Università degli Studi di Roma "La Sapienza"

Facoltà di Scienze Matematiche, Fisiche e Naturali

Dottorato di Ricerca in Fisica

XIX Ciclo

First observation of B_s mixing at the CDF II experiment with a newly developed Opposite Side b-flavour tagging algorithm using Kaons

Presented by Giuseppe Salamanna

Submitted in partial fulfillment of the requirements
for the degree of *Dottore di Ricerca*

Advisors:

Prof. Carlo Dionisi

Dott. Marco Rescigno

Varius, multiplex, multiformis...

Contents

Introduction

This thesis describes the development, calibration and performance evaluation of an Opposite-side b flavour tagger using K mesons at a $p\bar{p}$ hadron collider and the first observation of flavour oscillations in the B_s meson system at the Collider Detector at Fermilab (CDF). I report the value of the oscillation frequency Δm_s, measured using the Kaon tagger in conjunction with the other CDF flavour taggers to achieve the desired statistical significance. The work is performed using data collected by CDF during the Run II of the Tevatron hadron collider running at $\sqrt{s} = 1.96\ TeV$ at the Fermilab.

The measurement of Δm_s is one of the main goals of the physics programme of the Tevatron Run II. It has been long expected for in the past years in the context of flavour physics and is one of the most important experimental results recently achieved. Along with the precise knowledge of its equivalent in the B_d^0 sector, Δm_d, this measurement provides a very powerful constraint to the Cabibbo-Kobayashi-Maskawa (CKM) matrix governing quark coupling in weak interactions. The value of Δm_s in the Standard Model of particles is expected to be consistently larger than for Δm_d. Moreover, since new particles are predicted by *Beyond-the-Standard-Model* (BSM) scenarios to contribute to the flavour mixing loop diagram, an even larger value (or in general any different from SM predictions) would give a hint of new physics processes.

Provided the rapidity with which $B_s^0 - \overline{B}_s^0$ oscillations occur in the time domain, the experimental determination of Δm_s is a real challenge. A very good time resolution is needed to separate the various oscillation periods.

The $p\bar{p}$ Tevatron collider has a leading role in studying B_s mixing. In particular, the fact that CDF can select fully reconstructed hadronic decays of B_s mesons makes it a favourable experiment to detect the oscillations. The complete reconstruction of the initial kinematics in the transverse plane and the good precision with which CDF measures the proper decay length of the B_s provide the experiment with a suitable time resolution to observe the phenomenon, if it occurs at frequencies compatible with the Standard Model predictions ($\Delta m_s \approx 20\ ps^{-1}$). Infact, CDF measures Δm_s with a statistical significance $\geq 5\sigma$ and is able to claim a direct observation of the B_s oscillations.

Several experimental ingredients contribute to reach this level of significance. One fundamental point that determines the statistical power of the analysis is

flavour tagging. b flavour tagging consists of the determination of the flavour of the b quark contained within a meson. This information is vital to perform any time-dependent measurement involving flavour asymmetries in b meson decays and flavour oscillations. Here, in fact, it is necessary to know whether a b or $\bar{b}$ was present in the meson when it was produced.

The *effective* number of signal events used in the asymmetry measurement scales with the figure-of-merit of the combined flavour tagging used, εD^2. Given an initial number of events, ε is the fraction of them we can apply flavour tagging to. It is called *efficiency*. D is instead related to the probability P that the tag response is correct. It is defined as $D = 2P - 1$ and called *dilution*. An $\varepsilon D^2 = 1\%$ is equivalent to using only 1% of all the selected B_s signal sample, but with a perfect tagging.

This explains the importance of adding the Opposite Side Kaon Tagging for the B_s mixing analysis. It is also very important to combine its response with the other algorithms developed in the most refined way possible in order to enhance the overall effectiveness.

Several tagging algorithms have been put in place at the CDF experiment. Some of them look at the flavour of the *accompanying* b hadron to infer that of the B meson under study. At a hadron collider, two b quarks are likely to be produced in pairs; they are then hadronized incoherently into pairs of b hadrons which define ideally two distinct hemispheres. The class of taggers exploiting the accompanying b are called Opposite-side taggers (OST). They work on the principle that a correlation is established between the Opposite-side b and the electric charge of one or all of its decay daughters. CDF uses both leptons and the overall b-jet charge from b decays to tag the flavour on the Opposite-side, with performances of the order of $\varepsilon D^2 = 1\%$.

It is very likely that the b (in the Opposite side) decays through the chain $b \to c \to s$, originating a Kaon in the final state. If the Kaon is charged, a tracking system can detect it and reconstruct its trajectory back to the decay vertex of the b hadron. Consequently, the Kaon can be used to tag the b flavour.

I show that it is possible to develop a Kaon tagger with non-null performances in a high track multiplicity environment. I find a figure-of-merit of the same order of the other OST by exploiting two main lever arms. I select Kaons using the CDF II Particle Identification (PID) system, combining the information of the specific ionization of the Kaon in the CDF central tracker and that of the dedicated Time-Of-Flight detector. A reliable Kaon identification is obtained on a wide momentum range. It allows to suppress other charged particle species, in particular pions that dominate at momenta considered by flavour physics ($|\vec{p}| \leq \mathcal{O}(m(B))$.

The other useful feature is the displacement of signal Kaons from the Primary Vertex (PV), due to the non-null lifetime of its b-hadron parent. Using the information of the high-resolution CDF II silicon detector, I discriminate signal Kaons from other tracks likely to have been generated from the underlying interactions

of the $p\bar{p}$ collision. After applying a PID selection, the main component in the latter is represented by low-momentum Kaons.

Finally, one can extract information from the typical topology of heavy flavour decays by looking at the flight direction of the decaying b hadron. This is done by reconstructing the b-jet in the opposite side. Kaons close to b jets, where a secondary decay vertex was explicitly identified, are more likely to have originated in the $b \to c \to s$ chain. These are expected to carry more reliable information about the b flavour. This translates into a higher contribution to the final flavour tagging effectiveness.

Provided this additional information, CDF uses the OSKT in combination with the other existing flavour taggers of the Same-side and the Opposite-side. More than one OST may be present per event. Therefore, one needs to combine the various tags in a way to exploit the most of the correlations occurring among them. In this framework, I collaborate to develop a Neural Network (NN) based combination. I perform physics studies and show that the OSKT enhances the performances of the other tags when their responses are correlated. Such relationships are automatically taken into account by a NN. This is a processing algorithm that works based on the train and experience principle as the human brain. When establishing an overall NN tagger, it is proved to provide a consistent increase in the final performances with respect to exclusive approaches, when only one tagger at the time is considered.

At CDF, the introduction of the OSKT and the NN combination contribute as a fractional 20% of the overall CDF Opposite-side effectiveness and an enhancement of the statistical significance of the Δm_s measurement is achieved.

The layout of this thesis is as follows. In Chapter 1 I discuss the theoretical relevance of the measurement of Δm_s. I first introduce quark-quark weak interactions and the CKM matrix governing them. The features of the matrix and how CP violation is originated are briefly presented. This leads to the discussion of the UT. Then, I proceed to establish the connection between flavour oscillations and UT, by reviewing the theory of B meson mixing. The loop diagrams that govern B_d and B_s mixing present the coupling of (essentially) the t quark with a d and s quark respectively. Thus, I show how the combined measurement of the mixing frequencies determines the ratio of the amplitudes of the $t-d$ and $t-s$ couplings in the CKM. A survey the implications of a Δm_s value different from the SM expectations is also provided. Flavour oscillations are a good testing ground for BSM scenarios. Not only they can give hint of those, but they play a relevant role in discriminating among the various models.

In Chapter 2, the experimental scenario in the mixing of the B_d and B_s mesons is summarized. I review the world knowledge of Δm_d and run briefly through various techniques used for its measurement. My interest is mainly focused to the description of the various attempts that were done in the past to measure Δm_s. These lead to the evidence for B_s mixing at CDF early this year (2006). This

analysis has been performed using a maximum likelihood-based fit in the time domain. Nevertheless, when looking for oscillations of any kind it is convenient to move to the frequency domain, in order to look directly for a signal peak at the frequency of the phenomenon. This can be done by applying the Fourier transform to a sample of data. A scan is then performed at various frequencies to look for the amplitude of such a transform. Alternatively, an equivalent amplitude can be introduced in the likelihood, in a sort of "compromise" between the two methods. I show how the amplitude scan works and how it is useful to set a combined lower limit on Δm_s from the past experiments before the evidence at CDF.

Chapter 3 is devoted to describe the experimental apparatus and the trigger system of the CDF II experiment. The chapter is organized as follows: the way Tevatron ensures a high instantaneous luminosity through a multiple level of acceleration stages is presented first. I also show how it is possible to obtain a large number of anti-protons for collisions. The performances in terms of resolution and time-stability of the CDF II detector are also important. In particular, I review the tracking and vertexing systems and summarize their features, essential for the selection of tagging Kaons in the Opposite-side. Together with a good spatial resolution, in this study I also profit of the PID systems. So, I describe how the measurements of the specific ionization in the drift chamber, dE/dx, and a particle Time-Of-Flight are performed and calibrated on data. Provided their resolution, one can estimate how well Kaons can be identified from other charged particle species. The resolution is also used to simulate the TOF and dE/dx response. Finally, I describe how the trigger system is organized in three levels and review how each of them selects events and with which performances. The resulting inclusive data samples used for flavour tagging studies and exclusive samples for B_s oscillations are obtained by a subsequent set of more refined trigger requests that I present here.

In Chapter 4 a description of the signal selections for B_s mixing is provided, both for hadronic and semileptonic decays. The analysis requests and the relevant variables to identify and suppress the various kinds of backgrounds are introduced. The main difference between hadronic decays and semileptonic decays is given by the uncertainty with which the proper decay time of each B_s is measured. I show the differences among the two. Moreover, hadronic decays where one neutral particle was lost are proved to have an effective statistical weight comparable to the fully reconstructed modes, thanks to their good time resolution. Finally, I describe the flavour tagging algorithms already in place at CDF. The Same Side Kaon Tagger looks directly at the flavour of the neutral B meson under study. It uses the correlation between this and the electric charge of the leading product of the hadronization process that originated also the meson. I summarize the performances of the algorithm and how they are assessed using Monte Carlo simulation. On the other hand, the Opposite-side taggers are calibrated on a high-statistics inclusive sample of b semileptonic decays. I describe

the characteristics of this sample and how a background subtraction procedure is able to render it pure in b for an unbiased evaluation of the OSKT performances.

Chapter 5 is devoted to the description of the OSKT and the evaluation of its performances. First, I will explain how the algorithm works and the main ingredients of the physics analysis. The data samples used are presented; I also make use of Monte Carlo simulation for the comprehension of the various contributions to the dilution. Then, the requirements applied to the inclusive set of tracks for tagging are discussed, that lead to the selection of tagging candidates. A cut on the track displacement is introduced. All the possible tracking and vertexing effects are also evaluated to improve the rejection of background from the PV. The combined PID through a Likelihood Ratio is also fundamental and is discussed. The ranking of the Kaon tags based on the b-jet determination is also described. I also explain how a decision is taken when more than one tagging candidate is found. The performances on an independent semileptonic sample are presented. The assessment of such performances in different configurations and cross-checks are reported. I also discuss the origin of the tagger dilution from the various contributions in the tagging sample, which confirms my expectations based on physics considerations. Finally, I proceed to evaluate the effect on dilution of tracking asymmetry in the CDF detector and of the different absorption cross sections of positive and negative Kaons by the CDF detector material. These may, in fact, induce a bias when applying the tagger to a time-dependent asymmetry measurement. The chapter is concluded with a critical review of the main problems faced and of the perspectives for similar algorithms at future hadron colliders, like LHC. A dedicated Monte Carlo tuning based on Tevatron multiplicities and a thorough comprehension of the specific detector sub-systems will be necessary are shown to be very important.

The combination of the CDF Opposite-side taggers is the topic of Chapter 6. For that, an evaluation of the performances on exclusive data samples of the kind used for B_s mixing analysis is needed. I propose an example of how physics correlations among different Opposite-side tags increase their performances. These studies on tagger correlations culminate in the description of the NN-based OST. The informations on several physical quantities from each tagger are merged together and lead to a decision with an improved associated dilution.

Finally, Chapter 7 illustrates the results of CDF on the B_s mixing. The observation of $B_s^0 - \overline{B}_s^0$ oscillations at the CDF experiment is presented. An evidence for the phenomenon had been reported by CDF last spring. The present analysis is performed with the same set of data, corresponding to 1 fb^{-1}. The statistical significance necessary to claim an observation is reached by means of a significant increase in B_s decays selected both in semileptonic and hadronic modes and of the introduction of the more refined tagging algorithm in the Opposite-side. Here I summarize the yields for the different decays and the average proper time uncertainty. Finally, I show the result of the frequency scan and the value of Δm_s

as obtained from it. The way we evaluate the significance from the background fluctuation probability is also shown. The information inferred on the CKM elements and an interpretation of the measurement are given eventually. This uses a model-independent approach to NP contributions. Since the agreement of Δm_s with the SM prediction is at the 1σ level, room for NP is left only in a possible new phase concurring to the phenomenon. Its measurement could be performed at CDF using time-dependent decays of B_s into CP eigenstates, as $J/\psi\phi$. In this, the role of the CDF flavour tagging represents again a vital ingredient.

Chapter 1

Flavour oscillations and their impact on the Unitarity Triangle

Flavour oscillations are a quantum phenomenon originated by weak interactions, through which the flavour F of a meson is changed into its conjugate ($\Delta F = 2$ process). The quark weak coupling is predicted by the Standard Model (SM) to be entirely governed by the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements. The CKM matrix is complex and unitary. The complexity leads to the peculiar feature of weak interactions to violate the product *CP* of the Charge Conjugation C and the Parity P symmetries. The unitarity relations among the CKM elements define triangles in the complex plane, whose area is proportional to the amount of CP violation. I show here that the measurement of the frequency with which the flavour oscillation occur is a powerful constraint on the geometry of the triangles. As such, it provides information on the amount of CP Violation (CPV) in nature. Furthermore, new particles are predicted by models beyond the SM to contribute to $\Delta F = 2$ processes by coupling with SM quarks. In the most common scenarios this translates into values of the oscillation frequency larger than the SM predictions. Thus, the study of this phenomenon and the determination of the value of the frequency are a very powerful indirect search for New Physics (NP).

1.1 The CKM matrix and the Unitarity Triangle in the B sector

The 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix [1] [2] connects the *down* quark *weak eigenstates* (d', s', b') and the corresponding *mass eigenstates* d, s, b:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \hat{V}_{\rm CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

Thus, one can write the charged currents in the $SU(2) \times U(1)$ model of weak interactions as

$$\begin{aligned} J_\mu^{\mathcal{CC}} &= \overline{u}' \gamma_\mu (1-\gamma_5) d' \\ &= \overline{u} \gamma_\mu (1-\gamma_5) V_{\rm CKM}\ d \end{aligned} \tag{1.2}$$

where u, d are the spinors of the *up* and *down* quark fields respectively. As seen, the CKM matrix governs the V-A vector-axial charged currents and its elements give the strength of the various up-down quark couplings via the weak interactions. A 3×3 unitary and complex matrix has 4 independent parameters: 3 rotation angles and 1 complex phase. It can be shown that the presence of a complex phase leads to the non-conservation of the *CP* (Charge Conjugation C $\times$ Parity P) quantum operator in physics processes where the CKM is involved. Consequently, charged weak currents violate the CP symmetry.

Experimentally, the absolute values of the elements of the CKM matrix show a hierarchy with the diagonal elements being close to unity. The magnitude of the elements $|V_{us}|$ and $|V_{cd}|$ are of the order of 0.2. $|V_{cb}|$ and $|V_{ts}|$ are of the order of $4{\cdot}10^{-2}$, while $|V_{ub}|$ and $|V_{td}|$ are of the order $5{\cdot}10^{-3}$. This is explicitly visible when using the Wolfenstein parametrization [3] of the CKM matrix. Each element of the CKM matrix is expanded as a power series in the parameter $\lambda \equiv |V_{us}|$, where $\lambda \approx 0.22$ [1]. The parameterization is as follows:

$$\hat{V}_{\rm CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\varrho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\varrho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,, \tag{1.3}$$

In the Wolfenstein parametrization the four CKM independent parameters are:

$$\lambda, \qquad A, \qquad \varrho, \qquad \eta\,. \tag{1.4}$$

Given the smallness of λ, it is sufficient to keep only the first few terms in this expansion. Using this way of writing, the matrix complexity is expressed by η. A null η would mean the matrix is real and *no CPV* is present. A convenient way of expressing this concept is derived by writing the unitarity conditions of the CKM matrix

$$\sum_k V_{ik} V_{jk}^* = \delta_{ij} \tag{1.5}$$

in the Wolfenstein parameterization. Here δ_{ij} is the Kronecker symbol. Taking, for example, $j = d, k = b$, one finds for the B_d meson:

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0 \tag{1.6}$$

[1] For a detailed review on the measurements of $|V_{us}|$ from various observables, refer, e.g., to [4]

or, dividing by $V_{cd}V_{cb}^*$,

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \tag{1.7}$$

In the Wolfenstein parameterization, this becomes

$$\begin{aligned}(1-\frac{\lambda^2}{2})(\rho - i\eta) + 1 + (1-\rho-i\eta) &= \\ (\bar{\rho} - i\bar{\eta}) + 1 + (1-\bar{\rho}-i\bar{\eta}) &= 0\end{aligned} \tag{1.8}$$

having defined $\bar{\rho}(\bar{\eta}) = (1-\frac{\lambda^2}{2})\rho(\eta)$. This describes a triangle in the $\bar{\rho}-\bar{\eta}$ complex plane, called the *Unitarity Triangle* (UT). It is shown in Fig.1.1. Its sides are all

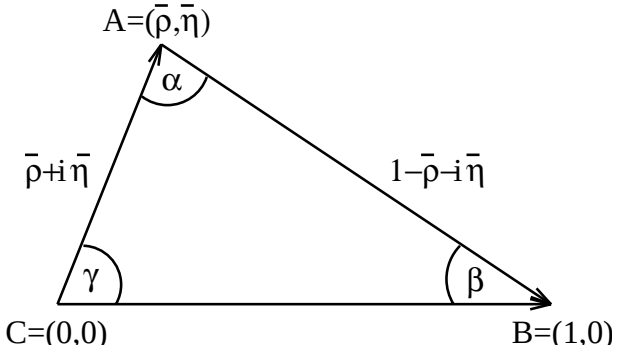

Figure 1.1: Example of Unitary Triangle.

$\mathcal{O}(1)$, which is *not* a general feature of the UT arising from the relations 1.5. For example, in the B_s triangle only two sides are $\mathcal{O}(1)$, while the remaining is $\mathcal{O}(\lambda^2)$. This results in a very acute triangle. Given this characteristics, the B_d triangle is expected to be more easily measured. Specific B_d "factories" have been built in order to perform a vast experimental programme in this sector. Measurements in the B_s system are useful to add more precise information to the former. The Tevatron hadron collider is a good "factory" for B_s, since it is able to produce such mesons in the hadronization of $b\bar{b}$ pairs.

In this picture, the triangle has an area $A_{UT} \neq 0$, which implies that $\eta \neq 0$ and CPV occurs in weak interactions. In fact, it can be shown [5] that A_{UT} is directly proportional to the amount of CPV in weak coupling. While the side CB has a unit length, the other sides are proportional in size to the CKM matrix elements:

$$\overline{CA} \equiv \frac{|V_{ud}V_{ub}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{\overline{\varrho}^2 + \overline{\eta}^2} = (1-\frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|, \tag{1.9}$$

$$\overline{BA} \equiv \frac{|V_{td}V_{tb}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{(1-\overline{\varrho})^2 + \overline{\eta}^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \tag{1.10}$$

The angles β and γ of the unitarity triangle are related directly to the complex phases of the CKM elements V_{td} and V_{ub}, respectively, through the expressions

$$\begin{aligned} V_{td} &= A\lambda^2(1-\bar{\rho}-i\bar{\eta}) = |V_{td}|e^{-i\beta} \\ V_{ub} &= A\lambda^2(\bar{\rho}-i\bar{\eta}) = |V_{ub}|e^{-i\gamma} \end{aligned} \tag{1.11}$$

According to the SM, all measurements of processes involving the CKM matrix have to be consistent with $A(\bar{\rho},\bar{\eta})$ (see e.g. Fig.1.2). The determination of the

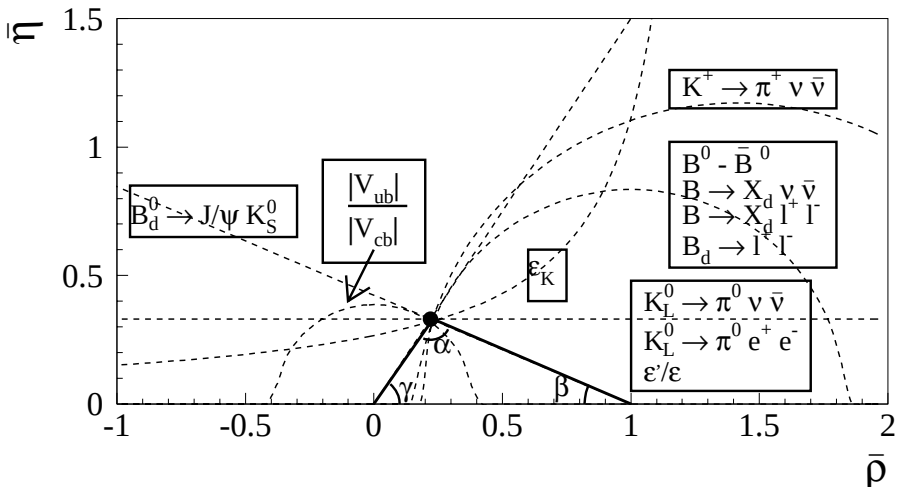

Figure 1.2: Example of perfect consistency on the values $\bar{\rho},\bar{\eta}$ of various measurements in the K and B sector.

vertex $A(\bar{\rho},\bar{\eta})$ through relations as 1.9, 1.10 is the main goal of experimental flavour physics. Together with the measurement of λ and the other parameter A, it leads to the complete knowledge of the CKM matrix. At present, the experimental strategy is to perform as many as possible different measurements of accessible quantities that involve the CKM matrix elements in the strange, charm and bottom sectors in order to reduce the uncertainties on the apex $A(\bar{\rho},\bar{\eta})$. All the informations from UT side measurements and CPV in the Kaon system (measurement of ε_K) are merged together and the current knowledge of the vertex reads

$$\begin{aligned} \bar{\rho} &= 0.188 \pm 0.036 \\ \bar{\eta} &= 0.371 \pm 0.027 \quad [69]. \end{aligned} \tag{1.12}$$

Any inconsistency in the $(\overline{\varrho},\overline{\eta})$ plane would represent a hint of physics beyond the SM. As one can see from Fig.1.2, the constraints coming from B decay amplitudes, B_d and B_s mixing and CP violation in the interference between decay and mixing can be used to derive values of η and ϕ. Those values are then

compared with processes involving Kaons, like $K^0 - \overline{K}^0$ mixing and rare decays (e.g. $K^+ \to \pi^+ \nu\bar{\nu}$). Since all these processes are governed by the CKM matrix, any disagreement among those would mean different New Physics (NP) contributions in the two systems. For example, NP could intervene in the loop diagrams through which B mixing occurs, as we will see later in this chapter.

1.2 Theory of flavour oscillations in the SM and constraints to the UT

Let's consider a general two-state system satisfying the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} B^0 \\ \overline{B}^0 \end{pmatrix} = \hat{H}\begin{pmatrix} B^0 \\ \overline{B}^0 \end{pmatrix} \tag{1.13}$$

where $\hat{H}$ is the hamiltonian operator with elements

$$\hat{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{1.14}$$

The hermitian matrix $\hat{H}$ can be written as the sum of two hermitian matrices $\hat{M}$ and $\hat{\Gamma}$

$$H_{ij} = M_{ij} - i\Gamma_{ij}/2 \tag{1.15}$$

Let's regard $|B^0\rangle = |\bar{b}d\rangle$ and $|\overline{B}^0 = b\bar{d}\rangle$ as two flavour eigenstates with eigenvalues $F = +1$ and $F = -1$. In general they may not be hamiltonian eigenstates. The hamiltonian eigenstates can be expressed in this complete base by diagonalizing the $\hat{H}$ matrix:

$$|B^0_L\rangle = p|B^0\rangle + q|\overline{B}^0\rangle \;\; ; \quad |B^0_H\rangle = p|B^0\rangle - q|\overline{B}^0\rangle \tag{1.16}$$

with $|q|^2+|p|^2 = 1$. As a result of the diagonalization of Eq.1.15, also the following relations are derived:

$$\begin{aligned} \Delta m^2 - 1/4\Delta\Gamma^2 = &\quad 4|M_{12}|^2 - |\Gamma_{12}|^2 \\ \Delta m\Delta\Gamma = &\quad 4Re(M_{12}\Gamma^*_{12}) \end{aligned} \tag{1.17}$$

where we define $\Delta m = M_H - M_L$ and $\Delta\Gamma = \Gamma_H - \Gamma_L$.

In the Standard Model, such $\mathrm{B}^0 - \overline{\mathrm{B}^0}$ mixing is ruled by a second-order loop process (called *box diagram*, Fig.1.3). Here d-type quarks couple with u-type quarks via weak interactions. It can be shown that the real part of the hamiltonian – $\mathrm{B}^0 - \overline{\mathrm{B}^0}$ *oscillation*, contained in M_{12} – is dominated by the coupling of the down quark with the top quark. Light quarks contribute to the imaginary

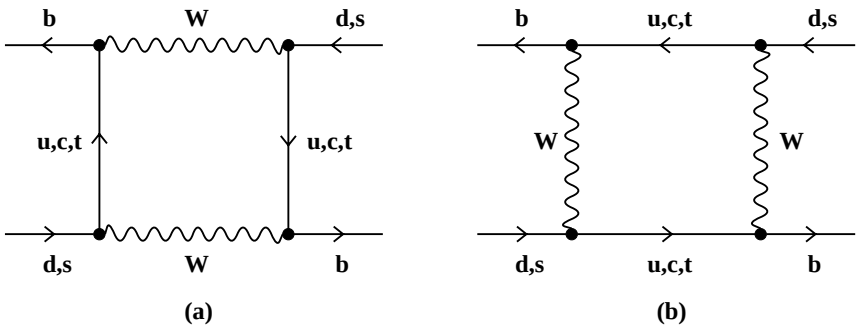

Figure 1.3: Box diagrams through which flavour oscillations of B_d^0, B_s^0 into their antiparticles occur.

part, relative to the decay (Γ_{12}). If one calculates the relative amplitudes of the two parts, one gets $\Gamma_{12}/M_{12} = m_b^2/m_t^2 << 1$. The relations 1.18 simplify to:

$$\begin{aligned} \Delta m = &\quad 2|M_{12}| \\ \Delta\Gamma = &\quad \tfrac{2Re(M_{12}\Gamma 12^*)}{|M_{12}|} \end{aligned} \tag{1.18}$$

In this case, $\Delta\Gamma << \Delta m$ proves to be a reasonable approximation. The time evolution of the flavour eigenstates is easily obtained from 1.14:

$$|B^0(t)\rangle_t = e^{-imt}\ e^{-\Gamma/2\ t}(cos\ \Delta m/2t\ |B^0\rangle\ +\ i\ \frac{q}{p}\ sin\Delta m\ t/2\ |\overline{B}^0\rangle)$$
$$|\overline{B}^0(t)\rangle_t = e^{-imt}\ e^{-\Gamma/2t}(cos\ \Delta m/2t\ |\overline{B}^0\rangle\ +\ i\ \frac{q}{p} sin\Delta mt/2\ |B^0\rangle)$$

which expresses the oscillation amplitude of the original state $|B^0\rangle \equiv |B^0(0)\rangle_t$ into its Charge Conjugate $|\overline{B^0}\rangle$. The probabilities that a meson created as $|B^0\rangle$ is found in the same state at a certain time t or that it has mixed to $|\overline{B}^0\rangle$ are obtained following the quantum mechanical rules:

$$\begin{aligned} Prob(|B^0\rangle \to |B^0(t)\rangle_t) &= |_t\langle B^0(t)|\hat{H}|B^0\rangle|^2 \\ &= \frac{1}{2}e^{-\Gamma t}(1 + cos\Delta mt) \qquad (1.19) \\ Prob(|B^0\rangle \to |\overline{B}^0(t)\rangle_t) &= |_t\langle \overline{B}^0(t)|\hat{H}|B^0\rangle|^2 \\ &= \frac{1}{2}e^{-\Gamma t}(1 - cos\Delta mt) \qquad (1.20) \\ & \qquad (1.21) \end{aligned}$$

From 1.21 one can also derive the time-integrated mixing probability $Prob(|B^0\rangle \to |\overline{B}^0(t)\rangle_t)$:

$$\chi = \frac{x^2}{2\,(1+x^2)} \qquad \text{with} \quad x = \Delta m \cdot \tau, \tag{1.22}$$

where τ is the B^0 meson lifetime ($\tau = 1/\Gamma$).

This can be viewed as a proper $\Delta F = 2$ *oscillation* between the two flavour eigenstates, occurring with a frequency Δm. Fig.1.4 displays the time dependence of the two probabilities in 1.21 for two different values of the parameter Δm. The SM expression for the frequency in the B_d^0 meson system, Δm_d, is derived

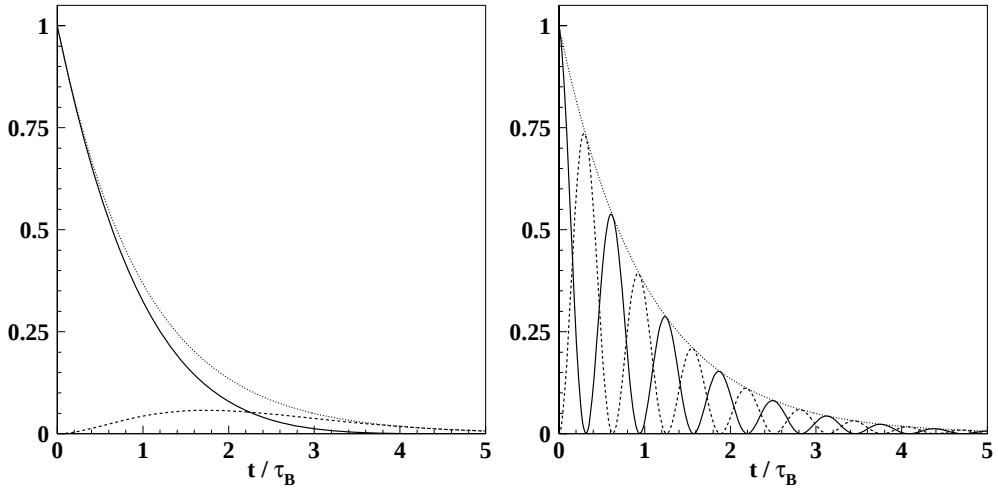

Figure 1.4: Time evolution of B^0–$\overline{B}^0$ oscillations; the non-mix (solid) and the mix (dashed) probability curves are displayed along with their sum (dotted) for (a) slow and (b) fast oscillation frequencies Δm.

calculating the amplitude for the box diagram and using the approximation in 1.24. We get

$$\begin{aligned} \Delta m_d &= \frac{G_F^2}{6\pi^2} m_W^2 \ \eta_b S_0(\frac{m_t^2}{m_W^2}) m_{B_d} \ f_d^2 \hat{B}_d \ |V_{tb}V_{td}^*|^2 \\ &= \frac{G_F^2}{6\pi^2} m_W^2 \ \eta_b S_0(\frac{m_t^2}{m_W^2}) \ m_{B_d} \ f_d^2 \hat{B}_d A^2 \lambda^6 [(1-\bar{\rho})^2 + \bar{\eta}^2] \end{aligned} \qquad (1.23)$$

Here:

- G_F is the Fermi coupling constant; m_W is the W boson mass, m_{B_d} is the B_d^0 meson mass;
- η_b is the correction at NLO correction from the short-distance perturbative QCD part;
- $S_0(\frac{m_t^2}{m_W^2})$ is the Inami-Lim function [6] accounting for the effect of the presence of heavy quarks within the low-energy mixing loop; here m_t is the measured top quark mass;

- the product $f_d^2 \hat{B}_d$ represents the non-perturbative contribution in the hadron matrix element. The f_d expresses the size of the meson wave function at the origin. The B-parameter $\hat{B}_d$ corrects for deviations from the vacuum saturation approximation,.This is introduced when calculating the mixing amplitude through the vacuum state insertion.

The CKM matrix enters the expression 1.23 through the product $A^2\lambda^6[(1-\bar{\rho})^2+\bar{\eta}^2]$. From Eq.1.10 we see that the measurement of the single mixing frequency Δm_d constraints the side BA of the UT, once also A and λ are measured. In other words, it fixes the apex $A(\bar{\rho},\bar{\eta})$ on a circumference centered in $(\bar{\rho},\bar{\eta}) = (0,1)$ and of radius given in 1.23. Such a requirement is depicted in Fig.1.5.

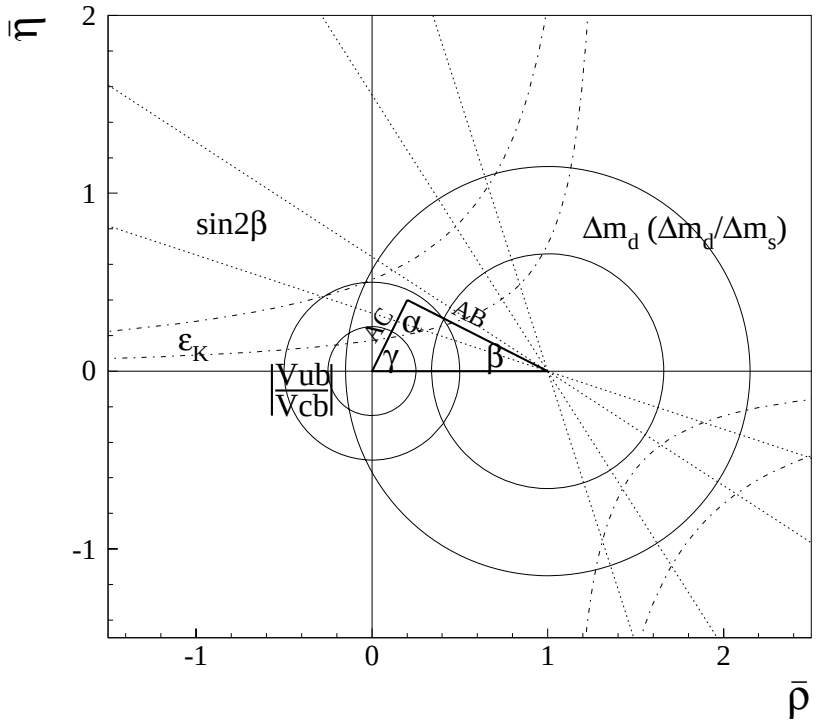

Figure 1.5: Constraint to the unitarity triangle side BA from the measurement of the B_d^0 mixing frequency δm_d and the ratio Δm_d Δm_s.

At present, the widest uncertainty in the Δm_d expression is given by the non-perturbative product $f_d\sqrt{\hat{B}_d}$: this is calculated using lattice techniques. Several approaches have been used to extract this value, e.g. [68]. The current uncertainty is $\approx 15\%$, which reflects directly onto the determination of $|V_{td}|$ from 1.23. To reduce this uncertainty it is convenient to use the ratio

$$\frac{\Delta m_d}{\Delta m_s} = \frac{m_{B_d} f_{B_d}^2 \hat{B}_{B_d}}{m_{B_s} f_{B_s}^2 \hat{B}_{B_s}} \frac{|V_{td}|^2}{|V_{ts}|^2}$$

$$= \frac{m_{B_d} f_{B_d}^2 \hat{B}_{B_d}}{m_{B_s} f_{B_s}^2 \hat{B}_{B_s}} \left(\frac{\lambda}{1-\frac{\lambda^2}{2}}\right)^2 \frac{(1-\bar{\rho})^2+\bar{\eta}^2}{\left(1+\frac{\lambda^2}{1-\frac{\lambda^2}{2}}\bar{\rho}\right)^2+\lambda^4\bar{\eta}^2} \,. \tag{1.24}$$

Infact, the theoretical uncertainty on the ratio $\xi = f_{B_s}\sqrt{\hat{B}_{B_s}}/f_{B_d}\sqrt{\hat{B}_{B_d}}$ is expected to be smaller than for the individual non-perturbative terms. Since V_{ts} is independent of $\bar{\rho}$ and $\bar{\eta}$ up to terms $\mathcal{O}(\lambda^4)$, the ratio provides the same constraint on the UT as the single mixing frequency Δm_d, but with a reduced error (4%).

On the other side, for the ratio 1.24 we have

$$\frac{\Delta m_d}{\Delta m_s} \propto \lambda^2 \approx 0.05. \tag{1.25}$$

Inserting the correct values for the expressions of the two amplitudes, the SM predicts for the $B_s^0 - \overline{B}_s^0$ oscillations to occur with a frequency $\Delta m_s \approx 40 \times \Delta m_d$. This is very relevant for the experimental determination of Δm_s. A better resolution on the B proper decay time is infact needed in the B_s^0 than for the B_d^0 case. This is the main reason why B_s oscillations could not be observed, up to know.

1.3 How does New Physics enter the picture?

The validity of the Standard Model has been confirmed by measurements of various processes. Nevertheless, it is believed to be a low energy ($E < 1\ TeV$) effective theory of a more general theory at higher energies. For instance, the origin of electroweak symmetry breaking (and the related hierarchy problem), the number of generations, the unification of gauge couplings at a certain energy scale are not explained by SM. Several models beyond the SM have been proposed.

Models contributing to flavour violation processes (e.g. $\Delta F = 2$ operators) can be divided into two main classes:

- those in which the only source of flavour violation remains the CKM matrix (Minimal Flavour Violation models, MFV). New interaction vertices are possible due to new particles;

- those which encode completely new sources of flavour violation.

For the first category the strength of quark-quark couplings is entirely given by CKM. Despite this, new operators can intervene in the Operator Product Expansion of the effective weak Hamiltonian for $\Delta F = 2$:

$$H_{\text{eff}}^{\Delta F=2} = \frac{G_F^2 M_W^2}{16\pi^2} \sum_i V_{\text{CKM}}^i C_i(\mu) Q_i. \tag{1.26}$$

Some models predict no further operators with respect to the SM ones. In these cases, the only non-standard contributions are given by new particles circulating within the box diagram 1.3, in addiction to *up*-type quarks. While the single SM predictions on Δm_d and Δm_s may change (see, e.g., [7]), their ratio is not expected to change. Consequently the constraint on the UT from these models is the same as in the SM. For example, in the case of the Minimal Supersymmetric Model (MSSM[2])

$$\begin{aligned} \Delta m_d &= \Delta m_d^{SM}[1 + f(m_{\chi_2^\pm}, m_{\bar{t}_R}, m_{H^\pm}, tan(\bar{\beta}))], \\ \Delta m_s &= \Delta m_s^{SM}[1 + f(m_{\chi_2^\pm}, m_{\bar{t}_R}, m_{H^\pm}, tan(\bar{\beta}))] \end{aligned} \tag{1.27}$$

The function $f(m_{\chi_2^\pm}, m_{\bar{t}_R}, m_{H^\pm}, tan(\bar{\beta})^3)$ describes the effect of supersymmetry as a function of the masses of the new particles circulating in the box diagram. f is positive-definite. Its maximum allowed value varies according to the model. E.g. in the case of SUpersymmetricGRAvity (SUGRA) models this is 0.45: this would result in an increased value for both Δm_d and Δm_s with respect to the expected SM value.

In the case of models with more operators added to the SM ones, New Physics (NP) contributions f of Eq.1.27 would no longer be universal. Infact, new operators act differently on the d and s quark spinors in the weak currents. Eq.1.27 becomes

$$\begin{aligned} \Delta m_d &= \Delta m_d^{SM}[1 + f_d(m_{\chi_2^\pm}, m_{\bar{t}_R}, m_{H^\pm}, tan(\bar{\beta}))], \\ \Delta m_s &= \Delta m_s^{SM}[1 + f_s(m_{\chi_2^\pm}, m_{\bar{t}_R}, m_{H^\pm}, tan(\bar{\beta}))] \\ f_d &\neq f_s \end{aligned} \tag{1.28}$$

In such a case, not only the Δm_d and Δm_s are changed separately, but also their ratio results affected by NP contributions. Thus, the position of the UT vertex $A(\bar{\rho}, \bar{\eta})$ is moved with respect to SM expectations. The relationship between the $\Delta m_s/\Delta m_d$ ratio and the UT angle γ is displayed in Fig.1.6, as evaluated in [9]. The value of $sin(2\beta)$ has been fixed to 0.6 . Its average known value is $sin2\beta = 0.687 \pm 0.032$ [37].

In this brief excursus we do not discuss any model where new sources of flavour transitions aside of the CKM matrix are concerned. In such cases further mixing phases may also emerge other than β in the CKM matrix. To express the effect of "non-minimal" flavor violation contributions in the measurable value of Δm, one can adopt a model-independent approach proposed in [10]. In this analysis, NP contributions are included in the general amplitude C_{B_q} and phase ϕ_{B_q} of the

[2] see e.g.[8]

[3] $tan(\bar{\beta}) \equiv v_u/v_d$, where $v_u(v_d)$ is the vacuum expectation value of the Higgs field coupling exclusively to up-type (down-type) quarks and leptons

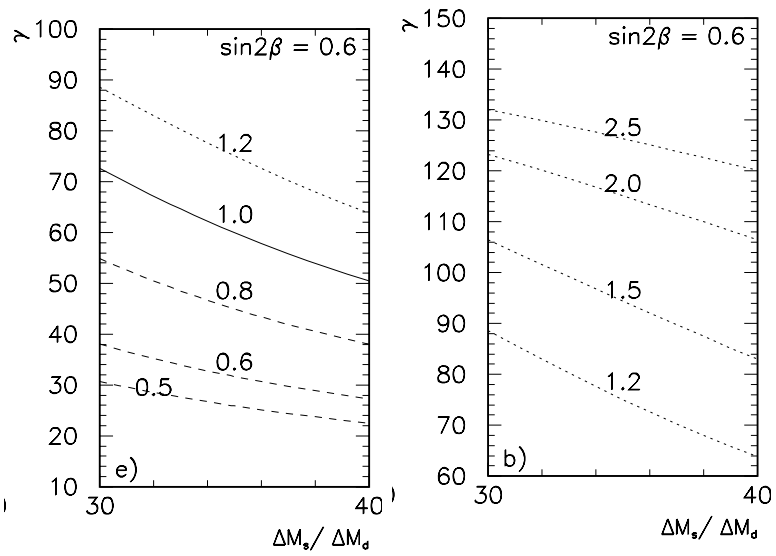

Figure 1.6: γ as a function of $\Delta m_s/\Delta m_d$ for $\sin 2\beta = 0.6$ ($sin2\beta = 0.687 \pm 0.032$ from charmonium [37]), for different values of $\frac{1+f_s}{1+f_d}$ (marked on the curves) and $\xi = 1.15$. From Ref.[9]

$B_q - \bar{B}_q$ mixing ($q = d, s$):

$$\begin{aligned} C_{B_q} e^{2i\phi_{B_q}} &= \frac{< B_q|H_{eff}^{SM+NP}|\bar{B}_q >}{< B_q|H_{eff}^{SM}|\bar{B}_q >} \\ &= \frac{A_q^{SM} e^{2i\phi_q^{SM}} + A_q^{NP} e^{2i(\phi_q^{SM}+\phi_q^{NP})}}{A_q^{SM} e^{2i\phi_q^{SM}}} \end{aligned} \tag{1.29}$$

where $H_{eff}^{SM+NP}(H_{eff}^{SM})$ is the $\Delta F = 2$ effective hamiltonian in the SM + NP extension (SM-only) model. If no SM constributions are involved $C_{B_q} = 1$, $\phi_{B_q} = 0$. The experimental observables are then related to the SM contributions by the following general expressions

$$\begin{aligned} \Delta m_q &= C_{B_q} \Delta m_q^{SM} \\ \beta_q^{exp} &= \beta_q + \phi_{B_q} \ (+2n\pi) \end{aligned} \tag{1.30}$$

where β_q has been defined in Eq.1.11. While C_{B_q} can be determined by comparing the measured Δm_q with the SM-only expectations, the phase is experimentally accessible through CP-violating processes. For example, one can look at time-dependent CP asymmetries induced by the interference of mixing in the decay of B^0 and $\overline{B^0}$ to CP eigenstates (e.g. $B_d^0 \to J/\psi K_s^0$ and $B_s^0 \to J/\psi\phi$). Since a new phase enhances the complex part of the interaction, also the decay-width (*i.e.* lifetime) difference $\Delta\Gamma$ of the mass eigenstates B_L^0 and B_H^0 can be used.

In conclusion, the measurement of **both** B_d^0 and B_s^0 flavour oscillations is a very powerful tool to improve our knowledge of the CKM matrix. This happens through a strong constraint to the free parameters ρ and η. Their determination provides a measurement of the amount of the CP symmetry violation in nature. The knowledge of Δm_d fixes the values of ρ and η only with a 15% uncertainty. This comes from theoretical calculations of the non-perturbative part of the mixing amplitude. The combined measurement of Δm_d and Δm_s reduces the uncertainty to 4%, instead. Furthermore, these quantities give indirect access to physics beyond the SM in the loop diagrams governing $\Delta F = 2$ processes. The two values are separately sensible to models that violate the flavour in a minimal way (no new sources aside the CKM), both with or without new effective hamiltonian operators. In the most common models higher frequencies will originate with respect to the SM expectations. The $\Delta m_d/\Delta m_s$ ratio is sensible to models where new operators contribute to flavour interactions. Finally, the time-dependent determination of flavour oscillations is also very useful to explore the presence of new phases that allow CP violation aside the CKM complex phase. This hypothesis arises in the case of a more general scenario where the CKM matrix is not the only source of flavour interactions.

Chapter 2

Experimental scenario in the B-mixing sector

A great effort to measure the mixing frequency both in the B_d and in the B_s sector has been undertaken in the past years. This includes contributions from both e^+e^- machines and the Tevatron hadron collider. In particular, the most precise determinations of Δm_d have come from the dedicated asymmetric *B-Factories* (PEP-II at SLAC and KEKB at KEK) running at the $Y(4S)$ resonance mass. The $Y(4S)$ decays to $B^0_d\bar{B}^0_d$ pairs with a probability $BR(Y(4S) \rightarrow B^0_d\bar{B}^0_d) = 49.4 \pm 0.8$ % [21]. The resulting B decays are analyzed by the BaBar and Belle detectors respectively. Mixing analyses have also been carried out at the e^+e^- SLC and LEP colliders, running at the Z boson mass and at higher centre-of-mass (c.o.m.) energies. The SLD detector on one side and the four LEP detectors on the other have performed both Δm_d and Δm_s analyses. These latter have resulted in a combined lower limit, established from the statistical tool of the *amplitude scan*. These analyses have been later complemented by the Tevatron CDF and DØ experiments. The DØ analysis resulted in an upper limit. The CDF study provided the first precise measurement of Δm_s this year, with a $> 3\sigma$ significance. Here I review the current experimental scenario in the B-mixing sector. This is done in order to introduce this thesis work about flavour tagging into the general effort of observing B_s oscillations at CDF with the necessary statistical significance.

2.1 The situation in the B_d sector

Several analyses have been published by different experiments at e^+e^- and hadron colliders. All the 4 experiments at LEP searched for B_d mixing. Results looking at (mainly) charmed semileptonic decays of the B_d mesons created in $Z^0 \rightarrow b\bar{b}$ decays have been published (e.g. [11], [12]). Several methods were used to establish the flavour of the candidate b at production. They used the lepton charge produced

in the decay of the accompanying b in the other hemisphere and the average charge of the whole opposite hemisphere. The proper decay time was measured for each B meson by looking at the three dimensional distance between the decay vertex of the daughter D meson and the primary vertex in the e^+e^- collision. All the informations were merged into an unbinned fit to the proper decay time using the maximum likelihood method. The value of Δm_d was extracted from the fit. The total uncertainty was at the 10% level, out of which $\approx 30\%$ was contributed by systematic uncertainties. The sytematics were dominated by the sample composition (essentially B^+ fraction). Also lifetime fit uncertainties and background tagging performances were considered among the systematics. Just as an example we report here the value in [11]: $\Delta m_d = 0.436 \pm 0.033\ ps^{-1}$.

The current world average is anyway largely dominated by the measurements of the B-Factory experiments Belle and BaBar. They have looked for both semileptonic and hadronic decays of the B_d mesons produced by the decays of the $Y(4S)$ resonance: several flavour eigenstates have been considered ($B^0_d \to D^{(*)-}\pi^+$ and other light meson resonances [13], [16]; $B^0_d \to J/\psi K^{*0}$ [13]; $B^0 \to D^{*-}\ell^+\nu_\ell$ [14] or dileptonic events [15]). At a B-Factory, the oscillations of the two B mesons are completely coherent. Thus, the flavour of the candidate B can be inferred by looking at the charge of the decay products of the other B in the event. Leptons and Kaons were searched for and the final performances of the algorithm were improved by introducing neural networks (e.g. see App.A) to combine information from soft pions and leptons and Kaons not selected before. The analyses have been performed in time based on the difference between the decay time of the tagged B and that of the candidate B. A likelihood-based fit was performed, that returned the mass difference together with several other parameters. Among them, the tagger purity for the various flavour taggers, for signal and background, and the time resolution. Systematic uncertainties were still related to the estimate of the B^+ background. Relevant systematics arose also from the knowledge of the time scale, given by silicon detector alignments, model and resolution of signal and background time difference distributions. Correlation effects between the mis-tag rate and the time difference resolution were also taken into account. I quote here the BaBar result from [13]: $\Delta m_d = 0.516 \pm 0.016(stat.) \pm 0.010(syst)\ ps^{-1}$. This has been obtained with an integrated luminosity $\int \mathcal{L} = 30\ fb^{-1}$. Later analyses with $\int \mathcal{L} = 80\ fb^{-1}$ [14] measured $\Delta m_d = 0.511 \pm 0.007(stat.)\ ^{+0.007}_{-0.006}(syst)\ ps^{-1}$.

Also the Tevatron experiments, CDF and DØ, have contributed to the measurement of Δm_d, both in Run I and Run II. For example, CDF published works during Run I using both semileptonic and J/ψ events, selected by the dedicated triggers [17]. The proper decay length was reconstructed from the distance between the primary vertex and the B decay vertex by correcting for the B Lorentz boost transverse component. CDF used the leading products of the b quark hadronization process and the consequent associated production of lighter mesons to tag the B_d flavour at production (*Same Side tagging*). Such technique

has been adopted again for B_d (pion tags) and B_s (Kaon tags) during Run II and is the leading contribution to the CDF flavour tagging. The most recent result from CDF II on Δm_d has been performed with a semileptonic sample of $\int \mathcal{L} = 1\ fb^{-1}$. This used only the decay products of the away b to tag the flavour at production. The calibration of these taggers on the specific exclusive semileptonic sample was also performed in the same likelihood-based fit returning the mass difference. The bias of the incoherent mixing of the away b on the prediction of the candidate B_d^0 was also accounted for from Monte Carlo simulations. CDF II measures $\Delta m_d = 0.509 \pm 0.010(stat.) \pm 0.016(syst.)\ ps^{-1}$ [18]. The systematic uncertainty was dominated by the sample composition and background parameterization. Relevant contributions also came from tagger performances on the background and the knowledge of the absolute decay time scale. Similar results have been obtained by DØ [19].

Assuming no CP Violation in the mixing of the B_d and using the available average of the time-integrated mixing probability χ_d and the constraint from $\Delta\Gamma_d$ and Γ_d, the PDG finds

$$\Delta m_d = 0.507 \pm 0.005\ [21]. \tag{2.1}$$

Fig.2.1 displays the different published measurements and the current world average: the dominance of the B-Factory measurements can be noted.

Using the model-independent approach and notation described in 1.3, one can evaluate the agreement of such a value with the SM expectations. The most probable value for C_{B_d} is

$$C_{B_d} = 1.25 \pm 0.43\ \ [69] \tag{2.2}$$

This shows the good compatibility of the measured value with what expected in the SM. Also the phase has been measured to be $\phi_{B_d} = -2.9 \pm 2.0\ ^\circ$, again in good agreement with the SM request $\phi_{B_d} = 0$.

2.2 The Amplitude Scan as a statistical tool for Δm_s

The experimental scenario in the B_s sector has been as much productive. LEP experiments and the SLD detector at SLAC have contributed in the past with several analyses in search for B_s oscillations. Since the mixing frequency in the B_s sector is much larger than in the B_d, a better time resolution is requested for the former to reconstruct a single oscillation period. The proper time resolution is essentially made-up of two terms:

$$\sigma_t \approx \frac{m_B}{p}\sigma_L \oplus ct\ \frac{\sigma_p}{p} \tag{2.3}$$

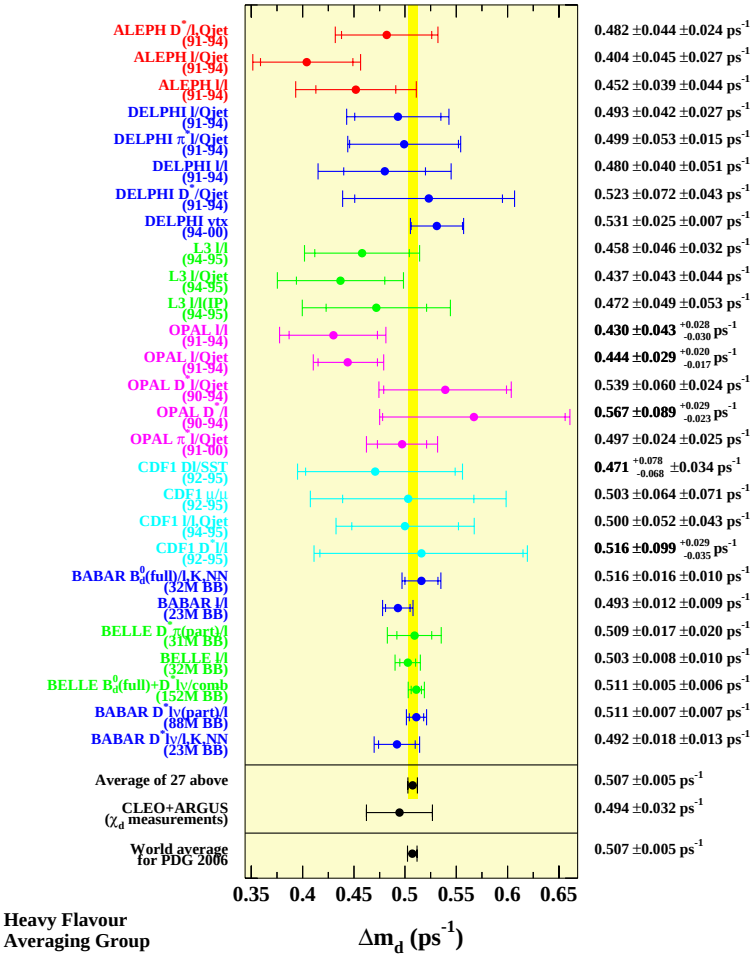

Figure 2.1: Table of the various published Δm_d measurements and World Average compiled by the Heavy Flavor Averaging Group [37]

where L is the decay length. Since most of the past analyses used semileptonic decays, where a fraction of the B momentum p is unmeasured due to the neutrino, the second term dominates and increases the final time uncertainty. Thus, despite several analysis with different statistics, were carried out, up to now it has been possible only to establish lower limits for Δm_s.

In Tevatron Run II, the SVT trigger provides CDF with the unique feature to select pairs of displaced tracks directly. This gives access to fully reconstructed hadronic decays having a much better proper time resolution than the semileptonic modes, so that a higher sensitivity to the phenomenon is achieved.

