Observation of B_s^0–$\overline{B}_s^0$ Oscillations and the Development and Application of Same-Side-Kaon Flavor Tagging

by

Alberto Belloni

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Author ....
Department of Physics
July 12, 2007

Certified by ....
Christoph M. E. Paus
Associate Professor of Physics
Thesis Supervisor

Accepted by ....
Thomas J. Greytak
Associate Department Head for Education, Professor

Observation of B_s^0–$\overline{B}_s^0$ Oscillations and the Development and Application of Same-Side-Kaon Flavor Tagging

Alberto Belloni

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

1861

Doctoral Dissertation

September 2007

Observation of B_s^0–$\overline{\mathrm{B}}_s^0$ Oscillations and the Development and Application of Same-Side-Kaon Flavor Tagging

by

Alberto Belloni

Submitted to the Department of Physics
on July 12, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

The search for flavor oscillations in the neutral $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ meson system constitutes a flagship analysis of the Tevatron proton–anti-proton collider physics program and an important probe for effects due to new physics beyond the Standard Model of particles and interactions. In particular, the precise measurement of a process such as B_s^0 oscillations sets a strong constraint on the parameters of quark-flavor mixing in the Standard Model.

This dissertation reports the analysis which, for the first time, observed $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations, using data collected with the CDF detector. The sub-percent precision of the measurement is also noteworthy. The data sample used for this analysis includes fully and partially reconstructed B_s^0 decays: $\mathrm{B}_s^0 \to \mathrm{D}_s^-\pi^+(\pi^-\pi^+)+c.c.$, $\mathrm{B}_s^0 \to \mathrm{D}_s^-\ell^+X+c.c.$, with $\mathrm{D}_s^- \to \phi^0\pi^-$, $\mathrm{K}^{*0}\mathrm{K}^-$, and $\pi^-\pi^+\pi^-$, and $\mathrm{B}_s^0 \to \mathrm{D}_s^-\rho^+$, $\mathrm{D}_s^{*-}\pi^+$, with $\mathrm{D}_s^- \to \phi^0\pi^-$.

The focus of this thesis is the algorithm of same-side-kaon tagging which provides a large fraction of the flavor-tagging power available to this analysis of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations. Flavor tagging consists in assessing whether a B_s meson is creates as a B_s^0 or a $\overline{\mathrm{B}}_s^0$ state, and constitutes an important ingredient in the analysis presented in this document. The algorithm which is here described combines particle identification information and kinematic characteristics of the B_s^0 event in an artificial neural network to provide improved tagging power $\epsilon\mathcal{D}^2$ of about 4.0 – 4.8%, depending on the data sample to which the algorithm is applied.

The search for B_s^0 oscillations is performed using an amplitude method based on a frequency scanning procedure. Applying a neural network-based combination of lepton, kaon and jet charge opposite-side tagging algorithms, with a total tagging power of $\epsilon\mathcal{D}^2 = 1.8\%$, and the same-side-kaon tagging algorithm to a data sample of 1 fb^{-1}, a signal of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations with a significance greater than 5 standard deviations is found. The oscillation frequency Δm_s is measured to be $17.77 \pm 0.10(stat) \pm 0.07(syst)\ \mathrm{ps}^{-1}$. The Δm_s measurement allows one to extract $|V_{td}/V_{ts}| = 0.2060 \pm 0.0007(exp)\ ^{+0.0081}_{-0.0060}(theor)$.

Thesis Supervisor: Christoph M. E. Paus
Title: Associate Professor of Physics

Contents

List of Figures

List of Tables

Introduction

The physics of the b quark represents one of the most lively research areas in high energy physics. Interest in the study of neutral meson oscillations has aroused since the first demonstration that matter-antimatter oscillations indeed take place in the $\mathrm{K}^0 - \overline{\mathrm{K}}^0$ system. The first evidence for neutral B meson oscillations was reported by the UA1 collaboration [1]. An anomalously large incidence of events with same-sign lepton pairs was found, possibly explained by B mixing. It is interesting to recall that prior to 1986 B^0 mixing was thought to be small, as the top quark was expected to be fairly light, in the $30 - 70$ GeV/c^2 range. Thus, the time-integrated measurement of an asymmetry presented by UA1 were interpreted as dominantly B_s^0 mixing, though $\mathrm{B}^0 - \mathrm{B}_s^0$ separation could not be established. In 1987, the Argus collaboration presented the first observation of B^0 oscillations in a time-integrated analysis [2]. This result was later confirmed by CLEO [3]. The evidence of slow $\mathrm{B}^0 - \overline{\mathrm{B}}^0$ oscillations, compared to the time scale of a B^0 decay, produced by the Argus, CLEO, and UA1 experiments required the top quark to be heavier than what was previously expected [4, 5]. The first time-dependent measurement of B^0 oscillations was performed in 1993 by the Aleph collaboration [6] and represents the first step in the technique that would be needed to perform the measurement of B_s^0 oscillations.

Altough the case of B_s^0 oscillations immediately became an important subject of research, all efforts for the B_s^0 system were baffled for almost twenty years. The nature of B_s^0 oscillations makes its study extremely challenging: the frequency of oscillations in the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ system is expected, within the Standard Model, to be so high that large samples and excellent tracking performance of a detector are required for a measurement. The importance of B_s^0 oscillations goes beyond the determination of a property of B_s^0 mesons. It provides a significant handle for testing the underlying model of flavor interactions and the possible presence of new physics.

The Collider Detector at Fermilab II (CDF II) is a general purpose detector which has been successfully collecting data for the last five years. It is installed at the Tevatron accelerator, which collides protons and anti-protons at a center-of-mass energy of about 2 TeV. This provides a unique environment for an immensely broad range of physics searches and measurements. A wealth of particles are produced and the most interesting heavy-quark states are currently available only at the Tevatron. The CDF II detector boasts an excellent tracking performance and a superlative trigger system allowing for the collection of data samples highly enriched in interesting physics content. These two aspects give CDF II an unequaled opportunity to perform a study of B_s^0 oscillations.

The organization of this document is as follows. The theoretical foundation of neutral B meson oscillations is presented in Chapter 1. The accelerator facilities and the CDF II detector are described in Chapter 2. The selection and reconstruction of B_s^0 candidates are contained in Chapter 3. The subsequent chapter concentrates on the presentation of an important technical aspects of the analysis: simulation of the data. The simulation of b events enters the analysis in many aspects. The Monte Carlo samples are described in Chapter 4. The ingredients and techniques for a mixing analysis are introduced in Chapter 5. The same-side tagger represents an important part of the analysis. Its development is presented Chapter 6. The final section, Chapter 7, contains the description of the likelihood fitting framework and the final results: the measurement of the B_s^0 oscillation frequency and the resulting constraints on the parameters of the flavor model.

Chapter 1

Theory and Measurement of Flavor Oscillations

The phenomenology of $B^0_s - \overline{B}^0_s$ *oscillations is presented in this chapter. The effects of the measurement of the oscillation frequency in the current model of particle physics are presented and its 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, $SU_{321} = SU(3) \otimes SU(2) \otimes U(1)$, 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:

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

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 gauge interactions in the SM. The right-handed counterpart of neutrinos ν is not

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 [7].
† Direct observation of top events.

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 [8, 9].

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 [10]. 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:

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

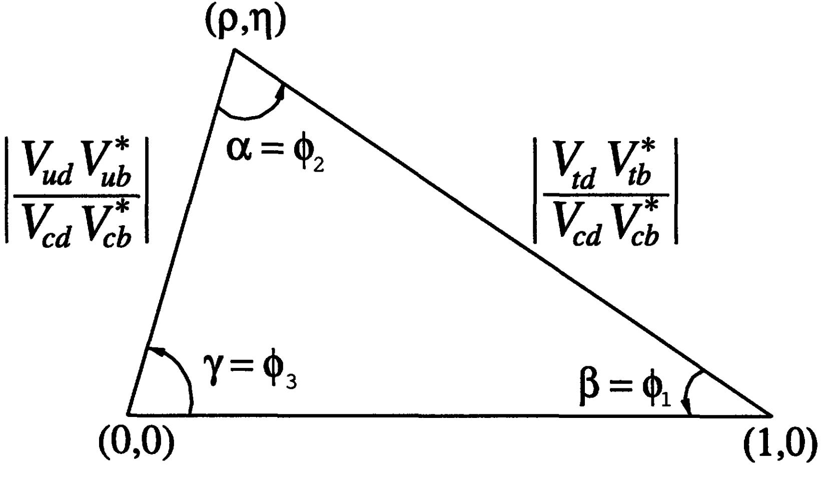

FIGURE 1.1: A sketch of the Unitarity Triangle.

Parameter	Value [7]
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.

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 presence of CP violating effects in the SM is indicated by any of the three angles being different from zero or π. 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.

1.2 The neutral B meson system

The description of the time evolution of neutral $B - \overline{B}$ systems is presented in this section. Starting from this section, the term "flavor" will be utilized to distinguish

Parameter	Value [7]
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.

particles from their respective antiparticles rather than among different quark types. For the sake of convenience, B and $\overline{\text{B}}$ will indicate the eigenstates of the strong interaction, i.e., the pure $\bar{b}q$ and $b\bar{q}$ states, while B_H and B_L will represent the mass eigenstates. Assuming that CP is a symmetry of the system, the latter will also have a definite CP-parity.

The Hamiltonian for free propagation, in the $\text{B} - \overline{\text{B}}$ basis, is expressed as follows in the Wigner-Weisskopf approximation [11, 12]:

$$\mathcal{H} = \begin{pmatrix} m & M_{12} \\ M_{12}^* & m \end{pmatrix} + \frac{i}{2} \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix} . \tag{1.2.1}$$

The diagonal elements of the Hamiltonian describe the mass and decay width of the flavor eigenstates. CPT invariance guarantees that the two eigenstates have the same mass m and decay width Γ, as indicated in Equation 1.2.1. 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). Bypassing the technical details of the diagonalization, the final results are:

$$|\text{B}\rangle = p|\text{B}_H\rangle + q|\text{B}_L\rangle , \tag{1.2.2}$$

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

where:

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} . \tag{1.2.4}$$

The time evolution of the B and $\overline{\text{B}}$ states is written as follows:

$$|\text{B}(t)\rangle = g_+(t)|\text{B}(0)\rangle + \frac{q}{p}g_-(t)|\overline{\text{B}}(0)\rangle , \tag{1.2.5}$$

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

where:

$$g_\pm(t) = \frac{1}{2}\left[e^{-\left(im_L + \frac{1}{2}\Gamma_L\right)t} \pm e^{-\left(im_H + \frac{1}{2}\Gamma_H\right)t}\right] . \tag{1.2.7}$$

It is of particular interest to determine the probability densities $\mathcal{P}_{\text{B}\to\overline{\text{B}}}(t)$ and $\mathcal{P}_{\text{B}\to\text{B}}(t)$ 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 $(\Gamma_L - \Gamma_H)/\Gamma$

is negligibly small, which hold to a good approximation for B^0 and B_s^0 mesons, the probability densities are given by:

$$\mathcal{P}_{B\to\bar{B}}(t) = \mathcal{P}_{\bar{B}\to B}(t) = \frac{\Gamma}{2}e^{-\Gamma t}\left[1-\cos\left(\Delta m t\right)\right], \tag{1.2.8}$$

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

where Γ is the inverse of the B_s^0 lifetime and Δm is equal to $m_H - m_L$. The expressions above are extremely useful. They are directly utilized in the fitting framework implemented in this thesis since they relate the parameter of interest, Δm_s, to the experimental observables, B flavor and decay time. Equation 1.2.8 describes the "mixed" case, where the B meson decays with the opposite flavor than the production flavor, while the "unmixed" case, where the B meson decays with the same flavor as at production, obeys to Equation 1.2.9.

The assumption that the eigenstates of the Hamiltonian for free propagation are also CP-eigenstates is justified in the case of B_s^0 mesons, where the CP-violating phase ϕ_s is expected to be very small. In the phase convention in which $V_{cb}V_{cs}^*$ is a real number, ϕ_s becomes equal to $\arg(M_{12})$. Introducing B_{even} and B_{odd}, eigenstates of the CP operator, the following relation is obtained:

$$|B_L\rangle = \frac{1+e^{i\phi_s}}{2}|B_{even}\rangle - \frac{1-e^{i\phi_s}}{2}|B_{odd}\rangle, \tag{1.2.10}$$

$$|B_H\rangle = -\frac{1-e^{i\phi_s}}{2}|B_{even}\rangle + \frac{1+e^{i\phi_s}}{2}|B_{odd}\rangle. \tag{1.2.11}$$

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 framework of the SM of electroweak interactions, neutral B meson mixing is described, at the lowest order, by the second order weak processes 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 [13]. 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:

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

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_s^0 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.

However, if the ratio between Δm_d and Δm_s is considered, most of the hadronic

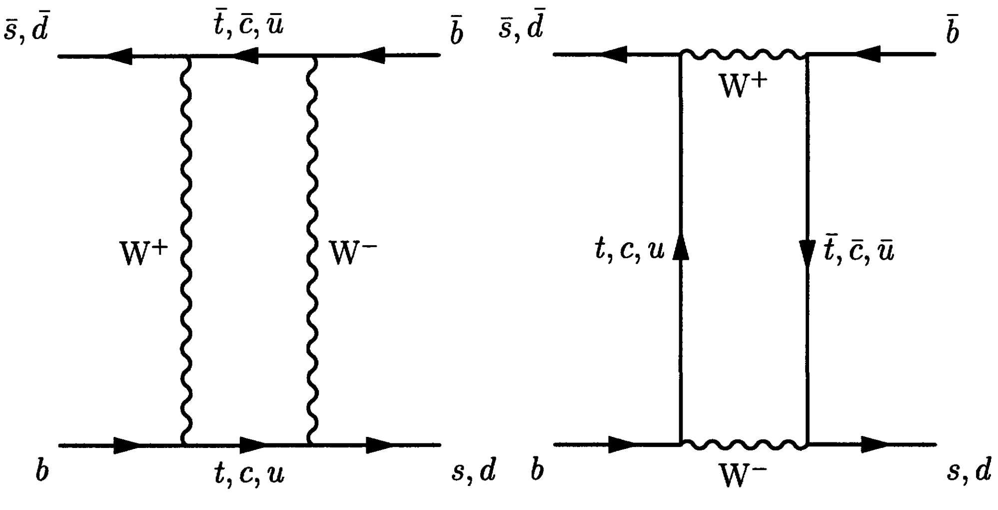

FIGURE 1.2: Lowest order diagrams for B mixing.

Parameter	Value [14]
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.

uncertainties that separately affect f_{B} and $\hat{B}$ cancel, and a more direct relation with elements of the CKM matrix is found:

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

where:

$$\xi = \frac{f_{\mathrm{B}^0_s}}{f_{\mathrm{B}^0}} \sqrt{\frac{\hat{B}_{\mathrm{B}^0_s}}{\hat{B}_{\mathrm{B}^0}}} = 1.210\ ^{+0.047}_{-0.039}\ [14]\,. \tag{1.3.3}$$

Thus, the measurement of the ratio $\Delta m_d/\Delta m_s$ allows for the precise estimation of $|V_{td}|/|V_{ts}|$, which, under the assumption that $|V_{cb}| = |V_{ts}|$[1], is equivalent to λ times the length of the side of the unitarity triangle opposed to the angle γ (or ϕ_3).

The status of the constraints in the $\bar{\rho}$–$\bar{\eta}$ plane as of the EPS 2005 [15] conference is illustrated by Figure 1.3. At that time, the combination of experimental inputs on Δm_s produced the 95% CL limit $\Delta m_s > 14.4\ \mathrm{ps}^{-1}$ which enters the picture. The fit by the CKM Fitter group [16] for Δm_s, which assumes the SM to be the fundamental theory and utilizes the currently available experimental results on CKM parameters, yields $\Delta m_s = 18.3\ ^{+6.5}_{-1.5}\ \mathrm{ps}^{-1}$.

The measurement of Δm_s provides a very stringent limit on the picture of flavor

---

[1] Next-to-leading order corrections in λ yield: $V_{cb} = A\lambda^2 + \mathcal{O}(\lambda^8)$ and $V_{ts} = -A\lambda^2 + A(1 - 2\rho)\lambda^4/2 - i\eta A\lambda^4 + \mathcal{O}(\lambda^6)$.

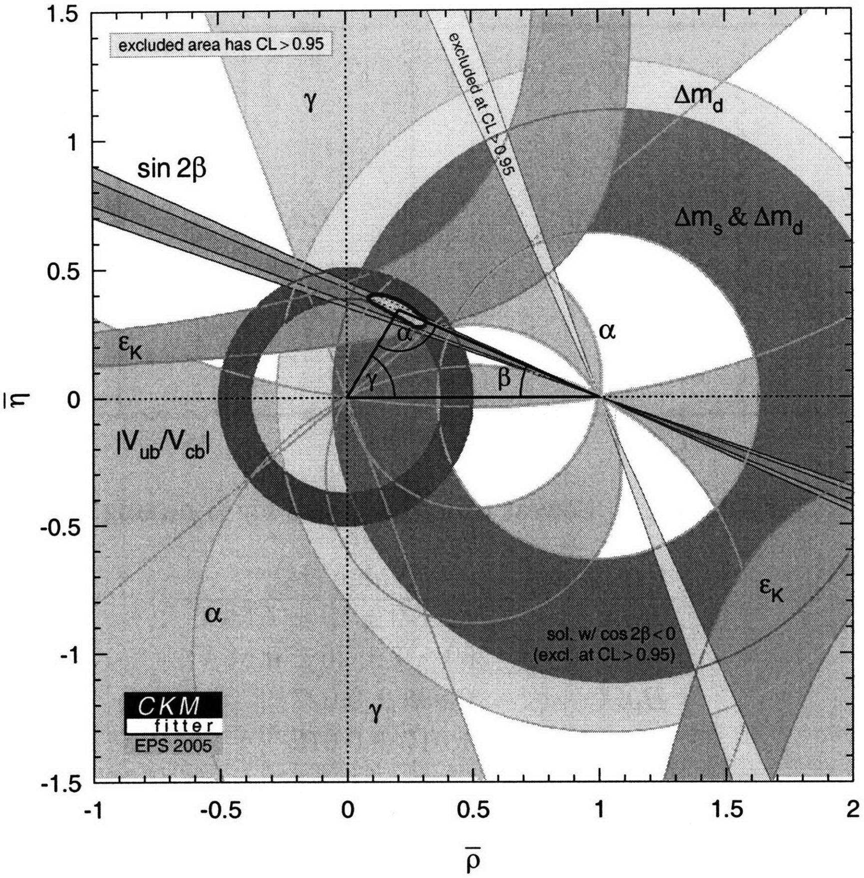

FIGURE 1.3: CKM fit [16] of the unitarity triangle presented at EPS2005, before the first measurement of Δm_s [17].

interactions in the SM, in which the CKM parameters are fundamental elements. Figure 1.3 shows that measurements overconstraining the CKM parameters start to strikingly limit the space for new contributions to flavor physics. The measurement of Δm_s plays an important rôle in the verification of the CKM picture by increasing the precision with which $\bar{\rho}$ and $\bar{\eta}$, the least known parameters of the CKM matrix, are determined. It is interesting to note that the measurement of the side of the unitarity triangle opposed to the angle γ (or ϕ_3) provides orthogonal information on $\bar{\rho}$ and $\bar{\eta}$ as opposed to the measurement of $\sin 2\beta$, which is precisely determined at B-factories.

1.4 Beyond the Standard Model

The measurement of the B_s^0 oscillation frequency provides a probe for New Physics (NP) beyond the SM. Extensions of the SM introduce additional fields which may mediate $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ transitions either directly or via box-diagrams similar to the ones shown in Figure 1.2. Existing experimental constraints anticipate tree-level contributions to $\mathrm{B}_s^0 \leftrightarrow \overline{\mathrm{B}}_s^0$ transitions to be small. In fact, the scale of NP is expected to be large compared to the mass of the W boson, with the exception of some specific models, such as supersymmetry with violation of R-parity. However, loop-mediated processes may provide large modifications to Δm_s.

Models which describe NP generally introduce new parameters, such as flavor changing couplings, short distance coefficients, matrix elements of new local operators, or CP violating phases. However, the mixing process is described by a single complex amplitude and is typically parameterized in terms of two parameters which quantify the difference of the complex amplitude with respect to that of the SM. The presence of NP modifies the SM contribution to the neutral B meson oscillation frequency Δm_q^{SM} as follows:

$$\Delta m_q^{NP} = \Delta m_q^{SM} |1 + h_q e^{i\sigma_q}|\,, \quad q = d \ or \ s\,, \tag{1.4.1}$$

where h_q and σ_q indicate the relative magnitude and phase of the NP contribution.

Two classes of SM extensions are identified by the modifications that they introduce to the effective Hamiltonian which describe $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ mixing. The phenomenology of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ mixing is described in terms of an effective Hamiltonian, the Operator Product Expansion [18] which is written as follows:

$$\mathcal{H}_{eff}^{\Delta B=2} = \frac{G_F^2 M_{\mathrm{W}}^2}{16\pi^2} \sum_i V_{CKM}^i C_i(\mu) Q_i\,, \tag{1.4.2}$$

where G_F is the Fermi constant, M_{W} the mass of the W boson, V_{CKM}^i the product of the appropriate CKM factors, $C(\mu)$ the Wilson coefficient, evaluated at the renormalization scale μ, and Q the local operator of the expansion.

In models with Minimal Flavor Violation (MFV [19, 20]), the source and strength of flavor violation is the CKM matrix only. No new operators are added to the expansion in Equation 1.4.2 and *non*-SM contributions come only from new particles circulating in the loop, which modify the Wilson coefficients. The principal consequence is *d–s* universality: the new Hamiltonian modifies both Δm_s and Δm_d of the same relative amount. The ratio between Δm_s and Δm_d is thus identical to the ratio in the SM and the constraint on the unitarity triangle is unchanged. This is the case for the Minimal Supersymmetric extension of the SM with flavor conservation (for example, Reference [21]), where the exchange of charged Higgs bosons takes place in the box diagram. In this model, the term $|1 + h_q \exp(2i\sigma_q)|$ is ≥ 1 and is expressed as a function of the masses of charginos $\widetilde{\chi}_2^{\pm}$, stop $\widetilde{t}$, charged Higgs bosons $H^{\pm}$ and $\tan\bar{\beta}$, the ratio of the vacuum-expectation-values of the Higgs bosons. The Feynman diagrams which involve the new particles are shown in Figure 1.4.

When additional operators Q are introduced in Equation 1.4.2, the effective Hamiltonian loses the *d–s* universality and the effect of NP is a change in Δm_s and Δm_d which do not preserve their ratio. The constraint on the unitarity triangle changes too and the position of the $(r\bar{h}o, \bar{\eta})$ apex is affected. The *d–s* universality is broken in models which change the structure of the CKM matrix, for example by adding a fourth generation of quarks or extra singlet quarks [22]. Generalized MFV models in which significant contributions from *non*-SM operators enter the effective Hamiltonian are presented in Reference [23].

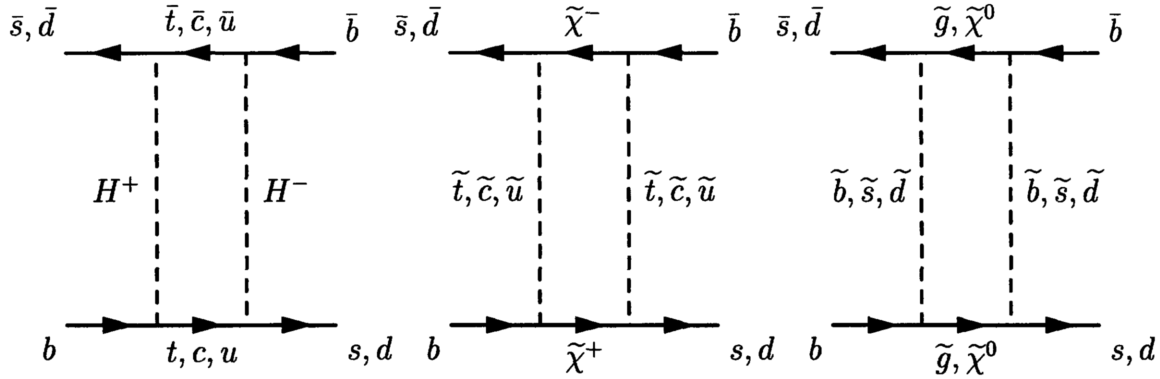

FIGURE 1.4: Contributions to B_s^0 mixing in MSSM with MFV.

The new precise measurement of Δm_s presented in this thesis reduces the available space for exotic extensions. The effect on some of the most interesting parameters of NP is presented in the last section of this document, Section 7.5.

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

In this section, an overview of the measurement of $B_s^0 - \overline{B}_s^0$ oscillations is presented, and some of the experimental issues are outlined.

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 (B^+, B^0, B_s^0, B_c^+, b-baryons) are produced, with ~ 10% of b quarks fragmenting into B_s^0 [7].

Figure 1.5 shows a stylized picture of a $p\bar{p}$ interaction in which a B is produced, and the subsequent decay of the B meson. The cartoon presents the steps of the analysis: reconstruction of a B candidate in a self-tagging final state (Section 3.4), measurement of proper decay-time (Section 5.2), and initial-state flavor tagging (Chapter 6 and Section 5.3). These three specific ingredients are briefly introduced below:

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. If the flavor of decay is different than (the same as) the flavor at production, the meson is classified as "mixed" ("unmixed").

3. **Proper decay-time:** the proper decay-time is the decay-time of the hadron in its rest frame. Since a B_s oscillates four times during its average lifetime ($\tau(B_s^0) = 1.466 \pm 0.059$ ps, or $c\tau(B_s^0) = 439.5 \pm 17.7$ μm [7]), the time dependent observation of $B_s^0 - \overline{B}_s^0$ oscillations requires excellent proper-time resolution. In principle, it is possible to perform a time-independent analysis, by measuring

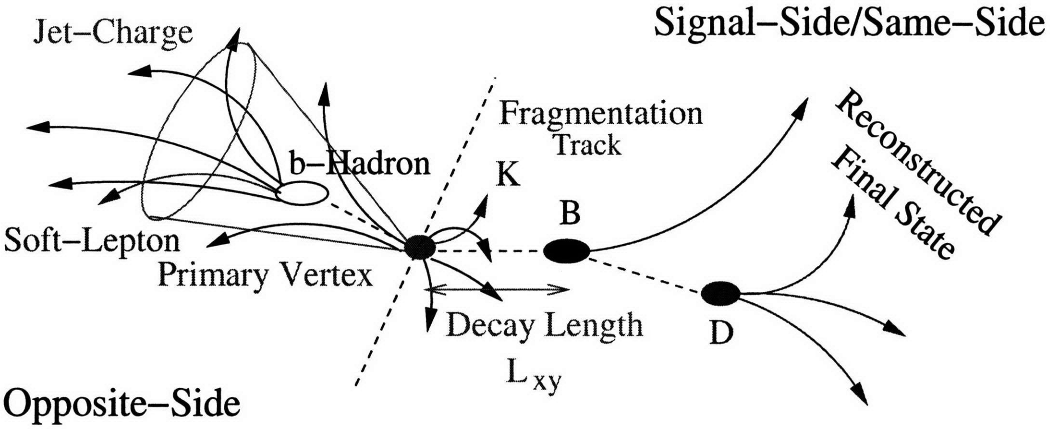

FIGURE 1.5: A schematic representation of a B event.

the residual $B_s^0 - \overline{B}_s^0$ asymmetry after integrating the dependence on the proper decay-time of a candidate. This approach is not feasible for an analysis of B_s^0 oscillations, because oscillations occur on a much shorter time-scale than B_s^0 decays. The residual asymmetry is thus so small that the sample of B_s^0 candidates currently available is not sufficient to measure it.

The first two items listed above refer to the flavor of the B meson at the time of production and decay. In this analysis, the flavor at the time of decay is known from the final state particles, since B_s^0 candidates are reconstructed in self-tagging final states. The reconstructed final states which enter this analysis are listed below (charge-conjugated modes are implied):

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

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

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

> 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 flavor of the B_s^0 candidate at decay 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 ascertain, and several techniques have been developed to perform this function. As shown in Figure 1.5, 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 in phase space. 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 perform the measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency, the oscillation probabilities in Equations 1.2.8 and 1.2.9 are mapped out as a function of decay time for at least a portion of the decay time spectrum. Because each of the three items listed above has experimental limitations, this analysis requires large samples of B_s decays with a good signal-to-noise.

The theoretical background to the phenomenon of $B_s^0 - \overline{B}_s^0$ *oscillations has been presented. The next chapter describes the Fermilab accelerator complex, and the CDF detector and trigger system.*

Chapter 2

Experimental apparatus

The focus of this chapter is on the accelerator complex at Fermilab and the CDF detector.

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 2 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 [7]:

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

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 $\sigma[\mathrm{cm}^2]$ and to $L[\mathrm{cm}^{-2}]$. 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, $N_{\bar{p}}$	$3 \cdot 10^{10}$
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}.

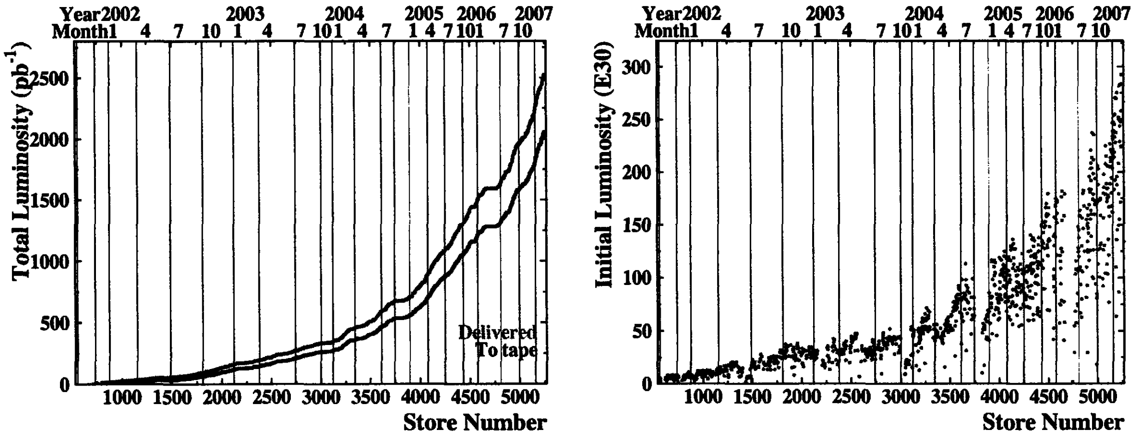

FIGURE 2.1: Integrated (left) and peak luminosity (right) delivered by the Tevatron. The plot covers the period between the beginning of 2002 and of 2007. The luminosity is shown as a function of store number.

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. For example, the total single B-hadron production cross section in a $p\bar{p}$ interaction, for the rapidity range $|y| < 0.6$ (defined in Equation 2.2.1), is $17.6 \pm 0.4(stat)\ ^{+2.5}_{-2.3}(syst)$ μb, as measured at the Tevatron [24]. Figure 2.1 shows the total integrated luminosity up to February 2007 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.

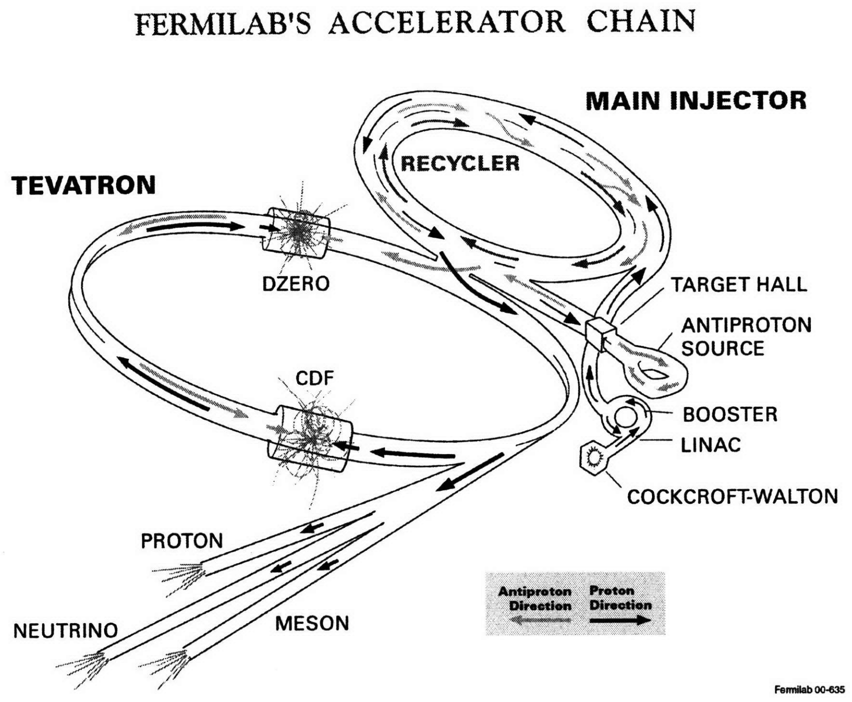

FIGURE 2.2: Accelerator complex at Fermilab.

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 [25]). 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 [26], 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 [27]) is a synchrotron with a circumference of 3319 m. It has the fundamental rôle of optimally connecting the Booster to the Tevatron. During the first run of the CDF detector, before the upgrades in the detector and the accelerator setup which started the Run II period of data-taking, the *Main Ring*, located in the same tunnel as the Tevatron, performed this duty. The setup needed an upgrade to overcome some restraints which limit the luminosity. Firstly, the Main Ring is not capable of accepting the protons that can be provided at injection by the Booster for the simple reason that the aperture of the Main Ring is significantly smaller than the beam delivered from the Booster at full intensity.

Secondly, the Main Ring shares the same tunnel with the Tevatron collider and this introduces additional backgrounds during data-taking.

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 [28] 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 [29] 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 [30]) 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.

2.1.5 Tevatron

The Tevatron collider [31] 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 B0 and D0. The locations are currently utilized by the CDF and D0 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

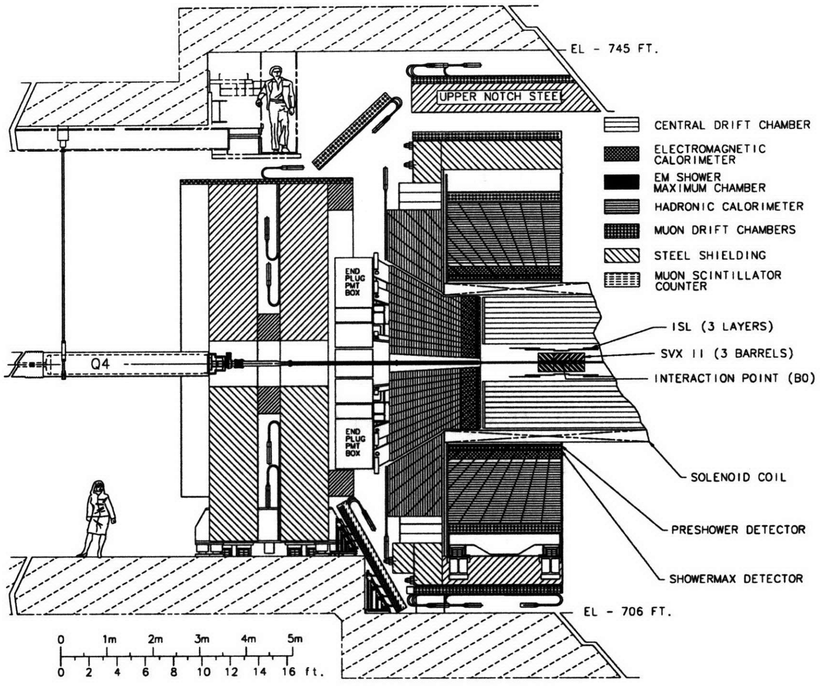

FIGURE 2.3: A cross-sectional view of one half of the CDF II detector.

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 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.

On a side note, the often mentioned point of closest approach between the trajectory and the z axis can belong to an extrapolation of the segment of helix which is reconstructed by the tracking system.

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 [32] and [33], and in specific references for each subdetector. The description of the trigger, tracking and TOF systems are emphasized, because they represent the aspects of the detector more critical to the analysis presented in this document.

2.2.1 Layer00

Layer00, L00 [34], 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,

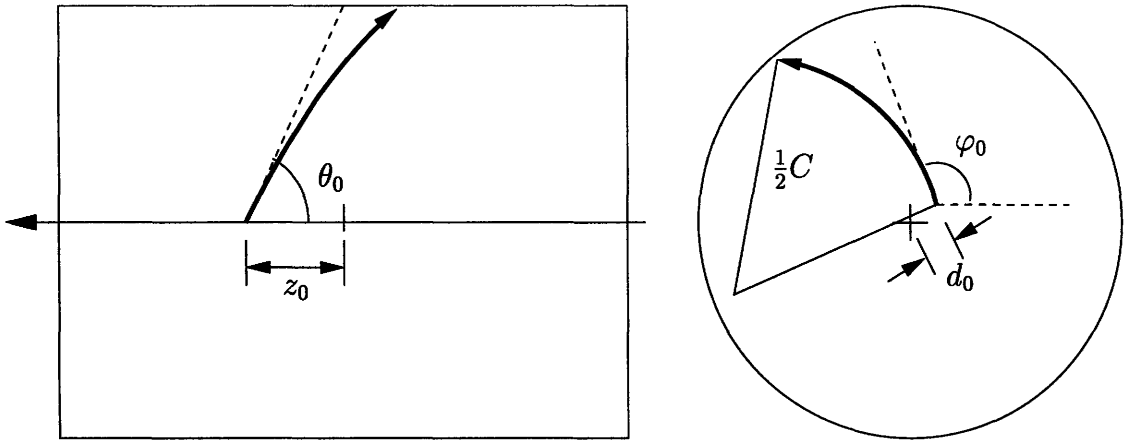

FIGURE 2.4: A pictorial representation of the parameters chosen to describe tracks in the CDF II detector, in the longitudinal, z-y (left), and transverse, x-y or ρ-φ (right), planes. The z axis is oriented in the right-left direction in the longitudinal view, while is indicated by the cross in the middle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

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 [35, 36], 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 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

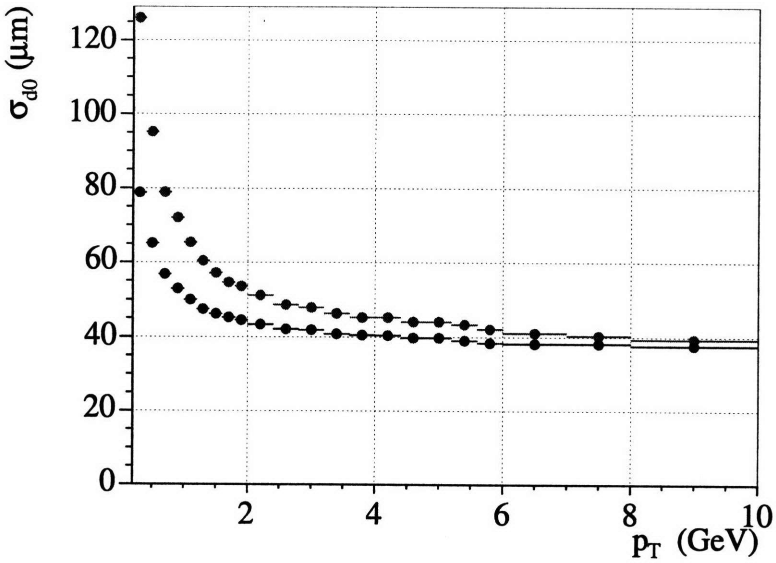

FIGURE 2.5: Impact parameter resolution of tracks with L00 hits (blue/dark) and without L00 hits (red/light), as a function of the transverse momentum of the tracks.

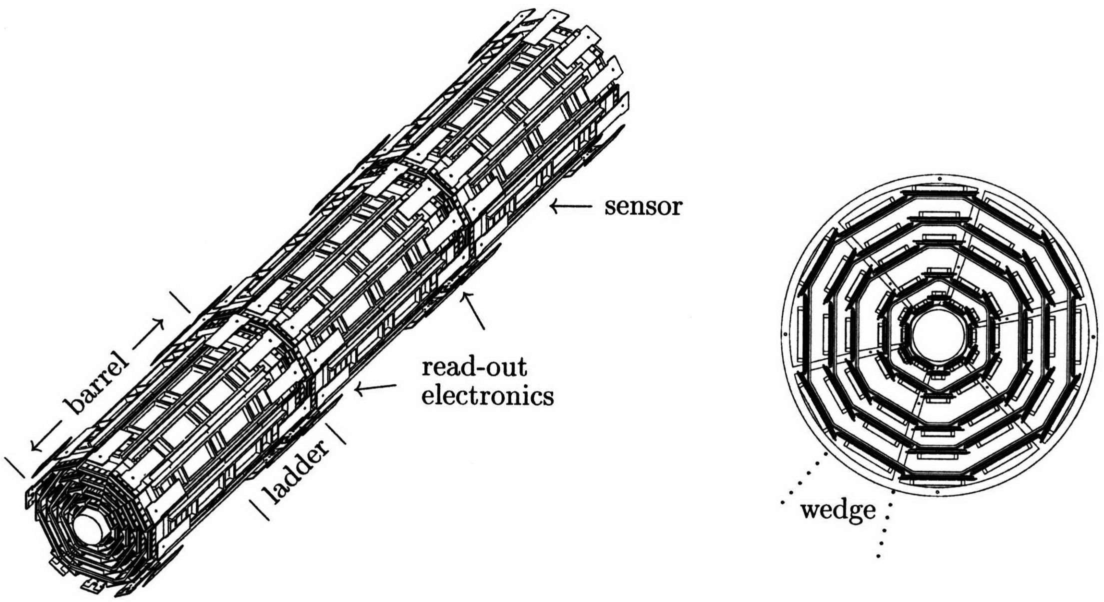

FIGURE 2.6: The Silicon VerteX detector SVXII. An illustration of the three instrumented barrels of SVXII (left) and a cross-sectional view of a barrel in the r–φ plane.

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.

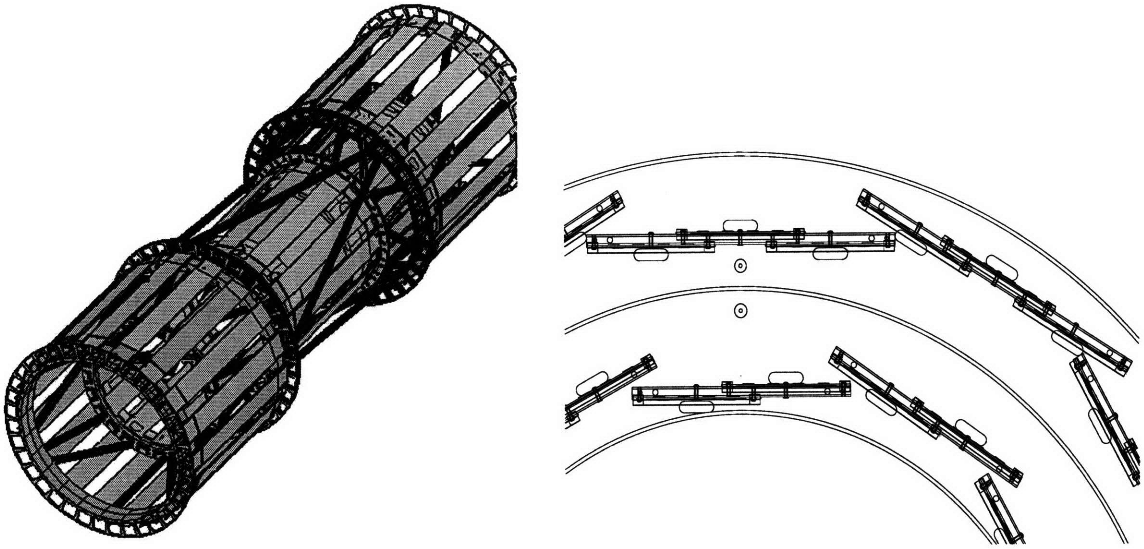

FIGURE 2.7: The Intermediate Silicon Later ISL: 3D view of the three barrels which compose the detector (left) and closeup of one section of the end view of an external barrel (right).

2.2.3 Intermediate Silicon Layer

The last silicon-based detector is the Intermediate Silicon Layer, ISL [37], 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° 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 [38]) is an open-cell drift chamber, with 8 *superlayers* consisting of 12 layers of wires each, for a total of 96 possible measurements per

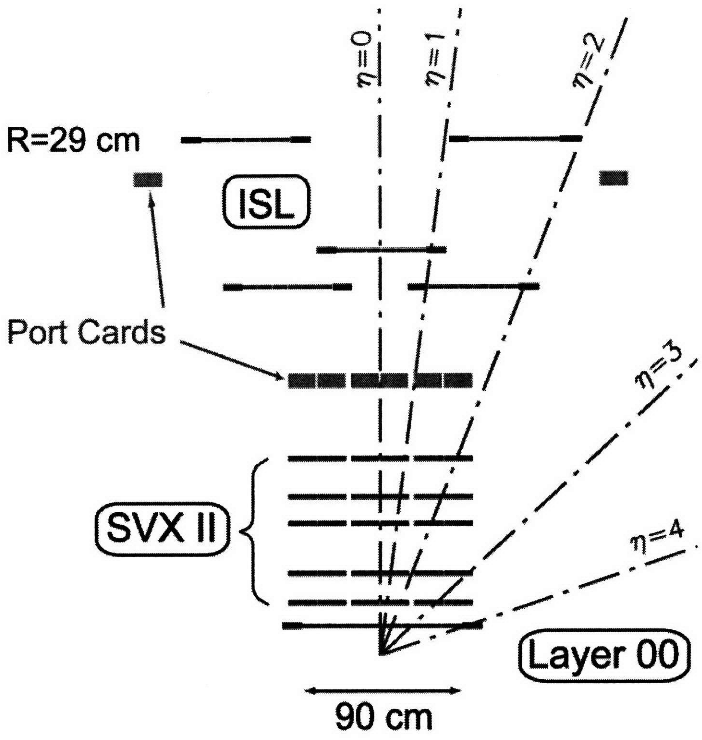

FIGURE 2.8: Coverage of the silicon subdetectors in the r–z plane. The scales of the z and r axes are different.

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 4 axial superlayers (wires parallel to the z axis) and the 4 stereo ones (± 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 [39]. 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 $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$. In addition, silicon measurements close to the beam allow precise

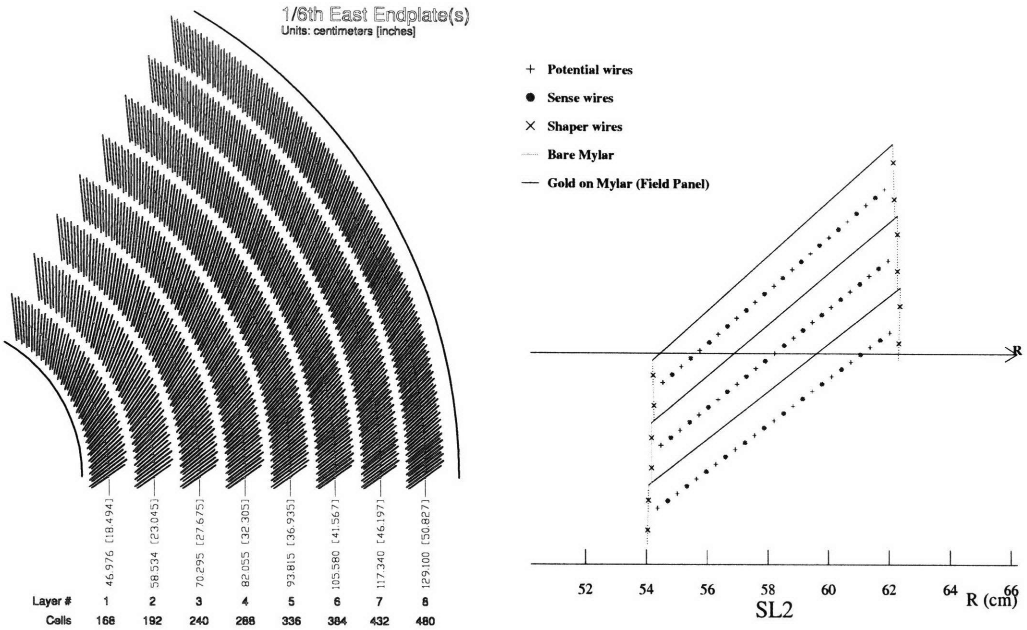

FIGURE 2.9: A view of a 1/6 section of the COT end-plate, in the r–φ plane, which shows the structure in superlayers (left), and a schematic draw of an axial cross-section of three cells in superlayer 2 (right).

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 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 $\mathrm{D}^*(2010)^+ \to \mathrm{D}^0\pi^+, \mathrm{D}^0 \to \mathrm{K}^-\pi^+$. The strong D^{*+} decay unambiguously defines the flavor of the D^0 meson, which dominantly decays in the Cabibbo-favored $\mathrm{K}^-\pi^+$ 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. A detailed description of the calibration procedure is presented in Reference [40]. More details on the use of dE/dx information for particle identification are reported in Section 4.3.4, which describes the tuning of the particle-

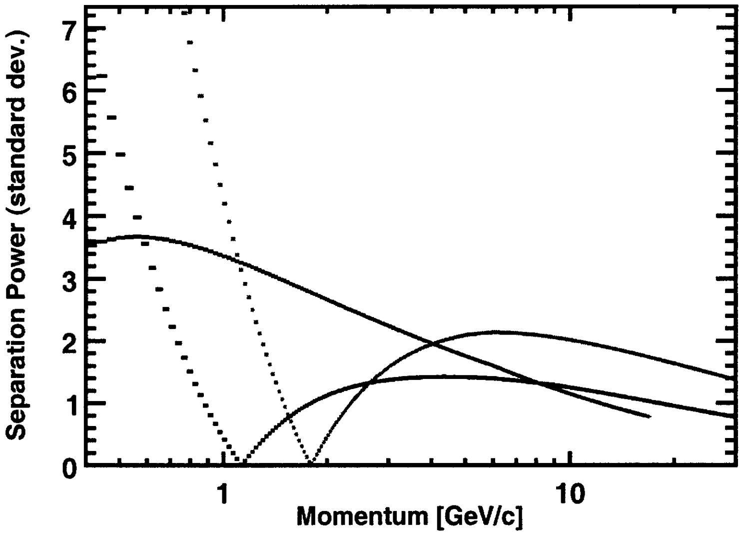

FIGURE 2.10: COT separating power in units of standard deviations *vs.* transverse momentum. The separation between pions and kaons, protons, and electrons are shown in black dashed, red dotted and blue solid line, respectively.

identification simulation in Monte Carlo, and in Section 6.5, where the application of particle-identification in the CDF same-side tagging algorithm is presented.

2.2.5 Time Of Flight

The Time of Flight (TOF [41]) detector was conceived and realized to provide particle identification capabilities for CDF expressly for the B_s^0 mixing analysis. It consists of 216 scintillator bars, approximately 280 cm long and with a cross-section of 4 cm × 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.

The TOF system plays a major rôle in the B_s^0 mixing analysis. 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:

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

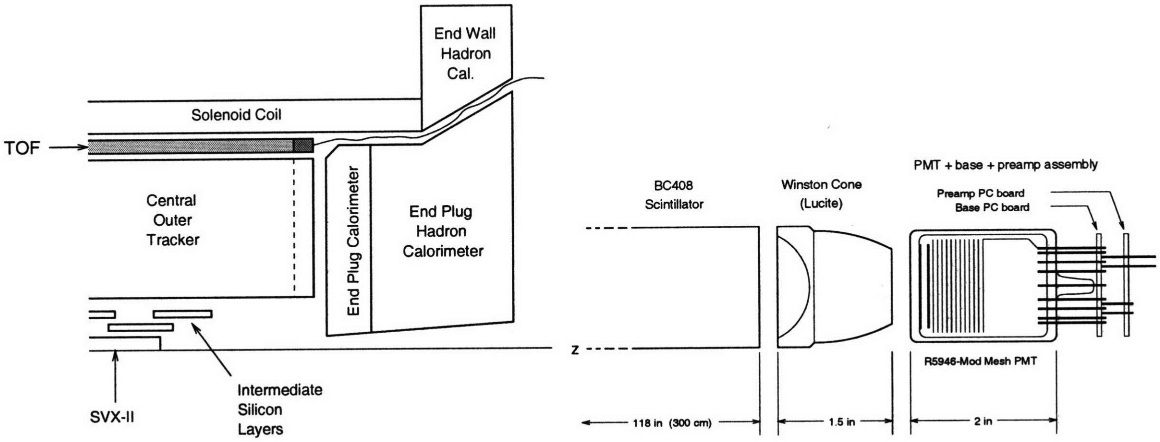

FIGURE 2.11: The Time Of Flight detector. On the left, the location of TOF is shown in a side view of CDF II. On the right, the arrangement of the scintillator, Winston cone and photomultiplier assembly. The Winston cone optimizes the optical coupling between the scintillator and the photomultiplier.

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 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 $\Delta\eta \times \Delta\varphi = 0.1 \times 15°$ in the η–φ plane. The Central ElectroMagnetic calorimeter (CEM [42]) 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

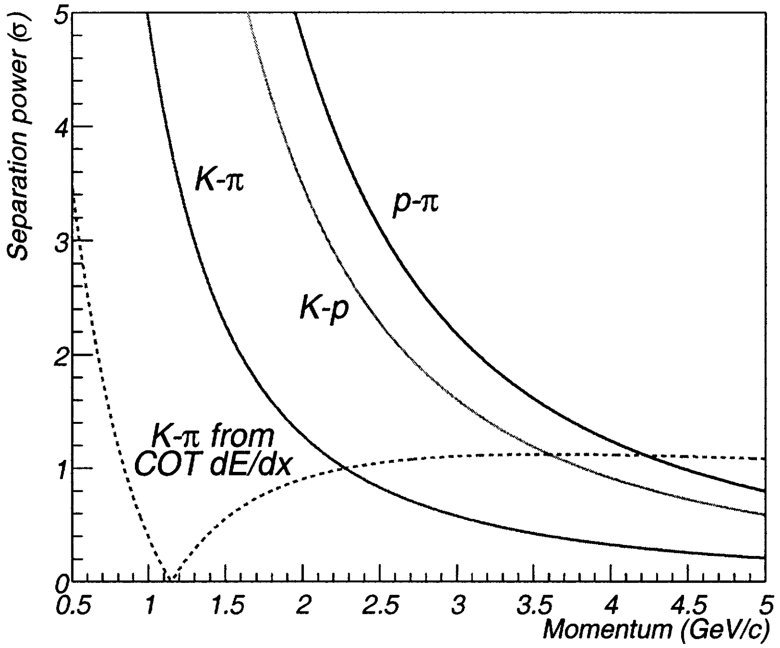

FIGURE 2.12: Expected TOF separating power, in units of standard deviations *vs.* momentum. The dashed line reports the K − π separation provided by specific ionization in the COT.

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

TABLE 2.2: Summary of the coverage, thickness and resolution of the CDF calorimeters [32]. The thickness is expressed in terms of the radiation length X_0 and the interaction length λ_0.

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 [43] 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

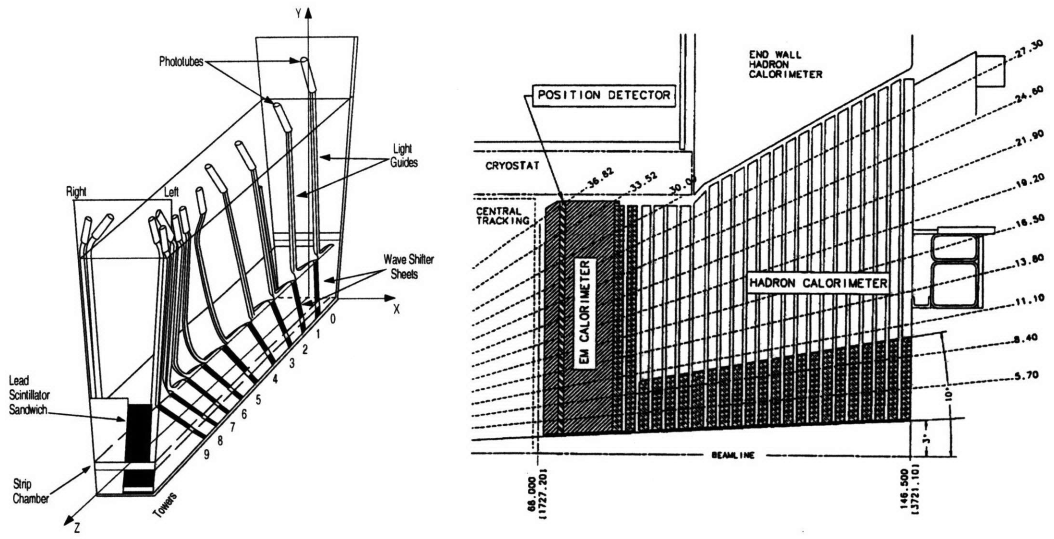

FIGURE 2.13: Schematic view of an azimuthal sector of central calorimeter (left) and elevation view of the upper part of the plug calorimeter (right). The elevation view on the right also indicates the location of the central calorimeters, above the cryostat, and the wall hadronic calorimeter, on the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

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 [44]), 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 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 $\Delta\eta \times \Delta\varphi = 0.1 - 0.16 \times 7.5°$ for $1.1 < |\eta| < 2.1$ and $\Delta\eta \times \Delta\varphi = 0.2 - 0.6 \times 15°$ for $2.1 < |\eta| < 3.6$. Their structure resembles the calorimeter in the central area. The electromagnetic section (PEM [45]) 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 [46]). 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 $7\lambda_0$. 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.

The Central MUon detector (CMU [47]) and Central Muon uPgrade detector (CMP [48]) 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

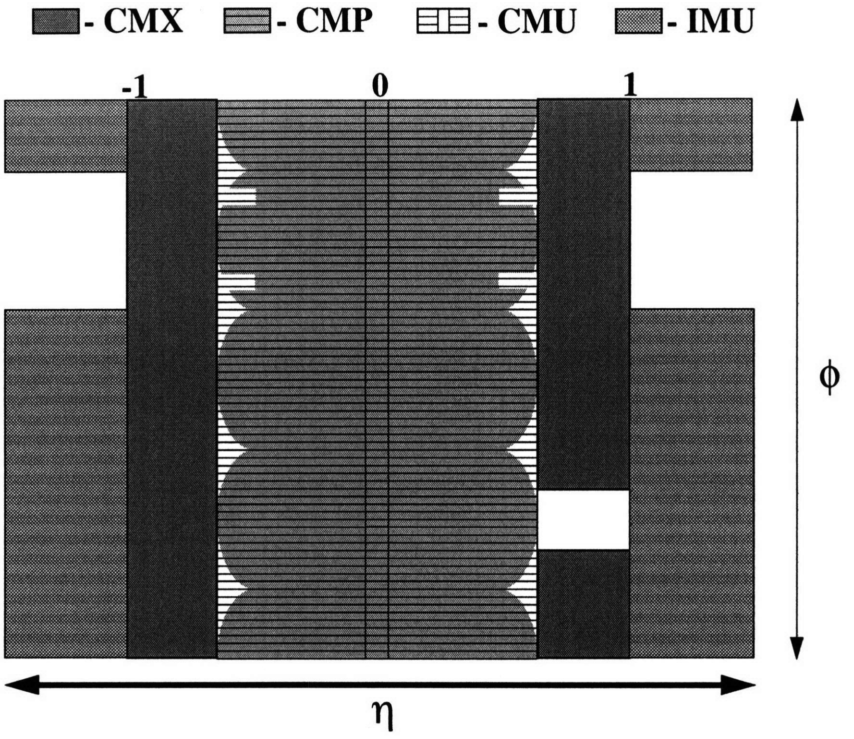

FIGURE 2.14: Coverage of muon chambers. The hatched and shaded areas represent the regions in η–φ which are instrumented by the subdetectors of the CDF muon system. The gap in the coverage of the CMX detector corresponds to the top area on the east side of CDF where the cryogenics system of the CDF solenoid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

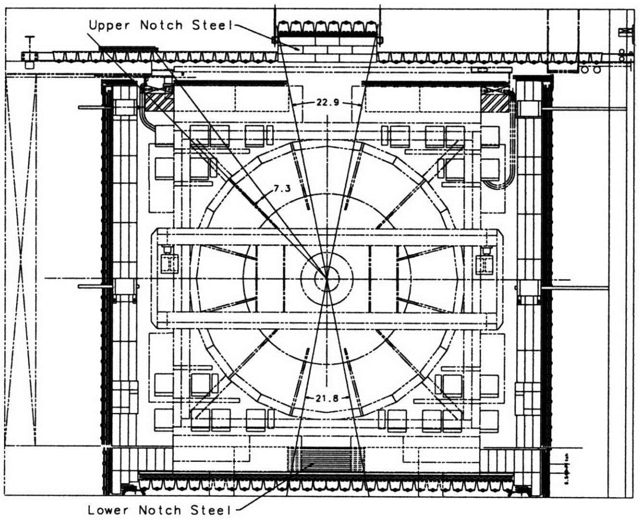

FIGURE 2.15: Central Muon uPgrade CMP. In this x–y section of the CDF detector, the CMP forms a rectangular box (dark shaded area) which surrounds the other parts of the CDF detector.

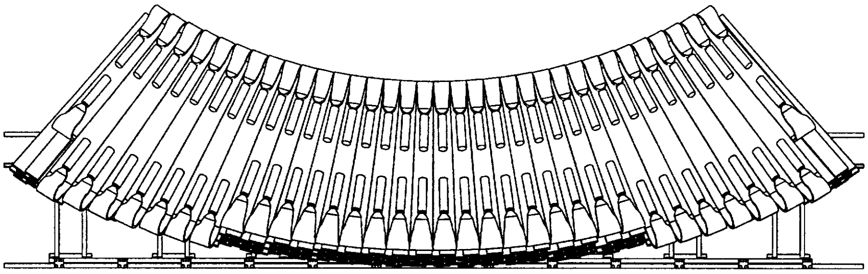

FIGURE 2.16: Central Muon eXtension. The drawing shows an elevation view in the x–y plane of a section of the CMX detector. The part depicted is referred to as "miniskirt" because it covers the lower section of the azimuthal range.

	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.

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 [49]) 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 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.

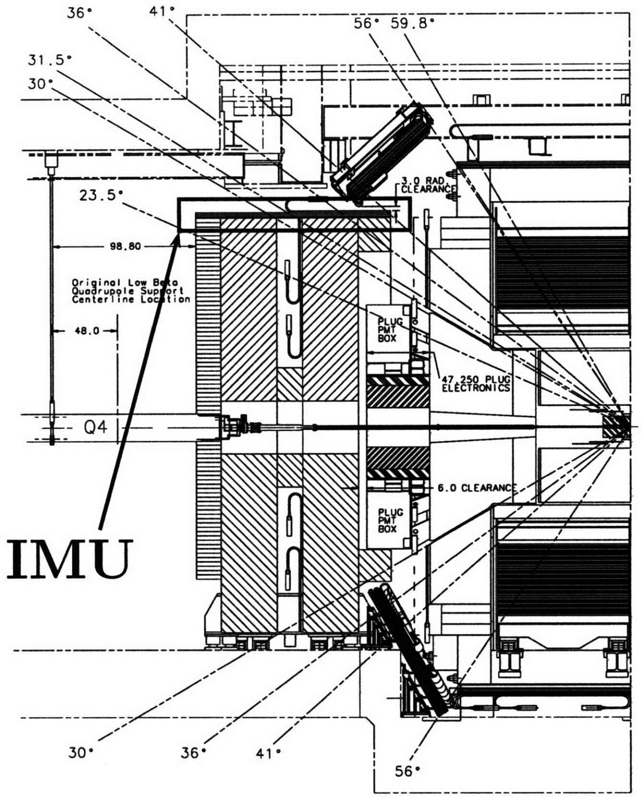

FIGURE 2.17: Intermediate MUon system IMU. The diagonally hatched area in the middle of the figure represents the sectional view of the toroids around which the IMU (dark shaded) is installed. The figure also shows a y–z view of the CMX detector, which corresponds to the two dark shaded areas in the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

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, 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.

