Scuola Normale Superiore di Pisa

CLASSE DI SCIENZE

Corso di Perfezionamento in Fisica

PH. D. THESIS

First observation of the $B_s^0 \to K^+K^-$ decay mode, and measurement of the B^0 and B_s^0 mesons decay-rates into two-body, charmless final states at CDF

Diego Tonelli

Advisor:
Prof. Giovanni Punzi

30 Novembre 2006

Abstract

We searched for decays of the type $B^0_{(s)} \to h^+ h^{'-}$ (where $h,\ h' = K$ or π) in a sample corresponding to 180 pb^{-1} of p$\bar{\text{p}}$ collisions at $\sqrt{s} = 1.96$ TeV, collected by the upgraded Collider Detector at the Fermilab Tevatron. A total signal of approximately 900 events was reconstructed, and the relative branching fractions ($\mathcal{B}$) of each decay mode were determined with a likelihood fit. We report the first observation of the decay mode $B^0_s \to K^+K^-$, with a yield of 236 ± 32 events, corresponding to a relative branching fraction

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.46 \pm 0.08\ (stat.) \pm 0.07\ (syst.),$$

where f_s/f_d is the ratio of production fractions of B^0_s and B^0 mesons from the hadronization of a b-quark in p$\bar{\text{p}}$ collisions. We measured the direct CP-violating asymmetry in the $B^0 \to K^+\pi^-$ decay,

$$A_{\mathsf{CP}}(B^0 \to K^+\pi^-) \equiv \frac{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) - \mathcal{B}(B^0 \to K^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) + \mathcal{B}(B^0 \to K^+\pi^-)} = -0.013 \pm 0.078\ (stat.) \pm 0.012\ (syst.),$$

and the ratio

$$\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.21 \pm 0.05\ (stat.) \pm 0.03\ (syst.),$$

in agreement with world average values. These measurements are performed for the first time in a hadron collider. No evidence for other modes was found, and we set the following upper limits at the 90% confidence level (CL):

$$\frac{\mathcal{B}(B^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} < 0.10, \quad \frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s \to K^-\pi^+)}{\mathcal{B}(B^0 \to K^+\pi^-)} < 0.08, \quad \text{and} \quad \frac{\mathcal{B}(B^0_s \to \pi^+\pi^-)}{\mathcal{B}(B^0_s \to K^+K^-)} < 0.05.$$

By normalizing the relative decay-rates to the world-average values of the $B^0 \to K^+\pi^-$ branching fraction and of the ratio f_s/f_d, we obtain the following absolute branching fractions:

$$\begin{aligned}
\mathcal{B}(B^0_s \to K^+K^-) &= [33 \pm 6\ (stat.) \pm 7\ (syst.)] \times 10^{-6}, \\
\mathcal{B}(B^0 \to \pi^+\pi^-) &= [3.9 \pm 1.0\ (stat.) \pm 0.6\ (syst.)] \times 10^{-6}, \\
\mathcal{B}(B^0 \to K^+K^-) &< 1.8 \times 10^{-6} \quad \text{at 90\% CL}, \\
\mathcal{B}(B^0_s \to K^-\pi^+) &< 5.6 \times 10^{-6} \quad \text{at 90\% CL, and} \\
\mathcal{B}(B^0_s \to \pi^+\pi^-) &< 1.7 \times 10^{-6} \quad \text{at 90\% CL.}
\end{aligned}$$

This analysis has been published in Physical Review Letters [1].

Contents

List of Figures

List of Tables

Introduction

The decays of b-hadrons are a fertile ground to investigate the flavor sector of the Standard Model and to look for first signals for new physics in the years preceding the Large Hadron Collider operations. In particular, the phenomenology of non-leptonic charmless two-body decays of b-mesons offers rich opportunities for increasing our understanding of the CP violation, i. e., the lack of symmetry of physical processes when all spatial coordinates are inverted and particles are replaced by their antiparticles. The precise measurements obtained recently at dedicated e^+e^- colliders already provided demonstration of the central role of these decay modes. To mention just one example, the first evidence of *direct* CP violation in decays of particles other than kaons was obtained from the measurement of partial rate-asymmetries between $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ decays [2].

Charmless hadronic b-meson decays proceed through an unique interplay of electroweak and low-energy strong interactions, allowing observation of CP-violating effects within and beyond the Standard Model. The problem is that, currently, no completely reliable theoretical prediction of decay rates is available, for most of them, because of the presence of strong interactions in non-perturbative regime, which introduce significant uncertainties in the predicted amplitudes. Several phenomenological models provide different predictions, none of them properly accounting for all observed decay-rates. This makes the interpretation of experimental observations difficult, since any discrepancy between predictions and measurements may be ascribed either to improper treatment of hadronic uncertainties, or to contributions of amplitudes not expected in the Standard Model.

Constraining quark-mixing (i. e., Cabibbo-Kobayashi-Maskawa, CKM) parameters from charmless hadronic b-meson decays requires, therefore, methods for addressing hadronic uncertainties. A variety of techniques has been proposed to overcome the difficulty [3, 4, 5]. One that recently received a lot of attention consists in combining the information from measurements of rates of $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ decays to *control* the effect of hadronic uncertainties to directly determine the phase of the V_{ub} element of the quark-mixing matrix (angle γ) or, alternatively, to test our understanding of dynamics of b-hadron decays, when compared with other determinations of γ. This method relies on the expected (partial) cancellation of hadronic amplitudes between decay-modes related by U-spin symmetry, a subgroup of the flavor SU(3) symmetry under which d quarks transform into s quarks. This symmetry is not exactly conserved in the Standard Model, introducing a residual uncertainty in the extraction of the CKM phases. However, the magnitude of the violation can be constrained using a combination of QCD calculations and experimental observables from other similar decays, and it is not expected to be large. This allows its treatment as a correction to the results obtained in the limit of exact validity. The large variety of two-body charmless decay modes supplies useful information

both on the weak phase and on the needed correction, provided that an experimental access to these modes is available.

While rich experimental data are available for B^0 and B^+ mesons from the ARGUS, CLEO, and LEP experiments, and, more recently, from the *BABAR* and Belle experiments, no hadronic charmless decays of B_s^0 mesons into pions or kaons has yet been observed.[1]

The upgraded Collider Detector at the Fermilab Tevatron (CDF II) is an ideal environment for studying hadronic charmless decays of B_s^0 (along with B^0 and B^+) mesons into charged final states. In this thesis, I describe a method developed for reconstructing such decays in the peculiar environment of a hadron collider, and the results of the analysis of the first sample, corresponding to $\int \mathcal{L}dt = 180\,\text{pb}^{-1}$, collected by the CDF experiment in the current period of operation.

The thesis is organized as follows. *Chapter 1* provides a concise description of the theory of non-leptonic b-hadron decays within the framework of the Standard Model, emphasizing the aspect related to the determination of CKM parameters. It includes a simplified and non-exhaustive discussion on the main phenomenological models currently employed in estimating non-perturbative strong contributions to the transition amplitudes. The first chapter contains also a short review of the current experimental situation and an outline of the analysis.

Chapter 2 describes the experimental apparatus. It contains a general description of the accelerator and of the CDF II detector. The subdetectors that reconstruct charged-particle trajectories and the trigger are described in greater detail, being the aspects of the detector more specific to the present analysis.

The remaining chapters cover the description of the data analysis I carried out for this thesis. Most of the discussions are self-contained, but certain issues are so inter-correlated that a few forward references could not be avoided. I tried to limit the technical descriptions to the minimum needed for a convincing description. Whenever possible, I moved them to the appendices.

Chapter 3 presents the data set used for the measurement and details the extraction of the signal from the background.

Chapter 4 discusses how the information from kinematics and particle identification was combined to achieve a statistical discrimination among the individual modes contributing to the signal, needed for the desired measurements of branching fractions.

The measurement of specific ionization induced by the passage of charged particles in the CDF II drift chamber (dE/dx) is an essential discriminant to infer the composition of the sample. *Chapter 5* contains a detailed description of the calibration of the dE/dx measurement and of the characterization of its performance, both necessary for this measurement.

Chapter 6 details the structure of the likelihood function used to fit the composition of the sample.

The results of the fit are presented in *chap. 7*.

The next chapter, *chap. 8*, details the evaluation of trigger and analysis efficiencies necessary to infer the desired measurements of branching fractions from the observed numbers of signal events in the sample.

Chapter 9 contains the discussion of systematic effects and their associated uncertainties.

[1] ARGUS stands for "A Russian German US" Collaboration, LEP for Large Electron Positron collider.

The final results of the measurement and their interpretation are discussed in *chap. 10*.

When the distinction between the different decay modes ($B^0 \to \pi^+\pi^-$, $B_s^0 \to K^-\pi^+$, $B_s^0 \to K^+K^-$, etc.) is not relevant, the expression $B^0_{(s)} \to h^+h^{'-}$ (h, $h' = \pi$ or K) is used to collectively indicate all of them. Unless otherwise stated, throughout this thesis, C-conjugate modes are implied and branching fractions ($\mathcal{B}$) indicate CP-averages, that is,

$$\mathcal{B}(B \to f) \equiv \frac{\Gamma(B \to f) + \Gamma(\overline{B} \to \overline{f})}{2\Gamma_B},$$

where $\Gamma(B \to f)$ is the partial decay-width of a particle B into a final state f, Γ_B represents the natural width, and over-lined quantities indicate CP-conjugated states.