In order to combine all the results from the various experiments by weighting properly all the different samples used, and assess a statistically consistent lower limit, a statistical tool has been developed [20], called the *Amplitude Scan*. Infact, since we are dealing with a periodic signal (mixing) analysis, it is useful to transform the probability of mixing (absence of mixing) calculated in the time domain (derived in Sec.1.2)

$$P_{\pm}(t) = \frac{\Gamma_s}{2} e^{-\Gamma_s t}[1 \pm cos(\Delta m_s t)] \tag{2.4}$$

into the frequency domain using the Fourier Transform. The cosine function is transformed into a Dirac delta

$$FT[cos(\omega t)](\nu) = \sqrt{\frac{2}{\pi}}\delta(\nu - |\omega|) \tag{2.5}$$

The proper time resolution model (essentially a Gaussian with mean 0 and sigma the detector resolution) and the effective dilution of flavour tagging modify the otherwise pure cosine-like dependence of $P_{\pm}(t)$. The convolution of these functions turns into a simple product when the transform is applied, such that the amplitude of the peak at $\nu = |\omega|$ is diluted. The idea proposed in [20] is to switch from oscillations of the B_s in the proper-decay time to a search for such a peak just as like as one would do for a new resonance. The sensitivity of an experiment to the oscillations at the frequency ω is obtained when the amplitude of the peak at ν_{sens}, $a(\nu_{sens})$ is such that

$$a(\nu_{sens}) = 1.645\ \sigma_a(\nu_{sens}) \tag{2.6}$$

In practice, [20] describes a new method, the amplitude scan, as an hybrid between the Fourier analysis and the method of the maximum likelihood fit. This consists of introducing an *amplitude* $\mathcal{A}$ in the time-dependent probability $P_{\pm}(t)$ of 2.4, such that it becomes:

$$P_{\pm}(t) = \frac{\Gamma_s}{2} e^{-\Gamma_s t}\mathcal{A} \cdot [1 \pm cos(\Delta m_s t)] \tag{2.7}$$

The likelihood fit is repeated at various probe frequencies: the value of $\mathcal{A}$ will be compatible with 1 if mixing is observed at a given frequency, else it is expected to be null. Since

$$\mathcal{A} = \frac{d(\nu)}{a(\nu)}, \tag{2.8}$$

where $d(\nu)$ is the amplitude of the FT at the frequency ν and $a(\nu)$ the amplitude of the peak at the mixing frequency $\nu = |\omega|$, one can use the above relation for the Fourier sensitivity to establish the sensitivity of the amplitude scan for a given experiment:

$$\mathcal{A}(\nu) = 1.645\ \sigma_{\mathcal{A}}(\nu) \tag{2.9}$$

From this, one can also derive the lower limit on ω for one or more experiments as the frequency ν_l at which

$$\mathcal{A}(\nu_l) + 1.645\ \sigma_{\mathcal{A}}(\nu_l) = 1 \tag{2.10}$$

In [20], the expression of $\sigma_{\mathcal{A}}$ is also given:

$$\frac{1}{\sigma_{\mathcal{A}}} = \sqrt{\frac{\mathcal{S}\varepsilon D^2}{2}} \times exp\left[-\frac{(\Delta m_s \sigma_t)^2}{2}\right] \sqrt{\frac{\mathcal{S}}{\mathcal{S}+\mathcal{B}}} \tag{2.11}$$

where $\mathcal{S}$ is the number of B signal events, $\mathcal{B}$ is the corresponding amount of background; the uncertainty reduces with the number of signal events with a flavour tag associated, $\mathcal{S}\varepsilon D^2$, where εD^2 is the overall flavour tagging effectiveness of an experiment (and is defined in more detail in Sec.4.3). The last term shows that the uncertainty increases with the oscillation frequencies and with the experiment's B proper decay time resolution σ_t.

The amplitude method has been used up to now to evaluate the sensitivity and lower limits of the B_s mixing analyses of the various experiments and to combine them into one single lower limit, summarized in the next section. It is also the method adopted by CDF to look at its data, although a pure likelihood fit is used to determine Δm_s in case $\mathcal{A} = 1$ is found.

2.3 Current knowledge of Δm_s

The combination of all the different (published) measurements of the amplitude $\mathcal{A}$ by LEP, SLC and the Tevatron experiments (mainly Run I) results in a lower limit as obtained with the above statistical requirements:

$$\Delta m_s > 14.4\ ps^{-1}\ [21] \tag{2.12}$$

at a 95% C.L. The combined sensitivity of all the measurements as defined in previous section is 18.2 ps^{-1}. The presence of possible biases in any of the above

measurements was checked by comparing the measured values of a common set of parameters [37].

This result has been recently superseded by the results of CDF and DØ using $\mathcal{L} = 1\ fb^{-1}$ of data collected with the Tevatron during Run II. Using a sample of B_s semileptonic decays of the kind $B_s^0 \to \ell^+ D_s^{(*)-} X$, and with a flavour tagging that looks at the charge of the leptons from decays of the accompanying b hadron and at the mean charge of the away side (tagging effectiveness $2.48 \pm 0.21\ ^{+0.08}_{-0.06}$ %) [19], DØ sets a two-sided bound on Δm_s:

$$17 < \Delta m_s < 21\ ps^{-1} \quad (90\%\ \text{C.L.}) \tag{2.13}$$

The corresponding probability that the related minimum in the negative log likelihood is generated by random background fluctuations is 5%[22].

Last spring, CDF published a first evidence of B_s flavour oscillations and measured Δm_s with a significance of $> 3\ \sigma$ [23]. CDF has performed an analysis of B_s oscillations based on a sample of fully reconstructed and semileptonic B_s decays. A Run I-like set-up was used for flavour tagging in the away side, complemented with a Same-side tagger using Kaons. CDF achieved a 25.8 ps^{-1} sensibility, thus exceeding the combined sensitivity of all previous experiments. The impact of the CDF measurement on the amplitude profile is shown in Fig.2.2. The right

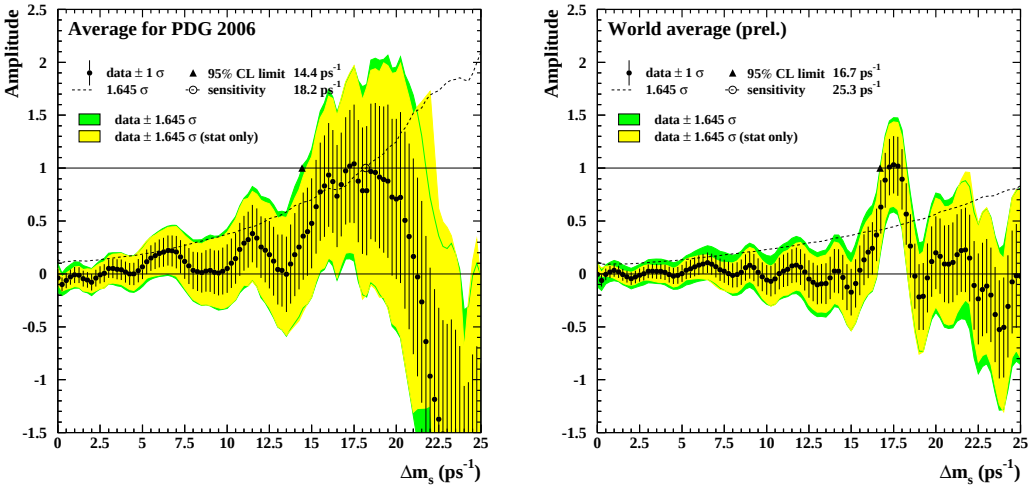

Figure 2.2:

plot corresponds to the most up-to-date world average on Δm_s: the sensitivity has increased to 25.3 ps^{-1} and the combined result is

$$\Delta m_s(W.A.) = 17.4^{+0.3}_{-0.2}\ ps^{-1} [21] \tag{2.14}$$

Provided its leading role in the search for B_s oscillations, CDF has undertaken an improvement of its analysis techniques in all its aspects, aimed at observing B_s

mixing with the necessary statistical confidence. In particular, the enhancement of flavour tagging is the main topic of this thesis and is achieved with the addiction of the Opposite Side Kaon Tagger and a refined combination of the Opposite-side taggers through a Neural Network technique.

Chapter 3

The experimental apparatus

The B_s[1] mesons are produced in $p\bar{p}$ collisions at the Fermilab[2] Tevatron collider. At present, this is the world most energetic particle collider. It collides protons and anti-protons at a centre-of-mass energy of $\sqrt{s} = 1.96\ TeV$, with a frequency of 2.5 MHz. The resulting events are analyzed by the CDF II and DØ detectors. Because of the large number of interactions per second and the wide range of physics processes originated in $p\bar{p}$ head-on collisions, event triggering is probably the biggest challenge for the two experiments, both in terms of process selectivity and decision time implied. Here I describe how collisions are generated and with which performances by the Tevatron acceleration complex. Then, I review briefly the various components of the CDF II detector. The sub-systems used for B_s mixing analyses, and in particular Particle Identification and flavour tagging, are treated in more detail. The 3-level system of on-line event selection is also described. In particular, the relevance of the SVT track-based trigger has to be stressed. It looks for decay products displaced from the collision point to select long-lived bottom and charm hadrons. This novelty, introduced in the upgrade of the CDF detector for the Run II of the Tevatron (started in the year 2000), provides CDF with the unique capability to trigger directly on fully reconstructed (hadronic) events, with a time uncertainty highly reduced with respect to semileptonic events. This feature provides CDF II with a prominent position to look for flavour oscillations.

3.1 Proton collisions

The Tevatron hadron collider is the final stage of a complex chain of a proton and anti-proton acceleration chain at the Fermilab. The whole complex is sketched in Fig.3.1. Protons are first produced in a ionization process of hydrogen gas contained in a Cockroft-Walton chamber. The H^- ions are accelerated by ap-

[1]Notation B_s means B_s^0 *and* $\overline{B}_s^0$ here and throughout this thesis
[2]Fermi National Accelerator Laboratory

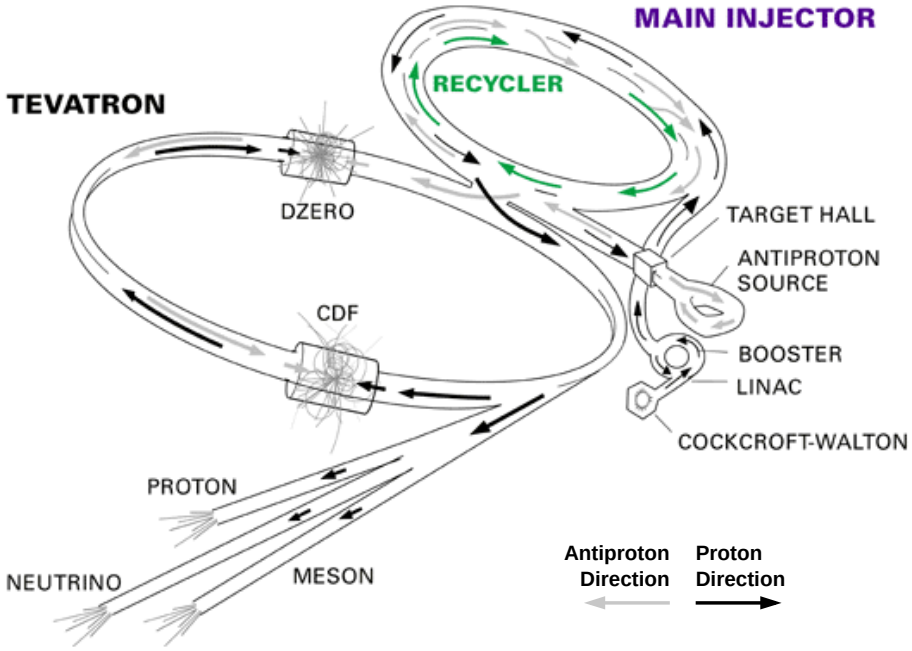

Figure 3.1:

plying a positive voltage to an average kinetic energy of 750 keV and selected by a further magnetic field. In the second stage of *pre-acceleration*, the ion energy is increased up to 400 MeV by a linear accelerator. The negative ions are made interact with carbon atoms in a foil to eliminate the electrons, eventually constituting a pure proton beam, which is then injected into a Booster. This is the first synchrotron in the chain, with a diameter of 150 meters, where the protons are further accelerated up to an energy of 8 GeV before reaching the Main Injector. This 500 meter radius synchrotron accelerates protons up to 150 GeV of energy. Six bunches containing each 6×10^{10} protons are accelerated to this kinetic energy and then merged together into a single bunch of 27×10^{10} p, ready to be injected to Tevatron. The cycle is repeated several times on protons until the Tevatron configuration of **36 proton bunches** is reached. At the same time, a portion of protons from the Main Injector is also used for $\bar{p}$ production.

An intense anti-proton supply is the largest challenge at a $p\bar{p}$ collider, as no obvious sources of anti-matter are available from nature. $\bar{p}$ are obtained by directing a bunch of $O(5 \times 10^{12})$ p onto a Nickel target. The products of these reactions are selected based on their momentum and charge using a system of magnets to form an 8 GeV energy $\bar{p}$ beam. These are stored in the Accumulator ring until an intense enough beam of anti-protons is gathered to be carried into the Recycler. For efficient acceleration and particle transfer from production to collision points, a strict phase-space confinement is desirable. The main source of

stochastic background eluding this requirement is thermal energy, so that different *cooling* techniques have been concieved at this stage in order to keep the $\bar{p}$ confined (see e.g. [24]). The most innovative is probably the *electron cooling* now in use, working on the principle of a heat exchanger: a beam of electrons with a small transverse kinetic energy (and approximately the same velocity as the $\bar{p}$'s) is made interact with the $\bar{p}$ beam by passing them collinearly with each other. The transverse kinetic energy of the warmer ($\bar{p}$) is transferred to the cooler (e), reducing the energy and, consequently, increasing the longitudinal momentum of the anti-protons. Thank's to the Liouville theorem on the conservation of the phase-space volume of a system in a conservative field, the trasverse distance among anti-protons is then reduced. It is in the Recycler that the electron cooling is performed. This important innovation introduced for Tevatron Run II exploits the same tunnel as the Main Injector. it can contain $\geq 150 \times 10^{10}$ $\bar{p}$ for as long as ≈ 600 hours. Here the energy of $\bar{p}$ is as well increased from 8 to 150 GeV and 36 anti-proton bunches with 33×10^9 particles each are sent to the Tevatron for last acceleration and collision.

At this last stage the energy of the two beams is increased up to 980 GeV each by circulating them in the 2 km diameter Tevatron synchrotron. Both beams contain 36 bunches, spaced 396 ns in time. Each of the bunches of protons (anti-protons) is 2.1 (1.8) ns long. The two beams cross each other in 72 points along the ring. Just before the $\bar{p}$ injection, a set of electrostatic separators are used to create a pair of non-intersecting closed helical orbits with the p and $\bar{p}$ circulating on the opposite strand of the helix. Thus, $p - \bar{p}$ only occur in the instrumented regions corresponding to CDF and DØ. The p-$\bar{p}$ interaction probability is enhanced by additional focusing performed with quadrupole magnets that reduce the beam spot. This increases the *instantaneous luminosity*. Approximating each beam to a cylinder (they have actually an elliptical section), we can define this quantity as

$$\mathcal{L} = \frac{N_B N_{\bar{p}} N_p f}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \tag{3.1}$$

where N_B is the number of bunches; $N_{p(\bar{p})}$ is the number of protons (anti-protons) in a single bunch; f is the bunch revolution frequency (at Tevatron $f = 47.7$ kHz); and $\sigma_{p(\bar{p})}$ is the effective width of the (anti-)proton beam. Since the number of collisions − and thus the number of events of a given physics process − is proportional to $\mathcal{L}$, it is desirable to have an as-small-as-possible beam section and to increase the number of particles in a single bunch. It has to be noted that, instead, N_B is upper-limited by the dead time of detector components, such as the drift time in the CDF central tracker. This requires for the interaction frequency $f_{Int} \leq 2.5MHz$, thus limiting the number of circulating bunches. The Tevatron *peak* instantaneous luminosity is $\mathcal{O}(1.5 \times 10^{32}\ cm^{-2}s^{-1})$, with a recent record of $\mathcal{L} = 2.05 \times 10^{32} cm^{-2}s^{-1}$.

The instantaneous luminosity degrades over time as particles are lost and beams begin to heat up, mostly due to the long range beam-beam interactions. The duration of each data-taking run is about 15 hours, after which it is terminated and a new colliding cycle is begun, which uses p and $\bar{p}$ accumulated in the mean time.

The number of signal events of a given physics process, assuming ideally that no background is selected and that the selection efficiency for that channel is 1, is given by

$$N_{\mathcal{S}} = \sigma \times \int \mathcal{L} dt \tag{3.2}$$

where t is time and σ the cross section for the given process. The instantaneous luminosity provided by the Tevatron accelerator complex proves, thus, to be one of the most relevant quantities for precision measurement and searches at the CDF and DØ experiments. It is displayed as a function of the running period in Fig.3.2, where the constant raise is evident.

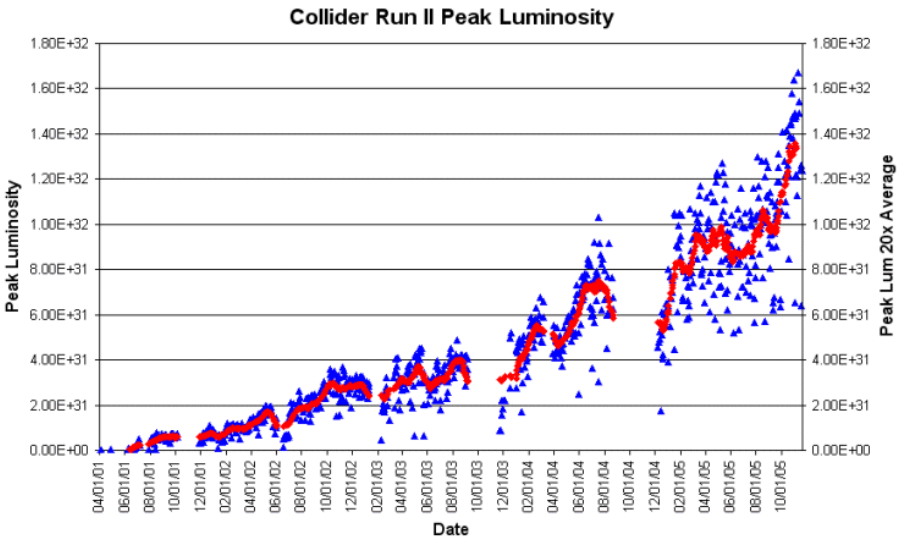

Figure 3.2: Instantaneous luminosity as a function of the operation year of the Tevatron Run II. Red points are the average over a period of blue points. Units of y axis are $cm^{-2}s^{-1}$.

The number of $p\bar{p}$ primary interactions is distributed according to a Poisson law as a function of the instantaneous luminosity. The average number of Primary Vertices in a collision as a function of $\mathcal{L}$ is shown in Fig.3.3, from which one can argue the trigger effort necessary to isolate a process originated at a given vertex when at high luminosity regimes.

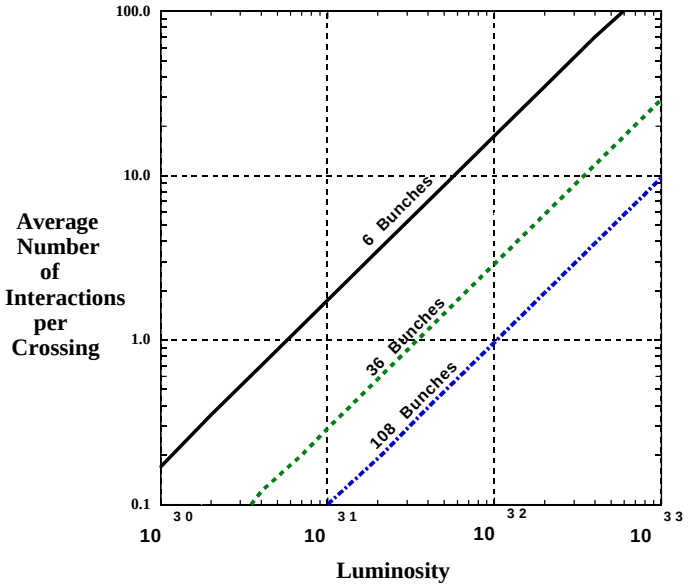

Figure 3.3: Average number of primary interactions as a function of the instantaneous luminosity, for different numbers of circulating bunches. The current configuration is the green curve.

3.2 The CDF II detector

CDF II is a general purpose solenoidal detector, with a standard for detectors of this type. From inside out, the following are present: tracking system (Silicon detector and drift chamber), magnet, electromagnetic (EM) calorimeter, hadron calorimeter and muon detectors. A Time-Of-Flight scintillator is also installed between the drift chamber and the solenoid magnet coil for Particle Identification at low momenta. An overview of the detector is pictured in Fig.3.4.

The coordinate system reflects the detector symmetries. The beam line determines the z axis. Positive z direction is given by the direction of the proton beam. The cartesian system is chosen with the y axis pointing up and the x axis completing this frame. Since the detector has a cylindrical symmetry, it is often convenient to use polar (r, θ, ϕ) or cylindrical (z, θ, ϕ) coordinates. The ϕ polar angle is defined as the angle in the transverse x, y plane, from the positive direction of the x axis and the θ polar angle is counted from the positive direction of the z axis. Nevertheless, the variable θ is usually replaced by the *pseudorapidity* $\eta = -log\ tan(\theta)$. This is the $E \to |\vec{p}|$ ultra-relativistic limit of the rapidity ζ, with the property that the quantity $\Delta\zeta$ is invariant under the Lorentz boost with which particles are produced in $p\bar{p}$ collisions. Using η has the advantage that it is a pure function of geometrical quantities.

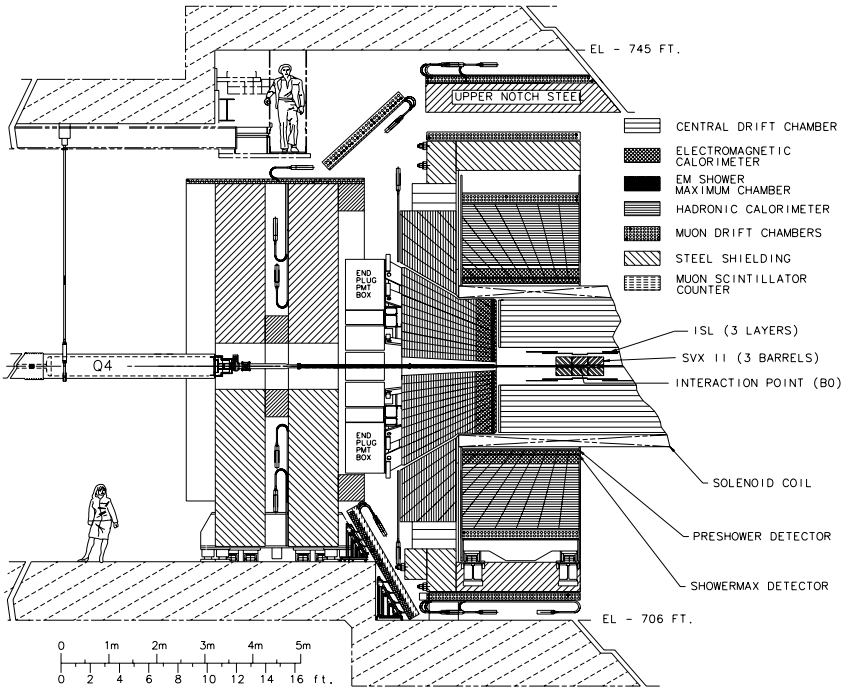

Figure 3.4: A section view of the CDF II detector.

3.2.1 Tracking

The trajectory of a charged particle is reconstructed exploiting its property to cause ionization as it traverses matter. Typically, in fact, the process is localized in little clusters, or *hits*, near the region of particle's passage which are used to form trajectories. The CDF II tracking volume is immersed in a $B = 1.4\ T$ uniform magnetic field directed along z. The consequent particle helical trajectories are described by 5 independent parameters:

- z_0, the z coordinate of closest approach to the beam line;
- d_0, the track **impact parameter**, defined as the projection of a track's distance of closest approach to the beam line onto the plane tranverse to the beam;
- ϕ_0, ϕ direction of the transverse component of the particle momentum at the point of closest approach to the beam line;
- $cot(\theta)$, the ratio of the helix step and its diameter;
- C, the helix curvature.

The determination of the trajectory in the magnetic field B provides a particle momentum measurement through the expression:

$$\begin{aligned} p_t &= \frac{cB}{2|C|} \\ p_z &= p_t cot(\theta) \end{aligned}$$

where p_t is the projection of the three-dimensional momentum $\vec{p}$ onto the xy plane and c the speed of light.

In order to determine the above parameters a track fit is performed. This is essentially a helical fit to a set of spatially distributed hits originated by the particle's passage through the CDF II tracking sub-detectors, the Central Outer Track (COT) and the Silicon detector. I describe the systems from the closest to the beam line outward and then illustrate how a track is reconstructed.

3.2.2 Vertexing

To perform a precise determination of the b hadron decay length one has to measure track parameters in the vicinity of the beam line with an excellent precision. On the other side, a detector so close to the interaction region is required to sustain large doses of radiation. Both these criteria are satisfied by silicon micro-strip detectors pioneered in a hadron collider environment by CDF [25]. Finely spaced strips of strongly doped p-type (p^+) silicon are implanted on a lightly doped n-type (n^-) silicon substrate ($\approx 300\ \mu m$ thick). On the opposite side a thin layer of strongly doped n-type (n^+) silicon is deposited. A positive voltage, applied to the n^+ side, depletes the n^- volume of free electrons and creates an electric field. A charged particle crossing the substrate leaves a trail of electron-hole pairs from the ionization, from whose drift a well-localized signal is originated. n^+ sides can also be manufactured, so to constitute a combined $r-\phi$ and $r-z$ positioning system. By weighting each of the adjacent strips on which a signal is generated by the collected amount of charge, a resolution of $\approx 12\ \mu m$ on the single hit is achieved at CDF. A sketch of a micro-strip is displayed in Fig.3.5. The overall CDF II Silicon detector is made up of 3 sub-systems: the Layer00 (L00), the Silicon VerteX detector (SVX) and the Intermediate Silicon Layer (ISL), from the closest to the farthest to the beam line.

Layer 00

The Layer 00 is the innermost layer of the CDF silicon detector. A series of micro-strips are assembled in six narrow (128 channels) and six wide (256 channels) groups in ϕ, as in Fig.3.6. The former are situated at a radius of 1.35 cm, the latter at 1.62 cm. There are six read-out modules in z, with two sensors bonded together in each module. These are made of light radiation-hard silicon and are single-sided, with a 25 (50) μm impant (read-out) pitch.

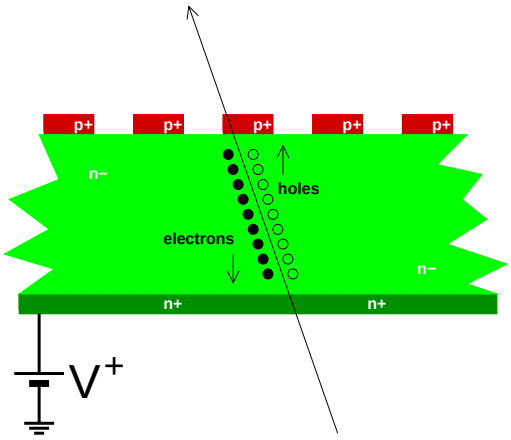

Figure 3.5: A silicon micro-strip detector.

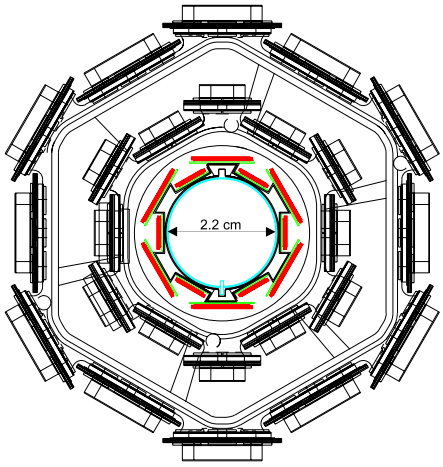

Figure 3.6: Transverse view of the Layer 00 along with the inner 2 layers of the SVX.

The importance of L00 resides in its proximity to the primary interaction region. This provides a substantial improvement in impact parameter resolution at low momenta, where multiple scattering originated from the passage through high-density regions of the remaining parts of the Silicon detector dominates. Consequently, this feature is particularly beneficial for the range of momentum of interest for B physics. The impact parameter performance of the tracking system is shown in Fig.3.7 with and without the inclusion of L00.

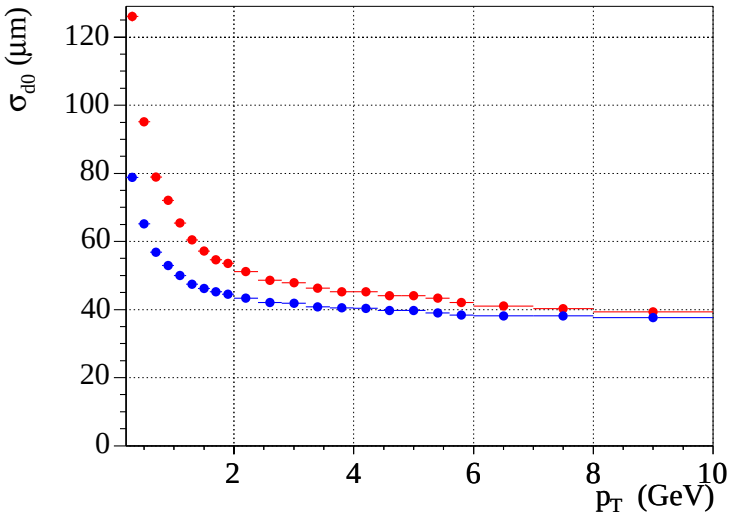

Figure 3.7: Impact parameter resolution for tracks with Layer 00 hits (blue) and without Layer 00 (red). The beam width contribution of $\approx 30\mu m$ is included in quadrature.

Silicon VerteX detector

The CDF SVX II is built in 3 cylindrical barrels each 29 cm long. Each of those supports 5 layers of double-sided silicon micro-strip detector, between radii of 2.5 and 10.7 cm. Layers 0, 1 and 3 combine $r-\phi$ measurement on one side with the 90° stereo $r-z$ detection on the other side, using the micro-strip mechanics described previously. Layers 2 and 4 are active in $r-\phi$ and at small stereo angle, 1.2°. The SVX layout is summarized in Tab.3.1. The actual silicon crystals (*wafers*) are supported by light substrates in assemblies called *ladders*. 12 ladders comprise a layer. The ladders are mounted bewteen two berillium bulkheads to form a single SVX barrel, shown in Fig.3.8. Each ladder is read out at both ends. The number of channels (> 400000) and the nature of the signals require that much of the electronics be mounted on the system instead of outside the detector, keeping the noise at a low level. On the other hand, this

SVX Layer	Radius, cm		N of strips		Stereo angle
	stereo	$r-\phi$	stereo	$r-\phi$	angle
0	2.55	3.00	256	256	90°
1	4.12	4.57	576	384	90°
2	6.52	7.02	6406	640	+1.2°
3	8.2	8.72	512	768	90°
4	10.10	10.65	896	896	−1.2°

Table 3.1: SVX mechanical layout.

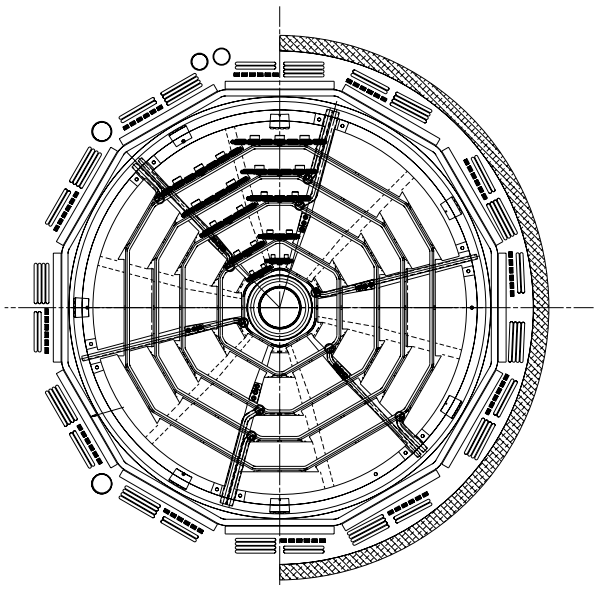

Figure 3.8: An $r-\phi$ plane cross section view of one SVX barrel.

causes additional multiple Coulomb scattering to affect a particle trajectory. For the purpose of vertexing and precision tracking, SVX is placed coaxial with the effective beam line, instead of the nominal direction given by the z axis. The asymptotic ($p_t \geq 8 - 10\ GeV/c$) impact parameter resolution provided by SVX is 40 μm, which includes also $\approx 30\ \mu m$ contribution from the beam width.

Intermediate Silicon Layer

The silicon system is completed by an outer detector that increases the lever arm for track reconstruction. In the central region ($|\eta| \leq 1.0$) a single silicon layer is placed at a radius of 22 cm. In the region $1.0 \leq |\eta| \leq 2.0$ two further layers are placed at radii 20 cm and 28 cm. Double sided silicon is used, where one of the two is stereo with an angle 1.2°. Only every other strip read out to reduce the total channel count. This affects the single hit resolution, which is ≈ 16 (23) μm on the axial (stereo) side, despite the fact that ISL and SVX layers are similar to each other in features.

Central Outer Tracker

The COT is a cylindrical multi-wire open cell drift chamber situated in the $|z| \leq 155\ cm$ region, between the radii of 44 and 132 cm. It operates with a 50:50 Ar-Ethane gas mixture as the active medium, which ensures a linear drift distance-time correlation, independent of the electric field applied. Ionization electrons extracted from gas by a particle are drifted toward the sense wires by cathode field panels and potential wires. The electrons move with an angle $\alpha \approx 35°$ with respect to the electric field lines due to the presence of the magnetic field. Since the best single-point resolution is obtained when the drift direction is perpendicular to that of the track, all cells are tilted by 35° away of the radial. The $r - \phi$ position of the track with respect to the sense wire is inferred from the arrival time on the wire of the signal produced by secondary electrons extracted in the ionization "avalanche". The maximum electron drift time is $\approx 100\ ns$. A charged particle traversing the entire COT in the radial direction ($|\eta| \leq 1$, $p_t \geq 400\ MeV/c$) ideally has 96 position measurements. The cells, each including 12 sense wires, are arranged in eight radially spaced superlayers (SL). A scheme of the SL is shown in Fig.3.9. These alternate between axial alignment (cell plane parallel to the z axis) and small angle *stereo* alignment (with cell planes rotated about radial planes of ±3°). Tab.3.2 explains the geometrical structure of the cells. Once the $r - \phi$ projection of the track is reconstructed from hit segments in the axial SL (only $\leq 5\%$ of tracks are estimated not to be matched as their circular projection is not recognized), we get a measurement of C, d_0 and ϕ_0 helix parameters. The remaining z_0 and $cot(\theta)$ are retrieved by matching the trajectory in the third dimension, as determined with the stereo SL, with the axial component. This completes the $\vec{p}$ momentum reconstruction.

COT Superlayer	Radius at center, cm	N of cells	Stereo angle
1	46	168	$+3°$
2	58	192	$0°$
3	70	240	$-3°$
4	82	288	$0°$
5	94	336	$+3°$
6	106	384	$0°$
7	119	432	$-3°$
8	131	480	$0°$

Table 3.2: COT geometrical summary. The number of cells per SL increses with the radius, in order for the drift of the distance to be approximately the same for any cell in any SL.

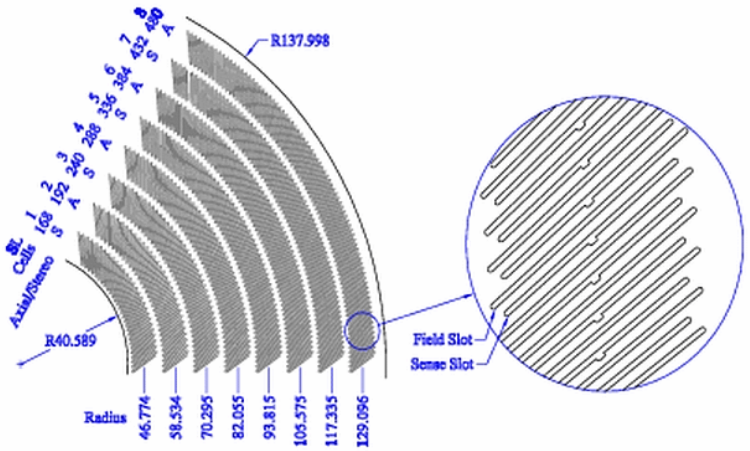

Figure 3.9: A sector of the COT endplate. Shown are the wire-plane slots grouped into 8 SL.

The COT has a single point resolution of

$$\sigma_{\text{single point}} \approx 140\ \mu m$$

This translates into a final resolution on the transverse momentum

$$\frac{\delta_{p_t}}{p_t} \approx 0.0017\ \frac{p_t}{GeV/c}$$

measured using cosmic ray events.

Track reconstruction

Since the Kaon tagger is essentially a track-based algorithm, it is convenient at this point to describe how the information from the silicon detector and the COT are merged together to reconstruct a particle trajectory. This is basically a helical fit to a set of spatially distributed hits preselected according to a patter recognition algorithm. As anticipated, the process is initiated with the reconstruction of the $r-\phi$ circular projection from the COT hits only and then the information from the stereo SL is added. Although this is already enough to determine the helix parameters, for a more precise tracking the SVX information is combined with the COT track into a unified track-fit (e.g. [26]). A precise inter-alignment between the two sub-detectors is crucial at this stage. In the standard tracking algorithm, the combination of the COT and SVX hits is performed by extrapolating the COT through to SVX. At each SVX layer, hits are matched to the track trajectory based on vicinity requirements. A new track candidate is generated when a new hit is appended and compared with the others based on the number of hits associated and the fit quality. The iterative comparison is performed until one final track is chosen.

Precise determination of track parameters is crucial for the Kaon tagger, as we will see in Ch.5. In fact, it allows for a separation of tracks coming from the Primary and Secondary Vertex, identifying tagging Kaons originated in long-lived b-hadron decays. The improvement on track parameters provided by combined Silicon-COT detector is listed in Tab.3.3: in particular the increase in d_0 resolution is beneficial to isolate signal Kaons displaced from the Primary Vertex.

Tracking asymmetry

A bias in the CDF tracking is induced by the asymmetrical track reconstruction efficiency in the COT. This is due to the 35° tilt of the COT drift cells (Fig.3.9): positively and negatively charged particles have opposite curvature in the CDF magnetic field, such that positive tracks traverse more of the inclined layer surface than negative tracks and are more likely to produce COT hits used in the track fit. A study of this effect has been undergone [27]. The effect is shown in Fig.3.10

	COT	COT + SVXII + ISL
$\delta p_t/p_t^2,\ (GeV/c)^{-1}$	2×10^{-3}	1×10^{-3}
$\delta d_0,\ \mu m$	600	40
$\delta z_0,\ \mu m$	5×10^3	70
$\delta cot(\theta)$	6×10^{-3}	4×10^{-4}

Table 3.3: Improvement in track parameter resolution using also the silicon detector information.

as a function of the track transverse momentum. The induced asymmetry is significant at lower momenta, as expected. Since the Kaon tagger uses the Kaon

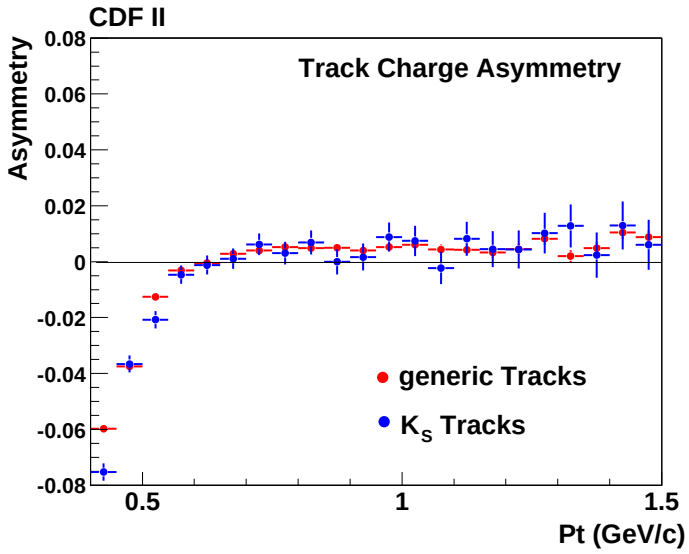

Figure 3.10: Tracking asymmetry induced by the COT cell asymmetric design as a function of the track transverse momentum. Red point are generic tracks, red points are K_S^0 tracks. From Ref.[27]

electric charge to infer the b flavour, possible effects of such a bias on the Kaon tagging sample are investigated, as discussed in sec.5.11.1.

dE/dx energy loss measurement

Not only the COT is used for tracking and momentum measurement. A particle energy loss per unit length dE/dx occurring in the COT volume is used in CDF to identify charged particle species. As known, the energy loss is described by

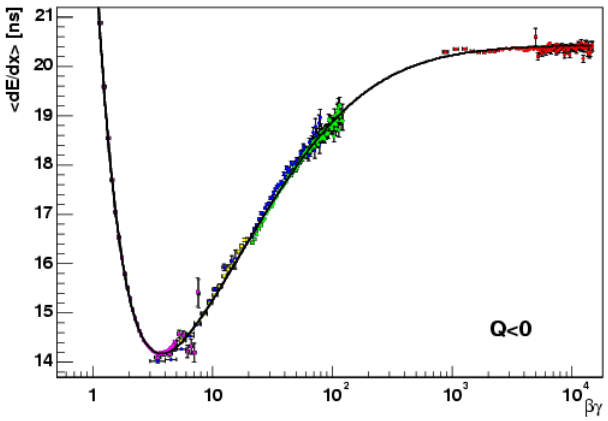

Figure 3.11: Measurement of the average energy loss per unit length as a function of particle $\beta\gamma$. The average is obtained from the charge collected at each of the drift cells of the COT. It is thus expressed in units of ns, as charge infromation is encoded in the pulse length of the COT electronics. From low to high $\beta\gamma$ protons (purple), K (yellow), π (blue), μ (green) and e (red) can be seen.

the Bethe-Bloch formula

$$-\frac{dE}{dx} = \kappa Q^2 \cdot \frac{Z}{A} \cdot \frac{1}{\beta^2}\left[\frac{1}{2}\, ln\left(\frac{2m_e c^2(\beta\gamma)^2}{I^2}\right)\, E_{\rm kin}^{\rm max} - \beta^2 - \frac{\delta}{2}\right] \qquad (3.3)$$

where $\beta\gamma = (|\vec{p}|/m)$, m is the particle rest mass and Q its electric charge; m_e is the electron mass; Z, A and I are the atomic number, atomic weight and mean excitation potential of the medium; $\kappa = 0.1535\ MeV\ cm^2/g$ accounts for other generic material features; δ is a correction factor for material density; and $E_{\rm kin}^{\rm max}$ the maximum kinetic energy transfer in a single particle-medium atom collision.

At CDF, the dE/dx is calibrated using data samples directly. We use physics samples pure in the different charged particle species to evaluate their specific ionization in the COT as a function of their $\beta\gamma$. Kaons and pions are extracted from a sample of $D^{*\pm} \to D^0\pi_*^\pm \to [K^\mp\pi^\pm]\pi_*^\pm$, where the K and π are identified exactly by their correlation with the charge of the π_*. Protons are obtained from $\Lambda \to p\pi^-$ decays, while pure lepton samples come from photon conversions (e) and $J/\psi \to \mu\mu$ decays. Particle-dependent corrections are applied in order to obtain a uniform response for any types with the track $\beta\gamma$. The result is shown in Fig.3.12, and the empirical fit (called "universal curve") to the various type distributions is overlaid.

Given the universal curve, one can predict a particle dE/dx for a given momentum and a mass hypothesis. By comparing the measured and the predicted value, particle identification probabilities and the most probable mass hypothesis can be thus assigned.

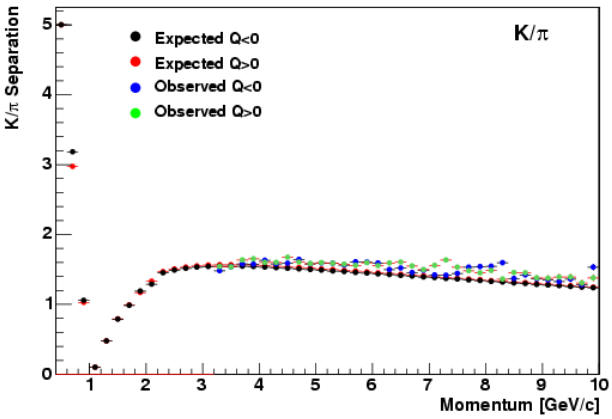

Figure 3.12: K/π separation power as a funtion of the track momentum: red and black dots are expected valued for $Q < 0$ and $Q > 0$ respectively; blue and green are the corresponding observed values.

I define *separation power* for two particle species A and B the distance, expressed in units of standard deviations, of the mean values of dE/dx distributions for the two species:

$$\text{Separation power} \equiv \frac{\text{dE/dx(A)} - \text{dE/dx(B)}}{\sigma(\text{dE/dx})_{\text{AB}}} \tag{3.4}$$

where $\sigma(dE/dx)_{AB}$ is the average resolution for the two particle species. This is an estimate of how well one is able to distinguish a given species based on the track specific ionization, in terms of the COT dE/dx resolution. This quantity is obtained from the distribution of the residuals $R(m_i)$ of the dE/dx for a given mass hypothesis (m_i):

$$R(m_i) = dE/dx_{MEAS} - dE/dx(m_i)_{i=\pi,K,p} \tag{3.5}$$

where $dE/dx(m_i)$ is the expected value for the given mass $m_i = |\vec{p}_i|/\beta\gamma$. The resolution results of the fitting of the residuals for each particle type with a suitable function (a sum of 3 gaussians) that accounts for the COT response to a particle energy loss. The $K - \pi$ separation power from Eq.3.4 is shown in Fig.3.12 above. This curve reflects the features of the "universal curve". A high separation is obtained at momenta $|\vec{p}_i| \leq 600 - 700\ MeV/c$: at such momenta, infact, while Kaons are already localized in the rising part of the curve due to the $1/\beta^2$, pions are already in the minimum of the ionization. As we get close to 1 GeV/c^2, both Kaons and pions are minimum ionizing and the separation reaches its minimum. Then, a smoother rise is found in the latter as pions exit the minimum of the universal curve and the Bethe-Bloch presents the logaritmic relativistic rise.

dE/dx simulation

A complete simulation of dE/dx is not available at present times in CDF. In order to include the effect of dE/dx into the Monte Carlo studies for the development of the Kaon tagger, in this work I use the predicted energy value for a given mass hypothesis retrieved from the 3.3 formula above. I generate a random number $b(dE/dx)$ from the available parameterizations of dE/dx residuals per particle type and charge, as obtained from 3.5. Then the residual is added to the predicted dE/dx by inverting Eq.3.5:

$$dE/dx(MC) = dE/dx(m_i) + b(dE/dx); \tag{3.6}$$

3.2.3 Time-Of-Flight

The Time-Of-Flight (TOF) system expands CDF Particle Identification capability in a momentum range complementary to the COT dE/dx. TOF measures a particle arrival time t_{TOF} with respect to the event production time t_0. The mass m of the particle can thus be inferred using the space-time relationship:

$$m = |\vec{p}| \times \sqrt{\frac{(ct_{TOF})^2}{L^2} - 1} \tag{3.7}$$

where L is the three-dimensional path length travelled by the particle to reach the TOF detector.

The TOF consists of 216 scintillator bars (using Bicron BC-408) situated at a radius 138 cm, in the 4.7 cm space between the outer shell of the COT and the the cryostat of the superconducting solenoid. Bars are 279 cm long and $4 \times 4\ cm^2$ in cross section. With cylindrical geometry, TOF provides 2π coverage in ϕ and roughly $|\eta| \leq 1$ on the longitudinal plane. Tracks hit TOF bars if they have $p_t \geq 350\ GeV/c$ and are central ($|\eta| \leq 1$). When a particle arrives on a TOF bar it originates scintillation photons that travel along the bar. These are eventually collected by two R5946 Hamamatsu PMT photo-multipliers (PMT) at each edge of the bar by use of Winston light-guides. The signal is passed to the fine-mesh amplification grid, made up of 19 dynode. This reduces the electron deflection in the CDF $B = 1.4\ T$ magnetic field and the consequent loss in signal gain. The observed signal is a factor 500 less than n the case when no magnetic field is applied.

The pulse on a TOF bar is associated with a reconstructed track by the TOF *matching* algorithm. This makes use of informations from both PMT's on the two sides of the bar. The z_{pulse} coordinate where the pulse occurred along the bar is, in fact, determined from the difference of the arrival times as recorded by the two PMT's using the expression:

$$\Delta t_{\text{east,west}} = \alpha + \beta z_{\text{pulse}} \tag{3.8}$$

where *east* and *west* refer to the two sides of the bar. α and β are calibrated using artificially generated signals. If one of the two PMT's has not been fired, no track is associated to that pulse. The track matching is based on the the agreement of charge and time measured by the two PMT's. These informations are combined into a χ^2

$$\chi^2 = \frac{(t_{east} - t_{west})^2}{\sigma^2_{east} + \sigma^2_{west}} + \frac{(Q_{east} - Q_{west})^2}{\sigma^2_{\Delta Q}} \tag{3.9}$$

and a cut $\chi^2 \leq 0.9$ is applied in order to associate the pulse to the track. We introduce the track-pulse matching efficiency, defined as the ratio

$$\varepsilon_{\rm match} \equiv \frac{N_{\rm tracks}({\rm matched})}{N_{\rm tracks}(|\eta| \leq 1,\ p_t \geq 350\ MeV/c)} \tag{3.10}$$

of the number of tracks matched with respect to the number of tracks within the TOF fiducial volume. We expect the efficiency to be a function of the track multiplicity. The high amount of tracks generated in $p\bar{p}$ collisions and in the subsequent secondary interactions produces an occupancy of TOF bars in the form of multiple hits. This worsens the above χ^2 and reduces the efficiency to find a match. Fig.3.13 shows the efficiency as calculated from pairs of muon tracks originated from a J/ψ decay and selected with the dedicated trigger ($p_t \geq 1.4\ GeV/c$): at higher luminosities the presence of additional vertices reduces the overall track matching efficiency, because the mean track occupancy per bar is increased. The efficiency to find a TOF match is also displayed as a function of the track p_t down to lower p_t ranges for an inclusive set of tracks, without any trigger bias ("minimum bias" sample). One possible explanation for the kinematical dependence is the following. At lower momenta, the population of the generic sample is dominated by pions from fragmentation and secondary vertices, like decays in flight or nuclear interactions of decay products onto the detector material. These particles are seldom associated with a proper tracks. Given the π lower mass with respect to, e.g., Kaons and protons, it is likely that these "spurious" particles have a higher velocity. Therefore, they reach TOF faster than reconstructed tracks from the primary interaction region and prevent the latter from being matched.

The uncertainty on the Time-Of-Flight measurement is made-up of two contributions: the detector resolution and the uncertainty assigned to the determination of the event production time t_0. The former is contributed by several effects: photostatistics; fluctuations in the time needed for photons to traverse the bar and reach the light guide; the travel time of electrons in the phototube; the jitter among the various front-end electronics channels. The t_0 is extracted as a parameter of a minimum likelihood-based fit to the Time-Of-Flight measured for the various tracks associated to the Primary Vertex. The two contributions can be separated using pairs of muons from J/ψ decays originating from the same

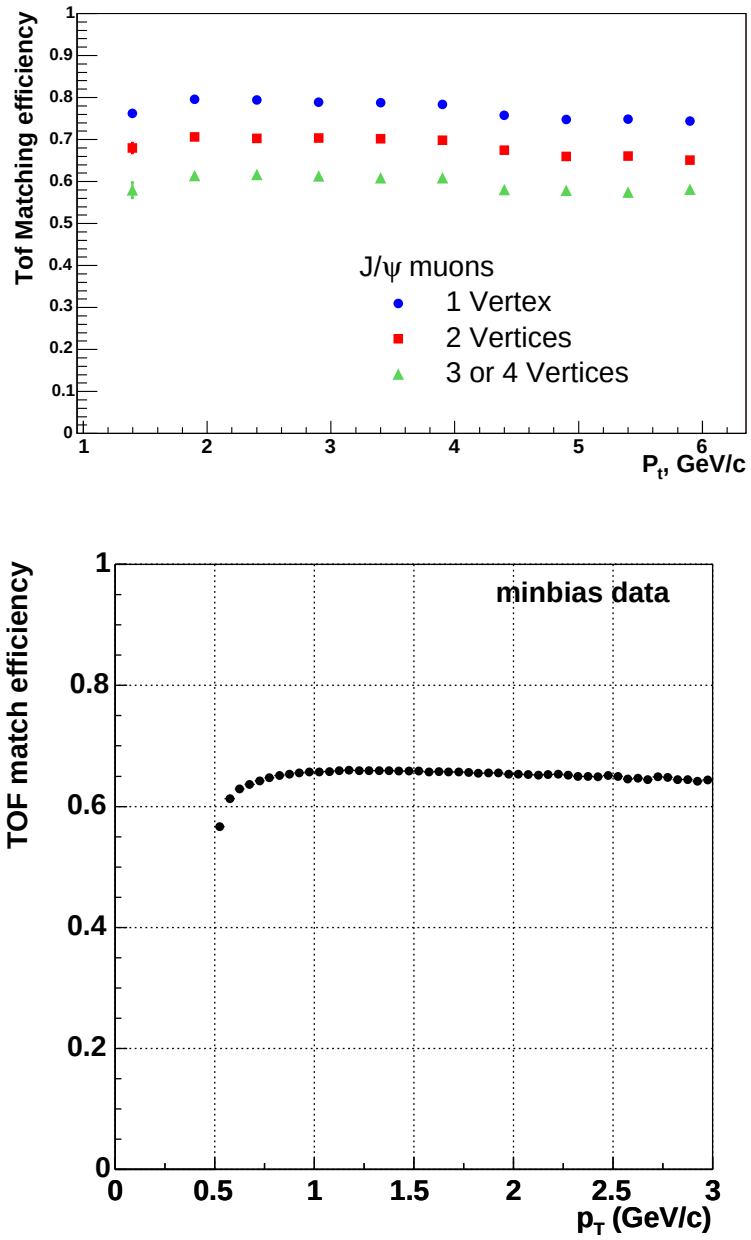

Figure 3.13: Upper plot: $\varepsilon_{\text{Match}}$ as a function of track p_t for muon track pairs from a J/ψ decay. The efficiency is shown as a function of 1 (blue), 2 (red) and 3 or 4 (green) Primary Vertices reconstructed along the z axis. Bottom plot: $\varepsilon_{\text{Match}}$ vs track p_t for generic tracks in a *minimum bias* sample (i.e. where no trigger bias is introduced).

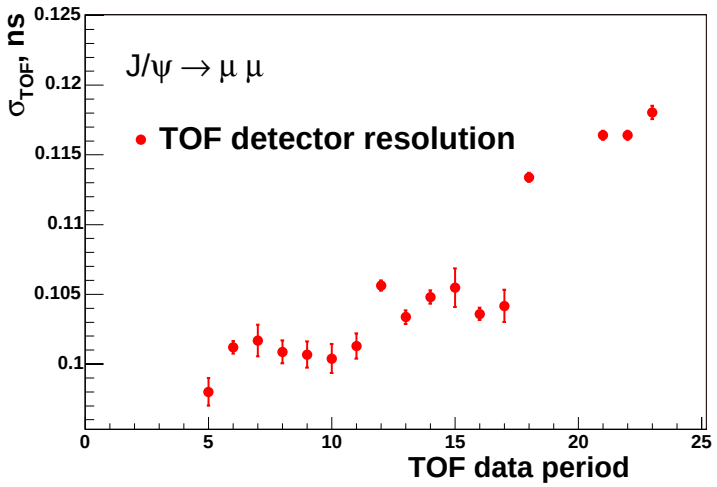

Figure 3.14: Instrumental component of the TOF resolution as a function of the data taking period. Each period corresponds to a new calibration table. The instrumental contribution is isolated from the t_0 component using pairs of muon tracks from J/ψ decays.

vertex [28]. For these, infact, the t_0 part cancels in the time difference

$$\Delta t = (t_{\mu_1} - t_{\mu_2}) = (T_{\text{measured},\mu_1} - t_0) - (T_{\text{measured},\mu_2} - t_0) \tag{3.11}$$

and the detector component of the uncertainty is extracted from the expression $\sigma_{\text{instrumental}} = \sigma_{\Delta t}/2$. We measure $\sigma_{\text{instrumental}} \approx 110\ ps$. Fig.3.14 shows the instrumental component of the resolution as a function of the data taking period. Note that the y axis scale starts at 95 ps. A degradation of the instrumental resolution of $\approx 20\%$ is observed in latest data with respect to the initial status. This is partially due to the fact that at present the PMT integrate less charge than in the past, as a result of the aging process. Consequently, a lower signal gain is obtained, resulting in a decrease of resolution.