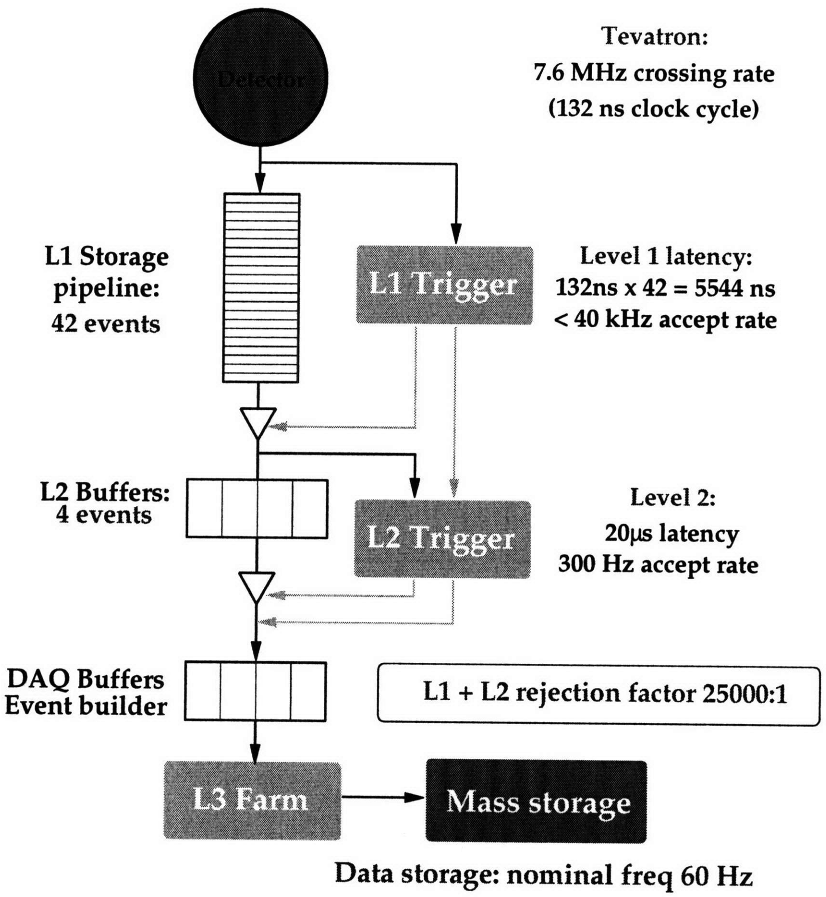

FIGURE 2.18: The CDF data acquisition system. The scheme emphasizes the timing information (latency, input rate and rejection factor) of the three levels of the trigger.

The 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.

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

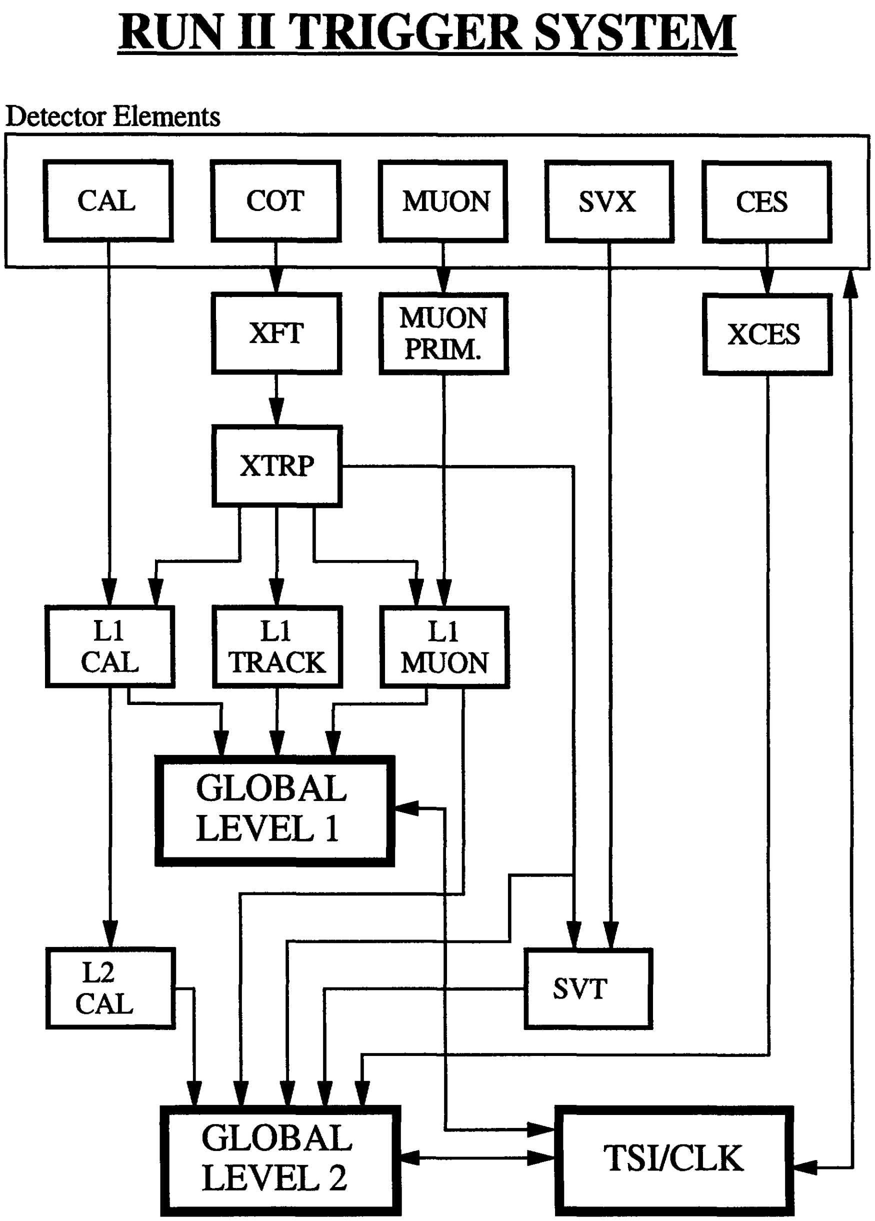

FIGURE 2.19: Scheme of the CDF trigger system. The connections between the subsystems of the CDF detector, in the upper part of the scheme, and the boards which constitute the trigger system are indicated. The various parts of the trigger system are described in the text in the relevant sections.

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

TABLE 2.4: Performance of XFT.

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 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 [50]). 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.

---

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

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 $\sin\theta$.

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

---

[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".

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 information 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 [51, 52]) is the most innovative part of the CDF trigger. It utilizes XFT tracks and SVXII hits, and reconstructs tracks, albeit 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

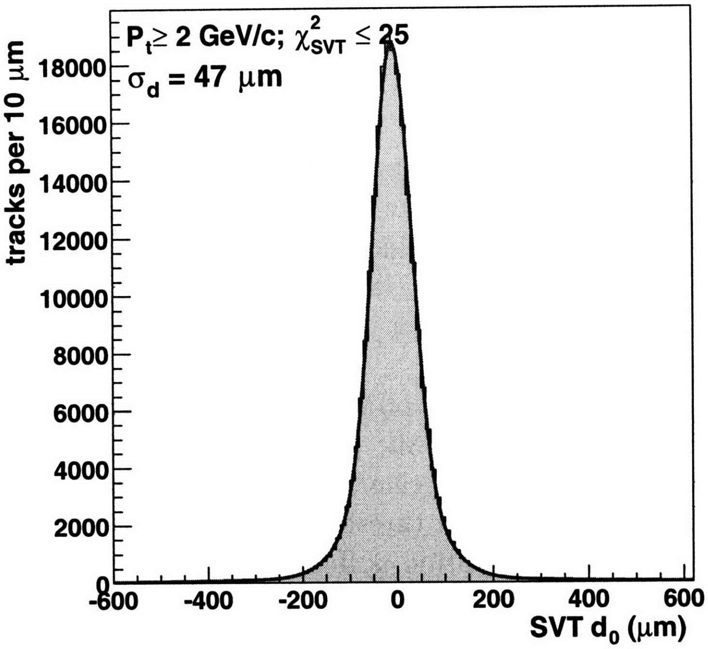

FIGURE 2.20: Impact parameter resolution as measured by SVT. The distribution includes the effect of the beam size. The SVT measures impact parameters with a r.m.s. width of 35 μm.

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 [53] 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 parameters 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 [54]) 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].

---

[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

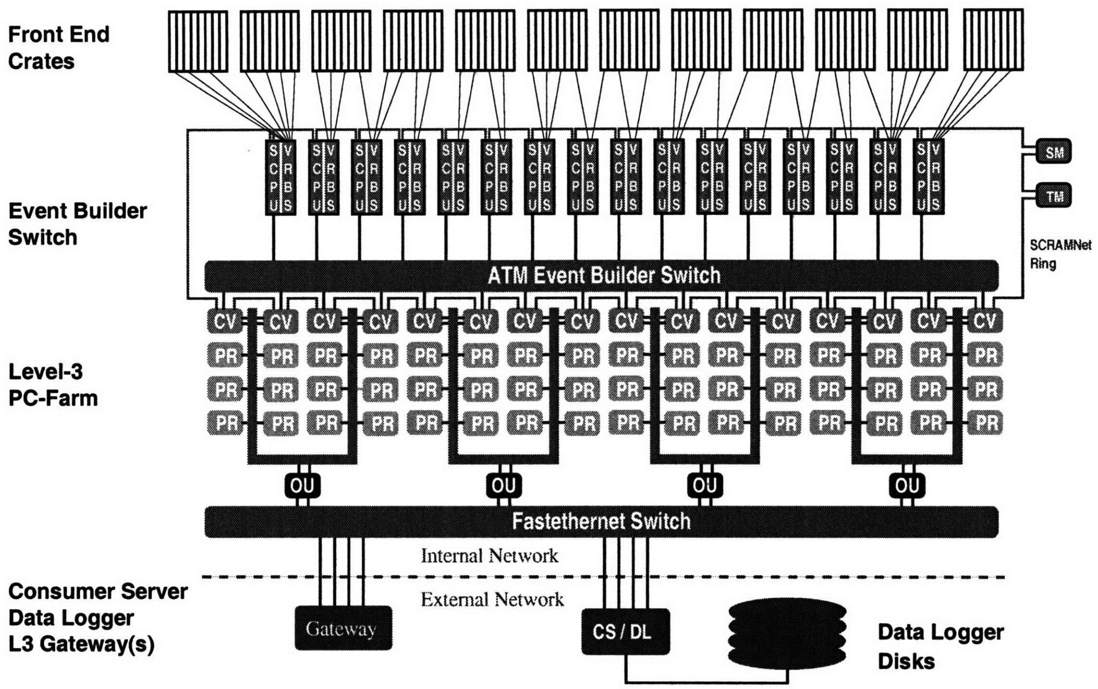

FIGURE 2.21: Scheme of the Level-3 farm. Data flow from top to bottom: from the front-end crates to the Event Builder crates, where data fragments are assembled, then to the Level-3 farm, through converter nodes (CV). Processor nodes (PR) produce the Level-3 trigger decision. Events which satisfy the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

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 rôle 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 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. After the description of how real $p\bar{p}$ interactions are produced in this chapter, the

---

subfarms with 21 to 22 nodes each.

next chapter will introduce the strategy for the online selection of B_s^0 *candidates, and present the reconstruction and selection of the* B_s^0 *candidates utilized for this analysis of* B_s^0 *oscillations.*

Chapter 3

B_s^0 data samples

This chapter presents the samples of data utilized for this analysis. The strategy adopted for the online selection of events enriched in B_s^0 *mesons is introduced in the first part of this chapter. The second part describes the reconstruction and offline selection of* B_s^0 *candidates.*

3.1 Triggers for the B_s^0 analysis

The set of the trigger prescriptions for the first, second and third level constitutes a trigger path. Triggers paths provide a logical, though not exclusive, classification among samples of events. The not-exclusiveness of the classification is present due to the possibility that events satisfy the requirements of multiple paths. The triggers utilized in this analysis are variations of paths belonging to two main categories: *two-track triggers* and *lepton + displaced-track triggers.*

3.1.1 Two-track triggers

The first class of triggers is characterized by the use of impact parameter cuts which enhance the long-lived particle content, and in particular the *b*-hadron content. The impact parameter of a track is correlated with the lifetime of the particle which produced it. B mesons fly, on average, 0.5 mm before decaying. This distance is significantly larger than the intrinsic beam size. B events are thus characterized by the presence of displaced tracks and vertices.

At Level-1, the trigger 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. The p_T cuts exploit the fact that the p_T spectrum of the particles produced in $p\bar{p}$ interactions follows p_T^{-8}, while the spectrum of the particles coming from a B decay is significantly harder. The cut on the angular separation eliminates the contribution from tracks inside back-to-back jets.

The Level-2 trigger prescribes the presence of two SVT tracks which confirm the Level-1 requirements. The angular cut, this time on $\Delta\varphi_0$, is tightened, with the application of a lower limit. Pairs of quasi-collinear tracks are found, for example, in hadronic jets, due to light quark fragmentation, while the opening angle between the products of a B decay is more evenly distributed between 0° and 180°. The trigger additionally includes the already mentioned cut on the impact parameter, d_0, and a lower cut on $\mathbf{r}_{xy} \cdot \sum \mathbf{p_T}$, where $\mathbf{r}_{xy}$ is the distance, in the transverse plane, between the beam position and the displaced vertex formed by the two trigger tracks, and $\sum \mathbf{p_T}$ is the vector sum of the transverse momenta of the two tracks. This quantity is symmetrically distributed around zero in the case of randomly chosen tracks and skewed toward positive values when the two tracks come from a B decay.

Level-3 applies the requirements of Level-2 utilizing the fits performed using the full detector information. Three-dimensional fits of the trigger tracks are available and, in particular, the points of closest approach to the z axis of the two tracks are required to be within 5 cm from each other.

Three trigger scenarios, denoted `BCHARM`, `LOWPT`, and `HIGHPT`, are identified by the different values of the cuts applied. The list of their requirements are summarized in Table 3.1. A variation of the `LOWPT` trigger adds the requirement of a muon with $p_T > 1.5$, or 2 GeV/c, depending on whether it is found in the CMU or CMX subdetectors, respectively. This muon is required to form an angle of at least 90° from the direction of each of the SVT tracks. This cut, which uses XFT tracks, and thus φ_6, was not imposed for some part of the data collected by this trigger path, leading to the collection of large semi-muonic B_s^0 decays with this trigger.

The various B triggers permit the full exploitation of the available trigger bandwidth. At high instantaneous luminosity in the Tevatron, higher purity selections, i.e., the `HIGHPT` scenario, are given priority, where purity is a measure of the signal-to-background ratio. As the instantaneous luminosity decreases during a store, trigger bandwidth becomes available and low purity triggers are utilized to maximize the B yield written to tape. Section 3.2 presents a summary of the trigger composition of the B_s^0 data sample utilized for this analysis of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations.

3.1.2 Lepton-plus-displaced-track triggers

The second class of triggers, the lepton-plus-displaced-track triggers, combines the request of a lepton in the final state and a track with impact parameter greater than a threshold. The type of lepton, either an electron or a muon, and the type of chambers which identified the lepton (in the case of muons, CMP, or CMU) label the trigger paths which fall in this category. Lepton identification is thoroughly explained in References [55] (muon identification) and [56] (electron identification).

Level-1 requires a muon or an electron object with $p_T > 4.0$ GeV/c and an XFT track with $p_T > 2.0$ GeV/c. The muon object is defined as an XFT track matched to a muon stub in both the CMU and the CMP chambers. The electron object is characterized by an XFT track matched to a calorimeter trigger tower with $E_T >$ 4 GeV. In addition, the ratio between the fraction of energy deposited in the HA section and in the EM one (E_{HA}/E_{EM}) is required to be smaller than a fixed threshold. An upper cut on the angular separation between the lepton object and the XFT trigger track is applied.

At Level-2, a Level-2 lepton and an SVT track must satisfy the Level-1 cuts. The

	Level-1 XFT tracks	Level-2 SVT tracks	Level-3 COT+SVX tracks
`HIGHPT`	$p_T > 2.46$ GeV/c opposite charge $\Delta\varphi_6 < 135°$ $\sum p_T > 6.5$ GeV/c	$p_T > 2.5$ GeV/c opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 6.5$ GeV/c $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m	$p_T > 2.5$ GeV/c opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 6.5$ GeV/c $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m $\lvert\Delta z_0\rvert < 5$ cm
`BCHARM`	$p_T > 2.04$ GeV/c opposite charge $\Delta\varphi_6 < 135°$ $\sum p_T > 5.5$ GeV/c	$p_T > 2.0$ GeV/c opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 5.5$ GeV/c $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m	$p_T > 2.0$ GeV/c opposite charge $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 5.5$ GeV/c $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m $\lvert\Delta z_0\rvert < 5$ cm
`LOWPT`	$p_T > 2.04$ GeV/c $\Delta\varphi_6 < 90°$	$p_T > 2.0$ GeV/c $\Delta\varphi_0 < 90°$ $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m	$p_T > 2.0$ GeV/c $2° < \Delta\varphi_0 < 90°$ $\sum p_T > 4.0$ GeV/c $120\ \mu\text{m} < d_0 < 1$ mm $L_{xy} > 200\ \mu$m $\lvert\Delta z_0\rvert < 5$ cm
†	CMU or CMP muon $p_T(\mu_{CMU}) > 1.5$ GeV/c or $p_T(\mu_{CMP}) > 2.0$ GeV/c	$\Delta\varphi_6(\mu, trk) > 90°$	

TABLE 3.1: Displaced two-track trigger requirements. The variable L_{xy} represents $\mathbf{r}_{xy} \cdot \sum \mathbf{p_T}$, defined in the text. The requirement for the tracks to be oppositely-charged in the two scenarios with higher purity optimizes the selection for $B^0_{d,s} \to h^+h^-$ decays, without majorly affecting the collection of B multi-body decays.
† The additional request of a muon object characterizes the μ+LOWPT path.

electron is additionally required to have at least 2 GeV energy measured by the CES. The SVT track is also required to have $120\ \mu\text{m} < d_0 < 1$ mm and, for electron triggers only, to be within an angle between 2° and 90° from the lepton track.

The processors of the Level-3 farm have access to better track fits and much more information which is used to define triggers. The opening angle between the lepton and SVT track is required to be in the $(2°, 90°)$ range for both electrons and muons. The lepton identification criteria are also tightened. In the electron case, the transverse (Δx) and longitudinal (Δz) distances between the shower centroid measured by CES and the extrapolated track position are required to be within a preselected upper value. The transverse and longitudinal profiles of the showers in the CES are also compared to default shapes produced by test beam electrons, and an

	Level-1	Level-2	Level-3
electron	$p_T > 4.0$ GeV/c	$p_T > 4.0$ GeV/c	$p_T > 4.0$ GeV/c
	$E_T > 4$ GeV	$E_{CES} > 2$ GeV	$\Delta x_{CES} < 5$ cm
			$\Delta z_{CES} < 3$ cm
	$E_{HA}/E_{EM} < 0.125$		$\chi^2_x < 10$, $\chi^2 < 15$
			$L_{shr} < 0.2$
		$2^\circ < \Delta\varphi_0(e, trk) < 90^\circ$	
muon	$p_T > 4.0$ GeV/c	$p_T > 4.0$ GeV/c	$p_T > 4.0$ GeV/c
	CMU and CMP stubs		$\Delta x_{CMU} < 15$ cm
			$\Delta x_{CMP} < 20$ cm
track	$p_T > 2.0$ GeV/c	$p_T > 2.0$ GeV/c	$p_T > 2.0$ GeV/c
		$120\ \mu\text{m} < d_0 < 1$ mm	$120\ \mu\text{m} < d_0 < 1$ mm
			$\text{m}(\ell, trk) < 5$ GeV/c^2
			$2^\circ < \Delta\varphi_0(\ell, trk) < 90^\circ$

TABLE 3.2: Lepton-plus-displaced-track trigger requirements.

appropriate χ^2 is defined by comparing the two sets. Level-3 is also able to enforce a cut on the reconstructed mass of the lepton-track pair, by using three-dimensionally-reconstructed tracks. Another quantity that is used for electron identification is the lateral shower sharing L_{shr} [57] which measures the difference between the observed sharing of energy deposition between towers in the CEM and the deposition expected from real electromagnetic showers. In the muon case, a cut is applied to the distance Δx between the CMU and CMP stubs and the extrapolated track. The requirements of the lepton-plus-displaced-track trigger are listed in Table 3.2.

A brief summary of the contribution of lepton-plus-displaced-track triggers to our data sample is presented in the next section.

3.2 Data samples for the analysis of B^0_s oscillations

The data samples utilized in the analysis presented in this document were recorded in the period from March 2002 to January 2006. They correspond to an integrated luminosity of about 1 fb^{-1}, after imposing the requirement that the full detector systems were properly functioning. As a side note, one of the biggest sources of data losses is the request that the SVXII is on. Stable beam conditions are necessary to reduce the possibility that the silicon detectors are damaged by beam losses.

Three periods of data-taking, usually referred to as `0d`, `0h`, and `0i`, are identified. The separation comes from the name of the files which contain the data corresponding to the March 2002 – September 2004 (355 pb^{-1}), December 2004 – November 2005 (410 pb^{-1}), and November 2005 – January 2006 (230 pb^{-1}) periods.

The integrated luminosity of the sample which passed the `BCHARM` trigger path is

Definition	Fraction
(BCHARM) / (TTT)	60%
(LOWPT *and not* HIGHPT *and not* BCHARM) / (TTT)	30%
(HIGHPT *and not* BCHARM *and not* LOWPT) / (TTT)	10%
(ℓ+SVT *and* TTT) / (ℓ+SVT)	60%
(ℓ+SVT *and not* TTT) / (ℓ+SVT *or* TTT)	10%

TABLE 3.3: Summary of overlaps among trigger paths and classes. In the table, trigger names indicate the number of events which passed the respective trigger: ℓ+SVT indicates the number of events which passed any of the lepton-plus-displaced-track triggers; TTT indicates the number of events which passed any of the three two-track triggers: BCHARM, HIGHPT, or LOWPT. The LOWPT tag includes the contribution of μ+LOWPT paths. The number of events which belong to the ℓ+SVT class is about 57 and 67 millions, for the muon and electron triggers, respectively. The TTT class contains about 560 million events.

about 642 pb^{-1}, while the `HIGHPT` and `LOWPT` paths collected 504 pb^{-1} and 418 pb^{-1}, respectively. These numbers include the effect of trigger prescaling. The three samples obviously overlap, because events may have been triggered by more than one trigger path. The fraction of the events, in the whole sample of two-track triggers, that passed the `BCHARM` trigger path is 60%. The `HIGHPT` trigger path exclusively selects 10% of the sample, and `LOWPT` the remaining 30%. In the sample selected with lepton-plus-displaced-track triggers, 60% of events are also selected by at least one of the two-track triggers, `BCHARM`, `HIGHPT`, or `LOWPT`. The fraction of events which exclusively passed a lepton-plus-displaced-track trigger is 10% of the union of the samples selected by a lepton-plus-displaced-track trigger and those selected by a two-track trigger. Table 3.3 presents a summary of overlaps among trigger paths and classes described in this paragraph.

3.3 Data format and analysis software

A framework, referred to as the `BStntuple` [58], has been implemented for efficiently storing and accessing the pieces of information which form the B_s^0 candidates. It shares the basic structure of the standard `stntuple` [59], which itself constitutes a more sophisticated ROOT ntuple [60] together with a set of convenient tools.

This framework contains structures to hold the reconstructed candidates information (stable and decaying objects) as well as tagging information, and particle identification information (TOF, $d\text{E}/d\text{x}$, muon and electron quantities). The actual data blocks correspond to instances of these classes for specific decay modes, which are implemented by cloning prototype modules which are appropriate for the decay topology.

This has revealed to be an efficient and uniforming framework, in terms of both CPU usage time and procedure sharing, which has facilitated candidate reconstruction and the process of producing ntuples for the various decays and data samples employed in the analysis.

3.4 B_s^0 data sample

The data sample consists of 1 fb^{-1} of data collected with the CDF II detector, between March 2002 and January 2006. According to the type of particles which take part in the decay chain, the sample of B_s^0 decays is naturally divided in two classes, which are usually referred to as "semileptonic" and "hadronic" modes,

For both hadronic and semileptonic modes, B_s^0 candidates are reconstructed from the final decay products. For example, $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$, $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$ candidates are reconstructed starting from a pair of oppositely charged tracks which are assigned the mass of a kaon (the positively charged track) and of a pion (the negatively charged one), producing a K^{*0} candidate. A D_s^- vertex is fit using another charged track, with kaon mass assigned, and the projected trajectory of the K^{*0}. Finally, the D_s^- candidate is associated with three pion tracks which are consistent with coming from the same vertex. The three pions are constrained to come from a single vertex, and consistency with this hypothesis is enforced by applying a cut on the χ^2 of the vertex fit. All fits of vertices are performed by the `CTVMFT` package [61]. The mass of the D_s^- candidate is fixed to its world average [62] in the final fit for the hadronic B_s^0 candidate. This is not the case for the semileptonic B_s^0 decays, where the mass values of D_s^- candidates are not constrained in the fit of the ℓD_s^- combination, because the unconstrained D_s^- mass serves as a discriminant in rejecting background.

3.4.1 Semileptonic B_s^0 decays

The sample of semileptonic decays consists of inclusively reconstructed $B_s^0 \to D_s^-\ell^+X$ candidates, where $\ell = e, \mu$, which are searched for in the sample collected with the two-displaced-track trigger and the lepton-plus-displaced-track trigger. The definition of these triggers are reported in Section 3.1. While not specifically reconstructed, the $B_s^0 \to D_s^-\tau^+X$ mode is part of the semileptonic signal sample. The D_s^- meson is reconstructed in the final states $\phi^0\pi^-$, $K^{*0}K^-$ and $\pi^-\pi^+\pi^-$, where $\phi^0 \to K^+K^-$, and $K^{*0} \to K^+\pi^-$. The sample is enriched in B_s^0 candidates by applying a cut-based selection.

The main advantage of these decays is the large branching ratio for semileptonic decays of B_s^0 mesons, which the Review of Particle Physics [7] reports to be equal to 7.9 ± 2.4 %, and the presence of a lepton which provides a clear signature. Conversely, the incompleteness of the reconstruction, which is caused, at least, by the participation of a neutrino, the energy and momentum of which are undetected, constitutes a challenge. In this case, it is customary to analyze both the ℓD_s^- mass distribution and the D_s^- one, which provides additional discriminating power between signal and background.

Lepton identification exploits the algorithms prepared for the study of lepton flavor taggers [55, 56]. Electrons and muons are separated from hadrons by means of a lower cut on a likelihood which is defined as follows:

$$\mathcal{L}_i = \frac{\mathcal{P}_i^{\mathcal{S}}}{\mathcal{P}_i^{\mathcal{S}} + \mathcal{P}_i^{\mathcal{B}}}\,, \quad i = e \ or\ \mu\,, \tag{3.4.1}$$

where $\mathcal{P}_{e,\mu}^{\mathcal{S}}$ and $\mathcal{P}_{e,\mu}^{\mathcal{B}}$ are, respectively, the likelihoods of the lepton candidate being a real lepton or a false one. These functions are simply the product of the probability density functions of the variables chosen to maximize the separating power between leptons and hadrons.

The distributions of the discriminating variables obtained in samples of pure muons, collected by reconstructing $J/\psi \to \mu^+\mu^-$, and electrons, from $\gamma \to e^+e^-$ conversions in the detector material, are fit to empirical functions. These parameterizations are treated as probability density functions of a real lepton candidate, and their product defines $\mathcal{P}_{e,\mu}^{\mathcal{S}}$. Electrons from conversions are identified by searching, around an electron candidate, as defined by the trigger cut in Table 3.2, for oppositely charged tracks which have a small opening angle with the candidate, and requiring the trajectories of the two tracks to be parallel to one another at their distance of closest approach.

The background likelihoods $\mathcal{P}_{\mu}^{\mathcal{B}}$ and $\mathcal{P}_{e}^{\mathcal{B}}$ are produced analogously to the signal likelihoods. The distributions of the discriminating variables, the parameterizations of which provide the background probability density functions, are constructed using samples of pions from $\mathrm{K}_s^0 \to \pi^+\pi^-$ and, in the case of $\mathcal{P}_{\mathcal{B}}^{\mu}$, also kaons and protons from $\mathrm{D}^0 \to \mathrm{K}^-\pi^+$ and $\Lambda^0 \to p\pi^-$, respectively.

The plots in Figure 3.1 show the electron and muon likelihood distributions $\mathcal{L}_e$ and $\mathcal{L}_\mu$ in the signal and background samples described above. The values for these likelihoods are bound between zero and one, with real electrons and muons populating the high likelihood region close to unity, while backgrounds preferentially occupy the low likelihood region close to zero. The requirement that $\mathcal{L}_e$ is smaller than 0.9 is about 90% efficient for conversion electrons with $p_T > 2.0$ GeV/c, while rejecting around 98% of the pions originating from K_s^0 decays and fulfilling the electron candidate requirements. Different values for cuts on $\mathcal{L}_\mu$ are utilized for candidate muons found in the various subdetectors of the CDF muon system. The cut values and efficiencies, for real and false muons, for the different detector components are reported in Table 3.4.

Various sources contribute to the background of a B_s^0 candidate reconstructed solely by matching a D_s^- candidate to a lepton candidate. Three sources are identified: "false lepton", "physics" and "combinatorial" backgrounds. Contributions to the background may contain a true, and correctly reconstructed, D_s^- meson. That is the case for B^0 and B^+ decays with a D_s^- in the final state, such as $\mathrm{B}^0/\mathrm{B}^+ \to \mathrm{D}_s\mathrm{D}$, $\mathrm{D} \to \ell\bar{\nu}X$, and B_s^0 candidates composed by a true D_s^- and a false lepton track. Another source of this type of background is constituted by $\mathrm{B}_s^0 \to \mathrm{D}_s^{+(*)}\mathrm{D}_s^{-(*)}X$, $\mathrm{D}_s^{(*)} \to \ell^+Y$ decays. These B_s^0 modes constitute a background, rather than a signal, because they are not self-tagging final states. These backgrounds, which are called "physics" backgrounds, are indistinguishable from signal candidates in the D_s^- mass plot, while the $\ell\mathrm{D}_s^-$ mass distribution provides some separation. Samples of simulated events are

Muon System	$\mathcal{L}_\mu >$	Eff. for real muons [%]	Eff. for false muons [%]
CMU	0.50	92.0	13.5
CMP	0.50	88.2	27.1
CMUP	0.05	98.8	55.0
CMX	0.50	91.8	22.2
IMU	0.70	78.8	9.6

TABLE 3.4: Efficiency of muon likelihood requirements for real and false muons matching the muon candidate requirements, compiled for different muon detector systems. The quantity $\mathcal{L}_\mu$ represents the likelihood that the particle is a real muon, and is defined in Equation 3.4.1. The efficiencies are calculated utilizing a sample of muons from $J/\psi \rightarrow \mu^+\mu^-$ (real muons), and pions, kaons, and protons from $K^0_s \rightarrow \pi^+\pi^-$, $D^0 \rightarrow K^-\pi^+$ and $\Lambda^0 \rightarrow p\pi^-$, respectively (false muons) [55].

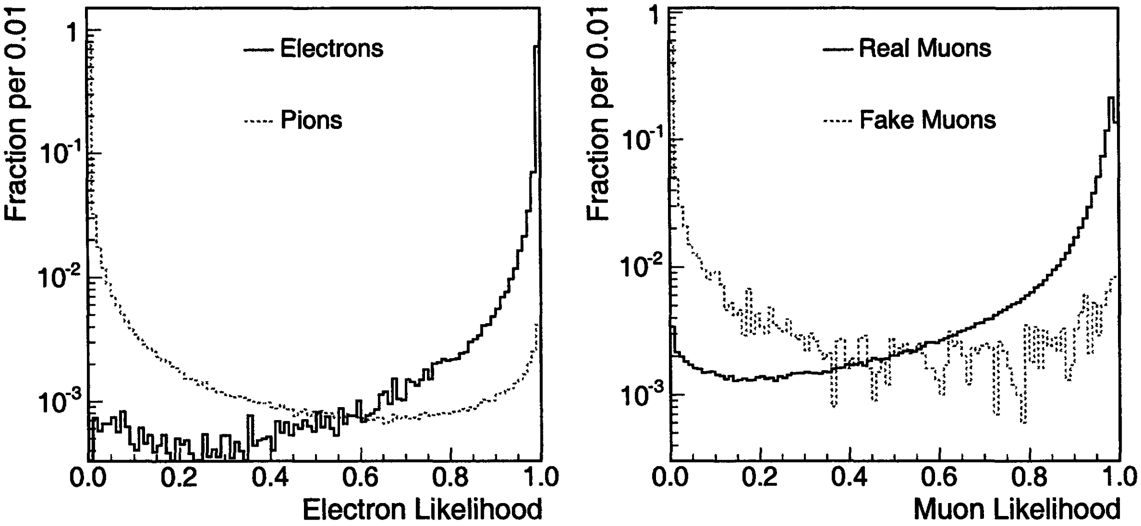

FIGURE 3.1: Electron likelihood distributions $\mathcal{L}_e$ for electrons and pions with $p_T >$ 2.0 GeV/c (left), and muon likelihood distributions $\mathcal{L}_\mu$ for real and false muons.

utilized to derive ℓD_s^- mass templates. Other backgrounds, which do not contain a true D_s^-, such as combinatorial background, are instead very well discriminated in all mass distributions. The mass distribution of D_s^- candidates provides an effective handle in separating combinatorial background from signal, as shown by the plots in Figure 3.3. In the ℓD_s^- mass distributions, the shape of the "false lepton" and "physics" component are obtained from Monte Carlo samples of $B \rightarrow D\ell X$ mesons, where ℓ is either a muon or an electron. The shape of the combinatorial background component is a template derived from the ℓD_s^- mass distribution of candidates which fall in the sideband regions of the D_s^- mass distribution.

The selection of B^0_s candidates is based on the cuts reported in Table 3.5. Each cut value has been optimized in order to maximize $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$. For the optimization, a sample of simulated semileptonic B^0_s candidates was utilized to estimate the efficiency of a set of cuts on the signal fraction. The distribution of D_s^- candidates in the sample of simulated events is fit with a Gaussian function. The mean μ_G and width σ_G of the Gaussian are used to define the signal region, $[\mu_G - 3\sigma_G, \mu_G + 3\sigma_G]$. The background fraction in the signal region is estimated by fitting the lower and upper sidebands in the D_s^- mass distribution in data. The sidebands are defined by excluding the signal region, as defined above, from the $[1.92, 2.02]$ GeV/c^2 mass range in which the fit of the background component is performed. Selection cuts are individually optimized. The cuts are divided in three classes:

- **fit quality**
  The quality of the vertex fits which composed the candidate is ensured by applying a lower cut on the fit probability of the B_s^0 vertex fit, $\mathcal{P}(B_s^0)$, and the two-dimensional $\chi^2_{r\phi}$ of the D_s^- vertex fit, $\chi^2_{r\phi}(D_s^-)$. The vertex position is obtained by constraining the D_s^- and the lepton candidate to a common vertex, in the three-dimensional space. The two-dimensional $\chi^2_{r\phi}$ of a vertex fit is calculated by removing the z component from the error matrix of the vertex fit.

- **kinematics**
  A cut on the transverse length significance $L_{xy}/\sigma_{L_{xy}}$ of the B_s^0 and D_s^- candidates exploits the large lifetime of B_s^0 mesons to discriminate between signal and those backgrounds which are mostly prompt. The transverse lengths $L_{xy}(B_s^0)$ and $L_{xy}(D_s^-)$, which are defined in Equation 4.3.1, are both calculated with respect to the $p\bar{p}$ interaction vertex.
  The cuts on the proper decay-lengths $ct^*(B_s^0)$, $ct(D_s^-)$ and proper-decay-length resolution $\sigma_{ct^*}(B_s^0)$ are meant to eliminate poorly reconstructed candidates. These quantities are defined in Equations 5.2.1 and 5.2.2. The * indicates that only the ℓD_s^- part of the candidate enters the calculation. The same reasoning justifies the minimum p_T required of the tracks that compose the B_s^0 candidate.
  The request for a minimum value of $|\cos\psi_H|$, the cosine of the helicity angle of the D_s^- in vector–scalar decays, such as $\phi^0\pi^-$ or $K^{*0}K^-$, rejects more flat-distributed background than signal, which peaks at large (absolute) values of $\cos\psi_H$. It is defined as the angle, in the reference frame in which the D_s^- candidate is at rest, between the transverse momenta of the B_s^0 and of the π^-, or K^-, candidates, for $D_s^- \to \phi^0\pi^-$ and $D_s^- \to K^{*0}K^-$ decays, respectively. Figure 3.2 presents the graphical definition of $\cos\psi_H(K^{*0})$, in the $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$ decay chain.

- **particle identification**
  Leptons are identified utilizing the likelihood functions which were developed for the soft lepton taggers. The likelihoods for lepton identification are briefly introduced at the beginning of this section.
  The work on the same-side tagger provides a combined likelihood ratio which allows for the separation of kaons from pions. This work is described in Section 6.5. The purity of the decay modes with kaons in the final state is enhanced by requiring the tracks which are assigned the mass of a kaon to pass a higher

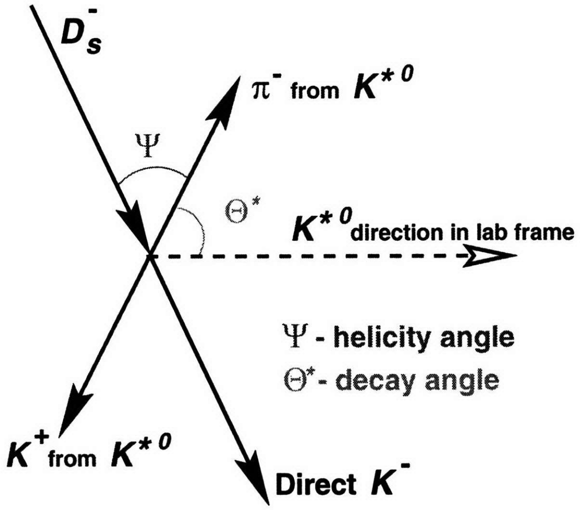

FIGURE 3.2: Definition of the helicity angle ψ_H (indicated by Ψ in the figure) and the decay angle θ^* (Θ^*, in the figure), in the $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$ decay chain. The momenta are drawn in the reference frame in which the K^{*0} is at rest.

Cut		$D_s^- \to \phi^0\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mathcal{P}(B_s^0)$	$>$	10^{-7}	10^{-5}	10^{-5}
$\chi^2_{r\phi}(D_s^-)$	$<$	20	20	20
$L_{xy}/\sigma_{L_{xy}}(D_s^-)$	$>$	5	8	11
$L_{xy}/\sigma_{L_{xy}}(B_s^0)$	$>$	—	2	2
$ct^*(B_s^0)$ [cm]	$>$	0.01	0.01	0.01
$\sigma_{ct^*}(B_s^0)$ [cm]	$<$	0.04	0.04	0.04
$ct(D_s^-)$ [cm]	$\in$	[-0.01,0.10]	[-0.01,0.10]	[-0.01,0.10]
$p_T(trk)$ [GeV/c]	$>$	0.4	0.4	0.7
$\lvert\cos\psi_H\rvert$	$>$	0.3	0.3	—
$m(\ell D_s^-)$ [GeV/c^2]	$\in$	[2.0,5.5]	[2.0,5.5]	[2.0,5.5]
$CLL(K_1)$	$>$	-2.5	-2.25	—
$CLL(K_2)$	$>$	-2.5	-1.1	—

TABLE 3.5: Selection criteria for $B_s^0 \to D_s^-\ell^+X$ candidates. K_1 and K_2 indicate, in the $D_s^- \to K^{*0}K^-$ decay, the kaon from the D_s^- decay and the kaon from the K^{*0} decay, respectively.

> cut on the combined likelihood ratio, *CLL* (Equation 6.5.3). In addition, the requirement on particle-identification information allows for looser kinematic selection. The purity of the sample thus increases without a loss in efficiency.

As mentioned above, Table 3.5 contains the value of the cuts utilized in this analysis. Finally, Table 3.6 summarizes the yields of B_s^0 candidates in the three reconstructed decay modes: $B_s^0 \to D_s^-\ell^+X$ with $D_s^- \to \phi^0\pi^-$, $K^{*0}K^-$, and $\pi^-\pi^+\pi^-$. The corresponding ℓD_s^- and D_s^- mass distributions are shown in Figure 3.3.

Decay Sequence	Yield
$D_s^- \rightarrow \phi^0\pi^-$	29,600 ± 800
$D_s^- \rightarrow K^{*0}K^-$	22,000 ± 800
$D_s^- \rightarrow \pi^-\pi^+\pi^-$	9,900 ± 700
Total	61,500 ± 1,300

TABLE 3.6: B_s^0 signal yields for the semileptonic modes in the various decay sequences. The quoted numbers corresponds to an integrated luminosity of ~ 1.0 fb^{-1}.

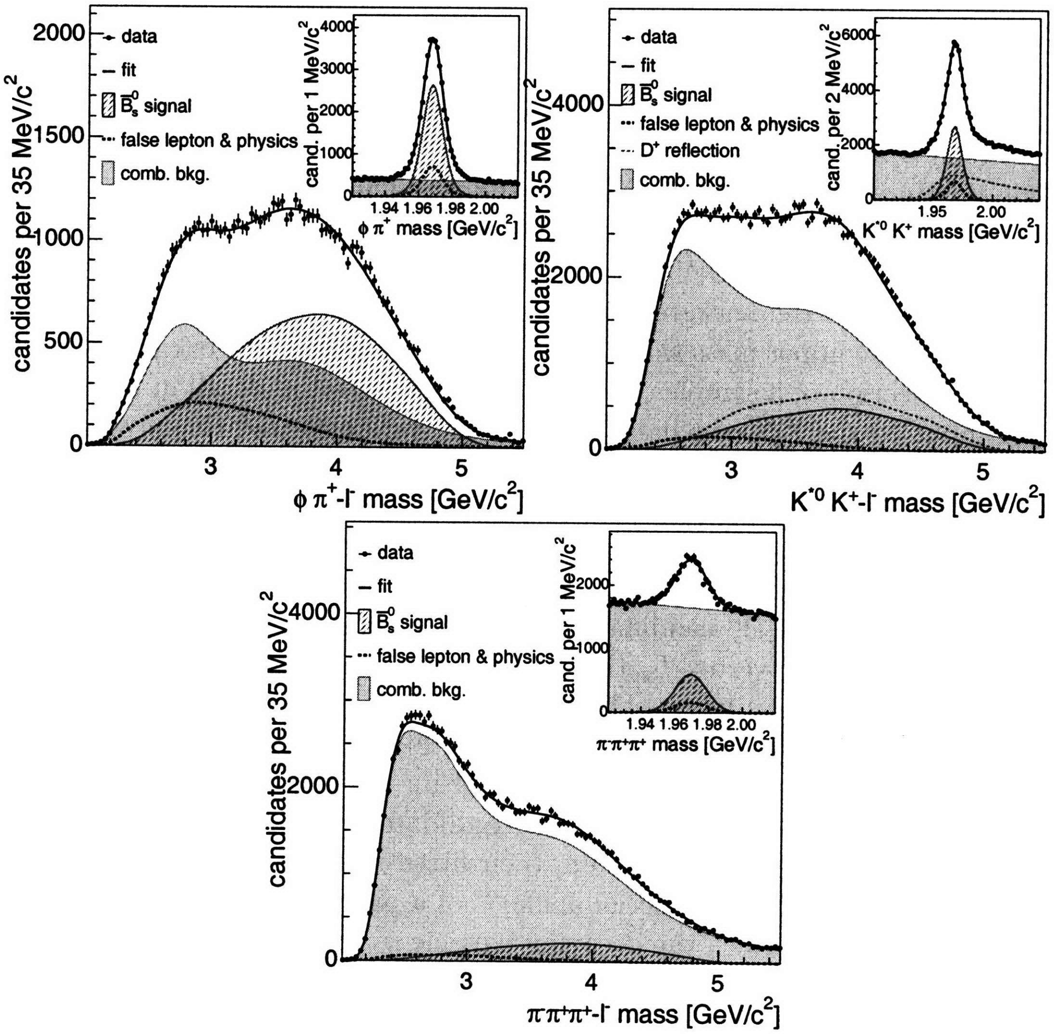

FIGURE 3.3: Mass distributions of the ℓD_s^- part of semileptonic B_s^0 candidates: $D_s^- \rightarrow \phi^0\pi^-$ (left), $D_s^- \rightarrow K^{*0}K^-$ (center) and $D_s^- \rightarrow \pi^-\pi^+\pi^-$ (right). In the insets are the corresponding D_s^- candidate mass distributions.

3.4.2 Fully hadronic B^0_s decays

Fully hadronic modes include $B^0_s \to D^-_s\pi^+(\pi^-\pi^+)$, with the same three D^-_s final states reconstructed as in the case of semileptonic decays. In addition, $B^0_s \to D^-_s\rho^+$ and $D^{*-}_s\pi^+$, with $D^-_s \to \phi^0\pi^-$, are included to the signal sample. Candidates in these modes are looked for in the set of events which satisfy the two displaced track trigger.

The samples of hadronic B^0_s decays suffer for the smallness of the branching ratios of the reconstructed decays, which are roughly an order of magnitude smaller than the semileptonic ones with which they share the same D^-_s final state. However, because B^0_s candidates are fully, or almost fully reconstructed[1], these samples really dominate the $B^0_s - \overline{B}^0_s$ oscillations analysis.

The selection of fully hadronic B^0_s candidates is performed by an Adaptive Neural Network (ANN). A concise presentation of the ANN framework utilized for the candidate selection is described in Reference [63]. On a side note, the application of an ANN-based selection of $B^0_s \to D^-_s\ell^+X$ candidates has been studied too, but because the observed improvement over the utilized cut-based selection was marginal, it has been decided to maintain the already implemented procedure for the selection of semileptonic B^0_s candidates. The ANN utilized in the hadronic B^0_s candidates assigns a single floating-point number to each candidate. The selection is based on a lower cut on the network output, which optimizes $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, where $\mathcal{S}$ is the total amount of signal in the $[5.31, 5.42]$ GeV/c^2 region, estimated from simulated events, and $\mathcal{B}$ is the total amount of the background in the same region estimated by extrapolating the mass fit of the upper mass sideband in data. The neural network package, and its use for the selection of a sample of B^0 decays, is described in detail in Reference [63].

The input to the ANN includes some of the variables traditionally used in cut-based selections. The kinematic of the B^0_s and D^-_s candidates and the quality of their vertex fits are represented by $\chi^2_{r\phi}$, the two-dimensional χ^2 of a fit for a candidate, transverse momentum p_T, impact parameter d_0, transverse decay length with respect to the primary vertex of the interaction L_{xy} and transverse decay length significance $L_{xy}/\sigma_{L_{xy}}$. In the case of D^-_s candidates, the mass m, the transverse decay length with respect to the B^0_s decay vertex $L_{xy}(D^-_s \to B^0_s)$, the decay angle $\cos\theta^*$ and the helicity angle $\cos\psi_H$ are also available. The mass of a D^-_s candidate is constrained to the world average of D^-_s mass measurements in the fit of B^0_s candidates, as described in the beginning of this section, but the result of its unconstrained mass fit is utilized as input to the ANN which performs the candidate selection. When subresonances (ϕ^0 or K^{*0}) are part of the decay chain, their fitted masses, transverse momenta and decay angles are utilized. The decay angle θ^* of a particle is defined as the angle, in the reference frame in which the decaying particle is at rest, between the momentum of a decay product (the pion in $B^0_s \to D^-_s\pi^+$, the D^-_s candidate in $B^0_s \to D^-_s\pi^+\pi^-\pi^+$) and the momentum of the decaying particle, in the laboratory reference frame. In a $B^0_s \to D^-_s \to \phi^0\pi^-$ ($K^{*0}K^-$) decay, the helicity angle ψ_H is the angle, in the reference frame in which the D^-_s candidate is at rest, between the B^0_s candidate and one of the D^-_s decay products. Analogously, $\psi_H(\phi^0, K^{*0})$ is the angle, in the reference frame in

---

[1] In the case of partially reconstructed hadronic decays, 96% of the momentum of a candidate is reconstructed on average.

which the ϕ^0 (K^{*0}) candidate is at rest, between the direction of the D_s^- candidate and one of the ϕ^0 (K^{*0}) decay products. The definitions of helicity and decay angles in the $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$ decay chain are shown in Figure 3.2.

The three reconstructed $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ modes offer additional useful quantities: the mass of the $\pi^-\pi^+\pi^-$ system, the $\chi^2_{r\phi}$ of the fit of the vertex defined by the three tracks and the minimum and maximum masses of opposite-charged track pairs. The last two variables are also available when the D_s^- candidate decays to $\pi^-\pi^+\pi^-$. The transverse-momentum imbalance between K^+ and K^- is used when the decay chain includes a ϕ^0 as intermediate state. Transverse momenta are, as usual, measured in the laboratory frame. Other variables characterize the set of tracks which reconstruct the B_s^0 candidate: transverse momentum p_T of the tracks, minimum and maximum transverse momentum, sum of the track impact parameter significances d_0/σ_{d_0}, maximum separation along the nominal beam direction max $|\Delta z_0|$.

The CLL quantity defined in Equation 6.5.3 provides particle-identification information to the network in the last four input variables: CLL of selected tracks, minimum and maximum CLL and the sum of CLL on all tracks, $\sum(CLL)$. Table 3.7 summarizes which variables have been used in the selection of each B_s^0 decay mode.

The mass distributions of B_s^0 candidates reconstructed in the fully hadronic decay chains are shown in Figure 3.4. The mass distribution of $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$ and partially reconstructed B_s^0 candidates is presented separately from the contribution of the other five decay modes: $B_s^0 \to D_s^-\pi^+$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$, and $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$, $D_s^- \to \phi^0\pi^-$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. In these distributions, the shape and normalization, with respect to the number of B_s^0 signal candidates, of the background contributions from B^0 and $\overline{\Lambda}_b^0$ decays are obtained from B^0 and $\overline{\Lambda}_b^0$ simulated events, as described in Section 4.2. The procedure adopted to calculate the normalization of these contributions is presented in Section 7.2.1.

The yields of B_s^0 candidates reconstructed in fully-hadronic decay chains are reported in Table 3.8. The ratios of signal-over-background in the $[5.32, 5.42]$ GeV/c^2 mass range is also indicated. This sample is by far the largest sample of fully hadronic B_s^0 in the world, which proves the impact of SVT in CDF B physics program.

The selection and reconstruction of B_s^0 *candidates in data are presented in this chapter. The next one will focus on the tools which provide simulated interactions, and on the efforts made to perfect the agreement between simulated events and real data.*

	$B_s^0 \to D_s^-\pi^+$ modes			$B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ modes		
Variable	$\phi^0\pi^-$	$K^{*0}K^-$	$\pi^-\pi^+\pi^-$	$\phi^0\pi^-$	$K^{*0}K^-$	$\pi^-\pi^+\pi^-$
$\chi^2_{r\phi}(B_s^0)$	√	√	√	√	√	√
$d_0(B_s^0)$	√	√	√	√	√	√
$L_{xy}(B_s^0)$	√	√	√	√	√	√
$L_{xy}/\sigma_{L_{xy}}(B_s^0)$	√	√	√	√	√	√
$p_T(B_s^0)$	√	√	√	√	√	√
$L_{xy}(D_s^- \to B_s^0)$	√	√	√	√	√	√
$\chi^2_{r\phi}(D_s^-)$	√	√	√	√	√	√
$d_0(D_s^-)$	√	√	√	√	√	√
$L_{xy}(D_s^-)$	√	√	√	√	√	√
$L_{xy}/\sigma_{L_{xy}}(D_s^-)$	√	√	√	√	√	√
$m(D_s^-)$	√	√	√	√	√	√
$p_T(D_s^-)$	√	√	√	√	√	√
p_T	all	all	all	all	1, 2, 3, 6†	all
$\cos\theta^*(B_s^0)$	√	√	√	√	√	√
$m(\phi^0 \text{ or } K^{*0})$	√	√	–	√	√	–
$p_T(\phi^0 \text{ or } K^{*0})$	√	√	–	√	√	–
$\min(d_0/\sigma_{d_0})$	√	√	√	√	√	√
$\max(d_0/\sigma_{d_0})$	√	√	√	√	√	√
$\min(p_T)$	√	√	√	√	√	√
$\max(p_T)$	√	√	√	√	√	√
$m(3\pi)$	–	–	–	√	√	√
$\chi^2_{r\phi}(3\pi)$	–	–	–	√	√	√
$m^{high}_{3\pi\to\pi^-\pi^+\pi^-}(\pi^-\pi^+)$	–	–	–	√	√	√
$m^{low}_{3\pi\to\pi^-\pi^+\pi^-}(\pi^-\pi^+)$	–	–	–	√	√	√
$m^{high}_{D_s^-\to\pi^-\pi^+\pi^-}(\pi^-\pi^+)$	–	–	√	–	–	√
$m^{low}_{D_s^-\to\pi^-\pi^+\pi^-}(\pi^-\pi^+)$	–	–	√	–	–	√
$\cos\psi_H(D_s^-)$	√	√	–	√	√	–
$\cos\psi_H(\phi^0 \text{ or } K^{*0})$	√	√	–	√	√	–
d_0/σ_{d_0}	all	all	all	all	all	all
$\max\lvert\Delta z_0\rvert$	√	√	√	√	√	√
$\max(CLL)$	√	√	√	√	√	–
$\min(CLL)$	√	√	√	√	√	√
$\sum(CLL)$	√	√	√	√	√	–
CLL	–	–	–	–	–	1, 4††
$\lvert p_T(K^+) - p_T(K^-)\rvert$	√	–	–	√	–	–

TABLE 3.7: ANN input variables for the selection of $B_s^0 \to D_s^-\pi^+(\pi^-\pi^+)$ candidates.
† 1,2,3, and 6 indicate the three pions from the direct decay of the B_s^0 candidate, and the kaon with the lower momentum between the two kaons produced by the decay of the D_s^- candidate.
†† 1 and 4 indicate the pions with the highest transverse momentum among the two sets of three pions produced in the decay of the B_s^0 and of the D_s^- candidates.

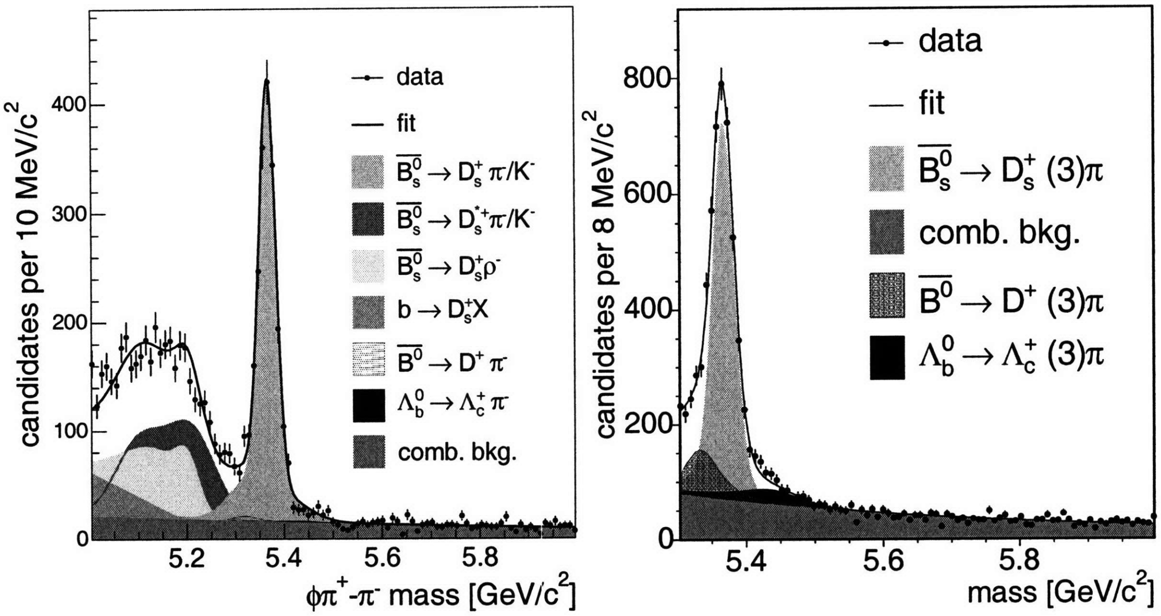

FIGURE 3.4: Mass distributions of fully hadronic B^0_s decay modes. The plot on the left contains the reconstructed $B^0_s \to D^-_s \pi^+, D^-_s \to \phi^0 \pi^-$ candidates. On the right is the sum of the mass distributions of the other hadronic modes: $B^0_s \to D^-_s \pi^+$ with $D^-_s \to K^{*0} K^-$ and $D^-_s \to \pi^- \pi^+ \pi^-$, and $B^0_s \to D^-_s \pi^+ \pi^- \pi^+$ with $D^-_s \to \phi^0 \pi^-$, $D^-_s \to K^{*0} K^-$, and $D^-_s \to \pi^- \pi^+ \pi^-$.

Decay Sequence	Yield	$\mathcal{S}/\mathcal{B}$
$B^0_s \to D^-_s \pi^+, D^-_s \to \phi^0 \pi^-$	1,900	11.3
$B^0_s \to D^-_s \pi^+, D^-_s \to K^{*0} K^-$	1,400	2.0
$B^0_s \to D^-_s \pi^+, D^-_s \to \pi^- \pi^+ \pi^-$	700	2.1
$B^0_s \to D^-_s \pi^+ \pi^- \pi^+, D^-_s \to \phi^0 \pi^-$	700	2.7
$B^0_s \to D^-_s \pi^+ \pi^- \pi^+, D^-_s \to K^{*0} K^-$	600	1.1
$B^0_s \to D^-_s \pi^+ \pi^- \pi^+, D^-_s \to \pi^- \pi^+ \pi^-$	200	2.6
Partially Reconstructed B^0_s Decays	3,300	3.4
Total	8,800	—

TABLE 3.8: Signal yields for the hadronic modes and signal to background ratio in the various decay sequences. The partially reconstructed B^0_s decays are $B^0_s \to D^-_s \rho^+$ and $D^{*-}_s \pi^+$, with $D^-_s \to \phi^0 \pi^-$. The π^0 from the $\rho^+ \to \pi^+ \pi^0$ decay, and the photon from the $D^{*-}_s \to D^-_s \gamma$ decay are not reconstructed. The $\mathcal{S}/\mathcal{B}$ ratio is evaluated in the $[5.32, 5.42]$ GeV/c^2 range of mass of fully reconstructed B^0_s candidates. In the case of partially reconstructed B^0_s decays, the selected mass range is $[5.0, 5.2]$ GeV/c^2. The quoted numbers corresponds to an integrated luminosity of ~ 1.0 fb^{-1}.

Chapter 4

Monte Carlo simulation

This chapter describes the procedure adopted to produce simulated data and defines the different sets of simulated data utilized in this analysis. Then, the additional steps performed to correct for differences between the simulation and the collected data are presented, and finally the comparison of data-simulation agreement is shown. Samples of simulated data which accurately reproduce the actual data collected by the detector are extremely important in many steps of the analysis presented. Most importantly, the calibration of the algorithm of same-side-kaon tagging that constitutes the central point of this dissertation, which is presented in Chapter 6, is shown, in the same chapter, not to be solely obtainable from detector data.

4.1 Overview of simulation procedure

The procedure adopted to simulate data tries to reproduce as closely as possible the various steps that lead from the $p\bar{p}$ interaction to the registration of data. The first step is the simulation of the production of primary particles, such as b-hadrons, which follow the $p\bar{p}$ hard collision. Then, the propagation of the produced particles, their decays, and interactions with the matter of the CDF detector are simulated. Secondary particles, i.e., not produced by the decay of the particles originated from the primary interaction, can be produced during these interactions with the detector. Finally, the detector simulation attempts to reproduce the response of the various parts of the CDF detector to the passage of the simulated particles These steps are divided in two broad categories, which are presented in the next sections.

4.1.1 Event generation

The analysis of B_s^0 oscillations presented in this document utilizes samples of simulated B-hadron samples for different purposes, listed in Section 4.2.

The first step in the Monte Carlo (MC) simulation is the production of primary particles, which are the ones that the simulation indicates as produced at the $p\bar{p}$ interaction point. The physical process of the $p\bar{p}$ hard interaction and b-hadron production is simulated by a so-called *generator*, the output of which is a set of particles with a defined identity (i.e., a b quark, a B^+ meson, a pion, a kaon, ...) and kinematic characteristics (the momentum vector is sufficient to fully describe the particle). The two generators of choice are BGENERATOR [64] and PYTHIA [65]. The former directly produces single b hadrons. The latter also includes the particles produced in association with the b hadrons in the simulated events, by reproducing the fragmentation process, starting from quark strings.

BGENERATOR produces events with one b hadron, according to the distributions of kinematic variables as measured in data. The symmetries in the geometry of the $p\bar{p}$ interaction and, consequentially, of b hadron production make the η and p_T distributions of the primary particle sufficient to completely describe the process. These distributions constitute the input to BGENERATOR for fast event generation. The input distributions utilized to generate MC samples of B mesons and $\overline{\Lambda}_b^0$ baryons are taken from References [24] and [66]. Samples of the different species of B meson, B^0, B^+, and B_s^0, are all generated with the same η–p_T distribution. The single-hadron MC samples produced with BGENERATOR are employed whenever the details of a $p\bar{p}$ interaction are not needed and it is only necessary to understand how the description of an ideal candidate is modified by detector effects and the trigger selection. The advantage of BGENERATOR over PYTHIA is the speed of event generation, which is about an order of magnitude faster.

PYTHIA aims at simulating the full range of particles produced in a $p\bar{p}$ interaction as faithfully as possible, and trying to reproduce experimental observations, such as the multiplicity distribution and type of charged particles within the leading (i.e., highest transverse momentum) charged jet in an event, within the limits of the current understanding of the underlying physics. The physical model adopted to describe the fragmentation process is the *string fragmentation* one [67, 68], as opposed to independent fragmentation [69, 70] and cluster fragmentation [71] models. The MC model for QCD hard scattering provided by PYTHIA in its default configuration describes fairly well the properties of the leading charged jet in an event. These properties include the multiplicity distribution and the p_T distribution of charged particles within leading charged jets, the size of leading charged jets, and the radial distribution of charged particles and transverse momentum around the direction of the leading charged jet. However, all the properties of the *underlying event*, which consists of the beam remnants after a $p\bar{p}$ interaction and possible additional contributions from multiple parton scatterings, are not correctly described. For example, the p_T dependence of the beam-beam remnant contribution to the underlying event is not correctly reproduced. The default configuration of PYTHIA is thus modified following the tuning described in References [72] and [73], in order to address the data-simulation disagreement in the description of the underlying event properties. The tuned configuration is defined in Appendix A. The additional modifications most relevant for the presented analysis are reported in Section 4.3.