Chapter 1

Two-body charmless $B^0_{(s)}$ decays into charged kaons and pions

This chapter contains a concise discussion on the theoretical and phenomenological framework associated to the study of strange and non-strange b-meson decays into pairs of charged pions and kaons ($B^0_{(s)} \to h^+h'^-$). The basic concepts and formalism of CP *violation in the Standard Model are introduced. Some relevant phenomenological models for predicting amplitudes of non-leptonic decays of b-mesons are illustrated, discussing their role in providing constraining information on both the* CP *violation mechanism and the low-energy strong dynamics involved in these processes. Lastly, the experimental part of this thesis is introduced: after a brief comparison between the different experimental environments exploited to study b-meson decays, we summarize the current experimental knowledge and we outline the structure of the data analysis described in the rest of the thesis.*

1.1 Flavor physics and CP violation

1.1.1 Introduction

In 1964, the observation of the decay of long-lived neutral K mesons in both two- and three-pion states showed that not all natural phenomena are symmetric under the mirror reversal of their spatial arrangement combined with the replacement of all involved particles with the corresponding antiparticles [6]. The observation of a $\mathcal{O}(10^{-3})$ branching fraction for the $K^0_L \to \pi^+\pi^-$ process was the discovery of the non conservation of the CP symmetry, i. e., the CP violation.

In particular, this is the manifestation of *indirect* CP violation, which originates from the fact that the mass eigenstates $K^0_{L,S}$ of the neutral kaon system are not eigenstates of the CP operator causing the small CP-even component of the K^0_L state to decay, through CP-conserving interactions, into the $\pi^+\pi^-$ final state.

It was only in 1999, after a search several decades long, that an alternative manifestation of the phenomenon, the *direct* CP violation, was established, again using neutral kaons [7]. The direct CP violation occurs in first order transitions, arising directly at the decay-amplitude level. Its observation

ruled-out the "super-weak" theory, proposed by Wolfenstein readily after the 1964 observation [8], and confirmed that the CP violation is an universal property of the weak interaction, thus being expected to appear in other phenomena governed by this interaction. In this respect, a large theoretical and experimental effort has been dedicated since the 1980s to further extend the exploration of the CP violation, following the first indications of Sanda, Carter, and Bigi on the the opportunity to detect CP violation in b-meson decays [9]. These have different features with respect to kaon decays, both from an experimental and theoretical point of view.

The b-mesons contain the b-quark, which belongs to the third quark generation and can directly decay into a first or a second generation quark; thus, the effects of three-quark generation are accessible already in lowest-order processes, yielding CP violation effects typically larger than in kaons. The larger mass of the b-quark with respect to the s-quark provides a two-fold advantage. From the theoretical standpoint, useful approximations in QCD calculations of (some) amplitudes allow more reliable predictions than in kaon decays. On the other hand, many decay modes are kinematically available, allowing multiple experimental access to CP-violating observables. However, the multiplicity of available channels results in small branching fractions for individual processes and the consequent need for high statistic samples. Also, the mixing properties are more difficult to extract, because physical $B^0_{(s)}$ meson states have similar lifetimes, as opposed to the kaon case. It is therefore convenient to use the flavor-eigenstates base, i. e., to study states whose flavor eigenstate at production is known. This further reduces the size of available samples, especially in the presence of inclusive incoherent $b\bar{b}$ production, as in hadron-hadron collisions.

The usual convention is to define as b-mesons the color-neutral bound states of a $\bar{b}$ valence antiquark and a valence quark:

$$B^0 \sim d\bar{b}, \quad B^0_s \sim s\bar{b}, \quad B^+ \sim u\bar{b}, \quad \text{and} \quad B^+_c \sim c\bar{b}. \tag{1.1}$$

Dedicated detectors installed at the KEKB and at the upgraded Positron Electron Project (PEP-II) e^+e^- colliders, operating at the center-of-mass energy (also referred to as the Mandelstam variable $\sqrt{s}$) corresponding to the $\Upsilon(4S)$ resonance with asymmetric beam energies (the so-called B-Factories), provided a great wealth of measurements related to B^0 and B^+ decays, with impressive precision. In 2001, the first observation of CP violation in decays of particles other than kaons was announced by the B-Factories: a time-dependent $B^0 \to J/\psi Ks$ decay-rate asymmetry was measured, due to the interference between the amplitude of the decay occurred after $B^0 - \overline{B}^0$ flavor mixing and the amplitude of the direct decay [10]. This was the observation of "mixing-induced" CP violation in the b-hadron system. In addition, in 2004, direct CP violation was also observed in $B^0 \to K^+\pi^-$ decays, thereby complementing the 1999 observation in the kaon system [2].

Understanding CP violation in Nature has fundamental implications related to topics as diverse as the microscopic time-reversibility of physical laws, or the origin of the baryonic asymmetry of the universe. In fact, one of the Sakharov conditions for explaining the $\mathcal{O}(10^{-10})$ cosmological baryon asymmetry is that elementary interactions do not conserve CP (and C) [11]. Interestingly, the estimates of the baryon asymmetry suggest that the magnitude of CP violation currently detected in the Standard Model is insufficient to cause the observed asymmetries, indicating the need for additional sources of CP violation. Studying the flavor sector of the Standard Model is a promising way to obtain information on new physics, through the virtual contributions that new particles may give in the transitions of Standard Model particles, thereby yielding abnormal amplitudes of processes. This

"indirect" approach for studying new physics complements the straightforward "discovery approach", consisting in producing new resonances in collisions of increasingly high energy. It generally probes higher energy scales (since non-Standard Model particles can participate as highly-virtual off-shell states) thereby giving earlier indication of new physics. Nearly all models of physics beyond the Standard Model imply new sources of flavor and of CP violation.

To recognize possible manifestations of CP violation beyond the Standard Model, it is necessary to first understand the Standard Model picture of flavor dynamics, governed by the quark-mixing (or CKM) matrix. The experimental phenomenology is extremely rich: the high number of *strange*, *beauty*, and *charmed* hadron decays accessible provides a wealth of CP-violating observables. The key issue is the theoretical interpretation of the results. Since the CP-violating effects are typically small, accurate predictions of the effects due to exchange of soft gluons between quarks are required. But the current understanding of QCD in non-perturbative regime does not allow to reach the desired accuracy; thus the predictions suffer from "hadronic" uncertainties which make it difficult to extract the short-distance CKM-related observables from the measured quantities.

1.1.2 CP violation in the Standard Model

In 1973, almost then years since its first observation, Kobayashi and Maskawa proposed an interpretation of the CP-violation which will be later established as an essential part of the Standard Model [12]. They extended the quark-mixing formalism, due to Cabibbo [13] and Glashow-Iliopoulos-Maiani [14], from two to three quark generations, and suggested that a physical phase in the associate quark-mixing matrix could have explained the observed phenomenon. The completion of the quark families, with the observation of the *top* quark in 1995, and the current level of experimental knowledge on CP violation confirm, so far, the Kobayashi-Maskawa theory.

The Standard Model with three generations of quarks does *not automatically* imply conservation of CP. It can accommodate the observed CP violation since it has a structure complex enough to support the existence of a physical weak phase in the CKM matrix. In the following, we assume that the combined operation CPT, in which T is the time-reversal operation, is a valid symmetry. This is a general property of Lorentz-invariant quantum field theories based on Hermitian local Lagrangian densities and was never found to be violated [15]. In addition, since no CP violation in processes governed by the strong or electromagnetic interactions has ever been detected, we assume that this phenomenon arises only in weak interactions. The flavor sector of the Standard Model is based on the spontaneously broken gauge group

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \longrightarrow \mathrm{U}(1)_{\mathrm{em}}. \tag{1.2}$$

The only structure of the Standard Model from which CP violation may originate is the charged-current interaction of quarks

$$d_i \longrightarrow u_j W^-, \tag{1.3}$$

where $d_i = (d, s, b)$ denote down-type quarks, $u_j = (u, c, t)$ up-type quarks and W^- is the $\mathrm{SU}(2)_L$ gauge boson. The set of coupling-magnitudes $V_{u_j d_i}$ are fundamental Standard Model parameters, usually represented in matricial form: the $\boldsymbol{V_{\mathbf{CKM}}}$ (Cabibbo-Kobayashi-Maskawa, or quark-mixing)

matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.4}$$

This couples the physical quarks (mass eigenstates, d, s, b) to the states that participate in the charged-current weak interaction (flavor eigenstates, d', s', b') through the above *unitary* transformation. As a consequence, the Lagrangian density of the non-leptonic charged current interaction is written as

$$L_{\mathrm{CC-int}} = -\frac{g_2}{\sqrt{2}} \left(\bar{u}_L, \bar{c}_L, \bar{t}_L\right) V_{\mathrm{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W^{\dagger}_{\mu} \quad + \quad \text{hermitian conjugates} \tag{1.5}$$

in terms of mass-eigenstates. The gauge coupling g_2 is related to the $\mathrm{SU}(2)_L$ gauge group and the $W^{\dagger}_{\mu}$ field describes the charged W bosons. If a CP transformation is applied to relation (1.3) one obtains

$$V_{u_j d_i} \longrightarrow V^*_{u_j d_i}. \tag{1.6}$$

Since $V_{u_j d_i} = V^*_{u_j d_i}$ (i. e., the interaction is CP-invariant) only if $V_{u_j d_i}$ is real, CP violation may be accommodated in the Standard Model through complex phases in the CKM matrix.