The t_0 calculation uncertainty is of the order of 60 ps and has to be added in quadrature with the instrumental component. Based on Eq.3.7, one can assign a particle a mass hypothesis with TOF and separate particle species. The expected TOF separation power for π, K and p, as defined in the previous section for dE/dx, is shown in Fig.3.15, assuming a total TOF resolution ($\sigma_{\text{instrumental}} \oplus \sigma_{t_0}$) of 125 ps. As seen, a $K-\pi$ separation of $> 3\sigma$ is obtained for a 1 GeV/c track, while at 2 GeV/c we are left with only 1σ separation, making TOF only marginally beneficial for Particle Identification above this threshold. The second plot in Fig.3.15 shows the particle β for π, K and p as a function of the momentum, for a generic sample of tracks.

As we will see in Ch.5, usage of TOF is one of the main ingredients to select

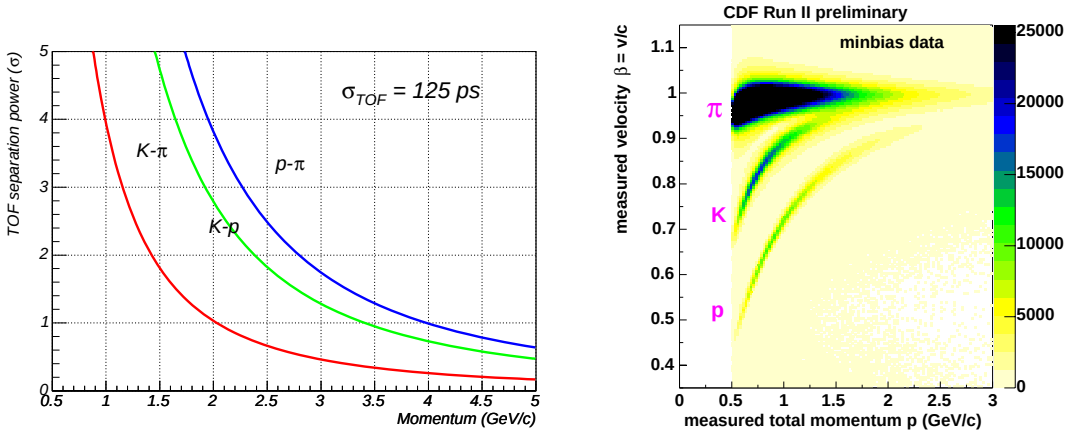

Figure 3.15: Left: Expected TOF separation power as a function of the track momentum $|\vec{p}|$, for various charged particle species. An overall TOF resolution of 125 ps is assumed. Right: $\beta \equiv |\vec{v}/c|$ distribution as a function of particle momentum $|\vec{p}|$ for a generic set of tracks, from which the effective TOF mass separation can be inferred.

signal Kaons for the OSKT. This is combined with dE/dx in a single Likelihood in order to extend the range of momentum where a good separation ($> 1\sigma$) is obtained. This is done using both the dE/dx and TOF Probability Density Functions (P.D.F.) as obtained from data samples. I already described how we obtain the P.D.F. for dE/dx in the previous section.
Concerning TOF, CDF parameterizes the detector response as a function of the difference $TOF - TOF(m_i)$, where m_i is the particle mass hypothesis assigned [30]. Pure samples of K from $D^0 \rightarrow K\pi$ decays, π from $K^0_s \rightarrow \pi\pi$ decays and protons coming from $\Lambda \rightarrow p\pi$ decays are used. The template function is given by the sum of two gaussians per each of the three particle species. One narrower gaussian with a $\sigma = \sigma_{instrumental}$ includes well-measured tracks; one broader gaussian ($\sigma \approx 3 \times \sigma_{instrumental}$), accounting for a $\approx 20\%$ component of worsely-measured tracks. The validity of the parameterization at lower momenta ($p_t \leq 2\ GeV/c$) is checked using pions π_* from D^* decays ($D^* \rightarrow D[K\pi]\pi_*$). Electrons and muons are treated as pions. The effective TOF separation power as obtained from this parameterized curves, taking into account also the effects of tails, is depicted in Fig.5.5 together with that of dE/dx and the combined TOF+dE/dx. I will discuss the impact of the combined Particle Identification on the Kaon tagger in the dedicated chapter.

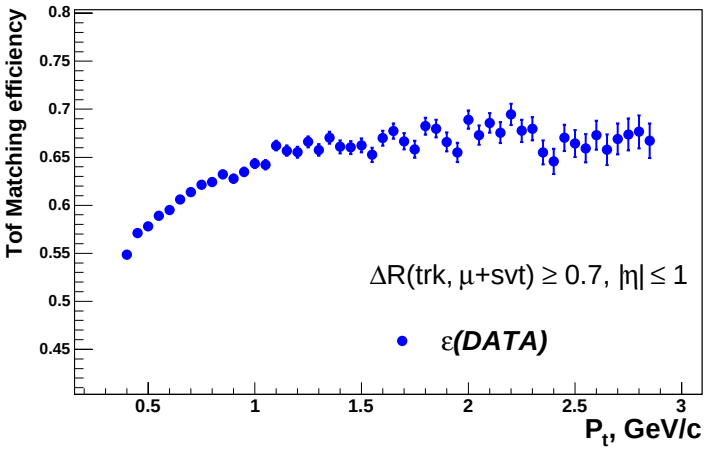

Figure 3.16: TOF matching efficiency for a set of tracks within the TOF fiducial volume ($|\eta| \leq 1$) and with $\Delta R < 0.7$.

TOF simulation

A full TOF simulation is available at CDF based on the `GEANT` package [31]. I proceed to evaluate the the agreement of Monte Carlo and data on the specific set of tracks selected for tagging and localized away from the trigger side (*Opposite side*). For this, I compare both the matching efficiency and time resolution distributions.

For efficiency, I find that the standard TOF Monte Carlo is not able to reproduce the large occupancy on TOF bars and generates too large effcencies at all transverse momenta. Consequently, it is necessary to scale it down to data. I use a p_t-dependent function:

$$\varepsilon_{\text{match}}(p_t) = A \times (1 - e^{-p_t/B}) \tag{3.12}$$

where A and B are retrieved from a binned fit to the $\varepsilon_{MC}/\varepsilon_{\text{data}}$ ratio of Fig.3.17. The result is displayed in the second panel of Fig.3.17. The absolute $\varepsilon_{\text{Match}}$ on data is displayed in Fig.3.16. The efficiency depends on the specific features of the sample used to evaluate it. Measuring it on the track sample used to calibrate the tagger (whose characteristics are described in Ch.5) I find the same trend with p_t as in the generic track sample of Fig.3.13; absolute values are also compatible with the latter, provided the kinematical differences of the two samples.

Regarding resolution, I compare data and Monte Carlo based on the distribution of the quantity $TOF - TOF(m_\mu)$ on a pure muon sample with $p_t(\mu) \geq 4\ GeV/c$. The result is depicted in Fig.3.18. Again I find that MC is not able to reproduce data in a satisfactory manner. In particular, the distribution from the standard simulation has a lower tail at lower measured TOF. This generates a

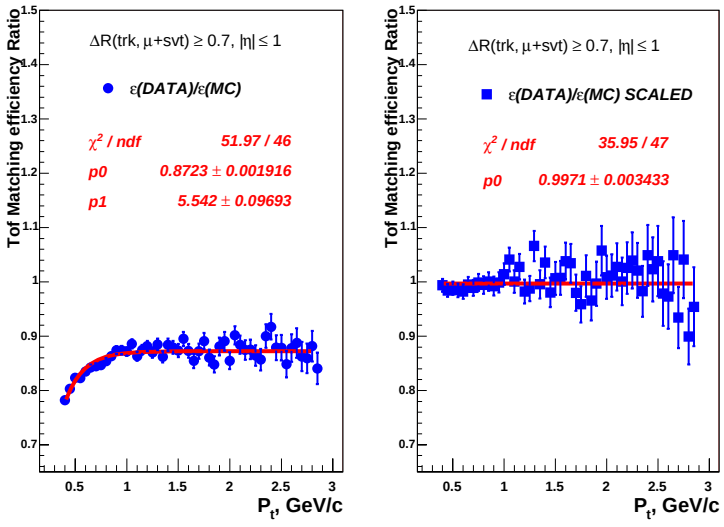

Figure 3.17: In the left panel the $\varepsilon_{MC}/\varepsilon_{\rm data}$ ratio is plotted and fit with the function of Eq.3.12. The right panel shows the same ratio after Monte Carlo efficiency has been scaled of this function.

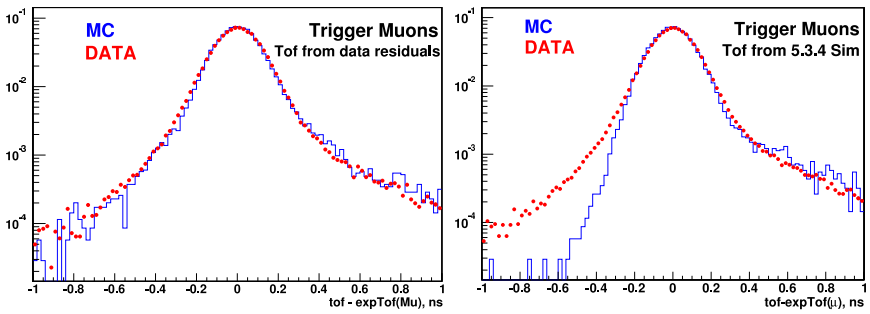

Figure 3.18: TOF resolution data-MC comparison using a pure muon sample. Data are red dots, MC is the blue line. The left plot is for TOF values generated from parameterization of TOF residuals on data. The right plot is for the outcome of CDF TOF standard simulation.

misidentification of real pions as Kaons. Even though I estimate that the effect on Kaon identification for the Opposite-side Kaon tagger would be tiny, I choose to generate TOF resolution based on the response of the detector on pure particle samples, obtained as in the previous section. A value of TOF for each matched track starting is generated from the expected TOF value for the given particle mass m_i, pathlength L_i and momentum $|\vec{p_i}|$. Then, a random number $a(TOF)$ extracted from the above parameterized curves is added:

$$TOF(MC) = TOF(M_i, L_i, p_i) + a(TOF). \tag{3.13}$$

The result of this operation is again shown in Fig.3.18 for muon tracks, which proves the goodness of the parametrization.

3.2.4 Calorimetry

CDF calorimetry has a uniformed pattern of matched towers of Electromagnetic (EM) and hadron calorimeters pointing back to the interaction region. The overall system is mechanically sub-divided into three regions: central, wall and plug with increasing pseudo-rapidity.

Central calorimeters cover 2π in ϕ angle and go as far as 1.1 (1.3) in $|\eta|$ for EM (hadron) calorimeter. The tower size is $0.1 \times 15^\circ$ in $\eta \times \phi$. Plug calorimeters extend η region coverage up to $|\eta| = 3.6$ and use variable segmentation optimized for physics targets. The EM calorimetry is performed using lead sheets interspersed with scintillator as the active medium; signal is read-out by PMT. Both calorimeters are equipped with shower maximum detectors, embedded at the depth of ≈ 6 radiation lengths X_0, which contribute to $e^{\pm}/\gamma$ identification using the position measurement to match with tracks.

A set of multiwire proportional chambers (CPR) is also present, in front of the central EM calorimeter, that enhance the CDF γ and soft electron identification by sampling the electromagnetic showers started in the solenoid magnet material ($1.075\ X_0$).

We quote here the resolution of the central EM and hadron calorimeter as a benchmark:

$$\begin{aligned} \sigma(E)/E &= 14\%/\sqrt{E\ sin\theta} \oplus 2\% \qquad \text{for EM calorimeter} \\ \sigma(E)/E &= 75\%/\sqrt{E\ sin\theta} \oplus 3\% \qquad \text{for Hadron calorimeter} \end{aligned}$$

3.2.5 Muon detectors

Muon identification in the CDF II detector is provided by four muon systems, situated within different $|\eta|$ regions around the outside of the hadron calorimeter. The Central Muon Detector (CMU) provides coverage at $|\eta| \leq 0.6$; the Central Muon Upgrade (CMUP) is un upgrade of the former, with the same $\phi - \eta$ coverage

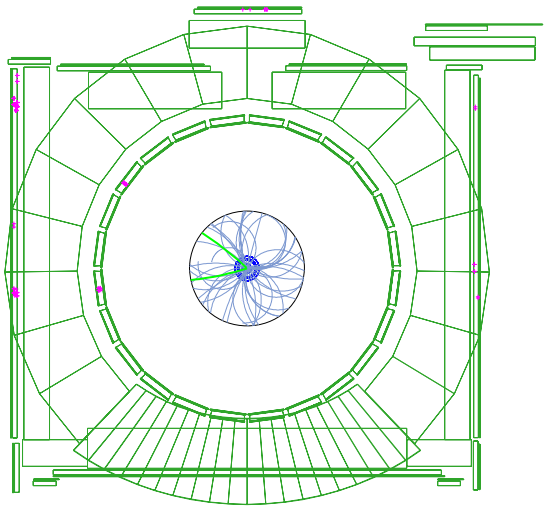

Figure 3.19: Di-muon event from a $J/\psi \to \mu\mu$ decay. Clusters of the purple dots outside the main tracking volume represent hits registered in the CMU (inner ring) and CMUP (thin outer square) muon systems.

as the CMU but is located further in radius, allowing for cleaner muon detection; the Central Muon Extenxion (CMX), covering the region $0.6 \leq |\eta| \leq 0.9$ and the Intermediate Muon Detector (IMU) ($0.9 \leq |\eta| \leq 1.5$). All of these are drift chambers that provide a reduced number of hits (e.g. CMU has 4 tracking layers with 4 cells each) per track. Drift time for muons can be as long as 1 μs, well beyond the time for one single collision. In order to assign stubs to the correct bunch-crossing, scintillators are placed next to the muon chambers.

Individual muon candidates are detected in the muon chambers from the track segment they leave (*muon stub*). A typical di-muon event is depicted in the Event Display, Fig.3.19. Hadronic punch-through and electronic noise can produce stubs as well. Only if a stub is matched with a track measured in the COT in both the ϕ direction and the z linear distance, the track is accepted as a muon. The rate of charged hadrons (π and K) misidentified as muons is given in Fig.3.20.

3.3 Triggers and online selection of data

At a hadron collider with a wide range of physics originated from $p\bar{p}$ collisions at a frequency of 2.5 MHz, a real-time selection of the events of interest is necessary and challenging at the same time. Furthermore, with event complexities and high

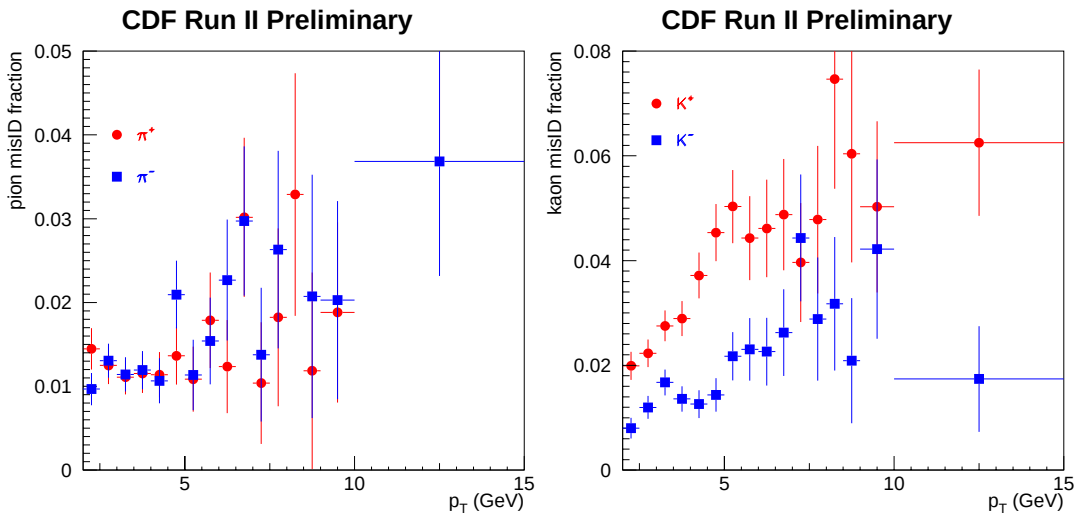

Figure 3.20: Rate of pion (left) and Kaon (right) tracks faking muon signal in the CDF II detector. A charge asymmetry for K is evident, due to different $K^+ - K^-$ absorption cross-section when interacting with the inner detector material.

multiplicities as at Tevatron, recording of all the detector information creates a problem of data storing. Thus, the CDF II triggering system has been designed based on three conditions:

- the trigger has to be able to analyze every single event, without losses due to dead time;
- the instrumentation (essentially the drift time in the COT) limits the track reconstruction for triggering to not less than 396 ns;
- the data logging system can write to tape about 100 events per second out of the 2.5 millions occurring, because of limited resources.

This is achieved by staging trigger decisions in three levels, as shown in Fig.3.22. Each level of the trigger is given a certain amount of time to reach a decision about accepting or rejecting an event. A higher level of complexity in the reconstruction tasks is reached at every stage, adding information from more detector components at each of them.

At Level 1, only a rough pattern recognition and filtering algorithm is performed. In order to do this in time, the Level 1 and Level 2 triggering mechanisms are implemented with custom electronics. The Level 3 is implemented with a PC farm of about 300 CPUs, instead. Using each CPU to merge the various subdetector informations allows for nearly one second to be allocated for the trigger decision. As a result, nearly an offline quality of event reconstruction is obtained

at the end of the triggering chain. The Level 3 rejection rate is about 10, resulting in 30 events/sec being accepted by the Level 3 trigger and written to tape.

The delay necessary to make a trigger decision is achieved by accumulating detector readout information into a storage pipeline. At Level 1, for every Tevatron clock cycle, the event is moved up one slot. By the time it reaches the end of the pipeline, the trigger will have reached a decision whether to accept or reject this event. If the event is accepted, its information will be sent over to the higher level of the trigger. Otherwise, the event is simply ignored. Since the Level 1 builder has 42 slots, the time allocated for making a trigger decision is about 5 μs. The rejection factor at this stage is ≈ 80, so that Level 1 accept rate is around 30 kHz. At Level 2 there are 4 buffers available, which results in 20 μs for the trigger decision. The accept rate is around 800 Hz.

3.3.1 Level 1

As said, the Level 1 is a synchronous system with an event read and an accept/reject decision made every bunch crossing. The decision is based on the transverse energy in the calorimeters, tracks in the COT and the stubs in the muon detector. Rudimentary versions of such objects are used at this level, called *primitives*.

In several cases track primitives are combined with calorimeter or muon primitives to form e, μ and jet objects as shown in Fig.3.22. The association is performed by two systems, that we describe hereafter.

XFT

EXtremely Fast Tracker (XFT), identifies track primitives, that is tracks with $p_t \geq 1.5\ GeV/c$, in the $r-\phi$ view using the four axial SL of the COT. Track identification is accomplished in two steps. First, all COT hits found are compared with predefined patterns and the XFT tries and find all segments likely to be part of the same track. Then, the ensemble of segments is compared to pre-defined roads to form a track. A crude estimate of the helix parameters is performed eventually and the track is stored for Level 2.

XTRP

A copy of an XFT track is also reported to the eXTRaPolator unit (XTRP), that extrapolates it to calorimeters and muon detectors, where a match with towers or stubs is attempted, respectively. This provides a more complex primitive, candidate to be a lepton, to be used for additional rejection at Level 1.

3.3.2 Level 2

The Level 2 is an asynchronous system which processes events accepted by the Level 1 in the time-ordered fashion. It uses Level 1 primitives as well as additional data coming from the shower maximum chambers in the central calorimeter and the $r-\phi$ strips of the SVX (Fig.3.22). There are 3 subsystems building primitives at this trigger stage: L2CAL, XCES and SVT.

The L2CAL hardware receives trigger towers from Level 1 calorimeter trigger and finds energy tower clusters by applying seed thresholds.

The XCES system generates bitmaps from shower maximum and, after track extrapolation, produces electron candidates.

SVT

The Silicon Vertex Trigger (SVT) represents probably the most important novelty of the CDF trigger system for the Tevatron Run II upgrade. In particular, its ability to select tracks identified as displaced from the Primary Vertex (PV) with great precision is one of the key ingredients for all the b and c analyses, among which the B_s mixing.

It uses SVX $r-\phi$ hits to extend XFT track primitives inside the SVX volume, closer to the beam line. This improves the XFT ϕ_0 and p_t resolution. Above all, it adds the measurement of the track impact parameter d_0 (original XFT track primitives are only beam line constrained). A significant impact parameter indicates a displaced track, providing a powerful handle to select long lived $b-$ and $c-$hadron decays. In particular, the SVT track reconstruction starts with the extrapolation of the primitive into the SVX, forming a *road*. Clusters of charge in the inner four $r-\phi$ layers of the given wedge have to be found inside this road. The silicon cluster information and the XFT segment information are fed into a linearized fitter which returns the measurements of p_t, ϕ_0 and d_0 of the track.

The resolution on d_0 from the SVT is depicted in Fig.3.21. Subtracting the $\approx 30\mu m$ beam width, the effective resolution on the impact parameter is $\approx 35\mu m$.

3.3.3 Level 3

In case an event is accepted by Level 2 the entire detector is read out, emptying a slot in the trigger pipeline for the next event. The read-out fragments are put in the proper order by the *Event Builder* system and the whole event is sent through to Level 3.

As anticipated, this last stage of the CDF trigger is software-based and consists of a computing farm with head nodes and several processor nodes. The head node receives the ordered sequence of informations on an event from the Event Builder and assembles them altogether in an *event record* in the form of

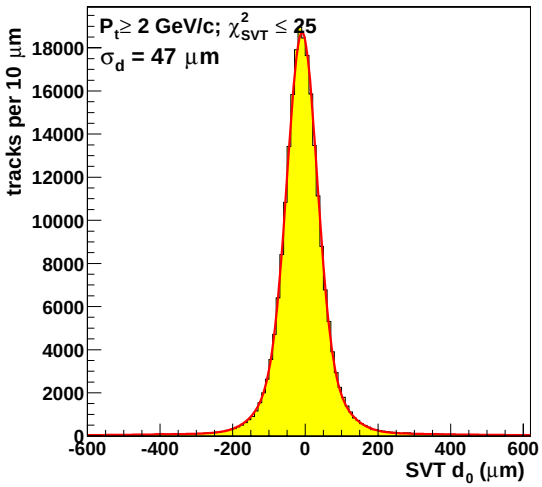

Figure 3.21: SVT impact parameter resolution for tracks with good-quality track-fit and $p_t \geq 2\ GeV/c$. $< \sigma(d_0) >= 47\ \mu m$.

database banks available for off-line analysis usage. Each event record is fed to one of the processor nodes, where event reconstruction is performed and final trigger requirements are applied. At this level, advantage is taken of all detector components and improved resolution.

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a *trigger path*. Requiring that an event be accepted through a well defined trigger path eliminates possible events which passed a higher level trigger requirement but did not satisfy lower level trigger selections.

The CDF II trigger system implements about 100 trigger paths. Here we describe those relevant for the selection of B_s hadronic and semileptonic samples used in flavour oscillations and for tagger calibrations.

3.3.4 Hadronic Trigger

The hadronic decays of b and c hadrons are selected at CDF II using the so-called *Two-Track-Trigger* (TTT). The strategy of this trigger path is as follows. At Level 1, fast measurements of track momenta are available from the XFT. By cutting on track momenta and angles, most of the inelastic background will be rejected.

For the Two-Track-Trigger, an event is accepted at Level 1 if two tracks are found in the event such that they have:

- opposite electric charge;
- $p_t \geq 2\ GeV/c$;

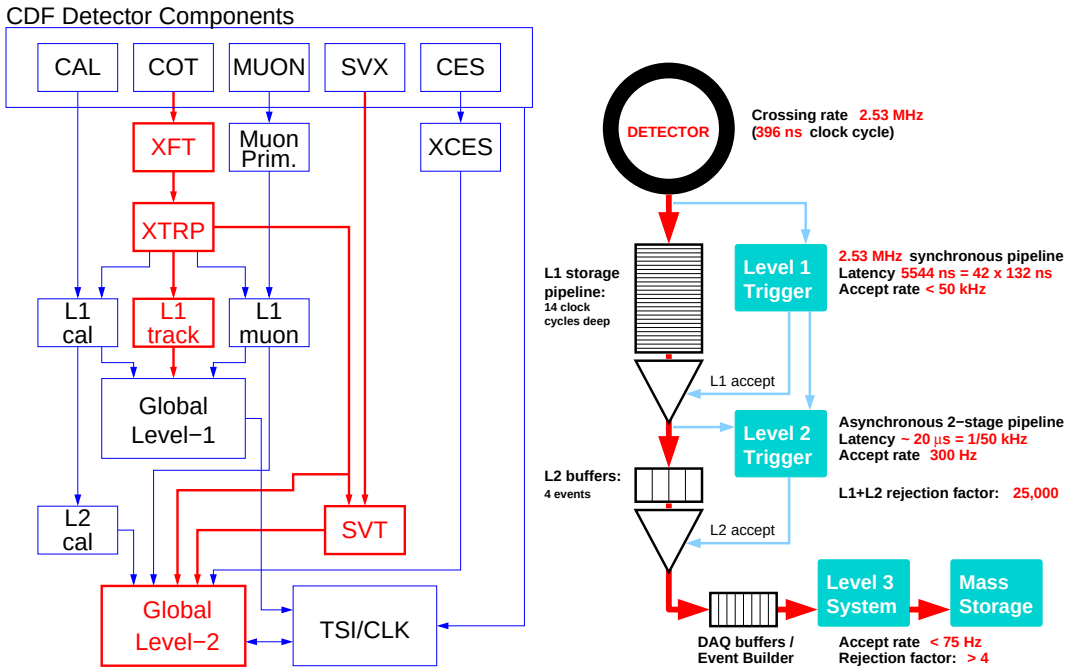

Figure 3.22: Left: Diagram of the trigger sequence at Level 1 and Level 2. The data flow for the trigger paths is outlined in red. Right: System of storage pipeline.

- the scalar sum of their transverse momenta is $p_t(1) + p_t(2) > 5.5\ GeV/c$;
- the angular separation between the tracks at COT SL 6 is $\Delta\phi_6 < 135°$.

At Level 2, the additional information on the impact parameter measurements is used from the SVT. An explicit requirement on a track impact parameter d_0 is applied at this stage, to select decays of long-lived $b-$ and $c-$hadrons. An event passes Level 2 selection if there is a track pair reconstructed in the SVT such that the tracks

- $120\ \mu m \leq |d_0| \leq 1000\ \mu m$;
- $2° \leq |\Delta\phi_0| \leq 90°$.

A further displacement requirement is applied, using the distance in the transverse plane from the beam line to the two-track vertex projected onto the two-track momentum vector, L_{xy}. Level 2 requires $L_{xy} > 200\ \mu m$. Selections on the pair of tracks performed at Level 1 are also confirmed.

The Level 1 and Level 2 conditions are repeated at Level 3 using improved track measurements and a complete event reconstruction is performed.

3.3.5 Semileptonic Trigger

Opposite-side taggers at CDF are calibrated using a high statistics sample of inclusive semileptonic b decays, as we will describe in sec.4.3.2. It is convenient at this point to describe how this sample is selected at the trigger level.

The semileptonic trigger requires a *lepton object* ℓ and an SVT track SVT, from which the path is called $\ell + SVT$. Two versions exist, for $\ell = e$ and $\ell = \mu$.

$e + SVT$

For the $e + SVT$ path the event selections is started by requiring at Level 1 that:

- a primitive e is found in the central calorimeter;
- $p_t \geq 4\ GeV/c$;
- hadronic to e.m. energy ratio in the calorimeter $E_{had}/E_{EM} < 0.125$.

At Level 2, the SVT information is required and an SVT track is looked for. The following requirements are set in addition to the ones of Level 1:

- $E_t(e) \geq 4\ GeV/c$;
- $p_t(\text{SVT track}) \geq 2\ GeV/c$;
- $120\ \mu m \leq |d_0(\text{SVT track})| \leq 1000\ \mu m$;

- SVT track with ≥ 4 r $-$ ϕ SVX hits;
- $2^\circ \leq |\Delta\phi_0| \leq 90^\circ$.

Several other quality requirements are applied at Level 3 in order to enhance the previous selection. The most relevant are:

- $M(e + SVT) \leq 5\ GeV/c^2$;
- ≥ 3 calorimetric towers for e;
- $E_{had}/E_{EM} < 0.080$.

where $M(e + SVT)$ is the combined mass of the two trigger tracks, where the SVT track is refit assigning it a pion mass. Th evalue was optimized from Monte Carlo for b and c semileptonic decay selection. The hadron to e.m. energy ratio cut is here tightened with respect to Level 1 requirement. A matching is also performed for the electron between the reconstructed COT track and the calorimeter deposit.

$\mu + SVT$

A similar logic is adopted for the selection of muon objects as it is for the electromagnetic objects. The information of the muon detectors is used instead of the e.m. calorimeter. The requirements on the SVT track are exactly the same as for the $e + SVT$ trigger path.

Eventually, at Level 3 an event is selected according to the $\mu + SVT$ path if

- one muon object is found with $|\eta| \leq 1$;
- the muon produced stubs in both CMU and CMUP;
- $p_t(\mu) \geq 4\ GeV/c$;
- p_t(SVT track) $\geq 2\ GeV/c$;
- $120\ \mu m \leq |d_0(\text{SVT track})| \leq 1000\ \mu m$;
- SVT track with ≥ 4 r $-$ ϕ SVX hits;
- $0^\circ \leq |\Delta\phi_0| \leq 90^\circ$.

Again, a match is performed between the reconstructed COT track and the muon stubs at Level 3.

Chapter 4

B_S mixing and flavour tagging at CDF

The value of the mixing frequency Δm_s can be extracted from the probability density of Eq.1.21, using the method of the maximum likelihood. The sample considered for the analysis corresponds to an integrated luminosity $\int \mathcal{L} = 1\ fb^{-1}$. In order to perform a likelihood based fit in the proper time domain, we:

- reconstruct the B_s final states in both hadronic ($\bar{B}_s^0 \to D_s^+\pi^-, D_s^+\pi^-\pi^+\pi^-$) and semileptonic ($\bar{B}_s^0 \to D_s^{+(*)}\ell^-\bar{\nu}_\ell, \ell = e, \mu$) decay channels, looking for charged particles in the final state;

- calculate the proper decay time for each B_s using the distance of the production and decay vertices in the transverse plane, the reconstructed momentum and the meson mass $m(B_s) = 5.3696\ GeV/c^2$ [21];

- determine whether the meson contained a b or $\bar{b}$ quark when it was produced, using both the correlation of the flavour with the leading products of the fragmentation that originated the B_s and the flavour of the other b created.

In this chapter we describe the various aspects of the CDF time-dependent analysis; the innovations in the flavour tagging sector introduced by this work are described more thoroughly in Ch.5 and 6. The yields for the various channels and the final performances are reported along with the results of the CDF analysis in Ch.7.

4.1 B_s final state selection

The hadronic and semileptonic modes are complementary: while the first have a better proper time resolution and thus provide us with a better sensitivity to

rapid oscillations, the semileptonic sample is several times larger in size than the hadronic. Nonetheless, its decay time resolution is worsened due to the unmeasured ν_ℓ momentum. Both hadronic and semileptonic decay modes are selected using the three-level Two Track Trigger (TTT) based on the SVT. The requirements applied to pairs of tracks to select the event are described in sec.3.3.4. Through the TTT, CDF has access to fully reconstructed decays, by exploiting the kinematics of long-lived charm and bottom hadron decays. As said, this is a unique feature inthe experimental approach to flavour oscillations.

The ensemble of events satisfying the trigger conditions is dominated by the decays of prompt charm mesons and requires an explicit B_s vertex reconstruction. This begins by selecting D_s meson candidates from several final states: $D_s^+ \to \phi\pi^+, K^*(892)^0K^+, \pi^+\pi^-\pi^+$, where the resonances decay as $\phi \to K^+K^-$ and $K^*(892)^0 \to K^+\pi^-$ respectively. These are required to be compatible with the known mass and width values [21]. The tracks forming a D_s candidate are constrained to originate from a common vertex in the transverse plane (r-ϕ coordinates) through an upper cut on the bi-dimensional χ^2 of the vertex fit, $\chi^2_{r\phi}$.

The resulting D_s mesons are then associated with other tracks to form $D_s^+\ell^-$, $D_s^+\pi^-$ and $D_s^+\pi^-\pi^+\pi^-$ vertices. A spatial constraint is introduced for them plus any other tracks associated to the B_s to originate from the same two-dimensional decay vertex, by imposing an upper cut on the B_s vertex fit $\chi^2_{r\phi}$. The D_s decay invariant mass is constrained to the known value for the D_s meson mass [21]. For the $D_s3\pi$ modes a requirement on the three pion tracks to originate from the same vertex is also applied.

4.1.1 Fully reconstructed hadronic decays

To reduce the contribution of the combinatorial background, a cut on the L_{xy} significance ($L_{xy}/\sigma(L_{xy})$) of both the B and the D vertex is applied. The B_s is also required to have a consistent p_t and to have originated from the primary vertex, by applying an impact parameter upper cut. The cut values are summarized in Tab.4.1.

After having established a set of hadronic B_s candidates in both the $D_s\pi$ and $D_s3\pi$ modes, the signal is separated from the remaining contribution of combinatorial background and prompt charm decays using a Neural Network (see Appendix A). The `ROOTSNNS` [36] package is chosen. The NN is trained using Monte Carlo samples of both the signal and the mass sidebands (regions of the mass spectrum away from the fully reconstructed signal) as an approximation of the background. A cut optimization is performed by maximizing the $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$ discriminant, where $\mathcal{S}$ and $\mathcal{B}$ are the number of signal and background events for a given mode, estimated from Monte Carlo and mass sidebands respectively. Eventually, a cut on the NN output among 0.954 and 0.996 is applied, depending on the different channels. A fit based on the maximum likelihood method is then performed to obtain the signal yields, B_s^0 mass and lifetime from the the various

	$B_s \to D_s\pi$ modes			$B_s \to D_s 3\pi$ modes		
	$D_s \to \phi\pi$	$D_s \to K^*K$	$D_s \to 3\pi$	$D_s \to \phi\pi$	$D_s \to K^*K$	$D_s \to 3\pi$
$\chi^2_{r\phi}(B)$	≤ 20	≤ 20	≤ 20	≤ 20	≤ 20	≤ 20
$\chi^2_{r\phi}(D)$	≤ 20	≤ 20	≤ 20	≤ 20	≤ 20	≤ 20
$\chi^2_{r\phi}(3\pi)$	≤ 50	≤ 50	≤ 50	≤ 50	≤ 50	≤ 50
$\frac{L_{xy}}{\sigma}(B)$	≥ 2.0	≥ 2.0	≥ 2.0	≥ 2.0	≥ 6.0	≥ 6.0
$\frac{L_{xy}}{\sigma}(D)$	–	–	–	≥ 2.0	≥ 6.0	≥ 6.0
$\lvert d_0(B)\rvert, \mu m$	≤ 200	≤ 200	≤ 200	≤ 200	≤ 200	≤ 200
$P_t(B), GeV/c$	≥ 5.5	≥ 5.5	≥ 5.5	≥ 5.5	≥ 4.0	≥ 4.0
$\lvert m_{K\pi\pi} - 1869.4\rvert, MeV/c^2$	–	≥ 16	–	–	≥ 19	–
$\lvert m_{K\pi\pi} - m_{K\pi}\rvert, MeV/c^2$	–	–	≥ 160	–	–	≥ 160

Table 4.1: Summary of preliminary selection cuts for all $B_s \to D_s(3)\pi$ decays, before passing events over to the NN. The following definitions are given: L_{xy} is the decay length of the B candidate in the $\rho\phi$ plane of the detector; $\chi^2_{\rho\phi}$ is the $\rho\phi$ component of the χ^2 of the B vertex fit.

modes. The invariant mass distribution for the $B^0_s \to D^-_s[\phi\pi]\pi^-$ *golden* mode is shown in Fig.4.1.
Several contributions are identified:

- signal events are modeled with a double Gaussian peak centered at the B_s mass;
- Cabibbo-suppressed $B_s \to D^-_s(2\pi)K^+$ decays are modeled with a wide peak centered just below the main signal mass and the shape is taken from Monte Carlo simulation;
- a background from Λ_b decays mis-reconstructed as B_s and described by a template Monte Carlo is present, mainly for K^*K modes;
- a background from mis-reconstructed B^0_d is also used under the signal peak;
- finally, a *combinatorial* background is found, due to the association of a real D meson with random tracks. This is the dominant background source and is modeled as the sum of decaying exponential and flat linear components.

4.1.2 Partially reconstructed hadronic decays

The "shoulder" at lower mass values is mainly produced by decays of the kind $\bar{B}^0_s \to D^{*+}_s(D^+_s\gamma)\pi^-$ or $\bar{B}^0_s \to D^+_s\rho^-(\pi^-\pi^0)$, where the neutral particle in the final state has not been detected. These modes have yields comparable with that of the fully reconstructed sample, although with a worse $\mathcal{S}/\mathcal{B}$, as may be seen

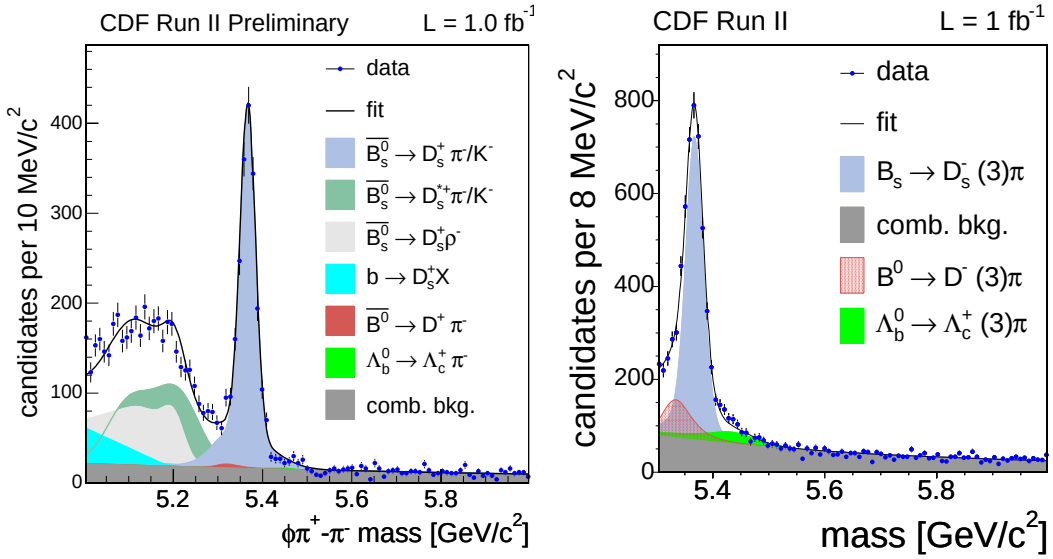

Figure 4.1: Left: Invariant mass distribution of $B_s^0 \to D_s^-[\phi\pi]\pi^-$ decays, including also the $\bar{B}_s^0 \to D_s^{*+}[D_s^+\gamma]\pi^-$ and $\bar{B}_s^0 \to D_s^+\rho^-[\pi^-\pi^0]$ contributions. Right: Invariant mass distribution of $B_s^0 \to D_s^-\pi^-$ decays, with the exclusion of the $D_s^- \to \phi\pi^-$ mode.

from Fig.4.1. Furthermore, since only one soft particle goes undetected in these decays, their M($KK\pi\pi$) invariant mass is very close to the fully reconstructed peak and the overall proper time uncertainty is of similar order as the latter. This feature provides them with an effective statistical power comparable with that of events in the main peak. Based on this consideration CDF decided to include these contributions into the B_s mixing analysis.

$B_s \to D_s(\phi\pi)\pi$ decay candidates selected with the hadronic TTT and having an invariant mass $M(KK\pi\pi) \geq 5.0\ GeV/c^2$ are accepted. From a detailed Monte Carlo simulation using the most up-to-date `EvtGen` [34] decay table, the fraction of the reconstructed sample not accounted for by the $D_s^{*+} \to D_s^+\gamma$ or the $\rho^- \to \pi^-\pi^0$ decay is at the 10% level and is predominantly composed of D^{**} and D semileptonic decays. A final fit based on the maximum likelihood method is performed in the above mass range, that returns the various yields and the B_s^0 mass and lifetime at the same time. The relative contributions are fixed to the fractions estimated on Monte Carlo, given the known Branching Ratios ($\mathcal{B}_i^{PDG}$) and the absolute trigger and selection efficiency for a generic signal B. Each fraction f_i is given by the expression:

$$f_i = \frac{\varepsilon_i \cdot \mathcal{B}_i^{PDG}}{(\sum_j \varepsilon_j \cdot \mathcal{B}_j^{PDG}) + \frac{N_{others}}{N_{generated}}} \tag{4.1}$$

The values for the branching fractions of the B_s are not well known and not

listed in the PDG [21]. We assume $SU(3)$ symmetry-induced invariance for the $\mathcal{B}_i$ under the $s \leftrightarrow d$ exchange of the spectator quark and use the better known values of the B_d in the corresponding modes.
As done for the fully reconstructed peak, the `ROOTSNNS` Neural Network is also used to identify the partially reconstructed hadronic decays.

4.1.3 Semileptonic decays

Aside of the dedicated semileptonic trigger (sec.3.3.5), applied to a fraction ($0.355\ fb^{-1}$) of its accumulated statistics, CDF identifies B_s semileptonic decays using the TTT over all the $1\ fb^{-1}$ sample. A 60% overlap is found among the two triggers. Use of the hadronic trigger is, in fact, profitable also for semileptonic modes, as it allows for identification of lower momentum leptons. To identify the lepton tracks a particle identification likelihood is used that incorporates information from several detector components; in particular CDF uses the muon spectrometer and the electromagnetic camorimeter informations. Since part of the initial momentum is carried away by the undetected neutrino, no B_s mass can be reconstructed. Nevertheless, the combined use of the D_s mass and the ℓD_s mass ($m(\ell D_s)$) proves to be effective in rejecting most of the background present. The identified sources of background are:

- sequential decays of the kind $B_s^0 \to D_s^{(*)} D^{(*)} X$, with one of the two charm mesons decaying semileptonically;
- the association of a real D_s meson and a fake lepton attached to the B_s production vertex;
- a real lepton from $c\bar{c}$ prompt events.

As one can see from the $m(\ell D_s)$ distribution in Fig.4.2, use of this variable allows to identify and model the false lepton and the physics (e.g. sequential decay) component significantly. For this, a cut $2.0 \leq m(\ell D_s) \leq 5.5\ GeV/c^2$ is applied. The following set of selections are also common for all modes in order to suppress the remaining contributions:

- $-0.01 \leq L_{xy}(B \leftarrow D) \cdot \frac{m(D_s)}{p_t(D_s)} \leq 0.10\ cm$;
- $\sigma_{ct^*}(B) \leq 0.04\ cm$;
- $ct^* \geq 0.01\ cm$ ($ct^* \geq 0.005\ cm$ for the semileptonic trigger).

Here ct^* is the *measured* proper decay length, not accounting for the missing neutrino momentum.
Other specific selections are applied to the different modes accounting for kinematics, quality of the B_s and the D_s vertex reconstruction and the displacement

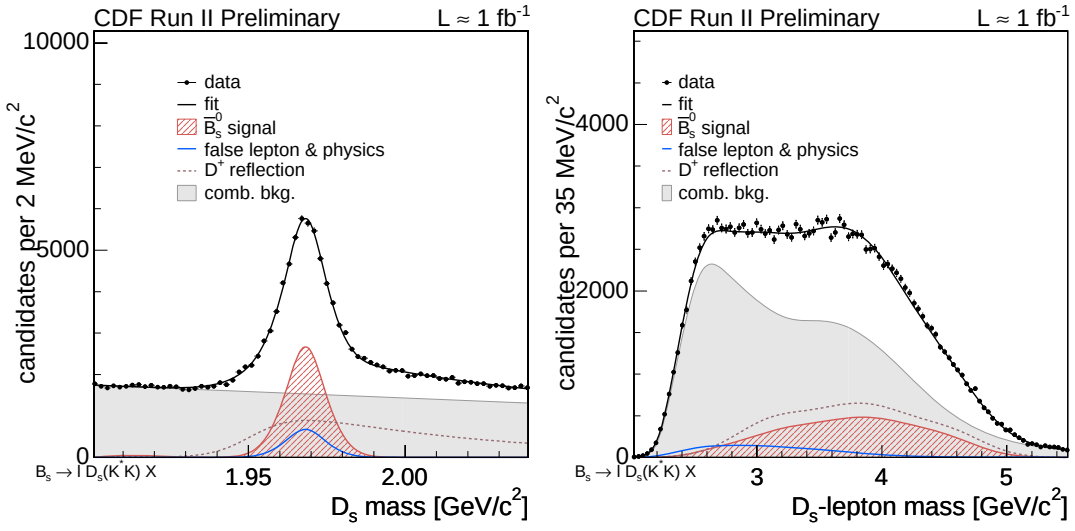

Figure 4.2: $\ell + D_s$ invariant mass distribution, with the different components overlaid. Also the D_s invariant mass distribution is shown (smaller panel).

of the D meson from the Primary Vertex. Also angular correlations between the D_s candidate momentum and the direction of the two decay products of the vector mesons (ϕ or K^*) are considered.

We imply Particle Ientification to select Kaons in the final states. By doing so, one can reduce the combinatorial background, dominated by pions. Moreover, in the case of the $D_s^0 \rightarrow K^{*0}[K^+\pi^-]K^-$ a substantial contribution from the $D^- \rightarrow K^{*0}\pi^-$ and the $D^- \rightarrow K^+\pi^-\pi^-$ non-resonant decays is found.
By combining a particle's Time-Of-Flight and its specific ionization in the COT via a Likelihood Ratio (in a way that is described in more detail in Sec.5.4) we get a $\approx 100\%$ increase in statistics for the $D_s^0 \rightarrow K^{*0}[K^+\pi^-]K^-$ channel, with respect to vetoing explicitly the D_s candidates consistent with a D^- reflection.

4.2 B_s candidate proper decay time measurement and resolution

4.2.1 Proper decay time

We want to measure the proper decay time of the B_s mesons, that is the decay time in the meson rest frame. This is done by measuring the distance between the B_s production vertex and its decay vertex in the xy detector plane, L_{xy}; and its transverse momentum p_t. The proper decay time is related to these quantities

by the formula:

$$t_{decay} = L_{xy} \cdot \left(\frac{m(B_s)}{p_t(B_s)} \right) \tag{4.2}$$

where $m(B_s)$ is the reconstructed B_s mass and the $\frac{m(B_s)}{p_t(B_s)} = (1/\beta\gamma)$ corrects for the B_s Lorentz boost in the laboratory frame.

An important feature of the proper time distribution in the CDF B_s samples is that they are sculpted by the hadronic trigger selections. We recall that the SVT trigger requires both a minimum (120 μm) and a maximum (1000 μm) $|d_0|$. This introduces a bias in the proper decay time, affecting mostly the short lived component of the distribution. This is further enhanced by the $L_{xy}/\sigma(L_{xy})$ requirement at the event reconstruction stage. In order to correct for this effect, we need to evaluate the trigger efficiency as a function of the proper decay length. This is done using a detailed Monte Carlo simulation of both the detector and the SVT trigger. After applying the selection cuts to the simulated sample,

$$\varepsilon(ct) = \frac{N_{SVT,cuts}(ct)}{\sum_i e^{-t/\tau} \otimes Gauss(0, \sigma^i_{ct})} \tag{4.3}$$

where $N_{SVT,cuts}(ct)$ is the number of events that passed the SVT and analysis cuts as a function of the proper decay length *ct in the laboratory system*; the denominator is the sum over all the generated events of their decay time distributions with a mean τ, multiplied by the event-by-event detector resolution. The efficiency curve for B^0_d mesons is displayed in Fig.4.3. The trigger efficiency function $\varepsilon(ct)$

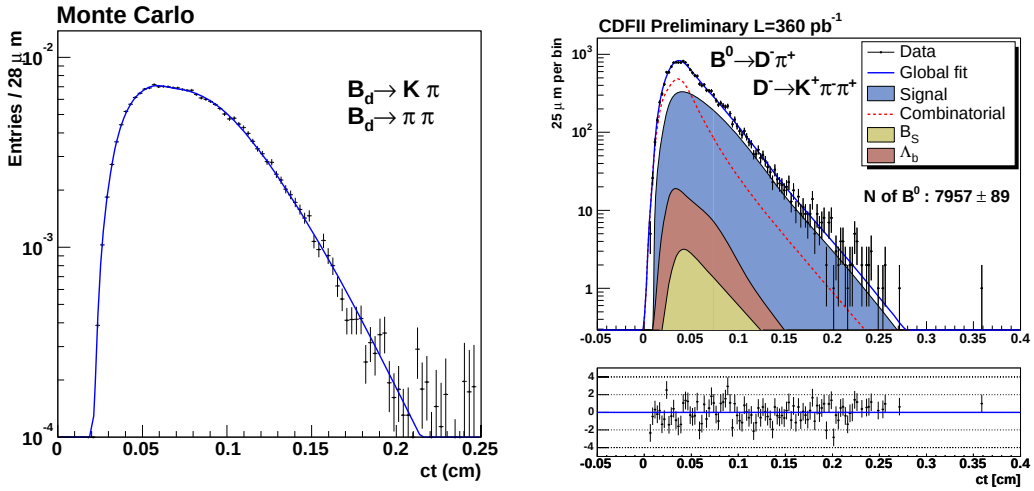

Figure 4.3: Left: An example of the dependence of trigger and selection efficiency on proper decay time, as obtained on the $B^0_d \to K\pi$ and $B^0_d \to \pi\pi$ decays. Right: $B^0_d \to D^-\pi^+$ lifetime fit and corresponding pulls.

is convoluted with the B_s lifetime and the detector resolution parameterization in the proper decay length part of the signal likelihood:

$$\mathcal{L}^{decay\ length} = \left(e^{-ct/c\tau} \otimes Gaus(ct, \sigma_{ct})\right) \cdot \varepsilon(ct) \tag{4.4}$$

An efficiency curve is derived for each B_s decay mode, given their kinematic features and different selection requirements. These also include the recently added partially reconstructed decays in the $B_s \to D_s(\phi\pi)\pi$ sample. No relevant difference is observed among the various modes. The trigger effect model is cross-checked within a lifetime measurement framework for high-statistics B^+ and B^0_d decays before being used in the measurement of B_s decay length for mixing. Infact, these modes provide a testing ground kinematically similar to the B_s modes. Results for the lifetime measurements are showed in Tab.4.2 and are found to be in good agreement with the world average values. The corresponding $c\tau$ for for the $B^0_d \to D^-\pi^+$ is also shown in Fig.4.3.
This indicates that our understanding of the trigger and selection effects is correct. The HFAG values are given in ref. [37].

Decay mode	CDF $c\tau, \mu m$	HFAG $c\tau, \mu m$
$B^0_d \to D^-\pi^+$	452.1 ± 5.1 (stat.)	458.7 ± 2.7 (stat.)
$B^- \to D^0\pi^-$	491.1 ± 5.1 (stat.)	491.1 ± 3.3 (stat.)
$B^0_s \to D^-_s(3)\pi^+$	461 ± 12 (stat.)	432 ± 20 (stat.)

Table 4.2: Lifetimes of the B mesons and comparison with the HFAG average values. The HFAG value for the B_s is measured from $B_s \to$ flavour specific modes only.

4.2.2 Proper decay time uncertainty

The uncertainty on the measurement of the B_s proper decay time, σ_{ct}, is made up of 3 terms:

$$\sigma_{ct} = \left(\frac{L_{xy}}{p_t}\right)\sigma_{m_B} \oplus \left(\frac{m_B}{p_t}\right)\sigma_{L_{xy}} \oplus \left(\frac{\sigma_{p_t}}{p_t}\right)ct \tag{4.5}$$

The uncertainty on the B hadron mass is very small compared to the others and the corresponding term can be neglected. The second term expresses how well we can determine the B_s decay length and receives contributions essentially from the detector vertex resolution and from the decay topology. We improve our determination of the decay length using CDF precise silicon tracking, including Layer 00. Furthermore, the Primary Vertex is calculated *for each event*, using tracks that can be associated to a single vertex along the beam. The improvement

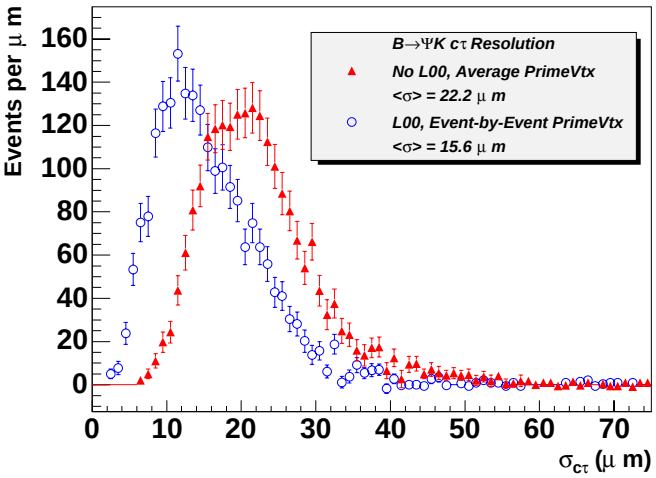

Figure 4.4: Comparison of the σ_{ct} distribution on the $B^+ \rightarrow J/\psi K^+$ sample, using the average beamline and no L00 (red), or with L00 and Event-by-Event Primary Vertex determination.

introduced by this algorithm, with respect to using the average beam position, is described in Fig.4.4. The third term is proportional to the uncertainty on the determination of the B_s transverse momentum; it has to be noted that this term increases linearly with the B_s proper time. This results in a degraded time resolution at longer times for classes of events with a poorly measured kinematics (*i.e. partially reconstructed events*).

The gaussian detector resolution estimated by CDF tracking is then calibrated in order to reproduce the behaviour of real data. A *scale factor* S_{ct} is applied to each event as a function of kinematical and topological quantities. The calibration is performed on a large data sample of prompt D^+ combined with one or three prompt tracks to mimic the B^0-like decay topologies. The prompt charm decays are selected with the TTT, where a charm meson plus a displaced track are identified. Their ct distribution is compared to the Primary Vertex position determined from all the other eligible tracks in the event. S_{ct} is obtained from the fit to the distribution obtained as such. Typical contributions to the sample from secondary D mesons are of the order of 7-15 % depending on selections. Also tracks not originating from the PV and entering the PV fit induce a bias in the evaluation of S_{ct}. While the prompt part is fitted with a single Gaussian function, the longer lived parts are modeled with two symmetric exponential tails. The resulting template is shown in Fig.4.5. Typical values for S_{ct} range from 1.1 to 1.5.