The framework for the implementation of the decay process of the produced B mesons is provided by the EvtGen [74] package. In the simulation of a decay chain, such as $B \to D^*[D\pi]\ell\nu$, the module uses decay amplitudes at each node of the chain, instead of probabilities, which allow for the correct simulation of all decay distributions. The physical properties of particles, such as the mass m and, for unstable particles, the lifetime τ and the branching ratios in possible decaying modes, are input to EvtGen. In order to focus on particular decay channels of interest, the user typically overrides the branching ratio settings of particles to force their decay.

4.1.2 Detector simulation

The GEometry ANd Tracking (GEANT [75]) framework allows for the simulation of the interaction of particles with the materials composing the detector. The response of the detector to the incident particles is described in terms of sets of hits and energy depositions in the various subdetectors. The package replicates the response of the CDF detector, and produces an output in the same format as obtained from the detector readout.

The simulation is very computing-power intensive. As a rule of thumb, the BGENERATOR production of 60 million events with a single B_s^0 being the primary particle, generated with $p_T > 0$ and $|\eta| < 10$ and the GEANT simulation of their interaction with the CDF detector require about 4 to 5 thousand hours on an individual CPU. This step yields, in the case of $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$, only 950 thousand events after the trigger selection. This sample is subsequently reduced to 60 thousand after B_s^0 candidates are reconstructed and analysis cuts applied.

All the MC samples used in many steps of the analysis presented in this document utilize the full-fledged detector GEANT simulation. The lists of the types of MC samples produced and of the aspects of the mixing analysis in which they have been necessary are presented in Section 4.2.

4.2 Monte Carlo samples

The MC samples utilized in the mixing analysis are organized in three separate categories. These classes and the rôle of the MC simulation in the mixing analysis are reviewed in this section. The MC categories are distinguished by the definition of the initial state:

- **Single-hadron**
  BGENERATOR is used to produce a single B meson or $\overline{\Lambda}_b^0$ baryon. The decay is simulated by EvtGen and the chain of decay products is completely defined by adequately setting the branching ratios for the decays of the particles in the chain.

- **Semi-generic**
  For the samples in this class, BGENERATOR produces a single hadron, the decay chain of which is not tightly defined as in the previous case. Each event still features a single hadron as parent particle, but all types of B mesons and $\overline{\Lambda}_b^0$ baryons are produced, according to the production fractions [62]: $f_u : f_d : f_s : f_{baryon} = 0.397 : 0.397 : 0.107 : 0.099$. A filter based on particle content is

applied, with the unique requirement that a D meson, the decay of which is forced to a specific mode, is produced in the decay of the primary b hadron.

- ***$b\bar{b}$***
  PYTHIA is the generator of choice for the samples which fall in the last category. The production consists of two steps. Firstly, sets of $q\bar{q}$ events are generated simulating three different production processes: flavor creation, flavor excitation and gluon splitting [76]. The simulated processes are $qq' \to qq'$, $q\bar{q} \to q'\bar{q}'$, $q\bar{q} \to gg$, $qg \to qg$, $gg \to q'\bar{q}'$, and $gg \to gg$, where q and q' are fermions, and g gluons. Then, events which contain $b\bar{b}$ pairs with at least one b quark with $p_T(b) > 4$ GeV/c^2 and $|\eta| < 3$ are selected. The output after this first step is a set of events containing b hadrons and the other particles produced during the fragmentation. The second and final step consists in the simulation of the time-evolution of the particles up to their decay, which is performed by the EvtGen package. Different configurations of the decay package are utilized, tailored to the diverse studies performed, starting from the same initial sample of particles obtained after the first step of the simulation.

The first class of MC samples allows the study of how the distributions of interesting quantities are modified by detector and selection effects for very specific decay chains. The B_s^0 final states included in the analysis have been simulated to study the mass and proper-decay-time distributions of background-free signal candidates. The case of proper decay-length *ct* is particularly important because the trigger and analysis selections peculiarly sculpt the *ct* distribution. The availability of an accurate detector simulation allows for the precise modeling of such sculpting effects, thus restoring the direct relation between the observed distribution of reconstructed *ct* and the expected *ct* distribution of a decaying particle with lifetime τ. A detailed description of the procedure adopted to correct for these sculpting effects is reported in Section 7.1.

Specific decays of B^0 and $\overline{\Lambda}_b^0$, namely $B^0 \to D^-\pi^+(\pi^-\pi^+)$, $D^- \to K^+\pi^-\pi^-$ and $\overline{\Lambda}_b^0 \to \Lambda_c^-\pi^+$, $\Lambda_c^- \to \bar{p}K^+\pi^-$, mimic the signature of the B_s^0 decays of interest. These modes contribute to backgrounds in the distributions of mass and proper decay-length of reconstructed B_s^0 candidates. Their contributions are modeled by template functions, the shape and normalization of which are defined by analyzing BGENERATOR-MC samples of B^0 and $\overline{\Lambda}_b^0$, which are forced to decay via the modes listed above.

The more generic samples in the second category play a major rôle in the characterization of the physics backgrounds arising from partially reconstructed decays of B mesons. The samples are constructed with the requirement that a D meson, which decays as a signal, is present in the final state. This requirement stems from the confidence that the largest fraction of the backgrounds which populate the low mass sideband in the reconstructed B-mass distributions is constituted by partially reconstructed B candidates, and that a real D meson is correctly fitted as part of the B candidate. The prescription adopted to produce MC samples in this category is generic enough to enable one to characterize and subsequently quantify the sources of background to a B decay mode. For example, a MC sample produced with the prescription that each event contains a $D_s^- \to \phi^0\pi^-$ meson allows one to study the physics backgrounds of the $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$, with $D_s^- \to \phi^0\pi^-$, decay modes. The first implementation of this method for the study of backgrounds in low mass sidebands is presented in References [77] and [78].

The PYTHIA-generated MC set contains the most complete simulation of a $b\bar{b}$ event and is thus the only resource for studies which require a description of the track environment in which B_s^0 candidate are searched for. This is particularly important for the study of the same-side tagger. In fact, this algorithm, which is utilized to determine the flavor, B_s^0 vs. $\overline{B}_s^0$, at production of a candidate, relies on the simulation of the other interaction products.

4.3 Monte Carlo tuning

The MC simulation produces results in good agreement with the available data, as will be shown in Section 4.4. The decay mechanism of the hadrons which take part to the reconstructed B_s^0 decay chains is well understood and the software does a fairly accurate job in reproducing the response of the detector to the passage of particles.

However, some additional efforts were needed to tune some aspects of the MC simulation that have direct effects on the analysis. Particularly critical is the part of the simulation which influences the calibration of the algorithm for same-side flavor tagging presented in Chapter 6, which is derived via MC events.

The various aspects which required tuning are separately discussed below. In Section 4.4, the success of the tuning is demonstrated by comparing data and simulation.

4.3.1 Trigger prescaling

The detector simulation framework does not reproduce the prescaling mechanisms utilized by the CDF trigger systems . Two alternative approaches have been adopted to account for the effects of prescaling in the simulation. The BGENERATOR-MC simulation implements prescaling by selecting events with a probability equal to the inverse of the prescaling factor. Such a prescaling factor is applied on a run-by-run basis to the sample and represents the effective prescale which accounts for different trigger paths with different prescaling methods. In the case of the PYTHIA-MC one, the need to maximize the statistical power by not throwing events away prescribes a different method. A statistical weight is associated to each event which depends on the trigger path (`BCHARM`, `HIGHPT`, or `LOWPT`, as defined in Section 3.1) to which it belongs and the period of data taking during which it was selected (`0d`, `0h`, or `0i`, Section 3.2). The weights utilized for the $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ decay mode are reported in Table 4.1.

4.3.2 Vertex position and resolution

The reproduction of the correct distribution of the positions of vertices and the resolutions of the position measurements enters the analysis at different levels. The

	0d	0h	0i
HIGHPT	1.431	1.268	1.416
BCHARM	1.037	1.050	1.158
LOWPT	0.477	0.625	0.578

TABLE 4.1: Statistical weights utilized to reproduce effective trigger prescales. The values in the table are applied to $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ MC events.

variables which are mostly influenced by these quantities are impact parameters of reconstructed particles (in this analysis, B and D mesons) d_0, transverse decay lengths L_{xy} and transverse decay length resolutions $\sigma_{L_{xy}}$. The transverse decay length L_{xy} of a B candidate is defined as follows:

$$L_{xy}(\mathrm{B}) = \frac{\mathbf{p_T} \cdot (\mathbf{r}_{xy}(SV) - \mathbf{r}_{xy}(PV))}{p_T}, \tag{4.3.1}$$

and indicates the projection on the transverse momentum of the particle of the distance between the primary vertex (PV, the vertex at which the $p\bar{p}$ interaction occurred) and the secondary vertex (SV, the vertex where the reconstructed particle decayed). In the case of a D meson produced in the decay of a B, the quantity $L_{xy}(\mathrm{D} \to \mathrm{B})$, which is defined as the distance, in the transverse plane, between the D decay vertex and the B decay vertex projected along the direction of the D transverse momentum, more properly defines a "decay" length. However, the notation $L_{xy}(\mathrm{D})$, throughout this document, defines the projection of the distance between the PV and the D decay vertex. The B impact parameter is utilized in the candidate selection, while $\sigma_{L_{xy}}$ contributes to the proper-decay-time error of B candidates, which is an important part of this analysis of B_s^0 oscillations (Sections 5.2 and 7.1).

The algorithm for PV reconstruction chosen for the mixing analysis utilizes some of the tracks which are believed to come from the same interaction point where the B_s^0 candidate was produced, and fits them to a common vertex using the beamline shape [79] as a fit constraint. This constraint conceptually represents an *a priori* probability for the position of primary vertices. The shape of the beamline around the interaction point at CDF is described as an hourglass, $\Delta x \simeq \Delta y \simeq 2$ mm for $|z| \simeq 30$ cm, and $\Delta x \simeq \Delta y \simeq 35$ μm for $z \simeq 0$ cm. The fitted vertex is referred to as the event-by-event PV [80]. This approach is not directly reproducible in the BGENERATOR-MC sets, because bare B_s^0 particles are generated. Therefore, no other tracks beside the ones which are produced by the B_s^0 decay are present in the event, and the algorithm utilized in data cannot be applied. Additional procedures have to be implemented to reproduce the distributions observed in data.

By default, BGENERATOR primary vertices are distributed according to the beamline shape. In the case of data, the distribution of primary vertices is dictated by the results of the event-by-event PV finding algorithm.

Residuals between the coordinates of the MC generated vertex and the reconstructed one of a B_s^0 candidate are distributed according to the the beam line covariance matrix $\mathcal{C}_{PV}$ evaluated at the z position of the B_s^0 candidate. This constitutes another discrepancy with respect to data. While in MC events the uncertainty with which each PV is determined depends only on its z coordinate, in data it depends also on the properties of the underlying event (i.e., the additional tracks produced at the interaction point).

It is thus necessary to tune MC candidates in order to correct for the differences introduced by using a different algorithm for PV finding in data and in MC simulation, i.e., event-by-event vertices *vs.* beamline vertices. The tuning of MC events is obtained by applying scale factors to $\mathcal{C}_{PV}$ and $\mathcal{C}_{SV}$ (i.e., the error matrix returned by the fit of a B secondary vertex). The values of the scale factors are obtained by scanning the space of their possible values while comparing distributions of quantities in data and in MC samples. The distributions of the quantities which are more directly affected by the scaling, such as impact parameters and transverse decay length resolutions, are utilized to tune the scale factors. The final values for the scale factors are chosen by minimizing the disagreement (indicated by a χ^2 test) between data and MC simulation. The result of the scaling is the modification of both the distributions of PV positions and of the covariance matrices of primary and secondary vertices.

The first type of scale factor, S_1, is meant to correct the PV distribution. The reconstructed position of the PV in the transverse plane is replaced with the transverse coordinates:

$$\mathbf{x}_{rec}^{new} = \mathbf{x}_{true} + S_1 \boldsymbol{\delta} \,, \tag{4.3.2}$$

where $\boldsymbol{\delta}$ is a 2D vector drawn from a 2D Gaussian distribution of variance $\mathcal{C}_{PV}^{xy}$, centered at $(0,0)$. Two different scale factors are defined, $S_1(\mathrm{B})$ and $S_1(\mathrm{D})$, to be used in the recalculation of quantities characterizing B and D candidates, respectively. The value of $S_1(\mathrm{B})$ is obtained by minimizing the χ^2 between the distributions of $d_0(\mathrm{B})$ in data and MC samples. The PV reconstruction method utilized to calculate quantities relative to D candidates does not use the other tracks present in the event, and is thus equally applicable to data and BGENERATOR-MC simulation. Therefore, the scaling factor $S_1(\mathrm{D})$ is equal to unity.

The second and third type of scale factors, S_2 and S_3, address the data–MC-simulation discrepancy in $\sigma_{L_{xy}}$, the resolution on the projection of the particle decay length on the transverse plane. This quantity is affected by the disagreement in the covariance matrix of both the primary (PV) and the secondary (SV, of the B or D decay) vertices. The error on L_{xy} is calculated as follows:

$$\sigma_{L_{xy}} = \begin{pmatrix} \mathbf{p_T}^T & -\mathbf{p_T}^T \end{pmatrix} \mathcal{M} \begin{pmatrix} \mathbf{p_T} \\ -\mathbf{p_T} \end{pmatrix} , \quad \mathcal{M} \equiv \begin{pmatrix} S_3 \mathcal{C}_{SV} & 0 \\ 0 & S_2 \mathcal{C}_{PV} \end{pmatrix} . \tag{4.3.3}$$

The quantities appearing in the above formula are defined as follows:

$$\mathbf{p_T} = \begin{pmatrix} p_x \\ p_y \end{pmatrix} , \quad \mathcal{C} = \begin{pmatrix} \mathcal{C}^{xx} & \mathcal{C}^{xy} \\ \mathcal{C}^{yx} & \mathcal{C}^{yy} \end{pmatrix} . \tag{4.3.4}$$

The tuning of $S_2(\mathrm{B})$, $S_2(\mathrm{D})$, and S_3 is performed in parallel by minimizing the disagreement between data and MC distributions of $\sigma_{L_{xy}}(\mathrm{B})$ and $\sigma_{L_{xy}}(\mathrm{D})$.

The $\mathrm{B}^0 \to \mathrm{D}^-\pi^+$, $\mathrm{D}^- \to \mathrm{K}^+\pi^-\pi^-$ sample was used to derive the tuning parame-

Scale factor	Description	Value
S_1(B)	re-smearing of the PV for B quantities	0.780
S_1(D)	re-smearing of the PV for D quantities	1
S_2(B)	uncertainty scale factor for B quantities	0.560
S_2(D)	uncertainty scale factor for D quantities	0.900
S_3	secondary vertex uncertainty scale factor	1.145

TABLE 4.2: Scaling factors for tuning of BGENERATOR-MC events, to address the event-by-event–PV algorithm used in data. The tuning has been performed using samples of $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$.

ters. The tuning was then verified on other decay modes such as $B^0 \to D^-\pi^+\pi^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ and $B^0_s \to D^-_s\pi^+(\pi^-\pi^+)$, $D^-_s \to \phi^0\pi^-$. Table 4.2 summarizes the numerical values of the tuning parameters.

4.3.3 Silicon hit resolution

The MC efficiency to find hits in L00 and match them to tracks is significantly larger than in data. In addition, the hit resolution in SVX layers is better in MC simulation than in data, while ISL hits present the same resolution. However, the hit resolutions in L00 and in the first layer of SVX, L0, dominate the error on the measurement of impact parameters and transverse decay lengths, because the information of the other layers is smeared by multiple scattering and transport uncertainties.

The procedure adopted to correct these discrepancies was developed utilizing data and MC samples of $B^0 \to J/\psi K^{*0}$, $J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$. The tracks in these events which are not part of the reconstructed B^0 candidate and satisfy some minimal quality requirements ($p_T > 450$ MeV/c, $d_0/\sigma_{d_0} < 10$, and at least 10 hits in the axial layers of the COT and 10 in the stereo layers) participate in the study [81].

The disagreement between the number of L00 hits assigned to tracks in data and in MC events is corrected for by randomly discarding 33% of the hits found in the innermost silicon layer. In fact, the fraction of tracks, in the sample described above, with a hit in L00 is 52% in MC events and only 35% in data. Because all tracks, both in data and in the MC sample, are refit after the addition of L00 hits at the analysis level, the operation of hit removal does not introduce any difference between the treatment of data and of MC events.

The single hit resolutions are modified by smearing hit positions according to Gaussian distributions. The widths of these distributions depend on the number of strips which constitute the hit signal, N_{strip}, and whether the hit is in L00 or in the first layer of SVXII, L0:

$$\begin{aligned}
\sigma(N_{strip}) &= \sigma_{L00 \text{ or } L0}\sqrt{\sigma^2_{data}(N_{strip}) - \sigma^2_{MC}(N_{strip})}\,, \\
\sigma_{L00} &= 1.2\,, \\
\sigma_{L0} &= 1.5\,,
\end{aligned} \tag{4.3.5}$$

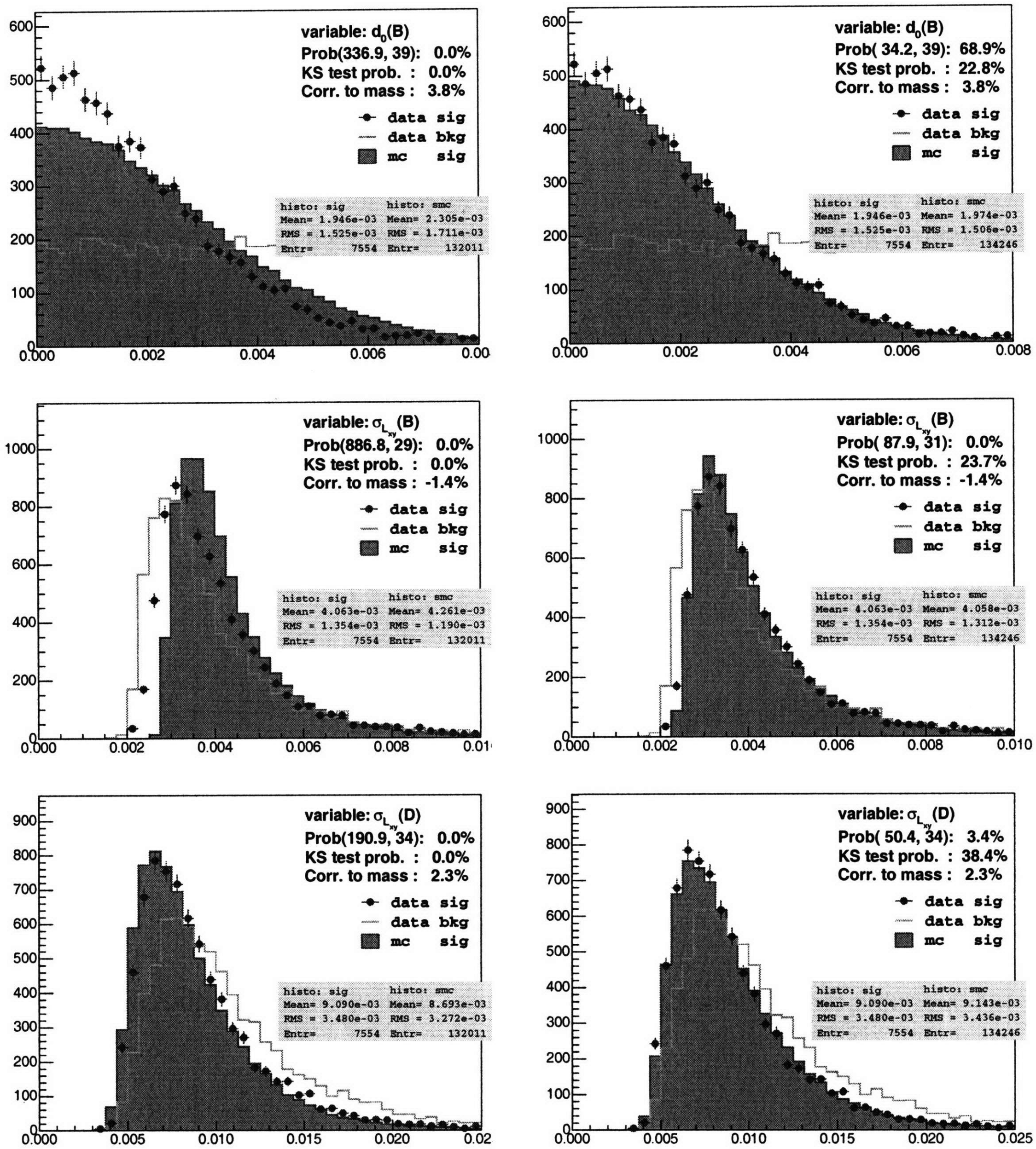

FIGURE 4.1: Data-simulation comparison before/after tuning. From top to bottom, the distributions of $d_0(\mathrm{B}^0)$, $\sigma_{L_{xy}}(\mathrm{B}^0)$ and $\sigma_{L_{xy}}(\mathrm{D}^-)$, in $\mathrm{B}^0 \to \mathrm{D}^-\pi^+, \mathrm{D}^- \to \mathrm{K}^+\pi^-\pi^-$, before (left) and after (right) the tuning described in Section 4.3.2.

$[10^{-4}$ cm$]$	Data				MC simulation			
N_{strip}	1	2	3	>3	1	2	3	>3
$L00$	17.3	15.0	18.5	30.3	13.5	9.4	11.6	29.8
$L0$	13.6	9.5	13.4	19.0	10.1	10.1	14.7	23.1

TABLE 4.3: Resolution of $r\varphi$ hits in data and MC simulation.

where $\sigma_{data,MC}$, reported in Table 4.3, are provided by the CDF tracking group [82] and the use of MC truth information (the intersection in the silicon layer found by GEANT is compared to the reconstructed hit position). The σ_{L00} and σ_{L0} factors are added to compensate for a residual disagreement between the distributions of $L_{xy}(\mathrm{B}^+)$ and $\sigma_{L_{xy}}(\mathrm{B}^+)$ in the data and PYTHIA-MC samples of $\mathrm{B}^0 \to J/\psi\mathrm{K}^{*0}$, where L_{xy} is the quantity defined in Equation 4.3.1. The presence of a residual inconsistency was not surprising as only single-Gaussian fits were performed to obtain the hit resolutions in the MC sample.

The tuning of the hit efficiency and the σ_{L00} and σ_{L0} factors has been cross-checked in an alternative B-meson decay mode, $\mathrm{B}^0 \to \mathrm{D}^-\pi^+\pi^-\pi^+$, $\mathrm{D}^- \to \mathrm{K}^+\pi^-\pi^-$. The data and BGENERATOR-MC samples of $\mathrm{B}^0 \to \mathrm{D}^-\pi^+\pi^-\pi^+$ are compared before and after the application of the tuning, which provides an improvement in the agreement between the two samples [83]. The correction derived in the tuning is applied to all MC samples.

4.3.4 Particle identification

Particle identification is based on specific ionization $d\mathrm{E}/d\mathrm{x}$ in the COT and information from the TOF system. Both subdetectors have been calibrated and the probability density functions of their response to different particles precisely modeled with data [40].

One necessary ingredient to simulate the particle identification in MC events is to know the truth information of the particles associated to the reconstructed tracks. It is possible to match $> 99.9\%$ of the tracks that satisfy the requirements for being a tagging track candidate to Monte Carlo truth information, using standard CDF matching tools [84]. However, only 98.5% of these tracks are actually associated to generator level particles[1], while the remaining 1.5% of tracks are associated to particles which have been produced inside the silicon detector. Because those particles entirely transverse the COT, their COT $d\mathrm{E}/d\mathrm{x}$ response is correctly simulated. The situation is different for the TOF simulation. Although the species of those particles are known, their production time is not, and thus their TOF response cannot be simulated properly. The default simulation calculates their response assuming these particles have the same production time as the main interaction, regardless of the fact that they are secondaries. (Section 6.8.5).

---

[1] generator level particles are those generated at the simulated $p\bar{p}$ interaction point, or part of the decay chain of particles produced at the interaction point.

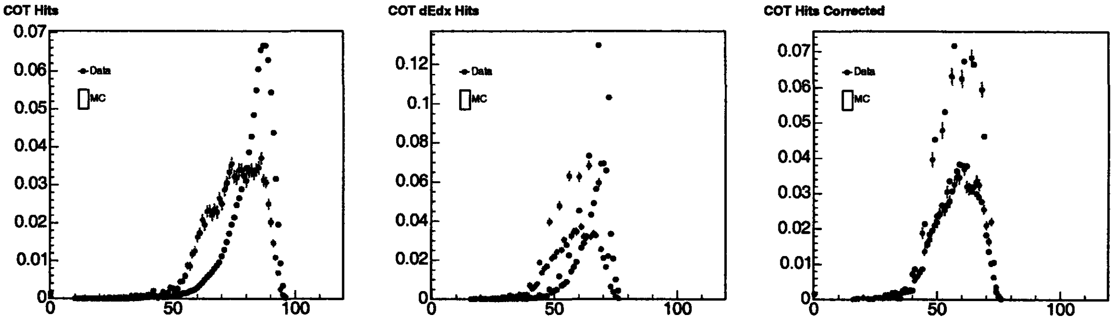

FIGURE 4.2: Number of COT hits (left) and COT hits utilized for the calculation of $d\mathrm{E}/d\mathrm{x}$ information (center) per track in data and simulation. The right plot shows the data-simulation comparison of the number of COT hits used to calculate a track's $d\mathrm{E}/d\mathrm{x}$ after correction. A fraction equal to 80% of the COT hits associated to a track is utilized to calculate $d\mathrm{E}/d\mathrm{x}$ of the track. The truncation of the number of COT hits produces the spikes in the distributions of data and MC events in the central and rightmost plots.

Energy loss in the COT

The simulation of the COT response is done in two steps. The number of COT hits is rescaled to match the distribution in data, correcting for the algorithm utilized in MC reconstruction which is more efficient in assigning hits. The efficiency with which COT $d\mathrm{E}/d\mathrm{x}$ information is made available to a track is corrected by a function of the transverse momentum p_T:

$$f(p_T) = a - e^{-bp_T}\,. \tag{4.3.6}$$

The functional form has been chosen by inspecting the distribution, as a function of the track transverse momentum, of the ratios of the number of COT hits with $d\mathrm{E}/d\mathrm{x}$ information in data and in MC samples. The parameters a and b are obtained by fitting the same distribution with the function in Equation 4.3.6. The distributions of COT hits with $d\mathrm{E}/d\mathrm{x}$ information in data and in MC events, before and after the correction described above, are shown in Figure 4.2.

Secondly, the MC track needs to be assigned a value of $(d\mathrm{E}/d\mathrm{x})^{raw}$, which indicates the value of $d\mathrm{E}/d\mathrm{x}$ before calibration. The formulae which implement the calibration of the measured $d\mathrm{E}/d\mathrm{x}$ of a track (the label "raw" indicates uncalibrated measurements) are reported here:

$$\begin{aligned} (d\mathrm{E}/d\mathrm{x})^{cor} = &\quad \alpha + \beta(d\mathrm{E}/d\mathrm{x})^{raw} && \texttt{0d}\ \textit{data}, \\ (d\mathrm{E}/d\mathrm{x})^{cor} = &\quad e^{\gamma}(d\mathrm{E}/d\mathrm{x})^{raw} && \texttt{0h}\ \textit{and}\ \texttt{0i}\ \textit{data}, \end{aligned} \tag{4.3.7}$$

where the α, β, and γ coefficients are provided by calibrations and depend on mass, momentum, φ_0, η, charge, number of COT hits with $d\mathrm{E}/d\mathrm{x}$ information, and run number of the tracks utilized in the calibration.

The $(d\mathrm{E}/d\mathrm{x})^{raw}$ value of the track is computed in the following steps:

- the reconstructed track is matched to a generated particle, thus permitting the assignment of the true particle type (i.e., a pion, a kaon, or a proton) to the track.

- the particle-type information is used to select the corresponding Z distribution, defined as:
$$Z \equiv \log\left[\frac{(d\mathrm{E}/dx)^{cor}}{(d\mathrm{E}/dx)^{pre}}\right], \tag{4.3.8}$$
where $(d\mathrm{E}/dx)^{cor}$ and $(d\mathrm{E}/dx)^{pre}$ are the calibrated and expected $d\mathrm{E}/dx$ of a track, with a specific particle-type hypothesis. The particle-type information is utilized to calculate the correct $(d\mathrm{E}/dx)^{pre}$, which depends on the particle's velocity. The variable Z is parameterized by a single Gaussian distribution, when $(d\mathrm{E}/dx)^{pre}$ is calculated with the correct particle-type hypothesis. The $(d\mathrm{E}/dx)^{pre}$ quantity is calculated using a variant of the Bethe-Bloch [85] curve, with the particle speed as input (β and γ):
$$(d\mathrm{E}/dx)^{pre} = \frac{1}{\beta^2}\left[c_1 \log\left(\frac{\beta\gamma}{b+\beta\gamma}\right) + c_0\right] + a_1(\beta-1) + a_2(\beta-1)^2 + C\,, \tag{4.3.9}$$
with a_i, b, c_i, and C parameters extracted from data, utilizing samples of kaons, protons, and pions obtained by reconstructing the $\mathrm{D}^*(2010)^+ \to \mathrm{D}^0[\mathrm{K}^-\pi^+]\pi^+$ and $\Lambda^0 \to p\pi^-$ decay modes.

- a random number δ_Z for the Z variable is generated, following a Gaussian distribution with width σ_Z obtained from the same $\mathrm{D}^*(2010)^+$ and Λ^0 calibration from above of the particle under study. This random number is defined as follows:
$$\delta_z = \mathrm{random}\left[\mathcal{G}(Z, \sigma_Z)\right]. \tag{4.3.10}$$
Thus, the corresponding $(d\mathrm{E}/dx)^{cor}$ assigned to the track is
$$(d\mathrm{E}/dx)^{cor} = e^{\delta_z}(d\mathrm{E}/dx)^{pre}, \tag{4.3.11}$$
where the particle identity, which the MC truth information revealed, is necessary to calculate $(d\mathrm{E}/dx)^{pre}$.

- utilizing the inverse of the functions in Equation 4.3.7, a new value of $(d\mathrm{E}/dx)^{raw}$ is computed and assigned to the track, with the $(d\mathrm{E}/dx)^{cor}$ obtained in the previous step used as input.

The last step, which may appear as a useless calculation, is required to be able to pass the MC sample through the same analysis code utilized for data events, which performs the $d\mathrm{E}/dx$ calibration in Equation 4.3.7.

Time-of-flight

The MC simulation of the TOF detector was still preliminary at the time of this analysis. Therefore, it was decided to develop a method for the generation of TOF information for MC tracks which provides a representative simulation of the TOF behavior and performance.

The available TOF simulation provides a good model for effects related to occupancy. It was decided to profit from that part of the simulation by recalculating the t_{flight} simulated by the MC executables for the tracks which are matched to a TOF pulse. Studies of the use of the TOF system for particle identification, such as in Reference [86], provide a parameterization of the t_{flight} residual defined as follows:

$$\Delta t_{\text{flight}} \equiv t_{\text{flight}}^{meas} - t_{\text{flight}}^{pre}, \tag{4.3.12}$$

where the predicted t_{flight} is a function of the particle mass m, its momentum p, and the path-length L traveled before reaching the TOF detector:

$$t_{\text{flight}}^{pre} = \frac{L}{pc}\sqrt{p^2 + m^2c^2}. \tag{4.3.13}$$

The distributions of t_{flight} residuals for different types of particles are produced utilizing the $\mathrm{D}^*(2010)^+$ and Λ^0 samples already utilized for the study of particle-identification with $d\mathrm{E}/dx$ information. The functional form adopted to fit the Δt_{flight} distribution is a sum of six Gaussian functions. The resolutions of the Gaussian used for the `0h` and `0i` data are multiplied by a factor 1.15 with respect to the `0d` data sample to account for the 15% worsening of the t_{flight} resolution observed in the more recent data samples [87].

The t_{flight} that is assigned to MC tracks is computed as the sum of a random number generated with a p.d.f. which reproduces the Δt_{flight} distribution from data plus the expected t_{flight} for the specific particle. The expected t_{flight} includes the pure theoretical value obtained from kinematic properties of the track and a correction, derived from data, which is dependent upon the particle species. This last correction is derived from the samples of kaons, pions, and protons which are used in many aspects of the studies of particle identification. Its introduction allows one to reduce the dependence on the particle type of the distributions of Δt_{flight}. Independence of the particle type is important because the combination of COT and TOF particle identification, introduced in Section 6.5, assumes that Δt_{flight} is distributed independently of the particle species.

Finally, the efficiency with which TOF information is matched to tracks is larger in simulated than in real data. The ratio of the efficiencies in data and simulation, R_ϵ, has a p_T-dependence which is modeled by a second order polynomial:

$$R_\epsilon(p_T) = a + bp_T + cp_T^2. \tag{4.3.14}$$

Three sets of values for the a, b, and c parameters have been fit, one for each of the three periods of data-taking `0d`, `0h`, and `0i`. The fitted sets of values are summarized in Table 4.4. Figure 4.3 shows the t_{flight} efficiencies in MC simulation and in data, and the distribution of R_ϵ, for all tagging track candidates in a sample of $\mathrm{B} \to \mathrm{D}\pi$ modes, in the `0d`, `0h`, and `0i` data and MC samples. The distributions of R_ϵ are fit with the function of Equation 4.3.14. The plots show the result of the fit and the

Parameter	`0d`	`0h`	`0i`
a	0.901 ± 0.049	0.743 ± 0.043	0.789 ± 0.056
b $[(\mathrm{GeV/c})^{-1}]$	0.010 ± 0.007	0.026 ± 0.007	$0.026^{\dagger}$
c $[(\mathrm{GeV/c})^{-2}]$	0.025 ± 0.022	0.008 ± 0.019	$0.008^{\dagger}$

TABLE 4.4: Values of parameters for t_{flight} efficiency correction.
† The b and c parameters were fixed to the value obtained in the fit of `0h` data.

curves utilized for the evaluation of systematic errors, which correspond to a ±8% (±10%) variation of the efficiency ratio for the `0d` (`0h` and `0i`) period of data-taking.

After the correction of the p_T dependence, no residual η dependence is seen. It has also been checked that R_ϵ is not dependent on the charge of the tracks, by separately comparing the distributions of the efficiency ratios for positively and negatively charged tagging track candidates. Moreover, because the performance of the TOF system is correlated to the occupancy, data-simulation efficiencies are compared for "early" and "late" runs in the `0d` period (run number < 169000, or > 169000, corresponding to September 2003), and for "low" and "high" luminosity (lower or higher than $25 \cdot 10^{30}$ cm^{-2}s^{-1}). The observed differences in the efficiency ratios are well covered by the band of variation chosen for systematics studies.

4.3.5 Additional tunings for Same-side tagging

This section contains the additional tunings that are needed to reproduce those features of data which are important for the same-side-kaon flavor tagging algorithm, which will be presented in Chapter 6.

Multiple interactions

The average number of $p\bar{p}$ interactions per bunch crossing is 2.3 for an instantaneous luminosity of 10^{32} cm^{-2}s^{-1}. The simulation contains only single $b\bar{b}$ events and thus lacks, by default, the possibility that the event actually contains additional tracks from another interaction, which is referred to as a *pile-up* event. Among these tracks, the algorithm for same-side flavor tagging described in Chapter 6, may select a tag candidate if the track satisfies the requirements in Section 6.3. Such selection provides a random tagging decision, because the flavor of the B meson produced in one $p\bar{p}$ interaction is uncorrelated with the charge of the particles produced in another $p\bar{p}$ interaction. To account for this effect, which reduces the performance of the same-side flavor tagger and becomes more important for increasing luminosity, a sample of candidate tagging tracks has been extracted from the data sample of reconstructed $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \to \phi^0 \pi^-$ decays and added to the PYTHIA-MC sample utilized to study the performance of same-side tagging algorithms.

Firstly, the number of additional potential tracks needs to be estimated. The number of tracks to be added to the N^{th} event is determined by counting the number of tracks in the $(N+1)^{th}$ event in the MC sample which are within ±1.2 cm, the width of the B_s^0 signal region, of the B_s^0 vertex in the N^{th} event, and satisfy the selection

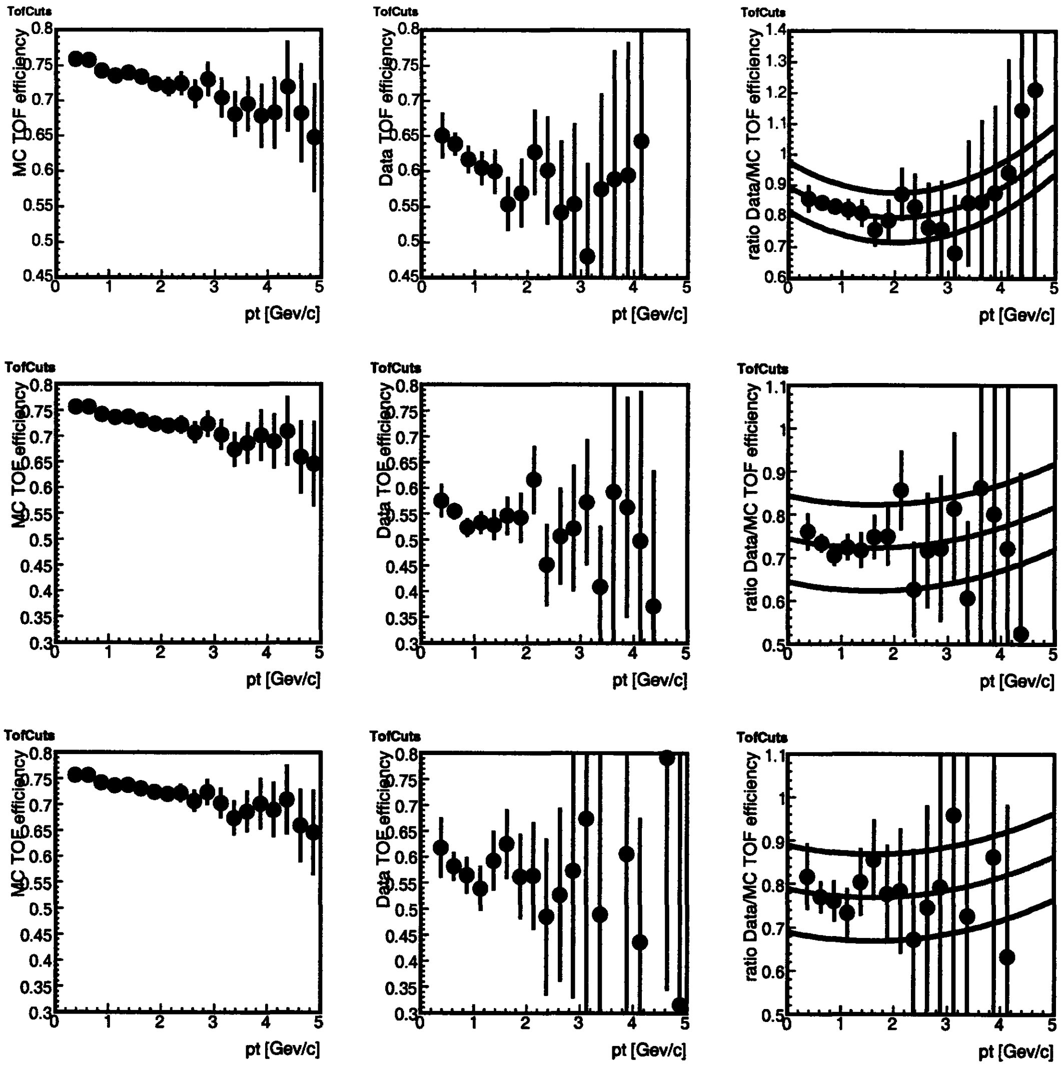

FIGURE 4.3: TOF efficiencies in simulation (left) and in data (center), and the data-simulation ratio R_ϵ (right), for Dπ modes as a function of p_T. From top to bottom, the plots for the 0d, 0h, and 0i periods of data-taking are shown. The red curve (middle) in the R_ϵ distributions is the derived correction function, the black curves represent the uncertainties used for later systematic analysis.

cuts defined in Section 6.3, where all the quantities are calculated with respect to the B_s^0 candidate of the N^{th} event. If the B_s^0 signal regions in events N and $N+1$ overlap, then event $N+2$ is utilized. This method allows to preserve the z distribution of the B_s^0 signal. By this method, a single additional tagging track candidate from pile-up events is counted in 0.22% of the events in 0d, 0.65% in 0h, and 0.72% in 0i data.

Then, a sample of potential tracks from pile-up events is retrieved from data, and is composed by all tracks which satisfy the condition $|\Delta z_0(\mathrm{B}_s^0, trk)| > 4$ cm, and the cuts in Section 6.3 (with the obvious exception of the $|\Delta z_0(\mathrm{B}_s^0, trk)| < 1.2$ cm requirement). Tracks from this sample are then mixed with the MC sample, according to the fractions determined with the method described above.

$B^{**}(B_J^*)$ production rate and characteristics

PYTHIA does not produce any excited B mesons in the simulation of the b-quark fragmentation in its default configuration, i.e., a $b\bar{q}$ (and charge-conjugate) state with angular momentum L equal to one. The absence of B^{**} states raises concerns on the ability of the MC simulation to accurately reproduce the performance of same-side tagging algorithms as found in data.

In fact, same-side tagging algorithms rely on the correlation between the flavor of the B meson and the charge of the particles generated in the fragmentation of a b quark [88]. The decay of a B^{**} meson predominantly produces a B^0 or B^+ meson and a charged pion. Due to the kinematic of the decay, the pion can be preferentially chosen as the tagging track candidate because it is often found in the proximity of the B candidate. Moreover, the charge-flavor correlation is the same as expected by same-side tagging algorithms for B^0 and B^+ mesons. Thus, any change in the production ratio of excited B mesons influences the measured performance in MC events of same-side tagging algorithms. However, when the tagging algorithm misses the decay pion and selects a fragmentation track produced with the B^{**}, the B flavor and the pion charge are anti-correlated. These two effects partially cancel, reducing the actual influence of B^{**} production ratios on the tagging properties as measured in MC samples.

In the case of B^{**} decays with a B_s^0 meson in the final state, the dominant decay is, due to the limited available phase-space, $B_s^0\gamma$. No charged particles are produced, thus the effects of the B^{**} tuning on the performance of same-side tagging algorithms are more subtle, and covered by statistical fluctuations of the available MC sample.

This study is important because it is possibile to check the effects of this tuning on B^+ and B^0 samples, where they are more significant, on data and simulated events. It thus contributes to building the confidence that simulated events can be used to calibrate a same-side tagger.

The default setup of PYTHIA has been modified by setting the production ratio of excited B mesons, equivalently, the fraction of B mesons originating from the decay of an $L = 1$ state, to 20%, which is aligned with experimental measurements (Reference [89], among others). Besides, the masses and widths of the states with u or d quarks (thus excluding the B_{sJ}^* ones) were replaced by the measurements published in Reference [62]. Table 4.5 contains the modified parameters of the PYTHIA configuration, which are also directly listed in Appendix A.

Fragmentation fractions

The PYTHIA-MC events are generated according to the Lund string fragmentation model, which requires a fragmentation function as input. The fragmentation describes the formation of hadrons out of the initial string. It thus affects the track multiplicity around the B meson, the momentum of the B meson, and the flavor and momentum of the tracks around the B. Two fragmentation functions are utilized: the Lund [90] and the Peterson [91] parameterizations. Their functional forms are reproduced in

Particle (+*c.c.*)	*LSJ*	Ratio [%]	Mass [GeV/c^2]	Width [GeV/c^2]
B_0^{*0}, B_0^{*+}	110	20	5.70	0.20
$B_1'^0$, $B_1'^+$	101	6.67	5.73	0.20
B_1^0, B_1^+	111	6.67	5.73	0.02
B_2^{*0}, B_2^{*+}	112	6.67	5.74	0.02

TABLE 4.5: Production ratio, mass and width of the B** states which were modified in PYTHIA-MC simulation. The states, all of which have angular momentum L equal to one, are identified by their spin S and total angular momentum J. Production ratios depend upon spin and total angular momentum of the B** states, but not on their charge.

the equation below, where L indicates the Lund function and P the Peterson one:

$$\begin{aligned} f_L(z|a,B) &\propto \frac{1}{z}(1-z)^a e^{-\frac{B}{z}} \\ f_P(z|\epsilon_b) &\propto \frac{1}{z}\left(1-\frac{1}{z}-\frac{\epsilon_b}{1-z}\right)^{-2} \\ z &\equiv \frac{E^B+p_L^B}{E^b+p^b}, \end{aligned} \tag{4.3.15}$$

where a, B and ϵ_b are the free parameters of the models. The variable z is defined as the ratio of the sum of the energy E and the longitudinal momentum p_L of the B meson, and the sum of energy and momentum p of the b-quark. Figure 4.4 shows an example of typical Lund and Peterson fragmentation functions.

MC events are produced using the Peterson fragmentation function with $\epsilon_b = 0.006$ for the strings with heavy quarks b and c, and the Lund function with $a = 0.3$ and $B = 0.58$ GeV^{-2} for the light strings u, d, and s.

The Peterson fragmentation function is known not to be the best choice for reproducing B data, as shown in Reference [93]. However, it has been decided to utilize the Peterson fragmentation function for the strings with heavy quarks. The reason for not directly utilizing the Lund function is that the Peterson one has a long tail toward low z values. This allows one to perform studies of systematic errors by reweighing the generated events according to different fragmentation functions, as shown in Section 6.8.2, with weights reasonably close to unity, thus reducing statistical fluctuations. The events containing a B meson, which are produced according to the Peterson function f_P with $\epsilon_b = 0.006$, are then reweighed with $w(z) = f_L(z|a = 1.68, B = 15.60)/f_P(z|\epsilon_b = 0.006)$, where the parameters chosen for the Lund function f_L follow from the prescription of Reference [93], which presents a tuning of a and B utilizing $e^+e^- \to b\bar{b}$ events registered at the Z^0 pole at LEP.

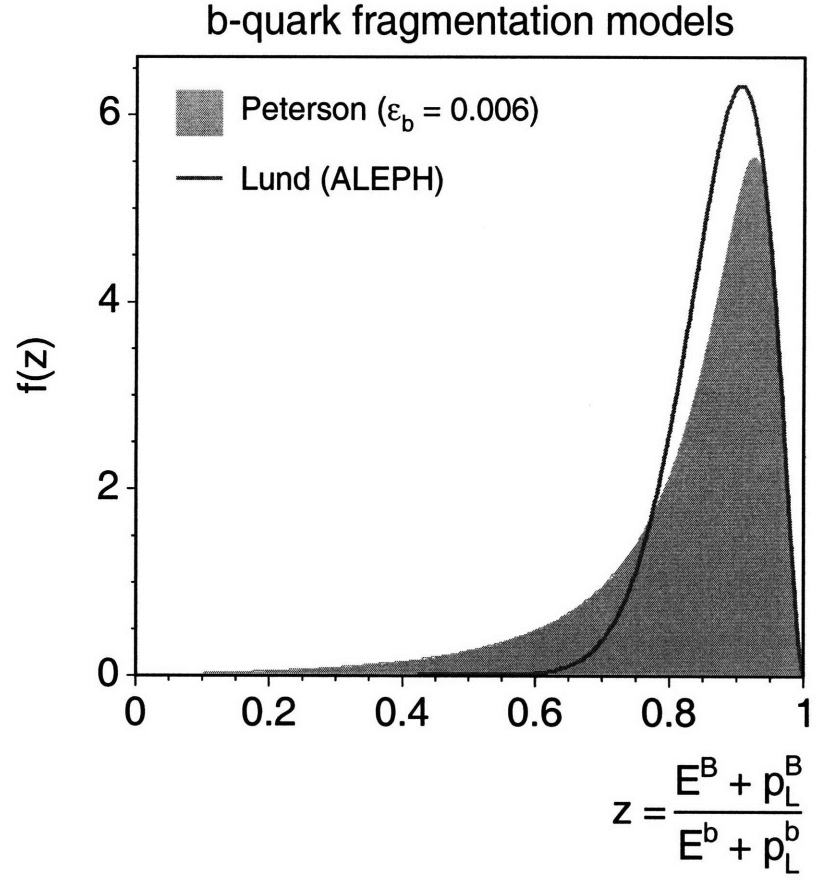

FIGURE 4.4: Fragmentation functions utilized for the production and reweighing of the MC events. The parameters of the Lund function correspond to the tuning which reproduces the measurement in ALEPH data [92].

4.4 Comparison with data

MC simulation is useful only if it accurately reproduces the characteristics of the data collected by the detector. This section presents comparisons between data and MC distributions which are relevant to various aspects of this analysis. These studies assess the accuracy of the simulation utilized and insure the applicability of the MC studies to data. Two main categories are identified:

- **B kinematics:** the most direct uses of the kinematic characteristics of B candidates are the candidate selection and its optimization, and the study of backgrounds from specific decays described in Section 4.2. The data–MC-simulation agreement is necessary to validate these studies.
- **Global event characteristics:** the data–MC-simulation comparisons in this class utilize the PYTHIA-MC set, which is a different type of samples in that it contains other primary particles beside the B meson of interest. The search for the best same-side flavor tagging algorithm and its calibration on B_s^0 MC

simulated data is a fundamental part of the analysis presented in this document, and relies on the simulation to provide a good representation of the data.

The distributions in the figures in the next sections contain some examples of the comparisons produced. The $B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$ data sample is often utilized to validate the MC simulation, because the large sample size allows for a more precise comparison.

In all the plots, the distributions indicated by the "data" label are obtained applying the following algorithm for sideband-subtraction:

- two mass ranges are selected to define the "signal" and "sideband" regions. Typical values for the $B^0_s \to D^-_s\pi^+, D^-_s \to \phi^0\pi^-$ decay mode are $[5.306, 5.425]$ GeV/c^2 and $[5.6, 5.9]$ GeV/c^2, for signal and sideband, respectively.
- the distribution of the mass of B candidates is fit with a function representing the signal (typically, a Gaussian function) and background (a linear or exponentially decaying function) components, and the ratio of the integrals of the background component in the signal and in the sideband regions defines the scale factor to use in the subtraction.
- finally, the "data" distribution, which the MC simulation is compared to, is obtained by scaling the distribution of the quantity of interest (i.e., p_T, η, ...) when the B mass is in the "sideband" region by the scale factor calculated in the previous step, and then subtracting the scaled distribution from the distribution in the "signal" region.

This algorithm relies on the assumption that the properties of the events in the sideband are similar to the background events in the signal region.

4.4.1 Data–PYTHIA-MC-simulation comparison

The figures in this section show the agreement with data achieved in the PYTHIA-MC simulation. The attention is turned to the quantities that will be utilized, as presented in Chapter 6, to select tagging track candidates and as an input to the tagging algorithm This section is meant to build the confidence that the conclusions of the next chapter, derived from MC studies, are robust.

The first set, in Figure 4.5, presents the comparison of the track quantities used to select tag candidates. These quantities include the impact parameter significance d_0/σ_{d_0}, the separation in η–ϕ space[2] ΔR and the longitudinal separation Δz_0 between the tagging track candidate and the reconstructed B^0_s meson, the pseudorapidity η, and the number of hits in the silicon detector. The "N-1 cuts" label indicates that the set of tag candidates which enter the distribution is selected by applying all cuts *except* the one on the quantity which is being tested.

An analysis of the kinematic properties of the B^0_s candidate completes the MC study. The distributions, in data and PYTHIA-MC, of transverse momentum, impact

---

[2] Distance in η–ϕ space is measured by $\Delta R \equiv \sqrt{\Delta\phi_0^2 + \Delta\eta^2}$, as defined in Equation 2.2.3.

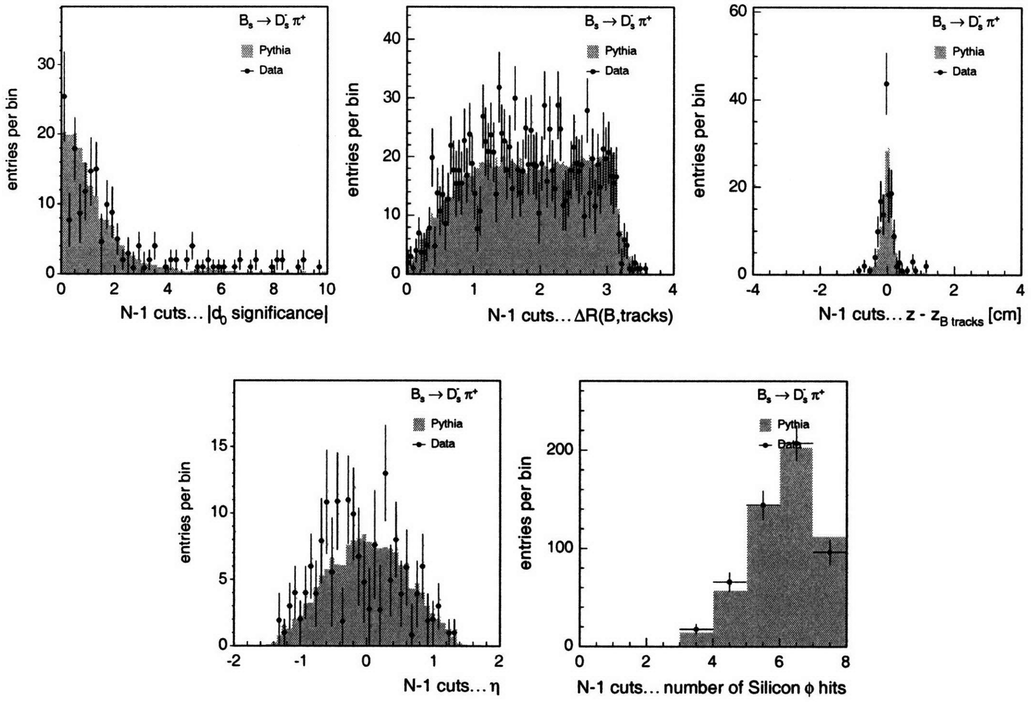

FIGURE 4.5: Data–PYTHIA-MC-simulation comparison of track variables. These track variables will be utilized to preselect tagging track candidates. From left to right, and top to bottom, are plotted the distributions for: impact parameter significance d_0/σ_{d_0}, angular separation $\Delta R(\mathrm{B}, trk)$, $\Delta z_0(\mathrm{B}, trk)$, pseudorapidity η, and number of hits in the silicon detectors (L00, SVX or ISL). Each plotted distribution is produced utilizing the sample of tracks which satisfy the requirements for being a tag candidate except for the cut on the variable shown. These distributions utilized the $\mathrm{B}_s^0 \rightarrow \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \rightarrow \phi^0 \pi^-$ data and MC samples.

parameter, transverse decay length and transverse decay length resolution are shown in Figure 4.6.

The degree of the data–MC-simulation agreement is quantified by a χ^2 test. In the calculation of the χ^2, bins with fewer than 20 entries are added to their next neighboring bins until the total number of entries is greater or equal to 20. Each of these bin clusters provides one degree of freedom. The results of the tests, which are presented in the title of each plot, show that, within the available statistics of MC events, the MC simulation adequately reproduces the distribution of variables as seen in data.

4.4.2 Data–BGENERATOR-MC-simulation comparison

The single-hadron BGENERATOR-MC simulation is utilized in many steps of the analysis, mainly concerning selection efficiencies and distributions of fundamental properties such as mass and proper decay-time for signal and background components.

The distributions presented in this section are obtained from the sample of B_s^0

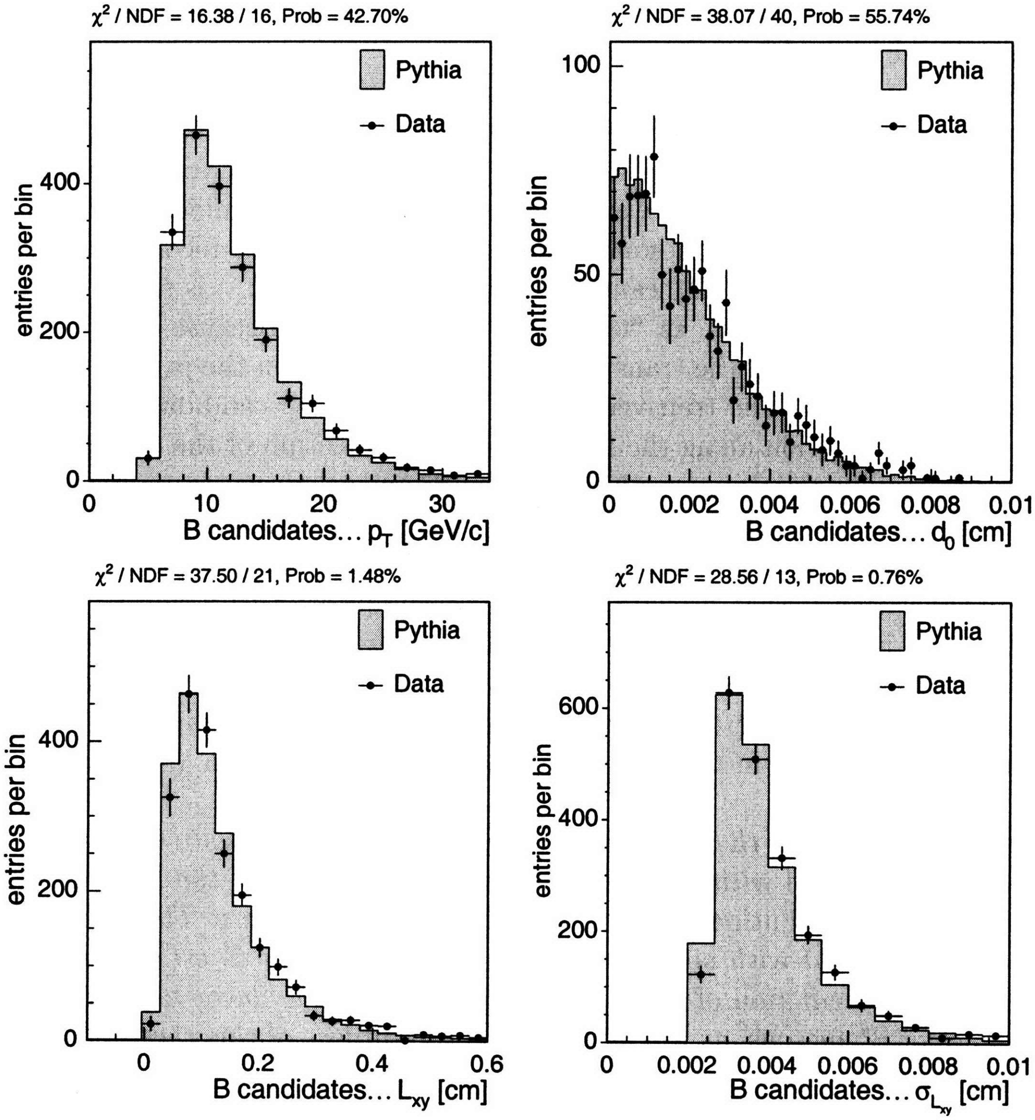

FIGURE 4.6: Comparison of distributions of the $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$ candidates between data (black markers) and PYTHIA-MC simulation (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse momentum, impact parameter, flight distance in the transverse plane L_{xy} and error in L_{xy}.

mesons reconstructed in the $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ decay mode, in data and in the MC simulation, and comprise the variables which are most crucial to the analysis. Figure 4.7 contain the distributions of variables such as the χ^2 of the fit of the B decay vertex, and the impact parameter d_0, the pseudorapidity η, the transverse decay length L_{xy}, the transverse decay length resolution $\sigma_{L_{xy}}$ and significance $L_{xy}/\sigma_{L_{xy}}$ of the reconstructed B candidate, and the impact parameter significance d_0/σ_{d_0} of the pion produced in the $B_s^0 \to D_s^-\pi^+$ decay.

The second set of comparisons, in Figure 4.8, presents the data–MC-simulation simulation agreement of quantities relative to the D_s^- candidate: transverse decay length, transverse decay length resolution and significance, transverse momentum p_T, and mass m of the D_s^- candidate, and the impact parameter significance d_0/σ_{d_0} of the two kaons which form the ϕ^0 meson in which the D_s^- candidate is reconstructed to have decayed. In the case of the D_s^- candidate, the quantity $L_{xy}(D_s^-)$ is rather improperly indicated, in the above list, as "transverse decay length" because it represents the distance, projected along the transverse momentum, between the primary vertex and the D_s^- decay vertex. The transverse decay length of a D_s^- candidate more properly indicates the projection along the D_s^- transverse momentum of the distance between the D_s^- production and decay vertices, i.e., the B_s^0 and D_s^- decay vertices, respectively. In each plot, the distribution obtained by utilizing background (i.e., selecting a B mass region far from the expected signal region) is also shown, which enables one to quickly identify the variables that provide signal–background separating power.

The plots in this section show a good level of agreement between data and MC simulation, which allows us to state that the selection optimized with MC data is truly close to the best possible one. As in the previous section, the data–MC-simulation agreement is expressed in terms of the probability returned by a χ^2 test.

The chapter presented the work devoted to produce an adequate simulation of the data sample collected with the CDF detector. The quality of the MC simulation is testified by the distribution comparisons shown in this last part. The confidence that the results obtained with simulated events are trustworthy is established here. In particular, the prediction of the performance of same-side flavor-tagging algorithms in samples of B_s^0 *mesons, as the one which is the subject of this thesis entirely relies on the MC simulation.*

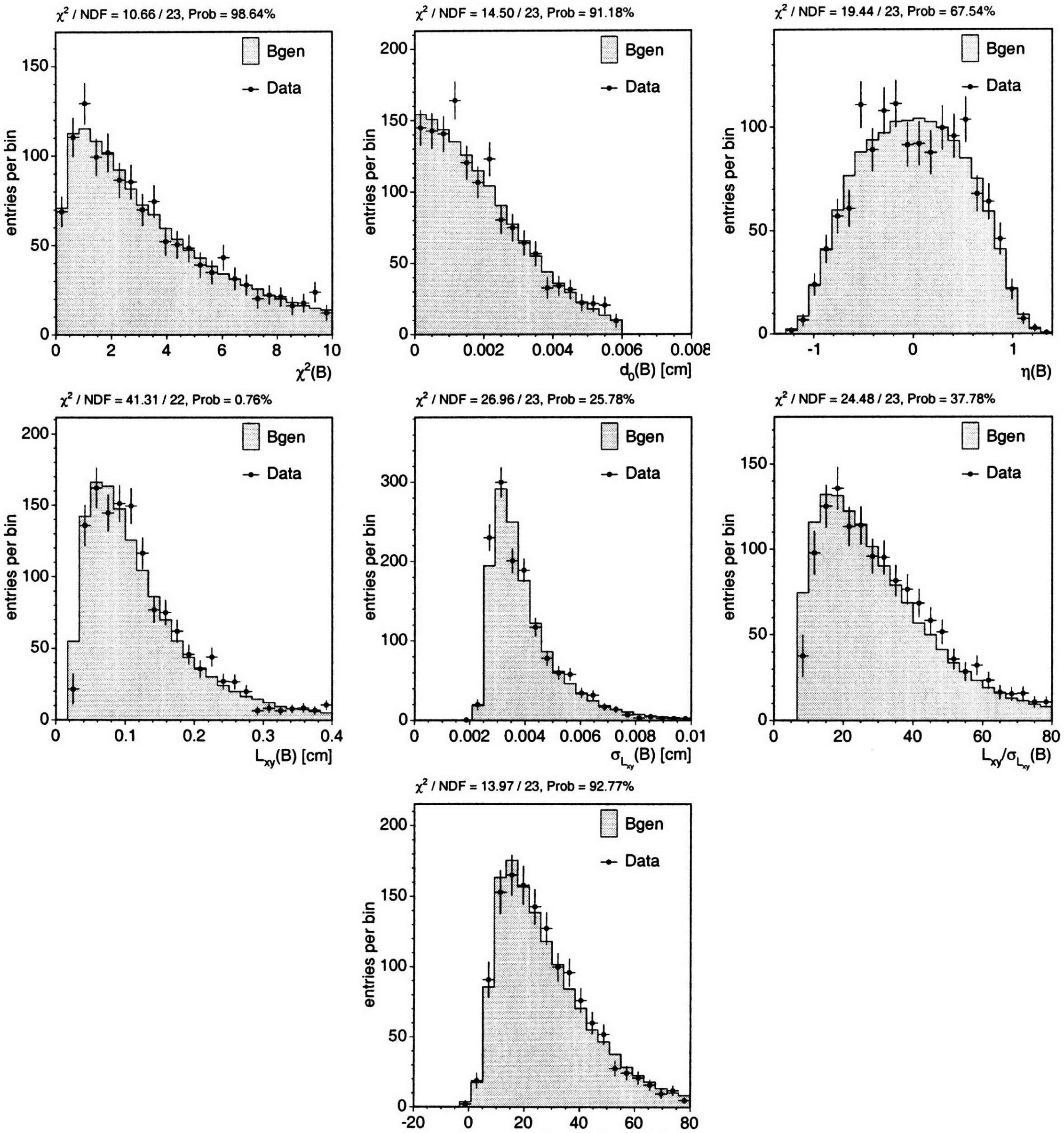

FIGURE 4.7: Comparison of distributions of the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates between detector (black markers) and BGENERATOR-MC data (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: the χ^2 of the fit of the B_s^0 decay vertex, the impact parameter, the pseudorapidity, the transverse decay length, the transverse decay length resolution and significance of the B_s^0 candidate, and the impact parameter significance of the pion which comes from the B_s^0 decay vertex.