Quark fields may be redefined according to phase transformations of the type

$$u_j \rightarrow e^{i\xi_u} u_j \quad \text{and} \quad d_i \rightarrow e^{i\xi_d} d_i. \tag{1.7}$$

The invariance of the charged-current interaction Lagrangian density under such transformations implies

$$V_{u_j d_i} \rightarrow e^{i\xi_u} V_{u_j d_i} e^{-i\xi_d} d_i. \tag{1.8}$$

These transformations may be used to eliminate unphysical phases from any $N \times N$ unitary matrix, thus reducing its parameterization to a set of $\frac{1}{2}N(N-1)$ Euler angles and $\frac{1}{2}(N-1)(N-2)$ complex phases.[1] With $N = 3$ quark generations this translates into three Euler angles and a single *physical*, observable phase. The degree of freedom represented by this irreducible complex phase permits integration of CP violation into the Standard Model [12]. However, given the *real* nature of physical observables, CP violation can be experimentally detected only when measurable quantities are affected by the quantum-mechanical interference of multiple terms which have relative phases different from 0 and π. In order to obtain observable CP violation effects within the framework of the Standard Model, some additional conditions need to be satisfied. A first theoretical condition prescribes non-degenerate quark generations (quarks with same charge must have different masses), otherwise the CP-violating phase of the CKM matrix could be removed through an appropriate unitary transformation of the quarks fields. The condition is experimental and states that the magnitude of the CP violation has to be sufficient to generate observable phenomena. These conditions are summarized in the following relation

$$(m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)(m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2) \times J_{\mathsf{CP}} \neq 0, \tag{1.9}$$

[1] Any $N \times N$ complex matrix has $2N^2$ parameters. Unitarity provides $\frac{1}{2}N(N-1)$ complex and N real constraints. This leaves $2N^2 - N(N-1) - N = N^2$ parameters of which $2N-1$ can be removed by redefining the corresponding number of quark phases, yielding $(N-1)^2$ parameters.

where m_i are masses of quarks, and

$$J_{\mathsf{CP}} = |\Im(V_{i\alpha}V_{j\beta}V_{i\beta}^*V_{j\alpha}^*)|, \quad i \neq j, \alpha \neq \beta, \tag{1.10}$$

is the Jarlskog invariant, independent of the quark-field parameterization, that defines the global scale of the CP violation [16].

In the following we use the approximated parameterization of the CKM matrix introduced by Wolfenstein [17]:

$$\begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \tag{1.11}$$

where A, ρ, and η are real parameters, along with the expansion parameter $\lambda \equiv \sin(\theta_C) \approx 0.22$, and $\mathcal{O}(\lambda^4)$ terms are neglected.The Wolfenstein parameterization is particularly convenient for phenomenological applications because it readily illustrates the observed hierarchy of magnitude of CKM parameters, represented by the following 90% confidence intervals [18]:

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.9736\text{–}0.9741 & 0.2262\text{–}0.2282 & 0.00387\text{–}0.00405 \\ 0.2261\text{–}0.2281 & 0.97272\text{–}0.9732 & 0.04141\text{–}0.04231 \\ 0.00750\text{–}0.00846 & 0.004083\text{–}0.004173 & 0.999096\text{–}0.999134 \end{pmatrix}, \tag{1.12}$$

currently known from unitarity-constrained global fits of the unknown CKM parameters based on experimental input from measurements as diverse as nuclear β decays ($|V_{ud}|$), semileptonic kaon decays ($|V_{us}|$), muon production from (anti)neutrino interactions ($|V_{cd}|$), etc. When next-to-leading corrections in λ play a role, the CKM elements can be calculated with increasing levels of accuracy through expansions in λ:

$$\begin{aligned}
V_{ud} &= 1-\lambda^2/2-\lambda^4/8+\mathcal{O}(\lambda^6), \\
V_{us} &= \lambda+\mathcal{O}(\lambda^7), \\
V_{ub} &= A\lambda^3(\rho-i\eta), \\
V_{cd} &= -\lambda+A^2\lambda^5[1-2(\rho+i\eta)]/2+\mathcal{O}(\lambda^7), \\
V_{cs} &= 1-\lambda^2/2-(1+4A^2)\lambda^4/8+\mathcal{O}(\lambda^6), \\
V_{cb} &= A\lambda^2+\mathcal{O}(\lambda^8), \\
V_{td} &= A\lambda^3\left[1-(\rho+i\eta)(1-\lambda^2/2)\right]+\mathcal{O}(\lambda^7), \\
V_{ts} &= -A\lambda^2+A(1-2\rho)\lambda^4/2-i\eta A\lambda^4+\mathcal{O}(\lambda^6), \\
V_{tb} &= 1-A^2\lambda^4/2+\mathcal{O}(\lambda^6).
\end{aligned} \tag{1.13}$$

These may be simplified using

$$\bar{\rho} \equiv \rho(1-\lambda^2/2) \quad \text{and} \quad \bar{\eta} \equiv \eta(1-\lambda^2/2), \tag{1.14}$$

as prescribed in a generalized Wolfenstein parameterization that yields $V_{td} = A\lambda^3(1-\bar{\rho}-i\bar{\eta})$, $V_{us} = \lambda$, $V_{cb} = A\lambda^2$, and $J_{\mathsf{CP}} = \lambda^6A^2\eta$ with excellent accuracy.

The unitarity of the CKM triangle corresponds to the condition $V^\dagger_{\rm CKM}V_{\rm CKM} = \mathbf{1} = V_{\rm CKM}V^\dagger_{\rm CKM}$, which leads to six normalization equations and six orthogonality relations. The latter can be graphically represented as six triangles in the complex plane, each of area $J_{\mathsf{CP}}/2$:

$$\underbrace{V_{ud}V^*_{us}}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V^*_{cs}}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V^*_{ts}}_{\mathcal{O}(\lambda^5)} = 0, \tag{1.15}$$

$$\underbrace{V_{us}V^*_{ub}}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V^*_{cb}}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{ts}V^*_{tb}}_{\mathcal{O}(\lambda^2)} = 0, \tag{1.16}$$

$$\underbrace{V_{ud}V^*_{ub}}_{(\rho+i\eta)A\lambda^3} + \underbrace{V_{cd}V^*_{cb}}_{-A\lambda^3} + \underbrace{V_{td}V^*_{tb}}_{(1-\rho-i\eta)A\lambda^3} = 0, \tag{1.17}$$

$$\underbrace{V^*_{ud}V_{cd}}_{\mathcal{O}(\lambda)} + \underbrace{V^*_{us}V_{cs}}_{\mathcal{O}(\lambda)} + \underbrace{V^*_{ub}V_{cb}}_{\mathcal{O}(\lambda^5)} = 0, \tag{1.18}$$

$$\underbrace{V^*_{cd}V_{td}}_{\mathcal{O}(\lambda^4)} + \underbrace{V^*_{cs}V_{ts}}_{\mathcal{O}(\lambda^2)} + \underbrace{V^*_{cb}V_{tb}}_{\mathcal{O}(\lambda^2)} = 0, \tag{1.19}$$

$$\underbrace{V^*_{ud}V_{td}}_{(1-\rho-i\eta)A\lambda^3} + \underbrace{V^*_{us}V_{ts}}_{-A\lambda^3} + \underbrace{V^*_{ub}V_{tb}}_{(\rho+i\eta)A\lambda^3} = 0. \tag{1.20}$$

In the above equations, only the leading non-vanishing terms in the Wolfenstein expansion are shown. Only eqs. (1.17) and (1.20) represent triangles with sides of comparable magnitude, $\mathcal{O}(\lambda^3)$, i. e., visibly non-degenerate. In other relations, one side in suppressed by factors λ^2 to λ^4 with respect to the others, resulting in almost degenerate triangles. Since eqs. (1.17) and (1.20) agree with each other at the λ^3 level, yielding

$$[(\rho + i\eta) + (-1) + (1 - \rho - i\eta)]A\lambda^3 = 0, \tag{1.21}$$

they can be taken as describing the same triangle, usually referred to as *the* unitarity triangle (UT) of the CKM matrix, or "Bjorken triangle".