The decay time resolution in the case of the semileptonic decays is expected to be worse on average due to the σ_{p_t}/p_t term that dominates at higher decay

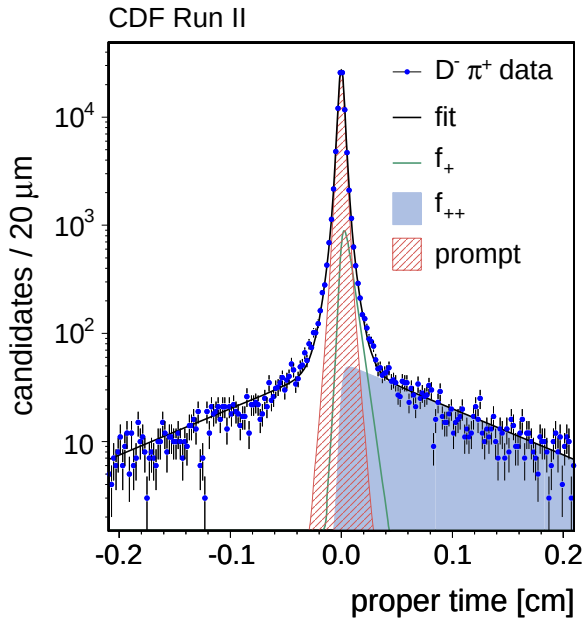

Figure 4.5: The proper time resolution function, as measured on the calibration sample. The Gaussian for prompt is in pink, green and full blue distributions model the short- (f_+) and long-lived (f_{++}) non-prompt contributions.

times. This reduces the statistical power of this large sample significantly.
One can express the rate of decay time lost due to unmeasured momentum as a κ factor:

$$ct(B_s^0) \equiv \kappa \cdot M_{B_s}^{PDG} \frac{L_{xy}(\ell D_s)}{p_t(\ell D_s)} \tag{4.6}$$

and derive this quantity's distribution from Monte Carlo simulation:

$$\kappa \equiv \left\langle \frac{p_t(\ell D_s)}{p_t(B_s)} \frac{L_{xy}(B_s)}{L_{xy}(\ell D_s)} \right\rangle_{MC} \tag{4.7}$$

The distribution of the κ factor is shown in Fig.4.6 for different ranges of $m(\ell D_s)$. By definition, the average gets closer to unity as the combined mass $m(\ell D_s)$ tends to the B_s mass. For such events, where little momentum is lost, the kinematic term in σ_{ct} is small: we weight these events more (and de-weight decays with smaller κ) by convoluting the signal likelihood for each partially reconstructed decay with the κ template corresponding to the value of $m(\ell D_s)$ for each event. From Fig.4.6 one can see that events with a better determined kinematics have a flatter time dependent decay time uncertainty (hadronic-like).

To the class of events with $\kappa \approx 1$, also belong the incompletely reconstructed hadronic modes. Their κ-factor distributions are depicted in Fig.4.7. The mean is $< \kappa > \approx 0.96$ for all modes, with an R.M.S. ≈ 0.02. Thus, since the σ_{ct} for

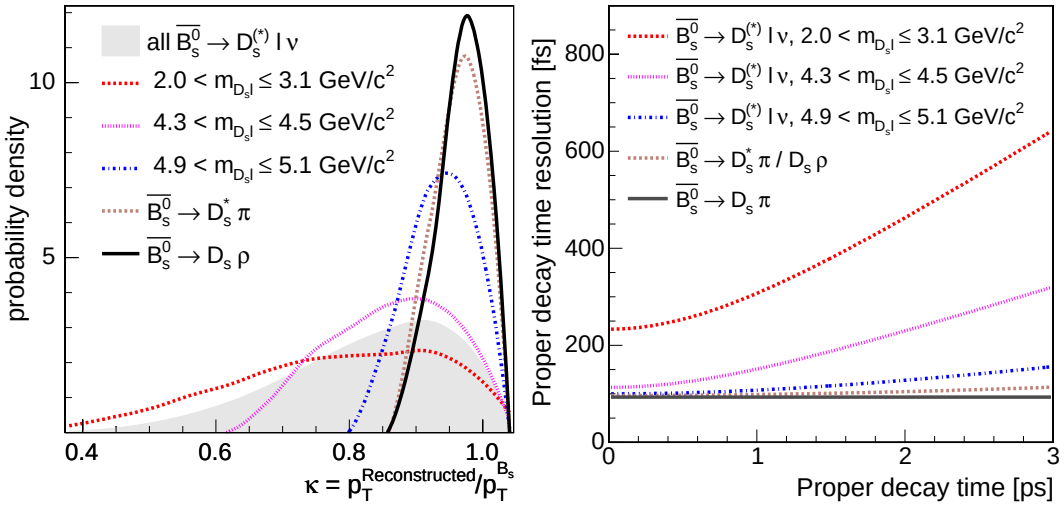

Figure 4.6: Left: Monte Carlo distribution of the κ-factor in semileptonic $B_s \to \ell D_s X$ decays, for three different ranges of the combined mass $m(\ell D_s)$. Also the partially reconstructed decays are displayed. Right: σ_{ct} as a function of ct for the same $m(\ell D_s)$ ranges. The solid black line represents hadronic decays.

these modes is comparable to the one of the fully reconstructed hadronic modes (see also Fig.4.6), the effective statistical power for mixing of the two kinds of samples is almost even.

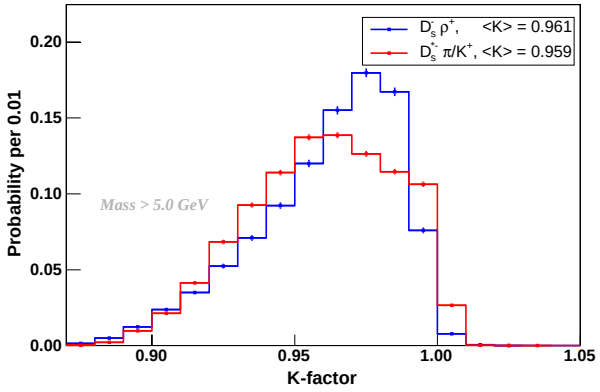

Figure 4.7: κ factor distribution for the dominant partially reconstructed hadronic modes: $< \kappa(B_s \to D_s \rho) >= 0.961$ (blue); $< \kappa(B_s \to D_s^* \pi / K) >= 0.959$ (red).

4.3 Tagging of the B_s candidate's flavour at production

The last part of the analysis is about determining whether at the meson was produced from a b or a $\bar{b}$ quark, in order to perform a time dependent asymmetry measurement

$$\mathcal{A} = \frac{N_{RS}(t) - N_{WS}(t)}{N_{RS}(t) + N_{WS}(t)} \tag{4.8}$$

where $N_{RS}(t)(N_{WS}(t))$ is the number of events decayed at time t with the same (opposite) flavour with respect to that at production. Two general strategies are available to tag the b-flavour:

- *Same Side Tags* (SST): these use the electric charge of associated particles produced in the fragmentation of the b quark that produces the reconstructed B_s;
- *Opposite Side Tags* (OST): they infer the initial flavour of the candidate B_s from the decay products of the hadron originated from the accompanying b quark in the event.

At the Tevatron hadron collider $b - \bar{b}$ pairs are created incoherently. While a B_s is found on the trigger side, any $b-$hadron can be present on the away side as a result of the hadronization of the other b. As already mentioned in Ch.2, the statistical significance of each point in the amplitude scan is related to the flavour tagging effectiveness εD^2 by the expression:

$$\frac{1}{\sigma(A)} \propto \sqrt{\frac{\mathcal{S}\varepsilon D^2}{2}} \tag{4.9}$$

where $\mathcal{S}$ is the number of B_s signal events. In this expression

- ε is the tagger **efficiency**:

$$\varepsilon = \frac{N_{\text{B tagged}}}{N_{\text{signal B}}} \tag{4.10}$$

- D is the tagger **dilution**:

$$D = \frac{N_{RIGHT\ TAG} - N_{WRONG\ TAG}}{N_{RIGHT\ TAG} + N_{WRONG\ TAG}} \tag{4.11}$$

where $N_{RIGHT\ TAG} + N_{WRONG\ TAG} = N_{\text{B tagged}}$

This last quantity is related by the following formula to the mistag rate $W \equiv N_{WRONG\ TAG}/(N_{RIGHT\ TAG} + N_{WRONG\ TAG})$, that is to the *purity* of the tagged sample:

$$D = 1 - 2 \cdot W \tag{4.12}$$

From Eq.4.9 one gets that the *effective* statistics accumulated for a mixing measurement is not given by the B_s sample available, rather by the fraction of it that has a flavour tag associated. The enhancement of tagging capabilities is thus an important task for an experiment's sensitivity to oscillations. This is particularly challenging at a hadron collider, given the event topology and the high rate of particles produced in a single collision. The OST suffer especially from the limited acceptance to the away b, as a large fraction of them are far from the triggered b. The SST 's challenge is to identify associated particles within a large ensemble of random tracks.

To profit the most of a tagger features, the dilution is usually evaluated in bins of one or more variables sensitive to signal. Infact, from the definition of *figure-of-merit* we have that binning the dilution increases the final εD^2 with respect to the unbinned case. For example, in the case of two bins we have:

$$\varepsilon_1 D_1^2 + \varepsilon_2 D_2^2 = \varepsilon D^2_{\text{unbinned}} + \frac{\varepsilon_1 \varepsilon_2 (D_1 - D_2)^2}{\varepsilon} \geq \varepsilon D^2_{\text{unbinned}} \tag{4.13}$$

As seen, a large gain can be achieved in cases when the dilution is strongly dependent on the chosen variable. An example of this is described in [55], where up to 25 bins are used in some cases, with a relative improvement of $\approx 10\%$ in their tagging effectiveness.

In the next sections the CDF tagging algorithms are described in more detail, starting from Same Side Kaon Tagging (SSKT), which has the highest tagging power, and then going through the existing Opposite Side Taggers.
The Opposite Side Kaon Tagger is given a separate chapter as this represents the main part of this thesis work and the latest novelty within the CDF tagging system.

4.3.1 The Same Side Kaon Tagger

Working principle

At a first order, we can establish a direct correlation between the flavour of a B meson and the electric charge of a particle produced in the same fragmentation process.
Infact, a B_s^0 is formed from a $\bar{b}$ and an s quark from an $s\bar{s}$ pair created in vacuum:

thus, the remaining $\bar{s}$ quark will likely originate a K^+ meson (Fig.4.8). In the same way, a $\bar{B}_s^0$ is accompanied by a K^-.

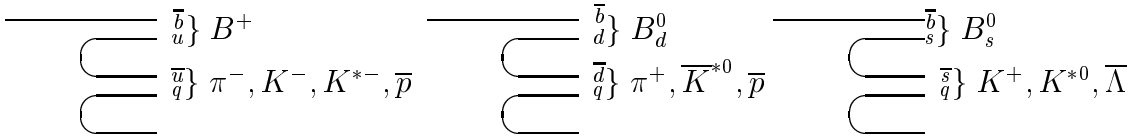

Figure 4.8: Fragmentation diagrams where a B meson is originated together with a charged particle, used for flavour tagging.

Thus, one needs to:

- select the leading fragmentation track;
- identify it as a Kaon.

For the first request, one has to look to tracks close to the B flight direction. This is implemented by applying a requirement on the quantity

$$\Delta R = \sqrt{|\eta(B) - \eta(K)|^2 + |\phi(B) - \phi(K)|^2} \tag{4.14}$$

where ϕ is the azimuthal angle and η is the pseudorapidity. Nevertheless, to distinguish Kaons from other particle types (mainly pions) also particle identification information is needed. We use the Time-Of-Flight and the measurement of the dE/dx and combine them in a single Kaon probability

$$LR(K) = \frac{P_{TOF}(K) \cdot P_{dE/dx}(K)}{0.9 \cdot P_{TOF}(\pi) \cdot P_{dE/dx}(\pi) + 0.1 \cdot P_{TOF}(p) \cdot P_{dE/dx}(p)} \tag{4.15}$$

The likelihood ratio technique is described in more detail later (Ch.5, Sec.5.4). The weights for the single particle probabilities are *a priori* fractions for the background composition. In the case of the SSKT, if either TOF or dE/dx is not available, the corresponding probability is set to 1 and the *a priori* probability is used.

Since the dilution of such a tagger depends strongly on the kind of B meson considered, its evaluation directly on B_s data could only be obtained by extracting the dilution as a parameter of the B_s oscillation fit. Thus, one has to rely on Monte Carlo simulation and measure dilution on it. Since this quantity is likely to depend on the decay kinematics as a result of the pecualiar fragmentation process, a detailed data-MC comparison of the main kinematic quantities and track multiplicity has been performed, in order to validate MC predictions.

Data and MC samples

A wide range of decays of B^+, B_d^0 and B_s^0 mesons has been used for the development of the SSKT. These include both J/ψ K and $D\pi$ channels ($B^+ \to J/\psi[\mu^+\mu^-]K^+$, $B^+ \to \bar{D}^0[K^+\pi^-]\pi^+\pi^+$ and corresponding flavour modes). The selections for these modes are listed in [43]. Data samples have been selected using both the dedicated J/ψ di-muon trigger and the SVT-based Two-Track-Trigger, with an integrated luminosity $\int \mathcal{L} = 355\ pb^{-1}$. MC samples were generated using the `PYTHIA` ([33]) package, with the activation of all the $b\bar{b}$ production mechanisms: flavour creation, flavour excitation and gluon splitting. The non-perturbative hadronization part is described by the default `PYTHIA` Lund function. The fraction of energy

$$z = \frac{(E + p_L)_B}{(E + p_L)_b} \tag{4.16}$$

of the initial b carried by the B meson is distributed using the corresponding fragmentation function

$$f(z) \propto \frac{1}{z}(1 - z)^\alpha \exp -\frac{\beta m_\perp^2}{z} \tag{4.17}$$

where $m_\perp$ is the tranverse mass of the resulting hadron; the values of the parameters α, β have been taken from fits to LEP data [38]. The contribution to the final B population given by the decay of the excited states B^{**} is fixed to 20%. The signal B is then forced to decay in one of the above specific modes, while the other one is decayed following the standard decay table. A full CDF detector simulation is also performed; MC sample was generated as to reproduce the features of the various data-taking periods in order to avoid MC-data discrepancies due to detector functional effects. The PID response is simulated as described in Secs.3.2.2 and 3.2.3.

Tagging algorithm

As anticipated, the development of the tagging algorithm, the optimization of the selections and the evaluation of the performances are done on the MC samples for the various B mesons.
Tracks are selected following the requests in Tab.4.3, looking at quality, proximity to the primary vertex that originated the signal B and vicinity in cone to the B momentum as reconstructed from its decay daughters. As one can see in Fig.4.9, in several events the number of tracks passing the above cuts is > 1. Different possible criteria were studied and their performances compared, in order to select the final tagging track. The chosen algorithm chooses the track most likely to be a Kaon (that is *with the highest PID value*). Its performances are measured on the B_s MC samples and the corresponding prediction is checked on the B^+ and B_d^0 data samples. The resulting data-MC disagreement is quoted as a systematics

quantity	cut
Silicon Hits r-ϕ	≥ 3
COT Hits	≥ 3
p_t	$\geq 450\ MeV/c$
$\|\eta\|$	≤ 1
$\Delta R(\ell + SVT, track)$	≤ 0.7
$\|d_0/\sigma(d_0)\|$	≤ 4
$\|z_0(trk) - z_0(B)\|$	$\leq 1.2\ cm$
lepton (e or μ) removal	Yes

Table 4.3: Selections applied to candidate tagging tracks.

on the value of εD^2 obtained when applying the algorithms to the B_s system.

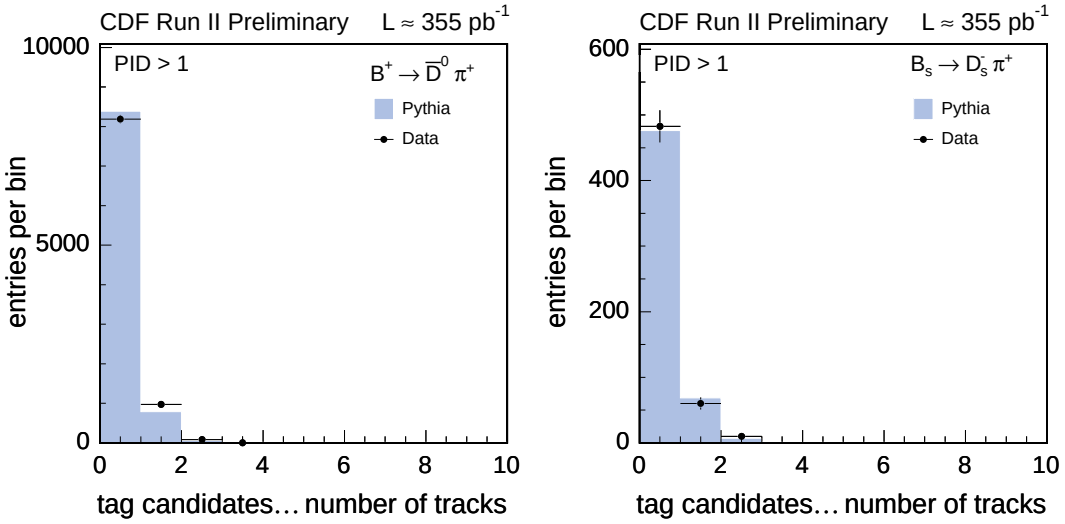

Figure 4.9: Data and MC comparisons of tagging candidate multiplicity; left plot is for the $B^+ \to \bar{D}^0\pi^+$ mode, right plot for $B_s^0 \to \bar{D}_s^0\pi^+$ mode. A $LR > 1$ is applied to all candidates.

Performances

Cases where > 1 candidate is present, and their responses disagree, are treated differently than the agreeing cases. Then, the dilution has been parameterized as a function of two relevant quantities, the LR(K) and the $p_T(K)$, separately for the agreeing and disagreeing cases. Thus, it is possible to assign a predicted

dilution to each event. The dependence of the dilution on the LR(K) is shown in Fig.4.10: as expected, the dilution increases with the Kaon probability. The trend is inverted at about $LR(K) \geq 2$, where the π/K separation is very good. This separation power is, in fact, only obtained at low momenta, where the fraction of Kaons from the underlying event is dominant with respect to the Kaons from the same $b-$fragmentation process, harder on average. At this point a maximum

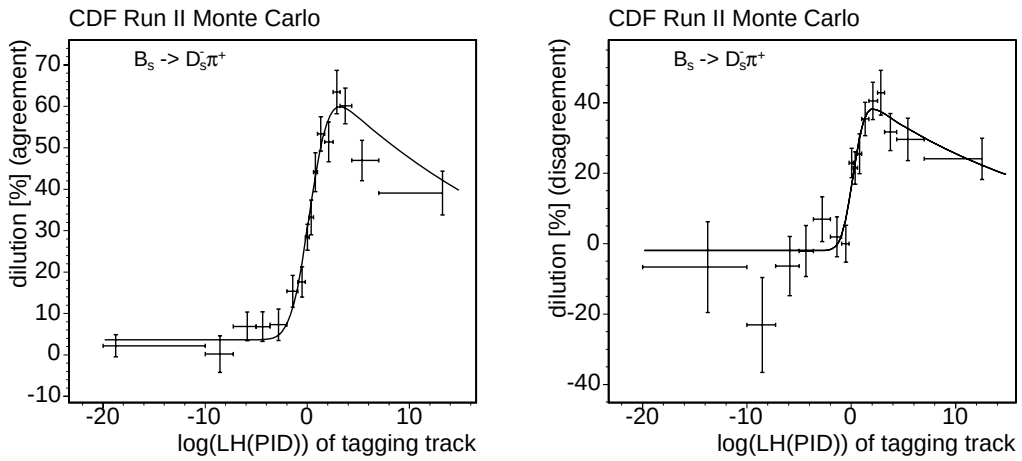

Figure 4.10: Dilution as a function of the log(LR(K)) for the $B_s \to \bar{D}_s^0 \pi^+$ channel. On the left, the case where candidates agree, on the right for the disagreeing case.

likelihood fit has been performed on the various modes to retrieve the tagger efficiency and dilution, where the dilution D appears in the likelihood for the signal:

$$\mathcal{L}_{signal,tagging} = \varepsilon_{signal} \frac{1 \pm D_{pred}\ cos(\Delta m_q t)}{2} + (1 - \varepsilon_{signal}) \tag{4.18}$$

The results on the different modes can be found in [43]. We report here the performances of the established algorithm on the $B_s \to \bar{D}_s^0 \pi^+$ channel.

Systematics

Since the evaluation of the tagger performances is essentially relying on MC, the range of validity of all the assumptions entering the simulation is verified and a subsequent systematic uncertainty is assigned. An extensive investigation of several sources has been undergone, regarding essentially the following aspects: the PID response parameterization; the relative fractions of the various production mechanisms as extrapolated from the most distinguishing variable ($\Delta\phi$ between

	$\varepsilon(\%)$	D(%)
Max average log(LR(K))	49.3 ± 2.3	22.8 ± 0.8
Max parameterized log(LR(K))	49.3 ± 2.3	$28.3^{+3.2}_{-4.2}$

Table 4.4: Performance of the max log(LR) on the $B_s \to \bar{D}^0_s \pi^+$ sample, with and without the Dilution parameterization with the LR(K), are shown. Efficiency is evaluated directly on data, dilution on MC.

the two b's); the various fragmentation-related features, like the behaviour of the function $f(z)$ in 4.17 and the resulting amount of the various particle species around the B meson; the effect of the pile-up of multiple interactions occurring in a high-luminosity regime, as is the case for present Tevatron running ($\mathcal{L} \sim 10^{32} cm^{-2} s^{-1}$). All of those have been assessed by reweighting or filtering out events from the original MC configuration and are described in [43].

Finally, the sum of all these effects is comparable with the systematic uncertainty accounting for the agreement between data and MC in the other B samples, B^+ and B^0_d, for which it is possible to measure the dilution directly on data. The overall systematics on the dilution D for the mixing golden mode $B_s \to \bar{D}^0_s \pi^+$ is 14%.

The final tagging power of the SSKT on the golden channel is thus:

$$\varepsilon D^2 = 4.0^{+0.9}_{-1.2} \ \% \tag{4.19}$$

4.3.2 The Opposite Side Taggers

Opposite Side Taggers complement the performances of the SSKT by tagging the away b flavour. This is done through the electric charge of its decay products:

- leptons from the $b \to \ell X$ inclusive decays: both a muon tag and an electron tag are implemented;
- jets originating from the b hadronic decays

Calibration sample

CDF Opposite-side taggers have been studied using an high statistics sample of *inclusive* semileptonic b decays collected with the dedicated $\ell + SVT$ trigger ($\mathcal{O}$(4 million) events compared to exclusive hadronic decays, $\mathcal{O}$(5 thousand)). The selections for the two separate configurations of this trigger, $e+SVT$ and $\mu+SVT$, have been described in sec.3.3.5. Further angular requirements are applied to the electron in conjunction with other tracks against photon conversions ($\gamma + N \to e^+e^- + N'$).

However, the $\ell + SVT$ data are not a pure sample of b decays. In addition to signal events, infact, they also include semileptonic charm decays, hadrons that fake the trigger lepton and other backgrounds. A background subtraction procedure has been established at CDF to enhance the B purity of the sample [32]. We distinguish among the components in the sample using two variables:

- the invariant mass $M(\ell, SVT)$;
- the signed impact parameter of the lepton or SVT track with respect to the $\ell + SVT$ momentum $\vec{P}$:

$$\delta = |d_0| \cdot sign(\vec{d_0} \cdot \vec{P}) \tag{4.20}$$

For the purposes of flavour tagging studies, only the range of masses $2 < M(\ell, SVT) < 4\ GeV/c^2$ is considered; this removes events in which both the lepton and SVT track are the decay products of a charm hadron.

The shape of the signed impact parameter distribution for the case where both the lepton and SVT track are decay products of a b-hadron can be calculated using a Monte Carlo model and assuming a particular p_t spectrum of b-hadrons and a model for b-hadron decay. The `Bgenerator` [35] was used here to describe the shape of the p_t spectrum and `EvtGen` [34] to model the b-hadron decays. The background subtraction procedure is based on the signed impact parameter of the SVT track. This subtraction procedure removes some of the signal, but is expected to eliminate most of the background. Infact, the signed I.P. will be symmetric for prompt tracks that produce fake lepton triggers and for cases where the lepton and SVT track are not the decay products of the same parent particle (Fig.4.11). The background subtracted distributions are then representative of those that would be obtained from a sample of pure b-decays.

The statistics available with this inclusive sample allows to optimize OS b-flavour tags with this sample. The charge of the trigger lepton provides an estimate of the B flavour on the trigger side at decay. The tagger efficiency, the dilution and the relative errors are calculated starting form the number of right (RS) and wrong tags (WS) out of N_{events} semileptonic events, as described by the formulas:

$$\begin{aligned}
\varepsilon &= \frac{(N^+_{RS} + N^+_{WS}) - (N^-_{RS} + N^-_{WS})}{N^+_{events} - N^-_{events}} \\
D &= \frac{(N^+_{RS} - N^+_{WS}) - (N^-_{RS} + N^-_{WS})}{(N^+_{RS} + N^+_{WS}) - (N^-_{RS} - N^-_{WS})} \\
\sigma^2_\varepsilon &= \frac{(N^+_{RS} + N^-_{RS} + N^+_{WS} + N^-_{WS})(N^+ - N^-)^2 + (N^+_{RS} - N^-_{RS} + N^+_{WS} - N^-_{WS})^2(N^+ + N^-)}{(N^+ - N^-)^4} \\
\sigma^2_D &= \frac{4[(N^+_{RS} - N^-_{RS})^2(N^+_{WS} + N^-_{WS}) + (N^+_{WS} - N^-_{WS})^2(N^+_{WS} - N^-_{WS})]}{(N^+_{RS} - N^-_{RS} + N^+_{WS} - N^-_{WS})^4}
\end{aligned} \tag{4.21}$$

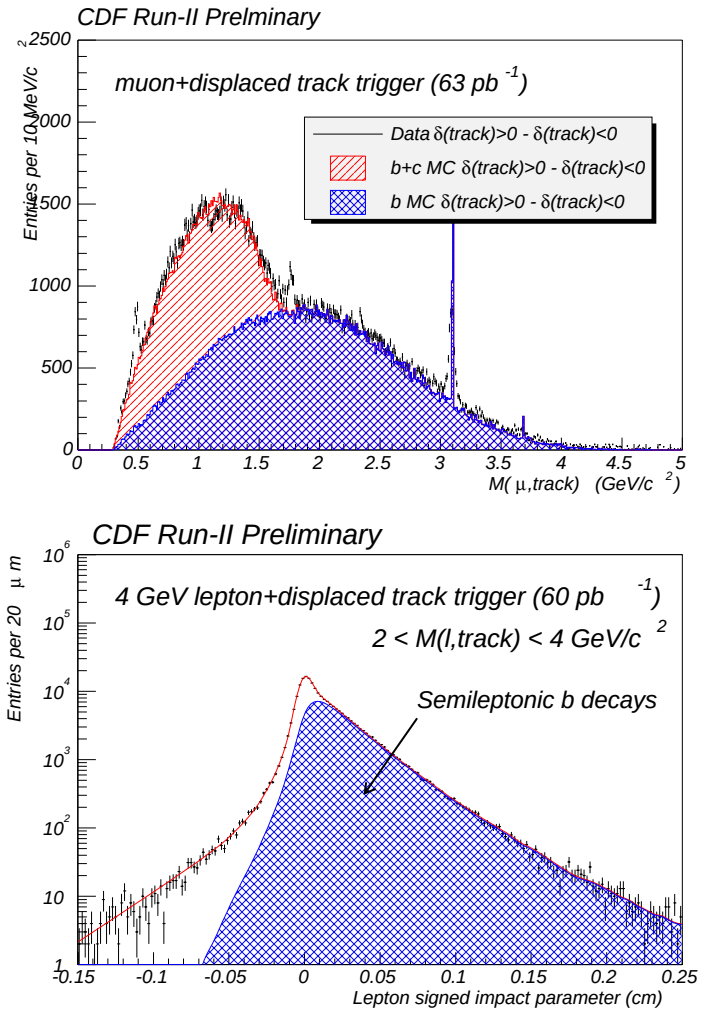

Figure 4.11: Upper plot: $M(\mu, SVT)$ distribution, after the signed I.P. background subtraction. In red the $b+c$ quark semileptonic contributions are shown; in blue, b decays are isolated. Several resonances can be noticed: at $\approx 0.5\ GeV/c^2$ the $K^0_s \to \pi^+\pi^-$ peak is present, where one of the pions fakes a μ by punching-through onto the muon chambers; the $J/\psi \to \mu\mu$ is visible where one of the muons triggered; finally, also a peak corresponding to the $D^0 \to K\pi$ decay is present, where one of the two tracks again faked a muon. Lower plot: lepton signed I.P. distribution for the $2 < M(\ell, SVT) < 4\ GeV/c^2$ range. The various fit contributions are displayed: real b decays are in blue.

Here the + (−) superscript indicates the positive (negative) signed impact parameter case. A flavour tag is counted as *Right Sign RS* if $Q_{tag} \cdot Q_\ell < 0$, with Q_{tag} being the tagger response ($Q_{tag} = +1$ (-1) for a $\bar{b}$ (b)). Ultimately, one can predict the dilutions that would result when the flavour tags are applied to a sample of B_s decays. The systematic uncertainty of this prediction is not expected to dominate a B_s mixing analysis.

Trigger-side Dilution

Cases when the B meson on the trigger side has mixed or has decayed via the chain $b \to c \to \ell$ may generate a wrong-sign lepton in the final state. On the other side, cases when a hadron from a real b decay fakes a lepton produce a random flavour-lepton charge correlation. All these effects reduce the *raw* dilution measured for an opposite side tag. The dilution on the trigger side is estimated using a detailed Monte Carlo simulation of the detector and the SVT trigger. The time-integrated mixing probability for B_d^0 is calculated using $\Delta m_d = 0.503 \pm 0.007$ (for B_s^0 the probability is assumed to be 0.5); the contribution of the neutral B mesons in the sample composition is then evaluated on Monte Carlo. Also the SVT trigger bias on the lower proper decay length of the b-hadrons is accounted for. The correction factor derived as such is:

$$D_{trigger} = 0.6412 \pm 0.0015(stat)^{+0.0141}_{-0.0079}(syst) \qquad \mu\text{+SVT} \tag{4.22}$$

$$D_{trigger} = 0.6412 \pm 0.0015(stat)^{+0.0215}_{-0.0367}(syst) \qquad \text{e} + \text{SVT} \tag{4.23}$$

The true dilution of opposite side flavour tags as assessed on a background subtracted semileptonic sample in the mass range $2 < M(\ell, SVT) < 4\ GeV/c^2$ is given by:

$$D_{true} = D_{raw}/D_{trigger} \tag{4.24}$$

where D_{raw} is the uncorrected dilution measured in the Opposite-side. The presence of hadrons that fake the trigger lepton is accounted for by adding a systematic uncertainty as derived from the fraction of false leptons. This is measured directly on data, relying on lepton identification tools and accounts for $\approx 70\%$ of the total systematic. The remaining contributions to the systematics are given by the Monte Carlo models of the meson fractions in the b sample, the SVT bias and from the knowledge of Δm_d as propagated into eq.4.23.

Soft Lepton Taggers

These taggers exploit the fact that, in semileptonic decays of the OS b-hadron, a correlation can be established between the b flavour and the charge of its lepton daughter: $b \longrightarrow \ell^-$ $(\ell = \mu, e)$, while $\bar{b} \longrightarrow \ell^+$.

An implementation of a muon and electron taggers was already in place during Run I [39]. For Run II a likelihood based approach is used that combines information from the muon detectors and both the electromagnetic and hadronic calorimeters in order to identify leptons from fakes without loss in efficiency [45],[46]. The distributions of the various identifying quantities for real leptons and for fakes are obtained by selecting pure samples of the various species. For the muon tagger, a pure sample of $J/\psi \rightarrow \mu\mu$ decays is considered, while fake muons are studied using pions from $K^0_s \rightarrow \pi^+\pi^-$, Kaons from $D^0 \rightarrow \pi^+K^-$ and protons from $\Lambda \rightarrow p\pi^-$. Electrons from photon conversions $\gamma \rightarrow e^+e^-$ are used to study the real electron features.

The resulting likelihood for signal and background can be written as

$$\mathcal{L}_{S(B)} = \prod_{i=quantities} \mathcal{L}_{i,S(B)} \tag{4.25}$$

where $L_{i,S(B)}$ is the PDF of the i-th identifying quantity for signal S (background B). The main background source for these taggers is given by sequential decays. A lepton originating from the latter has a wrong sign charge. To suppress this contribution and maximize the performances of the soft lepton taggers, the p_t of the lepton projected onto the b-hadron momentum, p_T^{rel}, is used. The dilution is binned in this variable, because fake leptons and lepton from sequential decays tend to have lower p_T^{rel}. This effect can be seen in Fig.4.12 for both lepton types.

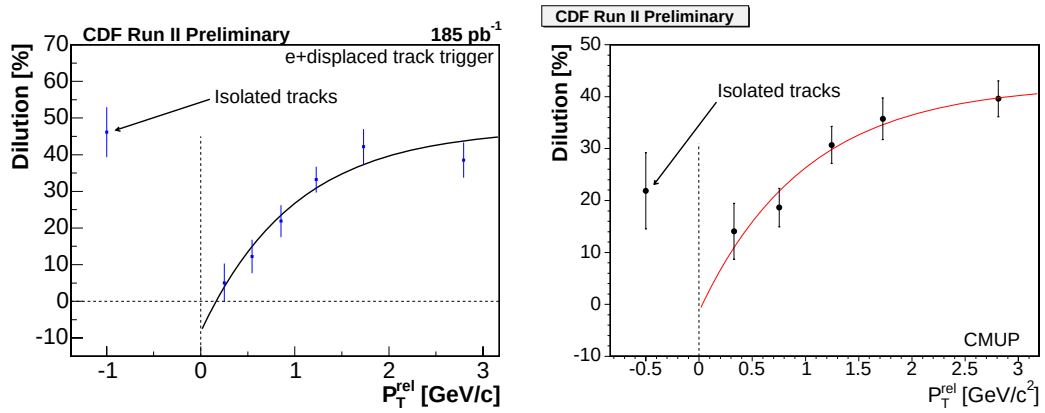

Figure 4.12: Dilution as a function of the lepton p_T^{rel} for both soft electron (left) and soft muon (right) taggers. $p_T^{rel} < 0$ indicates that the lepton is not within the b-jet.

A predicted dilution is then assigned to each event according to its likelihood and p_T^{rel}, by fitting for $D(p_T^{rel})$:

$$D(p_T^{rel}) = P_0 \cdot (1 - e^{-p_T^{rel}+P_1}) \tag{4.26}$$

and determining P_0, P_1 for different $\mathcal{L}$ bins.

From this we can evaluate the final performances of this class of taggers on the combined $\mu + SVT$ and $e + SVT$:

$$\varepsilon D^2 = 0.366 \pm 0.031^{+0.065}_{-0.056} \ \% \qquad \text{Soft Electron Tagger} \tag{4.27}$$

$$\varepsilon D^2 = 0.698 \pm 0.042^{+0.051}_{-0.027} \ \% \qquad \text{Soft Muon Tagger} \tag{4.28}$$

where the systematic uncertainty comes from the error on the mixing correction factor for the trigger side.

Jet Charge Tagger

The overall charge of the b-jet is also used to take a decision on the OS b flavour. To do that, we reconstruct jets in the Opposite-side using a track-based cone clustering algorithm. No calorimeter information is used. We start looking for tracks with $p_t \geq 1\ GeV/c$ as jet seeds and merge them into a new seed if their distance is $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 1.5$. Then, softer tracks within $\Delta R = 0.7$ are associated with the track clusters and the final jet momentum is calculated. The above selections are optimized on a `Pythia` [33] Monte Carlo simulated sample of $\ell + SVT$ events proved to reproduce the performances seen on data control samples. All $b\bar{b}$ production processes are generated. Since for a large number of events more than 1 jet is reconstructed, CDF uses a Neural Network to select the one most likely to originate from the away b decay [41]. In particular, the `NeuroBayes` package [61] is used. A two-stage NN first looks for tracks displaced from the b production vertex, more likely to come from a decay. A series of quantities are used for this purpose, among which: the track I.P. $|d_0|$, its p_t, its ΔR with respect to the B candidate and the rapidity with respect to the jet axis $y = \frac{1}{2}\ln\frac{E+p_L}{E-p_L}$ ($p_L = \frac{\vec{p}_{track}\cdot\vec{p}_{jet}}{|\vec{p}_{jet}|}$). A b-probability w_i is assigned to each track as the NN output. The second stage of the process uses these b-probabilities along with some kinematic variables and track multiplicity to the select the away b-jet. All tracks within the chosen jet are used to calculate the *jet charge*

$$Q_{jet} = \frac{\sum_i Q_i \ p_{t,i} \cdot (1 + w_i)}{\sum_i p_{t,i} \cdot (1 + w_i)} \tag{4.29}$$

starting from their electric charge Q_i, their transverse momentum and their b-probability w_i.

Three different types of jets are considered:

- Class 1: Jets with a secondary vertex as tagged by the `SecVtx` algorithm [54];
- Class 2: Jets not in Class 1 and with at least one track with probability greater w_i than 50%;

- Class 3: Jets not in Class 1 and with no tracks with probability greater than 50%

The dilution is calculated splitting the 3 tagging samples in bins of $|Q_{jet}| \cdot P_{nn}$, where P_{nn} is the jet probability. The behaviour of the dilution as a function of this variable for the 3 classes is showed in Fig.4.13.

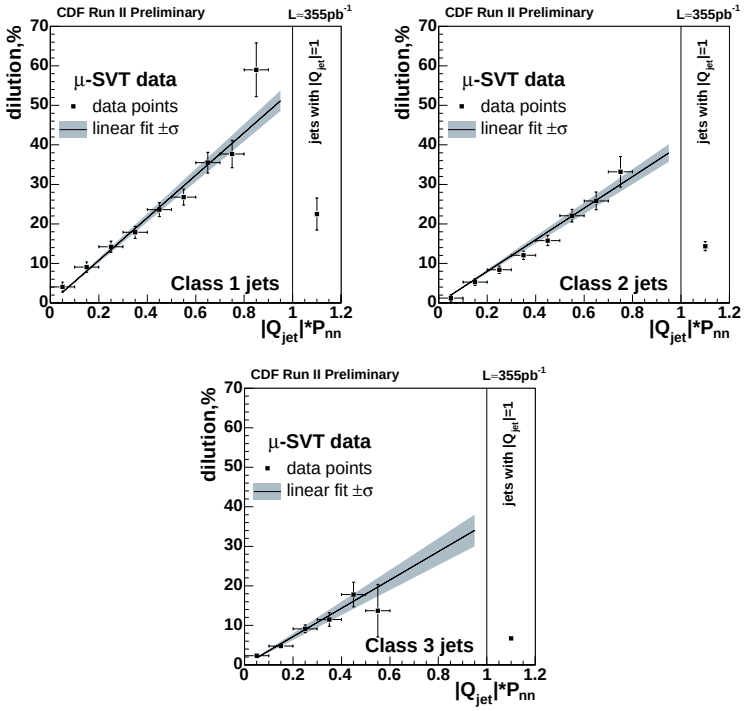

Figure 4.13: Dilution of the NN Jet Charge tagger as a function of the quantity $|Q_{jet}| \cdot P_{nn}$, for the 3 jet classes. It is evaluated on the $\mu + SVT$ sample.

The performances of the Jet Charge tagger are assessed on a $\ell + SVT$ sample of 355 pb^{-1} and can be found in Tab.4.5.

Jet type	ε (%)	$\sqrt{\varepsilon D^2/\varepsilon}$ (%)	εD^2 (%)
e+SVT data			
Class 1	10.56 ± 0.05	19.75 ± 0.43	0.412 ± 0.018
Class 2	28.20 ± 0.08	11.60 ± 0.33	0.380 ± 0.021
Class 3	56.72 ± 0.11	4.70 ± 0.29	0.125 ± 0.015
Combined	95.48 ± 0.15	9.80 ± 0.16	0.917 ± 0.031
$\mu + SVT$ data			
Class 1	10.51 ± 0.04	18.35 ± 0.37	0.354 ± 0.014
Class 2	28.76 ± 0.07	12.18 ± 0.29	0.426 ± 0.020
Class 3	56.45 ± 0.09	5.29 ± 0.26	0.158 ± 0.016
Combined	95.72 ± 0.12	9.90 ± 0.15	0.938 ± 0.029

Table 4.5: Performances of the NN Jet Charge tagger on the $\ell + SVT$ samples, for the various jet quality classes and combined. The quantity $\sqrt{\varepsilon D^2/\varepsilon}$ is the effective dilution as resulting form the binning in the $|Q_{jet}| \cdot P_{nn}$ variable.

Chapter 5

The Opposite Side Kaon Tagging algorithm

In this chapter I describe the development and the evaluation of the performances of the Opposite Side Kaon Tagger. This exploits the correlation between the *away* b and the electric charge of the Kaon originated in its decay. Given the high track multiplicity and the dominance of fragmentation tracks in the generic sample considered for tagging, a Kaon tagger at a hadron collider relies essentially on our capability to identify a track as a Kaon and to evaluate how displaced it is from the $p\bar{p}$ collision vertex. This is effective in selecting Kaons from b decays. In the following text I illustrate how such tools contribute to select a sample of signal Kaons. Furthermore, I show that knowledge of the flight direction of the away b is useful to further suppress background.

Using these informations, I develop a tagging algorithm and evaluate its efficiency and dilution. I demonstrate that a dilution of the same order as the other Opposite-side taggers can be obtained and study the behaviour of the tagger in different configurations.

Use of Monte Carlo simulation is also performed to understand the contribution of the different charged particle species to the overall tagger dilution.

5.1 The tagging principle and the main *ingredients*

Expectations at the beginning of CDF Run II were that flavour tagging using the Particle Identification (PID) of a Kaon both in the Same Side and in the Opposite Side would be the main source of tagging power and the CDF Particle Identification system was designed accordingly. The installation of the TOF detector is functional to this purpose. Now that a Same Side Kaon Tagger is in place and is substantially contributing to the CDF final εD^2, I undertake here the application of PID in the Opposite Side to complete the scenario of Kaon

taggers.

The Opposite Side Kaon Tagger exploits the correlation between the flavour of the accompanying b from the incoherent $b\bar{b}$ production in $p\bar{p}$ collisions and the electric charge of the Kaon originated by its decay through the chain:

$$b \to c \to s$$

In particular, a K^- comes with a b, while a K^+ comes from a $\bar{b}$. Also Cabibbo-suppressed and $b \to s$ decays are useful and are considered in this study, since they carry the same flavour-charge correlation.

Once the information on the accompanying b is available, the flavour of the mixing candidate B_s is automatically inferred. In order to develop an OSKT I undertake the following steps:

- find a track coming from a decay vertex, that is displaced from the Primary Vertex (PV) where the b was created;
- identify that as a Kaon.

Thus, Particle Identification (PID) and tracking do represent the main ingredients for this analysis.

At momenta interesting for B physics ($|\vec{p}| \leq \mathcal{O}(m_B)$), the generic charged particle population is dominated by pions. Consequently, PID plays an important role in suppressing background and enhance the Kaon component in the inclusive ensemble of tracks selected, distant from the candidate B_s.

Then, information on how much a track is likely to have originated from a Secondary Vertex (SV) is helpful to suppress the contribution of Kaons produced in the fragmentation process at the PV. These represent the main source of background after suppressing the pion component with PID.

In this study I also find that the tagger performances are consistently higher in the subsample of events where a b-jet has been explicitly identified in the *away hemisphere* by looking for a SV within the jet cone. Thus, the algorithm looks for such jets and the tags are classified based on whether or not the Kaon is close to a well-identified SV.

The performances of this algorithm can be completely assessed using a calibration sample *of data*. Infact, using a sample where the flavour of the b hadron is known, one can infer the flavour of the *accompanying* b; then he can compare it with the response given independently on the same quantity by the OSKT. By doing that, a completely data-driven measurement of the tagger dilution:

$$D = \frac{N_R - N_W}{N_R + N_W} \tag{5.1}$$

(introduced in Eq.4.11) is obtained.

5.2 Data and Monte Carlo samples used in the tagger development

5.2.1 Data

Following the CDF standard choice in the study of *Opposite Side* taggers, I use a high statistics inclusive sample of b semileptonic decays selected by the dedicated CDF $\ell + SVT$ trigger. This corresponds to an integrated luminosity of $\int \mathcal{L} = 355\ pb^{-1}$. The trigger selections are listed in Sec.3.3.5. The physics features of this sample have been described in detail in Sec.4.3.2. With ≈ 1.5 million events available after the background subtraction procedure, and given the tagger efficiency, the dilution D can be determined with a $\leq 10\%$ fractional uncertainty.

5.2.2 MC

I also make use of a Monte Carlo sample for a better comprehension of the various contributions entering the tagger dilution. This is a `PYTHIA`-generated [33] sample of 98 million events. All the flavour creation processes are simulated (setting the so-called $MSEL = 1$ configuration) and a filter on $b - \bar{b}$ events is then applied. $b-$hadrons are decayed by the `EvtGen` ([34]) package following the $b/\bar{b} \rightarrow \mu^{\mp} D_X(\rightarrow mK + n\pi)Y$, where μ^+ is for $\bar{b}$ decays; D_X represents a D meson going into m Kaons and n pions.

The mixing of the B_q $(q = d, s)$ mesons is not simulated at this stage and is accounted for later, by reproducing this effect analytically at the analysis level (as described in Sec.5.10). A full detector and $\ell + SVT$ trigger simulation is also performed. The same track-based requirements applied to data are also used for Monte Carlo. At the end of the analysis selection, a total of ≈ 100 thousand simulated events are available.

5.3 Tagging track selections

A preliminary set of requirements is applied to tracks in the $\ell + SVT$ sample, both for data and Monte Carlo. Such requirements are listed in Tab.5.1. The cuts on the number of hits in the tracking volume ensure a minimum quality in the track's reconstruction. In particular, the requirement on Silicon hits helps reducing the contribution of tracks produced in nuclear interactions away from the primary vertex. A minimum p_t cut is also applied to select only tracks that have enough momentum to reach TOF bars.

The $\Delta R(\ell + SVT, track)$ cut defines the *Opposite Side* and it is such that the complementary space is used to look for *Same* Side Kaon tags. Fig.5.1, taken from MC, shows the expected distance between the real trigger B direction and Kaons coming from the $b \rightarrow c \rightarrow s$ decays in the OS. This selection reduces the

quantity	cut
Silicon Hits r-ϕ	≥ 3
COT Hits	$\geq 10 + 10$
p_t	$\geq 350\ MeV/c$
$\Delta R(\ell + SVT, track)$	> 0.7
$\lvert d_0 \rvert$	$\leq 0.1\ cm$
$\lvert z_0(trk) - z_0(trigger\ \ell) \rvert$	$\leq 1.2\ cm$
dE/dx proton removal	Yes

Table 5.1: Fiducial and PID-related selections applied to tracks

efficiency mainly for events in which the $b\bar{b}$ pair was produced by gluon splitting (see Sec.5.7 for a recall on flavour production mechanisms) and the two b are collinear. I estimate from Monte Carlo simulation that the efficiency of this requirement on Kaons from away b decays is 88%.

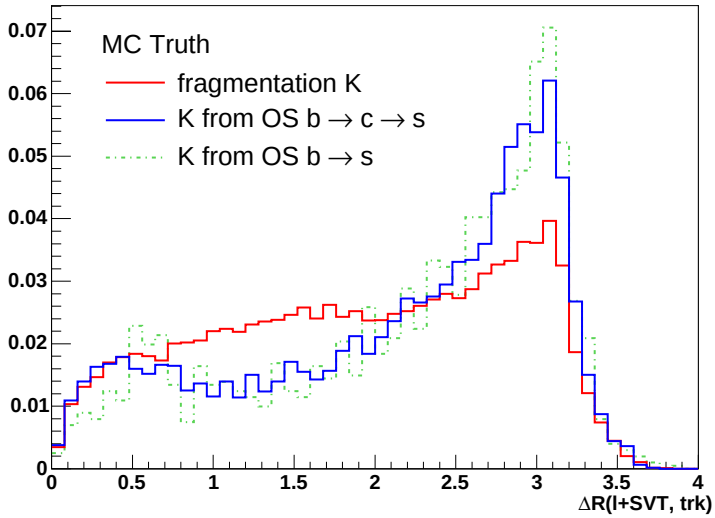

Figure 5.1: MC distribution of the ΔR between signal Kaons and the trigger pair. Signal Kaons from $b \to c \to s$ (blue) and from $b \longrightarrow s$ (green) are displayed together with fragmentation (red), showing a flatter cone distribution for the latter. The raise of signal at smaller cones is due to gluon splitting.

My last requirements concern tracking. The upper cut on the candidate track impact parameter, d_0, is introduced in order to suppress daughter tracks coming from decays of the long-lived components of the tagging sample (Λ and K_s^0). The impact parameter is calculated here with respect to the true position of the beam line.

The information on the transverse plan is complemented with the one along the beam axis z using the quantity:

$$\Delta z_0 = z_0(trk) - z_0(trigger\ \ell)$$

The B-vertex is estimated here using the z_0(trigger ℓ), that is the z_0 coordinate of the trigger lepton track. Both the z_0 of the candidate track and of the trigger lepton are corrected for the true position of the beam, as well. I also considered estimating the z_0 of the B from the average z_0 position of the two lepton and the SVT track. No sensible improvement with respect to the lepton-only case was found, given the stiff trajectory of the high-p_t lepton track. The cut along the beam coordinate also reduces the effect of additional PV produced in collisions at higher instantaneous luminosities, see Fig.3.3. The values of both the $|d_0|$ and $|\Delta z_0|$ cuts are subject to an optimization process, described in Sec.5.9.1.

Protons from the beam halo and those originated in the secondary nuclear interactions of particles from the PV with the detector material are also present in the tagging sample. On average they have a momentum $|\vec{p}| \leq 1\ GeV/c$. Given their β, a high ionization in the COT occurs. The distribution of the dE/dx at low momenta is shown in Fig.5.2.

Furthermore, such particles are produced at a different time with respect to the event's t_0, so that their TOF measurement is incorrect (see Fig.5.2). This results in their misidentification mainly as Kaons. Fig.5.2 displays the value of the Likelihood Ratio of such protons, where the Likelihood Ratio is the PID quantity defined in Eq.5.3 and described in Sec.5.4.

It has also to be considered that a charge asymmetry is introduced by such protons in the tagging sample. More p than $\bar{p}$ are present due to their relative abundance in the beam halo and different interaction cross sections with the detector material. This biases the measurement of the dilution D. The absolute number of p and $\bar{p}$ is shown in Fig.5.2 as a function of the proton momentum. The relative difference is $\approx$ 30% for $p_t \leq 800\ MeV/c$.

Thus, I discard tracks with an average specific ionization $< dE/dx > \geq\ 25\ ns$ and not compatible with the Kaon mass hypothesis.

5.3.1 Trigger side daughter removal

A removal of the daughters of the trigger side B passing the $\Delta R(\ell + SVT, track)\ >\ 0.7$ cut is also performed. Infact, Kaon tracks coming from the decay of the trigger side B have a charge-flavour correlation with the *away* b opposite to expectations. Thus, they reduce the dilution D. This is shown in Fig.5.23, from Monte Carlo.

Such tracks are expected to have a smaller average ΔR than tracks originated from the away b. Consequently, a combined cut on the invariant mass $M(\ell + SVT + trk)$ and the $\Delta R(\ell + SVT, track)$ is effective in rejecting the former. A

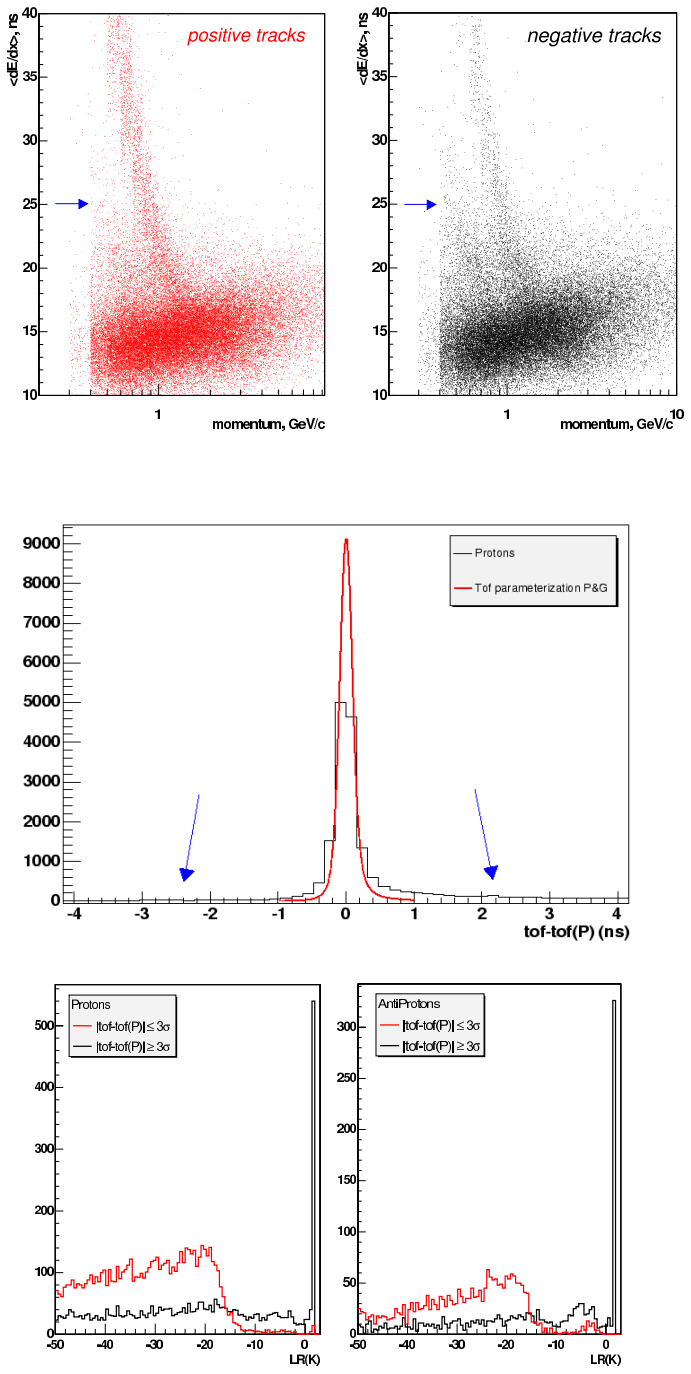

Figure 5.2: Above: TOF distribution for protons as identified using the dE/dx at low momenta. Tails are seen with respect to TOF parameterization (red) corresponding to particles having production time not compatible with the event t_0. Below: Likelihood Ratio in the Kaon hypothesis for protons and anti-protons again as selected according to their specific ionization. In red those having a TOF compatible with the proton hypothesis, in black those coming off-time (wrong TOF measurement).

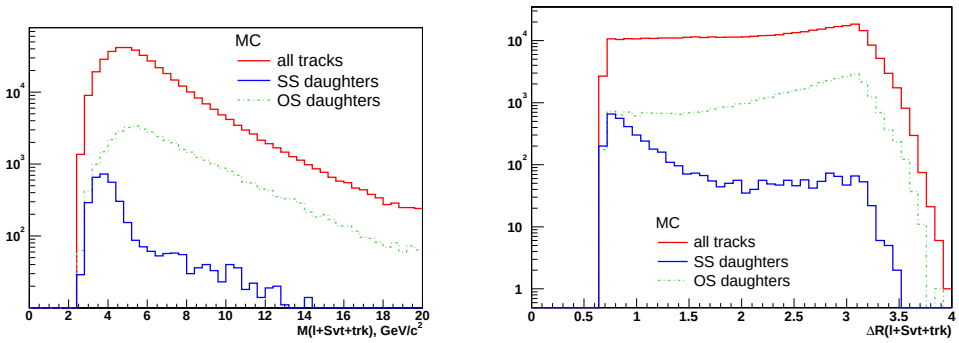

Figure 5.3: Left: Invariant mass distribution of $\ell + SVT$ plus any other track; the same for OS b daughters only (green) and SS b daughters only (blue) are overlaid. Right: the same for the $\Delta R(\ell + SVT, track)$ distribution is displayed.

track is rejected if it satisfies both the requirements:

$$\begin{aligned} M(\ell + SVT + trk) &\leq 5.4\ GeV/c^2 \\ \Delta R(\ell + SVT, track) &\leq 1.6 \end{aligned} \tag{5.2}$$

These values are obtained from a Monte Carlo optimization aimed at reducing the mistag rate maximizing the tagger εD^2. Monte Carlo distributions of $M(\ell + SVT + trk)$ and $\Delta R(\ell + SVT, track)$ and the scatter plot of one against the other are shown in Figs.5.3 and 5.4. A tail is visible for trigger B daughters at values of mass higher than the B meson masses, caused by random hadron tracks faking the trigger b lepton.

