take much to modify the SM predictions. For the description of the CP violation in the B_s^0 system two are the phases generally utilized: $2\beta_s$ and ϕ_s. They are correlated to each other and sometimes there is confusion in distinguishing them, thus some clarification is needed. The phase $2\beta_s$ appears in $b \to c\bar{c}s$ decays of neutral B-mesons when mixing is taken into account and is predicted to be $2\beta_s^{SM} = (2.2 \pm 0.6)^o = 0.04 \pm 0.01\ rad.$ Instead the phase ϕ_s appears in the expression of the decay-width difference $\Delta\Gamma = 2|\Gamma_{12}|cos\phi_s$ which expected value is $\Delta\Gamma_s^{SM} = 0.096 \pm 0.039\ ps^{-1}$. In the SM, the phase ϕ_s is predicted to be small, $\phi_s^{SM} = (0.24 \pm 0.04)^o = 0.0041 \pm 0.0008\ rad.$ A clear indication of NP could therefore come from the measurement of ample values for $2\beta_s$ or ϕ_s inconsistent with zero.

NP can similarly alter the phases $2\beta_s = 2\beta_s^{SM} - \phi_s^{NP}$ and $\phi_s = \phi_s^{SM} + \phi_s^{NP}$. If the NP contribution is sizeable, then in both cases only ϕ_s^{NP} survives and, since $2\beta_s^{SM}$ and ϕ_s^{SM} are significantly smaller than the current experimental resolution, the common used approximation is $2\beta_s = -\phi_s^{NP} = \phi_s$.

1.6 Analysis Strategies

In this section we present an overview of the measurements of the $B_s^0 - \overline{B}_s^0$ oscillation frequency (Δm_s), lifetime difference ($\Delta\Gamma_s$) and mixing induced CP violation phase (β_s), outlining some of the experimental issues.
The production of B hadrons at the Tevatron is dominated by processes that produce $b\bar{b}$ pairs. The b quark and $\bar{b}$ anti-quark are energetic enough that they are expected to fragment into B hadrons independently of one another. All B species are produced, with ∼10% of b quarks fragmenting into B_s^0 [6].

Measurement of $B_s^0 - \overline{B}_s^0$ Oscillations

A picture of a $p\bar{p}$ interaction where a B is produced, and the subsequent decay of the B meson is depicted in Figure 1.7. The cartoon presents the steps of the analysis:

1. **Flavor at the time of production:** knowledge of whether the meson was produced as a B_s^0 or a $\overline{B}_s^0$. This is referred to as "initial-state flavor tagging" or simply "flavor tagging."

2. **Flavor at the time of decay:** knowledge of whether the meson was a B_s^0 or $\overline{B}_s^0$ when it decayed. The meson is classified as "mixed" ("unmixed") if the flavor of decay is different than (the same as) the flavor at production.

3. **Proper decay-time:** the proper decay-time is the decay-time of the hadron in its rest frame. Since a B_s oscillates several times during its average lifetime the time dependent observation of $B_s^0 - \overline{B}_s^0$ oscillations requires excellent proper-time resolution.

The reconstructed final states which enter this analysis are listed below (the charge-conjugated modes are implied):

- $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \pi^+ \pi^- \pi^+$, with:
  - $D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$,
  - $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$,
  - $D_s^- \to \pi^-\pi^+\pi^-$.

  These modes are fully reconstructed, all tracks in the final state are included in the fit of the B_s^0 candidate.

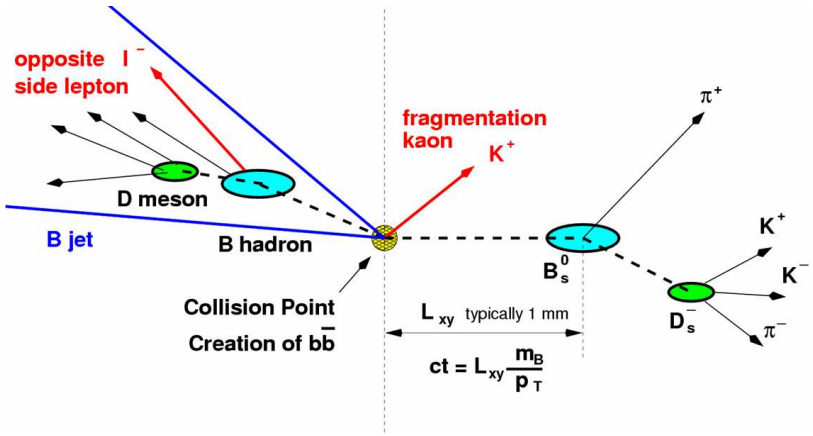

FIGURE 1.7: A schematic representation of a B_s^0 hadronic decay.

- $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$, with $D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$. The B_s^0 candidates in these samples are only partially reconstructed, because the γ, or π^0, which takes part to the $D_s^{*-} \to D_s^-\gamma$, or $D_s^-\pi^0$, decay and the π^0 of the $\rho^+ \to \pi^0\pi^+$ decay are not included in their fits.

- $B_s^0 \to D_s^-\ell^+X$, with:

  - $D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$,
  - $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$,
  - $D_s^- \to \pi^-\pi^+\pi^-$.

  These modes constitute the "semileptonic" samples. Only the lepton and the D_s^- candidate of the B_s^0 final state are utilized in the reconstruction of the B_s^0 candidate.

The B_s^0 candidates are reconstructed in self-tagging final states: the flavor is indicated by the charge of the D_s^- candidate which takes part in the reconstructed final state (D_s^- tags a B_s^0, whereas D_s^+ indicates a $\overline{B}_s^0$). The flavor at the time of production is more difficult to indentify, and several techniques have been developed to perform this function. As shown in Figure 1.7, flavor taggers are distinguished as being on the same-side or the opposite-side relative to the reconstructed B_s^0 candidate. Same-side flavor tagging algorithms explore flavor-charge correlations between the reconstructed B_s^0 and tracks nearby produced in the fragmentation process. Opposite-side flavor tagging algorithms are based on the identification of some property of the opposite-side B to determine its b quantum number, from which the production flavor of the trigger B_s^0 can be inferred. The proper decay-time is determined from the measurement of the momentum and the decay length of the B_s^0 candidate.

To finally perform the measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency, the oscillation probabilities in Equations 1.2.9 and 1.2.10 are exploited. Each of the three items listed above has experimental limitations, thus such analysis requires large samples of B_s decays with a good signal-to-noise ratio.

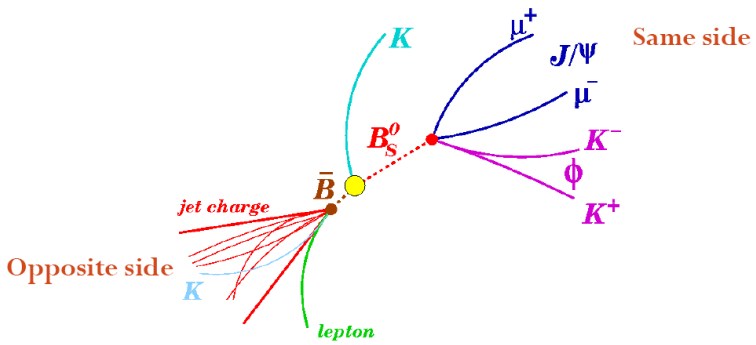

FIGURE 1.8: A schematic representation of a $B_s^0 \to J/\psi\phi$ decay.

Measurement of Lifetime Difference and CP Violation Phase in the B_s^0 System

For the measurement of lifetime difference and CP violation phase the Bs meson decay to $J/\psi\phi$ are analyzed. The relative stilized picture of the $p\bar{p}$ interaction is shown in Figure 1.8. This decay mode is particularly interesting for B-physics experiments at hadron machines because of its clear experimental signature, $J/\psi \to \mu^+\mu^-$ $\phi \to K^+K^-$. Several of the requirement described for the oscillation frequency measurement are still needed for this analysis and the same techniques are used:

1. **Flavor at the time of production**

2. **Proper decay-time**

This particular decay mode in not self-tagging, therefore we cannot infer the flavor of the B_s^0 meson at the decay time. The flavor tagging information at production time is necessary to separate the time evolution of the mesons produced as B_s^0 or $\overline{B}_s^0$. The sensitivity to the CP-violating phase is obtained relating this time development with the CP eigeinvalues of the final state. The decay mode $B_s^0 \to J/\psi\phi$ results to be an admixture of CP eigeinstates which are accessible through the angular distributions of the J/ψ and ϕ mesons reported in Equations 1.4.3 and 1.4.4.
Analyses of $B_s^0 \to J/\psi\phi$ decays without initial state tagging information can be performed [reference to PRL and Milnik thesis]. Such analyses are superior to the one using flavor tagging information in terms of efficiency, acceptance and purity but, being sensitive to $|cos(2\beta_s)|$ and $|sin(2\beta_s)|$, they lead to a four-fold ambiguity in the determination of $2\beta_s$. In this disseration we report the first flavor tagged determination of bounds on the CP violating phase $2\beta_s$.

The theoretical aspects and the phenomenology of $B_s^0 - \overline{B}_s^0$ system have been presented. The next chapter describes the Fermilab accelerator complex, the CDF detector and its trigger system.

Chapter 2

Experimental apparatus

This chapter is focused on the accelerator complex at Fermilab and the CDF detector descriptions.

2.1 Accelerators at Fermilab

The Fermi National Accelerator Laboratory (FNAL) is located 35 miles west of Chicago, IL. The set of accelerators hosted at FNAL allows for the production of the most powerful beams of particles currently available to experimentalists. Protons and anti-protons are produced and collide with center-of-mass energy equal to 1.96 TeV in the Tevatron, the main accelerator at Fermilab.

Apart from the collision energy, the instantaneous luminosity $\mathcal{L}$ is a key parameter in defining the quality of a collider, because it determines the production rate of physics processes. For the Tevatron, it is defined as follows [6]:

$$\mathcal{L} = f \cdot B \cdot \frac{N_p \cdot N_{\overline{p}}}{2\pi \left(\sigma_p^2 + \sigma_{\overline{p}}^2\right)} F\left(\frac{\sigma_l}{\beta^*}\right) , \tag{2.1.1}$$

where N_p are $N_{\overline{p}}$ are the number of protons and anti-protons, respectively, in each bunch, B the number of bunches circulating in the ring, f the rotation frequency, σ_p and $\sigma_{\overline{p}}$ the transverse size of the proton and anti-proton beams in the interaction point, F a form factor which corrects for the bunch shape and depends on the ratio of the bunch length σ_l to the value of the amplitude function β at the interaction point, β^*. The amplitude function β depends on the beam optics and represents a measure of the beam width. Thirty-six bunches of protons and an equal number of bunches of anti-protons are equidistantly accelerated. The time between bunch crossings, the *inter-bunch-separation*, is 396 ns. The peak value of $\mathcal{L}$ has been steadily increasing since the beginning of data-taking, in March 2002, reaching $2.8 \cdot\ 10^{32}$ cm^{-2}s^{-1} in the first months of 2007. The parameters of the Tevatron collider are summarized in Table 2.1.

The integrated luminosity L, defined as $L = \int dt\mathcal{L}$, is more relevant to physics analyses. The probability for interactions to occur is directly proportional to the cross section of the process σ[cm^2] and to L[cm^{-2}]. The unit adopted to measure cross

Parameter	Value
energy at center-of-mass	1.96 TeV
number of bunches, B	36
bunch σ_l	37 cm
inter-bunch spacing	396 ns
protons/bunch, N_p	$3 \cdot 10^{11}$
anti-protons/bunch, $N_{\overline{p}}$	$3 \cdot 10^{10}$
β^*	35 cm
interactions/crossing†	2
peak luminosity	$2.8 \cdot 10^{32}$ cm^{-2}s^{-1}

TABLE 2.1: Characteristic parameters of the Tevatron in early 2007.
† At a luminosity of $\mathcal{L} = 10^{32}$ cm^{-2}s^{-1}.

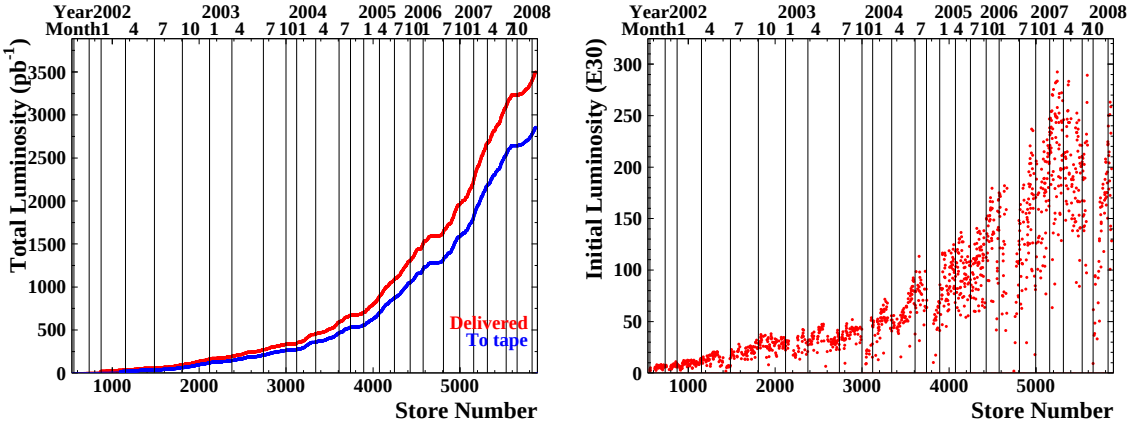

FIGURE 2.1: Integrated (left) and peak luminosity (right) delivered by the Tevatron. The plot covers the period between the beginning of 2002 and of 2008. The luminosity is shown as a function of store number.

sections observed in high energy collisions is the barn b, equivalent to 10^{-24} cm^2. Typical values in High Energy Physics are fractions of a barn. Figure 2.1 shows the total integrated luminosity up to January 2008 and the peak instantaneous luminosity in the same period.

The time period of stable circulation that the colliding $p\bar{p}$ beams are retained in the Tevatron is called *store*. The word indicates that protons and anti-protons are stored to fill the machine. Stores typically last $\mathcal{O}(10)$ hours and present stable colliding beams suitable for data taking. In the control rooms of the detectors, which are installed along the Tevatron, operators supervise the correct functioning of the respective detector and the registration of data in *runs* of variable length.

The following sections describe in more detail the various parts of the accelerator setup at Fermilab. A global picture of the accelerator complex of Fermilab is presented in Figure 2.2.

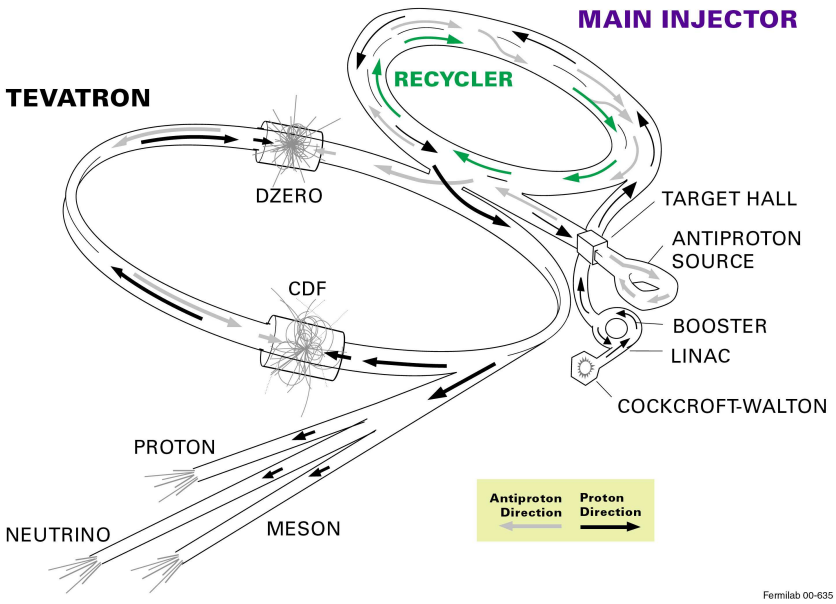

FIGURE 2.2: Accelerator complex at Fermilab.

2.1.1 LINear ACcelerator and Booster

In order to obtain beams of colliding protons and anti-protons, protons must first be obtained. Gaseous hydrogen is used, but rather than stripping off an electron to obtain protons, the H_2 is dissociated to obtain negatively charged H^- anions. They are subsequently accelerated in a Cockroft-Walton electrostatic machine up to a kinetic energy of 750 keV and then reach 400 MeV in the linear accelerator (LINAC [16]). The ions are finally directed on a carbon foil where their pairs of electrons are stripped off. The remaining protons are injected into the Booster [17], a circular synchrotron with a radius of 57 m. In the Booster, protons are grouped into 84 bunches, containing around $6 \cdot 10^6$ protons each, and are yet again accelerated, this time up to 8.9 GeV of total energy. Finally, the proton bunches are sent to the Main Injector.

2.1.2 Main Injector

The Fermilab Main Injector (FMI [18]) is a synchrotron with a circumference of 3319 m. It has the fundamental role of optimally connecting the Booster to the Tevatron. The Main Injector receives 8.9 GeV proton bunches from the Booster. Six injection cycles are necessary to fill it, with 498 proton bunches. The proton bunches, containing $2 \cdot 10^{11}$ protons each, reach 150 GeV, and three FMI cycles are necessary to transfer all the available protons to the Tevatron. In anti-proton–production mode, as opposed to the collider-injection mode described above, a single batch of protons, constituted by a set of 84 bunches (approximately $8 \cdot 10^{12}$ protons), is injected into the MI from the Booster. Protons are then accelerated up to 120 GeV and directed

to the Anti-proton Source.

2.1.3 Anti-Proton Source

The Anti-Proton Source [19] consists of three major components: the Target Station, the Debuncher, and the Accumulator. A proton pulse of 120 GeV is extracted from the Main Injector and focused on a nickel target. Anti-protons are thus produced, with a wide angular distribution, centered in the direction of the beam, and mean momentum of 8 GeV/c. On average, about 20 anti-protons are collected per one million incident protons. The anti-protons are collected and focused by a lithium lens and sent to the Debuncher, an 8 GeV triangularly shaped synchrotron, where the bunch structure is lost. The purpose of the Debuncher is to transform the anti-proton pulses in a continuous beam of monochromatic anti-protons, by applying the technique of bunch rotation, which transforms a beam with a large energy spread and a narrow time distribution (i.e., the beam is structured in bunches), into a beam with a large time spread and a narrow energy spread (i.e., a continuous, monochromatic beam), or vice-versa. De-bunching is necessary to reduce the large spread in energy of the produced anti-protons, which would make the transfer of anti-protons to subsequent accelerators difficult and inefficient. Stochastic cooling [20] is utilized to cool (in phase-space) the anti-proton beam before injecting it in the Accumulator Ring, another 8 GeV synchrotron. The anti-proton beam is then further cooled utilizing the same technique in the Accumulator Ring, where the division in bunches is also recovered. Finally, 8 GeV anti-proton bunches are injected in the Main Injector again, in the opposite direction than proton bunches, where they reach 150 GeV before extraction to the Tevatron.

2.1.4 Recycler Ring

The Recycler Ring (RR [21]) is a constant 8 GeV-energy storage ring, which shares the tunnel where the Main Injector is installed. A limiting factor of $p\bar{p}$ colliders is the availability of anti-protons. The RR has been conceived to exploit the anti-protons which are left in the Tevatron after the end of a cycle of collisions. Previously, left-over anti-protons, which amount to about 75% of the quantity originally injected, were discarded in lead beam-dumps. In the current phase of data-taking, their energy is reduced to 120 GeV in the Tevatron and they are then extracted and sent to the RR. Besides, the RR functions as a post-Accumulator ring. The content of the Accumulator Ring is periodically transferred in the RR, thus guaranteeing that the Accumulator Ring is always operating in its optimum anti-proton intensity regime. The RR can hold up to $5 \cdot 10^{12}$ anti-protons, which are efficiently cooled before being injected in the Main Injector for the preparation to a new cycle of collisions. The RR started operations in June 2004, resulting in one of the factors which contributed to the boost in integrated luminosity visible in Figure 2.1.

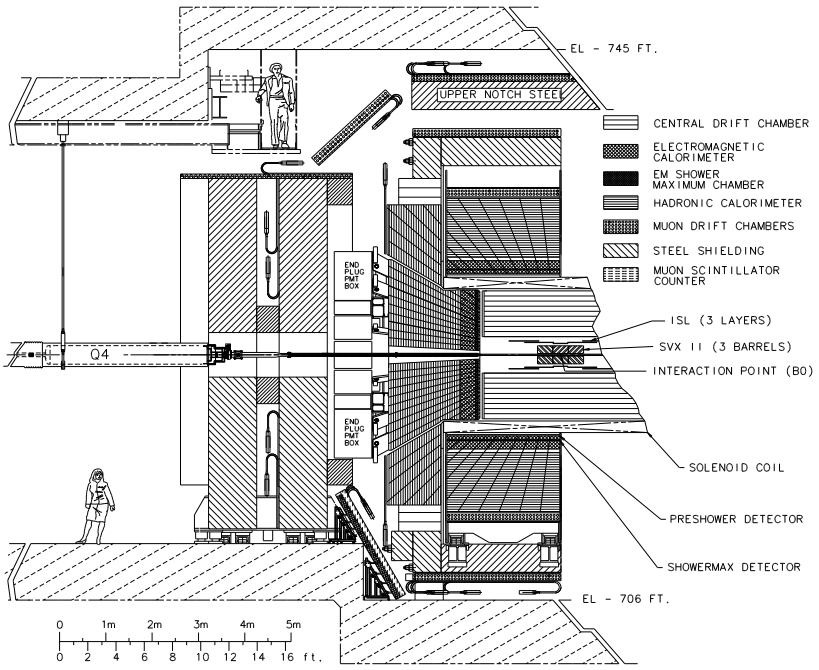

FIGURE 2.3: A cross-sectional view of one half of the CDF II detector.

2.1.5 Tevatron

The Tevatron collider [22] is the main accelerator in Fermilab. It contains 774 dipole (for steering) and 216 quadrupole (for focusing) superconducting magnets, distributed along a ring with a 1 km radius. Proton and anti-proton bunches are received from the Main Injector. The 4.5 T peak field in the Tevatron bending magnets allows the particles to be accelerated to an energy of 0.98 TeV. Protons are injected before anti-protons, and by means of electrostatic separators they are forced into a closed helicoidal orbit. The same prescription is applied to anti-protons, thus producing two strands with a transverse separation which prevents collisions outside the designed interaction points. The Tevatron has two interaction points, which are technically named BØ and DØ. The locations are currently utilized by the CDF and DØ experiments, respectively.

2.2 The CDF II detector

CDF II is a general purpose detector aimed at measuring the observables produced in $p\bar{p}$ collisions. It exhibits approximate cylindrical symmetry around the axis defined by the beamline. Furthermore, it is symmetrical with respect to the plane orthogonal to the beamline and containing the $p\bar{p}$ geometric collision point. The detector is shown in Figure 2.3.

The CDF II detector employs a Cartesian coordinate system which reflects the symmetries of the detector. It is a right-handed set of axes with the origin located in the geometrical center of the detector. The z axis is aligned with the proton direction,

while the y axis points upward and the x axis radially outward with respect to the center of the Tevatron accelerator. The symmetry of the detector also suggests the use of polar coordinates r, φ, and θ. The polar angle θ is defined relative to the z axis.

In hadron colliders, as an alternative to the polar angle, it is also useful to use the rapidity y, defined as follows:

$$y = -\frac{1}{2} \log \frac{E - p_T}{E + p_T}, \tag{2.2.1}$$

where p_T is the component of the momentum on the x–y (r–φ) plane. Differences in rapidity are invariant under Lorentz boosts along the z direction.

The pseudorapidity η is also often utilized as an approximation to rapidity. It is defined as follows:

$$\eta = -\log \tan \frac{\theta}{2}, \tag{2.2.2}$$

and well approximates the rapidity y when the energy of the particle is large as compared to its rest mass. The pseudorapidity is a convenient quantity because in the ultra-relativistic limit of a particle, in which it coincides with the rapidity, differences in pseudorapidity are Lorentz-invariant under $\hat{z}$ boosts. Besides that, the distribution of the light products of a $p\bar{p}$ interaction is roughly flat in η, with a density of about four charged particles per unit of rapidity, at the Tevatron.

Other convenient variables typically utilized are the transverse energy E_T and the approximately Lorentz-invariant angular distance ΔR:

$$\begin{aligned} E_T &\equiv E \sin\theta, \\ \Delta R &\equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2}. \end{aligned} \tag{2.2.3}$$

The innermost detector system is the tracking system. It consists of three Silicon microstrip detectors, Layer00 (L00), the Silicon VerteX detector (SVX), and the Intermediate Silicon Layer (ISL), and a multi-wire drift chamber, the Central Outer Tracker (COT). These detectors are cylindrically symmetric and are designed to record samples of the trajectories of charged particles. These trajectories are referred to as *tracks*.

The Time of Flight (TOF) system, which is designed to provide particle identification for low-momentum charged particles, is located immediately outside the tracking system.

The tracking system and the TOF detector are immersed in a 1.4116 T magnetic field, aligned with the beamline, provided by a superconducting solenoidal coil which is placed immediately outside of the TOF.

Charged particles follow helical trajectories inside a magnetic field, which are completely defined by five parameters, three of which are chosen to belong to the transverse plane of symmetry. These five parameters, illustrated in Figure 2.4, are:

d_0 The impact parameter d_0 measures the distance between the particle trajectory and the z axis at the point of closest approach between the trajectory and the

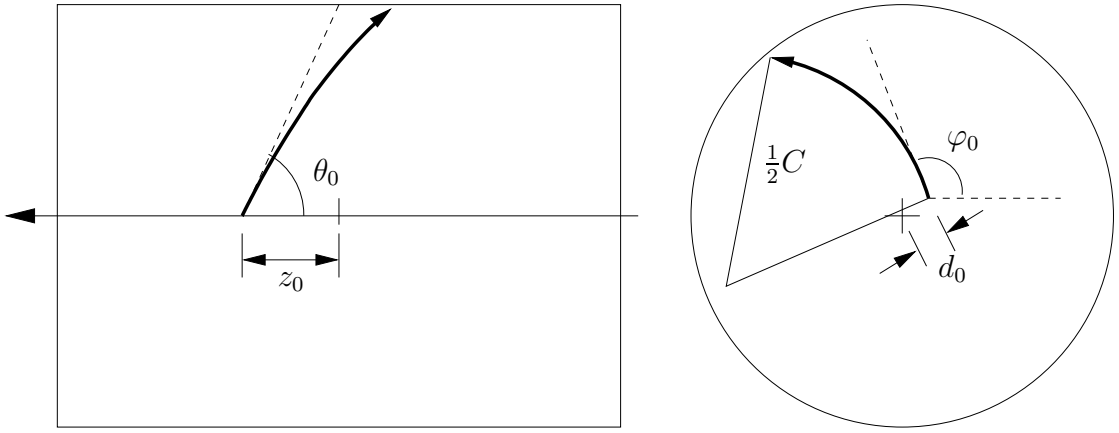

FIGURE 2.4: A pictorial representation of the parameters chosen to describe tracks in the CDF II detector, in the longitudinal, z-y (left), and transverse, x-y or ρ-φ (right), planes. The z axis is oriented in the right-left direction in the longitudinal view, while is indicated by the cross in the middle of the transverse view.

geometrical center of the detector. It is a signed quantity, and is defined as:

$$d_0 = q \cdot \left(\sqrt{x_c^2 + y_c^2} - R \right) , \tag{2.2.4}$$

where q is the charge of the particle, (x_c, y_c) the center of the helix, and R the radius of the circle obtained by projecting the helix on the r–φ plane.

C The curvature C is completely determined by the component of the particle momentum in the transverse plane. In fact, $C = a/p_T$, with $a = 2.115939 \cdot 10^{-3}$ cm^{-1} GeV/c at CDF II.

φ_0 The azimuthal angle φ_0 measures the direction, in the transverse plane, of the momentum of the particle at the point of closest approach to the center of the detector.

z_0 The z cylindrical coordinate of the point of closest approach between the particle track and the z axis defines the z_0 parameter.

λ The last parameter is defined as $\cot\theta_0/2$, where θ_0 is the angle between the z axis and the momentum vector of the particle.

The solenoidal magnet separates the tracking volume from the finely segmented electromagnetic and hadronic calorimeters, responsible for energy measurements of neutral and charged particles. Finally, the subdetectors of the Muon systems are located outside the calorimeters.

More information on the CDF II detector can be found in References [23] and [24], and in specific references for each subdetector. The description of the trigger, tracking and TOF and muon chamber systems are emphasized, because they represent the aspects of the detector more critical to the analyses presented in this document.

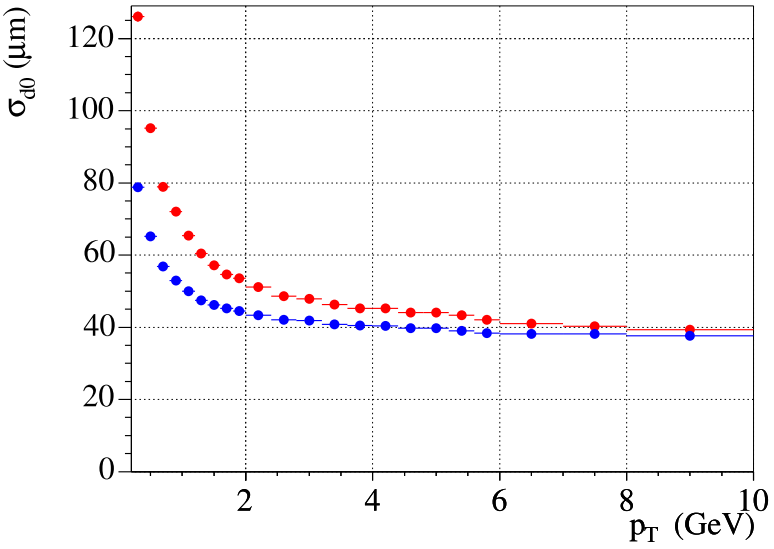

FIGURE 2.5: Impact parameter resolution of tracks with L00 hits (blue/dark) and without L00 hits (red/light), as a function of the transverse momentum of the tracks.

2.2.1 Layer00

Layer00, L00 [25], is the first detector that particles encounter after leaving the interaction point, and provides useful information for the two-dimensional reconstruction of tracks left by charged particles. It consists of a single layer of silicon microstrips, located at a radius of 1.6 cm from the beamline. It measures 80 cm in length. The basic readout elements are 10 cm long, single-sided axial strip sensors. The implant pitch is 25 μm with an alternate strip readout, giving a readout pitch of 50 μm. The single-hit resolution is 6 μm. The total number of channels readout is 13,824.

Figure 2.5 clearly shows the improvement in impact parameter resolution obtained by including L00 hits in the track fits, compared with fits which utilize only the information of the other subdetectors of the tracking system, SVX, ISL, and COT. Typical track momenta for B decay daughters are below 2 GeV/c, where the improvement in resolution is the greatest. The efficiency for adding a L00 hit to the other track hits is 65% and the effect is a 10 to 20% reduction of the impact parameter resolution.

2.2.2 Silicon VerteX detector II

The Silicon VerteX detector, SVXII [26, 27], shown in Figure 2.6, is made of five layers of double-sided silicon microstrip sensors. It extends radially from 2.5 cm to 10.6 cm and covers 87 cm along the z axis, guaranteeing a good geometric coverage up to $|\eta| \simeq 2.0$. Three layers have sensors which allow for the simultaneous measurement of the hit position in the transverse plane (the microstrips are parallel to the z axis) and along the z axis (the microstrips are orthogonal to the beamline direction). The

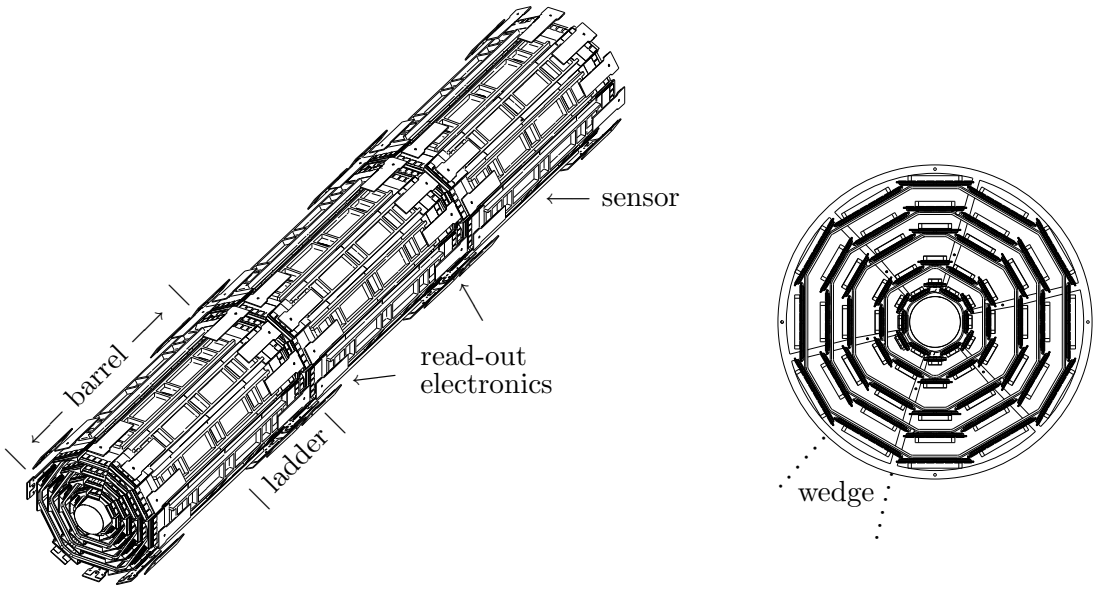

FIGURE 2.6: The Silicon VerteX detector SVXII. An illustration of the three instrumented barrels of SVXII (left) and a cross-sectional view of a barrel in the r–φ plane.

sensors of layers 2 and 4, instead, have microstrips which are orthogonal to a plane with a stereo angle of 1.2° with respect to the z axis, usually indicated with the notation of r–φ' plane. The readout pitch is 60 : 62 : 60 : 60 : 65 μm on the r–φ plane and 141 : 125.5 : 60 : 141 : 65 μm on the r–z or r–φ' planes. The readout pitch is larger for r–z strips to limit the total number of channels to read, which would be excessive for an almost 90 cm long detector. This design permits the three-dimensional reconstruction of tracks. The sensors are arranged in three barrel-shaped regions, each of which is divided into twelve wedges. The active area of silicon is about 2.5 m^2. The 400k channels of SVXII are read in 10 μs, which is fast enough to allow for their use in impact-parameter–based triggers in the second level of the CDF trigger.

2.2.3 Intermediate Silicon Layer

The last silicon-based detector is the Intermediate Silicon Layer, ISL [28], presented in Figure 2.7. It is installed between the SVX and the Central Outer Tracker drift chamber, and consists of three layers of double-sided silicon microstrip modules, with twelve wedges covering the entire azimuthal angle φ. The $|\eta| < 1$ region is covered by a single layer located at 23 cm of radius. This layer provides an additional position measurement which allows for a better extrapolation from the drift chamber to the SVX. Two lateral layers are installed 20 cm and 29 cm far from the beamline, extending longitudinally in the $1 < |\eta| < 2$ region. They permit three-dimensional reconstruction of tracks in a region where the coverage of the drift chamber is partial and allow for stand-alone silicon tracking. Figure 2.8 shows the location of the layers of the silicon subdetectors in the r–z plane.

The sensors have microstrips parallel to the z axis and with a stereo angle of 1.2°

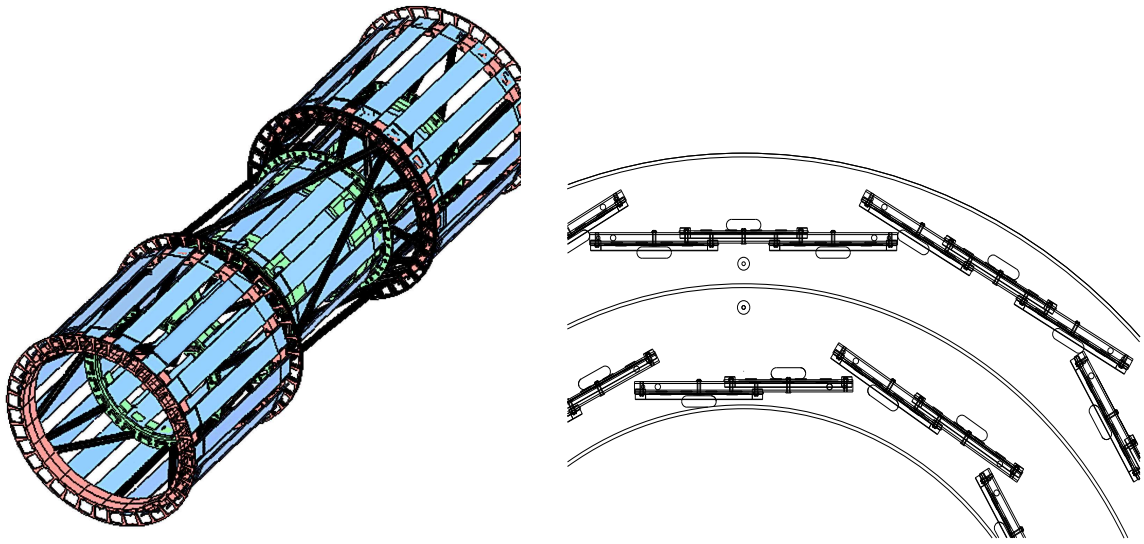

FIGURE 2.7: The Intermediate Silicon Later ISL: 3D view of the three barrels which compose the detector (left) and closeup of one section of the end view of an external barrel (right).

with respect to the same axis, for position measurements in the r–φ and r–φ' planes, respectively. The readout pitch is 112 μm (112 − 146 μm) for axial (stereo) strips, with an expected single-hit resolution of < 16 μm ($< 16 - 23$ μm).

Each readout module, called a *ladder*, consists of three sensors and their readout electronics. The ISL contains 296 ladders, which account for its more than 300k of readout channels. The detector is 174 cm long, with complete coverage in φ. The active area of silicon is 3.5 m^2.

2.2.4 Central Outer Tracker

The Central Outer Tracker (COT [29]) is an open-cell drift chamber, with 8 *superlayers* consisting of 12 layers of wires each, for a total of 96 possible measurements per track. A section of the r–φ view of the detector is shown in Figure 2.9. The active volume of the chamber extends radially from 43.4 cm to 132.3 cm and longitudinally in $|z| \lesssim 155$ cm. Tracks from the center of the CDF detector are completely contained in the COT when $|\eta| < 1.3$. The chamber is filled with a 50 : 50 Argon-Ethane gas mixture bubbled through Isopropyl alcohol (1.7%). In such an admixture, the drift velocity is equal to ~ 50 μm/s and hit signals are collected in less than 200 ns, which is shorter than the inter-bunch spacing of 396 ns. The drift field, the homogeneity of which is guaranteed by the 33k potential wires, is 3.5 KV/cm and the corresponding Lorentz angle is 35°. Instead of the usual field wires, 250 μm-thick gold-plated Mylar sheets separate the COT cells, shown in Figure 2.9. These field panels additionally provide mechanical isolation among cells, thus limiting the possible damages produced by broken wires. The COT contains 2520 cells, each of which has 12 active wires. Tracks are reconstructed in three dimensions exploiting information from the

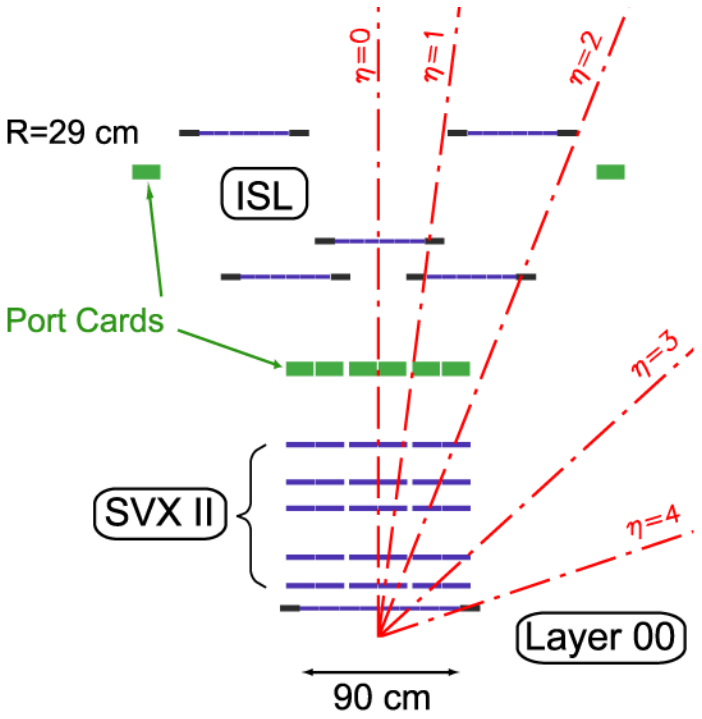

FIGURE 2.8: Coverage of the silicon subdetectors in the r–z plane. The scales of the z and r axes are different.

4 axial superlayers (wires parallel to the z axis) and the 4 stereo ones ($\pm$ 3° stereo angle between wires and z axis).

The tracking performance of the detector turned out to be better than expected. The tracking efficiency for tracks that transverse its entire volume radially is 99% for charged particles with $p_T \geq 2.0$ GeV/c, and falls to 95% when $p_T = 0.5$ GeV/c [30]. The hit resolution is about 140 μm. The transverse momentum resolution σ_{p_T}/p_T is approximately $0.15\% \cdot p_T$[GeV/c], which results in excellent mass resolution of completely reconstructed states. The mass resolution is typically 15 MeV/c^2 for $B_s^0 \rightarrow D_s^- \pi^+$. In addition, silicon measurements close to the beam allow precise reconstruction of decay vertices, with typical resolutions of 30 μm in the transverse plane and 70 μm along the beam direction.

The drift chamber provides important information for particle identification. The signal collected on the wires contains information from the primary ionization electrons, i.e., those directly produced by charged particles transversing the COT, and secondary ionization particles. The secondary ionization is generated by the primary electrons, which are strongly accelerated by the local electric field when they get close to the surface of the wires. The electronics attached to the end of each wire record the arrival time of the ionization charge, given by the leading edge of the measured pulse, and the width of the pulse. The former is utilized for tracking purposes, while the latter encodes the charge information used for energy-loss sampling. The ionization per unit track length (dE/dx) which a particle releases while transversing the COT is characteristic of the particle's velocity and is utilized

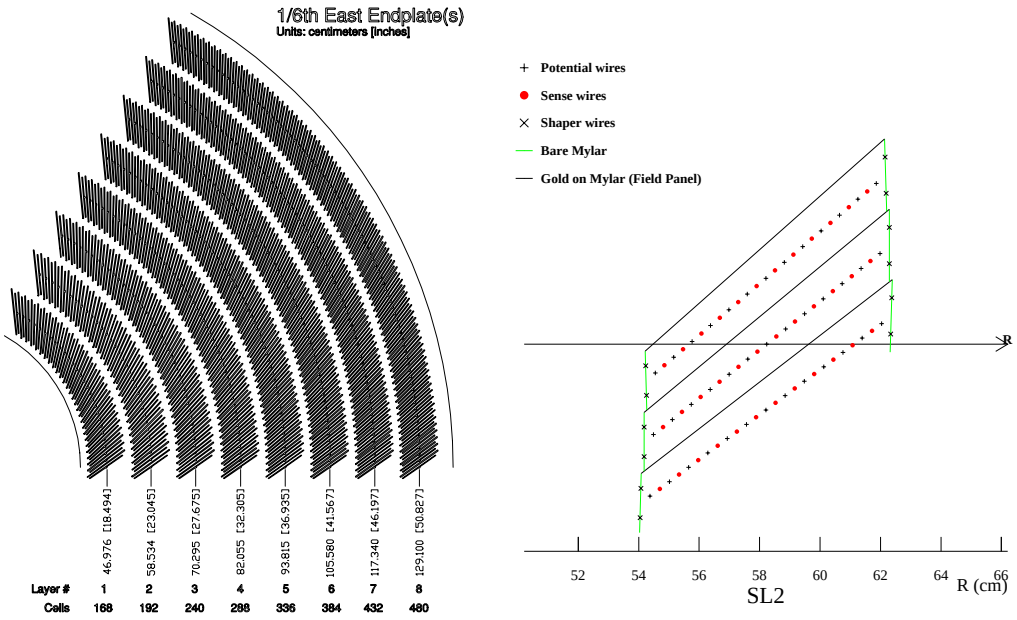

FIGURE 2.9: A view of a 1/6 section of the COT end-plate, in the r–φ plane, which shows the structure in superlayers (left), and a schematic draw of an axial cross-section of three cells in superlayer 2 (right).

to separate kaons, pions and protons. The separating power between kaons and pions, measured by comparing the dE/dx distributions of true kaons and pions, is 1.4 standard deviations in the range $p_T > 2.0$ GeV/c, as seen in Figure 2.10. The pure samples of kaons and pions utilized for the calibration of the pulse-width information are obtained by reconstructing $D^*(2010)^+ \to D^0\pi^+, D^0 \to K^-\pi^+$. The strong D^{*+} decay unambiguously defines the flavor of the D^0 meson, which dominantly decays in the Cabibbo-favored $K^-\pi^+$ mode. The reconstructed final state thus contain two like-sign pions and one oppositely-charged kaon. A sample of protons is obtained by reconstructing the $\Lambda^0 \to p\pi^-$ decay.

2.2.5 Time Of Flight

The Time of Flight (TOF [31]) detector was conceived and realized to provide particle identification capabilities for CDF expressly for B physics. It consists of 216 scintillator bars, approximately 280 cm long and with a cross-section of 4 cm $\times$ 4 cm, installed between the COT and the cryostat which contains the superconducting solenoid, at a radial distance of 140 cm from the interaction point (Figure 2.11). Each bar is equipped with photomultiplier tubes at both ends. The photomultipliers have a special design. The dynodes of a classic photomultiplier are replaced by aligned grids, "fine mesh" design, which allow the electron cascade to develop longitudinally, parallel to the magnetic field. This configuration permits the maintenance of an adequate gain even in the 1.4 T magnetic field in which the photomultipliers operate. The location of the TOF installation and the scintillator-photomultiplier assembly are shown in Figure 2.11.

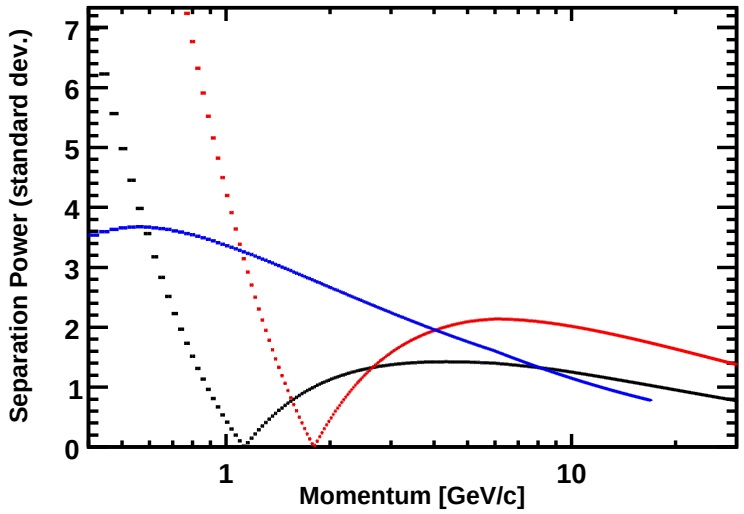

FIGURE 2.10: COT separating power in units of standard deviations *vs.* transverse momentum. The separation between pions and kaons, protons, and electrons are shown in black dashed, red dotted and blue solid line, respectively.

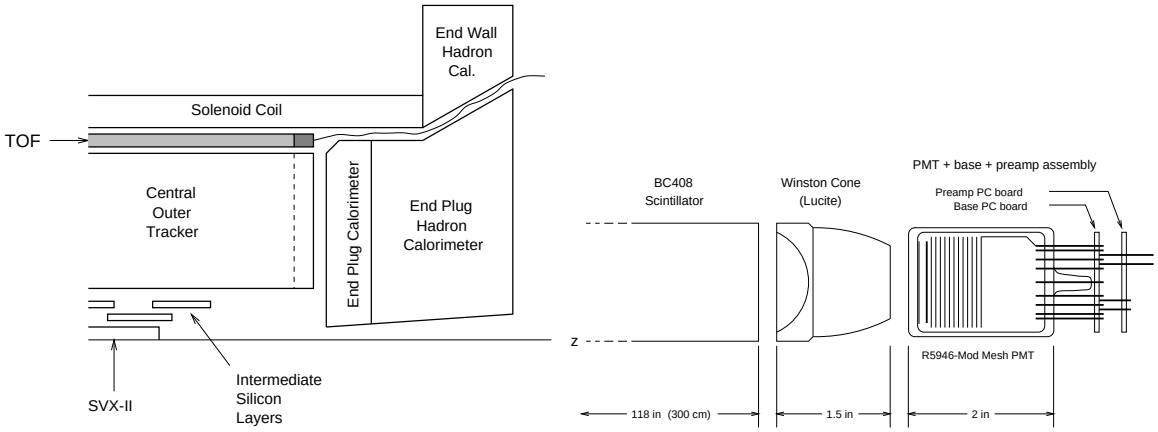

FIGURE 2.11: The Time Of Flight detector. On the left, the location of TOF is shown in a side view of CDF II. On the right, the arrangement of the scintillator, Winston cone and photomultiplier assembly. The Winston cone optimizes the optical coupling between the scintillator and the photomultiplier.

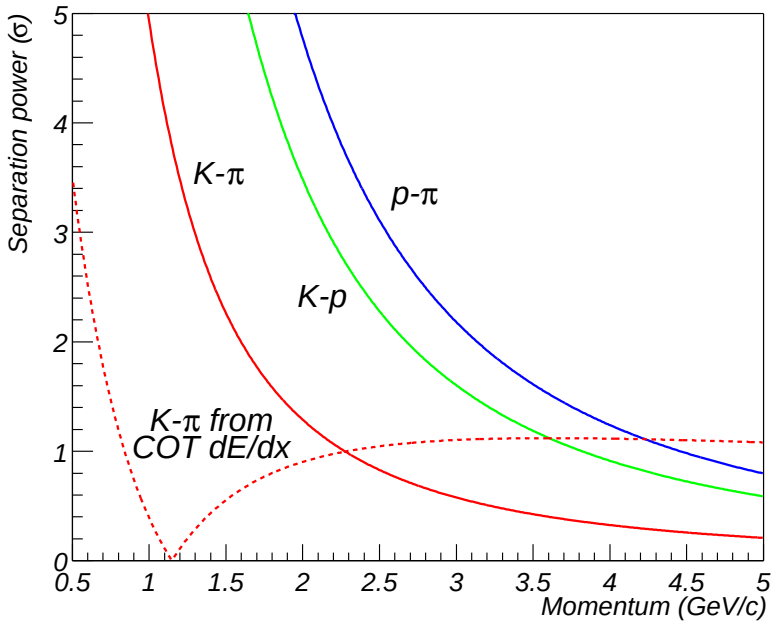

FIGURE 2.12: Expected TOF separating power, in units of standard deviations *vs.* momentum. The dashed line reports the $\mathrm{K}-\pi$ separation provided by specific ionization in the COT.

The TOF system plays a major role in analysis like the B_s^0 mixing one. The measurement of the arrival time (t_{flight}) to the TOF, with respect to the bunch-crossing time, of a particle allows one to infer the mass of the particle according to the following relation:

$$m = \frac{p}{c}\sqrt{\frac{c^2 t_{\mathrm{flight}}^2}{L^2} - 1}\,, \tag{2.2.5}$$

where p is the momentum of the particle and L is the path length, both precisely measured by the tracking system. The resolution of the measured t_{flight} of a particle is described by two Gaussians, the narrower of which has width between 100 ps and 120 ps, and contains 85% of the area of the resolution function. Such resolution allows for kaon-pion separation, which is fundamental for the same-side flavor tagging algorithm used in this thesis, at the $>$ 2–standard-deviations level for tracks with $p_T < 1.5$ GeV/c (Figure 2.12).

2.2.6 Calorimetry

All the calorimetric detectors in CDF are based on plastic scintillators. Layers of scintillator and absorbers are alternated to form sampling calorimeters in the shape of towers which subtend a portion of solid angle, segmented in rectangular cells in the η–φ plane. Each tower is divided into two compartments: on the inside is the electromagnetic calorimeter, using lead as absorber, which is followed by the hadronic

	η coverage	Thickness	Resolution [%]
Central EM (CEM)	$\lvert\eta\rvert < 1.1$	$19X_0$, $1\lambda_0$	$14/\sqrt{E[\text{GeV}]\sin\theta} \oplus 2$
Plug EM (PEM)	$1.1 < \lvert\eta\rvert < 3.6$	$21X_0$, $1\lambda_0$	$16/\sqrt{E[\text{GeV}]\sin\theta} \oplus 1$
Central HA (CHA)	$\lvert\eta\rvert < 1.1$	$4.5\lambda_0$	$50/\sqrt{E[\text{GeV}]\sin\theta} \oplus 3$
Wall HA (WHA)	$0.7 < \lvert\eta\rvert < 1.3$	$4.5\lambda_0$	$75/\sqrt{E[\text{GeV}]\sin\theta} \oplus 4$
Plug HA (PHA)	$1.3 < \lvert\eta\rvert < 3.6$	$7.0\lambda_0$	$74/\sqrt{E[\text{GeV}]\sin\theta} \oplus 4$

TABLE 2.2: Summary of the coverage, thickness and resolution of the CDF calorimeters [23]. The thickness is expressed in terms of the radiation length X_0 and the interaction length λ_0.

calorimeter on the outside, which instead contains iron and plastic scintillator. The coverage is complete in the azimuthal angle φ and up to $|\eta| < 3.6$. The η coordinate distinguishes two areas: Central and Plug.

In the next paragraphs, the different subdetectors of the CDF calorimeter system are discussed. A summary of their main characteristics is presented in Table 2.2.

Central calorimeters

The calorimeter in the Central region covers the $|\eta| < 1.1$ range in pseudorapidity. Each tower measures $\Delta\eta \times \Delta\varphi = 0.1 \times 15^\circ$ in the η–φ plane. The Central ElectroMagnetic calorimeter (CEM [32]) contains 5 mm-thick layers of scintillator and 3.4 mm-thick layers of lead, which corresponds to $0.6X_0$, where $X_0 = 0.56$ cm is the radiation length of lead. A particle incident normal to the detector transverses the detector encounters $19X_0$ and $1\lambda_0$ of matter (λ_0 is the nuclear interaction length, $\lambda_0(Pb) = 17.09$ cm).

The central electromagnetic calorimeter is integrated by two detectors which provide information about the position and shape of electromagnetic showers. A set of multi-wire proportional chambers (the Central Preshower Radiator, CPR) was installed between the solenoid and the first layer of the calorimeter to monitor photon conversions started in the tracker material or in the magnetic coil, which acts as a radiator. The CPR was replaced by a finely segmented layer of scintillators [33] during the programmed interruption of Tevatron operations in the fall of 2004. Another set of wire chambers (CES) is located at a radial depth of $6X_0$, where the peak of shower development is typically located. The transverse shower-shape is measured with 2.0 mm resolution (for 50 GeV electrons). The CPR and CES systems provide useful pieces of information for the identification of electrons.

The Central HAdronic calorimeter (CHA [34]), behind the CEM, contains 10 mm-thick layers of scintillator alternated with 2.54 cm-thick layers of steel. The total depth of the hadronic calorimeter, which contains 32 layers of absorber, is $4.5\lambda_0$.

The hadronic section is completed by the wall hadronic calorimeter (WHA), which imitates the structure of the central hadronic calorimeter, extending its coverage up to $|\eta| < 1.3$. The wall calorimeter contains only 15 layers of 5.1 cm-thick absorber, which explains its worse energy resolution, as shown in the summary in Table 2.2. Figure 2.13 presents a sketch of a sector of the Central calorimeter, while photographs

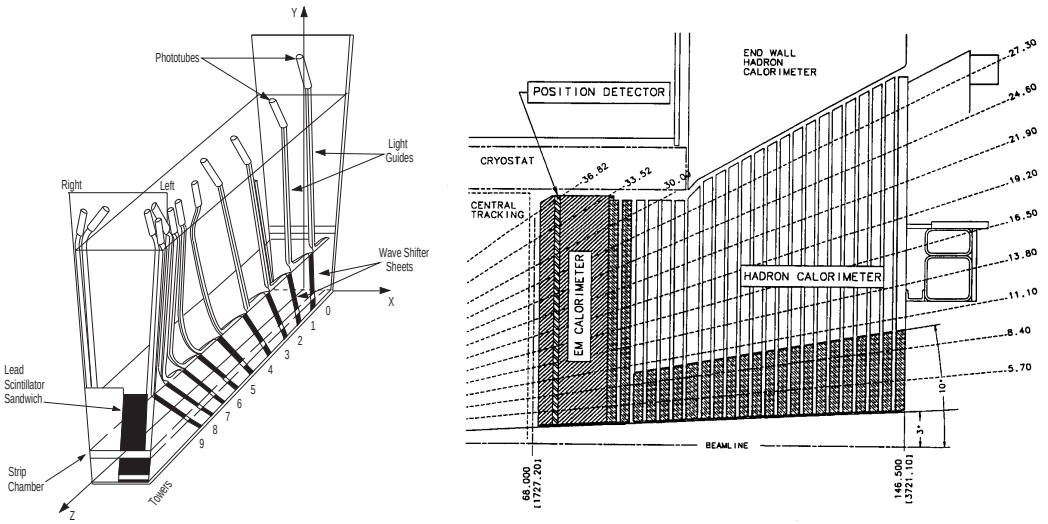

FIGURE 2.13: Schematic view of an azimuthal sector of central calorimeter (left) and elevation view of the upper part of the plug calorimeter (right). The elevation view on the right also indicates the location of the central calorimeters, above the cryostat, and the wall hadronic calorimeter, on the right of the central calorimeters and above the plug hadronic calorimeter. The plug shower-max detector is visible inside the plug electromagnetic calorimeter.

and further drawings of it may be found in the papers cited in this section.

Plug calorimeters

The towers of the plug calorimeter, which is shown in Figure 2.13, measure $\Delta\eta\times\Delta\varphi = 0.1-0.16\times 7.5^\circ$ for $1.1 < |\eta| < 2.1$ and $\Delta\eta\times\Delta\varphi = 0.2-0.6\times 15^\circ$ for $2.1 < |\eta| < 3.6$. Their structure resembles the calorimeter in the central area. The electromagnetic section (PEM [35]) is constituted by a sampling calorimeter. A PEM sampling unit is made by a layer of lead and one of scintillator, 4.5 mm ($0.8X_0$) and 4 mm-thick, respectively. The 23 samplings in each tower cover $21X_0$, $1\lambda_0$.

The position and shape of electromagnetic showers in the plug region are measured by a preshower detector (the Plug PReshower detector, PPR) and a shower-max position detector (the Plug Shower Max, PSM [36]). The first sampling unit of the PEM (i.e., the closest to the geometrical center of the CDF II detector) contains exceptionally thick scintillator layers (10 mm) which are individually read out and constitute the PPR. Incorporated in the plug calorimeter at a depth of $6X_0$ are the components of the PSM, designed to provide measurements at the nominal shower maximum. These consist of two layers of scintillator strips with 5 mm pitch and a 45° crossing angle between strips in the two layers, read out with wave-length shifting fibers. They measure the spatial position and profile with a resolution of 1 mm where the shower is at its greatest development.

The Plug HAdronic calorimeter (PHA) contains 23 sampling units, each of which has 6 mm of scintillator and 50 mm of iron. The depth of the detector measures $7\lambda_0$.

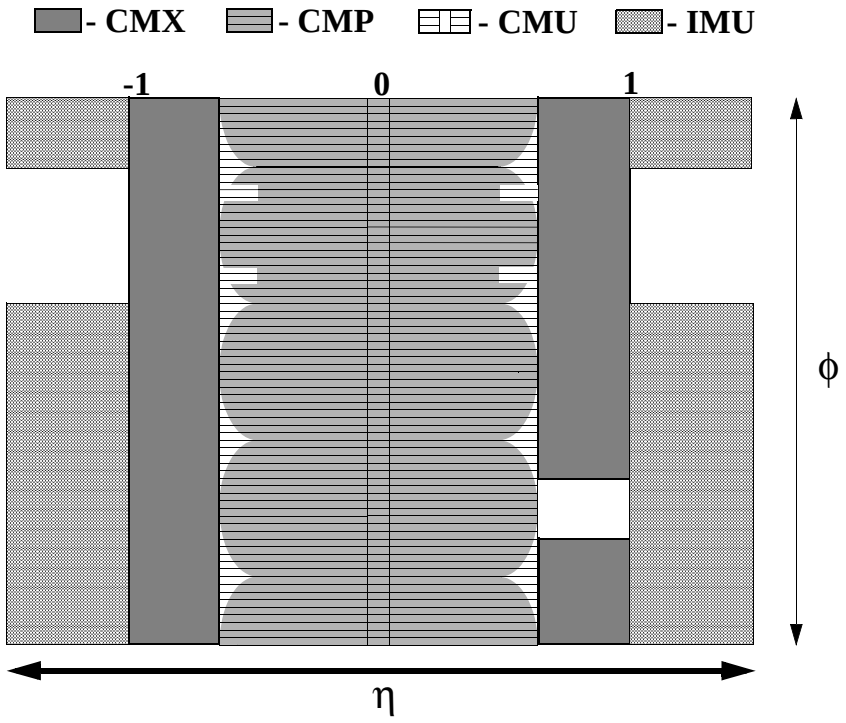

FIGURE 2.14: Coverage of muon chambers. The hatched and shaded areas represent the regions in η–φ which are instrumented by the subdetectors of the CDF muon system. The gap in the coverage of the CMX detector corresponds to the top area on the east side of CDF where the cryogenics system of the CDF solenoid is installed. The uninstrumented region of the IMU detector corresponds to the support structure of the toroids which hold the IMU muon chambers.

The layers of the Plug calorimeter have annular shape and the outer radius of each hadronic module increases with increasing $|z|$, producing the characteristic "plug" shape of the calorimeter.

2.2.7 Muon chambers

CDF II uses four independent systems of scintillators and drift chambers to detect muons in the $|\eta| < 1.5$ region. The subdetectors which compose the muon system are installed outside of the calorimeters and represent the last part of the CDF detector that a particle can interact with. Single-wire, rectangular drift chambers filled with a 50 : 50 gas mixture of Argon-Ethane compose the subdetectors. The chambers are arranged in staggered arrays with four layers, with various azimuthal segmentation, and are coupled to scintillators. Scintillators provide timing information to suppress backgrounds due to secondary interactions in the beam pipe material and cosmic rays. Hits in three matching radial layers constitute a *muon stub*. A muon stub corresponding to the extrapolation of a COT track identifies a muon candidate. The coverage of the CDF muon system in the η–φ space is shown in Figure 2.14.

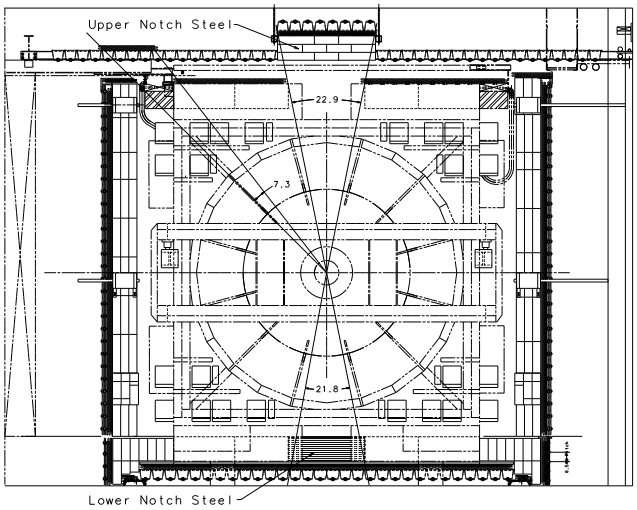

FIGURE 2.15: Central Muon uPgrade CMP. In this x–y section of the CDF detector, the CMP forms a rectangular box (dark shaded area) which surrounds the other parts of the CDF detector.

The Central MUon detector (CMU [37]) and Central Muon uPgrade detector (CMP [38]) cover the central region ($|\eta| < 0.6$) providing a measurement of the z and φ coordinates of the muon candidate. The CMU is installed at a radius of 347 cm from the beam axis, at a depth of $5.5\lambda_0$ from the interaction point. Each of the 144 modules of the CMU contains 16 cells, stacked four deep in the radial direction. The difference in arrival-time of the drift electrons between cells in different layers provide a resolution in the drift direction as good as 250 μm. Division of the charge collected at the two extremities of sense wires allows for the measurement of the z position of hits with up to 1.2 mm resolution.

The CMP is a second set of drift chambers, located behind an additional 60 cm of steel. The chambers are arranged to enclose the detector inside an approximately rectangular box (Figure 2.15). The wall drift chambers (i.e., the chambers which span the y–z plane) are coupled to a layer of scintillator counters, installed on the outside surface of the chambers. The purpose of CMP is to cover the φ gaps of CMU and enhance the rejection of penetrating high energy hadrons (fake muons).

The Central Muon eXtension detector (CMX) operates in $0.6 < |\eta| < 1.0$. Two layers of scintillator counters cover the internal and external surface of an eight-layer array of drift chambers. The CMX is installed at a radial distance of $400-600$ cm from the beam axis. Its chambers are arranged to form an arch, as shown in Figure 2.16. The azimuthal coverage is not complete in the east side of the CDF detector. The region which would contain the uppermost edges of the CMX detector is occupied by the cryogenics system of the CDF solenoid.

The Intermediate MUon system (IMU [39]) is used to identify muons in the $1.0 < |\eta| < 1.5$ region, with three-quarters of the azimuth instrumented. The incompleteness of the azimuthal coverage is due to the presence of support structures. The IMU consists of four staggered layers of drift chambers and a layer of scintillation

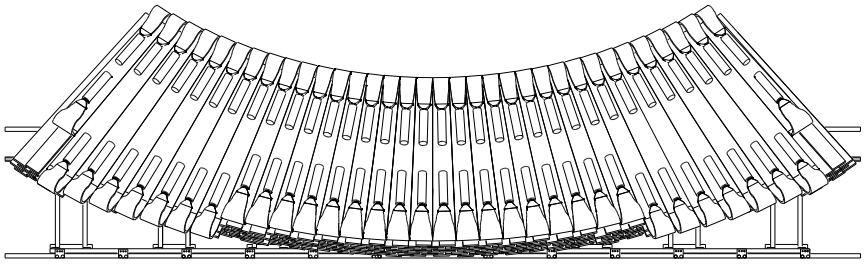

FIGURE 2.16: Central Muon eXtension. The drawing shows an elevation view in the x–y plane of a section of the CMX detector. The part depicted is referred to as "miniskirt" because it covers the lower section of the azimuthal range.

	η coverage	φ coverage	Depth	Minimum $p_T(\mu)$
CMU	$\lvert\eta\rvert < 0.6$	302°	$5.5\lambda_0$	1.4 GeV/c
CMP	$\lvert\eta\rvert < 0.6$	360°	$7.8\lambda_0$	2.2 GeV/c
CMX	$0.6 < \lvert\eta\rvert < 1.0$	360°	$6.2\lambda_0$	1.4 GeV/c
IMU	$1.0 < \lvert\eta\rvert < 1.5$	270°	$6.2 - 20\lambda_0$	$1.4 - 2.0$ GeV/c

TABLE 2.3: Summary of the coverage, thickness and minimum detectable p_T, on average, of the CDF muon detectors. The depth is expressed in pion interaction lengths and is quoted for a reference axial angle $\theta = 90°$ in CMU and CMP, and $\theta = 55°$ in CMX.

counters, mounted on the outer radius of two steel toroids. Due to the geometry of the installation, the amount of material that a particle has to cross before reaching the IMU chambers varies between 6.2 and 20 interaction lengths in the $|\eta|$ range covered by the subdetector. The IMU system is installed around the toroids (hatched shading) in the center of Figure 2.17.

The CDF calorimeter, the magnet return yoke, and additional steel shielding act as muon filters suppressing hadrons from reaching the muon chambers. The muon purity increases with the effective shielding, but at the expense of efficiency for low momentum muons, which do not have enough energy to fly through the shielding. The effective hadronic shielding and the minimum momentum that, on average, a muon must have to reach the muon detectors are summarized in Table 2.3, where the η–φ coverage of each muon subsystem is also reported.

2.2.8 CDF trigger system

The online selection of events with interesting physics content is crucial in the $p\bar{p}$ environment where CDF operates. The total cross-section of $p\bar{p}$ inelastic interactions is ~ 60 mb, which, at the luminosity of 10^{32} cm^{-1}s^{-1}, yields a rate of inelastic interactions of the order of 6 MHz. Moreover, because the average size of the information associated to each event is ~ 140 kbyte, an approximate throughput and storage rate of 840 Gbyte/s, unattainable with the currently available technology, would be needed to record all events. However, the cross-sections of interesting physics processes are many orders of magnitude smaller than the inelastic $p\bar{p}$ cross-section (for example,

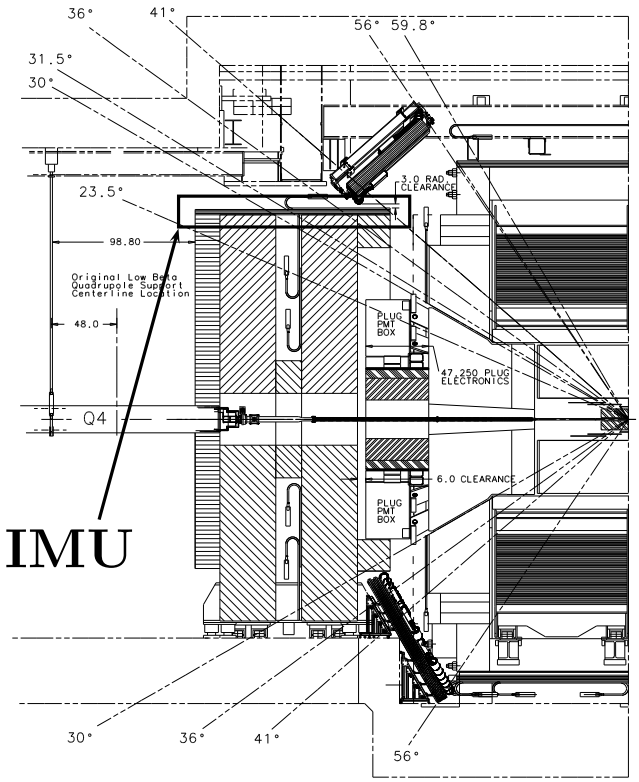

FIGURE 2.17: Intermediate MUon system IMU. The diagonally hatched area in the middle of the figure represents the sectional view of the toroids around which the IMU (dark shaded) is installed. The figure also shows a y–z view of the CMX detector, which corresponds to the two dark shaded areas in the middle of the picture that extend diagonally.

the total cross-section for $b\bar{b}$ production is about 0.1 mb), and the online preselection of events adapts the interaction rate to the storage rate of CDF.

The CDF detector utilizes a three-level trigger system which performs the online selection of events enriched in events with interesting physics. The input event rate is reduced at each level, providing increasing time for more complex and accurate reconstruction tasks. The rate of events which satisfy the trigger selection is ~ 75 Hz. The trigger system is designed to limit the *deadtime* to a minimum, during which events are discarded because no resources are available to process them. The schemes of the CDF data acquisition and trigger systems are shown in Figures 2.18 and 2.19.

Level-1 Trigger

The first level of the trigger (Level-1) utilizes custom designed hardware to find physics objects, such as tracks, or lepton candidates, based on a subset of the detector information. Events which satisfy selection criteria based on these objects are passed to the second level of the CDF trigger.

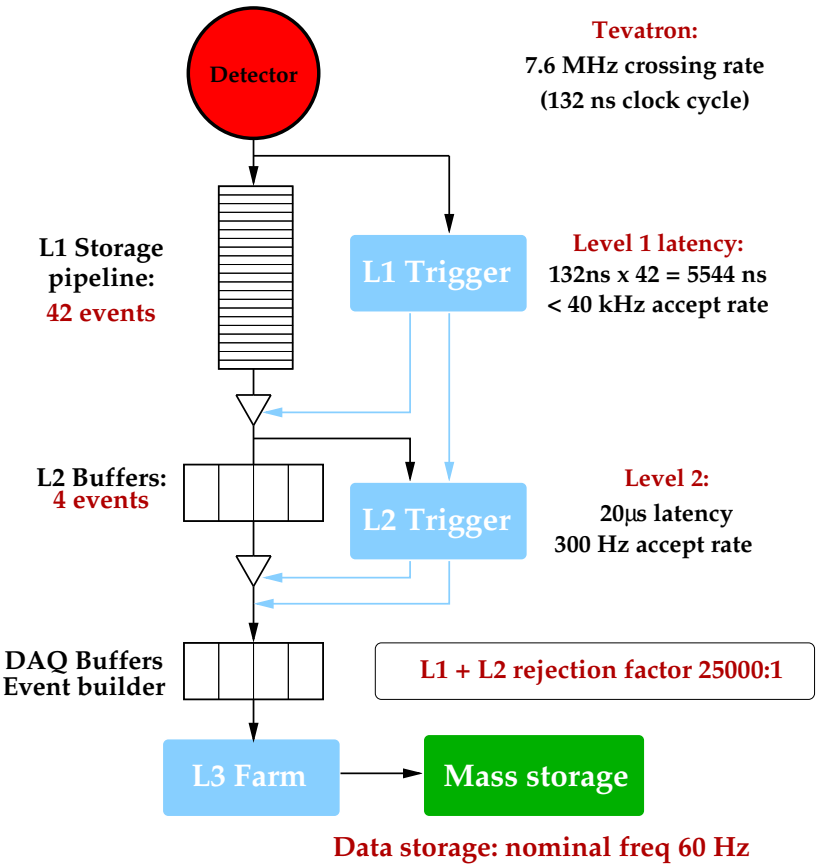

FIGURE 2.18: The CDF data acquisition system. The scheme emphasizes the timing information (latency, input rate and rejection factor) of the three levels of the trigger.

The first element of the Level-1 trigger consists of a FIFO pipeline with buffers for 42 events. The input rate is about 10 MHz and the output rate to the second level is approximately 20 kHz. Events are continuously fed to the pipeline at the rate of the Tevatron clock-cycle, i.e., 132 ns. Because the inter-bunch time is 396 ns, two thirds of cycles, corresponding to empty crossings, are automatically rejected. The pipeline thus collects a maximum of 14 bunch crossings.

The system has 5.5 μs at most to perform a trigger decision, before the pipeline is filled completely. Events which are not flagged by the system before they reach the end of the pipeline are rejected. Events are similarly lost if, even after a Level-1 accept, Level-2 is unable to process a new event because its four buffers are full. The latency of the Level-2 decision, which is 5.5 μs $\times 4 \simeq 20$ μs, is less than approximately 80% of the average time between Level-1 accepts, in order to minimize deadtime.

The input to the Level-1 system consists of a simplified subset of data coming from the COT, the calorimeters and the muon chambers, which are processed by custom-designed hardware to produce low-resolution physics objects, called *primitives*. The information from these objects is then combined into more sophisticated ones. For example, track primitives are matched with muon stubs to form muon objects, which

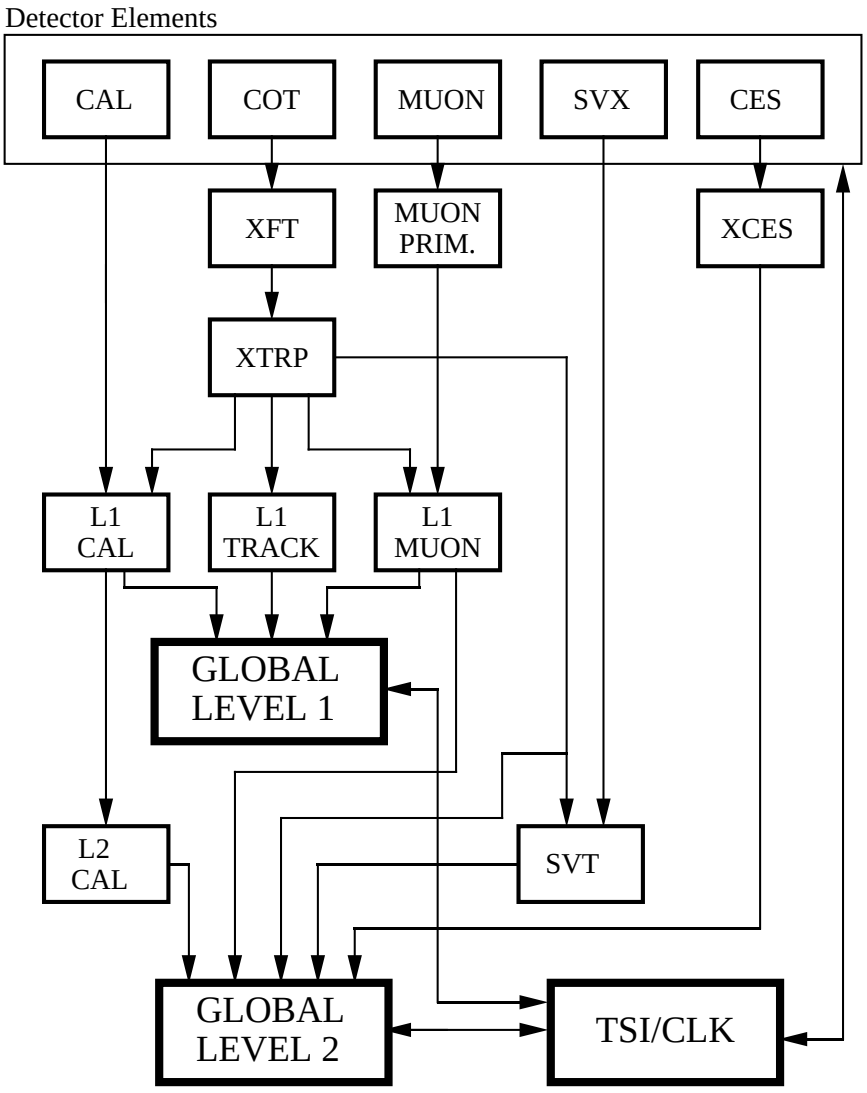

FIGURE 2.19: Scheme of the CDF trigger system. The connections between the subsystems of the CDF detector, in the upper part of the scheme, and the boards which constitute the trigger system are indicated. The various parts of the trigger system are described in the text in the relevant sections.

Parameter	Value
Track finding efficiency	96%
p_T resolution, σ_{p_T}/p_T^2	$1.7\%(\mathrm{GeV/c})^{-1}$
φ_6 resolution, σ_{φ_6}	3 mrad

TABLE 2.4: Performance of XFT.

are subjected to basic selections.

Track primitives constitute an important part of the trigger selections which are used to collect the B_s^0 data samples utilized in the analysis documented in this thesis. The online track processor which produces track primitives for the Level-1 trigger is the eXtremely Fast Tracker (XFT [40]). The XFT utilizes the hits on the four axial layers of the tracking chamber and produces 2D reconstruction of tracks in 2.7 μs, and thereby measuring the transverse momentum p_T and the azimuthal angle of the track on the sixth superlayer of the COT φ_6. These variables are used for track-based preselection of events. The performance of XFT is summarized in Table 2.4. The reconstruction proceeds by searching coincidences between the observed combinations of hits in each superlayer and a set of predetermined patterns. Each coincidence, which require a combination of hits with a minimum of 11 (out of 12) hits per superlayer[1], provides a track segment. Subsequently, a four-out-of-four match is sought among segments in the four superlayers, by comparing the segments with a set of about 2,400 predetermined patterns corresponding to all tracks with $p_T \gtrsim 1.5$ GeV/c originating from the beamline. The COT is logically divided by the XFT in 288 segments, with a unique track allowed per 1.25° segment. The pattern matching is performed in parallel in each of the 288 segments. If no track is found using all four superlayers, then the best track found in the innermost three superlayers is output.

The tracks found by the XFT are not uniquely utilized for track-based triggers, but are redistributed by the eXTRaPolation unit (XTRP), as shown in Figure 2.19, to the subsystems of the Level-1 trigger, which produce the objects of the trigger selection using the XFT track primitives. The XTRP is responsible for the extrapolation of the XFT tracks to the calorimeter and muon detector systems for matching with calorimeter towers and muon stubs. The XTRP also saves the XFT tracks in a buffer, ready to send them to the second level of the trigger in case the event is accepted.

The Level-1 subsystem that produces the calorimeter-based trigger is called L1CAL. Clusters of energy left in the calorimeters, formed by applying thresholds to individual calorimeter towers, are utilized to create primitives such as photons, jets[2], and electrons, the latter requiring an extrapolated XFT track to match with a calorimeter tower. The track extrapolation is done using look-up tables. The calorimeter trigger is also based on global event variables, such as the missing transverse energy $\not\!\!E_T$, and the total transverse energy $\sum E_T$. The transverse energy E_T is calculated by summing the calorimeter data into trigger towers weighed by $\sin\theta$.

[1] 10 hits out of 12 were required before October 2002.

[2] In a proton–anti-proton collision, a large transverse momentum outgoing parton manifests itself as a cluster of particles traveling roughly in the same direction. These clusters are referred to as "jets".

The L1MUON subsystem combines muon stubs in the muon chambers and track primitives into μ objects. The XFT-track primitive is extrapolated to the radii of the muon chambers by means of look-up tables. The presence in an event of objects of this type, Level-1 muons, characterizes a large class of trigger requirements.

Trigger decisions which are based solely on track information are produced by the L1TRACK subsystem. If more than six tracks are found by the XFT, an automatic Level-1 accept is generated. Otherwise, the p_T and φ_6 information is utilized to interrogate look-up tables to generate various Level-1 triggers.

Triggers are often in the awkward situation of requiring a reduction of their rate in order to accommodate them in the available bandwidth, but yet it is not possible to tighten the selection requirements without biasing the selected sample. The application of a randomized trigger rejection according to a *prescale* factor provides a solution. The prescale, a number larger than unity, represents the number of events which, though satisfying the trigger condition, are rejected for each accepted event, and thereby artificially reducing the trigger rate by the prescale factor. The CDF trigger system adopts three different types of prescale: *fixed*, in which the prescale factor does not change; *dynamic*, in which the prescale is reduced in integer steps as the instantaneous luminosity decreases and frees trigger bandwidth; and *uber-dynamic*, in which the trigger system feeds the Level-2 buffers with an event which passed the Level-1 trigger whenever they appear to be able to receive an additional event. Trigger prescales, as a function of time, are recorded in a database, together with the description of the run configuration, in order to allow physicists to precisely know the amount of luminosity which has been integrated.

Level-2 Trigger

The second level of the trigger (Level-2) consists of five subsystems which provide input to four programmable Level-2 processors in the Global Level-2 decision crate. These subsystems are represented, in Figure 2.19, by the five arrows which provide an input to the Global Level-2 decision board. Three of them are explicitly indicated in the scheme (L2CAL, XCES, and SVT), while the inputs from the XTRP and the L1MUON board feed the L2TRACKING and L2MUON modules, respectively.

L2CAL exploits the information from the calorimeters to define energy clusters, utilized for jet triggers. Due to time-constraints, it is not possible to perform cluster finding to reconstruct jets at Level-1. Thus, energy thresholds are applied to individual towers. Because jets are not fully contained by Level-1 trigger towers, these thresholds are set much lower than the energy of jets to provide an efficient trigger. This results in rates that are too high for readout into Level-3. Rates are reduced by performing the reconstruction of jets using clusters of towers, thus being able to base the trigger on more refined objects. The cluster finding algorithm starts from a tower with energy larger than a predefined threshold, which represents a seed for the cluster. All nearby towers with energy larger than a lower threshold, the "shoulder" towers, are then added to the seed tower. The reconstruction of a cluster is performed in parallel on all seed towers.

XCES refines the electromagnetic objects found at Level-1 utilizing the informa-

tion of the CES detector, located at the point of maximum development of electromagnetic showers in the central EM calorimeters. The signals in four adjacent CES wires are added and compared to a threshold to form a XCES bit, with azimuthal resolution equal to 2°. The resolution is finer than the one provided by the calorimetric towers and allows for a better discrimination of electrons from backgrounds by matching XFT tracks with CES information. The matching of an XFT track with an XCES cluster (i.e., the summed signals from four adjacent CES wires) is performed by Level-2 processors in the Global Level-2 decision crate.

The L2MUON processor is responsible for the construction of Level-2 muon candidates. The muon objects utilized by the Level-2 trigger have a more refined φ segmentation than Level-1 muons, 1.25° *vs.* 2.5°.

The Silicon Vertex Tracker (SVT [41, 42]) is the most innovative part of the CDF trigger. It utilizes XFT tracks and SVXII hits, and reconstructs tracks, although in the transverse plane only, with a resolution which is comparable with offline reconstruction algorithms. The revolutionary impact of the SVT consists in it performing online measurements of impact parameters of charged particles with a rate of 30 kHz. Their displacement with respect to the beamline is correlated with the lifetime of the decaying particle which produced them. The SVT is capable of discriminating $\mathcal{O}(100\ \mu\text{m})$ impact parameters from the $\mathcal{O}(10\ \mu\text{m})$ beam spot, fast enough to allow for the use of this information at Level-2. The speed of the SVT is largely due to a highly-parallelized architecture, which matches the geometrical segmentation of the SVXII. The twelve azimuthal sectors of each of the six half barrels of the SVXII are processed independently. The SVT requires the coincidence of an XFT track and hits in four axial SVXII layers. Track reconstruction consists of two stages. In the first, low-resolution, stage, adjacent detector channels are grouped together into super-bins, the width of which is programmable, with 250-700 μm typical values. A set containing about 95% of all super-bin combinations in four SVXII layers compatible with the trajectory of a charged particle with $p_T \gtrsim 2$ GeV/c originated from the beamline ("patterns") is calculated in advance from simulation and stored in the SVT internal memory. The combination of super-bins containing hits corresponding to the track which is being reconstructed is matched to a stored pattern. A low-resolution candidate track, called "road", consists of a combination of four excited super-bins plus the XFT track parameters. A maximum of 64 roads per event is retained for further processing after the first stage of pattern matching. The second stage of track reconstruction consists of a linearized fit. No exact linear relation holds between the transverse parameters d_0, C and φ_0 of a track in a solenoidal field and the coordinates of hits on a radial set of flat detector planes. It is shown in Reference [43] that for $p_T > 2$ GeV/c, $|d_0| < 1$ mm and $\Delta\varphi_0 < 15°$, a linearized fit biases the reconstructed d_0 by at most a few percent. The SVT 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. The constants which define the linear expansion are determined by the geometry of the detector and the beamline alignment. They are calculated in advance and stored in the internal memory of the SVT. The fit for the track is then reduced to the evaluation of a set of scalar products, which is performed within 250 ns per track. The distribution of SVT-measured impact pa-

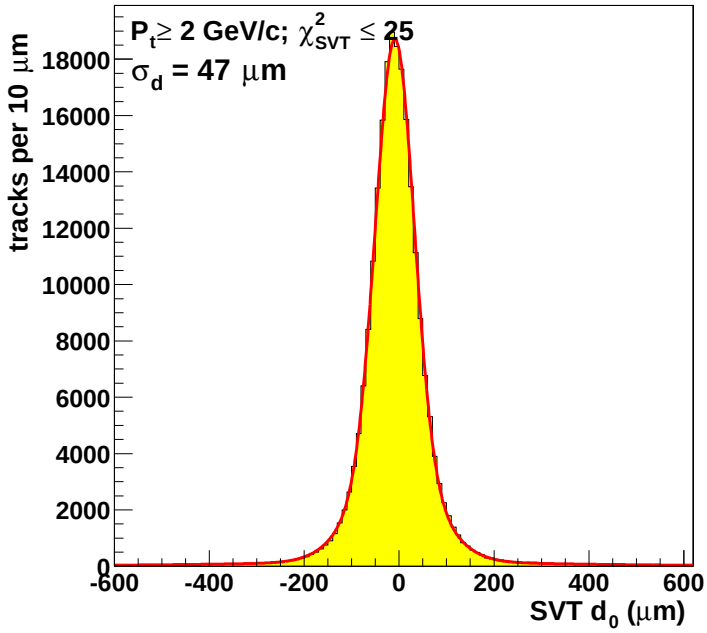

FIGURE 2.20: Impact parameter resolution as measured by SVT. The distribution includes the effect of the beam size. The SVT measures impact parameters with a r.m.s. width of 35 μm.

rameters of prompt tracks, i.e., those tracks associated to particles produced in the hard $p\bar{p}$ interaction, is shown in Figure 2.20. The r.m.s. width of the distribution, $\sigma \sim 47$ μm, includes the contribution of the transverse beam-spot size, while the SVT resolution is $\sigma_{SVT} \sim 35$ μm. The SVT efficiency is higher than 85%. This efficiency is defined as the ratio between the number of tracks reconstructed by the SVT and all the offline tracks of physics analysis quality which contain silicon hits and are matched to an XFT track.

Tracking information is collected by the L2TRACKING module, which receives the XFT tracks from the XTRP and the Level-2 tracks from the SVT, which include impact parameter information. The data from the SVT arrives later than the data from the other systems, because it takes on average 10 μs to process the SVXII, which is the total time allocated to collect Level-2 data. The L2TRACKING module — and the other Level-2 processors — starts analyzing the event before SVT data is complete. The impact parameter information is utilized only if it is required to make the Level-2 decision, while it is not tested if all the triggers which require SVT information are rejected by other cuts.

The system works as a two-stage pipeline with a design latency of 20 μs for an event. During the first stage, which takes 10 μs, events are loaded in the memory of the Level-2 processors. At the same time, L2CAL processes the calorimeter data and the SVT collects data from the SVXII. The last 10 μs are utilized by the Global Level 2 system to make the final Level-2 trigger decision. During the latter phase, the next event is loaded and analyzed. The Level-2 system uses four buffers to maintain the fraction of deadtime below a few percent. The output rate of the Level-2 trigger is

limited to the input capacity of the trigger of the third level, which is about 300 Hz.

Level-3 Trigger

The third level of the trigger (Level-3) is formed by a farm of commercial computers, running the `LINUX` operating system. The maximum input rate, which is identical to the output rate of the Level-2 trigger, is 300 Hz, and the Level-3 output rate is limited by a maximum mass storage rate of 20 Mbyte/s at which data are recorded to disk, and roughly corresponds to 75 Hz.

Upon Level-2 accept, the data from the whole detector are sent to the Level-3 farm by the EVent Builder (EVB [44]) system, as opposed to the Level-1 and Level-2 triggers, which only receive data from some subdetectors. The EVB assembles event fragments from the front-end crates of the CDF subdetectors in a unique event record, a block of data corresponding to a bunch crossing. As shown in Figure 2.21, data are first received by the VME Readout Boards (VRB), each of which is linked to a group of front-end crates. The VRB are grouped in 15 EVB crates, each of which is controlled by a single board processing unit, the Scanner CPU (SCPU). An Asynchronous Transfer Mode (ATM) network switch provides the connection between the EVB crates and the *converter* nodes (CV) of the Level-3 farm[3]. Converter nodes transfer event fragments from the EVB crates to the processing units of the Level-3 farm, as it will be explained in detail in the next paragraph. Data flow between SCPU's and the Level-3 farm is controlled by the Scanner Manager (SM), a process running in an additional EVB crate which constitutes the interface between the EVB system and the Trigger System Interface (TSI). The TSI is responsible for receiving the trigger decisions from Level-1 and Level-2, and supervising data flow until the EVB. When the TSI passes a Level-2 accept message to the SM, the SM instructs the SCPU's to read and combine the event fragments in their local crate, selects a converter node in the Level-3 farm among those which reported themselves available, and then directs the SCPU's to send the event fragments to the selected converter node.