When the experimental accuracy will be sufficient to probe next-to-leading order terms of the Wolfenstein expansion, the difference between eq. (1.17) and eq. (1.20) will become relevant. Including $\mathcal{O}(\lambda^5)$ terms, eq. (1.21) is generalized to

$$[(\bar\rho + i\bar\eta) + (-1) + (1 - \bar\rho - i\bar\eta)]A\lambda^3 + \mathcal{O}(\lambda^7) = 0. \tag{1.22}$$

It is convenient to divide the above relation by the normalization factor $A\lambda^3$ and to introduce two sides,

$$R_b \equiv \sqrt{\bar\rho^2 + \bar\eta^2} = \left(1 - \frac{\lambda^2}{2}\right)\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right| \quad \text{and} \quad R_t \equiv \sqrt{(1-\bar\rho)^2 + \bar\eta^2} = \frac{1}{\lambda}\left|\frac{V_{tb}}{V_{cb}}\right|, \tag{1.23}$$

and three angles,

$$\alpha \equiv \arg\left(\frac{V_{td}V^*_{tb}}{V_{ud}V^*_{ub}}\right), \quad \beta \equiv \arg\left(\frac{V_{cd}V^*_{cb}}{V_{td}V^*_{tb}}\right), \quad \text{and} \quad \gamma \equiv \arg\left(\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}}\right). \tag{1.24}$$

The notation $\phi_1 \equiv \beta$, $\phi_2 \equiv \alpha$, $\phi_3 \equiv \gamma$ is also used for the angles. These relations are represented in the complex plane by the triangle of fig. 1.1, which is a straightforward generalization of the leading-order

case of eq. (1.21), in which the coordinates of the apex are $(\bar{\rho}, \bar{\eta})$ rather than (ρ, η). Further convenient relations that will be used in the following are

$$V_{ub} = A\lambda^3 \left(\frac{R_b}{1 - \lambda^2/2} \right) e^{-i\gamma}, \quad \text{and} \quad V_{td} = A\lambda^3 R_t e^{-i\beta}. \tag{1.25}$$

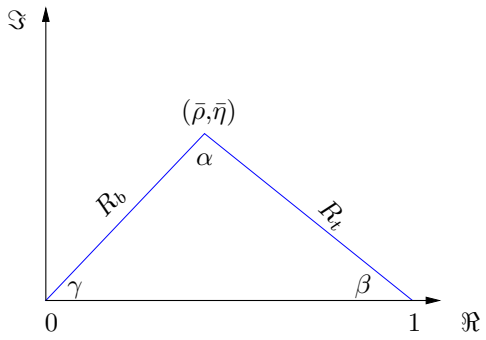

Figure 1.1: Representation of the unitarity triangle in the complex (Argand-Gauss) plane. The generalized Wolfenstein parameterization of the UT is used.

1.2 Phenomenological models of non-leptonic b-meson decays

Since in the Standard Model the mechanism of CP violation intervenes in the quark-mixing matrix, its phenomenology can be directly studied only in processes involving at least one quark pair. The presence of quarks in the interaction complicates the problem, because they are unobservable as physical states. In going from quarks to the observable hadrons, it is necessary to account for all the QCD processes that may arise between quarks with no alterations of the final state. However, strong interactions are not in a perturbative regime at energies $E \sim \Lambda_\chi \simeq 1$ GeV, owing to the infrared divergence of the strong coupling constant.[2] Among the processes involving quark decays, amplitudes of non-leptonic transitions of b-mesons are theoretically the most difficult to predict, because all particles in initial and final states are quarks, and the lack of theoretical information about the low-energy strong interactions affects both the initial bound meson state and the final states. These decays are mediated by $b \to q_1\ \bar{q_2}\ d(s)$ quark-level processes, with $q_1,\ q_2 = u,\ d,\ c, \text{or} s$.

Two topologies contribute to such decays: lower order "tree" topologies and internal-loop (i. e., "penguin") topologies, the latter being further divided in gluonic (QCD) and electroweak (EW) penguins. Figures 1.4–1.7 show the corresponding leading-order diagrams for $B^0_{(s)} \to h^+ h^{'-}$ decays. Depending on the flavor contents of their final states, non-leptonic decays are classified as

tree-only processes – where $q_1 \neq q_2$, and $q_2 = u,\ c$;

tree and penguin processes – where $q_1 = q_2$, and $q_1,\ q_2 = u,\ c$;

[2] With respect to the scale Λ_χ, quarks can be classified into two categories according to their mass: u, d, and s are "light" ($m_{u,d,s} < \Lambda_\chi$), whereas c, b, and t are "heavy" ($m_{c,b,t} > \Lambda_{\text{QCD}}$). Following this classification, we define as "heavy-flavored hadrons" (or "heavy-flavors" in short) hadrons containing quarks heavier than the strange.

penguin-only processes – where $q_1 = q_2$, and $q_1,\ q_2 = d,\ s$.

Several theoretical and numerical techniques has been developed to estimate the effect of low-energy QCD interactions and calculate the above transitions amplitudes. A comprehensive review of this activity is beyond the scope of this introduction to our measurement; we limit the discussion to those basic concepts that are useful for a better understanding of the motivation of the measurement and the interpretation of our final results.

1.2.1 Effective Hamiltonians

Currently, the most common conceptual framework in expressing amplitudes of heavy-flavor decays is connected with the use of low-energy, effective Hamiltonians. The problem of calculating hadronic amplitudes is simplified by considering that two energy scales are involved in the decays of pseudoscalar mesons: one $\mathcal{O}(100\ \text{GeV})$ scale characterizes the electroweak interaction of the quarks, another one, $\mathcal{O}(m_B)$, characterizes the hadronization of quarks into hadrons. The sizable difference between these two regimes allows the separation (i. e., factorization) of the problem, and derivation of a low-energy *effective* description of the phenomena. This effective theory is based on point-like interactions, in analogy with the early Fermi theory of genuine electroweak decays.

Given an energy scale μ appropriately chosen for the process of interest, effective Hamiltonians are calculated using the technique of *Operator Product Expansion* [19] to yield transition matrix elements of the following general form:

$$\langle f|\mathcal{H}_{\text{eff}}|i\rangle = \frac{G_F}{\sqrt{2}}\lambda_{\text{CKM}}\sum_k C_k(\mu)\langle f|Q_k(\mu)|i\rangle, \tag{1.26}$$

where G_F is the Fermi constant, λ_{CKM} contains the CKM-related quantities, and $|i\rangle$ and $\langle f|$ are initial and final states.

In eq. (1.26) one distinguishes a series of local operators (Q_k) multiplied by appropriate coefficients, acting as effective coupling constants. The perturbative short-distance contributions to the amplitude are separated form the long-distance ones, described by a product between perturbative quantities (the Wilson coefficient functions $C_k(\mu)$), and non-perturbative quantities (the hadronic matrix elements $\langle f|Q_k(\mu)|i\rangle$).

The interaction is considered as *effectively* governed by the local operators Q_k, generated by electroweak and strong interactions. The Wilson coefficients $C_k(\mu)$ are scale-dependent couplings related to the vertices described by the local operators and absorb all contributions of particles with mass $m > \mu$. Since the amplitude of a process can not depend on the scale μ chosen to separate the short-distance contributions from the long-distance ones, the dependence on μ of the Wilson coefficients ($C_k(\mu)$) has to be canceled by the dependence on μ of the operators ($Q_k(\mu)$). By imposing this condition, the equation of evolution of Wilson coefficients as a function of μ is obtained. The Wilson coefficient are then unequivocally determined by calculating them at a scale $\bar{\mu}$ high enough to allow a perturbative treatment (typically $\mu \approx m_W$), and then evaluating their value at the scale of the process of interest (typically $\mathcal{O}(\text{GeV})$) through the evolution equations. The construction of the effective Hamiltonian is made in a perturbative regime, i. e., the fact that the interacting hadrons are bound states of quarks does not affect the determination of C_k and Q_k factors. Calculations

are made considering only the quarks involved in the interaction. The problem of hadronization is therefore confined to the determination of the matrix elements of the Q_k operators. This requires non perturbative techniques (see sec. 1.4).

For a qualitative example of the extraction of a low-energy effective Hamiltonian the $\overline{B}^0 \to K^-\pi^+$ decay may be considered. Being induced by a quark level $b \to u\bar{u}s$ transition, it involves tree and penguin diagrams. For the moment we neglect penguin contributions, and assume a pure tree transi-

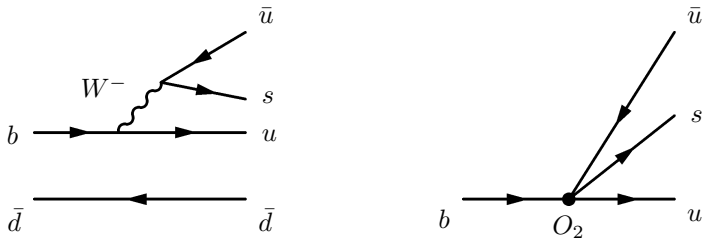

Figure 1.2: Color-allowed tree diagram contributing to the $\overline{B}^0 \to K^-\pi^+$ decay (left panel). Corresponding description in terms of the four-quark operator O_2 in the effective theory.

tion. In this case, the leading order diagram contributing to the $\overline{B}^0 \to K^-\pi^+$ decay is the one shown in the left panel of fig. 1.2, with the following transition amplitude:

$$T_{fi} = -\frac{g_2^2}{8} V_{us}^* V_{ub} [\bar{s}\gamma^\nu (1-\gamma_5) u] \left[\frac{g_{\nu\mu}}{k^2 - m_W^2}\right] [\bar{u}\gamma^\mu (1-\gamma_5) b]. \tag{1.27}$$

where g_2 is the $\mathrm{SU}(2)_L$ gauge coupling, k^2 is the transferred momentum, m_W is the W^+ mass, $g_{\mu\nu}$ is the metric tensor, γ indicate the usual 4×4 Dirac matrices, and s, u, b (and C-conjugates) indicate the four-spinors associated to quarks. Since the transferred momentum is much smaller than W mass, $k^2 \sim m_b^2 \ll m_W^2$, the W boson propagator can be assumed point-like, yielding

$$\frac{g_{\nu\mu}}{k^2 - m_W^2} \longrightarrow -\frac{g_{\nu\mu}}{m_W^2} \equiv -\left(\frac{8G_F}{\sqrt{2}g_2^2}\right) g_{\nu\mu}. \tag{1.28}$$