The fraction of events tagged by a trigger b daughter *after* applying the rejection cuts is estimated from Monte Carlo to be 1.7%. To confirm the MC optimization, I also look at data. I evaluate the performances of the OSKT in the case when no trigger b daughter removal is applied (*higher efficiency*). On the other hand, cases when the selection is tightened by rejecting *any* tracks with $\Delta R \leq 1.6$, regardless their $M(\ell + SVT + trk)$ are also considered (*higher dilution*). The overall tagger performances are lower in both cases and are listed in Sec.5.10.

5.4 Kaon Identification

Among tracks satisfying the preliminary selections of Tab.5.1, I select Kaons explicitly. Kaon identification at CDF is provided by the specific ionization dE/dx in the COT and the particle Time-Of-Flight measurement. I have described the features of these two tools separately in Sec.3.2.2 and 3.2.3, where I also provided an explanation of how their response to the various particle species is modeled

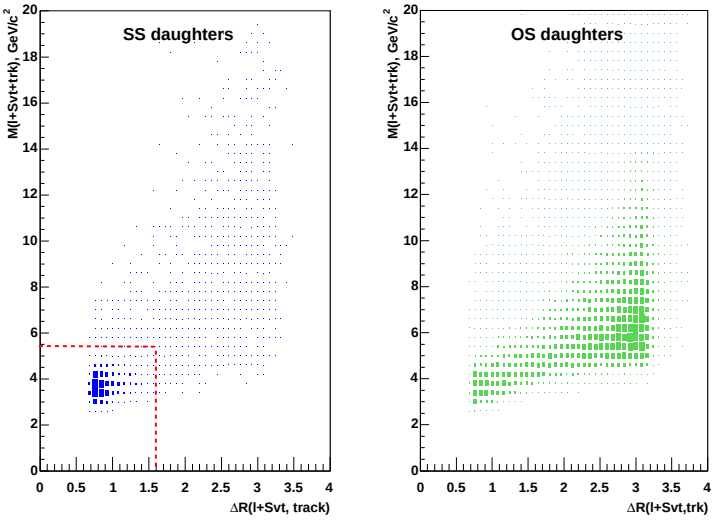

Figure 5.4: 2-D plots show the invariant mass versus ΔR for SS b daughters in the OS (left) and OS b daughters in the OS (right). Both plots use MC truth information

(*and simulated*). Given such parameterizations, one can evaluate the separation power as a function of track momentum, which is summarized in Figs.3.12, 3.15 for dE/dx and TOF respectively. As can be seen, a strong dependence of the separation power on momentum is present in both cases. On an inclusive sample like ours, where track momenta can span in a wide range due to their different origins, it is instead desirable to have a Kaon identification tool as momentum-independent as possible. Therefore, I combine the TOF and dE/dx decisions introducing the Likelihood Ratio:

$$LR(K) = \frac{\mathcal{L}(K)}{0.7 \cdot \mathcal{L}(\pi) + 0.2 \cdot \mathcal{L}(K) + 0.1 \cdot \mathcal{L}(p)} \tag{5.3}$$

where

$$\mathcal{L}(K) = P_{TOF}(K) \cdot P_{dE/dx}(K) \tag{5.4}$$

and P_{TOF} ($P_{dE/dx}$) is the TOF (dE/dx) Probability Density Function (P.D.F.). A ratio is used following Pearson-Neyman lemma on hypotheses separation. In the OSKT case the two hypotheses are that a given track is a signal Kaon (*numerator*) and that the track is part of the mixed pion, Kaon and proton generic background. Each term in the denominator is weighted by an overall relative fraction. These fractions have been determined in [29], by separating the various components in the sideband of the $\phi \to K^+K^-$ invariant mass distribution with TOF. The TOF and dE/dx P.D.F. are determined as described in Sec.3.2.3, based on the

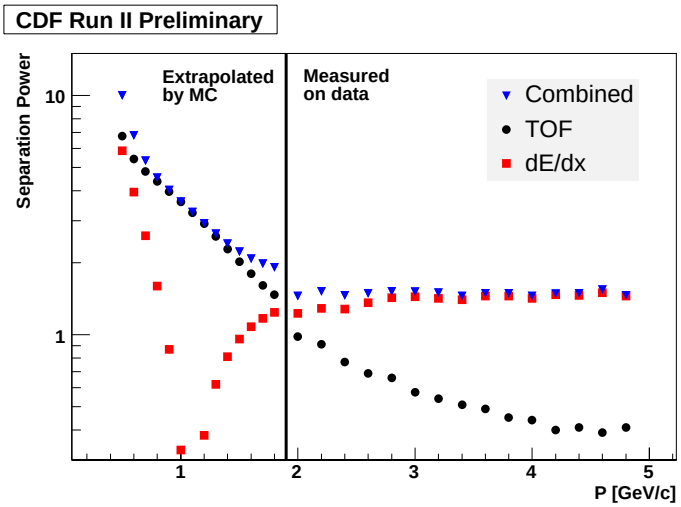

Figure 5.5: K/π separation power of TOF(black), dE/dx(red) and combined in the LR(K), as obtained on a sample of soft pions from D^*, and K from D mesons from D^*.

response of the two subdetectors on samples pure in the different particle species. In particular, $K-\pi$ separation is evaluated on a 355 pb^{-1} $D^{*\pm} \rightarrow D^0\pi^{\pm}_* \rightarrow [K^{\mp}\pi^{\pm}]\pi^{\pm}_*$, where the K and π are identified exactly by their correlation with the charge of the soft π_*. In the momentum range where no calibration samples are available, a toy Monte Carlo extrapolation is performed that uses the TOF and dE/dx P.D.F. parameterizations from data. I checked that no correlation occurs between the TOF and dE/dx P.D.F., so that no further terms are to be taken into account when building $\mathcal{L}$. The improvement provided by using a combined PID quantity instead of any of the two separately is shown in Fig.5.5.

In Fig.5.6 the distributions of log(LR(K)) separately for TOF and dE/dx and together in the combined LR are shown for tagging candidates. Given the definition of LR(K) adopted, in the optimal case all the Kaons are found in the right-most bin, where $log(LR(K)) = log(\frac{1}{0.2}) = 1.61$. Pions can be found well separated toward lower negative log LR values. Each of the Monte Carlo distributions presents a lower Kaon fraction than in data, probably reflecting the absence of a fraction tuning in simulation for the specific features of my physics sample. Nevertheless, since the tagger dilution is estimated relying *entirely* on data and no Monte Carlo model is used to assess the final performances, I do not investigate possible sources of disagreement more thoroughly.

In the specific case of the OSKT, tagging candidates have a p_t spectrum peaking at around 1 GeV/c, shown in Fig.5.7. In this region TOF is most powerful in separating K and π, as recalled from Fig.3.15. For instance, TOF is able to separate K from π at $> 3\sigma$ level for $p = 1.0$ GeV/c.

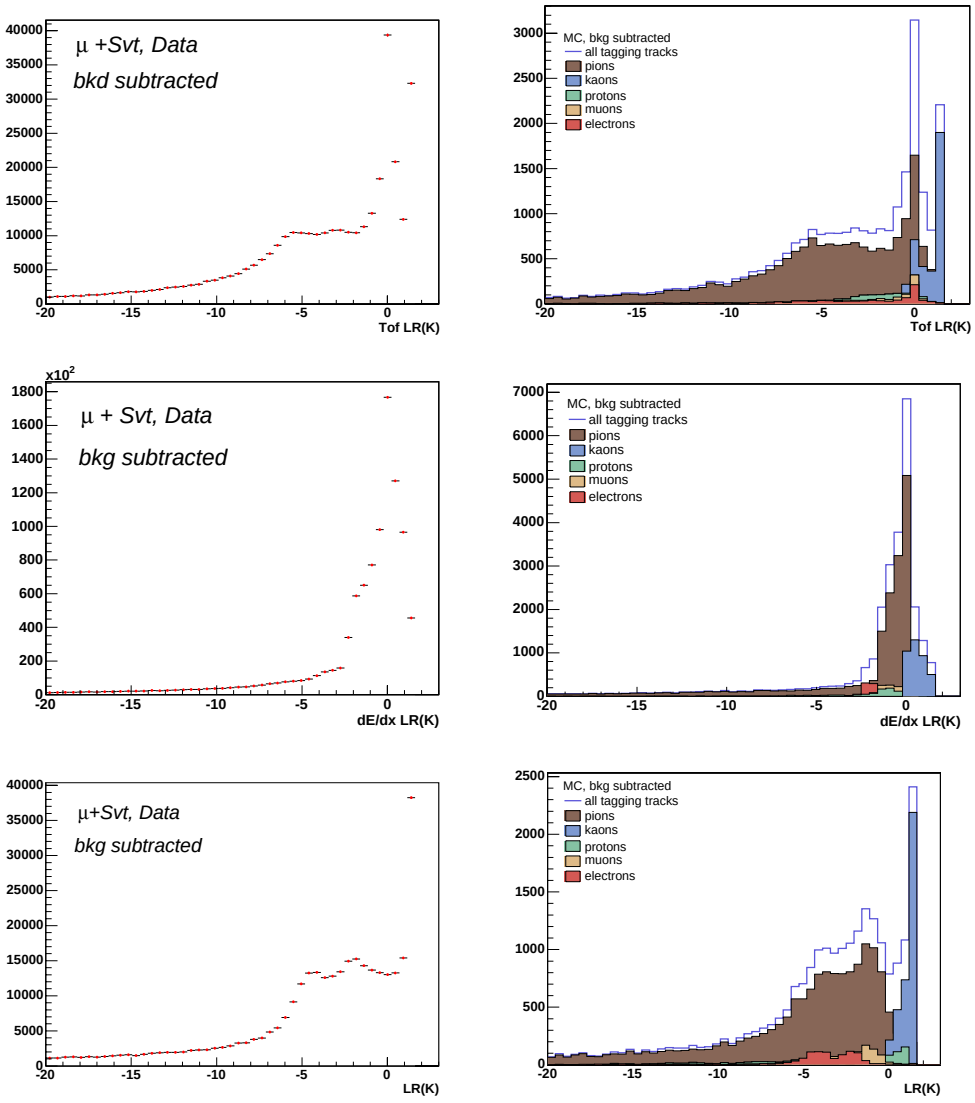

Figure 5.6: Distributions of LR(K) for TOF (above), dE/dx (middle) and combined TOF+dE/dx (below). Left column shows the distributions for data, right column for MC, with the contributions of the various particle types. Data and MC show roughly the same shape and from these I expect Kaons in the region LR(K) ≥ 0

Provided this, I choose to accept for tagging only tracks having a valid TOF measurement. Tracks without a TOF match are discarded. With the present TOF matching efficiency (Fig.3.16), this request represents the largest suppression to the OSKT tagging efficiency. From Monte Carlo I quantify the εD^2 improvement coming from the inclusion in the algorithm of tracks having dE/dx-only information as only a fractional $\approx 5\%$.

5.5 Track's displacement

I apply a cut to select tracks significantly displaced from the PV, most likely to come from the $b \to c \to s$ decay chain.

The main source of background in the inclusive semileptonic sample is represented by tracks generated at the PV by the underlying event in the $p-\bar{p}$ collision. In particular, after applying the PID selection this contribution is enhanced in fragmentation Kaons. To reject them, I use the significance of the the absolute impact parameter:

$$\frac{|d_0|}{\sigma(d_0)}$$

where the impact parameter is calculated with respect to the true position of the beam.

Usage of the significance rather than only d_0 provides a more effective tool to deweight poorly-measured tracks by the Silicon detector. With this respect, I refit all tracks in the Kaon mass hypothesis, using the information from L00 hits, if available. L00 is described in Sec. 3.2.2. The fraction of tracks having at least 1 L00 hit is $\approx 40\%$. A reweighting of both Monte Carlo L00 efficiency and single-hit resolution is applied in order to reproduce data (see Fig.5.8).

I also take into account the uncertainty coming from the beam width. This is $\approx 30\ \mu m$ and has a dependence on the PV coordinate along the z axis. Therefore, I add in quadrature the beam width measured for a data store with the $\sigma(d_0)$ coming from tracking. The d_0 is better measured for higher p_t tracks, less subject to multiple scattering in the Silicon detector. Thus, a bias toward higher p_t is introduced in the tagging sample by the I.P. significance cut. Nevertheless, this bias is eliminated when considering the beam width properly. Infact, while this contribution is small with respect to the $\sigma(d_0)$ from tracking at low p_t, it becomes more important the higher the p_t. As a result, Fig.5.9 shows the d_0 significance as a function of p_t: the trend is removed when introducing the beam width.

I introduce a beam width dependence on the z-coordinate of the primary vertex:

$$\sigma_{beam} = \sqrt{\varepsilon(\beta^* + (z - z_{min})^2/\beta^*)} \tag{5.5}$$

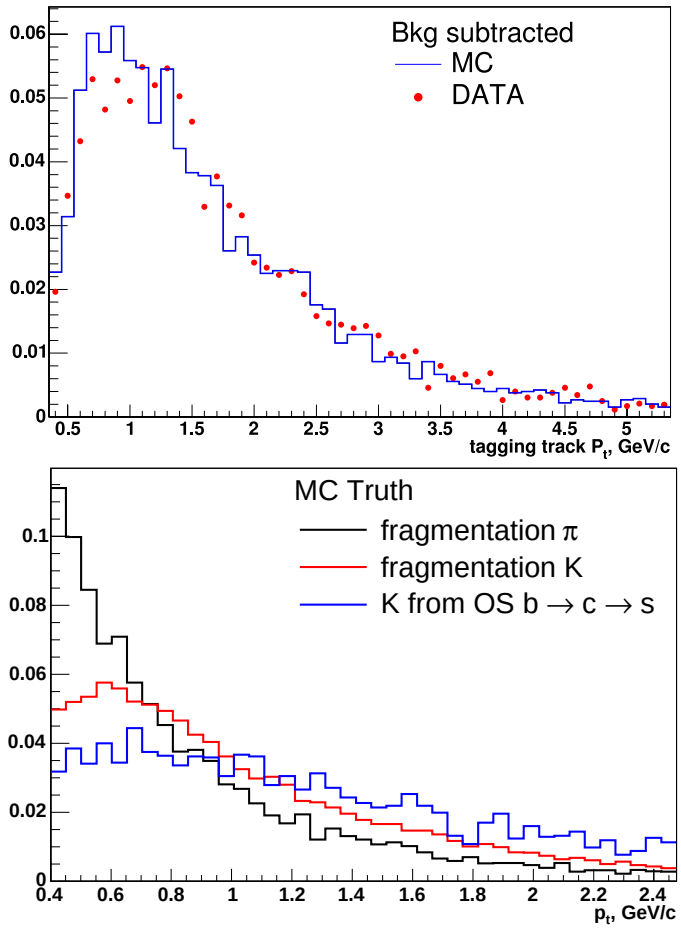

Figure 5.7: Upper plot: p_t spectrum of tracks passing all the tagging selections described, for data (red dots) and MC (blue line). Lower plot: Monte Carlo true p_t distribution for signal Kaons (blue), fragmentation Kaons (red) and fragmentation pions (black). As expected, fragmentation particles are softer than decay daughters.

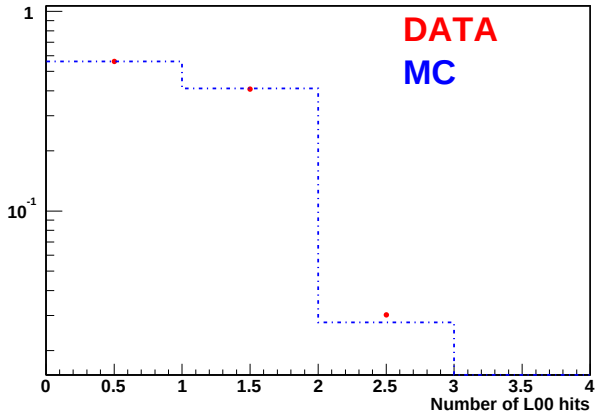

Figure 5.8: Number of L00 hits associated to a track, for Monte Carlo and data

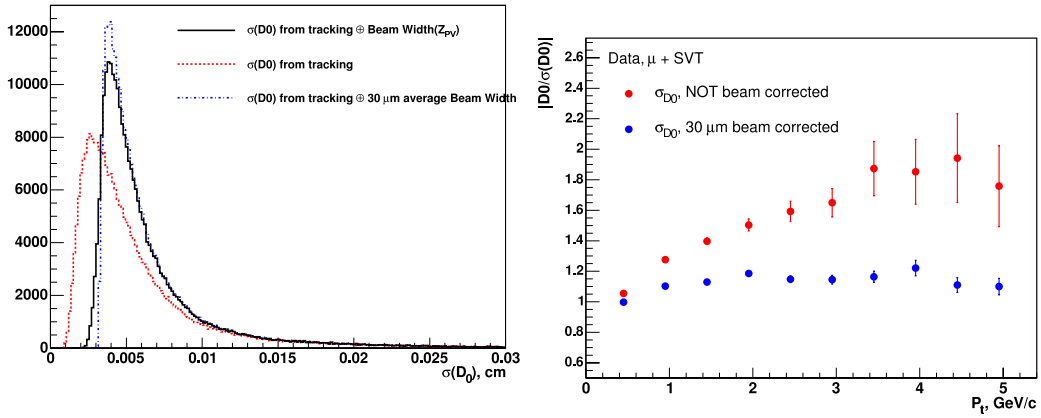

Figure 5.9: Left: Distribution of $\sigma(d_0)$ for various configurations: using only the CDF tracking output (red); with an average $30\mu m$ beam-width added (blue); and with the $z-$ dependence (black). Right: difference in the Gaussian fit to $d_0/\sigma(d_0)$ as a function of track's p_t, with and without the beam-width included. A bias toward higher p_t is introduced by cutting on I.P. significance, if I use the $\sigma(d_0)$ as-is from tracking.

where z_{min} is the location of the beam width minimum; ε is the tranverse emittance of the beam and β^* the amplitude function, related to bunch preparation and beam optics respectively. These parameters are retrieved from a fit performed for every data store and are also used in the simulation for consistency. The left plot in Fig.5.10 shows the impact of L00 on $\sigma(d_0)$ as a function of the track's z_0. Roughly speaking, a 15% improvement is found, as estimated from a gaussian fit to the impact parameter distribution. I also look for effects of high $|d_0|$ tracks by using exponential tails in the fit and find no sensible deviation from this result.
The model used for the z-dependence of the beam width is checked by fitting the $d_0/\sigma(d_0)$ distribution in bins of track's z_0: again I use a core gaussian for the resolution plus two side exponentials to account for any possible tails. The fit pulls are displayed as a function of z_0 in the right plot of Fig.5.10: as expected, they are consistent with the 1.4 average scale factor found for the proper decay time uncertainty [66]. The effect of the tails is negligible. After all these effects are considered, the resulting impact parameter significance is shown in Fig.5.11 for data and Monte Carlo simulation. A reasonable agreement is found. As expected, Monte Carlo shows that fragmentation tracks peak at zero, while decay products of the away b are more displaced and display an exponential tail toward longer decay lenghts as a result of the b hadron lifetime.

5.6 Signed I.P significance

To enhance the dilution I exploit the fact that, while tracks from PV have a symmetric d_0 around zero, the signal Kaons are distributed toward higher values of the impact parameter. Furthermore, these are originated ahead of the b decay vertex. Consequently, signing the I.P. with respect to the position of the b decay vertex is expected to reduce the prompt contributions significantly, at the expenses of a minimal signal efficiency loss.

For cases where the b decay vertex is explicitly identified I substitute the absolute I.P. in the expression of the significance with the d_0 signed with respect to the SV transverse distance from the PV:

$$IP_{sgn} = |d_0| \cdot \frac{(-d_0 sin(\phi_0) * (x_{SV} - x_{PV}) + d_0 cos(\phi_0) * (y_{SV} - y_{PV}))}{|-d_0 sin(\phi_0) * (x_{SV} - x_{PV}) + d_0 cos(\phi_0) * (y_{SV} - y_{PV})|} \qquad (5.6)$$

Fig.5.12 shows the signed I.P. distribution for fragmentation and signal Kaons. The relative improvement in dilution by using the signed I.P. instead of the absolute I.P. is found to be 5%. The corresponding decrease of tagging efficiency is negligible. I also compare this definition with the case when the identified b-jet momentum is used to sign d_0. I find a modest improvement in dilution ($\leq 5\%$), but a $\approx 30\%$ relative loss in signal efficiency. This is explained by the sketch in Fig.5.12, when the information on the d_0 sign for signal Kaons is smeared due to

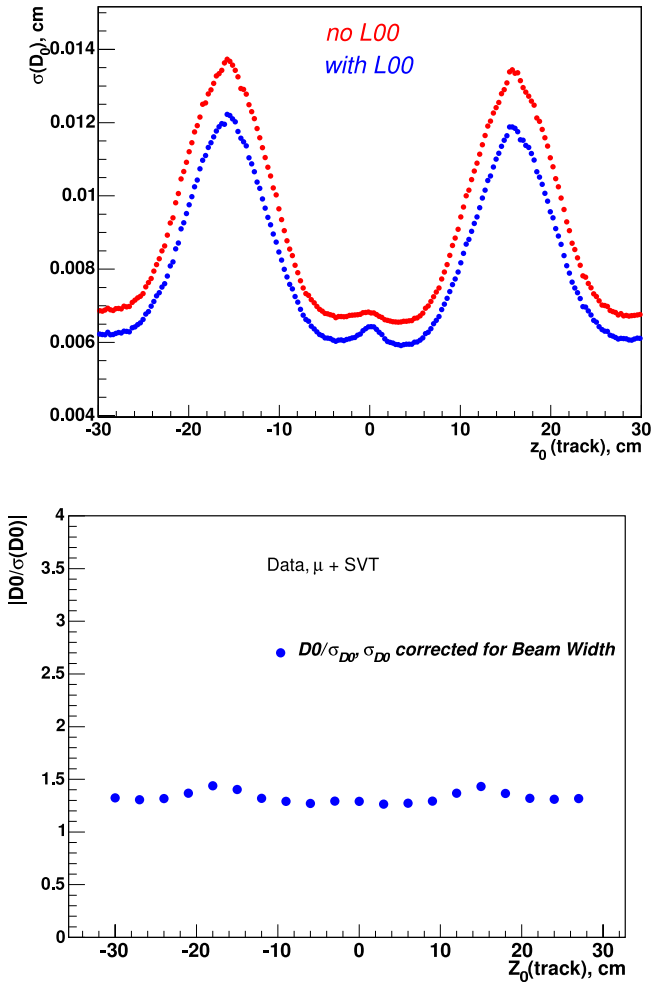

Figure 5.10: Left: Improvement on $\sigma(d_0)$ produced by L00 usage along the beam's z coordinate. Right: Pulls of the Gaussian+exponential tail fit to the impact parameter significance as a function of z. The 1.4 known scale factor is returned, as expected.

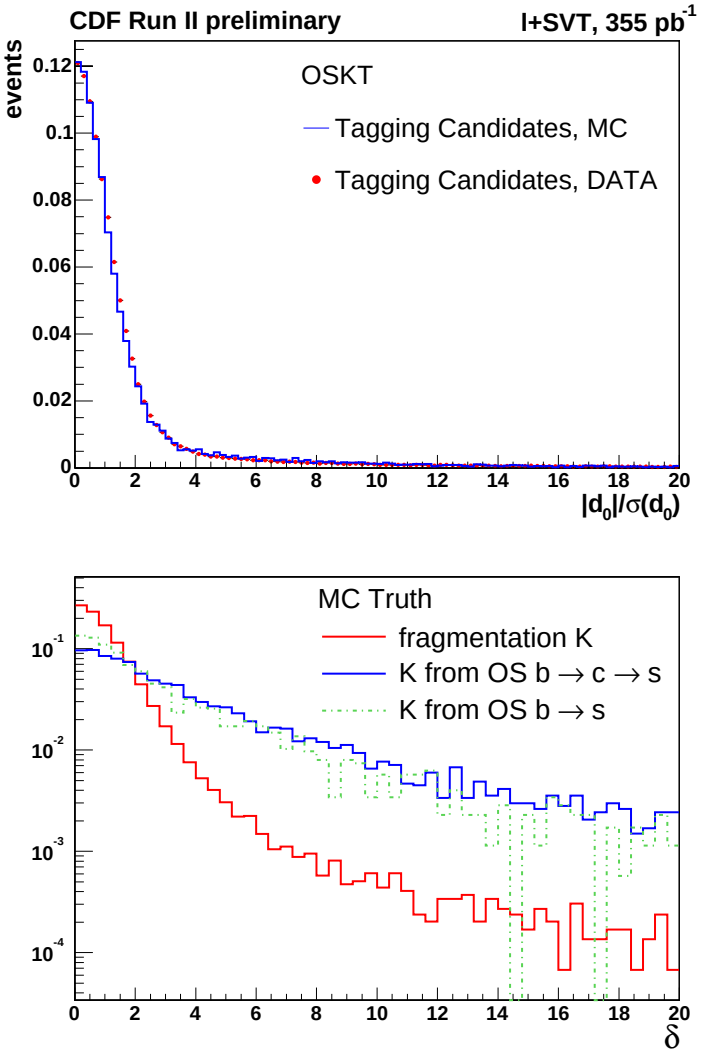

Figure 5.11: Top: Displacement $|d_0|/\sigma(d_0)$ in data (red dots) and agreement with MC (blue line) after reweighting the silicon hit efficiency and resolution. $\sigma(d_0)$ is corrected by the beam-width. Bottom: I.P. significance $|d_0|/\sigma(d_0)$ for OS signal Kaons and fragmentation Kaons, from MC truth.

the jet opening angle, resulting in an increase of cases when a Kaon from the b decay is assigned a negative d_0 and discarded.

It has to be noted that use of the signed I.P. instead of the absolute one does not produce any improvement in dilution when the no SV is identified within the jet is not identified explicitly, so that I prefer to use the latter for events when no SV is explicitly found.

5.6.1 ϕ removal

ϕ mesons can be originated in the opposite side by either $b \to c \to s$ decays or in the fragmentation process. Kaons from the $\phi \to K^+K^-$ decays have an equal probability to be correlated and anti-correlated with the flavour of the OS b and are not useful for tagging. Consequently, I look for combinations of a tagging candidate with any other opposite charge track in the event that passes the LR(K), and search for ϕ mass peaks around the PDG value ($1019.456 \pm 0.020\ MeV/c^2$, [21]) in the invariant mass distribution. With these requirements no significant peak is found.

Instead, a peak at the expected value is observed when combining the four-momenta of pairs of opposite charge *tagging* candidate tracks. In this case, both tracks satisfy the LR(K) *and* $d_0/\sigma(d_0)$ cuts. The invariant mass distribution is visible in Fig.5.13. I apply a ϕ removal cut in the mass window:

$$1010\ MeV/c^2 \leq M_{KK} \leq 1030\ MeV/c^2$$

for pairs of tagging candidates.

I also look for K^0_s and for $D^0 \longrightarrow K^+K^-$ decays, but no peak is observed in either cases.

5.7 Determination of the OS b direction

$b\bar{b}$ pairs are produced mainly by a process of the kind (see Fig.5.14):

$$q + \bar{q}\ (g + \bar{g}) \to b\ \bar{b}$$

This is called *flavour creation*. Other ways of producing $b\bar{b}$ pairs are from the collisions of $b/\bar{b}$ from the initial state onto gluons, generating $b/\bar{b}$ in the final state (*flavour excitation*) and from the conversion of a gluon into a $b\bar{b}$ pair (*gluon splitting*). In the case of flavour production, the two b are likely to be originated back-to-back. Fig.5.14 shows the $b - \bar{b}$ angular distribution for a $b\bar{b}$ pair. Signal Kaons are expected to be created close in angle with the parent b hadron. Consequently, their $\Delta R(\ell + SVT, trk)$ distributes as in Fig.5.1, with a peak at ≈ 3, corresponding to the back-to-back configuration. Fragmentation is isotropic with respect to the trigger b direction, instead. Using MC, I investigated the idea of

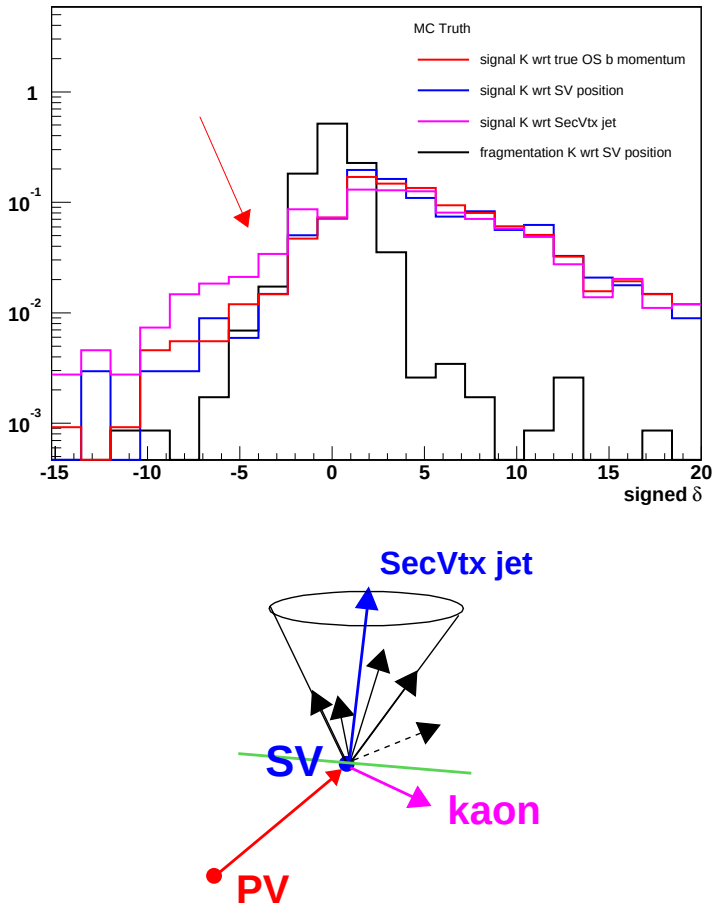

Figure 5.12: Left: MC signed I.P. significance with respect to the SV direction (blue) and momentum of the associated jet (violet) for the signal K. Also the MC truth with the OS b direction is plotted as a reference (red). Fragmentation (black) is overlaid to show how this variable discriminates signal and background. The arrow indicates the different tails. Right: sketch showing a situation when the sign given by the jet and the position of the SV are opposite for a signal track.

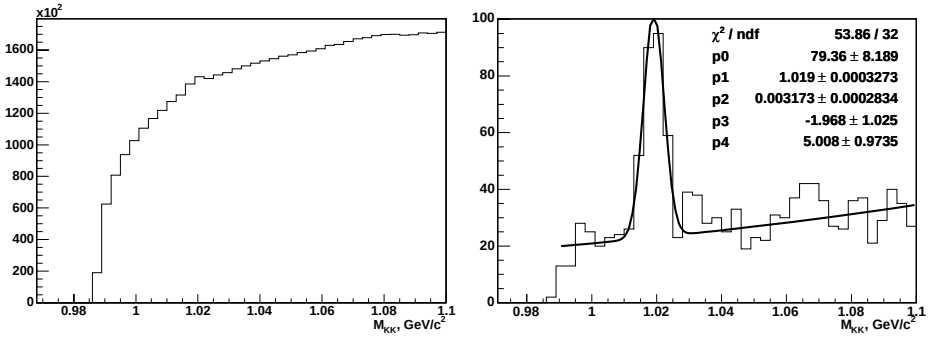

Figure 5.13: Left: invariant mass distribution in the ϕ mass window for a tagging track with any other track that passed the LR(K) requirement, such that their charges have opposite sign. No significant peak is found. Right: after applying also the requirement on the displacement, pairs of tagging candidates show a peak significant close to the expected PDG value. The fit is a Gaussian plus an exponential.

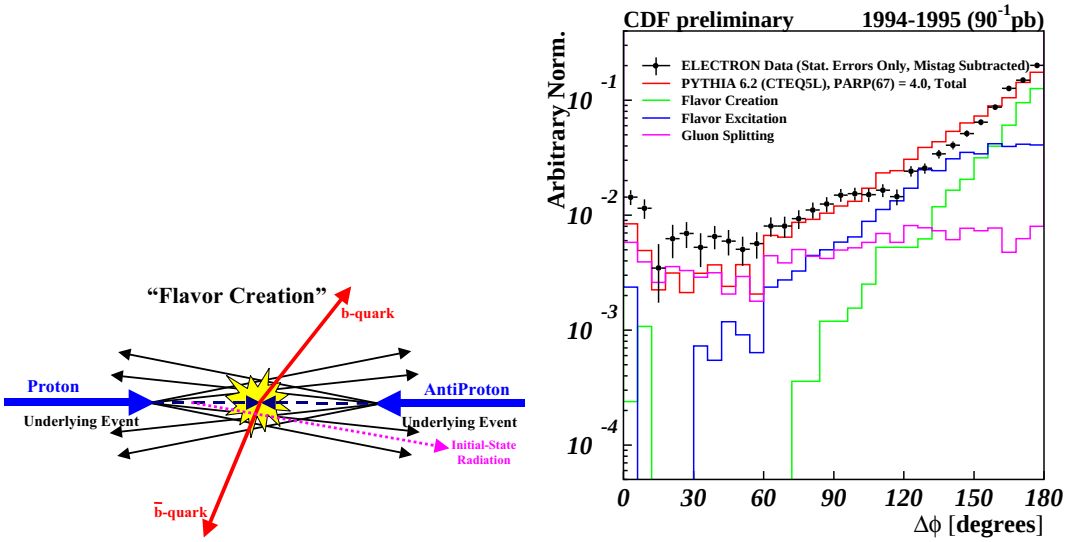

Figure 5.14: Sketch of $b\bar{b}$ pair production by flavour creation mechanism at a $p\bar{p}$ collider. This is the most relevant flavour production mechanism and the final pair is created back-to-back. The plot shows the $\Delta\phi$ angular difference between the two quarks as measured using incluseve electron and muon samples where a Secondary Vertex was reconstructed. From Ref.[48]. The fraction of b hadrons with $\Delta\phi \leq 90^\circ$ is $28.8 \pm 1.0(\text{stat}) \pm 3.1(\text{syst})\%$

exploiting this topology to reduce fragmentation further. Such studies indicated that closing the cone around the OS b would enhance the dilution D of the tagger significantly ($\approx \times 2$ factor), at a modest expense of the efficiency ε. Such studies were performed using the MC true flight direction of the away b.

5.7.1 Jet clustering

In order to identify the away b direction on data, I reconstruct candidate jets with an angular distance from the momentum of the $\ell + SVT$ pair. A cut $\Delta R(jet, \ell + SVT) \geq 1$ is applied to exclude the jet associated with the b-hadron on the trigger side.

I then look explicitly for a jet containing a decay vertex. If found, the jet axis is used to estimate the b flight direction. Otherwise, the highest momentum jet in the event or the jet containing the final tagging track are used as reference. I use a track-based jet clustering algorithm. A detailed description of the clustering algorithm can be found in [49]. In particular, I

- look for jet seeds using tracks with $p_t \geq 1\ GeV/c$;
- merge seeds that are closer than $\Delta R = 0.7$ to each other and form a new seed;
- when no more seed pairs to be merged are found, start associating softer tracks to them if they have:
  - $p_t \geq 400\ MeV/c$
  - $\Delta R(seed\ jet, trk) \leq 0.8$

No calorimeter information is used, since at this range of momenta the cone opening at the calorimeter level is too wide. More, the calorimeter energy resolution is not good enough to determine the momentum of the jet accurately. The final jet momentum is the sum of all tracks' momenta. Seed or softer tracks that was not possible to associate with others are considered *isolated*.

5.7.2 Jet selection

The multiplicity of the jets clustered in an event is shown in Fig.5.15. I check that the number of jets in an event is always smaller than the number of initial seeds, some of which may have been merged together during clustering. Moreover, I also verify that no seed is shared among two separate jets by checking that the ΔR among any two jets has always a minimum at 0.7 (*cone size*).

The average number of jets formed in an event is 6: for this reason, a selection of the jet most likely to come from the away b is required. I first look for jets containing a SV: the search is performed using the `SecVtx` (see e.g. [54]) algorithm

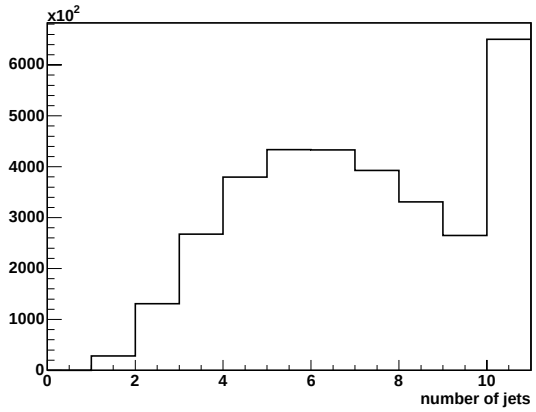

Figure 5.15: Number of jets clustered per event. Last bin contains events where ≥ 10 jets were clustered. The average of the distribution is at 6.

used for b-tagging purposes at CDF. I use a version with parameters optimized for the momentum range of interest for B-physics [49]. The rate of `SecVtx` tags in this sample is only 11%. Nevertheless, the distribution of the `SecVtx` tagged jets in Monte Carlo (Fig.5.16) shows how well such jets are able to estimate the away b direction. As can be noted from Fig.5.16, *right*, signal Kaons are distributed in ΔR close to the `SecVtx`-tagged jet axis. As expected, fragmentation tracks do not show any relationship with the jet direction, instead. Thank's to this different behaviour, I expect even this small amount of `SecVtx`-tagged events to provide the largest contribution to the final dilution.

If no `SecVtx` tag is found in the event, I consider two possible proxies to the OS b momentum: the highest momentum jet and the jet containing the candidate tagging track. The first one is used in the case of isolated tagging tracks, according to the definition given in 5.7.1. Considering this class of tracks independently proves on data to give the second best dilution of the tagging algorithm.

In no isolated tracks are found in the opposite side, choosing the jet to which a tagging track is associated guarantees that at least one of the tracks in the jet is displaced from the PV, due to the $d_0/\sigma(d_0)$ cut. This renders such jet more likely to be associated with a SV. This type of tags is expected to have the highest efficiency, as most of the tracks in an event are associated to a jet.

5.8 Classes of the OSKT

Following the event ranking based on the jet type found, the OSKT is divided into 3 classes, from the highest to the lowest in dilution:

1. **Class 1**: events with a `SecVtx` tagged jet;

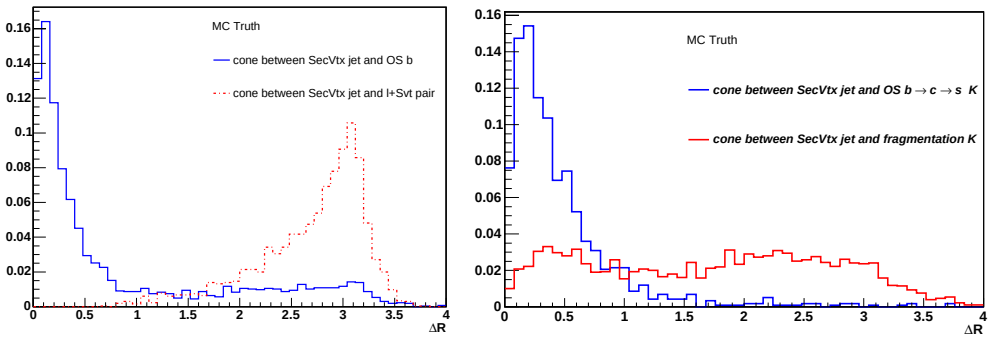

Figure 5.16: Left: Distance in cone between the SecVtx jet and the flight direction of the trigger (red) and OS (blue) b. Right: distribution of the ΔR between the signal OS $b \to c \to s$ K and the SecVtx jet (blue) and the same for fragmentation

2. **Class 2**: events with isolated candidate tracks;

3. **Class 3**: events with no isolated tracks; the reference jet is the one to which the candidate track is associated.

If ≥ 1 candidate tracks are found, associated to different classes, the highest class candidate is chosen.

5.9 Tagging algorithm

Finally, an algorithm based on track and jet quantities is required to choose among candidate tracks for events when > 1 track is selected. This happens for 11% of the selected events. The distribution of the tagging track multiplicity is shown in Fig.5.17 for Monte Carlo and data.

For this purpose, I consider three possible alternative algorithms:

- I choose the track with the highest I.P. significance;
- I choose the track most likely to be associated with a Kaon (*highest LR(K)*;
- I choose the one closest in ΔR to the reference jet.

While the first two options rely exclusively on track quantities, the last option is slightly more elaborate, as different mixes of the 3 types of jets may occur in the same event. In particular: if a `SecVtx` tag is found, the event belongs to Class 1 and the $\Delta R(jet, trk)$ is evaluated with respect to that jet. Otherwise, if one or more isolated tracks are found the event is flagged Class 2, and the highest momentum jet is taken as the reference vector. Finally, if the event is of Class 3, I calculate a $\Delta R(jet, trk)$ for each track with respect to its own jet. The one with the minimum ΔR is chosen.

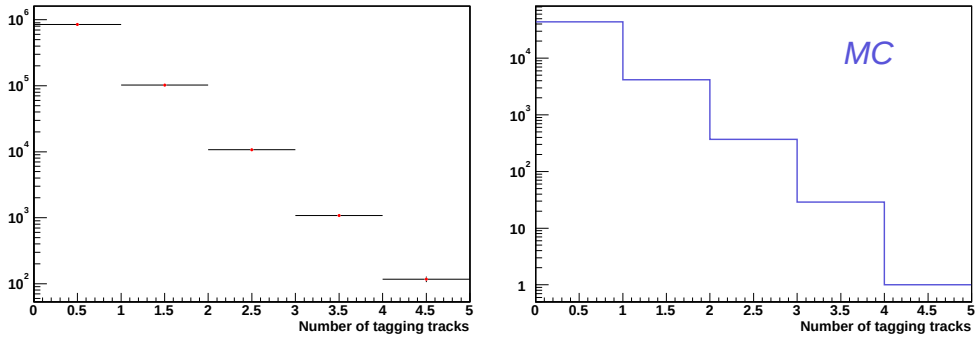

Figure 5.17: Multiplicity of candidate tagging tracks passing all the selection cuts with the optimizized values as of Sec.5.9.1; for data (left) and MC (right). The two distributions show a reasonable agreement. > 1 track is found in 11% of the cases fore data, 8% for MC.

Before evaluating the performances of the 3 above algorithms, I proceed to optimize the various cut values.

5.9.1 Cut optimization

The values of the various cuts applied for track selection are fixed following an optimization procedure. This is performed looking for a maximum of εD^2 in the 5-fold space of the quantities:

- $|d_0|$
- absolute I.P. significance
- signed I.P. significance
- LR(K)
- $|\Delta z_0|$

The signed I.P. significance cut is applied to Class 1 events only, while absolute I.P. significance selection is used for Class 2 and 3 only. The closest ΔR algorithm is run to obtain εD^2 at each point. I use half of the $\mu + SVT$ and half of the $e + SVT$ samples, reproducing the proportions of events for each data-taking period. The final performances are then evaluated on the other half of the total semileptonic sample in order to remove any bias.

The cuts are varied among the values listed in Tab.5.2. In the plots in Fig.5.19 the scans for the different variables are shown. A different tightness of the cuts for Class 1 on one side and Class 2 and 3 on the other side have been obtained, as expected from their different purity in b-jets. Nevertheless, I optimize on the

variable	cut values
$\|d_0\|$ (cm)	0.05, 0.1, 0.15, 0.20
abs IP signif	1.5, 2.0, 2.5, 3.0
LR(K)	-0.5, -0.3, -0.1, 0.0
$\|\Delta z_0\|$ (cm)	0.4, 0.8, 1.2, 1.6
signed IP signif	1.5, 2.0, 2.5, 3.0

Table 5.2: Ranges of variation for the cut optimization.

overall εD^2 and obtain a single set of cuts for the tagger.

The optimal values are reported in Tab.5.3. The corresponding values of efficiency, dilution and εD^2 are: 17.852 ± 0.067 %, 10.5 ± 2.0 % and 0.198 ± 0.014 % respectively.

Quantity	optimal cut value
$\|d_0\|$	0.1 cm
$\|d_0\|/\sigma(d_0)$	1.5
signed $d_0/\sigma(d_0)$	2.0
LR(K)	-0.3
Δz_0	1.2 cm

Table 5.3: Optimal values obtained for the *closest* ΔR algorithm d_0, $|d_0|/\sigma(d_0)$, LR(K) and Δz_0 and signed $d_0/\sigma(d_0)$.

I also estimate the signal efficiency and background rejection (defined as $1 -$ efficiency for background) of the various cuts. These are evaluated from Monte Carlo and are reported in Tab.5.4

The PID selection shows a very high rejection to pions, which is expected from the good separation observed in the LR(K) distributions in Fig,5.6. In fact, the optimized cut rejects $\geq 90\%$ of π, accepting essentially all the Kaons. The efficiency is not different for signal Kaons and fragmentation Kaons, due to the similarity of their p_t spectrum, from Fig.5.7. Although effective in rejecting prompt Kaons, the optimized values for the displacement variables have a lower acceptance to signal Kaons.

5.10 Performances

I apply the optimized cut values and proceed to evaluate the tagger performances. The *raw* dilution I find on such data sample has to be corrected for the factor 0.6412 accounting for the effects that the B mixing and sequential decays on the trigger side induce on dilution (Sec.4.3.2). The uncertainties on this factor

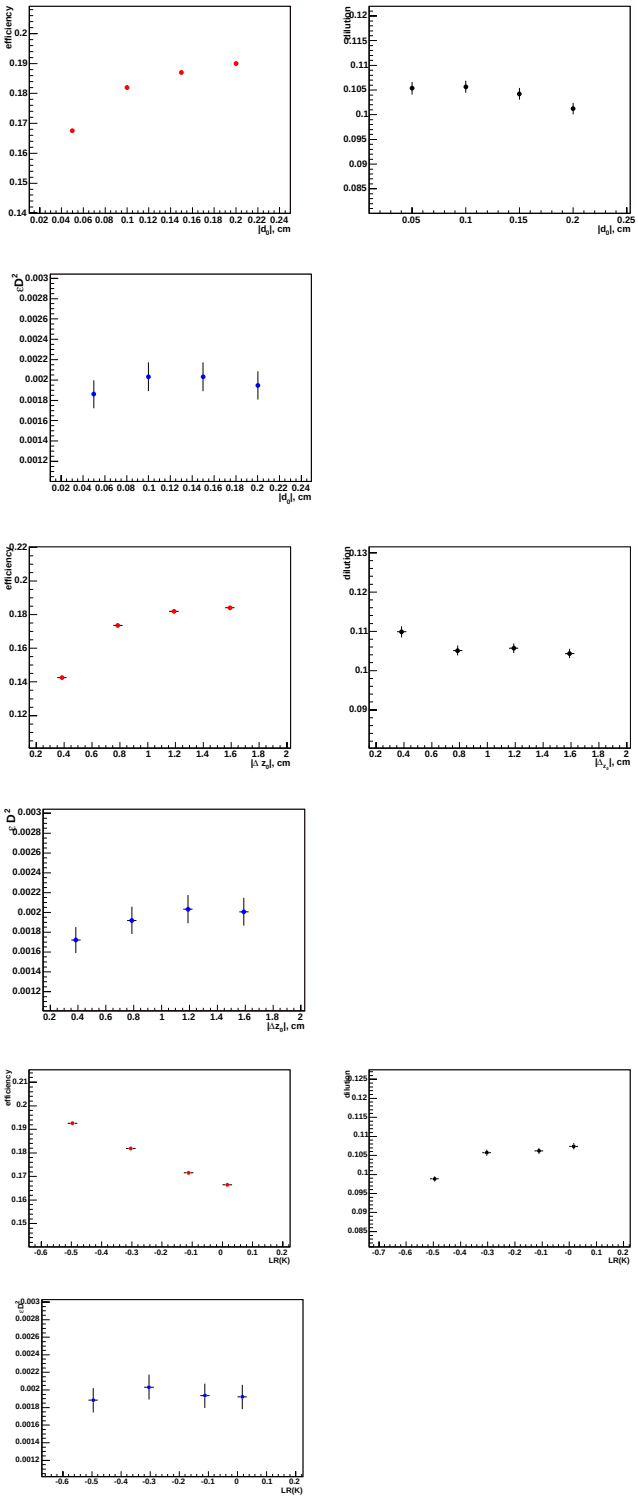

Figure 5.18: Scanning of the $|d_0|$, Δz_0 and LR(K) variables. Three plots each: ε (red), D (black), εD^2 (blue).

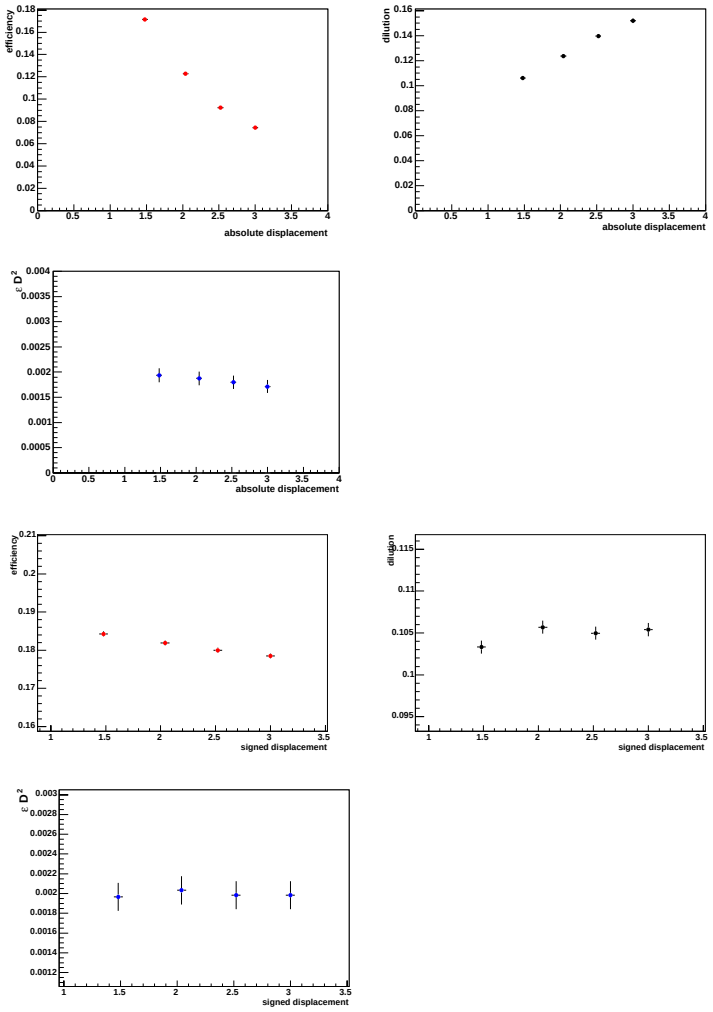

Figure 5.19: Scanning of the $|d_0|/\sigma(d_0)$ and signed $|d_0|/\sigma(d_0)$ variables. Three plots each: ε (red), D (black), εD^2 (blue).

Selection	Signal K eff. (%)	Fragm. π rej. (%)	Fragm. K rej. (%)
$\|d_0\| \leq 0.1$ cm	99.9	0.1	0.1
$\|d_0\|/\sigma(d_0) \geq 1.5$	59.2	79.9	81.7
signed $d_0/\sigma(d_0) \geq 2.0$	75.6	92.3	90.6
$LR(K) \geq -0.3$	97.9	91.2	2.0
$\Delta z_0 \leq 1.2$ cm	97.6	4.0	4.0

Table 5.4: Signal Kaon efficiency of the optimized track selections. The rejection to the main sources of background is also reported. These numbers are obtained from the Monte Carlo simulation of the algorithm, after having applied the preliminary cuts of Tab.5.1. The absolute I.P. cut efficiency and rejection are calculated on events without a `SecVtx` tag (Class 2 and 3) only; the signed I.P. cut quantities are calculated for events with a `SecVtx` tag (Class 1) only.

are then propagated as a systematic uncertainty on the tagger dilution. While correcting data for mixing of the B mesons on the trigger side, I have to reproduce the effect of the mixing in the away side on the Monte Carlo sample I use, since the effect is not simulated. I do this analytically in three steps:

- I check whether a track comes from a neutral B-meson, using the Monte Carlo generator-level information;
- if so, I calculate the probability

$$\mathcal{P}_{Mix} = \frac{1}{2}[1 - cos(\Delta m_d \cdot t)] \tag{5.7}$$

  that a B^0_d meson has mixed as a function of its proper decay time t; the probability is fixed to 0.5 for B^0_s;
- flip the K's charge accordingly.

5.10.1 Final performances

The performances of the OSKT are finally evaluated on the second half of the semileptonic sample and are shown in Tab.5.5 for the three algorithms considered. The combined εD^2 is the sum of the εD^2 for the three classes. The dilution D in the combined case is the average root mean square $D = \sqrt{\frac{\sum_{i=Class\ 1}^{Class\ 3} \varepsilon_i D_i^2}{\sum_{i=Class\ 1}^{Class\ 3} \varepsilon_i}}$. As can be seen, the algorithm which produces the highest figure of merit is the Maximum Likelihood Ratio, although only a 2% relative difference is observed between the best and worst algorithm.

For comparison, I also show the numbers obtained on MC for the best performing algorithm, the highest LR. A discrepancy is observed among data and

Monte Carlo. While the efficiency is in agreement for Class 1 and Class 2, I find a higher efficiency for Class 3 and a consistently higher dilution for all tag types in Monte Carlo. I interpret this as a hint for a lower track multiplicity in Monte Carlo due to the absence of dedicated fragmentation background tuning. Infact, while on one side this is expected to produce a general increase in dilution for simulated events, it also increases the rate of Class 3 events, less pure in specific b content and therefore more sensitive to the amount of fragmentation tracks.

Maximum I.P. significance	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.66 ± 0.03	16.07 ± 1.3	0.0945 ± 0.0097
Class 2	3.48 ± 0.03	10.35 ± 1.5	0.0373 ± 0.0068
Class 3	10.97 ± 0.06	6.05 ± 0.8	0.0402 ± 0.0072
Combined	18.11 ± 0.07	9.7 ± 0.8	$\mathbf{0.172 \pm 0.013}$
Closest ΔR	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.66 ± 0.03	16.2 ± 1.3	0.0960 ± 0.0097
Class 2	3.45 ± 0.03	10.4 ± 1.5	0.0374 ± 0.0068
Class 3	10.97 ± 0.06	6.0 ± 0.8	0.0402 ± 0.0072
Combined	18.11 ± 0.07	9.8 ± 0.8	$\mathbf{0.174 \pm 0.014}$
Highest LR(K)	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.66 ± 0.03	16.3 ± 1.3	0.0973 ± 0.0098
Class 2	3.48 ± 0.03	10.3 ± 1.5	0.0369 ± 0.0068
Class 3	10.97 ± 0.06	6.2 ± 0.9	0.0418 ± 0.0073
Combined	18.11 ± 0.07	9.9 ± 0.9	$\mathbf{0.176 \pm 0.014}$

Table 5.5: Performances of the OSKT on data for the various algorithms under study. Sample used is one half of the $\mu + SVT$ and $e + SVT$ samples mixed together.

MC, Highest LR(K)	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.42 ± 0.09	19.8 ± 4.0	0.134 ± 0.036
Class 2	2.96 ± 0.09	16.7 ± 4.5	0.082 ± 0.029
Class 3	7.92 ± 0.01	12.2 ± 2.8	0.117 ± 0.034
Combined	14.30 ± 0.18	15.3 ± 1.3	$\mathbf{0.333 \pm 0.057}$

Table 5.6: Performances of the OSKT on MC using the highest LR algorithm.