At the time that the data utilized in the mixing analysis had been collected, the 292 nodes of the Level-3 farm were divided in sixteen subfarms working in parallel[4]. A scheme of the Level-3 farm is shown in Figure 2.21. Each of the sixteen subfarms contains a converter node which is in direct contact with the EVB, as mentioned before. The converter node is placed at the head of a set of *processor* nodes (PR). Subfarms contain between 16 and 18 processor nodes. The converter has the role of assembling the event fragments from the VRB's in a unique *event record*, which constitutes the single and only piece of information about a particular event. It then selects the first available processor node in its subfarm, and sends the event record to it. The event reconstruction and the formation of trigger decisions are performed by processor nodes. At the stage of Level-3, the event reconstruction benefits from full detector information and improved resolution with respect to the preceding trigger

[3]The ATM network has been substituted by a Gigabit ethernet network in August 2005.

[4]The configuration of the Level-3 farm, as of June 2007, includes 384 nodes subdivided in 18 subfarms with 21 to 22 nodes each.

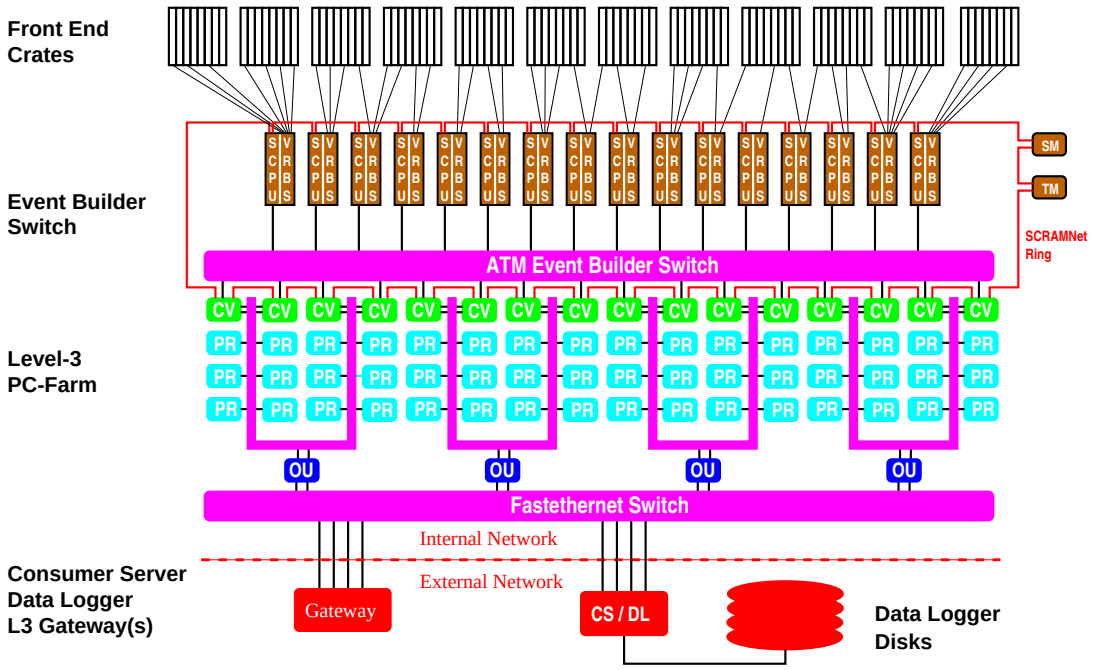

FIGURE 2.21: Scheme of the Level-3 farm. Data flow from top to bottom: from the front-end crates to the Event Builder crates, where data fragments are assembled, then to the Level-3 farm, through converter nodes (CV). Processor nodes (PR) produce the Level-3 trigger decision. Events which satisfy the Level-3 trigger requirements are sent to the data-storage system via output nodes (OU).

levels. In particular, three-dimensional reconstruction of tracks with code derived from offline and more precise matches between tracks and calorimeter or muon data are available. The events which pass the Level-3 trigger are sent by the processor nodes to *output* nodes (OU). The output nodes serve two subfarms each and host the software needed for the transmission of the reconstructed events from the processor nodes to the data-storage system.

This chapter presented the accelerator complex at Fermilab and the CDF II detector. In the next chapter the strategy for the online selection of B_s^0 *candidates is introduced. The reconstruction and selection of the* B_s^0 *candidates utilized for the* B_s^0 *oscillations search and the angular analysis for the mearuments of* $\Delta\Gamma_s$ *and* β_s *are presented.*

Chapter 3

B_s^0 Analyses Samples

This chapter describes the samples of data used to perform the analyses reported in this dissertation. The first part of the chapter is devoted to the online selection of the events. In the second part the offline reconstruction criteria are presented.

3.1 Triggers

At first order, the result of any measurement is based on the amount and quality of the data collected. The role of a trigger system at CDF is to identify interesting events within the large rate of $p\bar{p}$ collisions. The data used for the analyses described in this dissertation was recorded by the CDF detector from a period starting from February 2002 through October 2006. Different analyses require different selection criteria. The B_s^0 mixing analysis utilizes events gathered with the *two-track* trigger and *lepton-plus-displaced-track* trigger. The sample used for the B_s^0 mixing analysis corresponds to an integrated luminosity of 1 fb^{-1} (till January 2006). This amount of data was high enough to pass the 5 standard deviation to observation for the B_s^0 oscillation frequency. The angular analysis aiming at the measurements of the width difference, $\Delta\Gamma_s$ and the CP violation phase β_s used an integrated luminosity of 1.35 fb^{-1} (till October 2006). All these events were collected with the `di-muon` trigger.

3.1.1 The Two Track Trigger

With Two Track Trigger (`TTT`) it is intended a variation of paths which aims at identifying heavy flavor decays based upon kinematics, topology and decay time information. B_s^0 mesons have a relatively long lifetime due to the weakly decaying heavy-flavor hadrons. In fact they fly on average 0.5 mm before decaying, which is a larger distance than the intrinsic beam size. B_s^0 events are thus characterized by an impact parameter inconsistent with zero, namely displaced tracks and vertices. In Figure 3.1 the decay of a long-lived B hadron is depicted.

The `TTT` paths exploit the precise information on the track hits from the silicon vertex detector and the SVT electronics to look for displaced tracks at online level. Being extremely powerful in rejecting light flavor (u, d, s) backgrounds, they trigger

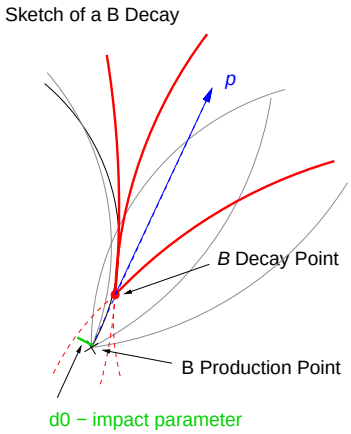

FIGURE 3.1: The decay of a long-lived B hadron. The impact parameter of the track measured by SVT system is the distance of closest approach to the primary vertex.

on long-lived particles, collecting data samples enriched in events containing bottom and charm hadrons.

Over the course of a Tevatron store, the available trigger bandwidth varies as the luminosity falls. In order to fully utilize the bandwidth, the trigger criteria have to vary accordingly to the luminosity variations. Higher trigger rates at high luminosity arise from both an increase in the real physics rate as well as an increase in fake triggers due to multiple $p\bar{p}$ interactions. As the luminosity decreases, the trigger bandwidth becomes under-utilized and lower purity triggers are thus used to fulfill it and maximize the B_s^0 yield written to tape. As reported in section 2.2.8 a prescaling system is used to prioritize triggers. The `TTT` is composed by a series of distinct trigger paths or "scenarios" applied simultaneously to efficiently use the bandwidth available. They are all characterized by the presence of at least a pair of displaced tracks to reconstruct a vertex. The three primary hadronic trigger paths are summarized in Table 3.1.

The Level-1 selection requires a pair of XFT tracks with a lower cut on the p_T of the tracks, on the scalar sum of the p_T of the tracks $\sum p_T$, and an upper cut on $\Delta\varphi_6$, the opening angle between the two tracks, as measured at COT superlayer 6. At Level-2 the tracks are matched with the hits by the SVT and the must satisfy the criteria from Level-1 in addition to cuts on the impact parameter d_0 and the transverse length L_{xy}, the distance from the primary $p\bar{p}$ collision point to the intersection of the two tracks measured in the plane perpendicular to the beamline. These last two cuts are very efficient at rejecting the light quarks background, but a selection based on the decay length information introduce a bias given that events with small values of the proper decay time are rejected. While performing any analysis using the `TTT`, the intrisic bias in it has to be properly incorporated. The Level-3 selection requires the same cuts using the fit based on the full detector information COT+SVX. At this stage three dimensional fits are available and it is possible to apply a cut on the distance between the closest approach to the z axis of the two tracks, Δz_0. The three different scenarios are defined by the different values of the cuts applied. The `HIGHPT` path

Scenario	Level-1	Level-2	Level-3
	XFT tracks	SVT tracks	COT+SVX tracks
HIGHPT	$p_T > 2.5\,\text{GeV/c}$ opposite charge $\Delta\varphi_6 < 135°$ $\sum p_T > 6.5\,\text{GeV/c}$	$p_T > 2.5\,\text{GeV/c}$ opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 6.5\,\text{GeV/c}$ $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$	$p_T > 2.5\,\text{GeV/c}$ opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 6.5\,\text{GeV/c}$ $80\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$ $\lvert\Delta z_0\rvert < 5\,cm$
BCHARM	$p_T > 2.0\,\text{GeV/c}$ opposite charge $\Delta\varphi_6 < 135°$ $\sum p_T > 5.5\,\text{GeV/c}$	$p_T > 2.0\,\text{GeV/c}$ opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 5.5\,\text{GeV/c}$ $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$	$p_T > 2.0\,\text{GeV/c}$ opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 5.5\,\text{GeV/c}$ $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$ $\lvert\Delta z_0\rvert < 5\,cm$
LOWPT	$p_T > 2.0\,\text{GeV/c}$ $\Delta\varphi_6 < 90°$	$p_T > 2.0\,\text{GeV/c}$ $\Delta\varphi_0 < 90°$ $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$	$p_T > 2.0\,\text{GeV/c}$ $2° < \Delta\varphi_0 < 90°$ $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$ $L_{xy} > 200\,\mu\text{m}$ $\lvert\Delta z_0\rvert < 5\,cm$
μ+LOWPT	CMU or CMP muon $p_T(\mu_{cMU}) > 1.5\,\text{GeV/c}$ or $p_T(\mu_{cMP}) > 2.0\,\text{GeV/c}$	$\Delta\varphi_6(\mu, trk) > 90°$	

TABLE 3.1: Trigger selections for the hadronic TTT paths.

provide a sample with higher purity (defined as signal-to-background ratio) and lower trigger rate compared to the BCHARM and LOWPT. At moderate istantaneous luminosity, bandwidth become available then a dynamic prescale is enabled to BCHARM. Finally LOWPT is active at low luminosity to maximize the yield of events written to tape. The trigger list in Table 3.1 is supplemented with an additional path that exploits the same cuts as LOWPT, but requiring a CMU (CMX) muon with $p_T > 1.5(2)\,\text{GeV/c}$. The "two-track plus muon" trigger was not originally implemented with the cut on the opening angle between the muon and the two tracks, in order to be efficient in recording semimuonic decay and in particular $\text{B}\overline{\text{B}}$ events, where one meson decays hadronically and the other one semimuonically. In such picture the muon carries the crucial information about the initial flavor tagging of the hadronically decayed B, so the attempt is to collect events which are more powerful in the analysis. With the increasing luminosity a cut on the angle $\Delta\varphi_6$ became necessary. Semileptonic B_s^0

decays, for example $B_s^0 \to D_s^- \ell^+ X$, are also selected by the TTT. In this case the lepton is recontructed in the offline selection because is not used in the trigger signature.

3.1.2 The Lepton-plus-displaced-track Trigger

The B_s^0 mixing analysis results to be the combination of two oscillation analyses on hadronic and semileptonic decay modes, respectively. To supplement the hadronic sample acquired with the hadronic trigger paths, there are additional semileptonic paths belonging to the "lepton-plus-displaced-track" trigger, $\ell + \text{SVT}$. In this trigger paths a lepton, either an electron or a muon, is required to have $p_T > 4\,\text{GeV/c}$ together with an SVT track with $p_T > 4\,\text{GeV/c}$ and $120\,\mu\text{m} < d_0 < 1000\,\mu\text{m}$. In the electron case the track is matched to the CEM electromagnetic cluster, while in the muon case it is matched with the segments in the CMU and CMP systems. The angle between the two tracks is required to be $2^\circ < \Delta\varphi_0 < 90^\circ$. The absence of a decay lenght cut makes the acceptance of the semileptonic B decays high. The function of the $\ell + \text{SVT}$ trigger is, in fact, two fold: not only it allows to collect high sample of B_s^0 semileptonic decays, but it is also utilized to record a large sample of inclusive B decays useful for calibrating the opposite side flavor algorithms. Additional details are given in publications devoted to the semileptonic analysis [45]

3.1.3 Di-muon Trigger

The data samples for the angular analysis on $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^{*0}$ decays are collected by the `di-muon` trigger paths. At Level-1, the `di-muon` trigger consists of a combination of the CMU-CMU (both muons from the CMU) and the CMU-CMX (one muon from the CMU and the other from the CMX) triggers. The criteria an event has to fulfill to be accepted by the `di-muon` trigger concern the following variables:

- The XFT reports a measure of the transverse momentum p_T and the angle between the two muons at the COT superlayer 6, $\Delta\varphi_6$. These information are obtained from the XTRP, which extrapolate the tracks from the XFT to the inner radius of the muon chambers. The extrapolation uncertainty comes from multiple scattering, thus XTRP determines, for each track, a "footprint", a ϕ window in which the track could end with 99.5% probability (3σ).

- If at least one footprint matches fired muon tower, then it is promoted into a muon tower. The tower is a logical unit composed of a couple of close stacks. The stack is an aggregate of four cells stacked on top of each other (Figure 3.2). The tower is said to have fired if at least cells 1 and 3 or 2 and 4 have hits separated in time by no more than the stub gate width of 396 ns. Otherwise is said to be empty.

- The two muon towers are required to be either separated by at least two other towers or on different sides of the detector.

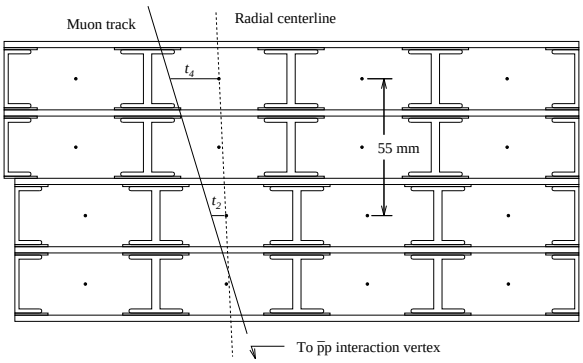

FIGURE 3.2: A single CMU chamber.

Because main inputs onm the online seection come from XFT traking and muon system which are already available at Level-1, the Level-2 trigger does not play an important role for the `di-muon` trigger. During the CDF operation period the `di-muon` trigger has been constantly revised and improved. The angular analysis reported in this disseration has been performed on the trigger path reported in Table 3.2. Each name reports several criteria which are encoded in the trigger name:

- **PTx:** p_T of the XFT track greater than x
- **DHIx:** The difference in ϕ of the two muons smaller than x
- **OPPQ:** Both muons need to have opposite charge
- **xMTYy:** The transverse mass for the two muons between x and y
- **PSx:** The trigger is prescaled by a factor x
- **DPS:** The trigger is dinamically prescaled

3.2 Data Format and Analysis Software

The most used standard for the data utilized to perfom B physics analyses at CDF is the BStntuple format [46]. Under the name BStntuple it is intended a framework developed to efficiently store the information of the B candidates. It is an extension of the Stntuple [46] providing a series of dedicate tools to access the kinematic and spacial variables for stable and decaying candidates, as well as particle identification and tagging information in a convenient data blocks structure. This framework has resulted to be very efficient in terms of CPU usage and extremely versatile, making possible candidate reconstruction for the various decays and data samples used in the analyses described in this dissertation.

Level-2	Level-3
L2_AUTO_L1_CMU1.5_PT1.5_&_CMX1.5_PT2	JPSI_CMU1.5_CMX20
L2_AUTO_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMU1.5_CMX22
L2_AUTO_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX_PS0	JPSI_CMU1.5_CMX23
L2_AUTO_L1_TWO CMU1.5_PT1.5	JPSI_CMU1.5_CMX2 ALLPHI
L2_PS100_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMU1.5_CMX2 NOL2
L2_PS100_L1_TWO_CMU1.5_PT1.5	JPSI_CMU1.5_CMX2 NOL2
L2_PS10_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMU1.5_CMX2
L2_PS10_L1_TWO_CMU1.5_PT1.5	JPSI_CMU2 CMX2_PS10
L2_TWO_CMU1.5_PT1.5_DPHI120_OPPQ	JPSI_CMU2 CMX2_PS2
L2_TWO_CMU1.5_PT1.5_DPHI120_OPPQ_DPS	JPSI_CMU2 CMX2_PS50
L2_TWO_CMU1.5_PT2_DPHI120_OPPQ	JPSI_CMUCMU1.5
L2_TWO_CMU1.5_PT2_DPHI120_OPPQ_PS10	JPSI_CMUCMU1.5_ALLPHI
L2_AUTO_L1_CMUP6_PT4	JPSI_CMUCMU1.5_NOL2
L2_CMU1.5_PT1.5_&_CMX1.5_PT2_DPHI120_OPPQ	JPSI_CMUCMU2
L2_CMU1.5_PT1.5_&_CMX1.5_PT2_DPHI120_OPPQ_DPS	JPSI_CMUCMU2_PS10
L2_CMU1.5_PT1.7_&_CMU1.5_PT3_1.7MT7_DPS	JPSI_CMUCMU2_PS2
L2_CMU1.5_PT1.7_&_CMU1.5_PT3_1.7MT7_LUMI_185	JPSI_CMUCMU2_PS50
L2_CMU1.5_PT1.7_&_CMX1.5_PT3_1.7MT7_DPS	JPSI_CMUCMU_MT_DPS
L2_CMU1.5_PT1.7_&_CMX1.5_PT3_1.7MT7_LUMI_185	JPSI_CMUCMU_MT_LUMI_185
L2_CMU1.5_PT2_&_CMX1.5_PT2_DPHI120_OPPQ_PS10	JPSI_CMUCMX3_MT_DPS
L2_CMU1.5_PT2_&_CMX1.5_PT2_DPHI120_OPPQ_PS2	JPSI_CMUCMX3_MT_LUMI_185
L2_CMU1.5_PT2_&_CMX1.5_PT2_DPHI120_OPPQ_PS50	JPSI_CMUP4_CMU1.5
L2_CMUP6_PT8	JPSI_CMUP4_CMU1.5_L2_DPS
L2_CMUP6_PT8 DPS	JPSI_CMUP4_CMU_L2_DPS
L2_CMX1.5_PT2_&_CMU1.5_PT3_1.7MT7_DPS	JPSI_CMUP4_CMX2
L2_CMX1.5_PT2_&_CMU1.5_PT3_1.7MT7_LUMI_185	JPSI_CMUP4_CMX2_L2_DPS
L2_DPS_L1_CMUP6_PT4	JPSI_CMUP4_CMX_L2_DPS
L2_PS200_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMUPCMU_HIGHPT
L2_PS200_L1_TWO_CMU1.5_PT1.5	JPSI_CMUPCMU_HIGHPT_DPS
L2_PS500_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMUPCMX_HIGHPT
L2_RL20HZ_L1_CMUP6_PT4	JPSI_CMUPCMX_HIGHPT_DPS
L2_TRK8_L1_CMUP6_PT4	JPSI_CMXCMU3_MT_DPS
L2_TWO_CMU1.5_PT2_DPHI120_OPPQ_PS2	JPSI_CMXCMU3_MT_LUMI_185
L2_TWO_CMU1.5_PT2_DPHI120_OPPQ_PS50	EXPRESS_JPSI_CMUCMU

TABLE 3.2: Summary of `di-muon` triggers used in the angular analysis.

3.3 Data Samples for the B_s^0 Analysis Oscillations

The data samples utilized for the B_s^0 oscillations analysis were recorded by the CDF detector in the period from March 2002 to January 2006. The corresponding integrated luminosity is about 1 fb^{-1}, with the imposed requirement of the full detector system working. The samples of B_s^0 decays are divided according to their decays modes in two main groups, the semileptonic and the hadronic one. The B_s^0 meson candidates are reconstructed in reverse order of their chain, starting from the event tracks to fit for B daughter resonances which are progressively combined in the expected topology of the B_s^0 signal. For instance in the hadronic mode $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$, every pair of oppositely-charged track hypothesized to be kaons is fit in three dimensions with requirements on the kinematics and quality of the vertex [47]. The resulting ϕ candidate vertices are then combined with an additional track, applying another set of selection criteria to form a high-level candidate. This new candidates are assumed to be D_s^- mesons. Finally they are matched with the remaining available tracks to form B_s^0 candidate vertices. At each step a new set of selections is applied. The following sections describe the offline criteria for the candidate reconstruction, the sample composition and the event yields.

$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mathcal{P}(B_s^0) > 10^{-7}$	$\mathcal{P}(B_s^0) > 10^{-5}$	$\mathcal{P}(B_s^0) > 10^{-5}$
$\chi^2_{r\phi}(D_s^-) < 20$	$\chi^2_{r\phi}(D_s^-) < 20$	$\chi^2_{r\phi}(D_s^-) < 20$
$L_{xy}/\sigma_{L_{xy}}(D_s^-) > 5$	$L_{xy}/\sigma_{L_{xy}}(D_s^-) > 8$	$L_{xy}/\sigma_{L_{xy}}(D_s^-) > 11$
—	$L_{xy}/\sigma_{L_{xy}}(B_s^0) > 2$	$L_{xy}/\sigma_{L_{xy}}(B_s^0) > 2$
$ct^*(B_s^0) > 0.01$ cm	$ct^*(B_s^0) > 0.01$ cm	$ct^*(B_s^0) > 0.01$ cm
$\sigma_{ct^*}(B_s^0) < 0.04$ cm	$\sigma_{ct^*}(B_s^0) < 0.04$ cm	$\sigma_{ct^*}(B_s^0) < 0.04$ cm
$ct(D_s^-) \in [-0.01, 0.10]$ cm	$ct(D_s^-) \in [-0.01, 0.10]$ cm	$ct(D_s^-) \in [-0.01, 0.10]$ cm
$p_T(trk) > 0.4$ GeV/c	$p_T(trk) > 0.4$ GeV/c	$p_T(trk) > 0.7$ GeV/c
$\lvert\cos(\psi_H)\rvert > 0.3$	$\lvert\cos(\psi_H)\rvert > 0.3$	—
$m(\ell D_s^-) \in [2.0, 5.5]$ GeV/c^2	$m(\ell D_s^-) \in [2.0, 5.5]$ GeV/c^2	$m(\ell D_s^-) \in [2.0, 5.5]$ GeV/c^2
$CLL(K^-) > -2.5$	$CLL(K^-) > -2.25$	—
$CLL(K^+) > -2.5$	$CLL(K^+) > -1.1$	—

TABLE 3.3: Selection criteria for $B_s^0 \to D_s^-\ell^+X$ decays in the analysis.

3.3.1 Semileptonic B_s^0 Decays

The three inclusive B_s^0 meson semileptonic final states reconstructed in the analysis are: $B_s^0 \to D_s^-\ell^+X$, where the D_s^- meson can decay in $D_s^- \to \phi\pi^-$ ($\phi \to K^+K^-$), $DsKSK$ ($K^{*0} \to K^+\pi^-$) and $D_s^- \to \pi^-\pi^+\pi^-$. To identify semileptonic B_s^0 decays both the TTT and the ℓ+SVT triggers are utilized. A significant overlap is present between the two triggers which amount to about 60%, hence the lepton-plus-displaced-track trigger gives only about 10% of additional semileptonic B_s^0 decays. These decays have relative high brancing ratio equal to $7.9 \pm 2.4\%$, according to the Review of Particle Physics [6]. Moreover the presence of a lepton provides a clear signature. On the other hand, the main challenge is represented by the incomplete reconstruction of at least the neutrino involved in the process. The final sample is obtained with a sequential cut-based selection. Each cut value has been optimized to maximize $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$. The selection criteria are then summarized in Table 3.3.

The quality of the vertex fits is ensured by applying a lower cut on the fit probability of the B_s^0 fit itself, $\mathcal{P}(B_s^0)$ and an upper cut on the two-dimensional $\chi^2_{r\phi}$ of the (D_s^-) vertex fit, $\chi^2_{r\phi}(D_s^-)$. Then a series of kinematical cuts are applied. A lower cut on the transverse lenght significance $L_{xy}/\sigma_{L_{xy}}$ of the D_s^- and B_s^0, based on the knowledge of the large B_s^0 mesons lifetime, helps to remove most of the background coming from prompt decays. The poorly reconstructed decays are eliminated applying cuts on the minimum p_T of the tracks which compose the B_s^0 candidate, on the proper decay length $ct^*(B_s^0)$, $ct(D_s^-)$ and on the proper decay length resolution $\sigma_{ct^*}(B_s^0)$. The * is indicating that only the ℓD_s^- part of the candidate enter the calculation. The cut on the minimum value of $\lvert\cos(\psi_H)\rvert$, the cosine of the helicity angle of the D_s^- in vector-scalar decays, such as $\phi\pi^-$ or $K^{*0}K^-$, rejects more flat-distributed background than signal, which peaks at large absolute values of $\cos\psi_H$. In the D_s^- rest frame, ψ_H is defined as the angle between the B_s^0 transverse momenta and of the π^- (K^-) candidates for the $D_s^- \to \phi\pi^-$ ($D_s^- \to K^{*0}K^-$) decays. The graphical definition is shown in Figure 3.3.

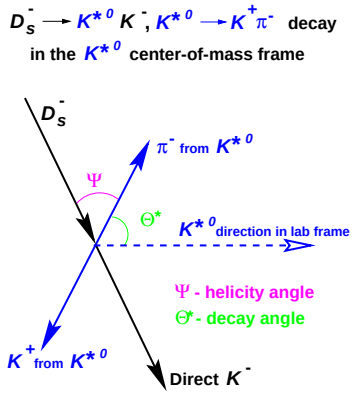

FIGURE 3.3: Definition of the helicity angle ψ_H (indicated as ψ in the Figure) and the decay angle θ^* (Θ^* in the Figure), in the $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$ decay chain. The momenta are drawn in the K^{*0} rest frame.

To identify leptons a multivariate approach has been developed to combine the various lepton identification variables into a global likelihood for higher efficiency and lower mis-identification [45] [48]. Electron and muons are thus separated by hadrons applying a lower cut on the relative likelihood defined as:

$$\mathcal{L} = \frac{\mathcal{P_S}^i}{\mathcal{P_S}^i + \mathcal{P_B}^i} \, i = e, \mu \tag{3.3.1}$$

where $\mathcal{P_S}^i$ describes the probability for a lepton to be indeed a real one and $\mathcal{P_B}^i$ the probability to be a fake one. In order to differentiate an electron from an hadron mimicking its signature, a sample of electrons coming from the conversion $\gamma \to e^+e^-$ is selected to describe the signal likelihood, $\mathcal{P_S}^e$. The same cuts used for ℓ + SVT trigger paths are applied with the requirement of two oppositely charged tracks with a small opening angle. The background likelihood is instead based on a sample of pions from the $K_s^0 \to \pi^+\pi^-$ decay mode. The signal muon likelihood, $\mathcal{P_S}^\mu$, is built on the information from muons reconstructed in the $J/\psi \to \mu^+\mu^-$. For the corresponding background likelihood not only pions from $K_s^0 \to \pi^+\pi^-$, but also kaons and protons from $D^0 \to K^-\pi^+$ and $\Lambda^0 \to p\pi^-$ are utilized. The plots in Figure 3.4 describe the lepton likelihood distributions in the signal and background samples described above.

The values for these likelihood are bound between zero and one with real electrons (muons) occupying the high likelihood region close to unity while the background populates the low region close to zero. A requirement on the electron likelihood to be higher than 0.9 is 90% efficient for conversion electrons with $p_T > 2.0\ GeV/c$, and it is able to reject almost 98% of the pions coming from K_s^0 decays. Instead the selection on the muon likelihood varies depending on the subdetector considered. A summary of cut values and efficiencies for the muon likelihood is reported in Table 3.4.

Finally the studies performed on the same-side taggers provided a likelihood ratio to discriminate kaons from pions and protons. This work will be described in Chapter 5. The purity of the sample is thus enhanced by requiring the tracks which are assigned to be kaon to pass a higher cut on the combined likelihood ratio CLL.

The incomplete reconstruction of the B_s^0 decays makes not possible to recontruct the B_s^0 mass. Therefore, the D_s^- mass and D_s^- *plus lepton* mass distributions are used

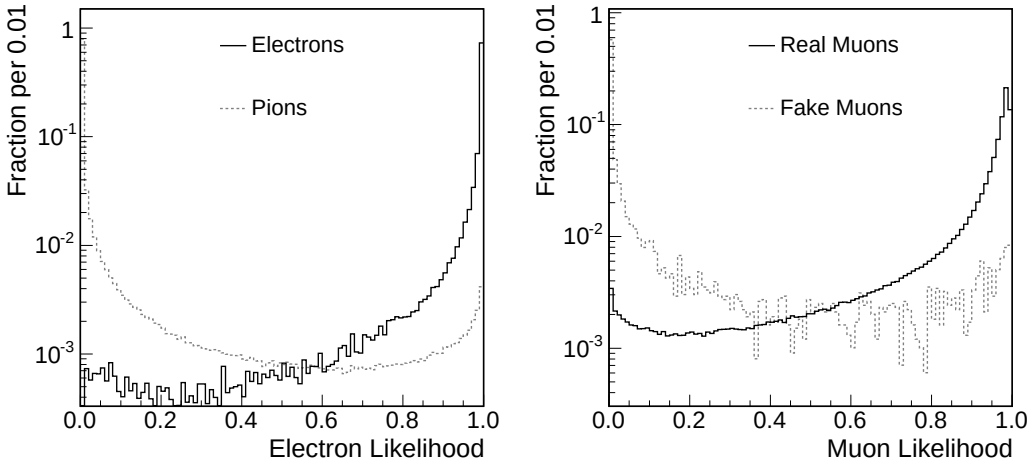

FIGURE 3.4: Left: Electron likelihood distributions $\mathcal{L}_e$ for electrons and pions with $p_T > 2.0$ GeV/c. Right: Muon likelihood distributions $\mathcal{L}_\mu$ for real and fake muons.

Muon System	$\mathcal{L}_\mu >$	Efficiency for real muons [%]	Efficiency for fake muons [%]
CMU	0.5	92.0	13.5
CMP	0.5	88.2	27.1
CMUP	0.05	98.8	55.0
CMX	0.5	91.8	22.2
IMU	0.7	78.8	9.6

TABLE 3.4: Efficiency of muon likelihood requirements for real and fake muons matching the muon candidate requirements, compiled for different muon detector systems.

Decay Sequence	S
$B_s^0 \to D_s^-(\phi\pi^-)\ell^- X$	29,600 ± 800
$B_s^0 \to D_s^-(K^{*0}K^-)\ell^- X$	22,000 ± 800
$B_s^0 \to D_s^-(\pi^-\pi^+\pi^-)\ell^- X$	9,900 ± 700
Total	61,500 ± 1,300

TABLE 3.5: Signal yields for the semileptonic modes (S) in the various decay sequences.

in the likelihood. The mass distributions are shown in Figure 3.5.

In addition to the B_s^0 signal candidates, two are the contributions which have real D_s^- meson in the final state: physics backgrounds and false leptons. With physics background it is intended real physics processes with the correct charge correlation between the lepton and the D_s^- meson. Sample of simulated events have been produced to compute their fraction in the signal region. The false lepton background mainly consists of a real D_s^- meson plus a wrongly identified lepton, so called false lepton. The evaluation of the shape of this contribution is obtained with a sample

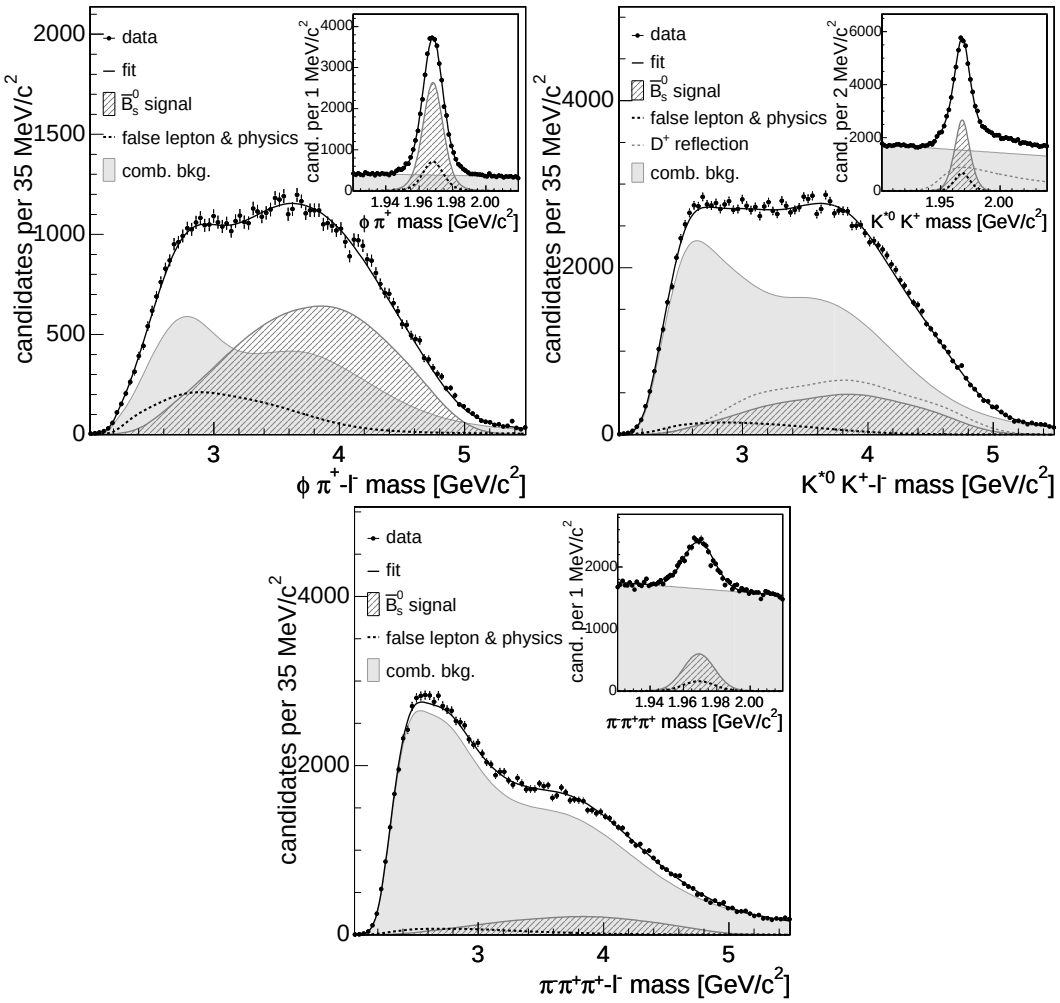

FIGURE 3.5: $D_s^- + \ell^-$ and D_s^- mass distributions for $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to \phi\pi^-$ (top left), $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to K^{*0}K^-$ (top right) and $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to \pi^-\pi^+\pi^-$ (bottom) decays. All candidates shown in the D_s^- mass distributions are also included in the $D_s^- + \ell^-$ mass distributions.

recorded applying the standard cut along with an anti-lepton selection. Finally in Table 3.5 the yields of B_s^0 candidates in the semileptonic decay modes $B_s^0 \to D_s^- \ell^+ X$ with $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ are reported.

3.3.2 Hadronic B_s^0 Decays

The fully hadronic sample considered in the B_s^0 mixing analysis consists of six topologies of $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^- \to \pi^-\pi^+\pi^-$. The charm meson D_s^- is reconstructed in one of the following final states: $D_s^- \to \phi\pi^-(\phi \to K^+K^-)$, $D_s^- \to K^{*0}K^-(K^{*0} \to K^+\pi^-)$ or $D_s^- \to \pi^-\pi^+\pi^-$. The branching ratio of the hadronic B_s^0 decay modes is roughly an order of magnitude smaller than the semileptonic ones with the same D_s^- final states. On the other hand the candidates are fully reconstructed, making of it the most powerful sample used for the B_s^0 oscillation analysis described in this

Mode	mass sideband	
	lower, GeV/c^2	upper, GeV/c^2
$B_s^0 \to D_s^-[\phi\pi^-]\pi^+$	–	5.50 - 5.80
$B_s^0 \to D_s^-[K^{*0}K^-]\pi^+$	–	5.50 - 5.80
$B_s^0 \to D_s^-[\pi^-\pi^-\pi^+]\pi^+$	–	5.50 - 5.60
$B_s^0 \to D_s^-[\phi\pi^-](3\pi)^+$	–	5.50 - 5.65
$B_s^0 \to D_s^-[K^{*0}K^-](3\pi)^+$	–	5.55 - 5.75
$B_s^0 \to D_s^-[\pi^-\pi^-\pi^+](3\pi)^+$	4.78 - 4.80	5.47 - 5.50

TABLE 3.6: Mass sidebands used to prepare training sample of background events.

dissertation. The sample is collected with the TTT: in order to have a kinematically well understood sample, it is explicitily required the both triggered tracks are present within the B_s^0 decay chain. The selection is performed in two steps. First, candidates are pre-selected applying loose sequential requirements, or rectangular cuts. Then, an Artificial Neural Network (ANN) performs the final selection. The use of an ANN for the semileptonic decays has been studied as well but the improvements with respect to the cut-based selection was marginal, so it has been decided not to include it in the semileptonic selection. An ANN can provide several advantages over a cut-based selection. It can successfully exploit the correlation among the variables discriminating signal from background. The use of a Neural Network allows one to aggregate the signal/background separation of the individual variables into a unique discriminator which consists of a floating-point number, the so called NN output. The ANN returns values close to 0 for background-like events and values close to 1 for signal-like events. Thus, while a candidate which fails a single rectangular cut would be discarded on a cut-based selection, in a NN scenario the same candidate would be likely assigned a value of the output closer to the signal like region and be above the threshold for accepting the event. The Stuttgart Neural Network Simulator(SNNS) [49] interfaced with ROOT package [50] has been used to perform the B_s^0 hadronic decay modes selection varying the cut on the ANN output and while monitoring the quantity $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$ for a global maximum. $\mathcal{S}$ is estimated on simulated events and represents the amount of signal in the region from 5.31 GeV/c^2 to 5.42 GeV/c^2, while $\mathcal{B}$ is the background amount estimated in the same region by extrapolating the fit of the upper mass sideband. The choice of the sideband is a compromise between avoiding signal-like B^0 and $\overline{\Lambda}_b^0$ reflections and desire to choose the sideband as close to the fully reconstructed signal as possible (such that the sideband describes the background under the signal peak most closely). The sideband regions are documented in Table 3.6. Detailed description of the principles and the framework are given in Reference [51].

The pre-selection cuts are listed in Tables 3.7 and 3.8. An additional clean-up before applying ANN selection is needed only for modes containing $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ decays where reflections from B^0 decays are present. To remove these reflections, a combination of particle identification and mass cuts is applied. In the $D_s^- \to K^{*0}K^-$ case, the contamination comes from $D^- \to K^{*0}\pi^-$ decays. Requiring

Variable	$B_s^0 \to D_s^- \pi^+$ modes		
	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^+$
$\chi^2_{r\phi}(B_s^0)$	< 20	< 20	< 20
$\chi^2_{r\phi}(D_s^-)$	< 20	< 20	< 20
$\chi^2_{r\phi}(\pi^-\pi^+\pi^+)$	N/A	N/A	N/A
$L_{xy}/\sigma_{L_{xy}}(B_s^0)$	> 2.0	> 2.0	> 2.0
$L_{xy}/\sigma_{L_{xy}}(D_s^-)$	> -99	> -99	> -99
$\lvert d_0(B_s^0)\rvert$, μm	< 200	< 200	< 200
$p_T(B_s^0)$, GeV/c	> 5.50	> 5.50	> 5.50

TABLE 3.7: Pre-selection cuts for the $B_s^0 \to D_s^- \pi^+$ hadronic decay modes.

Variable	$B_s^0 \to D_s^- \pi^-\pi^+\pi^+$ modes		
	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^-\pi^+$
$\chi^2_{r\phi}(B_s^0)$	< 20	< 20	< 20
$\chi^2_{r\phi}(D_s^-)$	< 20	< 20	< 20
$\chi^2_{r\phi}(\pi^-\pi^+\pi^+)$	< 50	< 50	< 50
$L_{xy}/\sigma_{L_{xy}}(B_s^0)$	> 2.0	> 6.0	> 6.0
$L_{xy}/\sigma_{L_{xy}}(D_s^-)$	> 2.0	> 6.0	> 6.0
$\lvert d_0(B_s^0)\rvert$, μm	< 200	< 200	< 200
$p_T(B_s^0)$, GeV/c	> 5.50	> 4.00	> 4.00

TABLE 3.8: Pre-selection cuts for the $B_s^0 \to D_s^- \pi^+\pi^-\pi^+$ hadronic decay modes.

Mode	Selection Requirement
$B_s^0 \to D_s^-[\phi\pi^-]\pi^+$	–
$B_s^0 \to D_s^-[K^{*0}K^-]\pi^+$	$\lvert m_{K\pi\pi} - 1869.4\,\mathrm{MeV/c^2}\rvert > 16\,\mathrm{MeV/c^2}\,\lvert\rvert\, CLL(K^-) > 0.4$
$B_s^0 \to D_s^-[\pi^-\pi^-\pi^+]\pi^+$	$\lvert m_{K\pi\pi} - m_{K\pi}\rvert > 160\,\mathrm{MeV/c^2}$
$B_s^0 \to D_s^-[\phi\pi^-]\pi^-\pi^+\pi^+$	–
$B_s^0 \to D_s^-[K^{*0}K^-]\pi^-\pi^+\pi^+$	$\lvert m_{K\pi\pi} - 1869.4\,\mathrm{MeV/c^2}\rvert > 19\,\mathrm{MeV/c^2}\,\lvert\rvert\, CLL(K^-) > 0.1$
$B_s^0 \to D_s^-[\pi^-\pi^-\pi^+]\pi^-\pi^+\pi^+$	$\lvert m_{K\pi\pi} - m_{K\pi}\rvert > 160\,\mathrm{MeV/c^2}$

TABLE 3.9: Additional pre-selection requirements, applied to reduce physics backgrounds from B^0 reflections.

that the supposed kaon from the D_s^- decay has a combined particle identification likelihood (CLL) consistent with a kaon hypothesis removes a fraction of these decays. Another fraction is removed by requiring that the $K^{*0}\pi$ invariant mass be inconsistent with the D^- invariant mass. For $D_s^- \to \pi^+\pi^-\pi^-$, the reflection comes from $D^{*-} \to D^0\pi^-, D^0 \to K^-\pi^+$ decays. The narrow phase space of the $D^{*-} \to D^0\pi^-$ decay is utilized to remove D^{*-} decays from the sample. The complete list of reflection-removing selection requirements is summarized in Table 3.9.

The different topology between the decay modes imply different sets of input variables to be used in the training process. The input variables used for the six hadronic modes are listed in Table 3.10 . The following information is used about the B_s^0 and

D_s^- decay vertex fits: the contribution from the transverse plane to the chi squared of the decay fit, $\chi^2_{r\phi}$, the impact parameter of the decay, $d_0(B_s^0)$, the transverse displacement of the vertex, L_{xy} along the direction of the particle transverse momentum, its significance, $L_{xy}/\sigma_{L_{xy}}$, the transverse displacement of the D_s^- decay with respect to the B_s^0 decay, $L_{xy}(D_s^- \to B_s^0)$, and the transverse momentum of the decay, p_T. All individual track transverse momenta, p_T(tracks), are used as input variables in most of the decay modes. In the $B_s^0 \to D_s^-\pi^-\pi^+\pi^+$, $D_s^- \to K^{*0}K^-$ decay, only the transverse momenta of the pion candidates from the B_s^0 decay and the kaon candidate from the D_s^- decay are used. For the D_s^- candidates the mass $M(D_s^-)$ and the helicity angle of the D_s^- decay, $\cos\psi_H$ (cosine of helicity angle as defined in Figure 3.3) are utilized as input variables. When resonances, such as the ϕ or $K^{*0}(892)$ are present, the mass, $M(\phi$ or $K^{*0})$, and transverse momentum, $p_T(\phi$ or $K^{*0})$ of the resonance decays are also used in the selection. Some amount of low-level track information is directly used. For instance, the minimum and maximum impact parameter significance ($\min(d_0/\sigma_{d_0})$ and $\max(d_0/\sigma_{d_0})$, respectively) of the candidate tracks are used as well as the minimum and maximum transverse momentum, ($\min(p_T)$ and $\max(p_T)$, respectively). For the $B_s^0 \to D_s^-\pi^-\pi^+\pi^+$ decays, more quantities are recorded about the three pions from the B_s^0 decay. The mass of the three pion decay, $m(\pi^-\pi^+\pi^+)$, the contribution from the transverse plane to the chi squared of the three pion vertex, $\chi^2_{r\phi}(\pi^-\pi^+\pi^+)$, and the two invariant mass combinations of the opposite charged tracks, $m^{high}_{\pi\pi\pi\to\pi^-\pi^+\pi^+}(\pi^+\pi^-)$ and $m^{low}_{\pi\pi\pi\to\pi^-\pi^+\pi^+}(\pi^+\pi^-)$. In $D_s^- \to \pi^-\pi^-\pi^+$ decays, the invariant mass combinations of opposite charged tracks, $m^{high}_{D_s^-\to\pi^-\pi^-\pi^+}(\pi^-\pi^+)$ and $m^{low}_{D_s^-\to\pi^-\pi^-\pi^+}(\pi^-\pi^+)$ are also used. The variable $\cos\theta^*(D_s^-)$ is defined as the decay angle of the D_s^- with respect to the B_s^0 direction. When applicable, the decay angle of the resonance, $\cos\theta^*(\phi$ or $K^{*0})$ with respect to the D_s^- direction, is also used (Figure 3.3). The sum of the track impact parameter significance, d_0/σ_{d_0}(tracks), is another useful variable, as is the maximum separation along the nominal beam direction, $\max|\Delta z_0|$. The combined particle identification likelihood, CLL, is used in form of the maximum CLL, $\max(CLL)$, minimum CLL, $\min(CLL)$, the scalar sum, $\Sigma(CLL)$ of track candidate likelihoods. For the $B_s^0 \to D_s^-\pi^-\pi^+\pi^+, D_s^- \to \pi^+\pi^-\pi^-$ decay, the CLL of the highest momentum pions from the B_s^0 and D_s^- decays, respectively, are also used. Finally, for decay sequences involving the $D_s^- \to \phi\pi$ decay, the transverse momentum difference of the two kaon candidates from the ϕ decay, $|p_T(K^+) - p_T(K^-)|$, is used in the selection.

Variable	$B_s^0 \to D_s^- \pi^+$ modes			$B_s^0 \to D_s^- \pi^- \pi^+ \pi^+$ modes		
	$\phi\pi^-$	$K^{*0}K^-$	$\pi^-\pi^+\pi^-$	$\phi\pi^-$	$K^{*0}K^-$	$\pi^-\pi^+\pi^-$
$\chi^2_{r\phi}(B_s^0)$	√	√	√	√	√	√
$d_0(B_s^0)$	√	√	√	√	√	√
$L_{xy}(B_s^0)$	√	√	√	√	√	√
$L_{xy}/\sigma_{L_{xy}}(B_s^0)$	√	√	√	√	√	√
$L_{xy}(D_s^- \to B_s^0)$	√	√	√	√	√	√
$p_T(B_s^0)$	√	√	√	√	√	√
$\chi^2_{r\phi}(D_s^-)$	√	√	√	√	√	√
$d_0(D_s^-)$	√	√	√	√	√	√
$L_{xy}(D_s^-)$	√	√	√	√	√	√
$L_{xy}/\sigma_{L_{xy}}(D_s^-)$	√	√	√	√	√	√
$M(D_s^-)$	√	√	√	√	√	√
$p_T(D_s^-)$	√	√	√	√	√	√
p_T(tracks)	all	all	all	all	select tracks	all
$\cos\psi_H$	√	√	√	√	√	√
$M(\phi$ or $K^{*0})$	√	√	–	√	√	–
$p_T(\phi$ or $K^{*0})$	√	√	–	√	√	–
$\min(d_0/\sigma_{d_0})$	√	√	√	√	√	√
$\max(d_0/\sigma_{d_0})$	√	√	√	√	√	√
$\min(p_T)$	√	√	√	√	√	√
$\max(p_T)$	√	√	√	√	√	√
$m(\pi^-\pi^+\pi^+)$	–	–	–	√	√	√
$\chi^2_{r\phi}(\pi^-\pi^+\pi^+)$	–	–	–	√	√	√
$m^{high}_{\pi\pi\pi\to\pi^-\pi^+\pi^+}(\pi^+\pi^-)$	–	–	–	√	√	√
$m^{low}_{\pi\pi\pi\to\pi^-\pi^+\pi^+}(\pi^+\pi^-)$	–	–	–	√	√	√
$m^{high}_{D_s^-\to\pi^-\pi^-\pi^+}(\pi^-\pi^+)$	–	–	√	–	–	√
$m^{low}_{D_s^-\to\pi^-\pi^-\pi^+}(\pi^-\pi^+)$	–	–	√	–	–	√
$\cos\theta^*(D_s^-)$	√	√	√	√	√	–
$\cos\theta^*(\phi$ or $K^{*0})$	√	√	–	√	√	–
d_0/σ_{d_0}(tracks)	all	all	all	all	all	all
$\max\lvert\Delta z_0\rvert$	√	√	√	√	√	√
$\max(CLL)$	√	√	√	√	√	–
$\min(CLL)$	√	√	√	√	√	√
$\Sigma(CLL)$	√	√	√	√	√	–
CLL(tracks)	–	–	–	–	–	select tracks
$\lvert p_T(K^+) - p_T(K^-)\rvert$	√	–	–	√	–	–
ANN cut	0.954	0.954	0.996	0.970	0.958	0.988

TABLE 3.10: Variables used in ANN training and the optimal value of the cut on ANN output.

The mass distributions of B_s^0 candidates reconstructed in fully hadronic decays are shown in Figure 3.6. The cleanest channel mass spectrum is found for the decay channel $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ and it is shown over a wider mass range. For this specific decay mode the data sample contains incompletely reconstructed hadronic decays, which are added to the fully reconstructed signal peak. Although in the case of partially reconstructed B_s^0 candidates one or more particles are excluded in the decay topology the prominent structure below the main peak is dominated by a small number of channels which are only missing a soft neutral particle. The primary source of partially reconstructed modes in the $B_s^0 \to D_s^- \pi^+$ topology are the $B_s^0 \to D_s^- \rho^+$, $\rho^+ \to \pi^+\pi^0$ and $B_s^0 \to D_s^{*-}\pi^+$ with $D_s^{*-} \to D_s^-\gamma$ or $D_s^{*-} \to D_s^-\pi^0$. For these decay modes, the soft neutral γ or π^0 leave no track in SVX or COT detector, thus they are neglected in the final reconstruction. In the case of partially reconstructed hadronic modes, 96% of the momentum of a candidate is reconstructed on average, therefore they represent a potential source of significant statistical power in the measurement.

In the other five decay modes most of the partially reconstructed candidates are instead excluded applying a cut at $5.3\,\mathrm{GeV/c^2}$, exploing then the excellent detector mass resolution. In the main hadronic mode $B_s^0 \to D_s^- \pi^+$ with $D_s^- \to \phi\pi^-$ the lower bound is decreased up to $5.0\,\mathrm{GeV/c^2}$. The mass models for the hadronic decays have several components, which most of them are in common. The fraction and the shape assigned to the Cabibbo-suppressed decay $B_s^0 \to D_s^- K^+(\pi^+\pi^-)$ or to the background coming from Λ_b and B^0 mis-recontructed decays or the generic $b \to D_s^- X$ decays are described with detailed simulated events which take into account trigger and reconstruction efficiencies. The remaining background is the combinatorial one and it is generally due to the pairing of real D_s^- meson with random tracks from underlying events. For the particular case of the $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi\pi^-$ the physics of the partially decay modes and their interaction with the CDF II detector are carefully simulated. The resulting mass spectra are parametrized with empirical models. Given precise measurements of exclusive B_s^0 decays are not yet available, SU(3) symmetry is assumed and the corresponding B^0 branching ratios are applied. Any uncertainty about the relative contributions of the partially reconstructed signals is addressed in the systematic error studies. The detailed analysis of the partially hadronic sample is described in a dedicated study reportef in [52]. Finally in Table 3.11 the signal yield and the ratio signal-to-background in the [5.32, 5.42] $\mathrm{GeV/c^2}$ mass range for each decay chain is reported.

3.4 Data Sample for the B_s^0 Angular Analysis

Sample Size and Reconstruction of Decays

The data sample used for the B_s^0 angular analysis for the width difference, $\Delta\Gamma_s$ and CP violation phase β_s measurements described in this dissertation was collected with the CDF detector between March 2002 and October 2006. The total integrated luminosity is 1.35 $\mathrm{fb^{-1}}$. The `di-muon` trigger paths were used exclusively for this analysis.

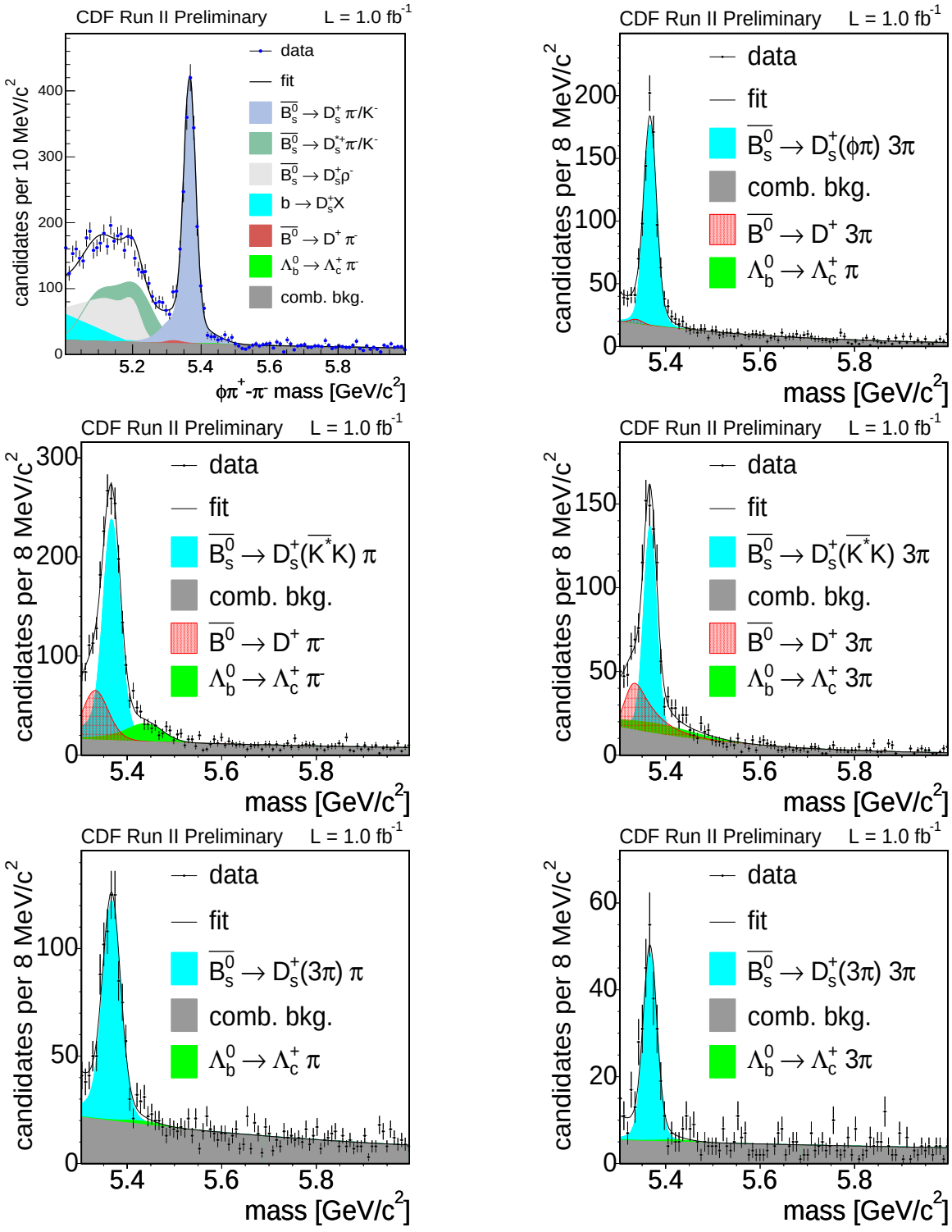

FIGURE 3.6: Hadronic mass spectra, showing all sample composition components. In the left column, $B_s^0 \to D_s^- \pi^+$ decays, in the right column, $B_s^0 \to D_s^- \pi^- \pi^+ \pi^+$ decays. Top to bottom, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to \pi^+\pi^-\pi^-$ decays.

Decay Sequence	S	S/B
$B_s^0 \to D_s^-(\phi\pi^-)\pi^+$	1,900	11.3
$B_s^0 \to D_s^-(K^*(892)^0K^-)\pi^+$	1,400	2.0
$B_s^0 \to D_s^-(\pi^+\pi^-\pi^-)\pi^+$	700	2.1
$B_s^0 \to D_s^-(\phi\pi^-)\pi^-\pi^+\pi^+$	700	2.7
$B_s^0 \to D_s^-(K^*(892)^0K^-)\pi^-\pi^+\pi^+$	600	1.1
$B_s^0 \to D_s^-(\pi^+\pi^-\pi^-)\pi^-\pi^+\pi^+$	200	2.6
Partially Reconstructed B_s^0 Decays	3,300	3.4
Total	8,800	—

TABLE 3.11: Signal yields for the hadronic modes (S) and signal to background ratio (S/B) in the various decay sequences.

Pre-Selection

The selection of $B_s^0 \to J/\psi\phi$ events, with $Jpsito\mu^+\mu^-$ and $\phi \to K^+K^-$, complies the same procedure discussed in section B. The use an Artificial Neural Network (ANN) for the event selection then follows some basic pre-selection cuts. These cuts are summarized in Table 3.12. Pre-Selection cuts serve several purposes:

- Certain regions of the event variable space are not well-modeled in Monte Carlo simulation. This is particularly true of tracks with low transverse momentum (p_T). We cut on these variables to ensure good Data-Monte Carlo agreement.
- Cut on certain variables to remove events that are clearly background. Indeed, polluting the training samples with events that are unequivocaly background events will simply add an additional burden on the network training. For instance this is the case for kaon tracks that have very pion-like values of particle ID variables.
- Reducing the number of events we use as input to the maximum likelihood fitter decreases the CPU time needed to perform the fit. Therefore, we include enough events from the mass sideband regions to be able to empirically describe the background angular distributions, but no more.

There are several variables used to pre-select the events. The quality of the vertex fits for the B_s^0 mesons ($\mathcal{P}(B_s^0)$) as well as for the two resonances ($\mathcal{P}(J/\psi)$, $\mathcal{P}(\phi)$). The number of silicon hits in ϕ associated with the muon tracks (μ_{nSiPhi}) to garanty good quality for the reconstructed tracks. Cuts on the minimum momentum of tracks and reconstruted particle from their decay product to ensure Monte Carlo simulation to better reproduce the data. Requirements on the reconstructed invariant masses for B_s^0, J/ψ and ϕ to peak around their PDG values [6] removes events which are clearly background. The upper cut on the lifetime error $\sigma_{c\tau}$ to cut out events which have a loose reconstruction. Finally the use of the combined particle identification likelihood (CLL) to be consistent with a kaon hypothesis for the track coming from the ϕ for removing fraction of the events where the kaon is misidentified.

Cut Variable	$B_s^0 \to J/\psi\phi$
$\mathcal{P}(B_s^0)$	$> 10^{-50}$
$\mathcal{P}(J/\psi)$	$> 10^{-50}$
$\mathcal{P}(\phi)$	$> 10^{-50}$
μ_{nSiPhi}	≥ 3
$p_T(B)$ [GeV/c]	> 4.0
$p_T(\phi)$ [GeV/c]	> 1.0
p_T^{μ} [GeV/c]	> 1.5
$p_T(K)$ [GeV/c]	> 0.4
$\lvert m(J/\psi\phi) - M_{PDG}^{B}\rvert$ [MeV/c^2]	< 250
$\lvert m(\mu^+\mu^-) - M_{PDG}^{J/\psi}\rvert$ [MeV/c^2]	< 80
$\lvert m(K^+K^-) - M_{PDG}^{\phi}\rvert$ [MeV/c^2]	< 14
$\sigma_{c\tau}$ [μm]	< 150
CLL^{K}	> -5.0

TABLE 3.12: Pre-Selection cuts on the B_s^0 sample collected with the di-muon trigger.

Neural Network Selection

The neural network is trained using the sidebands of the data as a background sample and, as a signal sample, realistic Monte Carlo simulation events that are passed through the standard GEANT-based [53] simulation of CDF II detector and are reconstructed as in real data. The Monte Carlo was generated using BGenerator [54] for event generation, and EvtGen [55] for the decays. In training, a neural network is given two samples, one background and one signal, and the same vector of variables for each event in the sample. Conceptually, the training consists of learning to discriminate between signal and background given the value of the vector for any given event. A trained neural network consists of weights given to each variable which will allow it, given an unknown event to assign it a value between 0 and 1 according to its probability of being background (0) or signal (1).

We use the following variables for B_s^0:

- χ^2 probability of the vertex fit for decay particles: B_s^0, J/ψ, ϕ
- χ^2 probability of the vertex fit in xy plane for the B_s^0
- Reconstructed mass of the vector particles: J/ψ, ϕ
- Transverse momentum p_T: B_s^0, J/ψ, ϕ
- Maximum and minimum of the $K^{+/-}$ p_T for each event
- Maximum and minimum of the $K^{+/-}$ combined log likelihood particle ID (CLL)
- Maximum and minimum of the $\mu^{+/-}$ likelihood, as described in [45] [48]

$B_s^0 \to J/\psi\phi$	NN Corr	KS dist
p_T^ϕ	0.63146	0.54423
$max(p_T^{K^+}, p_T^{K^-})$	0.62540	0.53913
$min(p_T^{K^+}, p_T^{K^-})$	0.62483	0.53448
$min(CLL^{K^+}, CLL^{K^-})$	0.37971	0.41272
p_T^{B}	0.35777	0.37895
$\chi^2_{xy}(\mathrm{B})$	0.28268	0.31295
$\mathcal{P}(\mathrm{B}_s^0)$	0.25047	0.26127
$min(lik^{\mu^+}, lik^{\mu^-})$	0.22534	0.22091
$max(CLL^{K^+}, CLL^{K^-})$	0.19340	0.52366
$max(lik^{\mu^+}, lik^{\mu^-})$	0.15854	0.18330
$p_T^{J/\psi}$	0.13714	0.18124
$m(\mu^+\mu^-)$	0.07343	0.15361
$\mathcal{P}(J/\psi)$	0.07180	0.07181
$m(K^+K^-)$	0.01451	0.21740
$\mathcal{P}(\phi)$	0.00463	0.02235

TABLE 3.13: Variables used to train the background-suppression neural network, with correlation to the neural network output value and Kolmogorov-Smirnov distance between signal and background distributions in the training samples.

These variables are shown in Table 3.13; also shown is the correlation of each variable to the neural network output value, as well as the Kolmogorov-Smirnov distance between the background and signal distributions in that variable.

Several tests are performed to verify that the neural network training proceeded as intended:

- **Output values for background and signal training events**. This is done in a set that is statistically independent from the one used for training. The result is shown as the upper left plot in Figure 3.7.

- **Purity of the sample vs neural network output**. This corresponds to the upper right plot in Figure 3.7. By definition, this plot should be monotonically increasing. If we generate the plot in a sample with the same signal to background ratio as the training samples, we also expect that it should be linear.

- **Generalization error**. We verify that we achieve the best possible generalization error as a function of training epoch, shown in the bottom left plot of Figure 3.7.

- **Correlations**. We check correlations among the training variables, in order to see that they match well-motivated physical expectations (bottom right plot in Figure 3.7).

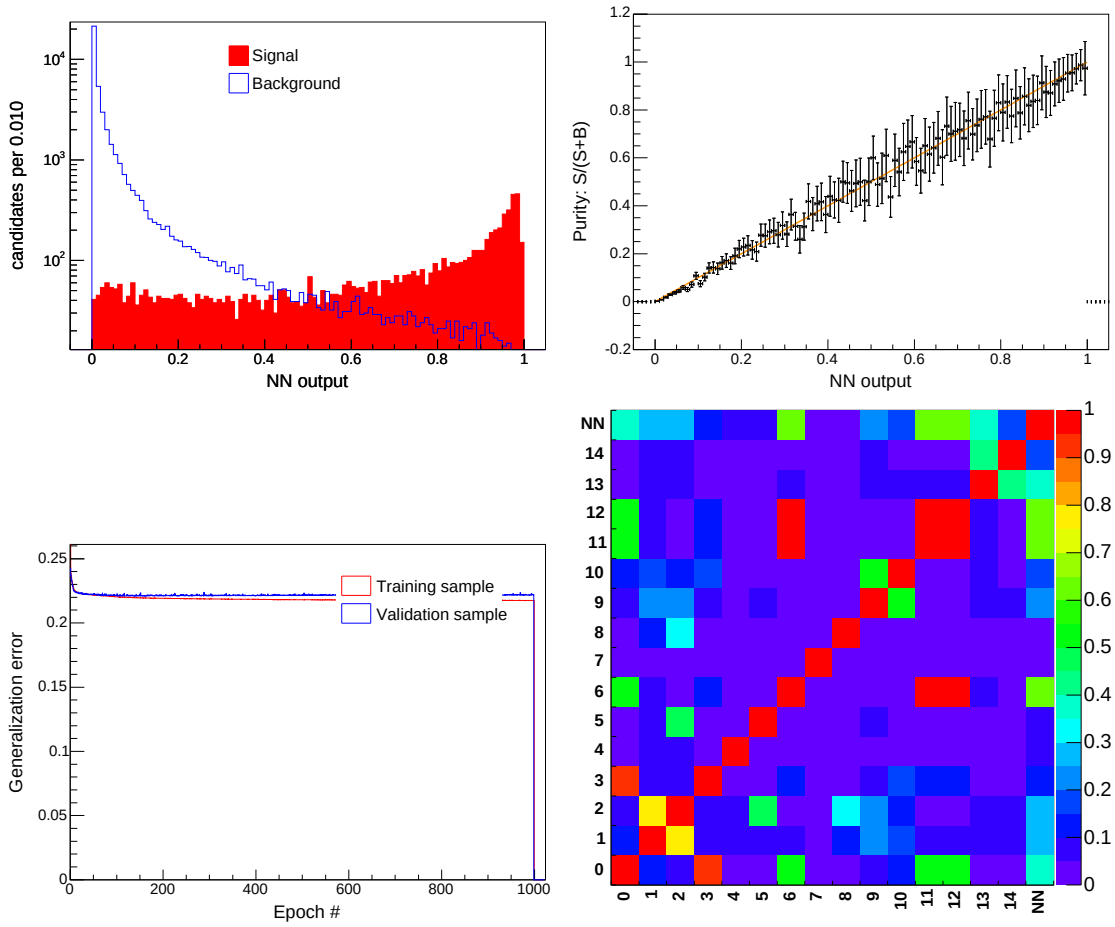

FIGURE 3.7: Neural network output plots. In order from top: Neural Network output values for background and signal, purity as function of Neural Network output, possible correlations of variables.

Once the neural network training is done, a decision has to be made about the output NN value cut. We use as a criterion the statistical significance, defined as $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, and choose the NN output value cut which would give our data sample with the greatest statistical significance. In order to obtain a good estimate of the signal and background fractions in out sample after the cut, we look as the mass distribution in the sample, which is an good determinant of the fraction of background. We use a simple fit model of Gaussian for signal and first order polynomial for background, then do a binned χ^2 fit.

In order to avoid biases when making the neural network cut, we do not compute the significance in the signal region using the number of data events in the Gaussian peak, but rather the number of Monte Carlo events in the Gaussian peak, divided by a factor to reflect the larger size of the Monte Carlo sample. We integrate the fitted signal Gaussian in $3\sigma_m$ around the PDG mass where σ_m is 0.0096 for B_s^0. A graph of the significance versus the neural network output is shown in the first plot of Figure 3.8.

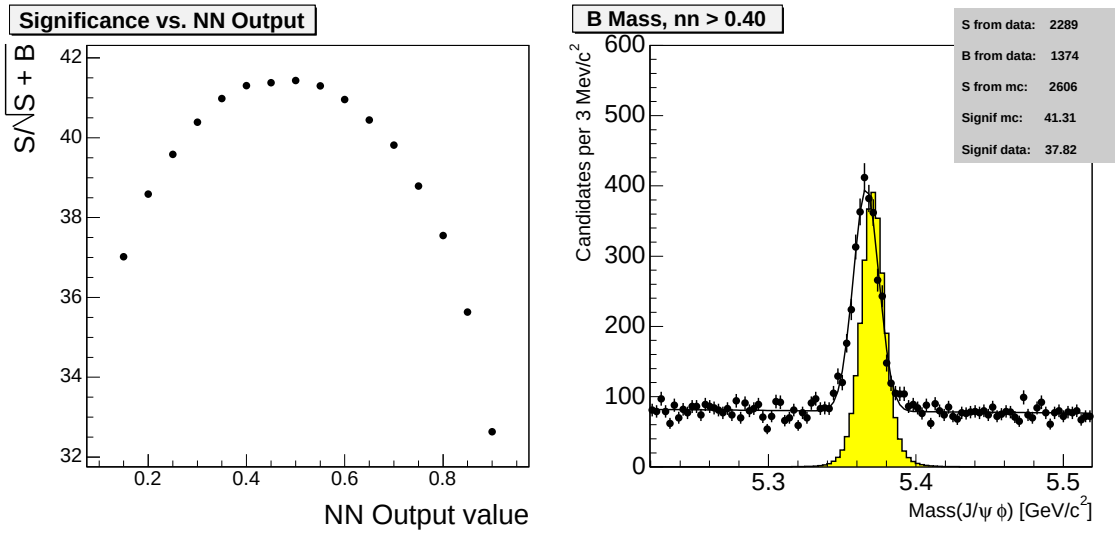

FIGURE 3.8: Significance plotted versus neural network output and mass distribution at chosen neural network cut of 0.40 for B_s^0. Monte Carlo is shown in yellow.

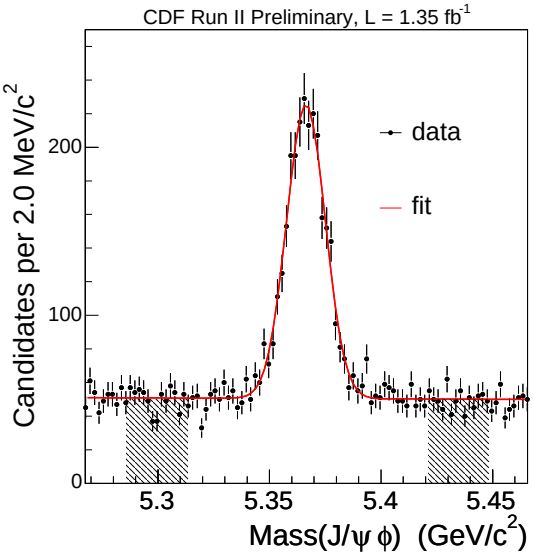

FIGURE 3.9: Invariant mass distribution after the chosen Neural Network output cut of 0.40 for B_s^0. The dashed regions corresponds to the sidebands used for the training $m(J/\psi\phi) \in [5.1280, 5.2142] \cup [5.3430, 5.3752]$.

Since we see that significance does not change very much for a range of neural network output cut in $\{0.40, 0.65\}$, we are faced with two possibilities: we can choose a tighter cut to reduce both signal and background, or a looser cut to enhance both signal and background. Motivated by the fact that our background is well modeled in the fit likelihood, we choose the looser cut to increase statistics and improve sensitivity.

In the second plot of Figure 3.8, we show the mass distribution and the result of the binned fit at the neural network cut value of 0.40. We obtain an expected signal yield of approximately 2000 signal events. The invariant mass distribution on the data after the Neural Network selection is shown in Figure 3.9.

The B_s^0 candidates reconstruction and selection for the $B_s^0 - \overline{B}_s^0$ oscillation frequency measurement as well as the angular analysis for the $\Delta\Gamma_s$ and β_s measurements has been presented in this chapter. The following chapter will focus on the description of the likelihood framework used.

Chapter 4

Elements for the B_s^0 Mixing and $\mathrm{B}_s^0 \to J/\psi\phi$ Angular Analyses

This Chapter is dedicated to the description of the key ingredients to perform the B_s^0 mixing analysis and $\mathrm{B}_s^0 \to J/\psi\phi$ angular analysis. The structure of the Chapter follows the time fashion of the analyses presented. Thus, primarily the B_s^0 mixing analysis elements are introduced. Many of this technicalities are in common with the $\mathrm{B}_s^0 \to J/\psi\phi$ angular analysis. Therefore, the description of the latter will mainly focus on the differences with respect to B_s^0 mixing analysis for the aspects concerning the proper time resolution and its calibration and the taggers performances.

4.1 Elements for the B_s^0 Mixing Analysis

This section presents the description of the method and the elements of the B_s^0 mixing analysis which resulted in the first observation of the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations.

4.1.1 The Amplitude Scan Method

The probability for a B_s^0 meson to decay at proper time t in the same ("unmixed" case) or opposite ("mixed" case) flavor with respect to the production flavor are (see Section [**?**]):

$$\mathcal{P}_{unmix}(t) = \mathcal{P}_{\mathrm{B}\to\mathrm{B}}(t) = \mathcal{P}_{\overline{\mathrm{B}}\to\overline{\mathrm{B}}}(t) = \frac{\Gamma}{2} e^{-\Gamma t} \left[1 + \cos\left(\Delta m t\right)\right] , \tag{4.1.1}$$

$$\mathcal{P}_{mix}(t) = \mathcal{P}_{\mathrm{B}\to\overline{\mathrm{B}}}(t) = \mathcal{P}_{\overline{\mathrm{B}}\to\mathrm{B}}(t) = \frac{\Gamma}{2} e^{-\Gamma t} \left[1 - \cos\left(\Delta m t\right)\right] , \tag{4.1.2}$$

with

$$\Gamma = \frac{\Gamma_L + \Gamma_H}{2} = \frac{1}{\tau_{B_S}} \quad \text{and} \quad \Delta m = m_H - m_L. \tag{4.1.3}$$

Figure 4.1 shows pictorially the probability density functions for a B_s^0 meson produced at time $t = 0$ to decay at the time t as a B_s^0 or $\overline{\mathrm{B}}_s^0$.

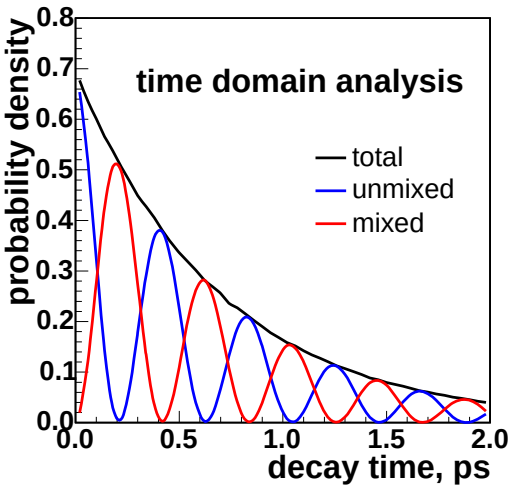

FIGURE 4.1: Probability density function for a B_s^0 meson produced at time $t = 0$ to decay at time t as a B_s^0 with the same flavor, "unmixed" case, or as a $\overline{B}_s^0$ with the opposite flavor, "mixed" case. The oscillation value used is equal to $\Delta m_s = 15$ pb^{-1}.

B_s^0 oscillations appear in the time dependence of the mixed asymmetry $\mathcal{A}(t)$:

$$\mathcal{A}(t) = \frac{\mathcal{P}_{unmix}(t) - \mathcal{P}_{mix}(t)}{\mathcal{P}_{unmix}(t) + \mathcal{P}_{mix}(t)} = \cos(\Delta m t) \tag{4.1.4}$$

A possible direct approach to measure the frequency Δm_s would be to fit for the asymmetry $\mathcal{A}(t)$ by counting, as a function of time, how many of the B_s^0 candidate decay with the same or opposite flavor compared to the production time. In SM, the B_s^0 mixing oscillation Δm_s is expected to be at least ~ 40 larger than Δm_d and in most of the SM extension it is even foreseen even to acquire higher values. Since it is not possible to know beforehand to have enough resolution for a direct oscillation observation, this method cannot be simply employed. Thus, historical mixing searches have been made possible introducing a Fourier-like coefficient, the amplitude $\mathcal{A}$, to scale the oscillation cosine term:

$$1 \pm \cos(\Delta m t) \rightarrow 1 \pm \mathcal{A} \cos(\Delta m t) \tag{4.1.5}$$

With the introduction of an amplitude $\mathcal{A}$, the search of the B_s^0 oscillation is performed in the frequency domain. The technique, called *amplitude scan* [56] [57] is described as follows. For a fixed oscillation frequency, the fit result for the parameter $\mathcal{A}$ is expected to be unit if the probed value Δm_s corresponds to the true value of the frequency. Otherwise it is expected to be consistent with zero. An example of an amplitude scan, produced with a toy Monte Carlo sample generated with $\Delta m_s = 15$ pb^{-1}, is shown in Figure 4.2.

The idea behind this approach is to scan increasing values of Δm_s, checking for the consistent values of the amplitude $\mathcal{A}$ with one. The possible scanned region depends on the sensitivity of the analysis, which is defined as the value of the frequency Δm_s

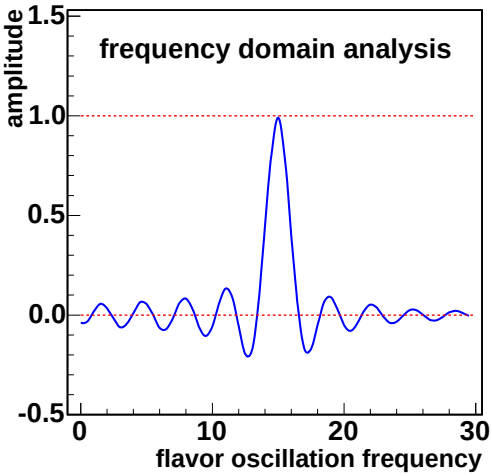

FIGURE 4.2: Amplitude scan for a toy Monte Carlo when the true value of the oscillation frequency is assumed to be equal to $\Delta m_s = 15$ pb^{-1}.

for which a measured null amplitude $\mathcal{A} = 0$ would imply the exclusion of $\mathcal{A} = 1$ at the desired confidence level, which is nominally chosen to be 95% in our analyses. The error on the amplitude $\sigma_{\mathcal{A}}$ is Gaussian distributed, so a Δm_s value can be excluded at 95% confidence level if the amplitude $\mathcal{A}$ respects the condition:

$$\mathcal{A} + 1.645 \cdot \sigma_{\mathcal{A}} \leq 1 \tag{4.1.6}$$

The exclusion limit is defined as the largest oscillation frequency that would be excluded in the ideal non-oscillation $\mathcal{A} = 0$:

$$1.645 \cdot \sigma_{\mathcal{A}} = 1 \tag{4.1.7}$$

The addition in utilizing a method which is not a direct search for the oscillation is the relative simplicity in combining two (or more) amplitude scans from two (or more) independent experiments [57]:

$$\mathcal{A} = \left(\frac{\mathcal{A}_1}{\sigma^2_{\mathcal{A},1}} + \frac{\mathcal{A}_2}{\sigma^2_{\mathcal{A},2}} \right) \text{ with } \frac{1}{\sigma^2_{\mathcal{A}}} = \left(\frac{\mathcal{A}_1}{\sigma^2_{\mathcal{A},1}} + \frac{\mathcal{A}_2}{\sigma^2_{\mathcal{A},2}} \right), \tag{4.1.8}$$

An example of combination between several experiments is reported in Figure 4.3, where the amplitude scans from LEP, SLD and CDF Run I are combined together. The lower limit on the B_s^0 frequency is $\Delta m_s > 14.5$ pb^{-1} at 95% Confidence Level with a sensitivity of 18.2 pb^{-1}.

It is manifested how the amplitude uncertainty $\sigma_{\mathcal{A}}$ is of particular importance because the analysis quality depends on its capability to discern unity from zero in the amplitude measurement. The estimator of $\sigma_{\mathcal{A}}$ is given by the following formula [56] [57]:

$$\frac{1}{\sigma_{\mathcal{A}}} \simeq \frac{\mathcal{S}}{\sqrt{\mathcal{S} + \mathcal{B}}} \, e^{-\frac{\Delta m_s^2 \sigma_{ct}^2}{2}} \sqrt{\frac{\epsilon \mathcal{D}^2}{2}} \tag{4.1.9}$$

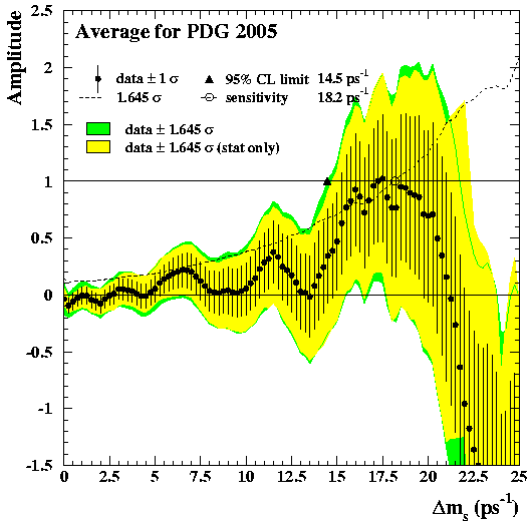

FIGURE 4.3: Combined measurement of the B_s^0 oscillation amplitude as a function of Δm_s, including all published results from LEP, SLD and the CDF Tevatron Run I.

Thus, the main obstacles to B_s^0 mixing analysis stem from three main contributions. The statistic power available expressed by the term $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$ where $\mathcal{S}$ and $\mathcal{B}$ are, respectively, the reconstructed events as signal and background. The dependence on the proper time resolution: the higher the resolution the higher will be the capability to search for larger oscillation values. Finally the ability to identify the B_s^0 flavor at the production time, which is defined by the figure of merit of the tagging algorithm used, $\epsilon\mathcal{D}^2$. The efficiency ϵ corresponds to the fraction of events to which the algorithm assigns a non null tag decision. The dilution $\mathcal{D}$ of the tagger is defined as $\mathcal{D} = 2\mathcal{P}_{tag} - 1$, where $\mathcal{P}_{tag}$ is the probability to tag correctly the B_s^0 candidate.

While the samples description and their optimized selections are reported in Chapter [?], the detailed discussion of the proper decay time resolution and the tagging algorithms are the main topics of the following sections.

4.1.2 Proper Decay Time

The measurement of the proper decay time is essential for the B_s^0 mixing observation. In the laboratory rest of frame it is defined as follows:

$$ct = \frac{L_{xy}(B_s) \cdot M_{B_s}}{p_T(B_s)} \tag{4.1.10}$$

where the observables for the event are the B_s^0 transverse momentum, $p_T(B_s)$ and the B_s^0 transverse decay length, $L_{xy}(B_s^0)$ defined as the transverse displacement of the B_s^0 decay vertex from the production vertex. The transverse decay length, L_{xy} is measured as the distance between the primary vertex of the $p\bar{p}$ collision and the secondary vertex position of the B_s^0 decay, projected onto the $p_T(B_s^0)$ direction:

$$L_{xy} = \frac{(\overrightarrow{r}_{SV} - \overrightarrow{r}_{PV}) \cdot \overrightarrow{p}_T(B_s)}{|\overrightarrow{p}_T(B_s)|} \tag{4.1.11}$$

The value of mass M_{B_s} used in the proper decay time formula comes, instead, from the world average computation [58].

As observed in Equation **??**, the statistical significance of the oscillation amplitude is strongly affected by the proper time resolution, via the relation $1/\sigma_A \propto e^{-\frac{\Delta m^2 \sigma_{ct}^2}{2}}$. Thus, achieving an good proper time resolution measurement in the B_s^0 mixing analysis is crucial. From the definition of Equation 4.1.10, the uncertainty on the ct measurement is obtained:

$$\sigma_{ct} = \left(\sigma_{L_{xy}} \cdot \frac{M_B}{p_T}\right) \oplus \left(ct \cdot \frac{\sigma_{p_T}}{p_T}\right) \tag{4.1.12}$$

where $\oplus$ implies the sum in quadrature of the terms and the uncertainty on M_B is not considered, because negligible. The resolution of the proper decay time depends on two contribution, the first of which arises from the uncertainty on the spatial position of the decay vertex, and the second one is relative to the uncertainty in the momentum reconstruction of the B_s^0 candidate. In the fully reconstructed decays, the σ_{p_T} contribution can be neglected and the precision measurement of the second vertex is the sole responsible for σ_{ct}. In the partially reconstructed modes it becomes an important addition to the global uncertainty. In fact in the semileptonic and partially reconstructed decay modes, the observables of one or more particle involved in the decay are not measured. Therefore, the proper decay time reconstructed using Equation 4.1.10 is not the proper decay time of the B_s^0 candidate and a correction factor, called "k-factor", is necessary to account for the missing momentum:

$$ct = ct^* \cdot k \,, k \equiv \frac{L_{xy}(B_s^0)}{L_{xy}^{reco}} \frac{p_T(B_s^0)}{p_T^{reco}} \tag{4.1.13}$$

where ct^* is traditionally called pseudo proper decay length and it is measured with only the information coming from the reconstructed candidates. Since it is not possible to determine the value of k event-per-event, this misknowledge is treated as a "probability density distribution" for an entire class of events. That is, an average distribution, $F(k)$, for the k-factor is obtained from BGENERATOR-MC simulation, and constitutes an important ingredient for the fit of proper decay length. The distributions of k-factors in the hadronic partially reconstructed and semileptonic decay modes is shown in the left plot of Figure 4.4.

It is observed that the broader the k-factor distribution, the greater is the uncertainty σ_{p_T}/p_T in Formula 4.1.12. Moreover, the resolution σ_{ct} worsens as the reconstructed lifetime increase. Thus once more appear evident the importance of having a well-defined momentum distributions for the incompleted reconstructed events.

The uncertainty returned by the CTVMFT vertex fitter [47] is in generally underestimated, therefore a correcting "scale factor" S_{ct} is applied to the resolution when performing a maximum likelihood fit on the data:

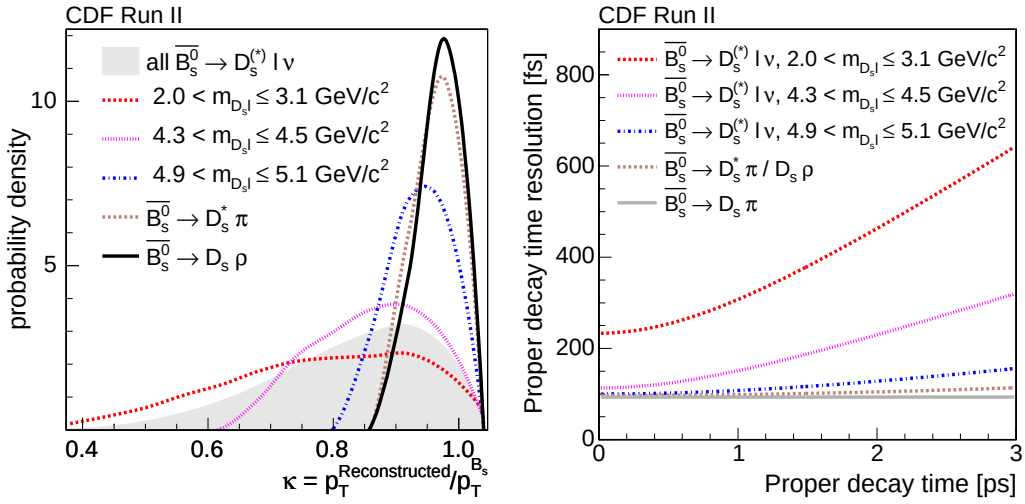

FIGURE 4.4: Left: k-factor distributions for different ℓD_s^- mass regions for the semileptonic and partially reconstructed hadronic decay mode. Right: the proper decay time resolution, σ_{ct}, as a function of the proper decay time lenght ct, extracted from Equation 4.1.10.

$$\sigma_{ct} \to S_{ct} \cdot \sigma_{ct} \tag{4.1.14}$$

The value of the scale factor is derived from data. In the $B_s^0 \to J/\psi\phi$ sample we will find out that the parameter adjustment of the Gaussian width describing the prompt component in the proper decay time distribution (see Chapter [?] is sufficient to obtained a corrected measurement of S_{ct}. However the prompt contribution is not a characteristic of lifetime-biased sample selected with the displaced track trigger. Then, a more detailed study has to be performed, with particular relevance in the B_s^0 mixing case.

4.1.3 Calibration of Proper Decay Time Resolution

The estimation of proper decay uncertainty is dramatically important when probing for high oscillation frequencies and for establishing reliable exclusion regions. The σ_{ct} expansion in Equation 4.1.12 shows that most of the complications arise from the determination of the error in the Secondary Vertex (SV) measurement. To perform a detailed study, an ensemble of B_s^0 candidates decaying at know positions would be needed. Since it is not the case, an "ad-hoc" calibration sample which mimics as closely as possible the kinematics and vertex topology of the signal samples has been selected. The main characteristic of the sample is to be "prompt": it contains events with a B-like vertex topology and expected to decay in the vicinity of the primary vertex of the $p\bar{p}$ interactions.

In the B_s^0 mixing analysis, the displaced track trigger (TTT), which selects the signal samples, is used to extract the calibration sample. The latter is obtained by pairing a real D meson with one (or three, depending on the channel topology) prompt

Cut		Dπ	Dπππ
$\chi^2_{r\phi}(D(3)\pi)$	$<$	15	15
$\lvert m_D - M_D^{PDG}\rvert$ [MeV/c^2]	$<$	8	8
$Mass$ [GeV/c^2]	$\in$	[5.4, 6.0]	[5.4, 5.8]
$\lvert m_{\pi\pi\pi}\rvert$ [GeV/c^2]	$<$	-	1.75
$p_T(D(3)\pi)$ [GeV/c]	$>$	5.5	6.0
$p_T(\pi)$ [GeV/c]	$>$	1.2	-
$\lvert d_0(D)\rvert$ [μm]	$<$	100	100

TABLE 4.1: Selection requirements for the Dπ and Dπππ samples for σ_{ct} calibration.

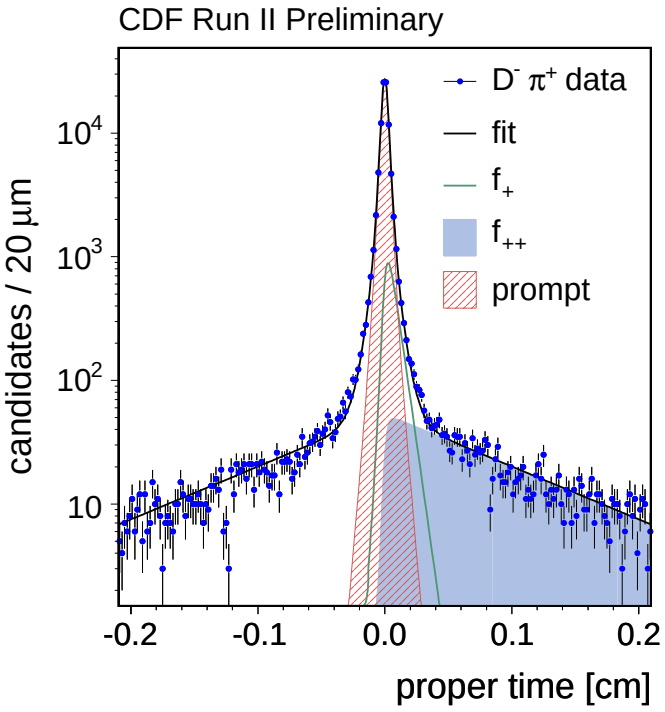

FIGURE 4.5: Representative ct distribution and fit projection for Dπ topology.

track(s) at the primary vertex. The selection cuts are listed in Table 4.1. In such way it is possible to select similar B_s^0 events topology where most of the candidates originates from the Primary Vertex (PV), and thus have $L_{xy} \sim 0$ by construction.