Thus,

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{us}^* V_{ub} [\bar{s}_\alpha \gamma_\mu (1-\gamma_5) u_\alpha][\bar{u}_\beta \gamma^\mu (1-\gamma_5) b_\beta] \equiv \frac{G_F}{\sqrt{2}} V_{us}^* V_{ub} O_2, \tag{1.29}$$

where α and β are the color indices of the $\mathrm{SU}(3)_C$ gauge group of QCD, and the current-current operator O_2 describes *effectively* the $b \to u\ \bar{u}\ s$ process (see fig. 1.2, right panel). The above description neglects QCD corrections. Their impact is two-fold: *factorizable* QCD corrections induce a renormalization-scale dependence on the Wilson coefficient C_2, i. e., $C_2(\mu) \neq 1$, whereas *non-factorizable* QCD corrections generate a second current-current operator, with mixed color indices, given by

$$O_1 \equiv [\bar{s}_\alpha \gamma_\mu (1-\gamma_5) u_\beta][\bar{u}_\beta \gamma^\mu (1-\gamma_5) b_\alpha]. \tag{1.30}$$

The scale μ appears as a consequence of the renormalization procedure necessary to eliminate the "artificial" ultraviolet divergences not present in the full theory, which arise from the transformation of the full W propagator into the point-like propagator (eq. (1.28)). The resulting low-energy

Hamiltonian has, therefore, the following structure:

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{us}^* V_{ub} [C_1(\mu) O_1 + C_2(\mu) O_2]. \tag{1.31}$$

The Wilson coefficients $C_1(\mu) \neq 0$ and $C_2(\mu) \neq 1$ are calculated by matching the full theory (the one with W exchange) with the effective theory (point-like, with W integrated out). This consists in calculating the QCD corrections in both the full and the effective theory, and then in expressing the QCD-corrected amplitude in terms of QCD-corrected matrix elements and Wilson coefficients. The results for the generic Wilson coefficient $C_k(\mu)$ contain terms of $\ln(\mu/m_W)$, which become large for $\mu = \mathcal{O}(m_b)$, i. e., at the scale governing the hadronic matrix elements of the O_k. Since the transition amplitude eq. (1.26) can not depend on the chosen renormalization scale μ, the renormalization group allows adding up the various terms of the Wilson coefficients:

$$\underbrace{\alpha_s^n \left[\ln\left(\frac{\mu}{m_W}\right)\right]^n}_{\text{leading order}}, \quad \underbrace{\alpha_s^n \left[\ln\left(\frac{\mu}{m_W}\right)\right]^{n-1}}_{\text{next-to-leading order}} \quad \dots \tag{1.32}$$

Relation (1.31) becomes more complicated when penguin topologies are taken into account. Application of the unitarity constraint

$$V_{ur}^* V_{ub} + V_{cr}^* V_{cb} + V_{tr}^* V_{tb} = 0 \quad (r = d,\ s) \tag{1.33}$$

allows integrating out the top quark entering the penguin loop processes, similarly to what done with the W (eq. (1.28)). The resulting effective Hamiltonian is

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \left\{ \sum_{j=u,c} V_{jr}^* V_{jb} \left[\sum_{k=1}^{2} C_k(\mu) Q_k^{jr} + \sum_{k=3}^{10} C_k(\mu) Q_k^r \right] \right\} \tag{1.34}$$

where another quark-flavor label has been introduced, $j = u,\ c$, and the effective operator basis has been expanded to contain the following operators:

Current-Current operators –

$$Q_1^{jr} = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) j_\beta][\bar{j}_\beta \gamma^\mu (1-\gamma_5) b_\alpha], \quad Q_2^{jr} = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) j_\alpha][\bar{j}_\beta \gamma^\mu (1-\gamma_5) b_\beta].$$

QCD-penguin operators –

$$Q_3^r = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\alpha] \sum_{q'} [\bar{q}'_\beta \gamma^\mu (1-\gamma_5) q'_\beta], \quad Q_4^r = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\beta] \sum_{q'} [\bar{q}'_\beta \gamma^\mu (1-\gamma_5) q'_\alpha]$$

$$Q_5^r = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\alpha] \sum_{q'} [\bar{q}'_\beta \gamma^\mu (1+\gamma_5) q'_\beta], \quad Q_6^r = [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\beta] \sum_{q'} [\bar{q}'_\beta \gamma^\mu (1+\gamma_5) q'_\alpha]$$

EW-penguin operators –

$$Q_7^r = \frac{3}{2} [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\alpha] \sum_{q'} e_{q'} [\bar{q}'_\beta \gamma^\mu (1+\gamma_5) q'_\beta], \quad Q_8^r = \frac{3}{2} [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\beta] \sum_{q'} e_{q'} [\bar{q}'_\beta \gamma^\mu (1+\gamma_5) q'_\alpha]$$

$$Q_9^r = \frac{3}{2} [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\alpha] \sum_{q'} e_{q'} [\bar{q}'_\beta \gamma^\mu (1-\gamma_5) q'_\beta], \quad Q_{10}^r = \frac{3}{2} [\bar{r}_\alpha \gamma_\mu (1-\gamma_5) b_\beta] \sum_{q'} e_{q'} [\bar{q}'_\beta \gamma^\mu (1-\gamma_5) q'_\alpha]$$

At the renormalization scale of interest, $\mu = \mathcal{O}(m_b)$, the Wilson coefficients of the current-current operators are $C_1(\mu) = \mathcal{O}(10^{-1})$ and $C_2(\mu) = \mathcal{O}(1)$, while those of the penguin operators are $\mathcal{O}(10^{-2})$. The small ratio $\alpha_{\rm QED}/\alpha_s = \mathcal{O}(10^{-2})$ would suggest a minor role for the EW penguin compared with the QCD penguins. However, the strong increase of the Wilson coefficient C_9 with the *top*-quark mass leads to important EW-penguin contributions in several b-meson decays, including $B^0_{(s)} \to h^+h'^-$.

1.2.2 Manifestation of new physics

The concept of low-energy effective Hamiltonian allows developing a phenomenological treatment of possible manifestations of new physics in b-meson decays. These are two-fold.

First, the new physics may modify the action of Standard Model operators through new, short-distance functions depending on new physics parameters, e. g., masses of charginos, squarks, etc. Virtual new physics particles may participate in second-order diagrams ("box" or penguin topologies) thereby being "integrated out" as the top quark and the W boson in the Standard Model. As a consequence, the initial conditions for the renormalization group evolutions become

$$C_k \longrightarrow C_k^{\rm SM} + C_k^{\rm NP}, \tag{1.35}$$

where $C_k^{\rm SM}$ and $C_k^{\rm NP}$ are the Wilson coefficients associated to the Standard Model and new physics amplitudes, respectively. The new physics-related $C_k^{\rm NP}$ coefficients may carry new CP-violating phases *not* related to the CKM matrix.

Secondly, new physics may enlarge the operator basis:

$$\{Q_k\} \longrightarrow \{Q_k^{\rm SM}, Q_l^{\rm NP}\}. \tag{1.36}$$

thus enhancing the role of operators otherwise absent or suppressed in the Standard Model. In this case, generally, new sources for flavor or CP violation arise.

The opportunity of measuring a relatively small set of CKM-related observables using a rich variety of observables associated to many distinct processes, some of which proceeding through very different dynamics, provides access to the detection of new physics: comparison between values of the same quantities measured through different processes, and experimental verification of the multiple correlations between CKM-parameters prescribed by the Standard Model, are sensitive to the possible virtual contributions of non-Standard Model particles in a large fraction of new physics models.

A straightforward evidence for new physics would be obtained, for instance, if decay amplitudes abnormally larger than expected were observed, indicating presence of non-Standard Model particles in penguin, box diagrams, or even tree diagrams.

1.3 CP-violating asymmetries

The low-energy effective Hamiltonians discussed so far are "universal", i. e., they apply to all b-meson decays governed by the same quark-level transition as, for instance, $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ decays. Consequently, the observed differences in rates between the various modes sharing the same quark-level processes are ascribed, in this formalism, to the hadronic matrix elements of the relevant

four-quark operators. The theoretical evaluation of such matrix elements is challenging and often associated to large uncertainties. In this respect, different methods have been devised to model the strong dynamics, reducing the number of hadronic matrix elements to be calculated. Before discussing these methods, it is convenient to introduce the CP-violating observables expressed in terms of the low-energy Hamiltonian. Following the relation (1.34) for the low-energy effective Hamiltonian, the amplitude of a $\overline{B} \to \overline{f}$ decay is

$$\mathcal{A}(\overline{B} \to \overline{f}) = \langle \overline{f}|\mathcal{H}_{\rm eff}|\overline{B}\rangle = \frac{G_F}{\sqrt{2}} \left\{ \sum_{j=u,c} V^*_{jr} V_{jb} \left[\sum_{k=1}^{2} C_k(\mu) \left\langle \overline{f}|Q^{jr}_k(\mu)|\overline{B}\right\rangle + \sum_{k=3}^{10} C_k(\mu) \langle \overline{f}|Q^r_k(\mu)|\overline{B}\rangle \right] \right\}, \tag{1.37}$$

while the amplitude of the CP conjugate process $B \to f$ is

$$\mathcal{A}(B \to f) = \left\langle f|\mathcal{H}^\dagger_{\rm eff}|B\right\rangle = \frac{G_F}{\sqrt{2}} \left\{ \sum_{j=u,c} V_{jr} V^*_{jb} \left[\sum_{k=1}^{2} C_k(\mu) \left\langle f|Q^{jr\dagger}_k(\mu)|B\right\rangle + \sum_{k=3}^{10} C_k(\mu) \left\langle f|Q^{r\dagger}_k(\mu)|B\right\rangle \right] \right\}. \tag{1.38}$$