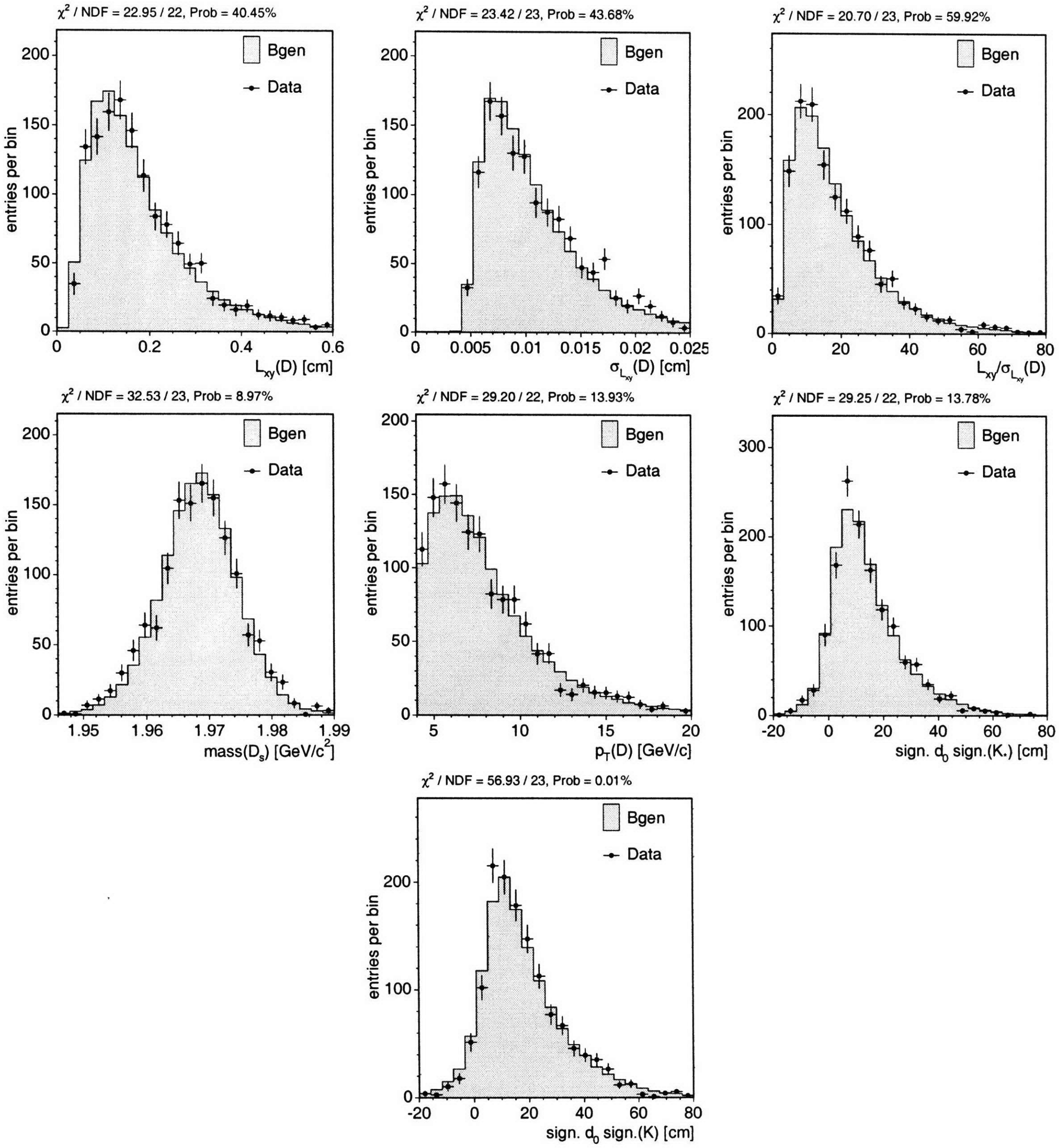

FIGURE 4.8: Comparison of distributions of the $D_s^- \to \phi^0\pi^-$ candidates from $B_s^0 \to D_s^-\pi^+$ decays between detector (black markers) and BGENERATOR-MC data (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse decay length, transverse decay length resolution and significance, mass and transverse momentum of the D_s^- candidate, and impact parameter significance of the kaons which form the ϕ^0 meson in the D_s^- final state.

Chapter 5

Elements of the B_s^0 mixing analysis

This chapter presents the description of the method and the ingredients for the analysis which resulted in the first observation of $B_s^0 - \overline{B}_s^0$ *oscillations.*

5.1 Time-dependent study of oscillations

The equations which describe the time evolution of B_s^0 mesons and relate their probability of decaying with the same ("unmixed" case) or the opposite ("mixed" case) flavor with respect to their production flavor are presented in Section 1.2. The probability density functions for a B_s^0 meson produced at time $t = 0$ to decay at time t as a B_s^0 or a $\overline{B}_s^0$ are drawn in Figure 5.1. An interesting quantity is represented by the asymmetry $\mathcal{A}$:

$$\mathcal{A}(t) = \frac{\mathcal{P}_{unmixed}(t) - \mathcal{P}_{mixed}(t)}{\mathcal{P}_{unmixed}(t) + \mathcal{P}_{mixed}(t)}, \tag{5.1.1}$$

where $\mathcal{P}_{unmixed}$ and $\mathcal{P}_{mixed}$ are the probabilities that a B_s^0 meson decays a time t after production with the same, or opposite, flavor with which it was produced. These two probability density functions are defined in Equations 1.2.9 and 1.2.8. Utilizing the expressions provided for $\mathcal{P}_{unmixed}$ and $\mathcal{P}_{mixed}$, the asymmetry $\mathcal{A}$ results in an expression which is directly proportional to $\cos \Delta m_s t$.

A direct approach to the measurement of a mixing frequency, which is the aim of an analysis of oscillations, consists in the fit of the asymmetry, obtained by reconstructing particle candidates and counting how many of them decay with the same, or the opposite, flavor as at production, and as a function of time. In the case of an analysis of B_s^0 mixing, however, a different method is required because, *a priori*, it is not known whether the available data are sufficiently sensitive to discriminate an oscillatory signal. The main obstacles are the ability to discern the candidate flavor at production, and the measurement of the proper decay-time, the resolution of which needs to be sufficiently precise to resolve the time-dependence of $B_s^0 - \overline{B}_s^0$ oscillations.

Because the analysis presented in this document aims at measuring a frequency, it is natural to consider performing a search in the frequency domain. The method is described in Reference [94] and is usually referred to as the amplitude scan method.

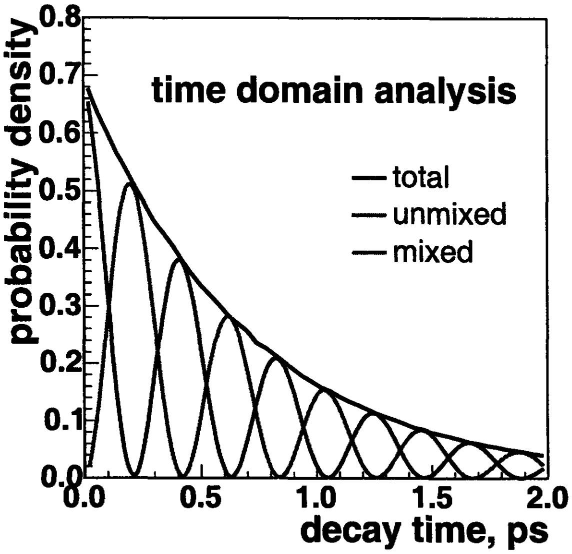

FIGURE 5.1: Probability density functions for a B_s^0 meson produced at time $t = 0$ to decay at time t as a B_s^0 ("unmixed", the probability density is $1/\tau_{B_s^0}$ at $t = 0$), or a $\overline{B}_s^0$ ("mixed", the probability density is zero at $t = 0$). The functions assume Δm_s equal to 15 ps^{-1}.

Equations 1.2.8 and 1.2.9 are modified by introducing the amplitude $\mathcal{A}$:

$$\mathcal{P}_{unmixed/mixed}(t) \propto [1 \pm \mathcal{A}\mathcal{D}\cos(\Delta m_s t)] \, , \tag{5.1.2}$$

where $\mathcal{D}$ is the dilution of the tagger utilized to determine the flavor of particles. The amplitude scan consists of a spectrum of the amplitude as a function of the oscillation frequency, obtained by performing a series of fits for $\mathcal{A}$ from the yields of B_s^0 candidates which are tagged as mixed/unmixed, as a function of proper time, while fixing Δm_s to a probe value. The signature of a mixing signal is represented by a region in the scan where the amplitude is consistent with unity and inconsistent with zero. An example of an amplitude scan, produced with a toy Monte Carlo sample generated with Δm_s = 15 ps^{-1}, is shown in Figure 5.2. Equation 5.1.2 also shows that the fit cannot distinguish between the amplitude $\mathcal{A}$ and the tagger dilution $\mathcal{D}$, which is the reason why the calibration of taggers is necessary.

The error on the amplitude, $\sigma_{\mathcal{A}}$, acquires a special interest, because the quality of an analysis is measured in terms of the ability to separate unity from zero in the scan. The following formula provides an estimator of $\sigma_{\mathcal{A}}$ [95, 94]:

$$\frac{1}{\sigma_{\mathcal{A}}} \simeq \frac{\mathcal{S}}{\sqrt{\mathcal{S}+\mathcal{B}}} e^{\frac{-\Delta m_s^2 \sigma_{ct}^2}{2}} \sqrt{\frac{\epsilon \mathcal{D}^2}{2}} \, , \tag{5.1.3}$$

which has been grouped into three terms.

The first term expresses the contribution of statistics, $\mathcal{S}$ and $\mathcal{B}$ being the number of B_s^0 candidates reconstructed as signal and background, respectively. The larger the

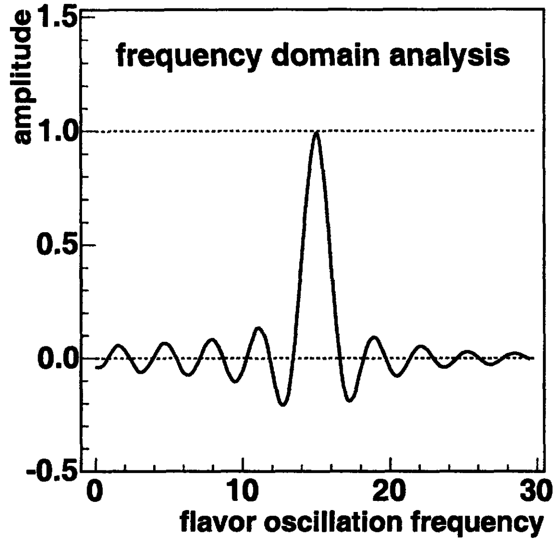

FIGURE 5.2: Amplitude scan in a toy Monte Carlo sample. The curve shown describes the expected $\mathcal{A}$ as a function of Δm_s, when the true value of Δm_s is equal to 15 ps^{-1}.

signal yield and the signal purity, the smaller the error on the amplitude.

The second factor represents the dependence of the amplitude error on the resolution with which proper decay-length and time are measured, σ_{ct}. The sensitivity for measuring Δm_s degrades exponentially with σ_{ct}, as shown in Equation 5.1.3. The B_s^0 mixing analysis is much more challenging than the study of B^0 oscillations because the SM expectation is $\Delta m_s/\Delta m_d \sim 40$ and most of the SM extensions foresee even higher values.

The last term contains the figure of merit of the flavor tagging algorithm utilized. The efficiency ϵ and dilution $\mathcal{D}$ are the traditional parameters with which the quality of a tagger is described. The efficiency ϵ of a tagging algorithm corresponds to the fraction of events to which the algorithm assignes a non-null tag decision. The dilution $\mathcal{D}$ is defined as $1 - 2\mathcal{P}_w$, where $\mathcal{P}_w$ is the probability that the assigned tag is incorrect. A perfect tagger ($\mathcal{P}_w = 0$) will have a dilution equal to unity, while one which randomly assigns flavor tags ($\mathcal{P}_w = 0.5$) will have a dilution equal to zero.

5.2 Proper decay-time and calibration of proper-decay-time resolution

The decay time in the B_s^0 rest frame is obtained as follows:

$$ct = L_{xy}(\mathrm{B}_s)\,\frac{m_{\mathrm{B}_s}}{p_T(\mathrm{B}_s)}\,, \quad L_{xy} \equiv \frac{\mathbf{r}\cdot\mathbf{p}_\mathrm{T}}{p_T}\,, \tag{5.2.1}$$

where the 2D vector $\mathbf{r}$ is the displacement of the B_s^0 decay point with respect to the production vertex, in the transverse plane. The quantity L_{xy} is referred to as the

B_s^0 transverse decay length (Equation 4.3.1). In the case of partially reconstructed hadronic and semileptonic B_s^0 decays, where the B_s^0 candidate is not fully reconstructed, a correction factor k has to be included to account for the missing momentum. The expression becomes:

$$ct = ct^*\, k, \quad ct^* = L_{xy}^{\ell \mathrm{D}_s}\, \frac{M_{\mathrm{B}_s}}{p_T^{\ell \mathrm{D}_s}}, \quad k \equiv \frac{L_{xy}(\mathrm{B}_s)}{L_{xy}^{\ell \mathrm{D}_s}}\, \frac{p_T^{\ell \mathrm{D}_s}}{p_T(\mathrm{B}_s)}, \tag{5.2.2}$$

where $L_{xy}^{\ell \mathrm{D}_s}$ and $p_T^{\ell \mathrm{D}_s}$ are the projected displacement and the transverse momentum of the reconstructed decay products, and M_{B_s} is the world average of B_s mass measurements. The quantity ct^* is traditionally called *pseudo*-proper decay-length and is constructed with only the information from the reconstructed lepton and D_s^- candidates. 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 determination of the proper-decay-time resolution is a critical part of the analysis, because it dramatically affects the sensitivity for observing an oscillation signal. From the definition of the proper decay-time in Equation 5.2.1, the following expression is obtained:

$$\sigma_{ct} = \sigma_{L_{xy}} \cdot \frac{M}{p_T} \oplus ct \cdot \sigma_{p_T}, \tag{5.2.3}$$

where the component due to the uncertainty on M is omitted, because it is negligible. The resolution of a proper-decay-time measurement thus contains two components, the first of which is independent of the proper time, while the second component is directly proportional to the decay time of reconstructed B_s^0 candidates.

The first term contains the contribution of the measurement of the transverse flight distance of the B_s^0 meson. It depends on the accuracy with which the PV, where the $p\bar{p}$ interaction occurred and the B_s^0 was produced, and the SV, the decay point of the B_s^0 meson, are measured. The position of the PV is determined for each event by fitting part of the tracks in the underlying event to a common origin, as described in Reference [80], and presented in Section 4.3.2.

The measurement of the error on the SV is more complicated, because it would require an ensemble of B_s^0 mesons decaying at known positions. Therefore, a calibration sample of *pseudo*-B_s^0 candidates was constructed by associating a track, which is prompt in most of the cases, to a D candidate that is selected by applying cuts which enhance the fraction of prompt D candidates. This method allows one to obtain topologies similar to B_s^0 decays, a large amount of which originates from the PV, and thus have $L_{xy} \sim 0$ by construction. The calibration sample is obtained by utilizing D candidates reconstructed in the two-displaced-track triggers, the requirements of which are presented in Table 3.1. Each D candidate is required to have been reconstructed using the two trigger tracks, its impact parameter must be less than 100 μm, and its reconstructed mass needs to be within 8 MeV/c^2 of the PDG value [7]. The additional cuts which are applied to the decay vertices of pseudo-B_s^0 candidates, in the D–single-track and D–three-tracks topologies, are reported in Table 5.1. The listed requirements allow for the selection of samples of pseudo-B_s^0 candidates which

Cut		Dπ	D$\pi\pi\pi$
Mass	[GeV/c^2]	[5.4,6.0]	[5.4,5.8]
p_T	[GeV/c]	$> 5,5$	> 6.0
$\chi^2_{r\phi}$		< 15	< 15
$p_T(\pi)$	[GeV/c]	> 1.2	—
Mass($\pi\pi\pi$)	[GeV/c^2]	—	< 1.75

TABLE 5.1: Selection cuts for Dπ and D$\pi\pi\pi$ vertex candidates. These cuts define the selection of the samples of pseudo-B^0_s candidates which are utilized to calibrate the proper-decay-time resolution. Pseudo-B^0_s candidates are produced by associating one or three tracks to a D candidate with $|d_0| < 100$ μm and reconstructed mass within 8 MeV/c^2 of its PDG value.

are enriched in candidates produced at the $p\bar{p}$ interaction point, and reproduce the topology of $B^0_s \to D^-_s\pi^+$ and $B^0_s \to D^-_s\pi^+\pi^-\pi^+$ decays.

The distribution of the proper decay-time measured in the calibration sample is shown in Figure 5.3. The small components which are fit with exponential functions (indicated by f_+, in the figure, plus a similar contribution symmetrical with respect to the origin) contain the non-prompt part of the sample of pseudo-B^0_s candidates. The width of the Gaussian function which fits the prompt component of the calibration sample is taken as the true resolution of the reconstructed decay time. From the comparison of our estimate of the error with the error resulting from the combination of the PV and SV fits a scale factor is obtained. This scale factor is applied to the $\sigma_{L_{xy}}$ returned by the fits of the B^0_s candidates reconstructed in data. The scale factor is parameterized as a function of several kinematic variables, to account for differences between the kinematic properties of the calibration sample and the B^0_s signal sample, and applied on a candidate-by-candidate basis. The parameterization is obtained by binning the sample of pseudo-B^0_s candidates with respect to a kinematic variable, and then measuring the scale factor in each of these subsamples. The scale-factor dependences are corrected one variable at a time, which assumes that the corrections factorize completely. After the final tuning step is applied, the scale factor shows a flat behavior centered around 1.0. In a realistic scenario, the factorization of the correction factor is not complete, and residual deviations are present. An additional global scale factor is applied to σ_{ct}, and accounts for the residual disagreement between measured and expected proper-decay-length resolution after the candidate-by-candidate calibration. A different global scale factor is utilized for each B^0_s decay mode. These scale factors are expected, and measured, to be close to unity. The determination of scale factors is detailed in Reference [96], where the procedure which has been briefly presented here is described in full detail.

The second factor in Equation 5.2.3 depends on the error in the p_T measurement, and increases linearly with the proper decay-length ct of a candidate. While in the case of fully reconstructed decays σ_{p_T} is negligible, for partially reconstructed ones it represents an important addition to the global uncertainty and is closely tied to the distribution of the k-factor defined in Equation 5.2.2. The k-factor distribution, $F(k)$, is obtained from MC simulation. The r.m.s. width of $F(k)$, defined as $\sqrt{\langle k^2\rangle - \langle k\rangle^2}$,

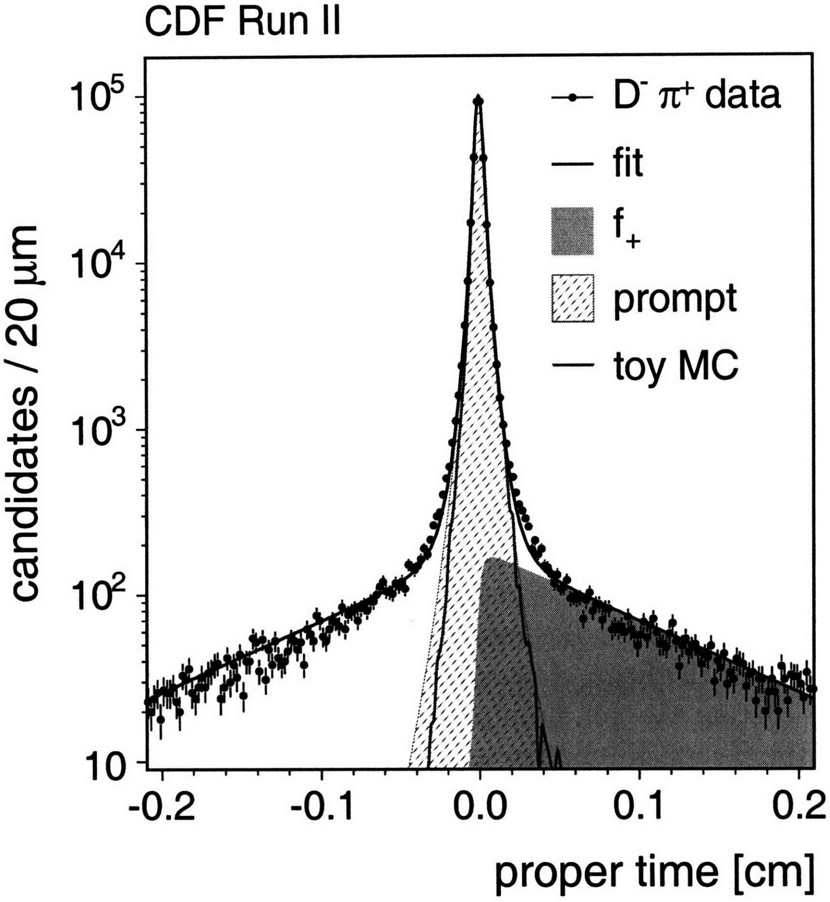

FIGURE 5.3: Proper-decay-time distribution of the calibration sample of D^-+track topology. The fitted width of the prompt component is assumed to represent the true resolution of the reconstructed decay time. A toy MC has been generated to test a simplified model for the vertex resolution.

strongly influences the proper-decay-time resolution. In fact, the effective proper-decay-time resolution for a partially reconstructed B_s^0 candidate is taken as the r.m.s. width of the following function:

$$\int e^{-(kt^*)^2/2\sigma_{t^*}} F(k)dk \,, \tag{5.2.4}$$

where t^* and σ_{t^*} are the pseudo-proper decay-time and decay time resolution of the candidate. The distributions of k-factors for partially reconstructed fully-hadronic B_s^0 decays, $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$, peak closely to unity and are very narrow (Figure 5.4), as a consequence of the softness of the lost particle in the decay chain. In the case of semileptonic decays, the distributions vary as a function of the reconstructed ℓD_s^- mass.

The distributions of k-factors in both hadronic and semileptonic cases are shown in the left plot of Figure 5.4. The right plot shows the proper-decay-length resolution,

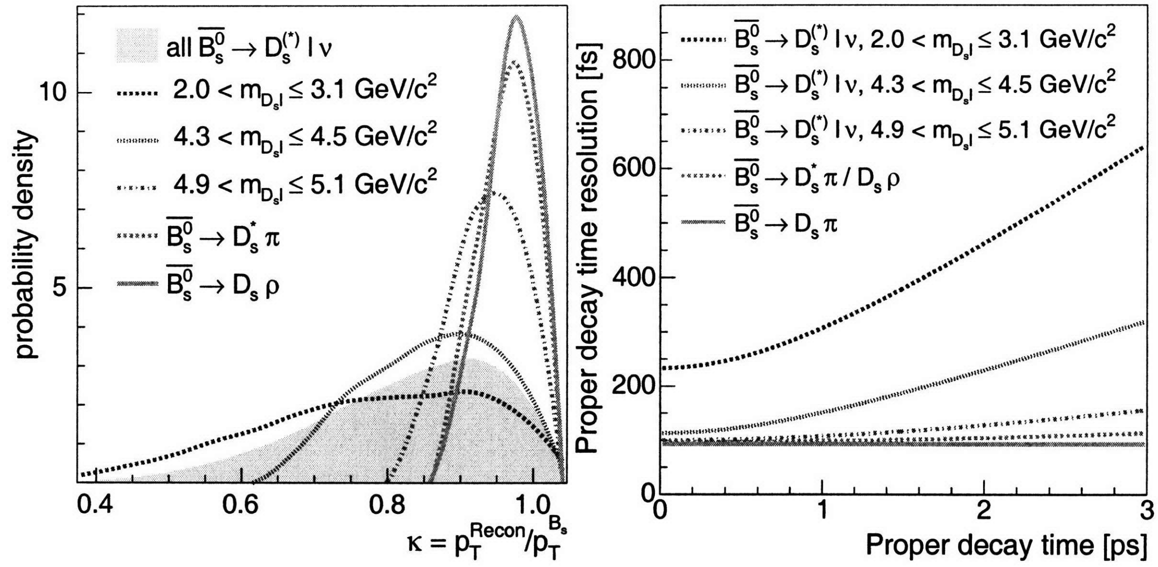

FIGURE 5.4: k-factor distribution for several ℓD_s^- mass regions (B_s^0 semileptonic decays, $D_s^- \to \phi^0\pi^-$) and for partially reconstructed hadronic decays [97] (left). The plot on the right shows the mean proper-decay-length resolution σ_{ct} as a function of the proper decay-length ct, derived from Equation 5.2.3.

as a function of the proper decay-length, derived from Equation 5.2.3. These plots show that the partially reconstructed hadronic samples provide a resolution which is comparable to the fully reconstructed ones. This is a consequence of the fact that the particles which are not reconstructed carry a small fraction of the momentum of a B_s^0 candidate. The proper-decay-length resolution of the partially reconstructed semileptonic samples is presented in bins of ℓD_s^- mass. The few candidates which populate the high mass bin are almost as good as fully reconstructed hadronic modes.

In Figure 5.5 the proper-decay-time distribution of $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$ candidates, both fully and partially reconstructed, is shown. A complete analysis of B_s^0 lifetime, which would include the evaluation of systematic uncertainties, has not been performed. However, a measurement of B_s^0 lifetime has been obtained as a by-product of an analysis of B_s^0 oscillations by performing a fit of proper decay-length as the one presented in Figure 5.5. The result is in agreement with the world average of B_s^0 lifetime measurements, which constitutes a cross-check that the absolute ct scale is correct. This issue is addressed in more detail in the discussion of the systematic uncertainties on Δm_s, in Section 7.2.2.

5.3 Flavor taggers: SST and OST

While the flavor of the B_s^0 candidate at the decay point is unambiguously defined by the charges of its daughter tracks, the flavor at production is inferred, with a certain degree of uncertainty using flavor tagging algorithms.

Two types of flavor tagging algorithms are utilized at CDF: opposite-side (OS)

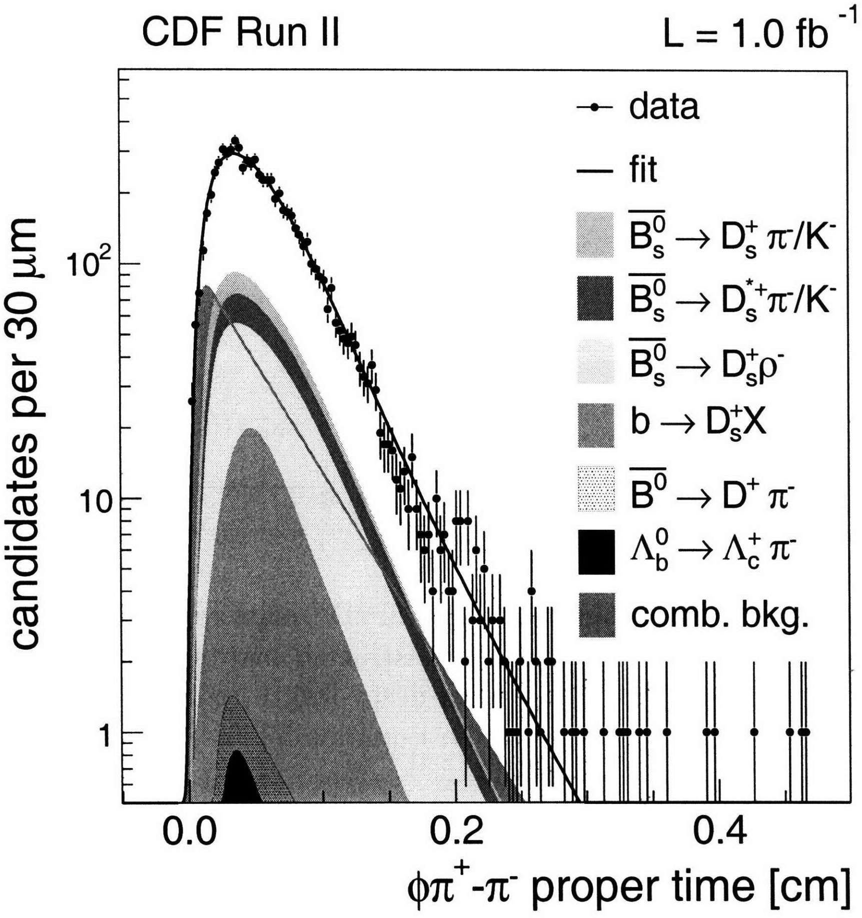

FIGURE 5.5: Proper-decay-time distribution of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates. The fitted value of B_s^0 lifetime is in agreement with the world average of B_s^0 lifetime measurements, and thus provides a test for the correctness of the global *ct* scale.

and same-side (SS) flavor taggers. The performance of a tagger is quantified by its efficiency ϵ, the fraction of candidates to which a tag is assigned, and dilution $\mathcal{D}$, defined as $1 - 2\mathcal{P}_w$, where $\mathcal{P}_w$ is the probability that the assigned tag is incorrect. The sensitivity for observing an oscillation signal is proportional to $\sqrt{\epsilon \mathcal{D}^2}$, which thus represents the figure of merit of a tagging algorithm. The decisions of the OS and SS taggers are combined by treating the two taggers as independent [98].

5.3.1 Opposite-side Flavor Tagging

Opposite-side taggers exploit the fact that at hadron colliders b quarks are mostly produced in $b\bar{b}$ pairs. Therefore, the flavor of the b quark in the opposite-side with respect to the reconstructed candidate is correlated to the flavor at production of the B_s^0 meson of interest. Limitations in opposite-side tagging algorithms arise because the second bottom hadron is inside the CDF detector acceptance in less than 40% of the events, or, it is also possible that the second B hadron is a neutral B meson that mixed into its anti-particle.

Soft-lepton taggers (SLT) are based on $b \to \ell^- X$ semileptonic decays. The charge of the lepton, either a muon or an electron, is correlated to the charge of the decaying B meson: an ℓ^- is produced in the transition $b \to c\ell^-\bar{\nu}X$, while an ℓ^+ signals a $\bar{b}$ flavor. The semileptonic B branching ratio is small, $\mathcal{BR}(\mathrm{B} \to \ell X) \simeq 20\%$, in terms of the tagger efficiency, but the lepton identification has a high purity. The tagger is thus expected to have low efficiency, but large dilution. A complete description of the soft-muon and soft-electron taggers utilized in this analysis are presented in References [55] and [56], respectively. A brief summary of the method adopted to identify lepton candidates is described in Section 3.4.1. Table 5.2 indicates the performance of the soft-muon and the soft-electron taggers, evaluated on `0d` data samples.

The opposite-side-kaon tagger (OSKT) is based on cascade decays $b \to c \to s$. The charge of a kaon from the charm decay $c \to sX$ is correlated with the B flavor: a K^- results from the decay chain $b \to c \to s$, a K^+ originates from a $\bar{b}$ quark. The challenge consists in identifying kaons among a vast background of pions and then finding the kaon candidate from the B hadron decay among all other kaons. The identification of kaons utilizes *pID* information from the TOF detector and specific ionization $d\mathrm{E}/d\mathrm{x}$ measured in the COT. The probabilities from the t_{flight} and $d\mathrm{E}/d\mathrm{x}$ measurements for a given track $\mathcal{P}(i)$ for the particle hypotheses $i = \mathrm{K}, \pi, p$ are combined in the likelihood ratio $\mathcal{LR}$:

$$\mathcal{LR}(\mathrm{K}) = \log\left(\frac{\mathcal{P}(\mathrm{K})}{f_{\mathrm{K}}\,\mathcal{P}(\mathrm{K}) + f_\pi\,\mathcal{P}(\pi) + f_p\,\mathcal{P}(p)}\right), \tag{5.3.1}$$

where $f_{\mathrm{K}} = 0.2$, $f_\pi = 0.7$ and $f_p = 0.1$ are the *a priori* fractions of kaons, pions and protons in the sample, as measured in Reference [99]. Kaons are identified by applying a lower cut on $\mathcal{LR}(\mathrm{K})$. The impact parameter significance d_0/σ_{d_0} is utilized to distinguish the kaons originating from B hadron decays. Kaons are separated in three classes to improve the effectiveness of the tagger. The first class contains tags where the identified kaon candidate is part of a jet produced at a SV. 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 includes all other tags. On average, tags of the first class are characterized by higher dilutions. As opposed to the soft lepton taggers, this method is characterized by a medium efficiency and small dilution. The opposite-side-kaon tagger is detailed in Reference [100]. The performance of this tagger on a sample of `0d` data is reported in Table 5.2, at the end of this section.

The jet-charge tagger (JQT) utilizes the correlation of the charge of a b-jet to the charge of the originating b quark. The charge of the b-jet is defined to be the momentum-weighed sum of the charges of the tracks associated with the jet. Track-based jets are fit using a cone-clustering algorithm utilized in Run I analyses and described, for example, in Reference [101]. Then, two ANN's, trained on $b\bar{b}$ `PYTHIA` MC, are utilized in the identification of tagging jets associated with B hadrons in the opposite-side. The first ANN, "trackNet", assigns each track in a jet the probability $\mathcal{P}_{trk}$ that it originates from a B hadron decay. The second ANN, "bJetNet" utilizes the track probabilities and additional jet related kinematic input to evaluate the probability that a jet is the tagging one. A comprehensive description of the

Flavor tagger	$\epsilon\mathcal{D}^2[\%]$
Soft-muon	$0.559 \pm 0.094 \pm 0.027$
Soft-electron	$0.264 \pm 0.054 \pm 0.022$
Jet-charge Sec. Vertex	$0.230 \pm 0.068 \pm 0.017$
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 5.2: Performance of opposite-side flavor taggers. The measured values of $\epsilon\mathcal{D}^2$ are followed by their statistical and systematic uncertainties. The taggers are applied to samples of B candidates corresponding to the `0d` period of data-taking, for a total integrated luminosity of 355 pb^{-1}. The reconstruction of the fully-hadronic B^0 and B^+ modes to which the soft-lepton and jet-charge taggers are applied is described in Reference [103]. The three classes of jet-charge taggers, described in the text, are separately presented. The opposite-side-kaon tagger is applied to a sample of semileptonic B decays, as described in Reference [100].

two ANN's is presented in Reference [101]. The jet with the highest probability, as calculated by bJetNet, is selected as the tagging jet. The jet charge Q_{jet}, from which one infers the opposite-side flavor, is defined as follows:

$$Q_{jet} = \frac{\sum_i Q_i \cdot p_T^i \cdot (1 + \mathcal{P}_{trk}^i)}{\sum_i p_T^i \cdot (1 + \mathcal{P}_{trk}^i)}, \tag{5.3.2}$$

where Q_i and p_T^i are, respectively, the charge and transverse momentum of a track in the jet, and $\mathcal{P}_{trk}^i$ is the probability that the track belongs to a b-jet. Jets are divided in three mutually exclusive classes to better utilize the statistical power of the tagging algorithm. The first class contains jets which are consistent with coming from a secondary vertex that has a decay length significance, $L_{xy}/\sigma_{L_{xy}}$, greater than 3. The second class includes all jets not in the first class, with at least one track in the jet such that $\mathcal{P}_{trk} > 50\%$. The third class contains all the remaining jets. The tagger purity decreases from the first to the third class. The jet tagger is expected to have high efficiency but lower dilution than the other algorithms. The performance of the jet-charge tagger applied to `0d` data samples is summarized in Table 5.2. The three classes of jets are separately presented.

Finally, an ANN combines the pieces of information provided by the three taggers [102]. The performance of opposite-side taggers is independent of the type of B meson produced on the signal side. It is thus possible to exploit high statistic B^+ and B^0 samples to calibrate opposite-side taggers.

The performances of the individual OST algorithms which contribute to the combined OST utilized in this analysis, evaluated on `0d` data samples, are summarized in Table 5.2. The figure-of-merit of the combined opposite-side tagger, in the hadronic and semileptonic data samples, is reported in Table 5.3, at the end of this section.

$\epsilon\mathcal{D}^2$	Semileptonic	Hadronic
OST	1.8±0.1 %	1.8±0.1 %
SST	4.8±1.2 %	3.4±0.9 %

TABLE 5.3: Performance of flavor taggers used at CDF in the hadronic and semileptonic samples. Statistical and systematic uncertainties are added in quadrature. The SST performance depends on the p_T of the B_s^0 candidates to which the tagger is applied. Semileptonic B_s^0 candidates have a harder p_T spectrum, which explains the difference in the performance of SST in the two samples.

5.3.2 Same-side Flavor Tagging

The same-side (kaon) tagger (SST) is based on the correlation between the b flavor and the charge of the particles produced in association with the B_s^0 candidate during the fragmentation process of the b quark. When a $B_s^0(\overline{B}_s^0)$ meson is formed, a $\bar{s}(s)$ quark is left at the end of the fragmentation chain and may form a $K^+(K^-)$. Thus, if a charged particle is found close to the B_s^0 meson and identified as a kaon, it is likely to be the leading fragmentation track, the charge of which is correlated to the charge of the b quark contained in the reconstructed candidate, at the time of its production.

The development of the algorithm for same-side flavor tagging utilized in this analysis of B_s^0 oscillations is presented in Chapter 6.

Table 5.3 reports a summary of the performance of the SST in the hadronic and semileptonic data samples, separately. The performance of opposite-side taggers is identical in both samples because these taggers utilize information which is independent of the behavior of the trigger-side B_s^0 candidate. On the other hand, the performance of the same-side tagger used in this analysis is dependent on the p_T of B_s^0 candidates. The semileptonic B_s^0 candidates reconstructed for this analysis have a harder p_T spectrum than the hadronic B_s^0 candidates. It is calculated using MC events that the harder the p_T spectrum of B candidates in a sample, the higher the performance of SST. This explains the difference in the performance of SST in the two samples.

The elements of the mixing analysis presented in this document are described in this chapter. The discussion on the flavor tagging algorithms utilized in this analysis is completed in the next chapter, which is entirely dedicated to same-side flavor tagging of B_s^0 candidates.

Chapter 6

Same-Side Flavor Tagging

The tagging of the flavor at production of B_s^0 *mesons is an indispensable ingredient for an analysis of* B_s^0 *oscillations. This chapter describes the development of the same-side-kaon tagger utilized in the analysis presented in this document. This algorithm provides most of the total flavor tagging power available to the* B_s^0 *mixing analysis presented in this document, being 2-3 times more powerful than the other available tagging algorithms.*

6.1 General description of tagging

Flavor tagging refers to the process of deciding whether a B meson at its birth contains a b or a $\bar{b}$ quark. The time-dependent analysis of mixing requires knowledge of the flavor of the meson at the time of its production and decay. The latter is readily known by reconstructing decay modes which unequivocally indicate the decay flavor of the B candidate, such as $B_s^0 \to D_s^- \pi^+$, where the charges of the pion and of the D_s^- candidate reveal the flavor of the decayed meson. Assessing the flavor at production is not as straightforward, and different algorithms have been developed. Some of these algorithms, among which the ones used in this and in previous CDF analyses, are described in the following sections.

As already explained in Section 5.1, the performance of a flavor tagger is traditionally measured in terms of its efficiency ϵ and dilution $\mathcal{D}$. The efficiency corresponds to the fraction of events to which a tag can be assigned. The dilution is defined as $1 - 2\mathcal{P}_w$, where $\mathcal{P}_w$ is the probability that the assigned tag is incorrect.

The dilution of a tagger is dependent on characteristics of the event. In an effort to extract as much information as possible from the available data, each candidate with a non-null tag decision is assigned a dilution which is parameterized as a function of various characteristic quantities of the candidate. The output of a tagging algorithm is thus a weighed decision, where the dilution constitutes the weight. The sample of candidates for which the tagging algorithm cannot determine the flavor at production is assigned a null decision. The candidate-by-candidate dilution is determined by parameterizing the average dilution in bins of characteristic quantities of the event. In each bin, the average dilution is equal to the ratio between the difference and the sum of the numbers of candidates, in that bin, which are correctly and wrongly tagged. These numbers are indicated by N_{RS} ("right" sign) and N_{WS} ("wrong" sign), respectively, in the following formula for the average dilution $\mathcal{D}$:

$$\mathcal{D} = \frac{N_{RS} - N_{WS}}{N_{RS} + N_{WS}} \, . \tag{6.1.1}$$

For example, the use of a MC sample allows one to exactly know whether tags are correct or not — more on this subject in Section 6.7. This procedure therefore provides the correct dilution of the candidate as a function of the chosen variables.

Finally, the candidate-by-candidate dilution provided by a flavor tagging algorithm is fine-tuned by calibrating the flavor tagger. The procedure adopted for calibrating a flavor tagger consists in the multiplication of the dilution which the tagger assigns to a candidate by a scaling factor, uniquely characterizing a tagging algorithm. This scale factor provides a global correction for candidate-by-candidate dilutions.

In the rest of this document, the parameter $S_{\mathcal{D}}$ indicates this global scale factor. The use of a single scale factor per tagging algorithm allows for the quantification of differences between the predicted and the actual dilutions of tagging algorithms when applied to the collected data sample. If the dilution parameterizations (both the functional forms and the constants of the parameterizations) are adequate and directly applicable to the samples to be fitted, the scale factors are expected to be consistent with unity.

The scaling factor of the same-side algorithms presented in this section, when applied to the B_s^0 samples utilized in this analysis, are determined by performing fits of MC events. The scaling factor is a free parameter of the fits of mass and proper decay-time of MC candidates, which are known to have been produced with Δm_s fictitiously set to zero. The B_s^0 candidates in the MC sample utilized for the tagger calibration are separated according to the tagger decision — mixed, unmixed, or untagged — and their mass and proper-decay-time distributions are fit simultaneously. The following equation presents a generalization of Equations 1.2.8 and 1.2.9, which assume the use of a perfect tagger, to the case in which a flavor tagger with dilution $\mathcal{D}$ is utilized:

$$\mathcal{P}_{unmixed/mixed}(t) \propto \left[1 \pm \mathcal{D} \cos\left(\Delta m_s t\right)\right] \, . \tag{6.1.2}$$

These equations describe the probability $\mathcal{P}$ that a B_s^0 meson produced at time $t = 0$ decays at time t with the same ("unmixed"), or the opposite ("mixed") flavor as at production. Because it is known that simulated B_s^0 candidates do not mix, the expressions in Equation 6.1.2, with $\Delta m_s = 0 \text{ ps}^{-1}$ and $\mathcal{D} \to S_{\mathcal{D}}\mathcal{D}$, allow for the measurement of $S_{\mathcal{D}}$. The calibration of flavor taggers is extremely important when a mixing analysis is expected to set a limit on Δm_s, as shown in Section 5.1, and when different flavor taggers are utilized, because it is necessary to know how to weigh them. The need for a correct scale factor for the same-side tagging algorithm constitutes one of the main reasons behind the MC-tuning efforts described in Chapter 4. The determination of scale factors for same-side taggers will be presented again in Section 6.7.

As a final remark, it is relevant to note that one does not need a perfect tagger, rather a well-calibrated one. The dilution provided by a tagging algorithm may not be optimal, which is the result, for example, of neglecting to account for any variable which is correlated with the dilution, thus averaging over it. However, when an algorithm is correctly calibrated, not having the optimum dilution does not undermine a mixing analysis any more than it does to not have the optimum event selection.

6.2 Principle of Same-Side Tagging

The Same-Side Taggers (SST) presented in the next sections are based on the correlation between the flavor of the b quark contained in the B meson of interest and the charge of the particles that are most likely produced during the hadronization process of the B meson itself [88, 104].

A $\overline{\mathrm{B}}_s^0$ meson, a $b\bar{s}$ bound state in terms of quarks, is produced when an $s\bar{s}$ pair is pulled out of the vacuum in proximity to the b quark. This leaves an s quark which can contribute to the formation of a kaon. The same-side tagger algorithm tries to recognize the leading fragmentation particle. In the fortunate case in which this particle is a light, charged, kaon, its charge indicates the b quark flavor: K^-'s follow $\overline{\mathrm{B}}_s^0$'s, while K^+'s are typically close to B_s^0's, as described in Figure 6.1.

From the point of view of the experimentalist, the tagger is expected to have good efficiency because the track which carries the flavor information is close to the candidate which triggered the event, and therefore has a high chance of ending up in the geometrical acceptance of the detector.

Nevertheless, there is a relevant issue in the case of same-side flavor tagging, which is that there is no straightforward way to measure its dilution on data, because the tagging characteristics depend on the particular B meson. That could only be possible if B_s^0 oscillations are observed, and the dilution is fitted from the data.

The study of same-side tagging in the environment of an hadronic collider presents many challenges. It is necessary to understand the production mechanism of b and $\bar{b}$ quarks, their hadronization in B mesons, the type of particles that are produced during the hadronization process. As mentioned above, the strangeness of B_s^0 mesons indicates strange particles as the best tagging particles. The algorithms for same-side flavor tagging for B_s^0 mesons are thus often referred to as same-side kaon taggers. Among the factors which modify their performance, it is worth mentioning the production of resonances such as excited B mesons, the decay of which produces the B_s^0 meson under study, and the production of B_s^0 mesons in association with resonances such as K^{*0}, or ϕ^0.

The following sections present the algorithms for same-side tagging which pave the way to the algorithm utilized in this analysis of B_s^0 oscillations. The problematics which derive from the physics of the process of B_s^0 production and the necessity to utilize a MC sample for the calibration of a tagging algorithm are treated in the section dedicated to systematic uncertainties on the calibration of same-side taggers.

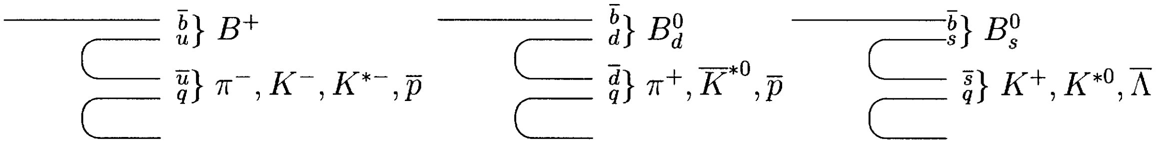

FIGURE 6.1: Schematic drawing of particles produced in association with different B mesons.

6.3 Selection of tag candidates

Different SST algorithms have been studied to select the track that is most likely to be the leading fragmentation track. The purpose of these studies was to find the algorithm which would have the best performance when applied to samples of B^0_s candidates. The next sections present some of these algorithms, with references to their first introduction.

The implementations of these algorithms for their use at CDF II share the same initial selection of tracks which form the ensemble of tag candidates. The selection cuts are divided in three main categories, which are presented in the next paragraphs.

The first set of cuts consists of requirements on the quality on the tagging track candidates, and that tag candidates are contained in a fiducial volume of the CDF detector:

- ***#Si hits* $\geq$ 3, *#COT stereo hits* $\geq$ 10 and *#COT axial hits* $\geq$ 10**
  The requirement on the number of hits used in the reconstruction of the track selects candidates with a reliable track fit. This selection test is widely utilized in CDF analyses. As a quality criterion, it enforces tracks to pass through the central region of the CDF detector. The distributions of the number of hits per track in the silicon-based detectors and in the COT are shown in Figures 4.2 and 4.5.

- **$p_T \geq$ 450 MeV/c**
  The tracking performance is asymmetrical with respect to charge for low momentum tracks. This is due to the design of the COT. In fact, the cells of the COT are tilted, with respect to the radius which connects a cell to the center of the detector, as it is visible in Figure 2.9. The section of a cell thus appears different to positively and negatively charged particles, which translates in a different tracking efficiency. This cut allows one to avoid this problem, without significantly affecting the performance of the tagger. The comparison of the p_T distributions in data and in PYTHIA-MC events of tag candidates is shown later in this chapter, in Figure 6.10.

- **$|\eta| \leq 1$**
  The cut on the pseudorapidity of the candidate track is strongly correlated to the requirements on ΔR and the number of hits, which prefer candidates in the central region. It removes a remaining 10% additional tracks above $|\eta| = 1$ which hardly have any TOF information. Moreover, the COT $d\mathrm{E}/dx$ performance is

well understood in the $|\eta| < 1$ range. In short, the cut removes preferentially tracks that are of low quality for tagging purposes. The distribution of η is presented in Figure 4.5.

The second set of cuts is introduce to enrich the pool of tag candidates with tracks that are close, in phase space, to the B candidate, and are likely to have been produced at the primary vertex of the $p\bar{p}$ interaction:

- **ΔR(B, *trk*) $\leq$ 0.7**
  ΔR is defined as the distance in the η–ϕ space between the reconstructed B meson and the tag candidate track:

  $$\Delta R \equiv \sqrt{[\phi_0(\mathrm{B}) - \phi_0(trk)]^2 + [\eta(\mathrm{B}) - \eta(trk)]^2}\,. \tag{6.3.1}$$

  The cut selects the tracks which are close to the B candidate. As a side note, the opposite-side taggers that will be used in this mixing analysis, described in Section 5.3.1, apply a complementary request ($\Delta R > 0.7$). The sets of tracks for opposite-side and same-side taggers are thus separated without overlap. The distribution of ΔR is presented in Figure 4.5.

- **$|\Delta z_0$(B, *trk*)$|$ $\leq$ 1.2 cm**
  The purpose of this cut is to remove tracks coming from $p\bar{p}$ interactions different from the one which produced the reconstructed B_s^0 candidate. The z resolution of tagging track candidates is shown in Figure 4.5, where the distributions of Δz_0 in data and PYTHIA-MC simulation are shown. The cut corresponds to about 3 standard deviations.

- **$|d_0/\sigma_{d_0}| \leq 4$**
  This impact parameter significance cut selects the tracks which come from the primary vertex of the interaction, where the B candidate is produced. The distribution of d_0/σ_{d_0} is shown in Figure 4.5.

The last set of cuts rejects tracks which are identified as not been possible tag candidates:

- **rejection of e, μ and conversions**
  A likelihood-based cut is applied to reject tracks that are likely to have been produced by an electron or a muon. In addition, tracks that are consistent with coming from a $\gamma \to e^+e^-$ conversion are removed. More details on the likelihood functions utilized to identify leptons are presented in Section 3.4.1.

- **rejection of B daughter tracks**
  The tracks which are used in the fit of the B candidate are explicitely excluded from the list of tagging track candidates.

Once all the selection cuts are applied, B candidates are left with zero, one or more tag candidates. The number of B candidates without any track passing the selection defines the efficiency of the tagging algorithm:

$$\epsilon = 1 - \frac{N_{0\ cands}}{N_{TOT}}. \tag{6.3.2}$$

When one or more tag candidates are present, events are naturally divided in two classes:

- **agreeing case:** if a single track is selected, or the charges of all tag candidates are identical;
- **disagreeing case:** if not all of the tag candidates have the same charge.

In the first case, the SST decision is the same for all SST algorithms, and correspond to the charge of the selected track(s), while for events in the second class each algorithm needs to provide a method to select the decision.

The distributions of the number of tagging track candidates in data and in PYTHIA-MC events are compared in Figure 6.2. The agreement between the two distributions represents an important confirmation of the goodness of the MC simulation.

Various tagging algorithms have been extensively studied, measuring their performance in different data samples and reconstructed B final states. The algorithms differ in the method adopted to select the tag candidate among the tracks which satisfy the previous selections. The first implementation of a same-side tagging algorithm in a CDF analysis is presented in Reference [105], which pioneered same-side tagging with an algorithm based on kinematics. The status of same-side tagging in CDF II before the introduction of the algorithm described in Section 6.6, and utilized in this mixing analysis, is summarized in Reference [106], which presents a review of various kinematic-based algorithms and an initial study of the particle-identification–based algorithm which would have been used in the analysis that resulted in the first measurement of Δm_s [17].

The algorithm of same-side tagging utilized in this analysis of B_s^0 oscillations utilizes a Neural Network to combine kinematic and particle-identification information of tagging candidate tracks. The next two sections present a review of the kinematic-based (Section 6.4) and particle-identification–based (Sec. 6.5) algorithms which were more accurately studied. These algorithms provided an excellent starting-point for the preparation of an improved tagger which would combine the pieces of information that they use. The combined tagger which is finally utilized in the analysis presented in this dissertation is documented in Section 6.6, while other attempts to combine kinematic and particle identification information, which were later discarded, are presented in Appendix C.

6.4 Kinematic based taggers

The kinematic characteristics of the leading fragmentation track are correlated with those of the B candidate, because the track is expected to be found close in phase space to the B meson. It is thus possible to exploit such correlations to select a tag

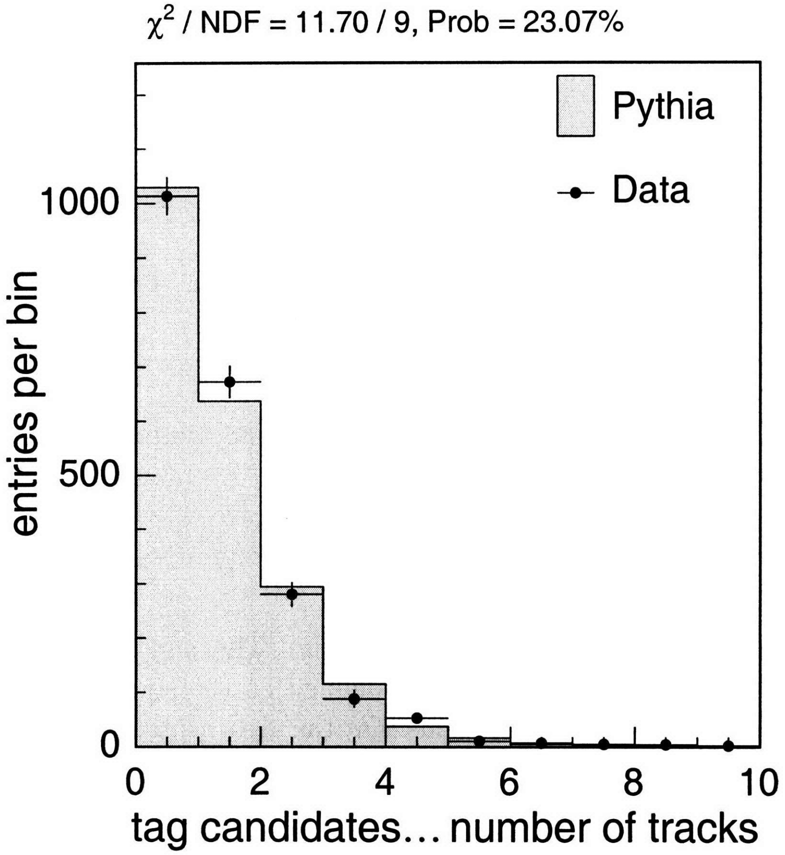

FIGURE 6.2: Distributions of the number of tagging track candidates in data and PYTHIA-MC events.

candidate, the charge of which will provide the decision of the corresponding flavor tag algorithm.

Different algorithms based on kinematic quantities have been studied in recent years, and are summarized in Reference [106]. They are conceptually similar. Each of them selects a kinematic variable which is correlated with the closeness in phase space that is, on average, expected between the B candidate and the best tag candidate. An SST algorithm thus chooses the tag candidate such that the selected variable is maximized, or minimized, depending on whether it is correlated, or anti-correlated, with its proximity in phase space . The decision of the tagging algorithm is given by the charge of the best tag candidate. Finally, the candidate-by-candidate dilution is parameterized in terms of a kinematic variable.

The algorithm which was most thoroughly studied selects the track with the maximum p_L^{rel} as the tag candidate, and is referred to as $\max p_L^{\text{rel}}$. The variable p_L^{rel} is graphically defined in Figure 6.3. The performance of this algorithm, when applied to a PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$, is expressed in terms of its *effective* dilution $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle}$:

$$S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 22.8 \pm 0.7\%\,,$$

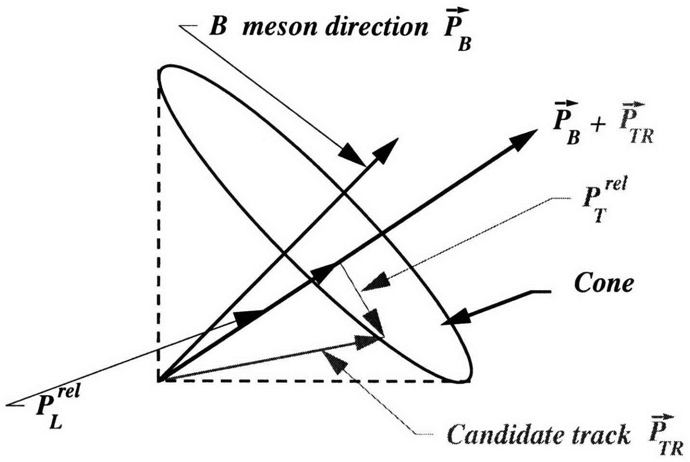

FIGURE 6.3: Definition of p_L^{rel}.

where the quoted uncertainty is statistical only. The *effective* dilution of a flavor tagging algorithm applied to a sample is defined as the square root of the average squared candidate-by-candidate dilutions that the algorithm assigns to each candidate in the sample, multiplied by the scale factor $S_{\mathcal{D}}$ obtained by calibrating the tagging algorithm, as described in Section 6.1. This scale factor depends on the flavor tagging algorithm which is utilized and on the sample to which the flavor tagger is applied. The candidate-by-candidate dilution is parameterized as a function of the transverse momentum p_T of the tag candidate. Because the pool of tag candidates utilized for all the same-side tagging algorithms presented in this chapter is selected by applying the same cuts, listed in Section 6.3, the effective dilution is sufficient to classify these algorithms in order of performance.

The $\max p_L^{\mathrm{rel}}$ algorithm for same-side tagging is found to perform worse than the particle-identification–based algorithm which will be described in the next section. However, the availability of this kinematic-based algorithm presented the opportunity to study the combination of kinematic- and particle-identification–based algorithms. The result of this study is the algorithm described in Section 6.6, which is the one finally used in this analysis of B_s^0 oscillations.

6.5 Particle-identification–based tagger

Most of the prompt tracks that are produced in a proton–anti-proton collision are pions. As seen in Figure 6.1, kaons are likely to be produced around B_s^0 mesons during the hadronization process. Prompt tracks which are identified as kaons thus carry, on average, important information about the flavor of the B_s^0 meson with which they were produced. It is not surprising that a tagging algorithm based on particle identification achieves better performance than the ones which do not use this piece of information.

The identification of particles is based on information provided by the COT and

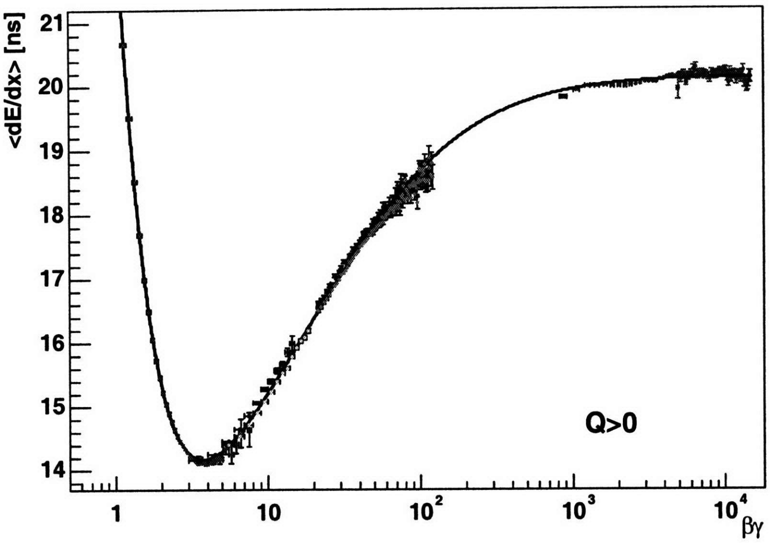

FIGURE 6.4: Average $d\text{E}/dx$ versus particle $\beta\gamma$ as measured for various particle species in CDF. The plot is produced utilizing only positively charged particles. From left to right, it is possible to distinguish the contributions of the calibration samples of protons, kaons, and pions.

TOF detectors. The descriptions of these detectors are presented in Sections 2.2.4 and 2.2.5. The variables utilized to assign particle-identification are introduced in Section 4.3.4, where the tuning of the MC samples is described, The procedure of particle-identification is summarized in the next paragraphs.

Specific energy loss per unit length, usually referred to as $d\text{E}/dx$, is correlated with the type of particle under consideration. The measured $d\text{E}/dx$ is calibrated, as described in Reference [40], and then utilized to evaluate the variable Z, already defined in Equation 4.3.8 as follows:

$$Z(i) \equiv \log\left[\frac{(d\text{E}/dx)^{cor}}{(d\text{E}/dx)^{pre}(i)}\right], \quad i = \pi, \text{K}, p, \tag{6.5.1}$$

where $(d\text{E}/dx)^{cor}$ refers to the calibrated $d\text{E}/dx$ and the predicted expectation is obtained from the universal curve in Figure 6.4, which expresses $d\text{E}/dx$ as a function of the particle speed. The probability that a particle is a kaon, a pion, or a proton, is evaluated by comparing the three ratios $Z(i)$ calculated for the calibrated $d\text{E}/dx$ of the particle with the distributions of the $Z(i)$ variables obtained in samples of pure kaons, pions, and protons. The $d\text{E}/dx$ information is available for essentially all the tracks utilized in this analysis and provides a constant 1.4–standard-deviation separation between kaons and pions with $p_T > 2$ GeV/c.

The TOF detector measures t_{flight} information. The variable utilized to assess the probability that a particle is a kaon, a pion, or a proton, is the t_{flight} residual, defined as follows:

$$\Delta t_{\text{flight}} \equiv t^{meas}_{\text{flight}} - t^{pre}_{\text{flight}}, \tag{6.5.2}$$

where the predicted t_{flight} is defined in Equation 4.3.13. The discriminating power of the CDF TOF system is > 2 standard deviations for kaons and pions with $p_T <$ 1.5 GeV/c. The efficiency with which t_{flight} is assigned to a particle, as shown in Figure 4.3, is dependent on the track transverse momentum and is measured to be about 65%. On the other hand, a large fraction of tag candidate tracks have momentum inferior to 1.5 GeV/c (Figures 6.10 and 6.11), which makes the TOF contribution of great importance.