The distribution of the proper decay-time measured in calibration sample for the Dπ sample is shown in Figure 4.5.

The dominant contribution is, as desired, a prompt Gaussian. In addition there are smaller contributions coming from secondary D mesons originating from B_s^0 candidates, mis-reconstructed D mesons and paired tracks not coming from the PV. These are found to be fitted by the sum of a short and long lived exponential tail, symmetrized about zero.

The spatial uncertainty on the secondary vertex depends on various characteristics

of the decay, as its kinematic and quality of the fit. The aim is to parametrize the scale factor, utilizing the observable from which it has a clear dependence. The achievement is to estimate the scale factor for the signal sample on a per-event basis. The following set of variables is used:

- $\Delta R(D\pi)$**:** angular distance between D and the pion(s),
- $I = p_T(D(3)\pi)/\sum_i p_T^i(\Delta R < 0.7)$**:** Isolation,
- η**:** Pseudorapidity,
- z **position:** beamline z axis position,
- χ^2**:** chi-square of the vertex fit in the transverse plane.

Due to the large calibration sample size the dependence of S_{ct} can be binned versus the observables listed above. Parabolic functions are sufficient to parametrize the scale factor dependencies. Assuming the several contribution to be factorisable, the scale factor can written as:

$$S_{ct} = F(\Delta R) \cdot F(I) \cdot F(\eta) \cdot F(z) \cdot F(\chi^2) \tag{4.1.15}$$

The parameters of the functions $F(i)$ - with the index i running through $i = \Delta R,\ I,\ \eta,\ z,\ \chi^2$ - are computed in an iterative way. The scale factor variation with respect to one quantity is fitted and then corrected for, and only then, the new dependence in the next variable in the list is fitted and the correction applied. After the final tuning it would be expected to have a scale factor with a distribution flat centered in one. In the realistic case, residual deviations are still present. The results of this technique are illustrated in Figure 4.6 which visualizes, for instance, how the original dependence on the isolation has been absorbed in the final parametrization. Since the calibration is extracted for each of the three data-taking period used for internal study and they are qualitatively similar we report the results from the longest data period corresponding to 410 pb^{-1}. Table 4.2 shows the improvement in the determination of the time resolution, when the correction is applied. A systematic uncertainty on the oscillation amplitude is evaluated to account for possible deficiency in this method. In the B_s^0 mixing analysis a residual variation on $\pm \sim 4\%$ from unity for the scale factor will be applied as gross shift to the overall scale factor in the data (more details in Chapter [**?**]). Different global scale factor are utilized for each B_s^0 decay mode.

Figure 4.7 depicts the distributions of the proper time resolution for hadronic and semileptonic B_s^0 decays, after the application of the above scale factor parametrization on an event-by-event basis. While for the fully reconstructed decay modes (the hadronic B_s^0 ones), the proper time uncertainty comes from this vertex resolution contribution, in the case of the partially reconstructed decays, the momentum uncertainty, characterized by the k-factor distribution $F(k)$, is an additional component which has to be accounted for. The description of how this will be incorporated in the whole fitter framework is given in Section [**?**].

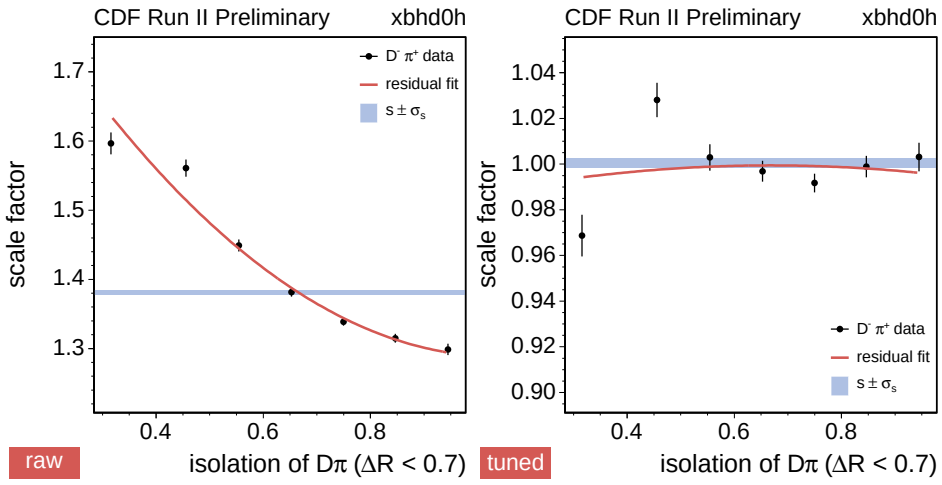

FIGURE 4.6: Scale factor dependence before (left) and after (right) parametrization for the isolation variable for the Dπ topology.

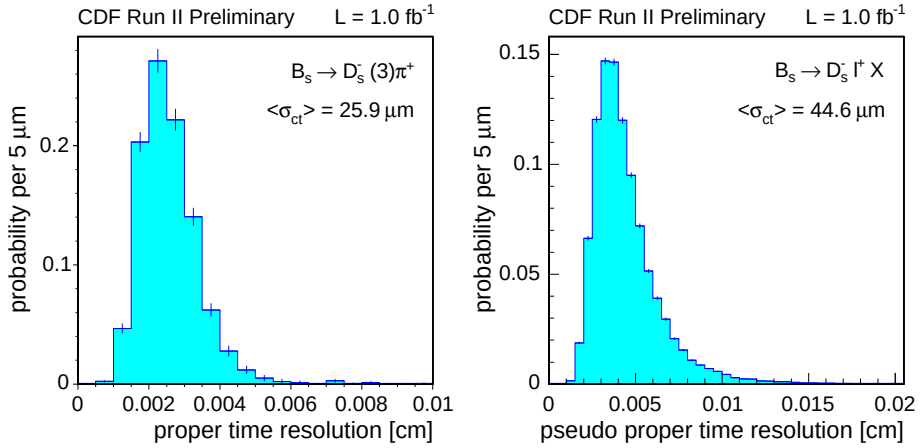

FIGURE 4.7: Distribution of the proper time resolution for hadronic (left) and semileptonic (right) B_s^0 decays.

Tuning	Average Scale Factor
raw	1.3807 ± 0.0030 (stat)
fully tuned	1.0006 ± 0.0021 (stat)

TABLE 4.2: Average scale factor in the Dπ sample for the period called `xbhd0h` using the internal CDF nomenclature , corresponding to a data-taking period of 410 pb^{-1}.

4.1.4 Flavor Tagging Algorithms

The flavor tagging, after the sample composition and the proper decay time resolution, constitutes another key ingredient of analyses involving the mixing phenomenon. While the flavor of the B_s^0 candidates can be unambiguously inferred by the charge sign of its daughter particles, the flavor tagging algorithms determine with a certain degree of uncertainty the B_s^0 flavor at production time. At the Tevatron accelerator b quarks are mostly produced in pairs in the hard parton interactions. Therefore the tagging algorithms are naturally divided in two main classes: the "Same Side" flavor taggers (SST) and the "Opposite Side" flavor taggers (OST). The former explores the flavor charge correlation between the triggered B_s^0 (which satisfy the trigger requirements) and the fragmentation tracks produced in the hadronization process nearby the meson. The opposite side algorithms examine the decay products of the other b-hadron produced in the $p\bar{p}$ collision: the flavor of the b-hadron is thus connected (opposite) to the flavor of the triggered B_s^0 meson. In the following sections we review in more details the technical development and study of the flavor tagging algorithms, being the area where my personal contribution to the B_s^0 mixing oscillation analysis has been more substantial.

Tagging Definitions

The performances of a tagger are quantified by two main quantities: the efficiency ϵ which is the measurement of the number of B_s^0 candidates we have tagged over the total number of events and the dilution $\mathcal{D}$ defined as:

$$\mathcal{D} \equiv 2\mathcal{P} - 1, \tag{4.1.16}$$

where $\mathcal{P}$ is the probability of a tagger to provide the correct decision. A perfect tagger is expected to have maximal dilution, $\mathcal{D} = 1$, whereas a completely random tagger - in case $\mathcal{P} = 0.5$ - would provide the worst dilution, $\mathcal{D} = 0$.

The sensitivity of B_s^0 mixing oscillation analysis depends explicitly from these two quantities (see Equation 4.1.9). The amplitude uncertainty $\sigma_{\mathcal{A}}$ is inversely proportional to $\epsilon\mathcal{D}^2$, which is thus generally used as a figure of merit of a tagging algorithm. Moreover the dilution $\mathcal{D}$, being an estimate of the flavor identification at production time enters directly in the oscillation measurement modulating the flavor asymmetry from Equation 4.1.4:

$$\begin{aligned} \mathcal{A}(t) &\rightarrow \mathcal{D}\cos(\Delta m t). \\ \mathcal{P}_{mix/unmix}(t) &\rightarrow \frac{1}{2}\big[1 \pm \mathcal{D}\cos(\Delta m t)\big]. \end{aligned} \tag{4.1.17}$$

Before going through the detailed description of the tagging algorithms it is important to stress that the taggers will be finally treated as independent [59]. In fact the SST and OST techniques are developed respectively within and outside an isolation domain which, for a given event, corresponds to a cone ΔR of 0.7 around the

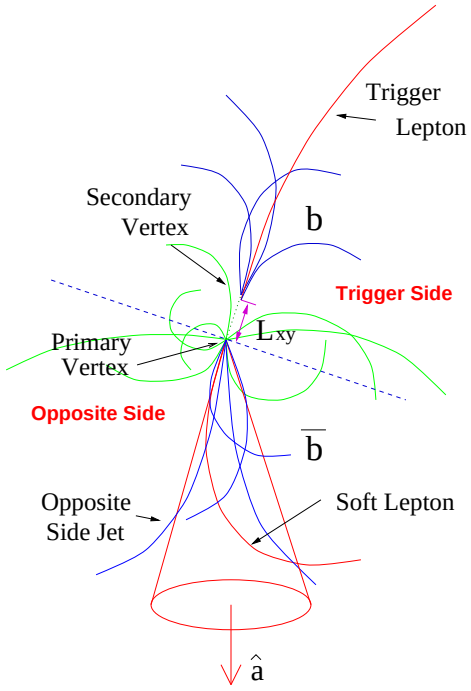

FIGURE 4.8: Illustration of a $b\bar{b}$ event accepted by the $\ell + SVT$ track trigger, where the *trigger lepton* defines the trigger side.

reconstructed B_s^0 candidate direction. In such way the information coming from the two flavor algorithms are kept uncorrelated.

4.2 Opposite Side Flavor Tagging

Opposite side taggers exploit the production of $b\bar{b}$ pairs at hadron colliders. Its decision is based on the analysis on the properties of the other b-hadron of the event, which is not triggered. This technique presents some limitations arising from the fact that the second b-hadron decay products are not within the fiducial volume of the detector acceptance or that the b-hadron is neutral B meson (e.g. B^0), thus subject to the oscillation phenomenon into its antiparticle. The addition of this technique is that the study of the attributes of the not triggered b-hadron makes the opposite side tagging performances independent by the species of the reconstructed B meson. Therefore it is possible to develop the OST and parametrize its dilution on a event-per-event fashion, based on the properties of the other b-hadrons directly on data. For this purpose at CDF it was utilized an high statistics sample of semileptonic B decays collected with the *lepton-plus-displaced-track* trigger described in Section [**?**]. A cartoon of an event accepted by the $\ell + SVT$ is shown in Figure 4.8.

After the dilution parametrization, we are allowed to simply transfer the information obtained on the large calibration sample to the other B_s^0 samples where a global factor is applied to account for kinematics differences between samples collected with different trigger paths.

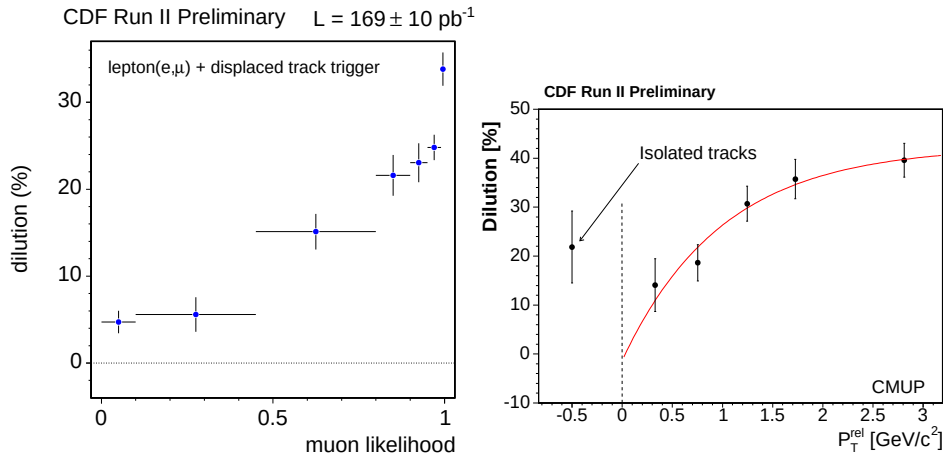

FIGURE 4.9: Left: muon tag dilution as a function of the muon likelihood. Right: dilution dependency of the soft muon tagger for CMUP muon type as a function of P_T^{rel}, the transverse momentum relative to the jet axis.

Soft Lepton Taggers

Soft lepton taggers look for lepton from the semileptonic decays of the opposite side b-hadrons, $b \to c\ \ell^-\ \bar{\nu}_\ell\ X$. The charge of the lepton, either an electron or a muon, is correlated with the b-flavor at production time: the ℓ^- originates from a b quark, while the ℓ^+ from a $\bar{b}$. It is the cleanest signature among the OST methods, because of the good purity in the lepton identification, so it is expected to have relatively high dilution. On the other hand the semileptonic B branching ratio is small, $\mathcal{BR}(\mathrm{B} \to \ell X) \simeq 20\%$, therefore the algorithm has low efficiency. Utilizing the information from several detector systems, the leptons are identified via likelihood discriminants. The latters have already been described in Section [**?**] and are extensively treated also in References [45] and [48]. The lepton tagging dilution is parametrized as a function of p_T^{rel}, which is the magnitude of the lepton momentum projected onto the perpendicular plane to the axis of the jet direction, where the jet momentum is recomputed without the lepton contribution. The semileptonic sample, due to its large size, is broken in bins of the relative lepton likelihood and for each bin the parametrization is determined. The electron and muon are treated as separated algorithms, where the higher purity of the muon chamber signature with respect to the electron calorimetric one results in a higher dilution for the soft muon tagger. In Figure 4.9 a pictorial example of variations of the muon tagger are shown.

Jet Charge Taggers

The jet charge method is based on the observation that the sum of all particle charges in a jet containing a B meson is correlated with the b quark charge, and thus to its flavor. An appropriately momentum-weighted sum gives a better result. The Jet Charge Tagger (JQT) uses a cone clustering algorithm with a maximum cone size of $\Delta R = 1.5$ to find the track-based jet [60]. The definition of the jet charge Q_{jet} is

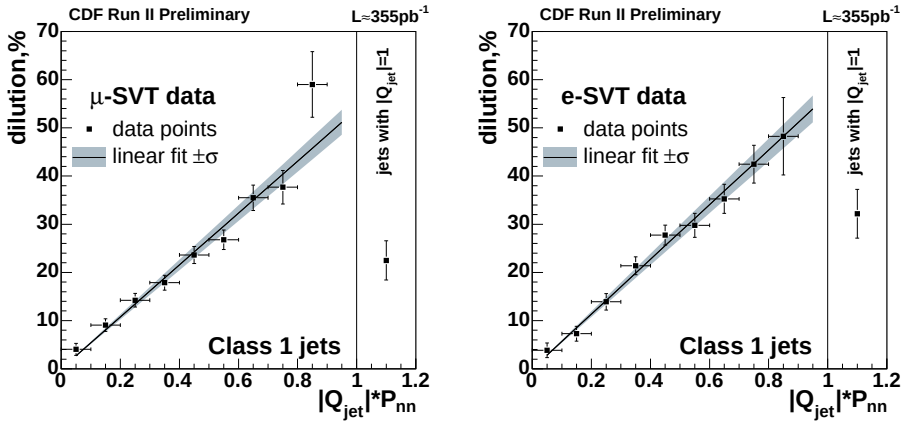

FIGURE 4.10: Dependence of the measured dilution on the quantity $|Q_{jet} \cdot \mathcal{P}_{nn}|$ for the $\mu + SVT$ dataset (left) and $e + SVT$ dataset (right) for the SecVtx class, labelled as Class 1.

given by

$$Q_{jet} \equiv \frac{\sum_{i=1}^{N} Q_i \cdot p_T^i \cdot (1 + \mathcal{P}_{trk}^i)}{\sum_{i=1}^{N} p_T^i \cdot (1 + \mathcal{P}_{trk}^i)} \tag{4.2.1}$$

where Q_i and p_T^i are the charge and the transverse momentum of a track in the jet, an $\mathcal{P}_{trk}^i$ is the probability that the track belongs to a b-jet and the sum run over all the N tracks belonging to the jet. Artificial Neural Networks, trained on a large PYTHIA Monte Carlo samples, are used to identify the most probable b-jets from a set of jets reconstructed with the mentioned cone clustering algorithm. The analysis in performed in two steps; initially an ANN is built to estimate the probability of each track to have originated from a b hadron decay, $\mathcal{P}_{trk}$. Afterwards a second ANN is fed with the probability $\mathcal{P}_{trk}$ for each track with additional kinematic variables to determine the probability $\mathcal{P}_{nn}$ for the associated jet to be that of a b quark. In order to better exploit the statistical power of this algorithm, the jets are splitted in three mutually exclusive classes. The first class contains jets satisfying the condition $L_{xy}/\sigma_{L_{xy}} > 3$ which are consistent with coming from a secondary vertex. In CDF jargon they are referred as *SecVtx* jets. If no SecVtx jet is found, the second class include jets with at least one track in the jet such that $\mathcal{P}_{trk} > 50\%$. Finally the third class contains all the remaining jets and the one having the highest p_T is chosen. Due to inclusive nature of jets, these algorithms are expected to have low dilution but high efficiency. The tagger purity and consequentially the dilution decreases from the first to the third class. Tagger dilution is parametrized using the quantity $|Q_{jet} \cdot \mathcal{P}_{nn}|$. In Figure 4.10 representative results are shown in the $\ell + SVT$ data for the SecVtx class.

Opposite Side Kaon Tagger

The opposite side kaon tagger (OSKT) exploits the knowledge that b quarks are likely to have the following sequential decay $b \to c \to s$. The OSKT [61] attempts to find

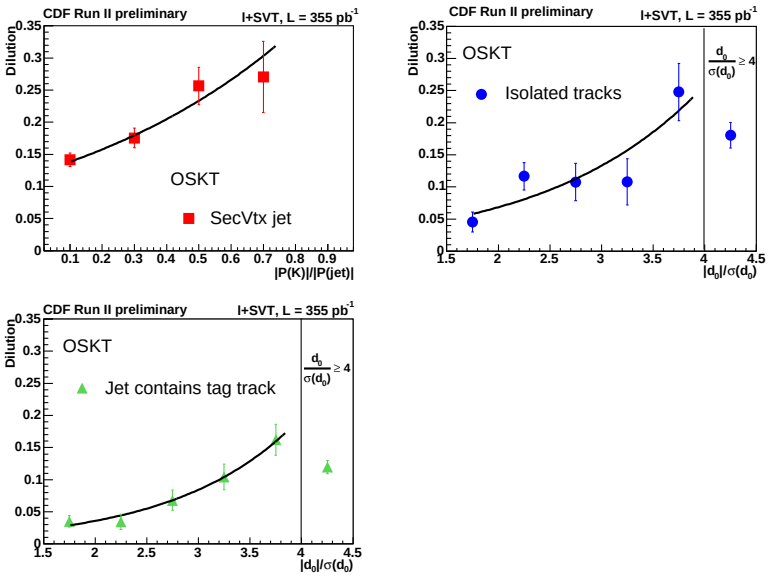

FIGURE 4.11: Dependence of the measured dilution on the fraction of jet momentum carried by SecVtx jet (upper left plot), and on the absolute impact parameter significance fr Class 2 (upper right and Class 3 (bottom left) tags.

this kaon and thereby correlate its charge with the B_s^0 flavor: a K^- originates from a b quark, while the K^+ from the $\bar{b}$ quark. The algorithm relies on particle identification utilizing the information from the TOF and COT drift chamber detectors. The challenge is, in fact, to discriminate kaons among the pions large background of pions. The PID information consist in a probability $\mathcal{P}(i)$ for a particle to be a kaon, a pion or a proton, $i = k, \pi, p$. All these information are combined via a likelihood ratio $\mathcal{LR}$:

$$\mathcal{LR}(K) = log\Big(\frac{\mathcal{P}(K)}{f_k\mathcal{P}(K) + f_\pi\mathcal{P}(\pi) + f_p\mathcal{P}(p)}\Big), \tag{4.2.2}$$

where $f_k = 0.2$, $f_\pi = 0.7$ and $f_p = 0.1$ are the *a priori* fractions for kaons, pions and protons in the sample, as measured in Reference [62]. As for the JQT the algorithm is broken into three exclusive classes according to increasing final dilution. In the first class the kaon is contained in a jet produced at a secondary vertex. The events in the second class do not have a SV identified, but isolated tracks. These tracks satisfy the requirements to be part of a jet, but no other track could be associated with them to form a jet. The third class contains all other tags. Because of the already mentioned difficulty to identify kaons within the large hadronic background, the dilution and the efficiency of the OSKT are both modest. In Figure [**?**] we report the dilution functional form for the different classes depending on different quantities,according to the class itself.

The performances of the three tagging algorithms, described so far, are summarized in Table 4.3. The figure of merit $\epsilon\mathcal{D}^2$ is used as parameter to compare the relative tagging power of each flavor tagger.

Flavor Tagger	$\epsilon\mathcal{D}^2[\%]$
Sotf-muon	$0.559 \pm 0.094 \pm 0.027$
Sotf-electron	$0.264 \pm 0.054 \pm 0.022$
Jet-charge SecVtx	$0.230 \pm 0.068 \pm 0.017$
Jet-charge Track Prob	$0.347 \pm 0.084 \pm 0.020$
Jet-charge Track p_T	$0.152 \pm 0.055 \pm 0.024$
Opposite side kaon	$0.229 \pm 0.016 \pm 0.001$

TABLE 4.3: Performances of the opposite-side flavor taggers. The measured values of $\epsilon\mathcal{D}^2$ are quoted with their statistical and systematic uncertainties.

Flavor Tagger	$\epsilon\mathcal{D}^2[\%]$
Hierarchical combination	1.5 ± 0.1
NN combination	1.8 ± 0.1

TABLE 4.4: Comparison of the opposite-side flavor taggers performances. The measured values of $\epsilon\mathcal{D}^2$ are quoted with their statistical and systematic uncertainties summed in quadrature.

Neural Network Combination of OST

Previous incarnation of the CDF B_s^0 mixing analysis [14] used an exclusive combination of the OST, based on the hierarchy of the average tag dilution. In the case of multiple tagging available, the choice of the algorithm was performed as following: muon tags were preferred to electron tag, which were in turn preferred to jet charge. At that time the OSKT was not yet available.
By applying a hierarchic choice we discard about all the possible correlations among the information coming from different taggers. Thus, all the multiple decision were finally combined with the use of a Adaptive Neural Network. Exploiting an ANN it has been possible to produce the "combined opposite side tagging" COST [63], a single tagging algorithm of higher average dilution which has the advantage to maximize the tagging information available for an event. The NN was trained on the $\ell + SVT$ data sample, using as input the tagging decisions, the predicted dilutions and various kinematic variables. The comparison between the COST and the mutually exclusive combination of the OSTs in terms of the tagging effectiveness $\epsilon\mathcal{D}^2$ are reported in Table 4.4. A relative improvement of approximately $\simeq 20\%$ is observed with respect the hierarchical combination.

4.2.1 Opposite Side Tagger Calibration

The advantage of opposite side tagging relies on its independence from the B species and decay mode analyzed. Therefore the opposite side tagging parametrization obtained on the $\ell + SVT$ data can be transferred to other trigger paths, for instance the TTT hadronic and semileptonic data samples. However, differences in the overall dilution may arise because of the different kinematic between the samples studied and the calibration one. To account for such differences, we introduce a global scale

factor for the tagging algorithm(s) to correct the predicted dilution:

$$\mathcal{D} \rightarrow S_{\mathcal{D}} \cdot \mathcal{D} \tag{4.2.3}$$

In fact, certain properties of the calibration sample (like $p_T(B)$ spectrum) do not match between the calibration sample and the signal one. Therefore, both the semileptonic and hadronic B_s^0 mixing analyses perform a final calibration on the scale factors using both B^+ and B^0 decays, as control samples. For the B_s^0 mixing analysis the $B^+ \rightarrow \bar{D}^0(3)\pi$ and $B^0 \rightarrow D^-(3)\pi$ decay modes have been used. Moreover, at the time this calibration was performed for the dataset `xbhd0i` corresponding to the last $\simeq 260$ pb^{-1} the $B^0 \rightarrow D^-3\pi$ was not available and only the remaining three $B^0 \rightarrow D^-\pi^+$ modes were used.

In this section, the flavor tagging performances are evaluated, through a full re-calculation of opposite tag calibration transfer scale factors, $S_{\mathcal{D}}$. At the same time, the fit for B^0 mixing frequency has been accomplished.

The aim of a scale factor is, in fact, to validate the tagger per-event dilution: if a tag is correctly calibrated, the transfer scale factor should be consistent with unity. Hence, the transfer scale factor is a measure of how well the per-event dilutions are estimating the true event dilution on average. Moreover, if tagger power is reduced, and the tagger is not correctly estimating the per-event dilutions, the transfer scale factor will absorb the discrepancies. Once the tagger is calibrated with a transfer scale factor derived from data, it is ready to be used in the B_s mixing analysis.

In the current B_s^0 mixing analysis, we use a combined opposite tagging algorithm (COST) which takes input from all the above mentioned opposite side taggers and produces a decision if at least one of the taggers provides flavor information.

For each event tagged by the combined opposite side tagger, not only the decision of the tagger is available, but also the predicted dilution that depends upon the quality of the tag. Since the combined opposite side tagger has an efficiency of the order of 95%, it is possible to split up the study and analyze the three samples apart corresponding to the three data periods. All the results concerning the efficiency, the scale factor, the tagging effectiveness for the tagger and the value of Δm_d are quoted in Table 4.5.

Further tests on the combined opposite side tagging performances can be done, by trying to understand the impact of the different decay modes in the final fit. For each dataset, the contribution of the two B^+ decay modes has been evaluated, by fitting for the scale factor with only these two modes. After, the same fitting has been repeated, but with the other two (B^0) decay modes combined. The results are shown in Table 4.6. As foreseen, the B^+ decay modes weigh more in the evaluation of the scale factor since they comprise a much larger fraction of the sample. In particular, this effect appears more evident, by looking at the error associated with the scale factor. It's also important to remember that for the `xbhd0i` dataset, the $B^+ \rightarrow \bar{D}^0 3\pi$ has not been used.

In the very first version of this analysis [64] using the hierarchic combination of OST and 355 pb^{-1}, a series of systematic effects were evaluated. The studies performed showed the uncertainties on the scale factors to be statistically dominated.

Dataset	ϵ[%]	Scale factor [%]	$\epsilon\mathcal{D}^2$ [%]	Δm_d [ps^{-1}]
0d	0.948 ± 0.002	1.115 ± 0.027	1.838 ± 0.119	0.523 ± 0.040
0h	0.959 ± 0.001	1.077 ± 0.052	1.831 ± 0.204	0.533 ± 0.037
0i	0.961 ± 0.002	1.040 ± 0.090	1.730 ± 0.336	0.504 ± 0.051

TABLE 4.5: Combined opposite side tagger performance for the three datasets, using event-by-event dilution.

Dataset	Scale factor (B^+)[%]	Scale factor (B^0)[%]	Scale factor ($B^+ + B^0$)
0d	1.115 ± 0.028	1.116 ± 0.126	1.115 ± 0.027
0h	1.088 ± 0.059	1.038 ± 0.109	1.077 ± 0.052
0i	1.114 ± 0.108	0.902 ± 0.153	1.040 ± 0.090

TABLE 4.6: Combined opposite side tagger performances for the $B^+ \to \bar{D}^0(3)\pi$ and $B^0 \to D^-(3)\pi^+$ separated. The last column quotes the final scale factors for the combination of all the modes. It is important to remember that only the $B^+ \to \bar{D}^0\pi^+$ has been used for the xbhd0i dataset.

The biggest source of systematic uncertainty is the potential correlation between the different opposite side taggers and the same side one (SSπT). Since the complete evaluation of the calibration scale factor is not part of the B_s^0 mixing analysis, these studies have not been repeated. Anyway, we are using a new tagger and so a good point is to understand the possible correlations between COST and SSπT. The likelihood construction for the combination of the two taggers is, in fact, predicated on the assumption that their decisions are uncorrelated. If that were not the case, the respective dilution would be biased. To address the potential correlation, we run the fit with SST+COST and again with only COST, combining the two B^+ decay modes, since they give us the main contribution to the scale factor fit (Table 4.6). We examine the shift in dilution scale factors between the two fits and take those values as a conservative estimate of possible correlation effects. The negligible change estimates are entered in Table 4.7.

Finally, the influence of the COST on the SSπT has been studied. The idea is the same as above. We compare the results for the SST Average Dilution coming out from the two B^+ decay modes fit with only SST activated, and the one with both COST and SST. The results are compiled in Table 4.8. It has to be emphasized that, while for the COST we use the parametrized dilution and introduce the relative scale factor, for the SSπT we use only the average one, because a parametrization has not been established.

4.3 Same Side Flavor Tagging

The Same side flavor tagging algorithm (SST) developed at CDF attempts to assess the flavor of the B meson at its birth, by looking at the associated particles produced around the meson during its hadronization process. In details, the algorithm infers

Dataset	Scale factor (COST)[%]	Scale factor (COST+SST)[%]	Relative Change [%]
0d	1.115 ± 0.028	1.115 ± 0.027	0.06
0h	1.088 ± 0.059	1.091 ± 0.058	0.28
0i	1.114 ± 0.108	1.107 ± 0.108	0.69

TABLE 4.7: Combined opposite side tagger performances without or with the same side tagging (SSπT) activated for the B^+ decay modes, combined together. The last column reports the relative change in the scale factor for the combined opposite side tagger due to the correlation with the SSπT.

Dataset	SST avg D (SST)[%]	SST avg D (SST+COST)[%]	Relative Change [%]
0d	0.260 ± 0.014	0.260 ± 0.014	0.04
0h	0.239 ± 0.012	0.238 ± 0.012	0.23
0i	0.220 ± 0.019	0.219 ± 0.019	0.28

TABLE 4.8: Same Side Pion Tagger performances without or with the Combined Opposite Side Tagger activated for the B^+ decay modes, combined together. The last column reports the relative change in the average dilution for the SSπT due to the correlation with the COST.

the flavor based on the correlation between the charge of the fragmentation tracks and the b quark [65].

A B_s^0 meson, which is $b\bar{s}$ bound state, is produced when a $s\bar{s}$ pair is pulled out of the vacuum in the proximity of the b quark. The left over s quark can contribute to the formation of a kaon. Thus for the B_s^0 mesons, due to their strangeness, strange particle can be good tagging particles. In the case in which the s quark left at the end of the fragmentation chain forms a charged kaon, its charge indicates the b flavor: K^- kaons follow $\overline{B}_s^0$ mesons, while K^+ kaons are associated with B_s^0 meson. The process is pictorially described in Figure 4.12.

The algorithm is also know as Same Side Kaon Tagging (SSKT) [66], because the strange particles, kaons, are the best tagging particles for the B_s^0. From the hadronization process just described it is clear how the tagger performances depend on the particular B species analyzed. This peculiarity raises the relevant issue of not having a straightforward way to measure the tagger dilution on data. This could be possible only if the B_s^0 oscillations were observed and the dilution treated as a free parameter of the likelihood and, then, fitted from the data. Since it is not known *a priori* if the data samples utilized have enough sensitivity to observe the oscillation, the study and parametrization of the tagging dilution and its calibration is achieved on a simulated events from Monte Carlo.

On the other hand, the same side tagger are expected to have good efficiency because the track which carries the flavor information is in the vicinity of the B_s^0 meson candidate, therefore it has an high chance to end up in the geometrical acceptance of the detector.

In the following session we introduce the different algorithms developed for the same-side tagging which lead to the conclusive algorithm used in the B_s^0 mixing anal-

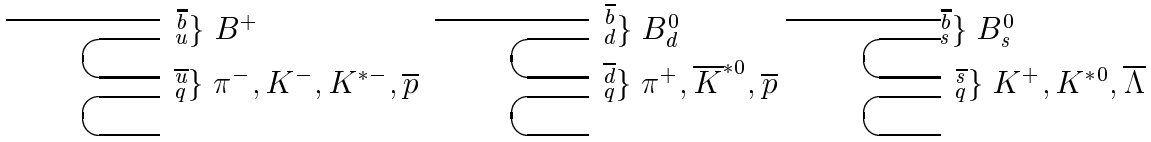

FIGURE 4.12: Sketch of fragmentation process for the particle produced in association with different B mesons.

b-quark	b-quark
$B^+ \to J/\psi K^+,\ J/\psi \to \mu^+\mu^-$	$\overline{B}^- \to J/\psi K^-,\ J/\psi \to \mu^+\mu^-$
$B^0 \to J/\psi K^{*0},\ J/\psi \to \mu^+\mu^-,\ K^{*0}$	$\overline{B}^0 \to J/\psi \bar{K}^{*0},\ J/\psi \to \mu^+\mu^-,\ \bar{K}^{*0}$
$B^0_s \to J/\psi\phi,\ J/\psi \to \mu^+\mu^-,\ \phi \to K^+K^-$	$\overline{B}^0_s \to J/\psi\phi,\ J/\psi \to \mu^+\mu^-,\ \phi \to K^+K^-$
$B^+ \to \bar{D}^0\pi^+,\ D^0 \to K^-\pi^+$	$\overline{B}^- \to D^0\pi^-,\ \bar{D}^0 \to K^+\pi^-$
$B^0 \to D^-\pi^+,\ D^- \to K^+\pi^-\pi^-$	$\overline{B}^0 \to D^+\pi^-,\ D^+ \to K^-\pi^+\pi^+$
$B^0_s \to D^-_s\pi^+,\ D^-_s \to \phi\pi^-,\ \phi \to K^+K^-$	$\overline{B}^0_s \to D^+_s\pi^-,\ D^+_s \to \phi\pi^+,\ \phi \to K^+K^-$

TABLE 4.9: Trigger side topologies of the 12 Monte Carlo samples.

ysis and in the angular analysis for the CP violation phase, β_s, measurement. This algorithm provides most of the total flavor tagging power available to the analyses, being $2-3$ times more powerful than the other available tagging algorithms.

4.3.1 Data and Monte Carlo Samples

The SSKT study has been done using good runs of an integrated luminosity of 355 pb^{-1}. The same reconstruction cuts have been applied for data and MC and the primary vertex is determined on an event-by-event basis.

One big `PYTHIA msel`=1 generation level Monte Carlo sample was produced, i.e., the following QCD high-$p_\perp$ processes where generated: $f_i f_j \to f_i f_j$, $f_i \bar{f}_i \to f_k \bar{f}_k$, $f_i \bar{f}_i \to gg$, $f_i g \to f_i g$, $gg \to f_k \bar{f}_k$ and $gg \to gg$, where $f_{i,j,k}$ are fermions and g gluons. The generated samples are then passed through a fully detector simulation. The underlying processes from the original sample are kept, but the particles are re-decayed using `EvtGen` [55]. This procedure was necessary to be able to produce large Monte Carlo samples in a reasonable amount of time. A set of twelve simulated samples were produced, which are summarized in Table 4.9.

The B of the event is forced to decay in a specified way, but the rest of the event are unbiased (`EvtGen` default tables). The B neutral meson mixing has not been simulated. This important feature is to be discussed later when describing the calibration procedure.

4.3.2 Selection of Flavor Tagging Candidates

Several are the SSKT algorithms studied to select the leading fragmentation track carrying the flavor information. The implementations of these algorithms share the

same initial selection on the possible B tagging tracks. All tracks which fulfill the following reconstruction requirements are considered as potential tagging candidates:

- number of Silicon ϕ hits ≥ 3,
- number of COT hits ≥ 10,
- $p_T \geq 450$ MeV/c,
- $|\eta| \leq 1$,
- $\Delta R(\text{track}, B) \leq 0.7$,
- $|\Delta z_0(\text{track}, B)| \leq 1.2$ cm,
- $|d_0/\sigma_{d_0}| \leq 4$,
- rejection of e,μ and conversions,
- rejection of B daughter tracks.

The requirements on the number of hits in the silicon and COT detectors select candidates with a reliable track fit and enforce the tracks to pass through the central region of the CDF detector. The minimum p_T cut is chosen to avoid the charge asymmetry for low momentum tracks in CDF detector. This is caused by the COT design: the cells are tilted, with respect to the radius which connects a cell to the center of the detector (see Figure [**?**]), thus the cell section appears different to positively and negatively charged particles, which translates in a different tracking efficiency. Additionally, it will be observed later that events tagged by low momentum tracks have weaker tagging performance than events tagged by high momentum tracks. The cut on the modulo of the pseudorapidity η is correlated to the ΔR cut and the COT and silicon hits requirements to establish a set of requirements which restrict the tagging track candidates to the central region. There are about 10% additional tracks above $|\eta| = 1$, but they hardly have any TOF information and their dE/dx response is purely determined. Thus they are potentially very low quality tracks for the tagging, in particular for the algorithms exploiting the PID information, so they have been removed.

The ΔR cut has been chosen to exclude as much as possible opposite side B daughter tracks, keeping at the same time most of the same side fragmentation tracks. As the opposite side taggers use tracks which are outside a cone of $\Delta R \geq 0.7$ we had to chose as maximum $\Delta R \leq 0.7$. With this cut we still have about 5% tagging track candidates originating from the opposite side B (fragmentation or decay), but their fraction among the selected tagging tracks is smaller than 0.5%. The purpose of the Δz_0 cut is to reject traks coming from multiple $p\bar{p}$ interactions different from the one produced in the B_s^0 candidate reconstruction, the so-called pile-up events. The impact parameter significance cut, $|d_0/\sigma_{d_0}| \leq 4$, is applied to select tracks consistent with coming from the primary vertex.

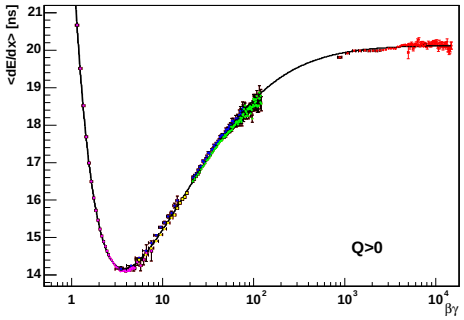

FIGURE 4.13: Average dE/dx versus particle $\beta\gamma$ as measured for various particle species in CDF. The plot is produced utilizing only positively charged particles. From left to right, it is possible to distinguish the contributions of the calibration samples of protons, kaons and pions.

The rejection cuts are meant to remove particles which are not explicitly coming from the fragmentation process and are likely to have been produced by muons or electrons or from $\gamma \to e^+e^-$ conversion. Finally, tracks used in the fit for B_s^0 candidate have to be clearly removed from the list of possible tagging tracks.

Figures 4.14 show data-Monte Carlo comparisons for some of the track variables used to select the tagging candidates. A general very good agreement is found.

Once these selection cuts are applied, we can have zero, one or more tagging candidates. If no extra cuts are applied, the number of events with zero candidates defines the efficiency ϵ of the same side tagger. When one or multiple tagging candidates are accepted, the events can be naturally divided - and treated - in two groups:

1. events either with one selected tagging candidate or with multiple tag candidates having identical charge: in this case no tagging decision has to be made,

2. events with the more than one tagging candidate differing in their charges: in this case we need a non trivial way to decide which charge corresponds to the same side tagger.

Various algorithms have been studied, which differ from the criteria used to select the tagging track among the possible candidates. The SSKT utilized in the final implementation of the B_s^0 mixing analysis and for the angular analysis, both presented in this dissertation, uses a NN approach to combine the kinematic and particle-identification information [66]. I collaborated, in my contribution to the B_s^0 mixing effort done by CDF, to the development and validation of the SSKT kinematic-based (Section [**?**]) and PID-based (Section [**?**]) algorithms which were more accurately studied. This PID-based algorithm, in particular, was the one used in the analysis that resulted in the first Δm_s measurement [14].

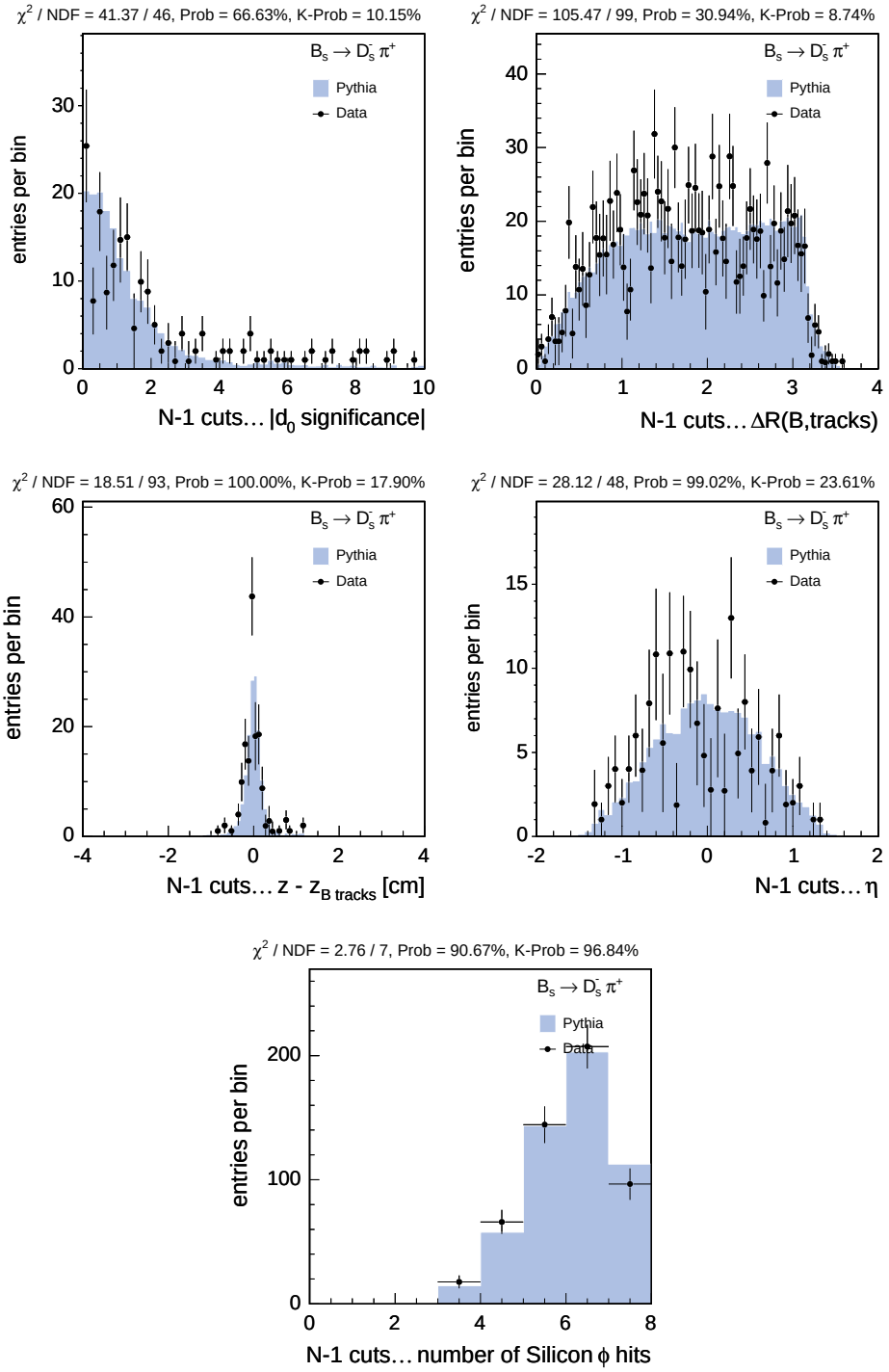

FIGURE 4.14: Data-PYTHYA-MC-simulation of track variables for the decay mode $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$. From left to right, and top to bottom the following distributions are shown: impact parameter significance $|d_0/\sigma_{d_0}|$, $\Delta R(\text{track}, B)$, $\Delta z_0(\text{track}, B)$, pseudorapidity $|\eta|$ and number of hits in the silicon detectors. Each distribution is plotted using tracks with satisfy all the criteria to be a tag candidate expect for the variable shown.

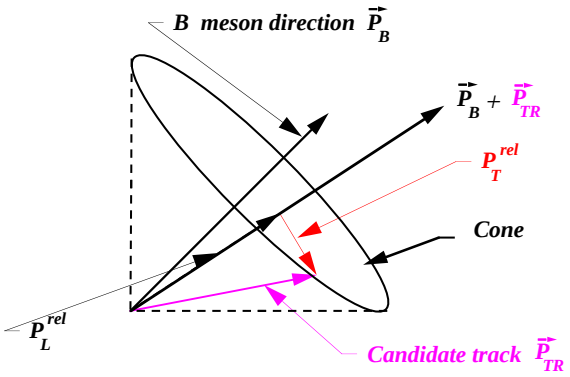

FIGURE 4.15: Graphical definition of p_L^{rel}.

4.3.3 Kinematic Same Side Kaon Tagger

Since the flavor tagging candidate is expected to be close to the vicinity of the B meson, the kinematics of the leading fragmentation tracks are supposedly correlated with those of the B candidate. Exploiting such correlations, several algorithms can be built to select the tagging candidate. Among all the possible methods developed, the most performing algorithm found, selects as flavor tag track, the one with the maximum p_L^{rel}, therefore it is also referred as *max* p_L^{rel}. The variable p_L^{rel} is pictorially defined in Figure 4.15.

In Figure 4.16 we report the candidate-by-candidate dilution parametrized as a function of the transverse momentum p_T of the tag candidate. Two independent parametrization are available, according to whether all the tags have the same charge (agreement case) or not (disagreement case). When applied on the Pythia-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$, the *effective* dilution $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle}$ obtained is:

$$S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 22.8 \pm 0.7(stat)\% \tag{4.3.1}$$

The error is statistical only. The effective dilution is defined as the square root of the average square candidate-by-candidate dilution for each candidate in the sample, multiplied by the scale factor $S_{\mathcal{D}}$ obtained by the calibration process. In general the figure of merit to classify a tagging algorithm is the tagging effectiveness $\epsilon\mathcal{D}^2$, but since the tag candidate sample is identical, independently from the algorithm studied and validated the effective dilution is sufficient to classify a tagger in order of performance.

4.3.4 Particle Identification Same Side Kaon Tagger

The fragmentation process for B_s^0, graphically described in Figure 4.12, foresees kaons to be likely produced around the meson and be the particles carrying the flavor information. The challenge in an hadronic collider, such as Tevatron, comes from the fact that most of the prompt tracks produced in $p\bar{p}$ collisions are pions. To separate kaons from other particles, mainly pions, a possibility is to develop a tagger which algorithm is based on the particle-identification information which may be available

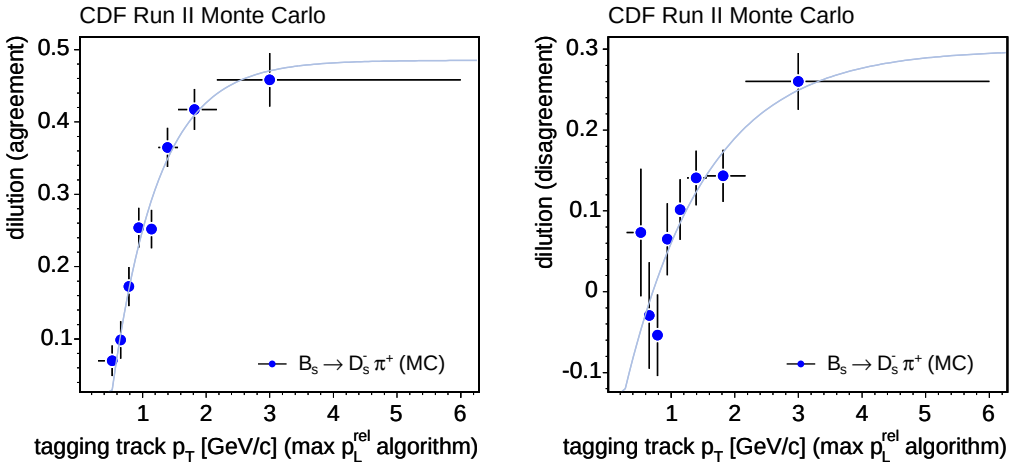

FIGURE 4.16: Dilution parametrization as a function of transverse momentum p_T of the flavor tagging candidate, for the *max* p_L^{rel} algorithm. The plot on the left contains all the cases when only on tagging candidate or more than one tagging candidate with the same charge are available. The right plot contains all the cases when more than one tagging candidate is available but they disagree in charge.

for the tracks. Particle identification information for charged tracks are provided by the COT and the TOF sub-detectors. The descriptions of these detectors are presented in Sections [**?**] and [**?**].

For the COT sub-detector, specific energy loss per unit length, referred as dE/dx is correlated with the type of particle considered. The measured dE/dx, once calibrated [], is used to evaluate the variable Z, defined as:

$$Z(i) \equiv log\Big[\frac{(dE/dx)^{cor}}{(dE/dx)^{pre}(i)}\Big], i = \pi, K, p, \tag{4.3.2}$$

where $(dE/dx)^{cor}$ represents the calibrated dE/dx obtained from the data and $(dE/dx)^{pre}$ refers to the predicted expectation obtained from the universal curve in Figure 4.13, where dE/dx is expressed as a function of the particle speed. The probability of a charged particle tested over the different hypotheses is extracted comparing their $Z(i)$ with the distributions of the $Z(i)$ variables obtained in sample of pure pions, kaons and protons.

The TOF detector measures the t_{flight} of a particles. The probability for the tested hypotheses is assessed utilizing the residual time of flight, defined as:

$$\Delta t_{flight} \equiv t_{flight}^{meas} - t_{flight}^{pre}, \tag{4.3.3}$$

where the predicted time of flight, t_{flight}^{pre}, is a function of the particle mass m, its momentum, p, and the lenght travelled before reaching the TOF detector:

$$t^{pre}_{flight} = \frac{L}{pc}\sqrt{p^2 + m^2c^2} \tag{4.3.4}$$

The two sub-detector measurements are complementary: COT provides 1.4 standard deviations separation between kaons and pions for particles with $p_T > 2$ GeV/c, while TOF has a separation power of 2 standard deviations for particles with $p_T < 1.5$ GeV/c. Although the TOF efficiency is low, about 65%, a large fraction of the tag candidate have a transverse momentum less than 1.5 GeV/c which makes the TOF contribution more important than the COT one.

An optimized separation power is obtained combining the information:

$$\mathcal{L}(i) = \mathcal{P}_{TOF}(i) \cdot \mathcal{P}_{COT}(i), \quad i = \pi, K, p. \tag{4.3.5}$$

For the PID-based tagger the information are combined in the following single variable, named CLL and defined as:

$$CLL = log\big(\frac{\mathcal{L}(K)}{f_p\mathcal{L}(p) + f_\pi\mathcal{L}(\pi)}\big), \tag{4.3.6}$$

where $f_p = 0.1$ and $f_\pi = 0.9$ are the prior probabilities for background composition. In the case the information from TOF or dE/dx are not available the respective probabilities are set to one. Figure 4.17 shows a representative example of the CLL distributions in data, for the first 355 pb^{-1} of integrated luminosity, and in PYTHIA-MC events, where the MC truth information is used to distinguish the pions, kaons and protons contributions.

From the CLL definition in Equation 4.3.6, it is implied that the higher the CLL value, the more probable the tag candidate is a kaon. Thus the natural choice for the tagging algorithm is to select the track which has the maximum CLL available, and it consequentially called the *max CLL algorithm*. As for the kinematic-based algorithm, we distinguish two scenarios when parametrizing the dilution on a event-by-event fashion, according to whether the tag candidates have the same charge or not. The dilutions for the separated cases are parametrized independently as function of CLL. Their functional forms are drawn in Figure 4.18.

The performance of the max CLL algorithm, quoted in terms of effective dilution on a PYTHYA-MC sample of B^0_s mesons decaying as $B^0_s \to D^-_s\pi^+$, $D^-_s \to \phi\pi^-$, is:

$$S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle} = 28.5 \pm 0.7(stat)\%, \tag{4.3.7}$$

which compared to the equivalent result on the kinematic-based algorithm, the max p^{rel}_L (Equation), shows a better performance: the particle-identification-based same side flavor algorithm was utilized in the analysis of the $B^0_s - \overline{B}^0_s$ oscillation presented in Reference [14].

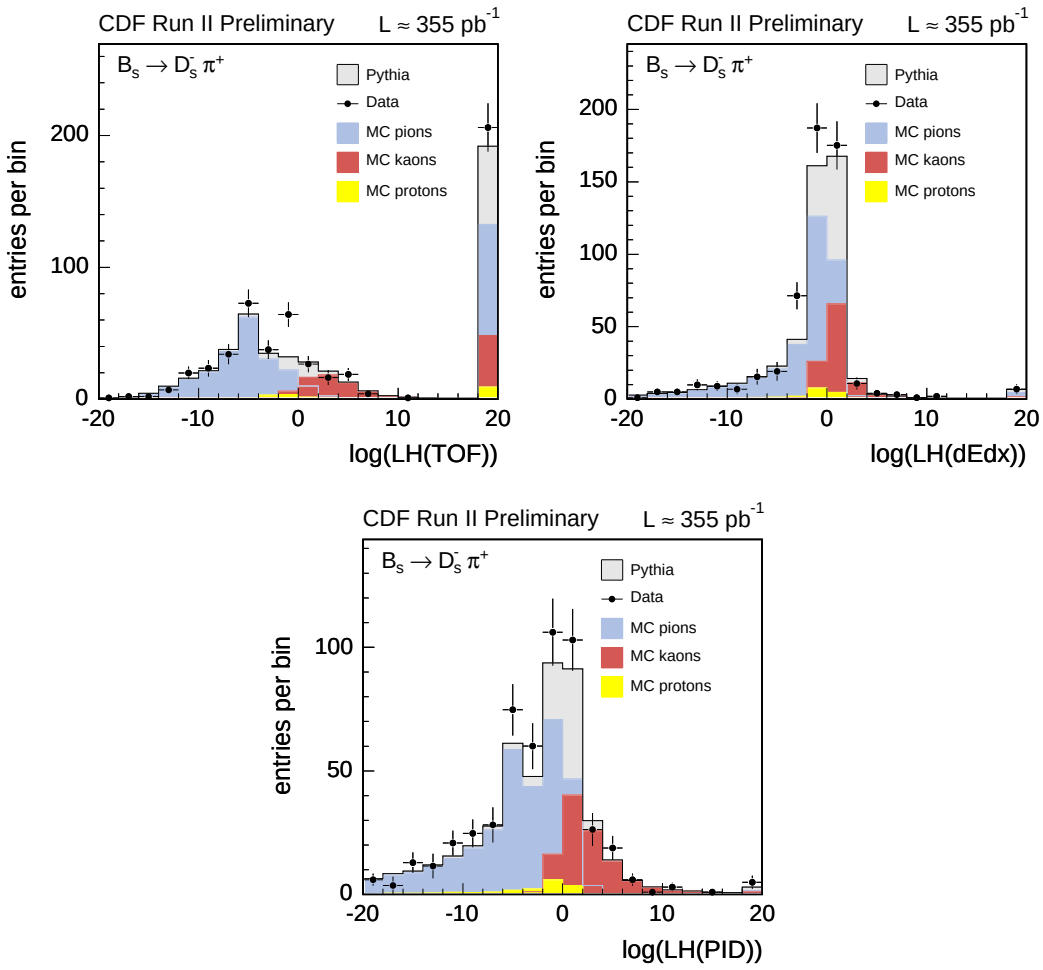

FIGURE 4.17: Distributions of CLL for tagging tracks (dots) and PYTHYA-MC events (histograms) for the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ decay mode. The distributions are for the TOF unique contribution (upper left), the COT unique contribution (upper right) and their final combination (bottom). The rightmost bin contains events in which either the t_{flight} or the dE/dx information is not available.

4.3.5 Neural Network Same Side Kaon Tagger

A natural extension of the selection tagging algorithms presented in the previous sections is to consider, in same way, their combination. The use of an Adaptive Neural Network provides the typical tool to perform such combination. The same ROOT-SNNS package used for the hadronic B_s^0 as well as for the $B_s^0 \to J/\psi\phi$ sample selections has been implemented. The ANN inputs are the CLL and various kinematic quantities, p_T, p_T^{rel}, p_L^{rel} and ΔR. In addition to these quantities it is also used a Boolean variable, set to true when all tag candidates have same charge and set to false otherwise. The PYTHYA-MC sample which contains $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ decay mode is used as signal sample for the training in order to select kaons with the correct charge correlation with the B_s^0 candidate flavor. As background, the

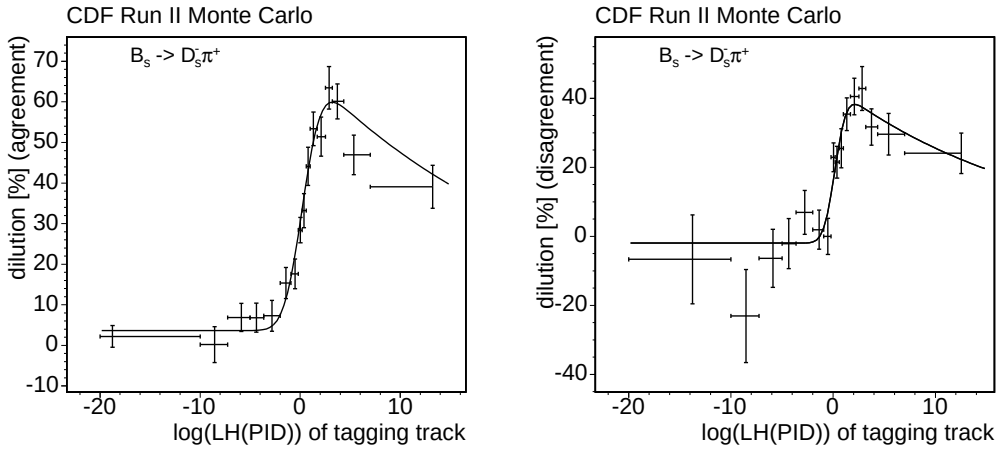

FIGURE 4.18: Dilution parametrization as a function of CLL of the flavor tagging candidate, for the max CLL algorithm. The plot on the left contains all the cases when only on tagging candidate or more than one tagging candidate with the same charge are available. The right plot contains all the cases when more than one tagging candidate is available but they disagree in charge.

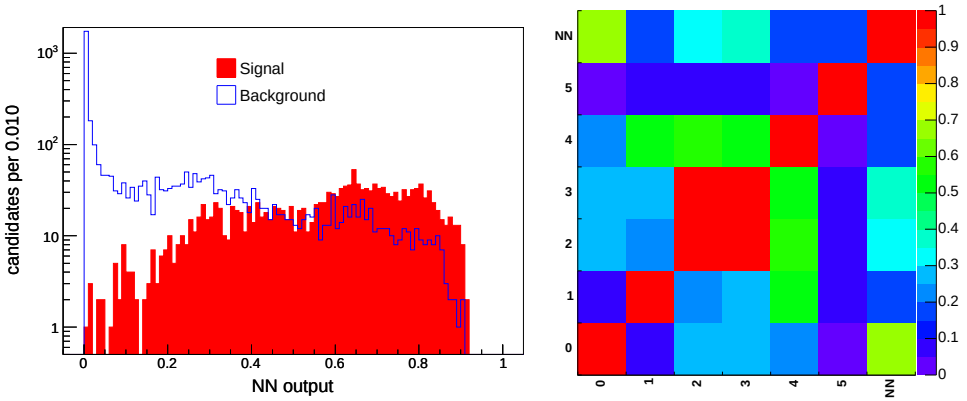

FIGURE 4.19: Neural Network distributions. Left: distribution of the ANN output for signal and background. Right: distribution of the correlations among the input variables; the number from 0 to 6 identify CLL, p_T, p_T^{rel}, p_L^{rel}, ΔR and the Boolean variable. The training is performed on PYTHYA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ decay mode candidates.

sample containing tag candidates which are pions, protons or kaons with the wrong correlation is used. The output of the neural network is maximized for the tag candidate which are kaons with the correct charge correlation. The distribution of the output for signal and background and the distribution of the correlations among the input variables are shown in Figure 4.19.

In order to train the Neural Network and expect a reliable response, the data-MC agreement for the kinematic and particle-ID related variables was verified for the tag candidates before and it showed a general good agreement. An example of these

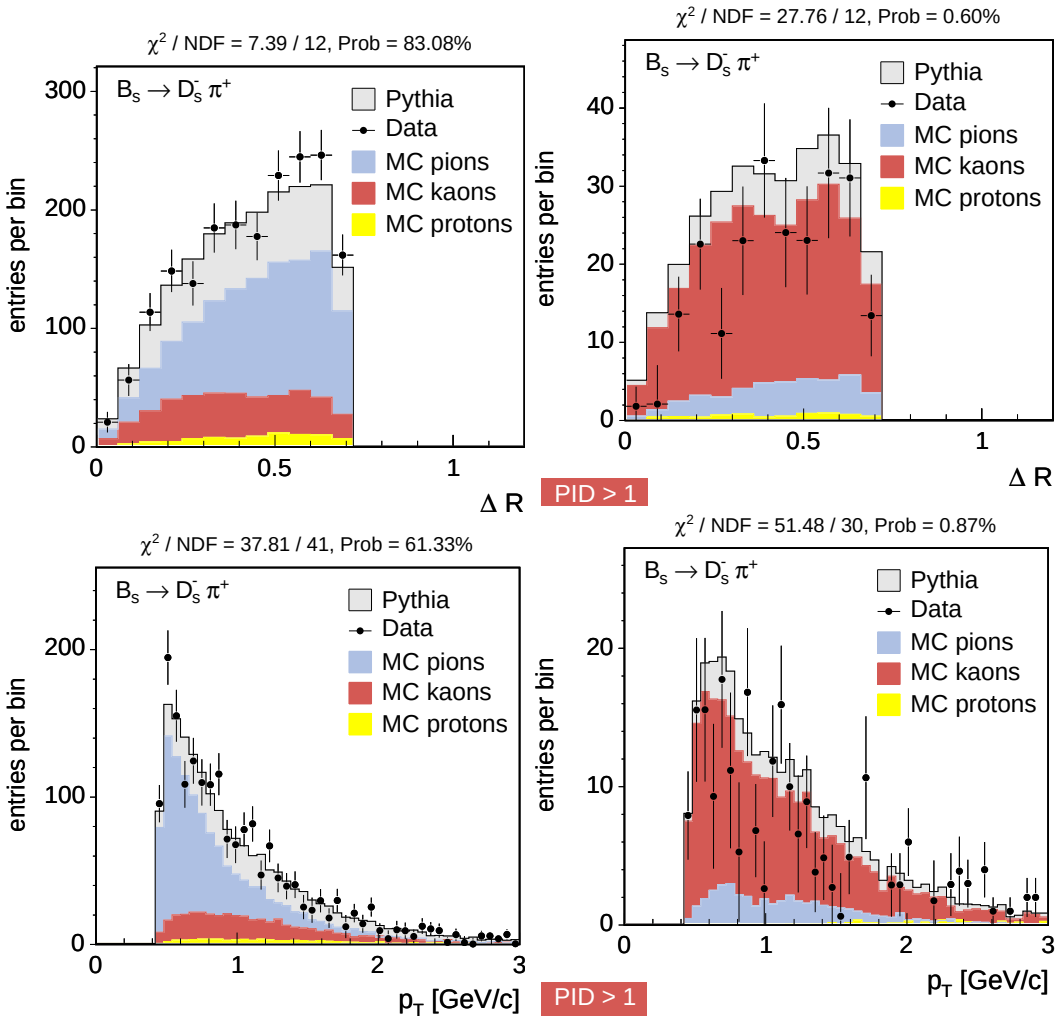

FIGURE 4.20: Distributions for ΔR (top) and p_T (bottom) of the tag candidates for the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ decay mode sample. The cut $CLL > 1$ applied on the distributions on the right side enriches the sample in kaons, as clearly shown by the separation of the PYTHYA-MC sample by particle species.

distributions is reported in Figure 4.20

The decision of the flavor tagger is given by the tag candidate which maximize the neural network output. The event-by-event dilution, parametrized as a function of the ANN output, is drawn in Figure 4.21

The effective dilution, calculated in a PYTHYA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$ is:

$$S_{\mathcal{D}} \sqrt{\langle \mathcal{D}^2 \rangle} = 30.2 \pm 0.7(stat)\%, \tag{4.3.8}$$

The error is statistical only. The performance of the ANN-based tagger is slightly better than the particle-identification tagger only. Nonetheless, the improvement with respect to the previous tagging algorithm is statistically significant if we consider that the same sample has been used to extract the performances of the different algorithms

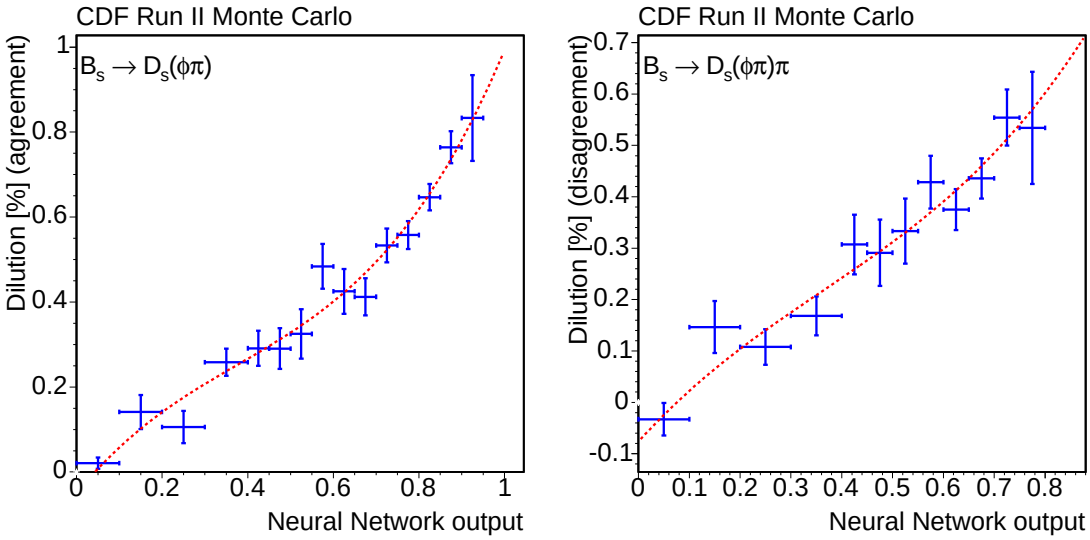

FIGURE 4.21: Dilution parametrization as a function of the Neural Network output of the flavor tagging candidate. The plot on the left contains all the cases when only on tagging candidate or more than one tagging candidate with the same charge are available. The right plot contains all the cases when more than one tagging candidate is available but they disagree in charge.

presented.

4.3.6 Calibration of the Same Side Kaon Tagger

The calibration of the Opposite Side Tagging was performed on data samples of B^+ and B^0 decay modes collected with the same trigger paths as the B_s^0 signal sample. This procedure was possible because the tagger was independent from the B species whose flavor had to be identified. On the other hand, the SST, by looking at the charge correlation between the tracks produced in the vicinity of the B meson considered, are expected to perform differently in the case of B^+, B^0 and B_s^0. For instance the SSKT performs better on B^+ with respect to B^0. Figure 4.12 shows pictorially that the correlations between the flavor of the B^+ meson and the same side kaons is the same as with same-side pions, while it is the opposite case for B^0. Since it is not known if the sample used for the analysis would have the sensitivity to the B_s^0 mixing oscillation measurement it is not possible to extract the calibration directly with the final fit. Moreover a correct calibration is even more crucial in the case there will be no such sensitivity and a limit should be set. The calibration is then performed on the PYTHYA-MC simulated sample. As usual, the calibration consists in a simultaneous fit of the mass and the proper decay time of the reconstructed B_s^0 candidates.
We remind that the PYTHYA-MC is generated without mixing which is equivalent to set $\Delta m_s = 0$. The Equation 4.1.4 would then become:

$$\mathcal{P}_{mix/unmix}(t) \propto \frac{1}{2}\left[1 \pm \mathcal{D}\cos(\Delta m t)\right] \to \frac{1}{2}\left[1 \pm S_{\mathcal{D}}\mathcal{D}\right]. \tag{4.3.9}$$

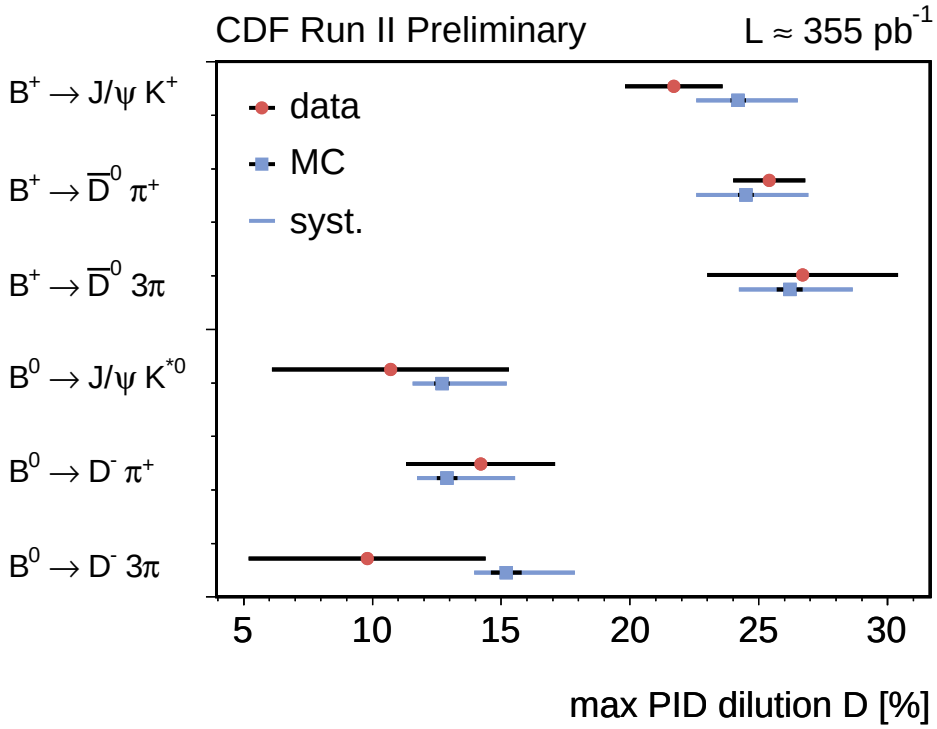

FIGURE 4.22: Data-MC dilution comparison for the max CLL algorithm applied to the B^+ and B^0 control sample and their relative PYTHYA generated samples. Data points are quoted with statistical error only, while the MC points have statistical and systematics errors.

The oscillation frequency is equal to zero by construction and the scale factor $S_{\mathcal{D}}$ is now the free parameter of the fit to be determined. The most important cross check of the validity of the procedure is to compare the results obtained on the B^+ and B^0 data with the respective MC samples generated with PYTHYA. This cross-check was performed in the first implementation of the the same side kaon tagging in the B_s^0 mixing analysis [14] when the PID-based tagger was used. Figure 4.22 shows the comparison of dilutions measured in data, with statistical uncertainty, and in the respective MC, with systematic uncertainties applied, for the B^+ and B^0 control samples.

4.3.7 Systematic Uncertainties

Systematic uncertainties are evaluated for SSKT dilution to account for the fact that its calibration relies on MC simulation. All systematics have been evaluated using the same MC sample but re-weighting MC events in order to simulate the effect under study which is suspected to contribute to the total systematic uncertainty. The extensive study of these systematics is treated in References [66] and [67]. The list of the analyzed systematics is:

Data Sample		$\sigma_{S_D}[\%]$
0d	$+\Sigma$	10.7
	$-\Sigma$	14.3
0h	$+\Sigma$	10.8
	$-\Sigma$	14.4
0i	$+\Sigma$	10.8
	$-\Sigma$	14.4

TABLE 4.10: Total systematic uncertainties for the ANN-based algorithm for the Same Side Kaon Tagging calculated on the PYTHYA-MC sample of $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-$ candidates.

- **b-production mechanism**, to account for the current uncertainty on the fraction of different processes contributing to the $b\bar{b}$ production: flavor creation, flavor excitation and gluon splitting.
- **Fragmentation process**, to test several different hypotheses and variation of the default fragmentation function chosen for the MC generation.
- **Particle-identification simulation**, varying the simulated TOF and COT dE/dx efficiency and resolution according to measured uncertainties in data.
- B^{**} **rates**, to account for different fractions of B meson candidates originating from the B^{**} with respect to the default 20%, chosen accordingly to the recent LEP measurements [68].
- **Multiple interactions** to study the tagging performances by varying the rate of additional potential tagging tracks which in turns varies with the increasing luminosity.
- **Data-MC agreement** which is very good in the variables studied. On the other hand this statement cannot be made more precise than our uncertainties on the data and MC samples. Therefore combined uncertainties from the higher statistic samples B^+ and B^0 have been added to the final result.

Figure 4.22 reports the data-Monte Carlo agreement including statistical and systematic uncertainties for the max CLL algorithm. There is a very good agreement in all modes which make us confident of the proper parametrization of the dilution and its associated uncertainty. It is important to stress that these uncertainties are important for the B_s^0 oscillation amplitude scan but not for the final Δm_s measurement. The total systematics uncertainties for the three data period considered are reported in Table 4.10.

4.3.8 Final Scale Factors

The performance of a tagging algorithm are described by its efficiency and dilution. In the previous Sections the dilution parametrization and the relative MC-

	max CLL SSKT [%]	ANN-based SSKT [%]
0d $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	3.9 ± 0.7	4.2 ± 0.7
0h $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	3.1 ± 0.5	2.9 ± 0.5
0i $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	3.3 ± 0.7	3.5 ± 0.7

TABLE 4.11: Tagging effectiveness comparison between the max CLL and ANN-based Same Side Kaon Tagging algorithms, as obtained from the PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-$ candidates.

based calibration has been summarized for the different tested algorithms. The systematic uncertainties which affect the scale factor are listed in Section **??**. The scale factors for the ANN-based algorithm extracted from a PYTHYA-MC sample of $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-$ candidates for the separated data taking periods are:

$$\begin{aligned} S_{\mathcal{D}}(\texttt{0d}; B_s^0 \to D_s^- \pi^+) &= 99.2^{+10.7}_{-14.3}\% \\ S_{\mathcal{D}}(\texttt{0h}; B_s^0 \to D_s^- \pi^+) &= 95.9^{+10.8}_{-14.4}\% \\ S_{\mathcal{D}}(\texttt{0i}; B_s^0 \to D_s^- \pi^+) &= 95.0^{+10.8}_{-14.4}\% \end{aligned} \tag{4.3.10}$$

where the errors are statistical and systematic. The MC used for the ANN-based same side tagging algorithm as well as the others is tuned to simulate the 0d period of data taking. The improvement in terms of tagging effectiveness is observed along 0i period of data taking while the reduced performances in 0h are expected to be given by statistical fluctuation. Table 4.11 reports the comparison in terms of $\epsilon\mathcal{D}^2$ between the max CLL and ANN-based algorithm as obtained from the PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-$ candidates.

4.4 Elements for the $B_s^0 \to J/\psi\phi$ Angular Analysis

This section presents the description of the elements of the $B_s^0 \to J/\psi\phi$ angular analysis which resulted in the first flavor tagged measurement of the CP violation phase β_s.

While in the $B_s^0 \to J/\psi\phi$ angular analysis an equivalent of the sensitivity Formula 4.1.9 for the B_s^0 analysis does not exist, we have seen at the beginning of the chapter how signals involving fast oscillations (which are present in B_s^0 decays) are highly susceptible to damping due to decay time resolution effects. One can roughly approximate the damping with a factor [57]:

$$\mathcal{A}' = \mathcal{A} \cdot e^{-\sigma_t^2 \Delta m_s^2/2} \tag{4.4.1}$$

where σ_t is the decay time resolution and Δm_s is the oscillation frequency.

The decay time resolutions used as input to this analysis are obtained from CTVMFT [47] fits to the decay topology. Since the track parameter resolution estimates are not completely calibrated, one does not expect fully correct decay time

resolution estimates out of CTVMFT either. Therefore we proceed to the estimation of proper decay uncertainty utilizing the same technique explained in Section 4.1.3, evaluating the potential improvement in the angular analysis.

In the $B_s^0 \to J/\psi\phi$ angular analysis the tagger impact is not easily quantified, as in B_s^0 mixing searches were the sensitivity Formula determines univocally to power of the analysis. Nevertheless the tagging techniques are of exceptional importance for the measurement of width-difference $\Delta\Gamma_s$ and CP violating phase β_s and their calibration will be described.

4.4.1 Calibration of Proper Decay Time Resolution

In the next sections, we study the systematic behavior of decay time resolution correction factors as functions of different decay variables. We repeat the same procedure already explained in Section 4.1.3, parametrizing an overall scale factor correction that, later, we want to use in the final $B_s^0 \to J/\psi\phi$ angular analysis. Thus, we evaluate the potential improvement to the result significance, and conclude with an estimate of the expected remaining systematic uncertainties.

Selection for Decay Time Resolution Study

In order to study resolution effects, it is needed to have a sample of decays which are consistent with coming from the beam-spot, which we can use for the calibration. This section describes the sub-sample of decays which are used in the study. The main issue in selecting such sample is that to look for a event of prompt decays which are kinematically very similar to the real signal. The sidebands of the mass distribution are a good starting point, but they contain too few events. However, these are events that kinematically most resemble true signal events. On the other hand, we have attempted to maximally increase the statistics by using events in the sidebands of the mass distribution, but removing the neural network selection criterion and using only pre-selection requirements. We found that the output of this calibration was sub-optimal (the behavior of the sample with tight selection cuts was significantly different from that of the sample with pre-selection cuts). We settled on using events in the sidebands of the mass distribution, but applying a loose neural network selection requirement ($nnOut > 0.05$). This gives us a reasonably large sample of prompt decays ($\sim$200k events for B^0, and $\sim$ 50k events for B_s^0 sidebands). The quality of the calibration and the transfer from the loose neural network selection to the tight neural network selection is explained in Section 4.4.1.

$B^0 \to J/\psi K^{*0}$ Background Decays

The sample of $B^0 \to J/\psi K^{*0}$ background decays constitute a large calibration sample with similar decay properties to those of B_s^0 decays. We use this large calibration sample to study the behavior of the scale factor as a function of different decay variables, and pick empirical functional forms to be used in the final calibration. The

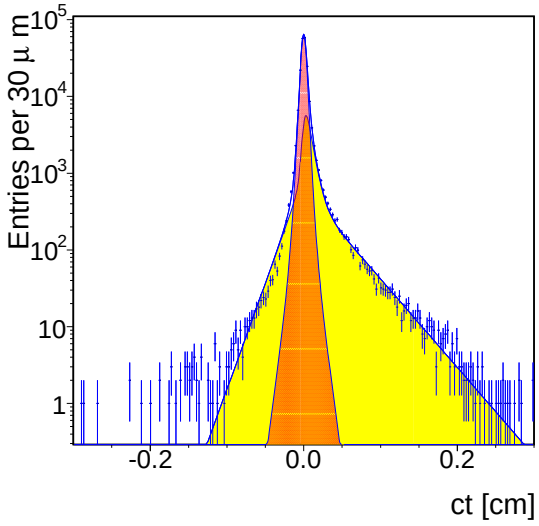

FIGURE 4.23: Projection of the likelihood into decay time space, for the calibration (sideband) sample for $B^0 \to J/\psi K^{*0}$.

actual parametrization of the scale factor, one wants to use in the final fit, will be derived using the $B_s^0 \to J/\psi\phi$ background sample.

The algorithm of the calibration is rather simple. Due to its large statistics, the sample is broken into several sub-samples in the variable that is being studied. For every bin, an unbinned likelihood fit for the decay time resolution scale factor is performed. The probability density function used in every bin is always the same:

$$\begin{aligned} p(t_i, \sigma_{t,i}|S_{ct}, f^+, f^-, c\tau^+, c\tau^-, mu) = \quad & (1 - f^+ - f^-) \cdot \mathcal{G}(t|\mu, S_{ct} \cdot \sigma_{t,i}) \\ & + f^+ \cdot G(t' - t|0, S_{ct} \cdot \sigma_{t,i}) \otimes \tfrac{1}{c\tau^+} e^{-t'/c\tau^+} \times \theta(t) \\ & + f^- \cdot G(t' - t|0, S_{ct} \cdot \sigma_{t,i}) \otimes \tfrac{1}{c\tau^-} e^{-t'/c\tau^-} \times \theta(-t) \end{aligned} \tag{4.4.2}$$

Figure 4.23 shows the projection of the likelihood into the decay time space for $B^0 \to J/\psi K^{*0}$. The crude model we are using here is good enough to describe the behavior of the narrow peak and the tails. In addition, the simple form above guarantees robustness when performing many fits in different bins. The fraction of events in the exponential tails is found to be low - around 4% of the decays participate in the long-lived positive tail and around 2% participate in the negative tail.

The dependencies of the uncalibrated scale factor on different decay variables are found and, using empirical parametrizations, we describe these behaviors reasonably well. After, we pick an order of variables and parametrize away one dependency at the time. The underlying assumption is that the dependencies in different variables will factorize. We test this assumption by repeating the study after the calibration is applied. If the effects do not factorize, there will be residual dependencies which

would need to be either calibrate with a second iteration or absorbed into a systematic uncertainty.

The chosen order of decay variables is $p_T(B)$, $\Delta R(B)$, isolation of the B meson, neural network output, and B meson pseudorapidity.

The parametrization of the scale factor correction for $B^0 \to J/\psi K^{*0}$ decays is given by:

$$\begin{aligned} S_{ct} = & (2.01791 \times 10^{-1} \cdot \log(p_T) + 0.921705) \times \\ & (-5.22451 \times 10^{-2} \cdot (\Delta R) + 9.94122) \times \\ & (1.01088 - 2.38921 \times 10^{-1} \cdot (iso - 0.6)) \times \\ & (-4.05768 \times 10^{-2} \cdot \log(nnOut) + 0.926753) \times \\ & (0.994482 + 3.28209 \times 10^{-2} \cdot (\eta - 0.4) + 9.16283 \times 10^{-2} \cdot (\eta - 0.4)^2) \end{aligned} \tag{4.4.3}$$

To validate the calibration, two tests are performed. First, we repeat the study of fitting in bins of different decay variables. No major dependencies are left after the calibration. Figure 4.24 shows an example of the proper time resolution dependency before and after the tuning for the $p_T(B)$ and the pseudorapidity η. All remaining dependencies are at the $\pm 4\%$ level.

Afterwards we test if the calibration behaves well in the sub-sample obtained by applying tight - same as used in signal selection - neural network selection cuts. The behavior as a function of the different decay variables after the tuning can be seen in Figures 4.25. Still, no significant dependencies are observed. The averaged scale factor is found to be 1.0001 ± 0.0024. We conclude that the large statistics sample has been successfully calibrated and we have vetted the calibration procedure. We now proceed to the sample of B_s^0 sideband decays.

$B_s^0 \to J/\psi\phi$ Background Decays

Having determined the functional shapes to be used in the scale factor parametrization on the high statistics sample, it is still preferred to use the B_s^0 background decays to obtain the final parametrization. The choice derives from the fact that the B_s^0 background events kinematically resemble the signal events even more than B^0 background decays do. The exact same procedure, as outlined in Section 4.4.1, is repeated. The comparison of examples of scale factor distributions before and after calibration are shown in Figure 4.26. After the calibration, we tighten the neural network cut to the signal selection value and repeat the decay time fit with the full statistics. The fit returns a ct scale factor of 0.997 ± 0.012, consistent with unit. We conclude that the calibration is successful at every step and infer that potential residual effects are at the order of $3 - 4\%$. The final parametrization of the scale factor correction is given by:

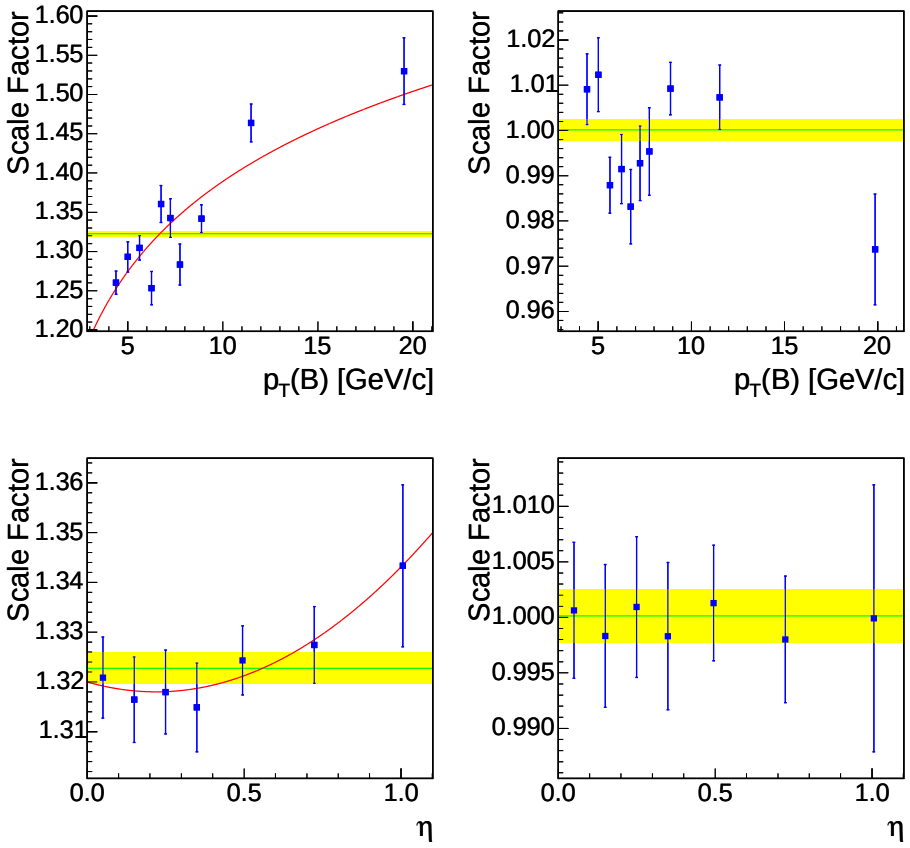

FIGURE 4.24: Proper decay time scale factor dependence on different decay variables, before (left) and after (right) tuning. In order, from top to bottom: $p_T(\mathrm{B}^0)$, $\eta(\mathrm{B}^0)$.

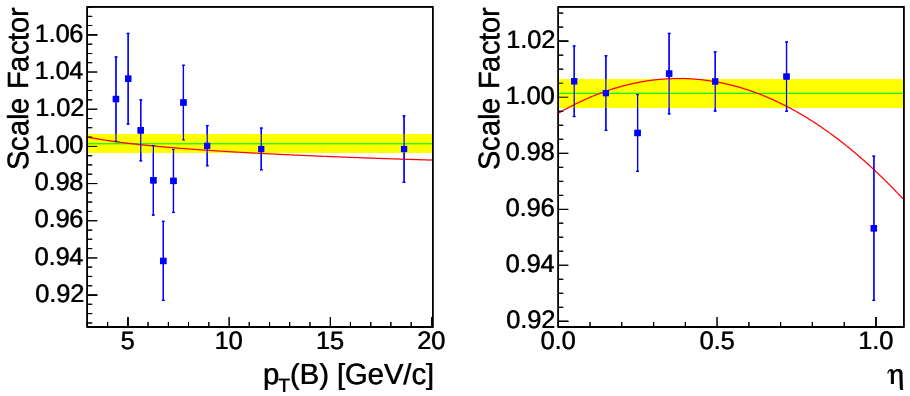

FIGURE 4.25: Proper decay time scale factor dependence on different decay variables, after tuning and when the tighter Neural Network selection cut is applied. In order: $p_T(\mathrm{B}^0)$, $\eta(\mathrm{B}^0)$.

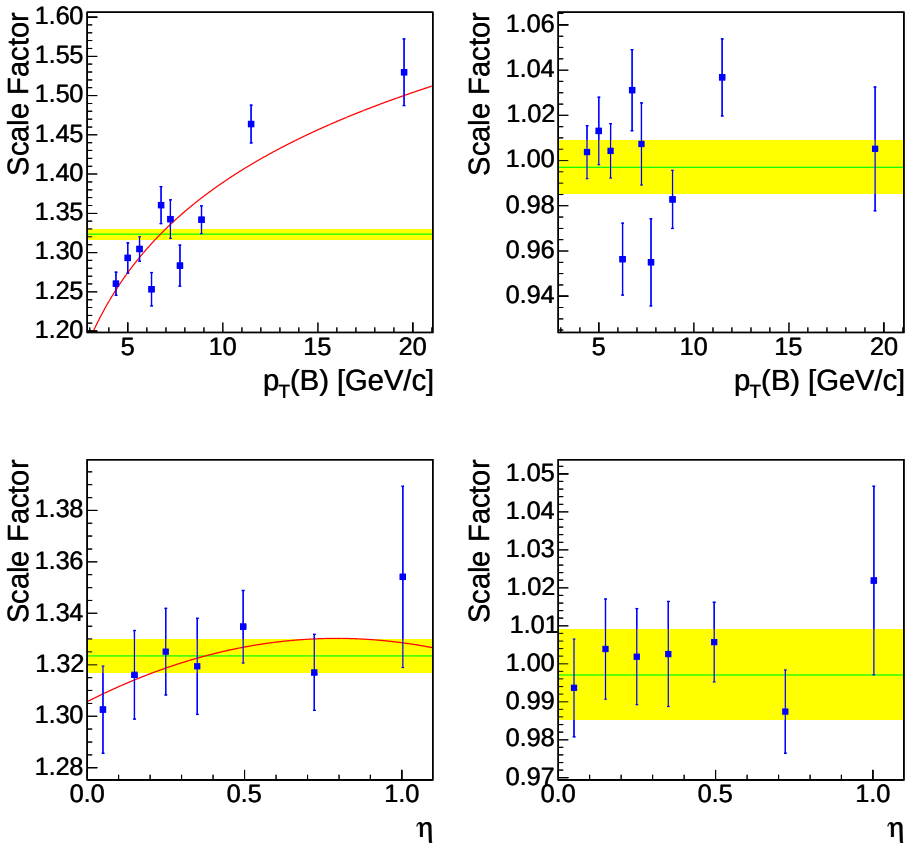

FIGURE 4.26: Proper decay time scale factor dependence on different decay variables, before (left) and after (right) tuning. In order, from top to bottom: $p_T(\mathrm{B}_s^0)$, $\eta(\mathrm{B}_s^0)$.

$$
\begin{aligned}
S_{ct} = \quad & (1.65414 \times 10^{-1} \cdot \log(p_T) + 1.00841) \times \\
& (2.30796 \times 10^{-2} \cdot (\Delta R) + 1.00331) \times \\
& (1.01301 - 2.38776 \times 10^{-1} \cdot (iso - 0.6)) \times \\
& (-2.86602 \times 10^{-2} \cdot \log(nnOut) + 0.948548) \times \\
& (0.997869 + 3.51456 \times 10^{-2} \cdot (\eta - 0.4) + 3.76758 \times 10^{-2} \cdot (\eta - 0.4)^2)
\end{aligned}
\tag{4.4.4}
$$

Performance Tests In CP Violation Fit

In order to test if we obtain improvements to the CP violation fit, we apply the decay time resolution calibration to the data being used as fit inputs. The fits are "blinded", meaning an unknown shift has been applied to the fit results for the width difference, $\Delta\Gamma_s$, and the CP violation phase β_s in order not to get biased when looking at the data. The selected criteria for observing an improvement are the following:

Parameter	Value	Uncertainty
$c\tau$	+0.045388	0.001532
$\Delta\Gamma$	-0.222917	0.061211
S_{ct}	+1.283806	0.021421
$\left\|A_0\right\|^2$	+0.530587	0.020713
$\left\|A_\parallel\right\|^2/(\left\|A_\parallel\right\|^2 + \left\|A_\perp\right\|^2)$	+0.549319	0.057381
ϕ_s	-1.850689	0.458526
$\delta_\parallel$	-2.477850	0.185345
$\delta_\perp$	+0.393692	0.787431

Likelihood minimum: -87956.54154622795

TABLE 4.12: Blinded fit results before applying the decay time resolution calibration.