As anticipated, the effective difference among the above algorithms is small. This is expressed by looking at the performances exclusively for cases when one needs to choose one track, that is for events when there is more than one track

available to tag and their responses disagree. Tab.5.7 summarizes this information, obtained by isolating events with > 1 candidate with a null sum of the charges $\sum Q$. Numbers for events when one tagging candidate was found and for events when > 1 candidate was found and the responses agree ($\sum Q \neq 0$) are also provided in the same table. From this one can also argue that no sensible improvement would be expected by implementing any further combination of these three basic algorithms.

1 tag candidate	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	2.85 ± 0.03	17.1 ± 1.4	0.084 ± 0.009
Class 2	2.87 ± 0.03	10.9 ± 1.6	0.034 ± 0.007
Class 3	9.646 ± 0.005	6.7 ± 0.9	0.043 ± 0.007
> 1 tag candidate, $\sum \mathbf{Q} \neq 0$	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	0.41 ± 0.01	24.8 ± 3.8	0.0255 ± 0.0050
Class 2	0.30 ± 0.01	13.1 ± 5.2	0.0051 ± 0.0026
Class 3	0.64 ± 0.01	9.0 ± 3.7	0.0051 ± 0.0027
> 1 tag candidate, $\sum \mathbf{Q} = 0$, LR algo	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	0.38 ± 0.01	1.5 ± 4.0	-
Class 2	0.25 ± 0.01	2.5 ± 5.6	-
Class 3	0.62 ± 0.01	-2.6 ± 3.6	-
> 1 tag candidate, $\sum \mathbf{Q} = 0$, $\mathbf{\Delta R}$ algo	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	0.38 ± 0.01	1.5 ± 4.0	-
Class 2	0.25 ± 0.01	1.8 ± 5.6	-
Class 3	0.62 ± 0.01	-2.2 ± 3.6	-
> 1 tag candidate, $\sum \mathbf{Q} = 0$, Displ algo	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	0.38 ± 0.01	2.0 ± 4.0	-
Class 2	0.25 ± 0.01	2.0 ± 5.6	-
Class 3	0.62 ± 0.01	-2.9 ± 3.6	-

Table 5.7: Performances of the OSKT for various sub samples of events when: only 1 candidate tagging track was found (first table); > 1 candidate tagging track is found, and $\sum Q \neq 0$ (second table); > 1 candidate tagging track is found, and $\sum Q = 0$, evaluated with the three different algorithms considered (bottom 3 tables).

As a cross-check, I also quote the numbers for electron and muon samples separately in Tab.5.10.1. It has to be noticed that this is taken from the full samples and therefore includes also the half of that already used for the optimization.

A 2.5 σ discrepancy is present between the electron and muon sample for the dilution of Class 2, that represents the highest contribution to the final discrepancy in εD^2. Aside of statistical fluctuations, this disagreement may also be

$\mu + \mathbf{SVT}$	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.63 ± 0.03	16.8 ± 1.2	0.102 ± 0.027
Class 2	3.43 ± 0.03	12.3 ± 1.3	0.052 ± 0.007
Class 3	10.94 ± 0.05	6.57 ± 0.5	0.047 ± 0.007
Combined	18.00 ± 0.06	$10.57 \pm 0.9.$	$\mathbf{0.201 \pm 0.029}$
e+SVT	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.76 ± 0.03	15.0 ± 1.2	0.084 ± 0.009
Class 2	3.53 ± 0.03	7.22 ± 1.5	0.018 ± 0.005
Class 3	10.95 ± 0.06	5.98 ± 0.8	0.039 ± 0.007
Combined	16.57 ± 0.06	9.9 ± 1.1	$\mathbf{0.162 \pm 0.013}$

Table 5.8: Performances of the OSKT on the full $\mu + SVT$ and $e + SVT$ data sample.

explained by the topological differences among the two samples due to the different kinematical acceptance of the detectors used to select them. For instance, different performances were found between $\mu + SVT$ and $e + SVT$ for all the other Opposite-side taggers ([41], [45], [46]). Quoting performances for a combination of the two samples, as I do, accounts automatically for statistical effects and topological differences. It may also be noticed that the application of the tagger to physics studies (as time-dependent CP asymmetry studies) is subject to a final calibration on the sample to which it is applied, such that the details of the semileptonic sample are not important to this regard.

Nevertheless, I check on data that no further physics effects are involved, by comparing the distributions of the quantities relevant for the tagger dilution (displacement, LR(K), track and jet multiplicity, $\Delta R(\ell + SVT, trk)$) between the two semileptonic samples. The plots are shown in Fig.5.20. No sensible differences are noticed.

5.10.2 Systematic uncertainties

The only source of systematics in the evaluation of the performances of an OST, performed entirely on data, comes from the trigger side effects on D, the 0.6412 factor. Thus, following previous studies on the other OST, I will quote this last effect as the systematics for the tagger, obtained by error propagation from Eq.4.23:

$$\varepsilon D^2 = 0.176 \pm 0.014\ (stat)\ \% \ ^{+0.001}_{-0.001}\ (syst)\ \% \tag{5.8}$$

I do not quote any systematics on the PID, which is the only tool I use where a parameterization is used. Infact, any systematically wrong assumption on the PID models would just make the tagger performances sub-optimal.

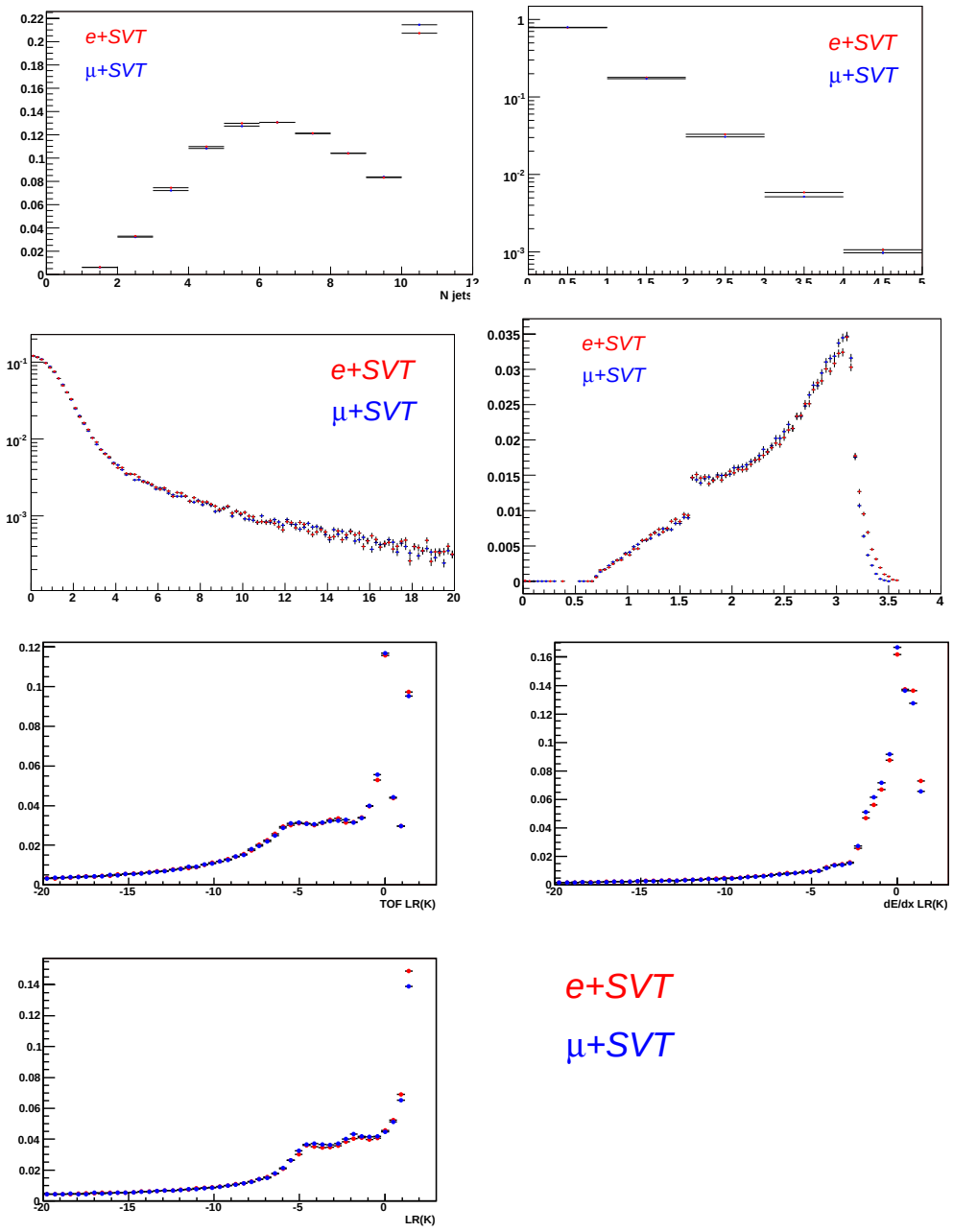

Figure 5.20: Comparison between $e+SVT$ and $\mu+SVT$ samples for the jet multiplicity, the tagging track multiplicity, the displacement, the $\Delta R(\ell + SVT, trk)$ and for the LR(K) distributions for TOF, dE/dx and combined.

5.10.3 Effect of additional vertices

Up to know I assumed that the inclusive sample of tracks considered is originated by one single primary interaction. If this is not the case and a track coming from a $p\bar{p}$ interaction other than the one that originated the away b is selected, a random charge-b flavour correlation enters that tag and the dilution is expected to be reduced. Furthermore, as we saw in Sec.3.2.3, TOF matching efficiency is a function of the number of primary vertices and decreases at higher vertex multiplicities due to a higher track occupancy. Consequently, I evaluate εD^2 in the sub-sample of events where only 1 primary vertex is reconstructed along the beam axis. I expect the Δz_0 cut to prevent a significant amount of tracks from other PV from contributing in the algorithm, anyway; so that only a small difference should be observed between this specific sub-sample and the full sample. The result is summarized in Tab.5.9.

$\mathbf{N_{PV} = 1}$	$\varepsilon(\%)$	$D(\%)$	$\varepsilon D^2(\%)$
Class 1	3.892 ± 0.039	16.9 ± 1.6	0.111 ± 0.013
Class 2	3.930 ± 0.044	9.0 ± 1.8	0.032 ± 0.008
Class 3	10.889 ± 0.071	5.6 ± 1.1	0.034 ± 0.008
Combined	18.71 ± 0.092	9.7 ± 1.3	$\mathbf{0.177 \pm 0.017}$

Table 5.9: Performances of the OSKT on data for the sub-sample of events where only 1 vertex is reconstructed along the beam.

As can be seen, $\leq 1\%$ fractional improvement is found with respect to the number obtained from the whole sample, result 5.8. In particular, the only improvement in the case with 1 vertex comes from the efficiency, as a result of better TOF efficiency.

5.10.4 Effect of trigger side daughter removal

As anticipated in Sec.5.3.1, I also evaluate the impact of the rejection of trigger b daughters distant in ΔR from the trigger b. As described in Sec.5.3.1, this is based on the combination of the invariant mass of the candidate track with the trigger pair and their ΔR. The following numbers correspond to the situation when no rejection at all is performed or, on the orther side, when a selection tighter than the standard one is used, such that *any* track with a $\Delta R(\ell + SVT, trk) \leq 1.6$ is discarded. As expected, in the first case a raise in efficiency corresponds to a higher amount of tracks originated from the trigger b decays contributing to the final dilution with a reverse charge-track correlation. The εD^2 is consequently lower than for the selection adopted in the standard algorithm, result 5.8. On

NO SS removal	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.95 ± 0.03	14.9 ± 1.2	0.0874 ± 0.0094
Class 2	5.46 ± 0.04	5.9 ± 1.2	0.0193 ± 0.0050
Class 3	12.36 ± 0.06	4.7 ± 0.8	0.0468 ± 0.0060
Combined	21.77 ± 0.08	8.4 ± 0.9	$\mathbf{0.1535 \pm 0.0120}$

$\mathbf{\Delta R > 1.6}$ **only**	$\varepsilon(\%)$	D(%)	$\varepsilon D^2(\%)$
Class 1	3.40 ± 0.03	16.0 ± 1.3	0.0867 ± 0.0092
Class 2	3.14 ± 0.03	10.2 ± 1.5	0.0330 ± 0.0064
Class 3	10.03 ± 0.05	6.5 ± 0.9	0.0424 ± 0.0073
Combined	16.57 ± 0.06	9.9 ± 1.1	$\mathbf{0.1621 \pm 0.0130}$

Table 5.10: Performances of the OSKT on data for the highest LR algorithm, without rejecting the SS daughters in the OS (upper table) or rejecting any track with $\Delta R \leq 1.6$, regardless their invariant mass with the $\ell + SVT$ pair.

the other side, a higher dilution is induced by tightening the selection, but at the exenses of the efficiency, since signal Kaons from the away b decays are discarded. This results again in a loss of tagging power.

5.10.5 Dependence of the dilution on the jet's momentum fraction and absolute displacement

Following the considerations in Sec.4.3 and in particular the result of Eq.4.13, I take advantage of the tagger characteristics by binning the dilution as a function of variables on which D shows a strong dependence. I use Monte Carlo simulation to look for variables useful to separate further signal Kaons from background and then calculate D in bins of such variables in order to establish a clear trend. This idea is based on the fact that the b carries a large part of the momentum of the accompanying jet, such that a kinematical correlation can be established between the away b daughters and the away jet. I considered several variables, among which, e.g., the $|\vec{p}|$ of the Kaon and the projection of its momentum along the jet axis, p^L_{rel}.

I find that the variable

$$P_{FRAC} = \frac{p(K)}{p(jet)}$$

is the most effective in separating signal from background. While b daughters are expected to carry a large fraction of the b jet, particles created in the primary interaction are uncorrelated with the b momentum and show a softer spectrum. Fig.5.21 summarizes these considerations separately for events when the jet sample is pure in b content (`SecVtx` tag) and for the other jets. While a separation

is present in the first class of events, it is less evident where an explicit b identification is missing. On the other side, from the optimization plots in Fig.5.19 I notice that the dilution for Class 2 and 3 is strongly dependent on the absolute I.P. significance.

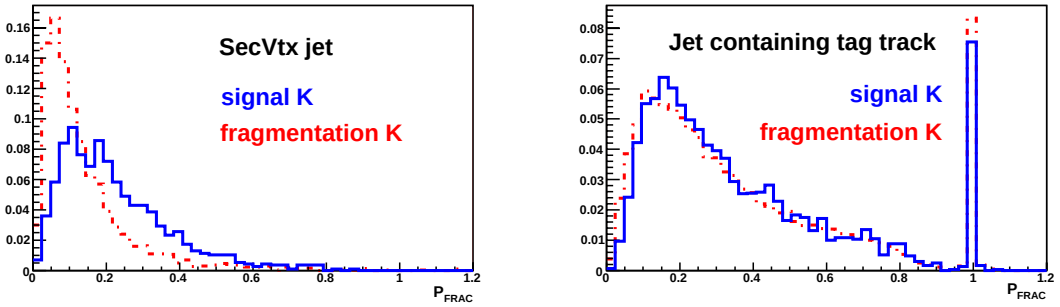

Figure 5.21: MC distributions of some kinematical variables considered, for fragmentation (red) and signal K (blue): upper left plot is the P_{FRAC} distribution with respect to a `SecVtx`-tagged jet. Upper right is the same but for the jet containing the track.

Thus, I bin the dilution as a function of P_{FRAC} for Class 1 tags and of $|d_0/\sigma(d_0)|$ for Class 2 and 3. In the former case, I bin from 0.0 to 0.8 with 0.2-wide bins. In the former cases, bins 0.5-wide are considered, starting from the lower cut value of 1.50. The last bin includes all the tracks with a displacement 4.0 to ∞. The result is shown in Fig.5.22. The point for each bin is at the bin centre.

I evaluate the resulting tagger performances summing over all the bins within each tag class and obtain the numbers and relative improvement in Tab.5.11: The dilution has increased from 9.9 ± 0.9 % to 11.1 ± 0.9 %. Thus, I evaluate the

	Binned εD^2 (%)	Average εD^2 (%)	$(\varepsilon D^2_{binned} - \varepsilon D^2_{average})/\varepsilon D^2_{average}$
Class 1	0.102 ± 0.010	0.097 ± 0.010	+ 5.1 %
Class 2	0.052 ± 0.008	0.037 ± 0.007	+ 40.5 %
Class 3	0.075 ± 0.009	0.040 ± 0.007	+ 87.5 %
Combined	0.229 ± 0.016	0.176 ± 0.014	+ 30.1 %

Table 5.11: Values of εD^2 in the binned case and in the case where the average dilution is considered, for each class separately. The combination is obtained by summing the signle class entries. The relative improvement is also given per each bin.

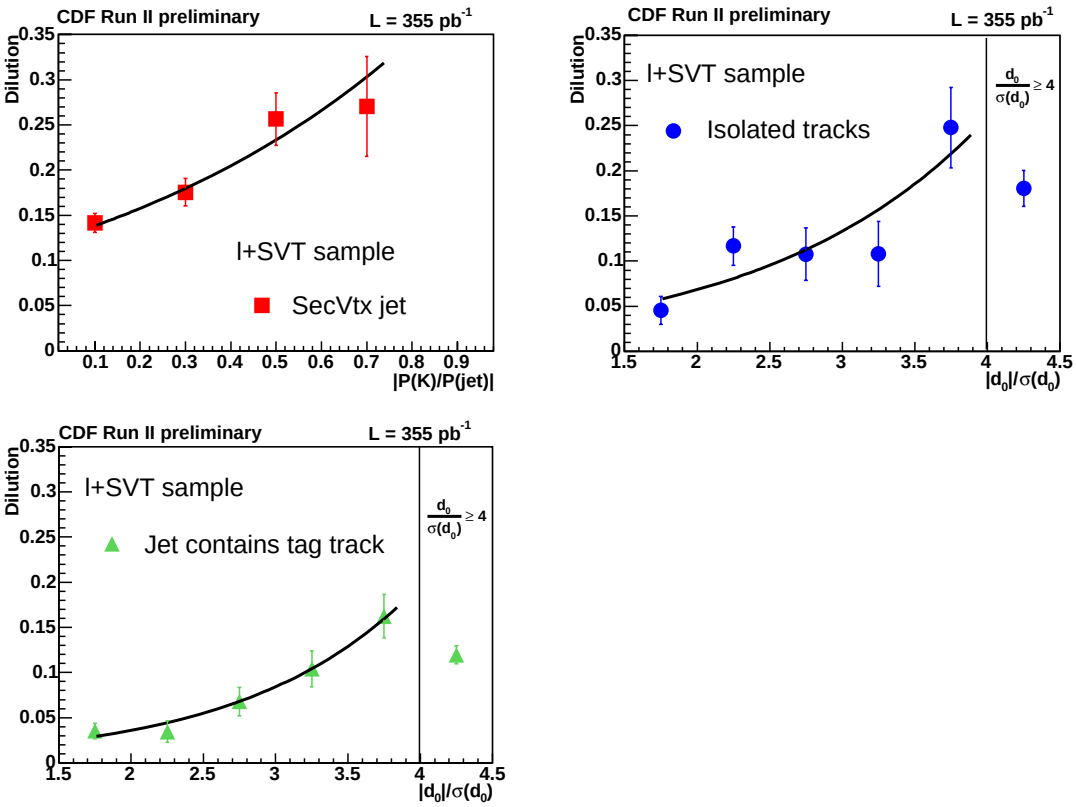

Figure 5.22: Dependence of the dilution on the P_{FRAC}, for class 1 jets. Dependence of the dilution on $|d_0/\sigma(d_0)|$, for class 2 and 3.

final OSKT tagging performances:

$$\varepsilon D^2 = 0.229 \pm 0.016\ (stat)\ \%\ ^{+0.001}_{-0.001}\ (syst)\ \% \tag{5.9}$$

where a net fractional improvement of 30% with respect to considering an average dilution is observed. Largest part of this increase is contributed by lower dilution classes. On one side, this means that the dilution of Class 1 is already at a high level before the binning is performed, thank's to its b purity; on the other side, this reflects the fact that a strong dependence of the dilution on the I.P. significance is present for lower purity classes, more sensitive to contributions from prompt fragmentation background.

Provided this trend, one can assign a predicted dilution value to any event to which the tagger is applied, based on its P_{FRAC} or its absolute I.P. significance. In order to be more refined and increase the overall dilution I use an analytic function rather then the average dilution for each bin. To do that a fit to the dilution versus each quantity is performed that returns the function parameters. For all the classes I find that an exponential function:

$$D(x) = e^{A+Bx}$$

where $x = (P_{FRAC}, |d_0/\sigma(d_0)|)$, reproduces well the dependence in the interval considered. The last bin for Class 2 and Class 3 is not considered in the fit: tagging tracks with a $|d_0/\sigma(d_0)| \geq 4$ are assigned the dilution measured in that bin. Tab.5.12 shows the vaues of the fit parameters A, B for the 3 different classes.

	A	B	D($\frac{\lvert d_0\rvert}{\sigma(d_0)} \geq 4$) (%)
Class 1	-2.11 ± 0.19	1.31 ± 0.45	
Class 2	-4.01 ± 0.45	0.66 ± 0.15	18.1 ± 2.0
Class 3	-5.03 ± 0.45	0.85 ± 0.14	11.9 ± 1.0

Table 5.12: Parameters of the exponential fit to the dilution dependence on P_{FRAC} (Class 1) and absolute I.P. significance (Class 2 and 3). The D for the last bin of class 2 and 3, that averages all the contributions with $\frac{|d_0|}{\sigma(d_0)} \geq 4$ is not included in the fit and is reported separately in the last column for the relative class.

5.11 Sources of Dilution

To understand the features and limitations of the dilution, I make use of MC truth and investigate the origin of tagging tracks for both right sign and wrong

source of flavour	fraction (%) in tagging sample
B^+	20.1
B_d	19.1
B_s	4.3
b baryons	3.0
SS b	1.7
Prompt charm	0.9
Fragmentation	50.2
MC particle unmatched	0.7

Table 5.13: MC fractions of the various sources of flavour contributing to the final D.

sign cases. In Tab.5.13 the composition of the MC tagging sample is shown. As one can expect, B_d and B^+ mesons are the main sources of signal Kaons and are reasonably equal in quantity. Only $\approx 50\%$ of the original sample of B mesons is tagged by the algorithm. Given a trigger B, infact, the away b (and so its decays products) suffers from a reduced acceptance due to its angular distance from the trigger side. Moreover, TOF matching efficiency reduces the efficiency significantly. Tracking acceptance is one of the main reasons for the different performances of the Kaon taggers in the Same-side and in the Opposite-side at CDF. Kaons on the trigger side are within the trigger B fiducial volume by definition and are much more likely to be detected ($\geq \times 2$ efficiency).

Despite a tight selection in terms of PID and displacement from the PV is applied, fragmentation tracks represent the highest source of tags and contribute half of the whole tagging sample. Nevertheless, the initial amount of pions, dominating the generic sample used, is very effectively reduced by the PID cut and provides only $\approx 20\%$ of the total tagging sample ($\approx 70\%$ initially). Tab.5.14 also shows that $\approx 70\%$ of tagging tracks are Kaons ($\approx 20\%$ of the initial unbiased sample).

These informations are combined together and shown in Fig.5.23. Here the absolute rate of Right Sign (RS) and Wrong Sign (WS) tags is shown for the Monte Carlo sample used, by parent type. The difference of the entries for blue and red point, divided by their sum, returns the dilution for each tag source. In particular, one can estimate the dilution in the case of a sample of pure signal Kaons from the $b \to c \to s$ decay chain. This is represented by the third bin in the upper right plot and returns $D(K) = 48.9 \pm 2.1\%$. The dilution in this case is not 100% because of several effects. First of all, cases when the away b mixed return a wrong dilution. Furthermore, $D^+ \to K^+X$ and $D^0 \to K^+X$ decays produce Kaons with an inverted flavour-charge correlation. These occur with a Branching Ratio of $(5.5 \pm 1.6) \times 10^{-2}$ and $(3.4^{+0.6}_{-0.4}) \times 10^{-2}$ respectively [21]. They have to be

particle type	Right Sign fraction (%)	Wrong Sign fraction (%)	from fragmentation (%)
π	11.3	11.7	13.4
K	44.3	23.5	30.4
p	4.1	4.1	5.0

Table 5.14: MC fractions of the main particle types in the tagging sample, for RS and WS. The sum of the first two columns over the three rows, plus the small lepton contribution not shown, is 1. In the third column, the absolute fraction coming from fragmentation is shown. E.g., on 23% of pions, an absolute 13.4% (a little bit more than half of tagging pions) come from fragmentation.

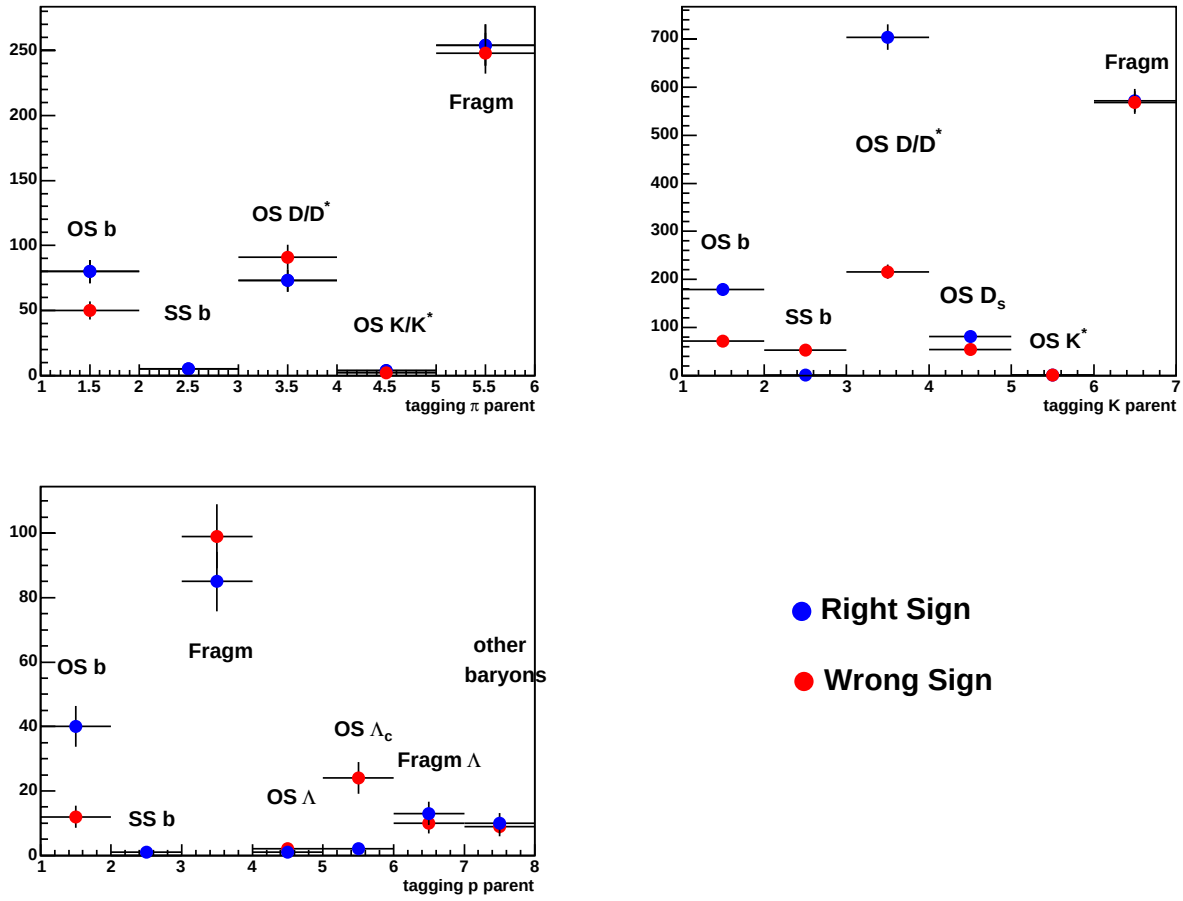

Figure 5.23: MC origin of the RS (blue dot) and WS (red dot) tagging pions, Kaons, protons. The various parents are stated on the relative bin.

compared with the corresponging right-sign $D^+ \to K^- X$ and $D^0 \to K^- X$, with a Branching Ratio of $(27.5 \pm 2.4) \times 10^{-2}$ and $(53 \pm 4) \times 10^{-2}$ respectively [21]. Cabibbo-suppressed and $b \to s$ transitions where the Kaon is originated directly from the b vertex produce a dilution $D(K \text{ from } b) = 52.0 \pm 3.9\%$. The tagging efficiency for signal only, including $b \to c \to s$, $b \to s$ and Cabibbo suppressed decays, is $\varepsilon = 4.67 \pm 0.10\%$. This corresponds to an overall $\varepsilon D^2 = 1.15 \pm 0.12\%$. As recalled several times, ε is affected by the efficiency of matching tracks to TOF pulses. We saw that, on average, this is $\approx 60\%$. Just as an example, a 100% matching efficiency would increase ε(signal K) to $\approx 7.8\%$ and εD^2 for a pure signal sample up to $\approx 1.9\%$.

Kaons from D_s decays present a lower dilution than for the other D mesons. In this case, infact, a wrong correlation is almost as probable as a right one, due to the similar Branching Ratios to charged Kaons: $\mathcal{B}(D_s^+ \to K^- X) = (13^{+14}_{-12}) \times 10^{-2}$, $\mathcal{B}(D_s^+ \to K^+ X) = (20^{+18}_{-14}) \times 10^{-2}$. Moreover, D_s are likely to be originated from a rapidly oscillating B_s, so that their overall dilution is expected to be around zero. I estimate $D(K \text{ from } D_s) = 12.4 \pm 6.1\%$, compatible with 0 within 2σ.

In the case of fragmentation Kaons, a dilution compatible with zero in expected, given their random correlation with the b flavour. Infact, I find $D(\text{fragmentation } K) = -0.7 \pm 2.1\%$. The efficiency is comparable to that of signal Kaons, as can also be inferred from Tab.5.14. In particular, Monte Carlo returns $\varepsilon = 4.73 \pm 0.10\%$.

Finally, a fraction of 1.7% of the whole tagging sample is represented by trigger side decays, even after the removal performed (Sec.5.3.1). As expected, Kaons coming from trigger side b decays of the kind $B^- \to \ell^- D^0 \to \ell^- [\pi^+ K^-]$ have a wrong sign with respect to the away b flavour. In particular, their negative dilution is estimated to be $D(\text{trigger side } K) = -95.8 \pm 2.9\%$, much higher than for Kaons from Opposite-side b decays. This is because mixing and charm decays into Kaons of inverted charge are not simulated for the trigger side b hadron.

Pions from trigger side decays provide a positive dilution, instead. The origin of pions can be seen in the top left plot. Such pions are likely to be originated in the above trigger b decays. As can be seen, the pion has a charge opposite to the Kaon and has a right sign for the OS tags, following the above argument. Other cases when a pion provides a positive dilution are trigger side decays of the kind $\bar{B}^0 \to \ell^- D^+ [\pi^+ \bar{K}^0]$ and the corresponding for $\bar{B_s}^0$. These result in a dilution $D(\text{trigger side } \pi) = 78 \pm 20\%$.

Also pions from Opposite-side b contribute to enhance the tagger dilution. A dilution $D(\text{OS side b } \pi) = 25.2 \pm 5.6\%$ is found, from decays of the kind $B^0 \to \pi^+ D^-$. These behave the same way as Cabibbo-suppressed Kaons.

I also quote a number for pions from charm mesons. In this case, a signal Kaon is most likely to be originated with the pion, in a decay of the kind $b \to D^+ \to K^- \pi^+$ and corresponding modes for D^0 and D_s. In fact, the pion is always Wrong Sign and a negative dilution is found: $D(\text{OS side c } \pi) = -24.3 \pm 5.2\%$.

Regarding protons, displayed in the bottom left plot, it is worth noticing the negative dilution evident in the Λ_c case. I evaluate $D(p \text{ from } \Lambda_c) = -83.7 \pm 7.8\%$. Infact, this baryon produces a $\bar{p}$, with a Branching Ratio $\mathcal{B}(\Lambda_c^+ \to pX) = (50 \pm 16) \times 10^{-2}$ [21]. Thus, the final state proton has a charge opposite than expected for a decay chain started by a Λ_b^0 baryon decay.

In the first bin, protons from away b decays of the kind $B^+ \to p\ \Lambda_c^- X$, and corresponding flavour modes, prove to be enhanced in right sign tags, so that they contribute with a positive dilution $D(\text{OS side b } p) = 43.9 \pm 9.9\%$ to the overall OSKT performances.

Protons produced directly from sea quark interactions in the hadronization process have $D(\text{fragmentation } p) = 3.7 \pm 5.5\%$, compatible with zero as expected.

Finally, Λ baryons and other remaining baryons prove to have null dilution within errors ($D(p \text{ from } \Lambda) = 10.5 \pm 16.1\%$ and $D(p \text{ from other baryons}) = -24 \pm 16\%$ respectively), since thay are produced mainly from fragmentation.

5.11.1 Dilution asymmetry

An asymmetry in how positive and negative charged tracks are reconstructed is present in the CDF detector due to the COT geometry. Positive tracks produce a larger number of hits in the COT than negative tracks on average, due to their opposite curvature. The effect is described in Sec.3.2.2.

As seen, the OSKT is essentially a track-based algorithm, in which the charge of the Kaon is the primary information for tagging. Thus, this effect from tracking may introduce a systematic bias in the tagger such that more positive than negative tracks may be present in the tagging sample. The resulting dilution, calculated from the Kaon charge with respect to the trigger ℓ charge, is expected to be asymmetric as well, in such conditions.

Moreover, more K^- than K^+ are absorbed in nuclear interaction with the detector material, such that the former are less likely to be reconstructed and to travel 140 cm to reach TOF. The two above effects separately and their combination are shown in Fig.5.24 as a function of the Kaon momentum. The overall effect is as big as a fractional 3% asymmetry where the tagging track momentum spectrum peaks ($\approx 1\ GeV/c$), such that in the case of the OSKT is not expected to play a relevant role.

The proton background in the tagging sample may also produce a dilution asymmetry: in fact, more p than $\bar{p}$ ar present, due to the absorption of the latter in nuclear interactions and the predominance of p in the beam halo.

An asymmetric tagger is a systematic effect that one has to study and measure in order to apply it to a time-dependent asymmetry measurement. For these reason, I investigate any systematic asymmetry in the tagger dilution by dividing the full data sample in events triggered by a ℓ^+ and events triggered by a ℓ^-. Since all these phenomena are expected to enhance the positively-charged population,

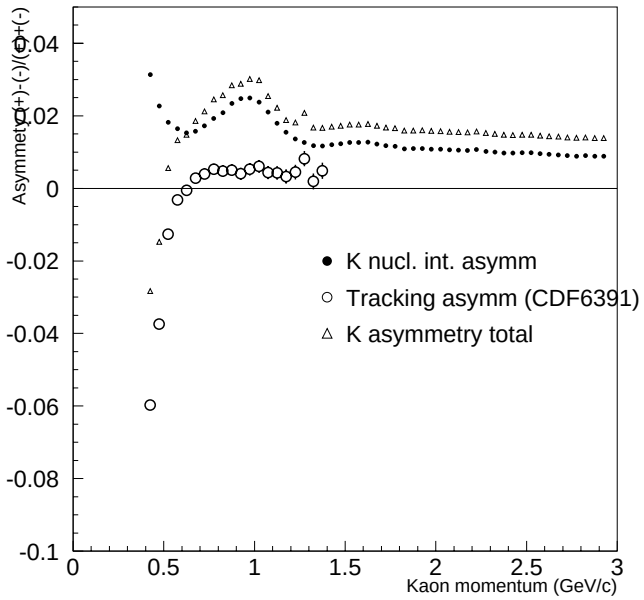

Figure 5.24: Fractional asymmetry ($\frac{N^+ - N^-}{N^+ + N^-}$) from tracking and for cross section of the process of Kaon capture by the detector, as a function of the Kaon momentum. The combined curve is overlaid. The tracking asymmetry curve is taken from [27], the cross section from [21].

a higher dilution is to be observed in the ℓ^- sample than in the ℓ^+[1].

On data, I measure an absolute dilution asymmetry $A \equiv D(\ell^-) - D(\ell^+) = 5.8 \pm 2.8$ %. Tab.5.15 shows the asymmetry as divided by class.

	$D(\ell^-) - D(\ell^+)$ (%)
Class 1	6.1 ± 2.6
Class 2	4.4 ± 3.0
Class 3	6.1 ± 2.6
Combined	5.8 ± 2.8

Table 5.15: Difference in dilution asymmetry on $\ell + SVT$ data for the different tagger classes.

The same effect is visible on Monte Carlo, where an asymmetry $A_{MC} \approx 4\%$ is observed. Since Monte Carlo simulation is not able to reproduce the COT tracking asymmetry, I use data to evaluate the contribution of the combined tracking and cross-section asymmetry on Kaon tracks:

- I evaluate the combined tracking and cross section asymmetry for Kaons in bins of p_t, starting from the curves in Fig.5.24;
- I weight the asymmetry for the p_t distribution of tagging tracks;
- for each p_t bin I measure the dilution D_i;
- I subtract D_i of the expected asymmetry from the above curves in bins of momentum and evaluate the corrected dilutions D'_i.

Applying this procedure I find that the dilution asymmetry is reduced of an absolute 1.5% with respect to the uncorrected case, compatible with the numerical difference of the asymmetries for data and Monte Carlo.

Also the effect of beam protons is studied using data directly. The constribution of such particles is found to be negligible since the asymmetry does not change with respect to the above result if the beam proton removal described in Sec.5.3 is not applied.

Since an asymmetry of the same order as in data is also visible on Monte Carlo, I use this to investigate further origins of the effect. In Fig.5.25 I plot the dilution for π, K, protons separately, according to their parent particle. I expect to see differences generating the asymmetry when plotting ℓ^+- and ℓ^--triggered samples separately.

A relevant contribution to a charge-dependent dilution ($\approx 40\%$ absolute asymmetry in this mode) comes from protons originated in decays of prompt baryons,

[1]Right Sign tags satisfy the condition $Q_K \cdot Q_\ell = -1$

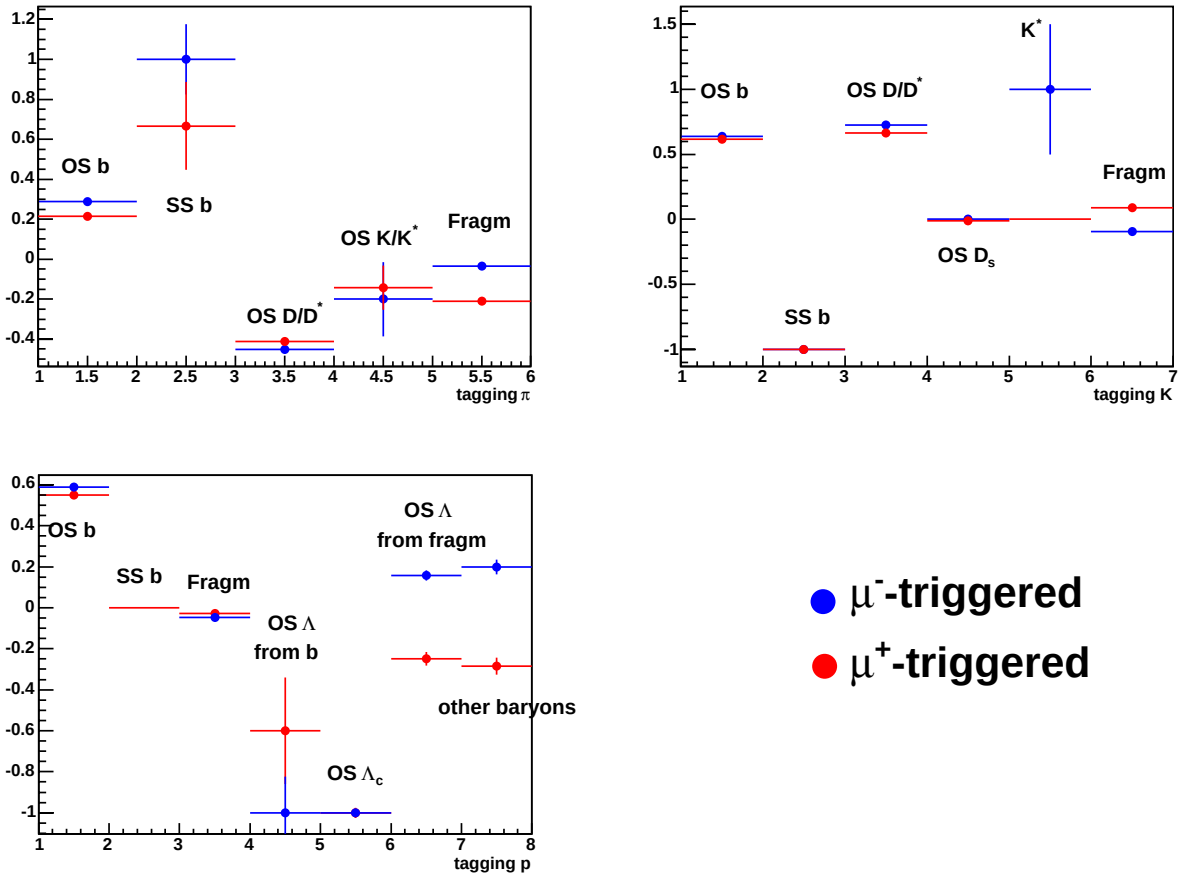

Figure 5.25: Dilution in MC for the various tagging particle types. Bins account for the origin of the track. Red dots are for positively triggered events, blue for negatively triggered ones. Here no OS mix is considered in order not to introduce any bias to the original simulation.

especially Λ's. In these cases, infact, while protons and anti-protons are produced with the same probability, according to the corresponding Branching Ratio, protons have a higher probability to be absorbed. The overall asymmetry results much smaller than $\approx 40\%$ due to the small fraction of tagging protons (Tab.5.14).

As expected, a small asymmetry is found for pions and Kaons coming from fragmentation. For both, this is originated by the different probability with which negative and positive particles interact with the detector nuclei.

In summary I find that the asymmetry is generated mainly by the different probability with which charged particles are absorbed by the detector material. From Monte Carlo simulation I find that this effect is relevant mostly for protons, while Kaons contribute for a smaller fraction ($\approx 1/4$ of the total effect), as well. An instrumental bias from COT reconstruction efficiency is convoluted with the cross section effect, as well. The beam halo protons are found to be negligible to this regard.

This physics effect can not be suppressed. While this is not relevant for a B_s mixing analysis, because no absolute asymmetry measurement is involved, this systematic effect is relevant when performing a CP measurement, like time-dependent CP asymmetry. In this case, infact, the asymmetry in the dilution introduces a bias in the final N^+/N^- proportion not due to the process in study. This results in a systematic uncertainty on the measurement from the tagger application.

5.12 Critical review of the results and main points for Kaon taggers at future experiments

In this chapter I have discussed how it is possible to develop an Opposite-side Kaon tagger at a hadron collider and I have demonstrated that dilutions of the same order of magnitudes as for the other away side flavour taggers are obtained. In particular, I have shown that events where the information of the away b are available by means of the explict identification of the b decays vertex have a dilution comparable with that of the best Jet Charge tags, which provide the largest contribution to the OST εD^2 at CDF.

Use of Particle Identification is crucial in this analysis. Since the generic sample I consider for tagging in the Opposite-side is dominated by pions ($70 - 80\%$ depending on the p_t range considered), a Kaon tagger would not be concievable without an effective capability to distinguish K mesons from such a huge background source. As seen, if I did not apply any PID at all, the dilution would be completely driven by fragmentation pions, and thus would be essentially null. As I have remarked several times throughout this work, the combination of TOF and dE/dx information allows to separate Kaons from such pions and from the remaing charged components in a wide momentum range. In particular TOF is

able to distinguish Kaons from pions with $\geq 1\ \sigma$ at momenta where $\approx 60\%$ of the signal Kaons are found and the final PID selection has $\geq 90\%$ pion rejection. This changes the contributions in the tagging sample, where eventually 70% of tags are due to Kaons.

Nevertheless, almost half of this fraction is still due to fragmentation. Kaons from the PV, created in the hadronization process that also originated the b hadron of interest or from the underlying event in the $p\bar{p}$ collisions, are ≈ 3 times the fraction of the signal Kaons. They are accepted with the same efficiency as the latter by the PID selection. Again, their flavour tag is totally randomic. The second important handle is thus a precise determination of the distance between the track and the PV. In this case, I have made use of a refined tracking system and of the CDF Silicon detector. In particular, it has to be remarked the contribution from the Layer 00, that allows a significant reduction of the uncertainty on the impact parameter, enhancing this way the separation between displaced signal tracks and prompt background. As seen, the efficiency of fragmentation Kaons using with the I.P. significance cut is only $\approx 20\%$, while its signal efficiency is $\approx 70\%$. Finally, it has also to be recalled that our Monte Carlo, without a more thorough tuning of QCD processes, has been found not able to reproduce the total Kaon multiplicity from data ($\approx 10\%$ less). It is reasonable to address this discrepancy to fragmentation, which accounts for (part of) the discrepancy between Monte Carlo and data predictions I observe.

I can thus conclude that the main "bottle-neck" for the dilution is represented by the separation of Primary and Secondary vertex. In particular, expectations about the effectiveness of the OSKT for the Run II were that the presence of Layer 00 and TOF would provide this tagger with an efficiency $\varepsilon = 11.2 \pm 0.3\%$, a dilution $D = 46.1 \pm 2.2\%$ and consequently a figure-of-merit $\varepsilon D^2 = 2.4 \pm 0.2\%$ [56]. This would represent a net improvement with respect to Run I, where neither of the two was present. This estimate had been produced using a Pythia Monte Carlo with the generation of $q\bar{q} \rightarrow b\bar{b}$ flavour creation mechanisms only (i.e. MSEL = 5). The efficiency for attaching a Layer 00 hit to a track was assumed of 100%, while I found that on the sample I used only 40% of tracks had a L00 hit associated. Estimates without Layer 00 were also produced at the time, with $\varepsilon D^2(\mathrm{N0\,L00}) = 1.9 \pm 0.1\%$.

Future experiments looking at higher $\sqrt{s}$ collisions (e.g. LHCb) can take advantage of the enhanced Lorentz boost of the b hadron in the Laboratory frame. In this case, infact, despite the impact parameter is invariant because in the transverse plane, the z information will probably be useful to separate primary and decay vertices. In my case the information along z proved not to be useful to discriminate displaced signal from prompt background for two main reasons: at Tevatron the b pairs are produced essentially central in pseudorapidity and do not travel much along z; the alignement of the tracking system (the Silicon detector, in particular) is not as well determined as in the trasverse plane and the uncertainty in the track z_0 is about the double as that on d_0. At LHCb the

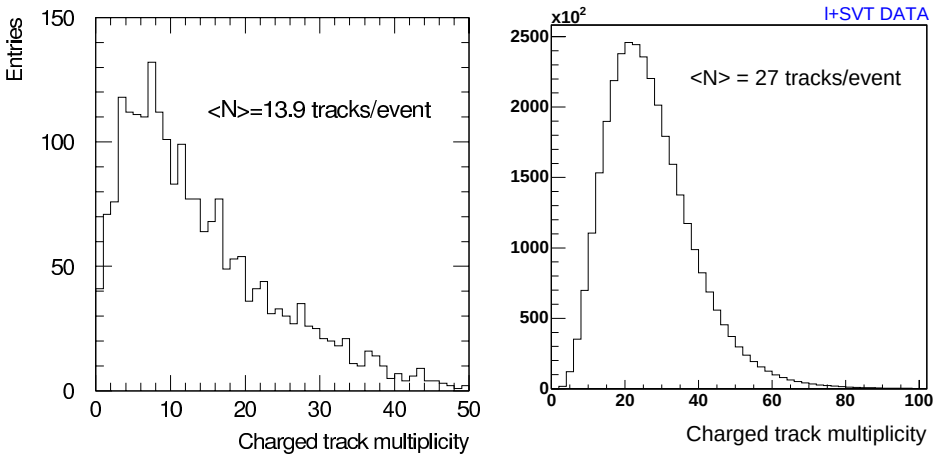

Figure 5.26: Charged track multiplicity: for the Pythia MSEL = 5 Monte Carlo used for projections on OSKT before CDF Run II (left, from Ref.[51]) and on the generic data sample used for OSKT in this study (right). The average number of charged tracks per event is 13.9 and 27.0 respectively. Note the different range of the x axis.

separation could be helped by such an information, provided that the detector is performing well enough along z, as well. This would also help reducing pollution in the tagging sample from the overlap with other primary interactions at higher luminosities ($\geq 2-3\cdot 10^{32} cm^{-2}s^{-1}$). LHCb preliminary Monte Carlo projections at about the same luminosity as Tevatron ($2\cdot 10^{32}\ cm^{-2}s^{-1}$) claim a relative reduction of $\approx 8\%$ in εD^2 ($\varepsilon D^2 = 2.31 \pm 0.21\%$) with respect to the performances for a sample with only 1 P.V ($\varepsilon D^2 = 2.48 \pm 0.18\%$). This is mainly due to a decrease in dilution of -20% between cases with single collision and cases with multiple collisions ($D = 30 \pm 1\%$ and $24.2 \pm 2.2\%$ respectively) [57]. In my case, I report a loss of less than 0.5% in εD^2, thanks to the Δz_0 cut.

Although PID information has proved very effective in rejecting pions and other species, this results in a significant efficiency drop. Infact, the request to have TOF information associated with a track subtracts relative $\approx 40\%$ of the initial acceptance to tracks and enters the final tagger efficiency *linearly*. The above numbers were produced with the assumption that TOF matching efficiency would be 100 %, and anyway not less than $\approx 90\%$. As seen, we have observed an average matching efficiency of 60%. The efficiency on Monte Carlo simulation results higher than real data due to the underestimation of the track multiplicity in the generic sample and, consequently, of the occupancy per TOF bar. The charged track multiplicity in the Monte Carlo used for initial projections and the one I measure on data are shown in Fig.5.26.

The expected occupancy is shown in Fig.5.27, as a function of the number of

additional PV in the collision. Also the measured hit occupancy per bar, defined as the number of events with a hit on the given bar divided by the total number of events, is shown as a function of the TOF bar number for a sample I selected using the J/ψ trigger and at a luminosity corresponding to an average number of PV $< N_{PV} >= 1$. From the comparison, I see that the real occupancy is ≥ 3 times the expected one.

In the case of LHCb, a two-stage Cherenkov detector is used to perform PID in a 1-100 GeV/c track momentum range. The Cherenkov pattern recognition is concieved as to associate a particle hypothesis to *each* of the reconstructed tracks. This is done extrapolating each track onto the Cherenkov. Then, a maximization of a likelihood over all the rings contained within the region of interest on the Cherenkov is performed. The algorithm considers all such rings and eventually assigns a mass. Criteria to perform the association are the ring radius and the ring center, as determined from the expected θ_C Cherenkov angle for each track and a given charged particle mass hypotesis h_j ($j = e, \mu, \pi, K, p$). This proved to be very performing on a realistic Monte Carlo simulation [58], with an efficiency to detect Kaons of 97% with respect to generation and a probability 96% to identify them correctly, averaged over the momentum range. This corresponds to a dilution of $D \approx 35\%$. This means that LHCb expects a very good rejection of fragmentation, as reproduced on Monte Carlo. Infact, such a predicted dilution is already close to the "natural" dilution $D \approx 50\%$ that we observed in simulation for a pure sample of signal Kaons. In case of an occupancy much higher than expected, a degradation of this performances would result due to an incorrect ring recognition.

From all these considerations, a correct estimation of the track multiplicities and simulation of QCD processes from Tevatron will represent the most important point in order to obtain a trustful prediction for the Kaon tagger dilution at future hadron collider experiments, like LHCb.

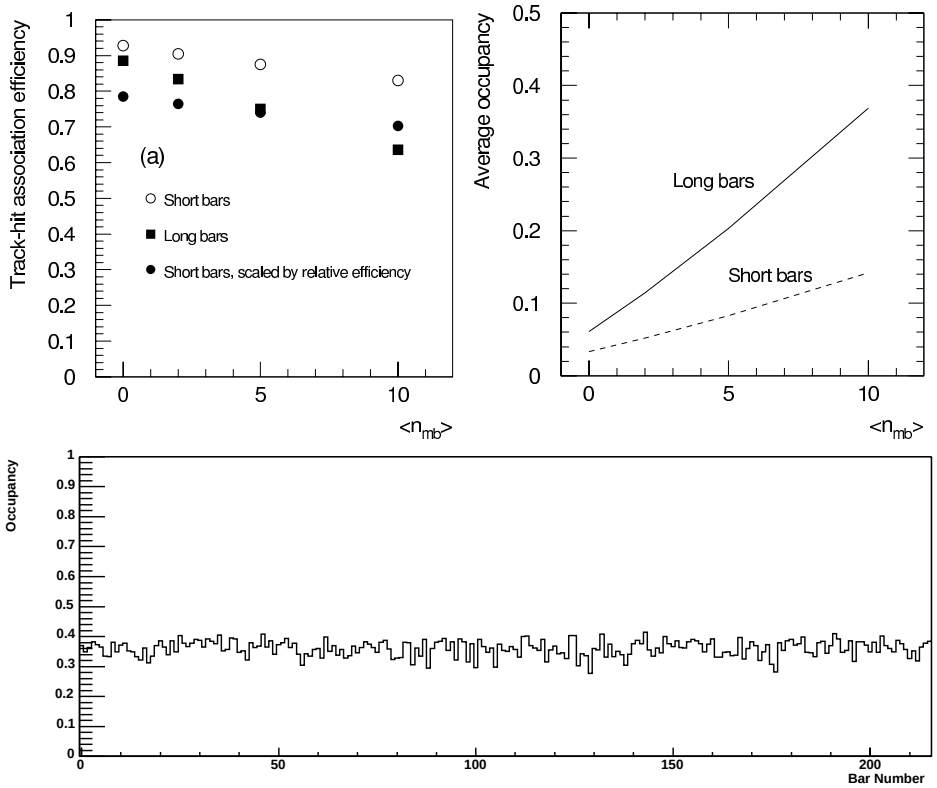

Figure 5.27: Top: Left - MC distribution of the TOF track matching efficiency as a function of the average number of *additional* primary interactions with respect to $N_{PV} = 1$. The distribution is shown for 3 different lengths of TOF bars: the actual length corresponds to the full squares. Right - MC occupancy as a function of the average number of *additional* primary interactions with respect to $N_{PV} = 1$ for long (solid) and short (dashed line) bars. From Ref.[51]. Bottom: Hit occupancy per each of the TOF bars, as obtained from a J/ψ trigger data sample, with $< N_{PV} > \approx 1$

Chapter 6

Combination of the CDF Opposite Side Taggers for B_s mixing analyses

In the past chapters I described the various Opposite-side taggers in use at CDF. In Ch.5 I have discussed a new Opposite-side algorithm to be added for CDF flavour tagging. The informations provided by the various OS b taggers need to be combined in one single response on the away b flavour. Furthermore, we saw that the dilution of each tagger, as obtained with the established algorithm on the $\ell + SVT$ sample, needs to be re-evaluated on the exclusive physics sample to which it is applied for a measurement. This is the case for the hadronic and semileptonic B_s decays. This results in the evaluation of a Scale Factor that accounts for kinematical differences among the given sample and the $\ell + SVT$ events on which taggers were calibrated.

In the present chapter I describe the way the Opposite-side taggers are combined in CDF. A Neural Network (NN) is used to take into account all the responses and the correlations between the various taggers for each event. This incorporates all the four CDF Opposite-side taggers and is used for the Δm_s measurement. A scale factor for each of the modes CDF used for mixing analyses is also applied to the NN combination. A 20% improvement in the Opposite-side tagging effectiveness is observed with respect to using the taggers hierarchically, with a pre-arranged dilution-based ranking.

6.1 Combination using a Neural Network

6.1.1 The OST Neural Network

Neural Networks (NN) are described in general in Appendix.A. In the combination of different flavour tags a NN is useful thanks to its ability to account

for correlations among several quantities involved in the final tagging decision. With this process we enhance the performances of the combined NN tagger with respect to choosing one tagger at the time, for instance the one with the highest predicted dilution for that event. At CDF the information of the muon, electron, jet charge and kaon taggers are merged together in a Neural Network (the `NeuroBayes` package, [61]) and a single response is returned.