The pieces of particle identification information from the COT and the TOF are combined in a single variable, CLL, which is defined as follows:

$$CLL = \log\left(\frac{\mathcal{L}(\mathrm{K})}{f_p\mathcal{L}(p) + f_\pi\mathcal{L}(\pi)}\right), \tag{6.5.3}$$

where:

$$\mathcal{L}(i) = \mathcal{P}_{t_{\text{flight}}}(i) \cdot \mathcal{P}_{d\mathrm{E}/d\mathrm{x}}(i)\,, \quad i = \pi, \mathrm{K}, p\,. \tag{6.5.4}$$

In the previous formula, $\mathcal{P}_{t_{\text{flight}}(d\mathrm{E}/d\mathrm{x})}(i)$ is the probability that the measured t_{flight} ($d\mathrm{E}/d\mathrm{x}$) is consistent with the hypothesis that the particle type is i, where $i = \pi, \mathrm{K}, p$. The variable CLL is thus the ratio of the likelihood $\mathcal{L}$ for the signal hypothesis divided by the one corresponding to the background hypothesis, which is constituted mainly by pions. The *a priori* fraction of background pions, f_π, is equal to 0.9, while protons constitute the remaining 10%, $f_p = 0.1$. The likelihood that a particle is of a certain type, $\mathcal{L}(i)$, where i is either π, K, or p, is the product of the probabilities $\mathcal{P}$ that the measured $d\mathrm{E}/d\mathrm{x}$ and t_{flight} are consistent with the hypothesized particle type, as shown in Equation 6.5.4.

The distributions of CLL in data and in PYTHIA-MC events, where MC truth information is utilized to separate the contributions of pions, kaons, and protons, are compared in Figure 6.5. The distributions produced by utilizing $d\mathrm{E}/d\mathrm{x}$ and t_{flight} each separately are shown in Figure 6.6. The definition of CLL in Equation 6.5.3 implies that the higher the value of CLL, the more probable the particle is a kaon. As expected, the greater separation provided by TOF, when its piece of information is available, with respect to $d\mathrm{E}/d\mathrm{x}$ is visible in the figures, where the CLL distribution of true kaons is more evidently shifted toward higher CLL values.

The particle-identification–based tagging algorithm presented in this section selects the tag from among the candidate tracks by taking the track which maximizes CLL. The tag decision is the charge of the selected track and the dilution is parameterized as a function of CLL. Two scenarios are distinguished, each of which adopts an independent parameterization, according to whether all the tag candidates have the same charge or not. In the first case, the tagging algorithms does not have to make a decision, while in the latter case the $\max CLL$ algorithm makes a non-trivial decision. The functional form of the two $\mathcal{D}$ *vs.* CLL curves is identical, but they have different final parameters, as visible in Figure 6.7. The details of the parameterizations, and the values of the parameters used, are presented in Section C.3.

The particle-identification–based same-side flavor tagging algorithm was utilized in the analysis of $\mathrm{B}^0_s - \overline{\mathrm{B}}^0_s$ oscillations presented in Reference [17]. The performance

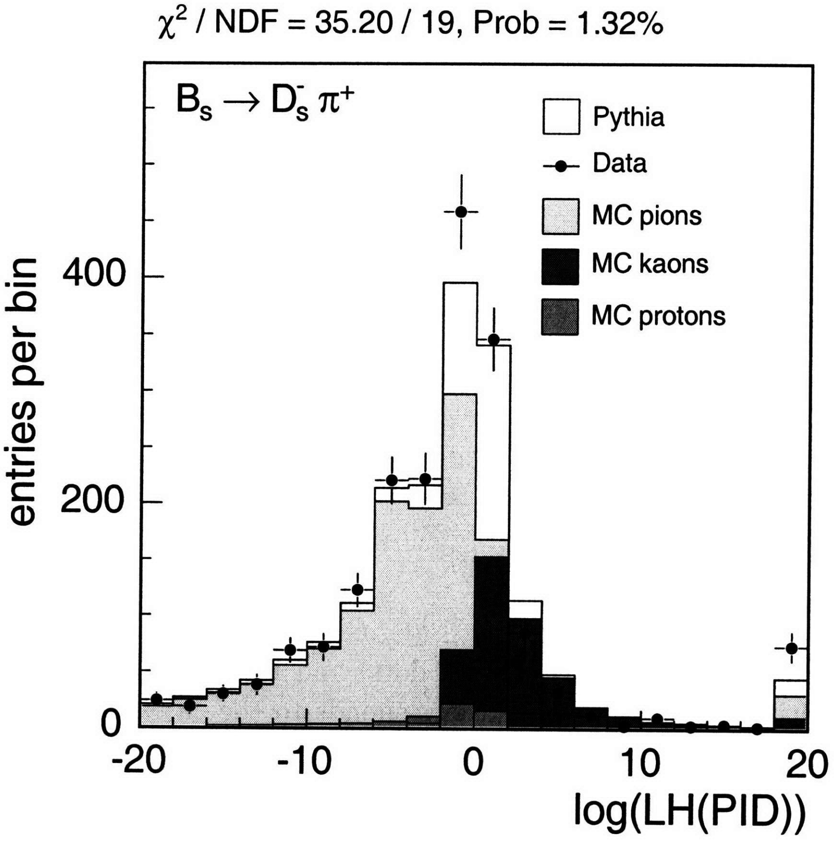

FIGURE 6.5: Distribution of CLL for tagging track candidates in data (black dots) and PYTHIA-MC events (histogram). The rightmost bin in the plot corresponds to the cases where neither $d\text{E}/d\text{x}$ nor t_{flight} information are available. The contributions of kaons, pions, and protons to the PYTHIA-MC plot are divided on the basis of MC truth information, and overlaid. The B_s^0 meson candidates are reconstructed in the $\text{B}_s^0 \to \text{D}_s^- \pi^+$, $\text{D}_s^- \to \phi^0 \pi^-$ decay mode.

of the algorithm is quoted in terms of the effective dilution $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle}$:

$$S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 28.5 \pm 0.7\% \,.$$

The effective dilution is calculated in a PYTHIA-MC sample of $\text{B}_s^0 \to \text{D}_s^- \pi^+$, $\text{D}_s^- \to \phi^0 \pi^-$. The quoted uncertainty is statistical only. This algorithm for same-side tagging constitutes the starting point for the algorithm used in the analysis of B_s^0 oscillations presented in this document.

6.6 Neural Network Same-Side Tagger

The previous sections presented a selection of tagging algorithms which have been studied in detail, and proficiently utilized in CDF analyses. It is natural to consider combining the particle identification variable with the kinematic description of the tag candidate. Many different approaches, which are detailed in Appendix C, have

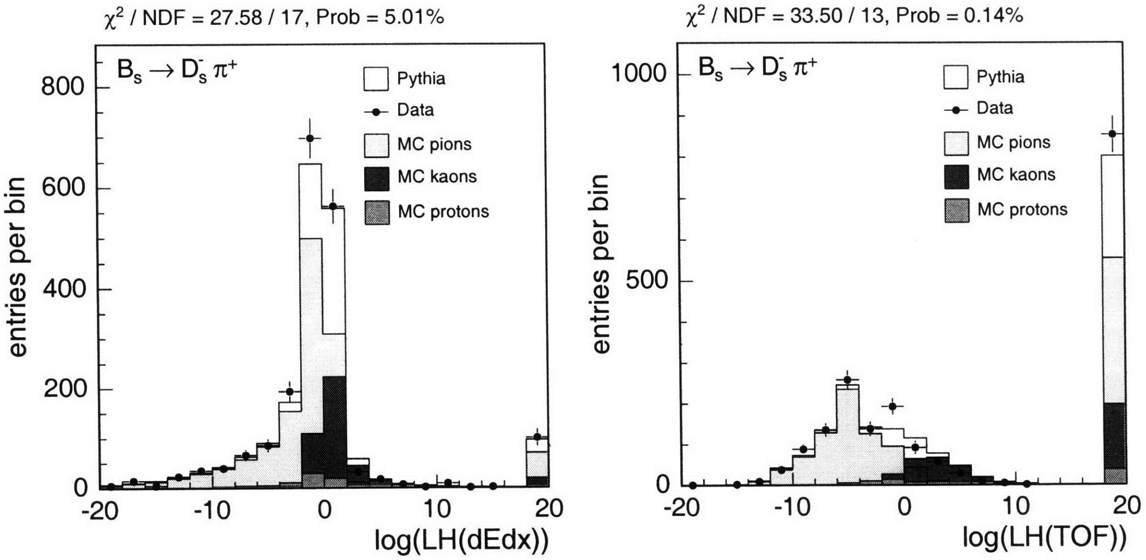

FIGURE 6.6: Distribution of CLL with $d\mathrm{E}/d\mathrm{x}$ only (left) and t_{flight} only (right) for tagging track candidates in data (black dots) and PYTHIA-MC events (histogram). The rightmost bin in the two plots, corresponds to the cases where no $d\mathrm{E}/d\mathrm{x}$ (left plot) or no t_{flight} (right plot) information is available. The contributions of kaons, pions, and protons to the PYTHIA-MC plots are divided on the basis of MC truth information, and overlaid. The B_s^0 meson candidates are reconstructed in the $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \to \phi^0 \pi^-$ decay mode.

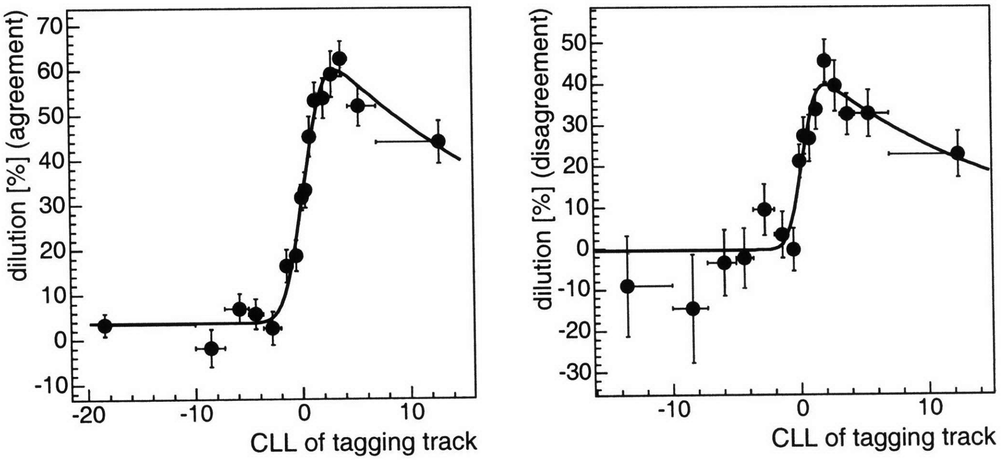

FIGURE 6.7: Parameterization of the dilution of the maxCLL SST algorithm as a function of CLL when the tagger is applied to $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$ candidates. The plot on the left contains the parameterization used when the tagging track is unique or all the tag candidates among which it has been chosen have the same charge. The plot on the right contains the parameterization used when the tagging track is chosen in an ensemble of tag candidates which have different charge. The parameterizations are obtained as described in Section 6.1.

been considered. Finally, an Adaptive Neural Network (ANN) was chosen to perform the combination. A brief summary of the principle of an ANN is presented in the next paragraph. The performance of the ANN-based tagger is only slightly superior to the particle-identification–only one. Nonetheless, the improvement with respect to the previous tagging algorithm is statistically significant because measured utilizing exactly the same events and the same tagging tracks. The comparison of the tagging performance of these two algorithms in a $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ data sample, is presented at the end of this chapter, in Table 6.8.

Neural Networks (NN) provide an algorithm for information processing which mimics biological neural systems. The physical quantities that are used by the ANN to discriminate signal from background are passed to the network through *input nodes*. These pieces of information are passed to the nodes which compose the *hidden layers* of the NN. The nodes in the hidden layers represent the neurons of the network, and are connected to the *output node(s)*. Each of the j-th nodes in a hidden layer receives a set of inputs x_{ij}, and calculates the weighed sum y_j:

$$y_j = \sum_i w_{ij} x_{ij} + c_j \,, \tag{6.6.1}$$

where w_{ij} are the weights assigned to the i-th quantity entering the j-th node and c_j is a bias characterizing the node, and independent of the input variables. The response of these internal nodes are modeled by an *activation function*, which is typically chosen to be a sigmoid function $g(y)$:

$$g(y) = \frac{1}{1 + e^{-y}} \,. \tag{6.6.2}$$

Similarly, each of the output nodes finally returns a value o obtained as follows:

$$o = g\left(\sum f_j g(y_j)\right) \,, \tag{6.6.3}$$

where f_j represents the weight applied to the output of the j-th node in the hidden layers connected to the output node. In the case of the ANN utilized by the same-side tagging algorithm presented in this section, the output node is unique, and returns a number between zero and unity. The training of a NN consists in the selection of the weights w_{ij}, c_j, and f_j for the internal and output nodes. The method adopted to train the ANN used in this analysis for same-side tagging is the *back propagation* method [107]. Neural Networks provide a conceptual advantage over a likelihood ratio, an example of which is in Equation 6.5.3, to discriminate signal from background. When more than one input variable are utilized, neural networks are able to exploit correlations between inputs by adjusting their weights, whereas likelihood ratios cannot because they are just built from the product of the individual probability density functions of the input variables.

The ANN input is constituted by CLL, the piece of information regarding particle identification (Equation 6.5.3), and various kinematic quantities which have been introduced in Section 6.4: p_T, $p_T^{\rm rel}$, $p_L^{\rm rel}$, and ΔR. Experience with the CLL-based tagger suggests the addition to the list of input variables a boolean value which is true when all the tag candidates have the same charge. The ANN is trained to select kaons with the correct charge correlation with the flavor of the B_s^0 candidate. The PYTHIA-MC sample which contains $B_s^0 \rightarrow D_s^- \pi^+$, $D_s^- \rightarrow \phi^0 \pi^-$, and charge-conjugate, decays is utilized for the training. The network maximizes its output on the tag candidates which are kaons with the correct charge correlation. In the case of a B_s^0 candidate, positively charged kaons receive a high network output. The subsample containing tag candidates which are pions, protons or kaons with opposite charge correlation represents background. The network is trained to minimize its output on these type of tracks. The distribution of the ANN output in signal and background candidates is shown in Figure 6.8. In the same figure, the plot of purity *versus* efficiency of the ANN and the distribution of correlations among the input variables listed above are shown.

The data–MC-simulation comparison of the kinematic characteristics that are directly utilized by the ANN-based tagging algorithm is presented in Figures 6.10 and 6.11. The comparisons are performed on the set of tracks which satisfy the requirements for being a tag candidate, and PYTHIA-MC truth information is exploited to separate the contribution of pions, kaons and protons. These components are separately shown to appreciate their different contribution to the total sample. By applying the cut $CLL > 1$, where CLL is the quantity defined in Equation 6.5.3 which contains the particle-identification information, it is possible to isolate a sample that the MC simulation shows to be highly enriched in kaons.

The comparisons of events in data and in MC simulation provide the confidence that the training of the ANN obtained with MC events is optimal for data, too. The plots presented there show the data–PYTHIA-MC agreement of transverse momentum p_T, ΔR, longitudinal and transverse momentum of the tag candidate, $p_T^{\rm rel}$ and $p_L^{\rm rel}$ (graphically defined in Figure 6.3), relative to the B candidate, in the complete sample, and after the $CLL > 1$ cut.

The decision of the tagging algorithm is the charge of the tag candidate which maximizes the output of the ANN trained as described in the previous paragraph. The dilution is parameterized as a function of this ANN output. Similar to the CLL tag case, the parameterization is different in the two cases defined by the charge-agreement among the tag candidates, as shown in Figure 6.9.

The ANN-based same-side tagging algorithm is the algorithm utilized in the analysis presented in this document. The next sections will describe the calibration of this tagger and the study of systematic uncertainties on the calibration scale factor $S_{\mathcal{D}}$. The presentation of the performance of this same-side tagging algorithm is postponed to Section 6.10.

6.7 Calibration of the Same-Side Tagger

It is important to stress that the correct calibration of a tagger is necessary when a limit on Δm_s is sought, and when the tagger is used together with other taggers, as stated in Section 6.1.

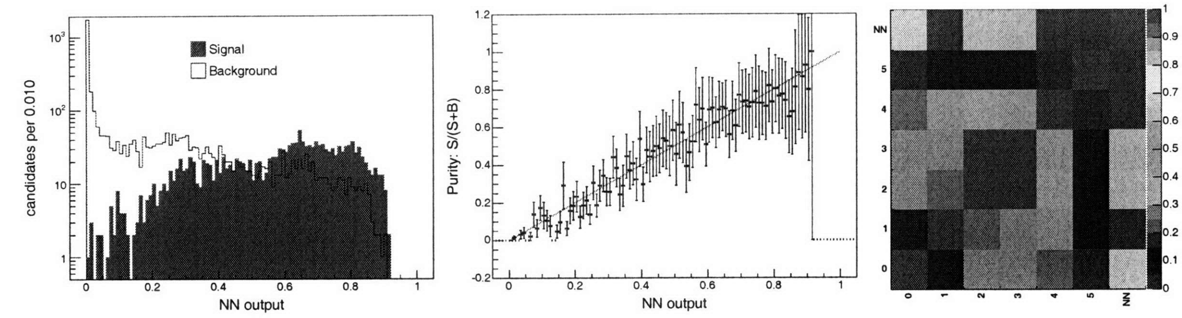

FIGURE 6.8: Plots of Neural Network training. Distribution of the ANN output for signal and background candidates (left), purity *vs.* ANN output (center), and correlations among the input variables of the ANN (right). The distribution of ANN output shows the signal-background separation achieved. For an optimally trained ANN, purity, defined as $\mathcal{S}/\mathcal{S}+\mathcal{B}$, is expected to be a linear function of the ANN output. The number from 0 to 6 identify CLL, p_T, p_T^{rel}, p_L^{rel}, and ΔR of the tag candidate, and the boolean variable which is true when the tag candidate is unique or selected among a pool of tracks with the same charge. The training is performed on a PYTHIA-MC sample of $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \to \phi^0 \pi^-$ candidates.

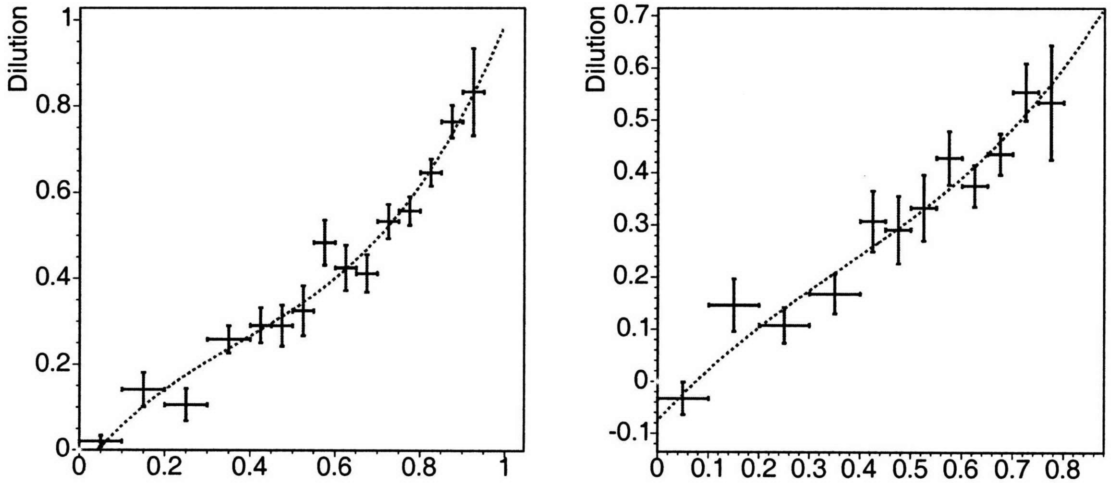

FIGURE 6.9: Parameterization of dilution as a function of ANN output in the following two cases: the tagging track is unique, or is extracted from a set of tag candidates with the same charge (left), and the tagging track is chosen among a set of tag candidates with different charges (right). The parameterizations are obtained as described in Section 6.1.

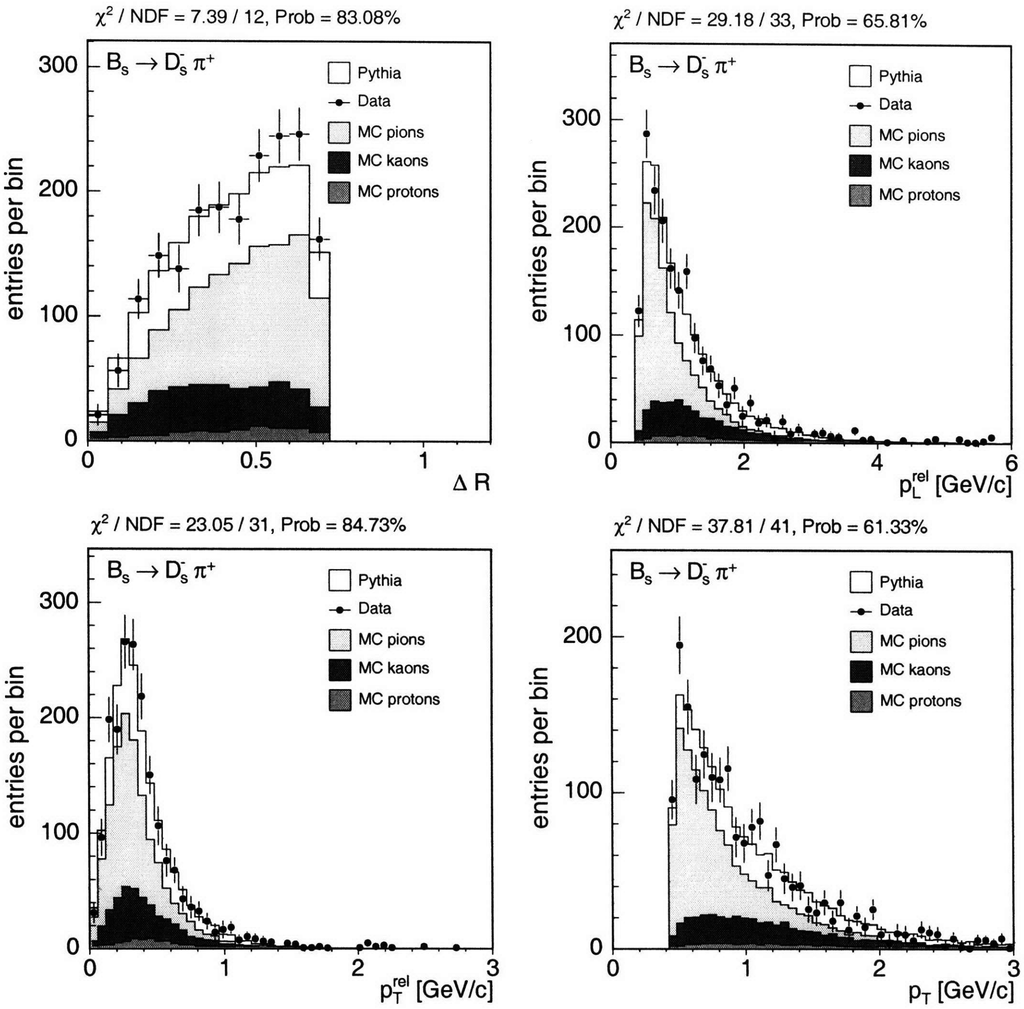

FIGURE 6.10: Data–PYTHIA-MC comparison of tag candidate variables. From left to right, and top to bottom, are plotted the distributions for: angular separation $\Delta R(\mathrm{B}, trk)$, longitudinal and transverse momentum p_T^{rel} and p_L^{rel} relative to the B candidate and transverse momentum of tagging track candidates. The contributions of kaons, pions, and protons to the PYTHIA-MC plots are divided on the basis of MC truth information, and overlaid, for the purpose of showing the relative population of different species.

The principle upon which the SST algorithms described in this section are based upon relates the B meson under study with the type of particles produced in association with it. It is thus expected that same-side taggers perform differently in the case of B^+, B^0 or B_s^0.

It is possible to calibrate flavor taggers for B^+ candidates directly on B^+ data. Because charged B mesons do not mix, the flavor of a candidate is directly indicated by the charges of the tracks in the candidate's final state, and the expressions in Equation 6.1.2 are utilized with the cosine term set equal to unity. The B^+ data sample constitute an extremely valuable source of information for the calibration of taggers because of the fact they do not mix and the correctness of a tagging decision

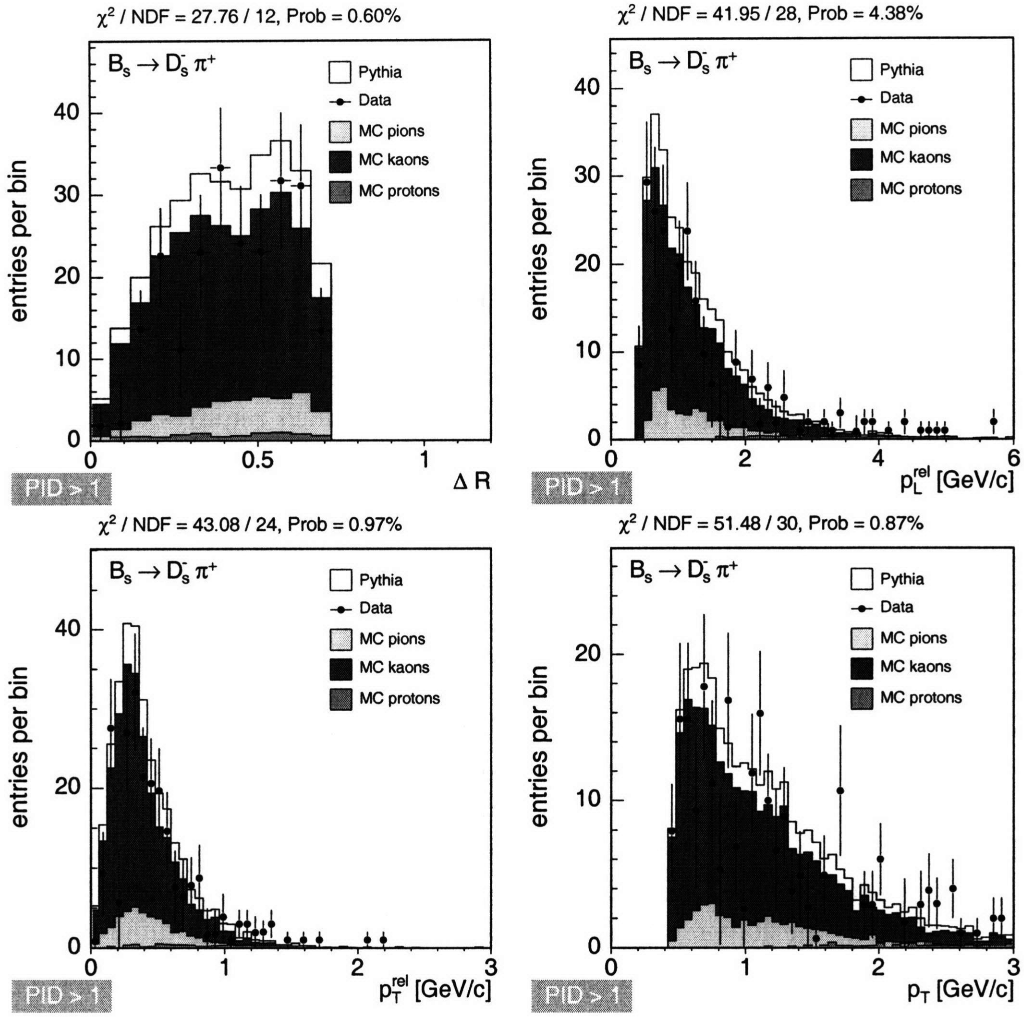

FIGURE 6.11: Data–PYTHIA-MC comparison of tag candidate variables with CLL cut. From left to right, and top to bottom, are plotted the distributions for: angular separation $\Delta R(\mathrm{B}, trk)$, longitudinal and transverse momentum p_T^{rel} and p_L^{rel} relative to the B candidate and transverse momentum of tagging track candidates. The label PID indicates CLL, defined in Equation 6.5.3. The cut $CLL > 1$ enriches the sample in kaons, as clearly shown by the separation of the PYTHIA-MC sample by particle species.

is readily known without uncertainty.

For the calibration of flavor taggers applied to B^0 candidates, B^0 data is also directly usable. In this case, the reason is that the CDF data samples are sensitive enough to measure the oscillation frequency Δm_d, a precise measurement of which is independently known [7]. It is thus possible to perform, at the same time, a fit for Δm_d *and* for the scale factor $S_{\mathcal{D}}$ of a tagger, and the value of Δm_d resulting from the fit can be compared to its world average. The fit utilizes Equation 6.1.2, with the substitution of Δm_s with Δm_d.

It is also notable that a same-side kaon tagger is expected to perform well on a B^+ sample. In fact, Figure 6.1 shows that the correlation between the flavor of the

B^+ meson and the charge of the same-side kaon is the same as with same-side pions, which are the next closest particles that the tagging algorithm may select. In the case of B^0 meson, the correlation between the charge of the same-side kaon and the flavor of the B^0 candidate is contrary to the correlation between the charge of the same-side pion and the flavor of the B^0 candidate. Same-side kaon taggers are thus expected to perform better on B^+ samples than on B^0 samples (additional information is provided in Appendix B).

In the case of the analysis of B_s^0 oscillations presented in this document, instead, same-side taggers have to be calibrated beforehand because, in case a measurement of Δm_s would not be possible, a limit would be set. Besides, other flavor tagging algorithms are utilized in this analysis (Section 5.3), which requires each of them to be calibrated to provide a correct combination. It is thus not correct to fit directly for Δm_s because it is not even known *a priori* if data contain enough information to be sensitive to that quantity.

The calibration of the SST for B_s^0 is performed on a MC sample which reproduces the fragmentation process in which the B_s^0 is generated. Chapter 4 is dedicated to the demonstration that the MC simulation accurately describes the properties of B_s^0 events which are of interest for this analysis. In the simulation the produced B_s^0 does not mix before decaying. The flavor at production is thus, by construction, identical to the flavor at decay, which is indicated without error by the charge of the decay products, because the reconstructed final states are self-tagging. The calibration consists, as anticipated in Section 6.1, in the simultaneous fit of mass and proper decay-time of the reconstructed B_s^0 candidates in the three subsamples which contains B_s^0 candidates which mixed (i.e., the flavor tagger indicates a production flavor different from the flavor as at decay), did not mix, or were not tagged. The following equations are utilized:

$$\mathcal{P}_{unmixed}(t) \propto [1 + S_{\mathcal{D}}\mathcal{D}] \,, \tag{6.7.1}$$

$$\mathcal{P}_{mixed}(t) \propto [1 - S_{\mathcal{D}}\mathcal{D}] \,. \tag{6.7.2}$$

These formulae derive from Equation 6.1.2. The oscillation frequency is equal to zero by construction (B_s^0 candidates in the MC simulation do not oscillate), and $S_{\mathcal{D}}$ is a free parameter in the fit.

The distributions of the ANN input variables in data and of the main characteristics of B_s^0 candidates and B_s^0 events have been thoroughly compared with the corresponding distributions obtained in simulated events, as shown in the plots in Section 4.4.

The most important cross-check of the validity of the procedure is the check that the results obtained in MC samples and in data are consistent, when using various B^0 and B^+ decay modes, where the SST scale factors can be measured directly in data and calculated utilizing MC events. This cross-check has been performed utilizing the particle-identification–based algorithm for same-side flavor tagging described in Section 6.5. Due to technical reasons, such comprehensive study could not be performed for the ANN-based SST algorithm. However, the level of agreement between distributions in data and MC simulation reached provides the confidence that the conclusions

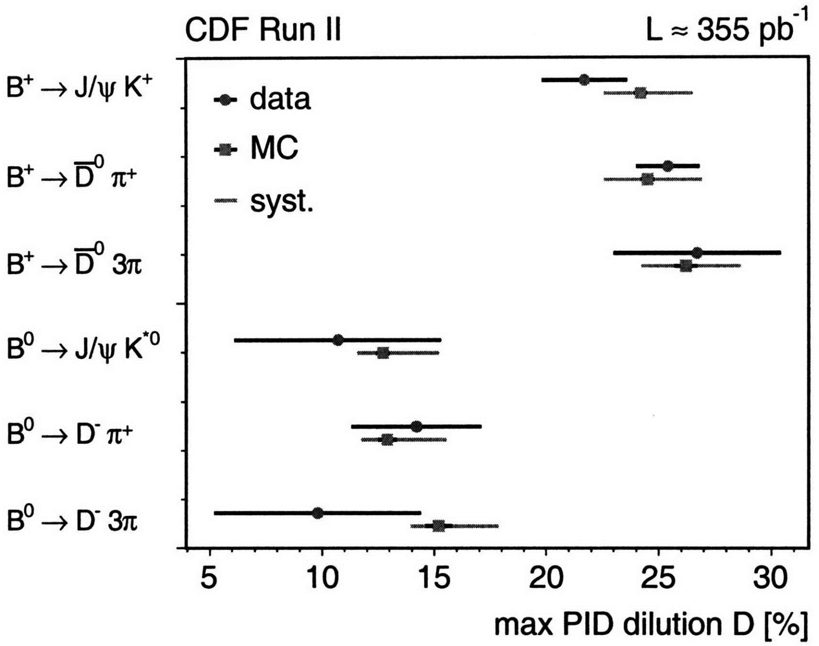

FIGURE 6.12: Data–PYTHIA-MC comparison of dilution of the maxCLL SST algorithm applied to B^0 and B^+ control samples. Data points include statistical uncertainties, while MC entries are drawn with systematic and statistical uncertainties.

of this cross-check are valid. Figure 6.12 shows the comparison of dilutions measured in data, with statistical uncertainty, and in the respective MC sample, with systematic uncertainties applied, in four B^0 and B^+ control samples. The SST algorithm utilized in the comparison is the particle-identification–based one described in Section 6.5. The control modes utilized in this analysis are $B^0 \to D^-\pi^+(\pi^-\pi^+)$, $D^- \to K^+\pi^-\pi^-$, $B^0 \to J/\psi K^{*0}, J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$, $B^+ \to \overline{D}^0\pi^+(\pi^-\pi^+)$, $\overline{D}^0 \to K^+\pi^-$, and $B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-$. The selection of candidates in these control modes is described in Appendix B. Additional plots of data–MC-simulation comparisons are presented as well.

6.8 Systematic uncertainties

The calibration of the tagger using MC samples introduces several possible sources of systematic uncertainties, most of which are connected with the model used to simulate the B_s^0 production process. Each of the next sections will present the evaluation of a source of systematic uncertainties. All systematic uncertainties have been evaluated by using the same MC sample, but reweighing MC events in order to simulate the effect under study which is suspected to contribute to the total systematic uncertainty. In some cases, such as the estimation of particle-identification–related systematic uncertainties, the properties of MC events are reproduced utilizing different algorithms.

The scale factor $S_{\mathcal{D}}$ fitted in the nominal MC configuration is compared to the scale factor fitted in the reweighed, or modified, MC sample. The discrepancy provides an estimate of the systematic uncertainty associate with the effect simulated by the reweighing, or modification. For some of the sources of systematic uncertainties, it has been chosen to directly utilize the uncertainty on the scale factor of the $\max CLL$ algorithm as an estimate of the uncertainty on the scale factor of the ANN-based algorithm. These sources are the b-production mechanism, the choice of a fragmentation function, the possibility of multiple $p\bar{p}$ interactions in the same event record, and particle-identification.

A summary of the studies performed, with an estimate of the total systematic uncertainty, is presented in the last section. The systematic uncertainty associated with each of the effects analyzed is indicated by σ_{S_D} in the tables at the end of this section.

6.8.1 b-production mechanism

Three different processes contribute to the production of $b\bar{b}$ pairs [76]: flavor creation (FC, $\sim$ 25%), flavor excitation (FE, $\sim$ 55%) and gluon splitting (GS, $\sim$ 20%). The current uncertainty in the fractions of these processes is large enough to have some influence in flavor tagging based on MC simulation. A way of estimating the uncertainty associated with the production fractions is to constrain the fractions from the data of this analysis, and then to see how much the MC results are affected by variations within the ranges permitted by the data.

The most discriminating variable for these processes is the angular difference $\Delta\phi$ between the signal and opposite-side B direction. While flavor creation and flavor excitation mainly produce B mesons back-to-back, B mesons from gluon splitting processes are more often directed in the same direction. In the context of same-side tagging, opposite-side B daughters and fragmentation tracks are more likely to disturb the tagger for gluon splitting events than for the two other processes. By fitting $\Delta\phi$ distributions from simulation for the different processes to the $\Delta\phi$ distribution in data the following ranges for the systematic variations have been determined: gluon splitting fraction within $[-68\%, +46\%]$, flavor excitation and creation within $[-50\%, +50\%]$ relative to their nominal appearance.

Several scenarios of the fit of the production mechanisms have been considered. The fits suffer from the limited statistics available, and no precise statement about the GS fraction in data could be made. Among the considered scenarios, two extreme cases are utilized to evaluate systematic effects. The two cases corresponds to fixing the ratio between FE and FC ratios to 1:1.5 and 1:0.5 relative to the nominal values in MC simulation. With these prescriptions, the fit results are FC $= 0.75 \pm 0.06$, FE $=$ FC$\times$1.5, GS $= 1.37 \pm 0.09$ and FC $= 1.13 \pm 0.08$, FE $=$ FC$\times$0.5, GS $= 0.81 \pm 0.13$. The systematic uncertainty is estimated by calculating the $\max CLL$ SST scale factor in the scenarios labeled "GS1", with FC $= 0.75 - 0.06$, FE $=$ FC$\times$1.5, GS $= 1.37 + 0.09$, and "GS2", with FC $= 1.13 + 0.08$, FE $=$ FC$\times$0.5, GS $= 0.81 - 0.13$, in Table 6.3. The difference between the results in these two scenarios and the calculated scale factor in the default configuration provides an estimate of the systematic uncertainty associated with the uncertainty on the relative fractions of the contributions of FC, FE, and GS to the production of b quarks.

6.8.2 Fragmentation process

The Lund string fragmentation model provided by the PYTHIA generator is utilized to produce the default MC sample. The numerous systematic studies related to the fragmentation that have been performed are described in the following paragraphs.

In Section 4.3.5 it has been explained that both Peterson and Lund fragmentation functions have been utilized in the generation of the PYTHIA-MC sample used to calibrate the ANN-based SST, for strings with heavy and light quarks, respectively. It is also relevant to remind that the use of the Peterson function, whereas Reference [93] shows that a Lund function better describes B data, is dictated by the presence of a long tail in the low-z region which characterizes this fragmentation function. In the framework utilized for this analysis, the simulation of the use of different fragmentation functions is implemented by reweighing the MC events in the produced sample. The tail in the Peterson fragmentation function allows for the reweighing of the produced MC sample with weights close to unity, and thus minimizing the effect of statistical fluctuations.

The allowed parameter space for the parameters of the symmetric Lund function which has been adopted to describe the fragmentation process has been determined from a simultaneous fit to several distributions in data and MC simulation which are sensitive to the fragmentation function, such as track multiplicity, transverse momentum of the B and of the fragmentation tracks. The data and MC samples of $\mathrm{B}^0 \to J/\psi \mathrm{K}^{*0}, J/\psi \to \mu^+\mu^-$, $\mathrm{K}^{*0} \to \mathrm{K}^+\pi^-$, $\mathrm{B}^+ \to J/\psi \mathrm{K}^+, J/\psi \to \mu^+\mu^-$, and $\mathrm{B}^0_s \to J/\psi\phi^0, J/\psi \to \mu^+\mu^-$, $\phi^0 \to \mathrm{K}^+\mathrm{K}^-$ were utilized for this study, in Reference [108]. Three alternative sets of parameters for a symmetric Lund function have been chosen to evaluate systematic uncertainties. As an additional cross check, three variations according to a Peterson function have been utilized, although not included in the computation of the final systematic uncertainty. The fragmentation functions which were utilized are shown in Figure 6.13. The systematic uncertainty on the scale factor $S_{\mathcal{D}}$ associated with the choice of a particular fragmentation function is calculated utilizing the $\max CLL$ algorithm for same-side tagging. This uncertainty is indicated by the labels "Peterson" and "Lund", followed by the value of the parameters utilized for these functions, in Table 6.3.

The fragmentation process determines the formation of hadrons out of the string. It thus essentially effects the track multiplicity around the B meson, the B momentum and the momenta of the fragmentation tracks.

In order to perform systematic variations of the SST scale factor, the MC events have been reweighed according to modifications of some characteristic distributions with one entry per (tagged) event. The following distributions have been chosen:

- transverse momentum of the B candidate;
- number of tagging track candidates;
- $\Delta R(\mathrm{B}, trk$ of the selected tagging track (Equation 6.3.1);

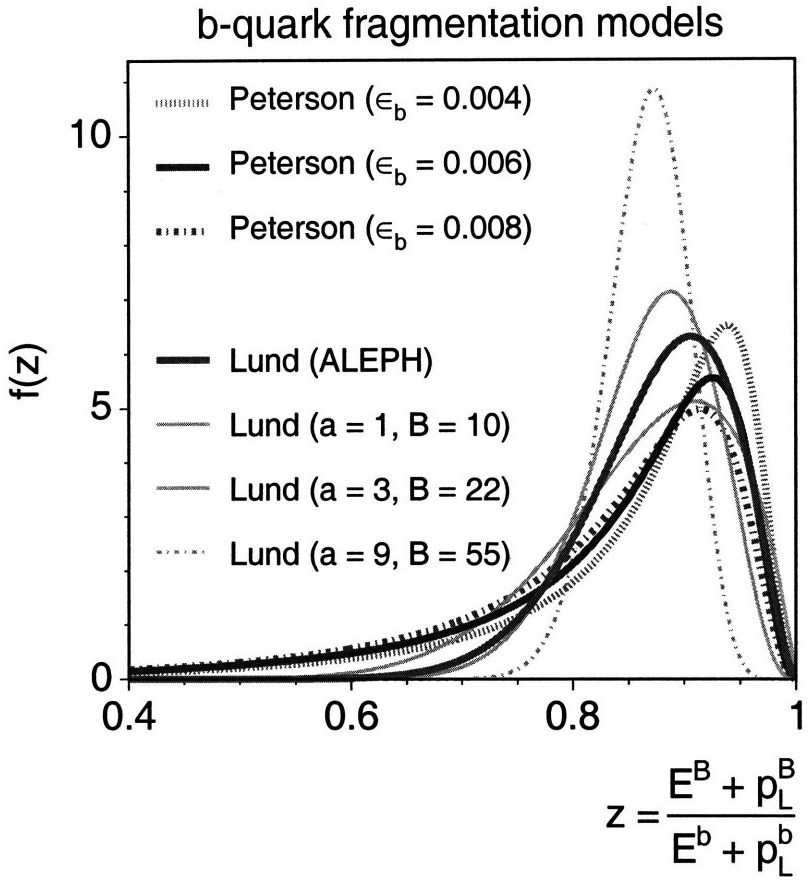

FIGURE 6.13: Fragmentation functions used to evaluate systematic uncertainties.

- CLL of the selected tagging track (Equation 6.5.3);
- p_T of the selected tagging track;
- p_T^{rel} of the selected tagging track (Figure 6.3);
- p_L^{rel} of the selected tagging track (Figure 6.3).

Two variations of the distributions for each of these variables are generated from the data as follows: the first bin of a distribution (bin 1) is modified by $+1\sigma_1$, where σ_1 is the uncertainty on the content of the first bin, and the last bin (bin N) is modified in the opposite direction, thus $-1\sigma_N$ down. The other bins (i = 2, ..., N-1) are modified by $(1-2i/N)\times\sigma_i$. The second set of distributions is obtained with the same algorithm with inverted sign, in order to produce a variation of the default distribution biased toward high mean values. The modified distributions are then normalized, and MC events are reweighed by the ratio of the MC distribution and the low (high) variation of the same distribution in data. Some examples of the modified distributions are reported in Figure 6.14.

By varying the MC sample within the ranges allowed by the statistical uncertainties on the corresponding distributions in the data, following the procedure described

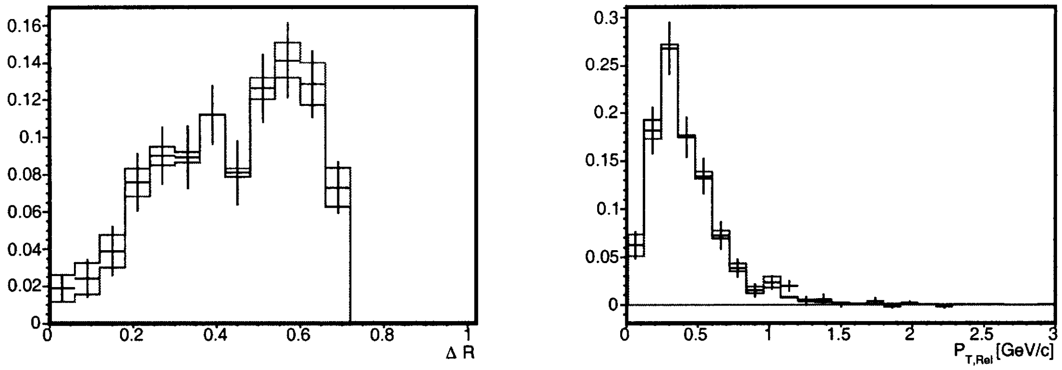

FIGURE 6.14: Reweighing distributions: $p_T^{\rm rel}$ (left) and ΔR (right). The black markers with error-bars indicate the reference distribution in data. The red/gray (blue/black) distribution is obtained following the algorithm described in the text to produce a low (high) variation of the distribution in data

in the previous paragraph, an estimate of how far the MC simulation can be maximally off is obtained. The quantities chosen for this test include some of the kinematic variables which are utilized as input to the ANN of same-side tagging. This study of systematic uncertainties has been performed using the ANN-based same-side tagging algorithm. The results are reported in Table 6.5.

6.8.3 Particle-content around the B meson

The particle species produced around the B_s^0 meson give us some insight into the fragmentation process. A measurement of the species of stable charged particles around B mesons has been performed in a high statistics sample of semileptonic B decays [99].

The fraction of kaons produced around B_s^0 mesons and selected as tagging tracks is found to be somewhat different between data and MC simulation, $20.2 \pm 1.4\%$ and $23.6 \pm 0.2\%$, respectively. The variation of $S_{\mathcal{D}}$, which is calculated using MC, that can be caused by having a different fraction of kaons around B_s^0 candidates in the MC simulation with respect to the one observed in data, is evaluated by studying the effect of reducing the kaon fraction to 19.5% in the MC sample. This new value for the kaon fraction represents a -1σ variation from the measured fraction in data. The corresponding weight applied to kaon-tagger MC candidates, $w_{\rm K}$, is calculated as follows:

$$w_{\mathrm{K}} = \frac{f_{\mathrm{K}}^{data} - \alpha \cdot \sigma_{f_{\mathrm{K}}}^{data}}{f_{\mathrm{K}}^{MC}} = \frac{20.2 - 0.475 \cdot 1.4}{23.6} = 0.828\,, \tag{6.8.1}$$

$$w_p = 1 + \frac{f_{\mathrm{K}}^{MC} - (f_{\mathrm{K}}^{data} - \alpha \cdot \sigma_{f_{\mathrm{K}}}^{data})}{f_p^{data}} = 1 + \frac{23.6 - (20.2 - 0.475 \cdot 1.4)}{4.2} = 1.960\,, \tag{6.8.2}$$

where the factor α defines the 1σ single-sided region in the kaon fraction. The weight for proton-tagged events, w_p, has been chosen to compensate for the drop in the kaon fraction, while pion-tagged events are left untouched. In the formula above, f_{K} and f_p represent the fraction of kaon-tagged and proton-tagged B_s^0 candidates, as measured in data and in MC events, while $\sigma_{f_{\mathrm{K}}}$ is the uncertainty on the measurement of the fraction of kaon-tagged events in data.

Two different scenarios have been considered, and modifications to the scale factor of the ANN-based tagging algorithm evaluated. Firstly, all events with a kaon as tagging track have been reweighed, thus bringing the total kaon fraction in MC events to match the one in data, and the performance of the tagger measured in the weighed MC sample. This approach assumes that the deficit in kaons which is indicated by the measurement in data is equally distributed among all kaons, independently if they are potentially good or bad tagging tracks.

The second scenario represents an extreme case: only candidates tagged by kaons originating from the string containing the b quark are randomly removed until the total kaon fraction is reduced to match w_{K} in Equation 6.8.2. These kaons carry, on average, more tagging power (i.e., higher dilution), and thus this scenario represents the worst possible case. Half of the deviation in $S_{\mathcal{D}}$ calculated in this scenario has been added in quadrature to the total systematic uncertainty.

In addition to studying the stable charged particles, the rate of kaons from resonances and vector particles such as ϕ^0, K_s^0 and K^{*0} has been checked. The mass distributions of the above resonances produced using the PYTHIA-MC sample are compared to the ones obtained with the high statistics $\mathrm{B}_s^0 \to \mathrm{D}_s^- \ell^+ X$ sample (Figure 6.15), in order to get an estimate of possible disagreements in the rates of kaons from such resonances between data and MC simulation. The statistical precision is not sufficient to make a precise statement. Therefore, the fraction of MC events where the actual tagging track originates from a ϕ^0, a K_s^0, or a K^{*0} has been varied by a factor of 2 lower and higher than nominal. The largest negative and positive deviations in the scale factor of the ANN-based same-side tagger obtained in these tests have been assigned as an additional systematic uncertainty.

These contribution to the total systematic uncertainty are indicated as "kaon fraction", "prompt kaon fraction", and "resonance/V0 content" in Table 6.5, respectively, and represent the largest part of the total systematic uncertainty. The contribution to the systematic uncertainty labeled with "prompt kaon fraction" in Table 6.5 is already corrected by a factor 0.5, as indicated in the description of this study.

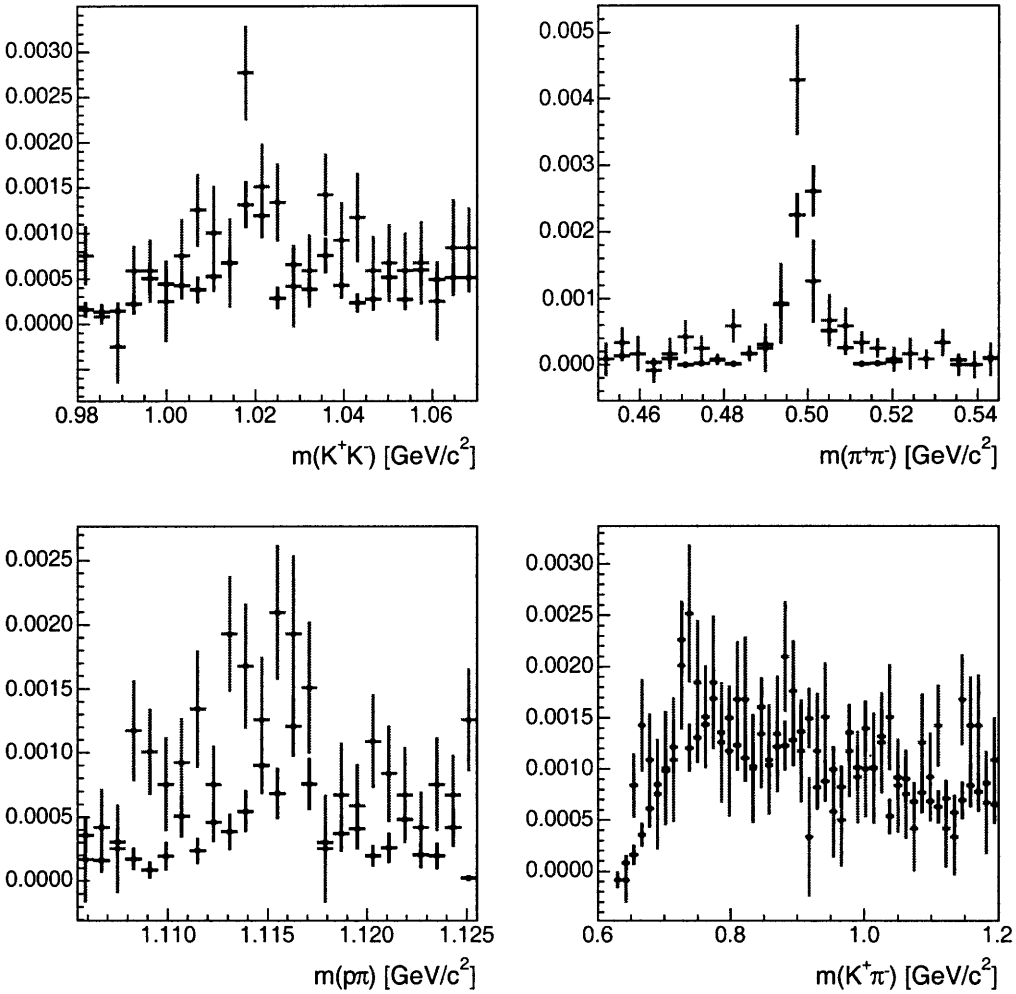

FIGURE 6.15: Reconstructed ϕ^0, Λ^0, K^0_s, and K^{*0} candidates. Black dots correspond to the $\mathrm{B}^0_s \rightarrow \mathrm{D}^-_s\pi^+$ PYTHIA-MC sample, red/gray dots to the $\mathrm{B}^0_s \rightarrow \mathrm{D}^-_s\ell^+X, \mathrm{D}^-_s \rightarrow \phi^0\pi^-$ sample in data. Plots are normalized to the number of B^0_s candidates.

6.8.4 Multiple interactions

The rate of additional potential tagging tracks coming from $p\bar{p}$ interactions other than the one which produced the B^0_s candidate, *pile-up* events, has been measured on data and accordingly added to our MC sample, as described in Section 4.3.5. Because this rate depends on luminosity, two scenarios have been studied: "high" and "low" luminosity, defined by the thresholds in Table 6.1. The numbers of events to be added to the MC sample have been measured in the different luminosity ranges. Table 6.1 contains the fractions of events to which tag candidates are added in the three periods of data taking and in the default, "high", and "low" luminosity scenarios, and the threshold in instantaneous luminosity utilized to define "high" and "low" scenarios.

The scale factor $S_{\mathcal{D}}$ of the particle-identification–based same-side tagger is then calculated in MC samples to which tagging tracks have been added, according to the "high" and "low" luminosity scenarios. The obtained values for the scale factor are

Period	$\mathcal{L}$ Threshold	Default	Low	High
`0d`	$25 \cdot 10^{30}$ cm^{-2}s^{-1}	0.22%	0.18%	0.48%
`0h`	$35 \cdot 10^{30}$ cm^{-2}s^{-1}	0.65%	0.5%	1.2%
`0i`	$35 \cdot 10^{30}$ cm^{-2}s^{-1}	0.72%	0.5%	1.1%

TABLE 6.1: Fractions of events with an additional tag candidate from pile-up events used for systematic uncertainties. The threshold separates the events which are counted in the high-luminosity and low-luminosity scenarios. "Default" indicates the fraction determined in the tuning procedure of the MC simulation.

finally compared to a MC simulation with the default rate. Discrepancies between the calculated values provide an estimate of a possible dependence of the SST scale factor on luminosity. As hard cuts on the impact parameter significance and the $\Delta z_0(\mathrm{B}, trk)$ are applied, the effect of multiple interactions on the scale factor is relatively small. The uncertainty estimated for the particle-identification–based algorithm has been utilized as an estimate of the uncertainty on the ANN-based algorithm for same-side tagging too.

6.8.5 Simulation of particle identification

Both t_{flight} and $d\mathrm{E}/dx$ measurements play a major role in this analysis. Therefore a fine tuning of the simulation was needed to ensure that the MC samples reproduce the data well, which was described in Section 4.3.4.

The evaluation of the systematic uncertainty related to particle identification has been performed by modifying the distributions which are utilized to simulate particle identification in MC events. Three sources of systematic uncertainty have been identified.

The first source of systematic uncertainty comes from the particular choice of a parameterization of the TOF resolution function. The systematic uncertainty associated with preferring one parameterization function over another one is investigated by utilizing a second parameterization. Two different TOF resolution parameterizations were developed. The parameterization which has been selected as default method is described in Reference [99], while the other one is described in Reference [109]. The main difference between the two methods is that the principal motivation for the technique developed in the latter reference is separating pions and kaons on a track-by-track basis, and the parameterization for the TOF resolution which is derived assumes p_T-independence of the TOF resolution function. The parameterization which is here utilized as default method, instead, was developed for the study of the species of charged particles produced in association with B mesons in ranges of track p_T. This required to focus on obtaining the correct statistical separation between particle species, which is attainable by simulating the tails in the TOF resolution function correctly. Moreover, a p_T-dependent resolution function allowed for the removal of systematic effects that would appear if all the p_T ranges investigated were treated uniformly. As systematic study, the first method is utilized to reconstruct the identity of particles, but the TOF response is simulated using the second method.

A second source of systematic uncertainty, which is still related with TOF, is due to the uncertainty on the efficiency function *versus* track p_T that describes the ratio between the data and the simulation. The efficiency ratio has been varied by $\pm 8\%$ for MC events which simulate `0d` data, and $\pm 10\%$ for MC events simulating `0h` and `0i` data, as shown in Figure 4.3.

The systematic uncertainties on the $d\mathrm{E}/dx$ measurement is estimated following the suggestions given by Reference [110]. While using the default distributions of Z, which is defined in Equation 4.3.8, to assess the identity of a particle, the distributions utilized in the simulation have been changed by varying their mean within the range $[-0.007, +0.002]$ (`0d` MC events), or $[-0.002, +0.004]$ (`0h` and `0i` MC events), and increasing their width by 3%. Finally, the number of COT hits with $d\mathrm{E}/dx$ information has been varied, in the simulation, by ± 5.

Three different scenarios have been considered for `0d` MC events, characterized by the variations of TOF efficiency, parameterizations of TOF resolution, number of COT hits with $d\mathrm{E}/dx$ information and distribution of the Z variable described in the previous paragraphs. The first two scenarios are characterized by a worsening of the performance of particle-identification, while the third scenario corresponds to an optimistic case. The prescriptions for the three scenarios are summarized below:

A different TOF p.d.f. used for simulation and reconstruction, -8% TOF efficiency, -5 COT $d\mathrm{E}/dx$ hits, $\sigma_Z \times 1.03$ and $\langle Z \rangle = -0.007$;

B different TOF p.d.f. used for simulation and reconstruction, -8% TOF efficiency, -5 COT $d\mathrm{E}/dx$ hits, $\sigma_Z \times 1.03$ and $\langle Z \rangle = +0.002$;

C $+8\%$ TOF efficiency, $+5$ COT $d\mathrm{E}/dx$ hits.

The differences between the scale factor $S_{\mathcal{D}}$ of the particle-identification–based tagger calculated in the MC sample with the default simulation of particle identification and the ones calculated with each of the modified scenarios for the particle-identification simulation are reported in Table 6.3. The differences corresponding to the three scenarios described in the list above are labeled t_{flight}+$d\mathrm{E}/dx$ A, B, and C, respectively.

One additional effect has to be taken into account for the TOF simulation. It is calculated in the MC simulation that 1.5% of the tracks in the TOF do not originate from the primary vertex, and thus their t_0 (i.e., the production time) is not known. One systematic study is to simulate them as if they were kaons from the primary interaction (OBSP K). The other systematic study is to simulate them as if they were pions from the primary interaction (OBSP π). The difference between the nominal $S_{\mathcal{D}}$ and the ones calculated with the two above assumptions are indicated by the labels OBSP K and OBSP π in Table 6.3.

In the case of `0h` and `0i` MC events, it has instead been suggested to separate the variations in TOF and COT performance. Fourteen different configurations have thus been defined:

- optimistic particle-identification: TOF resolution scale factor 1.05 (default 1.15)), $+10\%$ TOF efficiency, $+5$ COT $d\mathrm{E}/dx$ hits, $+1\sigma$ $d\mathrm{E}/dx$ efficiency;

- conservative particle-identification: TOF resolution scale factor 1.25 (default 1.15)), −10% TOF efficiency, −5 COT dE$/d$x hits, -1σ dE$/d$x efficiency;
- $\langle Z \rangle = -0.002$;
- $\langle Z \rangle = +0.004$;
- $\sigma_Z \times 1.03$ *or* 0.97 for pions with $p_T \leq 1$ GeV/c;
- $\sigma_Z \times 1.03$ *or* 0.97 for kaons with $p_T \leq 1$ GeV/c;
- $\sigma_Z \times 1.03$ *or* 0.97 for protons with $p_T \leq 1$ GeV/c;
- shift in TOF offset of ±15 ps for kaons only;
- shift in TOF offset of ±15 ps for protons only.

The last four configurations, which introduce a shift in the TOF offset, are meant to cover the additional effect described in the previous paragraph. The systematic shifts of the scale factor of the particle-identification–based tagging algorithm are reported in Table 6.4.

6.8.6 B^0 and B^+ Data–MC agreement

As explained in Section 6.7, it has been checked that the scale factors $S_\mathcal{D}$ of the particle-identification–based same-side tagger, described in Section 6.5, obtained from PYTHIA-MC and data samples of B^0 and B^+ mesons agree within their statistical and systematic uncertainties. However, this statement cannot be made more precise than the uncertainties on the data and MC sample. Therefore, the weighed mean of the uncertainties from the B^0 and B^+ control modes has been utilized as an estimate of the systematic uncertainty associated with a residual data–MC-simulation disagreement for the SST scale factor applied to B_s^0 decay modes. The following equation describes the evaluation of this component of the systematic uncertainty of $S_\mathcal{D}$:

$$\begin{aligned} \delta S_\mathcal{D} &= \frac{\sum_i (S_{data,i} - S_{MC,i}) \cdot w_i}{\sum_i w_i}, \\ w_i &= \frac{1}{\sigma^2_{S_{data,i}} + \sigma^2_{S_{MC,i}}}, \\ \sigma_{\delta S_\mathcal{D}} &= \frac{1}{\sqrt{\sum_i w_i}}, \end{aligned} \tag{6.8.3}$$

where S is the SST scale factor and the index i runs on the four B^0 and B^+ reconstructed modes, which are listed in Section 6.7. The weighed mean of the uncertainties $\delta S_\mathcal{D}$ and the "effective" variance $\sigma_{\delta S_\mathcal{D}}$ are utilized as estimates of the uncertainty associated with residual data–MC-simulation disagreement. In the case of data, $\sigma_{S_{data,i}}$ represents the pure statistical uncertainty. In the case of MC simulation, $\sigma_{S_{MC,i}}$ is the sum in quadrature of statistical and the other systematic uncertainties, where

algorithm [%]	δ	σ_δ	δ/σ_δ
$\mathcal{D}$, max p_L^{rel}	-2.1	1.4	1.5
$\mathcal{D}$, max CLL	-0.7	1.6	0.4
$S_\mathcal{D}$, max p_L^{rel}	-4.8	5.5	0.9
$S_\mathcal{D}$, max CLL	-7.6	4.7	1.6

TABLE 6.2: Systematic uncertainty from B^+ and B^0 data–MC simulation agreement. In the table, δ represents either $\delta\mathcal{D}$ or $\delta S_\mathcal{D}$. This study addresses the possibility that disagreements between the SST scale factors measured in data and calculated in the MC samples are covered by their statistical and systematic uncertainties. The significances of the discrepancies, indicated by δ/σ_δ, are consistent with the variations δ being statistical fluctuations.

positive systematic uncertainty is taken when the scale factor in data is higher than the one from MC simulation, while the negative uncertainty is taken when the scale factor in data is lower.