- adding the decay time resolution calibration supposedly increases our knowledge about the fitted data. Therefore the resulting fit should be a better fit to the data and the likelihood should become deeper at the minimum.

- consequently, the uncertainties on the quantities of interest: the CP violating phase and the width difference, should become smaller.

We test different configurations of the decay time resolution calibration. Table 4.12 shows the relevant fit output if no calibration is used. Table 4.13 shows the fit output if the full calibration is used. The expected improvement is not observed. The likelihood gets shallower, and the uncertainties on the parameters of interest increase.

We follow up our study by checking if there is an unphysical assumption in one of the calibrations which is driving the likelihood to shallower values and increasing the uncertainties. The CP violation fit results after removing the η, $\eta + bgNN$, $\eta + bgNN + iso$, and $\eta + bgNN + iso + \Delta R$ parametrizations, respectively, never restore the original likelihood depth. The uncertainties on the parameters of interest are always increased. It does not appear that there is a performance gain from using any part of the calibration.

In order to explore the possibility that this effect could be due to the order of parameters used or the parametrization of the curves we re-order the calibration and use a second order polynomial instead of the $a + b\log(x)$ parametrization wherever applicable (we observed the likelihood depth increasing for the variables for which the $a + b\log(x)$ parametrization is used). We also remove the parametrization in p_T altogether. Again, the performance is not restored. We conclude that, in the full CP violation fit, we do not observe a significant improvement to the analysis by using the decay time resolution parametrization in any form tested so far. The performance of the analysis changes only minutely, but for all relevant parameters involved, the performance is always inferior to that of the uncalibrated sample.

Parameter	Value	Uncertainty
$c\tau$	+0.045233	0.001524
$\Delta\Gamma$	-0.217805	0.062297
S_{ct}	+1.000813	0.016282
$\left\|A_0\right\|^2$	+0.528606	0.020777
$\left\|A_\parallel\right\|^2/(\left\|A_\parallel\right\|^2 + \left\|A_\perp\right\|^2)$	+0.544334	0.056901
ϕ_s	-1.844294	0.484346
$\delta_\parallel$	-2.459642	0.181452
$\delta_\perp$	+0.328856	0.788144

Likelihood minimum: -85411.26211783224

TABLE 4.13: Blinded fit results when applying the full decay time resolution calibration as described in Section 4.4.1.

Conclusion

We have studied the dependence of the decay time resolution scale factor on different decay variables in order to maximize the total power of the analysis. We have found embedded dependencies of the scale factor on several decay variables and removed them with a calibration procedure. However, after testing the calibration in full CP violation fits to the data, we find no improvement to the uncertainties on the relevant parameters or the depth of the likelihood. We therefore decide not to apply the calibration to our $B_s^0 \to J/\psi\phi$ angular analysis. At the same time, we have observed that the changes to the fit outputs vary very little when embedded dependencies are calibrated away. This suggests that the systematic uncertainty due to these remaining dependencies is mostly calibrated away by using an average decay time resolution scale factor, and remaining systematic uncertainties are very small, at the $1-2\%$ level.

4.4.2 Flavor Tagging Performances

In this section, we evaluate the flavor tagging performances for the angular analysis on $B_s^0 \to J/\psi\phi$. Most part of the aspects are in common with the description given for the B_s^0 mixing analysis already discussed in Section [**?**]. Consequently, we will avoid a repetition of the details to mainly focus on the differences between the treatment on the $B_s^0 \to J/\psi\phi$ sample with the respect to the B_s^0 decay modes used for the mixing oscillation search.

Opposite Side Tagging

We know tagging the production flavor of B mesons in $p\bar{p}$ collisions is in general a difficult task. These motivates the large variety of opposite side tagging algorithms developed: Soft Muon Tagger (SMT), Soft Electron Tagger (SET), Jet Charge Tagger (JQT) and Opposite Side Kaon Tagger (OSKT). All the available information were combined in a Neural Network fashion obtaining a single tagger, the combined

Dataset	ϵ [%]	Scale Factor [%]	Δm_d [ps^{-1}]
0d	0.951 ± 0.003	0.93 ± 0.11	0.37 ± 0.12
0h	0.963 ± 0.003	0.93 ± 0.09	0.57 ± 0.09
0i	0.961 ± 0.003	0.90 ± 0.12	0.37 ± 0.14
0d+0h+0i	0.958 ± 0.002	0.96 ± 0.06	0.48 ± 0.07

TABLE 4.14: Combined Opposite Side Tagger performances for the three datasets, using event-by-event dilution. The Δm_d value for the three datasets independently is consistent with the PDG average value inside the quoted statistical error.

opposite side tagger (COST), which produces a decision if at least one of the taggers provides flavor information. The purpose of combining information from various taggers is to exploit correlations between taggers, which allows us to improve the tagging effectiveness (ϵD^2) of roughly 20% compared to an exclusive combination (See Table 4.4).

Therefore, as was done in the B_s mixing analysis, we would like to use this combined opposite tagging algorithm (COST). While the calibration of all these taggers was already done on a very large $\ell + SVT$ triggered sample of semileptonic B decays, the aim now is to perform a calibration on the scale factor using both B^+ and B^0 as control samples.

We use the high statistic sample $B^0 \to J/\psi K^{*0}$ of 1.35 fb^{-1} available while for the $B^+ \to J/\psi K^+$ we did not have more than the ! fb^{-1} available at the time the calibration was performed.

Table 4.14 reports the performances of the COST for the three datasets available. The outcome of the fits for B^+ and B^0 mass and lifetime are shown in Figure 4.27. The lifetime results from the fits are consistent with the PDG average value.

As before, the separated B^+ and B^0 contributions have been evaluated. The results are shown in Table 4.15, where, it can be observed, by looking at the error associated with the scale factor, how the B^+ decay modes weigh more in the evaluation of the scale factor since they comprise a much larger fraction.
The non-mixing B^+ information is particularly valuable to perform precise studies on the possible asymmetry for the COST with respect to the b flavor of the particle to tag. In fact, we observed that the measured dilution dependence as a function of the per-event dilution does not correspond to the desired linear behaviour for B^+ and B^-. On the other hand, it is recovered when we repeat the exercise for the whole sample B^+ (see Figure 4.28).

First of all, this not-expected dependence on the B mesons flavor suggests that the COST could and should be improved and refined. More important, the use of such tagger (COST) in a CP violating analysis, where separating different particle flavors is crucial, could introduce a bias due to this asymmetry dilution behaviour. While this statement is not confirmed, we safely decided to drop the COST for the angular analysis and to combine the OS taggers in a more classic hierarchical fashion.

As already mentioned, in the previous B_s^0 mixing analysis [14] the taggers were utilized in a mutually exclusive combination. For the angular analysis we finally

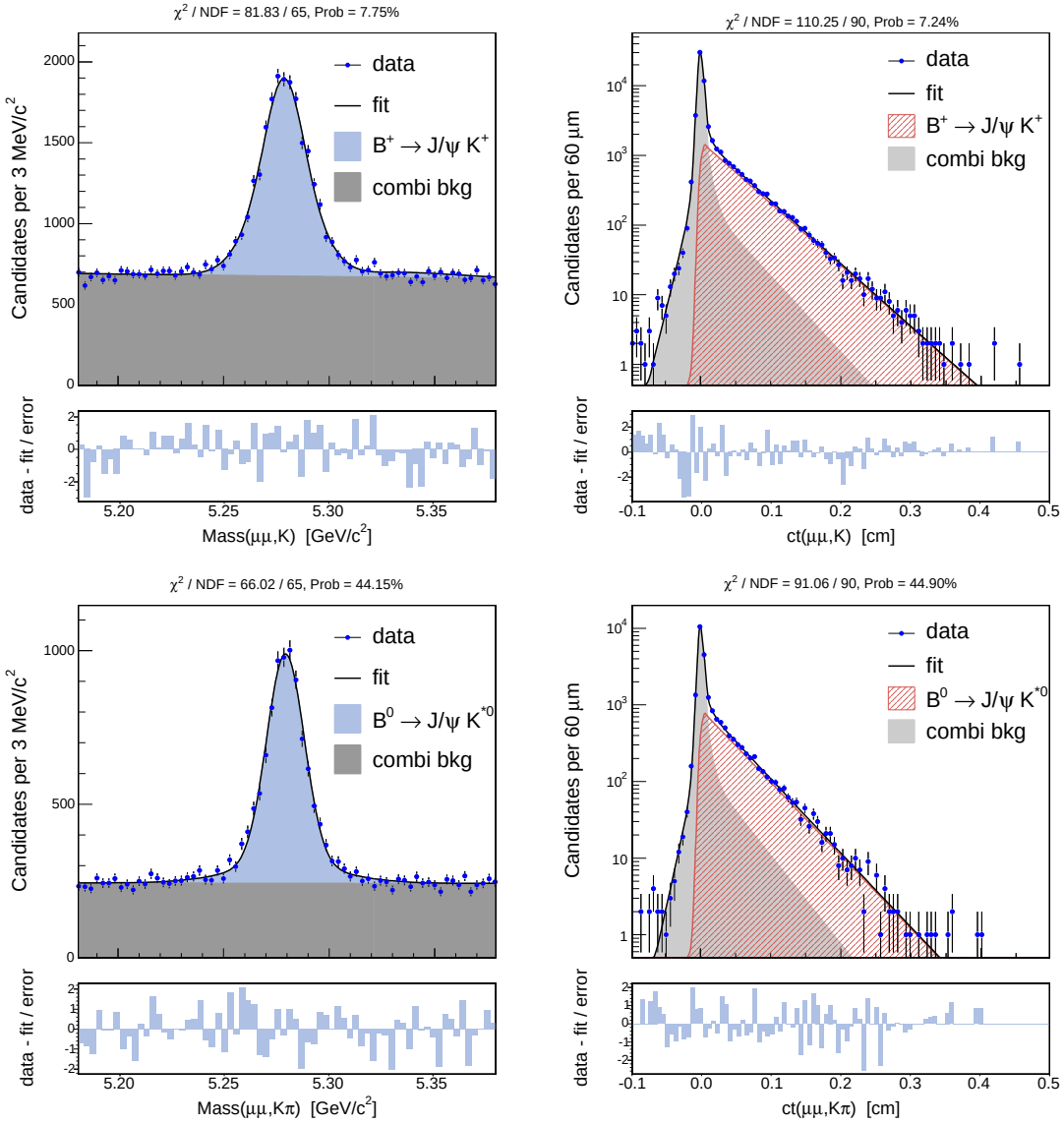

FIGURE 4.27: Mass and Lifetime projections for $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ decay modes for the whole datasets available, corresponding to a respective integrated luminosity of 1 fb^{-1} and 1.35 fb^{-1}.

Dataset	Scale Factor B^+	Scale Factor B^0	Scale Factor(B^++B^0)
0d	0.87 ± 0.13	1.12 ± 0.24	0.93 ± 0.11
0h	0.93 ± 0.10	0.91 ± 0.22	0.93 ± 0.09
0i	0.90 ± 0.15	0.89 ± 0.20	0.90 ± 0.12

TABLE 4.15: Combined Opposite Side Tagger performances for the three datasets, splitted accordingly to their decay mode, $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$. For B^+ decay mode only the first fb^{-1} has been used.

Decay Mode	Parameter	Value
$B^+ + B^0$	ϵ (UOST)	0.9589 ± 0.0015
$B^+ + B^0$	S_D (UOST)	0.853 ± 0.055
B^+	ϵ^{bkg} (UOST)	0.9773 ± 0.0007
B^+	$< D_{bkg} >$ (UOST)	0.0394 ± 0.0048
B^0	ϵ^{bkg} (UOST)	0.9775 ± 0.0012
B^0	$< D_{bkg} >$ (UOST)	-0.0149 ± 0.0082
	Δm_d (ps^{-1})	0.44 ± 0.07

TABLE 4.16: United Opposite Side Tagger performances for the whole datasets.

decided to treat the OS taggers in the same way but, differently from what done before, combine them in a single tagger, having to deal with only one efficiency, one dilution and one scale factor to extract. In such way we simplify the framework for handling multiple taggers and still benefit from the hierarchical combination. Once the "new" tagger, now called "United Opposite Side Tagger" (UOST), is defined, the following step is to calibrate it over the higher statistical samples B^+ and B^0. The results are quoted in Table 4.16. In Tables 4.17 and 4.18 the results of the calibration using respectively for single B^+ and B^0 decay modes are reported.

By using the taggers in this unified fashion we will have only one dilution distribution for signal and one for background to be plugged in our fitter framework. The predicted dilution distributions for B^+, B^0 and B_s^0 are shown in Figure 4.29.

As for the COST we checked in more details for possible worrisome asymmetries by analyzing the non-mixing B^+ decay mode. Again, we splitted the sample in several ranges of dilution and, this time, perform a linear fit while forcing the line to pass through the origin. The measured dilution versus the predicted dilution is reported in Figure 4.30 with the relative linear fits. The outcome of the fit is finally reported in Table 4.19. Conclusively, the strong asymmetry observed with the COST is soften while using the hierarchical taggers in this unified version. The definitive choice for the $B_s^0 \to J/\psi\phi$ angular analysis is therefore to utilize the UOST as Opposite Side Tagging algorithm.

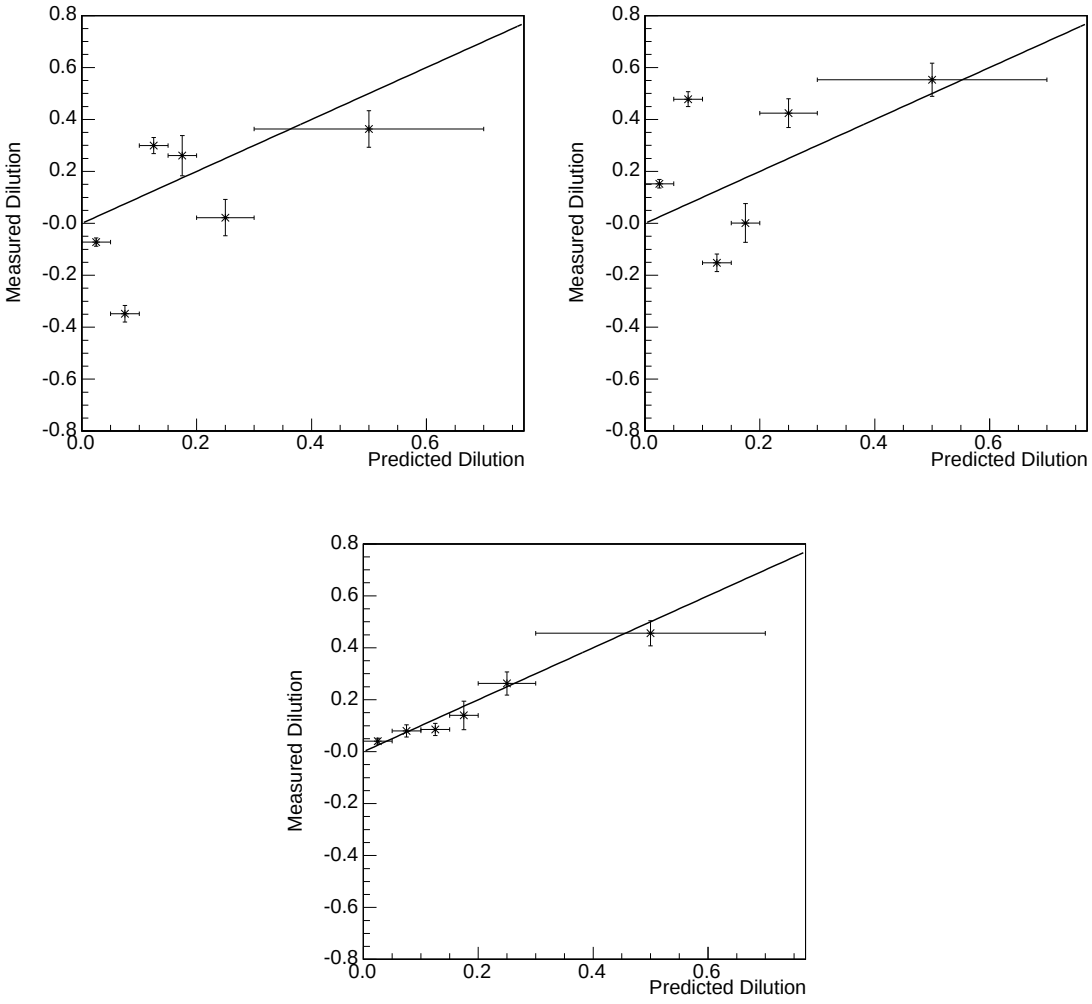

FIGURE 4.28: Measured dilution as a function of the per-event dilution for B^+ (top left), B^- (top right) and the entire sample B^+ (bottom). For reference, the linear function $y = x$ has been drawn.

Decay Mode	Parameter	Value
B^+	ϵ (UOST)	0.9577 ± 0.0019
B^+	S_D (UOST)	0.823 ± 0.064
B^+	ϵ^{bkg} (UOST)	0.9773 ± 0.0007
B^+	$< D_{bkg} >$ (UOST)	0.0394 ± 0.0048
B^+ only	ϵ (UOST)	0.9563 ± 0.0027
B^+ only	S_D^+ (UOST)	0.806 ± 0.091
B^+ only	ϵ^{bkg} (UOST)	0.9784 ± 0.0010
B^+ only	$< D_{bkg} >$ (UOST)	0.0295 ± 0.0067
B^- only	ϵ (UOST)	0.9591 ± 0.0027
B^- only	S_D^- (UOST)	0.839 ± 0.092
B^- only	ϵ^{bkg} (UOST)	0.9763 ± 0.0010
B^- only	$< D_{bkg} >$ (UOST)	0.0494 ± 0.0067

TABLE 4.17: United Opposite Side Tagger performances for the whole datasets for $B^+ \to J/\psi K^+$.

Decay Mode	Parameter	Value
B^0	ϵ (UOST)	0.9611 ± 0.0025
B^0	S_D (UOST)	0.94 ± 0.11
B^0	ϵ^{bkg} (UOST)	0.9773 ± 0.0012
B^0	$< D_{bkg} >$ (UOST)	-0.0154 ± 0.0082
	Δm_d (ps^{-1})	0.44 ± 0.07

TABLE 4.18: United Opposite Side Tagger performances for the whole datasets for $B^0 \to J/\psi K^{*0}$.

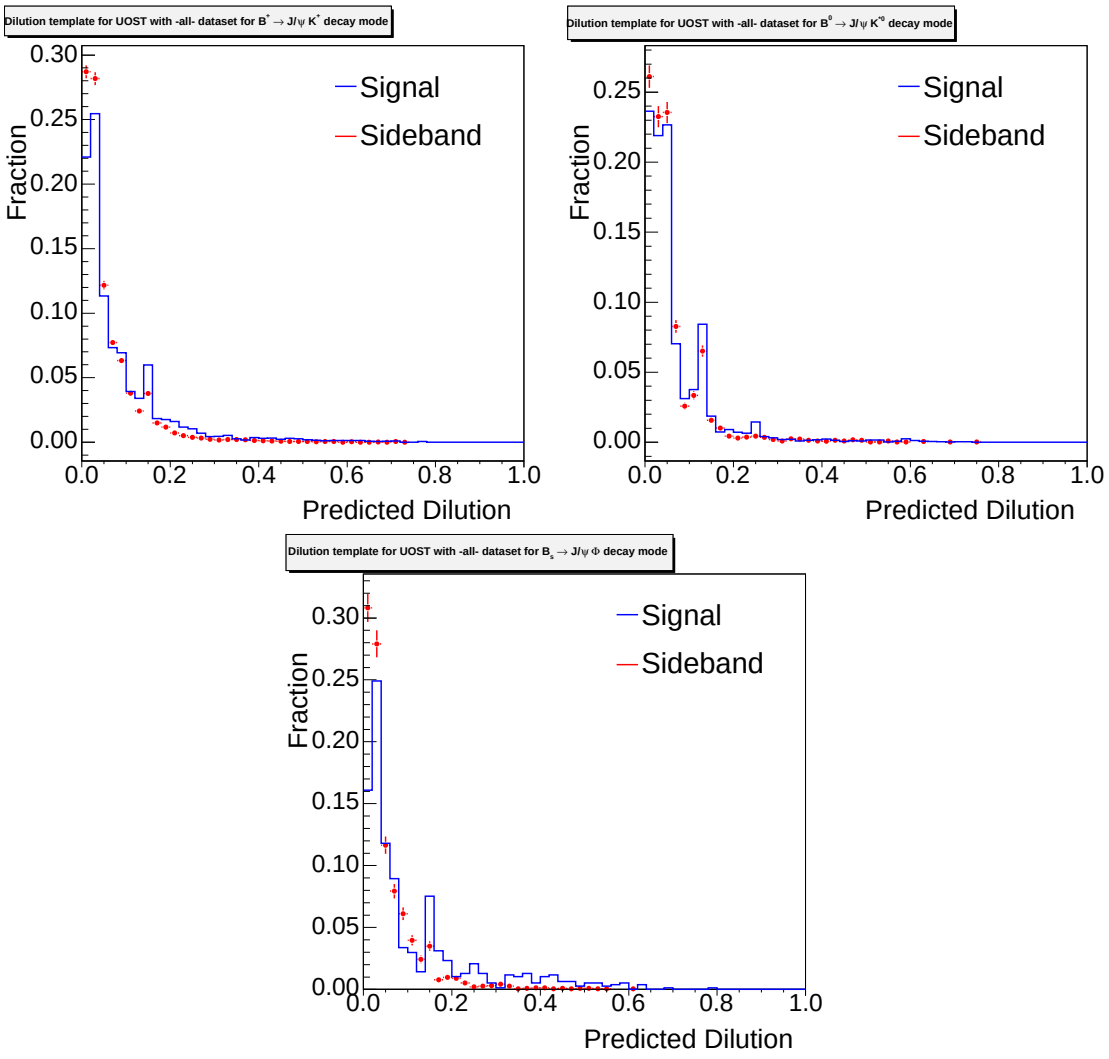

FIGURE 4.29: Distributions of the per-event predicted dilution respectively for the B^+, B^0 and B_s^0 for the UOST which combined in a single tagger the following tagging algorithm: SMT, SET, JQTVX, JQTJP, JQTPT.

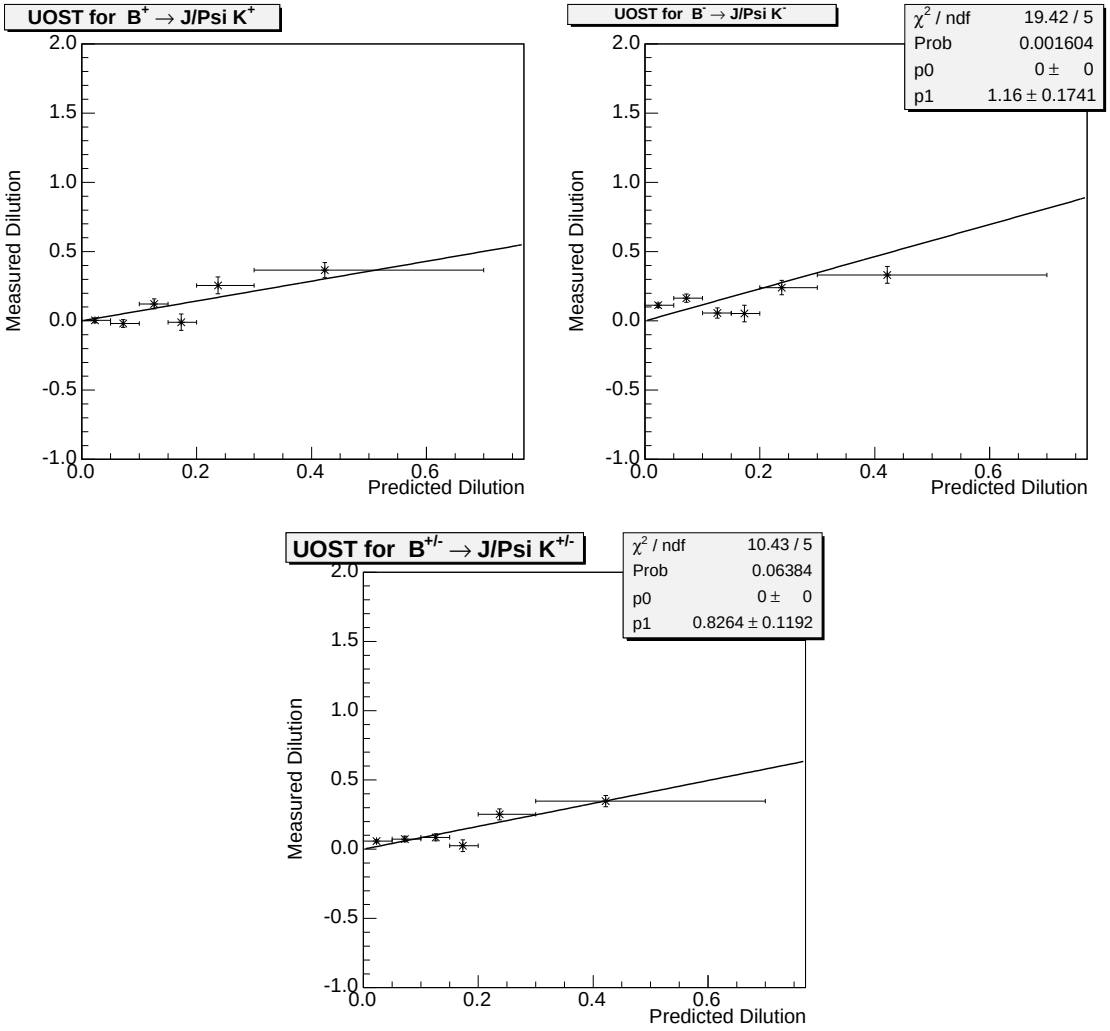

FIGURE 4.30: Measured Dilution as a function of the per-event dilution for B^+ (top left), B^- (top right) and the entire sample B^+ (bottom). The relative fit is drawn.

Decay Mode	Parameter	Value
B^+	S_D (UOST)	0.72 ± 0.15
B^-	S_D (UOST)	1.16 ± 0.18
B^+	S_D (UOST)	0.82 ± 0.12

TABLE 4.19: United Opposite Side Tagger fit results for $B^+ \rightarrow J/\psi K^+$ of the measured dilution vs the predicted dilution (see Figure 4.30). The fit function is $y = S_D * x$.

Same Side Kaon Tagger

For the Same Side Kaon Tagging (SSKT) we use the ANN-based tagger which combines the PID and kinematic information. The description of the development of this tagger algorithm is given in Section [**?**] and its results, in terms of calibration scale

NN SSKT	
ϵ (Data)	50.0 %
$< D >$(Data)	25.4 %
$\epsilon S_D < D^2 >$	3.0 ± 0.9 %

TABLE 4.20: Overall Same Side Kaon Tagger performance.

factor, are summarized in Equation 4.3.10. The overall SSKT performance on the sample $B_s^0 \to J/\psi\phi$ is reported in Table 4.20. For the final fit we will use an average scale factor weighted over the luminosity of 0.964.

4.4.3 Conclusions

We have reviewed the key elements utilized to perform the analyses presented in this dissertation. For the B_s^0 mixing search it is not known *a priori* if the samples available have enough sensitivity to a direct measurement in the time domain. Therefore an indirect approach in the frequency domain, called "amplitude scan", was employed. Additionally, this method guarantees the possibility to establish limits when enough sensitivity to observation is missing. The contributions affecting the power of this analysis are the proper decay resolution and the need of an event-by-event parametrization for it and the tagging performances. The σ_{ct} parametrization is motivated by the non-correct estimation of the resolution from the `CTVMFT` fitter. We defined the technique for extracting the σ_{ct} resolution. We also reviewed all the taggers employed to identify the B_s^0 flavor at its production time, the Opposite Side Tagging (OST) and the Same Side Kaon Tagging (SSKT). These algorithms provide a candidate-by-candidate weight for the correctness of their decision. Their calibrations, as a crucial aspect of the mixing analysis, were also discussed. It was observed how the SSKT, performing 2/3 better with respect to the OST, provides most of the tagging power of the analysis.

The same contributions were studied for the $B_s^0 \to J/\psi\phi$ analysis drawing sometimes different conclusion. The proper decay time resolution parametrization showed no improvements in a blind test on the data and thus it was dropped in the final analysis: all the possible dependencies were mostly calibrated away utilizing a simpler average scale factor. Concerning the taggers, like-wise to the B_s^0 mixing analysis the same SSKT algorithm was used. On the other hand it was observed a worrisome dilution asymmetry in the ANN-based OST (COST) used in the mixing analysis. Therefore the OS taggers were combined in hierarchical fashion, under the name "United Opposite Side Tagging" (UOST) which showed a complete recovery of this asymmetry.

In this chapter we described all the elements to perform the analyses presented in this dissertation. The next chapter introduces the maximum likelihood fitter which combines the mass, proper decay-time and tagging information to extract the parameters of interest, namely the B_s^0 mixing oscillation, Δm_s and the CP violation phase β_s.

Chapter 5

B_s^0 Analysis Fitter Frameworks

In the following chapter we describe the fitter frameworks utilized for the measurements of the oscillation frequency, Δm_s, and the decay width-difference, $\Delta\Gamma_s$, and the CP violating phase, β_s. The base principles of a maximum likelihood fit will be reported, underlying the common aspects and the differences between the two analyses. The likelihood description for the B_s^0 oscillation analysis will be not treated extensively since it was not the focus of my work for this analysis [1]; more emphasis will be, then, given to the likelihood description for the angular analysis aiming at the first mixing-induced CP violation measurement in the B_s^0 system, where my contribution to the likelihood development was more substantial.

5.1 Observables

As we reconstruct an event, several are the measured quantities, called observables, which describe a decay mode and can be used as input variables to the fitter.

For the B_s^0 oscillation analysis the quantities reconstructed are:

- The mass: m.
- The proper decay length and its uncertainty: ct, σ_{ct}.
- The flavor tagging information: ξ, $\mathcal{D}$.

Here, m is the reconstructed mass of the B_s^0 candidate. We remind that the proper decay length ct is calculated from the transverse decay length measurement, L_{xy}, the transverse momentum measurement, p_T, and the particle mass from the PDG [6] as

$$ct = \frac{L_{xy}}{p_T} \cdot m_{PDG} \tag{5.1.1}$$

Analogously, the uncertainty on the proper decay length is defined as:

$$\sigma_{ct} = \frac{\sigma_{L_{xy}}}{p_T} \cdot m_{PDG} \tag{5.1.2}$$

[1] The B_s^0 oscillation measurement is also described in Ref. [69] [66] [70] among others.

A tagger is defined as an algorithm which attempts to determine the flavor of the B_s^0 meson at the production time (as opposed to the decay time). It is characterized by its decision ξ which is a discrete variable taking values in $\{-1, 0, 1\}$, respectively when the flavor at production time is different from the flavor at decay time ($\xi = -1$), when no decision can be made ($\xi = 0$) and when the flavor is the same at production and decay time ($\xi = 1$), by convention. In addition to a decision, for each event, the tagger also returns a dilution $\mathcal{D}$, defined as:

$$\mathcal{D} \equiv 2P^{RS} - 1 = 1 - 2P^{WS}, \tag{5.1.3}$$

where P^{RS} is the probability of having made the correct tagging decision, and $P^{WS} = 1 - P^{RS}$ (RS stands for "right sign" and WS for "wrong sign"). Finally, a tagger is characterized by its efficiency $\epsilon \equiv P(\xi \neq 0)$. This efficiency is treated as a parameter to be determined in the maximum likelihood fit and differs for each tagger used.

For the $B_s^0 \to J/\psi\phi$ angular analysis another set of observables is added to the already mentioned quantities (from mass, lifetime and tagging):

- The angles $\vec{\omega} \equiv (\cos\theta_T,\ \phi_T,\ cos\psi_T)$.

The angles forming the vector $\vec{\omega}$ are defined in the transversity basis, chosen for its convenience in theoretical formulations (note: $\vec{\omega}$ is not a vector in the proper sense that it transforms as one under rotation. Rather, it is a convenient shorthand for an object with three components). The component of $\vec{\omega}$ have been already introduced in Chapter 1. We recall here their definitions. Considering the decay $B_s^0 \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$, the first two angles are calculated in the rest frame of the J/ψ, and the third one in the rest frame of the ϕ. In the rest frame of the J/ψ, the ϕ direction defines the x axis. The plane of K^+K^- defines the $x - y$ plane with $p_y(K^+) > 0$. From there:

- θ_T is the angle between the μ^+ and the $x - y$ plane in the frame of the J/ψ,
- ϕ_T is the angle between the x axis and the projection of p_{μ^+} onto the $x - y$ plane in the frame of the J/ψ,
- ψ_T is the angle between the K^+ and the negative of the flight direction of the J/ψ In the frame of the ϕ

The three angles are shown pictorially in Figure 5.1.

5.2 Maximum Likelihood Framework

All the results presented in this dissertation are obtained utilizing a classic unbinned likelihood estimation method [71] that analyzes mass, lifetime, time-dependent flavor asymmetry and, when needed, angular aspects of the data samples. By using such

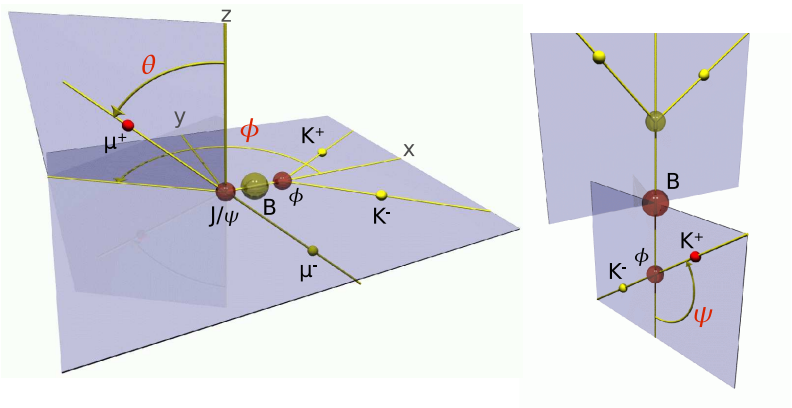

FIGURE 5.1: Transversity basis and angles definition for the $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decay mode.

technique, it is indeed possible to combine different information coming from different decay modes and their fits in order to maximize the statistical power of an analysis. The general definition for the likelihood of a candidate i is written as

$$\mathcal{L}_i = \sum_j f_j \cdot \mathcal{P}_i^j(\overrightarrow{x}_i; \overrightarrow{y}), \tag{5.2.1}$$

where the index j runs over the various signal and background components in the samples, f_j is the fraction of the $j-th$ component and $\mathcal{P}_i^j(\overrightarrow{x}_i, \overrightarrow{y})$ is the corresponding probability density, function of the $i-th$ candidate observables, which for simplicity of notation are condensed in a vector $\overrightarrow{x}_i$, and the parameters of interest $(\overrightarrow{y})$ which one wants to extract. By construction $\sum_j f_j = 1$. The global likelihood $\mathcal{L}$ is naturally defined as the multiplication of all the single-candidate likelihoods $\mathcal{L}_i$:

$$\mathcal{L} = \prod_i \mathcal{L}_i. \tag{5.2.2}$$

All aspects of interest of the data samples are described with a model and then the model parameters are found by minimizing the value $-2log\mathcal{L}$ using MINUIT [72]. The likelihood terms describing each component, $\mathcal{P}_i^j$ are given by the joint PDFs (Probability Density Functions) of the fit input variables. The general formulations which have been used are

$$\mathcal{P}_i^j = \mathcal{P}_m^j \mathcal{P}_{ct}^j(ct\ |\sigma_{ct}, \mathcal{D}) \mathcal{P}_{\sigma_{ct}}^j(\sigma_{ct}) \mathcal{P}_{\mathcal{D}}^j(\mathcal{D}) \qquad \text{for the } B_s^0 \text{ oscillation analysis} \tag{5.2.3}$$

$$\mathcal{P}_i^j = \mathcal{P}_m^j \mathcal{P}_{ct}^j(ct, \vec{\omega}\ |\sigma_{ct}, \mathcal{D}) \mathcal{P}_{\sigma_{ct}}^j(\sigma_{ct}) \mathcal{P}_{\mathcal{D}}^j(\mathcal{D}) \qquad \text{for the angular analysis} \tag{5.2.4}$$

Analyzing the above formulas, the reader can observe that the PDFs for signal and background in the mass m are not coupled to the functional dependence of probabilities on the remaining observables. However, tagging information, lifetime and, in

the angular analysis, the angles are interconnected through the mixing and the mistags. They cannot be factorized among themselves as easily as it can be done for the mass PDF from the rest. Nevertheless, the Theory of Probability helps us providing a theorem that allows us to do factorization via the conditional probability [73]

$$\mathcal{P}(a,b) = \mathcal{P}(a\,|b) \cdot \mathcal{P}(b), \tag{5.2.5}$$

where $\mathcal{P}(a|b)$ is the probability to observe a given b and $\mathcal{P}(b)$ is the decoupled probability distribution for b. The PDFs $\mathcal{P}_m$ and $\mathcal{P}_{ct}$ are parametric models based on the known physics of the system, while the background ones are empirically chosen to describe the data in the mass and lifetime (and angular) space at satisfactory level. The dilution (σ_{ct}) terms $\mathcal{P}^j_{\mathcal{D}}(\mathcal{D})$ ($\mathcal{P}^j_{\sigma_{ct}}(\sigma_{ct})$), also referred as "dilution (σ_{ct}) templates" are simply normalized histograms prepared from the data separately for background, from the sideband regions, and for signal, using the sideband subtraction technique. The algorithm for the sideband-subtraction is defined from the following steps:

1. For each decay mode considered in the analysis, two mass ranges are selected to define the signal and the sideband regions.

2. The invariant mass distribution is fit with a function representing the signal and the background components (see next Section). The ratio of the integrals of the background component extrapolated under the signal region and in the sideband region determines the scale factor to use in the subtraction.

3. The distribution for a quantity of interest (i.e. p_T, σ_{ct}, ...) obtained from the data sideband region is rescaled by the scale factor previously calculated. Thus, the "sideband-subtracted" distribution for this quantity is obtained by subtracting the sideband distribution rescaled from the distribution of the same quantity in the signal region.

The parametrizations for the different signal and background components are described in the following sections.

5.3 Mass Model

The signal mass PDF is usually described by a Gaussian, associated with the detector resolution. However for higher statistics hadronic modes used in the B^0_s mixing analysis or for the $B^0_s \to J/\psi\phi$ mode used in the angular analysis, a combination of a narrower and a broader width Gaussian is commonly used. The two Gaussians have the same mean (M). Note that these Gaussians need actually to be normalized in the fitting mass range rather than in the full real axis. The following parameters are let floating during the fit minimization: the mean of the Gaussians (M), the width of the narrower Gaussian (σ_1) and the ratio of the two Gaussians (R, where $\sigma_2 = R \cdot \sigma_1$), and the relative fraction of the wider Gaussian (f_m). The mass signal probability for an event with reconstructed B^0_s mass m_j is then

$$\mathcal{P}_m^{sig}(m_j;\ M,\sigma_1,R,f_m) = (1-f_m)\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(m_j-M)^2}{2\sigma_1^2}} + f_m\frac{1}{\sqrt{2\pi}R\sigma_1}e^{-\frac{(m_j-M)^2}{2R^2\sigma_1^2}}. \quad (5.3.1)$$

In the single Gaussian case, the fraction f_m is fixed to zero (the ratio R is obviously not allowed to float). In the likelihood for the semileptonic decay modes, for the B_s^0 oscillation analysis, $\mathcal{P}_m^{sig}$ contains terms for both the D_s^- candidate mass and ℓD_s^- mass distributions (see Chapter 3).

The background is as well described by functions which are normalized in the fitting range. Combinatorial background is generally modeled as the sum of a decaying exponential and a flat linear component. For simplest cases, as in the $B_s^0 \to J/\psi\phi$ decay mode, this a first-degree polynomial is sufficient:

$$\mathcal{P}_m^{bkg}(m_j;\ A) = A\cdot m + \frac{1}{M_{max}-M_{min}}\left[1-\frac{A}{2}(M_{max}^2-M_{min}^2)\right], \quad (5.3.2)$$

where A is the slope of the polynomial, and M_{min}, M_{max} are the boundaries of the chosen invariant mass distribution.

The mass model, while being entirely separable from the lifetime (and angular) portions of the global likelihood, is primarily used to discriminate between signal and background events. For most of cases several additional components need to be taken into account. These are often necessary to describe partially reconstructed and Cabibbo-suppressed decays as well as mis-reconstructed decays. The shapes are in general extracted from study on the simulated events and a systematic error will be associated to potential mis-modelings. As an example, the B^0 or $\overline{\Lambda}_b^0$ background contributions to the B_s^0 hadronic decay modes are derived from studies on simulated events produced using the $b\bar{b}$ Monte Carlo generator `BGenerator` [54].

5.4 Lifetime Model

If we momentarily exclude the flavor tagging component, the proper decay length distribution for the B mesons is modeled with an exponential PDF convolved with a Gaussian function that accounts for the detector resolution effects. If ct_j is the measured proper decay length per event and σ_j^{ct} is its uncertainty, the signal PDF for the completely reconstructed B_s^0 candidate is given by

$$\mathcal{P}_{ct}^{sig}(ct_j,\sigma_{ct_j};\ c\tau,S_{ct}) = E(ct_j;\ c\tau)\otimes G(ct_j,\sigma_{ct_j};\ S_{ct}), \quad (5.4.1)$$

where E and G are defined as follows:

$$E(ct \mid c\tau) = \begin{cases} 0 & , \quad ct < 0 \\ \dfrac{1}{c\tau} e^{-\frac{ct}{c\tau}} & , \quad ct \geq 0 \end{cases}$$

$$G(ct, \sigma;\ S_{ct}) = \frac{1}{\sqrt{2\pi} S_{ct} \sigma} e^{-\frac{c^2 t^2}{2(S_{ct}\sigma)^2}}.$$

S_{ct} is the scale factor used to correct over- or under-estimated proper decay length uncertainties (σ_{ct}). The above Formula, as reported, works for the $B_s^0 \to J/\psi\phi$ data sample which is recorded with the `di-muon` trigger paths which does not apply any on-line requirement on ct. In fact, for the decays such as $B_s^0 \to D_s(3)\pi$, the ct distribution of real B mesons is biased by the displaced track trigger (`TTT`) selection cuts. In these cases, the bias is corrected with the introduction of the so-called efficiency function, $\epsilon(ct)$, and the signal PDF for the lifetime model becomes

$$\mathcal{P}_{ct}^{sig}(ct_j, \sigma_{ct_j};\ c\tau, S_{ct}) = \frac{1}{\mathcal{N}} \left[\frac{1}{c\tau} e^{-\frac{ct_j}{c\tau}} \otimes G(ct_j, \sigma_{ct_j};\ S_{ct}) \right] \cdot \epsilon(ct_j), \tag{5.4.2}$$

where $\mathcal{N} = \int \left[\frac{1}{c\tau} e^{-\frac{ct_j}{c\tau}} \otimes G(ct_j, \sigma_{ct_j};\ S_{ct}) \right] \cdot \epsilon(ct_j) dt$ is the normalization factor.

The efficiency curve, $\epsilon(ct)$, depends only on the kinematics of the decay under study and it is defined as

$$\epsilon(ct) = \frac{\text{ct} - \text{distribution after reconstruction and selection}}{\sum_{i=1}^{N} \frac{1}{c\tau} e^{-\frac{ct_i}{c\tau}} \otimes G(ct_i, \sigma_{ct_i};\ S_{ct})}. \tag{5.4.3}$$

The numerator distribution is obtained from a Monte Carlo signal only sample which passes all the analysis selection cuts. For each event i accepted, the expected ct distribution is given by an exponential smeared with a Gaussian resolution function as in Equation 5.4.1, where the width is given by σ_{ct_i}, the ct error of the event. The denominator corresponds to the sum of these N distributions. For each B_s^0 decay mode reconstructed with the `TTT`, the mode specific efficiency curve has been derived utilizing a `BGenerator`-MC sample. The parameter τ represents the world average of B_s^0 lifetime measurements [58]. The shape of the decay-length efficiency curve is parametrized using the following template:

$$\epsilon(ct) = \sum_{j=1}^{3} \alpha_j (ct - \beta_j)^2 e^{-\frac{ct}{\gamma_j}} \theta(ct - \beta_j), \tag{5.4.4}$$

where the parameters α_j, β_j and γ_j are obtained from the fit of the Equation 5.4.4. This functional form allows for the analytical normalization of the proper decay time signal PDF. In Figure 5.2 examples for the efficiency curve dependence are shown.

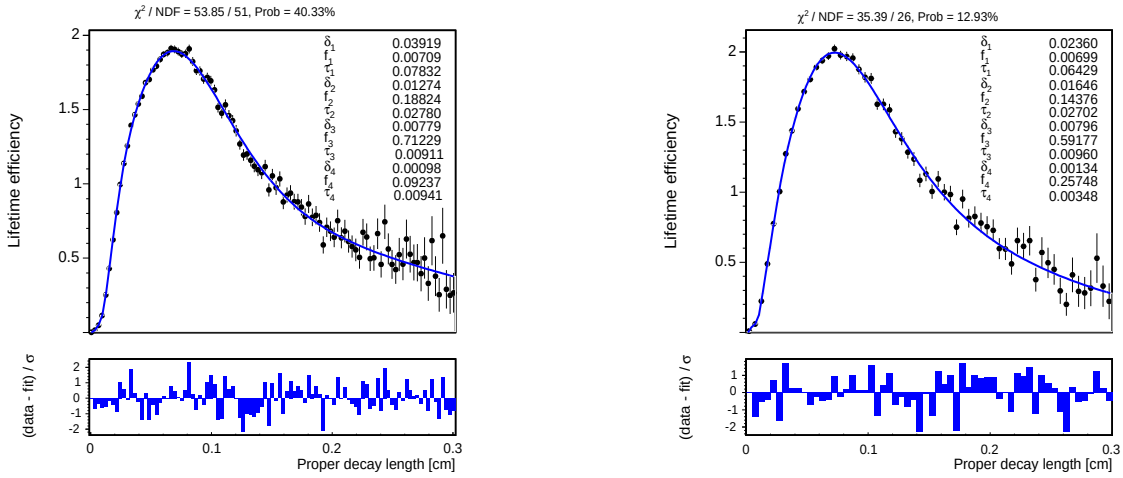

FIGURE 5.2: Dependence of trigger and selection on the proper decay time for $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi\pi^-$ and $\phi \to K^+K^-$, (left) and $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to K^{*0}K^-$, with $K^{*0} \to K^+\pi^-$ (right).

The rapid turn-on of the efficiency is due to the cut on the minimum impact parameter and L_{xy} significance, while the turn-off at larger proper decay length comes from the upper cut on the impact parameter in the trigger.

The case of incompletely reconstructed B_s^0 candidates (hadronic and semileptonic) involves the additional complication of a convolution with the k-factor distribution $F(k)$ (Section 4.1.2):

$$\mathcal{P}_{ct}^{sig}(ct_j^*, \sigma_{ct_j^*};\ c\tau, S_{ct}) = \frac{1}{\mathcal{N}} \int dk \left[\frac{k}{c\tau} e^{-\frac{k\ ct_j^*}{c\tau}} \otimes G(ct_j, \sigma_{ct_j^*};\ S_{ct}) \right] \cdot \epsilon(ct_j^*) \cdot F(k), \quad (5.4.5)$$

where ct_j^* is the pseudo-proper decay-time. Analogously, the efficiency curve has to take into account this new degree of complexity:

$$\epsilon(ct) = \frac{\text{ct} - \text{distribution after reconstruction and selection}}{\int dk \sum_{i=1}^{N} \frac{1}{\tau} e^{-\frac{ct_i}{c\tau}} \otimes G(ct_i, \sigma_{ct_i};\ S_{ct}) \cdot F(k)}. \quad (5.4.6)$$

Introducing the term F(k) and integrating over it account for the missing momentum in partially reconstructed decays. For the fully reconstructed decay, the k-factor distribution is assumed to be a delta function $\delta(k-1)$, therefore the Formula is reduced to the one in Equation 5.4.3.

The background distribution for the proper decay length is empirically described as the sum of several contributions. Sometimes the contributions in the background description can appear in terms of templates which shape is taken from proper Monte Carlo simulations. We can consider as an example the case of the $B_s^0 \to J/\psi\phi$

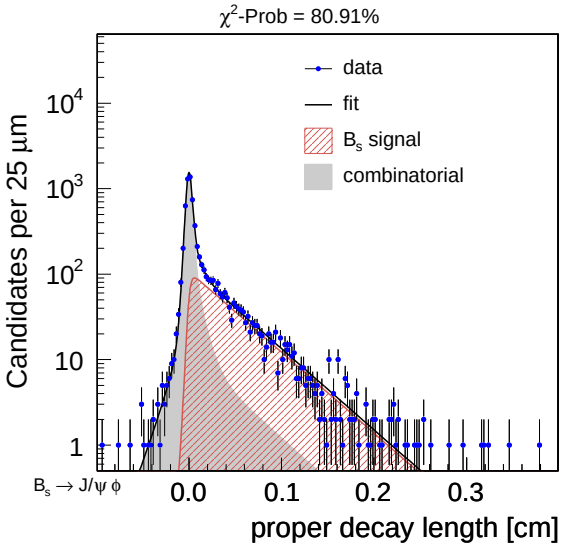

FIGURE 5.3: Proper decay lenght projection for the $B_s^0 \to J/\psi\phi$ decay mode.

decay mode, whose proper decay time distribution is reported in Figure 5.3. The combinatorial background is described as a sum of the following contributions.

- A δ-function centered at zero to model the large background candidates constructed with tracks coming from the primary vertex (prompt J/ψ paired with prompt track(s)).
- A short-lived exponential at negative lifetime, $c\tau_-$, and another at positive lifetime, $c\tau_+$, to handle the mis-measured candidates. They include tracks that makes use of erroneous hits, or tracks that in reality belong to different displaced vertexes.
- A long lived exponential at positive lifetime, $c\tau_{--}$ to describe the long lived background component. It accounts for background contributions from residual physics backgrounds that are not explicitly parametrized in the likelihood, as true displaced J/ψ mesons paired with a random track or two sequential muonic decays $b \to c\mu^-\bar{\nu_\mu}$, $c \to s\mu^+\nu_\mu$, where the muon pair can fake a J/ψ meson. We convolve all terms with the Gaussian resolution function.

Combining all these, one obtain

$$\begin{aligned}
&\mathcal{P}_{ct}^{bkg}(ct_j, \sigma_{ct_j};\ f_-, f_+, f_{++}, c\tau_-, c\tau_+, c\tau_{++}, S_{ct}) = \\
&\quad (1 - f_- - f_+ - f_{++}) \otimes G(ct_j, \sigma_{ct_j};\ S_{ct}) \\
&\quad + f_- \cdot E(-ct_j;\ c\tau_-) \otimes G(ct_j, \sigma_{ct_j};\ S_{ct}) \\
&\quad + f_+ \cdot E(ct_j;\ c\tau_+) \otimes G(ct_j, \sigma_{ct_j};\ S_{ct}) \\
&\quad + f_{++} \cdot E(ct_j;\ c\tau_{++}) \otimes G(ct_j, \sigma_{ct_j};\ S_{ct})
\end{aligned} \tag{5.4.7}$$

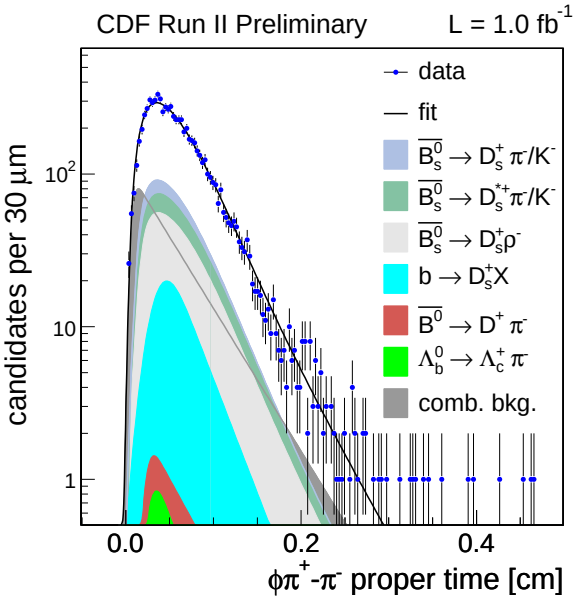

FIGURE 5.4: Proper decay lenght projection for the $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi\pi^-$, decay mode.

where $c\tau_-, f_-$ are the decay constant and fraction of the negative-lifetime tail, $c\tau_+, f_+$ are the decay constant and fraction of the short-live positive-lifetime tail, and $c\tau_{++}, f_{++}$ are the decay constant and fraction of the long-lived positive-lifetime tail, all of which float in the fit. For the $B_s^0 \to J/\psi\phi$ sample, the scale factor S_{ct} can be extracted directly from the data adjusting the width of the prompt Gaussian component which describes the ct distribution of combinatorial background events. However for the other sample affected by the SVT trigger requirements (i.e. the B_s^0 hadronic decays) such prompt component is greatly reduced and ct scale factor cannot be determined with the fit as for t-unbiased samples. Thus, a more detailed approach has been developed whose description is given in Section 4.1.3. An example of the proper decay time distribution of $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi(\to K^+K^-)\pi$, candidates, both fully and partially reconstructed, is shown in Figure 5.4.

Defining correctly the conditioned probability [73], the signal and the background components of the proper decay length PDFs are finally multiplied by their σ_{ct} distributions, $\mathcal{P}_{\sigma_{ct}}(\sigma_{ct})$ which are extracted from data. As we can see in Figure 5.5, which refers to the case of $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi\pi^-$, the σ_{ct} distributions for signal, obtained using data sideband-subtracted, and for background, from the event in the sideband regions of the mass space, are distributed differently. Thus, the introduction of such a term in the likelihood gives more discriminating power to the fitter. A complete analysis of the B_s^0 lifetime, which would include the evaluation of the systematic uncertainties, has not been performed. However, a measurement of the B_s^0 lifetime has been obtained as a by-product of the B_s^0 oscillation analysis or the $B_s^0 \to J/\psi\phi$ angular analysis by performing a fit of the proper decay lenght as the one presented in Figures 5.3 and 5.4. The results are in agreement with the world average of the B_s^0 lifetime measurements, which constitute a cross-check that the absolute ct scale

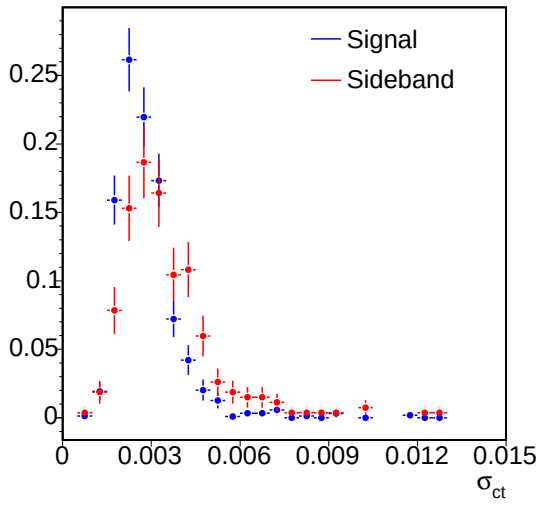

FIGURE 5.5: Distributions for the σ_{ct} obtained from the data in the signal mass region sideband subtracted (blue crosses) and in the sideband mass region (red crossed) for the decay mode $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi\pi^-$.

is correct.

5.5 Likelihood with Tagging

The introduction of the taggers increases the information available for each event and it adds up new degrees of complexity. The description on how the tagger presence enters the likelihood construction is given in this section. We will start with the implementation of one tagger to move to the general two taggers construction. In fact, in the final incarnation of each analyses two are the tagging algorithms employed: the opposite and the same side tagger (details are given in Chapter 6). At this point, the discussion will be naturally splitted according to the specific analysis considered.

5.5.1 General Construction with One Tagger

The mass portion of the likelihood is not affected by tagging since we do not construct it differently for matter and antimatter. Therefore we are interested in the likelihood part which combine altogether the information of the tagging (efficiency and dilution), of the lifetime and, in the $B_s^0 \to J/\psi\phi$ angular analysis, of the angles $\vec{\omega}$. Being connected through the mixing and the mis-tags, these contributions cannot be factorized, so what we are interested in defining now has the general form

$$\mathcal{P}(ct, \sigma_{ct}, \vec{\omega}, \mathcal{D}, \xi) \equiv \mathcal{P}(ct, \sigma_{ct}, \vec{\omega}, \mathcal{D} \,|\xi) \cdot \mathcal{P}(\xi), \tag{5.5.1}$$

where

$$\mathcal{P}(\xi) = (1-\epsilon)\cdot\delta(\xi-0) + \epsilon\cdot\delta(|\xi|-1). \tag{5.5.2}$$

Because of the mixing, a positive tag decision can be obtained either from a correct tag of a B or an incorrect tag of a $\overline{\mathrm{B}}$. Likewise, a negative tag can come either from a correct tag of a $\overline{\mathrm{B}}$ or an incorrect tag of a B:

$$\mathcal{P}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}\mid+1) = \mathcal{P}^{RS}\cdot\mathcal{P}_{\mathrm{B}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}) + \mathcal{P}^{WS}\cdot\mathcal{P}_{\overline{\mathrm{B}}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}), \tag{5.5.3}$$

$$\mathcal{P}(ct,\sigma_{ct},\vec{\omega}\,\mathcal{D};\mid-1) = \mathcal{P}^{RS}\cdot\mathcal{P}_{\overline{\mathrm{B}}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}) + \mathcal{P}^{WS}\cdot\mathcal{P}_{\mathrm{B}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}). \tag{5.5.4}$$

Utilizing Equation (5.1.3), we can combine the two expressions as well as include the case of no decision and get the general construction

$$\mathcal{P}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}|\xi) = \frac{1+\xi\mathcal{D}}{1+|\xi|}\cdot\mathcal{P}_{\mathrm{B}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}) + \frac{1-\xi\mathcal{D}}{1+|\xi|}\cdot\mathcal{P}_{\overline{\mathrm{B}}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}). \tag{5.5.5}$$

5.5.2 General Construction with Two Taggers

With more than one tagger, Equation (5.5.2) changes to accommodate independent decisions by the various taggers. Defining the decision ξ_j and the dilution $\mathcal{D}_j$ of a tagger j the general expression for the probability of obtaining a vector of tag decisions $\vec{\xi}$ can be written as

$$\mathcal{P}(\vec{\xi}) = (1-\sum_j \epsilon_j)\cdot\delta(\sum_j \xi^j - 0) + \sum_j \epsilon_j\cdot\delta(|\xi^j|-1),$$

$$\mathcal{P}(\vec{\xi}) = \begin{cases} (1-\epsilon_1)(1-\epsilon_2) & (\xi_1=0,\xi_2=0) \\ \epsilon_1(1-\epsilon_2) & (\xi_1=\pm1,\xi_2=0) \\ (1-\epsilon_1)\epsilon_2 & (\xi_1=0,\xi_2=\pm1) \\ \epsilon_1\epsilon_2 & (\xi_1=\pm1,\xi_2=\pm1) \end{cases} \tag{5.5.6}$$

Moreover, we can rewrite Equation (5.5.5) to account for independent decisions made by more than one tagger:

$$\mathcal{P}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}|\vec{\xi}) = \frac{1+\xi_1\mathcal{D}_1}{1+|\xi_1|}\cdot\frac{1+\xi_2\mathcal{D}_2}{1+|\xi_2|}\cdot\mathcal{P}_{\mathrm{B}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}) + \frac{1-\xi_1\mathcal{D}_1}{1+|\xi_1|}\cdot\frac{1-\xi_2\mathcal{D}_2}{1+|\xi_2|}\cdot\mathcal{P}_{\overline{\mathrm{B}}}(ct,\sigma_{ct},\vec{\omega},\mathcal{D}) \tag{5.5.7}$$

Note that Equation (5.5.7) reduces to Equation (5.5.5) if only one tagger makes a decision. The general Formula presented includes the case of correct decisions by both taggers, incorrect decisions by both or a combination of correct and incorrect

decisions. The taggers are thus independently combined.

5.6 B_s^0 Mixing Likelihood

5.6.1 Signal Likelihood

When looking for mixing oscillation, we are interested in expressing the probability of observing a flavor changing ($\xi = -1$) or not changing ($\xi = 1$). At the proper time t, this probability can be expressed using the Formulas 1.2.9 and 1.2.10 defined in Chapter [**?**]

$$\mathcal{P}(\xi, t;\ \Delta m) = \frac{1 + \xi cos(\Delta m t)}{2}, \tag{5.6.1}$$

Starting with the simpler one tagger case, the mis-tagging is accounted using Equation 5.5.5:

$$\mathcal{P}(\xi t;\ \Delta m) = \frac{1 + \xi cos(\Delta m t)}{2}\frac{1 + \mathcal{D}}{2} + \frac{1 - \xi cos(\Delta m t)}{2}\frac{1 - \mathcal{D}}{2} = \frac{1 + \xi \mathcal{D} cos(\Delta m t)}{2} \tag{5.6.2}$$

It is thus possible to report the final construction for the B_s^0 mixing analysis combining flavor tagging and proper decay time for a tagger i, which has a dilution $\mathcal{D}_i$, efficiency ϵ_i and tagging decision ξ_i, for the cases

- **untagged:**

$$\int dk \big(1 - \sum_i \epsilon_i\big) e^{\frac{-kct'}{c\tau}} \otimes G(kct' - kct;\ \sigma_{ct}, S_{ct}) \cdot F(k) \cdot \epsilon(ct), \tag{5.6.3}$$

- **single tag:**

$$\int dk \frac{\epsilon_i}{2}[1 + \xi_i \mathcal{D}^i cos(\Delta m_s kct')] e^{\frac{-kct'}{c\tau}} \otimes G(kct' - kct;\ \sigma_{ct}, S_{ct}) \cdot F(k) \cdot \epsilon(ct), \tag{5.6.4}$$

  where $\xi_i = \pm 1$ is the sign of a single (same-side or opposite side) tagger,

- **double tag:**

$$\int dk \frac{\epsilon_i \epsilon_j}{2} \left[\frac{(1 + \xi_i \xi_j \mathcal{D}^i \mathcal{D}^j) + (\xi_i \mathcal{D}^i + \xi_j \mathcal{D}^j) cos(\Delta m_s kct')}{2}\right] e^{\frac{-kct'}{c\tau}} \otimes G(kct' - kct;\ \sigma_{ct}) \cdot F(k) \cdot \epsilon(ct), \tag{5.6.5}$$

  where $\xi_i = \pm 1$ and $\xi_j = \pm 1$ are the signs of the two tagger.

5.6.2 Background Likelihood

A possible tagging asymmetry is also allowed in the case of background which is treated as non-mixing. Without mixing the terms containing flavor tagging information are independent from the lifetime parts. This is the case for combinatorial background, for $\overline{\Lambda}_b^0$ decays and for the partially reconstructed modes which are not signal ($b \to \mathrm{D}_s^- X$). As for the signal component, the final formulation is splitted according to different tagging cases

- **untagged:**
$$\big(1 - \sum_i \epsilon_i\big)\mathcal{P}_{ct}(ct), \tag{5.6.6}$$

- **single tag:**
$$\frac{\epsilon_i}{2}[1 + \xi_i \mathcal{D}_i]\mathcal{P}_{ct}(ct), \tag{5.6.7}$$

- **double tag:**
$$\frac{\epsilon_i \epsilon_j}{2}\left[\frac{(1 + \xi_i \mathcal{D}_i)(1 + \xi_j \mathcal{D}_j)}{2}\right]\mathcal{P}_{ct}(ct), \tag{5.6.8}$$

where $\mathcal{P}_{ct}(ct)$ is the probability density function for the proper time background components as described in Section 5.4. As for the lifetime component, in the final likelihood construction it is necessary to include the probability distribution function of the candidate-by-candidate dilution for signal and background, $\mathcal{P}_{\mathcal{D}}(\mathcal{D})$. The exclusion of such terms could lead to biases for the fitted values of the likelihood parameters. The distributions for the combinatorial background are obtained from the candidates in the mass sideband, while for the signal component the templates are extracted with the sideband subtraction technique. In Figures 5.6 the same-side tagger dilution templates are shown for the $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, with $\mathrm{D}_s^- \to \phi\pi^-$, decay mode and for $\mathrm{B}_s^0 \to J/\psi\phi$ decay mode.

5.7 B_s^0 Angular Analysis

5.7.1 Signal Angular Model

As for the B_s^0 oscillation analysis, the construction of the full signal likelihood for the angular one as a function of ct and $\vec{\omega}$ comes directly from Equation 5.5.7

$$\mathcal{P}(ct, \sigma_{ct}, \vec{\omega}, \mathcal{D}\,|\vec{\xi}) = \frac{1 + \xi_1\mathcal{D}_1}{1 + |\xi_1|}\cdot\frac{1 + \xi_2\mathcal{D}_2}{1 + |\xi_2|}\cdot\big|\langle J/\psi\phi|\mathrm{B}_s^0\rangle\big|^2 + \frac{1 - \xi_1\mathcal{D}_1}{1 + |\xi_1|}\cdot\frac{1 - \xi_2\mathcal{D}_2}{1 + |\xi_2|}\cdot\big|\langle J/\psi\phi|\overline{\mathrm{B}}_s^0\rangle\big|^2 \tag{5.7.1}$$

From there we can use the work in Reference [15], where the time-dependent angular distribution of the transition $\mathcal{P}(\mathrm{B}_s^0 \to J/\psi\phi)$ is expressed as a sum over the possible bi-linear combinations of decay amplitudes

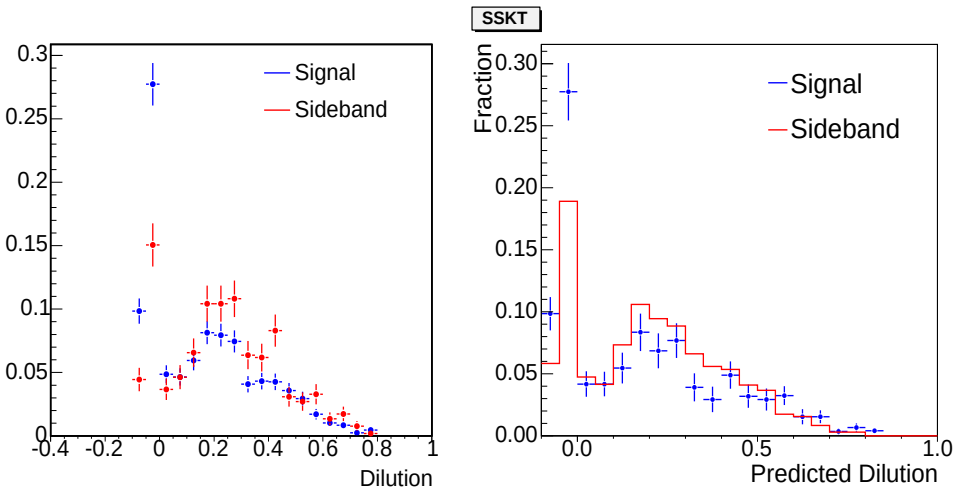

FIGURE 5.6: Distributions for the dilution $\mathcal{D}$ obtained from data for the signal sideband subtracted (red line) and the sideband (blue line) in the $B_s^0 \to D_s^- \pi^+$, with $D_s^- \to \phi\pi^-$ (left) and for $B_s^0 \to J/\psi\phi$ decay mode (right).

$$\begin{aligned}\sum_k \mathcal{O}^{(k)}(t) f^{(k)}(\theta,\phi,\psi) = &\left|A_0(t)\right|^2 f_1(\vec{\omega}) + \left|A_\parallel(t)\right|^2 f_2(\vec{\omega}) + \left|A_\perp(t)\right|^2 f_3(\vec{\omega}) \\ &+ Im\{A_\parallel^*(t)A_\perp(t)\} f_4(\vec{\omega}) + Re\{A_0^*(t)A_\parallel(t)\} f_5(\vec{\omega}) \\ &+ Im\{A_0^*(t)A_\perp(t)\} f_6(\vec{\omega}).\end{aligned} \tag{5.7.2}$$

The amplitudes time evolutions $|A_\alpha(t)|$, with $\alpha = \{0, \parallel, \perp\}$, were previously reported in Equations 1.4.3 and 1.4.4. The time angular probabilities for B_s^0 can be expressed as

$$\begin{aligned}\frac{d^4P(t,\vec{\omega})}{dtd\vec{\omega}} \propto &\left|A_0\right|^2 \mathcal{T}_+ f_1(\vec{\omega}) + \left|A_\parallel\right|^2 \mathcal{T}_+ f_2(\vec{\omega}) \\ &+ \left|A_\perp\right|^2 \mathcal{T}_- f_3(\vec{\omega}) + \left|A_\parallel\right|\left|A_\perp\right| \mathcal{U}_+ f_4(\vec{\omega}) \\ &+ \left|A_0\right|\left|A_\parallel\right| \cos(\delta_\perp) \mathcal{T}_+ f_5(\vec{\omega}) \\ &+ \left|A_0\right|\left|A_\perp\right| \mathcal{V}_+ f_6(\vec{\omega}),\end{aligned} \tag{5.7.3}$$

where the functions $f_1(\vec{\omega})...f_6(\vec{\omega})$ are defined in 1.4.2. The probability $\bar{P}$ for $\overline{B}_s^0$ can be obtained by substituting $\mathcal{U}_+ \to \mathcal{U}_-$ and $\mathcal{V}_+ \to \mathcal{V}_-$. The time-dependent term $\mathcal{T}_\pm$ is defined as

$$\begin{aligned}\mathcal{T}_\pm = e^{-\Gamma t} \times\ &[\cosh(\Delta\Gamma t/2) \mp \cos(2\beta_s)\sinh(\Delta\Gamma\ t/2) \\ &\mp \eta \sin(2\beta_s)\sin(\Delta m_s\ t)],\end{aligned} \tag{5.7.4}$$

where $\eta = +1$ for P and -1 for $\bar{P}$. The other time dependent terms $\mathcal{U}_\pm$ and $\mathcal{V}_\pm$ are defined as

$$\begin{aligned}\mathcal{U}_{\pm} = \pm e^{-\Gamma t} \times [&\sin(\delta_{\perp} - \delta_{\parallel}) \cos(\Delta m_s\ t) \\ &- \cos(\delta_{\perp} - \delta_{\parallel}) \cos(2\beta_s) \sin(\Delta m_s\ t) \\ &\pm \cos(\delta_{\perp} - \delta_{\parallel}) \sin(2\beta_s) \sinh(\Delta\Gamma_s\ t/2)], \\ \mathcal{V}_{\pm} = \pm e^{-\Gamma t} \times [&\sin(\delta_{\perp}) \cos(\Delta m_s\ t) \\ &- \cos(\delta_{\perp}) \cos(2\beta_s) \sin(\Delta m_s\ t) \\ &\pm \cos(\delta_{\perp}) \sin(2\beta_s) \sinh(\Delta\Gamma_s\ t/2)].\end{aligned} \tag{5.7.5}$$

In the case of a perfect detector we could assume an homogeneous acceptance distribution for the angles. Since CDF does not have uniform efficiency and performances in the tri-dimensional space, we expect not to have uniform acceptance for the transversity angles reconstructions. So it is crucial for the analysis to be able to unfold this effect, which is usually referred as "sculpting", using CDF jargon. In order to study the detector impact on the angles reconstructions, we generate a flat phase space Monte Carlo, uniform in all its angular distributions. The simulated events, generated with `BGenerator` and decayed using `EvtGen`, are then passed through the full-fledged detector simulation. Thus, we obtain a tri-dimensional acceptance curve $A(\vec{\omega})$, which accounts for the detector acceptance in the transversity angles reconstruction. The projections of $A(\vec{\omega})$ onto the three axes of the transversity basis are shown in Figure 5.7. The sculpting effect is finally taken into account by the multiplicative term $A(\vec{\omega})$ in the signal angular PDF. As for the lifetime bias introduced by the `TTT` when the signal likelihood is multiplied by an efficiency curve $\epsilon(t)$ (see Equation 5.4.2), in the angular analysis the signal likelihood is multiplied by the three dimensional acceptance curve $A(\vec{\omega})$:

$$\mathcal{P}(ct, \sigma_{ct}, \vec{\omega}, \mathcal{D}\ |\vec{\xi}) \cdot \mathcal{P}(\mathcal{D}) \rightarrow \frac{1}{\mathcal{N}} \mathcal{P}(ct, \sigma_{ct}, \vec{\omega}, \mathcal{D}\ |\vec{\xi}) \cdot \mathcal{P}(\mathcal{D}) \cdot A(\vec{\omega}) \tag{5.7.6}$$

Since the distribution were generated flat, the acceptance function can be interpreted as the probability to find an event at each position in the $\vec{\omega}$ space. The presence of this additional term bring some complication in the normalization process: both the angular distribution and the acceptance distribution are independently normalized, thus it is not assured their product will be. Finally, the integral of the signal angular likelihood can be calculated analytically integrating over the acceptance curve in a way that most part of the calculations can be performed only once. The details of this normalization procedure for the calculation of $\mathcal{N}$ are described in Appendix B.

5.7.2 Acceptance Function Correction

The simulation used to reproduce signal events and estimate the acceptance curve used does not reproduce the trigger dynamic prescaling. Moreover at the time the analysis was performed the complete description of the beam line alignment covered the data taking period up to 760 pb^{-1}. The data sample used for the angular anal-

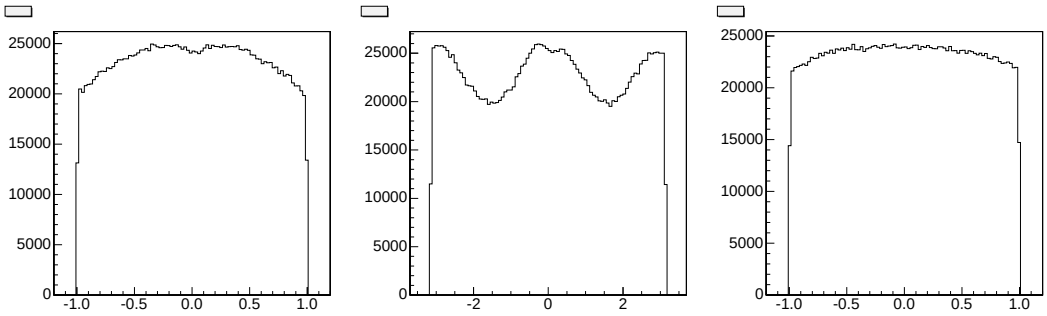

FIGURE 5.7: The distributions of the three angles describing the $B_s^0 \to J/\psi\phi$ candidates after reconstruction of the events simulated. The three distributions were generated flat.

ysis corresponds to an integrated luminosity of 1.35 fb^{-1}. To correct for this lack of knowledge, we compare different sensitive trigger variables for data and Monte Carlo and re-weight the simulation to reproduce the data before deriving the acceptance shape [74]. The distributions for the data in the signal region are obtained by subtracting the background using the sidebands events in the mass distribution. The procedure re-weights the following variables:

- Ratio of di-muon pairs triggered by CMU-CMU and CMU-CMX triggers, to correct the distribution of the muons in pseudorapidity η, which affects the acceptance shape.
- The number of candidates distributed in three different classes of muon p_T: both muons with $p_T > 3$GeV/c, both with $p_T > 2$GeV/c but at least on below $p_T < 3$GeV/c and everything else. The choice of this particular distribution is done in order to account the use of simultaneous triggers with different transverse momentum thresholds.
- The B_s^0 meson transverse momentum spectrum is taken from a histogram, to correct for the global kinematic of the events.

The re-weighting is applied in a consecutive fashion: we first correct for the trigger ratio, then for the transverse momentum classes and finally for the B_s^0 meson p_T spectrum. At each step the weights are calculated only after correcting for the previous effects. Although this correction procedure improves the agreement between data and simulated events, the choice of variables and the methodology could still be not exhaustive, therefore the remaining discrepancies are accounted in the evaluation of the systematics (see Appendix A.

5.7.3 Background Angular Model

Since the transversity angles do not have any intrinsic meaning for combinatorial background, we use a purely empirical model derived by analysing the angular distributions in the mass sidebands. The absence of mixing in the background makes the angular component separable from the lifetime likelihood part.

The background angular probability density function is written as:

$$\mathcal{P}_{\vec{\omega}}^{bkg}(\vec{\omega}) = P^{\theta_T} \cdot P^{\phi_T} \cdot P^{\psi_T}, \tag{5.7.7}$$

where:

$$\begin{aligned}
P^{\theta_T} &= \frac{1}{N_\theta} \left(1 + p_2^\theta \cos^2\theta + p_4^\theta \cos^4\theta\right) \\
P^{\phi_T} &= \frac{1}{N_\phi} \left(1 + \frac{1 - p_1^\phi \cos\phi}{2} + \frac{1 + p_2^\phi \cos 2\phi}{2} + \frac{1 + p_4^\phi \cos 4\phi}{2}\right) \\
P^{\psi_T} &= \frac{1}{N_\psi} \left(1 + p_1^\psi \cos\psi\right)
\end{aligned}$$

The parameters p_2^θ, p_4^θ, p_1^ϕ, p_2^ϕ, p_4^ϕ, and p_1^ψ then enter the fit as background parameters. Though this is a simple description, we find that it is adequate to describe the angular distributions in the sidebands. Figure 5.8 shows a binned fit to the sideband angular distributions using the above equations. The assumption of no correlation among the three angular variables is a source of possible concern. To test our model hypothesis, we plot two dimensional histograms of the angular distributions in the mass sidebands. In Figure 5.9, we show those histograms for the mass sidebands after the pre-selection cuts, and in Figure 5.10, we show the same distributions after Neural Network selection. We observe that after the Neural Network selection any dependence is gone within statistical fluctuation.

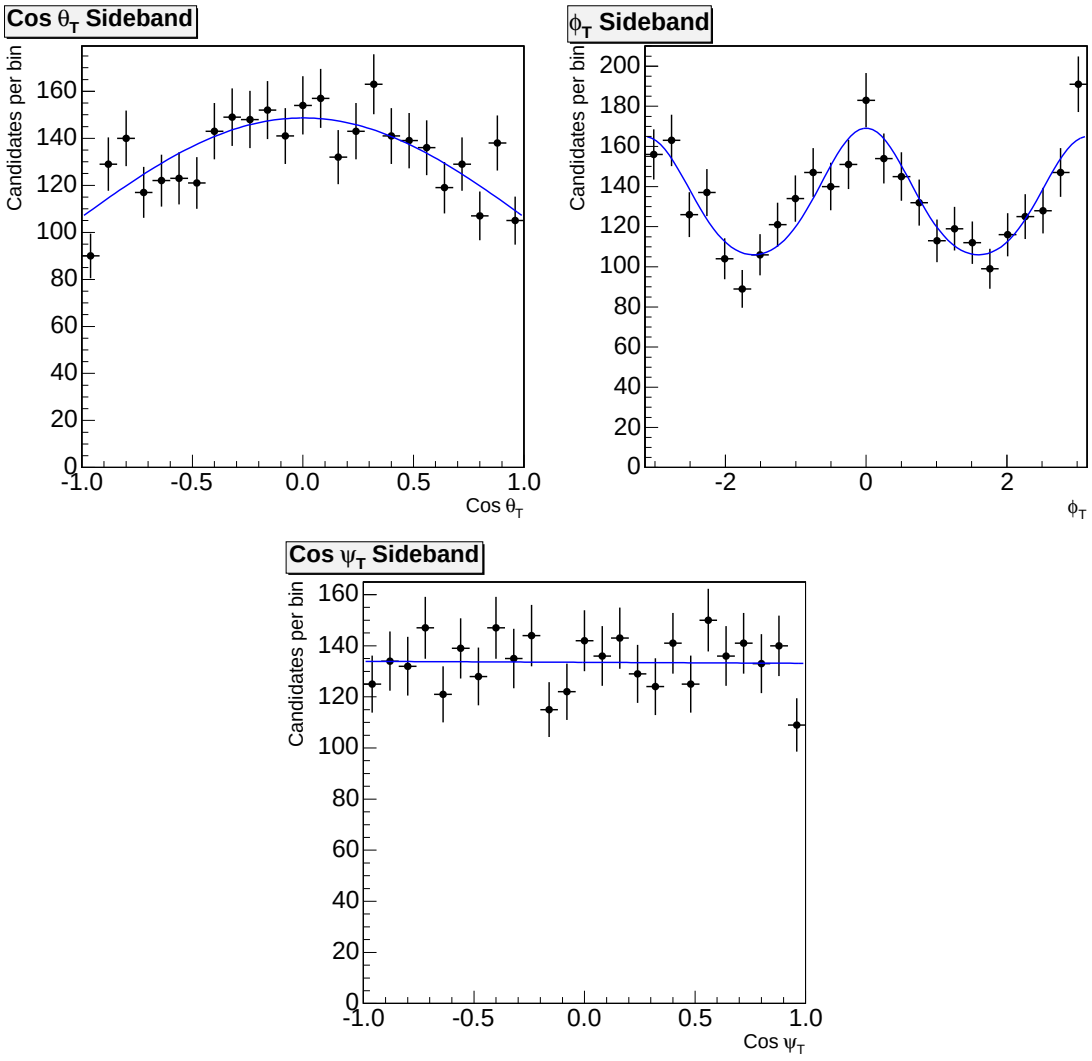

FIGURE 5.8: Angular distributions in the B_s^0 mass sidebands after Neural Network selection.

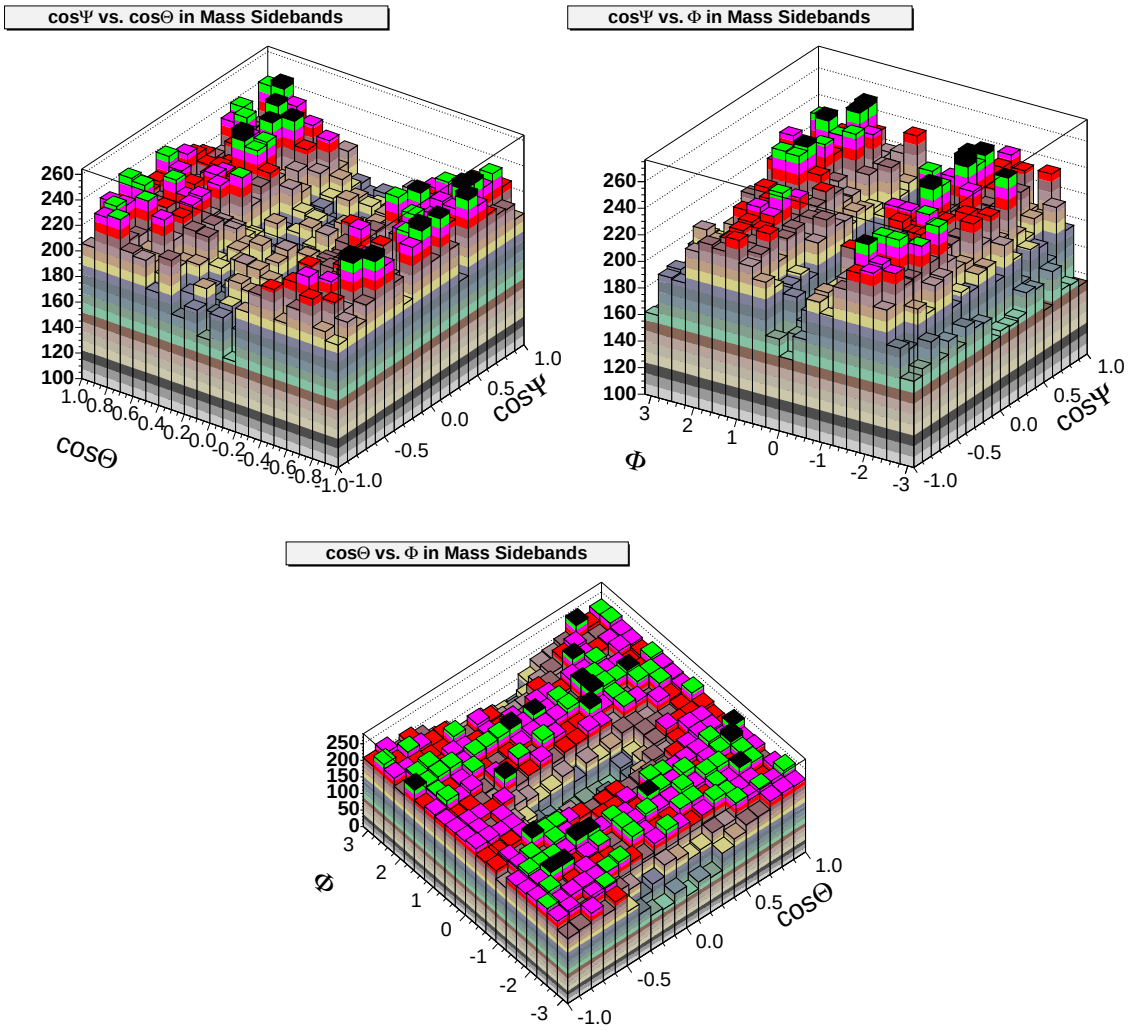

FIGURE 5.9: Two dimensional angular distributions in the B_s^0 mass sidebands after preselection for $B_s^0 \to J/\psi\phi$ decay mode.

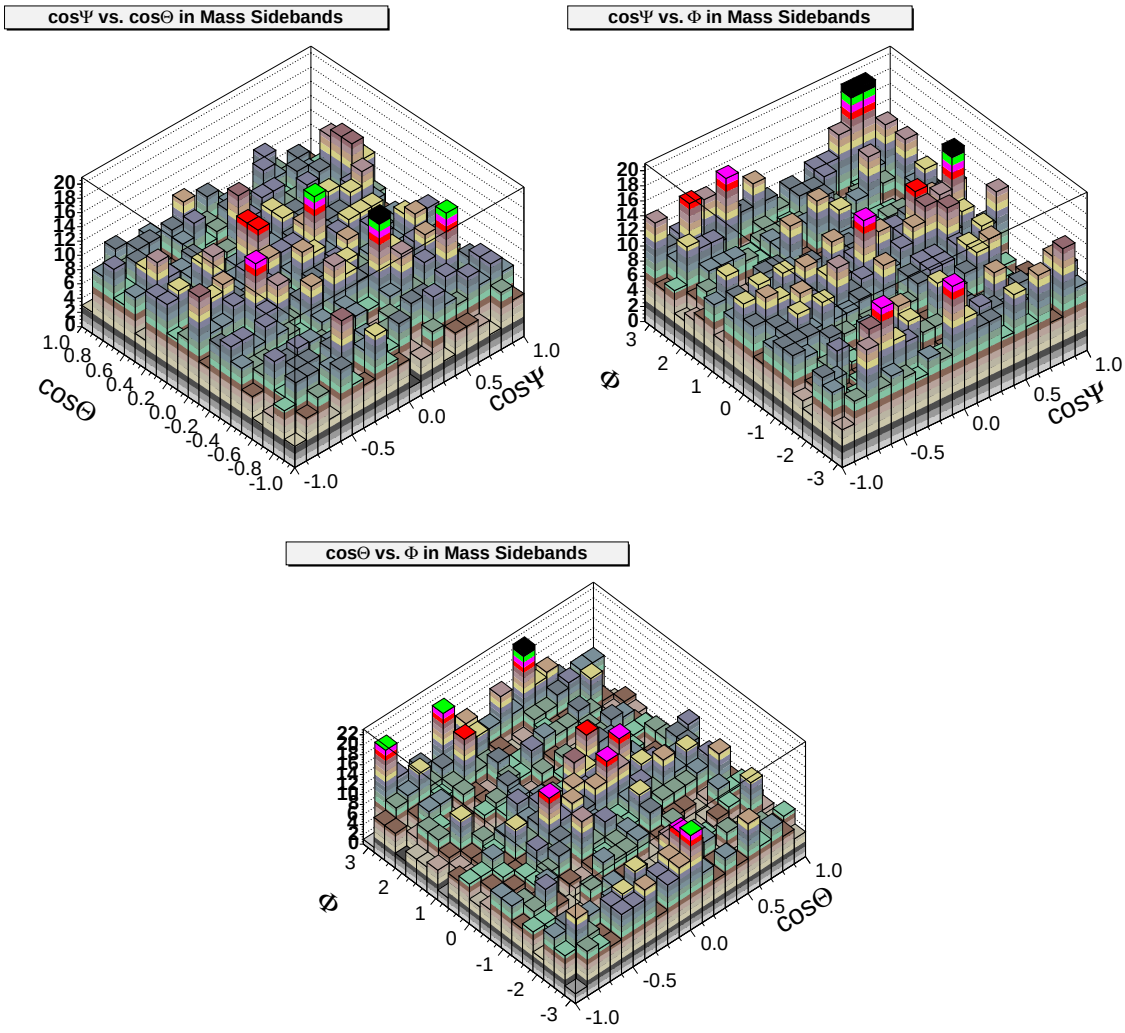

FIGURE 5.10: Two dimensional angular distributions in the B_s^0 mass sidebands after Neural Network selection for $B_s^0 \to J/\psi\phi$ decay mode.

5.7.4 Fitter Validation

Once the fitter framework is built, the likelihood has to be validated. The test of the maximum likelihood fitter is performed in order to accomplish several goals:

- validate the correctness of its implementation,
- test the statistical limits of the sensitivity to the fit parameters,
- investigate the behavior of the likelihood under various circumstances,
- detect any potential fit biases.

Pull distributions are a commonly used method to answer several of these questions. The procedure is as follows: we generate a large set of pseudo-experiments (also referred to as toy Monte Carlo), randomly polling the likelihood probability density function itself to assign event variables. For each of these pseudo-experiments, we perform a fit as we would on data. For each parameter A that is allowed to float in the fit, the pull distribution P is defined as

$$P = \frac{A^{fit} - A^{input}}{\sigma^A}. \tag{5.7.8}$$

In order to be able to claim that the likelihood fitter can measure a given quantity, P for that quantity should follow a Gaussian distribution with a mean at 0 and a width of 1.

Naturally, several complications arise. In some cases, there can be a hard physical limit on one side or both of the allowed parameter range. In these cases, we use asymmetric error evaluations in our data fit to account for this constraint. In our pull studies, we then adopt the convention that the pull for a given pseudo-experiment is calculated using the positive error in the denominator if the fit value is lower than the input value and the negative error otherwise.

In other cases, a degeneracy in the likelihood expression itself can lead to a fit value that is simply another equivalent minimum without being a physically different value (as far as the likelihood itself would be able to measure). In order to detect these instances, we then plot the fit values themselves from the sample of pseudo-experiments, to see if this is the cause of the imperfect pull distribution. Biases in the fit likelihood have been already observed in certain scenarios of the $B_s^0 \to J/\psi\phi$ angular analysis when performed without using the flavor tagging information, utilizing a sample of 1.7 fb^{-1} [74]. The next section describes with more accurate details the studies performed on toy Monte Carlo experiments to investigate for possible biases in the tagged $B_s^0 \to J/\psi\phi$ angular analysis.