The network is based on the output of the Opposite-side taggers already in place at CDF. In particular, we consider the following quantities:

- (1) b flavour
- muon tagger

  (2) type
  (3) decision $\times$ dilution
  (4) decision $\times$ likelihood
  (5) decision $\times p_T^{rel}$
- electron tagger

  (6) type
  (7) decision $\times$ dilution
  (8) decision $\times$ likelihood
  (9) decision $\times$ p_T^{rel}
- jet charge tagger

  (10) type
  (11) decision $\times$ dilution, Class 1
  (12) decision $\times$ dilution, Class 2
  (13) decision $\times$ dilution, Class 3
- opposite side Kaon tagger

  (14) type
  (15) decision $\times$ dilution, Class 1
  (16) decision $\times$ dilution, Class 2
  (17) decision $\times$ dilution, Class 3
  (18) decision $\times$ LR(K)
  (19) decision $\times$ p_{FRAC}
  (20) decision $\times$ $|d_0/\sigma(d_0)|$

As said, in this configuration the NN acts as a combination of several already established algorithms, whose physics effects have been studied on a standalone base. The correlation matrix of the various inputs is shown in Fig.6.1. Obviously, some of the variables belonging to one single tagger are strongly inter-correlated (i.e. tag-type with others). In particular, since we are most interested in the OSKT, we can notice that Class 1 dilution is strongly dependent on the p_{FRAC}, while Class 2 and 3 show a clear link to the I.P. significance. This is not surprising, since these are the variables on which the dilution is binned. The Soft Lepton Taggers also show that the lepton likelihood and p_T^{rel} are not completely independent of each other, as expected. It can be also seen that the jet charge quantities are related to the other taggers variables, as one can expect given the usage of $b-$jet quantities made in the other algorithms. For instance, the dilution of OSKT Class 1 and NNJQT Class 1 are correlated; OSKT Class 3 shows a relationship with NNJQT Class 2, where in both cases we look for at least one displaced track within a clustered jet. As a confirmation, both also prove to be dependent on the OSKT displacement variable, $|d_0/\sigma(d_0)|$.

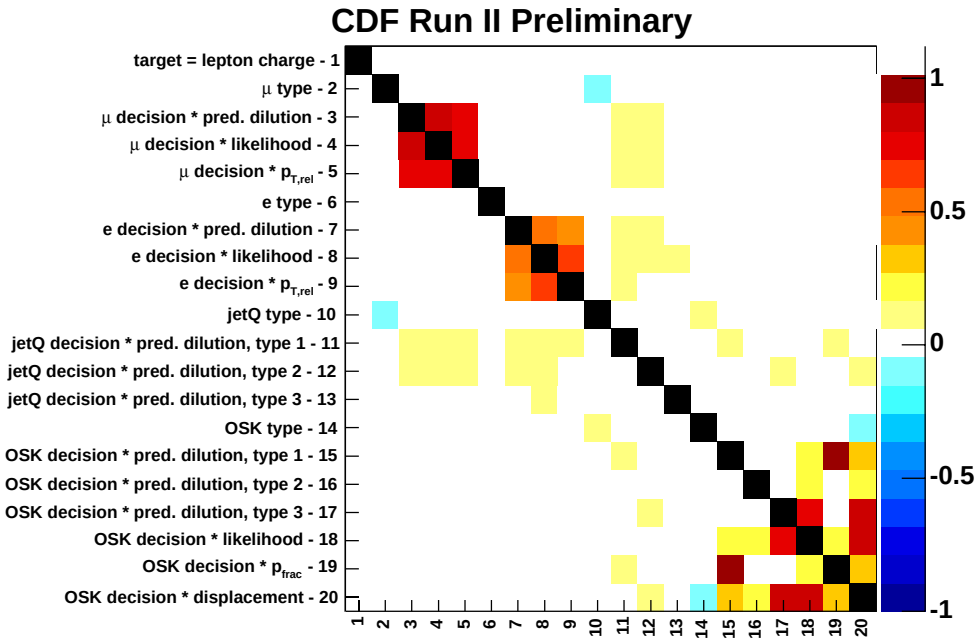

Figure 6.1: Correlations among the various network input variables with each other and with the training target (= b flavour). The variables corresponding to the indices are indicated for each bin.

The training patterns are obtained from the $\ell + SVT$ data sample, with the usual background subtraction via the signed $d_0(SVT)$. Given the complexity of

the physics environment and the high multiplicities at Tevatron, CDF chooses to perform the training directly on data. Also, calibrating the Network on the b flavour from trigger ℓ charge is consistent with was has been done for the basic taggers. Again, we account for B mixing and sequential decays in the trigger side on data by correcting the dilution found for the 0.6412 factor (see Sec.4.3.2). An integrated luminosity of $\int \mathcal{L} = 1\ fb^{-1}$ for both the $\mu+SVT$ and $e+SVT$ has been used. A first half of the total sample has been used for training, the second half for testing. The learning curve for the training stage is shown in Fig.6.2, where it can be seen that the minimum in the difference Δ between O and $\overline{O}$ (Eq.A.5) was not reached, meaning that the NN is not overtrained. This is confirmed at the testing stage, as well.

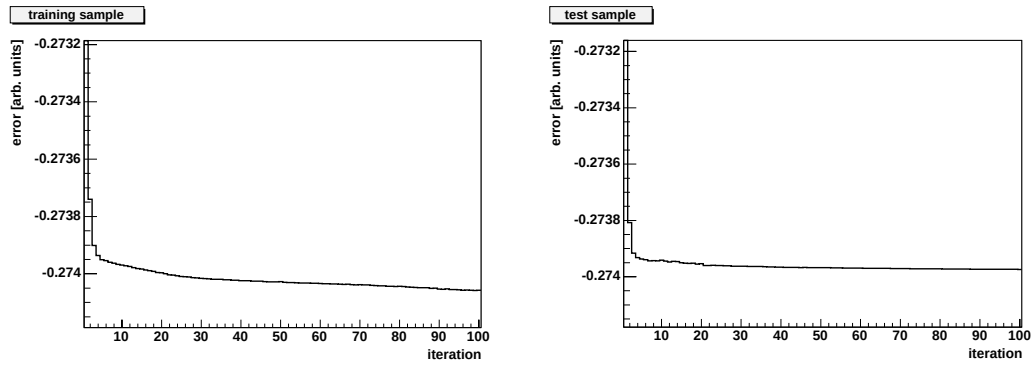

Figure 6.2: Learning curves for the training stage (left) and after testing (right): on the x axis the ordinal number of the step of minimization of the function Δ is displayed. On the y axis the resulting value for Δ is shown.

In Fig.6.3 the NN output for the combined tagger is displayed for samples triggered by positively and negatively charged leptons. The shift of the two responses returns the effectiveness of the tagger and also proves that the NN has not learnt the specific features of any of the two samples.

The output here ranges within $[-1, +1]$. -1 corresponds to 100% negative tagged events (or, as in the plot, 0% probability that the event is positive lepton triggered), $+1$ to 100% positive tagged events. As it may be seen from the right plot in Fig.6.3, a clear relation stands between the tag purity rate:

$$\mathcal{P} = \frac{D+1}{2}$$

and the NN output, as expected in the case of a well-trained network. Thanks to this, a unified tagger is established that:

- takes a decision on the b-flavour following the sign of the output O;
- associates to each event a dilution $D_{PRED} = |O|/0.6412$.

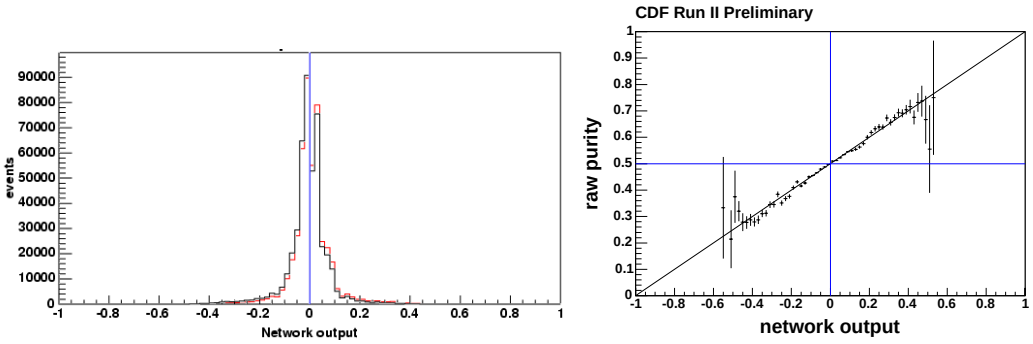

Figure 6.3: NN output O for events with a $Q = +1$ lepton trigger (red) and with a $Q = -1$ lepton trigger (black). Right plot shows the relation between the *measured* purity vs the NN output O.

6.1.2 Establishing the improvement with the NN

We want to measure the improvement of the tagging performances in the OS produced by the NN combination, using the semileptonic sample. This corresponds first to validating the final output of the combined tagger on the $\ell + SVT$ sample, that is to verify that the NN predicted dilution is right. Then, we measure the combined dilution on samples presenting the same characteristics as the B_s exclusive samples used for the mixing analysis. The preliminary check of the calibration is performed on a mixed $\mu + SVT$ and $e + SVT$ sample corresponding to $\approx 765\ pb^{-1}$. Enough statistics is available to divide the predicted dilution in 10 bins and compare it with the measured dilution for each interval:

$$\begin{aligned} D_{PRED,i} &= \frac{1}{N_i}\sum_{k\in i} D_{pred,k} \\ D_{MEAS,i} &= \frac{1}{0.6412} \times \frac{(N_{RS,i} - N_{WS,i})_{\delta>0} - (N_{RS,i} - N_{WS,i})_{\delta<0}}{(N_{RS,i} + N_{WS,i})_{\delta>0} - (N_{RS,i} + N_{WS,i})_{\delta<0}} \end{aligned} \quad (6.1)$$

where i runs on the number of bins, k on the number of *tagged* events with a predicted dilution in the range of the i-th bin. The scale factor is determined with a linear fit to the distribution of $D_{MEAS,i}$ as a function of $D_{PRED,i}$. This is done for both the hierarchical combination and the NN combination. In order to assess the improvement with the introduction of the OSKT and the use of NN, the hierarchical combination does not include the OSKT in the exclusive cascade. The fit outcome is displayed in Fig.6.4 and 6.5.

The numerical results for the scale factors are:

$$\begin{aligned} \mathcal{S}_{hierarchy} &= 0.89 \pm 0.01 \\ \mathcal{S}_{NN\ comb} &= 0.99 \pm 0.01 \end{aligned}$$

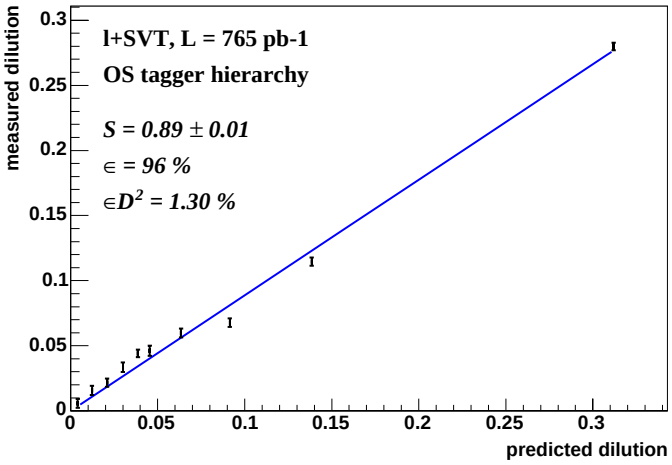

Figure 6.4: Measured dilution vs predicted dilution and corresponding linear fit for the SF overlaid in the case of the hierarchical combination of the OST.

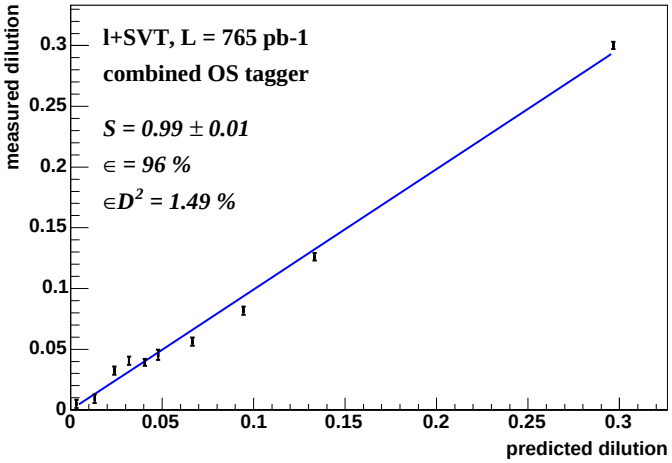

Figure 6.5: Measured dilution vs predicted dilution and corresponding linear fit for the SF overlaid in the case of the NN tagger.

Once one corrects the expected dilution for these numbers, the final performances on $\ell + SVT$ compare like this:

$$\begin{aligned} \varepsilon D^2(\text{hierarchy}) &= 1.30\% \\ \varepsilon D^2(\text{NN tagger}) &= 1.49\% \end{aligned}$$

and the relative improvement provided by the NN on the $\ell + SVT$ is as large as

$$\frac{\Delta \varepsilon D^2}{\varepsilon D^2(\text{hierarchy})} = 15\% \tag{6.2}$$

This is a preliminary estimate of how the OST effectiveness increases with:

- the addition of the OSKT;
- the NN combination.

In the following we discuss this improvement.

6.1.3 Understanding the improvement

As said, the improvement in performances produced by the NN is due to correlations occurring among the various taggers. With respect to the case when a hierarchical combination is used, that looks for the highest dilution tags (Soft Lepton Tags) first, and then the Jet Charge and Kaon tags, we have to consider that:

- the LSMT has its εD^2 improved of a fractional $\approx 20\%$;
- the LSET improves of a relative $\approx 40\%$.

Before using the NN within the mixing framework, a thorough investigation has been performed in order to ensure the improvement is completely understood. The plot in Fig.6.6 shows the different distribution of event dilution between the two methods. From there, we can see that the difference is driven especially by higher dilution events.

For the lepton tags, an improvement is found thanks to double lepton tags and by the cross-talk of each lepton tagger with the Kaon and jet tags. For the first case, the double-tag plot in Fig.6.7 proves that the Right-Wrong quadrants have consistently a $\approx 10\%$ more events than the Right-Right or Wrong-Wrong tags. This is expected, since these are decays of the kind $b \to \nu_\ell X \ell D[\ell' \nu'_\ell Y]$, where ℓ and ℓ' have an opposite charge due to the b and c vertex respectively. The hypothesis that these are real leptons and not random hadrons faking a lepton is demonstrated by the plots in Fig.6.8. Here I show the lepton likelihoods $\mathcal{L}(\mu)$ vs $\mathcal{L}(e)$ for events where the muon tag is right sign and the electron tag is wrong sign are shown. A $\mathcal{L}(\ell) > 0.05$ cut is applied. With thes, I prove that, even

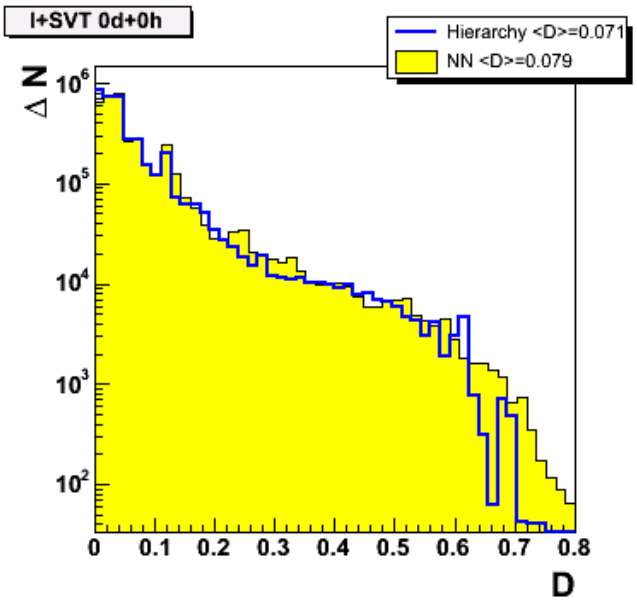

Figure 6.6: Distribution of the predicted dilution D for the NN combination (yellow) and the hierarchical combination (blue line)

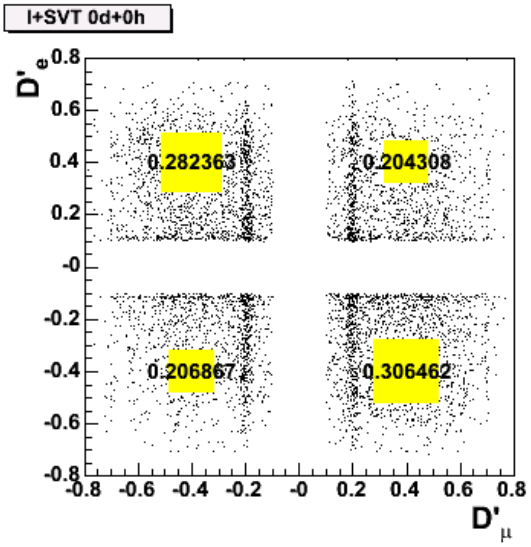

Figure 6.7: $D' = decision \cdot D$ for the LSET vs LSMT, for the four possible Right/Wrong sign combinations. The area of each yellow square is proportional to the fraction of events in a given quadrant.

when giving a wrong answer, the electron tagging sample is essentially made up of real electrons. In the same Fig.6.8 we also show the corresponding $p_T^{rel}(\mu)$ vs $p_T^{rel}(e)$ plot. Since no particular trend is seen and the two leptons have a similar p_T^{rel}, they are likely to come from the same decay chain. In conclusion, one sees that the lepton taggers are correlated with each other and their performances are likely to be enhanced when a double-tag is found. Nevertheless, it has to be considered that the efficiency for a double lepton tag is $\approx 10\%$ *relative* to the muon tag efficiency ($\approx 7\%$). The relative improvement $\Delta\varepsilon D^2/\varepsilon D^2$ from double lepton tags is measured to be 1.5%.

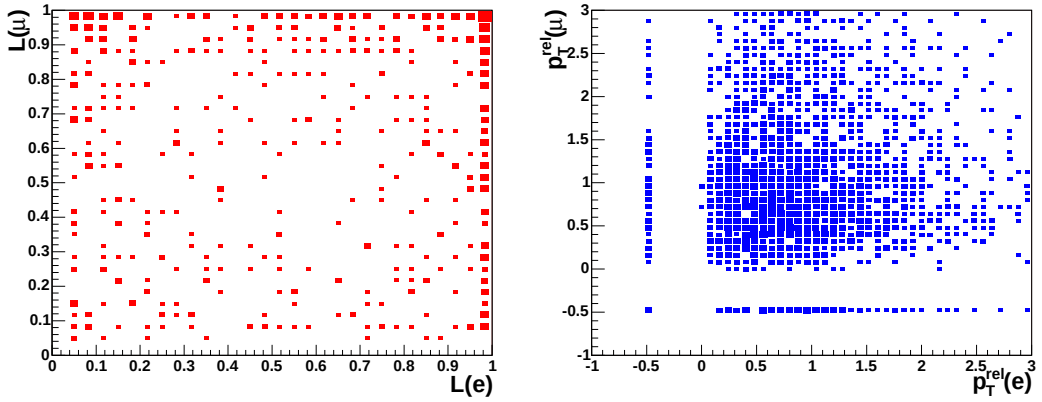

Figure 6.8: Right: $\mathcal{L}(\mu)$ vs $\mathcal{L}(e)$ for events where a double lepton tag is present; the muon tag is RS while the electron tag is WS. Left: $p_T^{rel}(\mu)$ vs $p_T^{rel}(e)$ for the same RS-WS combination. A cut $\mathcal{L}(\ell) > 0.05$ is applied for both plots to be consistent with what is done for the NN input. Both histograms are background subtracted via signed I.P.

The overall improvement for the Soft Lepton Taggers as a function of the p_t^{rel} and the contribution of the Kaon tags to such improvement, as a function of the lepton tag D, are shown in Fig.6.9.

A correlation is visible between the Lepton and the Kaon tagger, comparing these plots with the distribution in Fig.6.10. This is the result of a preliminary study I performed to investigate possible physical correlations among the taggers, using the $\ell + SVT$ sample on which the OSKT has been developed. At low $p_T^r el$, when the effect of the sequential decays reduces the standalone Muon tagger dilution, this is increased by the presence of an agreeing Kaon tag, since the two are both originated by the away b decay. The opposite happens when the two taggers disagree. The largest contribution from the Kaon tagger lies in the p_T^{rel} range where the largest improvement from the NN combination is found, and corresponds to values of $D_{SLT} \approx 20\%$, where the effective contribution from the Kaon tag is found, in Fig.6.9. This confirms that the largest part of the increase

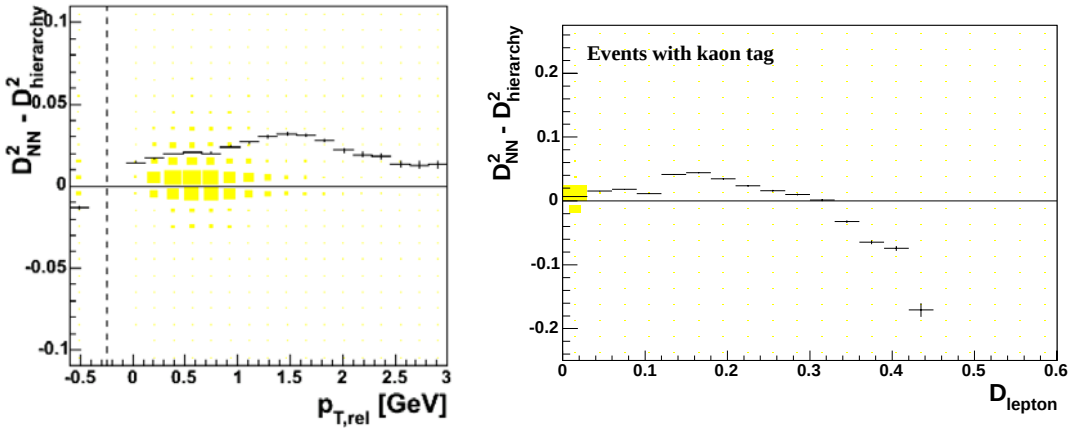

Figure 6.9: Left: Squared dilution difference $D^2_{NN} - D^2_{hie}$ as a function of the lepton p_t^{rel}, for Soft Muon tags (black dots); the overlaid yellow box scatter plot shows the event distribution of the 2 variables. Right: Squared dilution difference $D^2_{NN} - D^2_{hie}$ as a function of the lepton tag dilution D, for Soft Lepton tags when also a Kaon tag is present(black dots); the overlaid yellow box scatter plot shows the event distribution of the 2 variables. Note that the two plots have a different y axis scale.

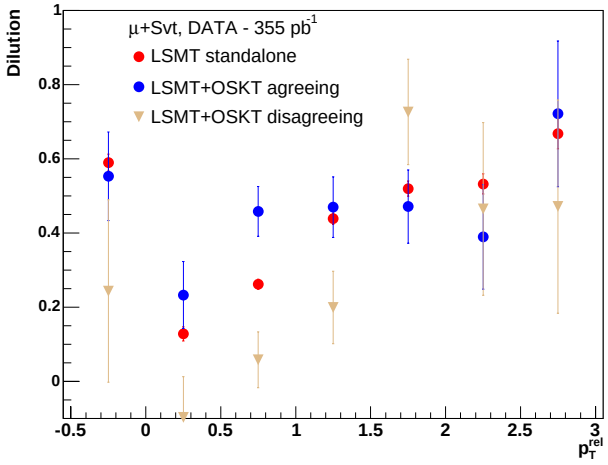

Figure 6.10: Dilution of the Soft Muon tags as a function of the p_t^{rel}, for the tagger standalone (red dots); for OS Kaon-Muon tags where the two responses agree (blue dots) and for cases when they disagree (brwon dots). The double Kaon-Muon tag rate is $\approx 5\%$ of all the Muon tags.

of lepton tagger dilution is introduced by its physics correlation with the Kaon tagger.

The largest contribution to the increased performances is anyway from events when a Jet Charge tagger is involved, due to its high efficiency. In particular, the improvement is pictured in Fig.6.11 for events where both a lepton tag and a jet charge tag are found ($\approx$ 20% absolute tagging efficiency). In this case, the improvement increases with the jet charge dilution. Infact, for low dilution jet charge tags the performances of both the NN and the hierarchical tagger are anyway determined largely by the lepton tags, and the difference between the two methods is small. At higher dilution the information from the jet charge becomes more important: while, by construction, it is considered in the hierarchy only if no lepton tag is found, in the NN combination it contributes significantly to increase the final performances.

We also look at the events tagged by both the Jet Charge tagger and the Kaon tagger. In such cases, a small improvement is observed in the low NNJQT dilution region, when the correlations with the Kaon tags help enhance the dilution of Class 3 Jet Charge. At higher dilutions, the hierarchical method is preferred, since it always chooses the Class 1 Jet Charge with a higher expected dilution than the OSKT Class 1. While this measurements and comparisons have been

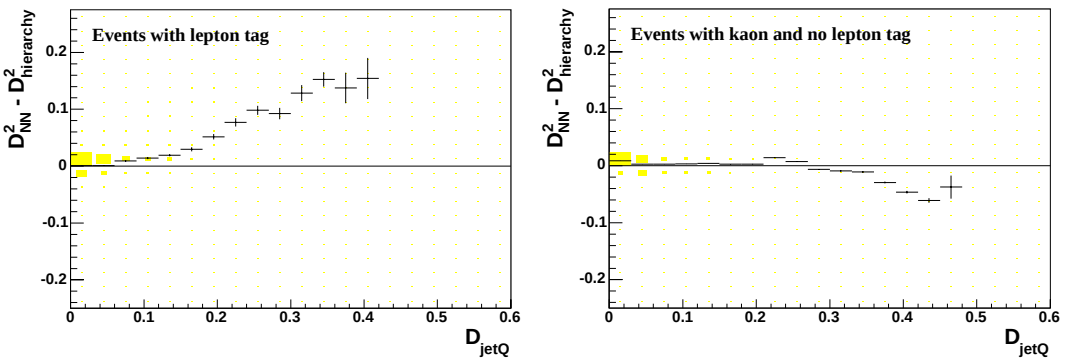

Figure 6.11: Squared dilution difference $D^2_{NN} - D^2_{hie}$ as a function of the NNJQT dilution D(black dots), in the events where a lepton tag is also present; the D_{NNJQT} event distribution is overlaid (yellow boxes).

made on the inclusive semileptonic sample, the determination of the effective contribution of the NN tagger to the CDF sensitivity to B_s oscillations has to be performed on the dedicated samples used for this analysis. This is described in the next section.

6.2 Impact on CDF sensitivity to B_s mixing

When developing the NN combination of the Opposite-side taggers, we already checked that the dilution is well calibrated by comparing the predicted and the measured dilution on the $\ell + SVT$ sample corresponding to 1 fb^{-1} (Fig.6.5). A final absolute scale factor is needed for the NN OST to be applied onto the B_s fully reconstructed hadronic and exclusive semileptonic decays ([63], [64]). In the case of the hadronic sample, this has been done using a maximum-likelihood based fit on B^+ and B_d^0 decays into different final states, selected using both the hadronic trigger ($D\pi$ modes) and the dedicated CDF J/ψ trigger ($J/\psi K$ modes), corresponding to $\int \mathcal{L} = 1\ fb^{-1}$.

The unbinned maximum likelihood fitter is built starting from signal and background's Particle Density Functions (PDF) and fractions: the PDF is the factorization of 4 terms describing: mass distribution, candidate's proper decay time, proper decay time resolution and flavour tagging

$$\mathcal{P} = \mathcal{P}_M \cdot \mathcal{P}_{ct} \cdot \mathcal{P}_{\sigma(ct)} \cdot \mathcal{P}_{D_{OST}} \tag{6.3}$$

The different modes are fit separately for mass and proper decay time. Then a combined fit for all the modes is performed, that returns, among the others, the value of the dilution Scale Factor and of Δm_d from the B_d^0 component. The combinatorial background in the sample is assumed not to mix and is assigned an independent tagger efficiency and average dilution. Similarly, the physics background contributions (see Sec.4), entering the proper time analysis at the level of a few percent, is also assumed to have a constant tagging asymmetry. A yet different set of tagger parameters is assigned to this component. Together with the hadronic analysis, also a semileptonic one in performed to evaluate the specific effectiveness of the NN OST. This is based on 3 exclusive decay modes of the kind $B^0 \to \ell^+ D^{(*)-} X \to \ell^+ [K^+\pi^-\pi^-] X$ and corresponding modes for the B^+. The same likelihood-based fitter as above is used. The results are reported in Tab.6.1. The asymmetry fit to the combined B^+ and B_d^0 semileptonic modes, tagged with the NN tagger, is displayed in Fig.6.12. While most of the systematic effects

decays	SF	$\varepsilon D^2, \%$
hadronic	1.103 ± 0.023 (stat)	1.80 ± 0.12 (stat)
semileptonic	1.079 ± 0.010 (stat) ± 0.033(syst)	1.82 ± 0.04 (stat) ± 0.11 (syst)

Table 6.1: Results for the combined NN OST scale factor on the hadronic and semileptonic decay modes. The corresponding tagging effectiveness is also reported.

entering the scale factor determination are common to hadronic and semileptonic modes [64][63], in the latter case the highest contribution ($\approx 95\%$) is given by

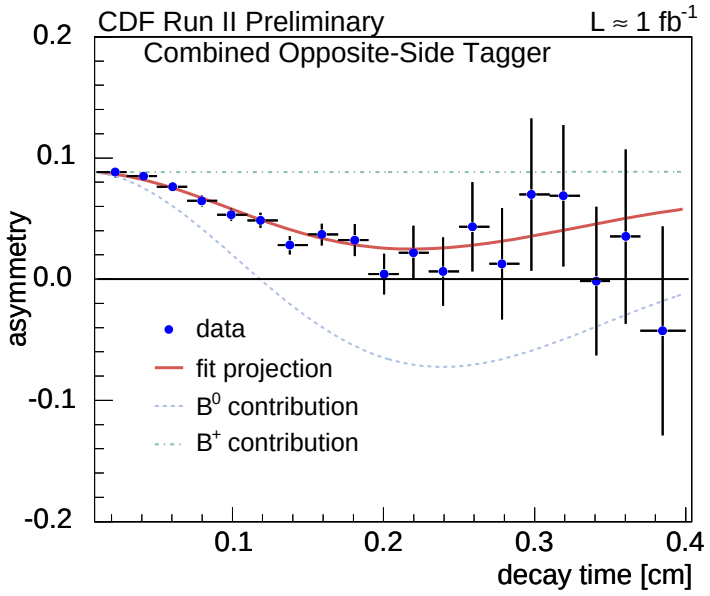

Figure 6.12: Asymmetry distribution for semileptonic B^+ and B_d^0 candidates, fitted with the maximum likelihod method to return Δm_D and the combined NN tagger scale factor.

semileptonic fake background (dilution and fraction estimate). The hadronic scale factor measurement is affected by negligible systematics, instead, which are not quoted in Tab.6.1.

Given the above numbers and the old estimates from the CDF mixing analyses ($\varepsilon D^2_{hie} = 1.50 \pm 0.10\%$, [23]), we can quantify the effective increase in tagging effectiveness introduced by the OSKT and the NN combination. The relative improvement in εD^2 is found to be $\approx 20\%$.

6.3 Tagger Application

Each event can be tagged by both a Same-side and an Opposite-side tag. As seen, on the trigger side only one tagging algorithm is applied. The combination of the Opposite-side taggers has just been discussed and leads to a single response on the flavour of the away b as well. Events where a response could be found from both sides have their final decision determined by the power of each tag, as well as whether or not the two tags agree. In case the two decisions *disagree*, the flavour tag is provided by the higher dilution tagger. Assuming that the two sides are completely uncorrelated in taking their decisions, the final dilution with which the event enters the mixing likelihood is given by

- tags agree: $D = \frac{D_1+D_2}{1+D_1 \cdot D_2}$

- tags disagree: $D = \frac{D_1 - D_2}{1 - D_1 \cdot D_2}$

where D_1 and D_2 are the dilutions of the two tags, ordered such that $D_1 > D_2$. As can be noted, if the two tags agree, the event is assigned a dilution higher than it would with a single tag. Else, the resulting dilution is lower than with a single tag. A systematic uncertainty is provided on the amplitude $\mathcal{A}$ in the amplitude scan (see Ch.7) due to correlations among the two sides.

Chapter 7

The CDF results on Δm_s

The three parallel analyses carried by CDF on fully reconstructed hadronic modes, the fully and partially reconstructed $B_s \to D_s[\phi(KK)\pi]\pi$ (*golden*) decay mode and the semileptonic modes have been described in Ch.4. There, the flavour taggers established from the past have been surveyed. In the following chapters 5,6 a detailed analysis of the Opposite-side Kaon tagger and its combination with the other Opposite-side taggers has been given. If both an Opposite-side tag and a Same-side tag are present, they are combined using the formula 6.3. Here we show the results of the CDF analyses together and evaluate the statistical significance of such a result. This also produces an evaluation of the CKM matrix element ratio $|V_{ts}/V_{td}|$ which constrains the relative Unitarity Triangle, as described in Ch.1.

7.1 Likelihood

The *amplitude scan* technique [20] recalled in Ch.2 is used for the analysis of B_s oscillations. We examine the values of the amplitude A within the range $\Delta m_s \in [0.0, 35.0]\ ps^{-1}$. Data are fit using an unbinned likelihood function defined as follows:

$$\mathcal{L} = f_{signal}\mathcal{L}_{signal} + f_{comb}\mathcal{L}_{comb} + f_{prompt}\mathcal{L}_{promp} + f_{phys}\mathcal{L}_{phys} \tag{7.1}$$

where f_i ($i = signal, comb, prompt, phys$) are the relative contributions of signal, combinatoric background, prompt background and physics background to the physics sample, as described in Ch.4 for the various analyses. Each of the likelihood functions is the product:

$$\mathcal{L}_i = \mathcal{L}_i^m \mathcal{L}_i^{ct} \mathcal{L}_i^{\sigma_{ct}} \mathcal{L}_i^D \tag{7.2}$$

where $\mathcal{L}_i^m$ represents the mass distribution; $\mathcal{L}_i^{ct}$ and $\mathcal{L}_i^{\sigma_{ct}}$ are the proper decay time and proper decay time uncertainty terms; finally, $\mathcal{L}_i^D$ is the event-by-event

tagging dilution PDF.
The amplitude term $\mathcal{A}$ is inserted in the $\mathcal{L}_i^{ct}$ term for signal, where the time-dependent mixing probability Eq.1.21 is used. In particular, as seen in Sec.6.3, an event can be tagged by 0, 1 or 2 flavour taggers. The corresponding terms are:

- untagged:

$$\mathcal{L}^{ct}_{signal} = \left(1 - \sum_{j=1}^{N_{taggers}} \varepsilon_j\right) \frac{\kappa}{c\tau} e^{-ct/c\tau} \otimes \mathcal{G}(ct) \otimes F(\kappa) \otimes \varepsilon_{SVT}(ct) \qquad (7.3)$$

- single tag:

$$\mathcal{L}^{ct}_{signal} = \frac{\varepsilon_j}{2}\left[1 + \mathcal{A}S^jT^jD^j cos(\Delta m_s ct)\frac{\kappa}{ct}e^{-ct/c\tau}\right] \otimes \mathcal{G}(ct) \otimes F(\kappa) \otimes \varepsilon_{SVT}(ct) \qquad (7.4)$$

- double tag:

$$\mathcal{L}^{ct}_{signal} = \left[\frac{(1 + T^jT^lS^jD^jS^lD^l) + (T^jS^jD^j + T^lS^lD^l)\mathcal{A}cos(\Delta m_s ct)\frac{\kappa}{ct}e^{-ct/c\tau}}{(1+T^j)(1+T^l)}\right] \otimes$$
$$\otimes \mathcal{G}(ct) \otimes F(\kappa) \otimes \varepsilon_{SVT}(ct) \qquad (7.5)$$

Here ε_j indicates the efficiency of the j-th tagger ($N_{taggers} = 2$), T^j is the j-th tagger decision, D^j its predicted dilution, S^j the relative scale factor, as described in Ch.6; κ is the k-factor accounting for unmeasured momentum in the partially reconstructed decays (e.g. $\kappa = 1$ in the case of the hadronic decays), $F(\kappa)$ its distribution as obtained from Monte Carlo simulation; $\mathcal{G}(ct)$ is the detector proper decay length Gaussian resolution, $\varepsilon_{SVT}(ct)$ the SVT trigger efficiency curve. The hadronic modes, the *golden* mode, both fully and partially reconstructed, and the semileptonic modes are fit separately due to the differences in sample composition and the proper decay length features.

The application of the flavour taggers on both sides for the partially reconstructed hadronic modes is treated in exactly the same manner as for the completely reconstructed decay $B_s \to D_s(\phi(KK)\pi)\pi$. In this case, in fact, only one *neutral* particle is lost: this means its presence cannot affect the charge-flavour correlation used to tag. The same predicted dilution and relative scale factors as for the main signal peak are thus applied in the signal likelihood.

The fitting framework has been tested extensively in order for it not to respond in a biased way at various Δm_s, using toy Monte Carlo experiments with the same statistics and sample features as the real B_s data samples. It has been also verified that the amplitude scan obtained from the above likelihood is able to return a value of Δm_d compatible with the world average. The result is shown in Fig.7.1, for the first 355 pb^{-1} of fully reconstructed B_d decays. CDF reproduces the world average result as expected, with a statistical uncertainty consistent with our expectations for this data sample.

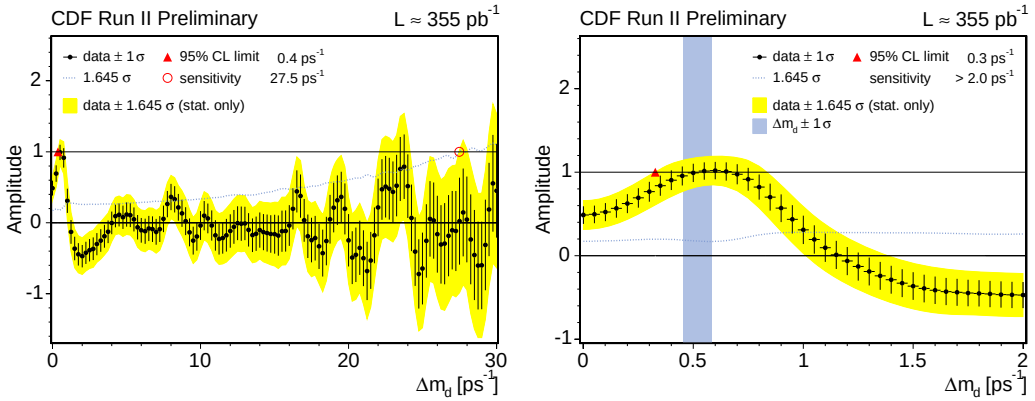

Figure 7.1: Example of amplitude scan as obtained from the likelihood fitting technique, applied to neutral B_d meson. It uses 355 pb^{-1} of fully reconstructed decays. It is displayed here as a sanity check for the framework and does not represent teh most up-to-date analysis configuration.

7.2 Statistics for the different modes

Using the process described in Ch.4 for the various decay modes, we select ≈ 7600 B_s hadronic and 61500 semileptonic B_s signal events. As said, the hadronic decays, both fully and partially reconstructed, are identified using a Neural Network. This technique provides an increase in the overall number of events with respect to the cut-based approach used in the previous analysis [23]. $\approx 20\%$ more *fully reconstructed* events are added to signal in the fully reconstructed decays, out of the same set of data. The partially reconstructed modes provide almost as many signal events as the fully reconstructed, but with a lower $\mathcal{S}/\mathcal{B}$. The yields for the various modes and their $\mathcal{S}/\mathcal{B}$ are shown in Tab.7.1

The effective statistical power of each of these samples as far as B_s mixing is concerned does not correspond barely to their yields. It is also determined by

B_s decay mode	Yield	$\mathcal{S}/\mathcal{B}$
$D_s(\rightarrow \phi\pi)\ \pi$	1900	11.3
partially reconstructed	3300	3.4
$D_s(\rightarrow K^*K\ )\ \pi$	1400	2.0
$D_s(\rightarrow 3\pi\ )\ \pi$	700	2.1
$D_s(\rightarrow \phi\pi)\ 3\pi$	700	2.7
$D_s(\rightarrow K^*K\ )\ 3\pi$	600	1.1
$D_s(\rightarrow 3\pi\ )\ 3\pi$	200	2.6

Table 7.1: Signal yield, and $\mathcal{S}/\mathcal{B}$ for the various hadronic modes, as obtained from a fit to the invariant mass distributions.

the proper time uncertainty. In the case of the fully reconstructed decays, the distribution of σ_{ct}, shown in Fig.7.2, has an average of 26 μm, corresponding to $\approx 1/4$ of an oscillation period at the value of $\Delta m_s = 17.3$ measured by CDF. The partially reconstructed hadronic decays have an average σ_{ct} of 29 μm: as expected, their time resolution is similar to the fully reconstructed modes, as only one soft neutral particle is lost for these decays. For semileptonic decays, the average time resolution is 45 μm (Fig.7.2). As explained in Sec.4.2.2, events with a higher $m(\ell D_s)$ enter the mixing likelihood with a higher weight as their σ_{ct} is lower.

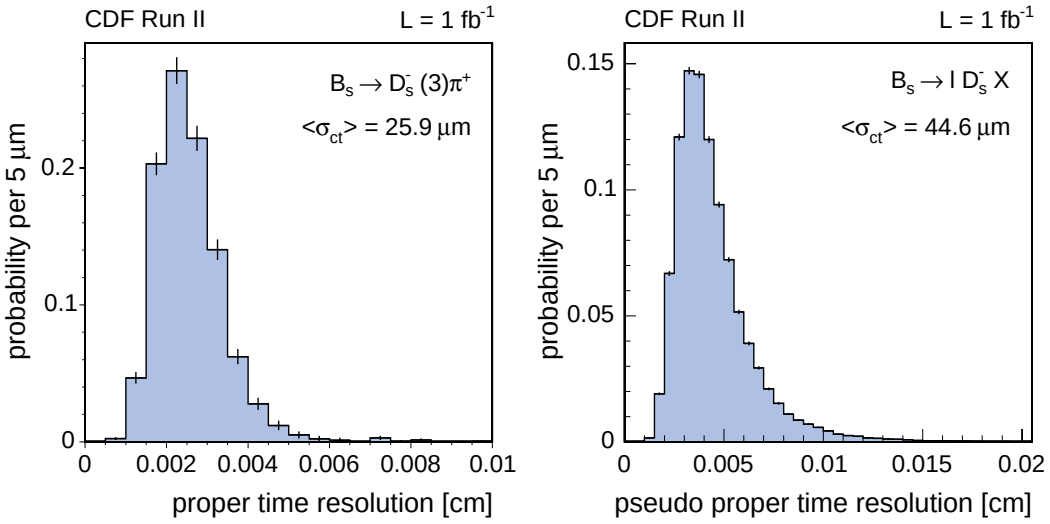

Figure 7.2: Distribution of proper time resolution for fully reconstructed hadroni (left) and partially reconstructed semileptonic (right) decays. The mean in the first case is $< \sigma_{ct} >_{had}$= 25.9 μm, while in the second it is $< \sigma_{ct} >_{semi}$= 44.6 μm

The Opposite-side tagger performances are reported in Tab.6.1 for the hadronic and semileptonic modes. A Neural Network-based version of the Same Side Kaon Tagger has been introduced for this analysis in order to improve the signal Kaon selection. This combines PID with kinematical information, like the projection p_L^{rel} of the Kaon momentum along the reconstructed B_s momentum [44]. The use of a NN has slightly improved the overall effectiveness (fractional improvement +2%) with respect to the one quoted in Eq.4.19.

7.3 Amplitude Scan and extraction of Δm_s

I report here the latest results of the amplitude scan and the extraction of the value of Δm_s from CDF. These results include the above improvements on the signal selection. Also, they are performed with the improved flavour tagging system, comprising: the Opposite Side Kaon Tagger and its combination with the other Opposite-side taggers, representing the main contribution of this thesis work, and the Same Side Kaon Tagger.

First we report the results of the different analyses separately, to illustrate the different statistical weights of the various contributions. Finally, we show the final result, obtained as a combination of the three likelihood functions for fully and partially reconstructed hadronic decays and for semileptonic decays.

7.3.1 The hadronic scan

The hadronic amplitude scan, **without** the *golden* mode $B_s \to D_s(\phi(KK)\pi)\pi$ contribution, is shown in Fig.7.3 for the full 1 fb^{-1} of data. A significant peak is observed at $\Delta m_s \approx 17.75\ ps^{-1}$, where $\mathcal{A} = 1.29 \pm 0.29$. The sensitivity of the hadronic analysis, again defined as the frequency value at which $\sigma(A) = 1/1.645$, is 28.3 ps^{-1}. It has to be noted that, thanks to the increased yield produced by the improvement in the selections and the flavour tagging improvements, this number overcomes the *combined* sensitivity of the whole previous result from CDF, that was 25.8 ps^{-1}. Two undershoots are present aside the main peak at $\Delta m_s \approx 17.75\ ps^{-1}$: these structures are expected and can be explained as the Fourier tranform of the time-dependent SVT trigger efficiency curve into the Δm_s frequency domain.

The measurement is statistics-limited since very low frequency values ($\leq$ 5 ps^{-1}). The evaluation of the various contributions to the final systematic uncertainty of this analysis, CDF uses toy Monte Carlo samples mimicking the data sample qualities and amplitude scan sensitivity. To estimate any given systematic uncertainty, a dedicated toy MC sample is generated and it is fit twice. The first one using parameterizations reproducing the situation at generation. In the second fit, the systematic effect is simulated by introducing an imperfection in

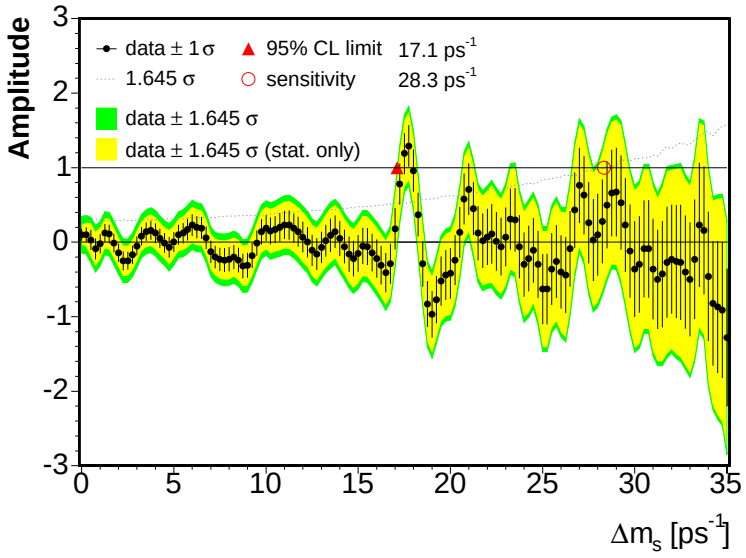

Figure 7.3: The amplitude scan for the hadronic analysis, $\int \mathcal{L} = 1\ fb^{-1}$. The lighter band is for statistical uncertainty, the darker band describes stistical+systematical uncertainties.

the fitter. The value for the given i-th MC sample is given by:

$$\sigma^i_{syst} = (\mathcal{A}^i_1 - \mathcal{A}^i_0) + (1 - \mathcal{A}^i_0) \cdot \frac{\sigma^i_1 - \sigma^i_0}{\sigma^i_0} \quad [20] \tag{7.6}$$

The value of σ^i_{syst} is extracted from a distribution obtained from 1000 toy MC.

For the hadronic scan, the dominant contribution at lower sampled frequencies ($\Delta m_s \leq 20\ ps^-1$) is found to be the SSKT dilution absolute scale. Another important contribution, dominant at higher frequencies, is provided by the determination of the absolute scale of ct resolution. Part of this is due to the fact that mixing of the B_d is ignored in the PDF used in the fit for oscillations of B_s, although the B_0 decay component reflected into the $D_s^- \to K^*K^-$-related decays is not negligible. Another source, about constant all over the probe range, is the correlation in the tagging algorithms between the two sides. The combined systematic uncertainty is depicted in Fig.7.4 as a function of Δm_s.

Under the hypothesis that the significant peak we observe in the amplitude scan is determined by a real obscillation signal, we measure Δm_s from the minimum of the quantity:

$$\Lambda \equiv \log \left[\frac{\mathcal{L}^{\mathcal{A}=0}}{\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)} \right] \tag{7.7}$$

where $\mathcal{L}$ is the likelihood of eq.7.1. Following the Pearson-Neyman lemma, we use

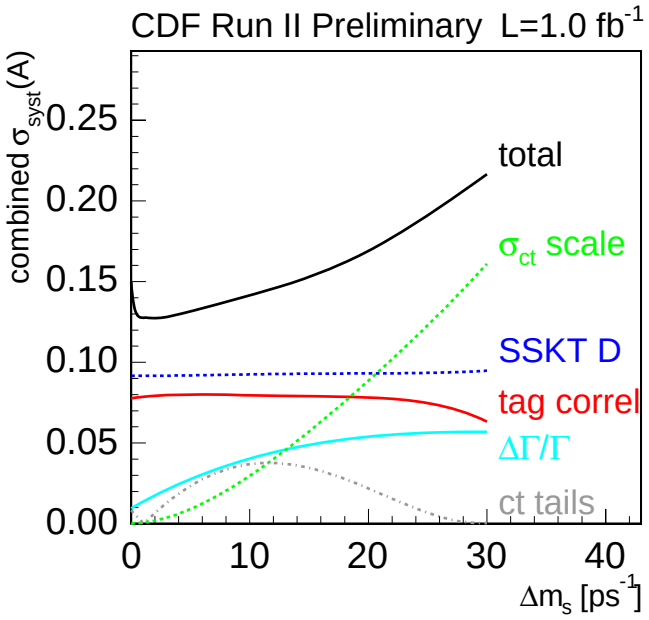

Figure 7.4: Systematic uncertainties on $\mathcal{A}$ as a function of the probe frequency Δm_s for the hadronic channels.

the logarithm of the ratio of likelihoods for the hypothesis of oscillations ($\mathcal{A} = 1$) at the probe value Δm_s and the frequency-independent hypothesis of random fluctuations ($\mathcal{A} = 0$). The profile of the likelihood ratio for the hadronic analysis is displayed in Fig.7.5. At the minimum of the likelihood, we have:

$$\Delta m_s(hadronic) = 17.66 \pm 0.11\ (stat.)\ ps^{-1}$$

7.3.2 The *golden* mode scan

As said, this analysis uses the fully reconstructed and partially reconstructed events in the $B_s \rightarrow D_s(\phi(KK)\pi)\pi$ mass range. The relative amplitude scan is shown in Fig.7.6 for the full 1 fb^{-1} of data. Also in this case a significant peak is observed around $\Delta m_s \approx 17.75\ ps^{-1}$, with $\mathcal{A} = 1.27 \pm 0.34$. In this case, a sensitivity of 22.7 ps^{-1} is obtained. The scans for completely and partially recosntructed decays separately are also shown in Fig.7.7. These show the $B_s \rightarrow D_s(\phi(KK)\pi)\pi$ signal has a better sensitivity than the $B_s \rightarrow D_s\rho$ and $B_s \rightarrow D_s^*\pi$ decays as exptected, given its better decay time resolution.

The systematic uncertainties on the amplitude are evaluated the same way as for the hadronic analysis and are depicted in Fig.7.8 as a function of Δm_s, together with the various contributions considered. In this case, the systematic uncertainties on the fully reconstructed decays are still to be applied for the $B_s \rightarrow$

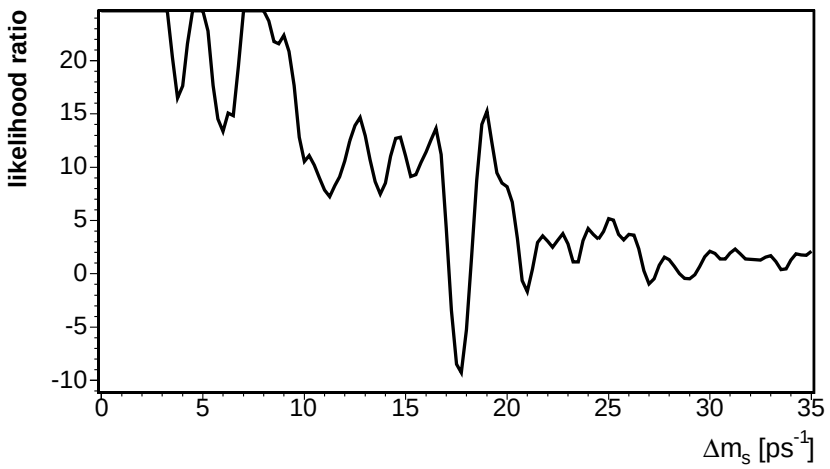

Figure 7.5: Likelihood Ratio distribution for the 5 fully reconstructed hadronic modes.

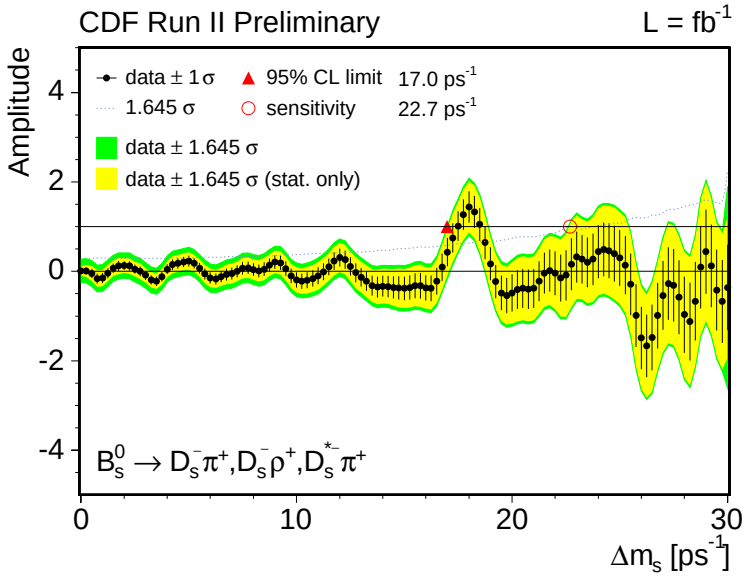

Figure 7.6: The amplitude scan for the golden channel analysis, fully and partially reconstructed decays together, $\int \mathcal{L} = 1\ fb^{-1}$. The lighter band is for statistical uncertainty, the darker band describes stistical+systematical uncertainties.

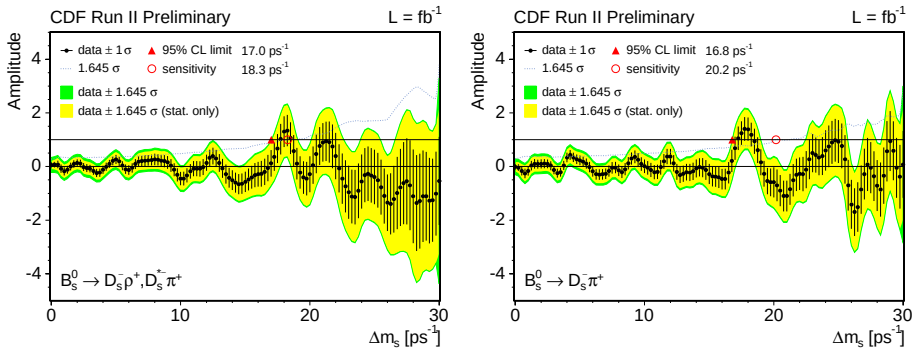

Figure 7.7: The amplitude scan for the partially (left) and fully (right) reconstructed decays of the *golden mode* separately.