The formulae in Equation 6.8.3 have been used to estimate the residual data–MC-simulation disagreement for the scale factor and average dilution of the max CLL SST algorithm. The average dilution of a same-side tagging algorithm is calculated in a MC sample and measured in B^+ and B^0 data utilizing Equation 6.1.1. As a cross-check, the same calculation has been performed for the scale factor and average dilution of the the max p_L^{rel} algorithm presented in Section 6.4. The results of the comparison are presented in Table 6.2. The significance of the deviations, $\delta S_\mathcal{D}(\mathcal{D})/\sigma_{\delta S_\mathcal{D}(\mathcal{D})}$, are between 0.5 and 1.5, completely consistent with a statistical fluctuation. Therefore, the maximum value between $\delta S_\mathcal{D}$ and $\sigma_{\delta S_\mathcal{D}}$ for the particle-identification–based tagging algorithm is chosen as estimate of the uncertainty and added to both the negative and positive total systematic uncertainty of the ANN-based tagging algorithm.

The component of the total systematic uncertainty evaluated in this section, which is a rather important one, conservatively estimates the possibility that large statistical and systematic uncertainties in the SST scale factors, as measured in data and calculated in the MC samples, cover residual disagreements between the scale factors measured in data and calculated using MC simulated events. As such, this systematic uncertainty is expected to decrease when larger MC samples will be available.

6.8.7 Total systematic uncertainty

The complete list of the analyzed sources of systematic uncertainties which affect the scale factor of the ANN-based same-side tagging algorithm applied to B_s^0 samples is detailed in Tables 6.3, 6.4, and 6.5. The first two tables contain systematic uncertainties which were calculated adopting the particle-identification–based same-side tagging algorithm. These uncertainties have been transferred, unmodified, to the ANN-based algorithm.

The total systematic uncertainty is indicated by $\pm\Sigma$ in Table 6.6. It is calculated by picking the largest positive and the largest negative deviations for each effect, and adding them in quadrature separately for positive and negative deviations. The

Source	$\sigma_{S_\mathcal{D}}$ [%]
GS1	-1.5
GS2	$+0.6$
Peterson (0.004)	-3.1
Peterson (0.006)	$+1.0$
Peterson (0.008)	-0.5
Lund (1,10)	$+0.5$
Lund (3,22)	$+1.8$
Lund (9,55)	$+4.4$
- pile-up	$+0.2$
+ pile-up	-0.2
t_{flight}+dE/dx A	-4.0
t_{flight}+dE/dx B	-3.0
t_{flight}+dE/dx C	$+3.6$
OBSP K	-2.6
OBSP π	-0.3

TABLE 6.3: Systematic uncertainties $\sigma_{S_\mathcal{D}}$ on the scale factor of the maxCLL algorithm for same-side tagging. These estimates are utilized, unmodified, as uncertainties on the scale factor of the ANN-based algorithm. The systematic uncertainties associated with pile-up events and with the TOF and COT dE/dx simulation refer to MC events which simulate `0d` data only. The details of the evaluation of these systematic uncertainties are reported in Sections 6.8.1, 6.8.2, 6.8.4, and 6.8.5. All the uncertainties are referred to the scale factor calculated in a PYTHIA-MC sample of $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$ candidates.

unique exception is the systematic uncertainty related with fragmentation functions, where the variations of the Peterson function, which were required as a cross-check, are reported for reference only.

The largest component of systematic uncertainty is the one which estimates the possibility that a residual disagreement between scale factors measured in B^0 and B^+ data and calculated in MC simulation is covered by their uncertainties. As mentioned before, this component is expected to decrease if larger MC samples are utilized. Besides that, the dominant source of systematic uncertainties is the kaon fraction around B mesons, followed by the variation of the tagging related distributions within the statistical uncertainties from data. All other systematic uncertainties are small compared to those two sources. Adding each variation separately from the other systematic uncertainties is a very conservative approach.

6.9 Transfer Between Samples

The only difference between the performance of a SST on different B_s^0 decay modes is their different momentum spectra due to different trigger and reconstruction cuts. The scale factor $S_\mathcal{D}$ of the ANN-based same-side tagger has been calculated using a

Source	$\sigma_{S_\mathcal{D}}$ [%]
- pile-up	$+0.1$
+ pile-up	-0.6
optimistic particle-identification	$+3.9$
conservative particle-identification	-4.1
$\langle Z \rangle = -0.002$	< 0.1
$\langle Z \rangle = +0.004$	< 0.1
$\sigma_Z \times 1.03$, kaons with $p_T \leq 1$ GeV/c	< 0.1
$\sigma_Z \times 0.97$, kaons with $p_T \leq 1$ GeV/c	-0.2
$\sigma_Z \times 1.03$, pions with $p_T \leq 1$ GeV/c	-0.6
$\sigma_Z \times 0.97$, pions with $p_T \leq 1$ GeV/c	$+0.5$
$\sigma_Z \times 1.03$, protons with $p_T \leq 1$ GeV/c	< 0.1
$\sigma_Z \times 0.97$, protons with $p_T \leq 1$ GeV/c	< 0.1
TOF response +15 ps, protons	$+0.1$
TOF response -15 ps, protons	< 0.1
TOF response +15 ps, kaons	< 0.1
TOF response -15 ps, kaons	-0.3

TABLE 6.4: Systematic uncertainties $\sigma_{S_\mathcal{D}}$ in the `0h` and `0i` data samples on the scale factor of the maxCLL algorithm for same-side tagging. These estimates are utilized, unmodified, as uncertainties on the scale factor of the ANN-based algorithm. The total uncertainty associated with particle identification is $^{+3.9}_{-4.2}$, while the particle-identification–related uncertainty in the `0d` data sample, from Table 6.3, is $^{+3.6}_{-4.0}$. The details of the evaluation of these systematic uncertainties are reported in Sections 6.8.4 and 6.8.5. All the uncertainties are referred to the scale factor calculated in a PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates.

MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates. It may thus be necessary to apply a correction to this scale factor when the tagger is applied to a different B_s^0 decay mode. In order to estimate the correction factor to utilize when analyzing the additional modes that will be used in this analysis, which are listed in Section 1.5, the $p_T(B)$ distribution of the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ MC sample utilized to calculate the default value of $S_\mathcal{D}$ has been reweighed to match the distribution observed in data from the additional modes. The calculated correction factors are reported in Table 6.7.

6.10 Final scale factors

The performance of a tagging algorithm is summarized by its efficiency and dilution. Section 6.7 described the calibration of candidate-by-candidate dilution of same-side tagging algorithms, which requires the use of a MC sample when the tagger is applied to B_s^0 candidates. The calibration consists in the calculation of a scale factor for the dilution provided by the algorithm for flavor tagging. The systematic uncertainties which affect the value of this scale factor are analyzed in Section 6.8. This final

Source	$\sigma_{S_\mathcal{D}}$ [%]
var. $-N$	$+0.3$
var. $+N$	-0.7
var. $-p_T(\mathrm{B})$	-0.3
var. $+p_T(\mathrm{B})$	$+0.4$
var. $-p_L^{\mathrm{rel}}$	-0.3
var. $+p_L^{\mathrm{rel}}$	$+0.1$
var. $-CLL$	$+1.6$
var. $+CLL$	-0.3
var. $-p_T$	$+0.3$
var. $+p_T$	-0.1
var. $-p_T^{\mathrm{rel}}$	$+0.6$
var. $+p_T^{\mathrm{rel}}$	$+0.2$
var. $-\Delta R$	-0.8
var. $+\Delta R$	$+0.2$
kaon fraction	-6.6
prompt kaon fraction	-5.3
resonance/V0 content	$^{+4.3}_{-3.8}$

TABLE 6.5: Systematic uncertainties $\sigma_{S_\mathcal{D}}$ on the scale factor of the ANN-based algorithm for same-side tagging. The details of the evaluation of these systematic uncertainties are reported in Sections 6.8.2 and 6.8.3. All the uncertainties are referred to the scale factor calculated in a PYTHIA-MC sample of $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \to \phi^0 \pi^-$ candidates.

Data sample		$\sigma_{S_\mathcal{D}}$ [%]
`0d`	$-\Sigma$	10.7
	$+\Sigma$	14.3
`0h`	$-\Sigma$	10.8
	$+\Sigma$	14.4
`0i`	$-\Sigma$	10.8
	$+\Sigma$	14.4

TABLE 6.6: Total systematic uncertainty $\sigma_{S_\mathcal{D}}$ on the scale factor of the ANN-based algorithm for same-side tagging. This uncertainty is referred to the scale factor calculated in a PYTHIA-MC sample of $\mathrm{B}_s^0 \to \mathrm{D}_s^- \pi^+$, $\mathrm{D}_s^- \to \phi^0 \pi^-$ candidates.

Decay mode	Correction [%]
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-$	0
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^{*0} K^-$	+1.3
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^- \pi^+ \pi^-$	+1.1
$B_s^0 \to D_s^- \pi^+ \pi^- \pi^+,\ D_s^- \to K^{*0} K^-$	+6.1
$B_s^0 \to D_s^- \pi^+ \pi^- \pi^+,\ D_s^- \to \phi^0 \pi^-$	+5.4
$B_s^0 \to D_s^- \pi^+ \pi^- \pi^+,\ D_s^- \to \pi^- \pi^+ \pi^-$	+0.0
$B_s^0 \to D_s^- \ell^+ X$	+6.3

TABLE 6.7: $p_T(\mathrm{B})$ transfer corrections for different B_s^0 decay modes.

section presents the value of the scale factors which are utilized in the analysis of B_s^0 oscillations presented in this dissertation.

The scale factors for the ANN-based SST algorithm calculated in a PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates, for the three different periods of data-taking, with their total uncertainties, are reported here:

$$\begin{aligned} S_{\mathcal{D}}(\texttt{0d};\ B_s^0 \to D_s^- \pi^+) &= 99.2\ ^{+10.7}_{-14.3}\%\,, \\ S_{\mathcal{D}}(\texttt{0h};\ B_s^0 \to D_s^- \pi^+) &= 95.9\ ^{+10.8}_{-14.4}\%\,, \\ S_{\mathcal{D}}(\texttt{0i};\ B_s^0 \to D_s^- \pi^+) &= 95.0\ ^{+10.8}_{-14.4}\%\,. \end{aligned} \qquad (6.10.1)$$

These scale factors are obtained as a result of the fit described in Section 6.7. The knowledge of the true flavor of the B_s^0 candidates in the MC sample allows one to fit for $S_{\mathcal{D}}$. These scale factors are used as an input to the fit for $B_s^0 - \overline{B}_s^0$ oscillations presented in the last chapters of this document. The small difference in the uncertainty for the 0h and 0i data samples, with respect to the 0d uncertainty, is due to the different contribution to the total systematic uncertainty of particle identification ($^{+3.9}_{-4.2}$ *vs.* $^{+3.6}_{-4.0}$) and pile-up events ($^{+0.2}_{-0.2}$ *vs.* $^{+0.1}_{-0.6}$).

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

$$S_{\mathcal{D}} \sqrt{\langle \mathcal{D}^2 \rangle} = 30.2 \pm 0.7\%\,. \qquad (6.10.2)$$

The quoted uncertainty is statistical only. The figure-of-merit for a tagger is $\epsilon \mathcal{D}^2$, as it will be clearly shown in Section 5.1. The figure-of-merit of the ANN-based SST algorithm presented in Section 6.6 is reported in Tables 6.8. Efficiency is evaluated on data only, while the the scale factor $S_{\mathcal{D}}$ is calculated in simulated events. The measured performance of the ANN-based SST algorithm is presented separately in the three periods of data-taking 0d, 0h, and 0i.

The MC sample utilized to train the ANN which performs flavor-tagging is tuned to simulate the 0d period of data-taking. The same improvement over the particle-identification–based SST algorithm (Section 6.5) is observed in the 0i period of data-taking, while the *reduced* performance in 0h data is expected to be due to a statistical fluctuation.

This chapter presented, in Section 6.6, the same-side algorithm utilized in this anal-

[%]	maxCLL algorithm	ANN-based algorithm
`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 6.8: Performance of particle-identification and ANN Same-Side Taggers. Efficiency ϵ and average dilution $\sqrt{\langle \mathcal{D}^2 \rangle}$ are measured in the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ data sample utilized in this mixing analysis. The scale factor $S_{\mathcal{D}}$ is calculated by applying the SST algorithm in a PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ candidates. The numbers in the table are correlated because the two algorithms are applied to the same data samples. A second source of correlation is the use of the same PYTHIA-MC sample utilized to calculate the scale factors for the two algorithms. Thirdly, part of the systematic uncertainties estimated for the scale factor of the particle-identification–based algorithm have been assigned, unchanged, to the scale factor of the ANN-based algorithm.

ysis of $B_s^0 - \overline{B}_s^0$ *oscillations. The algorithm provides a candidate-by-candidate weight for the correctness of its decision. The calibration of the tagging algorithm, a crucial aspect of mixing analyses which may be in the situation of setting a lower limit for* Δm_s*, instead of making a measurement, consists in calculating a scale factor for the weights returned by the tagger. The calculation of the SST scale factor to be used in this analysis is presented, and the evaluation of systematic uncertainties reported. The calculated scale factor is utilized as an input to the analysis for* $B_s^0 - \overline{B}_s^0$ *oscillations. The next chapter introduces the maximum likelihood fitter which combines mass, proper decay-time, and tagging information. The final result of this analysis, the observation of* $B_s^0 - \overline{B}_s^0$ *oscillations and a precise measurement of* Δm_s*, is presented, with its implications on the Standard Model, and constraints on parameters which describe new physics beyond the SM.*

Chapter 7

Analysis framework and observation of $B_s^0 - \overline{B}_s^0$ oscillations

In this chapter, the fit framework utilized for amplitude scans and the measurement of Δm_s are described in this chapter. The study of systematic uncertainties on the measured value of Δm_s and the impact of its measurement on the Standard Model picture of flavor interactions are presented, too.

7.1 Maximum Likelihood framework

A fitting framework based on the unbinned maximum likelihood estimation method[1] has been developed for this analysis, and used to extract the parameters of interest from data. It allows for the straightforward combination of various pieces of information coming from different decay modes and their simultaneous fit, maximizing the statistical power of the samples. In its most generic formulation, the contribution to the global likelihood of a candidate i is written as follows:

$$\mathcal{L}_i = \sum_j f_j \, \mathcal{P}_i^j(m_i, ct_i, \sigma_{ct_i}, \mathcal{D}_i, T_i)\,, \tag{7.1.1}$$

where the index j indicates the various signal and background components that are present in the samples, f_j is the fractions of the j-th component and $\mathcal{P}_i^j$ is the probability that a candidate has mass m_i, decay time ct_i, decay time uncertainty σ_{ct_i} and predicted dilution $\mathcal{D}_i$, under the assumption that it belongs to the j-th component. By construction, $\sum_j f_j = 1$. The global likelihood $\mathcal{L}$ is naturally defined as the multiplication of all the single-candidate likelihoods $\mathcal{L}_i$:

$$\mathcal{L} = \prod_i \mathcal{L}_i\,. \tag{7.1.2}$$

---

[1] The maximum likelihood estimation method is described in Ref. [111], among others

The minimum of $-\log \mathcal{L}$ indicates the best fit values for the parameters of the likelihood. For each contribution j, $\mathcal{P}_i^j$ is factorized as:

$$\mathcal{P}_i^j(m_i, ct_i, \sigma_{ct_i}, \mathcal{D}_i, T_i) = \mathcal{P}_m^j(m_i)\mathcal{P}_{ct}^j(ct_i|\sigma_{ct_i}, \mathcal{D}_i, T_i)\mathcal{P}_{\sigma_{ct}}^j(\sigma_{ct_i})\mathcal{P}_{\mathcal{D}}^j(\mathcal{D}_i)\,. \tag{7.1.3}$$

For background components of the data sample, empirical descriptions of the factors in Equation 7.1.3 are sufficient, while a physics model is used to describe signal contributions. Each component will be described in detail in the next paragraphs.

The mass component is simply the probability density function for the mass of each candidate. It is completely separated from the rest of the likelihood and depends exclusively on the reconstructed mass of the B_s^0 candidate. For the semileptonic likelihood, $\mathcal{P}_m^j(m_i)$ contains terms for both the D_s^- candidate mass and the $\ell\mathrm{D}_s^-$ mass distribution. In the hadronic samples, fully reconstructed $\mathrm{B}_s^0 \to \mathrm{D}_s^-\pi^+(\pi^-\pi^+)$ signal candidates are modeled with a double Gaussian peak centered at the B_s^0 mass. Combinatorial background, generally due to the pairing of a real D_s^- meson to random tracks from the underlying event, is modeled as the sum of decaying exponential and flat linear components. Templates derived from BGenerator-MC events describe other background components, such as B^0 or $\overline{\Lambda}_b^0$ contributions. In the case of semileptonic decays, signal candidates are fit with a Gaussian peak and combinatorial background with a linear function in the D_s^- mass distribution, while the shape of the other background components is extracted from a study of simulated events. The $\ell\mathrm{D}_s^-$ mass distributions of all signal and background components are obtained from the study of BGenerator-MC events.

Temporarily disregarding the flavor tagging part, the proper-decay-time component depends on ct and σ_{ct}. When a B_s^0 candidate is completely reconstructed, it is possible to write:

$$\mathcal{P}_{ct}(ct, \sigma_{ct}; \tau) = \left[\frac{1}{c\tau}e^{-\frac{ct'}{c\tau}} \otimes \mathcal{G}(ct - ct', \sigma_{ct})\right] \cdot \epsilon(ct)\,. \tag{7.1.4}$$

The exponentially decaying function, which describes the probability that a particle with lifetime τ decays after t from the production time, is convoluted with a Gaussian resolution function to account for the uncertainty on the measured proper decay-time. The calibration of the proper-decay-time resolution is discussed in Section 5.2. The last term is an efficiency function which corrects for the bias in the proper-decay-time distribution introduced by the displaced track trigger and the candidate selection, as shown in Reference [106]. It depends only on the kinematics of the decay under study and is defined as follows:

$$\epsilon(ct) = \frac{ct \textit{ after reconstruction and final selection}}{\sum_{i=1}^{N} \frac{1}{\tau}e^{-t'/\tau} \otimes G(t - t'; \sigma_i)}\,. \tag{7.1.5}$$

The distribution in the numerator is obtained from all the candidates in a signal-only MC sample which pass all the analysis selection cuts. For each accepted event i, the expected ct distribution without any bias is an exponential smeared by a Gaussian resolution function, where the width is the ct error (σ_{ct_i}) of that event. The denominator

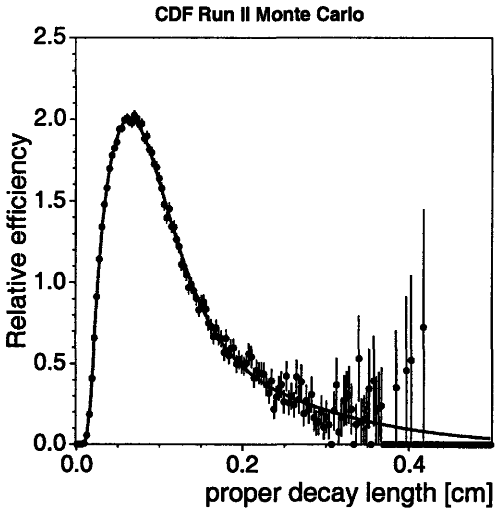

FIGURE 7.1: A representative example of the dependence of trigger and selection efficiency on proper decay time. This curve is for $B^0 \to D^-\pi^+$, with $D^- \to K^+\pi^-\pi^-$. The vertical scale is in arbitrary units.

is the sum of the N expected distributions without any bias of the same events which entered the distribution in the numerator. A different efficiency function is prepared for each B^0_s decay mode by utilizing a BGENERATOR-MC sample which reproduces the B^0_s decay chain of interest, as described in Section 4.2. The parameter τ represents the world average of B^0_s lifetime measurements [62]. The proper-decay-length efficiency curve is parameterized by the following template:

$$\epsilon(ct) = \sum_{j=1}^{3} \alpha_j \,(ct - \beta_j)^2 \, e^{-\frac{ct}{\gamma_j}} \, \theta(ct - \beta_j)\,, \tag{7.1.6}$$

where the parameters α_i, β_i, and γ_i are obtained from the fit of the expression in Equation 7.1.5. This functional form for the efficiency function allows for the analytical normalization of the proper-decay-time signal probability density function. Figure 7.1 shows a representative example of the efficiency dependence. The rapid turn-on of the efficiency is due to minimum impact parameter and L_{xy} significance requirements, while the turn-off at larger proper decay length is due to an upper cut on impact parameter of 1 mm in the triggers. The efficiency function is defined as a function of the reconstructed proper decay-lengths because it tries to correct for the effect of selection criteria which are themselves applied to reconstructed observables.

The case of incompletely reconstructed B^0_s candidates is slightly more complicated and involves the introduction of the k-factor distribution $F(k)$ (Section 5.2):

$$\mathcal{P}_{ct}(ct^*, \sigma_{ct^*}; \tau) = \int dk \left[\frac{k}{c\tau} e^{-\frac{kct^{*\prime}}{c\tau}} \otimes \mathcal{G}(ct^* - ct^{*\prime}, \sigma^*_{ct}) \right] \cdot \epsilon(ct^*) \cdot F(k)\,, \tag{7.1.7}$$

where ct^* is the *pseudo*-proper decay-time, calculated as indicated in Equation 5.2.2.

The definition of the efficiency function is analogously modified:

$$\epsilon(ct) = \frac{ct \textit{ after reconstruction and final selection}}{\int dk \sum_i \frac{k}{\tau} e^{-kt'/\tau} \otimes \mathcal{G}(t - t'; \sigma_i) \cdot F(k)} \,. \tag{7.1.8}$$

The integration over the k-factor distribution, $F(k)$, accounts for the missing momentum in partially reconstructed decays. For fully reconstructed decays, which have no missing momentum, this is not necessary as $F(k) = \delta(1)$.

The component of the likelihood which contains flavor tagging information is closely tied to the proper-decay-time component, because they share some pieces of information. In the case of signal candidates, the two components are actually indivisible, as indicated by Equation 5.1.2. Two independent tagging algorithms are available to the analysis. Each event can be tagged by neither, one or both algorithms, thus distinguishing three classes. The flavor tagging and proper-decay-time likelihood factors for tagger m, which has a dilution $\mathcal{D}^m$, efficiency ϵ_m, and tagging decision T_m, are combined as follows:

- **untagged:**

$$\int dk \left(1 - \sum_m \epsilon_m\right) e^{-k\,ct'/c\tau} \otimes \; \mathcal{G}(k\,ct' - k\,ct; \sigma_{ct}) \cdot F(k) \cdot \epsilon(ct)\,, \tag{7.1.9}$$

- **single tag:**

$$\int dk \frac{\epsilon_m}{2} \left[1 + \mathcal{A} T^m \mathcal{D}^m \, \cos(\Delta m_s \, k\,ct')\right] e^{-k\,ct'/c\tau} \otimes \; \mathcal{G}(k\,ct' - k\,ct; \sigma_{ct}) \cdot F(k) \cdot \epsilon(ct)\,, \tag{7.1.10}$$

  where $T^m = \pm 1$ is the sign of a single (same-side or opposite-side) tag,

- **double tag**:

$$\int dk \frac{\epsilon_m \epsilon_n}{2} \left[\frac{(1 + T^m T^n \mathcal{D}^m \mathcal{D}^n) + \mathcal{A}(T^m \mathcal{D}^m + T^n \mathcal{D}^n) \cos(\Delta m_s \, k\,ct')}{2}\right] e^{-k\,ct'/c\tau} \otimes \; \mathcal{G}(k\,ct' - k\,ct; \sigma_{ct}) \cdot F(k) \cdot \epsilon(ct)\,, \tag{7.1.11}$$

  where $T^m = \pm 1$ and $T^n = \pm 1$ are the signs of the two tags.

The combination of decisions and dilutions in the double tagged case explicitly accounts for whether the two tags agree or disagree [98]. For each candidate i and tagger m, the calibrated dilution $\mathcal{D}^i_m$ which enters the above equations is obtained by multiplying the candidate-by-candidate predicted dilution returned by the tagging algorithm m by the global scale factor $S^m_{\mathcal{D}}$ calculated in the calibration of the tagger, as explained in Section 6.1.

In the case of backgrounds which are treated as non-mixing, a potential global tagging asymmetry is allowed. This is the case of combinatorial background, partially reconstructed B^0_s decays (excluding the signal modes $\mathrm{D}^-_s \rho^+$ and $\mathrm{D}^{*-}_s \pi^+$, with $\mathrm{D}^-_s \to$

$\phi^0\pi^-$), mis-reconstructed $\overline{\Lambda}_b^0$ decays, and mis-reconstructed B^0 decays, which do mix in a much longer time-scale than B_s^0 mesons. The effect of B^0 mixing is included by scaling the dilution of the candidates in this background component by $1-2\chi_d$, where $\chi_d = 0.186 \pm 0.004$ [62] is the measured time-integrated probability for a B^0 meson to oscillate into a $\overline{B}^0$ meson, effectively integrating B^0 oscillations. The flavor tagging term is simpler than in the previous case and separates completely from the proper-decay-time component. For the sake of consistency with the description adopted for signal, the combination of the two factors is implemented:

- **untagged:**

$$\left(1 - \sum_m \epsilon_m\right) \mathcal{P}_{ct}(ct)\,, \tag{7.1.12}$$

- **single tag:**

$$\frac{\epsilon_m}{2}\left[1 + \mathcal{T}^m \mathcal{D}^m\right] \mathcal{P}_{ct}(ct)\,, \tag{7.1.13}$$

- **double tag:**

$$\frac{\epsilon_m \epsilon_n}{2}\left[\frac{(1+\mathcal{T}^m\mathcal{D}^m)(1+\mathcal{T}^n\mathcal{D}^n)}{2}\right] \mathcal{P}_{ct}(ct)\,, \tag{7.1.14}$$

where $\mathcal{P}_{ct}$ represents the proper-decay-time component, which, in the case of backgrounds, is a template derived from simulated events, or, for combinatorial background, candidates in the sidebands of the B_s^0 (D_s^-) mass distributions of hadronic (semileptonic) B_s^0 decays.

The last two components of the single-event likelihood in Equation 7.1.3, $\mathcal{P}_{\sigma_{ct}}(\sigma_{ct})$ and $\mathcal{P}_{\mathcal{D}}(\mathcal{D})$, are simply the probability distribution functions of proper-decay-time resolution and candidate-by-candidate dilution. It is necessary to explicitly include these terms when the distributions of σ_{ct} and $\mathcal{D}$ in signal and background are different, because their exclusion could lead to biases for the fitted values of the likelihood parameters [112]. The distributions of σ_{ct} and $\mathcal{D}$ in B_s^0 candidates reconstructed in data are utilized. The templates for signal components are obtained with the sideband-subtraction technique defined in Section 4.4, while candidates in the mass sidebands are used to produce the templates for the combinatorial background component. The σ_{ct} and same-side-tagger $\mathcal{D}$ templates for signal B_s^0 candidates and combinatorial background in the $B_s^0 \to D_s^-\pi^+$, $D_s^- \to \phi^0\pi^-$ decay mode are presented in Figure 7.2.

7.2 Systematic uncertainties

Systematic uncertainties differently affect the significance of an oscillation signal, obtained via the amplitude scan, and a measurement of the oscillation frequency Δm_s. The effect of uncertainties in the amplitude is always a reduction of the sensitivity, and, in case it is not possible to measure Δm_s, likewise on the limit on the oscillation frequency. Systematic uncertainties on Δm_s are evaluated when a measurement is performed.

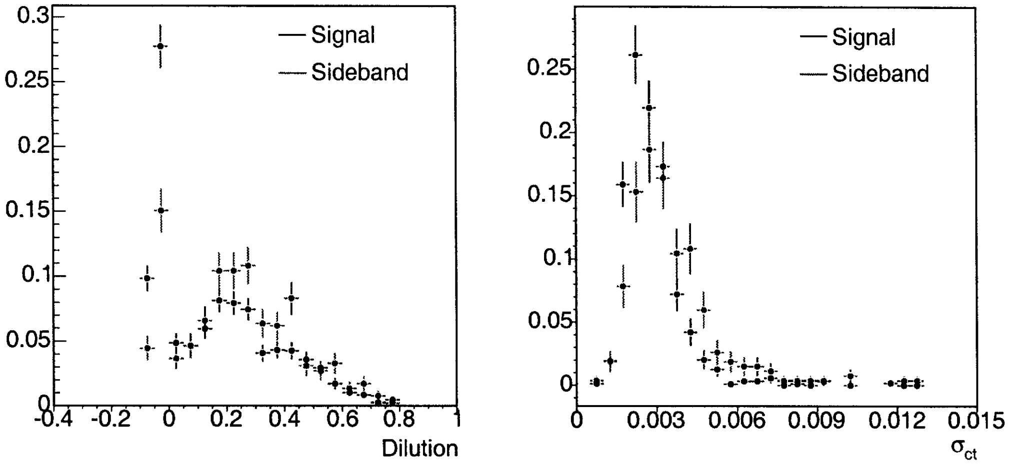

FIGURE 7.2: SST dilution (left) and σ_{ct} (right) templates, for B_s^0 signal and combinatorial background in the $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi^0 \pi^-$ decay mode.

It should be noted that both the amplitude scan in the Δm_s region of sensitivity relevant for this analysis, and the fit for the mixing frequency are dominated by statistical uncertainty.

7.2.1 Systematic uncertainties on the amplitude

Systematic uncertainties are evaluated following the original formulation of the amplitude method in Reference [94]. Toy MC samples, which are generated reflecting the characteristic of the data, are extensively utilized to estimate systematic effects. For each point in the Δm_s spectrum of the amplitude $\mathcal{A}$ (Section 5.1), a set of one thousand "default" toy experiments is generated with the signal oscillating at that frequency. The fit of each of these experiments produces a $(\mathcal{A}_0, \sigma_{\mathcal{A},0})$ pair. For each systematic effect, unless differently specified, "biased" toy MC samples which simulate the potentially mis-modeled aspect the data are generated from the same sequence of random numbers which was utilized to generate the elements of the default set. The fits of the biased experiments produce the set of $(\mathcal{A}_1, \sigma_{\mathcal{A},1})$ pairs. For a given pair of toy MC samples, composed of a default experiment "0" and a biased experiment "1", which are both generated with the same sequence of random number i, the systematic uncertainty is obtained by [94]:

$$\sigma^i_{syst} = (\mathcal{A}^i_1 - \mathcal{A}^i_0) + (1 - \mathcal{A}^i_0) \cdot \frac{\sigma^i_{\mathcal{A},1} - \sigma^i_{\mathcal{A},0}}{\sigma^i_{\mathcal{A},0}}. \tag{7.2.1}$$

The distribution of σ^i_{syst} provides the estimate for the systematic uncertainty σ_{syst}.

In case of a binary effect, i.e., the effect is represented by a binary shift in fitting models, the mean value of the σ^i_{syst} distribution is taken as systematic uncertainty. An example of this type of effect is the possibility that a potentially large value of $\Delta\Gamma/\Gamma$, the widths of the two mass-lifetime eigenstates in the neutral B_s meson system, could bias the results of an amplitude scan. In fact, the expressions in Equation 5.1.2, which are the basis of the probability density functions utilized to define our likelihood function, are valid under the assumption that $\Delta\Gamma/\Gamma$ is negligibly small, as explained in Section 1.2. The possibility of a bias deriving from neglecting a possibly large value of $\Delta\Gamma/\Gamma$ is studied by comparing fit results in toy MC samples with $\Delta\Gamma/\Gamma$ equal to 0.2 ("biased" sample) and 0.0 ("default").

For effects in which a fit parameter or effect is continuously varied across a range of values, the width of the σ^i_{syst} distribution is utilized as the estimate of a systematic uncertainty. For example, toy experiments are generated with a SST $S_\mathcal{D}$ extracted from a Gaussian distribution which is centered in the nominal value of the SST scale factor and has width equal to the statistical uncertainty of the scale factor (Equation 6.10.1). The fit of these samples performed by utilizing the nominal value of the SST scale factor allows one to study the systematic uncertainty related to the incomplete knowledge of the value of the same-side-tagger scale factor.

The main sources of systematic uncertainties which affect the measurement of the amplitude are reviewed in the list below, separated depending on the type of likelihood term which they affect, and ordered, within each class, according to their contribution to the total uncertainty.

- **Flavor tagging**
  Systematic uncertainties related to flavor tagging contribute uniformly to the total uncertainty across the entire Δm_s spectrum. The dominant contribution comes from the scale factor of the same-side-kaon tagger, which is determined with a 14% precision (Equation 6.10.1). The size of this contribution is an uncertainty of $\sim$ 10% on the amplitude for any value of Δm_s in the range chosen for amplitude scans. On the other hand, the scale factor of the opposite-side tagger is known with a very good precision: $S_\mathcal{D}(OST) = 0.99 \pm 0.01$ [102]. Toy MC samples are generated with scale factors extracted from Gaussian distributions, each of which is centered in the nominal value of a scale factor and has width equal to the statistical uncertainty of the scale factor which is being analyzed. These samples are then fit with the nominal scale factors. The tag decisions are combined under the assumption that they are uncorrelated. The bias that would result from a correlation is estimated by introducing a correlation between OST and SST at various levels. The effect of the possible presence of correlations between same-side and opposite-side taggers is an uncertainty of about 8% on the amplitude for $\Delta m_s = 15$ ps^{-1}.

  A few additional studies have been performed for the hadronic data sample. The probability distributions for the dilution which are used in the fit model for the signal and background ($\mathcal{P}(\mathcal{D})$) are known with finite statistical precision. Toy MC samples are generated with variations of the distributions within their statistical uncertainties and fit with the nominal set of distributions for the signal and background components. While the $\mathrm{B}^0_s \to \mathrm{D}^-_s\pi^+(\pi^-\pi^+)$ decay is self-tagging, the Cabibbo-suppressed $\mathrm{B}^0_s \to \mathrm{D}^-_s\mathrm{K}^+(\pi^-\pi^+)$ and $\mathrm{B}^0_s \to \mathrm{D}^+_s\mathrm{K}^-(\pi^-\pi^+)$ decays both receive contributions from tree-level amplitudes with the same order of magnitude, in terms of CKM parameters. The nominal model has the dilution of the Cabibbo-suppressed component equal to the Cabibbo-favored counterpart. Two toy MC samples are produced varying the dilution of the Cabibbo-suppressed component by ±100%, to mimic the possibility that a $D_s^- K^+ (\pi^- \pi^+)$ final state tags, or does not tag, the decay of a B_s^0 meson. Each ensemble is fitted with the nominal model and the larger variation between +100% and −100% is taken as the systematic uncertainty estimate. The same method is utilized to study the effect of assigning wrong dilutions to the candidates which enter the $\overline{\Lambda}_b^0$ component. Among the effects described in this paragraph, the largest contribution to the systematic uncertainty on the amplitude comes from the possible mis-modeling of the dilution of Cabibbo-suppressed B_s^0 decays. The size of the contribution of this effect is a $\sim 4\%$ uncertainty on the amplitude for $\Delta m_s = 15$ ps^{-1}.

- **proper decay-time**

  The most significant systematic uncertainty derives from the global scale factor assigned to proper-decay-time uncertainties. This systematic uncertainty, which increases steeply with the sampled Δm_s, has been evaluated by generating toy experiments with a scale factor on σ_{ct} calculated by adding (or subtracting) the 1σ uncertainty obtained by the calibration of σ_{ct} to the nominal value of the scale factor. This modification of the scale factor simulates systematic over- or under-estimations of the uncertainty on proper decay-time of the B_s^0 signal. The size of this contribution to the total systematic uncertainty is about 5% for $\Delta m_s = 15$ ps^{-1}, and increases with Δm_s.

  The fit does not include the effect of a lifetime difference between $B_{s,H}$ and $B_{s,L}$, the eigenstates of the Hamiltonian, with widths Γ_H and Γ_L. The formulae in Equations 1.2.9 and 1.2.8 assume a negligible value of $\Delta\Gamma = \Gamma_L - \Gamma_H$. The possibility that a potentially large value of $\Delta\Gamma/\Gamma$ introduced a bias in the amplitude has been studied by generating a set of toy experiments with $\Delta\Gamma/\Gamma = 0.2$. Theoretical calculations indicate $\Delta\Gamma/\Gamma = 0.12 \pm 0.06$ [113], while the world average is $\Delta\Gamma/\Gamma = 0.121\ ^{+0.083}_{-0.090}$ [7]. The size of the systematic uncertainty associated with neglecting the effects of a potentially large value of $\Delta\Gamma/\Gamma$ is about 5% for $\Delta m_s = 15$ ps^{-1}.

  Another systematic uncertainty derives from the incomplete description of the detector resolution in the fitter framework. The nominal fit assumes that the detector resolution function is a Gaussian for the B_s^0 signal. An alternative model, a double Gaussian plus exponential tails, is used to evaluate the uncertainty. The systematic uncertainty associated with the use of a single Gaussian to model the detector resolution is about 4% for $\Delta m_s = 15$ ps^{-1}.

  The proper-time efficiency curves $\epsilon(t)$ are of primary importance when performing a lifetime analysis. The mixing of B_s^0 candidates occurs on such a short time scale that the mixing analysis is largely insensitive to the efficiency parameterization. Negligible biases are found to be introduced by this aspect of the model. The systematic uncertainty is evaluated by fitting the same toy MC sample with the default efficiency function $\epsilon(ct)$, defined by Equation 7.1.5, and with a modified efficiency function $\epsilon'(ct)$, obtained by changing τ, the world average of the B_s^0 lifetime, in Equation 7.1.5, within its uncertainty, and accordingly reweighing the BGENERATOR-MC sample used to calculate the efficiency function to simulate the modified lifetime. The values utilized for the B_s^0 lifetime and its uncertainty are $c\tau(B_s^0) = 438\ \mu m$ and $\sigma_{c\tau}(B_s^0) = 17\ \mu m$ [62]. The size of this contribution to the systematic uncertainty is $\sim 1\%$ for $\Delta m_s = 15\ ps^{-1}$.

The likelihood for the hadronic sample does not utilize separate $\mathcal{P}_{\sigma_{ct}}$ for signal and background, because the signal and background σ_{ct} distributions are very similar. This means that small biases may be introduced, which are estimated by fitting the standard toy MC sample with versions of the model which either do or do not include the $\mathcal{P}(\sigma_{ct})$ terms. The size of this contribution to the systematic uncertainty is $\sim 1\%$ for $\Delta m_s = 15\ ps^{-1}$.

The default fit model assumes that the contributions of B^0 and partially reconstructed candidates do not oscillate. Toy MC samples are generated such that these components do mix. These samples are fit with the nominal version of the fitter, and with a version of the fitter which accounts for the oscillation. The difference between the results of the two fits is taken as estimate of the bias introduced by neglecting the mixing. This effect gives a negligible contribution to the total systematic uncertainty on the amplitude.

In the analysis of partially reconstructed hadronic decays, an additional possible source of bias is studied. It is known that the distribution of proper decay-time for combinatorial background has a slow dependence on the mass. This dependence is due to the use of the world average of B_s^0 mass measurements in Equation 5.2.1 for candidates which populate the sideband region far from the signal region. The nominal fit utilizes a single template for the background across the full mass range, effectively averaging over the small variations. A toy MC sample is generated with a background *ct* template obtained by using the reconstructed mass of the B_s^0 candidates in the upper mass sideband in Equation 5.2.1 instead of the world average of B_s^0 mass measurements. This toy MC sample is then fit with the nominal model to conservatively estimate the effect of the choice of a single template. The size of this contribution to the systematic uncertainty is consistent with zero for $\Delta m_s = 15\ ps^{-1}$.

- **Sample composition and mass models**

  The studies of systematic uncertainties in this section are split in three cases: fully reconstructed hadronic B_s^0 decays, partially reconstructed hadronic B_s^0 decays, and semileptonic B_s^0 decays.

  Several uncertainties are assigned to the incompleteness of the knowledge of the sample composition. These uncertainties address the uncertainty in the levels of the Cabibbo-suppressed $B_s^0 \to D_s^- K^+ (\pi^-\pi^+)$, which is treated as a signal component, and the contributions of $\overline{\Lambda}_b^0$ and B^0 decays to the background. The ratio of the number of B_s^0 candidates arising from mis-reconstructed $\overline{\Lambda}_b^0$ and B^0

Reflection [%]	$D_s^- \to \phi^0\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$	—	0.22 ± 0.06	0.08 ± 0.04
$B_s^0 \to D_s^-\pi^+, D_s^- \to K^{*0}K^-$	0.44 ± 0.10	—	0.01 ± 0.00
$B_s^0 \to D_s^-\pi^+, D_s^- \to \pi^-\pi^+\pi^-$	0.00 ± 0.00	0.00 ± 0.00	—
$B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$	2.30 ± 0.17	32.7 ± 2.4	0.44 ± 0.03
$\overline{\Lambda}_b^0 \to \Lambda_c^-\pi^+, \Lambda_c^- \to \bar{p}K^+\pi^-$	1.18 ± 0.20	18.1 ± 3.0	0.22 ± 0.04
$\overline{\Lambda}_b^0 \to \Lambda_c^-\pi^+, \Lambda_c^- \to \bar{p}\pi^+\pi^-$	0.02 ± 0.00	0.16 ± 0.03	3.48 ± 0.58

TABLE 7.1: Reflection ratios for $B_s^0 \to D_s^-\pi^+$ decay modes. The table contains the amount of B^0, $\overline{\Lambda}_b^0$ and B_s^0 hadrons that are erroneously reconstructed as a B_s^0 candidate, relative to the amount of $B_s^0 \to D_s^-\pi^+$ signal candidates. The decay chain of the D_s^- candidate is indicated on top of each column.

Reflection [%]	$D_s^- \to \phi^0\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$B_s^0 \to D_s^-\pi^+\pi^-\pi^+, D_s^- \to \phi^0\pi^-$	—	0.65 ± 0.19	4.77 ± 1.35
$B_s^0 \to D_s^-\pi^+\pi^-\pi^+, D_s^- \to K^{*0}K^-$	0.51 ± 0.15	—	0.51 ± 0.16
$B_s^0 \to D_s^-\pi^+\pi^-\pi^+, D_s^- \to \pi^-\pi^+\pi^-$	0.01 ± 0.01	0.53 ± 0.21	—
$B^0 \to D^-\pi^+\pi^-\pi^+, D^- \to K^+\pi^-\pi^-$	3.35 ± 0.38	50.2 ± 5.6	30.0 ± 3.4
$\overline{\Lambda}_b^0 \to \Lambda_c^-\pi^+\pi^-\pi^+, \Lambda_c^- \to \bar{p}K^+\pi^-$	1.83 ± 0.30	31.4 ± 5.2	14.2 ± 2.4
$\overline{\Lambda}_b^0 \to \Lambda_c^-\pi^+\pi^-\pi^+, \Lambda_c^- \to \bar{p}\pi^+\pi^-$	0.07 ± 0.01	0.31 ± 0.05	5.10 ± 0.85

TABLE 7.2: Reflection ratios for $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ decay modes. The table contains the amount of B^0, $\overline{\Lambda}_b^0$ and B_s^0 hadrons that are erroneously reconstructed as a B_s^0 candidate, relative to the amount of $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ signal candidates. The decay chain of the D_s^- candidate is indicated on top of each column.

decays with respect to the number of signal B_s^0 candidates is fixed in the fits which produce the amplitude scan. The expected value is calculated using $\overline{\Lambda}_b^0$, B^0, and B_s^0 BGENERATOR-MC samples, and the relative branching ratios and production cross-sections published in References [114], [77], and [115]. The resulting normalization ratios are reported in Tables 7.1 and 7.2. A systematic uncertainty is evaluated for all components of the contribution which is larger than 1%, with respect to the signal fraction. The normalization ratios of each of these components are varied within their uncertainties to estimate their contribution to the total systematic uncertainty. The same procedure is utilized to estimate the expected amount of Cabibbo-suppressed B_s^0 decays. In this case, the $B_s^0 \to D_s^-K^+(\pi^-\pi^+)$ branching ratios are assumed to be equal to 5%. The normalization ratio of the Cabibbo-suppressed component is varied by a factor of 2 lower and higher than nominal, and the larger systematic variation in the two scenarios is added to the total systematic uncertainty. The size of the combination of these contributions to the systematic uncertainty is about 1% for $\Delta m_s = 15$ ps^{-1}.

Partially reconstructed hadronic components additionally require a study of the effect of incomplete knowledge of the relative signal fractions. Maximal confu-

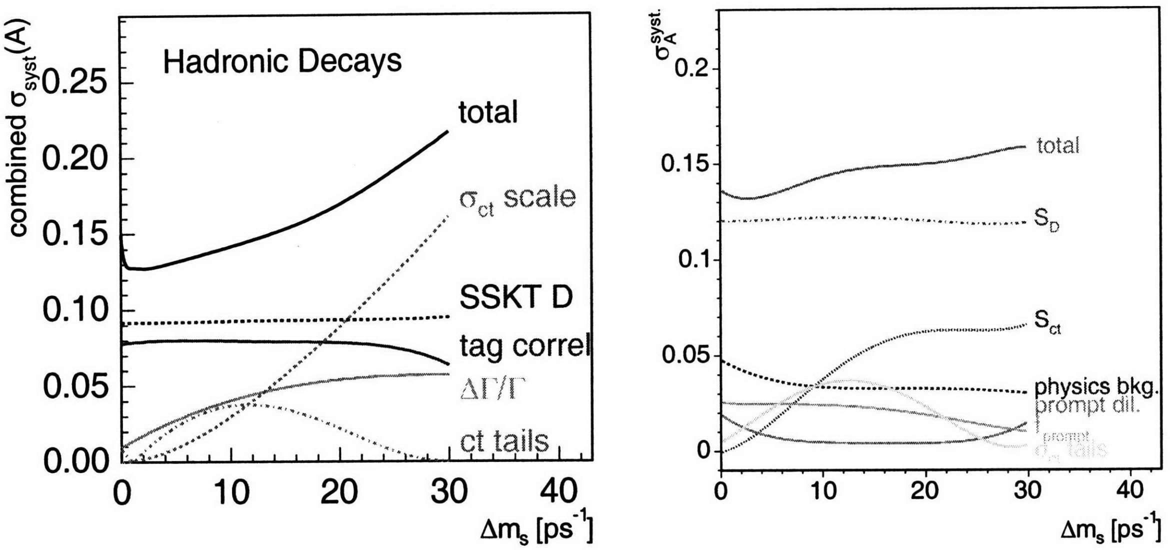

FIGURE 7.3: Systematic uncertainties on $\mathcal{A}$ in hadronic (left) and semileptonic (right) amplitude scans, as a function of Δm_s.

> sion between the components is simulated by generating a toy MC ensemble composed entirely of one signal component and performing the fit as though it were the other. Moreover, a systematic uncertainty is assigned to address a possible bias produced by wrongly modeling the combinatorial background. In the nominal fit, the background is modeled as a smooth exponential function. Toy MC samples are generated with backgrounds which rise more rapidly in the region of partially reconstructed signal candidates. The fit with the nominal model effectively treats the background events as signal. The size of this contribution to the systematic uncertainty is about 2% for $\Delta m_s = 15$ ps^{-1}.
>
> In the case of the semileptonic analysis, two additional sources of systematic uncertainties are considered. The combinatorial background parameterization is derived from the D_s^- mass sidebands. A different set of sidebands, obtained by shifting the bounds of the sideband window by ±50 MeV/c^2 from the nominal values, is utilized to assess systematic uncertainty deriving from a particular choice. The second source of uncertainty comes from the uncertainty in the fraction of the false lepton background. The fraction of the false lepton background has been obtained from a fit of the $m_{\ell^+ D_s^-}$ distribution. The uncertainty from this fit is used to study the related systematic uncertainty. A toy MC sample is generated with the value of this fraction fixed to its nominal value. The default fit of this sample is compared to a fit of the same sample which has the fraction of the false lepton background shifted of $\pm 1\sigma$ from its nominal value, where σ is the uncertainty in this fraction obtained as described above. The size of all these contributions to the systematic uncertainty is below 5% for $\Delta m_s = 15$ ps^{-1}.

The plots of systematic uncertainties *vs.* Δm_s are shown in Figure 7.3, for the amplitude scans produced with hadronic and semileptonic B_s^0 decays.

7.2.2 Systematic uncertainties on Δm_s

The uncertainties presented in the previous section are relevant for the amplitude scan. These uncertainties affect the calculation of the experimental sensitivity and the limit on the oscillation frequency, but they are not relevant to the extraction of Δm_s. For completeness, all the sources of systematic uncertainty in the amplitude scan have been analyzed as possible causes of uncertainty in the Δm_s fit too. They proved to be negligible, while the main systematic uncertainties come from the proper-time scale. Four effects have been studied and presented below. The first three effects are analyzed in detail in the analysis which resulted in the measurements of $B_{s,H}$, $B_{s,H}$, and $\Delta\Gamma_s$, presented in Reference [116].

- **Silicon detector alignment**
  Imperfect alignment of the silicon detector could affect the measurement of proper decay-lengths. A test of possible effects has been performed by introducing distortions into the simulation of silicon detectors and then measuring lifetimes of B mesons using the standard alignment. The B mesons which were utilized for this check are reconstructed in the $B^0 \to J/\psi K^{*0}, J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$, $B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-$, and $B_s^0 \to J/\psi\phi^0, J/\psi \to \mu^+\mu^-$, $\phi^0 \to K^+K^-$ decay modes. The distortions introduced in the simulation include radial displacements and bowing of silicon tensors within tolerances from a physical survey of the detector. The maximum lifetime bias is found to be 1.0 μm, which corresponds to a 0.2% uncertainty on the proper-time scale.

- **Track-fit bias**
  Mis-measurements of track curvature introduce mis-measurements of the transverse decay length, via the location of track vertices reconstructed in the laboratory frame, and the proper decay-time, which is boosted into the B_s^0 reference frame using its transverse momentum. The sign of the bias depends on whether the tracks involved curve toward or away from each other. The bias has been reproduced and studied in the simulation of the COT and silicon detector. It introduces an overall systematic shift in measured lifetimes which is found to be 1.3 μm, corresponding to 0.3% in proper-time scale.

- **Primary vertex bias**
  Mis-measurements of the primary vertex position lead to mis-measurements of the transverse decay length and, therefore, of the proper decay-time. The bias is studied by comparing the primary vertex position with the average beam position in a large sample of fully reconstructed B candidates. The maximum bias is found to be 1.0 μm in the reference frame of the detector, which corresponds to 0.02 ps^{-1} mean bias to Δm_s in the toy MC samples.

- **Hadronic *k*-factors**
  The dominant partially reconstructed $B_s^0 \to D_s^-\rho^+$ and and $D_s^{*-}\pi^+$ channels inhabit the same phase space and have the same qualitative models. For the purposes of this analysis, the only significant difference in the modeling of their

Source	Value [ps^{-1}]
Silicon detector alignment	0.04
Track fit bias	0.05
Primary vertex bias	0.02
Hadronic k-factors	0.03

TABLE 7.3: Systematic uncertainties in the fit for Δm_s.

proper-time components is in the k-factor distributions, which have slightly different widths and mean values. This raises the concern that not using the correct relative fractions of partially reconstructed components could produce a shift in the fitted value of Δm_s has been addressed. The use of incorrect relative fractions is equivalent to applying the wrong likelihood weights to each of these k-factor distributions. The maximum effect is obtained by fitting a component with the k-factor distribution $F(k)$ and the ct efficiency $\epsilon(ct)$ of another one. The result of this fit compared with the result of a second fit, in which the correct weight functions are utilized. In practice, because there is no basis for thinking that the modeling might be so incorrect, 50% of the induced bias is utilized as a very conservative systematic uncertainty, for a final error contribution of 0.03 ps^{-1}. This error is assigned only for the measurement from the partially reconstructed hadronic sample, and does not contribute significantly to the systematic uncertainty for the overall measurement.

Table 7.3 summarizes the systematic uncertainties on the measurement of Δm_s. The total systematic uncertainty is 0.07 ps^{-1}.

7.3 Amplitude scans

This section presents the amplitude scans obtained from data. As defined in Section 5.1, an amplitude scan consists of a set of fits of the amplitude $\mathcal{A}$. Each of these fits is performed at a (different) fixed value of Δm_s. The following expression for the probability that a B_s^0 candidate decays with the same (opposite) flavor with which it was produced is utilized:

$$\mathcal{P}_{unmixed/mixed}(t) \propto [1 \pm \mathcal{A}\mathcal{D}\cos(\Delta m_s t)] \ , \qquad (5.1.2)$$

where the + (−) sign indicates the case of a B_s^0 meson decaying at time t with the same (opposite) flavor as at production, when the flavor tag of the candidate has dilution $\mathcal{D}$.

The amplitude scans obtained from data are reported in Figures 7.4 and 7.5, in the different data samples and separately utilizing the SS and OS taggers. The signature of a mixing signal is an amplitude value consistent with unity and inconsistent with zero. The scans present this behavior in the proximity of the bin corresponding to $\Delta m_s = 17.75$ ps^{-1}. While the scan in the fully hadronic sample has a striking

Sample	Sensitivity
Hadronic	30.7 ps^{-1}
Semileptonic	19.4 ps^{-1}
Same-Side Tagger[†]	30.3 ps^{-1}
Opposite-Side Tagger[†]	25.5 ps^{-1}
Combined	31.3 ps^{-1}

TABLE 7.4: Sensitivity of a mixing analysis in different data subsamples, and in the combination of all data samples, with same-side and opposite-side taggers, separately. [†] Computed using statistical uncertainty only; however, this is the dominant uncertainty.

signature, the one in the semileptonic sample alone is sufficient to set a 95% C.L. double-sided limit on Δm_s.

The sensitivity of an analysis is defined as the value of the frequency for which a measured null amplitude value $\mathcal{A} = 0$ would imply the exclusion of $\mathcal{A} = 1$ at the desired confidence level. From this definition, it derives that the sensitivity of a mixing measurement is defined as the Δm_s value for which $1.645\sigma_{\mathcal{A}} = 1$. The factor which multiplies $\sigma_{\mathcal{A}}$, 1.645, defines a confidence level of 95% for the sensitivity [2]. The sensitivity of an analysis of B_s^0 oscillations in the semileptonic sample is 19.4 ps^{-1}, while an analysis in the hadronic one reaches 30.7 ps^{-1}. The power of the same-side tagger is evident in the amplitude scans reported in Figure 7.5, where the SS and OS taggers achieve a sensitivity, based on the statistical uncertainty only, of 30.3 ps^{-1} and 25.5 ps^{-1}, respectively.

The combined amplitude scan, which includes all data samples and all tagging algorithms, is shown in Figure 7.6 and reaches a sensitivity of 31.3 ps^{-1}. The signature of a mixing signal around $\Delta m_s = 17.75$ ps^{-1} is striking. It is important to point out that this signal lies well within the reach of the sensitivity, and is consistent with unity, which indicates that all components of the analysis are correctly calibrated.

Table 7.4 summarizes the sensitivity of mixing analyses in different subsamples, hadronic and semileptonic B_s^0 decays, and separately utilizing same-side and opposite-side taggers.

7.4 Δm_s fit

The amplitude scans presented in the previous section show the clear signature of a mixing signal. The point in the amplitude scan which is most inconsistent with $\mathcal{A} = 0$ is in the bin corresponding to $\Delta m_s = 17.75$ ps^{-1}, in an amplitude scan performed in steps of 0.25 ps^{-1}. The next natural step is the estimation of the significance of the signal observed, and the measurement of Δm_s.

The significance of the signal observed in an amplitude scan measures how likely it is that random fluctuations produce a signal of oscillations as large or larger than the one observed. The quantity $\Lambda(\Delta m_s)$ has been chosen to estimate the significance

---

[2] The factor 1.645 derives from the following formula: $\int_{\mu-1.645\sigma}^{+\infty} \mathcal{G}(x;\mu,\sigma)dx = 0.95$. It allows one to set a lower limit with a confidence level of 95%.

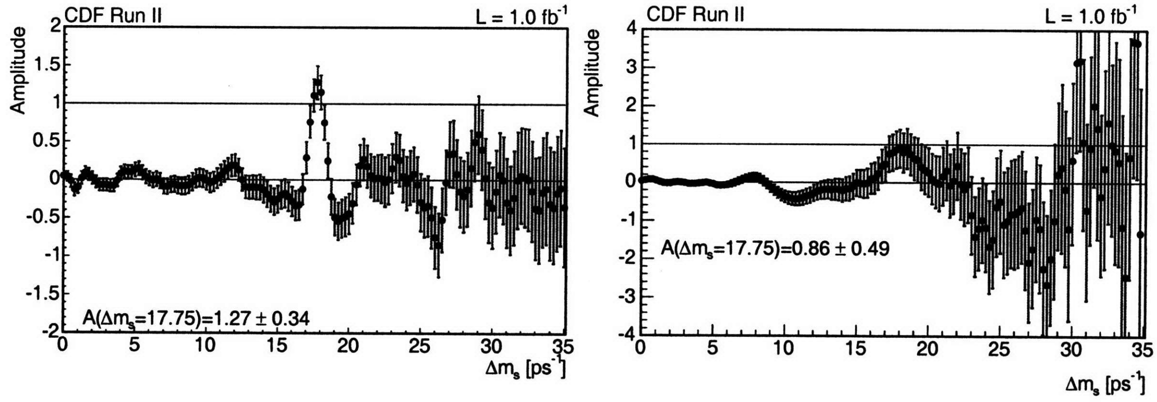

FIGURE 7.4: Amplitude scan, hadronic (left) and semileptonic (right) candidates.

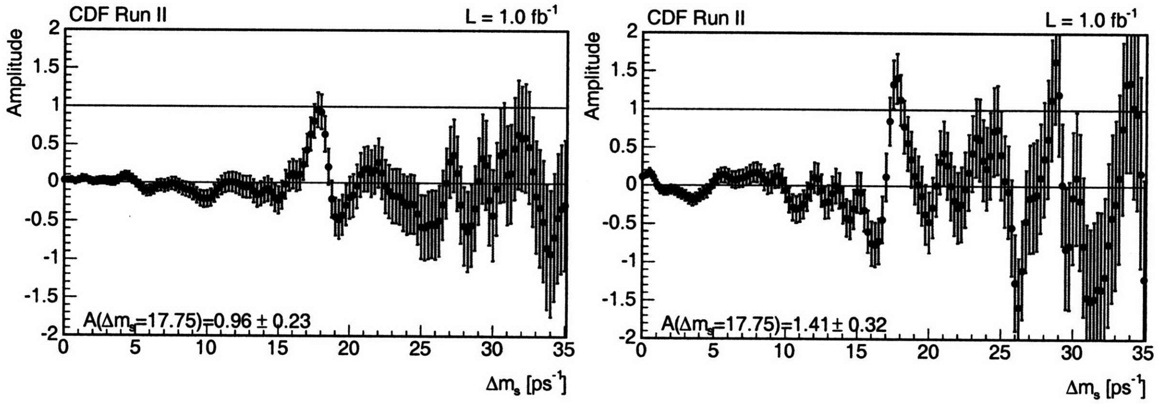

FIGURE 7.5: Amplitude scan: SST (left) and OST (right) only.

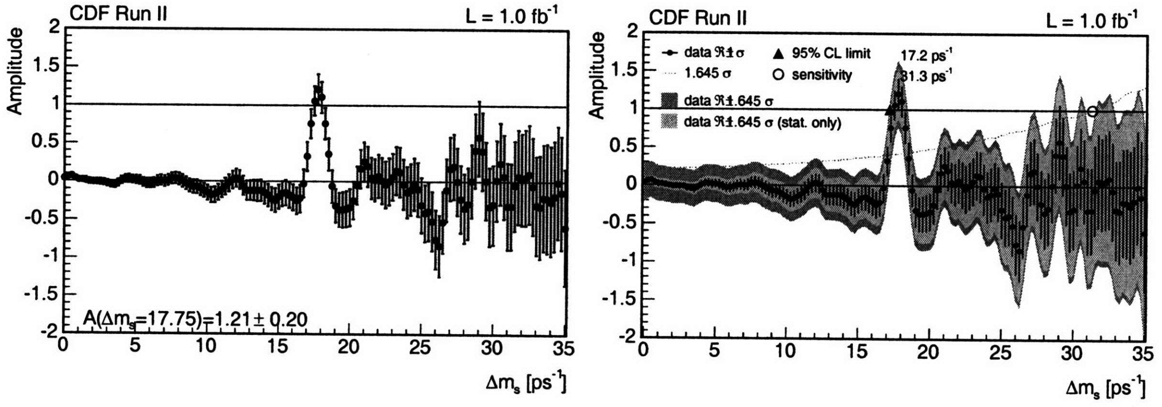

FIGURE 7.6: Amplitude scan with the data of all samples and taggers combined. The version on the left includes the (dominant) statistical uncertainty only. This choice emphasizes the signature of a signal for $\Delta m_s = 17.75$ ps^{-1}. The version on the right includes also systematic uncertainties.

of this analysis. This quantity is defined as the negative logarithm of the ratio of the likelihood calculated with $\mathcal{A} = 1$, which corresponds to the hypothesis that Δm_s is the true mixing frequency, and $\mathcal{A} = 0$, which is independent of Δm_s and corresponds to the hypothesis that there are no oscillations, i.e.:

$$\Lambda \equiv -\log \frac{\mathcal{L}(\mathcal{A}=1)}{\mathcal{L}(\mathcal{A}=0)} . \tag{7.4.1}$$

The significance is determined as a function of the minimum value assumed by Λ in the Δm_s range chosen for the amplitude scan, Λ_{min}. Random fluctuations are simulated by randomly assigning flavor tags to the data sample utilized for this analysis. Therefore, the significance of the analysis presented 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 in data, at any value of Δm_s. Such probability is referred to as *p*-value. The distribution of Λ_{min} in a set of amplitude scans performed after repeatedly and differently randomizing flavor tag decisions in the data sample is shown in Figure 7.7. The cumulative distribution function of Λ_{min} (right plot in Figure 7.7) directly provides the *p*-value as a function of Λ_{min}:

$$p(\Lambda_{\text{min}}^{observed}) = \int_{-\infty}^{\Lambda_{\text{min}}^{observed}} d\Lambda_{\text{min}} f(\Lambda_{\text{min}}) , \tag{7.4.2}$$

where $f(\Lambda_{\text{min}})$ is the distribution of Λ_{min}, in the left plot in Figure 7.7.

The range of amplitude scans has been arbitrarily chosen to be $0 < \Delta m_s[\text{ps}^{-1}] < 35$. The estimation of the *p*-value does not incur any significant bias by selecting a finite window in Δm_s because the likelihood ratio converges rapidly to zero for $\Delta m_s > 35$ ps^{-1}. A cross-check is performed by extending the search range up to $\Delta m_s = 50$ ps^{-1}, with no effect on the *p*-value distribution.

The plot in Figure 7.8 shows the value of the likelihood ratio Λ as a function of Δm_s, for the hadronic and semileptonic B_s^0 decays, separately, and for all data samples combined. The minimal observed value of Λ is $\Lambda_{\text{min}} = -17.26$. A quick look back to the right plot in Figure 7.7 allows one to obtain the *p*-value corresponding to $\Lambda_{\text{min}} = -17.26$. The distribution of *p*-value *vs.* Λ_{min} shows that, out of 3.5×10^8 entries, only 28 scans have a value of Λ_{min} smaller than -17.26. This means that the probability for random scans to produce a signal as significant as the one seen in data, i.e., the *p*-value, is 8×10^{-8}. This is well below the 5–standard-deviations threshold which corresponds to 5.7×10^{-7}.