Parameter	Initial value	Pull Mean	Pull Width	χ^2 Fit Prob
p_1^m	-0.4000	+0.046 ± 0.032	1.012 ± 0.023	0.97
S_{ct}	+1.3000	-0.090 ± 0.033	1.033 ± 0.023	0.79
f_+	+0.1000	+0.070 ± 0.032	1.009 ± 0.023	0.95
f_{++}	+0.0370	+0.028 ± 0.032	1.006 ± 0.023	0.02
f_-	+0.0250	+0.090 ± 0.032	1.002 ± 0.022	0.62
$c\tau_+$	+0.0082	-0.100 ± 0.032	1.022 ± 0.023	0.08
$c\tau_{++}$	+0.0455	-0.062 ± 0.033	1.052 ± 0.024	0.00
$c\tau_-$	+0.0140	-0.084 ± 0.033	1.028 ± 0.023	0.62
p_2^θ	-0.5800	+0.052 ± 0.032	1.006 ± 0.023	0.03
p_4^θ	+0.3200	-0.042 ± 0.032	1.003 ± 0.022	0.10
p_1^ϕ	+0.0330	+0.041 ± 0.033	1.056 ± 0.024	0.00
p_2^ϕ	+0.5400	+0.009 ± 0.032	1.020 ± 0.023	0.07
p_4^ϕ	+0.0230	-0.003 ± 0.032	1.028 ± 0.023	0.53
p_1^ψ	+0.0150	+0.002 ± 0.032	1.025 ± 0.023	0.03
ϵ_{SST}^B	+0.7100	-0.000 ± 0.031	0.985 ± 0.022	0.48
$\mathcal{D}_{SST}^B$	+0.0150	-0.032 ± 0.031	0.972 ± 0.022	0.18
ϵ_{OST}^B	+0.9800	+0.032 ± 0.031	0.985 ± 0.022	0.93
$\mathcal{D}_{OST}^B$	+0.0100	-0.045 ± 0.031	0.993 ± 0.022	0.08

TABLE 5.1: Background pulls fit results for the $B_s^0 \to J/\psi\phi$ angular analysis with $10,000$ events, $1,000$ pseudo-experiments.

Tagged Toy Monte Carlo

As general attitude, while validating the fitter, we move from optimistic to pessimistic tests of the fitter. We begin with the area of least a priori concern: the background parametrization. Table 5.1 shows the results of pulls studies with $10,000$ background events and no signal. We note that the parameters that are least well described are the ones describing the various lifetime components in the background. This is to be expected, especially when the fraction of the respective component is small. Nevertheless, the various empirical background models hold up well in these pull studies, with the worst bias being 0.1σ away from 0. Figure 5.11 shows the pull distributions for each parameter.

We then evaluate pulls with high statistics ($10,000$ events per pseudo-experiment) for a signal-only sample. We expect that the input value of ϕ_s will make a difference here. We begin our investigation of this effect by generating pseudo-experiments at a large, and then a small value of β_s, respectively for a New Physics or Standard Model hypothesis. Table 5.2 shows the results for $\beta_s = 0.45$ rad. When generating pseudo-experiments with $\beta_s = 0.45$, we see that the pulls behave relatively well. The strong phases $\delta_\parallel$ and $\delta_\perp$ display some non-Gaussian behavior, but this was expected. This is caused by the non-parabolic behavior of the likelihood in those parameters, which was already observed in untagged analysis at CDF [74].

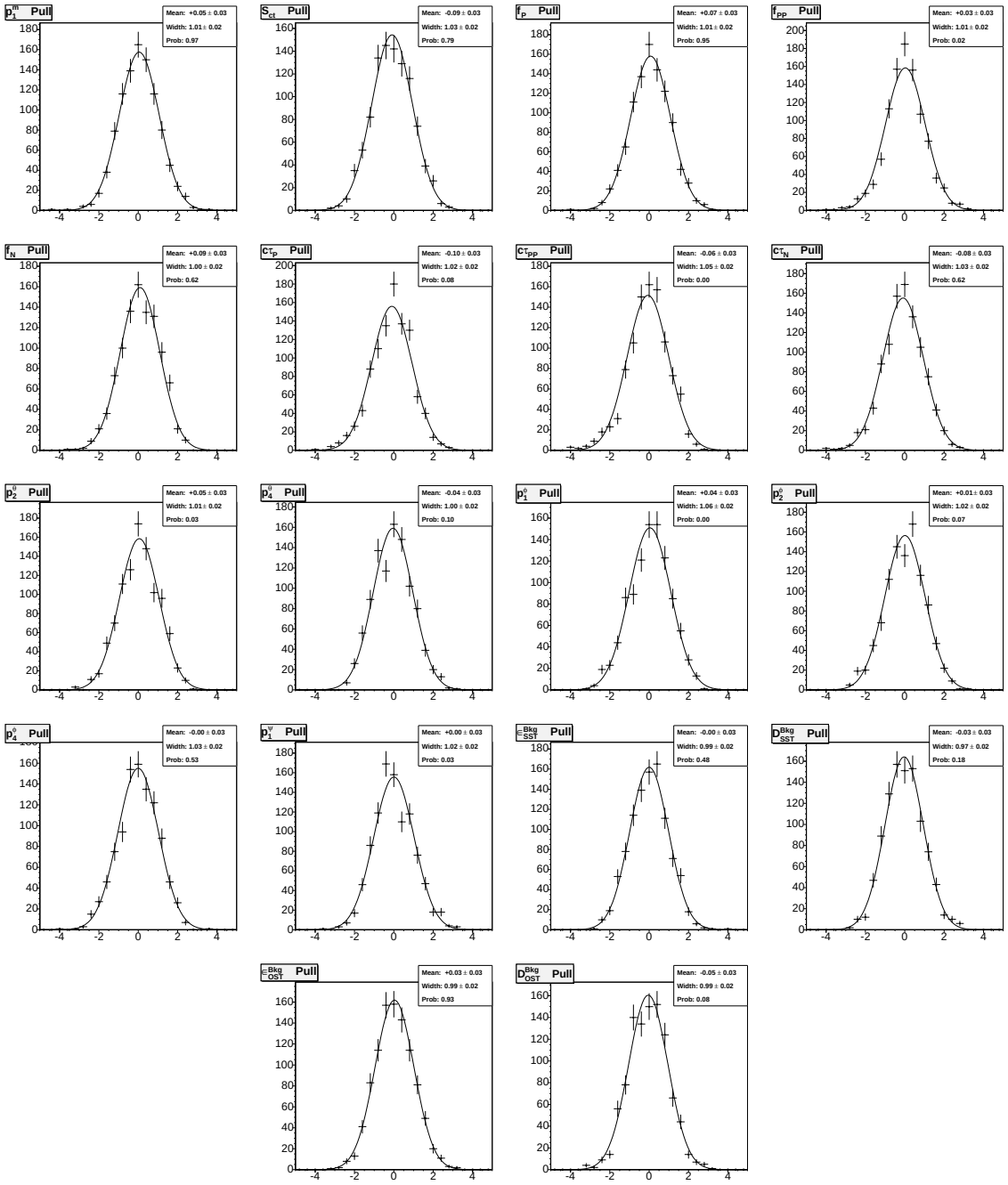

FIGURE 5.11: Background pulls distributions with $10,000$ events.

Parameter	Initial Value	Pull Mean	Pull Width	χ^2 Fit Prob
$c\tau$	+0.0450	+0.019 ± 0.039	1.022 ± 0.028	0.01
$\Delta\Gamma$	+0.2200	-0.003 ± 0.042	1.086 ± 0.030	0.46
$\left\|A_0\right\|^2$	+0.5500	+0.034 ± 0.039	1.018 ± 0.028	0.33
$\left\|A_\parallel\right\|^2/(\left\|A_\parallel\right\|^2+\left\|A_\perp\right\|^2)$	+0.5555	-0.025 ± 0.038	0.985 ± 0.027	0.62
ϕ_s	-0.9000	+0.089 ± 0.041	1.055 ± 0.029	0.01
$\delta_\parallel$	+2.7500	+0.036 ± 0.043	1.111 ± 0.030	0.01
$\delta_\perp$	+0.8000	-0.036 ± 0.044	1.149 ± 0.031	0.04

TABLE 5.2: Signal pulls fit results for the $B_s^0 \to J/\psi\phi$ angular analysis with $10,000$ events, input $\beta_s = 0.45$.

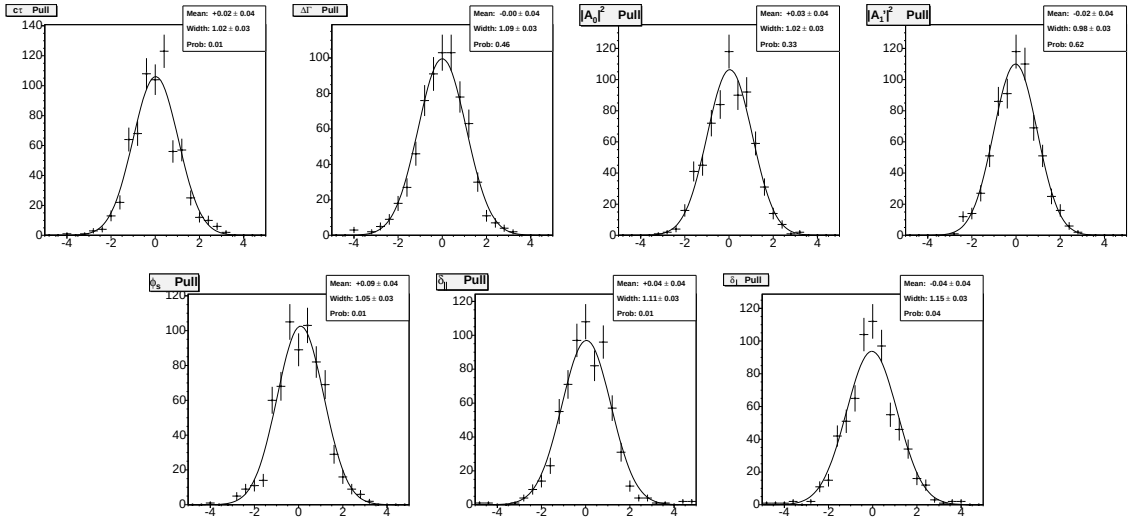

FIGURE 5.12: Signal pulls distributions for the $B_s^0 \to J/\psi\phi$ angular analysis with $10,000$ events, input $\beta_s = 0.45$.

Next we generate pseudo-experiments with $\beta_s = 0.05$. Table 5.3 and Figure 5.13 summarize the results for $\beta_s = 0.05$. We see that this introduces several biases and some additional non-Gaussian behavior for the likelihood. Additional problem arises observing that the bias magnitude in the tested parameters varies as the input values in the simulated experiments vary, therefore it is not possible to simply subtract it from the final results returned by the minimization process of the likelihood.

Parameter	Initial Value	Pull Mean	Pull Width	χ^2 Fit Prob
$c\tau$	+0.0450	-0.251 ± 0.035	0.967 ± 0.025	0.74
$\Delta\Gamma$	+0.2200	+0.245 ± 0.037	1.015 ± 0.026	0.41
$\left\|A_0\right\|^2$	+0.5500	-0.007 ± 0.036	0.994 ± 0.025	0.90
$\left\|A_\parallel\right\|^2/(\left\|A_\parallel\right\|^2+\left\|A_\perp\right\|^2)$	+0.5555	+0.027 ± 0.036	0.998 ± 0.025	0.64
ϕ_s	-0.1000	+0.033 ± 0.043	1.200 ± 0.031	0.01
$\delta_\parallel$	+2.7500	-0.058 ± 0.038	1.067 ± 0.027	0.00
$\delta_\perp$	+0.8000	+0.209 ± 0.048	1.323 ± 0.034	0.00

TABLE 5.3: Signal pulls fit results for the $B_s^0 \to J/\psi\phi$ angular analysis with $10,000$ events, input $\beta_s = 0.05$.

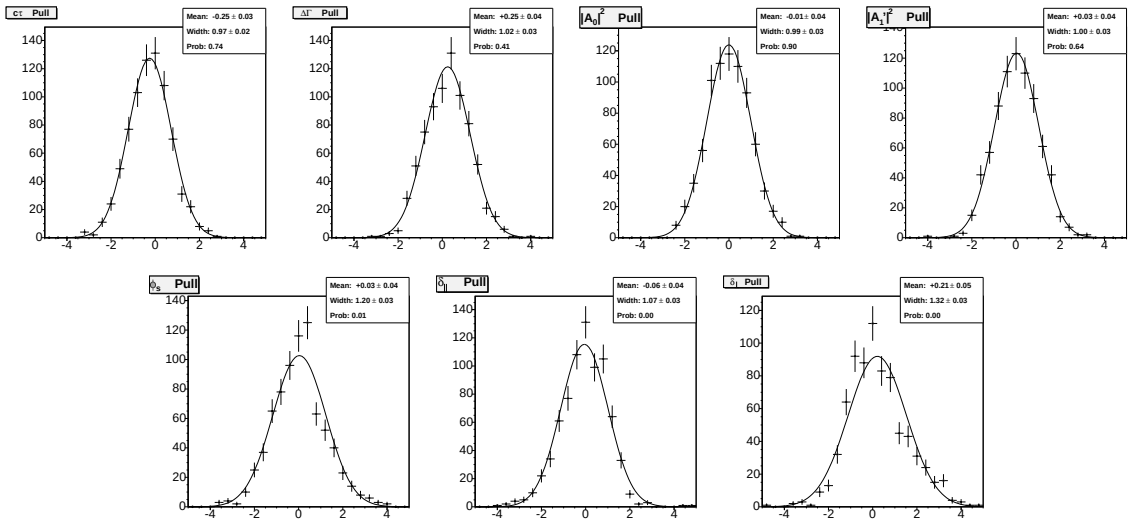

FIGURE 5.13: Signal pulls distributions for the $B_s^0 \to J/\psi\phi$ angular analysis with $10,000$ events, input $\beta_s = 0.05$.

Likelihood Invariances

Analyzing carefully the likelihood, one exact symmetry is observed to be present in it, while performing the angular analysis. The transformation under which the likelihood is invariant is:

$$\begin{aligned} \beta_s &\to \frac{\pi}{2} - \beta_s \\ \Delta\Gamma_s &\to -\Delta\Gamma_s \\ \delta_\parallel &\to 2\pi - \delta_\parallel \\ \delta_\perp &\to 2\pi - \delta_\perp \end{aligned} \tag{5.7.9}$$

This symmetry holds regardless of the statistical power of the data. Since this ambiguity is not yet resolved, the result should be given in terms of the two possible minima. At first approximation, one possible approach would be to remove the two-fold symmetry and separately study the two mentioned minima. This can be performed by limiting the variation of some parameters in proper ranges. For instance a possibility exploited at CDF was to box $\delta_\parallel$ in the range $\delta_\parallel \in [0, \pi)$. Using such prescription, several tests have been performed in order to study the likelihood behaviour, if the symmetry is removed. In Figures 5.14 and 5.15 we report the results from the contour scan on simulated experiments generated with a realistic mixture of signal and background. The toys Monte Carlo are generated with different input values for $\beta_s = 0.8,\ 0.4,\ 0.02$ rad, using a sample of $12,000$ signal events with realistic tagging efficiencies and dilutions. As we observe from the Figures, a problematic set of pseudo-invariances appears in the likelihood, which is manifested as a second local minimum in the likelihood contour. These pseudo-invariances are a direct consequence of a limited statistical power for the analysis considered. In fact these are not exact invariances, which were expected to be removed with the boxing in $\delta_\parallel$, but approximate ones. In conclusion the limited statistical power available makes not possible to resolve the interference terms well enough to distinguish these approximate invariances.

Conclusions on the Angular Analysis

In the process of validating the signal likelihood, a possible inherent ambiguity appeared in the likelihood, in addition to biases for certain input values of the parameters. The likelihood for the angular analysis on the $\mathrm{B}_s^0 \to J/\psi\phi$ decay mode shows a *non-Gaussian* behaviour in the parameters estimations. The immediate consequence is the impossibility to quote a reliable point estimate for the measurements of the quantities of interest, among which the CP violating phase β_s. This is one of the crucial point of the whole angular analysis: the results will be given in terms of a confidence region and not using the classical point estimate returned by the fit minimization. The exact prescription followed to determine these contours, and thus the final results, are discussed in Chapter **??**.

In this Chapter we has been reviewing the fitter frameworks used for the two analyses reported in this dissertation. By combining mass, lifetime, tagging and, in the angular analysis case, angles information, the statistical power of the analyses is maximed. The results obtained with such frameworks are described in the next Chapter. The first observation of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ *oscillation and a precise measurement of* Δm_s *will be there presented. Furthermore the first angular analysis on the* $\mathrm{B}_s^0 \to J/\psi\phi$ *decay mode using flavor tagging for the determination of the CP violating phase* β_s *is reported: the result is given in terms of a confidence region in the space* $\Delta\Gamma_s - \beta_s$*, because of the non-Gaussian likelihood behaviour with the current level of statistics which does not allow to reliably quote a point estimate for the parameters of interest.*

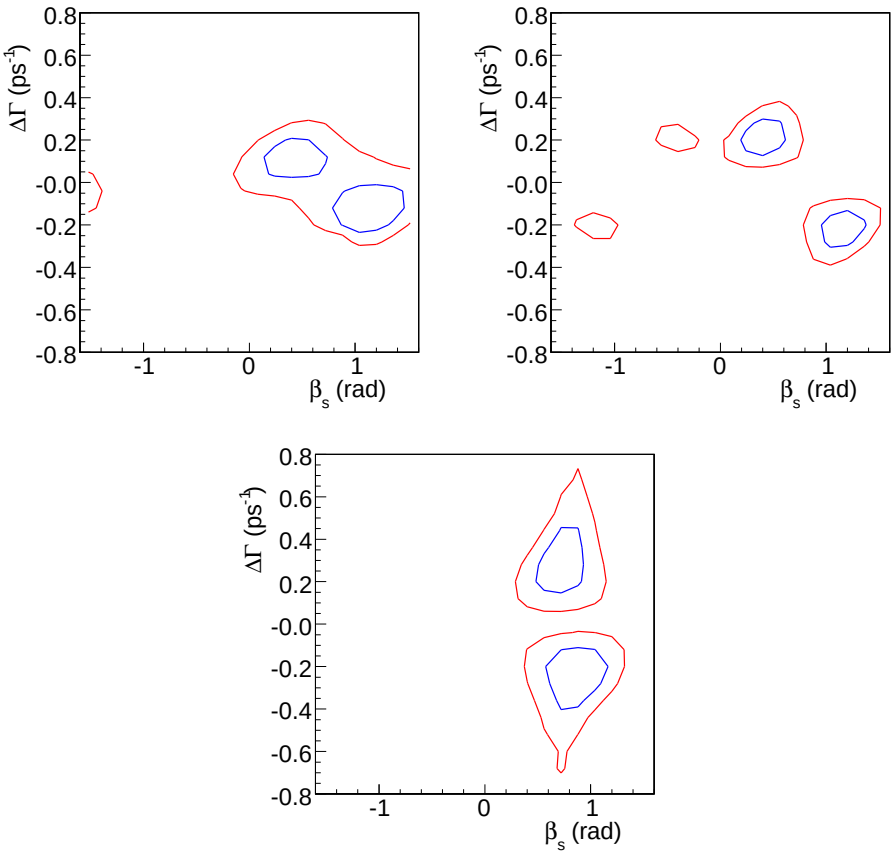

FIGURE 5.14: Two toy MC likelihood contours in the plane $\beta_s - \Delta\Gamma_s$ with β_s generated at 0.40 rad (up) and one toy MC likelihood contour with $\beta_s = 0.80$ rad (bottom). Only the 7 signal parameters (amplitudes, strong phases, lifetime, $\Delta\Gamma_s$ and β_s) are allowed to float in the fit. The blue curves correspond to 2.31 units above the minimum in the likelihood ($\sim 68\%$ confidence level region) and the red curves to 5.99 units above the same minimum ($\sim 95\%$ CL region). The fit is performed boxing $\delta_{\parallel} \in [0, \pi)$.

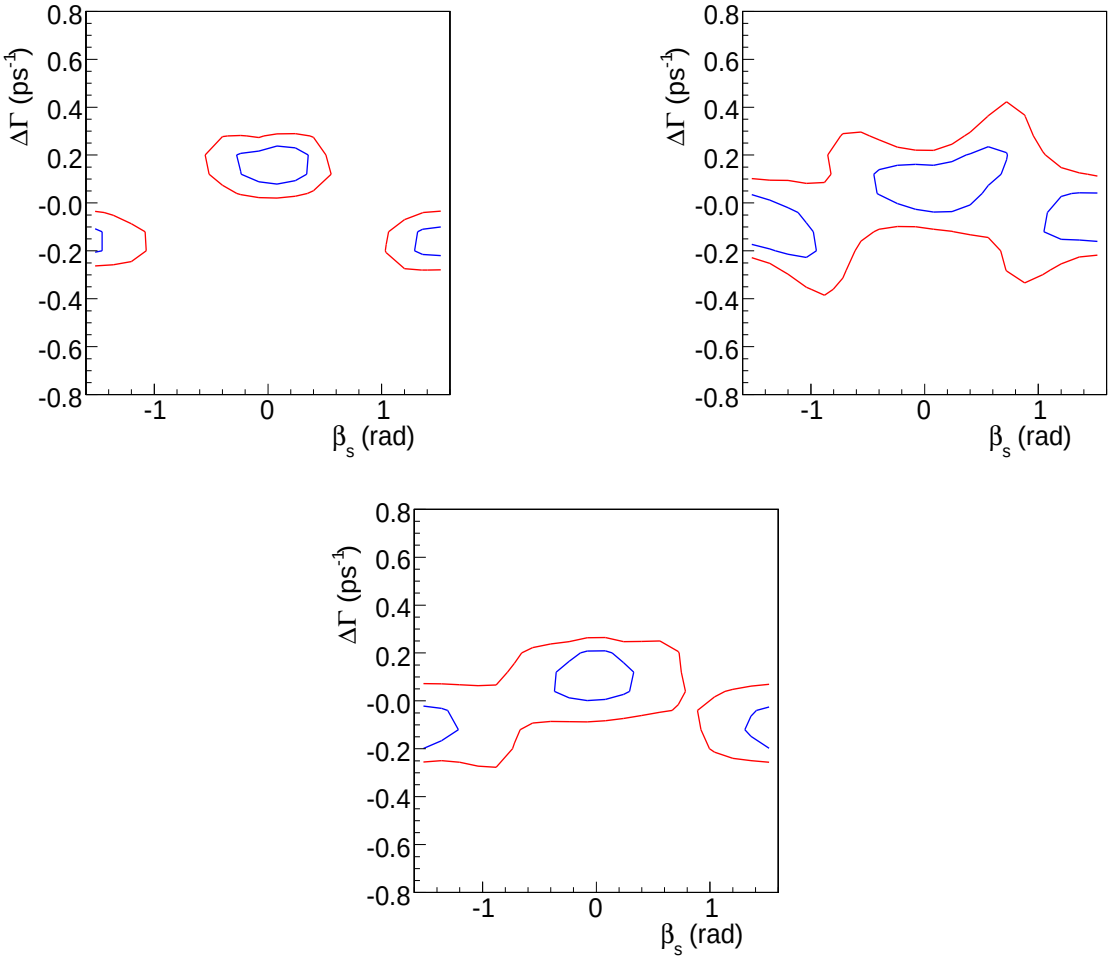

FIGURE 5.15: Toy MC likelihood contours in the plane $\beta_s - \Delta\Gamma_s$ with β_s generated at the Standard Model expectation value $\beta_s = 0.02$ rad [75]. Only the 7 signal parameters (amplitudes, strong phases, lifetime, $\Delta\Gamma_s$ and β_s) are allowed to float in the fit. The blue curves correspond to 2.31 units above the minimum in the likelihood ($\sim 68\%$ confidence level region) and the red curves to 5.99 units above the same minimum ($\sim 95\%$ CL region). The fit is performed boxing $\delta_{\parallel} \in [0, \pi)$.

Chapter 6

Final Results

In this Chapter the validated fit frameworks are applied to the B_s^0 samples. The final results for the B_s^0 oscillation frequency measurement as well as the $\Delta\Gamma_s$ and β_s measurements for the angular analysis on $B_s^0 \to J/\psi\phi$ decays, are presented. The systematics uncertainties studies and the impact of the analyses on the Standard Model flavor interaction sector are also described.

6.1 Observation of $B_s^0 - \overline{B}_s^0$ Oscillations

In this section we review the combined results of the fully, partially reconstructed hadronic and semileptonic samples with the amplitude scans in the frequency domain and evaluate the significance of the oscillation signature. The precise measurement of Δm_s is finally described.

6.1.1 Systematic Uncertainties

Systematic uncertainties concern the significance of the amplitude scan for the search on an oscillation signal, as well as the measurement of the oscillation frequency, Δm_s. The studies of the systematic effects on the amplitude scan are important for setting lower limits on Δm_s, while they become almost negligible when there is significant sensitivity for the oscillation measurement. The systematic effects studied for the amplitude scan and for the Δm_s measurement are reviewed in the next Sections. It is relevant to notice that the results are statistically limited, both in the amplitude scan sensitivity region and in the fit for the mixing frequency.

Systematic Uncertainties on Amplitude

The methodology used for the systematic uncertainties evaluation on the amplitude scan follows its original formulation [57]. For each effect two large ensembles of toy MC samples are generated: one to simulate the potential mismodel and another with the nominal model which is applied on data. The fit of each pseudo-experiment return an amplitude measurement with its uncertainty. If we use the label "0" for the biased

mismodel results of the "i-esim" toy MC $(\mathcal{A}_0^i, \sigma_{\mathcal{A},0}^i)$ and "1" for the results of the default model $(\mathcal{A}_1^i, \sigma_{\mathcal{A},1}^i)$, the systematic uncertainty is estimated to be

$$\sigma_{syst}^i = (\mathcal{A}_1^i - \mathcal{A}_0^i) + (1 - \mathcal{A}_0^i) \cdot \frac{\sigma_{\mathcal{A},1}^i - \sigma_{\mathcal{A},0}^i}{\sigma_{\mathcal{A},0}^i}. \tag{6.1.1}$$

The distribution of σ_{syst}^i is used for the evaluation of the systematic uncertainty. In case of binary shift effects in the fitting model, the mean value of the σ_{syst}^i is taken as systematic uncertainty. Otherwise, when the fit parameters are continuously varied across a range of values, the width of the σ_{syst}^i distribution is utilized for the systematic evaluation. The various source of systematic effects estimated are:

- **Dilution Scale Factor:** The uncertainty on the magnitude of the predicted dilution for each event is a source of systematic uncertainty. The potential variation is taken from the scale factor calibration analyses for OST and SSKT, illustrated in Sections 4.2.1 and 4.3.6. These variations are propagated into the toy MC samples and fitted with the nominal dilution modeling. The dominant contribution is then given by the SSKT scale factor which is determined with a 14% precision, while the OST scale factor is known with a very good precision. The size of its contribution is $\simeq 10\%$ on the amplitude for any value of Δm_s.
- **Correlated Taggers:** The underlying assumption for the B_s^0 mixing analysis is the tagging decisions for the OST and SSKT are uncorrelated. The bias that would result from taggers whose decisions are correlated at 10% level is studied and introduce an uncertainty of about 8% on the amplitude for $\Delta m_s = 15$ ps^{-1}.
- **Dilution Templates:** The probability distributions for dilution which are used in the fit model for signal and background are known with a finite statistical precision. Toy MC samples are generated with variations within their statistical uncertainties and fit with the nominal templates applied in the amplitude scan.
- **Dilution of Cabibbo-Suppressed Component:** While the $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+ (\pi^- \pi^+)$ decays are self-tagging, the Cabibbo-suppressed $\mathrm{B}_s^0 \to \mathrm{D}_s^- \mathrm{K}^+ (\pi^- \pi^+)$ and $\mathrm{B}_s^0 \to D_s^+ \mathrm{K}^- (\pi^- \pi^+)$ decays receive contributions from two different tree-level amplitudes which have opposite sign in the final state particles. This component is therefore not self tagging. The nominal model uses the same dilution of Cabibbo-favored component as well as for the Cabibbo-suppressed counterpart. A biased MC ensemble is generated where each experiment has the dilution damped by a random factor between zero and one. The same technique is used to study the effect of assigning wrong dilution to Λ_b component. The large contribution comes from the Cabibbo-suppressed B_s^0 decays and corresponds to about 4% uncertainty for the amplitude at $\Delta m_s = 15$ ps^{-1}.
- σ_{ct} **Scale Factor:** The global scale factor σ_{ct} parametrization results in a calibration which has a residual of few percent (see Section 4.1.3). The possible over- or under-estimation of the uncertainty on the proper time of the B_s^0 signal is accounted by generating toy experiments with $\pm 4\%$ variation in the σ_{ct} scale factor. This systematic uncertainty increases with the sampled Δm_s and it is of about 5% for the amplitude at $\Delta m_s = 15$ ps^{-1}.

- **Non-Negligible $\Delta\Gamma_s/\Gamma_s$:** The fit on the B_s^0 mixing does not include the effect on the lifetime difference between the heavy and light mass $B_{s,H/L}$ eigenstates. On the other hand the value of $\Delta\Gamma_s/\Gamma_s$ different from zero affects the proper decay time fit and can introduce a bias. Toy MC are generated with a $\Delta\Gamma_s/\Gamma_s = 0.02$ and fitted with the nominal fit, assuming it is zero. The effect is of about 5% for the amplitude at $\Delta m_s = 15$ ps^{-1}.
- **Detector Resolution Function:** The default fit model assumes that the detector resolution function for proper decay time is a Gaussian. The high statistic samples used for the calibration indicate that a double Gaussian plus symmetric exponential tails provide a more precise resolution model. The systematic associated with the use of a single Gaussian to model the detector resolution is about 4% for the amplitude at $\Delta m_s = 15$ ps^{-1}.
- **Selection Bias Curves:** When performing a lifetime analysis using the displaced track trigger (TTT), the proper-decay-time efficiency curves $\epsilon(t)$ are of primary importance. The systematic is evaluated by fitting the same toy MC sample with the default efficiency curve $\epsilon(t)$ and with a modified efficiency function $\epsilon'(t)$, obtained by changing τ, the world average for the B_s^0 lifetime, within its uncertainty and properly re-weighting the BGENERATOR-MC sample to calculate the function. B_s^0 occur on such a short time scale to make the analysis insensitive to the $\epsilon(t)$ parametrization.
- **Probability Templates for σ_{ct}:** The likelihood for the hadronic sample does not utilize separate σ_{ct} probability distributions for its signal and background components, because they are very similar. This could introduce a bias which is estimated by fitting the toy MC sample with and without the σ_{ct} distributions. The systematic associated is of about $\simeq 4\%$ for the amplitude at $\Delta m_s = 15$ ps^{-1}.
- **Mixing of B^0 and Non-Signal Satellites:** The nominal fit model assumes that the small contribution of B^0 and partially reconstructed modes which are not treated as signal do not oscillate. Toy MC samples are generated including the mixing for these components. These samples are fit with the default version and with a version which accounts for oscillations. The effect gives a negligible contribution to the total systematic uncertainty on the amplitude.
- **Sample Composition and Mass Models:** Several systematic uncertainties are assigned for sample composition. Various toy MC ensembles are generated to address the uncertainty on the relative fractions of Cabibbo-suppressed $B_s^0 \to D_s^- K^+$ and the so-called "reflection" components of Λ_b^0 and B^0. The size is about 1% for the amplitude at $\Delta m_s = 15$ ps^{-1}. In the partially reconstructed mode an additional study is performed on the effect of the incomplete knowledge of relative signal fractions $B_s^0 \to D_s^- X$. Toy MC are generated as composed by only one signal component and performing the fit as it were composed only by the other component. In such way the procedure simulates a 100% confusion between the components. Moreover, in the hadronic partially reconstructed case toy MC are generated with backgrounds which rise more rapidly in the region of partially reconstructed signal candidates. The fit with the nominal method treats this background events as signal. The effect is responsible for

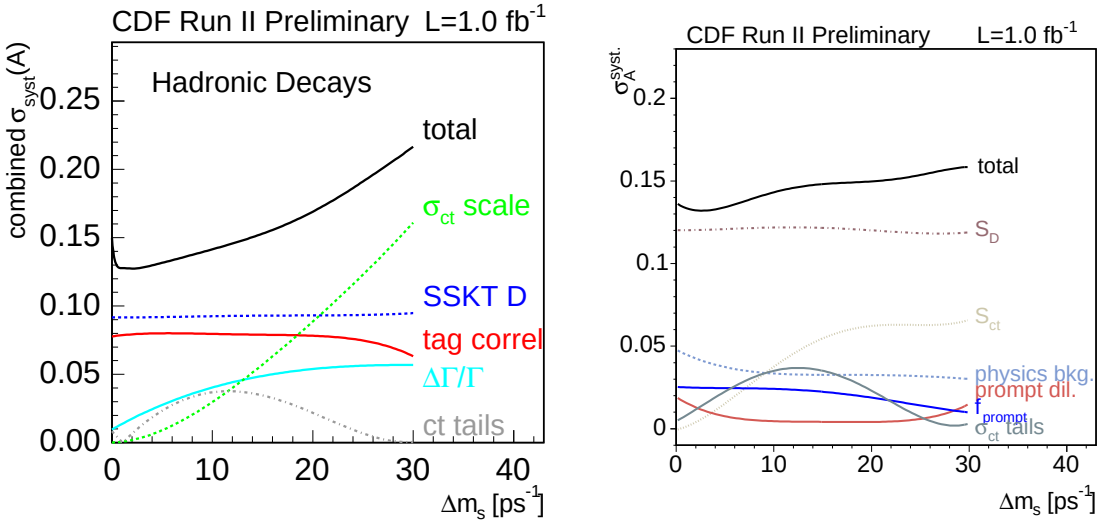

FIGURE 6.1: Systematic uncertainties on $\mathcal{A}$ as a function of Δm_s in the hadronic (left) and semileptonic (right) amplitude scans.

a 2% uncertainty for $\Delta m_s = 15$ ps^{-1}. For the semileptonic B_s^0 decays, two additional sources of systematics are considered. A systematic effect arising from the combinatorial background parametrization derived from the D_s^- mass sideband: a different set of sidebands, obtained by shifting the bounds of the sideband windows in ± 50 MeV/c^2 from the nominal values. The second source of bias comes from the uncertainty in the fraction of false lepton background, which is obtained from a fit to the $m_{\ell D_s^-}$ distribution. A toy MC ensemble is generated with the nominal value for the false lepton fraction. The results from the default fit are compared to a fit of the same sample having the fraction for the false lepton in the background shifted of $\pm 1\sigma$ from its nominal value. The size of all these systematics is below 5% for $\Delta m_s = 15$ ps^{-1}.

Figures 6.1 shows the behaviour of the systematics for increasing value of Δm_s, for the amplitude scan method.

Systematic Uncertainties on Δm_s

Systematic effects on Δm_s are estimated after the Δm_s measurement is performed. The systematic uncertainties presented in the previous section are relevant for the amplitude scan method, but not for the extraction of Δm_s. In fact the same potential source of uncertainties have been tested for an associated bias in the oscillation frequency, but they are found to be negligible. The three primary sources of systematic uncertainty on Δm_s are related to the absolute proper time scale of the CDF detector. The only non-negligible contribution coming from the partially reconstructed modes arises from the relative signal fractions. All the systematic effects are summarized as follow:

- **Silicon Detector Alignment:** The proper-decay time measurement is affected by the imperfect alignment of silicon detector. This effect is tested by introducing distortions of the order of $\simeq 50\mu m$ applied to radial displacement, bowing and rotation of silicon sensors within tolerances from a physical survey of the detector. The B mesons decays utilized for the study are $B^0 \to J/\psi K^{*0}$, with $J/\psi \to \mu^+\mu^-$ and $K^{*0} \to K^+\pi^-$, $B^+ \to J/\psi K^+$, with $J/\psi \to \mu^+\mu^-$, and $B_s^0 \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$. The maximum lifetime bias is found to be of $0.1\mu m$, which corresponds to 0.2% uncertainty on the proper decay time. A systematic error contribution of 0.04 ps^{-1} is therefore assigned to the measured Δm_s value.
- **Track Fit Bias:** Mis-measurements of the track curvature may be propagated into the calculations of proper decay time, via the transverse decay length L_{xy} of track vertices or via the measured B momentum. The sign of the bias depends on whether the tracks curve towards or away from each other. the bias has been studied in the simulation of the COT and silicon detectors: an overall systematic shift is found to be $1.3\mu m$, corresponding to 0.3% in proper-time scale or 0.05% for the Δm_s measurement.
- **Primary Vertex Bias:** Mis-measurements of primary vertex position results in mis-measurements of the transverse decay length and, therefore, of the proper decay time. The bias is estimated in data samples of fully reconstructed B^+ and B^0 events by comparing the primary vertex position to that of the average beam position. The maximum bias is found to be $1.0\mu m$ in the reference frame of the detector, which corresponds to 0.02 ps^{-1} shift to Δm_s in toy MC.
- **Hadronic k-Factors:** The dominant partially reconstructed $B_s^0 \to D_s^-\rho^+$ and $B_s^0 \to D_s^{*-}\pi^+$ channels have similar kinematics and qualitatively similar models. For the Δm_s measurement, the only significant difference is the modeling of the k-factor distributions. The not correct use of the relative fractions of partially reconstructed components could produce a shift in the fitted value of Δm_s. The maximum effect can be studied by fitting a component with the k-factor distribution and the ct efficiency curve of the other. Afterwards, this fit is compared with the one having the correct weight functions. Since there is no practical belief for thinking that the model might be so incorrect, 50% of the induced bias is used as a conservative systematic uncertainty. The final uncertainty contribution is 0.03 ps^{-1}. This error is assigned only for the measurements from the partially reconstructed hadronic sample.

Table 6.1 reports the values for the individual systematic effects on the Δm_s measurement.

6.1.2 Measurement of Δm_s

Amplitude Scan

The traditional utility of the amplitude scan lies in the possibility to combine independent scans. In the B_s^0 mixing analysis described in this dissertation we combine

Source	Value [ps^{-1}]
Amplitude Scan Effects	< 0.01
Silicon Detector Alignment	0.04
Track Fit Bias	0.05
Primary Vertex Bias	0.02
Hadronic k-Factor	0.03
Total	0.07

TABLE 6.1: Summary of systematic uncertainties on the Δm_s measurement.

the likelihoods of the independent analyses on each sample and examine the variations of the total likelihood:

$$ln\mathcal{L}_{TOT}(\mathcal{A}, \Delta m_s) = \sum_{samples} ln\mathcal{L}_s(\mathcal{A}, \Delta m_s) \tag{6.1.2}$$

The amplitude scan method results in a series of fits performed at increasing Δm_s values. At each fixed Δm_s the likelihood $-2ln\mathcal{L}_{TOT}(\mathcal{A}, \Delta m_s)$ is minimized providing, as results, the measurement of an amplitude with its uncertainty ($\mathcal{A}$, $\sigma_{\mathcal{A}}$). We remind the definition of sensitivity for this method: the value of the frequency for which a measured null amplitude $\mathcal{A} = 0$ would imply the exclusion of $\mathcal{A} = 1$ at desired confidence level (CL). The chosen confidence level is 95% which corresponds to set the condition $1.645\sigma_{\mathcal{A}} = 1$. The signature for a mixing signal is an amplitude consistent with unity and inconsistent with zero. Figure 6.2 depicts the amplitude scan for the hadronic and semileptonic samples with only the statistical error reported. The hadronic amplitude scan shows a clear signature, while the semileptonic sample is sufficient to set a double-sided limit on Δm_s at 95% C.L. The sensitivity of the analysis in the hadronic samples reaches 30.7 ps^{-1}, while in the semileptonic samples is 19.4 ps^{-1}. Figure 6.3 shows the amplitude scans in the different data samples and separately utilizing the Same Side or the Opposite Side Tagging. The sensitivity, evaluated only on statistical uncertainty, is 30.3 ps^{-1} for the Same-Side Tagging only and 25.5 ps^{-1} for the Opposite Side Tagging only. Finally, in Figure 6.4, the definitive amplitude scan, utilizing all data samples and all taggers combined together is shown. The signature for the signal around $\Delta m_s = 17.75$ ps^{-1} is striking. The signal signature lies in the sensitivity region which arrives up to 31.3 ps^{-1}, which is the highest in the world. The amplitude is consistent with unity ($\mathcal{A}(\Delta m_s = 17.75 \text{ ps}^{-1}) = 1.21 \pm 0.20$ $(stat)$) which indicates that all components are well calibrated. A lower limit is set $\Delta m_s > 17.2$ ps^{-1} at 95% C.L. which, although meaningful, it is superseded by the significance of the oscillation. The latter justifies a direct measurement of the frequency.

Signal Significance for Δm_s Fit

A clear oscillation signature is present with the amplitude scan, showing a consistency with unity and inconsistency with zero for $\mathcal{A}$ at $\simeq 6\sigma$ level. The natural step is the estimation of the significance of the signal observed and the measurement of Δm_s.

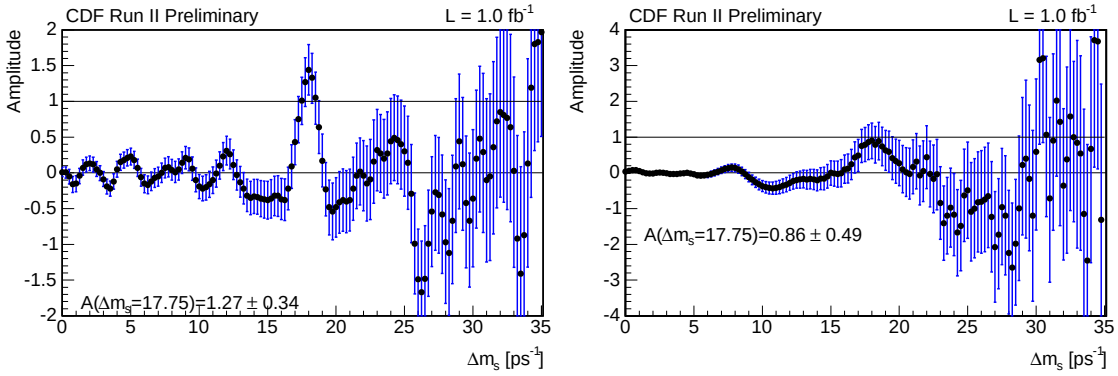

FIGURE 6.2: Left: amplitude scan for the hadronic modes. Right: amplitude scan for the semileptonic modes. The errors are statistical only.

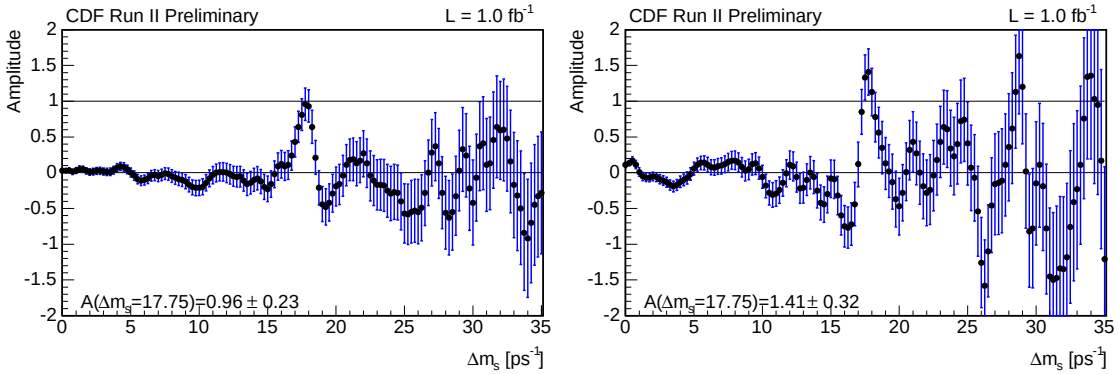

FIGURE 6.3: Left: amplitude scan for the SST only. Right: amplitude scan for the OST. The errors are statistical only.

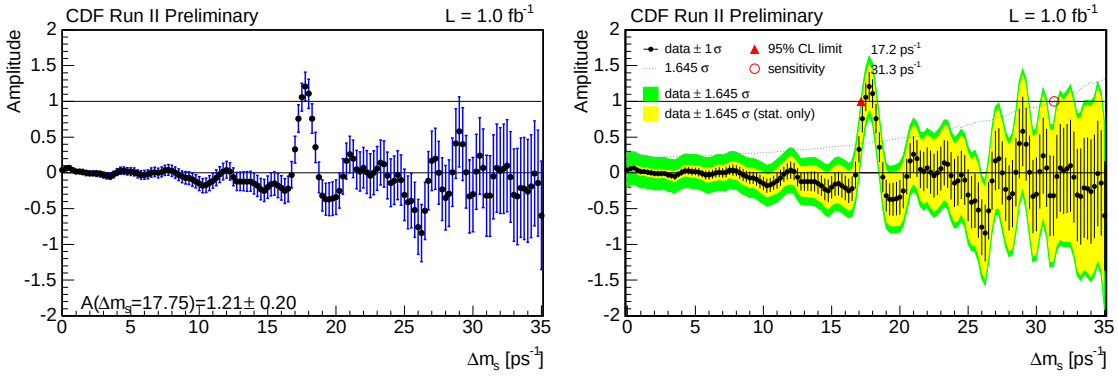

FIGURE 6.4: Amplitude scan with the data of all samples and taggers combined together. The plot on the left emphasize the statistical only uncertainty which is the dominant one. The version on the right includes also systematic uncertainties.

The significance of the B_s^0 mixing oscillation is quantitatively defined by the p-value, the probability for a random fluctuation to produce the observed signal. For the estimation of the p-value the likelihood ratio is used:

$$\Lambda = -log\frac{\mathcal{L}(\mathcal{A}=1)}{\mathcal{L}(\mathcal{A}=0)} \tag{6.1.3}$$

where $\mathcal{L}(\mathcal{A}=1)$ is the likelihood of the data under the hypothesis that Δm_s is the true mixing frequency, and $\mathcal{L}(\mathcal{A}=0)$ corresponds to the hypothesis of no oscillations ($\Delta m_s \to \infty$). From the likelihood ratio Λ we extract the likelihood profile defining the response of the data in the frequency space. The minimum value of Λ, (Λ_{min}) is related to the significance estimation, while its position determines the mean value and statistical uncertainty in the Δm_s measurement. Figure 6.5 shows the likelihood ratio as a function of Δm_s. The likelihood ratio deeper minimum at $\Lambda_{min} = -17.26$ is obtained for $\Delta m_s = 17.77$ ps^{-1}.

The next step is to estimate the p-value. A natural method to reproduce the statistical power of the data while simulating the null hypothesis is to randomize the flavor tags in the data itself. In such way, the significance is quantified by the probability that a data sample with randomly assigned flavor tags achieves a value of Λ_{min} smaller or equal than the observed one at *any value of* Δm_s ($\Lambda_{min} = -17.26$). The likelihood profile is inspected in the range from $[0-35]$ ps^{-1}. Statistical fluctuation at high value of Δm_s are increasingly small at around zero, therefore examination of the likelihood profiles for values outside the decided window will not affect the p-value distribution (see left plot on Figure 6.5).

Figure 6.6 shows the p-value distribution as a function of the Λ_{min} obtained from all the randomized experiments. Only 28 scans out of 3.5×10^8 entries have a Λ_{min} smaller than -17.26, which corresponds to a p-value$= 8 \times 10^{-8}$. This value is well below the 5-standard-deviations threshold which corresponds to a p-value$= 5.7 \times 10^{-7}$. Therefore the combined CDF datasets allow the world first definitive observation of $B_s^0 - \overline{B}_s^0$ oscillation. The best fit on data allows the determination of the Δm_s value which corresponds to the value that minimizes Λ, while its statistical uncertainty is determined by the value of Δm_s where Λ changes by 0.5 from the minimal value.

The final result obtained, including the systematics, is the following:

$$\Delta m_s = 17.77 \pm 0.10 \text{ (stat) } \pm 0.07 \text{ (syst) } \text{ ps}^{-1}. \tag{6.1.4}$$

The likelihood also determines the confidence intervals

$$17.56 < \Delta m_s < 17.96 \text{ ps}^{-1} \text{ at } 90\% \text{ CL}, \tag{6.1.5}$$

and

$$16.51 < \Delta m_s < 18.00 \text{ ps}^{-1} \text{ at } 95\% \text{ CL}, \tag{6.1.6}$$

including both statistic and systematic uncertainties.

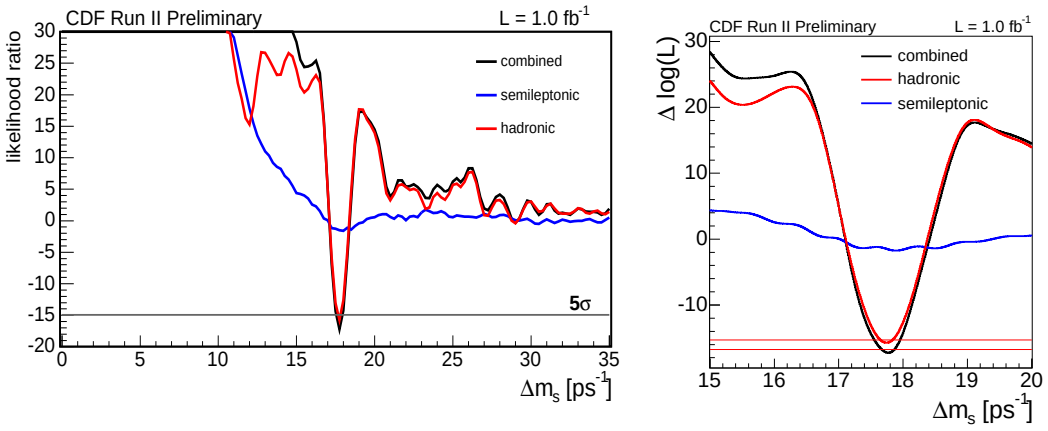

FIGURE 6.5: Likelihood profile: likelihood ratio as a function of Δm_s (left) with a zoom-in of the $[15, 20]$ ps^{-1} region. The contribution of hadronic (red), semileptonic (blue) and their combination (black) are shown separately.

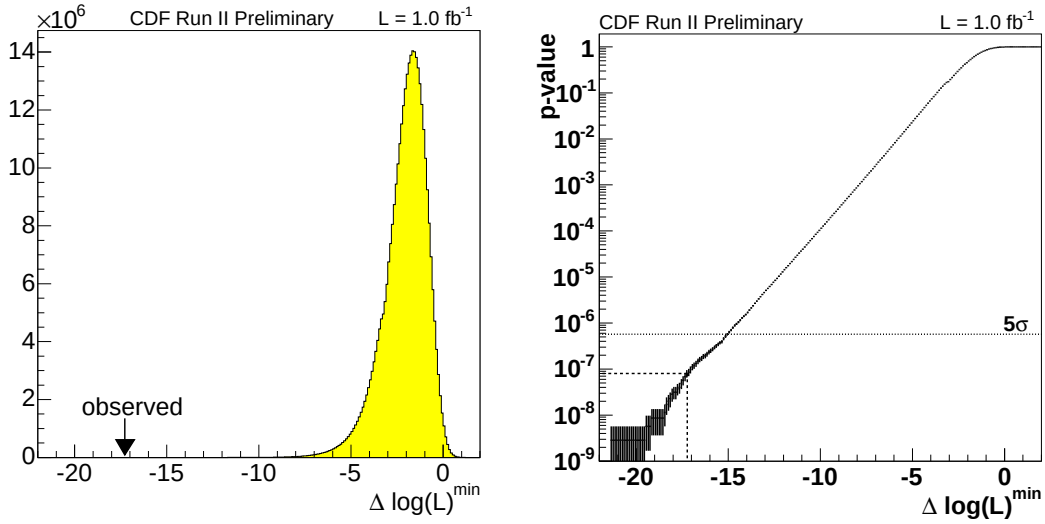

FIGURE 6.6: Distribution of the Λ_{min} obtained from the randomized data(left) and p-value as a function of the Λ_{min}, which is the cumulative of the distribution of Λ_{min} shown on the left.

Moreover, the sensitivity of the analysis makes feasible to study the flavor asymmetry in the B_s^0 proper time domain. The result, using the fully and partially reconstructed hadronic decays, is shown in Figure 6.7. The asymmetry in each time bin is calculated with the same likelihood used for the Δm_s measurement. To obtain the asymmetry in each bin, the value of Δm_s is set to zero in the fit and we fit for the mixing amplitude. In this way, event-by-event dilution, signal fractions and probability distributions for proper-time and dilution are automatically taken into account.

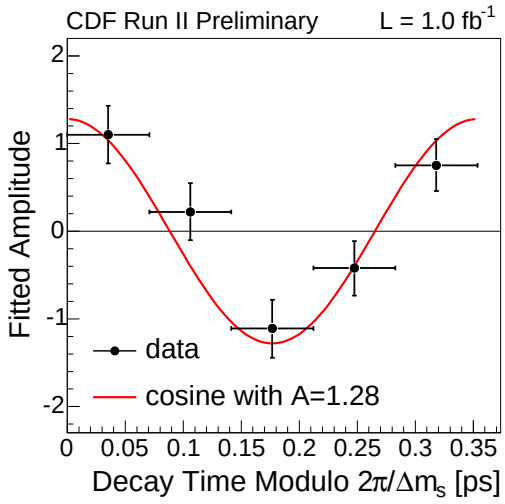

FIGURE 6.7: Flavor asymmetry as a function of the B_s^0 proper-decay-time. The value of each point is calculated from the mixing likelihood, modulo $T = 2\pi/\Delta m_s$ ps, and is distributed with a cosine of amplitude and frequency corresponding to the overlaid curve. $\chi^2/ndf = 4.77/5$.

However, the damping effect on the asymmetry due to the proper-time resolution is not properly considered, since $\Delta m_s = 0$. In order to correct for this effect, the dilution of each event is scaled by a term $e^{(\sigma_{ct}\Delta m_s)^2/2}$. By dividing the events in multiples of the measured oscillation period $T = [2\pi/17.77]$ ps, we enhance the visible asymmetry in the data, which would otherwise be distributed over about ten oscillation periods. The expected shape is a cosine with an amplitude equal to $\mathcal{A} = 1.28 \pm 0.22$, which is consistent with the amplitude obtained from the amplitude scan method.

6.1.3 Measurement of $|V_{td}/V_{ts}|$ and impact on Unitarity Triangle

The precise measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency allows to constrain the apex of the Unitarity Triangle. In Section 1.3 we showed how it is possible to derive the ratio $|V_{ts}|^2/|V_{td}|^2$ from the ratio $\Delta m_s/\Delta m_d$ via the equation 1.3.2

$$\frac{\Delta m_s}{\Delta m_d} = \xi^2 \frac{m_{B_s^0}}{m_{B^0}} \frac{|V_{ts}|^2}{|V_{td}|^2}\,, \tag{6.1.7}$$

where:

$$\xi = \frac{f_{B_s^0}}{f_{B^0}} \sqrt{\frac{\hat{B}_{B_s^0}}{\hat{B}_{B^0}}} = 1.210\ ^{+0.047}_{-0.039}\ [11]\,. \tag{6.1.8}$$

Combining the CDF measurement $\Delta m_s = 17.77 \pm 0.10\ (stat) \pm 0.07(syst)$ ps^{-1} with the world average $\Delta m_d = 0.507 \pm 0.005$ ps^{-1} [6] and the CDF measurement of $m_{B^0}/m_{B_s^0} = 0.98390$ [76] which has a negligible uncertainty ($\mathcal{O}(10^{-4})$), the ratio

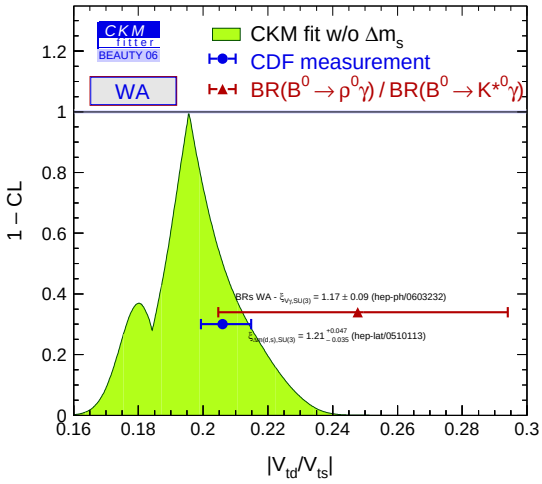

FIGURE 6.8: CKM fit and two experimental measurements of $|V_{ts}|^2/|V_{td}|^2$. Each horizontal section of the green area corresponds to interval in which $|V_{ts}|^2/|V_{td}|^2$ lies with the level of confidence read on the vertical axis. The two measurements represent the CDF result and the average of the Belle and BaBar measurement of the ratio $|V_{ts}|^2/|V_{td}|^2$.

extracted is:

$$\frac{|V_{ts}|^2}{|V_{td}|^2} = 0.2060 \pm 0.0007\ (\text{exp})\ ^{+0.0081}_{-0.0060}\ (\text{theory})\ . \tag{6.1.9}$$

The first uncertainty, referred as "exp", is only related to the Δm_s measurement, while the second includes all the other sources, dominated by the theoretical uncertainty on the parameter ξ. The ratio $|V_{ts}|^2/|V_{td}|^2$ is known with an experimental uncertainty which is an order of magnitude smaller than the theoretical one. The immediate consequence is the necessity for a work on the improvement of the lattice calculation aiming at the determination of the parameter ξ, in order to better exploit the information provided by the B_s^0 mixing analysis. The importance of CDF measurement of $|V_{ts}|^2/|V_{td}|^2$ is pictorially shown, using the CKM fitter framework [13], in Figure 6.8, where the CDF measurement is compared to the theoretical expectation and the average of the measurements from Belle [77] and BaBar [78], Figure 6.9 shows the direct effect of the Δm_s measurement on the unitarity triangle. The striking result is to squeeze the $(\bar{\rho}, \bar{\eta})$ 95% CL ellipse corresponding to the $\Delta m_d/\Delta m_s$ measurement. The central Δm_s value of 17.77 ps^{-1} is consistent with the Standard Model theoretical expectation of $18.3^{+6.5}_{-1.5}$ ps^{-1} from the CKM fitter group, thus this observation is consistent with the unitarity of CKM matrix.

6.1.4 Conclusions

The first part of this dissertation reports the analysis resulting in the first observation of the time-dependent B_s^0 flavor oscillations which concludes a twenty years

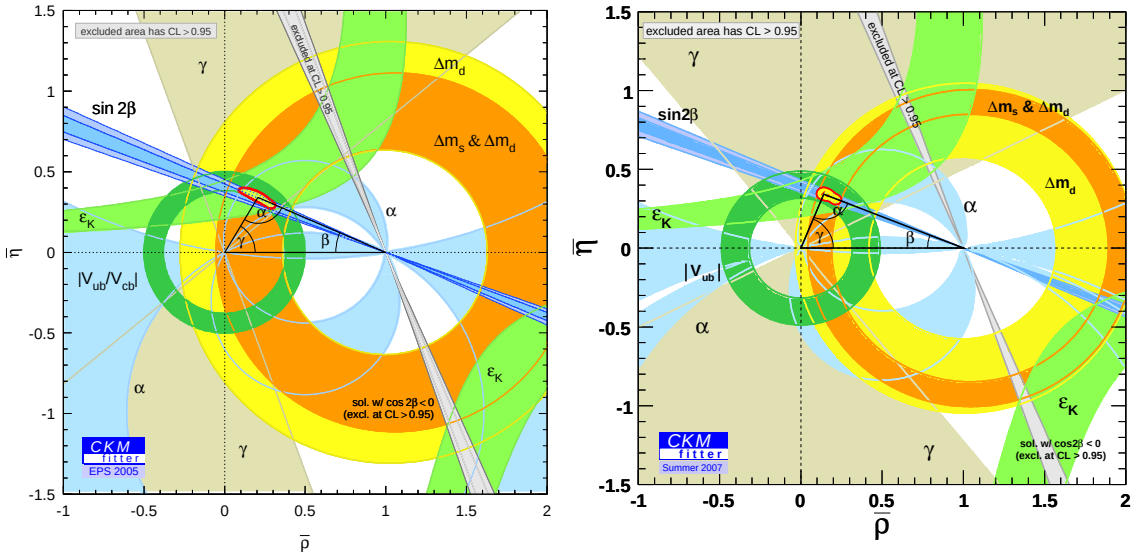

FIGURE 6.9: CKM fit of the Unitarity Triangle at EPS2005 (left) compared with the most recent one, which includes the CDF observation (right).

long search in the flavor interaction field. My personal contribution to the analysis was focused in the flavor tagging sector, in order to provide and calibrate the algorithms for the B meson flavor identification at the time of its production. More specifically, I worked in the Opposite Side Tagging calibration, in its final Neural Network implementation, and in the early Same Side Kaon Tagging development of the kinematic-based and particle-identification-based algorithms. The latter, integrating the particle-identification information provided by the CDF Time-Of-Flight detector and the measurement of ionization energy loss in the Central Outer Tracker, was the first implemented SSKT algorithm in the B_s^0 mixing analysis which provided the first evidence for the B_s^0 measurements [14]. On the other hand, the final version of the Same-Side Kaon Tagging used for the B_s^0 mixing world's first observation consists in a major upgrade of the algorithm combining together the kinematic and particle-identification information in a Neural Network fashion [66]. The tagging power of the Same Side Kaon Tagger described in this thesis is about 4%, when applied to the B_s^0 samples reconstructed, to be compared with the about 1.8% for the Opposite Side Tagger. The analysis has been performed of three different samples: the semileptonic B_s^0 decay modes [45], the hadronic partially reconstructed B_s^0 decay modes [70] and the hadronic fully reconstructed B_s^0 decay modes. The latters represent the world largest B_s^0 hadronic sample collected, made possible thanks to the great performances of CDF tracking and trigger system, in particular the Layer00 and the Secondary Vertex Tracker. The fully reconstructed hadronic samples have been additionally used for the proper-decay time resolution calibration given the excellent precision provided by CDF and they represent the strongest contribution to the statistical power of the B_s^0 mixing analysis.

The $B_s^0 - \overline{B}_s^0$ oscillation frequency measurement achieved combining all the dataset

corresponding to an integrate luminosity of 1 fb^{-1} is:

$$\Delta m_s = 17.77 \pm 0.10 \text{ (stat)} \ \pm 0.07 \text{ (syst)} \ \text{ ps}^{-1}. \tag{6.1.10}$$

with a significance superior to 5 standard deviations. The consistency with unity of the amplitude scan, $\mathcal{A} = 1.21 \pm 0.20$ is a considerable achievement for the analysis, because it is a clear indication that all components have been well calibrated.

Although the definitive observation of the B_s^0 mixing frequency is an important measurement by itself, it also provides an handle to get extremely precise measurements of the CKM parameters. In particular, it was possible to extract the ratio $|V_{ts}|^2/|V_{td}|^2$:

$$\frac{|V_{ts}|^2}{|V_{td}|^2} = 0.2060 \pm 0.0007 \text{ (exp)} \ ^{+0.0081}_{-0.0060} \text{ (theory)} \ . \tag{6.1.11}$$

where the first uncertainty - referred as "exp" - is only related to the Δm_s measurement and the second includes all the other sources, dominated by theoretical uncertainty. The Δm_s is well consistent with the Standard Model unitarity for the CKM matrix and, at this point in time, the test is only limited by theoretical calculations. The oscillation frequency determination indicates that the presence of New Physics beyond the Standard Model, if any, in the magnitude of the mixing amplitude are extremely small. Nevertheless, New Physics could modify the phase of the mixing amplitude β_s, by contributing with additional processes. Indeed, global fits on experimental data constrain the CP violation phase to a small value $2\beta_s \simeq 0.04$ [75]. Therefore, New Physics could lead to significantly larger value. After the Δm_s measurement, the natural continuation of the B_s^0 system studies is to measure the phase β_s. Thus, the second part of my thesis, which will be discussed in the following section, was focused on the angular analysis of $B_s^0 \to J/\psi\phi$ decays for the world's first measurement of β_s using flavor tagging information.

6.2 Angular Analysis on $B_s^0 \to J/\psi\phi$ Decay Modes

This Section is devoted to the description of the results for the angular analysis on $B_s^0 \to J/\psi\phi$ utilizing, for the first time, flavor tagging information. Such an analysis allows to measure the width difference, $\Delta\Gamma_s \equiv \Gamma_s^L - \Gamma_s^H = 1/\tau_{B_s^L} - 1/\tau_{B_s^H}$ between the two mass eigenstates and then to extract the CP violating phase β_s. The recent results of a similar analysis from the DØ collaboration, in comparison with the CDF one, is discussed. The impact of the new inputs, in the attempt to infer if CP violation in the quark sector is entirely explained within the Standard Model, are also analyzed.

6.2.1 Feldman-Cousins Confidence Region

As already reported in Section [**?**], the likelihood used for the angular analysis on the $B_s^0 \to J/\psi\phi$ decay mode exhibits an exact symmetry which holds regardless of the true values of the parameters of interest or the statistical power available:

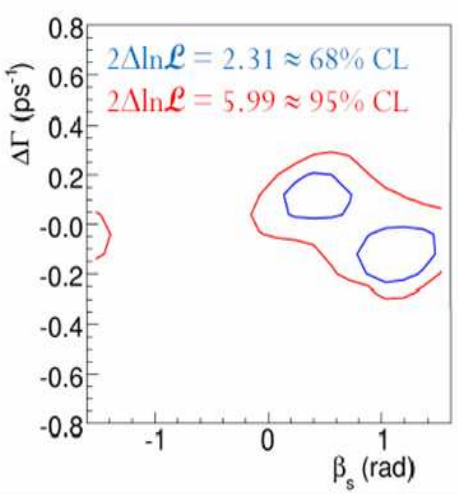

FIGURE 6.10: Toy MC likelihood contour for β_s generated with 0.40 where only the 7 signal parameters (amplitudes, strong phases, lifetime, $\Delta\Gamma_s$ and β_s) are allowed to float in the fit. The blue curve corresponds to 2.30 units above the minimum in the likelihood ($\sim 68\%$ probability) and the red curve to 5.99 units above the same minimum ($\sim 95\%$ probability). The fit is performed boxing $\delta_\parallel \in [0, \pi)$.

$$\begin{aligned} \beta_s &\to \frac{\pi}{2} - \beta_s \\ \Delta\Gamma_s &\to -\Delta\Gamma_s \\ \delta_\parallel &\to 2\pi - \delta_\parallel \\ \delta_\perp &\to 2\pi - \delta_\perp \end{aligned} \tag{6.2.1}$$

The studies performed on toys Monte Carlo showed a problematic set of pseudo-invariances appearing in the likelihood. Although the exact symmetry was removed, by limiting the allowed range of $\delta_\parallel$, a second local minimum manifested in the likelihood. Effects like this are a consequence of the fairly limited statistical power available, which prevents us from the possibility to resolve the interference terms well enough to distinguish these approximate invariances. As a representative example, Figure 6.10 shows one the expected behaviour for these pseudo-invariances. The toy Monte Carlo uses a data set of 12000 signal events (6 times the current available data), with realistic mixture of signal and background plus realistic efficiency and dilution for the tagging algorithms. In this example the local pseudo invariance is manifested as a second ambiguity according to the transformation $\beta_s \to \pi/2 - \beta_s$ and $\Delta\Gamma_s \to -\Delta\Gamma_s$. Additionally, this test also shows that even when the statistics in the data are significantly increased over the present level, it may be not possible to quote a point estimate for certain "true" value of β_s.

Having establish the possibility of an ambiguity in our likelihood with the current statistical power in the analysis, despite the exact symmetry removal, we have checked the likelihood contour scans in the data. Before going through the description of the

likelihood results on data, it is mandatory to mention that all these studies, and the approach which was decided as a consequence of them, have been performed in a "blind" fashion. In fact, in order to maintain an unbiased search for the $\Delta\Gamma_s$ and β_s, a constant value between $[-2.5, 2.5]$ to the fitted value of $\Delta\Gamma_s$ and a different unknown constant between $[0-100]$ to the fitted value of β_s. The following technique preserves all likelihood scans and errors returned by the fit, providing at the same time all the information needed to determine the fit performances while remaining unbiased about the central value estimation. Moreover, it is important to underline some additional features belonging to the fitter. First of all we constraint Δm_s to the value determined in the B_s^0 mixing oscillation previously described in the Chapter. In fact the single sample of 2000 signal events in $B_s^0 \to J/\psi\phi$ does not have enough sensitivity to the oscillation measurement which could in principle be determined by the fitter. The dilution scale factors are also constrained within their total errors, which include both statistical and systematic errors (see Chapter [**?**]).

The scans obtained with the fit re-minimized at each point in $\Delta\Gamma_s - \beta_s$ space with all other parameters floating are depicted in Figure 6.11. Three likelihood scans are presented: a contour obtained with $\delta_\parallel$ boxed in $[0, \pi)$, another contour with $\delta_\parallel$ boxed in $(\pi, 2\pi]$, which were meant to remove the exact symmetry, and the final contour without any boxing in $\delta_\parallel$. These scans clearly do not show the required parabolic behaviour for the likelihood. This is by far the most crucial point of the whole analysis on $B_s^0 \to J/\psi\phi$: the classical point-estimate will not provide a reliable estimation of the fit parameters, thus the final result will be quoted in terms of a confidence region in the bi-dimensional space $\Delta\Gamma_s - \beta_s$.

Confidence Region Determination

If we were guaranteed of a correct coverage of the fit for the two-dimensional contour scans in the $\Delta\Gamma_s - \beta_s$ plane, then the bottom plot in Figure 6.11 would be our final result. Since we are not assured this is the case of our likelihood, we decided to check the result in a more robust frequentist approach according to the Feldman-Cousins likelihood ratio ordering in which nuisance parameters are included [79].
The prescription followed to determine the confidence region starts with the definition of the likelihood ratio for a point $(\Delta\Gamma_s,\ \beta_s)$ in the two dimensional space:

$$\mathcal{LR}(\Delta\Gamma, \beta_s) = log\frac{\mathcal{L}(\hat{\Delta\Gamma}, \hat{\beta}_s, \hat{\theta})}{\mathcal{L}(\Delta\Gamma, \beta_s, \hat{\theta}')} \tag{6.2.2}$$

where θ represents the vector of all fit parameters but the tested $\Delta\Gamma$ and β_s and defined as "nuisance" parameters. The "hat" $\hat{\Delta\Gamma},\ \hat{\beta}_s,$ and$\hat{\theta}$ are the parameters values which minimize the likelihood, when all parameters are floating in the fit. On the other hand, given the pairs $(\Delta\Gamma,\ \beta_s)$, $\hat{\theta}'$ represents the values of the nuisance parameters obtained by minimizing the likelihood when $\Delta\Gamma$ and β_s fixed to the tested value.

Utilizing the likelihood ratio, for a given pair $(\Delta\Gamma,\ \beta_s)$, a p-value can be extracted to evaluate the agreement of the data with the model predicting the chosen point in

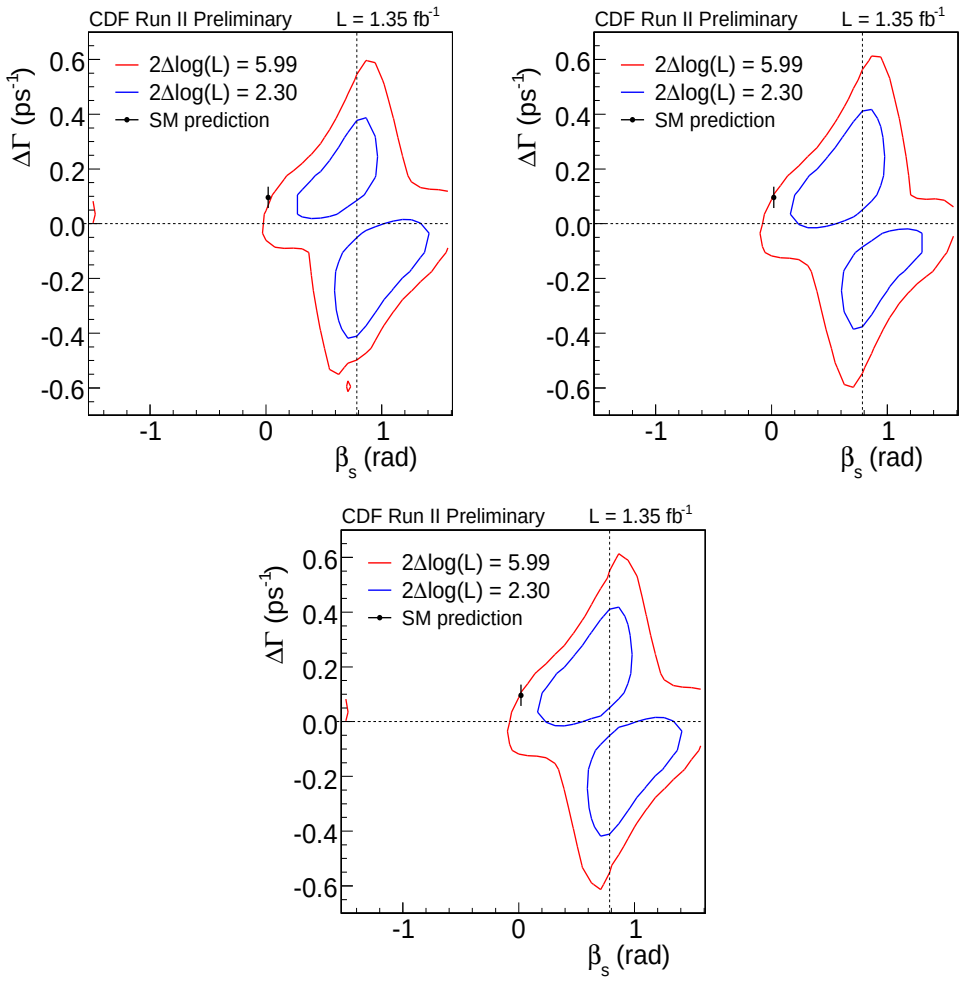

FIGURE 6.11: Comparison between the likelihood profile with $\delta_{\parallel} \in [0, \pi)$ (upper left), $\delta_{\parallel} \in [0, \pi)$ (upper right) and without any boxing in $\delta_{\parallel}$. The blue curve corresponds to 2.30 units above the minimum in the likelihood ($\sim 68\%$ probability) and the red curve to 5.99 units above the same minimum ($\sim 95\%$ probability). The black point corresponds to the Standard Model expectation $\Delta\Gamma = 0.096$ ps^{-1} and $\beta_s = 0.02$ [75].

the grid. The procedure is the following:

1. For the fixed pair ($\Delta\Gamma$, β_s), we generate realistic toy Monte Carlo samples. The initial values for the width-difference and the CP violation phase are the ones of the selected pair ($\Delta\Gamma$, β_s). As best possible estimate for the remaining fit parameters θ, we use the values obtained from a fit to the data, while fixing $\Delta\Gamma$ and β_s to the tested values we are testing. This methodology is also known as plug-in method.

2. We fit the toy Monte Carlo samples twice: once with all parameters floating and once with the pair ($\Delta\Gamma$, β_s) fixed to the chose initial value. We can construct the likelihood ratio as defined in Equation 6.2.2

3. We compare the normalized distribution of $\mathcal{LR}$ values from the pseudo-experiments, $f(\mathcal{LR},\ \Delta\Gamma,\ \beta_s)$ to the value of the likelihood ratio obtained from the data, $\mathcal{LR}_{data}(\Delta\Gamma,\ \beta_s)$. Finally the p-value is defined as:

$$p\text{-value}(\Delta\Gamma,\ \beta_s) = \int_{\mathcal{LR}_{data}(\Delta\Gamma,\ \beta_s)}^{\infty} f(\mathcal{LR},\ \Delta\Gamma,\ \beta_s)\ d\mathcal{LR}. \tag{6.2.3}$$

By utilizing the described method, we are able to determine the probability to observe a result with $\mathcal{LR} \geq \mathcal{LR}_{data}$, if the pair $(\Delta\Gamma,\ \beta_s)$ is the one predicted by some model.

For a defined confidence level $X\%$, the confidence region is now defined by the pairs $(\Delta\Gamma,\ \beta_s)$ which satisfy the condition p-value $> (1 - X)\%$:

$$(\Delta\Gamma,\ \beta_s) \in X\%\ \ C.L.region \iff p\text{-value}(\Delta\Gamma,\ \beta_s) > X\% \tag{6.2.4}$$

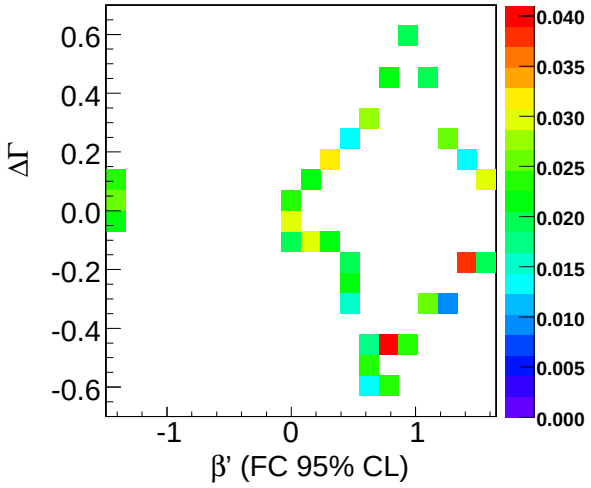

FIGURE 6.12: Points along the 95% C.L. region. The colors correspond to the RMS (Root Mean Square) of the 16 p-values corresponding to the different random nuisance parameters.

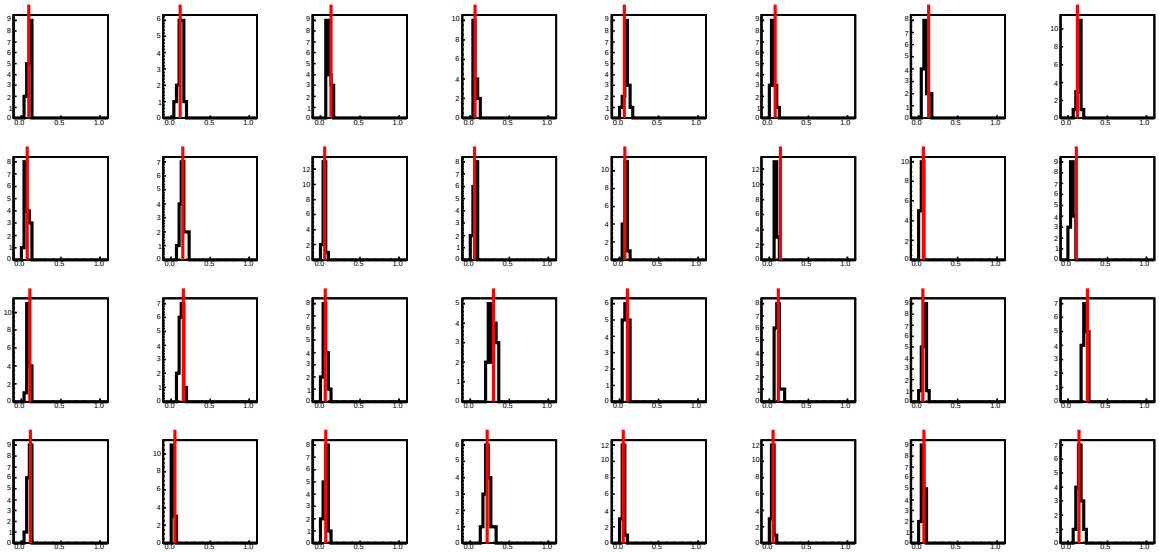

FIGURE 6.13: Distributions of the 16 p-values corresponding to the different random nuisance parameters for each point along the 95% C.L. contour. The red line indicate the p-value observed in data.

The confidence region obtained with this approach is only an approximation, since its construction assumes that the contour shape depends essentially on the two parameters of interest, $\Delta\Gamma$ and β_s. In fact, to get the correct confidence region, we should equally vary all the other nuisance parameters. Although this methodology is, in principle, the correct one, it would require extremely long time and unavailable computational resources. Nevertheless, we checked the stability of the confidence regions against variations of the nuisance parameters. For each point along the 95% C.L. region we generated 16 pseudo-experiments with the initial values for the nuisance parameters θ uniformly distributed within $\pm 5\sigma$ from their central value in the data. Figure 6.12 shows the points that are investigated along the 95% C.L. region. The distributions of the 16 p-values corresponding to the 16 "alternative universe" compared to the p-value obtained in data are reported in Figure 6.13.

Although an excellent agreement is observed between the p-value in data and the p-values distributions in the 16 alternative nuisance parameters hypotheses, we accounted for the any possible under-coverage by increasing the confidence region obtained with the average, over all tested points, of the RMS of the 16 p-values distributions, which corresponds to the different random nuisance parameters choices. This average RMS results to be 2.3% , which implies that, for example, what we quote to be the 95% C.L. is, in fact, the 97.3% obtained on the data.

Final Results

By utilizing the Feldman-Cousins likelihood ratio ordering with a rigorous frequentist inclusion of systematic uncertainties, we determine the confidence levels for a 20×40 grid evenly spaced in $\beta_s \in [-\pi/2,\ \pi/2]$ and $\Delta\Gamma_s \in [-0.7, 0.7]$. The 68% and 95% confidence regions obtained are shown in Figure 6.14. The coverage is tested against the deviations of the nuisance parameters. In such way, we exclude a specific point in the $\Delta\Gamma_s - \beta_s$ plane only if it can be excluded for any other possible value of the remaining nuisance parameters (inside a 5σ deviation from the nominal value returned by the fit on the data). This rigorous approach is equivalent to project the multi-dimensional confidence region onto the bi-dimensional space of interest, $\Delta\Gamma_s - \beta_s$. As a side note, several sources of systematic uncertainties have been studied in the case of a point estimate. While this is not the situation we are facing, the list of systematic effects addressed is described in Appendix **??** to give a reader a scale of the possible systematic uncertainties.

The contour obtained with the Feldman-Cousins prescription is wider compared to the one resulting from the likelihood profile. However, the general shape is very similar between the two methods. Assuming the Standard Model predicted values at $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.096$ ps^{-1} [75], the probability to observe a likelihood ratio equal to or higher than what it is observed in data is 15%. A deeper investigation of the two minima from showed how the solution centered in $0 \leq \beta_s < \pi/4$ and $\Delta\Gamma_s > 0$ corresponds to $\cos(\delta_\perp) < 0$ and $\cos(\delta_\perp - \delta_\parallel) > 0$, while the opposite is true for the solution centered in $\pi/4 \leq \beta_s < \pi/2$ and $\Delta\Gamma_s < 0$.

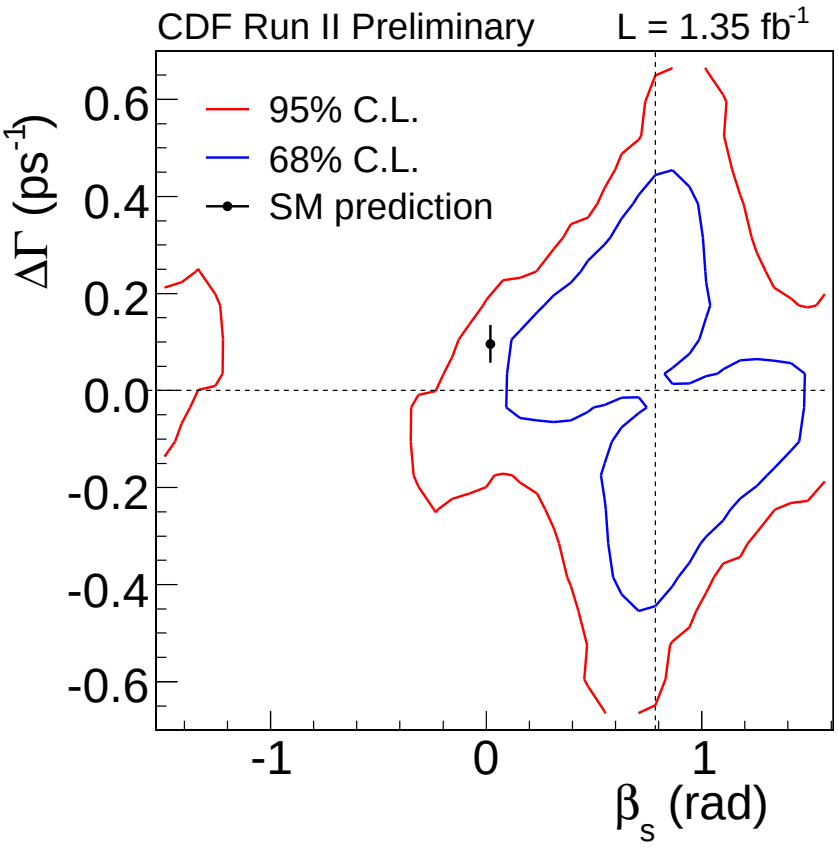

FIGURE 6.14: Final Feldman-Cousins confidence region after the coverage adjustment. The blue curve corresponds to 2.30 units above the minimum in the likelihood ($\sim 68\%$ probability) and the red curve to 5.99 units above the same minimum ($\sim 95\%$ probability). The black point corresponds to the Standard Model expectation $\Delta\Gamma = 0.096$ ps^{-1} and $\beta_s = 0.02$ [75].

In Section 4.4, while describing the different ingredients for the angular analysis on $B_s^0 \to J/\psi\phi$, we underlined the difficulty to quantify the tagging impact, in comparison to the B_s^0 mixing searches where the sensitivity Formula 4.1.9 determines univocally the tagging power of the analysis, through its figure of merit ϵD^2. Nevertheless, via the comparison between the Feldman-Cousins confidence region of the untagged analysis [74] and the one described in this dissertation, we can draw some visual conclusions. Figure 6.15 shows how the use of the flavor tagging, although does not strongly help to reduce the errors, does reduce the four-fold ambiguity in the likelihood for the untagged analysis to a two-fold ambiguity. The flavor tagging algorithms reduce by half the space available for β_s to float.

Additionally to the two-dimensional confidence region, we report the Feldman-Cousins interval of β_s where $\Delta\Gamma_s$ is treated as a nuisance parameter, and find that $\beta_s \in [0.16,\ 1.41]$ at the 68% confidence level.

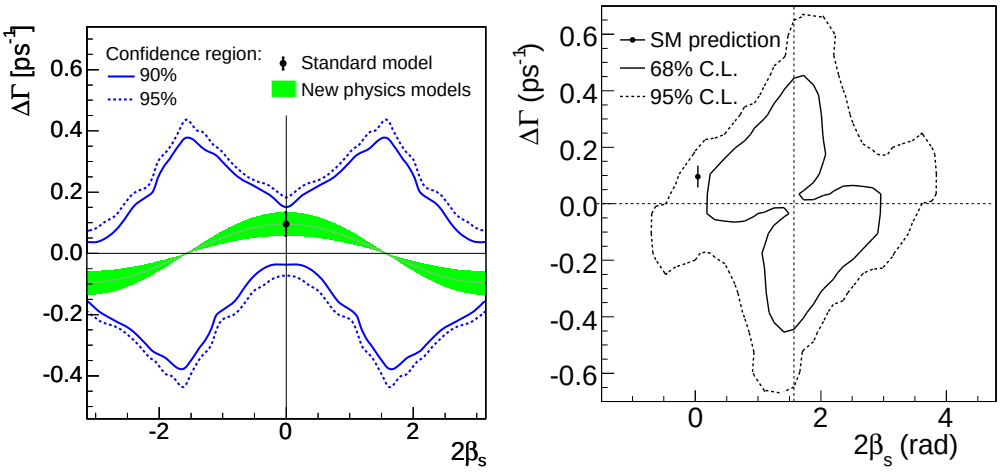

FIGURE 6.15: Feldman-Cousins $\Delta\Gamma_s - 2\beta_s$ confidence regions comparison between the untagged analysis [80] (left) and the tagged analysis [81], described in this dissertation. The confidence region on the left quotes the 90% and 95%; the green band corresponds to $\Delta\Gamma_s = 2|\Gamma_{12}|\cos(2\beta_s)$, with $|\Gamma_{12}| = 0.048 \pm 0.018$ [75], which connects the phase $2\beta_s$ to the width-difference $\Delta\Gamma_s$. The confidence region on the right is identical to the one in Figure 6.14, but drawn in the $\Delta\Gamma_s - 2\beta_s$ plane instead of the original $\Delta\Gamma_s - \beta_s$, for the sake of a better comparison to the untagged one; it quotes the 68% (continous line) and 95% (dashed) C.L. In both regions the black point corresponds to the Standard Model expectation: $\Delta\Gamma = 0.096$ ps^{-1} and $\beta_s = 0.02$ [75].

We also exploit the current experimental and theoretical information to extract tighter bounds on the CP violation phase β_s. In fact we know that $\Delta\Gamma_s$ cannot float all over the available space but it is connected to the CP violation phase through the formula $\Delta\Gamma_s = 2|\Gamma_{12}|\cos(\phi_s) = 2|\Gamma_{12}|\cos(2\beta_s)$, if New Physics is the dominating contribution ($\phi_s = \phi_s^{SM} + \phi_s^{NP}$ and $2\beta_s = 2\beta_s^{SM} - \phi_s^{NP}$). Therefore we use the mentioned Formula for $\Delta\Gamma_s$, imposing the constraint on $|\Gamma_{12}| = 0.048 \pm 0.018$ [75]. We obtain $\beta_s \in [0.12,\ 0.68] \cup [0.89,\ 1.45]$ at the 68% confidence level.

Moreover, in the limit of $SU(3)$ being a symmetry of the SM, we add up the constraints on the strong phases $\delta_\parallel$ and $\delta_\perp$ from the $B^0 \to J/\psi K^{*0}$ decays [82] measured at BaBar and the constraint on the B_s^0 mean width from the world average B^0 width [6], to find $\beta_s \in [0.20,\ 0.60]$ at the 68% confidence level. In Figure 6.16 we show the comparison of the default likelihood contour on the data with respect to the modified contours when constraints on the B_s^0 lifetime and the strong phases $\delta_\parallel$ and $\delta_\perp$ are applied.

We finally report in Figure 6.17 the angular sideband-subtracted distributions on the data for the transversity basis, to be compared with the fit projections. The good agreement found confirms the correct angular treatment in the likelihood construction.

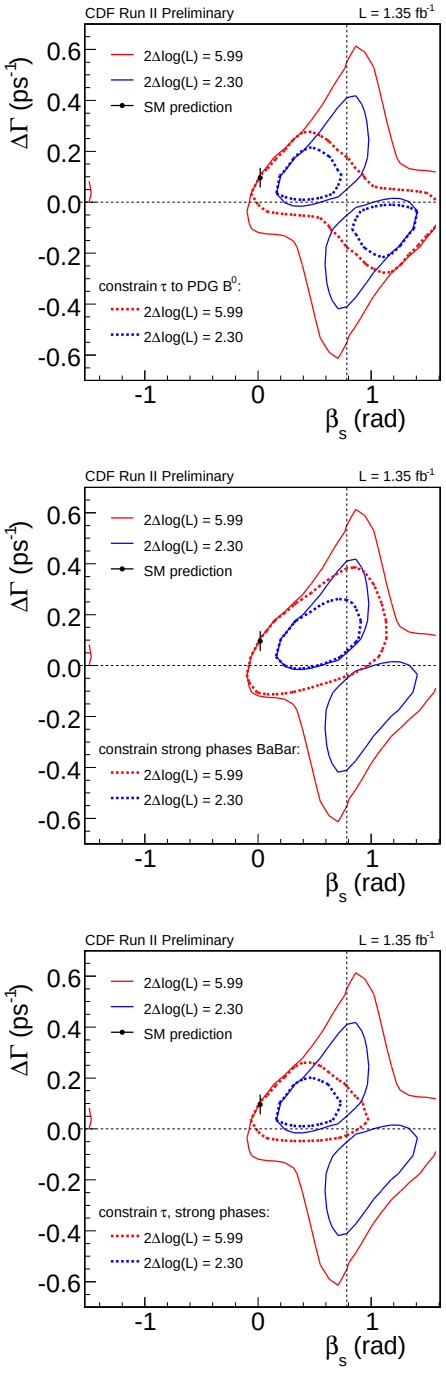

FIGURE 6.16: Comparison between the likelihood contours in the $\Delta\Gamma_s - \beta_s$ plane obtained on the data sample $B_s^0 \to J/\psi\phi$ with external constraint (dashed line) and without any constraint (continuous line). The blue curves correspond to 2.30 units above the minimum in the likelihood ($\sim 68\%$ probability) and the red curves to 5.99 units above the same minimum ($\sim 95\%$ probability). The black point corresponds to the Standard Model expectation $\Delta\Gamma = 0.096$ ps^{-1} and $\beta_s = 0.02$ [75]. In order, from top to bottom the constrained applied are: a constraint on the lifetime from the world average B^0 lifetime [6], a constraint on the strong phases $\delta_{\|}$ and $\delta_{\perp}$ from the $B^0 \to J/\psi K^{*0}$ decays [82], a constraint on both the lifetime and the strong phases.