Using the CP invariance of the strong interaction, $(\mathsf{CP})Q^{jr\dagger}_k(\mathsf{CP})^\dagger = Q^{jr}_k$, and applying the identity operator $\mathbf{1}= (\mathsf{CP})^\dagger(\mathsf{CP})$ to each $\langle f|$ and $|B\rangle$, we obtain

$$\mathcal{A}(B \to f) = e^{i[\phi_{\mathsf{CP}}(B)-\phi_{\mathsf{CP}}(f)]} \frac{G_F}{\sqrt{2}} \left\{ \sum_{j=u,c} V_{jr} V^*_{jb} \left[\sum_{k=1}^{2} C_k(\mu) \left\langle \overline{f}|Q^{jr\dagger}_k(\mu)|\overline{B}\right\rangle + \sum_{k=3}^{10} C_k(\mu) \left\langle \overline{f}|Q^{r\dagger}_k(\mu)|\overline{B}\right\rangle \right] \right\}, \tag{1.39}$$

where the convention-dependent phases $\phi_{\mathsf{CP}}(B)$ and $\phi_{\mathsf{CP}}(f)$ are defined by

$$\mathsf{CP}|B\rangle = e^{i\phi_{\mathsf{CP}}(B)}|\overline{B}\rangle, \quad \mathsf{CP}|f\rangle = e^{i\phi_{\mathsf{CP}}(f)}|\overline{f}\rangle. \tag{1.40}$$

The most general expression of a non-leptonic b-meson decay amplitude in the Standard Model is, therefore

$$\begin{aligned} \mathcal{A}(\overline{B} \to \overline{f}) &= e^{i\varphi_1}|\mathcal{A}_1|e^{i\delta_1} + e^{i\varphi_2}|\mathcal{A}_2|e^{i\delta_2} \\ \mathcal{A}(B \to f) &= e^{i[\phi_{\mathsf{CP}}(B)-\phi_{\mathsf{CP}}(f)]} \left[e^{-i\varphi_1}|\mathcal{A}_1|e^{i\delta_1} + e^{-i\varphi_2}|\mathcal{A}_2|e^{i\delta_2}\right] \end{aligned} \tag{1.41}$$

where φ_1 and φ_2 are CP-violating phases originated from the CKM factors $V_{jr}V^*_{jb}$ (or from non-Standard Model sources of CP-violation), and the CP-conserving (or strong or rescattering or final-state interaction) amplitudes $|\mathcal{A}_1|e^{i\delta_1}$ and $|\mathcal{A}_2|e^{i\delta_2}$ involve the hadronic matrix elements of the four-quark operators.

Using eq. (1.41), the following CP asymmetry is defined:

$$\begin{aligned} A_{\mathsf{CP}} &\equiv \frac{\Gamma(\overline{B} \to \overline{f}) - \Gamma(B \to f)}{\Gamma(\overline{B} \to \overline{f}) + \Gamma(\overline{B} \to \overline{f})} = \frac{|\mathcal{A}(\overline{B} \to \overline{f})|^2 - |\mathcal{A}(B \to f)|^2}{|\mathcal{A}(B \to f)|^2 + |\mathcal{A}(\overline{B} \to \overline{f})|^2} \\ &= -\frac{2|\mathcal{A}_1||\mathcal{A}_2|\sin(\delta_1-\delta_2)\sin(\varphi_1-\varphi_2)}{|\mathcal{A}_1|^2 + 2|\mathcal{A}_1||\mathcal{A}_2|\cos(\delta_1-\delta_2)\cos(\varphi_1-\varphi_2) + |\mathcal{A}_2|^2}. \end{aligned} \tag{1.42}$$

The phenomenon that yields a non vanishing value of the asymmetry in eq. (1.42) is referred to as direct CP violation (or CP violation in the decay). It originates directly from the amplitude level of

the decay, being a general property of the flavor-changing weak interactions at first order [20]. This phenomenon was observed for the first time in the neutral kaon system [7], and recently it has been established in the B sector through the asymmetries of $B^0 \to K^+\pi^-$ decay-rates at the B-Factories [2]. The direct CP violation may arise in decays of both charged and neutral b-hadrons.

Any single phase in an amplitude is irrelevant, being subject to removal by a redefinition of the phase convention. Only phase *differences* are important. In fact, a non vanishing value of the asymmetry arises when (at least) two amplitudes interfere, provided a non-trivial CP-violating phase-difference $\varphi_1 - \varphi_2$ and a non-trivial CP-conserving phase-difference $\delta_1 - \delta_2$ are present. The latter provide the "reference" phases with respect to which the change in sign of the weak phase can be detected. Equation (1.42) shows also that the more similar the magnitudes of the interfering amplitudes, the larger the resulting asymmetry. As a consequence, the decays that are experimentally more suited for measuring direct CP violation are those in which the small inter-generation mixing ("Cabibbo suppression") reduces the amplitudes of dominant, lowest-order, diagrams (tree) to a level comparable with the amplitudes of higher-order ones (penguin).

The $\varphi_1 - \varphi_2$ weak phase difference is in general given by one of the UT angles (usually γ). Unfortunately, hadronic uncertainties affect its extraction from asymmetries measured in data, because of the poorly known hadronic matrix elements in eq. (1.37).

1.3.1 The role of neutral b-mesons

Owing to the possibility of flavor-oscillations, neutral meson decays provide an enriched phenomenology of CP violation with respect to charged mesons. In the framework of the Standard Model, $B_q^0 - \overline{B}_q^0$ flavor-mixing ($q = d, s$) is among the few measurable manifestations of second-order electroweak interactions, occurring through box diagrams (see fig. 1.3). For time-scales significantly larger than the

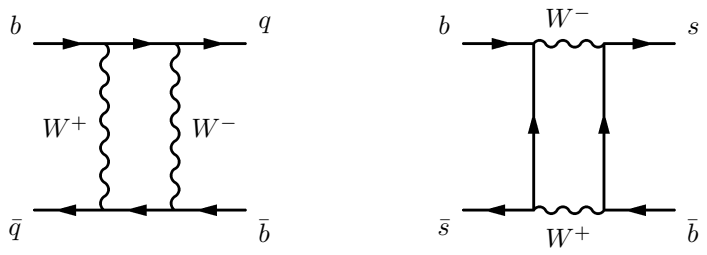

Figure 1.3: Box diagrams contributing to B_q^0-$\overline{B}_q^0$ flavor mixing.

strong-interaction scale, the mutual weak coupling between meson and antimeson induces a virtual transition between the two states, occurring as a function of time so that an initially ($t = 0$) pure B_q^0 meson state evolves into a time-dependent, linear combination of B_q^0 and $\overline{B}_q^0$ states.

$$|B_q^0(t)\rangle = a(t)|B_q^0\rangle + b(t)|\overline{B}_q^0\rangle. \tag{1.43}$$

In the Weisskopf-Wigner approximation for coupled unstable systems [21], the coefficients $a(t)$ and $b(t)$ are governed by a non-Hermitian quasi-Hamiltonian $\mathcal{H}$ through the following Schrödinger equation:[3]

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathcal{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \equiv \underbrace{\left[\begin{pmatrix} M_0^{(q)} & M_{12}^{(q)} \\ M_{12}^{(q)*} & M_0^{(q)} \end{pmatrix}\right.}_{\text{mass matrix}} - \frac{i}{2}\underbrace{\left.\begin{pmatrix} \Gamma_0^{(q)} & \Gamma_{12}^{(q)} \\ \Gamma_{12}^{(q)*} & \Gamma_0^{(q)} \end{pmatrix}\right]}_{\text{decay matrix}}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{1.44}$$

in which the $\mathcal{H}_{11} = \mathcal{H}_{22}$ condition derives from the assumption of conservation of the CPT symmetry. The eigenstates and eigenvalues are

$$|B^0_{q,\pm}\rangle = \frac{1}{\sqrt{1+|\alpha_q|^2}}\left(|B^0_q\rangle \pm |\overline{B}^0_q\rangle\right) \quad \lambda_{q,\pm} = \left(M_0^{(q)} - \frac{i}{2}\Gamma_0^{(q)}\right) \pm \left(M_{12}^{(q)} - \frac{i}{2}\Gamma_{12}^{(q)}\right)\alpha_q. \tag{1.45}$$

The eigenstates have masses $m_{q,H}$ ("heavy") and $m_{q,L}$ ("light") and widths $\Gamma_{q,H}$ and $\Gamma_{q,L}$, satisfying

$$\Delta m_q \equiv m_{q,H} - m_{q,L} = 2|M_{12}^{(q)}| > 0 \quad \text{and} \quad \Delta\Gamma_q \equiv \Gamma_{q,H} - \Gamma_{q,L} = \frac{4\Re(M_{12}^{(q)}\Gamma_{12}^{(q)*})}{\Delta m_q} < 0. \tag{1.46}$$