The plot of $\Lambda(\Delta m_s)$ in Figure 7.8 allows for the determination of the value of Δm_s that best fits the data, which corresponds to the value that minimizes Λ, and its statistical uncertainty, which is determined by the value of Δm_s where Λ changes by 0.5 from the minimal value. The following measurement is obtained:

$$\Delta m_s = 17.77 \pm 0.10(stat.) \pm 0.07(syst.)\ \text{ps}^{-1} . \tag{7.4.3}$$

Finally, the plot in Figure 7.9 shows the comparison between the CDF measurement of Δm_s and the result of the fit for Δm_s performed by the CKM Fitter group [16]

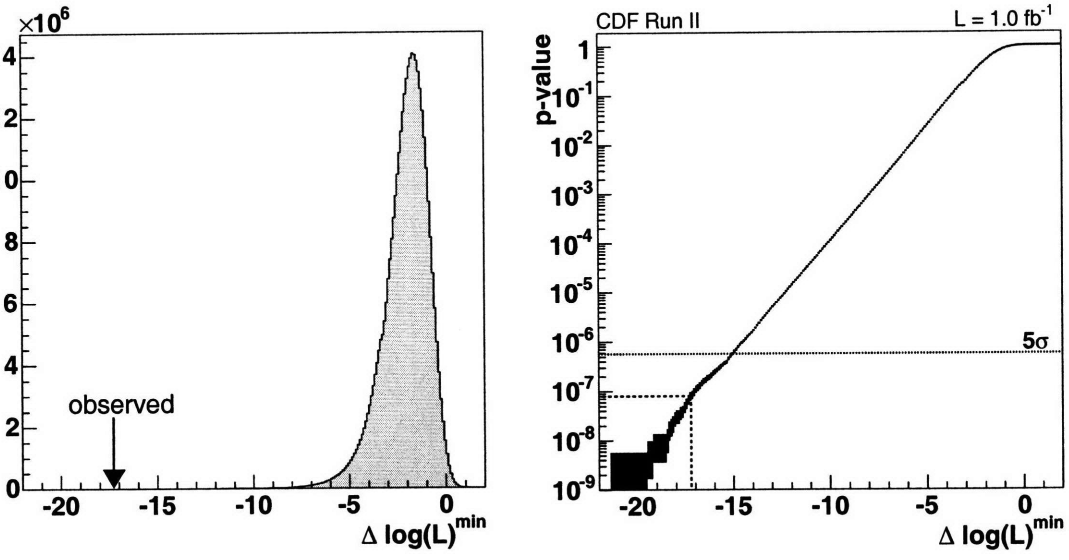

FIGURE 7.7: Distribution of Λ_{min} (left) and *p*-value *vs.* Λ_{min} (right). Following the definition of *p*-value in Equation 7.4.2, the plot on the right is obtained as the cumulative distribution of Λ_{min}, which is shown on the left.

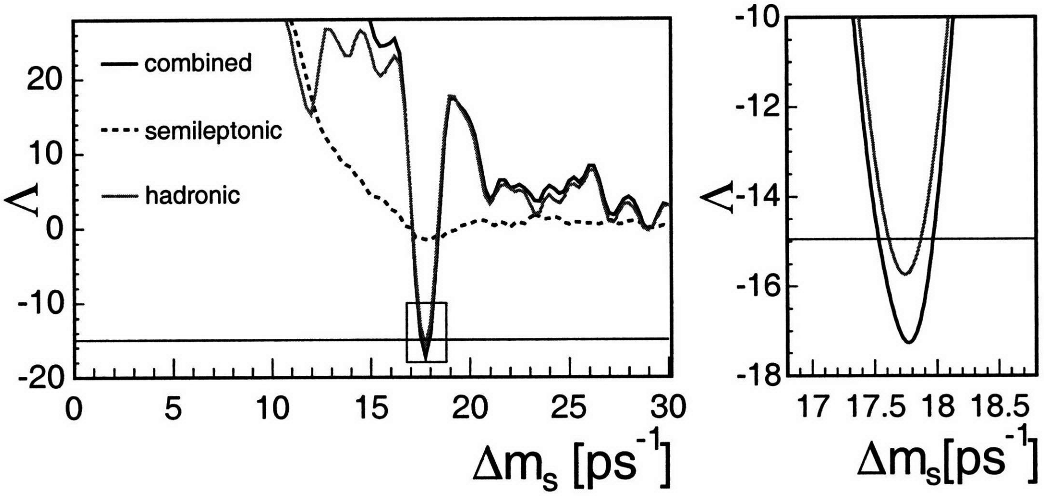

FIGURE 7.8: The likelihood ratio Λ *vs.* Δm_s, where the left plot is the full scan and the right plot is a zoom-in of the $[15, 20]$ ps^{-1} region in Δm_s. The contribution of hadronic and semileptonic B_s^0 decays are shown separately (solid gray and dashed black) and combined (solid black). The horizontal line in the right plot indicates $\Lambda = -15$, which corresponds to a probability of 5.7×10^{-7} (5 standard deviations) in the case of randomly tagged data.

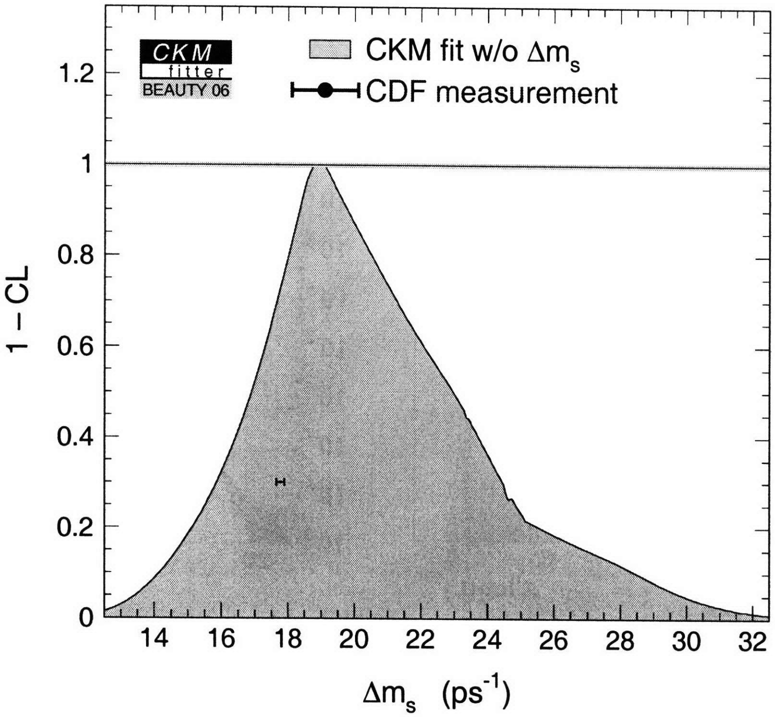

FIGURE 7.9: CKM [16] fit and CDF experimental measurement of Δm_s. Each horizontal section of the shaded area represents the interval in which Δm_s lies with the level of confidence read on the vertical axis of the plot. Intervals are determined by a fit which assumes that flavor interactions are completely described by the SM. The fit does not include any Δm_s analysis among its inputs. The CDF measurement is drawn at 1-CL= 31.7%, which corresponds to a 1σ interval.

without input from Δm_s measurements. The 1σ confidence interval of the CDF measurement is completely contained within the corresponding interval defined by the CKM fit. However, the size of the intervals for Δm_s, at different levels of confidence, indicated by this CKM fit with no Δm_s information makes this a weak test for the validity of the SM of flavor interactions.

7.5 Measurement of $|V_{td}/V_{ts}|$ and impact on Unitarity Triangle

The previous section presented the measurement of Δm_s, and the observation of B_s^0 oscillations with a significance larger than 5 standard deviations. This measurement represents a significant achievement *per se*, but it also provides a way to get a constraint on CKM parameters with unprecedented precision.

The measurement of $\Delta m_d/\Delta m_s$, under the assumption that $|V_{cs}| = |V_{tb}|$, allows for the determination of the side of the unitarity triangle opposed to the angle γ with

Parameter	Value
$m_{\mathrm{B}^0}/m_{\mathrm{B}^0_s}$[117]	$0.98390 \pm \mathcal{O}(10^{-4})$
Δm_d [7]	$0.507 \pm 0.005\ \mathrm{ps}^{-1}$
Δm_s	$17.77 \pm 0.10 \pm 0.07\ \mathrm{ps}^{-1}$
ξ [14]	$1.210\ ^{+0.047}_{-0.035}$

TABLE 7.5: Parameters used to determine $|V_{td}|/|V_{ts}|$ in Equation 1.3.2.

a 3 – 4% resolution. As was derived in Section 1.3 this ratio can be expressed as:

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

where ξ is a parameter from lattice calculations. With the measurements reported in Table 7.5, the following determination is obtained:

$$\frac{|V_{td}|}{|V_{ts}|} = 0.2060 \pm 0.0007\ ^{+0.0081}_{-0.0060}. \tag{7.5.1}$$

The first uncertainty refers to the contribution of the Δm_s measurement only, while the second includes all other sources, dominated by the theoretical uncertainty of the parameter ξ.

To put the impact of this measurement in perspective, the constraint on the unitarity triangle obtained in Equation 7.5.1 can be compared to the status as of EPS 2005 [15] in Figure 7.10. This result nicely complements the measurements of $\sin 2\beta$ (or $\sin 2\phi_1$) and $|V_{ub}|$ from $\mathrm{B}^0 \to \pi^- \ell^+ \nu_\ell$. The importance of the CDF measurement of $|V_{td}/V_{ts}|$ is also clearly shown by Figure 7.11, where the result of the theoretical expectation for $|V_{td}/V_{ts}|$, the CDF measurement, and the average of Belle [118] and BaBar [119] measurements are compared.

The resolution of the experimental inputs to Equation 1.3.2 contributes a negligible part of the total uncertainty on $|V_{td}|/|V_{ts}|$. It thus appears necessary to work on the improvement of the determination of the parameter ξ in order to completely exploit the information provided by our analysis. It is also interesting to notice that the relative precision with which Δm_s is measured, $\sigma_{\Delta m_s}/\Delta m_s$, is superior to the relative precision of the Δm_d measurement: $\sigma_{\Delta m_s}/\Delta m_s \sim 0.5\%$ *vs.* $\sigma_{\Delta m_d}/\Delta m_d \sim 1\%$. The precise measurement of Δm_s presented in this document will thus not be the limiting factor in $|V_{td}|/|V_{ts}|$, even if the precision with which the ξ term, which is obtained from lattice QCD calculations, were to greatly improve.

A review of the implications of the measurement of Δm_s in the large class of models in which the 3 × 3 CKM matrix is unitary and tree-level decays are dominated by SM contributions is presented in Reference [120]. This measurement imposes strict constraints on the phase space available to the parameters h_s and σ_s, introduced in Equation 1.4.1, which describe NP contributions to the frequency of B^0_s oscillations in a model-independent fashion.

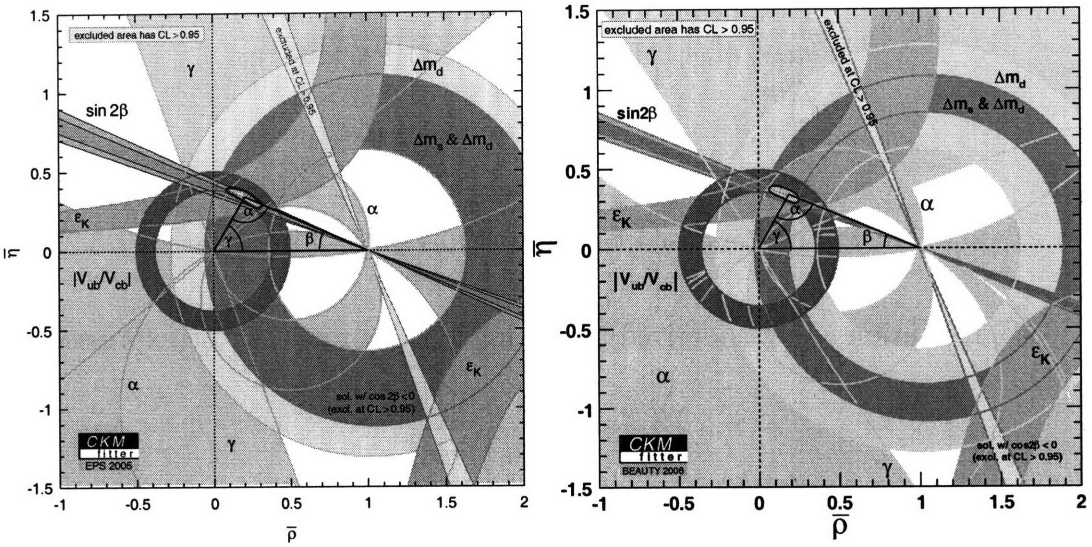

FIGURE 7.10: CKM fits of the unitarity triangle at EPS2005 (left) and after the CDF observation (right) [16].

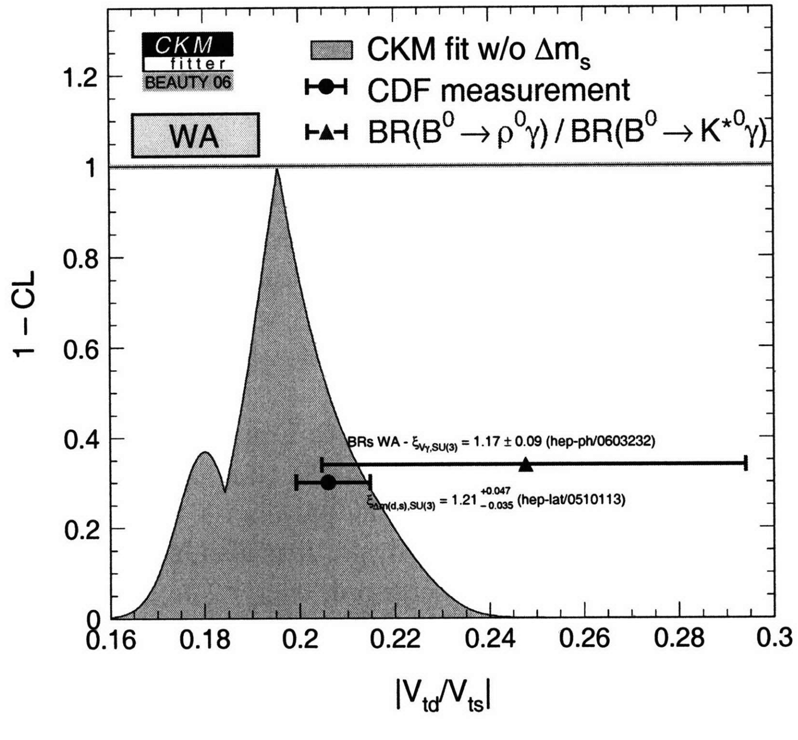

FIGURE 7.11: CKM [16] fit and two experimental measurements of $|V_{td}/V_{ts}|$. Each horizontal section of the shaded area represents the interval in which $|V_{td}/V_{ts}|$ lies with the level of confidence read on the vertical axis of the plot. Intervals are determined by a fit which assumes that flavor interactions are completely described by the SM. The fit does not include any Δm_s analysis among its inputs. The experimental measurements correspond to the CDF measurement and the average of Belle and BaBar measurements of $|V_{td}/V_{ts}|$.

Conclusion

The analysis which resulted in the first observation of time-dependent B_s^0 flavor oscillations is reported in this dissertation, which focuses on the development of a neural-network–based same-side tagging algorithm.

This algorithm for same-side flavor tagging exploits the particle-identification and kinematic information provided by the CDF detector to separate charged hadrons, such as pions, kaons, and protons. This analysis is one of the first CDF analyses to fully integrate particle-identification information provided by the CDF Time-Of-Flight system and the measurement of ionization energy loss in the Central Outer Tracker. Besides same-side flavor tagging, this information is utilized in the selection of B_s^0 candidates, where it allows for a better separation between signal and background than in previous selection schemes. The use of particle-identification is one of the factors which contributed to the significant increase of B_s^0 statistics and sensitivity with respect to analyses which utilized the same sample of CDF data.

The tagging power of the same-side tagging algorithm described in this thesis is about 4%, when applied to the B_s^0 samples reconstructed for this analysis. The same-side tagger is combined with an opposite-side flavor tagger with a tagging power of about 1.8%.

It is also interesting to note the importance of the sample of fully hadronic B_s^0 decays, which contribute 90% of the statistical power available to this analysis. The proper-decay-time resolution achieved in the reconstruction of these candidates is the reason for their superiority. This points to the great performance of CDF tracking and trigger systems, in particular Layer00 and the Secondary Vertex Tracker, which allowed for the collection of large samples of hadronic B_s^0 decays, with excellent proper-decay-time resolution, providing CDF with a great advantage over competing experiments.

The reported result is obtained with a dataset corresponding to an integrated luminosity of about 1 fb^{-1}. The frequency of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations is measured to be:

$$\Delta m_s = 17.77 \pm 0.10(stat) \pm 0.07(syst)\ \mathrm{ps}^{-1}\,,$$

with a significance superior to 5 standard deviations. The signal of B_s^0 oscillations is characterized by an amplitude equal to 1.21 ± 0.20. The consistency of the amplitude with unity indicates that all components of the analysis are correctly calibrated, which is a remarkable achievement.

The measurement of Δm_s provides a stringent constraint on the determination of CKM parameters. In particular, it is possible to derive the following measurement of the Standard Model quantities:

$$\frac{|V_{td}|}{|V_{ts}|} = 0.2060 \pm 0.0007\ ^{+0.0081}_{-0.0060},$$

where the first uncertainty refers to the contribution from the Δm_s measurement only. The second term of the uncertainty is completely dominated by the theoretical uncertainty. The observation of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations concludes a twenty-year long search, and provides an important demonstration of the SM of flavor-interactions.

Appendix A

PYTHIA Setting for MC Generation

The PYTHIA version which is used in CDF is `6.216`. The PYTHIA input parameters for the MC sample described in Chapter 4, which differ from the default settings, are briefly described in this Appendix. The parameters which differ are:

- Tune A (Rick Field tuning) for the underlying event has been used [72, 73].
- B^{**} rate of 20% has been chosen.
- Default values for B^{**} masses and widths were replaced by recent measurements [62].
- The Lund string fragmentation model has been used. As input to this fragmentation model a so-called z variable (Equation 4.3.15) is needed, which describes the ratio of momenta taken by the B meson from the string. For high z values the B meson tends to have higher momentum and the average number of particles formed out of the string is accordingly lower. There are various different z variable distributions on the market. The default shape for this distribution (symmetric Lund [90]) has been used for the light (u,d,s) strings, while for the heavy quarks (c,b) the Peterson fragmentation function [91] with the tuning parameter $\epsilon = 0.006$, has been used. This is not the PYTHIA default but it is needed later for tuning of the z variable distribution of the b string.

The following tcl switches were used to generate the PYTHIA-MC sample described in Chapter 4:

```
module enable Pythia
module talk   Pythia
  PythiaMenu
    msel set 1
    cmEnergy set 1960
    commonMenu
```

```
     set_ckin -index=3 -value=5.
     set_ckin -index=4 -value=-1.
// Tuning Pythia for Underlying event
// Prescription "A" from Rick Field
// ---------------------------------------
//     PDFs - CTEQ Set 5L (LO)
//     These settings are only valid for CTEQ5L
//
     set_mstp -index=51 -value=4046
     set_mstp -index=52 -value=2
// --------------------------------------
// Set ISR max scale factor parameter
// Old ISR setting with more initial-state radiation
     set_parp -index=67 -value=4.0
// ---------------------------------------
// Multiple Interaction parameters
// turn m.i. ON
     set_mstp -index=81 -value=1
// ---------------------------------------------------------
// assume single gaussian hadronic matter distr. turn off at
     set_mstp -index=82 -value=4.0
// ---------------------------------------------------------
// turn-off parameters
     set_parp -index=82 -value=2.0
//     Warm-Core: 50% of matter in radius 0.4
     set_parp -index=83 -value=0.5
     set_parp -index=84 -value=0.4
// -----------------------------------------------------
// probability of gg interaction with colour connection
// Almost Nearest Neighbor
     set_parp -index=85 -value=0.9
// total probability of gg interactions
     set_parp -index=86 -value=0.95
// reference energy scale for m.i.
     set_parp -index=89 -value=1800.
     set_parp -index=90 -value=0.25
// -------------------------------------------------------
// set top mass
     set_pmas -masscode=6 -mass=175.
// -------- setup the fragmentation function in PYTHIA ---
// Peterson with epsilon = 0.006
     set_mstj -index=11 -value=3
     set_parj -index=55 -value=-0.006
//--------------- setup the B** rates -------------------
     set_parj -index=14 -value=0.2
```

```
      set_parj -index=15 -value=0.0666667
      set_parj -index=16 -value=0.0666667
      set_parj -index=17 -value=0.0666667
 // --------------- override B** and D** masses and widths ---------
 // B**
      set_pmas -masscode=10521 -mass=5.70 -width=0.200 -maxdev=0.10
      set_pmas -masscode=10511 -mass=5.70 -width=0.200 -maxdev=0.10
      set_pmas -masscode=20523 -mass=5.73 -width=0.200 -maxdev=0.10
      set_pmas -masscode=20513 -mass=5.73 -width=0.200 -maxdev=0.10
      set_pmas -masscode=10523 -mass=5.73 -width=0.020 -maxdev=0.05
      set_pmas -masscode=10513 -mass=5.73 -width=0.020 -maxdev=0.05
      set_pmas -masscode=525   -mass=5.74 -width=0.020 -maxdev=0.05
      set_pmas -masscode=515   -mass=5.74 -width=0.020 -maxdev=0.05
 // D**
      set_pmas -masscode=10421 -mass=2.31 -width=0.300 -maxdev=0.10
      set_pmas -masscode=10411 -mass=2.31 -width=0.300 -maxdev=0.10
      set_pmas -masscode=20423 -mass=2.43 -width=0.300 -maxdev=0.10
      set_pmas -masscode=20413 -mass=2.43 -width=0.300 -maxdev=0.10
      set_pmas -masscode=10423 -mass=2.42 -width=0.020 -maxdev=0.05
      set_pmas -masscode=10413 -mass=2.42 -width=0.020 -maxdev=0.05
      set_pmas -masscode=425   -mass=2.46 -width=0.020 -maxdev=0.05
      set_pmas -masscode=415   -mass=2.46 -width=0.020 -maxdev=0.05
 // Ds**
      set_pmas -masscode=10433 -mass=2.536 -width=0.002 -maxdev=0.0001
      set_pmas -masscode=435   -mass=2.572 -width=0.015 -maxdev=0.0005
    exit
  exit
exit
```

Appendix B

B^+ and B^0 data and MC samples

This appendix complements the information presented in Chapter 3 and 4. The data and MC samples of B^+ and B^0 mesons which are described in this appendix are utilized to cross-check the work performed on the same-side flavor tagger presented in Chapter 6. They provide an important test because of their large number of events, compared to the B_s^0 data samples.

In particular, this section concentrates on the samples which are utilized to verify the validity of the work on same-side flavor tagging. The reconstructed B^+ and B^+ decay chains are the following:

- $B^+ \to \overline{D}^0\pi^+$, $\overline{D}^0 \to K^+\pi^-$;
- $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$;
- $B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-$;
- $B^0 \to J/\psi K^{*0}$, $J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$.

The samples collected correspond to the `0d` period of data-taking, which is defined in Section 3.2, for a total integrated luminosity of 355 pb^{-1}.

The trigger paths utilized to collect the fully-hadronic decay modes are the `BCHARM`, `LOWPT`, and `HIGHPT` scenarios of two-track triggers. The trigger requirements are explained in Section 3.1. Candidates decayed in modes with a J/ψ in the final state are instead collected using the `DIMUON` trigger path, which is defined as follows:

Level-1

- two XFT tracks with opposite charge;
- each XFT track is matched to two muon stubs;
- $p_T^{XFT} > 1.5(2.2)$ GeV/c for each CMU (CMX) muon;
- $\Delta\varphi_6(CMU, CMU) < 135°$, no cut on $\Delta\varphi_6(CMU, CMX)$;

Level-2

- no cuts: events which pass Level-1 are automatically accepted by Level-2;

Cut		$B^+ \to \overline{D}^0\pi^+$	$B^0 \to D^-\pi^+$
$\chi^2_{r\phi}$(B)	$<$	15	15
$\chi^2_{r\phi}$(D)	$<$	15	15
$\lvert d_0(\mathrm{B})\rvert$ [μm]	$<$	80	110
$L_{xy}/\sigma_{L_{xy}}$(B)	$>$	7	11
$L_{xy}(\mathrm{D} \to \mathrm{B})$ [μm]	$>$	-150	-300
p_T(B) [GeV/c]	$>$	5.5	5.5
$p_T(\pi_\mathrm{B})$ [GeV/c]	$>$	1.0	1.2
$\Delta R(\mathrm{D}, \pi_\mathrm{B})$	$<$	2.0	1.5

TABLE B.1: Selection criteria for fully hadronic B^+ and B^0 candidates. The label π_B indicates the bachelor pion produced in the decay of a B^+ or B^0 meson.

Cut		$B^+ \to J/\psi K^+$	$B^0 \to J/\psi K^{*0}$
$\mathcal{P}$(B)	$>$	10^{-3}	10^{-4}
$\chi^2_{r\phi}$(B)	$<$	—	225
$L_{xy}/\sigma_{L_{xy}}$(B)	$>$	4.5	4.5
$\sigma_{L_{xy}}$(B) [cm]	$<$	0.04	0.04
p_T(B) [GeV/c]	$>$	5.0	5.0
$p_T(K^{+,*0})$ [GeV/c]	$>$	1.0	1.0
$\lvert m_{K\pi} - m_{K^{*0}}\rvert$ [MeV/c^2]	$<$	—	50

TABLE B.2: Selection criteria for di-muon B^+ and B^0 candidates.

Level-3

- mass $m_{\mu\mu}$ between 2.7 GeV/c^2 and 4.0 GeV/c^2.

The triggers that belong to the family of `DIMUON` triggers are differentiated by the type of muon-pair which they require: (CMU,CMU) or (CMU,CMX).

B.1 Selection of data samples

The selection of B^+ and B^0 candidates is performed by applying rectangular cuts. The method is analogous to the one adopted to select B^0_s semileptonic decays (Section 3.4.1), while the selection of hadronic B^0_s candidates is based on a Neural Network (Section 3.4.2). The value of the cuts is chosen by optimizing $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, where the amount of signal $\mathcal{S}$, in a predefined signal region, is evaluated in a BGENERATOR-MC sample of signal events, while the number of background events $\mathcal{B}$ is measured extrapolating the mass fit of the sidebands in data. The upper sideband only is utilized in the case of $B^+ \to \overline{D}^0\pi^+$ and $B^0 \to D^-\pi^+$ decays, while both upper and lower sidebands are used in the case of the $J/\psi K^{+,*0}$ decay modes. The selection cuts utilized are summarized in Table B.1 and B.2, for fully hadronic and di-muon modes, respectively.

Decay Sequence	Yield
$B^+ \to \overline{D}^0\pi^+$	9270
$B^0 \to D^-\pi^+$	8040
$B^+ \to J/\psi K^+$	5240
$B^0 \to J/\psi K^{*0}$	2360

TABLE B.3: B^+ and B^0 signal yields. The quoted numbers corresponds to an integrated luminosity of ~ 355 fb^{-1}.

The yields of B^+ and B^0 candidates collected in 355 pb^{-1} of integrated luminosity are reported in Table B.3. The yields of B_s^0 candidates, in the same data sample, are one order of magnitude smaller.

B.2 Monte Carlo samples of B^+ and B^0 mesons

The PYTHIA-MC samples utilized for the study of same-side tagging are prepared and tuned following the directions explained in Chapter 4. This section presents the data–MC-simulation comparison of the quantities that are expected to have the greatest influence on the performance of an algorithm for same-side flavor tagging.

The plots with the comparison of data and simulated events are divided in three sets. The plots in the first set present the comparison of the distributions of the track variables utilized to select tagging track candidates: impact parameter significance d_0/σ_{d_0}, the separation in η–φ space ΔR, the longitudinal separation Δz_0 between the tagging track candidate and the reconstructed B meson, the pseudorapidity η, and the number of hits in the silicon detector. Each distribution is produced by applying all the cuts used for the selection of tag candidates, which are presented in Section 6.3, except the one on the quantity which is being tested. The comparison of the distributions in data and simulated events of the number of tag candidates found per B candidate is shown too. Figures B.1 to B.4 present the distributions relative to $B^+ \to \overline{D}^0\pi^+$, $B^0 \to D^-\pi^+$, $B^+ \to J/\psi K^+$, and $B^0 \to J/\psi K^{*0}$, in the same order.

The second set of plots contains the comparison between the distributions, in data and PYTHIA-MC, of transverse momentum, impact parameter, transverse decay length and transverse decay length resolution of the reconstructed B candidates. The plots for B^+ and B^0 candidates, reconstructed in fully hadronic modes and in decay chains containing a J/ψ, are presented in Figures B.5 to B.8.

The plots in the third set compare the distribution of CLL (Equation 6.5.3), the variable utilized to perform particle identification. The four plots in Figure B.9 show the comparison of the distributions of CLL in data and PYTHIA-MC simulation for the B^+ and B^0 decay modes utilized throughout this appendix.

The distributions presented in this section show a good agreement between data and simulated events for all the characteristics that are important to assess the performance of a same-side tagging algorithm. Figure 6.12 completes the data–MC-simulation comparison by showing a good agreement between the performance of the maxCLL algorithm for same-side tagging measured in data and calculated in

Sample	$S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$ [%]	
	Data	Simulation
$B^+ \to \overline{D}^0\pi^+$	30.4 ± 1.3	29.0 ± 0.3
$B^0 \to D^-\pi^+$	19.0 ± 2.5	17.2 ± 0.4
$B^+ \to J/\psi K^+$	26.4 ± 2.1	28.5 ± 0.2
$B^0 \to J/\psi K^{*0}$	13.6 ± 5.4	16.8 ± 0.3

TABLE B.4: maxp_L^{rel} algorithm of Same-Side Tagging. The dilution is measured in data and PYTHIA-MC samples. The quoted uncertainty is statistical only.

Sample	$S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$ [%]	
	Data	Simulation
$B^+ \to \overline{D}^0\pi^+$	25.7 ± 1.3	27.5 ± 0.2
$B^0 \to D^-\pi^+$	17.6 ± 2.3	17.4 ± 0.4
$B^+ \to J/\psi K^+$	23.9 ± 2.2	27.0 ± 0.2
$B^0 \to J/\psi K^{*0}$	13.7 ± 5.5	17.9 ± 0.3

TABLE B.5: Particle-identification–based algorithm of Same-Side Tagging. The dilution is measured in data and PYTHIA-MC samples. The quoted uncertainty is statistical only.

PYTHIA-MC samples of reconstructed B^+ and B^0 candidates.

B.3 Performance of same-side taggers

The data and MC samples of B^+ and B^0 candidates presented in this appendix allows one to perform an important test for same-side taggers. Because B^+ mesons do not mix, and B^0 mesons mix with a known frequency which is measurable with precision with CDF data, it is possible to measure tagging performance directly on these data samples, and compare these results with the ones obtained on simulated events. The agreement between the results in data and in MC samples confirm the validity of utilizing a B_s^0 MC sample to calibrate same-side taggers, and then apply the results of the calibration on MC events to B_s^0 data.

Tables B.4 and B.5 show the level of agreement achieved between the performances of the maxp_L^{rel} and maxCLL algorithms for same-side tagging measured in data and calculated in PYTHIA-MC samples of B^+ and B^0 candidates. The agreement between the effective dilution measured in data and calculated in PYTHIA-MC samples is always better than two standard deviations.

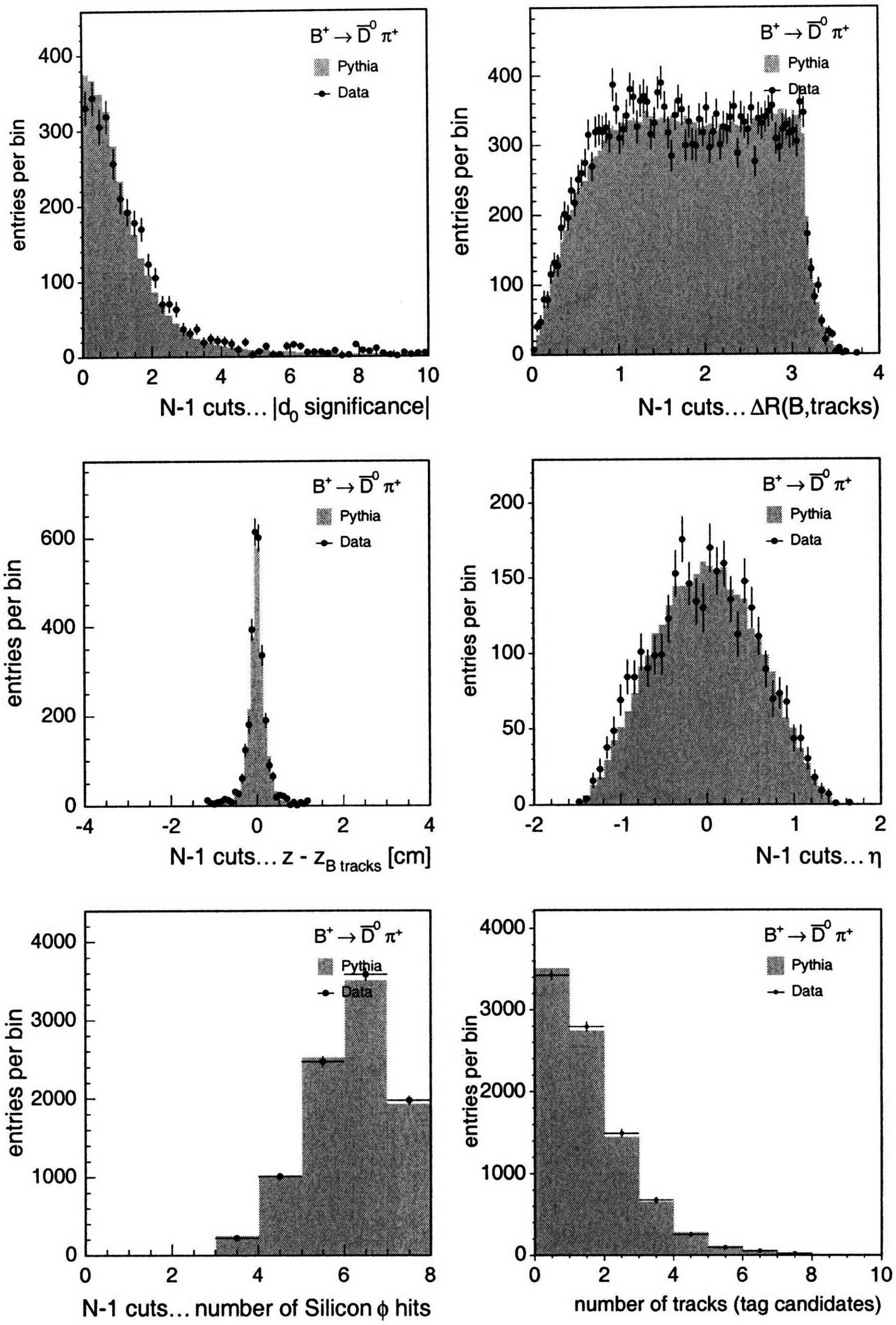

FIGURE B.1: Data–PYTHIA-MC-simulation comparison of track variables. These track variables will be utilized to preselect tagging track candidates. From left to right, and top to bottom, are plotted the distributions for: impact parameter significance d_0/σ_{d_0}, angular separation $\Delta R(\mathrm{B}, trk)$, $\Delta z_0(\mathrm{B}, trk)$, pseudorapidity η, number of hits in the silicon detectors (L00, SVX or ISL), and number of SST tag candidates, after all cuts. Each plotted distribution is produced utilizing the sample of tracks which satisfy the requirements for being a tag candidate except for the cut on the variable shown. These distributions are from the $\mathrm{B}^+ \to \overline{\mathrm{D}}^0\pi^+$, $\overline{\mathrm{D}}^0 \to \mathrm{K}^+\pi^-$ data and MC samples.

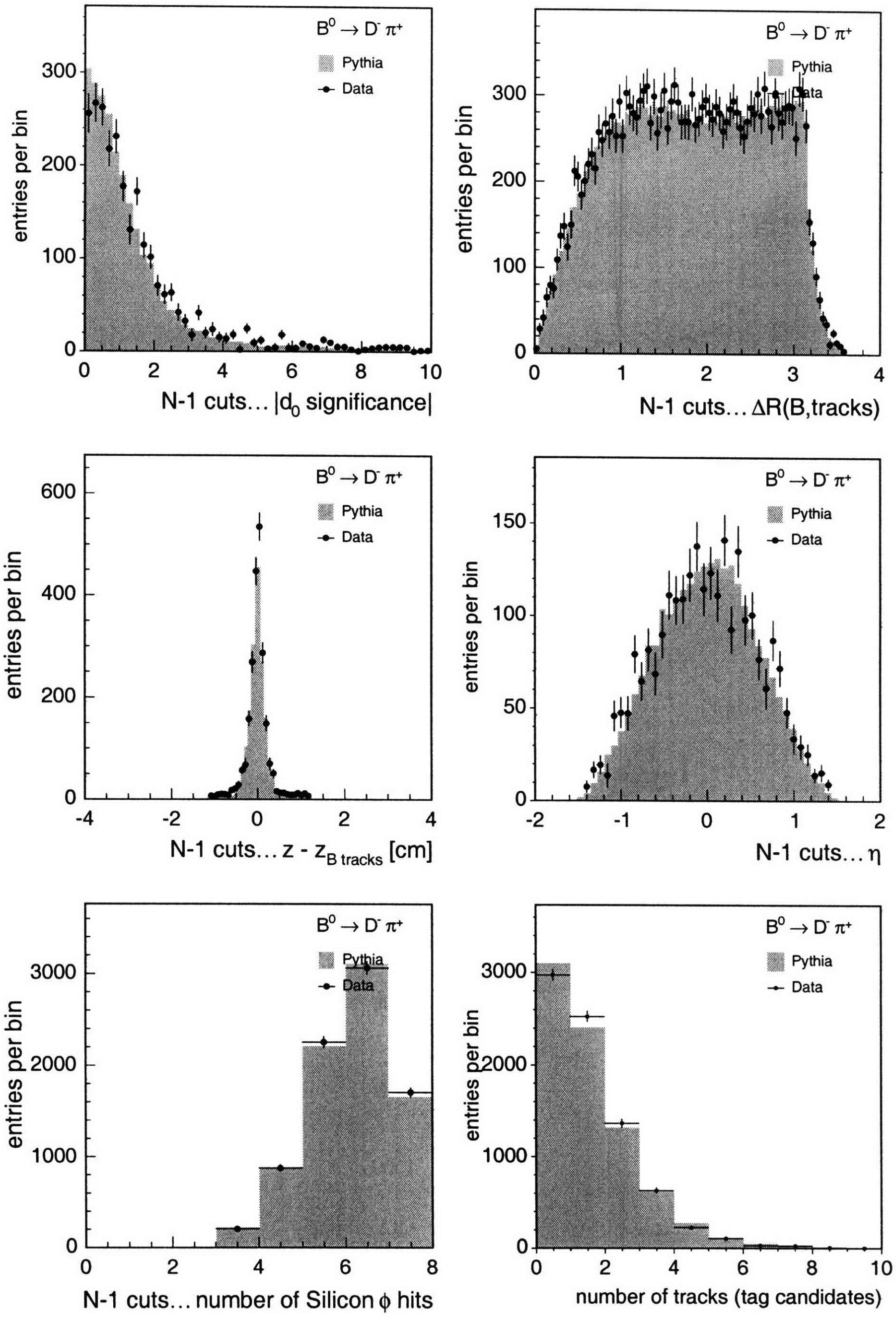

FIGURE B.2: Data–PYTHIA-MC-simulation comparison of track variables. These track variables will be utilized to preselect tagging track candidates. From left to right, and top to bottom, are plotted the distributions for: impact parameter significance d_0/σ_{d_0}, angular separation $\Delta R(\mathrm{B}, trk)$, $\Delta z_0(\mathrm{B}, trk)$, pseudorapidity η, number of hits in the silicon detectors (L00, SVX or ISL), and number of SST tag candidates, after all cuts. Each plotted distribution is produced utilizing the sample of tracks which satisfy the requirements for being a tag candidate except for the cut on the variable shown. These distributions are from the $\mathrm{B}^0 \to \mathrm{D}^-\pi^+$, $\mathrm{D}^- \to \mathrm{K}^+\pi^-\pi^-$ data and MC samples.

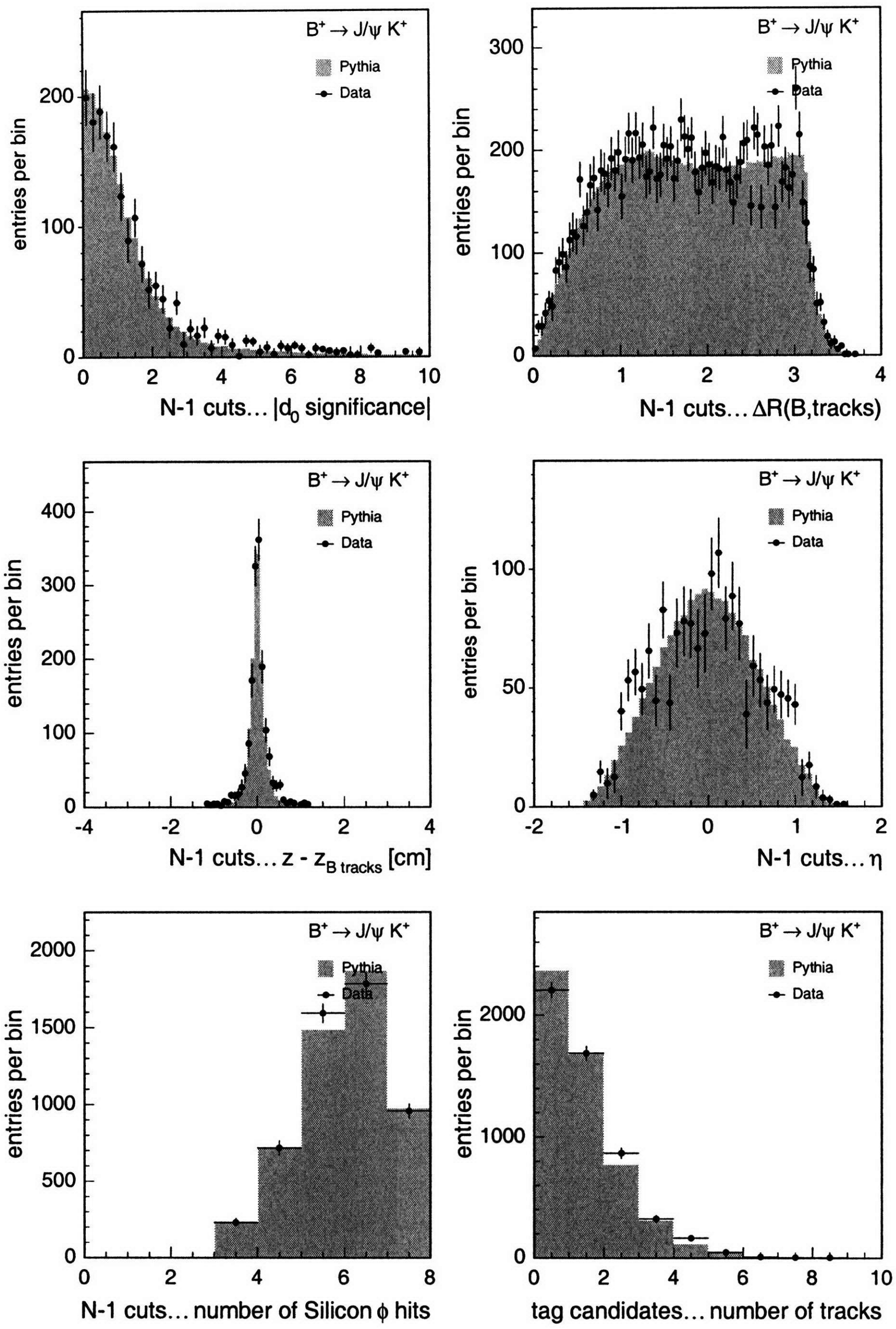

FIGURE B.3: Data–PYTHIA-MC-simulation comparison of track variables. These track variables will be utilized to preselect tagging track candidates. From left to right, and top to bottom, are plotted the distributions for: impact parameter significance d_0/σ_{d_0}, angular separation $\Delta R(\mathrm{B}, trk)$, $\Delta z_0(\mathrm{B}, trk)$, pseudorapidity η, number of hits in the silicon detectors (L00, SVX or ISL), and number of SST tag candidates, after all cuts. Each plotted distribution is produced utilizing the sample of tracks which satisfy the requirements for being a tag candidate except for the cut on the variable shown. These distributions are from the $\mathrm{B}^+ \to J/\psi \mathrm{K}^+, J/\psi \to \mu^+\mu^-$ data and MC samples.

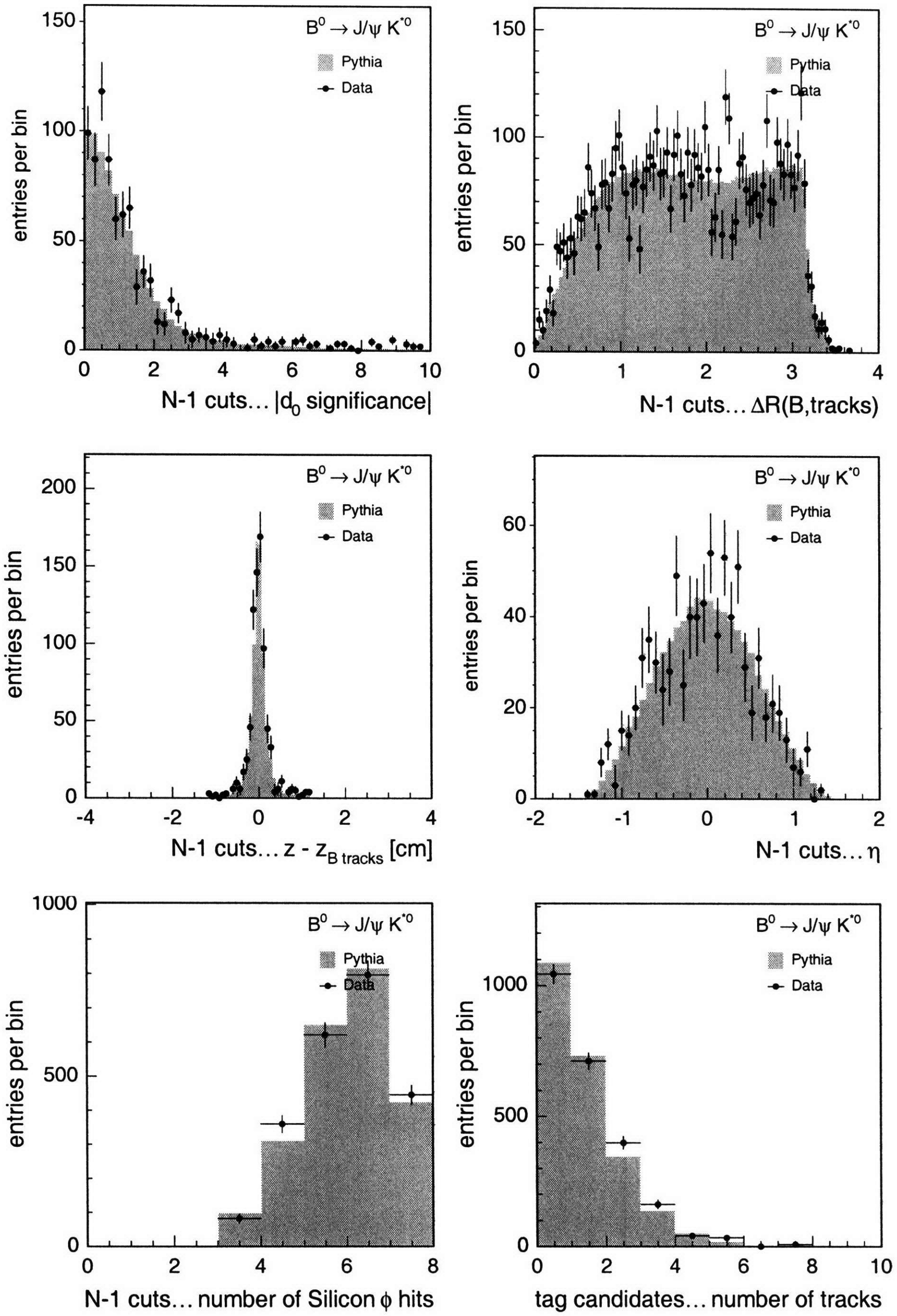

FIGURE B.4: Data–PYTHIA-MC-simulation comparison of track variables. These track variables will be utilized to preselect tagging track candidates. From left to right, and top to bottom, are plotted the distributions for: impact parameter significance d_0/σ_{d_0}, angular separation $\Delta R(\mathrm{B}, trk)$, $\Delta z_0(\mathrm{B}, trk)$, pseudorapidity η, number of hits in the silicon detectors (L00, SVX or ISL), and number of SST tag candidates, after all cuts. Each plotted distribution is produced utilizing the sample of tracks which satisfy the requirements for being a tag candidate except for the cut on the variable shown. These distributions are from the $\mathrm{B}^0 \to J/\psi \mathrm{K}^{*0}$, $J/\psi \to \mu^+\mu^-$, $\mathrm{K}^{*0} \to \mathrm{K}^+\pi^-$ data and MC samples.

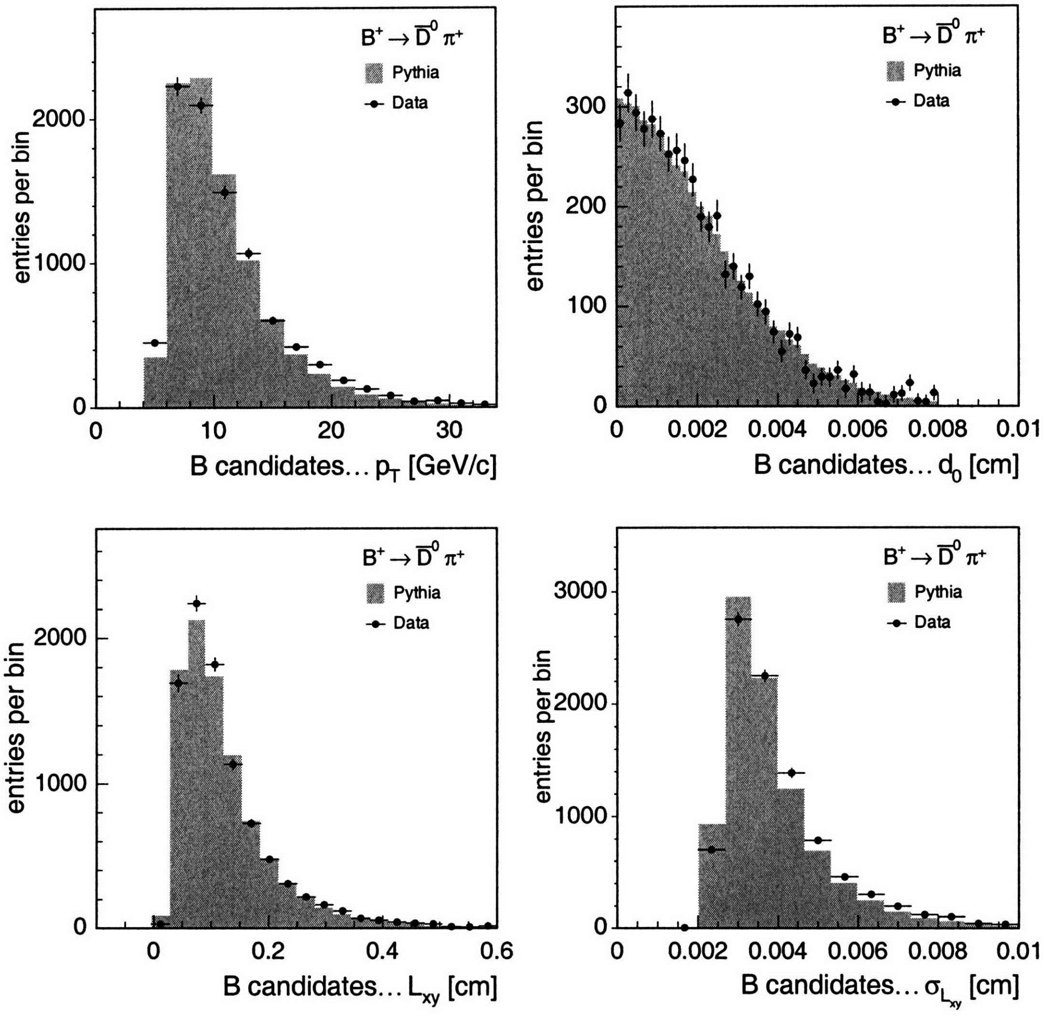

FIGURE B.5: Comparison of distributions of the $B^+ \to \overline{D}^0\pi^+$, $\overline{D}^0 \to K^+\pi^-$ candidates between data (black markers) and PYTHIA-MC simulation (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse momentum, impact parameter, flight distance in the transverse plane L_{xy}, and uncertainty in L_{xy}.

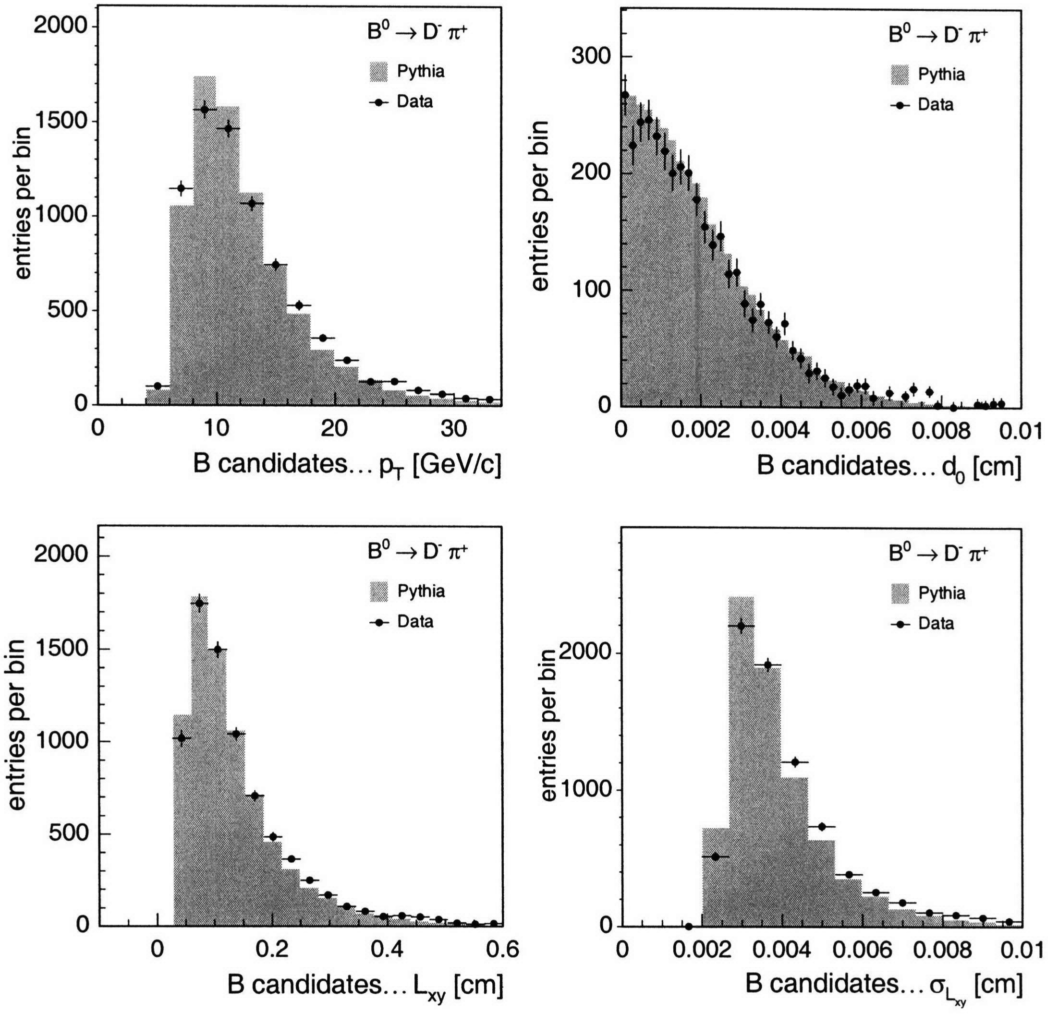

FIGURE B.6: Comparison of distributions of the $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ candidates between data (black markers) and PYTHIA-MC simulation (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse momentum, impact parameter, flight distance in the transverse plane L_{xy}, and uncertainty in L_{xy}.

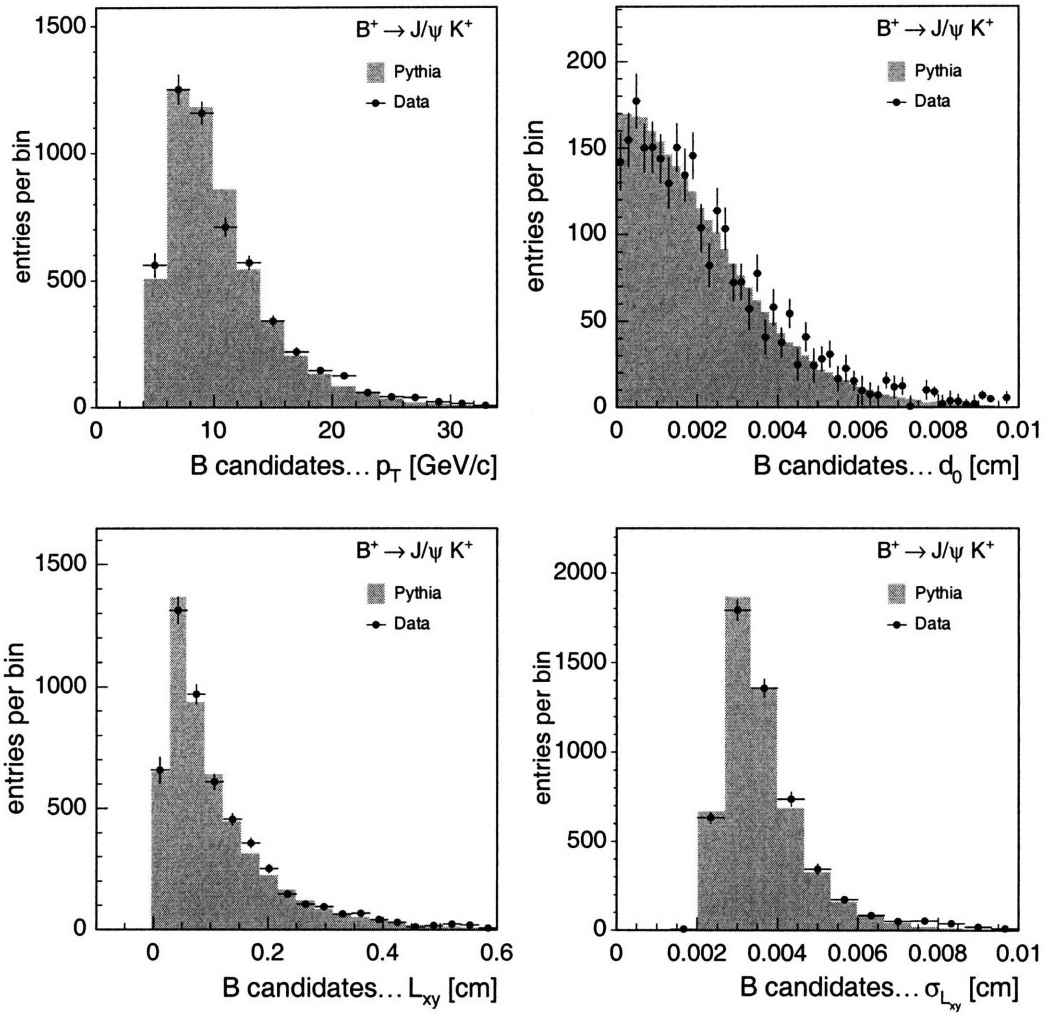

FIGURE B.7: Comparison of distributions of the $B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-$ candidates between data (black markers) and PYTHIA-MC simulation (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse momentum, impact parameter, flight distance in the transverse plane L_{xy}, and uncertainty in L_{xy}.

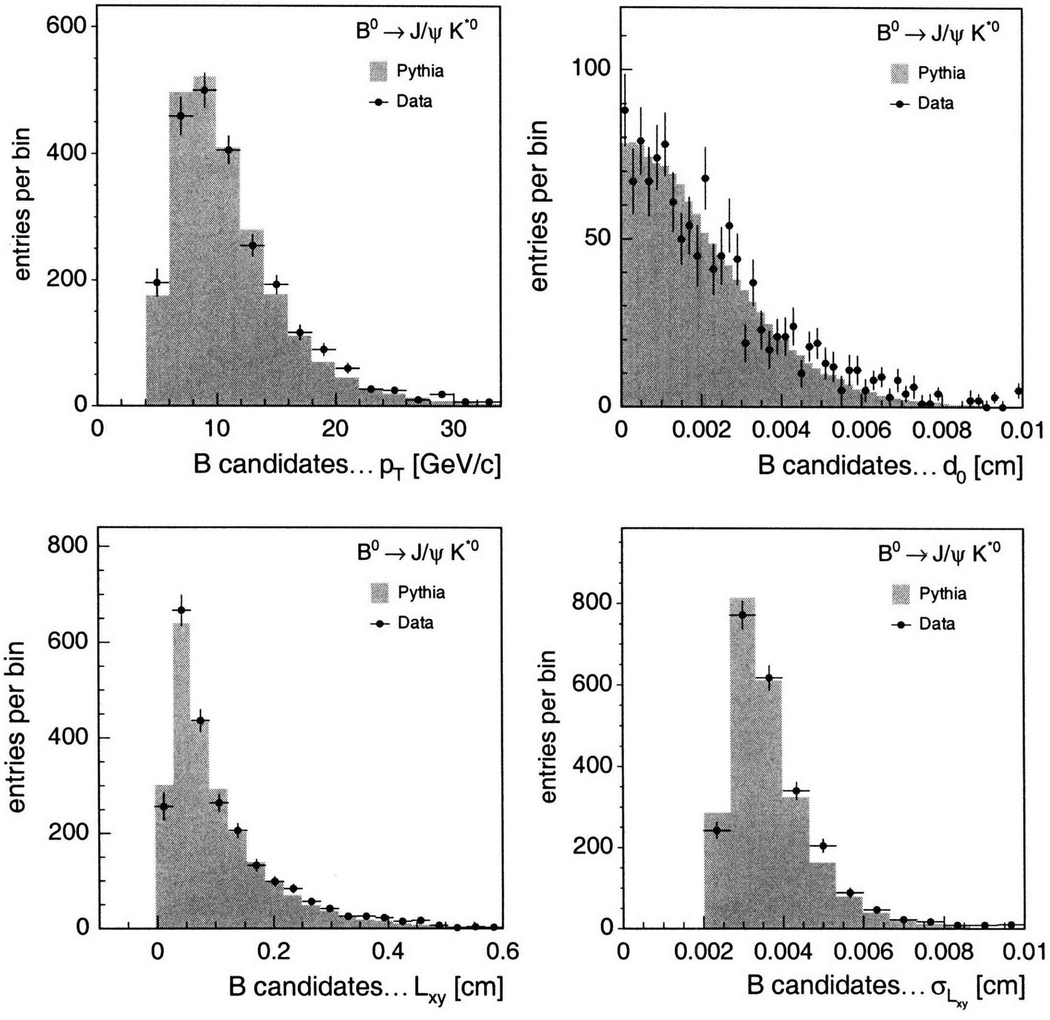

FIGURE B.8: Comparison of distributions of the $B^0 \to J/\psi K^{*0}$, $J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$ candidates between data (black markers) and PYTHIA-MC simulation (solid gray histogram). From left to right, and top to bottom, are plotted the distributions for: transverse momentum, impact parameter, flight distance in the transverse plane L_{xy}, and uncertainty in L_{xy}.