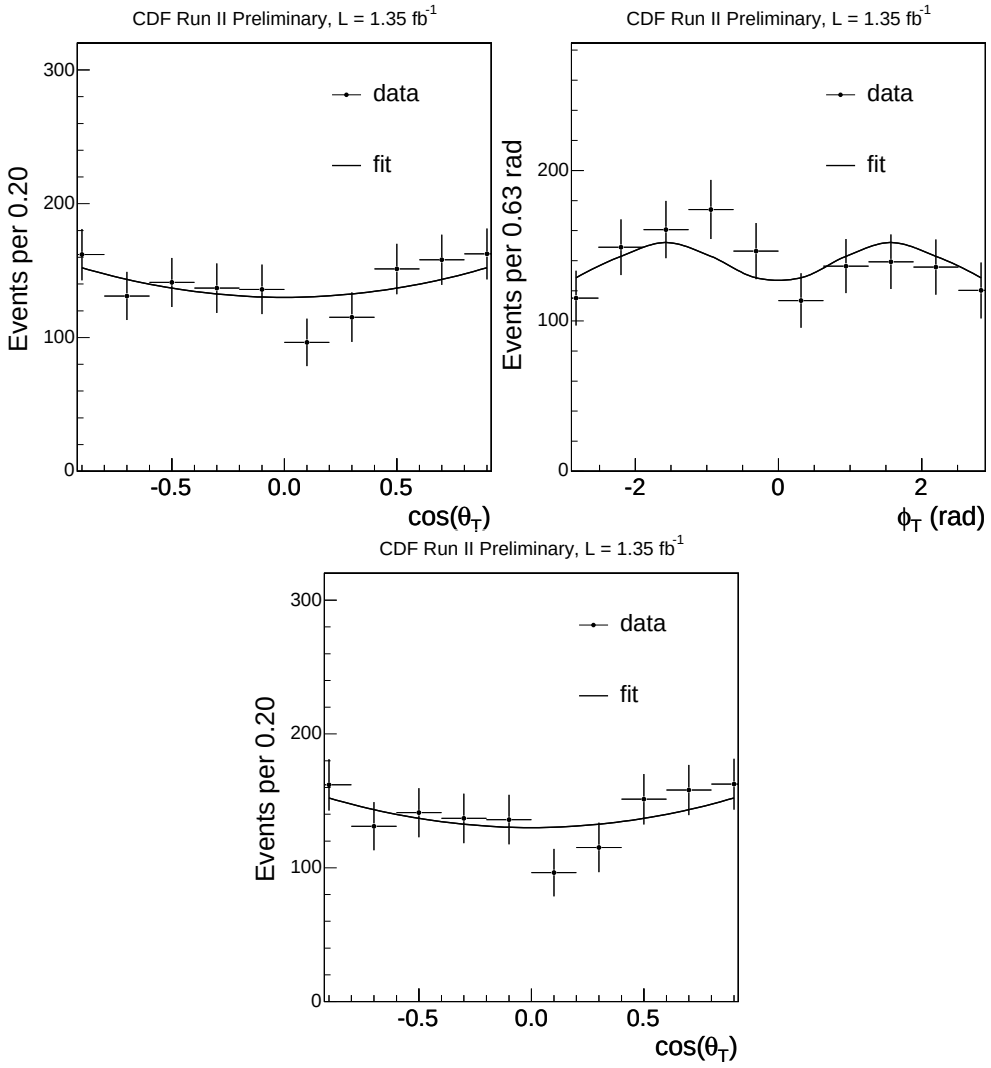

FIGURE 6.17: Sideband-subtracted angular distributions for the transversity basis on data compared to the fit projections. From left up to bottom: $cos(\theta_T)$, ϕ_T and $\cos(\psi_T)$.

6.2.2 Comparison With DØ Results

During the preparation of this dissertation, the DØ experiment released a similar analysis on 2.8 fb^{-1} of data [83] for the measurement of B_s^0 mixing parameters from the decay $B_s^0 \to J/\psi\phi$ utilizing flavor tagging. The signal yield they obtained from the fit is 1967 ± 65 B_s^0 events.

The approach to the analysis is different from CDF one, and needs some explanation. First of all some nomenclature: the CP violation phase ϕ_s quoted by the DØ analysis correspond to the CDF $-2\beta_s$.

As a response to the high degree of correlation between Δm_s, β_s and the two CP-

Parameter		Value
$\bar{\tau}_s$	$[ps]$	$+1.52 \pm 0.06$
$\Delta\Gamma_s$	$[\mathrm{ps}^{-1}]$	$+0.19 \pm 0.07$
$\lvert A_\perp(0)\rvert$		$+0.41 \pm 0.04$
$\lvert A_0(0)\rvert^2 - \lvert A_\parallel(0)\rvert^2$		$+0.34 \pm 0.05$
$\delta_1 = \delta_\perp - \delta_\parallel$		-0.52 ± 0.42
$\delta_2 = \delta_\perp$		$+3.17 \pm 0.39$
ϕ_s		$-0.57^{+0.24}_{-0.30}$

TABLE 6.2: Summary of the likelihood fit DØ results when all parameters floating.

conserving strong phases, $\delta_1 = \delta_\perp - \delta_\parallel$ and $\delta_2 = \delta_\perp$, compromising the stability of the fit, with the current statistical power, DØ collaboration decided to apply several constraints to the fit.

They fixed Δm_s to 17.77 ps^{-1}, which is the value obtained by CDF, and described in this dissertation. There is no big difference from what we did at CDF. Because of the lack of sensitivity to the Δm_s measurement in the $\mathrm{B}_s^0 \to J/\psi\phi$ sample only, we Gaussianly constrained the Δm_s value within its statistical and systematic uncertainties. DØ, instead, by fixing the Δm_s value, properly evaluated a systematic effect varying the Δm_s input, which resulted to be their higher systematic uncertainty on the ϕ_s measurement.

The strong difference lies in the constraining of the δ_1 and δ_2 phases from the world-average results on the $\mathrm{B}^0 \to J/\psi K^{*0}$ sample, $\delta_1 = -0.46$ and $\delta_2 = 2.92$ for their main results. The width of the Gaussian constraint was arbitrarily chosen to be $\pi/5$, in order to allow some degree of violation of the $SU(3)$ symmetry, while still effectively constraining the signs of $\cos\delta_i$, consistent with the factorization estimate.

All these constraints induce a more Gaussian-like behaviour of the likelihood. The DØ results given in terms of point estimates are presented in Table 6.2.

In Figure 6.18 we also report the DØ"expected" confidence level contour in the $\Delta\Gamma_s - \phi_s$ plane. The p-value, estimated in pseudo-experiments, to obtain a value inferior to the fitted $\phi_s = -0.57$ is found to be 6.6%.

In Figure 6.19 we compare the Tevatron results: they are undoubtedly consistent and, more interesting, they both show a deviation from the SM in the same directions. Nevertheless there are important distinctions to be discussed. DØ point estimates are strongly dependent on theoretical assumptions which limit the universality of their result. Instead, at CDF we just applied the minimal required constraints, e.g. Δm_s, and decided to cope directly with the instability of the likelihood renouncing and quoting confidence regions after a deep check of the coverage. These differences, and the consequential inconsistency in the two approaches, have been a source of debate inside the physics community. In particular, these issues were raised when trying to encompass the two results inside a global fit framework such as the CKM fit [13] and **UT**fit [84].

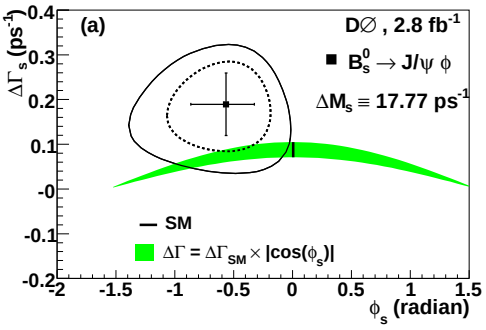

FIGURE 6.18: DØ confidence level contours in the $\Delta\Gamma_s - \phi_s$ plane. The curves correspond to expected 68% C.L. (dashed) and 90% (solid). The green band corresponds to the Equation $\Delta\Gamma_s = \Delta\Gamma_s^{SM} \cdot |\cos(\phi_s)|$.

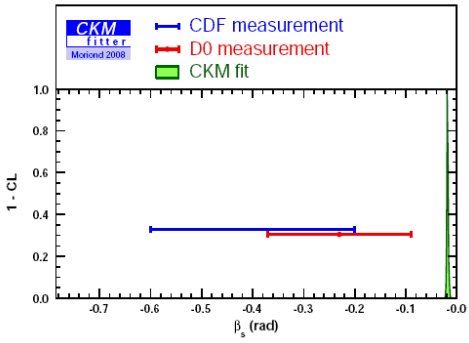

FIGURE 6.19: Comparison of the CDF and DØ results as presented at the conference Moriond EWK 08 by the CKM fitter collaboration. The blue line corresponds to the CDF Feldman-Cousins one-dimensional confidence region when all the constraints on Γ_{12}, $\delta_\parallel$, $\delta_\perp$ and B^0 width. The red line corresponds to the DØ result $\phi_s = -0.26 \pm 0.14$ obtained when $\Delta\Gamma_s$ is constrained by the expected relation $\Delta\Gamma_s^{SM} \cdot \cos(\phi_s)$. The green distribution corresponds to the SM expectation for the CKM fit $\beta_s = -0.0183^{+0.0009}_{-0.0008}\ rad$. In the CKM collaboration the nomenclature adopted is ${\beta_s}^{CKM} = -{\beta_s}^{CDF} = {\phi_s}^{D\emptyset}/2$.

6.2.3 Impact on the Standard Model

The β_s and $\Delta\Gamma_s$ measurements performed at the Tevatron experiments provide new inputs to over-constrain the Unitarity Triangle in the B_s^0 system. In this Section, we review the prescription defined inside the CDF collaboration for combining the Feldman-Cousins confidence region in a global fit framework and we analyze their first results produced with the use of these very recent measurements.

How To Combine CDF Results in Global Fits

The final result of our analysis is not coming from a likelihood profile, but from a two-dimensional bounds in the $\Delta\Gamma_s - \beta_s$ space using the Feldman-Cousins likelihood

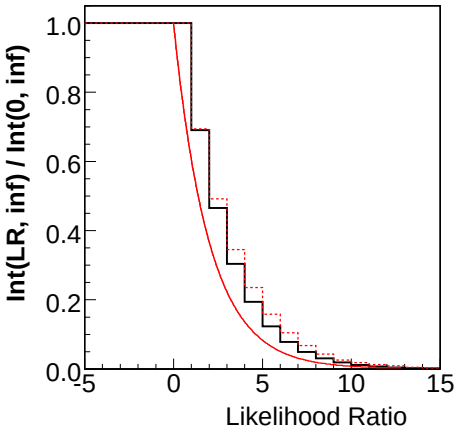

FIGURE 6.20: Likelihood ratio distributions used as reference to correct the likelihood ratio obtained on data assuming the parabolic behaviour of the likelihood. The continuous red line represents the 2-degrees-of-freedom χ^2 distribution compared to the nominal Feldman-Cousins (black line) and the Feldman-Cousins with the correction due to the nuisance parameters variation (dashed red line).

ratio ordering. The choice of confidence regions was dictated by the strongly non-Gaussian likelihood, with multiple minima that prevented the possibility to quote a reliable point-estimate. Feldman-Cousin was preferred to the "profile likelihood": the entity of the non-linearity of the problem led to suspect that the asymptotic limit was not reached with the current statistics.

For the nuisance parameters treatment, we use a "projection method": our confidence region in the $\Delta\Gamma_s - \beta_s$ space is the result of a projection of a multidimensional confidence region in the larger space that includes all nuisance parameters. In fact, we exclude a specific value of the pair $(\Delta\Gamma_s, \beta_s)$ only if it can be excluded for any assumed values of the nuisance parameters (within 5σ from their nominal values). To do this in practice, we evaluate for each $(\Delta\Gamma_s - \beta_s)$ point the most conservative p-value for our observed profile-$\mathcal{LR}$ statistics, out of a number of possible choices for the values for nuisance parameters.

We are found to be in a situation where the distribution of profile-$\mathcal{LR}$, while it is not a 2-degree-of-freedom-χ^2 distribution, is very nearly independent of the value of the pair $(\Delta\Gamma_s, \beta_s)$.

The distribution obtained with nominal values for the nuisance parameters, and the most conservative distribution obtained by random sampling 16 alternative choices is shown in Figure 6.20. As can be seen from this plot, showing the tail integrals, all distributions have longer tails than a χ^2, shown for reference in red. The effect on the results is not negligible: if you use the most conservative of these 16 distributions, that has a 90%-point at $\mathcal{LR} \simeq 6$, to be compared with the nominal χ^2-value of 4.6 (In deriving our 90% confidence regions, we have used an estimated 2.3% of additional tail in the "most conservative" distribution with respect to the distribution for nominal nuisance parameters (shown in black)).

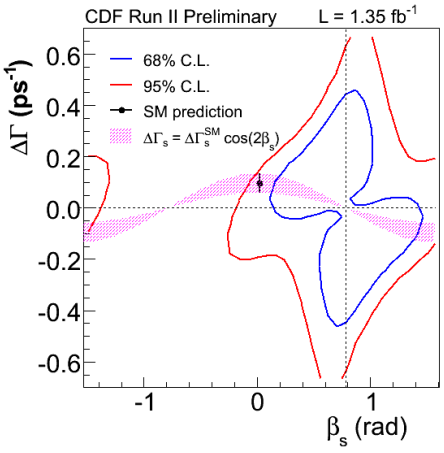

FIGURE 6.21: Likelihood ratio on data corrected according to the most conservative probability distribution of the current statistics depicted in Figure 6.20. The drawn confidence region is meant to reproduce the final Feldman-Cousins in Figure 6.14 to allow the incorporation of the $B_s^0 \to J/\psi\phi$ angular analysis in global fit frameworks. The pink band corresponds to the Equation $\Delta\Gamma_s = 2|\Gamma_{12}|\cos(2\beta_s)$, with $|\Gamma_{12}| = 0.048 \pm 0.018$ [75].

For the final inclusion inside a global fit CDF provides the values of the observed profile-$\mathcal{LR}$ in our data, sampled on a grid in the $\Delta\Gamma_s - \beta_s$) space and the "most conservative" probability distribution of this statistics, over the possible variations in the nuisance parameters (see Figure 6.20). The likelihood ratio obtained applying this correction is shown in Figure 6.21.

Global Fits Results

In presence of non asymptotic effects, The likelihood by itself is not a sufficient information to perform a global fit analysis which combine several analyses assuming the parabolic behaviour of the likelihoods involved.

Nevertheless, the procedure provided by CDF allows to properly include the angular analysis on $B_s^0 \to J/\psi\phi$ in a global framework, while accounting for all the peculiarities of the likelihood. Moreover, these new constraints will presumably have limited numerical impact on CKM fits within SM, while being particularly useful for NP scenario studies. In such analyses the typical need of many more hypotheses (both theoretical and experimental) than the ones needed for traditional SM fits - some of them being questionable - makes the analysis results to be understood in a more qualitative than quantitative way.

The CKM fitter group recently reported, for the first time, the B_s^0 Unitarity Triangle resulting from the equation:

$$\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*} + 1 + \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} = 0 \qquad (6.2.5)$$

As we can see from Figure 6.22 the triangle is quite squashed. This can be

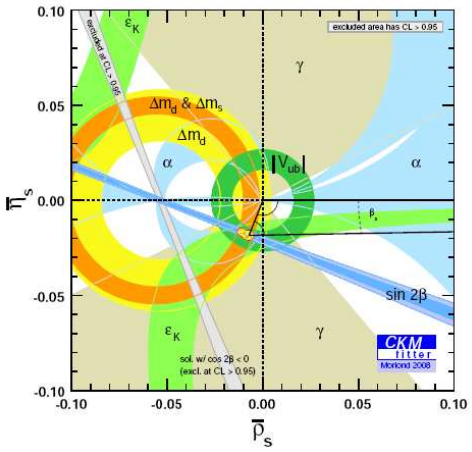

FIGURE 6.22: CKM fit of the B_s^0 Unitarity Triangle as presented in the Moriond Electroweak conference of 2008.

understood by looking at the order of magnitude of its sides from equation 6.2.6, $\mathcal{O}(\lambda^2) + \mathcal{O}(1) + \mathcal{O}(1) = 0$. The apex $(\bar{\rho}_s, \bar{\eta}_s)$ is determined via:

$$\bar{\rho}_s + i\bar{\eta}_s = -\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*} \tag{6.2.6}$$

The **UT**fit collaboration, on the other hand had a more direct approach to CDF and DØ results. By combining all the available experimental information, included the latest tagged analyses of $B_s^0 \to J/\psi\phi$ discussed in this dissertation, they found, out of their global fit framework, the B_s^0 mixing phase amplitude to deviate more than 3σ from the Standard Model prediction; accordingly they claimed the "first evidence of New Physics in $b \leftrightarrow s$ transitions" [85]. This result disfavors New Physics models with Minimal Flavor Violation (MFV) with the same significance. In fact the MFV models, being governed by the CKM matrix, expect a tiny value for the CP violation phase, as in the Standard Model case. Therefore, any experimental observation of sizable CP violation in B_s^0 mixing would clearly indicate New Physics as well as rule out the MFV paradigm.

The **UT**fit collaboration explored, with model-independent approach, the possible contributions of NP effects to $B_s^0 - \overline{B}_s^0$ mixing. The mixing process can be parametrized in terms of only two new parameters, which are chosen to quantify the difference of the amplitude, in absolute value and phase, with respect to the SM one:

$$C_{B_s} e^{2i\phi_{B_s}} = \frac{\langle B_s^0 | H_{eff}^{full} | \overline{B}_s^0 \rangle}{\langle B_s^0 | H_{SM}^{full} | \overline{B}_s^0 \rangle} = \frac{A_s^{SM} e^{-2i\beta_s} + A_s^{NP} e^{2i(\phi_{B_s}^{NP} - \beta_s)}}{A_s^{SM} e^{-2i\beta_s}} \tag{6.2.7}$$

where H_{eff}^{full} is the effective Hamiltonian generated by both SM and NP, while H_{eff}^{SM} only contains SM contributions. The interesting aspect of this formulation is the possibility to connect the experimental measurements to the SM expectations via:

Observable	68% Prob.	95% Prob.
ϕ_{B_s}	-19.9 ± 5.6	$[-30.45, \ -9.29]$
	-68.2 ± 4.9	$[-78.45, \ -58.2]$
C_{B_s}	$+1.07 \pm 0.29$	$[0.62, \ 1.93]$

TABLE 6.3: Fit results for the NP parameters from Formulas 6.2.8. For the ϕ_{B_s} there are two solutions corresponding to the ambiguity of the measurements from the $B_s^0 \to J/\psi\phi$ tagged analyses (see right plot in Figure 6.23).

$$\Delta m_s^{exp} = C_{B_s} \cdot \Delta m_s^{SM}$$
$$\beta_s^{exp} = \beta_s^{SM} - \phi_{B_s} \tag{6.2.8}$$

In Figure 6.23 we show the comparison of the 68% (dark) and the 95% (light) probability regions in the $\phi_{B_s} - C_{B_s}$ space, between the former analysis from the **UT**fit collaboration [85] and the previous result. The latter (left plot in the Figure 6.23) was not including the recent tagged analyses performed at the Tevatron experiments, but only the DØ untagged analysis on $B_s^0 \to J/\psi\phi$ [86]: the tighter bound on C_{B_s} is related to the very precise Δm_s measurement described along this dissertation. The phase ϕ_{B_s}, instead, presented originally a large range for the solution and the 4 bumps come from the 4-fold ambiguity in the likelihood for the angular analysis on $B_s^0 \to J/\psi\phi$ decay modes in its untagged version. When introducing the information from the tagged analyses on $B_s^0 \to J/\psi\phi$, the probability region along the ϕ_{B_s} axis is strongly reduced: the two 4-fold ambiguity is now reduced to a 2-fold, as already observed in the Feldman-Cousins confidence region reported by the CDF experiment. The results for the C_{B_s} and ϕ_{B_s} estimations are reported in Table 6.3.

Despite all the improvements, the power of the **UT**fit analysis performed is bound to the assumption done in the treatment of the input parameters used, and in particular of the DØ result. In DØ angular analysis, the two-fold ambiguity for arbitrary strong phases was removed exploiting the results on the phases measurement on the $B^0 \to J/\psi K^{*0}$ sample from BaBar [], with the underlying assumption of $SU(3)$ symmetry. This makes the likelihood strongly dependent on theoretical constraints, and it is often a source of criticism. Moreover, the strong phases in $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi\phi$ cannot be exactly related in the $SU(3)$ limit. In order to be conservative, the **UT**fit collaboration reintroduced the ambiguity in the DØ measurement, by taking DØ error as Gaussian and symmetrizing the likelihood with the application of the discrete ambiguity. What we have been observing in the CDF $B_s^0 \to J/\psi\phi$ analysis, and shown again in Figure 6.24, is that the application of the constraints on the phases alter the likelihood shape and a symmetrization will not be able to restore the original shape as it would be in absence of constraints. An attentive reader could still argue that the overlap between the likelihood with and without the strong phases constraints is very good in the region of interest to claim compatibility or not with the SM.

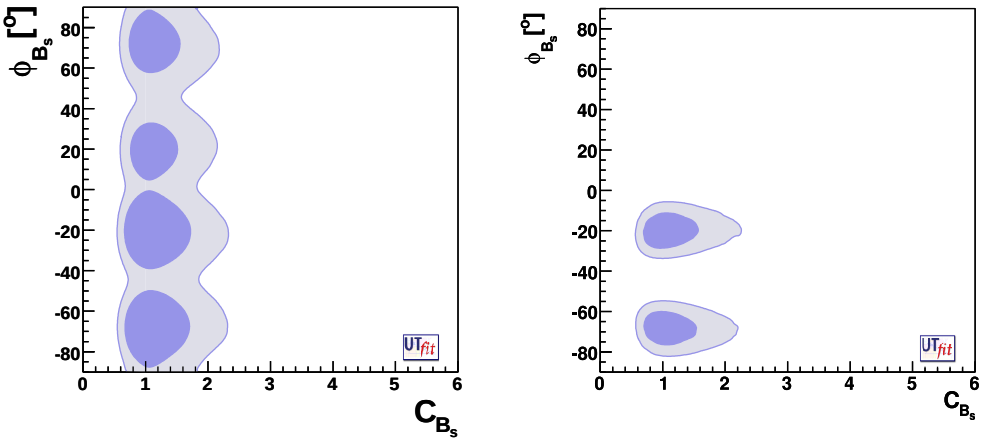

FIGURE 6.23: Comparison of the 68% (dark) and the 95% (light) probability regions in the $\phi_{B_s} - C_{B_s}$ space, between the former analysis from the **UT**fit collaboration (right) [85] and the previous result (left). The latter (left plot in the Figure) was not including the recent tagged analyses performed at the Tevatron experiments, but only the DØ untagged analysis on $B_s^0 \to J/\psi\phi$ [86].

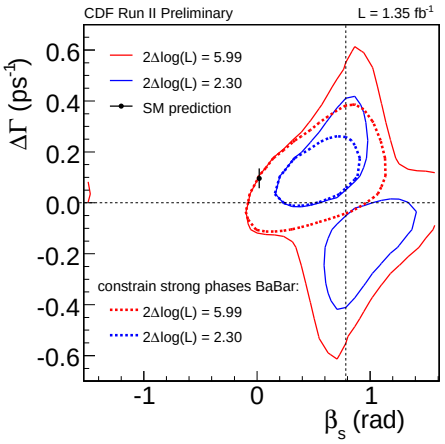

FIGURE 6.24: Comparison between the likelihood contours in the $\Delta\Gamma_s - \beta_s$ plane obtained on the data sample $B_s^0 \to J/\psi\phi$ with external constraint (dashed line) and without any constraint (continuous line). The blue curves correspond to 2.30 units above the minimum in the likelihood ($\sim 68\%$ probability) and the red curves to 5.99 units above the same minimum ($\sim 95\%$ probability). The black point corresponds to the Standard Model expectation $\Delta\Gamma = 0.096$ ps^{-1} and $\beta_s = 0.02$ [75]. In order, from top to bottom the constrained applied are: a constraint on the lifetime from the world average B^0 lifetime [6], a constraint on the strong phases $\delta_\parallel$ and $\delta_\perp$ from the $B^0 \to J/\psi K^{*0}$ decays [82], a constraint on both the lifetime and the strong phases.

Since the treatment of DØ result is not unique, given the available information, the **UT**fit collaboration tested the significance of the NP signal against different modeling of the probability density function. They used the 90% CL range for $\phi_s = [-0.06,\ 1.20]^\circ$ to estimate the standard deviation, resulting in using a $\phi_s = (0.57 \pm 0.38)^\circ$. In a second test, they implemented the likelihood profiles for ϕ_s and $\Delta\Gamma_s$ given by DØ, without accounting for correlations, but restoring the phase ambiguity. The two tests show still a deviation more than 3σ from zero for ϕ_{B_s}, although the exact number of σs depends on the procedure adopted. Additionally to the complex handling of DØ result, the likelihood ratio provided by CDF was used without the corrections required and discussed at the beginning of the Section, which does not account for the proper coverage heavily tested inside the CDF experiment and artificially increases the constraining power of the likelihood. All these discussions are a clear manifestation of the debate born around this interesting paper. Everybody hopes in the future to have a common approach to the method of quoting the result between the Tevatron collaboration, in a way to facilitate the combination in global fits, without too much workaround. In this direction the increasing integrated luminosity collected by the CDF and DØ will increase the size of the data samples as well, adding information which will help to cope with the very unstable likelihood and could strengthen or rule out the claimed evidence. In fact, although all the inputs used in the **UT**fit analysis have a remarkable agreement with the combined result, no single measurement has a 3σ yet. Nevertheless it does not take a complicated fit to see that both measurements are in the same direction. Thus, in conclusion, the results on the analyses on $\mathrm{B}_s^0 \to J/\psi\phi$ give a promising hint for the presence of New Physics beyond the SM and they will certainly need further investigations to support an evidence.

6.2.4 Conclusions

The second part of this dissertation describes the recent world's first angular analysis on $\mathrm{B}_s^0 \to J/\psi\phi$ for the measurement of the B_s^0 system parameters $\Delta\Gamma_s$ and β_s using the flavor tagging information. The result is based on approximately 2000 $\mathrm{B}_s^0 \to J/\psi\phi$ decays reconstructed in a 1.35 fb^{-1} data sample collected with the CDF II detector using $p\bar{p}$ collisions produced at the Fermilab Tevatron. Differently from the B_s^0 mixing case, where my contribution restricted to the flavor tagging area, I contributed with considerable effort to this analysis, taking part to all the stages required to finalize it. The analysis has been performed on a large sample of $\mathrm{B}_s^0 \to J/\psi\phi$ decay mode, selected with the use of a Neural Network. Thus, I was offered the opportunity to learn the principles of a Neural Network for the optimization of the selection. With the expertize gathered in the previous B_s^0 mixing analysis, I took care of the tagging performances understanding as well as their calibration. In particular for the Opposite Side Tagging we discarded the Neural Network based algorithm in favor of a hierarchical combination to avoid an observed asymmetry of the performances in matter with respect to anti-matter. The Neural Network based SSKT was inherited from the B_s^0 mixing analysis and applied unmodified in the likelihood. The initial step of the analysis was the performance of an untagged mass-time-space analysis [74] and,

at this stage, I directly contributed to the development of the control sample on B_d. A more detailed description is given in Chapter 7. After testing the robustness of the framework, we moved to the performance of multidimensional fit with the tagging information in there, to get involved in the statistical treatment of the result.

In summary, we presented a confidence bounds on the CP-violation parameter β_s and the width difference $\Delta\Gamma$ from the measurement of $B_s \to J/\psi\phi$ decays using flavor tagging. The confidence region was obtained using the Feldman-Cousins likelihood ratio ordering. The choice of confidence regions was driven by the non-Gaussian likelihood, with multiple minima that prevented the possibility to quote a reliable point-estimate. Feldman-Cousin was preferred to the "profile likelihood" because the entity of the non-linearity of the problem induced to support the hypothesis that the asymptotic limit was not reached with the current level of statistics. Assuming the Standard Model predicted values of $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.096\ ps^{-1}$ [75] ,the probability to observe a likelihood ratio equal to higher than what is observed in data is 15%, which corresponds to 1.5 Gaussian standard deviations. Treating $\Delta\Gamma_s$ as a nuisance parameter and fitting only for β_s, we find $\beta_s \in [0.16,\ 1.41]$ at the 68% confidence level. The presented experimental bounds restrict the knowledge of β_s to two of the four solutions allowed in measurements that do not use flavor tagging [74] [**?**] and improve the overall knowledge of this parameter.

We also exploit current experimental and theoretical information to extract tighter bounds on the CP violating phase. By applying the constraint $|\Gamma_{12}| = 0.048 \pm 0.018$ [75] in the relation $\Delta\Gamma_s = 2|\Gamma_{12}|\cos(2\beta_s)$, we obtain $\beta_s \in [0.12,\ 0.68] \cup [0.89,\ 1.45]$ at the 68% confidence level. If we additionally constrain the strong phases $\delta_\parallel$ and $\delta_\perp$ to the results from $B^0 \to J/\psi K^{*0}$ decays [82] and the B_s^0 mean width to the world average B^0 width [6], we find $\beta_s \in [0.20,\ 0.60]$ at the 68% confidence level.

This result gives a very promising hint for the presence of New Physics beyond the Standard Model, therefore it needs further investigations. Improvements are expected by refining the analysis and by simply including more data collected by the CDF II detector.

Chapter 7

Angular Analysis of $\mathrm{B}^0 \to J/\psi K^{*0}$ Decays

In this Chapter, we illustrate and discuss the angular analysis on the $\mathrm{B}^0 \to J/\psi K^{*0}$ *decay mode, also named* B^0 *transversity analysis, aiming at the measurement of the transversity amplitudes and strong phases which describe the decay mode itself.*

7.1 Introduction

When performing any analysis, a common procedure is to develop methods to validate the analysis itself. While studying the unknown properties of a specific decay mode, a general way to achieve such validation is to apply the developed framework on a kinematically equivalent decay mode, which properties are, on the other hand, very well known. Thus, an experimentalist can compare the results of what it is called *control sample* with the ones obtained independently from other experiments, which may have even a better statistical accuracy. Therefore, the control sample serves the purpose of improving the reliability of the main analysis. In our context, the angular studies on the $\mathrm{B}^0 \to J/\psi K^{*0}$ decay mode for the measurement of the transversity amplitudes and strong phases play a key role in the validation of the entire framework used for the $\mathrm{B}_s^0 \to J/\psi\phi$ angular analysis for the measurement of the decay-width difference and CP violating phase β_s. The analysis of the $\mathrm{B}^0 \to J/\psi K^{*0}$ decay properties represents my major contribution to the $\mathrm{B}_s^0 \to J/\psi\phi$ analysis performed in its first incarnation without the use of the flavor tagging information [74]. For its intrinsic nature of control sample, many of the technical aspects are in common with the $\mathrm{B}_s^0 \to J/\psi\phi$ analysis (performed with or without utilize flavor tagging information) and, whenever possible, useless repetitions will be avoided.

The Chapter is structured in the following way: the selections and reconstruction of the B^0 candidates will be detailed firstly, then followed by the likelihood description. Only after the validation of the likelihood fitter and the illustration of the sources of systematic uncertainties considered, the final results are discussed.

7.2 Data Sample

The selection of the events is performed with the use of an Artificial Neural Network which is anticipated by a loose Pre-selection, to remove clearly background-like events.

7.2.1 Sample Size and Reconstruction of Decays

The analysis of B^0 decays was performed on 1.35 fb^{-1} of data collected at the CDF Run II experiment. Though an additional 0.4 fb^{-1} of data was available we stopped at this point in integrated luminosity. The changes to the COT system, made after the first 1.35 fb^{-1}, required additional calibration which had not been performed at the at the time the analysis was implemented; we remind that the COT subsystem is heavily utilized for gathering the PID information, so crucial to distinguish pions and kaons in the B^0 decays, as well as for tagging algorithms (SSKT among others).

The `di-muon` trigger paths, detailed in Section 3.1.3, were used exclusively. While additional statistics would be gained by adding the displaced track trigger paths (`TTT`), this would complicate the analysis by introducing a non-trivial lifetime bias. The event yield in $B^0 \to J/\psi K^{*0}$ is comparatively low in the Two Track Trigger, but the background is strongly reduced. Nevertheless, at this stage the analysis is already competitive in terms of precision with the B factories results [87, 88], which makes possible a satisfying comparison. Thus, there was little to be gained by adding the `TTT` data sample at that point in time.

We analyzed data processed in the `BStntuple` framework [46]. Among salient details of the processing, we used:

- Event-by-event primary vertex reconstruction,
- 4-track vertexing for the B^0,
- dE/dx calibration and simulation as described in Reference [67],
- TOF efficiency and simulation as described in Reference [67].

7.2.2 Pre-selection

As anticipated, we use an Artificial Neural Network (ANN) for event selection after applying some basic Pre-selection cuts. These cuts are summarized in Table 7.1. They ensure a good Data-Monte Carlo agreement indispensable for the Neural Network training and reliability.

7.2.3 Neural Network Selection

In order to separate signal from background $B^0 \to J/\psi K^{*0}$ candidates we use an Artificial Neural Network. As already described, Neural Networks work by learning patterns in and correlations among event variables that are either signal-like or background-like (see Chapter **??**). In such way, they primarily improve the statistical

Cut Variable	Value of Cut
$\mathcal{P}(\mathrm{B}_s^0)$	$> 10^{-50}$
$\mathcal{P}(J/\psi)$	$> 10^{-50}$
$\mathcal{P}(\mathrm{K}^{*0})$	$> 10^{-50}$
μ_{nSiPhi}	≥ 3
p_T^{B} [GeV/c]	> 4.0
$p_T^{\mathrm{K}^{*0}}$ [GeV/c]	> 2.0
p_T^{μ} [GeV/c]	> 1.5
p_T^{K} [GeV/c]	> 0.4
$\lvert m(J/\psi \mathrm{K}^{*0}) - M_{PDG}^{\mathrm{B}} \rvert$ [MeV/c^2]	< 250
$\lvert m(\mu^+\mu^-) - M_{PDG}^{J/\psi} \rvert$ [MeV/c^2]	< 80
$\lvert m(K\pi) - M_{PDG}^{\mathrm{K}^{*0}} \rvert$ [MeV/c^2]	< 80
$\sigma_{c\tau}$ [μm]	< 150
CLL^{K}	> -50.0
CLL^{π}	> -50.0

TABLE 7.1: Pre-selection cuts on the $\mathrm{B}^0 \rightarrow J/\psi K^{*0}$ sample collected with the `di-muon` trigger paths.

significance of the sample by reducing background and enhancing signal. We choose the NeuroBayes package [89] which is one of the most advanced implementations of a Neural Network. Assuming that the background events in the B^0 mass regions have similar properties as under the mass signal region, the network is trained using the sidebands from both sides of the mass peak of the data as a background sample. The lower sideband mass regions is $m \in [5.13 - 5.2094]$ GeV/c^2, while the upper one is $m \in [5.3294 - 5.43]$ GeV/c^2. A realistic Monte Carlo, using `BGenerator` for event generation and `EvtGen` for the decays, is utilized as a signal sample. As for the data, the Monte Carlo sample is processed using the `BStntuple` framework.

The discriminating variables used for the training, we used are listed in Table 7.2 together with the plot showing their relative correlations.

Figure 7.1 shows the results of the training, in terms of the Neural Network output for signal and background separation and purity dependence as a function of the NN output itself. No deviation from the linear dependence indicates an optimal behaviour of the NN. The network output is the chosen variable used to make the final candidate selection. On an event-by-event basis, the ANN assign a value between -1 and 1 to the reconstructed B^0 candidate, according to its probability of being background (-1) or signal (1). We use as unbiased criterion to optimize the selection the statistical significance, defined as $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, and choose the ANN output value cut which would give the greatest statistical significance. In Figure 7.2 we show the significance as a function of the Neural Network output cut. The selected cut corresponds to 0.5: $nnOut > 0.5$.

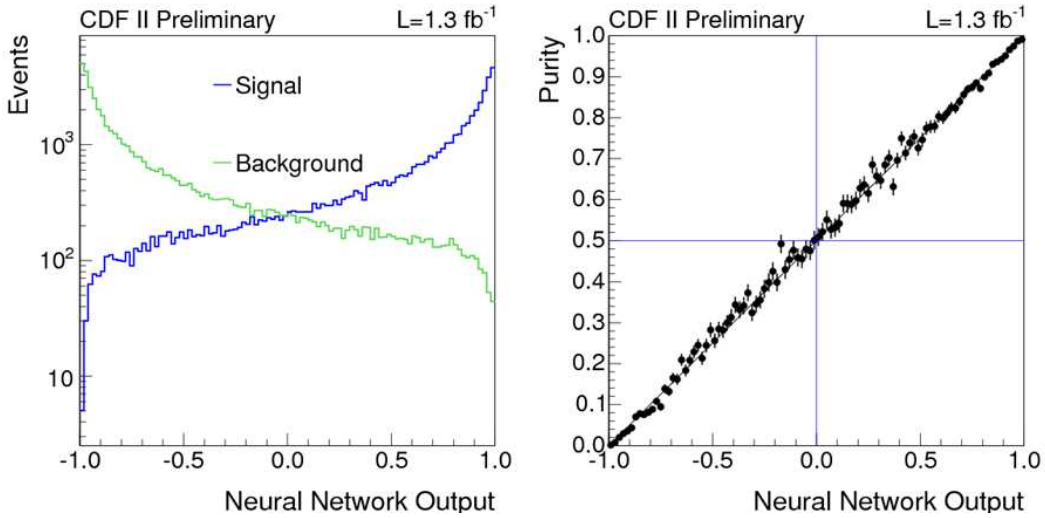

FIGURE 7.1: Left: Neural Network output distributions for simulated B^0 meson as signal (blue line) and mass sidebands events for background (green line). Right: purity of the training sample as a function of the Neural Network output.

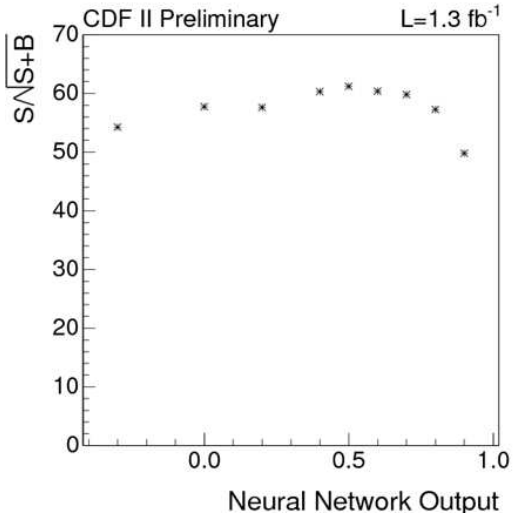

FIGURE 7.2: Significance $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$ as a function of the Neural Network cut for the $B^0 \to J/\psi K^{*0}$ decay mode.

#	Variable B^0
3	p_T^B
10	$p_T^{J/\psi}$
5	$p_T^{K^{*0}}$
8	p_T^K
9	p_T^π
15	$max(p_T^{\mu^+}, p_T^{\mu^-})$
16	$min(p_T^{\mu^+}, p_T^{\mu^-})$
12	p_T^B
7	$\lvert m(\mu^+\mu^-) - M_{PDG}^{J/\psi}\rvert$
4	$\mathcal{P}(B_s^0)$
11	$\mathcal{P}(J/\psi)$
6	$\mathcal{P}(K^{*0})$
2	$\chi^2_{xy}(B)$
17	$CLL_k(K)$
18	$CLL_\pi(K)$
19	$CLL_k(\pi)$
20	$CLL_\pi(\pi)$
13	$max(lik^{\mu^+}, lik^{\mu^-})$
14	$min(lik^{\mu^+}, lik^{\mu^-})$

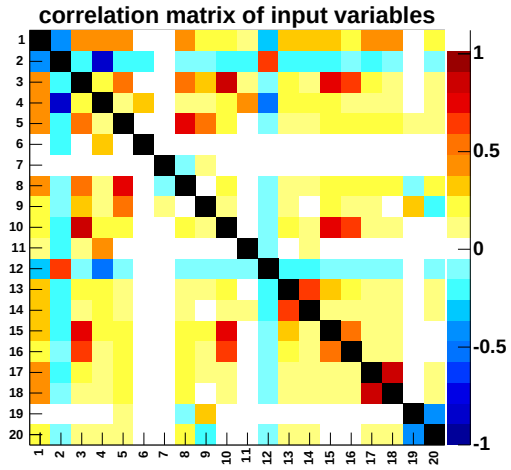

TABLE 7.2: Left: list of the variables used in the Neural Network training for the $B^0 \to J/\psi K^{*0}$ decay mode. Right: Correlations of the variables.

7.2.4 K^{*0} Swap Suppression

For the $B^0 \to J/\psi K^{*0}$ decay mode the mass association of the K^{*0} products, kaon and pion, could be ambiguous. The direct consequence is that a certain percentage of our B^0 candidates is reconstructed with the wrong kaon-pion hypothesis. We usually refer to them as the "swapped" candidates. Previous studies for the B_s^0 mixing analysis on the B^0 sample showed this percentage to be at the level of 10 %. Although it was used a less performing sequential selection (called "rectangular"), this component cannot be discarded also with a general Neural Network selection; in fact it was initially confirmed while performing the fit without an additional selection, which we are going to describe in this section.

In fact, to cope with the swap component of the sample, we are faced with two possible paths to follow. A first solution could be to parametrize this contribution in the likelihood fit and separate two components for the signal: one for the events reconstructed with the correct mass assignment and another one for the wrong mass assignment. While this is valid in principle, its practicality requires an increasing level of complexity in the likelihood description which may easily result in a large systematic contribution. For instance to parametrize the mass model of the swapped component we have to rely on a Monte-Carlo simulation as well as for the fraction of events reconstructed with the mass mis-assignment. Thus we decided to move towards

another approach which consists in the training of a dedicated Neural Network to remove the fraction of candidates reconstructed with the misidentified kaon and pion mass assignment. The list of variables chosen to train the NeuroBayes Neural Network is reported in Table 7.3 as well as their correlations plot. The Neural Network output (nnOutSwap) and purity distributions are depicted in Figure 7.3. For the training, the same sample of simulated events has been used as signal, reconstructing the events with the correct mass assignment, and as background, reconstructing the events with the wrong mass assignment (the swapped candidates). The ANN can distinguish very well the correct reconstructed candidate from the swapped one and by applying a cut at 0.8 - nnOutSwap > 0.8 - the fraction of swapped candidates is reduced to 0.5%, as measured in our MC sample. The uncertainty on the determination of this fraction from simulated events will be treated as a systematic error.

#	Variable B^0
2	p_T^{K}
3	$\sigma(p_T^{\mathrm{K}})$
4	p_T^{π}
5	$\sigma(p_T^{\pi})$
6	$\lvert M_{K\pi}\rvert$
7	$CLL_K(K)$
8	$CLL_\pi(K)$
9	$CLL_\mu(K)$
10	$CLL_e(K)$
11	$CLL_p(K)$
12	$CLL_K(\pi)$
13	$CLL_\pi(\pi)$
14	$CLL_\mu(\pi)$
15	$CLL_e(\pi)$
16	$CLL_p(\pi)$

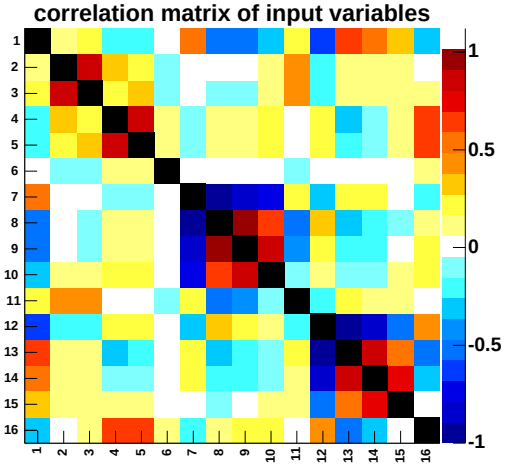

TABLE 7.3: Left: list of the variables used in the swap suppression Neural Network training for the $B^0 \to J/\psi K^{*0}$ decay mode. Right: Correlations of variables.

In the final selection for the $B^0 \to J/\psi K^{*0}$ before applying the Neural Network selection cut (nnOut > 0.5), we first remove the swap candidates using the dedicated Neural Network for the swap suppression (nnOutSwap > 0.8). Figure 7.4 compares the invariant mass distribution after the loose Pre-selection cuts and after the two Neural Network selections. A yield of about 7800 signal events is obtained. The ensemble of candidates will be treated as a sample of reconstructed candidates with the correct mass assignment for the kaon and the pion coming from the decay chain $K^{*0} \to K^+\pi^-$. Again, the validity of this assumption will be accounted in the systematics evaluation.

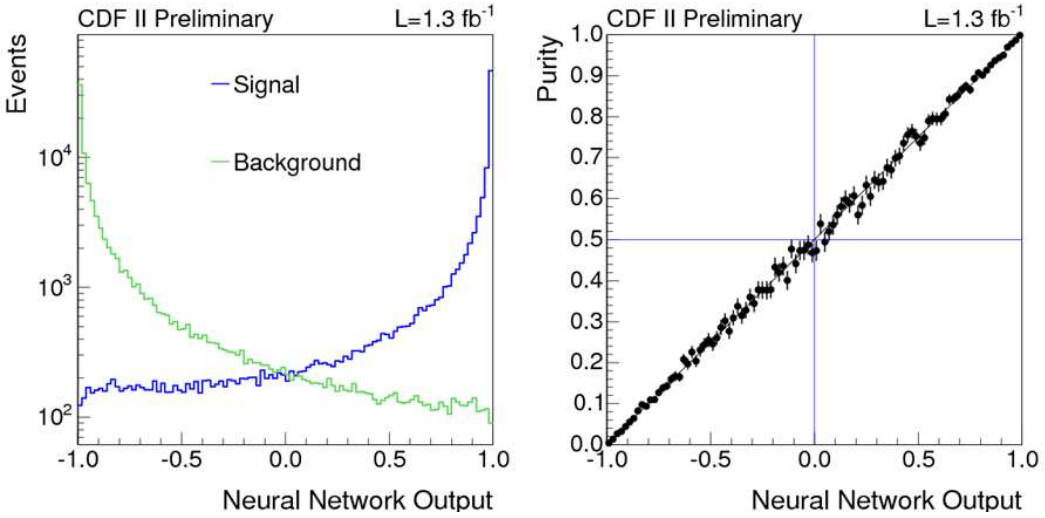

FIGURE 7.3: Left: Neural Network output distributions for simulated B^0 meson reconstructed with the correct mass assignment for the kaon and pion produced in the K^{*0} decay and treated as signal (blue line) and for the same simulated events but reconstructed with the wrong mass assignment as background (green line). Right: purity of the training sample as a function of the swap suppression Neural Network output.

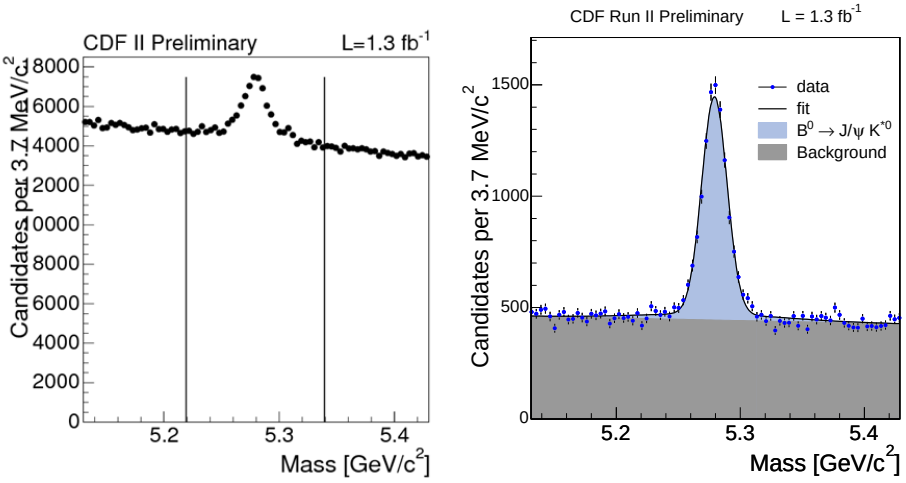

FIGURE 7.4: Invariant mass distribution after the loose pre-selection cuts (left) and after the Neural Network cuts which include the swap suppression selection.

7.3 Likelihood

In the following Section we describe the fitter frameworks utilized to obtain the measurements of the amplitude and strong phases for the $B^0 \to J/\psi K^{*0}$ decay mode. In order to extract the parameters of interest, a simultaneous fit in mass, lifetime and angles has been performed.

7.3.1 Observables

The reconstructed variables describing the decay mode and used as input to the fitter are

- Mass: m,
- Proper Decay Length: ct, σ_{ct},
- Angles: $\vec{\omega} \equiv (\cos\Theta_T, \phi_T, cos\Psi_T)$,

where m is the reconstructed mass of the B^0 candidate and ct is the proper decay length calculated from the measurements of the transverse decay length L_{xy} and the transverse momentum p_T, combined with the B^0 mass from the PDG world average [58]:

$$ct = \frac{L_{xy}}{p_T} \cdot m^{B^0}_{PDG} \tag{7.3.1}$$

Analogously, the uncertainty on the proper decay length is defined as:

$$\sigma_{ct} = \frac{\sigma_{L_{xy}}}{p_T} \cdot m^{B^0}_{PDG} \tag{7.3.2}$$

The angles forming the vector $\vec{\omega}$ are defined in the transversity basis where a convenient description of the decay is given: the angles have been already introduced in Section 5.1 and pictorially shown in Figure 5.1.

7.3.2 Mass Model

The signal mass distribution is parametrized using two Gaussians with the same mean and two different widths, associated with the detector resolution. The following parameters are let floating in the fit: the mean of the Gaussians (M), the width of the narrower Gaussian (σ_1) and the ratio of the two Gaussians (R, where $\sigma_2 = R \cdot \sigma_1$), and the relative fraction of the wider Gaussian (f_m). The B^0 mass signal probability for an event with reconstructed mass m_j is

$$\mathcal{P}^{sig}_m(m_j \big|\, M, \sigma_1, R, f_m) = (1 - f_m)\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(m_j - M)^2}{2\sigma_1^2}} + f_m \frac{1}{\sqrt{2\pi} R \sigma_1} e^{-\frac{(m_j - M)^2}{2R^2\sigma_1^2}}, \tag{7.3.3}$$

The background model is described by a first-degree polynomial function normalized in the fitting range:

$$\mathcal{P}_m^{bkg}(m_j \mid A) = A \cdot m + \frac{1}{M_{max} - M_{min}}[1 - \frac{A}{2}(M_{max}^2 - M_{min}^2)] \tag{7.3.4}$$

where A is the slope of the polynomial, and M_{min}, M_{max} are the boundaries of the mass window, respectively 5.13 GeV/c^2 and 5.43 GeV/c^2. The inclusion of mass model in the fit is primarily used to discriminate between signal and background events.

7.3.3 Lifetime Model

The proper decay length signal distribution for the B^0 mesons is modeled by an exponential PDF convolved with a Gaussian function that accounts detector resolution effects. If ct_j is the measured proper decay length and σ_{ct_j} is its uncertainty, then the PDF for the signal B_d is given by:

$$\mathcal{P}_{ct}^{sig}(ct_j, \sigma_{ct_j} \mid c\tau, S_{ct}) = E(ct_j \mid c\tau) \otimes G(ct_j, \sigma_{ct_j} \mid S_{ct}), \tag{7.3.5}$$

where E and G are defined as follows:

$$E(ct \mid c\tau) = \begin{cases} 0 & , \quad ct < 0 \\ \frac{1}{c\tau} e^{-\frac{ct}{c\tau}} & , \quad ct \geq 0, \end{cases}$$

$$G(ct, \sigma \mid S_{ct}) = \frac{1}{\sqrt{2\pi} S_{ct} \sigma} e^{-\frac{c^2 t^2}{2(S_{ct}\sigma)^2}}.$$

S_{ct} scale factor is used to correct over- or under-estimated proper decay length uncertainties (σ_{ct_j}).

The background distribution for the proper decay length is empirically described as the sum of several contributions:

- A δ-function centered at zero to model the large background candidates constructed with tracks coming from the primary vertex, e.g. prompt J/ψ paired with prompt track(s).
- A short-lived exponential at negative lifetime and another at positive lifetime to handle the mis-measured candidates.
- A long lived exponential at positive lifetime to describe the long lived background component. We convolve all terms with the Gaussian resolution function.

Combining all these, we obtain:

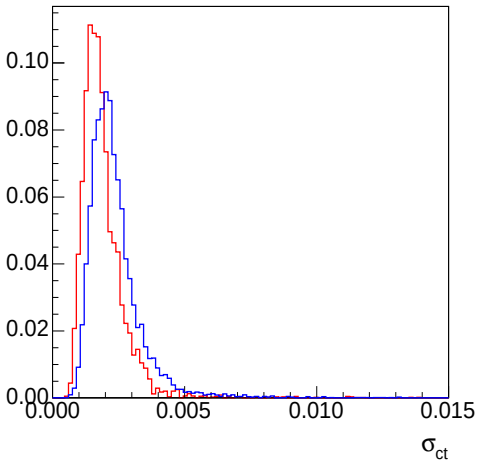

FIGURE 7.5: Distributions for the σ_{ct} obtained from data for the signal sideband subtracted (red line) and the sideband (blue line) in the $B^0 \to J/\psi K^{*0}$ decay mode.

$$\begin{aligned}
&\mathcal{P}^{ct}_{bkg}(ct_j, \sigma_{ct_j} \mid f_-, f_+, f_{++}, \lambda_-, \lambda_+, \lambda_{++}, S_{ct}) = \\
&(1 - f_- - f_+ - f_{++}) \otimes G(ct_j, \sigma_{ct_j} \mid S_{ct}) \\
&+ f_- \cdot E(-ct_j | c\tau_-) \otimes G(ct_j, \sigma_{ct_j} \mid S_{ct}) \\
&+ f_+ \cdot E(ct_j | c\tau_+) \otimes G(ct_j, \sigma_{ct_j} \mid S_{ct}) \\
&+ f_{++} \cdot E(ct_j | c\tau_{++}) \otimes G(ct_j, \sigma_{ct_j} \mid S_{ct})
\end{aligned} \tag{7.3.6}$$

where $c\tau_-, f_-$ are the decay constant and fraction of the negative-lifetime tail, $c\tau_+, f_+$ are the decay constant and fraction of the short-live positive-lifetime tail, and $c\tau_{++}, f_{++}$ are the decay constant and fraction of the long-lived positive-lifetime tail. All these parameters, as well as the scale factor S_{ct} are allowed to float in the fitter.

Using the conditioned probability definition, as reported in Ref [73], the signal and the background proper-decay-length PDFs are finally multiplied by their σ_{ct} distributions, $\mathcal{P}_{ct}(\sigma_{ct})$ which are extracted from data. Figure 7.5 refers to the $B^0 \to J/\psi K^{*0}$, where the σ_{ct} distributions for signal (obtained using the data sideband-subtracted technique) and for background (from the event in the sideband regions of the mass space) are distributed differently. The introduction of such term gives more discriminating power to the fitter.

7.3.4 Angular Model

Signal Angular Model

The time-integrated probability density used to describe the angular distribution for the signal is taken from the theoretical calculation of the differential decay rate in angles and time reported in Equation [**?**]. It is important to stress that the formula has been obtained summing over the initially produced B^0 and $\overline{B}^0$, without taking into account the flavor identification at the production time. In fact the analysis of the $B^0 \to J/\psi K^{*0}$ as control sample was performed in the context of the $B_s^0 \to J/\psi\phi$ untagged angular analysis [74], which anticipated the introduction of the tagging information described in this dissertation. The signal angular PDF is:

$$\begin{aligned}\mathcal{P}^{sig}_{\vec{\omega}}(\vec{\omega}\,|A_0, A_\parallel, A_\perp) = \frac{1}{Z_n} g_P(\vec{\omega}\big|A_\alpha) \times A(\vec{\omega}) = \\ \frac{1}{Z_n} \cdot \{|A_0|^2 \cdot f_1(\vec{\omega}) \; + \big|A_\parallel\big|^2 \cdot f_2(\vec{\omega}) + \\ \big|A_\perp\big|^2 \cdot f_3(\vec{\omega}) \pm Im(A_\parallel^* A_\perp) \cdot f_4(\vec{\omega}) \; + \\ Re(A_0^* A_\parallel) \cdot f_5(\vec{\omega}) \pm Im(A_0^* A_\perp) \cdot f_6(\vec{\omega}) \; \} \times A(\vec{\omega})\end{aligned} \tag{7.3.7}$$

where the sign $+(-)$ is used for the flavor-specific final state of the $K^{*0} \to K^+\pi^-$ ($K^-\pi^+$) and Z_n is the normalization factor. In the likelihood construction the decay rate formula is multiplied by tri-dimensional acceptance curve $A(\vec{\omega})$. In fact, the CDF detector does not have uniform efficiency and performances in the tri-dimensional space. Thus we expect not to have uniform acceptance for the transversity angles reconstructions. So it is crucial for the analysis to be able to unfold this effect, which is usually referred as sculpting, in CDF jargon. In order to study the detector impact on the angles reconstructions, we generate a flat phase space Monte Carlo, uniform in all its angular distributions. The simulated events, generated with `BGenerator` and decayed using `EvtGen`, are then passed through the full fledged detector simulation. Thus, we obtain a tri-dimensional acceptance curve $A(\vec{\omega})$, which describes the effect of the detector reconstruction. The projections of $A(\vec{\omega})$ onto the three angles are shown in Figure 7.6. The so-called sculpting effect is taken into account by this multiplicative term in the signal angular PDF. The normalizing is performed accordingly. The implementation of normalization factor Z_n is discussed in Appendix B.

S-wave interference

In the first stage of the analysis, the time-integrated angular distributions, used to describe the decay channel $B^0 \to J/\psi K^{*0}$, are assuming that the $K\pi$ system may be described entirely in terms of P-wave amplitudes. The differential rate, as obtained from Equation 7.3.7, is invariant under the transformation:

$$(\delta_\parallel - \delta_0,\; \delta_\perp - \delta_0) \leftrightarrow (\delta_0 - \delta_\parallel,\; \pi + \delta_0 - \delta_\perp) \tag{7.3.8}$$

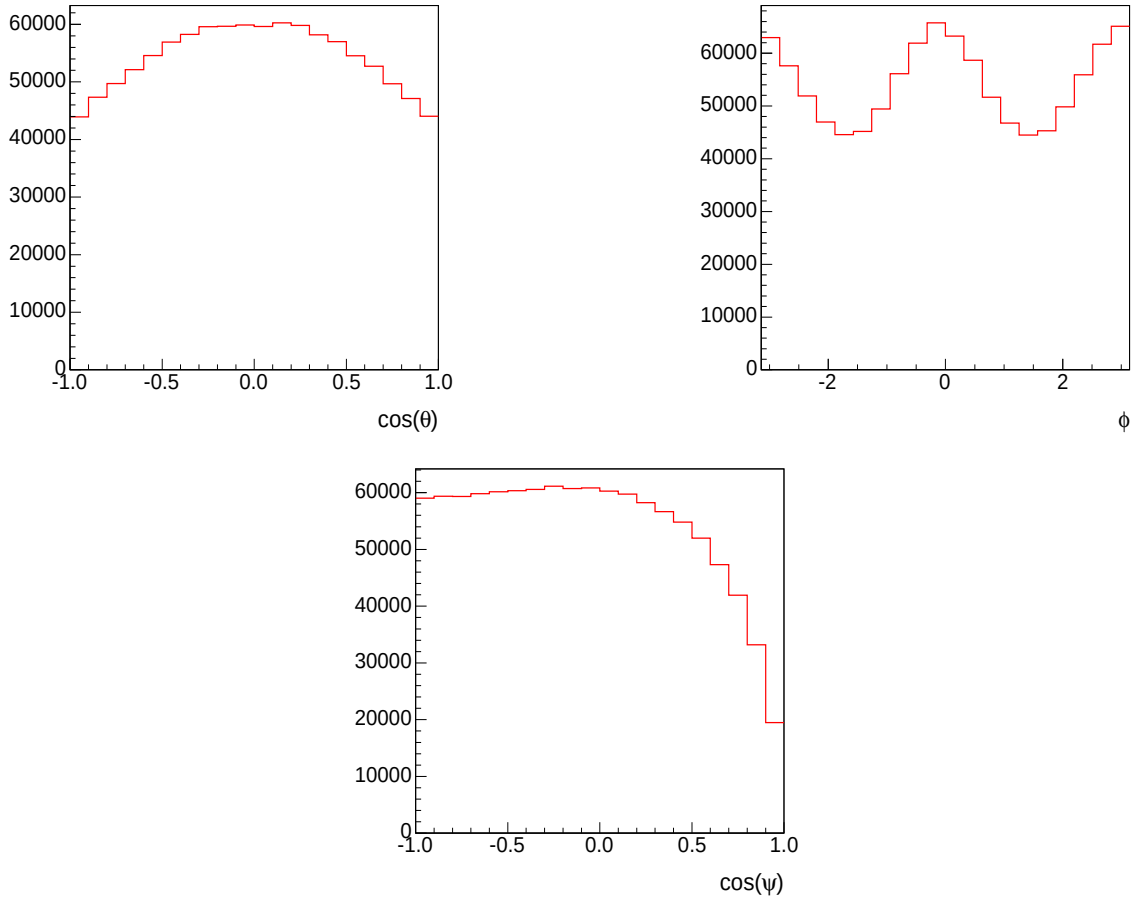

FIGURE 7.6: Distributions of the three angles describing the $B^0 \to J/\psi K^{*0}$ candidates after reconstruction. From top left to bottom: $\cos(\theta_T)$, ϕ_T, and $\cos(\psi_T)$. The three distributions are generated flat.

At the BaBar experiment this phase ambiguity was solved, including the $K\pi$ S-wave contribution and then measuring the $K\pi$ mass-dependence of its phase difference with respect to the P-waves [82]. It was shown that the inclusion of a $K\pi$ S-wave with a significant S-P interference is required to describe the BaBar data. This is also the case for the CDF $B^0 \to J/\psi K^{*0}$ angular analysis with the current available statistics. To properly account for this effect present in our data but not in our MC simulation, one should introduce in the likelihood the presence of the S-wave and its complete dependence for the K^{*0} mass. With this addition, the new signal angular likelihood, given the mass value $m_{K\pi}$, becomes

$$\begin{aligned} g_{S+P}(\vec{\omega}\big|m_{K\pi}, A_\alpha, \lambda) = cos^2(\lambda)\cdot g_P(\vec{\omega}\big|A_\alpha) + sin^2(\lambda)\cdot f_7(w) + \\ \frac{1}{2} sin2\lambda \cdot [f_8(w)cos(\delta_\parallel - \delta_S)\big|A_\parallel\big| + \\ f_9(w)cos(\delta_\perp - \delta_S)\big|A_\perp\big| + f_{10}(w)cos(\delta_S)|A_0|], \end{aligned} \tag{7.3.9}$$

where

$$\begin{aligned} f_7(w) &= \frac{3}{32\pi} 2[1 - sin^2\theta cos^2\phi], \\ f_8(w) &= -\frac{3}{32\pi} \cdot \sqrt{6} sin\psi \cdot sin^2\theta \cdot sin2\phi, \\ f_9(w) &= \frac{3}{32\pi} \cdot \sqrt{6} sin\psi \cdot sin2\theta \cdot cos\phi, \\ f_{10}(w) &= \frac{3}{32\pi} 4\sqrt{3} cos\psi [1 - sin^2\theta cos^2\phi], \end{aligned} \tag{7.3.10}$$

and at any given mass $K\pi$, the normalization is obtained by introducing the parametrization

$$\begin{aligned} cos\lambda &= \frac{A_P}{\sqrt{A_P^2 + |A_S|^2}}, \\ sin\lambda &= \frac{|A_S|}{\sqrt{A_P^2 + |A_S|^2}}, \end{aligned} \tag{7.3.11}$$

where λ is in the range $[0, \pi/2]$. The term $\cos^2\lambda$ ($\sin^2\lambda$) represents the fraction of the P-wave (S-wave) intensity at that value of mass $K\pi$.

In Equation 7.3.9, the dependence of g_{S+P} on mass $K\pi$ follows from that of λ and the strong phases $\delta_i (i = 0, \|, \perp, S)$. The S-P interference contributions essentially average out over the broad ($\pm$80 MeV/c^2) $K\pi$ mass interval centered on the K^{*0}(892) (from the analysis pre-selection cut) and, since the S-wave intensity is proven to be only a few percent of that of the P-wave, the presence of the S-wave is in general accounted for a small additional source of systematic uncertainty (as in BaBar). The path followed in the CDF analysis is now defined. Initially, the results will be given only in terms of P-wave, exploiting the solution of the phase ambiguity solved by BaBar Collaboration [82]. At this point we will observe how the data indeed needs an additional contribution to be explained completely and how the S-wave accounts for this lack in the description. This term will be only considered as an average contribution over the $K\pi$ mass, therefore a systematic error will be evaluated.

Background Angular Model

For the background description of the angle distributions we use an empirical model obtained extracting the distributions of the angles from the sideband regions defined as $[5.13 - 5.21]$ GeV/c$^2 \cup [5.35 - 5.43]$ GeV/c^2. With the underlying assumption of the three distributions not to being correlated among each other, the PDF introduced in the fitter framework are

$$\mathcal{P}_{\vec{\omega}}^{bkg}(\vec{\omega}) = \mathcal{P}_{\theta_T} \cdot \mathcal{P}_{\phi_T} \cdot \mathcal{P}_{\psi_T}, \tag{7.3.12}$$

where

$$\mathcal{P}_{\theta_T} = \frac{1}{N_\theta} \left(1 + A_2 cos^2\theta_T + A_4 cos^4\theta_T\right) \tag{7.3.13}$$

$$\mathcal{P}_{\phi_T} = \frac{1}{N_\phi} \left(1 + B_1 cos(2\phi_T + B_{shift}) + B_2 cos^2(2\phi_T + B_{shift})\right) \tag{7.3.14}$$

$$\mathcal{P}_{\psi_T} = \frac{1}{N_\psi} \Big(1 + C_1(1 - cos\psi_T) + C_2(1 - cos^2\psi_T) + C_3(1 - cos^3\psi_T) + C_4(1 - cos^4\psi_T) + C_5(1 - cos^5\psi_T)\Big) \tag{7.3.15}$$

Figure 7.7 shows the projections for the lifetime and the angular distributions when fitting only in the sideband regions. The good agreement between the data and the fit results makes us confident to have implemented a correct description of the background events.

7.3.5 Complete Likelihood for the $B^0 \to J/\psi K^{*0}$ Decays

After the formulation of each component we can combine all the information together and obtain the final likelihood construction:

$$\mathcal{L}(m, ct, \sigma_{ct}, \vec{\omega}) = (1 - f_B)\mathcal{P}_m^{sig} \cdot \mathcal{P}_{ct}^{sig} \cdot \mathcal{P}_{\vec{\omega}}^{sig} + f_B \mathcal{P}_m^{bkg} \cdot \mathcal{P}_{ct}^{bkg} \cdot \mathcal{P}_{\vec{\omega}}^{bkg}, \tag{7.3.16}$$

Differently from what happens in the $B_s^0 \to J/\psi\phi$ case, where the angular and lifetime components cannot be separated one from the other, all the components in the likelihood for the $B^0 \to J/\psi K^{*0}$ decay mode are disentangled.

7.3.6 Fitter Validation

Before carrying out the fit over the data sample, we perform several tests to the maximum likelihood fitter. The aims is to validate the correctness of the implementation and test the statistical limits of our sensitivity to the fit parameters. Moreover, investigating our assumption and the behavior of the likelihood under various circumstances, we may detect any potential fit bias.

Likelihood factorization

As observed, for the $B^0 \to J/\psi K^{*0}$ angular analysis the signal lifetime and angular PDFs are not interconnected among each other. While this assumption is true at theoretical level, it is important to verify that the detector does not introduce any unexpected correlation. In order to perform such check, we restricted to the events in the signal region around the peak. We then divide each angle variation in 4 sub-ranges. For each of these sub-ranges, we produce the lifetime sideband subtracted

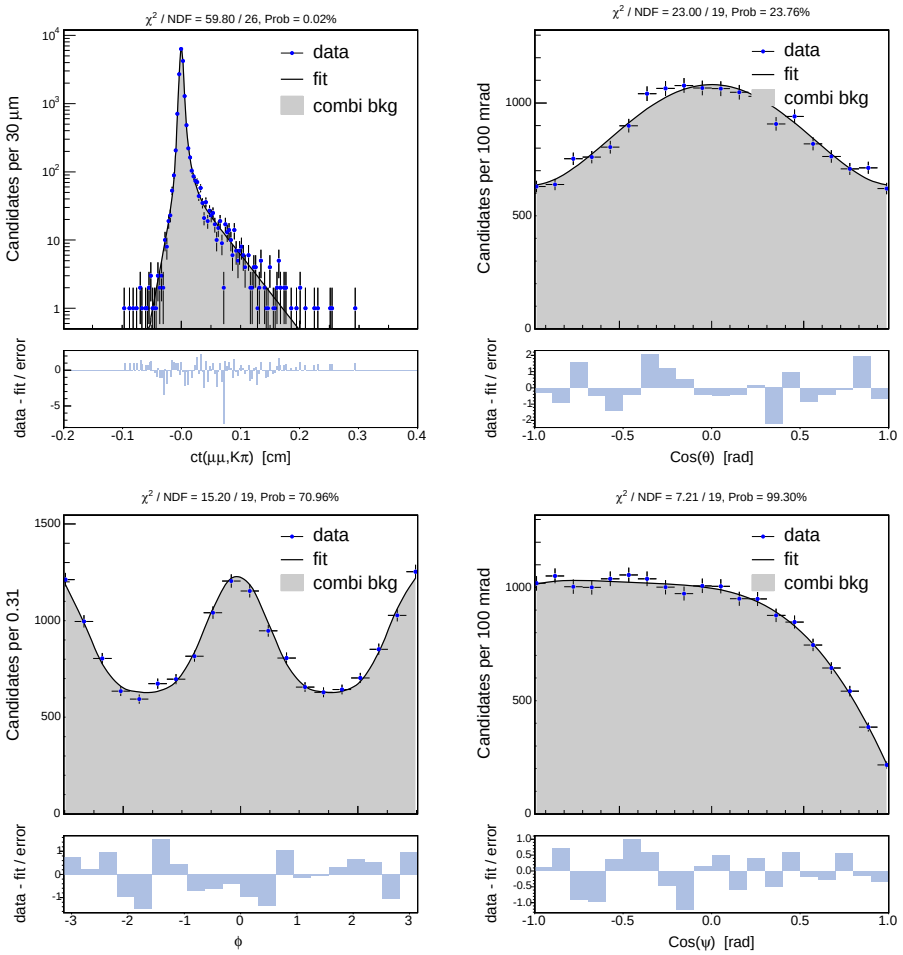

FIGURE 7.7: Projections obtained fitting the sideband region defined as $[5.13 - 5.21] \cup [5.35 - 5.43]$. From top left to bottom right: proper decay time ct, $\cos(\theta_T)$, ϕ_T, and $\cos(\psi_T)$.

distribution. We then compare the normalized lifetime distributions among themselves to check if there is any dependency for different values of the angles considered. Figure 7.8 shows no unexpected discrepant behaviour in the lifetime distributions as we vary the angles ranges considered. Thus, we conclude that CDF detector does not introduce additional correlations between the lifetime and the angles and the likelihood factorization test has been successful.

Test with Monte Carlo Toys

Pull distributions are a commonly used method to address several of questions, especially if testing the likelihood for potential biases. The procedure is implemented as follows: we generate a large set of pseudo-experiments, also referred as "toy Monte Carlo", randomly polling the probability density functions of the likelihood to assign event variables. For each of these pseudo-experiments, we perform the fit as we would on data. For each parameter "A" allowed to float in the fit, the pull distribution P

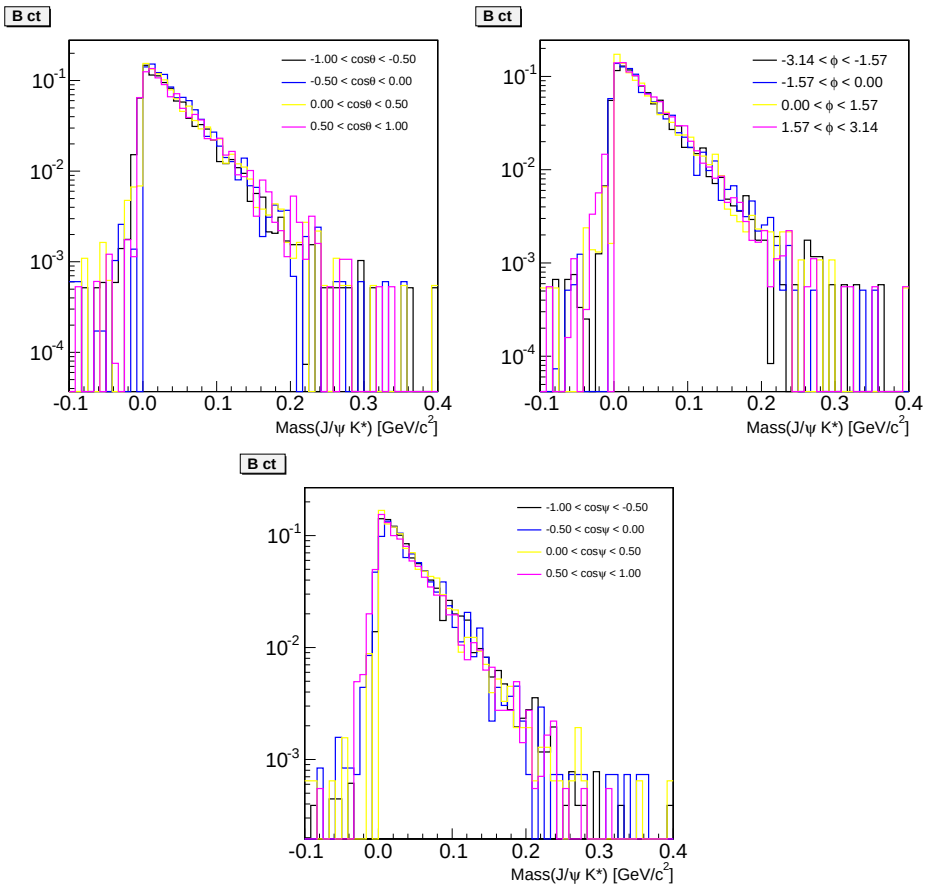

FIGURE 7.8: Normalized sideband subtracted lifetime distributions for even spaced angles ranges.

is defined as follows:

$$P = \frac{A^{fit} - A^{input}}{\sigma^A}. \tag{7.3.17}$$

In order to be able to claim that the likelihood fitter can measure a given quantity, P for that specific quantity should follow a Gaussian distribution with a mean at 0 and a width of 1. In our pull studies, we adopt the convention that the pull for a given pseudo-experiment is calculated using the positive error in the denominator if the fit value is lower than the input value and the negative error otherwise. Naturally, several complications arise. In some cases, there can be a hard physical limit on one or both sides of the allowed parameter range. In other cases, a degeneracy in the likelihood expression itself can lead to a fit value that is simply another equivalent minimum without being a physically different value (as far as the likelihood itself would be able to measure).

We begin testing the likelihood by producing 1000 toy MC samples with a 1.35 fb^{-1} equivalent statistic, using the same signal fraction we would get from data. The results are shown in Table 7.4

Parameter		MEAN	RMS
M(mass)	$[GeV]$	-0.079 ± 0.030	0.949 ± 0.021
σ_1	$[GeV]$	-0.010 ± 0.031	0.989 ± 0.022
f_m		-0.021 ± 0.033	1.028 ± 0.023
R$(=\sigma_1/\sigma_2)$		-0.038 ± 0.033	1.024 ± 0.023
f_{Bkg}		-0.017 ± 0.032	1.012 ± 0.023
$A(= slope_{bkg})$		-0.026 ± 0.031	0.982 ± 0.022
$c\tau$	$[\mu m]$	$+0.012 \pm 0.034$	1.059 ± 0.024
S_{ct}		-0.020 ± 0.031	0.975 ± 0.022
f_P		$+0.072 \pm 0.033$	1.023 ± 0.023
f_{PP}		-0.008 ± 0.033	1.035 ± 0.023
f_N		-0.019 ± 0.031	0.986 ± 0.022
$c\tau_P$	$[\mu m]$	-0.073 ± 0.033	1.038 ± 0.023
$c\tau_{PP}$	$[\mu m]$	-0.026 ± 0.033	1.042 ± 0.023
$c\tau_N$	$[\mu m]$	-0.046 ± 0.032	1.022 ± 0.023
$\lvert A_0 \rvert^2$		$+0.003 \pm 0.033$	1.027 ± 0.023
$\lvert A_\parallel \rvert^2$		$+0.020 \pm 0.031$	0.986 ± 0.022
$\delta_\parallel$		-0.001 ± 0.034	1.069 ± 0.024
$\delta_\perp$		$+0.033 \pm 0.032$	0.999 ± 0.022
A_2		-0.053 ± 0.034	1.073 ± 0.024
A_4		$+0.048 \pm 0.034$	1.065 ± 0.024
B_1		$+0.031 \pm 0.031$	0.968 ± 0.022
B_2		-0.007 ± 0.031	0.984 ± 0.022
B_{shift}		$+0.004 \pm 0.031$	0.984 ± 0.022
C_1		$+0.029 \pm 0.040$	1.256 ± 0.028
C_2		$+0.023 \pm 0.036$	1.143 ± 0.026
C_3		-0.039 ± 0.044	1.368 ± 0.031
C_4		-0.006 ± 0.036	1.138 ± 0.026
C_5		$+0.024 \pm 0.044$	1.371 ± 0.031

TABLE 7.4: Fit results for the pull distributions for $B^0 \to J/\psi K^{*0}$ simulated pseudo-experiments. The toy Monte Carlo are generated with the same amount of statistic and signal fraction we would expect from data.

On average we are satisfied with the results obtained. One could just argue about the $cos\psi$ background parameters that show consistently values for the RMS out of 3σs from one. First of all to remove any doubt we have to say the functional shape of the pulls is Gaussian. All the pulls distributions, overlapped with the relative Gaussian fits, are shown from Figure 7.9 to Figure 7.11. To investigate the source of this problem, the background parametrization for $cos\psi$ has been tested in several configurations. Increasing the statistics to $100,000$ events did not remove the issue. So it has been tried to use a different parametrization removing every time one of the parameter among C_3, C_4 or C_5. These three sets of toy MC have pulls with the RMS consistent with one. So this leads us to the final conclusion that the problem

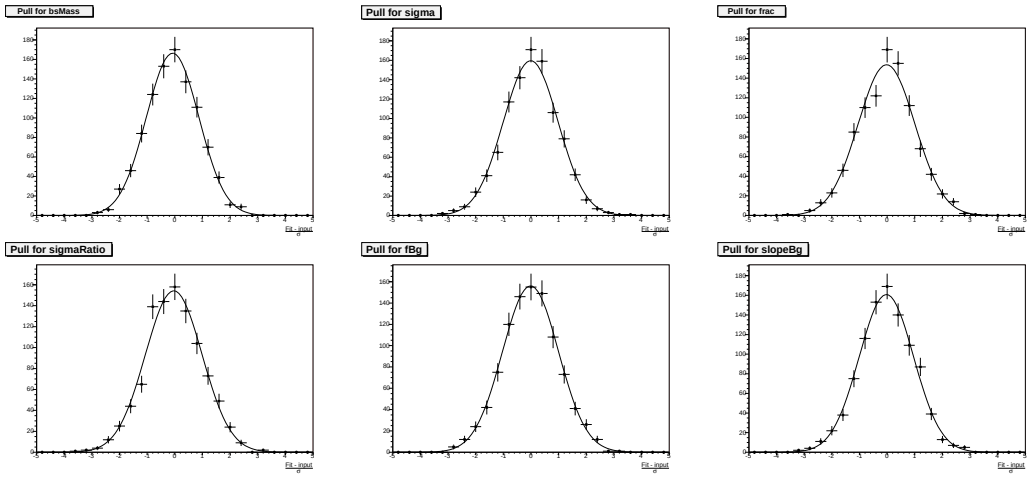

FIGURE 7.9: Pull distributions for the mass parameters for $B^0 \to J/\psi K^{*0}$ decay mode.

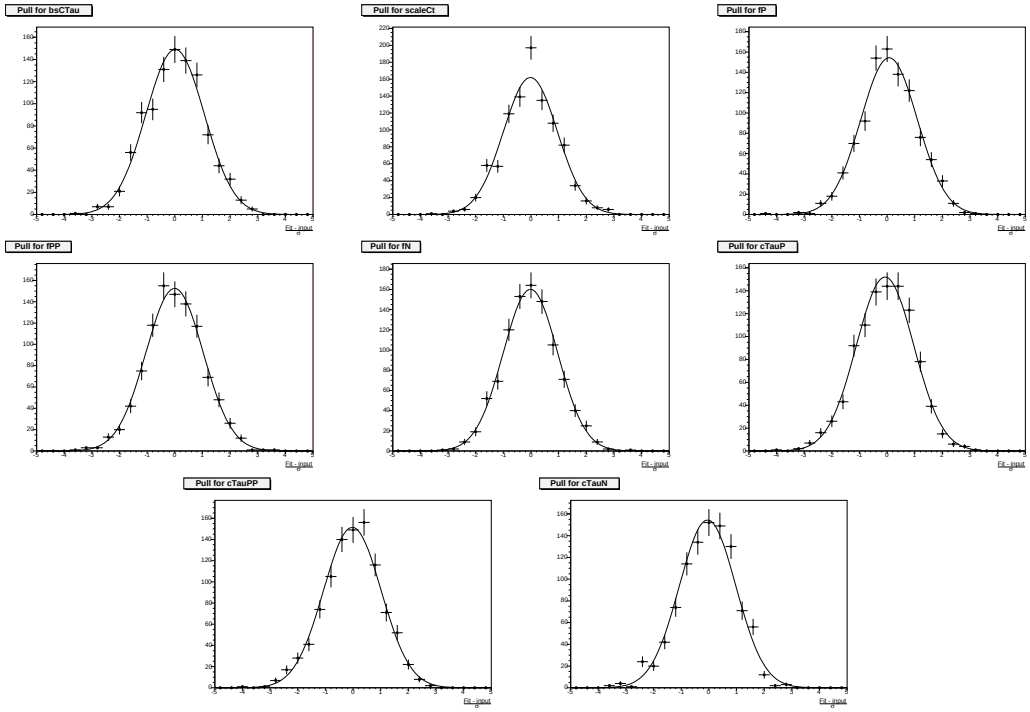

FIGURE 7.10: Pull distributions for the lifetime parameters for $B^0 \to J/\psi K^{*0}$ decay mode.

comes at the generator level. The package we use to generate events have difficulties in reproducing all the features of a five order grade polynomial. If we add the observation that the pulls results showing some troubles are for the background parameters, while the signal parameters of interest for the analysis (the amplitudes $\left|A_{\|}\right|$, $\left|A_{\perp}\right|$ and the strong phases $\delta_{\|} - \delta_0$, $\delta_{\perp} - \delta_0$), which we want to measure, do not have any bias, we can strongly rely on the solidity of the fitter results.

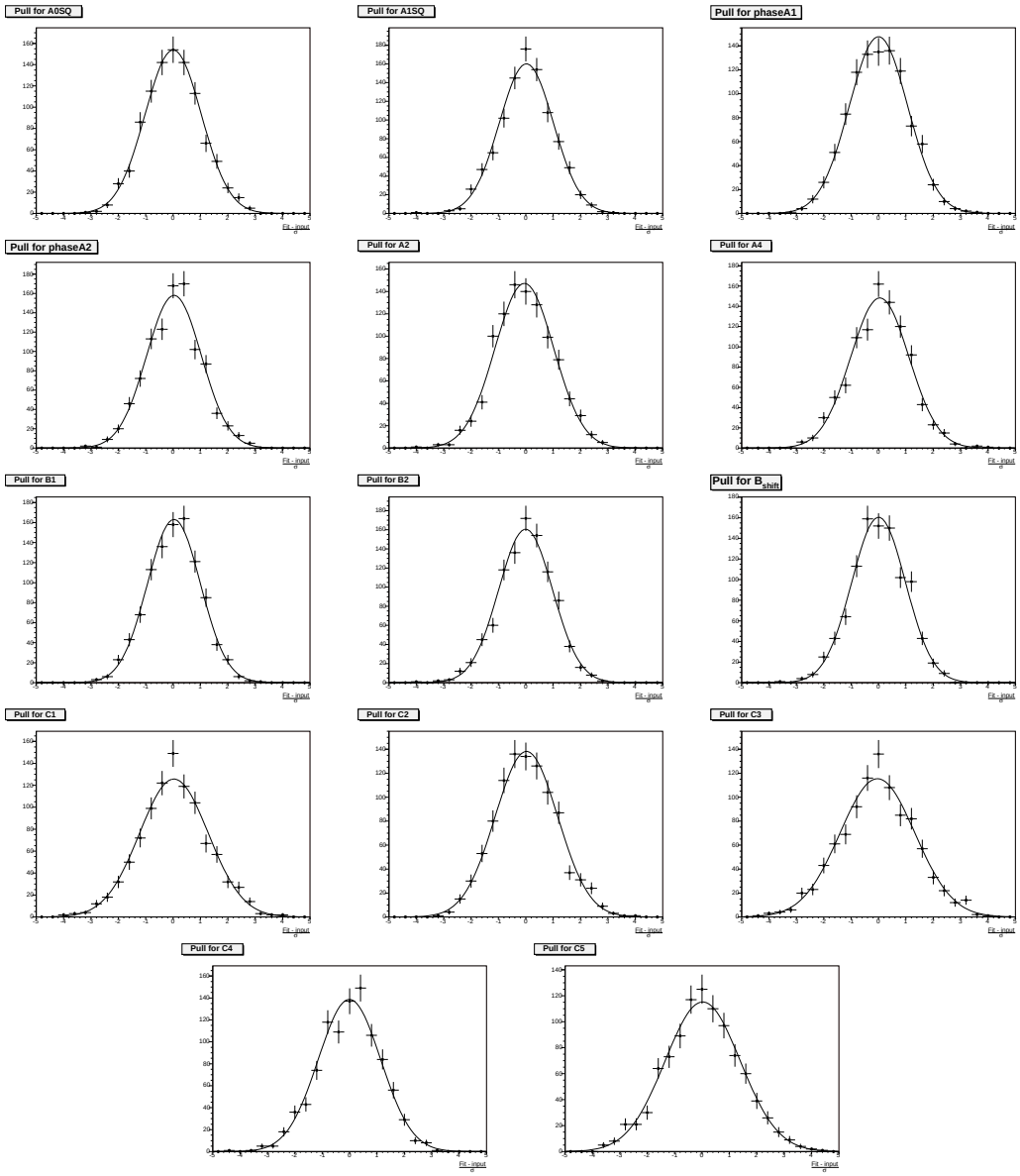

FIGURE 7.11: Pull distributions for the angular parameters for $B^0 \to J/\psi K^{*0}$ decay mode.

7.4 Systematic Uncertainties

Systematic uncertainties estimated from the likelihood fit are meant to cover the effects which may have not been properly incorporated in our modelization and could precisely lead to systematic effects. These uncertainties are mostly evaluated using pseudo-experiments. For a parameter "A", the magnitude of a systematics is evaluated from the difference of the mean A^{syst} from the toy MC with the systematic effect included and the mean A^{ref} of the toy MC when the effect is removed. Only the S-wave contribution and the systematic arising from the proper decay time uncertainty distributions are evaluated directly on data.

Signal Mass Model

The model used for the signal mass PDF is a double Gaussian. The data seems to be well described by it, but it could be possible to have missed another additional component. With the underlying assumption that the unknown third component will contribute less than the absence of the second Gaussian, we produce pseudo experiments with the double Gaussian structure, but fitting allowing only one Gaussian distribution for the mass mmodel. The mass mis-model introduces only small uncertainties on the fit results.

Angular Background Model

We observe a good agreement between the data in the sideband mass regions and the angular background model, but we cannot exactly know if we used too simple or too complex models. Thus, we tested it on pseudo-experiments generated with the complete empirical model used on the data sample. Two fits were performed. One with the complete model, used in the generation, and the second one with a reduced model where some of the parameters were fixed to zero. The difference between the fits defines the estimation of the related systematic uncertainty.

Proper Decay Time Resolution Model

The signal decay time distribution used in the likelihood is an exponential convolved with a Gaussian which accounts for the detector resolution effect (see Section 7.3.3). The negative tail in the proper decay time distribution of the the background indicates that it might be not enough. To estimate the influence of an additional positive and negative long-lived term, we simulate events where we add exponential tails to the resolution function. The fraction and lifetime of the exponential tails for the signal events is taken from the fraction and lifetime of the negative lifetime tail of the background component.

Swapped Mass Assignment

After the dedicated Neural Network selection for the swap removal - Section 7.2.4 - we estimated to 0.5% the residual of the percentage of swapped events. The final fit was performed by treating all the candidates as having the correct mass assignment for the K^{*0} decay products. The effect on this assumption is studied by replacing some of the B^0 events with swapped reconstructed decays when generating the pseudo-experiments. We exchange twice as much of the corrected reconstructed signal events with the swapped candidates. The mass for this candidates is simulated having three time wider width than the narrower signal peak Gaussian. For the lifetime we used the same distributions for the simulated B^0 candidate with the correct mass assignment, while the angular distribution is taken from the simulated B^0 events with swapped $K\pi$ assignment.

SVX Alignment

The position of the silicon layers are known only with a limited precision. Recent studies at CDF [90] estimated the mis-alignment of the SVX detector, concluding that the uncertainty in the lifetime measurements is about 2 μm. We assume to have similar uncertainties, but we have also to evaluate the impact of this mis-alignment on the other parameters. Thus, we generated pseudo-experiments with a lifetime value changed of ± 2 μm from the measured value. As final systematic, we take the largest effect found for the positive and negative shift.

S-wave Contribution

Although the S-wave component is needed to better describe the data, it affects the P-wave only results just slightly. Thus the difference of the fit with and without this effect is considered as a conservative estimate of any additional contribution not yet included in the fitter.

Proper Decay Time Resolution Uncertainty

One of the study performed was to test how important is the modeling of the σ_{ct} distributions (see Figure 7.5) and how it affects the final results. The fits with and without their inclusion was performed. For the $B_s^0 \to J/\psi\phi$ sample in the untagged analysis [74] no significance deviation for the main quantities has been observed. On the other hand for the $B^0 \to J/\psi K^{*0}$ we see an important shift in the final results, as shown in Table 7.5. As we could expect there is a clear shift in the lifetime; what was less easy to foresee was the shift in $|A_0|^2$. A toy Monte Carlo has been performed in order to eventually reproduce this effect. The toy MC has been generated starting from all the PDFs, including the σ_{ct} distribution terms, with the additional feature to generate two different background angular shapes for the prompt and not prompt component respectively. The simulated sample has been fitted twice. Once including the σ_{ct} distributions we get from data and the second time without their inclusion. The toy MC by itself does not describe the shift we see in data: the results compared with the one obtained from data are in Table 7.6. We know from theory [73] that discarding the σ_{ct} distributions is not the correct way to proceed and the shift observed without them is already well covered inside the statistical plus already evaluated systematic uncertainties (summed in quadrature). We anyway decide not to use the result from the toy MC and assign a more conservative systematic using half of the shift we see in data.