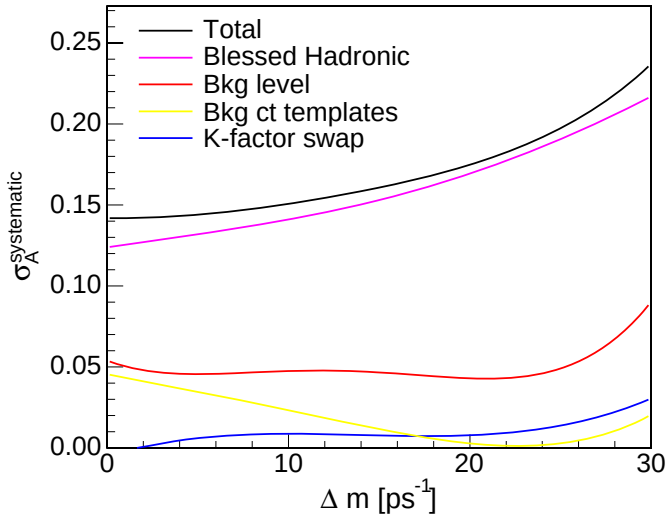

Figure 7.8: Systematic uncertainties on $\mathcal{A}$ as a function of the probe frequency Δm_s for the golden channel.

$D_s(\phi(KK)\pi)\pi$ part. The systematic effects associated with the reconstruction of the $B_s \to D_s\rho$ and $B_s \to D_s^*\pi$ decays are dominated by the background model, in particular the combinatorial background in the low mass region. Both the effects of the background in the mass and proper decay time are considered, for which a single average template is adopted. Also the absolute model for the κ factor of each decay and the possible mis-assignment ("swap") of part of the κ-factor population due to a wrong estimate of the various decay contributions are investigated. CDF finds the decay model adequate and in agreement with the PDG numbers for the various resonances implied; therefore we choose not to quote any uncertainty on the amount of momentum unmeasured.

The likelihood ratio profile shows a minimum of -7.4 at

$$\Delta m_s(golden) = 18.01^{+0.17}_{-0.18}\ (stat.)\ ps^{-1}$$

and is shown in Fig.7.9, combined and for the two contributions separately.

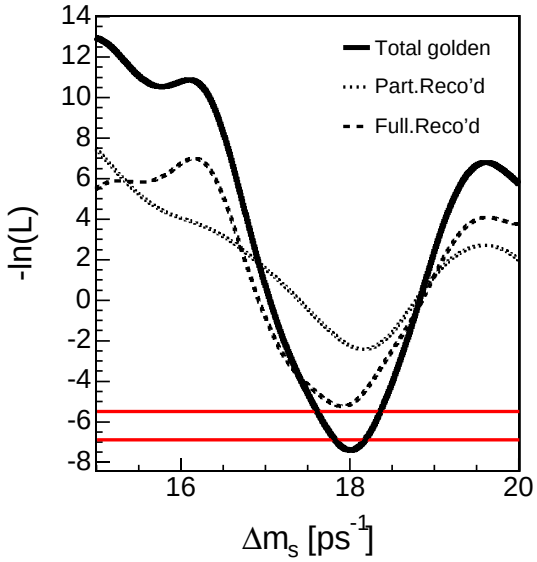

Figure 7.9: Likelihood Ratio distribution for the golden channel analysis. The solid line is the combined profile of partially and fully reconstructed contributions. The dashed line is for full reconstruction only; the dotted one for partial reconstruction only. The lower flat line indicates the 1σ level, the upper flat line shows the 95% C.L. bounds.

7.3.3 The semileptonic mode scan

Finally, also the scan for the semileptonic modes is displayed in Fig.7.10 for the full 1 fb^{-1} of data. As expected from the rate of momentum unmeasured in these

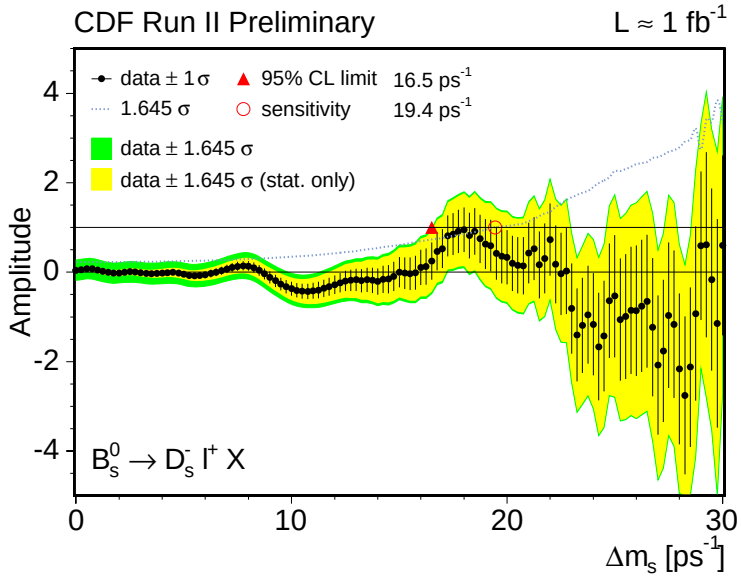

Figure 7.10: The amplitude scan for the semileptonic analysis, $\int \mathcal{L} = 1\ fb^{-1}$. The lighter band is for statistical uncertainty, the darker band describes stistical+systematical uncertainties.

decays, the sensitivity of this sample is lower than for the previous hadronic cases, including the *quasi*-reconstructed decays in the golden mode. In this case, the sample is sensitive up to 19.4 ps^{-1}. For comparison with the hadronic modes, $\mathcal{A}(\Delta m_s = 17.75\ ps^{-1}) = 0.86 \pm 0.49$.

As for the previous cases, a summary plot (Fig.7.11) for the various constributions in the total systematic uncertainty on $\mathcal{A}$ is shown, together with a comparison of statistical and systematic uncertaintites, which dominate in this case up to $\approx 10\ ps^{-1}$, a frequency double the hadronic cross-over. The dilution scale factor uncertainty term is predominant, mainly due to the composition and the dilution of the combined Opposite-side tagger for the fake background. Other contributions include modeling of the physics background, where the uncertainty is dominated by our comprehension of the prompt background; and of the proper decay time resolution distribution and Scale Factor determination.

The value at the likelihood minimum is

$$\Delta m_s(semileptonic) = 17.89\ ps^{-1}$$

We choose not to quote any errors on this number, since the likelihood has not a really gaussian minimum and the defintion of "error" would not be the same as for the hadronic results.

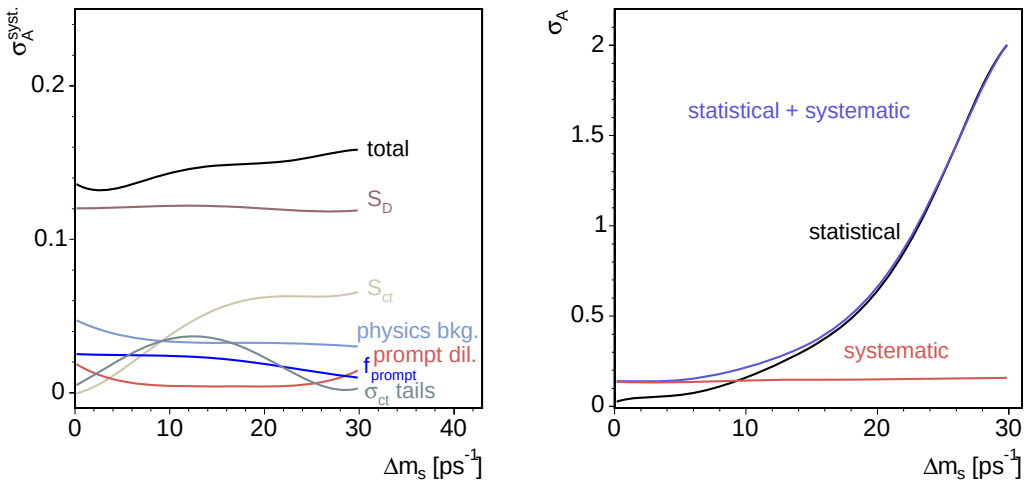

Figure 7.11: Systematic uncertainties on $\mathcal{A}$ as a function of the probe frequency Δm_s for the semileptonic channels.

7.3.4 The combined amplitude scan

The amplitude scan obtained from the combined likelihood of all the above analyses is displayed in Fig.7.12. Also the statistical-only scan is shown (Fig.7.13), to disentangle visually the effects of the systematic uncertainties.

Around $\Delta m_s = 17.75\ ps^{-1}$ CDF observes an amplitude $\mathcal{A} = 1.21 \pm 0.20$, that is $\approx 6\sigma$ inconsistent with 0. The combined sensitivity of all modes is 31.3 ps^{-1}, compared to the old combined sensitivity of 25.8 ps^{-1}. As one can notice, the scan is dominated up to $\approx 10\ ps^{-1}$ by the semileptonic sample, which is equivalent to say that the systematic part uncertainties are predominant up to this value. The three dimensional plot depicting the likelihood profile in the $\Delta m_s - \mathcal{A}$ space is shown in Fig.7.14

The separate effectiveness of Opposite-side taggers and Same-side Kaon tagger are depicted in Figs.7.15. In both cases, a peak in $\mathcal{A}$ is visible. The value of the amplitude around 17.75 ps^{-1} is compatible with 1 for both tags. In particular, the value $\mathcal{A} = 1$ for the SSKT proves that the evaluation of the tagger dilution has been performed with a very good accuracy. It can be noted that Opposite-side taggers are less statistically significant in general. Nevertheless, they produce a visible peak at $\Delta m_s \approx 18\ ps^{-1}$. It has a significance $\mathcal{A}/\sigma(\mathcal{A}) \approx 4.4$. The SSKT-tagged events have less fluctuations at low probe values. An asymmetry can be noticed in the peak area, for both taggers. This means that the dilution has a dependence on the B_s proper time, that translates into an asymmetric tail when passing to the frequency domain. A possible explanation concerns the

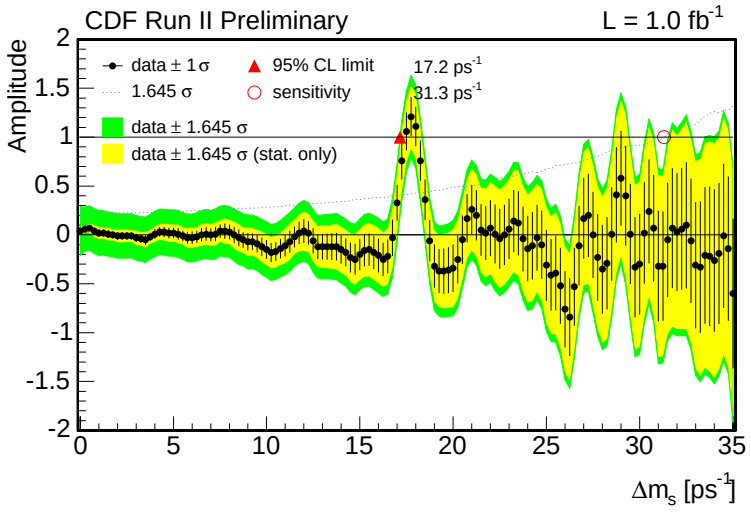

Figure 7.12: The combined amplitude scan for all the mixing analysis, $\int \mathcal{L} = 1\ fb^{-1}$. The lighter band is for statistical uncertainty, the darker band describes stistical+systematical uncertainties.

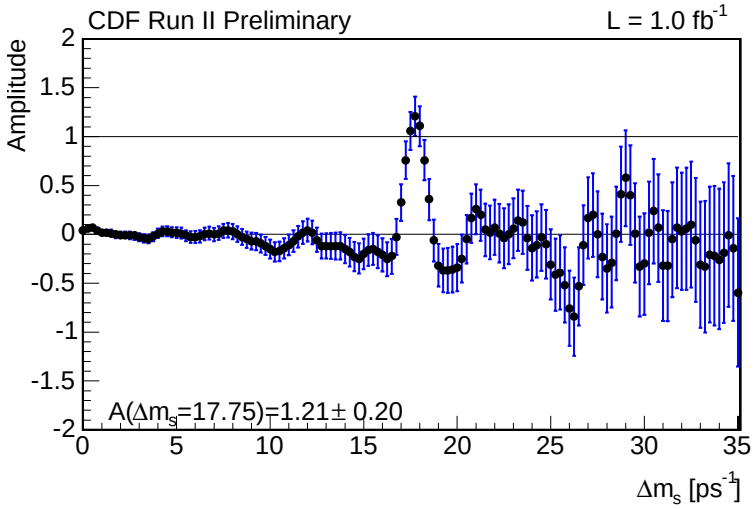

Figure 7.13: The combined amplitude scan for all the mixing analysis, $\int \mathcal{L} = 1\ fb^{-1}$. Only the statistical uncertainty is reported for each point.

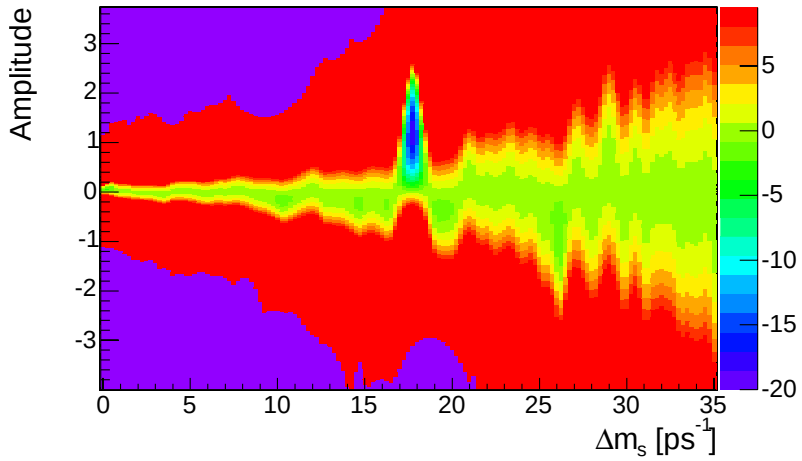

Figure 7.14: Combined negative log likelihood as a function of amplitude and Δm_s.

composition of the calibration sample for the SSKT and the OST separately. The purity of the sample could be reduced at shorter B_s proper decay times by the contributions of short-lived B_c and partially reconstructed hadronic decays in the calbration samples. Therefore, the resulting dilution estimate will depend on the proper decay time.

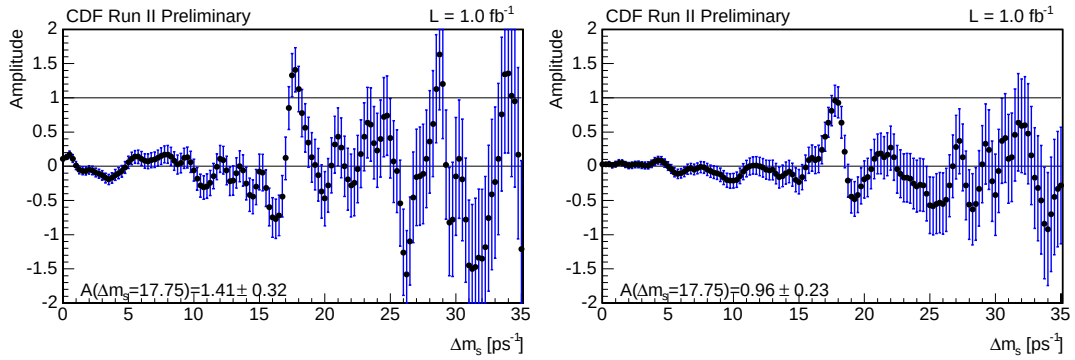

Figure 7.15: Amplitude scan with only Opposite-side tags (left) and Same-side tags (right). The error is statistical only.

A steep minimum is observed for the likelihood at the frequency where the amplitude is compatible with unity. The likelihood profile is better observed in the one dimensional plot in Fig.7.16 The minimal value observed for the negative log likelihood ratio is -17.26.

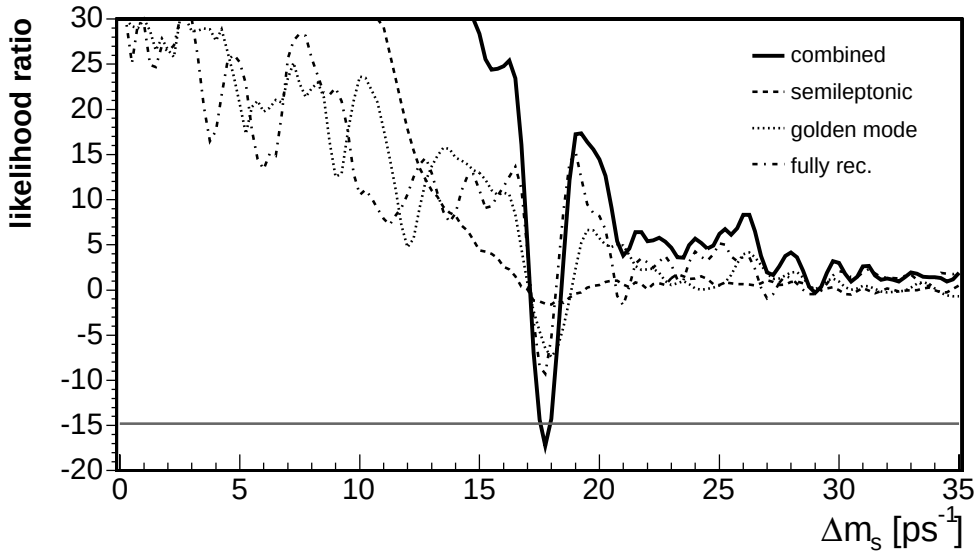

Figure 7.16: Combined negative log likelihood and profiles of the various analyses.

7.4 Systematic uncertainties on Δm_s

The overall systematic uncertainty on the measurement of Δm_s is the combination of the separate systematic uncertainties of the various analyses. Both the hadronic and semileptonic analyses have a symmetric systematic error $\sigma^{syst}(\Delta m_s) = 0.07\ ps^{-1}$. As with the uncertanties on the amplitude scan, the magnitude of the each systematic effect is evaluated by incorporating it into a detailed toy Monte Carlo and assessing the potential shift in the measured value of Δm_s. All the effects affecting the amplitude scan are also probed for Δm_s and found to be negligible. Three primary sources related with our knowledge of the absolute lifetime measurement scale are also investigated:

- **Silicon detector alignment** This includes any possible effects introduced in our knowledge of the absolute time scale by a misalignment of the silicon detector; they are tested by creating artificial distortions in the Monte Carlo simulation of the detector compatible with the radial displacements and the bowing of silicon sensors within tolerances as established from a physical survey. The maximum bias found at the end of the standard lifetime measurement procedure is 1.0 μm;

- **Track-fit bias** The effect of the mis-measurement of the tranverse momentum p_t of tracks on the evaluation of the B meson decay length in the transverse plane L_{xy} and eventually on its lifetime. Again the bias is stud-

Source of systematic	uncertainty (ps^{-1})
Amplitude scan systematic effects	< 0.01
Silicon detector alignment	0.04
Track fit bias	0.05
Primary vertex bias	0.02
total	0.07

Table 7.2: Summary of systematic uncertainties on the measurement of Δm_s. The total uncertianty is derived as the sum in quadrature of the various contributions and is 0.07 ps^{-1}.

ied introducing variations of the measured momentum around its true value in Monte Carlo. The bias is evaluated to be 1.3 μm;

- **Primary Vertex bias** Our knowledge of the absolute time scale of the B meson is also mis-lead by how well we determine the position of the primary vertex. This is in particular the case when displaced tracks originated from the away b-hadron decay are included in the primary vertex. For instance, events when an Opposite-side flavour tag is present are enriched in this kind of tracks and the bias is enhanced. It is found to be 1 μm after comparing the primary vertex position reconstructed event by event with the average beam position.

Tab.7.2 summarizes the systematic uncertainties on Δm_s derived from the above contributions. In the case of the golden mode, a further systematic uncertainty of 0.0256 ps^{-1}, that accounts for our knowledge of the background model and the assignment of the k-factors to the partially reconstructed decays.

The uncertainty on the combined Δm_s measurement is evaluated by smearing each of the likelihood curves with a Gaussian of the appropriate width. By doing so, CDF measures the $B_s^0 - \bar{B}_s^0$ oscillation frequency

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

From the combined log likelihood ratio, including systematics, also the following CL intervals are obtained:

$$\Delta m_s \in [17.60, 17.92]\ ps^{-1} \text{ at } 90\% \text{ CL}$$
$$\Delta m_s \in [17.56, 17.95]\ ps^{-1} \text{ at } 95\% \text{ CL}$$

7.5 Evaluation of the statistical significance

To evaluate the statistical significance of the measurement, we compute the *false-alarm* probability p. Given two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ (in this case that the log

likelihood minimum *is* or *is not* produced by flavour oscillations at that frequency), p is the probability that we declare true the wrong hypothesis $\mathcal{H}_0$ (also called *null-hypothesis*) when $\mathcal{H}_1$ is actually in effect. p is calculated as

$$p = \int_{\Omega_1} f_0(x)dx \tag{7.8}$$

where x is the observable (the minimal value of Λ in Eq.7.7 in our case), $f_0(x)$ its P.D.F. according to (false) hypothesis $\mathcal{H}_0$ and Ω_1 is the *decision region*, that is the interval in x to accept $\mathcal{H}_1$. For the Δm_s measurement, the probability density of the the minimal value of Λ is determined directly from data, randomizing the tagging decision many times independently. We also checked that toy Monte Carlo experiments return the same answer. From Eq.7.8, by defining $\Omega_1 = [-min(\Lambda)_{data}, +\infty] \; = \; [17.26, +\infty]$, we get

$$p = (8.0 \pm 1.5) \times 10^{-8} \tag{7.9}$$

that corresponds to $(5.367 \pm 0.034)\ \sigma$ in the Gaussian approximation of the log likelihood minimum. This corresponds to the observation of B_s flavour oscillations.

Fig.7.17 depicts the log likelihood minimum P.D.F. generated from $O(350\ M)$ randomized tag experiments. With the above likelihood we also evaluate the

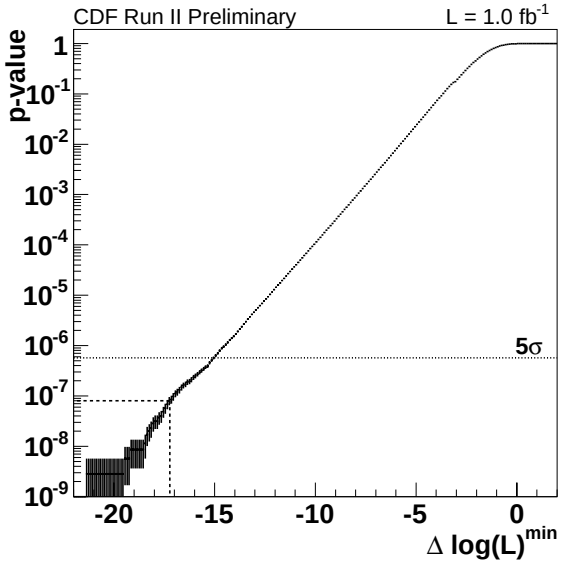

Figure 7.17: Value of false alarm probability p as a function of the negative log likelihood ratio, Λ, as obtained after $O(350M)$ tagging decision randomizations of the CDF data sample. The point where the dashed lines cross indicates the minimum of Λ and the corresponding p value.

flavour asymmetry in the B_s proper time domain for the hadronic sample, both partially and fully reconstructed. The result is shown in Fig.7.18. The asymmetry in each time bin in calculated with the likelihood used to measure Δm_s. Per each bin we turn off the oscillating term and measure the asymmetry, $A = \frac{N-\bar{N}}{N+\bar{N}}$, where $N(\bar{N}) = N(B_s^0)(N(\overline{B}_s^0))$. For each event, we subtract a multiple of the measured oscillation period $T = 2\pi/17.77\ ps^{-1}$ from the proper decay time. This enhances the visible asymmetry in the data, which would otherwise be distributed over about ten oscillation periods. The expected shape of this figure is a cosine with an amplitude equal to the measured mixing amplitude at $\Delta m_s = 17.77\ ps^{-1}$, which is $\mathcal{A} = 1.28 \pm 0.22$. This cosine is plotted in the figure. The data are compatible with the expectation: the χ^2 is 4.77, giving a probability of 44% for 5 degrees of freedom.

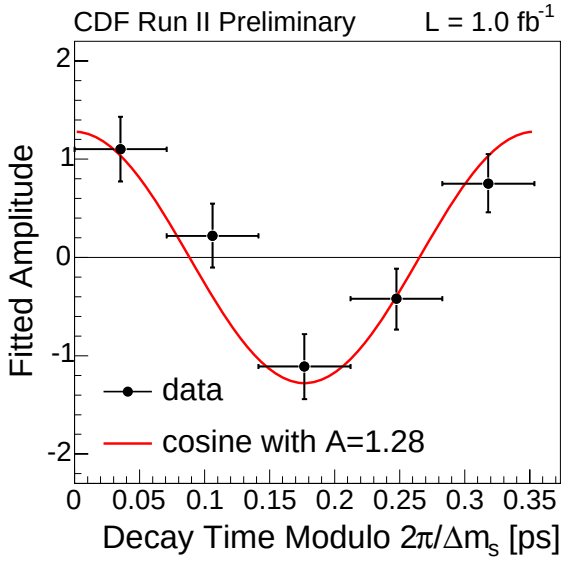

Figure 7.18: Flavour asymmetry as a function of the B_s proper decay time. The value of each point is calculated from the mixing likelihood and is expected to be distributed with a cosine of amplitude and frequency corresponding to the overlaid curve. $\chi^2/ndf = 4.77/5$

7.6 $|V_{td}/V_{ts}|$ and measurement interpretation

7.6.1 Value of $|V_{td}/V_{ts}|$

Using the expression 1.24, CDF infers the value of the ratio of the CKM elements $|V_{td}/V_{ts}|$. Aside of the measured value of Δm_s, the following input values are used:

- $\Delta m_d = 0.507 \pm 0.005$ [21];
- $\frac{m_{B^0}}{m_{B_s}} = 0.98390 \pm 0.00019$ [67];
- $\xi = 1.210\ ^{+0.047}_{-0.035}$ [68].

CDF finds

$$|V_{td}/V_{ts}| = 0.2060 \pm 0.0007(exp)\ ^{+0.0081}_{-0.0060}(theo)$$

As expected, the error is at the 4% level and is already almost entirely due to theoretical lattice calculations.

7.6.2 Interpretation of the results

It is of primary interest to establish whether the value of Δm_s measured by CDF is compatible with the SM expectations. Infact, as illustrated in Ch.1, $\Delta F = 2$ flavour transitions are sensible to New Physics contributions that may result in an enhancement of the mixing frequency with respect to the SM predictions. The latter can be retrieved performing a global unbiased fit to all the *other* constraints to the Unitarity Triangle and using the output CKM element values in the SM Δm_s expression.

This has been done, for example, by the **UT**Fit Collaboration [69] and the indirect prediction for Δm_s is shown in Fig.7.19. The most probable value is

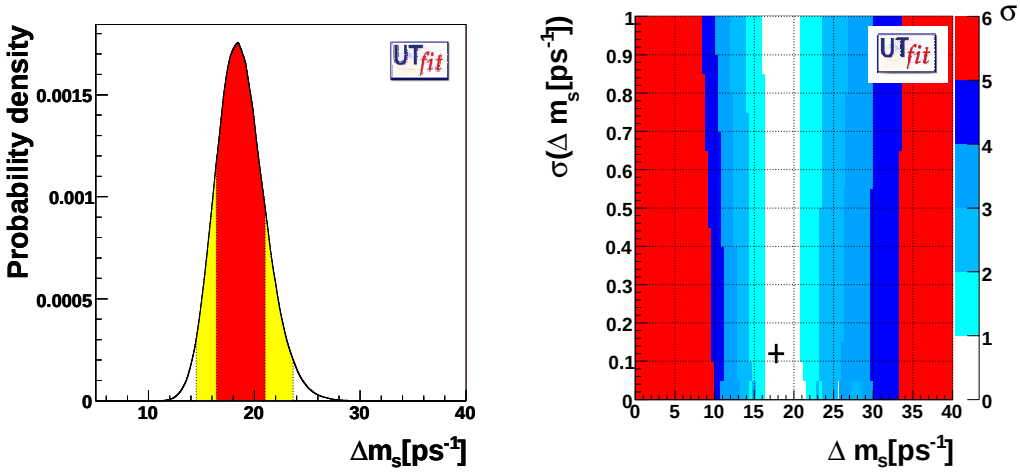

Figure 7.19: Left: P.D.F. of Δm_s from the global fit to all the UT quantities ([69]) but Δm_s itself. The red band represents the 68% CL, the yellow bands the 95% CL. Right: $\sigma(\Delta m_s)$ vs Δm_s distribution as obtained from the same fit. The experimental point represents the CDF measurement.

$\Delta m_s = 18.7 \pm 2.3\ ps^{-1}$. The agreement between this value and the CDF measurement is also depicted in the pull plot in Fig.7.19, where one can notice that the measured value is compatible with what is expected by the SM-only contributions well within 1σ.

The impact on the allowed values for $\bar{\rho}$ and $\bar{\eta}$ as costrained by the Δm_s measurement are also shown in Fig.7.20, along with the situation before the CDF measurement (2005).

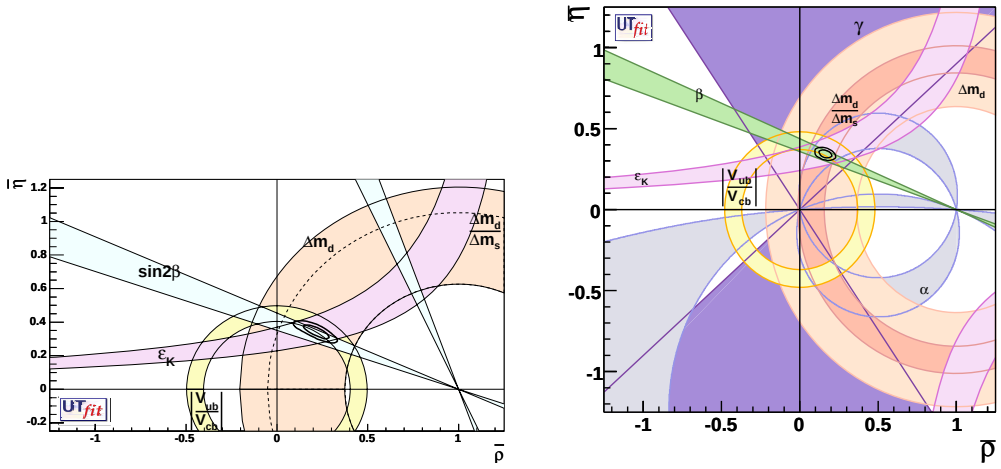

Figure 7.20: The fit to the UT by the **UT**Fit collaboration, with all the constraints used indicated with different colours. The 2005 (left) and present (right) constraints to $\bar{\rho}$ and $\bar{\eta}$ are indicated, from which the improvement by the Δm_s measurement is visible.

Provided this degree of agreement on Δm_s between the measured value and the SM predictions, they have also evaluated [70] the phase space left for New Physics contributions using the model-independent approach described in Sec.1.3. The CDF measurement of Δm_s provides the first constraint to the NP amplitude, and the value of the parameter is

$$C_{B_s} = 1.15 \pm 0.35. \tag{7.10}$$

As one can notice, the uncertainty is already smaller than in the thoroughly-explored B_d sector (see Eq.2.2). The recent improvements in the precision on the measurement of the lifetime difference $\Delta\Gamma_s$ [71], [72], also result in a constraint on both the B_s amplitude and phase, given by the (kilometric!) relation:

$$\frac{\Delta\Gamma_q}{\Delta m_q} = -2\frac{\kappa}{C_{B_q}}\left\{\cos\left(2\phi_{B_q}\right)\left(n_1 + \frac{n_6 B_2 + n_{11}}{B_1}\right) - \frac{\cos\left(\phi_q^{\mathrm{SM}} + 2\phi_{B_q}\right)}{R_t^q}\left(n_2 + \frac{n_7 B_2 + n_{12}}{B_1}\right) +$$

$$+\frac{\cos\left(2(\phi_q^{\rm SM}+\phi_{B_q})\right)}{R_t^{q^2}}\left(n_3+\frac{n_8B_2+n_{13}}{B_1}\right)+\cos\left(\phi_q^{\rm Pen}+2\phi_{B_q}\right)C_q^{\rm Pen}\cdot$$
$$\cdot\left(n_4+n_9\frac{B_2}{B_1}\right)-\cos\left(\phi_q^{\rm SM}+\phi_q^{\rm Pen}+2\phi_{B_q}\right)\frac{C_q^{\rm Pen}}{R_t^q}\left(n_5+n_{10}\frac{B_2}{B_1}\right)\Bigg\} \tag{7.11}$$

Here $\kappa = -2\pi m_b^2/(3M_W^2\eta_B S_0(x_t))$, B_i are the B parameters; n_i are explained in [10]; $R_t^q = |V_{tq}V_{tb}^*|/|V_{cq}V_{cb}^*|$; finally $C_q^{\rm Pen}$ and $\phi_q^{\rm Pen}$ parametrize possible NP contributions to $\Delta F = 1$ penguins. The inclusion of the lifetime difference, sensible to NP constributions, constrains the mixing phase to the interval is $\phi_{B_s} \in (-3 \pm 19) \bigcup (94 \pm 19)^\circ$. Fig.7.21 depicts the resulting P.D.F. for C_{B_s} vs ϕ_{B_s} and the phase space left for NP contributions as a function of the phase ϕ_{B_s}.

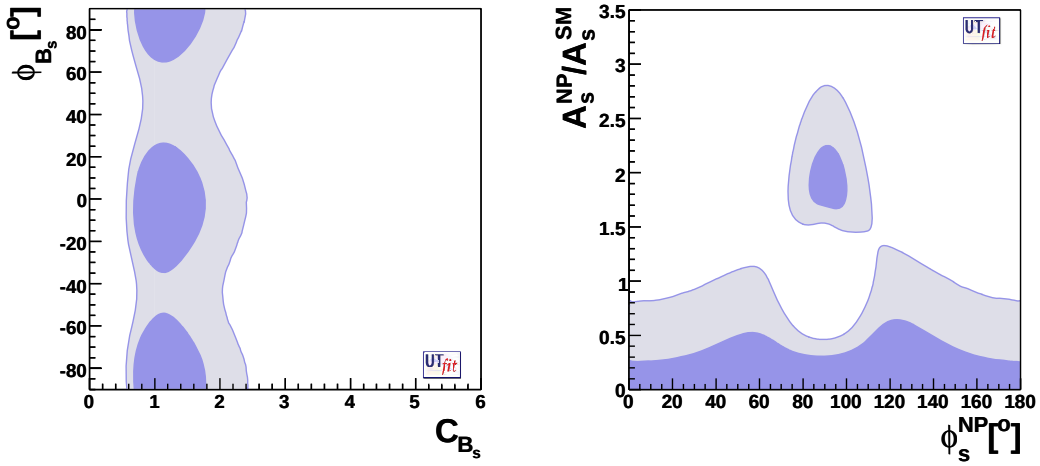

Figure 7.21: Left: The ϕ_{B_s}, C_{B_s} as constrained by the Δm_s and $\Delta\Gamma_s$, especially thank's to the Tevatron experiments. Right: A_s^{NP}/A_s^{SM} vs ϕ_{B_s}. For both plots, the darker region indicates the 68% C.L., the lighter one the 95% C.L.

While the value of C_{B_s} is already well compatible with the SM-only expectations, the phase $\phi_{B_s} = \beta_s^{SM} - \beta_s^{exp}$ still shows an ambiguity that allows for NP contributions to enter the picture. Experimentally, this can be probed by improving the $\Delta\Gamma_s$ measurement and looking for a value larger then the SM value, close to 0. This has been done recently by the DØ experiment, looking directly at the B_s lifetime difference in the $J/\psi\phi$ final state and at the CP-violating asymmetry in semileptonic decays

$$A_{SL} = \frac{N(b\bar{b} \to \ell^+\ell^+X) - N(b\bar{b} \to \ell^-\ell^-X)}{N(b\bar{b} \to \ell^+\ell^+X) + N(b\bar{b} \to \ell^-\ell^-X)}$$

and combining the two results with the (previous) CDF result on Δm_s [23], to give

$$\beta_s^{exp} = -0.56^{+0.44}_{-0.41}[73] \tag{7.12}$$

The mixing phase can also be measured both at the Tevatron and, with better sensitivity, by the LHCb experiment, from the time-dependent CP asymmetry in the decay of the B_s mesons to the CP eigenstate $J/\psi\phi$

$$\begin{aligned} a_{J/\psi\phi}(t) &= \frac{\mathcal{P}(B_s^0 \to J/\psi\phi) - \mathcal{P}(\bar{B}_s^0 \to J/\psi\phi)}{\mathcal{P}(B_s^0 \to J/\psi\phi) + \mathcal{P}(\bar{B}_s^0 \to J/\psi\phi)} \\ &= C_{J/\psi\phi} cos(\Delta m_s t) + S_{J/\psi\phi} sin(\Delta m_s t) \\ &\approx -2 sin(2\beta_s^{exp}) sin(\Delta m_s t) \end{aligned} \tag{7.13}$$

by neglecting the contribution of direct CP violation with respect to mixing-decay interference. This measurement requires knowledge of the flavour of the B_s at the creation (*tagged* measurement), such that the CDF flavour tagging system, to which this work contributes, may again play an important role in the process of constraining NP constributions in the flavour sector.

Conclusions

The focus of this work has been the development of an Opposite-side b flavour tagger which uses Kaons and, from the time dependent measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency Δm_s, the first observation of the particle-antiparticle transformations in neutral B_s mesons.

Using data collected by the Collider Detector at Fermilab from $p\bar{p}$ collisions at the Tevatron collider I have studied the feasibility of an algorithm to determine the flavour of a b hadron. This uses the correlation between the latter and the electric charge of the Kaon originated in the $b \to c \to s$, $b \to s$ and Cabibbo suppressed decays. The tagger has been used to infer the flavour of the neutral B_s meson in the observation of flavour oscillations originated by $\Delta F = 2$ hamiltonian operators at CDF. The flavour of the resulting B_s meson under study has been inferred from the flavour of the accompanying b hadron produced incoherently from $b\bar{b}$ pairs at the Tevatron.

At a hadron collider, the main source of background for a Kaon tagger is represented by tracks produced in the fragmentation process that originated also the b hadron and in the underlying interactions in the $p - \bar{p}$ collisions. The main contribution to this population at the momenta of interest ($|\vec{p}| \leq \mathcal{O}(m_B)$) is represented by pions. To select signal Kaons and thus retrieve the information on the b flavour I have made use of the CDF combined Particle Identification technique. It merges the measurement of a particle Time Of Flight and of its specific ionization in the central drift chamber to separate Kaons from the other charged particle species.

Moreover, the Kaon population in the generic sample I have considered is dominated by fragmentation tracks, associated to the event Primary Vertex. To reject those I have looked at candidates displaced from the PV in the plane transverse to the beam. For that, the track impact parameter divided by its uncertainty has been used. Encoding the information on the position of the track with respect to the b decay vertex has also proved to be beneficial. Infact, it enhances the signal in the tagging sample.

Finally I have also made use of the information on the flight direction of the jet associated to the Opposite-side b. This determines the quality of the Kaon tag, based on the corresponding dilution.

The tagger has been calibrated using an inclusive sample of b semileptonic

decays. A background subtraction procedure has been performed to enhance its b-purity. I have measured the performances of the Opposite Side Kaon Tagger on such events, that are reported in this work. I find

$$\begin{aligned} \varepsilon &= 18.11 \pm 0.07\ \% \\ D &= 9.9 \pm 0.9\ \% \end{aligned}$$

and, for the final figure-of-merit,

$$\varepsilon D^2 = 0.176 \pm 0.014\ (stat)\ \%\ ^{+0.001}_{-0.001}\ (syst)\ \%$$

The dilution is binned with variables for which signal Kaons and background behave differently. This results in a $\approx 30\%$ gain in εD^2 with respect to using the average dilution. The dilution is then:

$$D = 11.1 \pm 0.9\ \%$$

and for the tagger figure of merit:

$$\varepsilon D^2 = 0.229 \pm 0.016\ (stat)\ \%\ ^{+0.001}_{-0.001}\ (syst)\ \%$$

The first conclusion of this work is, then, that an Opposite Side Kaon Tagger is feasible at a hadron collider. I find a dilution of the same order of the other Opposite-side taggers at CDF. In particular, the best tags have a dilution comparable with that of the best tags of the Jet Charge tagger, which provides the largest relative contribution to the overall OST εD^2.

This result has also been analyzed in terms of most useful variables and limiting factors. The initial Monte Carlo simulations for Run II suggested that the OSKT effectiveness would be a leading force in the CDF II flavour tagging system, thanks mainly to the Time-Of-Flight detector and the silicon Layer 00 used for vertexing. In this work I have shown that a significant reduction of such expectations is observed on data, due to the large track multiplicity from underlying and fragmentation processes in $p\bar{p}$ collisions which the CDF II Monte Carlo simulation could not reproduce completely. Flavour tagging with Kaons will still play an important role in time-dependent analyses at future colliders (e.g. flavour mixing and CP asymmetry in tagged $B_s \to D_s K$ decays at LHCb). A trustful simulation of physics processes and occupancy-prove pattern recognition and PID algorithms will be necessary and are the main ingredients for a reliable prediction on the OSKT effectiveness.

The second part of this work is devoted to describe the observation of the $B_s^0 - \overline{B}_s^0$ flavour oscillations at CDF and the measurement of the corresponding frequency Δm_s. CDF observes a mixing signal with a statistical significance of 5.4 σ. Such significance is contributed by three components: signal-to-background ratio; proper time uncertainty; flavour tagging effectiveness. Although reconstructed,

signal events without a flavour tag are useless for a Δm_s measurement. CDF has used different flavour taggers, that have been described in the text. Four types of tags are present for the Δm_s analysis. The Opposite Side Kaon Tagger adds to the lepton and jet charge taggers and is combined with them. Its use and the combination of the OSKT with the other Opposite-side taggers have been shown in this work to represent 20% of the total effectiveness of the CDF Opposite Side taggers. I have shown that the OSKT contributes both standalone and in the enhancement of the lepton and jet tag dilution by means of their correlations. I have also proved that agreement between signal Kaons and leptons provides the largest contribution to the increase of the lepton tags in the NN.

Therefore, the relevance of this work is the increase in the CDF sensitivity to B_s flavour oscillations. It contributed to make it possible for CDF to perform the observation of this phenomenon after several decades of attempts at various colliders. On one side, such a result represents the discovery of a quantum phenomenon expected in the B_s sector, after it had been observed in the B_d system. On the other hand, it provides a direct information on the amount of violation of the CP symmetry in nature. Its importance resides, finally, also in its capability to provide hints of new physics scenarios beyond the Standard Model in the flavour sector.

In conclusion, the observation of the B_s mixing achieved at CDF is reported in this work. The OSKT has been used for this result and the contribution of this thesis work has been found to have an important impact in enhancing the statistical significance of the CDF measurement.

CDF measures

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

obtained with a significance of 5.4 σ. This allows to infer the ratio $|V_{td}/V_{ts}|$ of the amplitudes of the CKM matrix elements:

$$|V_{td}/V_{ts}| = 0.2060 \pm 0.0007\ (exp)\ {}^{+0.0081}_{-0.0060}\ (theo)$$

Given the precise measurement, the uncertainty on this quantity is already fully driven by the theoretical calculations. This constraints the values of the parameters of the CKM matrix, ρ and η, reducing the uncertainty on these parameters by $\approx 15\%$.

Appendix A

How does a Neural Network work

An artificial neural network (NN) is an information processing algorithm mimicking the biological neural systems.
Physical quantities are passed to the network through a certain number of input nodes; these are used by the NN to discriminate a signal from a background, for which it can be used as a practical alternative to constructing a likelihood from many correlated variables. This is the main aim for which a NN is used in high-energy physics.
After the input nodes, some *hidden* layers (that is not immediately accessible/settable by the user) are activated by the incoming information, as shown in fig.A.1. In practice, each j-th "neuron" (= node) in a hidden layer receives a

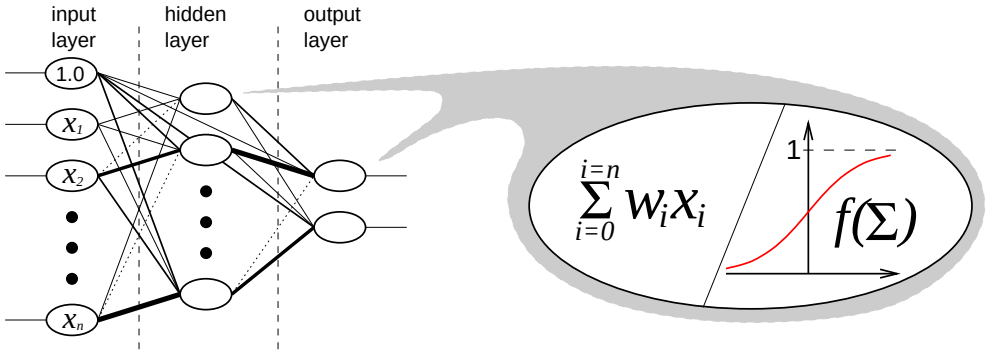

Figure A.1: An example structure of a NN with a single hidden layer and 2 output nodes. The thickness and the type of a line represent the weight with which the output of a node is carried forward. On the right, a zoomed node is shown, with its activation function.

certain number of inputs x_{ij}, with $i = 1, 2, 3, ..., n$ and calculates a weighted sum

$$X_j = c_j + \sum_{i=0}^{n} w_{ij} \cdot x_{ij} \tag{A.1}$$

where w_{ij} are the weights assigned to the i-th quantity entering the j-th node and the c is a given bias, independent of the variables.
The range of variation of the x_{ij} is, in general, $[-\infty, \infty]$: to obtain a finite output, the X is translated by each node into a sigmoid function

$$S(X) = \frac{1}{1 + e^{-X}} \tag{A.2}$$

whose shape is shown again in fig.A.1. As a matter of fact, this is a space transformation

$$[-\infty, \infty] \rightarrow [0, 1] \tag{A.3}$$

where 0 is usually for the background and 1 for the signal (similar to a likelihood). Eventually, a NN made up of n input nodes and m hidden nodes will generate an output number O:

$$O = S(\sum_{j=0}^{m} f_j \cdot X_j) \tag{A.4}$$

Fig.A.2 shows an example output for a sample of background and of signal. The

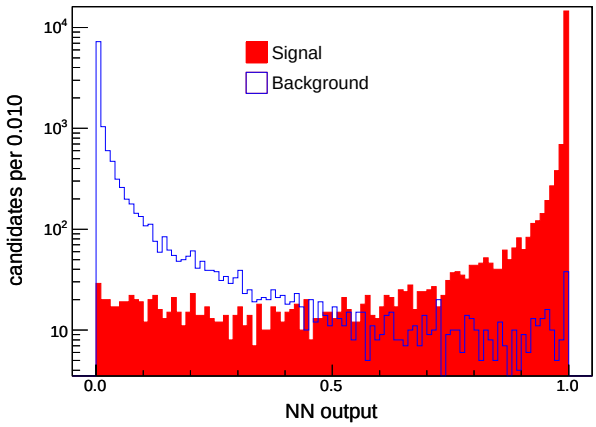

Figure A.2: An example output of a NN, with signal (red) tending to 1 and background (blue) to 0 by construction of the sigmoid eq.A.2

optimal values for the weights and biases are found in a process called *training*. Given a number of events tagged as of k different kinds (e.g. MC simulation of a physical process and its backgrounds), the training consists of passing the discriminating quantities onto the network, with random starting values for the various parameters. Different iterations are advisable, through which the NN minimizes the function

$$\Delta = \sum_{evts} W_k [O - \overline{O}]^2 \tag{A.5}$$

where the weights W_k are associated to each different k-th sample and the $\overline{O}$ is the corresponding expected value (e.g. 1 for the signal MC sample). The values

for the final parameters in A.4 are obtained from the last training iteration, established as to avoid that the NN gets overtrained on that particular samples (i.e. learns the peculiar features of *that* ensemble of events), loosing its generalization.

The motivation to use a NN instead a Likelihood Ratio to discriminate a signal from background lies is both a:

- conceptual reason: when using 2 or more input physics variables, the LR will not consider their *correlations*, being just built from the *product* of the different one-dimensional P.D.F.'s; the NN accounts for such correlations automatically, instead: the weights in the sigmoid are adjusted starting from a whole sample of different events;

- practical reason: even if one wanted to introduce the multi-dimensional distributions of the different variables, the larger is the number of inputs, the more difficult it is to construct a LR.

Acknowledgements

I would like to thank the CDF B physics group and the CDF B_s mixing group for the support, the interest and the suggestions while this work was carried on. All this, and especially the last months, has been an extremely exciting time to stay at Fermilab and take part in the activities, meetings and crucial moments that lead to the important discovery CDF has performed.

Il mio ringraziamento più grande va a Marco (Rescigno), per la sua guida, la sua presenza costante e la sua amicizia, che hanno fatto di questa tesi un periodo di lavoro veramente straordinario. Davvero gran parte di quel (poco) che so lo devo a lui. Insieme a lui vorrei ringraziare Carlo (Dionisi), Stefano (Giagu) e il gruppo di Roma, per le discussioni e (le ore di) meeting con una visione critica e attenta di questo lavoro. E anche per il supporto *del* morale. E a Paolone per il supporto logistico, la cucina e il rassettamento di casa ad 11 Potawatomi, durante il blessing.

Grazie a tutti quelli che ho incontrato e ai miei amici di sempre, grazie ai quali ho trovato casa ovunque sia stato. A Roma, dove hanno pazientato sempre cercando di capire quando sarei tornato; e a Fermilab, dove gli ultimi tempi sono stati tra i più divertenti e fantastici che ricordi, grazie ad un insieme di persone che una strana alchimia ha reso davvero difficile da ritrovare altrove. Grazie Paolo, Anna, Gian Piero, Roberto, Peter e Julia, Stephane, Kai, Mia, Elena, er Nardecchia, er Pannarale, Chiara e Michele. Grazie *Piccioncini*, Valentina, Simona, Sebastiano, Marcone e Margherita (anche se un po' da lontano), Francesco. Grazie a quelli che ho visto ogni giorno, e con i quali ho condiviso un sacco di sere e la magia della musica in sala prove. Grazie ai miei *Gauge*. E alla saletta, la mia theory division. Grazie ai grandi fisici sparsi in giro: il Bomben, Matteo, Stefano.

Grazie a France, per tutto quello che abbiamo vissuto assieme in questi anni.

Grazie soprattutto ai miei. Di tutto.

Bibliography

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 631 (1963)

[2] M. Kobayashi and T.Maskawa, Prog. Theor. Phys.**49**, 652 (1973)

[3] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983)

[4] E. Blucher and W. J. Marciano, review in [21]

[5] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985)

[6] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981), *Erratum-ibid.* **65** 1772 (1981)

[7] A Ali and D. London, hep-ph/0002167

[8] M. Drees, hep-ph/9611409

[9] A. J. Buras, P. H. Chankowski, J. Rosiek, L. Slawianowska, Nucl. Phys. **B619** 434-466 (2001)

[10] M. Bona *et al.* (UTFit Collaboration), JHEP **0603**, 080 (2006)

[11] D. Buskulic *et al.*, The ALEPH Collaboration, Z. Phys. **C75**, 397 (1997)

[12] P. Abreu *et al.*, The DELPHI Collaboration, Z. Phys. **C76**, 579 (1997)

[13] B. Aubert *et al.*, The BABAR Collaboration, Phys. Rev. **D66**, 032003 (2002)

[14] B. Aubert *et al.*, The BABAR Collaboration, Phys. Rev. **D67**, 072002 (1997)

[15] N. C. Hastings *et al.*, The BELLE Collaboration, Phys. Rev. **C67**, 092004 (1997)

[16] K. Abe *et al.*, The BELLE Collaboration, Phys. Rev. **D71**, 072003 (2005)

[17] F. Abe *et al.*, The CDF Collaboration, Phys. Rev. **D59**, 032001 (1999)

[18] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/060406.blessed-semi_B0mix/

[19] The DØ Collaboration, DØ Note 5029

[20] H.G. Moser and A. Roussarie, Nucl. Instr. Methods Phys. Res. **A384**, 491 (1997)

[21] W.-M. Yao *et al.*, The Particle Data Group, J. Phys. **G33**, 1 (2006)

[22] V. M Abazov *et al.*, The DØ Collaboration, Phys. Rev. Lett. **97**, 021802 (2006)

[23] A. Abulencia *et al.*, The CDF Collaboration, Phys. Rev. Lett. **97**, 062003 (2006)

[24] Fermilab Beams Division,
http://www-bdnew.fnal.gov/pbar/documents/PBAR_Rookie_Book.PDF
(1999)

[25] D. Amidei *et al.*, Nucl. Instr. Methods Phys. Res. **A350**, 73 (1994)

[26] W. M. Yao and K. Bloom, CDF Note 5991

[27] S. D'Auria *et al.*, CDF Note 6391

[28] G. Salamanna, CDF Note 6810

[29] R. Carosi *et al.*, CDF Note 7488

[30] R. Carosi *et al.*, CDF Note 7866

[31] J. Allison *et al.*, Nucl. Sci. **53**, 1 (2006)

[32] M. Jones *et al.*, CDF Note 6480

[33] T. Sjöstrand *et al.*, Computer Phys. Commun. 135 (2001) 238

[34] D. J. Lange, Nucl, Instr. Meth. A, Vol. 462, Issue 1-2, p.152-155.

[35] P. Sphicas, CDF Note 2655

[36] K. Anikeev *et al.*, CDF Note 7999

[37] The Heavy Flavor Averaging Group, http://www.slac.stanford.edu/xorg/hfag

[38] E. Ben-Haim *et al.*, Phys. Lett. B **580**, 108 (2004)

[39] F. Abe *et al.*, The CDF Collaboration, Phys. Rev. **D60**, 051101 (1999)

[40] The B_s Mixing Group, CDF Note 7482

[41] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/040812.blessed-jqt-lsvt/

[42] The B_s Mixing Group, CDF Note 7979

[43] The CDF Collaboration, CDF Note 8206

[44] The CDF Collaboration, CDF Note 8344

[45] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/081204.blessed-like-el-tag/

[46] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/072204.blessed-like-mu-tag/

[47] D. Usynin *et al.*, CDF Note 6483

[48] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/030320.blessed-bb-corr/

[49] I. Kravchenko *et al.*, CDF Note 6951

[50] G. Giurgiu *et al.*, CDF Note 7043

[51] M. Jones *et al.*, CDF Note 4695

[52] The B_s Mixing Group, CDF Note 8169

[53] H. Budell and F. Toilet, Anal. Part. Phys. **666**, 1 (1979)

[54] D. Acosta *et al.*, The CDF Collaboration, Phys. Rev. **D71**, 052003 (2005)

[55] H. Kakuno *et al.*, hep-ex/0403022 v1 (2004)

[56] C. Grozis *et al.*, Nucl. Phys. Proc. Suppl. **120**, 219 (2003)

[57] J. Babel, M. Calvi, O. Leroy, M. Musy, internal LHCb talk, August 30th, 2005

[58] R. Forty and O. Schneider, LHCb Note LHCB/98-040

[59] The CDF Collaboration,
http://www-cdf.fnal.gov/physics/new/bottom/050804.hadr_B0mix/

[60] The B_s Mixing Group, CDF Note 8172

[61] Phi-T Physics Information Technologies, www.phi-t.de

[62] D. B. MacFarlane, Contributed to 30th SLAC Summer Institute on Particle Physics: Secrets of the B Meson (2002)

[63] The B_s Mixing Group, CDF Note 8333

[64] The B_s Mixing Group, CDF Note 8313

[65] The B_s Mixing Group, CDF Note 8314

[66] The CDF Collaboration, CDF Note 7944

[67] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 202001 (2006)

[68] M. Okamoto, PoS LAT2005 (2005) 013 (hep-lat/0510113)

[69] M. Bona *et al.* (UTFit Collaboration), `http://www.utfit.roma1.infn.it`

[70] M. Bona *et al.* (UTFit Collaboration), hep-ph/0605213

[71] The CDF Collaboration, `http://www-cdf.fnal.gov/physics/new/bottom/060126.blessed-BsKK_lifetime/`

[72] The DØ Collaboration, DØ Conference Note 5052

[73] The DØ Collaboration, DØ Conference Note 5189