In addition, $\Gamma_q \equiv \Delta\Gamma_q/2$. We introduced the quantity α_q such that

$$\alpha_q e^{i\left(\Theta_{\Gamma_{12}}^{(q)} + n'\pi\right)} = \sqrt{\frac{4\left|M_{12}^{(q)}\right|^2 e^{-2i\left(\Theta_{\Gamma_{12}}^{(q)} - \Theta_{M_{12}}^{(q)}\right)} + \left|\Gamma_{12}^{(q)}\right|^2}{4\left|M_{12}^{(q)}\right|^2 + \left|\Gamma_{12}^{(q)}\right|^2 - 4\left|M_{12}^{(q)}\right|\left|\Gamma_{12}^{(q)}\right|\sin\left(\Theta_{M_{12}}^{(q)} - \Theta_{\Gamma_{12}}^{(q)}\right)}}, \tag{1.47}$$

in which $n' = 0, 1$ and

$$M_{12}^{(q)} \equiv \left|M_{12}^{(q)}\right| e^{i\Theta_{M_{12}}^{(q)}} \quad \text{and} \quad \Gamma_{12}^{(q)} \equiv \left|\Gamma_{12}^{(q)}\right| e^{i\Theta_{\Gamma_{12}}^{(q)}}. \tag{1.48}$$

Analytical (QCD sum-rules) and numerical (lattice simulation) calculations of mixing amplitudes allow estimating values of the mass and decay matrix elements:

$$M_{12}^{(q)} \propto (V_{tq}^* V_{tb})^2 e^{i\left[\pi - \phi_{\mathsf{CP}}(B_q^0)\right]}. \tag{1.49}$$

It is found that $\Gamma_{12}^{(q)}/M_{12}^{(q)} \approx \mathcal{O}(m_b^2/m_t^2) \ll 1$. Consequently, the expansion of eq. (1.47) at first order yields

$$\alpha_q = \left[1 + \frac{1}{2}\left|\frac{\Gamma_{12}^{(q)}}{M_{12}^{(q)}}\right|\sin\left(\Theta_{M_{12}}^{(q)} - \Theta_{\Gamma_{12}}^{(q)}\right)\right] e^{-i\left(\Theta_{\Gamma_{12}}^{(q)} + n'\pi\right)}. \tag{1.50}$$

The deviation of $|\alpha_q|$ from unit measures the CP violation in the *mixing* transition. The expected magnitude of the corresponding asymmetry, $\frac{|\alpha_q|^4-1}{|\alpha_q|^4+1} = \mathcal{O}(10^{-4})$, is strongly suppressed in the Standard Model. It is difficult to measure, but represents also an interesting probe of new physics, that may significantly enhance the asymmetry.

The time evolution of initially ($t = 0$) pure B_q^0- and $\overline{B}_q^0$-meson states is given by

$$|B_q^0(t)\rangle = f_+^{(q)}(t)|B_q^0\rangle + \alpha_q f_-^{(q)}(t)|\overline{B}_q^0\rangle \quad \text{and} \quad |\overline{B}_q^0(t)\rangle = \frac{f_-^{(q)}(t)}{\alpha_q}|B_q^0\rangle + f_+^{(q)}(t)|\overline{B}_q^0\rangle, \tag{1.51}$$

[3] Small corrections to the exponential decay laws affecting very short and very long decay times are irrelevant for the current level of experimental accuracy of mixing- and CP-related measurements.

in which

$$f_{\pm}^{(q)}(t) = \frac{1}{2}\left(e^{-i\lambda_{q,+}t} \pm e^{-i\lambda_{q,-}t}\right). \tag{1.52}$$

It is convenient to introduce the following quantities:

$$|g_{\pm}^{(q)}(t)|^2 = \frac{1}{4}\left[e^{-\Gamma_{q,L}t} + e^{-\Gamma_{q,H}t} \pm 2e^{-\Gamma_q t}\cos(\Delta m_q t)\right], \tag{1.53}$$

$$g_{-}^{(q)}(t)g_{+}^{(q)*}(t) = \frac{1}{4}\left[e^{-\Gamma_{q,L}t} + e^{-\Gamma_{q,H}t} + 2ie^{-\Gamma_q t}\sin(\Delta m_q t)\right], \tag{1.54}$$

and two observables that are independent of the phase convention chosen:

$$\xi_f^{(q)} = e^{-i\Theta_{M_{12}}^{(q)}}\frac{\mathcal{A}(\overline{B}_q^0 \to f)}{\mathcal{A}(B_q^0 \to f)} \quad \text{and} \quad \xi_{\overline{f}}^{(q)} = e^{-i\Theta_{M_{12}}^{(q)}}\frac{\mathcal{A}(\overline{B}_q^0 \to \overline{f})}{\mathcal{A}(B_q^0 \to \overline{f})}. \tag{1.55}$$

In fact, from eq. (1.49):

$$\Theta_{M_{12}}^{(q)} = \pi + 2\text{arg}(V_{tq}^* V_{tb} - \phi_{\mathsf{CP}}), \tag{1.56}$$

and the chosen CKM and CP phase conventions cancel out in the ratios of amplitudes. By combining the observables described, we arrive at the following transition rates for initially ($t = 0$) present B_q^0 or $\overline{B}_q^0$ mesons:

$$\Gamma(B_q^0(t) \to f) = \left\{|g_{-}^{(q)}(t)|^2 + |\xi_f^{(q)}|^2|g_{+}^{(q)}(t)|^2 - 2\Re\left[\xi_f^{(q)}g_{+}^{(q)}(t)g_{-}^{(q)*}(t)\right]\right\}\Gamma_f(B_q^0 \to f), \tag{1.57}$$

and

$$\Gamma(\overline{B}_q^0(t) \to f) = \left\{|g_{+}^{(q)}(t)|^2 + |\xi_f^{(q)}|^2|g_{-}^{(q)}(t)|^2 - 2\Re\left[\xi_f^{(q)}g_{-}^{(q)}(t)g_{+}^{(q)*}(t)\right]\right\}\Gamma_f(\overline{B}_q^0 \to f), \tag{1.58}$$

in which $\Gamma_f(B_q^0 \to f)$ and $\Gamma_f(\overline{B}_q^0 \to f)$ are obtained from the decay-amplitudes of non-evolved states.

1.3.2 Decays into CP eigenstates

If the study is restricted to decays of neutral B_q^0 and $\overline{B}_q^0$ mesons into CP eigenstates, i. e., $(\mathsf{CP})|f\rangle = \pm|f\rangle$), the equivalence $\xi_f^{(q)} = \xi_{\overline{f}}^{(q)}$ holds, and the expression for the CP-violating asymmetry takes on the following form:

$$A_{\mathsf{CP}}(t) \equiv \frac{\Gamma[\overline{B}_q^0(t) \to f] - \Gamma[B_q^0(t) \to f]}{\Gamma[B_q^0(t) \to f] + \Gamma[B_q^0(t) \to f]} \simeq \frac{A_{\mathsf{CP}}^{\text{dir}}\cos(\Delta m_q t) + A_{\mathsf{CP}}^{\text{mix}}\sin(\Delta m_q t)}{A_{\Delta\Gamma}\sinh(\Delta\Gamma_q t/2) - \cosh(\Delta\Gamma_q t/2)}, \tag{1.59}$$

in which terms of second order in the CP-violating parameters have been neglected the asymmetries are expressed as follows:

$$A_{\mathsf{CP}}^{\text{dir}} \equiv \frac{1 - |\xi_f^{(q)}|^2}{1 + |\xi_f^{(q)}|^2}, \quad A_{\mathsf{CP}}^{\text{mix}} \equiv \frac{2\Im(\xi_f^{(q)})}{1 + |\xi_f^{(q)}|^2}, \quad \text{and} \quad A_{\Delta\Gamma} \equiv \frac{2\Re(\xi_f^{(q)})}{1 + |\xi_f^{(q)}|^2}. \tag{1.60}$$

The direct CP term, $A_{\mathsf{CP}}^{\text{dir}}$, measures the CP violation in the decay and corresponds to the quantity shown in eq. (1.42). The $A_{\mathsf{CP}}^{\text{mix}}$ term is a feature of only neutral decays and represents the *mixing-induced* CP violation. The third asymmetry, $A_{\Delta\Gamma}$, is not independent of the other two, but may be experimentally accessible in case of a sizable width-difference ($\Delta\Gamma_q$).

The phase-convention independent CP-violating observable $\xi_f^{(q)}$ can be calculated inserting eq. (1.41) into eq. (1.55), and noting that $\phi_{\mathsf{CP}} = \pm 1$ because we are considering CP eigenstates:

$$\xi_f^{(q)} = \mp e^{-i\phi_q} \left(\frac{e^{i\varphi_1}|A_1|e^{i\delta_1} + e^{i\varphi_2}|A_2|e^{i\delta_2}}{e^{-i\varphi_1}|A_1|e^{i\delta_1} + e^{-i\varphi_2}|A_2|e^{i\delta_2}} \right), \tag{1.61}$$

where

$$\phi_q \equiv 2\arg\left(V_{tq}^* V_{tb}\right) = \begin{cases} 2\beta & (q = d) \\ -2\delta\gamma & (q = s) \end{cases} . \tag{1.62}$$

The angles β, γ, refer to the UT shown in fig. 1.1; the (small) angle δ derives from the last orthogonality relation of the CKM matrix in eq. (1.20). In analogy with eq. (1.42), the extraction of $\xi_f^{(q)}$ is in general affected by large hadronic uncertainties. However, if a single weak amplitude has a dominant role in the $B_q^0 \to f$ transition one obtains

$$\xi_f^{(q)} = \mp e^{-i\phi_q} \left(\frac{e^{i\varphi_f/2}|A_f|e^{i\delta_f}}{e^{-i\varphi_f/2}|A_f|e^{i\delta_f}} \right) = \mp e^{-i(\phi_q - \phi_f)}. \tag{1.63}$$

In the above expression, the hadronic matrix element $|A_f|e^{i\delta_f}$ cancels. The requirements needed for a non-zero direct CP violation are no longer valid in this case thereby $A_{\mathsf{CP}}^{\mathrm{dir}} = 0$, but the mixing-induced asymmetry becomes simply related to the CP-violating phase difference: $A_{\mathsf{CP}}^{\mathrm{mix}} = \pm\sin(\phi_q - \phi_f)$. This is the case of the $B^0 \to J/\psi K_S^0$ decay which, being strongly dominated by the tree level amplitude, provides direct access to the CKM parameter $\sin(2\beta)$. In the case of mixing induced CP violation, the mixing amplitude plays the role of the strong amplitude in the direct CP violation case, providing the "reference phase" through which the change in sign of the weak phase can be detected.