Summary

Table 7.7 summarizes the value estimated for each systematics effect. The total systematic uncertainties are obtained by summing in quadrature the individual contributions. While the error of the lifetime measurement is mostly influenced by the proper SVX alignment which, indeed, determines the proper decay length measurement, the results on the transversity amplitudes and phases are mostly affected by

Parameter	With σ_{ct} dist.	Without σ_{ct} dist.
$c\tau$ $[\mu m]$	$+455 \pm 6$	$+452 \pm 6$
$\|A_0\|^2$	$+0.569 \pm 0.009$	$+0.559 \pm 0.009$
$\left\|A_\parallel\right\|^2$	$+0.211 \pm 0.012$	$+0.215 \pm 0.012$
$\delta_\parallel - \delta_0$	-2.956 ± 0.083	-2.945 ± 0.083
$\delta_\parallel - \delta_0$	$+2.971 \pm 0.056$	$+2.965 \pm 0.056$

TABLE 7.5: Fit results for $B^0 \to J/\psi K^{*0}$ using the NeuroBayes Neural Network selection and including the S-wave interference for the two cases: with and without the σ_{ct} distribution included in the fitter.

Parameter	Data Shift	Toy MC
$c\tau$ $[\mu m]$	1	2
$\|A_0\|^2$	0.005	$---$
$\left\|A_\parallel\right\|^2$	0.002	$---$
$\delta_\parallel - \delta_0$	0.006	$---$
$\delta_\perp - \delta_0$	0.003	0.001

TABLE 7.6: Systematic uncertainty as obtained from the studies of data with respect to the toy Monte Carlo on the σ_{ct} distribution terms.

Systematic Source	$c\tau[\mu m]$	A_0^2	$A_\parallel^2$	$\delta_\parallel - \delta_0$	$\delta_\perp - \delta_0$
Signal mass model	2	0.001	$---$	0.001	0.003
ct resolution model	3	$---$	0.002	0.005	0.002
K/π swap	2	0.001	0.001	0.001	0.003
SVX alignment	4	$---$	$---$	0.001	0.001
Angular bkg model	2	0.005	0.004	$---$	$---$
only P-wave	1	0.006	0.004	0.031	0.002
ct uncertainty dist.	1	0.005	0.002	0.006	0.003
Total	6	0.009	0.006	0.032	0.006

TABLE 7.7: Final systematic Table for the $B^0 \to J/\psi K^{*0}$ transversity analysis.

the correct modelization of the signal angular distribution, i.e. the difference between the presence or absence of the S-wave interference.

7.5 Final Results

Table 7.8 reports the results of the fit when using the P-wave only modelization for the signal angular distributions. Since the two fold ambiguity has been solved by BaBar Collaboration, we quote only the results with a physical meaning for the strong phases. The angular distributions are shown in Figure 7.12 and they refer to the distribution for the signal when removing the detector effect. A forward-backward asymmetry which is not reproduced in our fit but is present in the data points is clearly visible comparing the $\cos\psi$ distributions for a pure P-wave. This

Parameter		fit value		error
M(mass)	$[GeV]$	$+5.2791$	$\pm$	0.0002
σ_1	$[GeV]$	$+0.0102$	$\pm$	0.0001
f_m		$+0.143$	$\pm$	0.010
R$(=\sigma_1/\sigma_2)$		$+5.82$	$\pm$	0.24
f_{Bkg}		$+0.8207$	$\pm$	0.0025
$A(= slope_{Bkg})$		-0.86	$\pm$	0.20
$c\tau$	$[\mu m]$	$+455$	$\pm$	6
S_{ct}		$+1.2924$	$\pm$	0.0075
f_P		$+0.1015$	$\pm$	0.0049
f_{PP}		$+0.0412$	$\pm$	0.0039
f_N		$+0.0194$	$\pm$	0.0012
$c\tau_P$	$[\mu m]$	$+65$	$\pm$	5
$c\tau_{PP}$	$[\mu m]$	$+355$	$\pm$	24
$c\tau_N$	$[\mu m]$	$+169$	$\pm$	9
$\lvert A_0\rvert^2$		$+0.574$	$\pm$	0.009
$\lvert A_\parallel\rvert^2$		$+0.207$	$\pm$	0.011
$\delta_\parallel$		-2.987	$\pm$	0.086
$\delta_\perp$		$+2.969$	$\pm$	0.057
A_2		-0.660	$\pm$	0.045
A_4		$+0.234$	$\pm$	0.052
B_1		$+0.353$	$\pm$	0.009
B_2		$+0.139$	$\pm$	0.017
B_{shift}		$+0.237$	$\pm$	0.022
C_1		$+0.788$	$\pm$	0.034
C_2		$+0.529$	$\pm$	0.059
C_3		-0.15	$\pm$	0.12
C_4		-0.806	$\pm$	0.065
C_5		-1.19	$\pm$	0.10

TABLE 7.8: Fit results for $B^0 \to J/\psi K^{*0}$ using the NeuroBayes Neural Network selection with the P-wave only included in the signal angular component.

is due to S–P interference missing description. In fact, if we repeat the fit including the S-wave contribution (Section 7.3.4), the agreement between the fit and the data for the angular distributions is largely improved. The results are summarized in Table 7.9 and the relative angular distributions are shown in Figure 7.13. Figure 7.14 shows the lifetime and angular distributions containing both signal and background contributions, as well as the detector sculpting effects for the angular projections.

Likelihood Scans

The results we quote for the $B^0 \to J/\psi K^{*0}$ transversity analysis are given in terms of point estimations of the likelihood fit. Therefore the likelihood parabolic behaviour for the relevant parameters has to be checked. The likelihood scans for lifetime,

Parameter		fit value		error
mass		$+5.2791$	$\pm$	0.0002
σ_1		$+0.0102$	$\pm$	0.0001
frac	$[GeV]$	$+0.135$	$\pm$	0.010
$R(=\sigma_1/\sigma_2)$	$[GeV]$	$+5.67$	$\pm$	0.26
f_{Bkg}		$+0.8224$	$\pm$	0.0025
$A(=slope_{bkg})$		-0.86	$\pm$	0.20
$c\tau$	$[\mu m]$	$+456$	$\pm$	6
S_{ct}		$+1.2922$	$\pm$	0.0075
f_P		$+0.1016$	$\pm$	0.0048
f_{PP}		$+0.0432$	$\pm$	0.0039
f_N		$+0.0194$	$\pm$	0.0012
$c\tau_P$	$[\mu m]$	$+65$	$\pm$	5
$c\tau_{PP}$	$[\mu m]$	$+359$	$\pm$	23
$c\tau_N$	$[\mu m]$	$+170$	$\pm$	9
$\lvert A_0\rvert^2$		$+0.569$	$\pm$	0.009
$\lvert A_\parallel\rvert^2$		$+0.211$	$\pm$	0.012
$\delta_\parallel - \delta_0$		-2.956	$\pm$	0.083
$\delta_\perp - \delta_0$		$+2.971$	$\pm$	0.056
$\cos\lambda$		$+0.995$	$\pm$	0.003
δ_S		$+2.29$	$\pm$	0.26
A_2		-0.660	$\pm$	0.045
A_4		$+0.235$	$\pm$	0.051
B_1		$+0.352$	$\pm$	0.009
B_2		$+0.138$	$\pm$	0.017
B_{shift}		$+0.235$	$\pm$	0.022
C_1		$+0.797$	$\pm$	0.034
C_2		$+0.536$	$\pm$	0.059
C_3		-0.16	$\pm$	0.12
C_4		-0.812	$\pm$	0.065
C_5		-1.19	$\pm$	0.10

TABLE 7.9: Fit results for $B^0 \to J/\psi K^{*0}$ using the NeuroBayes Neural Network selection with the P- and S-wave included in the signal angular component.

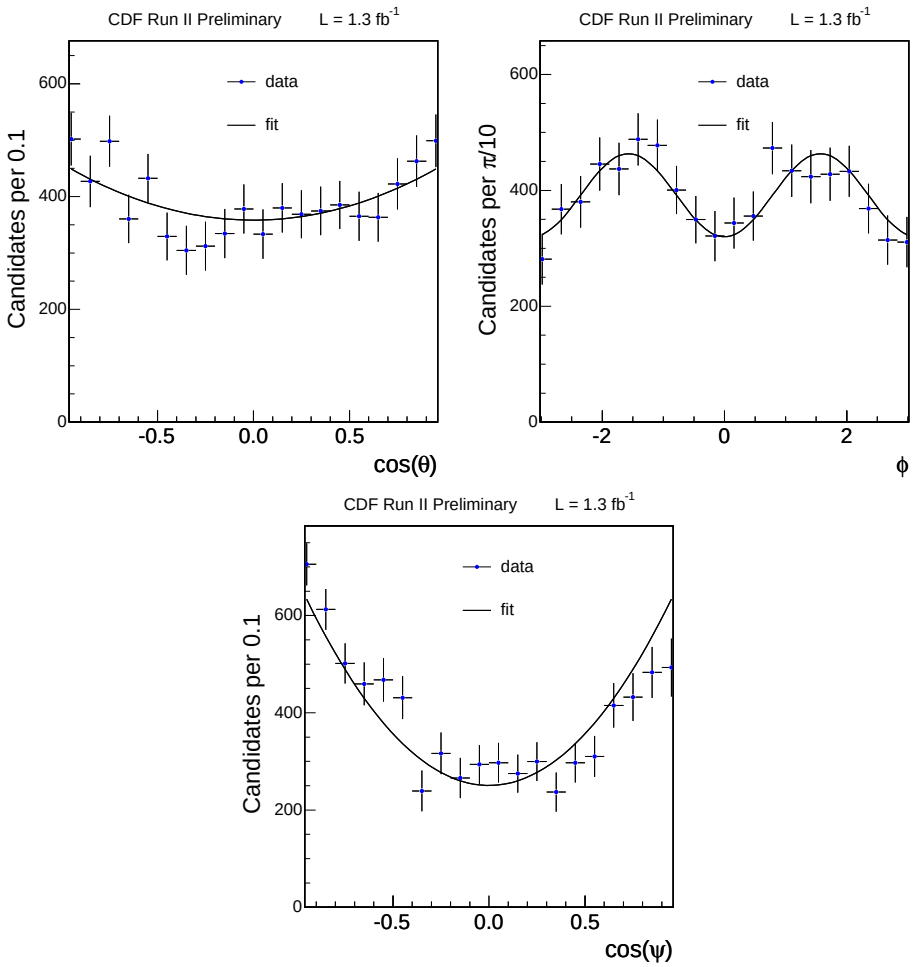

FIGURE 7.12: Projections for the three angular distributions removing the detector sculpting effect. The black line is the theoretical curve.

amplitudes and strong phases are reported in Figure 7.15, showing the expected parabolic behaviour mandatory to be able to have a reliable point estimate.

7.6 Conclusions

We measured the decay amplitudes for the $B^0 \to J/\psi K^{*0}$ flavor eigenmodes and observe the strong-phase differences. We used the interference between the K^{*0} S-wave and P-wave amplitudes in the region of the $K^{*0}(892)$ to properly describe data. Figure 7.16 shows how the addition of the S-wave component is necessary to account for the backward-forward asymmetry in the angular variable $\cos(\psi_T)$ observed in data. Then, as final result, we quote the amplitudes and phases obtained with the likelihood which includes the S-P interference. The values as estimated from the fit are summarized in Table 7.10. By analyzing the table content, we observe how the level of statistic and systematic accuracy reached at CDF is not only comparable but competitive with the most recent analyses performed at the B factories[87, 88]. In

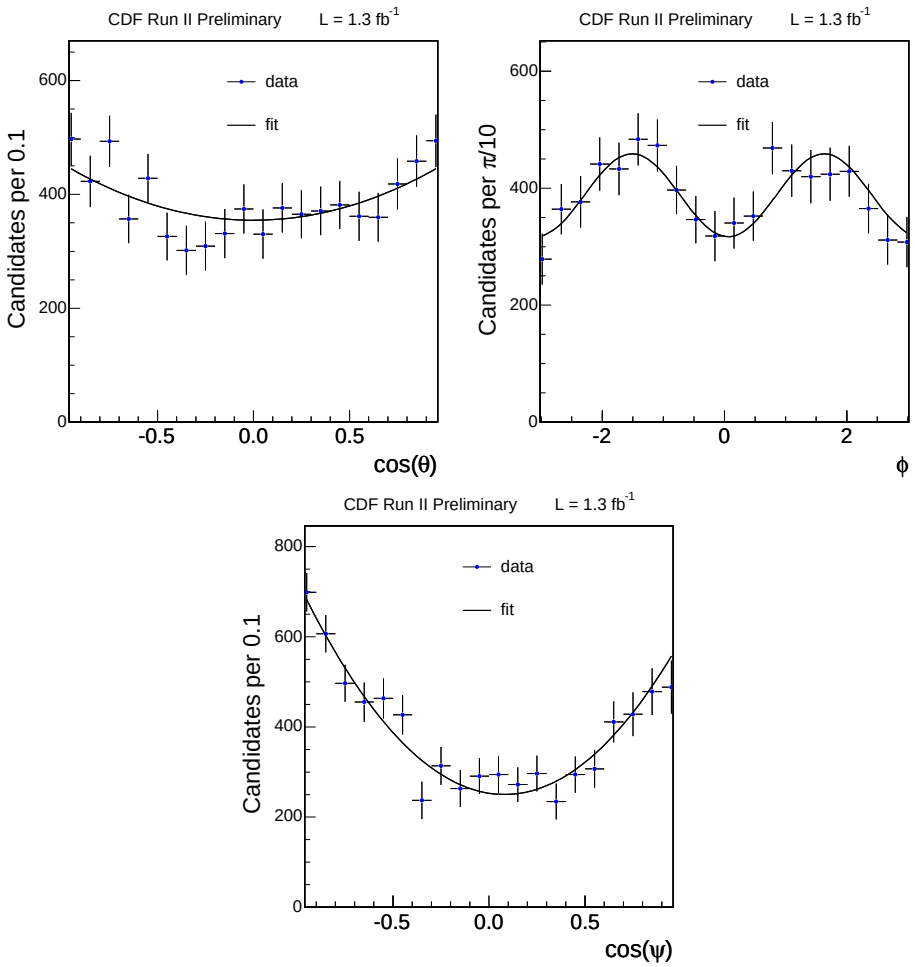

FIGURE 7.13: Projections for the three angular distributions removing the detector sculpting. The black line is the theoretical curve.

Table 7.10 the CDF analysis is, in fact, compared with the results from Babar and Belle. This represent a remarkable result by itself, especially if we account for the high sources of background and the busy environment tipical of an hadronic collider, such as Tevatron, as compared to the B factories analyses. In conclusion, all three experiments have compatible results among each other which proves successful the check of our framework on the kinematically equivalent data sample $B^0 \to J/\psi K^{*0}$: these results contribute to enforce the reliability of the angular analysis results obtained on the $B_s^0 \to J/\psi\phi$ decay mode sample and discussed in detail in Chapter **??**.

We presented the results on $B^0 \to J/\psi K^{*0}$ *decays using* 1.35 fb^{-1} *of data collected with the CDF run II detector. We measured the transversity amplitudes and strong phases of the decay providing a result compatible and competitive with the most recent ones from the B factories. This validate completely the* $B_s^0 \to J/\psi\phi$ *angular analysis which is one of the central topic of this dissertation.*

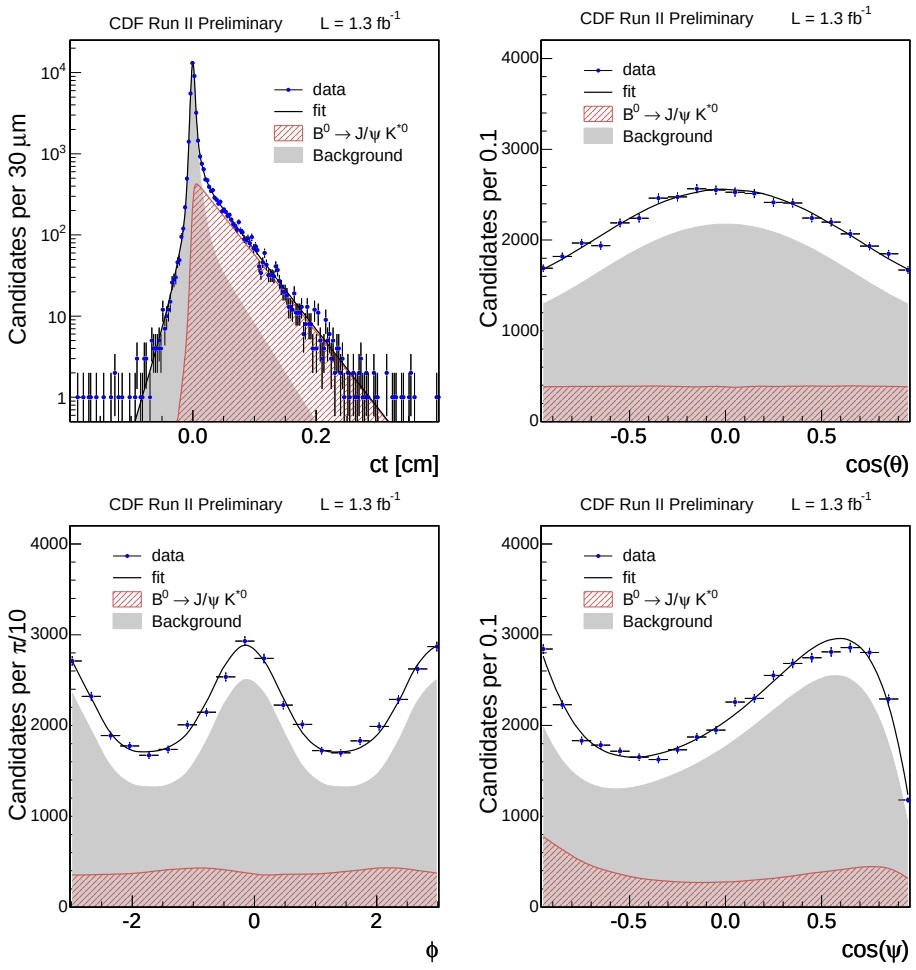

FIGURE 7.14: Projections for the proper decay lenght and for the angular distributions including the detector sculpting using the NeuroBayes Neural Network selections with the S wave interference in the likelihood.

Par.	CDF	BaBar [87]	Belle [88]
$c\tau$	$456 \pm 6 \pm 6 \mu m$	$- - -$	$- - -$
$\|A_0\|^2$	$+0.569 \pm 0.009 \pm 0.009$	$+0.556 \pm 0.009 \pm 0.010$	$+0.574 \pm 0.012 \pm 0.009$
$\|A_\parallel\|^2$	$+0.211 \pm 0.012 \pm 0.006$	$+0.211 \pm 0.010 \pm 0.006$	$+0.231 \pm 0.012 \pm 0.008$
$\delta_\parallel - \delta_0$	$-2.96 \pm 0.08 \pm 0.03$	$-0.293 \pm 0.08 \pm 0.04$	$+0.2887 \pm 0.090 \pm 0.008$
$\delta_\perp - \delta_0$	$+2.97 \pm 0.06 \pm 0.01$	$-0.291 \pm 0.05 \pm 0.03$	$+0.2938 \pm 0.064 \pm 0.010$

TABLE 7.10: Final results for the $B^0 \to J/\psi K^{*0}$ angular analysis performed at CDF. The first uncertainty is statistical, while the second one corresponds to the systematic one. For comparison, in the third and forth column the most recent B factories results on the measurement of the amplitudes and strong phases are quoted.

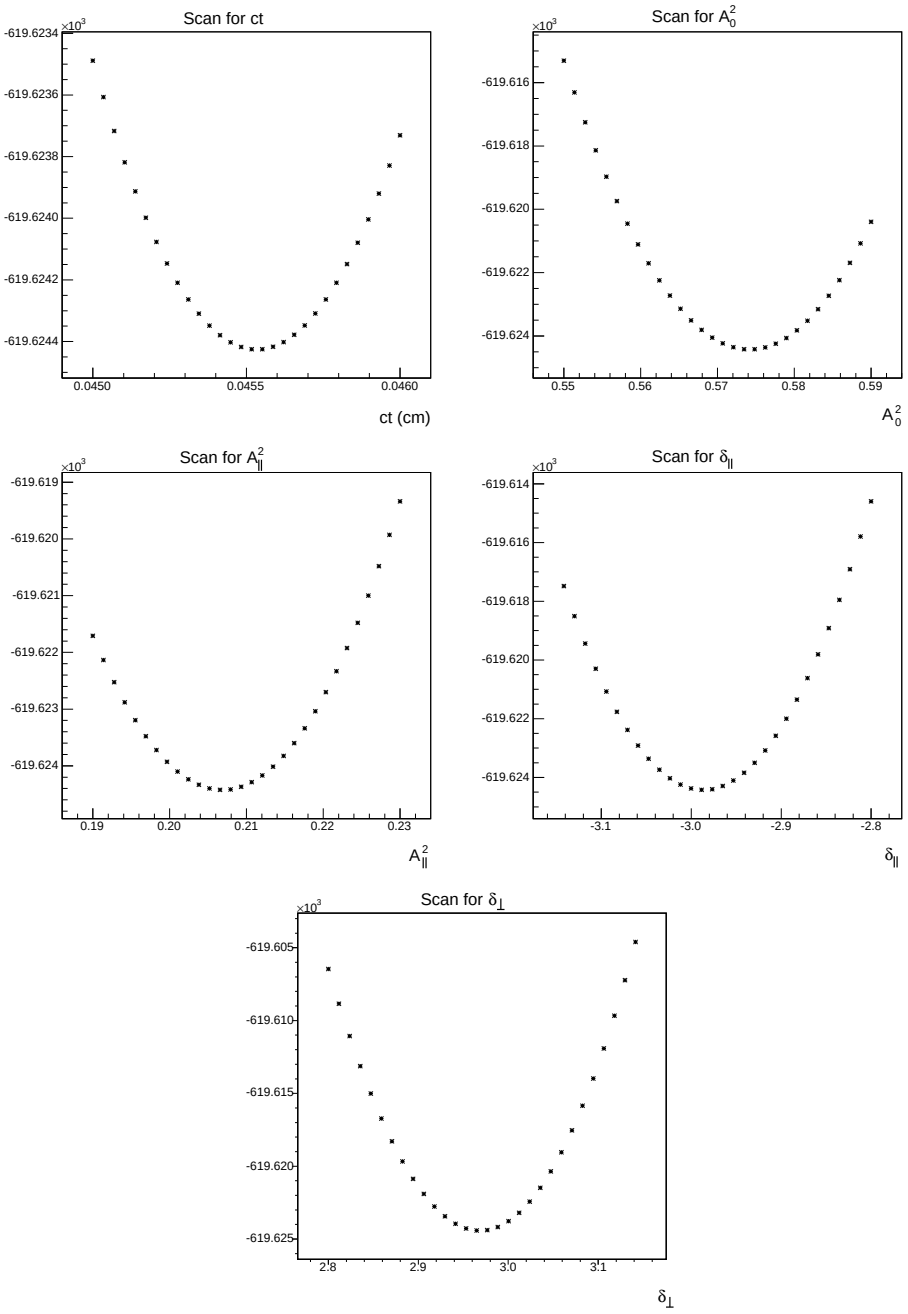

FIGURE 7.15: Likelihood scans for ct, $|A_0|^2$, $\left|A_\parallel\right|^2$, $\delta_\parallel$ and $\delta_\perp$, respectively.

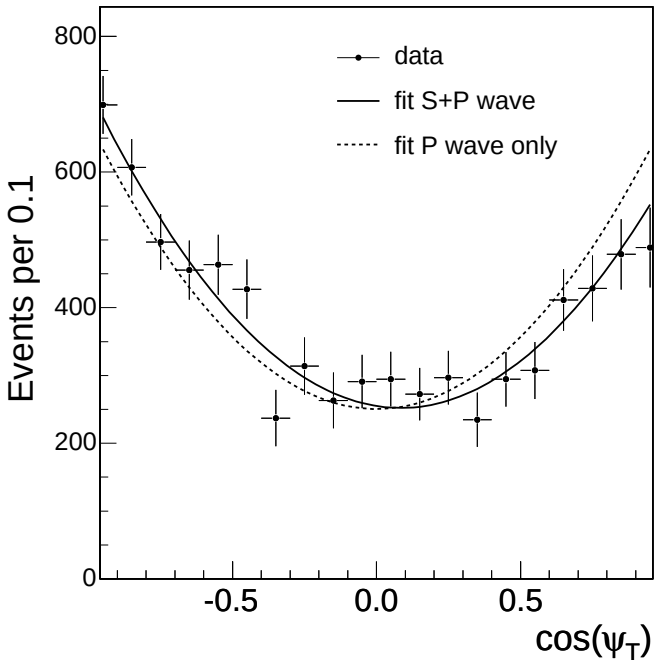

FIGURE 7.16: Angular distributions with fit projection overlaid for $\cos(\psi_T)$. Dashed line: projection for the nominal fit with the S–P interference included in the fitter. Continuous line: projection for the fit with only P-wave included.

Conclusions And Perspectives

Conclusions

Along this dissertation we discussed two analyses focusing on the studies of the B_s^0 neutral mesons system properties. The B_s^0 physics is, in fact, a rich branch of the heavy flavor physics which offers several possibilities to test physics beyond the Standard Model.

We devoted part of the thesis to the first observation of the time-dependent B_s^0 flavor oscillations. The $B_s^0 - \overline{B}_s^0$ oscillation frequency measurement was achieved through a combination of several datasets of 1 fb^{-1} in integrated luminosity, which resulted in:

$$\Delta m_s = 17.77 \pm 0.10(\text{stat.}) \ \pm 0.07(\text{syst.}) \ \ \text{ps}^{-1}. \tag{7.6.1}$$

with a significance superior to 5 standard deviations. The definitive observation of the B_s^0 mixing frequency also provided an handle to measure with high precision CKM parameters. In particular, it was possible to extract the ratio $|V_{ts}|^2/|V_{td}|^2$:

$$\frac{|V_{ts}|^2}{|V_{td}|^2} = 0.2060 \pm 0.0007(\text{exp.}) \ ^{+0.0081}_{-0.0060}(\text{theory}) \ . \tag{7.6.2}$$

where the first uncertainty, referred as "exp", is only related to the Δm_s measurement performed at CDF, and the second includes all the other sources, which are dominated by theoretical uncertainty.

The New Physics contributions, if present in the magnitude of the mixing amplitude, are small. Nevertheless, New Physics could also modify the phase of the mixing amplitude, β_s, by contributing with additional processes in the electro-weak loop diagrams. The global fits on experimental data constrain the CP violation phase to a small value $2\beta_s \simeq 0.04$ rad [75]: New Physics could then appear in a significantly large deviation from the Standard Model tiny expected value. After the Δm_s measurement, the natural step, in the B_s^0 system studies, is to measure the phase β_s. Thus, the second part of my thesis was focused on the angular analysis of $B_s^0 \to J/\psi\phi$ decays for the world first measurement of β_s using flavor tagging information. In this analysis, the non-Gaussian likelihood uncertainties and biases, observed in simulated experiments with the available statistics, did not allow a reliable point estimate of the parameters of interest. Therefore, we presented a confidence bounds on the CP-violation parameter β_s and the width difference $\Delta\Gamma_s$ obtained using the Feldman-Cousins according to the Feldman Cousins criterion with rigorous inclusion of systematics uncertainties.

In fact, any $\Delta\Gamma_s - \beta_s$ pair was excluded at a given CL only if it could be excluded for any choice of *all* other fit parameters, sampled uniformly within ± 5 σ of the values determined in their estimate on data. Assuming the Standard Model predicted values of $\beta_s = 0.02$ and $\Delta\Gamma_s = 0.096$ ps^{-1} [75], the probability of a deviation as large as the observed data is 15%, which corresponds to 1.5 Gaussian standard deviations.

Perspectives

The precise measurement of the oscillation frequency Δm_s allows to stringently test the the CKM matrix unitarity assumption. As shown in the determination of the ratio $|V_{ts}|^2/|V_{td}|^2$ (Equation 7.6.2), this test is only limited by theoretical calculations. Although the analysis could be pursued adding more data or refining the technical tools, it will not dramatically improve the knowledge we have already aquired. At least, until the moment we will obtain more precise lattice QCD calculations.

On the other hand, the angular analysis on $B_s^0 \to J/\psi\phi$ is at its first iteration. As already observed, the results shows an interesting fluctuation pointing to a promising hint for the presence of New Physics beyond the Standard Model and, thus, it needs further investigations. The interest lies also in the observation that other independent measurements, such as the tagged angular analysis on $B_s^0 \to J/\psi\phi$ decay from DØ collaboration [83] or asymmetry measurements in the semileptonic B_s^0 decays [91, 92, 93] point consistently in the same direction. Additionally, the Belle Collaboration recently reported a measurement in the difference of the direct charge-parity violation between charged and neutral B meson decays [94]. It was established a 4.4 σ level discrepancy from zero in the difference between the direct CP violation asymmetry measurements in $B^+ \to K^+\pi^0$ and $B^0 \to K^+\pi^-$. The two asymmetries were expected to be rather close and their difference almost zero. This anomaly may originate from the same $b \to s$ transition involved in the $B_s^0 \to J/\psi\phi$ decay mode analyzed along this thesis. Again, this result underlines the importance of pursuing this analysis with more data.

The $B_s^0 \to J/\psi\phi$ angular analysis presented in this dissertation exploits the data collected with the CDF II detector up to an integrated luminosity of 1.35 fb^{-1}. So far, the Tevatron accelerator registered excellent performances and collected already 3 fb^{-1} of data. Although its running has been extended till the end of the fiscal year 2009, the whole collaboration is intended to prolong the Tevatron working of an additional year, to conclude its performances in 2010. Most of the undergoing searches at Tevatron would certainly profit from this additional amount of data. According to the projections, CDF could record, at the end of its last data taking run, an integrated luminosity spanning from 6 to 8 fb^{-1}. By simply including more data, the analysis will profit from 4 times the actual amount of data, in the most conservative scenario. Therefore, the angular analysis on $B_s^0 \to J/\psi\phi$ decay, while being a milestone, is far from its last conclusion.

In our realistic Monte Carlo studies we observed the presence of local invariances when using 6 times the signal yield of the current data (see Section 5.7.4). Therefore,

the likelihood could keep manifesting the same non-Gaussian behaviour in the future: the understanding of its response has been performed by observing how the likelihood profiles change as we increase the dataset available. In Figure 7.17 we report the likelihood profiles for two scenarios: one corresponding to 5 fb^{-1} and another one, more "stretched", corresponding to 10 fb^{-1} of data available. The Monte Carlo toys assume the analysis to have the same performances in terms of yield per fb^{-1}, tagging dilutions and tagging efficiencies. The input value for the CP violating phase is set to $\beta_s = 0.4$ rad which is the average of the interval obtained from the one-dimensional Feldman-Cousins contour if all the constraints are applied ($\beta_s \in [0.20,\ 0.60]$ rad, see Section 6.2.1). To better visualize the improvements, the contours are drawn for the $1, 3, 5\ \sigma$(s) error. We observe that, if these first indications are confirmed in the future, CDF could support an evidence, which would be possible exploiting the full Run II data sample. Following the same trend, in Figure B.15 we show the probabilities to observe a certain number of sigma effect $(3, 4, 5)$ as a function of the luminosity, with the underlying assumption of β_s hypothetically large (0.4 rad). The same data taking efficiency and no improvements in the analysis are assumed.

So far, we do not have considered possible improvements of the analysis but just the data addition, where the only limitations come from the necessity to re-calibrate the detector performances. An example over all is the $d\text{E}/d\text{x}$ calibration from the Central Outer Tracker system. In fact, after the integrated luminosity of 1.35 fb^{-1} the first two superlayer of the COT (see Chapter 2) have been turned off, not allowing the direct use of the SSKT in its present implementation for further data. This was the reason which limited the use of less than half the data available at the time the analysis was performed.

In the future, another possible way to proceed would be to refine the tagging algorithms, in order to use a unique tagger which combines the OST and SST information in a Neural Network fashion. This could provide better performances in terms of efficiency and dilution as well as account for the correlations among the two classes algorithms (OST and SST). Moreover, with the Δm_s precisely measured it would be desirable to calibrate the SSKT directly on data using, for instance, the B_s^0 hadronic decays which provide the highest statistical power in the mixing analysis. During the discussion of the final $\text{B}_s^0 \to J/\psi\phi$ results we observed that the tagging major impact on the analysis was not to reduce drastically the uncertainties for the parameters interest with respect to the untagged analysis, but rather to reduce the likelihood ambiguities. Thus, a more effective action would be to exploit other trigger paths to get additional signal data, necessary to increase the probability of observing significant effects in the β_s measurement (e.g. bottom curves in Figure B.15). An idea could be to utilize the `TTT` trigger paths already used for the mixing analysis: in Table B.5 we show an example of the $\text{B}_s^0 \to J/\psi\phi$ invariant mass distribution reconstructed with event selected with the `TTT`. We used the data collected with the `TTT` paths for a total integrated luminosity of 640 pb^{-1} obtained summing up the two data taking periods `xbhd0h` ($\simeq$ 410 pb^{-1}) and `xbhd0i` ($\simeq$ 230 pb^{-1}). The $\text{B}_s^0 \to J/\psi\phi$ sample has been collected applying a simple sequential selection, reported in Table B.5, and removing the events which fires the `di-muon` trigger paths as well. The yield obtained is of 296 ± 15 signal events and it would already ensure, at this point, an

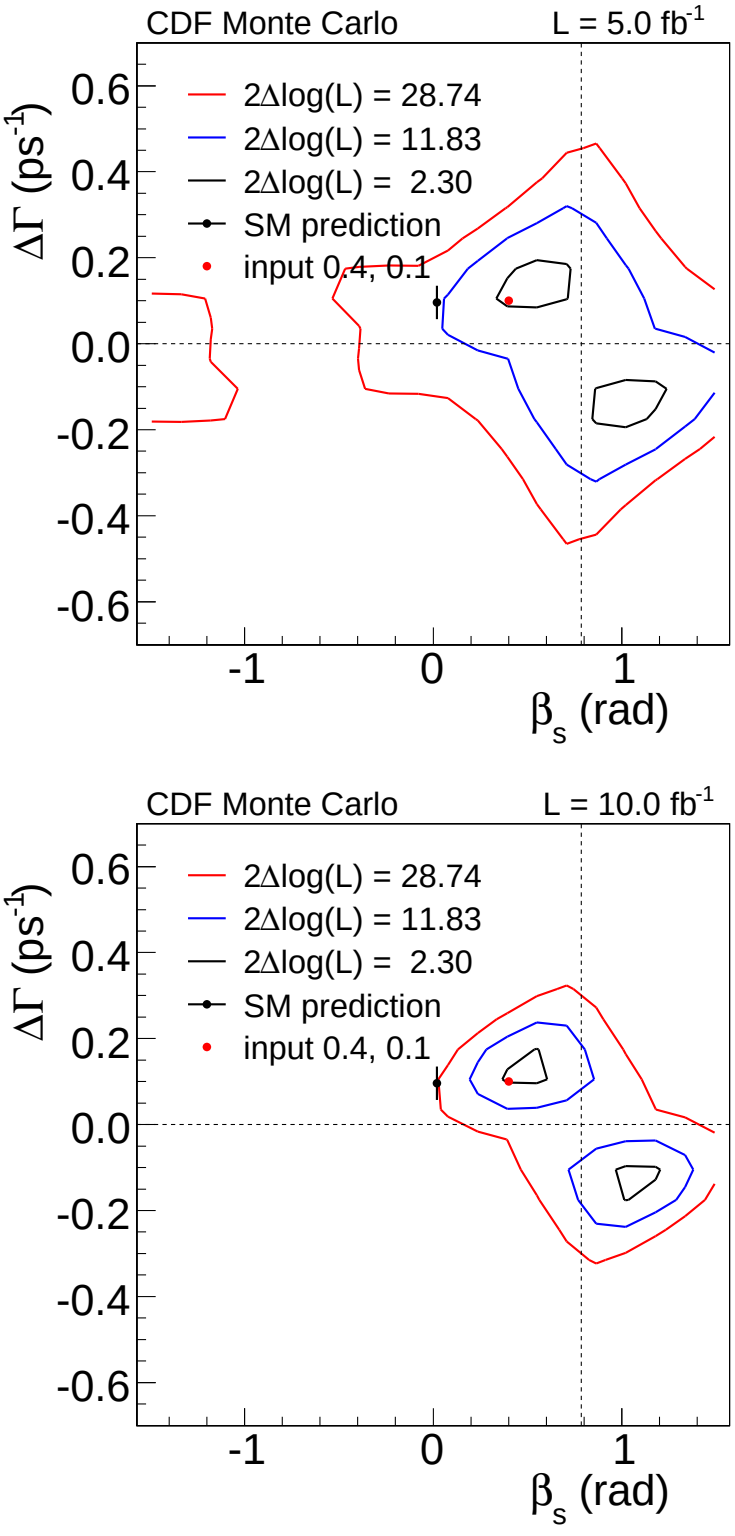

FIGURE 7.17: Likelihood count-ours projections for 5 and10 fb^{-1}. The pseudo-experiments are simulated with $\beta_s = 0.4$ and $\Delta\Gamma_s = 0.1$. From the most internal to the most external, the count-ours drawn correspond to $1, 3, 5\ \sigma$(s) error.

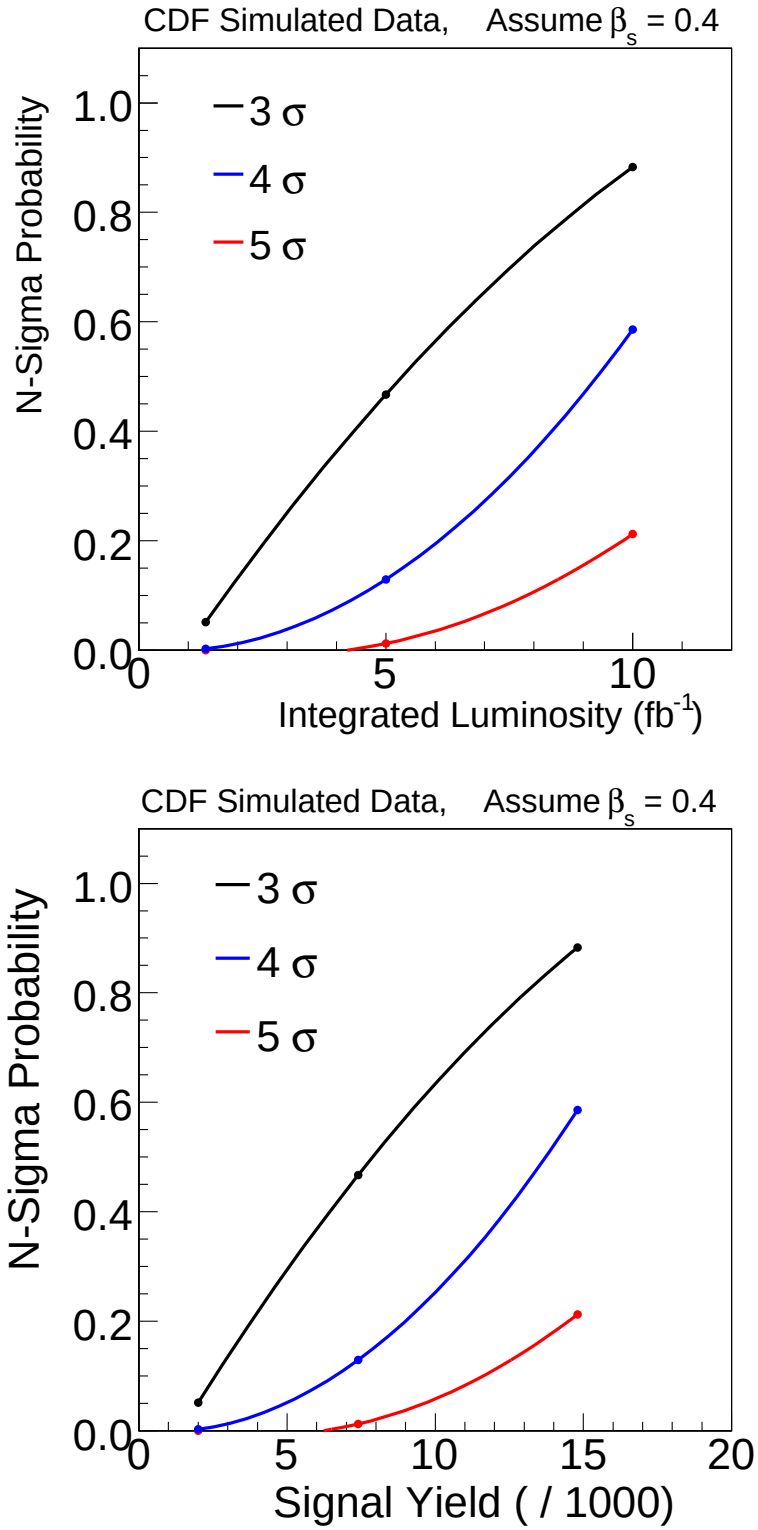

FIGURE 7.18: Probabilities to observe a $3, 4, 5\ \sigma$(s) effect, if $\beta_s = 0.4$ rad, as a function of the luminosity (top) and as a function of the yield of data (bottom). The N-sigma probability curves assumes the same data taking efficiency a no improvements in the analysis.

additional $\simeq 30\%$ of signal independent events. The most interesting aspect of these events is their reduced background with respect to the candidates reconstructed with the `di-muon` trigger paths. Moreover the selection could be improved with a Neural Network approach. The drawback lies in the lifetime bias introduced by the trigger online selections. This indeed would complicate a time-dependent analysis such this and the main challenge would be to properly account the bias (e.g. efficiency curves from MC). Nevertheless, this would offer a non-negligible way to include more data in the analysis finalization.

Variable		Cut Value
μ_{nSiPhi}		$\geq$ 3
p_T^{B}	[GeV/c]	$>$ 5.0
$L_{xy}/\sigma_{L_{xy}}(\mathrm{B}_s^0)$		$>$ 4.9
$d_0(\mathrm{B}_s^0)$	cm	$<$ 0.007
$p_T(\phi)$	[GeV/c]	$>$ 1.4
$\chi^2_{r\phi}(\mathrm{B}_s^0)$		$<$ 23
$p_T(\mu)$	[GeV/c]	$>$ 0.450
σ_{ct}	[μm]	$<$ 400
$\lvert m(\mu^+\mu^-) - M^{J/\psi}_{PDG}\rvert$	[MeV/c^2]	$<$ 80
$\lvert m(K^+K^-) - M^{\phi}_{PDG}\rvert$	[MeV/c^2]	$<$ 14

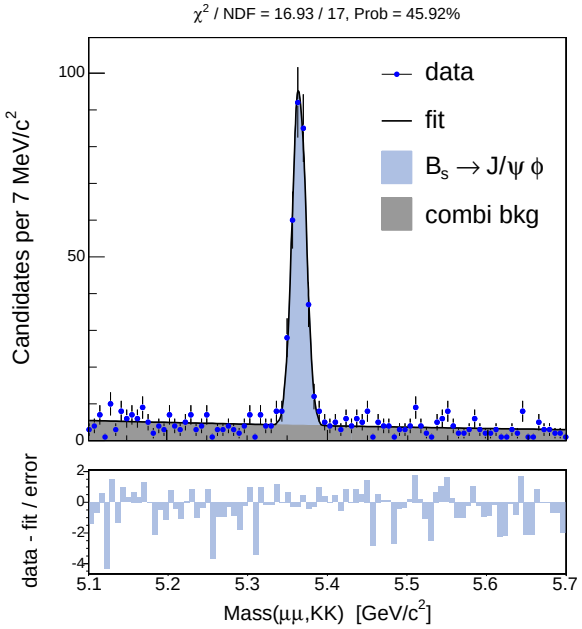

TABLE 7.11: Left: Sequential selection criteria to reconstrut $\mathrm{B}_s^0 \to J/\psi\phi$ events with the `TTT`. Right: B_s^0 invariant mass distribution for events reconstructed with `TTT`on $\simeq 640$ pb^{-1} of data. The events which fire the `TTT` and `di-muon` triggers are removed. The signal yield obtained is 296 ± 15.

By analyzing the curves in Figure B.15 and considering possible improvements, we realize that exciting times await the B_s^0 physics at Tevatron, especially considering a desirable synergy between both CDF and DØ Collaborations. All the scientific community hopes in the future a common strategy in quoting the final result, to have a consistent combination without any external constraint. In conclusion, the Tevatron experiments offer a great opportunity to study the B_s^0 flavor sector, crucial in the perspective of the LHCb experiment starting at CERN Large Hadronic Collider (according to the projections, LHC will allow LHCb to reach a resolution of $\sigma(\beta_s) = 0.023$ with the first 2 fb^{-1} recordable in a year of data taking).

Appendix A

Systematic For The $\mathrm{B}_s^0 \to J/\psi\phi$ Angular Analysis

Several sources of systematic uncertainties for the tagged measurement have been addressed in case it would be possible to quote a point estimate. While this is not the present situation, we leave this discussion here in order to give the reader a sense of the scale of the systematic uncertainties under those conditions. The following is a list of systematic uncertainties that we considered:

Tagger dilution scale factor

Both the same side and opposite taggers are calibrated to within a certain precision, which is encapsuled in the uncertainty on the dilution scale factor. In our final fit, we let the dilution scale factor float with a Gaussian constraint. In such way, any systematic effect due to imperfect knowledge of the scale factors is already included in the final errors assigned to the fit parameters from the fit minimization.

Decay time resolution

In Section 4.4.1 we present a detailed study of the embedded dependencies of decay time resolution on different decay variables. We expect that the remaining uncalibrated effects could change the decay time scale factor by less than $\pm 4\%$. In our final fit, we calibrate the decay time scale factor by floating it in the fit. The propagation of the statistical uncertainty on the decay time scale factor is then a conservative estimate of the remaining uncertainty due to resolution effects.

Alignment and uncertainty on Δm_s

We utilize a Gaussian constraint on the B_s^0 oscillation frequency. The width of the constraint corresponds to a combined statistical plus systematic uncertainty on the

oscillation frequency. We allow the fitter to propagate this information into the output paramters. We do not assign a separate systematic uncertainty via a Monte Carlo study so as not to double-count the uncertainty (the alignment uncertainty is the dominant uncertainty on the measurement of the B_s^0 oscillation frequency).

Background Polarization

We evaluate the systematic uncertainty using toy Monte Carlo experiments in which the events are generated with modified angular distributions and fit them with the default angular distributions. We assume that the influence of any missed order of the polynomial is equal or smaller than the last one we used. Therefore we simulate data with one order higher for each angle than the one we use for fitting our data.

Signal Mass Model

For the Mass model we have been using a double Gaussian. This model describe with a good level of accuracy our mass peak. Anyway we try to evaluate the effect of missing a possible addtional component. Therefore we produce pseudo-experiment with a double Gaussian model and fit using only one Gaussian. By following this procedure we are intrisically assuming that the systematic effect of missing a third component is smaller than what we get removing one of the two Gaussian pdfs.

Background Lifetime Model

As for the mass model systematic described above we try to evaluate the effect of missing a possible component in the background modeling for the lifetime. We then generate using the two exponential for the positive tail and fit removing one of the positive exponential.

Signal Angular Acceptance

In order to evaluate a possible systematic due to the modeling of the signal angular acceptance coming from the Monte Carlo, we generate toys with the reweighted angular acceptance and we fit with the nominal model (the reweighted one) and using the acceptance without the reweight applied. We take the mean of the distribution of the difference of the two fit as evaluation of the systematics.

B^0 reconstructed as B_s^0

We find from a realistic MC sample of 250M events generated B^0 only 2461 survive in the region around the peak of $m \in [5.33\ \mathrm{GeV/c^2} - 5.41\ \mathrm{GeV/c^2}]$. The fraction

$\frac{N(B^0)}{N(B_s^0)} \simeq 0.12\%$ which correspond to a negligible contamination of 0.84%. Therefore no systematic uncertainty is assigned to it.

Finally we report the results for the systematics evaluation in table A.1.

Systematic Source	$c\tau[\mu m]$	$\Delta\Gamma$	$\left\|A_0\right\|^2$	$\left\|A_\parallel\right\|^2/(\left\|A_\parallel\right\|^2+\left\|A_\perp\right\|^2)$	$-2\beta_s$	$\delta_\parallel$	$\delta_\perp$
Signal mass model	$+0.6$	$+0.004$	-0.002	-0.001	-0.004	$+0.02$	$+0.005$
Angular bkg model	$+0.02$	-0.0002	$+0.0001$	$+0.0004$	-0.0009	$---$	$---$
Lifetime bkg model	-5.4	-0.004	$+0.001$	-0.002	$+0.029$	-0.027	-0.015
Signal Angular Eff	$+0.9$	-0.005	$+0.006$	$+0.004$	-0.031	$+0.007$	$+0.004$

TABLE A.1: Final systematic uncertainties list for the $B_s^0 \to J/\psi\phi$ angular analysis.

Appendix B

Normalization of the $f_i(\vec{\omega})$ functions

In the final likelihood formulation the angular part is expressed as a sum over the possible bilinear combinations of decay amplitudes (see equation 5.7.2). In this appendix we are going to discuss the numerical technique used for the normalization of the $f_i(\vec{\omega})$ functions.

First of all we know by construction the $f_i(\vec{\omega})$ functions have the following integrals over all the angular space:

$$\int_{-1}^{1}\int_{-\pi}^{\pi}\int_{-1}^{1} f_i(\vec{\omega})\ dcos\theta\ d\phi\ dcos\psi = \begin{cases} 1 & , \quad i = 1,2,3 \\ 0 & , \quad i = 4,5,6 \end{cases}$$

The integral results, and the consequent normalizations for each term, are no longer valid whenever we introduce the detector sculpting in the likelihood. In fact the new integrals we have to compute are of the kind:

$$\int f_i(\vec{\omega})\ A(\vec{\omega})\ d\vec{\omega},\ i = 1,2...6$$

where $A(\vec{\omega})$ is the acceptance curve describing the angular sculpting due to CDF detector. We use a three-dimensional histogram normalized to unity and derived from Monte Carlo to model it. Therefore we find ourselves in the situation of computing integrals of a mathematical function multiplied by an histogram. We cannot anymore compute them analytically, but we need a numerical computation.

In order to explain the formula used, let's start with the simplest one dimensional case. Given an integrable function $\int g(x)\ dx = G(x)$ and a normalized (to one) histogram, $\gamma(x)$, with a number of bins N_{Bins} along the "x" axis, the numerical computation of $\int g(x)\ \gamma(x)\ dx$ would be approximated to a sum over the histogram bins:

$$\int g(x)\ \gamma(x)\ dx \simeq \sum_{j=1}^{j=N_{Bins}} \{[G(x_j^{max}) - G(x_j^{min})]\ \cdot \gamma(x_j^{med})/binWidth\}$$

where

$$\begin{aligned} x_j^{max} &= \text{value of x at the up edge of j-esim bin} \\ x_j^{min} &= \text{value of x at the low edge of j-esim bin} \\ x_j^{med} &= (x_j^{min} + x_j^{max})/2 \\ binWidth &= \text{Width of a bin} \end{aligned}$$

The extension to our tridimensional case is then straightforward because for each function $f_i(\vec{\omega})$ we can compute the three dimensional angular integral analytically.

$$\int f_i(\vec{\omega})\ d\vec{\omega} = F_i(\vec{\omega}),\ i = 1, 2...6$$

In our case we will sum over the bins along $cos(\theta)$, ϕ and $cos(\psi)$ and the $binWidth$ would become the bin volume $binVol$. Finally, the integrals result to be approximated as:

$$\int f_i(\vec{\omega})\ A(\vec{\omega})\ d\vec{\omega} \simeq \sum_{j=1}^{j=N_{Bins}} \{[F_i(\vec{\omega}_j^{max}) - F_i(\vec{\omega}_j^{min})]\ \cdot A(\vec{\omega}_j^{med})/binVol\},\ i = 1, 2...6 \tag{B.0.1}$$

RESUME

Introduction

Le Modèle Standard (MS) est le cadre théorique développé qui decrit les particules élémentaires connues et leurs interactions. La physique du quark b est un domaine de recherche très actif pour tester les prévisions de MS ainsi que les limitations. C'est en ce sens un probe unique pour tester au delà du MS. Les mesures précises des désintégrations des mésons B^0 et de B^+, réalisées aux usines à B pendant la décennie passée, ont amélioré la conneissance de la dynamique de saveurs, prouvant que la description du MS decrit bien les mesures réalisées dans le domanie de précision et d'énergie correspondant. Une connaissance expérimentale comparable des désintégrations B_s^0 manquait jusqu'à peu. Les mésons B_s^0 sont des particules très rares et l'accélérateur Tevatron, où des protons et des antiprotons entrent en collision à une énergie du centre de masse de 1.96 TeV, fournit un accès simultané aux mésons b étranges et non-étranges, donnant une excellente occasion d'étudier le secteur de saveur du méson neutre B_s^0, même avant le démarrage du Large Hadronic Collider au CERN (LHC). A cet égard, les expériences au Tevatron peuvent avoir un impact significatif dans la mesure des propriétés du système du B_s^0 dans les années précédant les opérations de LHCb.

Phénoménologie du Système de Mésons B_s^0

Les mésons neutres B_s^0 se composent d'anti-quark $\bar{b}$ et de quark s:

$$|\mathrm{B}_s^0\rangle = |\bar{b}s\rangle \qquad |\bar{B}_s^0\rangle = |b\bar{s}\rangle$$

En l'absence de l'interaction faible avec changement de saveur les états $|\mathrm{B}_s^0\rangle$ et $|\overline{\mathrm{B}}_s^0\rangle$ seraient des états propres de l'Hamiltonien. Avec l'inclusion des termes d'interaction faible, l'Hamiltonien n'est plus diagonale dans la base $|\mathrm{B}_s^0\rangle$, $|\overline{\mathrm{B}}_s^0\rangle$, et des transitions entre les états sont possibles. Le phénomène est connu comme le mélange de méson B_s^0. L'évolution en temps du système $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ est réglé par l'équation de Schrödinger:

$$i\frac{d}{dt}\begin{pmatrix} |\mathrm{B}_s^0(t)\rangle \\ |\overline{\mathrm{B}}_s^0(t)\rangle \end{pmatrix} = \mathcal{H}\begin{pmatrix} |\mathrm{B}_s^0(t)\rangle \\ |\overline{\mathrm{B}}_s^0(t)\rangle \end{pmatrix} \tag{B.0.2}$$

$$\text{where} \quad \mathcal{H} = \begin{pmatrix} m & M_{12} \\ M_{12}^* & m \end{pmatrix} + \frac{i}{2} \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix} \tag{B.0.3}$$

avec des matricès hermitiennes (2×2) de masse et de désintégration. Les éléments diagonaux de l'Hamiltonien décrivent la masse et la largeur de désintégration des états propres de saveur. L'invariance CPT garantit que les deux états propres ont la même masse m et largeur de désintégration Γ. Les termes non-diagonaux représentent des transitions particule-antiparticule virtuelles (M_{12}) et vraies (Γ_{12}) et, si elles différent de zéro, impliquent que les états propres de masse et de saveur ne sont pas identiques. L'hamiltonien $\mathcal{H}$ est diagonalisé, par définition, dans la base de ses états propres, B_s^H et B_s^L, qui ont masse et largeur de désintégration définies ($\Gamma = 1/\tau$, où τ indique la vie). Les indices supérieurs L et H sont les abbréviations pour "light" (légère) et "heavy" (lourd). Les états propres de masse au temps $t = 0$ sont une combinaison linéaire des états propres de saveur $|\mathrm{B}_s^0\rangle$ et $|\overline{\mathrm{B}}_s^0\rangle$:

$$|B_s^L\rangle = p|\mathrm{B}_s^0\rangle + q|\overline{\mathrm{B}}_s^0\rangle\,, \tag{B.0.4}$$

$$|B_s^H\rangle = p|\mathrm{B}_s^0\rangle - q|\overline{\mathrm{B}}_s^0\rangle\,, \text{ où } |p|^2 + |q|^2 = 1 \tag{B.0.5}$$

A partir des calculs détaillés sur l'évolution de temps en l'état de saveur, il est possible de déterminer les densités de probabilité d'observer qu'un état propre de saveur produit à $t = 0$ se désintégre avec la même saveur ou son opposé, respectivement, au temps t :

$$\mathcal{P}_{\mathrm{B}_s^0 \to \overline{\mathrm{B}}_s^0}(t) = \mathcal{P}_{\overline{\mathrm{B}}_s^0 \to \mathrm{B}_s^0}(t) = \mathcal{P}_{mix}(t) = \frac{\Gamma_s}{2} e^{-\Gamma_s t} \left[1 - \cos\left(\Delta m_s t\right)\right]\,, \tag{B.0.6}$$

$$\mathcal{P}_{\mathrm{B}_s^0 \to \mathrm{B}_s^0}(t) = \mathcal{P}_{\overline{\mathrm{B}}_s^0 \to \overline{\mathrm{B}}_s^0}(t) = \mathcal{P}_{unmix}(t) = \frac{\Gamma_s}{2} e^{-\Gamma_s t} \left[1 + \cos\left(\Delta m_s t\right)\right]\,, \tag{B.0.7}$$

avec

$$\Gamma_s = \frac{\Gamma_s^L + \Gamma_s^H}{2} = \frac{1}{\tau_{\mathrm{B}_s^0}} \quad \text{et} \quad \Delta m_s = m_s^H - m_s^L. \tag{B.0.8}$$

La fréquence des oscillations de saveur correspond à la différence de masse entre les deux valeurs propres de masse du système, Δm_s. Avec la convention $\hbar = c = 1$, Δm_s est décrit dans les unités de temps inverse, typiquement ps^{-1}. Dans le MS, les diagrammes au premiere ordre pour décrire le mélange des mésons neutres B_s^0 sont représentés par les deux diagrammes dans la Figure B.1, dont la phénoménologie dépend de la matrice de Cabibbo-Kobayashi-Maskawa (CKM) de mélange des quarks.

La contribution aux boucles dans la Figure B.1 est calculé pour être proportionnelle à la masse des quarks qui apparaissent dans la boucle [10]. La masse du quark top est d'un facteur de l'ordre de 100 plus grande que la masse des quarks charm et up, et donc la contribution à la boucle du quark top domine.
Un autre paramètre important du système du B_s^0 est la différence des largeurs de désintégration définie comme:

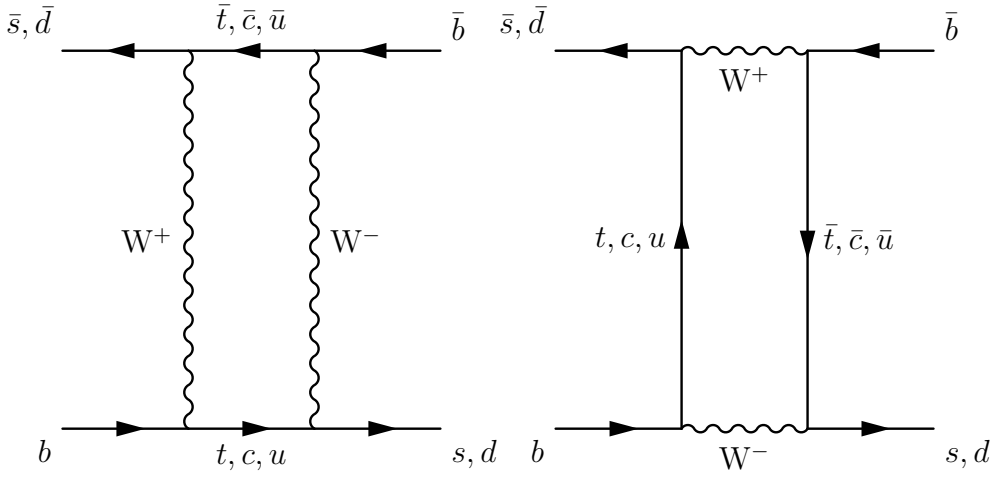

FIGURE B.1: Diagrams de plus bas ordre pour le mélange de B.

$$\Delta\Gamma_s = \Gamma_s^L - \Gamma_s^H = 1/\tau_{B_s^L} - 1/\tau_{B_s^H}. \tag{B.0.9}$$

Δm_s et $\Delta\Gamma_s$ sont définis comme étant positifs. L'information expérimentale $\Delta m_s \gg \Gamma_s$ implique indépendamment $|\Gamma_{12}| \ll |M_{12}|$, ainsi par l'expansion en termes de $\Gamma_{12}/|M_{12}$, nous obtenons:

$$\Delta m_s = 2|M_{12}| \left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \simeq 2|M_{12}| \tag{B.0.10}$$

$$\Delta\Gamma_s = 2|\Gamma_{12}|cos\phi_s \left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \simeq 2|\Gamma_{12}|cos\phi_s \tag{B.0.11}$$

où

$$\phi_s \equiv arg(-\frac{\Gamma_{12}}{M_{12}}) \tag{B.0.12}$$

Comme observé dans les équations ci-dessus, la phénoménologie du mélange des mésons B_s^0 et $\overline{\mathrm{B}}_s^0$ est caractérisée par la différence de masse des deux états propres de masse, Δm_s, et par la différence des largeurs de désintégration $\Delta\Gamma_s \equiv \Gamma_s^L - \Gamma_s^H$. Ce dernier dépend de la phase de violation CP, ϕ_s, par la relation B.0.12. Tandis que le MS prédit une petite valeur pour le phase, ϕ_s environ 4×10^{-3} rad [75], la Nouvelle Physique (NP) pourrait modifier la valeur de cette phase de manière significative par la contribution de processus additionnels, $\phi_s = \phi_s^{MS} + \phi_s^{NP}$. De plus, la même phase ϕ_s^{NP} devrait changer la phase entre le mélange et les transitions $b \to c\bar{c}s$, $2\beta_s = 2\beta_s^{MS} - \phi_s^{NP}$, où la contribution du MS est définie par $-2\beta_s^{MS} = -2\arg(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}) \approx \mathcal{O}(0.04)$, avec V_{ij} éléments de la matrice CKM. Dans notre discussion, nous pouvons approcher ϕ_s par $-2\beta_s$ puisque ϕ_s^{MS} et β_s^{MS} sont très petites par rapport à la résolution expérimentale actuelle. Ainsi, la mesure de

l'importante valeur de $2\beta_s$ (ϕ_s) serait une indication claire de Nouvelle Physique.

Cette thèse décrit le travail consacré aux mesures des paramètres qui définissent de manière cruciale le système du B_s^0. Tandis que l'oscillation de B^0 a été déterminée avec une haute précision aux usines B [2, 3], l'oscillation du B_s^0 a été récemment observé à CDF [4], après une recherche d'une vingtaine d'années. Une partie de cette thèse est consacrée à la description de l'analyse qui a permit la mesure finale de Δm_s. Dans cette dissertation je présente également les premières analyses angulaires dépendant du temps des désintégrations $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ fautes par l'expérience de CDF en utilisant les informations d'étiquetage de la saveur, qui permettent de séparer l'évolution en temps des mésons produits à l'origine comme un B_s^0 ou un $\overline{B}_s^0$. Le mode considéré ici est la désintégration d'un pseudo-scalaire à un état intermédiaire vecteur-vecteur ($P \to VV$), qui se développe par des transitions $b \to c\bar{c}s$, et génèré détats finaux impaires et pairs en CP. Dans ces désintégrations on peut statistiquement distinguer la parité P en regardant la corrélation angulaire entre les particules d'état final. Pour $B_s^0 \to J/\psi\phi$, J/ψ et ϕ sont des états propres paires en C, ainsi les propriétés de l'état $J/\psi\phi$ sous la parité P sont les mêmes que celles sous CP. Par conséquence il est possible de séparer les états de masse légèr et lourd et, donc, de mesurer la différence de largeur de désintégration $\Delta\Gamma_s$ et de déterminer la phase β_s. Ces quantités sont extrêmement utiles pour contraindre le Triangle d'Unitarité et elles expliquent de ce fait si la violation de CP dans le secteur des quarks est entièrement expliqué dans le Modèle Standard. L'étude de la désintégration $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}(\to K^+\pi^-)$, semblable du point de vue cinématique, est employée comme échantillon de calibrage pour l'analyse principale de la désintégration $B_s^0 \to J/\psi\phi$. Les résultats obtenus sont au départ prévus pour être vérifiés avec ces des usines à B plus précise, et permettent de tester examiner la robustesse de notre travail.

Echantillons de Données pour L'Analyse des Oscillations du Méson B_s^0

Les systèmes de déclenchement sont d'une importance cruciale par extraire des signatures de physique du B dans un environnement hadronique. Le système de déclenchement de CDF permet d'identifier les événements intéressants parmi le grand taux de collisions $p\bar{p}$. Les données utilisées pour l'analyse des oscillations du B_s^0 décrite dans cette dissertation ont été enregistrées par le détecteur de CDF II sur une période commençant à partir de Février 2002 et allant jusqu'à Janvier 2006.

Les différentes analyses exigent différents critères de sélection. L'analyse du mélange du B_s^0 utilise des événements recueillis avec le *two-track trigger* (TTT, déclenchement sur deux traces) et le *lepton-plus-displaced-track trigger* (ℓ + SVT, déclenchement sur un lepton plus une trace deplacée). En particulier, le déclenchement TTT , en utilisant le fait que les mésons B_s^0 ont un long temps de vie propre, est defini pour

rechercher des traces déplacées du point d'interaction primaire $p\bar{p}$. Le déclenchement TTT utilise l'information précise sur les traces à partir du détecteur de vertex en Silicium et du systéme électronique de SVT. Les déclenchements ont extrêmement puissants pour rejetter les fonds des saveurs légères (u, d, s); ils déclenchent donc sur les particules à longs temps de vie, permettant de selectionner des échantillons de données enrichis dans les événements contenant des hadrons du fond et de charme. Ces déclenchements, donnant un accès unique à B-decays hadronique, jouent un rôle principal dans la mesure d'oscillation du B_s^0 à CDF. Les échantillons du B_s^0 sont separés en deux groupes principaux, selon leur modes de désintégrationes: les désintégrations semileptoniques et hadroniques.

Désintégrations Semileptoniques du B_s^0

Les trois états finaux semileptoniques de mésons B_s^0 reconstruits dans l'analyse sont : $B_s^0 \to D_s^- \ell^+ X$, où le méson D_s^- peut se désintégrer en $D_s^- \to \phi\pi^-$ ($\phi \to K^+K^-$), $D_s^- \to K^{*0}K^-$ ($K^{*0} \to K^+\pi^-$) et $D_s^- \to \pi^-\pi^+\pi^-$. Pour identifier les désintégrations semileptoniques du B_s^0 tous les déclenchements TTT et ℓ + SVT sont utilisés.

Un accouvrement significatif existe entre les deux déclenchements de l'ordre de à 60%, par conséquent le déclenchement *lepton-plus-track* donne seulement environ 10% de désintégrations semileptoniques additionnelles du B_s^0. Selon le Review of Particle Physics [6] ces désintégrations ont un rapport d'embranchement relatif élevé et égal à $7.9 \pm 2.4\%$. De plus la présence d'un lepton fournit une signature claire. D'une part, le défi principal est représenté par la reconstruction incomplete d'au moins le neutrino impliqué dans le processus. L'échantillon final est obtenu avec une série de coupures séquentielles, optimisées pour maximiser $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$. Les critères de sélection sont basés sur la qualité des ajustements du vertex, sur les variables cinématiques et les variables d'identification des particules (PID), provenant de l'information de particule-identification fournie par le détecteur de Temps-De-Vol (TOF) de CDF II et la mesure de la perte d'énergie d'ionisation (dE/dx) dans le trajectomètre externe central (COT). En plus, pour identifier des leptons, une approche multivariable a été développée en combinant les diverses variables leptoniques d'identification dans une probabilité globale [45] : l'électron et les muons sont ainsi séparés par des hadrons appliquant une coupure inférieure sur la probabilité relative. La reconstruction inachevée de certains produits de désintégrations du B_s^0 (par example le neutrino) ne rend pas possible de reconstruire la masse invariante des candidats B_s^0. Par conséquent, la masse du D_s^- et les distributions de masse du D_s^- *plus lepton* sont employées dans la probabilité. Les distributions de masse sont montrées dans Figure B.2.

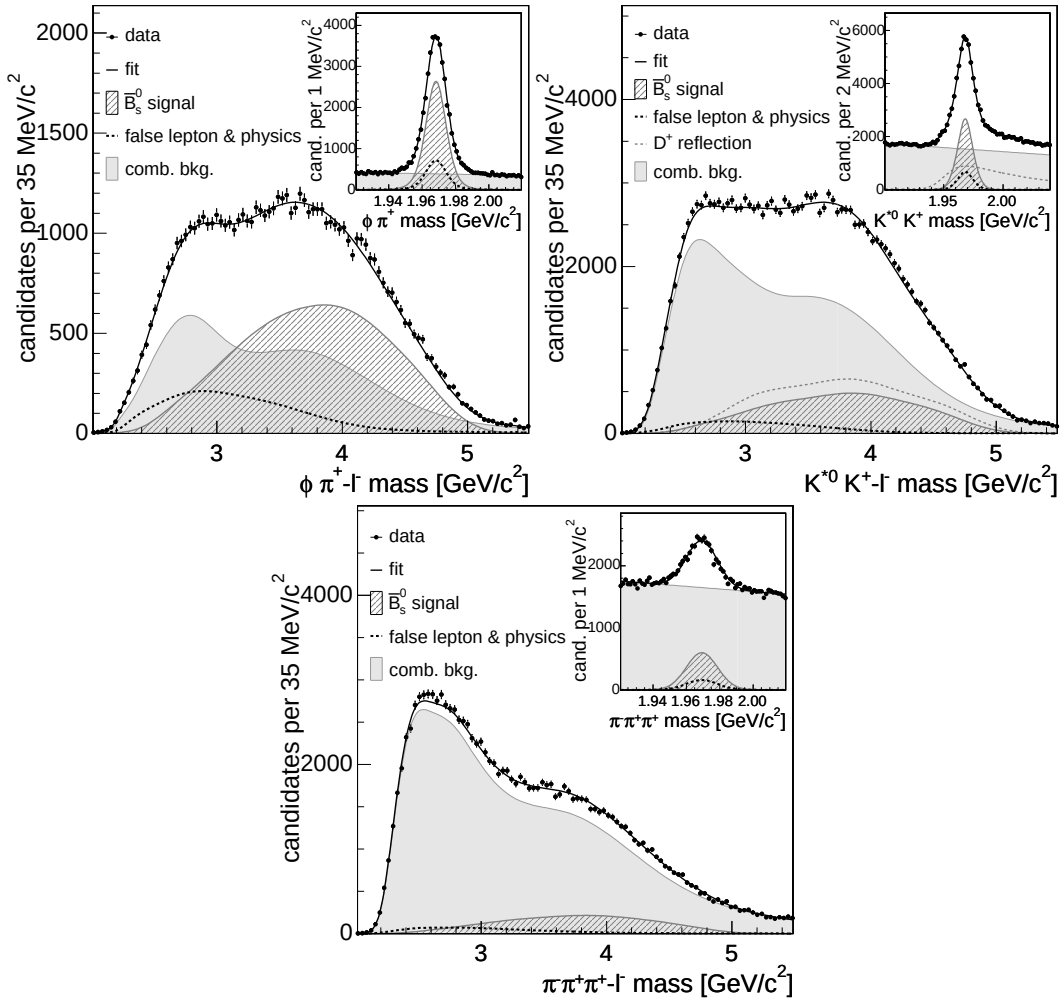

FIGURE B.2: $D_s^- + \ell^-$ et D_s^- distributions de masse pour $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to \phi\pi^-$ (en haut à gauche), $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to K^{*0}K^-$ (en haut à droite) et $B_s^0 \to D_s^- \ell^- X$, $D_s^- \to \pi^-\pi^+\pi^-$ (fond) désintégrations. Tous les candidats montrés dans les distributions de masse invariante du D_s^- sont également inclues dans la distribution de masse invariante $D_s^- + \ell^-$.

Decay Sequence	S
$B_s^0 \to D_s^-(\phi\pi^-)\ell^- X$	29,600 ± 800
$B_s^0 \to D_s^-(K^{*0}K^-)\ell^- X$	22,000 ± 800
$B_s^0 \to D_s^-(\pi^-\pi^+\pi^-)\ell^- X$	9,900 ± 700
Total	61,500 ± 1,300

TABLE B.1: Evénements du signal (S) pour les différentes désintégrations semileptoniques.

La Tableau B.1 contient les rendements des candidats B_s^0 en modes de désintégration semileptonique $B_s^0 \to D_s^- \ell^+ X$ avec $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ et $D_s^- \to \pi^-\pi^+\pi^-$.

Désintégrations Hadroniques du B_s^0

L'échantillon entièrement hadronique considéré dans l'analyse de mélange du B_s^0 se compose de six topologies du $B_s^0 \to D_s^-\pi^+$ et de $B_s^0 \to D_s^-\pi^-\pi^+\pi^-$. Le méson D_s^- charmé est reconstruit dans un des états finaux suivants: $D_s^- \to \phi\pi^-(\phi \to K^+K^-)$, $D_s^- \to K^{*0}K^-(K^{*0} \to K^+\pi^-)$ ou $D_s^- \to \pi^-\pi^+\pi^-$. Le rapport d'embranchement des modes de désintégrations hadroniques du B_s^0 est à peu près un ordre de grandeur plus petit que le cas semileptonique avec les mêmes états finaux du D_s^-. D'autre part les candidats sont entièrement reconstruits, faisant que l'échantillon hadronique est le plus puissant de ceux utilisés pour l'analyse d'oscillation du B_s^0 décrite dans cette dissertation. L'échantillon est recuillé avec le déclenchement `TTT`: afin d'avoir un échantillon cinématiquement bien compris, on exige explicitement que les deux traces déclenchées soient présentes dans la chaîne de désintégration du B_s^0. La sélection s'effectue en deux étapes. Les candidats sont d'abord pré-sélectionnés par des coupures séquentielles relâchées, ou des coupures rectangulaires. Puis, un réseau des neurones artificiel (ANN) est utilisé pour effectuer le choix final. L'utilisation d'un réseau ANN pour les désintégrations semileptoniques a été aussi étudiée mais les améliorations par rapport aux coupures séquentielles étaient marginales; on a donc décidé de ne pas l'inclure dans le cas semileptonique. Une réseau ANN peut aussi fournir plusieurs avantages par rapport à un choix d'utilisation de coupures séquentielles. Elle peut exploiter avec succès la corrélation parmi les variables en distinguant le signal du fond. Le simulateur de réseau des neurones de Stuttgart (`SNNS`) [49] connecté par interface au package de ROOT [50] a été employé pour effectuer la sélection des modes hadroniques de désintégration du B_s^0 en optimisant la quantité $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$. Les variables d'entrée de NN utilisées pour les six modes hadronique sont liées au B_s^0 et à l'ajustements des vertexes de désintégration, à des variables cinématiques aussi bien que des variables d'identification des particules (PID). Les distributions de masse des candidats B_s^0 entièrement reconstruits dans des désintégrations hadroniques sont montrées dans la Figure B.3. Le spectre de masse le plus claire est celui trouvé pour le canal de désintégration $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi\pi^-$. Pour ce mode de désintégration spécifique, l'échantillon des données contient aussi des désintégrations hadroniques incomplètement reconstruites. Bien que dans le cas des candidats B_s^0 semileptoniques, une ou plusieurs particules soient exclues dans la topologie de désintégration et la

Decay Sequence	S	S/B
$B_s^0 \to D_s^-(\phi\pi^-)\pi^+$	1,900	11.3
$B_s^0 \to D_s^-(K^*(892)^0K^-)\pi^+$	1,400	2.0
$B_s^0 \to D_s^-(\pi^+\pi^-\pi^-)\pi^+$	700	2.1
$B_s^0 \to D_s^-(\phi\pi^-)\pi^-\pi^+\pi^+$	700	2.7
$B_s^0 \to D_s^-(K^*(892)^0K^-)\pi^-\pi^+\pi^+$	600	1.1
$B_s^0 \to D_s^-(\pi^+\pi^-\pi^-)\pi^-\pi^+\pi^+$	200	2.6
Partially Reconstructed B_s^0 Decays	3,300	3.4
Total	8,800	—

TABLE B.2: Evénements de signal (S) et rapport signal sur fond (S/B) pour les differents désintégration hadroniques.

structure au-dessous du peak principal est dominée par un nombre restreint de canaux qui manquent seulement une particule neutre de faible impulsion. La source primaire des modes partiellement reconstruits dans la topologie de $B_s^0 \to D_s^-\pi^+$ sont le $B_s^0 \to D_s^-\rho^+$, $\rho^+ \to \pi^+\pi^0$ et $B_s^0 \to D_s^{*-}\pi^+$ avec $D_s^{*-} \to D_s^-\gamma$ or $D_s^{*-} \to D_s^-\pi^0$. Pour ces modes de désintégration, la particule neutre γ ou π^0 ne laisse aucune trace dans le trajectomètre, le COT et le detector au Silicium; ils sont donc négligés dans la reconstruction finale. Dans le cas des modes hadroniques partiellement reconstruits, 96% de l'impulsion d'un candidat est reconstruite en moyenne, donc ils représentent une source potentielle statistiquement significatif dans la mesure.

Le Tablau B.2 montre le rendement du signal et le rapport du signal-au-fond dans la gamme de masse [5.32, 5.42]GeV/c^2 pour chaque canal de désintégration.

Eléments pour L'Analyse du Mélange du B_s^0

La résolution de l'oscillation rapide entre B_s^0 et $\overline{B}_s^0$ et la mesure précise de la fréquence Δm_s d'oscillation ont été des buts importants des expériences LEP au SLC et au Tevatron. Les oscillations du B_s^0 apparaissent dans la dépendence en temps de l'asymétrie mélangée $\mathcal{A}(t)$

$$\mathcal{A}(t) = \frac{\mathcal{P}_{unmix}(t) - \mathcal{P}_{mix}(t)}{\mathcal{P}_{unmix}(t) + \mathcal{P}_{mix}(t)} = \cos(\Delta m t)\,. \tag{B.0.13}$$

Une approche directe possible pour mesurer la fréquence Δm_s serait d'ajuster l'asymétrie $\mathcal{A}(t)$, en fonction du temps, sur le nombre des désintégrations de candidats B_s^0 avec la même saveur ou la saveur opposée, comparé au temps de production. Dans le MS, on s'attend à ce que l'oscillation de mélange Δm_s du B_s^0 soit au moins environ 40 fois plus grande que Δm_d; en plus, dans la majeure partie des extensions du MS la frequence d'oscillation est prévu d'acquérir des valeurs plus élevées. Puisqu'il n'est pas possible de savoir à l'avance si l'analyse aura assez de résolution pour une observation directe de l'oscillation, cette méthode ne peut pas être utilisée de maniere simple. Historiquement, les recherches de mélange ont été rendues possible grâce à l'insertion d'un coefficient de Fourier, l'amplitude $\mathcal{A}$, pour ajuster l'échelle du terme

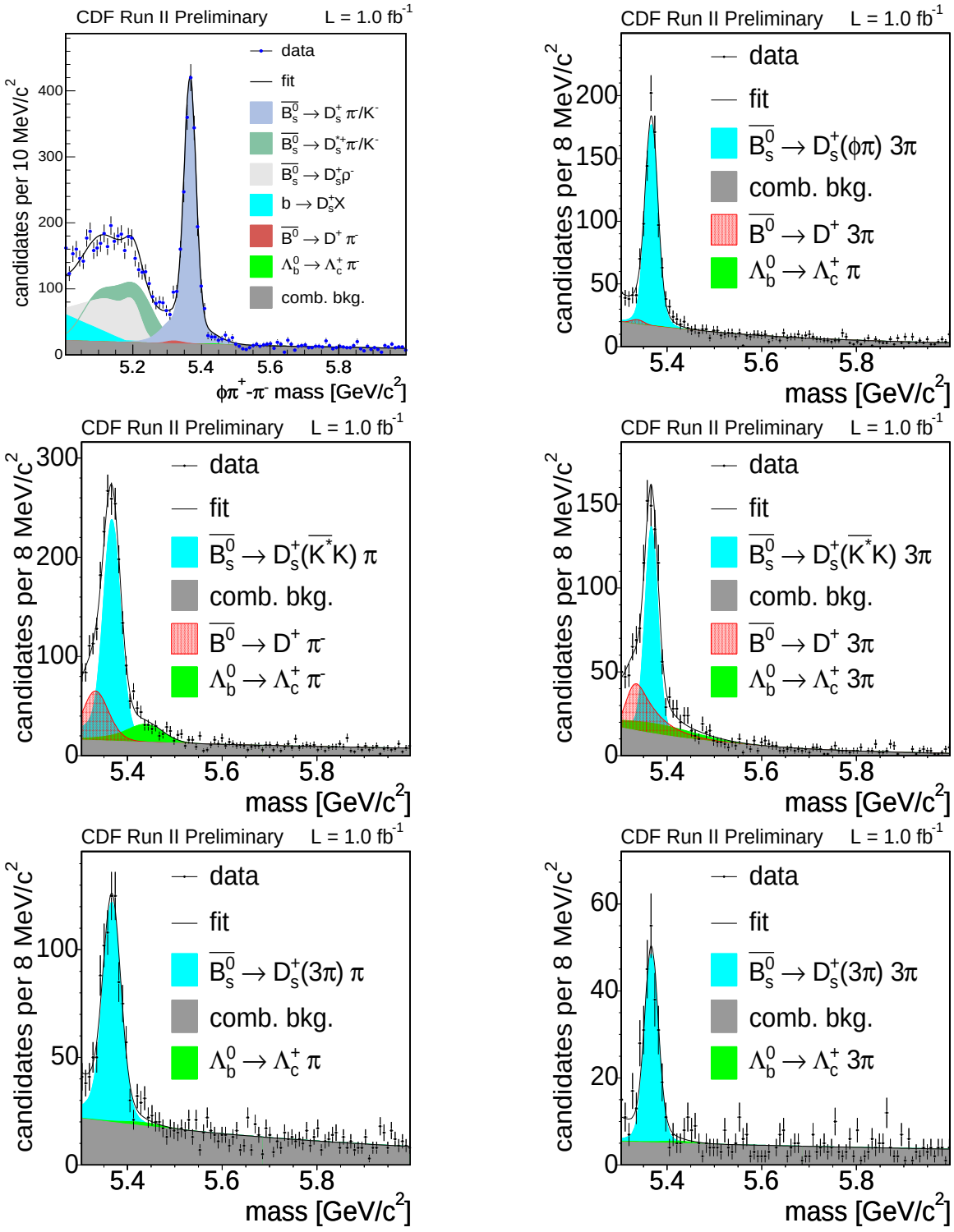

FIGURE B.3: Spectres de masse hadronique, qui montrent tous les composants des échantillons. Dans la colonne à gauche les désintégrations $B_s^0 \to D_s^- \pi^+$, dans la colonne à droite les désintégrations $B_s^0 \to D_s^- \pi^- \pi^+ \pi^+$. De haut au bas les désintégrations $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$, $D_s^- \to \pi^+ \pi^- \pi^-$.

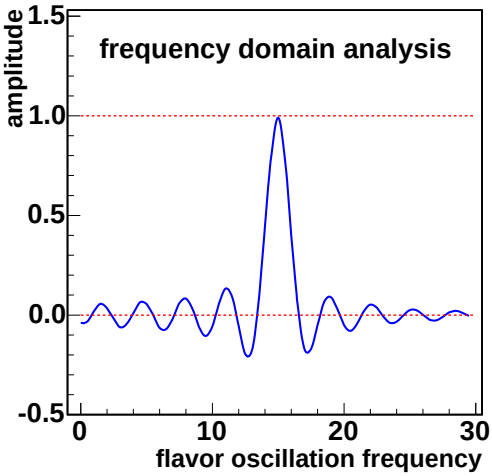

FIGURE B.4: *Amplitude scan* pour une simulation Monte Carlo qui assume la valeur vraie de la fréquence d'oscillation égale à $DMS = 15$ pb^{-1}.

de cosinus de l'oscillation

$$1 \pm \cos(\Delta m t) \rightarrow 1 \pm \mathcal{A} \cos(\Delta m t) . \tag{B.0.14}$$

Avec l'introduction d'une amplitude $\mathcal{A}$, la recherche de l'oscillation du B_s^0 se fait dans le domaine de fréquence. La technique, appelée *amplitude scan* [56, 57] est décrite comme suit. Pour une fréquence fixe d'oscillation, le résultat de l'ajustement du paramètre $\mathcal{A}$ est prévu être compatible avec l'unité si la valeur sondée Δm_s correspond à la valeur vraie de la fréquence. Autrement, on s'attend à qu'il soit compatible avec zéro. Un exemple d'*amplitude scan*, produit avec un échantillon de Monte Carlo qui simule $\Delta m_s = 15$ pb^{-1}, est montré dans la Figure B.4.

L'idée derrière cet approche est de sonder des valeurs d'augmentation de Δm_s. La région sondée dépend de la sensibilité de l'analyse, qui est définie comme la valeur de la fréquence Δm_s pour laquelle une amplitude nulle mesurée $\mathcal{A} = 0$ impliquerait l'exclusion de $\mathcal{A} = 1$ au niveau de confiance désiré, qui est nominalement choisi être de 95% dans nos analyses. L'erreur sur l'amplitude $\sigma_{\mathcal{A}}$ a une distribution Gaussienne, ainsi une valeur de Δm_s peut être exclue au niveau de confiance de 95% si l'amplitude $\mathcal{A}$ respecte la condition

$$\mathcal{A} + 1.645 \cdot \sigma_{\mathcal{A}} \leq 1. \tag{B.0.15}$$

La limite d'exclusion est définie comme le plus grande fréquence d'oscillation qui serait exclue dans la non-oscillation idéale $\mathcal{A} = 0$:

$$1.645 \cdot \sigma_{\mathcal{A}} = 1. \tag{B.0.16}$$

L'estimateur de $\sigma_{\mathcal{A}}$ est donné par la formule suivante [56, 57]

$$\frac{1}{\sigma_{\mathcal{A}}} \simeq \frac{\mathcal{S}}{\sqrt{\mathcal{S}+\mathcal{B}}}\, e^{-\frac{\Delta m_s^2 \sigma_{ct}^2}{2}} \sqrt{\frac{\epsilon \mathcal{D}^2}{2}}. \tag{B.0.17}$$

Les obstacles principaux à l'analyse du mélange du B_s^0 proviennent de trois contributions principales. La statistique disponible, exprimée par le terme $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$ où $\mathcal{S}$ et $\mathcal{B}$ sont, respectivement, les événements reconstruits comme signal et fond. La dépendance de la résolution en temps de vie propre: plus la résolution sera éleveé plus les possibilités de rechercher de grandes valeurs d'oscillation seront élevées. Enfin la capacité d'identifier la saveur B_s^0 au temps de production est définie par le facteur de mérite de l'algorithme d'étiquetage utilisé, $\epsilon\mathcal{D}^2$. L'efficacité ϵ correspond à la fraction des événements auxquels l'algorithme d'étiquetage assigne une décision non nulle. La dilution $\mathcal{D}$ du marquer (*tagger*) est définie comme $\mathcal{D} = 2\mathcal{P}_{tag} - 1$, où $\mathcal{P}_{tag}$ est la probabilité d'étiqueter correctement le candidat B_s^0.

Temps de Vie Propre

Pour résoudre les oscillations rapides, une détermination précise de la période de désintégration des mésons B_s^0 est cruciale. CDF grâce à son système de trajectomètres en Silicium détermine precisement la distance de vol dans le plan transverse, L_{xy}. Le temps de désintégration est $ct = L_{xy}(B_s^0) \cdot M_{B_s^0}/p_T(B_s^0)$, où les observables pour l'événement sont l'impulsion transverse du B_s^0, $p_T(B_s^0)$, et la longueur transversale de désintégration du B_s^0, $L_{xy}(B_s^0)$ définie comme le déplacement transverse du vertex de désintégration du B_s^0. La valeur de la masse M_{B_s} utilisée dans la formule appropriée du temps de désintégration provient du calcule des moyennes mondiales [58]. Dans les désintégrations hadroniques, la mesure de précision du deuxième vertex est la contribution principale à σ_{ct}. Dans les désintégrations semileptoniques et hadroniques partiellement reconstruits cela devient une addition importante à l'incertitude globale. En fait, dans ce cas là, les observables d'une ou plusieurs particules ne sont pas mesurées. Par conséquent, le temps de vie propre reconstruit n'est pas le temps propre de désintégration du candidat B_s^0, et un facteur de correction, appelé le " k-factor", est nécessaire pour expliquer l'impulsion manquante

$$ct = ct^* \cdot k,\ k \equiv \frac{L_{xy}(B_s^0)}{L_{xy}^{reco}} \frac{p_T(B_s^0)}{p_T^{reco}}. \tag{B.0.18}$$

La longueur de désintégration ct^* est traditionnellement appellée pseudo-longueur de désintégration propre et elle est mesurée avec seulement l'information venant des candidats reconstruits. Puisqu'il n'est pas possible de déterminer la valeur du facteur k événement-par-événement, on le traite comme une "distribution de densité de probabilité" pour une classe entière des événements. C'est-à-dire, une distribution moyenne, $F(k)$, parce que le facteur k est obtenu à partir de la simulation par `BGenerator-MC`. Les distributions des facteurs k des modes de désintégration hadroniques partiellement reconstruits et semileptoniques est montrées dans la Figure B.5.

La distribution du temps de désintégration du signal est modélisée par l'equation

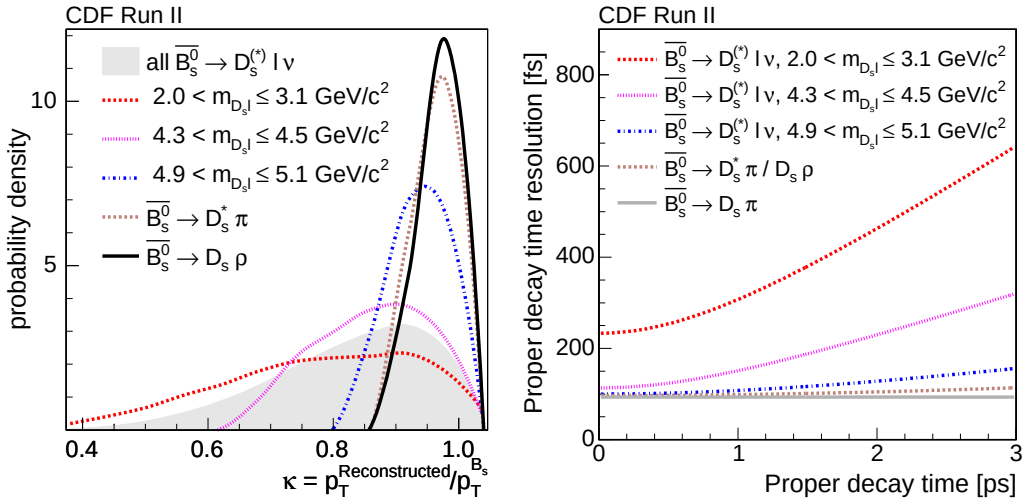

FIGURE B.5: A gauche: Distributions du facteur k pour differents regions de masse du ℓD_s^- pour des modes de désintégration semileptonique et hadronique partiellement reconstruits. A droite: la resolution de temps de vie propre, σ_{ct}, en fonction de la longueur du temps de désintégration ct.