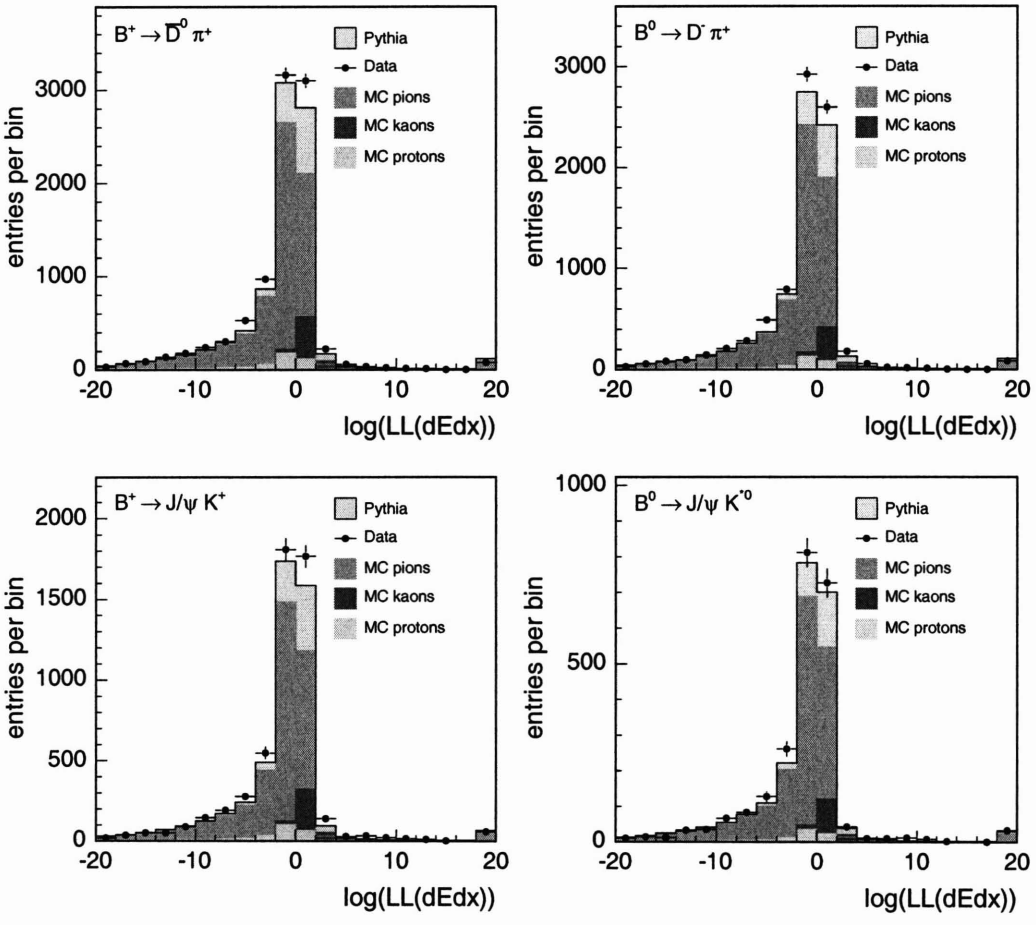

FIGURE B.9: Distribution of CLL for tagging track candidates in data (black dots) and PYTHIA-MC events (histogram). The rightmost bin in each plot corresponds to the cases where neither $d\mathrm{E}/d\mathrm{x}$ nor t_{flight} information are available. The contributions of kaons, pions, and protons to the PYTHIA-MC plot are divided on the basis of MC truth information, and overlaid. From left to right, and top to bottom, are plotted the comparisons for: $\mathrm{B}^+ \to \overline{\mathrm{D}}^0\pi^+$, $\mathrm{B}^0 \to \mathrm{D}^-\pi^+$, $\mathrm{B}^+ \to J/\psi\mathrm{K}^+$, and $\mathrm{B}^0 \to J/\psi\mathrm{K}^{*0}$.

Appendix C

Additional SST algorithms studied

Two algorithms for same-side flavor tagging have been studied into detail: $\max p_L^{\mathrm{rel}}$ (Section 6.4), based on kinematic characteristics of tag candidates, and $\max CLL$, which utilizes particle-identification information (Section 6.5). The analysis of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations which provided the first measurement of Δm_s [17] used the particle-identification–based algorithm, which had proved to perform better than the $\max p_L^{\mathrm{rel}}$ one. This appendix documents the attempts to combine the two algorithms that have been discarded in favor of the neural-network–based algorithm described in Chapter 6.

C.1 Combination of $\max p_L^{\mathrm{rel}}$ and $\max CLL$ decisions

The simplest way, in terms of the available quantities, to combine the kinematic information contained in the $\max p_L^{\mathrm{rel}}$ algorithm to the particle-identification–based one, $\max CLL$, is to consider the two algorithms as independently providing two tag decisions with their respective dilutions. It is natural to derive the dilution of the combined tag decision as follows [98]:

$$\mathcal{D}' = \frac{\mathcal{D}_1 + \mathcal{D}_2}{1 + \mathcal{D}_1 \mathcal{D}_2}, \tag{C.1.1}$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ indicate the candidate-by-candidate dilution of the $\max p_L^{\mathrm{rel}}$ and $\max CLL$ tagging algorithms, respectively. The tag decision of the combination corresponds to the decision of the tagger with the larger dilution. The quantity $\mathcal{D}'$ directly represents the true dilution of the event only in the case of two independent taggers. Correlations between the two algorithms which enter the combination in Equation C.1.1 are expected. However, $\mathcal{D}'$ still represents a useful approximation of the true dilution, and is chosen to parametrize the true candidate-by-candidate dilution.

Events have been split in two samples, whether the decisions of the two algorithms in Equation C.1.1 agree or disagree, and in both samples the relation between $\mathcal{D}'$ and the true dilution is polynomial, as shown in Figures C.1. The parameters of the fits

Parameter	Agreement	Disagreement
a_0	-0.009 ± 0.005	0.083 ± 0.026
a_1	-0.079 ± 0.106	0.404 ± 0.139
a_2	1.973 ± 0.481	—
a_3	-1.060 ± 0.515	—

TABLE C.1: Result of the fit for the parameterization of the dilution as a function of $\mathcal{D}' = \frac{\mathcal{D}_1+\mathcal{D}_2}{1+\mathcal{D}_1\mathcal{D}_2}$, where $\mathcal{D}_1$ and $\mathcal{D}_2$ indicate the candidate-by-candidate dilution of the maxp_L^{rel} and maxCLL tagging algorithms, respectively. In the case of disagreement between the decisions of the two taggers, the following definition holds: $\mathcal{D}' = \frac{\mathcal{D}_1-\mathcal{D}_2}{1-\mathcal{D}_1\mathcal{D}_2}, \mathcal{D}_1 > \mathcal{D}_2$.

[%]	0d	0h	0i
ϵ	52.1 ± 0.3	52.2 ± 0.3	52.6 ± 0.3
$S_{\mathcal{D}}$	99.1 ± 2.3	93.5 ± 2.3	93.2 ± 2.3
$S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$	28.9 ± 0.7	27.4 ± 0.7	27.2 ± 0.7

TABLE C.2: Results of the fit on MC events for the scale factor of the same-side tagger which uses the combination of maxp_L^{rel} and maxCLL algorithms, described in Sections 6.4 and 6.5. The tagging algorithm is applied to a PYTHIA-MC sample of $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ candidates.

of the distributions of the true dilution in bins of $\mathcal{D}'$ are shown in Table C.1.

Using the described parameterizations, the tagger is applied to a MC sample, divided in the three periods of data-taking defined in Section 4.3. Following the tagger-calibration procedure described in Section 6.1, a scale factor $S_{\mathcal{D}}$ is calculated for each of the three periods of data-taking. The obtained scale factors are reported in Table C.2. The scale factor is very close to unity in the case of 0d MC events, because that was the sample used to derive the parameterization in Table C.1. The degradation of SST performance in 0h and 0i reflects the decrease in the power of the particle identification described in References [121] and [87]. In particular, a degradation in the performance of the TOF detector is observed, quantified in a $\sim 20\%$ reduction of the arrival time resolution in 0h data with respect to 0d data, and a reduction of the TOF efficiency of about 10%.

The algorithm for same-side tagging presented in this section performs worse than the ANN-based algorithm described in Section 6.6. Quoting from Equation 6.10.2 and Table C.2, the parameterized dilutions $S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$ of the ANN-based and of the CLL-based tagging algorithms are 30.2 ± 0.7 and 28.9 ± 0.7, respectively. These figures are calculated on the same PYTHIA-MC sample of $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ decays. The ANN-based algorithm has thus been favored for this analysis of $B_s^0 - \overline{B}_s^0$ oscillations.

C.1.1 Study of correlations

In this section the correlations between the maxp_L^{rel} and maxCLL algorithms (Sections 6.4 and 6.5) are analyzed in an attempt to understand whether it is actually

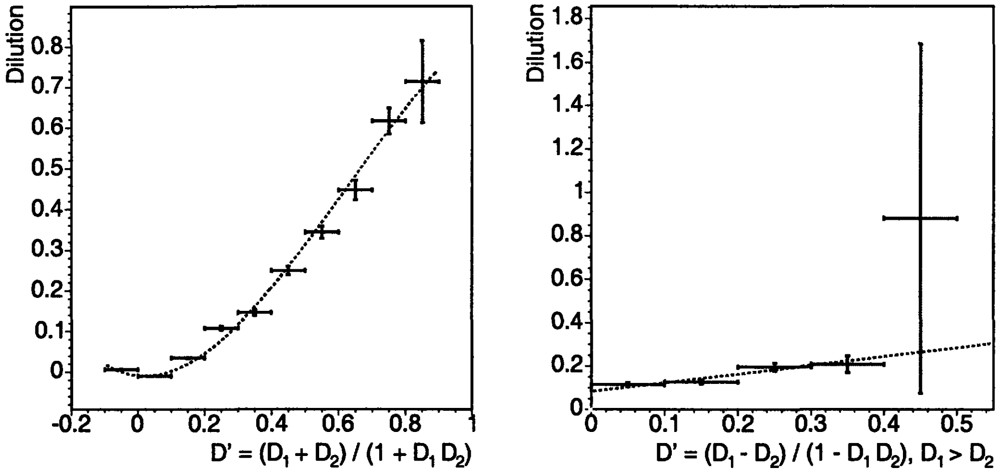

FIGURE C.1: Parameterization for the dilution of the combined max$p_L^{\rm rel}$ and maxCLL algorithms as a function of the raw dilution $\mathcal{D}'$ for cases of agreement and disagreement between the individual decisions of the max$p_L^{\rm rel}$ and maxCLL algorithms, described in Sections 6.4 and 6.5.

possible to produce an algorithm by combining the two algorithms, as described in Equation C.1.1, which performs better than the maxCLL one.

The number of events in which the decision of the max$p_L^{\rm rel}$+maxCLL algorithm is different from the maxCLL-only decision is 1.3%, with a very small uncertainty (the number of tagged MC candidates is about 15 thousand). The effect of combining the taggers thus consists in the increased dilution of the event when the two taggers agree. Events can be thus divided in two samples, according to the agreement between the decisions of the maxCLL and the max$p_L^{\rm rel}$ algorithms, which will be analyzed separately.

The maxCLL and max$p_L^{\rm rel}$ decisions agree in $89.0 \pm 0.3\%$ of the cases in which a tag is assigned to both. The classification of candidates is extended by further subdividing the classes of tag decisions by other characteristics of the candidate. The first classification is based on the agreement between the decisions of maxCLL and max$p_L^{\rm rel}$ algorithms. The first class is further subdivided in three distinct samples:

- the tag candidate track is unique,
- the tag candidate tracks have the same charge,
- the tag candidate tracks do not have the same charge,
    - maxCLL and max$p_L^{\rm rel}$ decision agree,
    - maxCLL and max$p_L^{\rm rel}$ decision disagree.

Tag candidate tracks are defined by the selection cuts presented in Section 6.3. Each of these four classes is finally divided in three subclasses, whether the tag candidate with the maxCLL is strongly identified as a kaon ($CLL > 2$), it is very likely a pion

Class	Cut	Fraction [%]
Unique tag candidate	$CLL > 2$	2.0 ± 0.1
	$CLL < -2$	31.1 ± 0.4
	$\|CLL\| < 1$	11.8 ± 0.3
Tag cands w/ same charge	$CLL > 2$	0.4 ± 0.1
	$CLL < -2$	3.7 ± 0.2
	$\|CLL\| < 1$	3.0 ± 0.2
Tag cands w/ diff. charges max CLL and max p_L^{rel} agree	$CLL > 2$	4.7 ± 0.2
	$CLL < -2$	4.2 ± 0.2
	$\|CLL\| < 1$	8.9 ± 0.3
Tag cands w/ diff. charges max CLL and max p_L^{rel} disagree	$CLL > 2$	4.8 ± 0.1
	$CLL < -2$	1.1 ± 0.1
	$\|CLL\| < 1$	4.4 ± 0.1

TABLE C.3: Fraction of events in the different max CLL and max p_L^{rel} classes defined in the text.

($CLL < -2$) or there is small particle-identification information ($|CLL| < 1$). The distribution of CLL, defined in Equation 6.5.3, for kaons, pions, and protons in a PYTHIA-MC sample, is shown in Figure 6.5. Events with $1 < |CLL| < 2$ are not included in the set of plots in favor of plots which present classes with well defined particle-identification characteristics (kaons, pions and little particle-identification information). Twelve subclasses are thus defined and the scatter plots in Figures C.2-C.5 show interesting correlations between the raw dilution $\mathcal{D}'$, which is almost linearly correlated to the true dilution, and the parameterized dilutions of the max CLL and max p_L^{rel} algorithms. The populations of the various classes are reported in Table C.3.

The largest improvement in candidate-by-candidate dilution over the max CLL-only algorithm is expected in the case of the classes which contain tag candidates with $|\max CLL| < 1$. In fact, when CLL information tags the track as a kaon ($CLL > 2$), max CLL constitutes a very powerful tagger which dominates the combination in Equation C.1.1, as shown by the leftmost plots in Figures C.2-C.5. When CLL information is weak, the max p_L^{rel} algorithm contributes to the total dilution. This behavior is proved by the central plots in the same set of figures. In the last case, when max $CLL < -2$ (rightmost plots in Figures C.2-C.5), the max p_L^{rel} algorithm completely drives the assignment of the event dilution, but the scale factor, in this bin of CLL, is very close to zero, because tag candidates are likely to be pions and thus have no tagging information.

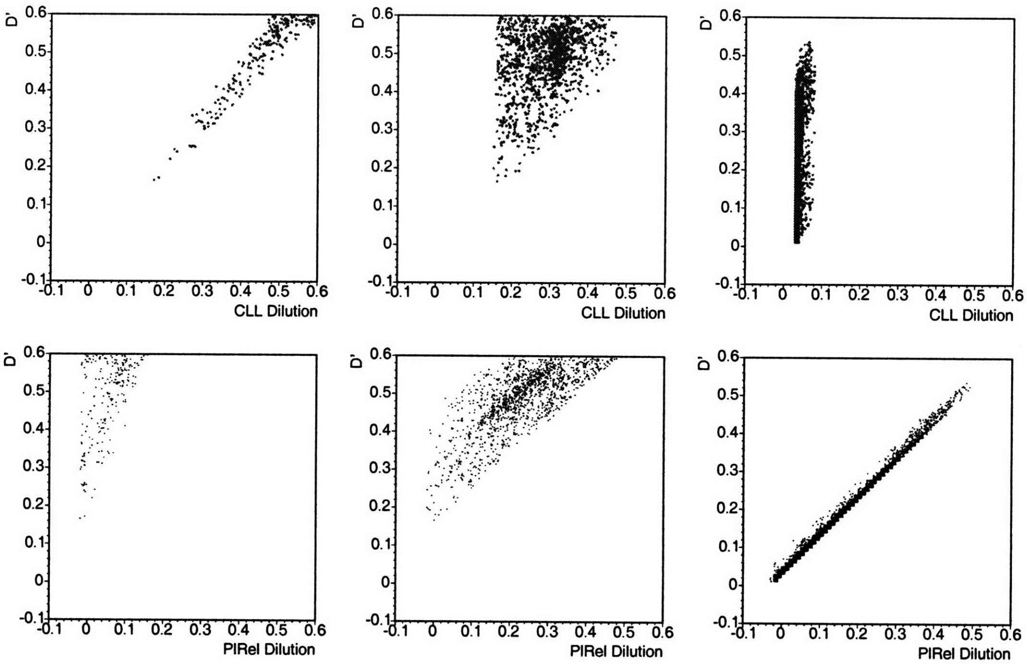

FIGURE C.2: Raw dilution $\mathcal{D}'$ versus true dilution $\mathcal{D}$ of $\max CLL$ (top) and $\max p_L^{\mathrm{rel}}$ (bottom) algorithms, when there is a unique tag candidate track with $CLL > 2$ (left), $|CLL| < 1$ (center), $CLL < -2$ (right). The decisions of $\max CLL$ and $\max p_L^{\mathrm{rel}}$ algorithms agree.

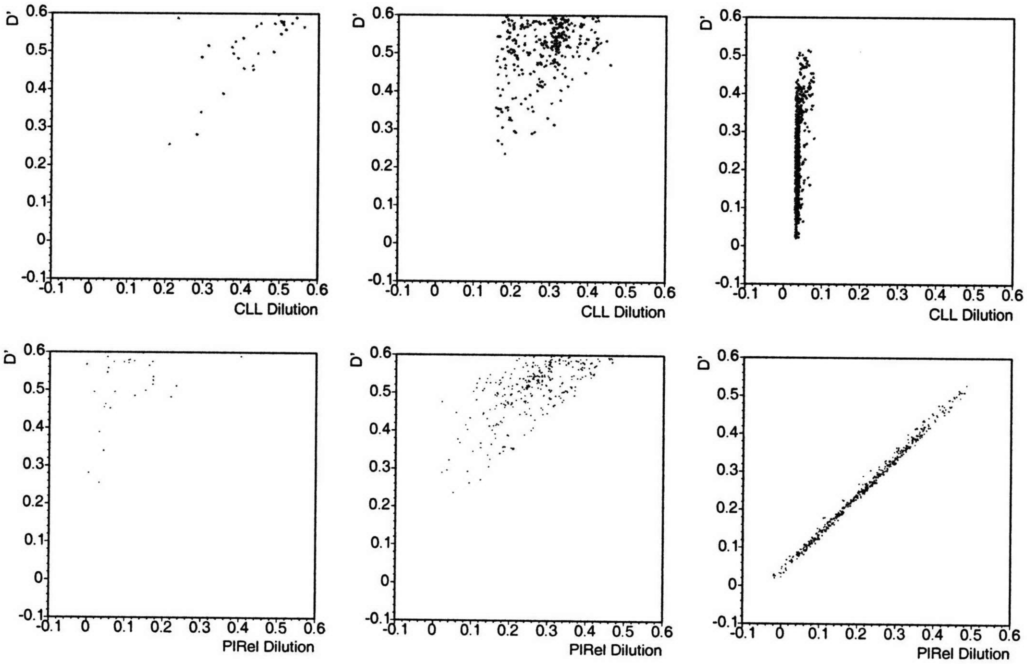

FIGURE C.3: Raw dilution $\mathcal{D}'$ versus true dilution $\mathcal{D}$ of $\max CLL$ (top) and $\max p_L^{\mathrm{rel}}$ (bottom) algorithms, when there are multiple tag candidate tracks with the same charge, and $\max CLL > 2$ (left), $|\max CLL| < 1$ (center), $\max CLL < -2$ (right). The decisions of $\max CLL$ and $\max p_L^{\mathrm{rel}}$ algorithms agree.

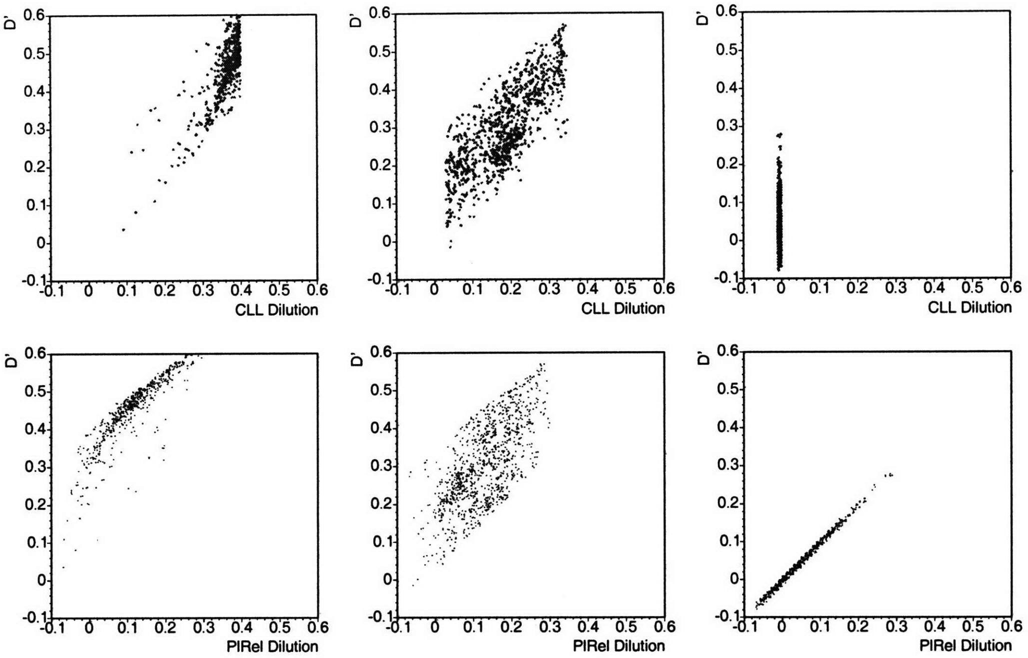

FIGURE C.4: Raw dilution $\mathcal{D}'$ versus true dilution $\mathcal{D}$ of $\max CLL$ (top) and $\max p_L^{\mathrm{rel}}$ (bottom) algorithms, when there are multiple tag candidate tracks with different charges, and $\max CLL > 2$ (left), $|\max CLL| < 1$ (center), $\max CLL < -2$ (right). The decisions of $\max CLL$ and $\max p_L^{\mathrm{rel}}$ algorithms agree.

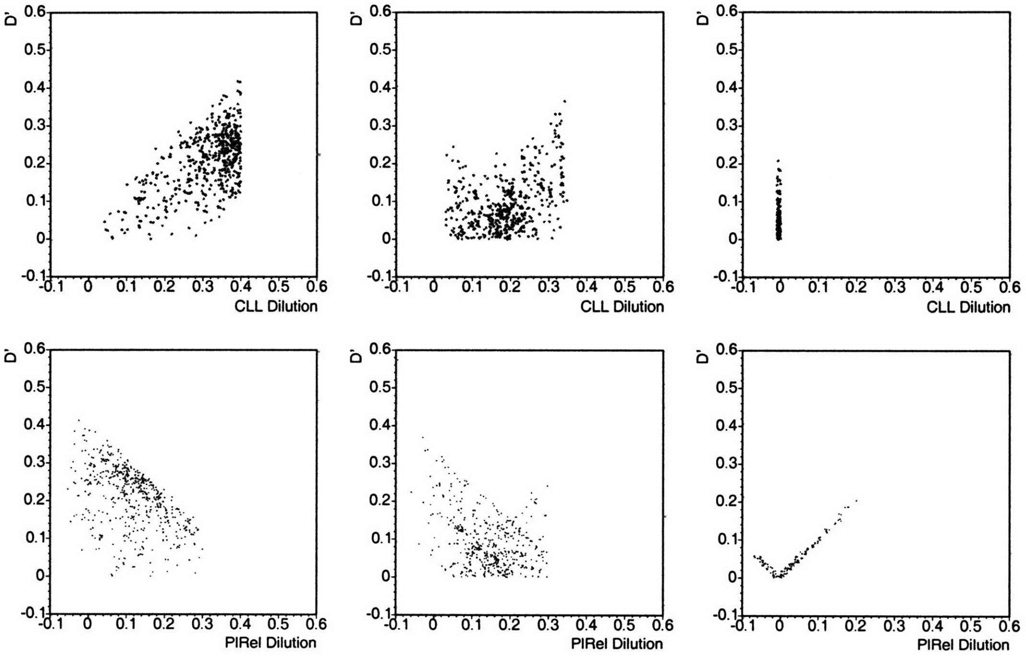

FIGURE C.5: Raw dilution $\mathcal{D}'$ versus true dilution $\mathcal{D}$ of $\max CLL$ (top) and $\max p_L^{\mathrm{rel}}$ (bottom) algorithms, when there are multiple tag candidate tracks with different charges, and $\max CLL > 2$ (left), $|\max CLL| < 1$ (center), $\max CLL < -2$ (right). The decisions of $\max CLL$ and $\max p_L^{\mathrm{rel}}$ algorithms disagree.

[%]	0d
ϵ	52.1 ± 2.8
$S_{\mathcal{D}}$	104.1 ± 2.8
$S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle}$	29.8 ± 0.8

TABLE C.4: Performance of the SST on MC events, upgraded with the parameterization of the dilution in terms of both CLL and p_T of the tag candidate track, in a PYTHIA-MC sample of $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi^0 \pi^-$ decays.

C.2 Parameterization of $\mathcal{D}$ with CLL and p_T

The more traditional way to extract useful information from a MC sample is to introduce a parameterization for candidate-by-candidate dilution $\mathcal{D}$ in two variables, rather than using a single-variable parameterization. The PYTHIA-MC sample of tag candidates for $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi^0 \pi^-$ decays produced for this analysis is divided in bins of CLL and, subsequently, in bins of p_T, because these two variables showed, in previous studies, strong correlations with the candidate-by-candidate dilution. The tag decision is based on the charge of the track with the maximum CLL. Thus, the new tag algorithm is identical to the maxCLL one used before, except for the dilution of the B_s^0 candidate, which is now parameterized in terms of the CLL and the transverse momentum of the tagging track. With respect to the maxCLL-only SST algorithm, the decision is identical but the weight (i.e., the dilution) of the candidate is different, including also information from the kinematic of the tagging track. No additional parameterization in p_T has been introduced in the lowest bin in CLL, $CLL < -2$. For these candidates, the dilution is parameterized only in terms of CLL of the selected tagging track. In fact, the tag track in events of this class is well identified as being a pion and thus have no tagging power. This assumption has been tested by fitting for the scale factor of the SST algorithm in the set of events with max$CLL < -1$ and finding the scale factor consistent with zero ($S_{\mathcal{D}} = 3.0 \pm 1.2\%$). The results of the parameterizations are shown in Figure C.6. Events are divided in two major classes: whether all the tag candidate tracks have the same charge or not. In the former case, no decision has to be made, while in the latter the tag decision corresponds to the charge of the track with the maximum CLL. The resulting scale factor and the performance of the tagger based on the CLL and p_T of the tag candidate track are reported in Table C.4.

Despite offering the best performance among the algorithm presented in this appendix (the results for $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle}$ of the three algorithms are presented in Tables C.2, C.4, and C.6), it has been decided not to utilize the combined CLL–p_T parameterization in the upgrade of the SSKT. The statistics of the MC sample, while large enough to provide an accurate parameterization of the dilution in terms of a single variable (maxCLL or p_T of the track with the maximum p_L^{rel} in the previous study), is not sufficient to derive a robust and stable parameterization in two variables, as shown by some of the plots in Figure C.6.

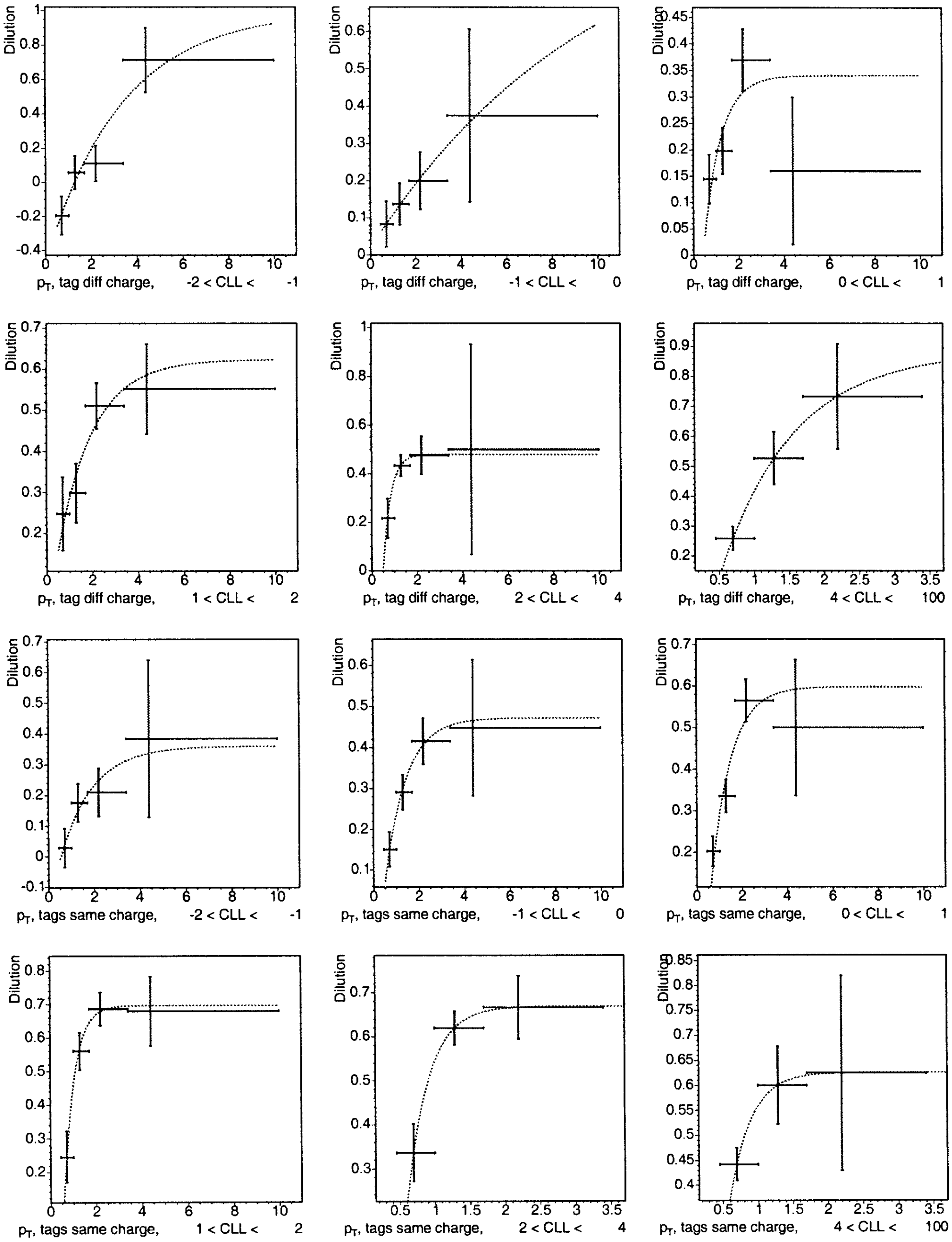

FIGURE C.6: Dilution versus p_T in bins of CLL. The decisions of the tagger is indicated by the charge of the track with $\max CLL$.

Parameter	B^+	B^0	B^0_s
a	1.631	1.642	1.385
b	1.499	1.813	1.870
c [(GeV/c)$^{-1}$]	0.801	1.011	1.496

TABLE C.5: Parameters used to describe p_T-dependent dilution correction factor.

C.3 $CLL + p_T$ Parameterization

The idea developed in this section is to use a parametrization of tagging track CLL and of tagging track p_T for the predicted event-by-event dilution. This new algorithm keeps the tagging decision according to the maxCLL algorithm, but the parameterization of the predicted dilution is improved.

By using max$p_L^{\rm rel}$ to select the tagging track, a dependence of the dilution on the track p_T is found, as seen in Figure C.7. This has been exploited already, as summarized in Section 6.4, to improve the tagging performance of this algorithm. The functional form used to describe the shape of the p_T dependence of the dilution has been parametrized as:

$$\mathcal{D}(p_T) = \alpha - \beta e^{-\gamma p_T} \,. \tag{C.3.1}$$

This expression is found to appropriately model the distribution of average dilution in bins of p_T of tag candidates. As expected, a similar dependency of the dilution as a function of tagging track p_T is seen when maxCLL is utilized to select the tagging track (Figure C.8).

Unfortunately, the statistics of the available MC sample was too small to derive reliable dilution parameterizations as a function of p_T in different bins of CLL. This limitation is dictated by the large computing power required to produce PYTHIA-MC events. Therefore, the unbinned maximum-likelihood fitter was utilized to determine a unique overall p_T-dependent term that can be interpreted as a correction function to be applied to the original candidate-by-candidate dilution returned by the maxCLL algorithm, which is parameterized with a function of CLL only. The predicted dilution is thus described as:

$$\begin{aligned} \mathcal{D}(CLL, p_T) &= \mathcal{D}(CLL) \cdot \mathcal{D}(p_T) \,, \\ \mathcal{D}(CLL) &= \alpha + \frac{1}{2}\beta \cdot e^{\frac{\gamma^2}{2\beta^2} - \frac{x}{\beta}} \cdot \left[1 - erf\left(\frac{1}{\sqrt{2}}\left(\frac{\beta}{\gamma} - \frac{x}{\beta}\right)\right)\right] \,, \quad x = \delta - CLL \,, \\ \mathcal{D}(p_T) &= a - b e^{-c p_T} \,. \end{aligned} \tag{C.3.2}$$

$\mathcal{D}(CLL)$ is the very same parameterization which has already been derived in the previous study and shown in Figure 6.7. The parameters a, b and c are the same for all events. No separation between events with agreeing or disagreeing charges of tagging track candidates is made. The parameters a, b, and c, found by fitting the PYTHIA-MC sample of $B^0_s \to D^-_s \pi^+, D^-_s \to \phi^0 \pi^-$ decays in the 0d configuration, are listed in Table C.5. The correction functions obtained fitting PYTHIA-MC samples of $B^0 \to D^- \pi^+$, $B^+ \to \overline{D}^0 \pi^+$, and $B^0_s \to D^-_s \pi^+$ decays, are displayed in Figure C.9.

	[%]	B^+	B^0	B^0_s
MC (`0d`)	$S_{\mathcal{D}}$	100.0± 0.9	100.0± 1.9	100.0± 2.4
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	28.3± 0.2	18.2± 0.4	29.3± 0.8
data (`0d`)	$S_{\mathcal{D}}$	100.1± 4.9	102.7± 15.1	—
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	29.3± 1.3	18.6± 2.7	—
MC (`0h`)	$S_{\mathcal{D}}$	97.8± 0.9	98.8± 1.9	96.8± 2.4
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	27.9± 0.3	18.0± 0.4	28.4± 0.7
data (`0h`)	$S_{\mathcal{D}}$	99.2± 4.2	97.5± 13.3	—
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	29.0± 1.1	17.8± 2.4	—
MC (`0i`)	$S_{\mathcal{D}}$	98.4± 0.9	96.1± 1.9	97.9± 2.4
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	28.0± 0.2	17.5± 0.4	28.6± 0.7
data (`0i`)	$S_{\mathcal{D}}$	93.7± 6.6	79.2± 24.0	—
	$S_{\mathcal{D}}\sqrt{<\mathcal{D}>^2}$	26.5± 1.7	14.3± 4.3	—

TABLE C.6: Performance of parameterized maxCLL+p_T algorithm in data and MC.

The tagging performances, as measured in data and calculated in MC samples, using this p_T-dependent correction factor for the predicted dilution are listed in Table C.6. As the parameterization has been derived using the unbinned fitter, the dilution scale factor has to be 100% by definition for the `0d` MC sample. Good data–MC-simulation agreement between the scale factors and effective dilutions measured in data and MC samples of $B^+ \to \overline{D}^0\pi^+, \overline{D}^0 \to K^+\pi^-$ and $B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$ decays is shown in Table C.6.

For a more complete comparison, the results of the maxCLL algorithm only applied on the $B^+ \to \overline{D}^0\pi^+, \overline{D}^0 \to K^+\pi^-$ and $B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$ data and MC samples are reported in Table C.7. The comparison of Table C.6 and Tab. C.7 shows that the improvement provided by the introduction of a p_T-dependent correction function, compared with the use of a CLL-only parameterization, which is observed in $B^+ \to \overline{D}^0\pi^+$ and $B^0 \to D^-\pi^+$ decays in PYTHIA-MC samples is as well confirmed in data.

An absolute gain in dilution between 0.8 and 1.1% on the B^0_s MC sample is observed. This transform to a relative gain of 5-8% in $\epsilon\mathcal{D}^2$ (Table C.8), depending on the sample (`0d`, `0h` or `0i`). This improvement over the original maxCLL algorithm is smaller than the one provided by the ANN-based algorithm described in Section 6.6, which has been finally chose for this analysis of $B^0_s - \overline{B}^0_s$ oscillations.

In order to check that the improvement found on the $B^0_s \to D^-_s\pi^+, D^-_s \to \phi^0\pi^-$ PYTHIA-MC sample is properly estimated and does not come from over-tuning on the MC sample, the following test has been performed. The sample is split up in two halves, a dilution correction function is derived on one half with the following parametrization:

$$\mathcal{D}(p_T) = 1.405 - 1.611e^{-1.432*p_T}\,. \qquad \text{(C.3.3)}$$

[%]		B^+	B^0	B_s^0
MC (0d)	$S_\mathcal{D}$	100.0± 0.9	98.5± 2.1	98.8± 2.5
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	27.5± 0.2	17.4± 0.4	28.5± 0.7
data (0d)	$S_\mathcal{D}$	98.5± 5.1	101.0± 15.7	—
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	27.8± 1.3	18.3 ± 2.8	—
MC (0h)	$S_\mathcal{D}$	96.1± 0.9	95.3± 2.1	94.4± 2.5
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	27.2± 0.3	16.6± 0.4	27.3± 0.7
data (0h)	$S_\mathcal{D}$	95.6± 4.4	93.1± 13.9	—
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	27.4± 1.2	16.8± 2.5	—
MC (0i)	$S_\mathcal{D}$	97.1± 0.9	96.1± 2.1	96.4± 2.5
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	27.2± 0.2	17.4± 0.4	27.9± 0.7
data (0i)	$S_\mathcal{D}$	94.5± 6.7	76.2± 24.3	—
	$S_\mathcal{D}\sqrt{<\mathcal{D}>^2}$	26.3± 2.0	13.7± 4.3	—

TABLE C.7: Performance of parameterized maxCLL algorithm in data and MC.

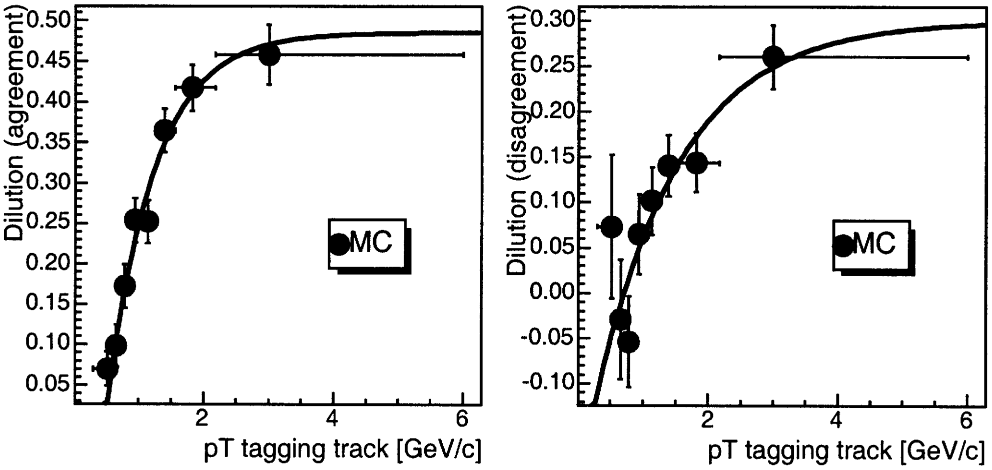

FIGURE C.7: Parameterization for the dilution of maxp_L^{rel} algorithm as a function of the p_T of the tagging track for cases of agreeing tagging candidate charges (left) and disagreeing ones (right), for $B_s^0 \to D_s^-\pi^+, D_s^- \to \phi^0\pi^-$ candidates.

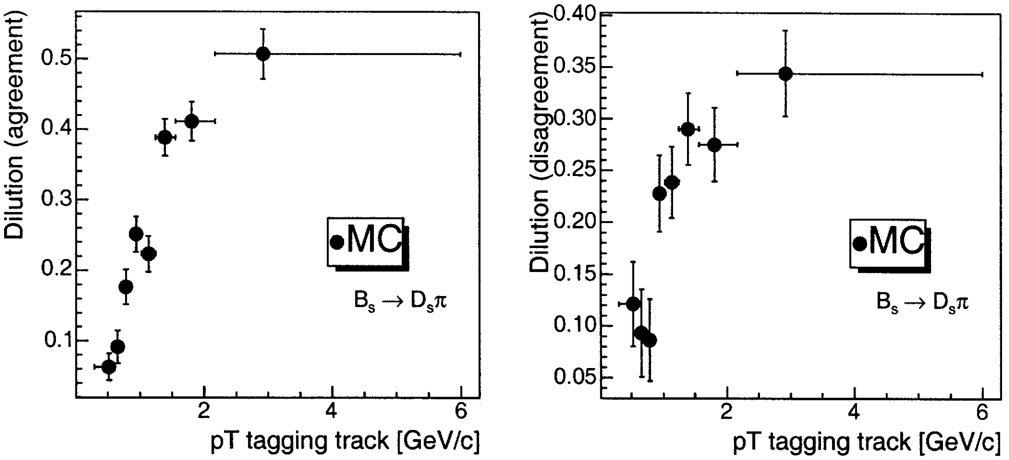

FIGURE C.8: Distribution of the dilution of maxCLL algorithm as a function of the p_T of the tagging track for cases of agreeing tagging candidate charges (left) and disagreeing ones (right) for $\mathrm{B}_s^0 \to \mathrm{D}_s^-\pi^+, \mathrm{D}_s^- \to \phi^0\pi^-$ candidates.

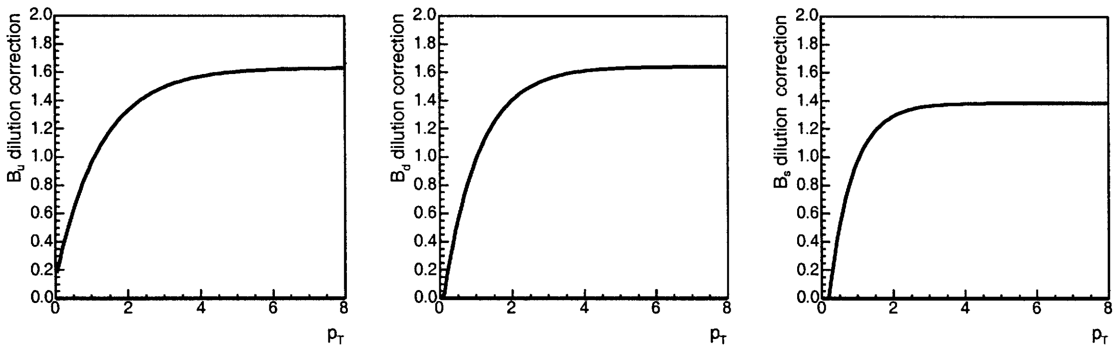

FIGURE C.9: Correction function for the predicted dilution as a function of the p_T of the tagging track for B^+ (left), B^0 (center) and B_s^0 (right). The parameters have been derived fitting `0d`-like MC samples.

This parametrization yields an effective dilution of $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 30.3 \pm 0.7\%$ on the subsample it has been derived on, and $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 28.3 \pm 0.7\%$ on the other one. The effective dilution on those two subsamples without the correction factor was $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 29.7 \pm 0.7\%$ and $S_{\mathcal{D}}\sqrt{\langle \mathcal{D}^2 \rangle} = 27.2 \pm 0.7\%$, respectively. The gain is as well present in the control sample, thus proving that introducing the tagging track p_T dependent dilution correction gives a small but real improvement.

Using the predicted dilution $\sqrt{\langle \mathcal{D}^2 \rangle}$ and the efficiency ϵ from the $\mathrm{B}_s^0 \to \mathrm{D}_s^-\pi^+, \mathrm{D}_s^- \to \phi^0\pi^-$ data sample, and the scale factor $S_{\mathcal{D}}$ from the MC study, the performances reported in Table C.8 are found.

[%]	maxCLL algo	maxCLL + p_T algo
`0d` data $S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$	28.3 ± 2.4	29.2 ± 2.4
`0h` data $S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$	24.2 ± 1.9	25.6 ± 1.9
`0i` data $S_{\mathcal{D}}\sqrt{\langle\mathcal{D}^2\rangle}$	25.8 ± 2.8	26.1 ± 2.8
`0d` data $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	4.0 ± 0.7	4.2 ± 0.7
`0h` data $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	2.9 ± 0.5	3.2 ± 0.5
`0i` data $\epsilon S_{\mathcal{D}}^2\langle\mathcal{D}^2\rangle$	3.3 ± 0.7	3.4 ± 0.7

TABLE C.8: Performance of maxCLL and maxCLL + p_T algorithms on B_s^0 data. Statistical errors only are quoted.

Bibliography

[1] ALBAJAR, C. AND OTHERS. *Search for B0 anti-B0 Oscillations at the CERN Proton - anti-Proton Collider. 2.* Phys. Lett., **B186:247**, 1987.

[2] ALBRECHT, H. AND OTHERS. *Observation of B0 - anti-B0 Mixing.* Phys. Lett., **B192:245**, 1987.

[3] ARTUSO, M. AND OTHERS. *B0 anti-B0 Mixing at the Upsilon (4S).* Phys. Rev. Lett., **62:2233**, 1989.

[4] ELLIS, JOHN R. AND HAGELIN, J. S. AND RUDAZ, S. *Reexamination of the Standard Model in the Light of B Meson Mixing.* Phys. Lett., **B192:201**, 1987.

[5] ALTARELLI, GUIDO AND FRANZINI, PAULA J. *B0 - anti-B0 Mixing Within and Beyond the Standard Model.* Z. Phys., **C37:271**, 1988.

[6] BUSKULIC, D. AND OTHERS. *Observation of the time dependence of B(d)0 - anti-B(d)0 mixing.* Phys. Lett., **B313:498–508**, 1993.

[7] YAO, W. -M. AND OTHERS. *Review of particle physics.* J. Phys., **G33:1–1232**, 2006.

[8] CABIBBO, N. *Unitary Symmetry and Leptonic Decays.* Phys. Rev. Lett., **10:531–532**, 1963.

[9] KOBAYASHI, MAKOTO AND MASKAWA, TOSHIHIDE. *CP Violation in the Renormalizable Theory of Weak Interaction.* Prog. Theor. Phys., **49:652–657**, 1973.

[10] WOLFENSTEIN, LINCOLN. *Parametrization of the Kobayashi-Maskawa Matrix.* Phys. Rev. Lett., **51:1945**, 1983.

[11] WEISSKOPF, V. AND WIGNER, EUGENE P. *Calculation of the natural brightness of spectral lines on the basis of Dirac's theory.* Z. Phys., **63:54–73**, 1930.

[12] WEISSKOPF, V. AND WIGNER, E. *Over the natural line width in the radiation of the harmonius oscillator.* Z. Phys., **65:18–29**, 1930.

[13] 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.

[14] OKAMOTO, MASATAKA. *Full determination of the CKM matrix using recent results from lattice QCD.* PoS, **LAT2005:013**, hep-lat/0510113, 2006.

[15] HEP2005 EPS CONFERENCE WEBPAGE. `http://www.lip.pt/events/2005/hep2005/`.

[16] 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.

[17] 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.

[18] WILSON, KENNETH G. *Nonlagrangian models of current algebra.* Phys. Rev., **179:1499–1512**, 1969.

[19] BURAS, A. J. AND GAMBINO, P. AND GORBAHN, M. AND JAGER, S. AND SILVESTRINI, L. *Universal unitarity triangle and physics beyond the standard model.* Phys. Lett., **B500:161–167**, hep-ph/0007085, 2001.

[20] CIUCHINI, MARCO AND DEGRASSI, G. AND GAMBINO, P. AND GIUDICE, G. F. *Next-to-leading QCD corrections to B -> X/s gamma in supersymmetry.* Nucl. Phys., **B534:3–20**, hep-ph/9806308, 1998.

[21] BURAS, A. J. AND GAMBINO, P. AND GORBAHN, M. AND JAGER, S. AND SILVESTRINI, L. *ϵ'/ϵ and rare K and B decays in the MSSM.* Nucl. Phys., **B592:55–91**, hep-ph/0007313, 2001.

[22] EYAL, GALIT AND NIR, YOSEF. *Constraining extensions of the quark sector with the CP asymmetry in B -> psi K(S).* JHEP, **09:013**, hep-ph/9908296, 1999.

[23] BURAS, ANDRZEJ J. AND CHANKOWSKI, PIOTR H. AND ROSIEK, JANUSZ AND SLAWIANOWSKA, LUCJA. *Delta(M(s))/Delta(M(d)), sin 2beta and the angle gamma in the presence of new Delta(F) = 2 operators.* Nucl. Phys., **B619:434–466**, hep-ph/0107048, 2001.

[24] ACOSTA, D. AND OTHERS. *Measurement of the J/ψ meson and $b-$hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s}$ = 1960 GeV.* Phys. Rev., **D71:032001**, hep-ex/0412071, 2005.

[25] *Fermilab Linac upgrade conceptual design revision 4A.* FERMILAB-LU-CONCEPTUAL-DESIGN.

[26] POPOVIC, M. AND AKENBRANDT, C. *Performance and measurements of the Fermilab booster.* FERMILAB-CONF-98-175.

[27] ASHER, T. *The Main Injector Rookie Book.* 2003.

[28] Morgan, J. *The Antiproton Source Rookie Book.* 1999.

[29] Marriner, John. *Stochastic cooling overview.* Nucl. Instrum. Meth., **A532:11–18**, physics/0308044, 2004.

[30] Jackson, Gerry. *The Fermilab recycler ring technical design report. Rev. 1.2.* FERMILAB-TM-1991.

[31] Group, TeVI. *Design Report Tevatron 1 project.* FERMILAB-DESIGN-1984-01.

[32] Blair, R. and others. *The CDF-II detector: Technical design report.* FERMILAB-PUB-96-390-E.

[33] Abe, F. and others. *The CDF detector: an overview.* Nucl. Instr. Meth., **A271:387–403**, 1988.

[34] Hill, Christopher S. *Operational experience and performance of the CDFII silicon detector.* Nucl. Instrum. Meth., **A530:1–6**, 2004.

[35] Azzi, P. *The CDF silicon detector upgrade.* Nucl. Instrum. Meth., **A419:532–537**, 1998.

[36] Sill, A. *CDF Run II silicon tracking projects.* Nucl. Instrum. Meth., **A447:1–8**, 2000.

[37] Affolder, Anthony A. and others. *Intermediate silicon layers detector for the CDF experiment.* Nucl. Instrum. Meth., **A453:84–88**, 2000.

[38] Affolder, Anthony A. and others. *CDF central outer tracker.* Nucl. Instrum. Meth., **A526:249–299**, 2004.

[39] 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.

[40] Tonelli, Diego. *First observation of the B/s0 -> K+ K- decay mode, and measurement of the B0 and B/s0 mesons decay-rates into two-body charmless final states at CDF.* FERMILAB-THESIS-2006-23.

[41] Acosta, D. and others. *A time-of-flight detector in CDF-II.* Nucl. Instrum. Meth., **A518:605–608**, 2004.

[42] Balka, L. and others. *The CDF Central Electromagnetic Calorimeter.* Nucl. Instrum. Meth., **A267:272**, 1988.

[43] 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.

[44] BERTOLUCCI, S. AND OTHERS. *The CDF Central and Endwall Hadron Calorimeter.* Nucl. Instrum. Meth., **A267:301**, 1988.

[45] ALBROW, M. G. AND OTHERS. *The CDF plug upgrade electromagnetic calorimeter: Test beam results.* Nucl. Instrum. Meth., **A480:524–546**, 2002.

[46] 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.

[47] ASCOLI, G. AND OTHERS. *CDF Central Muon Detector.* Nucl. Instrum. Meth., **A268:33**, 1988.

[48] FERMILAB AND OTHERS. *The 1992 CDF Muon System Upgrade.* 1994. CDF2858 public note.

[49] GINSBURG, C. M. *CDF Run 2 muon system.* Eur. Phys. J., **C33:s1002–s1004**, 2004.

[50] THOMSON, EVELYN J. AND OTHERS. *Online track processor for the CDF upgrade.* IEEE Trans. Nucl. Sci., **49:1063–1070**, 2002.

[51] ASHMANSKAS, BILL AND OTHERS. *The CDF silicon vertex trigger.* Nucl. Instrum. Meth., **A518:532–536**, physics/0306169, 2004.

[52] DELL'ORSO, M. *The CDF Silicon Vertex Trigger.* Nucl. Phys. Proc. Suppl., **156:139–142**, 2006.

[53] 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.

[54] GOMEZ-CEBALLOS, G. AND OTHERS. *Event builder and Level 3 at the CDF experiment.* Nucl. Instrum. Meth., **A518:522–524**, 2004.

[55] 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.

[56] TIWARI, VIVEK. *Measurement of the B_s $\bar{B}_s$ oscillation frequency using semileptonic decays.* FERMILAB-THESIS-2007-09.

[57] PROUDFOOT, JAMES. *Electron Identification in the CDF Central Calorimeter.* For Proc. of Workshop on Calorimetry for the Superconducting Supercollider, Tuscaloosa, Alabama, Mar 13- 17, 1989.

[58] PAUS, CH. AND OTHERS. *BottomMods.* `http://cdfcodebrowser.fnal.gov/CdfCode/source/BottomMods/`.

[59] Murat, P. and others. *Stntuple.* `http://cdfcodebrowser.fnal.gov/CdfCode/source/Stntuple/`.

[60] Brun, R. and others. *Root – an object-oriented data analysis framework.* `http://root.cern.ch`.

[61] Marriner, J. *Secondary Vertex Fit with mass and pointing constraints (CTVMFT).* 1993. CDF1996 public note.

[62] Eidelman, S. and others. *Review of particle physics.* Phys. Lett., **B592:1**, 2004.

[63] 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.

[64] Anikeev, K. and Murat, P. and Paus, Ch. *Description of Bgenerator II.* 1999. CDF5092 internal note.

[65] Sjostrand, Torbjorn and Lonnblad, Leif and Mrenna, Stephen. *PYTHIA 6.2: Physics and manual.* hep-ph/0108264, 2001.

[66] Yu, Shin-shan. *First measurement of the ratio of branching fractions* $B(\Lambda_b^0 \to \Lambda_c^+\mu^-\bar{\nu}_\mu)/B(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ *at CDF II.* hep-ex/0504059, 2005.

[67] Andersson, Bo and Gustafson, G. and Ingelman, G. and Sjostrand, T. *Parton Fragmentation and String Dynamics.* Phys. Rept., **97:31**, 1983.

[68] Andersson, Bo. *The Lund model.* Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol., **7:1–471**, 1997.

[69] Hoyer, P. and Osland, P. and Sander, H. G. and Walsh, T. F. and Zerwas, P. M. *Quantum Chromodynamics and Jets in* e^+e^-. Nucl. Phys., **B161:349**, 1979.

[70] Ali, Ahmed and Pietarinen, E. and Kramer, G. and Willrodt, J. *A QCD Analysis of the High-Energy* e^+e^- *Data from PETRA.* Phys. Lett., **B93:155**, 1980.

[71] Marchesini, G. and Webber, B. R. *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation.* Nucl. Phys., **B310:461**, 1988.

[72] Affolder, Anthony A. and others. *Charged jet evolution and the underlying event in* $p\bar{p}$ *collisions at* 1.8 TeV. Phys. Rev., **D65:092002**, 2002.

[73] Field, Rick. *CDF run II Monte-Carlo tunes.* To appear in the proceedings of TeV4LHC 2006 Workshop 4th meeting, Batavia, Illinois, 20-22 Oct 2006.

[74] Lange, D. J. *The EvtGen particle decay simulation package.* Nucl. Instrum. Meth., **A462:152–155**, 2001.

[75] Brun, R. and Bruyant, F. and Maire, M. and McPherson, A. C. and Zanarini, P. *GEANT3.* CERN-DD/EE/84-1.

[76] Field, R. D. *The sources of b quarks at the Tevatron and their correlations.* Phys. Rev., **D65:094006**, hep-ph/0201112, 2002.

[77] Abulencia, A. and others. *Measurement of the ratios of branching fractions* $B(B_s^0 \to D_s^- \pi^+)/B(B^0 \to D^- \pi^+)$ *and* $B(B^+ \to \overline{D}^0 \pi^+)/B(B^0 \to D^- \pi^+)$. Phys. Rev. Lett., **96:191801**, hep-ex/0508014, 2006.

[78] Furic, Ivan K. *Measurement of the Ratio of Branching Fractions* $Br(B_s^0 \to D_s^- \pi^+)/Br(B^0 \to D^- \pi^+)$. 2004. CDF7352 public note.

[79] Lai, S. and others. *An Updated Measurement of the Beam Width at CDF.* 2003. CDF6492 internal note.

[80] Rappoccio, Salvatore and others. *First Look at SecVtx Using Event Primary Vertex Finder PrimeVtx.* 2003. CDF6417 public note.

[81] Menzemer, Stephanie. *L00 Hit-Resolution/Efficiency in MC/Data.* July 2005. Talk at BPAK Meeting.

[82] Herndon, M. private communication, 2005.

[83] Menzemer, Stephanie. *L00(L0) Hit-Resolution/Efficiency in MC/Data.* July 2005. Talk at Tracking Meeting.

[84] CDF Tracking Group Webpage. `http://www-cdf.fnal.gov/upgrades/computing/projects/reconstruction/tracking/user-docs/matching.html`.

[85] Bethe, H. *Theory of the passage of fast corpuscular rays through matter.* Annalen Phys., **5:325–400**, 1930.

[86] Punzi, Giovanni and Squillacioti, Paola. *Particle identification by combining TOF and dE/dx information.* 2005. CDF7488 internal note.

[87] Giagu, Stefano and others. *dEdx, TOF validations studies on 0h/0i data for Bs mixing analyses.* 2006. CDF8169 internal note.

[88] Gronau, Michael and Nippe, Alex and Rosner, Jonathan L. *Method for flavor tagging in neutral B meson decays.* Phys. Rev., **D47:1988–1993**, hep-ph/9211311, 1993.

[89] Albrecht, Z. and others. *A study of excited b-hadron states with the DELPHI detector at LEP.* Prepared for International Europhysics Conference on High- Energy Physics (HEP 2003), Aachen, Germany, 17-23 Jul 2003.

[90] Andersson, Bo and Gustafson, G. and Soderberg, B. *A General Model for Jet Fragmentation.* Z. Phys., **C20:317**, 1983.

[91] Peterson, C. and Schlatter, D. and Schmitt, I. and Zerwas, Peter M. *Scaling Violations in Inclusive e^+e^- Annihilation Spectra.* Phys. Rev., **D27:105**, 1983.

[92] Heister, A. and others. *Study of the fragmentation of b quarks into B mesons at the Z peak.* Phys. Lett., **B512:30–48**, hep-ex/0106051, 2001.

[93] Ben-Haim, E. and Bambade, P. and Roudeau, P. and Savoy-Navarro, A. and Stocchi, A. *Extraction of the x-dependence of the non-perturbative QCD b-quark fragmentation distribution component.* Phys. Lett., **B580:108–118**, hep-ph/0302157, 2004.

[94] Moser, H. G. and Roussarie, A. *Mathematical methods for B0 anti-B0 oscillation analyses.* Nucl. Instrum. Meth., **A384:491–505**, 1997.

[95] 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.

[96] Miles, Jeffrey and others. *Scale Factors for Proper Time Uncertainties at CDF.* 2005. CDF7944 public note.

[97] Miles, J. PhD thesis, Massachusetts Institute of Technology, 2007. thesis in preparation.

[98] Sphicas, P. *Combining Flavor Taggers.* 1995. CDF3425 internal note.

[99] 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.

[100] Salamanna, G. and others. *Opposite Side Kaon Tagger.* 2006. CDF8241 public note.

[101] 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.

[102] Khur, T. and others. *Combined Opposite Side Flavor Tagger.* 2006. CDF8460 public note.

[103] The CDF Collaboration. *Updated Measurement of Δm_d and Calibration of Flavor Tagging in Fully Reconstructed Decays.* 2005. CDF7920 public note.

[104] Gronau, Michael and Rosner, Jonathan L. *Identification of neutral B mesons using correlated hadrons.* Phys. Rev., **D49:254–264**, hep-ph/9308371, 1994.

[105] Abe, F. and others. *Measurement of the $B_d^0-\overline{B}_d^0$ flavor oscillation frequency and study of same side flavor tagging of B mesons in $p\bar{p}$ collisions.* Phys. Rev., **D59:032001**, hep-ex/9806026, 1999.

[106] Leonardo, Nuno T. *Analysis of B/s flavor oscillations at CDF.* FERMILAB-THESIS-2006-18.

[107] Rumelhart, David E. and Hinton, Geoffrey E. and Williams, Ronald J. *Learning Internal Representations by Error Propagation*, volume 1. MIT Press, 1986.

[108] Menzemer, S. and others. *Same Side Kaon Tagging Studies in Fully Reconstructed Decays.* 2006. CDF7979 internal note.

[109] Squillacioti, Paola. *Measurement of the branching fraction ratio* $BR(B^+ \to \bar{D}^0K^+ \to [K^+pi^-]K^+)/BR(B^+ \to \bar{D}^0\pi^+ \to [K^+\pi^-]\pi^+)$ *with the CDF II detector.* FERMILAB-THESIS-2006-27.

[110] Giagu, S. private communication, 2006.

[111] Lyons, L. *Statistics for Particle and Nuclear Physicists.* Cambridge University Press, 1986.

[112] Punzi, Giovanni. *Comments on likelihood fits with variable resolution.* physics/0401045, 2004.

[113] Anikeev, K. and others. *B physics at the Tevatron: Run II and beyond.* hep-ph/0201071, 2001.

[114] Abulencia, A. and others. *Measurement of* $\sigma_{\Lambda_b^0}/\sigma_{B^0} \times \mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\pi^-)/\mathcal{B}(B^0 \to D^-\pi^+)$ *in* $p\bar{p}$ *Collisions at* $\sqrt{s} = 1.96$ TeV. Phys. Rev. Lett., **98:122002**, 2007.

[115] Abulencia, A. and others. *Measurement of the Ratios of Branching Fractions* $B(B_s \to D_s\pi\pi\pi)/B(B_d \to D_d\pi\pi\pi)$ *and* $B(B_s \to D_s\pi)/B(B_d \to D_d\pi)$. Phys. Rev. Lett., **98:061802**, hep-ex/0610045, 2007.

[116] Anikeev, Konstantin. *Measurement of the lifetimes of B meson mass eigenstates.* FERMILAB-THESIS-2004-12.

[117] 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.

[118] 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.

[119] 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.

[120] Ligeti, Zoltan and Papucci, Michele and Perez, Gilad. *Implications of the measurement of the B/s0 - anti-B/s0 mass difference.* Phys. Rev. Lett, **97:101801**, hep-ph/0604112, 2006.

[121] MENZEMER, S. AND OTHERS. *Same side Kaon Tagging Studies in Fully Reconstructed Decays on the New Data.* 2006. CDF8182 internal note.