$\mathcal{P}(ct_i, \sigma_{ct_i} = \epsilon(ct_i) \times \int \Gamma_s e^{-\Gamma_s t'} \Gamma(t - t_i, \sigma_{ct_i}) dt'$, où ct_i est la longueur mesurée de désintégration du candidat i, Γ_s est la largeur de désintégration du B_s^0, $\Gamma(x-\mu, \sigma)$ est une distribution gaussienne de la variable aléatoire x avec la moyenne μ et la largeur σ, et σ_{ct_i} est la résolution prévue du temps de désintégration-temps du candidat. Pour les désintégrations $B_s^0 \to D_s(3)\pi$ la distribution en ct de vrais mésons B_s^0 est biasée par les coupures sur le temps de vie du système de déclenchement TTT. Dans ces cas, le biais peut être corrigé avec l'introduction d'une fonction d'efficacité du temps de désintégration, $\epsilon(ct)$, qui est déterminée par la simulation Monte Carlo. Nous estimons la résolution σ_{ct_i} du temps de désintégration pour chaque candidat en utilisant les paramètres de mesures des traces et leurs incertitudes telle que prévues. En fait, l'incertitude spatiale sur le vertex secondaire dépend de diverses caractéristiques de la désintégration, comme sa cinématique et la qualité de l'ajustement. Le but est de paramétriser la résolution σ_{ct_i} en utilisant les observables dont il y a dépendance claire. Pour réaliser une telle étude, un ensemble de désintégrations de candidats B_s^0 aux positions connues est nécessaire. Puisque ce n'est pas le cas, un échantillon "ad hoc" de calibrage qui reproduit possiblement la topologie de cinématique et du vertex des échantillons du signal a été choisi. La caractéristique principale de l'échantillon est d'être "*prompt*" : elle contient des événements avec la topologie du vertex comme un méson B et supposé se désintégrer en proximité du vertex primaire dans les interactions $p\bar{p}$. Le déclenchement TTT est employé pour extraire l'échantillon de calibrage. Ce dernier est obtenu en groupant un vrai méson D avec une (ou trois, selon la topologie du canal) trace(s) "prompt" du vertex primaire. Après la calibration finale, des déviations résiduelles sont encore présentes. Dans l'analyse du mélange du B_s^0, une variation résiduelle d'environ $\pm 4\%$ est observée et incluse dans les erreurs systematiques de la mesure du Δm_s. Les événements semileptoniques ont une incertitude

intrinsèque de p_T dûe à l'impulsion manquante du neutrino. Ceci se traduit dans une incertitude sur le temps de vie, qui porte préjudice à la sensibilité pour les oscillations rapides (grand Δm_s). Pour cette raison, les événements hadroniques sont particulièrement importants. Le temps de désintégration moyen est de 87 fs et de 150 fs respectivement pour des événements hadroniques et semileptoniques. Ceci devrait être comparé à la période de la fréquence d'oscillation du B_s^0 environ de 300 fs ($\Delta m_s = 18$ ps^{-1}). La résolution du temps de désintégration, σ_{ct}, en fonction de la longueur du temps de désintégration ct est montrée dans la Figure B.5.

Les Algorithmes d' Etiquetage de la Saveur

L'étiquetage de la saveur, après la sélection de léchantillon et la résolution en temps de désintégration apropriée, constitue un autre ingrédient clé des analyses incluant le phénomène de mélange. Tandis que la saveur des candidats B_s^0 peut être expliquée par le signe de la charge des particules filles, les algorithmes d' étiquetage de la saveur déterminent avec un certain degré d'incertitude la saveur du B_s^0 au temps de production. Au Tevatron les quarks b sont la plupart du temps produits en paires dans les interactions fortes de partons. Par conséquent les algorithmes d'étiquetage sont naturellement divisés en deux classes principaux : "Same Side flavor taggers" (SST) et "Opposite Side flavor taggers (OST)". Le premier explore la corrélation de charge de saveur entre les B_s^0 et les traces de fragmentation produites dans le processus d'hadronisation tout près du méson B. Les algorithmes d'étiquetage OST examinent les produits de désintégration de l'autre hadron b produit dans la collision $p\bar{p}$: la saveur du hadron b est ainsi reliée (opposée) à la saveur du méson B_s^0 sélecitoné par le déclenchements. L'efficacité $Q = \epsilon\mathcal{D}^2$ de ces techniques est mesurée avec une efficacité ϵ, la fraction de candidats signal avec une étiquette de saveur, et une dilution $\mathcal{D}^2 = 1 - 2w$, où w est la probabilité que l'étiquetage est incorrecte. Pour les *Opposite side taggers* nous employons la charge leptonique (e et μ) et la charge de jet pour l'étiquetage à partir de techniques développées dans CDF I pour la mesure de Δm_d. Nous ajoutons un étiquetage de saveur sur la charge des kaons identifiés, et puis nous combinons l'information du kaon, du lepton, et des étiquetages de charge des jets en utilisant un réseau de neurones artificiel (ANN). La dilution est mesurée dans les données en utilisant de grands échantillons de mésons $B^{\pm}$, qui ne changent pas de saveur, et de B^0, qui peuvent être employés en tenant compte de leur fréquence d'oscillation bien connue. L'efficacité combinée d'étiquetage par OST est $Q = 1.8 \pm 0.1\%$. On s'attend à ce que la dilution des étiquetages de saveur OST soit indépendante du type du méson B qui produit les désintégrations hadroniques ou semileptoniques. Les *Same-side flavor tags* sont basées sur les charges des particules produites dans la fragmentation du quark b qui produit le B_s^0 reconstruit. Dans l'image la plus simple de la fragmentation, un π^+ (π^-) accompagne la formation d'un B^- (B^+), un π^- (π^+) accompagne le $\bar{B}^0$ (B^0), et un K^- (K^+) accompagne un $\overline{B}_s^0$ (B_s^0). Dans le Run I, CDF a établi cette méthode d'identification de saveur de production dans les mesures de Δm_d et du paramètre de violation CP, $sin(2\beta)$. Nous avons employé un étiquetage du SST basé sur notre probabilité d'identification de kaons, basée sur la mesure du dE/dx et l'information du temps-de-vol, combinée

avec un réseau de neurones artificiel (ANN) à des quantités cinématiques du candidat kaon dans une variable d'étiquetage simple T. Les traces à côté de l'espace de phase du candidat B_s^0 sont considérées comme candidates d'étiquetage du kaon de même-côté, et la trace avec la plus grande valeur de T est choisie comme trace d'étiquetage. Nous prédison la dilution de l'étiquetage du même-côté en utilisant les échantillons de données simulés produits par le programme de Monte Carlo `PYTHIA`. Des échantillons de contrôle B^+ et B^0 sont utilisés pour valider les prévisions de la simulation. L'efficacité de cet étiquetage de saveur augmente avec le montant p_T du B_s^0 ; nous trouvons $Q = 3.7\%$ (4.8%) dans l'échantillon (semileptonique) hadronique de désintégration. L'incertitude partielle sur Q est approximativement de 25%. S'il y a à la fois un étiquetage de même-côté et un étiquetage du côté opposé, nous combinons l'information des deux étiquetages en supposant qu'ils ne sont pas corrélés. Les algorithmes du même-côté, en étant à peu près trois fois plus efficace que les algorithmes du côté opposé, augmente sensiblement la puissance statistique de l'analyse.

Résultats des Oscillations de Méson B_s^0

Nous employons un ajustement sans "*binning*" avec une fonction de vraisemblance pour rechercher des oscillations du B_s^0. La fonction de vraisemblance combine la masse, le temps de désintégration, la résolution en temps de désintégration et l'information sur l'étiquetage de la saveur pour chaque candidat, et inclut des termes pour le signal et chaque type de fond. L'ajustement est fait en trois étapes. D'abord, un ajustement combiné de masse et de temps de désintégration est realisé permettant de séparer le signal du fond et de fixer des modèles de la masse et du temps de désintégration. Les ajustements combinés pour la masse du B_s^0 et la largeur de désintégration dans les échantillons hadroniques et pour la largeur de désintégration dans les échantillons semileptoniques conduisent à des mesures compatibles avec les moyennes mondiales du PDG. Puis les asymétries de saveur sont mesurées pour les composantes du fond. La dernière étape est un ajustement pour des oscillations $B_s^0 - \overline{B}_s^0$; les modèles de la masse et du temps de désintégration et les asymétries du fond sont fixés à partir des deux étapes précédentes. La distribution de densité de probabilité de signal a la forme générale:

$$S_{\pm}(t_i, \sigma_{ct_i}, \mathcal{D}_i) = \epsilon(t_i) \int \frac{\Gamma_s}{2} e^{-\Gamma_s t'} [1 \pm \mathcal{A}\mathcal{D}_i cos(\Delta m_s t')] \mathcal{G}(tit', \sigma_{ct_i}) dt', \qquad \text{(B.0.19)}$$

où $\mathcal{D}_i$ est la dilution du i-ème candidat, et t_i, σ_{ct_i}, $\mathcal{G}$, et$\epsilon(t)$ ont été définis précédemment. Suivant la méthode décrite, nous ajustons l'amplitude d'oscillation $\mathcal{A}$ en fixant Δm_s à une valeur testée. Sprès la calibration de tous les effets du détecteur ($\mathcal{D}_i$, σ_{ct_i}), on s'attend à ce que l'amplitude d'oscillation soit compatible avec $\mathcal{A} = 1$ quand la valeur du test est la véritable fréquence d'oscillation, et compatible à $\mathcal{A} = 0$ quand la valeur du test est loin de la véritable fréquence d'oscillation. La Figure B.6 montre la valeur de l'amplitude en fonction de la fréquence d'oscillation. La sensibilité est 19.3 ps^{-1} pour les désintégrations semileptoniques, 30.7 ps^{-1} pour

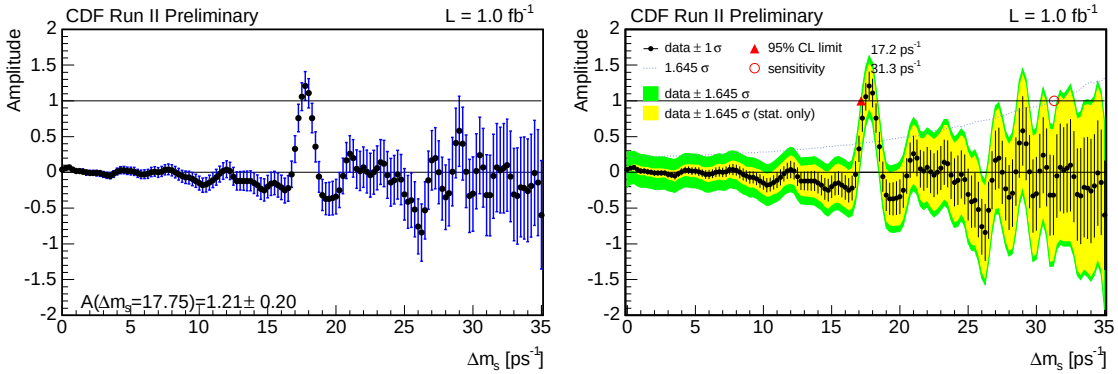

FIGURE B.6: L'Amplitude de scan pour les données de tous les échantillons et les algorithmes d'étiquetage combinés. Le graphe à gauche souligne seulement l'incertitude statistique qui est la partie dominante de l'erreur sur l'amplitude. La version à droite inclut également les incertitudes systématiques.

les désintégrations hadroniques, et 31.3 ps^{-1} pour tous les désintégrations combinés. Pour $\Delta m_s = 17.75$ ps^{-1}, l'amplitude observée $\mathcal{A} = 1.21 \pm 0.20$(stat.) est compatible avec l'unité, ce qui indique que les données sont compatibles avec des oscillations $B_s^0 - \overline{B}_s^0$ avec cette fréquence, alors que l'amplitude n'est pas compatible avec zéro: $\mathcal{A}/\sigma_{\mathcal{A}} = 6.05$, où $\sigma_{\mathcal{A}}$ est l'incertitude statistique sur $\mathcal{A}$. La petite incertitude sur $\mathcal{A}$ à $\Delta m_s = 17.75$ ps^{-1} est dûe à la résolution supérieure du temps de désintégration des modes hadroniques. Ces derniers représentent le plus grand échantillon du B_s^0 hadronique au niveau mondial, obtenu grâce aux grandes exécutions de CDF dépistant et de système de déclenchement, en particulier le Layer00 et le systéme de trajectographie du vertex secondaire.

Nous évaluons la signification du signal en utilisant $\Lambda(\Delta m_s) = log[\mathcal{L}^{\mathcal{A}=0}/\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)]$; c'est le logarithme du rapport des probabilités correspondant à l'hypothèse d'oscillations ($\mathcal{A} = 1$) à la valeur du test Δm_s et de l'hypothèse que $\mathcal{A} = 0$, qui équivant à des étiquetages de production de saveur aléatoires. La Figure B.7 montre Λ en fonction de Δm_s. Des courbes séparées sont montrées seulement pour les données semileptoniques (pointillées), les données hadroniques (trait plein léger), et seules les données combinées (trait plein foncé).

Pour mesurer Δm_s, nous fixons $\mathcal{A} = 1$ et nous ajustons pour la fréquence d'oscillation. Nous trouvons $\Delta m_s = 17.77 \pm 0.10$ (stat.) ± 0.07 (syst.) ps^{-1}. La seule incertitude systématique non-négligeable sur Δm_s est l'incertitude sur l'échelle absolue de la mesure de temps de désintégration. Les contributions à cette incertitude incluent des biais dans la reconstruction du vertex primaire dûe à la présence du hadron b dans le côté opposé, aux incertitudes sur l'alignement du détecteur Silicium, et aux biais sur la determination des traces. L'incertitude sur le facteur k pour les candidats hadroniques avec un photon ou un π^0 est inclue mais a un effet négligeable. Au minimum, $\Delta m_s = 17.77$ ps^{-1}, $\Lambda = -17.26$. La signification du signal est la probabilité que des données étiqueté d'un façon aléatoire produisent une valeur de Λ plus petite que -17.26 à n'importe quelle valeur de Δm_s. Nous répétons le scan de la fonction de

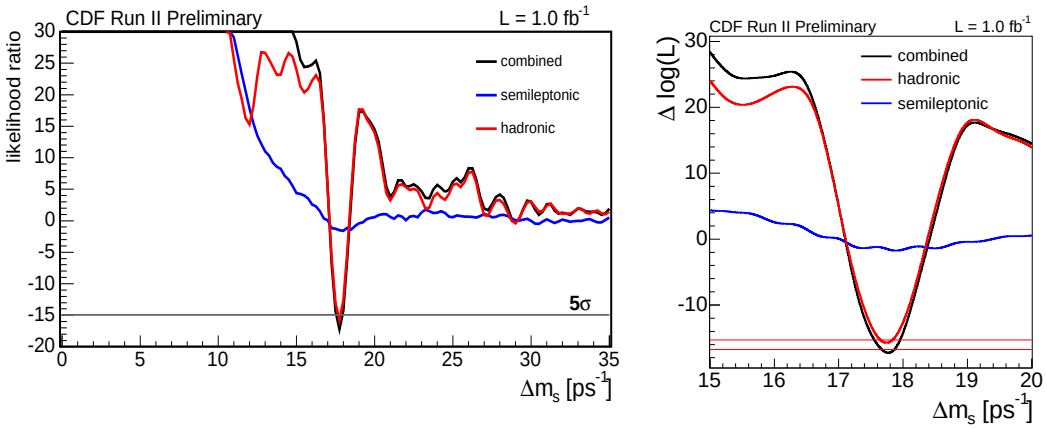

FIGURE B.7: Profil du rapport de probabilité: rapport de la fonction de vraisemblance en fonction de Δm_s (à gauche) avec un zoom sur la region $[15, 20]$ ps^{-1}. La contribution hadronique (rouge), semileptonique (bleu) et leur combinaison (noir) sont montrées séparément .

vraisemblance 350 million de fois avec des déterminations aléatoires de l'étiquetage; 28 ont une valeur de Λ inferieure à -17.26, ce qui correspond à une probabilité de 8×10^{-8} (5.4σ), bien en-dessous de 5.7×10^{-8} (5σ). Bien que l'observation définitive de la fréquence de mélange du B_s^0 soit une mesure importante en elle-même, elle fournit en autre le moyen d'obtenir des mesures extrêmement précises des paramètres CKM. En particulier, il est possible d'extraire le rapport $|V_{ts}|^2/|V_{td}|^2$

$$\frac{|V_{ts}|^2}{|V_{td}|^2} = 0.2060 \pm 0.0007\ (\text{exp})\ ^{+0.0081}_{-0.0060}\ (\text{theory})\ , \qquad \text{(B.0.20)}$$

où la première incertitude - indiquée par "exp" - est uniquement liée à la mesure de Δm_s et la deuxième inclut toutes les autres sources, et est dominées par l'incertitude théorique. La valeur de Δm_s est tout à fait compatible avec l'unitarité de la matrice de CKM dans le Modèle Standard et, actuellement le test est seulement limité par les calculs théoriques. La conséquence immédiate est la nécessité d'un travail d'améliorer le calcul de QCD sur le réseau pour la détermination du paramètre ξ, pour optimiser l'utilisation de l'information fournie par l'analyse de mélange du B_s^0. L'importance de la mesure par CDF de $|V_{ts}|^2/|V_{td}|^2$ est montré, en utilisant le cadre du "CKM fitter" [13], dans la Figure B.8, où la mesure de CDF est comparée à la prediction théorique et à la moyenne des mesures de Belle [77] et BaBar [78], et dans la Figure B.9 qui montre l'effet direct de la mesure de Δm_s sur le triangle d'unitarité. Le résultat saisissant est le reserrement de l'ellipse de 95% de taux de confiance dans le plan $(\bar{\rho}, \bar{\eta})$ correspondant à la mesure de $\Delta m_d/\Delta m_s$. La valeur centrale de Δm_s de 17.77 ps^{-1} est compatible avec la prediction théorique du Modèle Standard de $18.3^{+6.5}_{-1.5}$ ps^{-1} donnée par le groupe "CKM fitter"; ainsi cette observation est compatible avec l'unitarité de la matrice CKM.

La détermination de fréquence d'oscillation indique que la présence de Nouvelle Physique au delà du Modèle Standard, si il y a physique nouvelle, est extrêmement

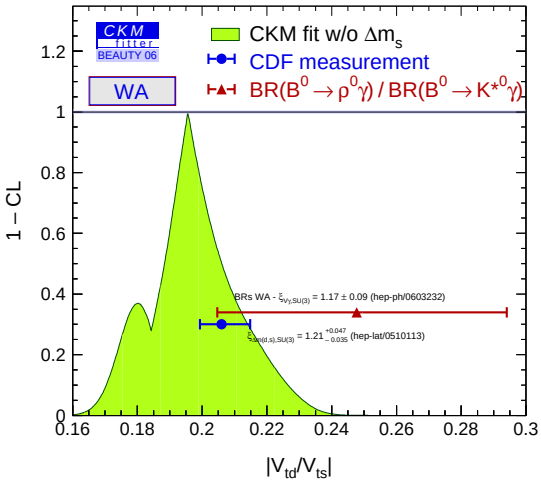

FIGURE B.8: Ajustement du groupe "CKM fitter" et deux mesures expérimentales de $|V_{ts}|^2/|V_{td}|^2$. Chaque section horizontale du secteur vert correspond à l'intervalle où $|V_{ts}|^2/|V_{td}|^2$ se trouve avec le niveau de confiance lu sur l'axe vertical. Les deux mesures représentent les résultats de CDF et la mesure moyenne de Belle et de BaBar du rapport $|V_{ts}|^2/|V_{td}|^2$.

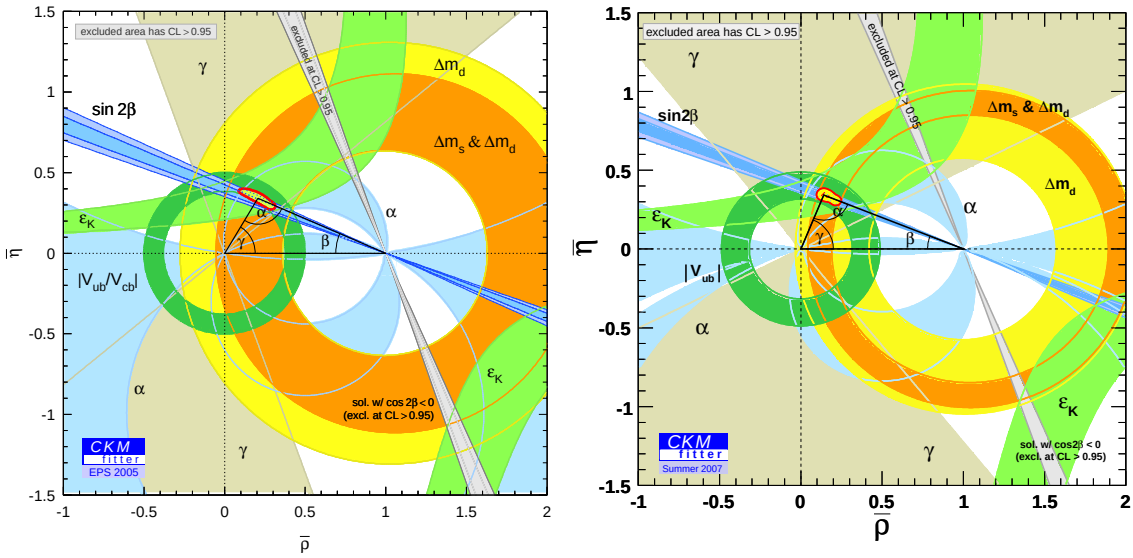

FIGURE B.9: Ajustement du Triangle d'Unitarité par le "CKM fitter" à EPS2005 (gauche) comparé à les mesures plus récents, qui inclut l'observation de CDF (droite).

petite par rapport à la magnitude de l'amplitude de mélange. Néanmoins, la Nouvelle Physique peut modifier la phase de l'amplitude de mélange β_s, par le contribution avec des processus additionnels. En effet, les ajustements globaux sur des données expérimentales contraignent la phase de violation de CP à une petite valeur $2\beta_s \simeq 0.04$ [75]. Par conséquent, la Nouvelle Physique peut amener une valeur sensiblement plus grande. Après la mesure de Δm_s, la suite normale des études de système du B_s^0 est la mesure de la phase β_s. La deuxième partie de ma thèse, qui sera discutée dans les sections suivantes, a donc été centrée sur l'analyse angulaire de désintégration $B_s^0 \to J/\psi\phi$ pour la première mesure mondiale de β_s en utilisant l'information d'étiquetage de la saveur.

Ma contribution personnelle à l'analyse du mélange de méson B_s^0 s'est focalisée dans le secteur d'étiquetage de la saveur, afin de fournir et de calibrer les algorithmes pour identifier la saveur du méson de B au moment de sa production. Plus spécifiquement, j'ai travaillé au calibrage des étiquetage de la saveur du côté opposé, dans sa rèalisation finale avec le réseau des neurones, et dans les premières études d'étiquetage avec les kaons du même côté dans le développement des algorithmes basés sur la cinématique et l'identification des particules. Ce dernier a été le premier algorithme de SSKT mis en application dans l'analyse de mélange du B_s^0 et a fourni la première évidence pour les measures de B_s^0 [14]. D'autre part, la version finale de l'étiquetage de même côté avec les kaons utilisé pour la première observation mondiale du mélange du B_s^0 consiste en une mise à jour importante de l'algorithme combinant ensemble la cinématique et l'information d'identification des particules avec un réseau des neurones [66]. La puissance de l'étiquetage du *same side tagger* décrit dans cette thèse est environ de 4%, une fois appliquée aux échantillons du B_s^0 reconstruits, ceci est à comparer avec environ le 1.8% pour le *opposite side tagger*.

Echantillons de Données pour L'Analyse Angulaire des Désintégrations $B_s^0 \to J/\psi\phi$

La sélection des événements $B_s^0 \to J/\psi\phi$, avec $J/\psi \to \mu^+\mu^-$ et $\phi \to K^+K^-$, suit le même procédé discuté pour la sélection hadronique du B_s^0. Une serie de coupures relâchées de Pré-Sélection est suivi par l'utilisation d'un réseau de neurons artificiel (ANN) pour la sélection des événements. Les coupures de pré-sélection poursuivent plusieurs objectifs:

- Certaines régions de l'espace des variables de l'événement ne sont pas bien modelisées par la simulation Monte Carlo. C'est particulièrement vrai pour les traces avec basse impulsion transverse (p_T). Nous coupons sur ces variables pour assurer le bon accord Données-Monte Carlo.
- Couper sur certaines variables pour supprimer les événements qui sont clairement du fond. En effet, la pollution des échantillons de formation du réseau de neurons par des événements qui sont des événements de fond ajoutera simplement une complication additionnelle au réseau de neurons.

- Réduire le nombre d'événements que nous employons pour la fonction de vraisemblance et diminuer le temps de CPU pour réaliser l'ajustement. Par conséquent, nous incluons assez d'événements provenant des régions de masse sur les bords latéraux pour pouvoir décrire empiriquement les distributions angulaires du fond, mais pas plus.

Le réseau des neurones est exercé en utilisant, comme échantillon de fond, des données qui viennent des bordes latéraux dans la distribution de masse invariante du $J/\psi\phi$ et, comme échantillon de signal, les événements qui viennent d'une simulation réaliste Monte Carlo, avec la simulations detaillé du détecteur de CDF II basée sur GEANT [53] et reconstruits comme de vraies données. Le Monte Carlo a été produit en utilisant `BGenerator` [54] pour la génération d'événement, et `EvtGen` [55] pour les désintégrations. Nous employons les variables suivantes pour la formation du réseau de neurons:

- La probabilité d'ajustement du vertex χ^2 pour les particules B_s^0, J/ψ et ϕ.
- La masse reconstruite des particules vecteurs J/ψ et ϕ.
- L'impulsion transverse p_T du B_s^0, J/ψ et ϕ.
- L'impulsion transverse maximum et minimum du $\mathrm{K}^{+/-}$ pour chaque événement
- Les variables relatives a l'identification des particules.

Une fois exercíle réseau de neurons est faite, une décision doit être prise sur la coupure de la valeur de sortie. Nous utilisons comme critère la signification statistique, définie comme $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, et choisissons la valeur de coupure qui nous donnerait l'échantillon de données avec la plus grande signification statistique. Un graphique de la signification statistique contre la valeur du réseau de neurons est montré dans la Figure B.10.

Puisque nous voyons que la signification ne change pas trop pour une gamme de coupures entre $\{0.40, 0.65\}$, nous avons deux possibilités : nous pouvons choisir une coupure plus stricte pour réduire le signal et le fond, ou une coupure plus lâche pour augmenter le signal et le fond. Motivé par le fait que notre fond est bien modelisé pour la vraisemblance d'ajustement, nous choisissons la coupure plus lâche pour augmenter la statistique et pour améliorer la sensibilité. Nous obtenons approximativement $2,019 \pm 74$ événements $\mathrm{B}_s^0 \to J/\psi\phi$ signal. La distribution de masse invariante sur les données après la sélection finale du réseau de neurons est montrée dans la Figure B.10.

Eléments pour L'Analyse Angulaire des Désintégrations $\mathrm{B}_s^0 \to J/\psi\phi$

Temps de Vie Propre

Comme dans le cas de l'analyse de mélange du B_s^0, nous étudions le comportement de la résolution du temps de vie propre en fonction de différentes variables de

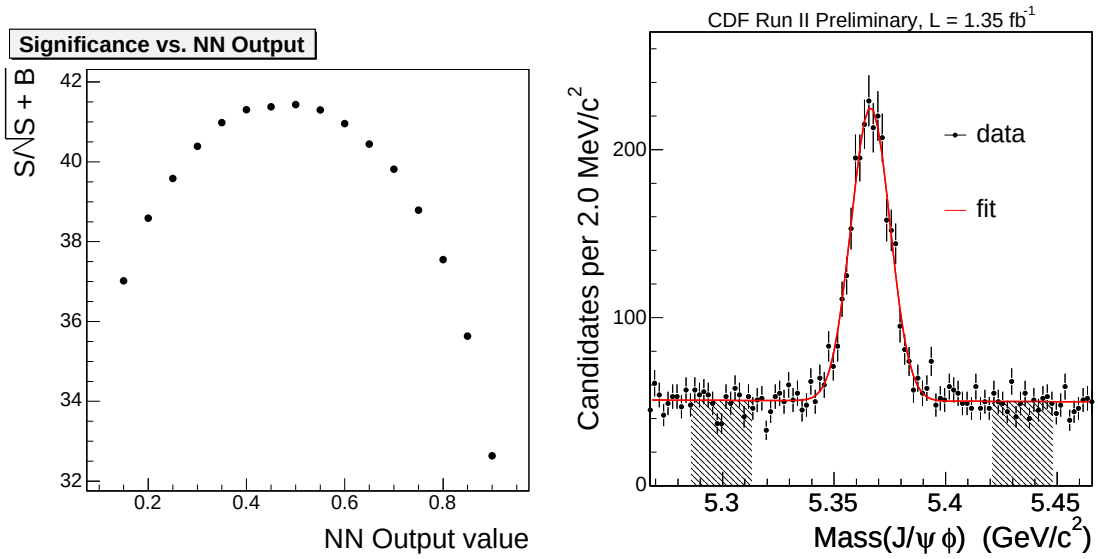

FIGURE B.10: Gauche: Signification contre le rendement de réseau de neurons. Droite: Distribution de masse invariante après la coupure du réseau de neurons à 0.40 pour les événements $B_s^0 \to J/\psi\phi$. Les régions à tiret correspondent aux bandes latérales de masse invariante utilisées pour la formation du réseau de neurons, $m(J/\psi\phi) \in [5.1280, 5.2142] \cup [5.3430, 5.3752]$.

désintégration. Nous répétons la même procédure, en paramétrisant la résolution σ_{ct} que, plus tard, nous voulons employer dans l'analyse angulaire finale sur les désintégrations $B_s^0 \to J/\psi\phi$. Afin d'étudier les effets d'une paramétrisation pour la résolution, il est nécessaire d'avoir un échantillon qui soit compatible comme provenant du *beamspot*, que nous pouvons employer pour la calibration. L'enjeu majeur dans le chois d'un tel échantillon c'est d'avoir des événements qui soient cinématiquement très semblables au vrai signal et de nature *prompt*. Les événements des bords latéraux de la distribution de masse invariante sont un bon point de départ, mais sont trop peu nombreux. Cependant, ces sont des événements qui, pour la plupart, ressemblent cinématiquement à de véritables événements du signal. Nous avons essayé d'augmenter au maximum la statistique en employant des événements dans les bords latéraux de la distribution de masse, mais en supprimant le critère de sélection de réseau de neurons et en employant seulement des conditions de pré-sélection. Nous avons constaté que le rendement de cette calibration était sous-optimal (le comportement de l'échantillon avec les coupures strictes était sensiblement différent de celui de l'échantillon avec des coupures de pré-sélection). Nous avons utilisé des événements des bords latéraux avec une condition de sélection lâche sur le réseau de neurons ($nnOut > 0.05$). Ceci nous donne un échantillon raisonnablement grand d'événements ($\simeq$ 50k) qui ont la caractéristique d'être *prompt*s. La qualité de la calibration est testée puis transférée de la sélection lâche à la sélection stricte sur le réseau de neurons utilisé comme selection finale pour l'échantillon $B_s^0 \to J/\psi\phi$. Pour examiner si nous obtenons des améliorations à l'ajustement final, nous appliquons la calibration de la résolution

en temps de désintégration aux données. Les ajustements sont définis comme "aveugle", pour signifier qu'un décalage inconnu a été appliqué aux résultats d'ajustement pour la différence des largeurs de désintégration, $\Delta\Gamma_s$, et pour la phase de violation CP β_s. Cette methode préserve l'incertitude sur les paramètres de l'ajustement, mais empêche d'être biaisé par un regard sur les données. Si la calibration complete est employé, on n'observe pas l'amélioration prévue sur les incertitudes pour les quantités d'intérêt. De plus le minimum de la fonction de vraisemblance ne s'accentue pas, et les incertitudes sur les paramètres d'intérêt augmentent. Nous concluons que, dans l'ajustement pour les paramètres de violation CP, nous n'observons pas une amélioration significative de l'analyse en utilisant la paramétrisation de résolution du temps de désintégration sous aucune des formes examinées jusqu'ici. La performance de l'analyse change seulement de pas un minimale, mais pour tous les paramètres relevants impliqués, le résultat est toujours inférieur à celui de l'échantillon non-calibré; c'est la raison pourqoi la calibration de σ_{ct} n'est pas utilisée.

Algorithmes d' Etiquetage de la Saveur

Nous savons que l'étiquetage de la saveur de production des mésons B dans des collisions $p\bar{p}$ est, en général, une tâche difficile. Cela ci motive la grande variété des algorithmes d'étiquetage du côté opposé développés pour l'analyse de mélange des B_s^0: *Soft Muon Tagger* (SMT), *Soft Electron Tagger* (SET), *Jet Charge Tagger* (JQT) et *Opposite Side Kaon Tagger* (OSKT). Toutes les information disponibles ont été combinées avec un réseau de neurons obtenant un seul algorithme d' étiquetage, *Combined Opposite Side Tagger* (COST), qui produit une décision si au moins un des taggers fournit des informations de saveur. Le but de combiner l'information de divers taggers est d'exploiter des corrélations entre les taggers, ce qui permet d'améliorer l'efficacité de l'étiquetage (ϵD^2) approximativement de 20% par rapport à une combinaison exclusive. Pour évaluer les performances du COST dans l'analyse angulaire $B_s^0 \to J/\psi\phi$, nous employons l'échantillon statistiquement plus grand de données $B^+ \to J/\psi K^+$ et $B^0 \to J/\psi K^{*0}$. Tout en estimant les exécutions de COST, nous observons que la dépendance de la dilution mesurée en fonction de la dilution événement-par-événement (qui vient de la calibration des algorithmes d'étiquetage de saveur) n'a pas le comportement linéaire désiré, si nous analysons séparément la matière, le B^+, et l'antimatière, B^-. D'une part, ce mauvais comportement est récupéré quand nous répétons l'exercice pour l'échantillon entier $B^\pm$ (voir la Figure B.11).

Par conséquent, l'utilisation du COST, dans une analyse où la séparation des différentes saveurs de particules est cruciale, pourrait présenter un biais dûe à cette asymétrie de la dilution. Nous avons laissé tomber COST pour l'analyse angulaire et avons combiné les taggers du côté opposé d'un mode hiérarchique plus classique. Le nouveau tagger, maintenant appelé *United Opposite Side Tagging* (UOST), a été de nouveau examiné pour des asymétries possibles en utilisant le mode de désintégration $B^\pm$, qui ne mélange pas matière et antimatière. La dilution mesurée par rapport à la dilution prédite, reportée sur la Figure B.12 avec les ajustements linéaires relatifs, montre le comportement linéaire prévu. Par conséquent, le choix définitif

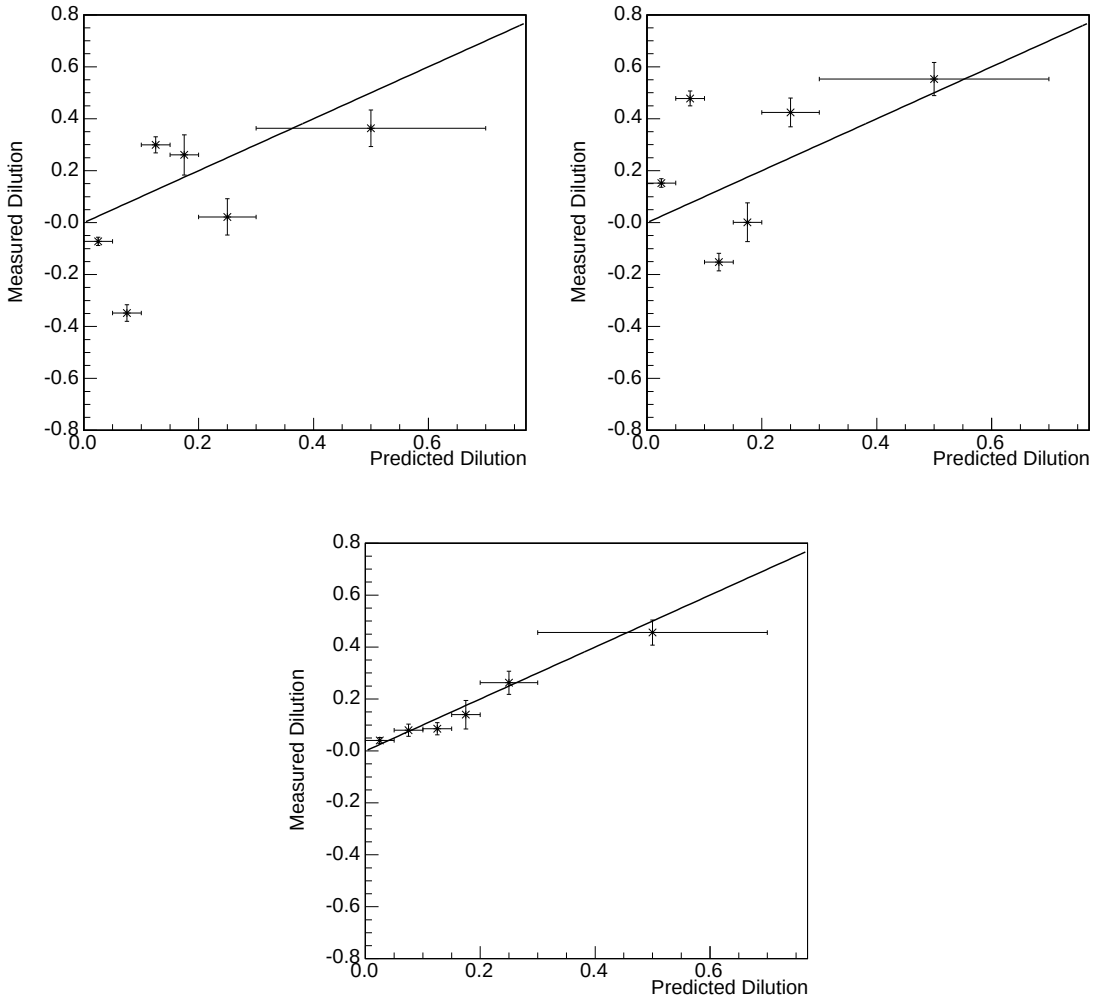

FIGURE B.11: Dilution mesurée pour le COST en fonction de la dilution événement-par-événement pour B^+ (en haut à gauche), B^- (en haut à droite) et l'échantillon entier $B^\pm$ (bas). Pour la référence, la fonction linéaire $y = x$ a été dessinée.

pour l'analyse angulaire de $\mathrm{B}_s^0 \to J/\psi\phi$ est d'utiliser l'UOST comme algorithme d'étiquetage du côté opposé. Pour le *Same Side Kaon Tagging* (SSKT) nous employons le tagger basé sur le ANN, qui combine l'information d'identification des particules et l'information cinématique. Les performances globales de l'UOST et de SSKT pour l'analyse $\mathrm{B}_s^0 \to J/\psi\phi$ sont rapportées dans le Tableau **??**.

Résultats de L'Analyse Angulaire des Désintégrations $\mathrm{B}_s^0 \to J/\psi\phi$

La deuxième partie de cette dissertation décrit l'analyse récente angulaire sur les désintégrations $\mathrm{B}_s^0 \to J/\psi\phi$ pour la mise à jour de la mesure de $\Delta\Gamma_s$ et la première

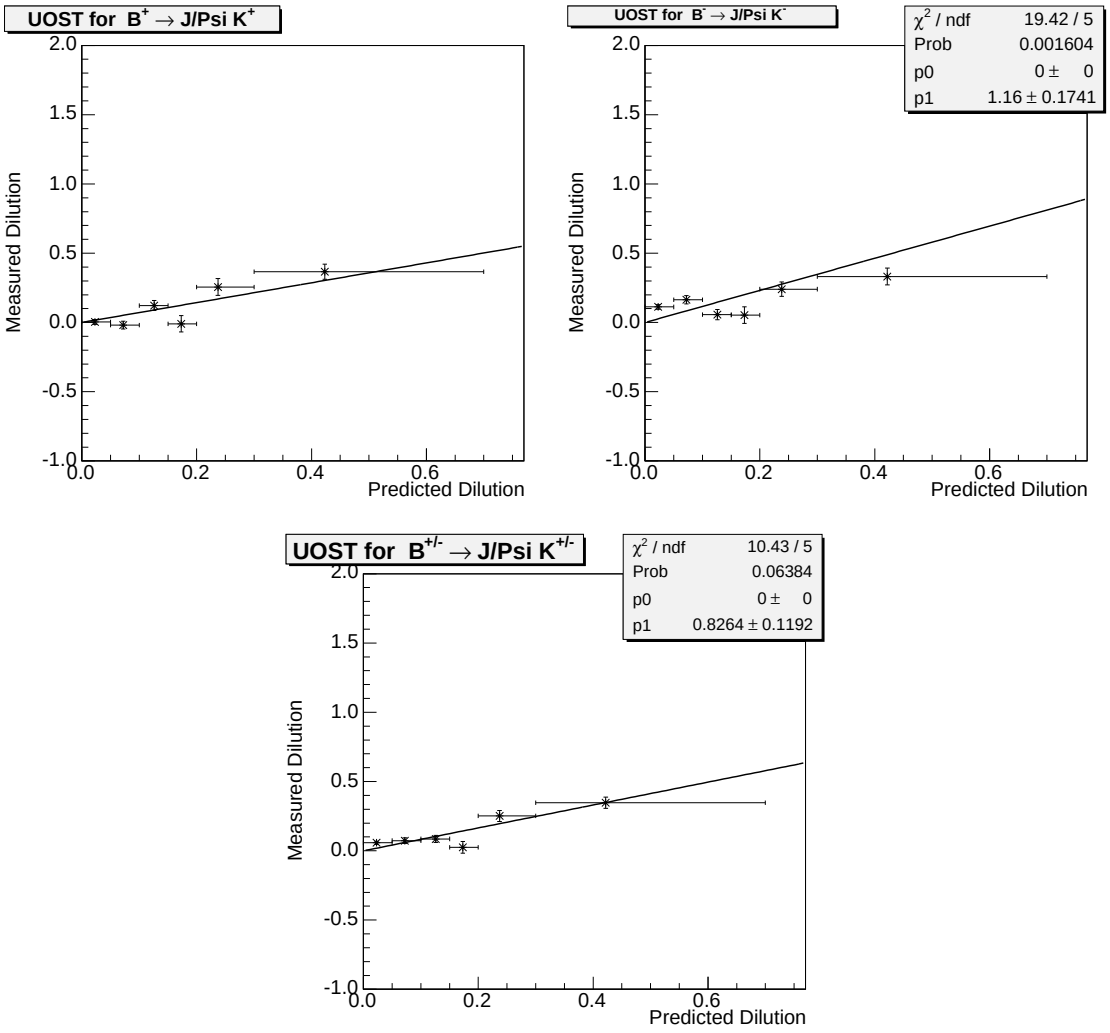

FIGURE B.12: Dilution mesurée en fonction de la dilution événement-par-événement pour B^+ (en haut à gauche), B^- (en haut à droite) et l'échantillon B^+entiers (en bas). L'ajustement relatif est dessiné.

	UOST	SSKT
ϵ (Data)	96 ± 1 %	50.0 %
$< \mathcal{D} >$(Data)	11 ± 2 %	27 ± 4 %
$\epsilon S_{\mathcal{D}} < \mathcal{D}^2 >$	1.2%	3.4 %

TABLE B.3: Rendements des algorithmes d'étiquetage de la saveur. Le facteur d'échelle $S_{\mathcal{D}}$ vient de la calibration des algorithmes d'étiquetage sur les échantillons de haute statistique, $S_{\mathcal{D}} = 0.853$ pour l'UOST, ou sur les échantillons Monte Carlo, $S_{\mathcal{D}} = 0.964$ pour le SSKT.

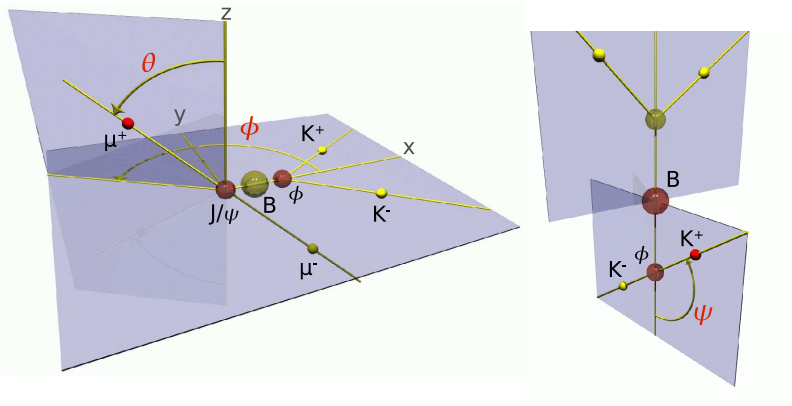

FIGURE B.13: Base de transversité et définition des angles dans le cas de la désintégration $B_s^0 \to J/\psi\phi$.

mesure de β_s en utilisant l'information des algorithmes d'étiquetage de la saveur. Cette phase décrit le temps du développement avec des termes proportionnels à la fois à $|\cos(2\beta_s)|$ et à $sin(2\beta_s)$. Les analyses de $B_s^0 \to J/\psi\phi$ qui n'emploient pas l'étiquetage de saveur sont principalement sensibles à $|\cos(2\beta_s)|$ et $|\sin(2\beta_s)|$, menant à une quadruple ambiguïté dans la fonction de vraisemblance. Le résultat est basé sur environ $2,000$ candidats signal reconstruits dans un échantillon de données de 1.35 fb^{-1} recuillé par le détecteur CDF II. Les moments angulaires orbitaux des mésons vecteurs (spin 1), J/ψ et ϕ, produits dans la désintégration du méson pseudoscalaire B_s^0 (spin 0), sont employé pour distinguer les états finaux impaires en CP, ondule-S et ondule-D, et paires en CP, ondule-P. Nous mesurons les angles θ_T, ϕ_T, ψ_T de désintégration (raison pour laquelle nous définissons cette analyse comme angulaire), définis dans la "base de transversité", pour étudier les corrélations angulaires parmi les particules d'état final dans les désintégration $P \to VV$. C'est un système de cordonnées approprié d'un point de vue expérimental et théorique, pour simplifier le calcul du taux différentiel pour les désintégrations $B_s^0 \to J/\psi\phi$, et il est la base pour la construction de notre fonction de vraisemblance. Les trois angles θ_T,ϕ_T, ψ_T sont définis comme suit. Considérons la désintégration $B_s^0 \to J/\psi\phi$, avec $Jpsi \to \mu^+\mu^-$ et $\phi \to K^+K^-$. Dans le système de référence de $Jpsi$, la direction de vol de ϕ définit l'axe x, alors que le plan du système de K^+K^- définit le plan $x-y$ avec l'axe de y orienté tels que $p_y(K^+) > 0$. En adoptant un système *right-handed*, l'ambiguïté dans le choix de l'axe z est résolue. L'angle θ_T est défini comme l'angle entre la direction de vol du μ^+ et la direction positive de l'axe z. L'angle ϕ_T est l'angle entre l'axe des x et la projection du μ^+ sur le plan $x-y$. En conclusion, ψ_T est l'angle du K^+ dans le système de repos de ϕ relativement à la direction négative de J/ψ dans cet système de référence. Dans toute cette dissertation nous dénoterons comme $\vec{\omega} = \{\cos\theta_T, \phi_T, \cos\psi_T\}$ cet ensemble de trois variables angulaires.

Les amplitudes de polarisation linéaire transversales à $t = 0$, à $A_\parallel$ et $A_\perp$, correspondent à des états finaux CP-even et CP-odd, respectivement. L'amplitude de polarisation longitudinale A_0 correspond à un état final CP-even. Les amplitudes de polarisation sont exigées de respecter la condition $|A_0|2 + \left|A_\parallel\right|^2 + \left|A_\perp\right|^2 = 1$.

Un ajustement de maximum vraisemblance est effectué pour extraire les paramètres d'intérêt, $2\beta_s$ et $\Delta\Gamma_s$, plus les paramètres de *nuisance*, qui incluent la fraction f_s de signal, la largeur moyenne du B_s^0, $\Gamma_s = (\Gamma_s^L + \Gamma_s^H)/2$, la fréquence de mélange Δm_s, la valeur des amplitudes de polarisation $|A_0|^2$, $\left|A_\parallel\right|^2$, et $\left|A_\perp\right|^2$, et les phases fortes $\delta_\parallel = \arg(A_\parallel^* A_0)$ et $\delta_\perp = \arg(A_\perp^* A_0)$. L'ajustement emploie l'information sur la masse reconstruite m du candidat B_s^0, le temps de désintégration propre du candidat B_s^0, ct, et son incertitude σ_{ct}, les angles $\vec{\omega}$, et l'information des algorithmes d'étiquetage $\mathcal{D}$ et ξ. $\mathcal{D}$ est la dilution événement-par-événement et $\xi = -1, 0, +1$ est la décision d'étiquetage. $\xi = +1$ correspond à un candidat étiqueté comme B_s^0, $\xi = -1$ à un $\overline{B}_s^0$, et $xi = 0$ à un candidat sans étiquetage. La probabilité pour chaque événement est décrite en fonction de distributions de probabilité (PDF) du signal (P_s) et du fond (P_b) comme

$$\begin{aligned} f_s P_s(m) P_s(t, \vec{\omega}, \xi | \mathcal{D}, \sigma_{ct}) P_s(\sigma_{ct}) P_s(\mathcal{D}) + \\ (1 - f_s) P_b(m) P_b(t|\sigma_{ct}) P_b(\vec{\omega}) Pb(\sigma_{ct}) P_b(\mathcal{D}). \end{aligned} \tag{B.0.21}$$

La distribution de probabilté de la masse pour le signal $P_s(m)$ est paramétrisé avec une simple Gaussienne, tandis que la distribution de probabilté de la masse pour le fond, $P_b(m)$, est paramétrisé comme un polynôme de premier ordre. Les distributions de l'incertitude du temps de désintégration et de la dilution événement-par-événement sont observées être différentes entre le signal et le fond; c'est pourquoi nous incluons leur distribution de probabilté explicitement dans la fonction de vraisemblance. Les distributions de probabilité du signal, $P_s(\sigma_{ct})$ et $P_s(\mathcal{D})$, et du fond, $P_b(\sigma_{ct})$ et $P_b(\mathcal{D})$, sont déterminé à partir des distributions des donneés. La dépendance angulaire et en temps de la distribution de probabilité du signal, $P_s(t,\ \vec{\omega},\ \xi\ |\mathcal{D}, \sigma_{ct})$, pour une seul algorithme d'étiquetge de saveur peut être écrit en fonction de deux PDFs, P pour le B_s^0et $\bar{P}$ pour $\overline{B}_s^0$:

$$\begin{aligned} P_s(ct, \sigma_{ct}, \vec{\omega}, |\mathcal{D}, \xi) = \frac{1 + \xi\mathcal{D}}{2} \cdot P(ct, \vec{\omega}, |\sigma_{ct}) \cdot \epsilon(\vec{\omega}) \\ + \frac{1 - \xi\mathcal{D}}{2} \cdot \bar{P}(ct, \vec{\omega}, |\sigma_{ct}) \cdot \epsilon(\vec{\omega}), \end{aligned} \tag{B.0.22}$$

ce qui est prolongé de maniere triviale dans le cas de deux algorithmes d'étiquetage de saveur indépendants (OST et SSKT). Les effets dûes à l'acceptance du détecteur pour les distributions des angles dans la base de transversité, $\epsilon(\vec{\omega})$, sont modelisés avec des données $B_s^0 \to J/\psi\phi$ simulées. Des distributions à trois dimensions des angles de transversité sont employées pour déterminer $\epsilon(\vec{\omega})$, afin de tenir correctement compte de toutes les dépendances parmi les angles. Les probabilités angulaires et en temps du B_s^0 peuvent être exprimées comme

$$\begin{aligned}\frac{d^4P(t,\vec{\omega})}{dtd\vec{\omega}} \propto & \left|A_0\right|^2\mathcal{T}_+f_1(\vec{\omega}) + \left|A_\parallel\right|^2\mathcal{T}_+f_2(\vec{\omega}) \\ & + \left|A_\perp\right|^2\mathcal{T}_-f_3(\vec{\omega}) + \left|A_\parallel\right|\left|A_\perp\right|\mathcal{U}_+f_4(\vec{\omega}) \\ & + \left|A_0\right|\left|A_\parallel\right|\cos(\delta_\perp)\mathcal{T}_+f_5(\vec{\omega}) \\ & + \left|A_0\right|\left|A_\perp\right|\mathcal{V}_+f_6(\vec{\omega}),\end{aligned} \tag{B.0.23}$$

où les fonctions $f_1(\vec{\omega})...f_6(\vec{\omega})$ sont définies dans la Réf. [82]. La probabilité $\bar{P}$ pour $\overline{\mathrm{B}}{}^0_s$ peut être obtenue en substituant $\mathcal{U}_+ \to \mathcal{U}_-$ et $\mathcal{V}_+ \to \mathcal{V}_-$. Les termes dépendant du temps $\mathcal{T}_\pm$ est définie par

$$\begin{aligned}\mathcal{T}_\pm = e^{-\Gamma t}\ x\ [&\cosh(\Delta\Gamma t/2) \mp \cos(2\beta_s)\sinh(\Delta\Gamma\ t/2) \\ &\mp \eta\sin(2\beta_s)\sin(\Delta m_s\ t)],\end{aligned} \tag{B.0.24}$$

où $\eta = +1$ pour P et -1 pour $\bar{P}$. Les autres termes dépendant du temps $\mathcal{U}_\pm$ et $\mathcal{V}_\pm$ sont définis comme

$$\begin{aligned}\mathcal{U}_\pm = \pm e^{-\Gamma t}\ x\ [&\sin(\delta_\perp - \delta_\parallel)\cos(\Delta m_s\ t) \\ &- \cos(\delta_\perp - \delta_\parallel)\cos(2\beta_s)\sin(\Delta m_s\ t) \\ &\pm \cos(\delta_\perp - \delta_\parallel)\sin(2\beta_s)\sinh(\Delta\Gamma_s\ t/2),\end{aligned}$$

$$\begin{aligned}\mathcal{V}_\pm = \pm e^{-\Gamma t}\ x\ [&\sin(\delta_\perp)\cos(\Delta m_s\ t) \\ &- \cos(\delta_\perp)\cos(2\beta_s)\sin(\Delta m_s\ t) \\ &\pm \cos(\delta_\perp)\sin(2\beta_s)\sinh(\Delta\Gamma_s\ t/2).\end{aligned} \tag{B.0.25}$$

Ces relations supposent qu'il n'y a pas de violation de CP directe dans le système. La dépendence en temps est convolvuée avec une fonction Gaussienne avec σ_{ct} comme déviation standarde qui tiens compte de la résolution en temps du detector; σ_{ct} est ajusté par un facteur global de calibration déterminé à partir de l'ajustement basé sur des candidats de fond qui se désintégrent promptement. Nous modelisons les distributions de probabilité pour le temps de vie du fond, $P_b(ct|\sigma_{ct})$, avec une fonction de Dirac à $t = 0$ plus une fonction exponentielle négatif et deux fonctions exponentielles positives, qui sont convolvués avec la fonction Gaussienne de résolution. Les distributions de probabilité du fond angulaire sont factorisés par $P_b(\vec{\omega}) = Pb(\cos\theta_T)P_b(\phi_T)P_b(\cos\psi_T)$, et obtenus en utilisant des événements de bords latéraux de la distribution de masse invariante du B^0_s. La fréquence de mélange $\Delta m_s = 17.77 \pm 0.12\ \mathrm{ps}^{-1}$ est contrainte par l'ajustement dans les incertitudes expérimentales. Les incertitudes systématiques venant de l'alignement, des effets du détecteur, des distributions angulaires du fond, de la désintégration d'autres mésons B, de la modélisation du signal et du fond s'avèrent avoir un effet négligeable sur la détermination de $\Delta\Gamma_s$ et β_s par rapport aux incertitudes statistiques. Une symétrie exacte est présente dans la distribution probabilité du signal, qui est invariante par la transformation simultanée $\beta_s \to \frac{\pi}{2} - \beta_s, \Delta\Gamma_s \to -\Delta\Gamma_s, \delta_\parallel \to 2\pi - \delta_\parallel, \delta_\perp \to 2\pi - \delta_\perp$. La fonction de probabilité

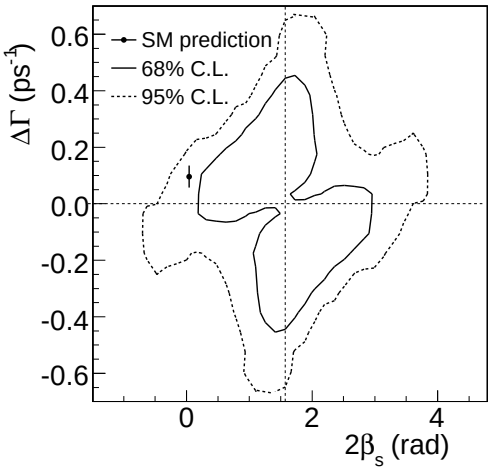

FIGURE B.14: A gauche: Le contours du niveau de confiance dans le plan $\beta_s - \Delta\Gamma_s$ basé sur les algorithmes d'étiquetage de la saveur. Le niveaux de confiance a 68% sont indiqués en trait plein, ceux à 95% sont en trait à tiret. La solution avec $\Delta\Gamma_s > 0$ correspond à $\cos(\delta_\perp) < 0$ et $\cos(\delta_\perp - \delta_\parallel) > 0$. Le contraire est vrai pour la solution avec $\Delta\Gamma_s < 0$. Le point noir correspond à la prediction du Modèle Standard, $\Delta\Gamma_s = 0.096$ ps^{-1} et $\beta_s = 0.02$ [75].

a deux minima. La symétrie peut être supprimée en contraignant l'un quelconque des paramètres que l'on vient de citer dans des limites appropriées. Cependant, même après la suppression de la symétrie exacte, des symétries approximées persistent, produisant des minimum locaux. Comme la fonction de vraisemblance est non-parabolique, nous ne pouvons pas de manière evidente donner des évaluations ponctuelles avec la statistique disponible. Nous présentons donc des régions de confiance pour les paramètres β_s, phase de violation de CP, et $\Delta\Gamma_s$, différence de largeur, pour la mesure des désintégrations $B_s^0 \to J/\psi\phi$ en utilisant l'étiquetage de saveur. La région de confiance est construite selon le critère de Feldman-Cousins avec l'inclusion rigoureuse des incertitudes systématiques. En fait, n'importe quel couple $\beta_s - \Delta\Gamma_s$ est exclu à un niveau de confiance (CL) donné seulement et seulement s'il peut être exclu pour *n'importe quel choix* de tous les autres paramètres d'ajustement, prélevé uniformément dans le domaina à ± 5 σ des valeurs déterminées avec les données. En supposent des valeurs de β_s et $\Delta\Gamma_s$ comme prédites par le Modèle Standard, c'est à dire $2\beta_s = 0.04$ rad et $\Delta\Gamma_s = 0.096$ ps^{-1} **??**, la probabilité d'une déviation comme observée dans les données est de 15%, ce qui correspond à déviation standarde gaussiene de 1.5σ. La Figure B.14 montre les régions de confiance obtenues selon le critère Feldman-Cousins à 68% et à 95% de niveau de confiance.

De plus, si $\Delta\Gamma_s$ est consideré comme un paramètre de nuisance, CDF trouve $2\beta_s \in [0.31, 2.82]$ rad pour un niveau de confiance de 68%. En exploitant l'information expérimentale et théorique actuelle, CDF extrait des limites plus strictes sur la phase β_s de violation CP. En imposant la contrainte $|\Gamma_{12}| = 0.048 \pm 0.018$ ps^{-1} [75] dans $\Delta\Gamma_s = 2|\Gamma_{12}|\cos(2\beta_s)$, on obtient $2\beta_s \in [0.24,\ 1.36] \cup [1.78,\ 2.90]$ rad avec 68% de degré de confiance. De plus, en contraignant les phases fortes $\delta_\parallel$ et $\delta_\perp$ avec les résultats

des usines à B sur $B^0 \to J/\psi K^{*0}$ [82] et la largeur moyenne du B_s^0 à la largeur de la moyenne B^0 mondiale [6], on trouve $2\beta_s \in [0.40,\ 1.20]$ rad au niveau de confiance de 68%.

La Collaboration DØ a rapidement fourni une analyse semblable qui a confirmé le résultats de CDF [83]. L'accord des analyses de $B_s^0 \to J/\psi\phi$ montre des fluctuations intéressantes dans la même direction à CDF et à DØ, et les analyses auront besoin certainement de poursuivre ce travail d'analyse pour une mise en évidence, qui serait possible en exploitant le plein échantillon de données fourni par le Tevatron, et si ces premières indications seront confirmées à l'avenir.

Contrairement au cas de l'etude du mélange du B_s^0, où ma contribution s'est porté essentiellement sur l'étiquetage de la saveur, j'ai apporté une contribution essentielle à cette deuxième phase de l'analyse de la violation de CP dans le secteur B_s^0 avec une partecipation importante à toutes les étapes nécessaires pour la mener à bonne fin. L'analyse a été exécutée sur un grand échantillon de modes de désintégrations du $B_s^0 \to J/\psi\phi$, sélectionné en utilisant un réseau de neurons. J'ai eu, ainsi, la possibilité d'apprendre les principes d'un réseau de neurons pour l'optimisation de la sélection. Avec l'expertise acquisé dans l'analyse de mélange du B_s^0, j'ai pris en charge la calibration des algorithmes d'étiquetage. En particulier, pour l'algorithme du côté opposé, nous avons rejeté l'algorithme basé sur le réseau de neurons et choisi plutôt une combinaison hiérarchique pour éviter une asymétrie observée entre la matière et l'antimatière. Le SSKT basé sur le réseau de neurons est hérité de l'analyse de mélange du B_s^0 et appliqué sans modification à la fonction de vraisemblance. La mesure initiale à été basée sur une analyse sans les algorithmes d'étiquetage (*untagged*) [74]; j'ai, à ce stade, apporté une contribution directe au développement de l'échantillon de calibration et de contrôle sur les désintégrations $B^0 \to J/\psi K^{*0}$. Les valeurs obtenues à partir de l'ajustement de $B^0 \to J/\psi K^{*0}$ sont résumées dans le Tableau B.4. Nous observons un niveau d'uncertitude statistique et systématique à CDF non seulement comparable mais même competitif avec les analyses plus récentes faites aux usines à B. Dans le Tableau B.4 les résultats de l'analyse de CDF sont comparés aux résultats de Babar et de Belle. Chacune des ces trois expériences a des résultats compatibles les uns avec les autres: le contrôle de notre système d'analyse au un échantillon cinématiquement équivalent $B^0 \to J/\psi K^{*0}$ de données en a prouvé la robustesse. Cela rend l'analyse angulaire faite sur l'échantillon de $B_s^0 \to J/\psi\phi$ d'autant plus solide.

Perspectives

La mesure précise de la fréquence d'oscillation Δm_s permet de vérifier l'hypothèse d'unitarité de la matrice CKM. Comme montré dans la détermination du rapport $|V_{ts}|^2/|V_{td}|^2$ (Equation B.0.20), ce test est maintenant limité seulement par les calculs théoriques. Bien que l'analyse pourrait être poursuivie en ajoutant plus de données ou en raffinant encore les outils techniques, elle n'améliorera pas spectaculairement

Par.	CDF	BaBar [87]	Belle [88]
$c\tau$	$456 \pm 6 \pm 6\mu m$	$- - -$	$- - -$
$\|A_0\|^2$	$+0.569 \pm 0.009 \pm 0.009$	$+0.556 \pm 0.009 \pm 0.010$	$+0.574 \pm 0.012 \pm 0.009$
$\|A_\parallel\|^2$	$+0.211 \pm 0.012 \pm 0.006$	$+0.211 \pm 0.010 \pm 0.006$	$+0.231 \pm 0.012 \pm 0.008$
$\delta_\parallel - \delta_0$	$-2.96 \pm 0.08 \pm 0.03$	$-0.293 \pm 0.08 \pm 0.04$	$+0.2887 \pm 0.090 \pm 0.008$
$\delta_\perp - \delta_0$	$+2.97 \pm 0.06 \pm 0.01$	$-0.291 \pm 0.05 \pm 0.03$	$+0.2938 \pm 0.064 \pm 0.010$

TABLE B.4: Les résultats finaux pour l'analyse angulaire $B^0 \to J/\psi K^{*0}$ exécutée à CDF. La première incertitude est statistique, alors que la seconde correspond à l'erreur systématique. Pour comparaison, dans la troisième et la quatrième colonne on montre les plus récents résultats des usines à B sur la mesure des amplitudes et des phases fortes.

la connaissance que nous avons déjà acquise. Au moins, jusqu'au moment où seront obtenus les calculs plus précis de QCD sur le réseau.

L'analyse angulaire du processus $B_s^0 \to J/\psi\phi$ est à sa première phase. Comme déjà observé, les résultats montrent une fluctuation intéressante sur la présence possible de la Nouvelle Physique au delà du Modèle Standard. Ceci doit être maintenant poursuivi. L'intérêt se situe également dans l'observation que d'autres mesures indépendantes, telles que l'analyse angulaire avec les algorithmes d'étiquetage de la saveur sur les désintégrations $B_s^0 \to J/\psi\phi$ faite par DØ [83] ou les mesures d'asymétrie en désintégrations semileptoniques du méson B_s^0 [91, 92, 93], pointent vers la même direction. En plus, la Collaboration Belle a récemment rapporté une mesure dans la différence de violation de CP directe entre les désintégrations chargées et neutres du méson B [94]. Belle a établi une anomalie au niveau de 4.4 σ de zéro dans la différence entre la mesure d'asymétrie directe de violation CP des désintégrations $B^+ \to K^+\pi^0$ et $B^0 \to K^+\pi^-$. Les deux asymétries étaient attendus être étroites et la différence presque nulle. Cette anomalie peut provenir de la même transition $b \to s$ impliquée dans le mode de désintégration du $B_s^0 \to J/\psi\phi$ analysé dans cette thèse. Ce résultat souligne l'importance de poursuivre cette analyse dans CDF. L'analyse angulaire du mode $B_s^0 \to J/\psi\phi$ présentée en cette dissertation exploite les données rassemblées avec le détecteur CDF II jusqu'à une luminosité intégrée de 1.35 fb^{-1}. Jusqu'ici, le Tevatron a excellents rendements et a recuillé déjà plus de 3 fb^{-1} des données. Bien que son fonctionnement soit dejà prolongé jusqu'à la fin de l'année budgétaire 2009, il est prévu de prolonger le fonctionnement du Tevatron d'une année additionnelle, c'est à dire jusqu'à fin 2010. Les expériences au Tevatron profiteraient certainement de cette quantité additionnelle de données. Selon les projections, CDF pourrait enregistrer, à la fin de sa dernière prise de données, une luminosité intégrée entre les 6 et 8 fb^{-1}. L'analyse bńéficiera des 4 fois plus de données que actuellement, dans le scenario le plus conservateur. Par conséquent, l'analyse angulaire sur le mode de désintégration $B_s^0 \to J/\psi\phi$, tout en étant une étape importante, est encore loin de sa conclusion. Dans nos études Monte Carlo réalistes on a observé la présence des

invariances locales en utilisant 6 fois la quantité de données de signal actuelle (voir la Section 5.7.4). Par conséquent, la fonction de vraisemblance pourrait continuer à manifester le même comportement non-gaussien à l'avenir. Dans la Figure B.15 nous montrons les probabilités d'observer un certain nombre de 3, 4, 5 sigmas à CDF, alors que β_s est hypothétiquement grand ($\beta_s = 0.4$ ce qui correspond à la moyenne de notre mesure actuelle) en fonction de la luminosité. Le même taux de données par luminosité et aucune amélioration de l'analyse sont supprimées.

Jusqu'ici, nous n'avons pas considéré des améliorations possibles pour l'analyse en discussion que juste l'addition de données où les seules limitations viennent de la nécessité de recalibrer les rendements du détecteur. Un exemple sur touts est le calibrage du dE/dx obtenu grace au système de trajectographie central (COT). En fait, après la luminosité intégrée de 1.35 fb^{-1} les deux premiers superlayer (voir le Chapter 2) ont été arrêtés, ne permettant pas l'utilisation directe du SSKT dans son exécution actuelle pour d'autres données. C'était la raison qui a limité l'utilisation de moins que moitié des données disponibles alors que l'analyse a été exécutée. A l'avenir, une autre manière possible de procéder serait de raffiner les algorithmes d'étiquetage, afin d'employer un *tagger* unique qui combine l'information d'OST et de SST avec un réseau de neurons. Ceci devrait fournir un meilleur rendement en termes d'efficacité et dilution aussi bien qu'expliquent les corrélations parmi les deux classes d'algorithmes. D'ailleurs, avec le Δm_s mesuré précisément, ca serait souhaitable de calibrer le SSKT directement sur des données, par exemple, en utilisant les désintégrations hadroniques du B_s^0 qui fournissent la puissance statistique la plus élevée dans l'analyse de mélange. Pendant la discussion des résultats finaux de $B_s^0 \to J/\psi\phi$ nous avons observé que l'impact principal des algorithmes d'étiquetage sur l'analyse n'était pas de réduire l'erreur par rapport à une analyse sans l'étiquetage de la saveur, mais de réduire plutt les ambiguïtés dans la fonction de vraisemblance. Par conséquent, une action plus efficace serait d'exploiter d'autres chemins de déclenchement pour obtenir des données additionnelles de signal, nécessaires pour augmenter la probabilité d'observer des effets significatifs dans la mesure de β_s (par exemple, voir les courbes inférieures dans Figure B.15). Une idée est d'utiliser les chemins du `TTT` trigger utilisés dans l'analyse de mélange: dans le Tableau B.5 nous montrons un exemple de distribution de masse invariante de $B_s^0 \to J/\psi\phi$. Nous avons employé les données rassemblées avec le `TTT` pour une luminosité intégrée totale de 640 pb^{-1}. L'échantillon de $B_s^0 \to J/\psi\phi$ a été rassemblé appliquant une sélection séquentielle simple, rapporté dans Tableau B.5, et enlevant les événements qui passe les sélections du `di-muon` trigger. Le rendement obtenu est de 296 ± 15 événements de signal et il déjà garantit, en ce moment, un additionnel $\simeq 30\%$ d'événements indépendants de signal. L'aspect intéressant de ces événements est leur fond réduit par rapport à de candidats reconstruits avec `di-muon` trigger. D'ailleurs la sélection peut être amélioré avec un réseau de neurons. L'inconvénient se situe dans la biais présente a cause de les coupures sur le temps de vie du système de déclenchement `TTT`. Ceci en effet complique une analyse dépendant du temps et le défi principal serait de rendre compte correctement de ce biais (par exemple pour le courbes d'efficacité comme dans l'Equation B.0.19). Néanmoins, ceci

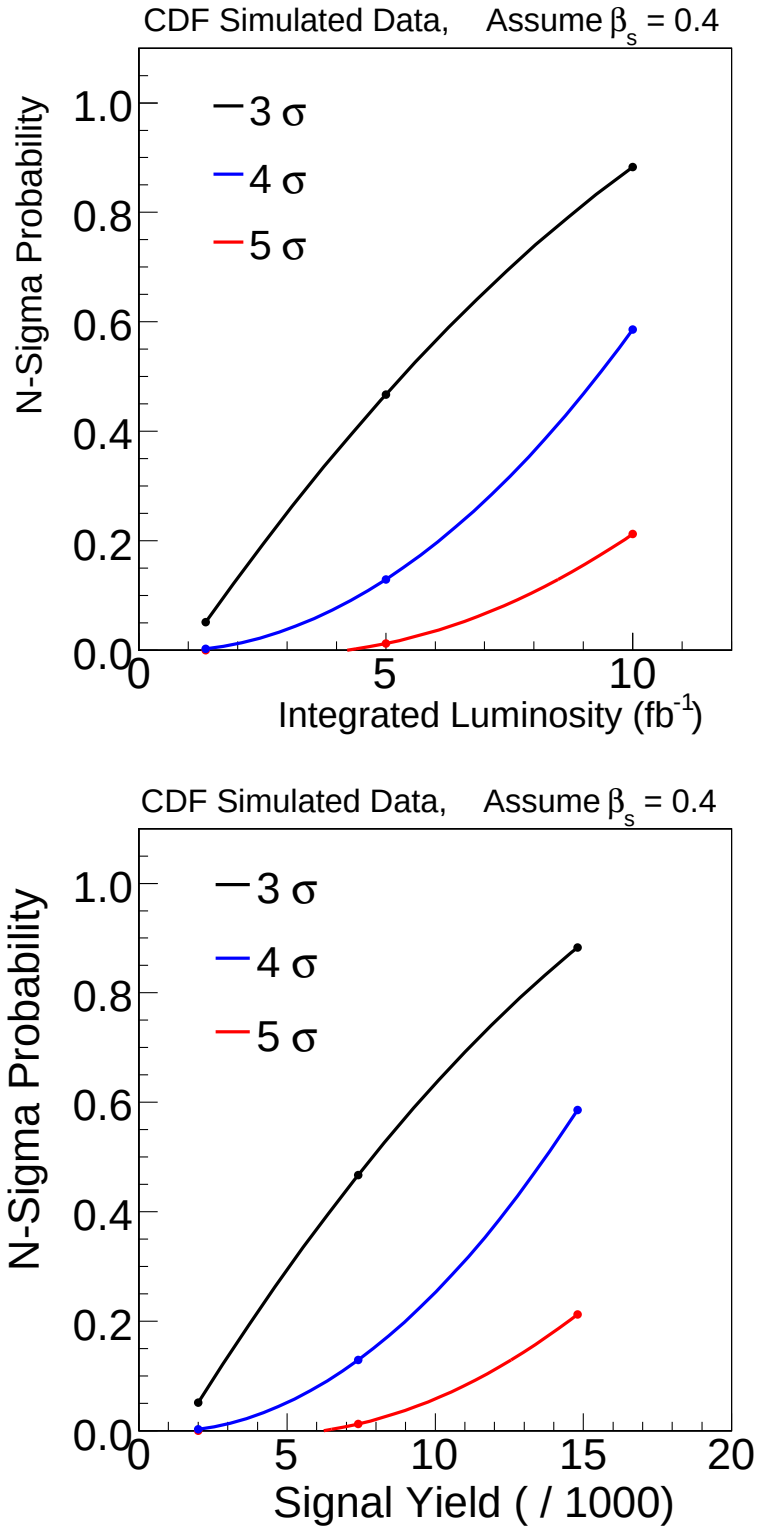

FIGURE B.15: La Probabilité d'observer l'effet à $3, 4, 5\ \sigma$, si $\beta_s = 0.4$, en fonction de la luminosité (en haut) et en fonction du taux de données (en bas). La courbe de probabilité N-sigma suppose la même quantité de données par luminosité et aucune amélioration de l'analyse.

offrirait une manière non-négligeable d'inclure de nouvelles données dans l'analyse.

Variable		Cut	Value
μ_{nSiPhi}		$\geq$	3
p_T^{B}	[GeV/c]	$>$	5.0
$L_{xy}/\sigma_{L_{xy}}(\mathrm{B}_s^0)$		$>$	4.9
$d_0(\mathrm{B}_s^0)$	cm	$<$	0.007
$p_T(\phi)$	[GeV/c]	$>$	1.4
$\chi^2_{r\phi}(\mathrm{B}_s^0)$		$<$	23
$p_T(\mu)$	[GeV/c]	$>$	0.450
σ_{ct}	$[\mu m]$	$<$	400
$\lvert m(\mu^+\mu^-) - M^{J/\psi}_{PDG}\rvert$	[MeV/c^2]	$<$	80
$\lvert m(K^+K^-) - M^{\phi}_{PDG}\rvert$	[MeV/c^2]	$<$	14

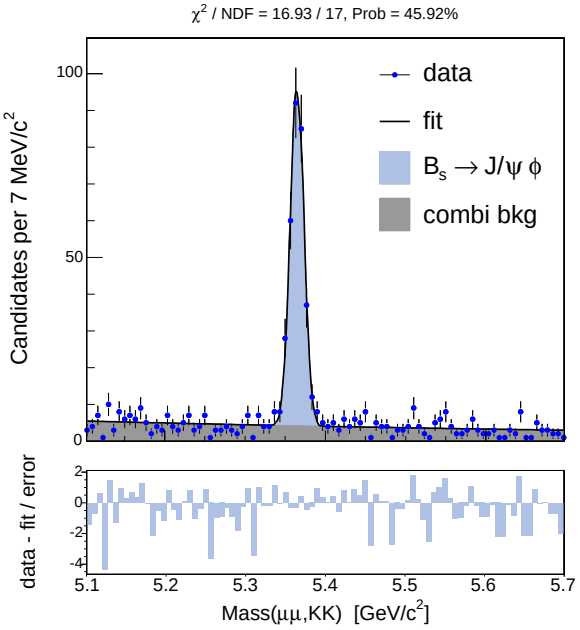

TABLE B.5: A gauche: Critères de sélection pour reconstruire des événements $\mathrm{B}_s^0 \to J/\psi\phi$ avec le déclenchement TTT. A droite: la distribution de masse invariante du $\mathrm{B}_s^0 \to J/\psi\phi$ pour des événements reconstruits avec le déclenchement TTT sur environ 640 pb^{-1} de données. Le rendement du signal obtenu est de 296 ± 15 événements.

En analysant les courbes dans Figure B.15 et les améliorations possibles, on se peut rendre compte que des temps intéressants attendent la physique du B_s^0 au Tevatron, particulièrement avec la perspective d'une synergie souhaitable entre les Collaborations CDF et DØ. Toute la communauté scientifique espère, à l'avenir, une stratégie commune pour obtenir les résultats finaux, par une combinaison cohérente sans contraintes externes. Les expériences au Tevatron offrent une grande opportunité d'étudier le secteur de saveur du B_s^0, importante pour les expériences qui vont démarrer au Large Hadronic Collider au CERN, et en particulier LHCb (selon les projections, LHC permettra à LHCb d'atteindre une résolution de $\sigma(\beta_s) = 0.023$ avec le premiers 2 fb^{-1} enregistrables pendant une année de prise de données).

Bibliography

[1] Lande, K. and Booth, E. T. and Impeduglia, J. and Lederman, L. M. and Chinowsky, W. *OBSERVATION OF LONGLIVED NEUTRAL V PARTICLES*. Phys. Rev., **103:1901–1904**, 1956.

[2] Aubert, B. and others. *Measurement of $B^0 - \overline{B}^0$ flavor oscillations in hadronic B^0 decays*. Phys. Rev. Lett., **88:221802**, hep-ex/0112044, 2002.

[3] Tomura, T. and others. *Measurement of the oscillation frequency for B0 - anti-B0 mixing using hadronic B0 decays. ((B))*. Phys. Lett., **B542:207–215**, hep-ex/0207022, 2002.

[4] Abulencia, A. and others. *Observation of B/s0 anti-B/s0 oscillations*. Phys. Rev. Lett., **97:242003**, hep-ex/0609040, 2006.

[5] Abazov, V. M. and others. *First direct two-sided bound on the B_s^0 oscillation frequency*. Phys. Rev. Lett., **97:021802**, hep-ex/0603029, 2006.

[6] Yao, W. -M. and others. *Review of particle physics*. J. Phys., **G33:1–1232**, 2006.

[7] Cabibbo, N. *Unitary Symmetry and Leptonic Decays*. Phys. Rev. Lett., **10:531–532**, 1963.

[8] Kobayashi, Makoto and Maskawa, Toshihide. *CP Violation in the Renormalizable Theory of Weak Interaction*. Prog. Theor. Phys., **49:652–657**, 1973.

[9] Wolfenstein, Lincoln. *Parametrization of the Kobayashi-Maskawa Matrix*. Phys. Rev. Lett., **51:1945**, 1983.

[10] Buras, A. J. and Slominski, W. and Steger, H. *B0 anti-B0 Mixing, CP Violation and the B Meson Decay*. Nucl. Phys., **B245:369**, 1984.

[11] Okamoto, Masataka. *Full determination of the CKM matrix using recent results from lattice QCD*. PoS, **LAT2005:013**, hep-lat/0510113, 2006.

[12] HEP2005 EPS Conference Webpage. `http://www.lip.pt/events/2005/hep2005/`.

[13] Charles, J. and others. *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories.* Eur. Phys. J., **C41:1–131**, hep-ph/0406184, 2005.

[14] Abulencia, A. and others. *Measurement of the B_s^0 - $\bar{B}_s^0$ Oscillation Frequency.* Phys. Rev. Lett., **97:062003**, hep-ex/0606027, 2006.

[15] Dunietz, Isard and Fleischer, Robert and Nierste, Ulrich. *In pursuit of new physics with B/s decays.* Phys. Rev., **D63:114015**, hep-ph/0012219, 2001.

[16] *Fermilab Linac upgrade conceptual design revision 4A.* FERMILAB-LU-CONCEPTUAL-DESIGN.

[17] Popovic, M. and Akenbrandt, C. *Performance and measurements of the Fermilab booster.* FERMILAB-CONF-98-175.

[18] Asher, T. *The Main Injector Rookie Book.* 2003.

[19] Morgan, J. *The Antiproton Source Rookie Book.* 1999.

[20] Marriner, John. *Stochastic cooling overview.* Nucl. Instrum. Meth., **A532:11–18**, physics/0308044, 2004.

[21] Jackson, Gerry. *The Fermilab recycler ring technical design report. Rev. 1.2.* FERMILAB-TM-1991.

[22] Group, TeVI. *Design Report Tevatron 1 project.* FERMILAB-DESIGN-1984-01.

[23] Blair, R. and others. *The CDF-II detector: Technical design report.* FERMILAB-PUB-96-390-E.

[24] Abe, F. and others. *The CDF detector: an overview.* Nucl. Instr. Meth., **A271:387–403**, 1988.

[25] Hill, Christopher S. *Operational experience and performance of the CDFII silicon detector.* Nucl. Instrum. Meth., **A530:1–6**, 2004.

[26] Azzi, P. *The CDF silicon detector upgrade.* Nucl. Instrum. Meth., **A419:532–537**, 1998.

[27] Sill, A. *CDF Run II silicon tracking projects.* Nucl. Instrum. Meth., **A447:1–8**, 2000.

[28] Affolder, Anthony A. and others. *Intermediate silicon layers detector for the CDF experiment.* Nucl. Instrum. Meth., **A453:84–88**, 2000.

[29] Affolder, Anthony A. and others. *CDF central outer tracker.* Nucl. Instrum. Meth., **A526:249–299**, 2004.

[30] Acosta, D. and others. *Measurement of prompt charm meson production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*. Phys. Rev. Lett., **91:241804**, hep-ex/0307080, 2003.

[31] Acosta, D. and others. *A time-of-flight detector in CDF-II*. Nucl. Instrum. Meth., **A518:605–608**, 2004.

[32] Balka, L. and others. *The CDF Central Electromagnetic Calorimeter*. Nucl. Instrum. Meth., **A267:272**, 1988.

[33] Gallinaro, Michele and others. *A new scintillator tile/fiber preshower detector for the CDF central calorimeter*. IEEE Trans. Nucl. Sci., **52:879–883**, physics/0411056, 2005.

[34] Bertolucci, S. and others. *The CDF Central and Endwall Hadron Calorimeter*. Nucl. Instrum. Meth., **A267:301**, 1988.

[35] Albrow, M. G. and others. *The CDF plug upgrade electromagnetic calorimeter: Test beam results*. Nucl. Instrum. Meth., **A480:524–546**, 2002.

[36] Apollinari, G. and Goulianos, Konstantin and Melese, P. and Lindgren, M. *Shower maximum detector for the CDF plug upgrade calorimeter*. Nucl. Instrum. Meth., **A412:515–526**, 1998.

[37] Ascoli, G. and others. *CDF Central Muon Detector*. Nucl. Instrum. Meth., **A268:33**, 1988.

[38] FERMILAB and others. *The 1992 CDF Muon System Upgrade*. 1994. CDF2858 public note.

[39] Ginsburg, C. M. *CDF Run 2 muon system*. Eur. Phys. J., **C33:s1002–s1004**, 2004.

[40] Thomson, Evelyn J. and others. *Online track processor for the CDF upgrade*. IEEE Trans. Nucl. Sci., **49:1063–1070**, 2002.

[41] Ashmanskas, Bill and others. *The CDF silicon vertex trigger*. Nucl. Instrum. Meth., **A518:532–536**, physics/0306169, 2004.

[42] Dell'Orso, M. *The CDF Silicon Vertex Trigger*. Nucl. Phys. Proc. Suppl., **156:139–142**, 2006.

[43] Belforte, S. and Dell'Orso, M. and Donati, S. and Gagliardi, G. and Galeotti, S. and Giannetti, P. and Labanca, N. and Morsani, F. and Passuello, D. and Punzi, G. and Ristori, L. and Sciacca, G. and Turini, N. and Zanetti, A. M. *Silicon Vertex Tracker Technical Design Report*. 1995. CDF3801 public note.

[44] Gomez-Ceballos, G. and others. *Event builder and Level 3 at the CDF experiment*. Nucl. Instrum. Meth., **A518:522–524**, 2004.

[45] Tiwari, Vivek. *Measurement of the B_s $\bar{B}_s$ oscillation frequency using semileptonic decays*. FERMILAB-THESIS-2007-09.

[46] Paus, Ch. and others. *BottomMods*. `http://cdfcodebrowser.fnal.gov/CdfCode/source/BottomMods/`.

[47] Marriner, J. *Secondary Vertex Fit with mass and pointing constraints (CTVMFT)*. 1993. CDF1996 public note.

[48] Giurgiu, Gavril A. *B Flavor Tagging Calibration and Search for B(s) Oscillations in Semileptonic Decays with the CDF Detector at Fermilab*. FERMILAB-THESIS-2005-41.

[49] A.Zell et al. *SNNS, Stuttgart Neural Network Simulation, User manual, Version 3.2*. (University of Stuttgart, Comp. Science Dept. Report No. 6/94, 1994) `http://wwww-ra.informatik.uni-tuebingen.de/SNNS/`.

[50] Brun, R. and others. *Root – an object-oriented data analysis framework*. `http://root.cern.ch`.

[51] Casal, B. and Ruiz, A. *Preliminary study of performances of Artificial Neural Networks in $B_d \to D^-[K^+\pi^-\pi^-]\pi^+$ process*. 2006. CDF8067 public note.

[52] Miles, J. *Observation of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ Oscillations Using Partially Reconstructed Hadronic B_s decays*. PhD thesis, Massachusetts Institute of Technology, 2007. FERMILAB-THESIS-2007-42.

[53] Brun, R. et al. CERN-DD-78-2REV (1978).

[54] Anikeev, K. and Murat, P. and Paus, Ch. *Description of Bgenerator II*. 1999. CDF5092 internal note.

[55] Lange, D. J. *The EvtGen particle decay simulation package*. Nucl. Instrum. Meth., **A462:152–155**, 2001.

[56] Moser, Hans Gunther. *The possibility to measure the time dependence of B(s)0 - anti-B(s)0 oscillations using Fourier analysis*. Nucl. Instrum. Meth., **A295:435–442**, 1990.

[57] Moser, H. G. and Roussarie, A. *Mathematical methods for B0 anti-B0 oscillation analyses*. Nucl. Instrum. Meth., **A384:491–505**, 1997.

[58] Eidelman, S. and others. *Review of particle physics*. Phys. Lett., **B592:1**, 2004.

[59] Sphicas, P. *Combining Flavor Taggers*. 1995. CDF3425 internal note.

[60] Lecci, Claudia. *A neural jet charge tagger for the measurement of the B/s0 anti-B/s0 oscillation frequency at CDF*. FERMILAB-THESIS-2005-89.

[61] Salamanna, Giuseppe. *Study of B/s mixing at the CDFII experiment with a newly developed opposite side b-flavour tagging algorithm using kaons.* FERMILAB-THESIS-2006-21.

[62] Usynin, Denys. *Study of charged particle species produced in association with anti-B0, B-, and anti-B/s0 mesons in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV.* FERMILAB-THESIS-2005-68.

[63] Kuhr, T. and others. *Combined Opposite Side Flavor Tagger.* 2006. CDF8460 public note.

[64] CDF B_s Mixing Group. *Updated Measurement of Δm_d and Calibration of Flavor Tagging in Fully Reconstructed Decays.* 2005. CDF7920 public note.

[65] Gronau, Michael and Rosner, Jonathan L. *Identification of neutral B mesons using correlated hadrons.* Phys. Rev., **D49:254–264**, hep-ph/9308371, 1994.

[66] Belloni, Alberto. *Observation of Bs0 - anti-Bs0 oscillations and the development and application of same-side-kaon flavor tagging.* FERMILAB-THESIS-2007-36.

[67] Menzemer, S. and others. *Same Side Kaon Tagging Studies in Fully Reconstructed Decays.* 2006. CDF7979 internal note.

[68] DELPHI Collaboration. *A Study of Excited b-Hadron Sates with the DELPHI Detector at LEP.* `http://delphiwww.cern.ch/pubxx/conferences/summer05/paper_lp56.ps.gz`.

[69] Leonardo, Nuno T. *Analysis of B/s flavor oscillations at CDF.* FERMILAB-THESIS-2006-18.

[70] Miles, Jeffrey Robert. *Observation of $\boldsymbol{B_s^0 - \overline{B}_s^0}$ Oscillations Using Partially Reconstructed Hadronic $\boldsymbol{B_s}$ Decays.* FERMILAB-THESIS-2007-42.

[71] Lyons, L. *Statistics for Particle and Nuclear Physicists.* Cambridge University Press, 1986.

[72] James, F. and Roos, M. *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations.* Comput. Phys. Commun., **10:343–367**, 1975.

[73] Punzi, Giovanni. *Comments on likelihood fits with variable resolution.* physics/0401045, 2004.

[74] Milnik, Michael. *Measurement of the lifetime difference and cp-violating phase in $B_s^0 \to J/\psi\phi$ decays.* FERMILAB-THESIS-2007-38.

[75] Alexander Lenz and Ulrich Nierste. *Theoretical update of Bs-Bs-bar mixing*, 2006.

[76] Acosta, D. and others. *Measurement of b hadron masses in exclusive J/ψ decays with the CDF detector*. Phys. Rev. Lett., **96:202001**, hep-ex/0508022, 2006.

[77] Abe, K. and others. *Observation of $b \to d\gamma$ and determination of $|V(td)/V(ts)|$*. Phys. Rev. Lett., **96:221601**, hep-ex/0506079, 2006.

[78] Aubert, B. and others. *Measurement of the branching fractions for the decays $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$*. hep-ex/0607099, 0700.

[79] Gary J. Feldman and Robert D. Cousins. *A Unified Approach to the Classical Statistical Analysis of Small Signals*. Physical Review D, **57:3873**, 1998.

[80] Aaltonen, T. and others. *Measurement of Lifetime and Decay-Width Difference in B0s → J/psi phi Decays*. Phys. Rev. Lett., **100:121803**, 0712.2348, 2008.

[81] Aaltonen, T. and others. *First Flavor-Tagged Determination of Bounds on Mixing- Induced CP Violation in Bs → J/psi phi Decays*. 0712.2397, 2007.

[82] Aubert, B. and others. *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**, hep-ex/0411016, 2005.

[83] Abazov, V. M. and others. *Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$*. 0802.2255, 2008.

[84] Bona, M. and others. *The UTfit collaboration report on the unitarity triangle beyond the standard model: Spring 2006*. Phys. Rev. Lett., **97:151803**, hep-ph/0605213, 2006.

[85] Bona, M. and others. *First Evidence of New Physics in $b \leftrightarrow s$ Transitions*. 0803.0659, 2008.

[86] Abazov, V. M. and others. *Lifetime difference and CP-violating phase in the B_s^0 system*. Phys. Rev. Lett., **98:121801**, hep-ex/0701012, 2007.

[87] Aubert, B. and others. *Measurement of decay amplitudes of B → J / psi K*, psi(2S K*, and chi(c1) K* with an angular analysis*. Phys. Rev., **D76:031102**, 0704.0522, 2007.

[88] Itoh, R. and others. *Studies of CP violation in B –¿ J/psi K* decays*. Phys. Rev. Lett., **95:091601**, hep-ex/0504030, 2005.

[89] Michael Feindt. *A Neural Bayesian Estimator for Conditional Probability Densities*, 2004. arXiv.org:physics/0402093.

[90] J. Boudreau, J. Cranshaw, L. Flores-Castillo, V. Papadimitriou, M. Spezziga, and K. Carrel. *Measurement of exclusive B lifetimes in the modes: $J/\psi K^+$, $J/\psi K^{*+}$, J/ψK^{*0} and $J/\psi K_s$.* 2004. CDF6387 internal note.

[91] Abazov, V. M. and others. *Measurement of the CP-violation parameter of B0 mixing and decay with p anti-p –¿ mu mu X data.* Phys. Rev., **D74:092001**, hep-ex/0609014, 2006.

[92] Abazov, V. M. and others. *Measurement of the charge asymmetry in semileptonic B_s decays.* Phys. Rev. Lett., **98:151801**, hep-ex/0701007, 2007.

[93] CDF collaboration. *Measurement of CP asymmetry in semileptonic* B *decays*, 2007.

[94] *Difference in direct charge-parity violation between charged and neutral B meson decays.* Nature, **452:332–335**, 2008.