1.4 Hadronic matrix elements

Except in special cases, the presence of strong (CP-conserving) phases is necessary for obtaining observable CP-violating asymmetries. These phases need to be taken into account in the extraction of CKM parameters from observed decay-rates. However, knowledge of strong phases implies theoretical evaluation of matrix elements, which is challenging and often associated to large uncertainties.

In this respect, different phenomenological methods have been devised to model the strong dynamics, reducing the number of hadronic matrix elements to be calculated. All involve systematic expansions of QCD in ratio of quark masses and in the scale $\Lambda \simeq \Lambda_{\mathrm{QCD}}$ associated with the hadronization.

The commonly used phenomenological approaches can be classified into three general categories: methods relying on *QCD-based calculations* from first principles, methods based on *general amplitude parameterizations*, and methods combining specific aspects of both. A detailed discussion of the technical aspects of each approach is beyond the scope of this thesis, further information can be found in Ref. [22]. In the following, we briefly compare the relevant features of these approaches in the context of the measurement described in this work.

1.4.1 Factorization

Among the methods based on QCD calculations, the use of *factorization theorems* in the heavy b-quark limit of QCD provided the most complete theoretical analysis of two-body non-leptonic b-meson decays. The principle subtending the factorization approach is that, in a decay of an heavy $B^0_{(s)}$ meson into lighter mesons, the fast meson produced by a point-like source decouples from soft QCD interactions ("color transparency"). Amplitudes are systematically expanded in powers of $\Lambda/m_b \ll 1$ (heavy b-quark limit), yielding factorization theorems for the hadronic matrix elements in terms of universal hadronic functions such as light-cone distribution amplitudes of the mesons and matrix elements describing a heavy-to-light transition. This reduces the number of hadronic parameters.

A simplified example of factorization, based on the assumption of a pure tree transition for the $\overline{B}^0 \to K^-\pi^+$ decay (see fig. 1.2, left panel), illustrates the basic principle subtending the factorization of hadronic elements of four-quark operators into the product of hadronic matrix elements of quark currents [23]. Using the $\mathrm{SU}(3)_C$ color-algebra relation,

$$T^a_{\alpha\beta}T^a_{\gamma\delta} = \frac{1}{2}\left(\delta_{\alpha\delta}\delta_{\beta\gamma} - \frac{1}{N_C}\delta_{\alpha\beta}\delta_{\gamma\delta}\right), \tag{1.64}$$

to rewrite the operator O_1 (see eq. (1.30)), we obtain

$$\begin{aligned}\left\langle K^-\pi^+|\mathcal{H}_{\text{eff}}|\overline{B}^0\right\rangle &= \frac{G_F}{\sqrt{2}}V^*_{us}V_{ub} \\ &\times \Big\{a_1\left\langle K^-\pi^+|[\bar{s}_\alpha\gamma_\mu(1-\gamma_5)u_\alpha][\bar{u}_\beta\gamma^\mu(1-\gamma_5)b_\beta]|\overline{B}^0\right\rangle \\ &+ 2C_1\left\langle K^-\pi^+|[\bar{s}_\alpha T^a_{\alpha\beta}\gamma_\mu(1-\gamma_5)u_\beta][\bar{u}_\gamma T^a_{\gamma\delta}\gamma^\mu(1-\gamma_5)b_\delta]|\overline{B}^0\right\rangle\Big\}, \end{aligned}\tag{1.65}$$

where $a_1 = C_1/N_C + C_2 \sim 1$ is a combination of Wilson coefficients representing a phenomenological color factor that governs decays in which all color indices run through the whole diagram ("color-allowed"). When this is not the case, the color factor is smaller ("color-suppressed"). Since the second term of eq. (1.65) is zero, the hadronic matrix element is factorized as

$$\begin{aligned}\left\langle K^-\pi^+|[\bar{s}_\alpha\gamma_\mu(1-\gamma_5)u_\alpha][\bar{u}_\beta\gamma^\mu(1-\gamma_5)b_\beta]|\overline{B}^0\right\rangle &= \left\langle K^-|[\bar{s}_\alpha\gamma_\mu(1-\gamma_5)u_\alpha]|0\right\rangle \\ &\times \left\langle\pi^+|[\bar{u}_\beta\gamma^\mu(1-\gamma_5)b_\beta]|\overline{B}^0\right\rangle \\ &= if_K \times F_0^{B^0\pi}(m_K^2)\times(m_B^2 - m_\pi^2)\end{aligned}\tag{1.66}$$

where f_K is a decay constant, $F_0^{B^0\pi}(m_K^2)$ is a form factor, and $(m_B^2 - m_\pi^2)$ is a kinematic factor. In this simple case, the resulting amplitudes are rarely reliable, because non-factorizable and annihilation amplitudes are neglected and final state interactions are assumed absent. Different choices of the relevant expansion scales and of the input parameters lead to different models. Among the most commonly used in describing non-leptonic b-meson decays are the *QCD Factorization* (QCDF) approach, the *Perturbative QCD* (PQCD) approach, and the *Soft Collinear Effective Theory* (SCET).

QCD factorization

The QCDF-based predictions [24] represent the most complete theoretical analysis of two-body non-leptonic b-meson decays. This approach combines $1/m_b$ expansions with the perturbation theory in

α_s to achieve scale separation at leading power in $\Lambda_{\rm QCD}/m_b$ for a large class of exclusive hadronic decays in the limit $m_b \gg \Lambda_{\rm QCD}$. In a $B \to m_1 m_2$ decay, $m_b, E_m \gg \Lambda, m_m$, thus the amplitudes can be factorized into simpler non-perturbative objects, and a factorization theorem can be verified at one-loop. The resulting amplitudes use elements of the simplified factorization illustrated above. If at least one of the final state mesons is light, QCDF yields an amplitude with the following general structure:

$$A(\overline{B} \to m_1 m_2) = [\text{“naive model”}] \times [1 + \mathcal{O}(\alpha_s) + \mathcal{O}(\Lambda_{\rm QCD}/m_b)]. \tag{1.67}$$

Several input parameters are needed, including quark masses, heavy-to-light form factors, light-cone distributions amplitudes etc. The radiative, non-factorizable corrections $\mathcal{O}(\alpha_s)$ can be calculated systematically, whereas the main limitation originates from the $\mathcal{O}(\Lambda_{\rm QCD}/m_b)$ terms. Infrared end-point logarithmic and linear divergences arise when calculating certain contributions (e. g., annihilation diagrams). These $1/m_b$ divergences are usually chirally enhanced and need to be regularized empirically, thus introducing large uncertainties in the predictions. The benefits of the QCDF approach are that some of the hadronic parameters, namely ratios of tree-to-penguin amplitudes, strong phases, and corrections to the form factors, are obtained from first principles and independently of models. However, the form factors (evaluated at a point) need to be determined from QCD sum-rules [25] or from data. QCDF predicts that most strong phases, being expansions in α_s, are suppressed. As a consequence, factorization predicts small direct-CP asymmetries, possibly in contrast with experimental data.[4] In addition, QCDF predictions for weak annihilation amplitudes suffer from large uncertainties.

Perturbative QCD

An alternative approach that factorizes hard components from a QCD process which can be treated by perturbation theory was proposed in Ref. [26]. Form factors are assumed to have a perturbative expansion and the meson wave functions depend on transverse momenta. Non-perturbative parts are organized as universal hadron light-cone wave functions, which can be extracted from experimental data or constrained by lattice calculations or QCD sum-rules. It allows more stable treatment of end-point singularities arising when calculating non-factorizable and annihilation amplitudes with the QCDF approach.

Soft collinear effective theory

The SCET is an effective theory that contains both collinear and soft degrees of freedom allowing removal of divergences for highly-energetic particles (products of decays of b-mesons into light mesons of energies $E_m \sim m_b/2$) [27]. The accuracy of this expansion is $\Lambda/E_m \sim 2\Lambda/m_b \sim 0.2 \ll 1$, where the condition on Λ/E_m applies only to non-leptonic decays to two light mesons, $B^0_{(s)} \to m_1 m_2$, with energies E_m $m_b/2$. The factorization theorem derived using SCET agrees with the structure of the QCDF proposal if perturbation theory is applied at the scales m_b^2 and $m_b\Lambda$. It improves and generalizes QCDF allowing each of the scales m_b^2, $E_m\Lambda$, and Λ^2_{QCD} to be treated independently and factorization to be generalized to all orders in α_s. As a result, the contributions of long-distance $c\bar{c}$

[4] Equation (1.42) shows that an observable CP-violating asymmetry requires two interfering CP-violating and two interfering CP-conserving (i. e., strong) amplitudes.

penguin amplitudes ("charming-penguins"), left unfactorized, are predicted of the same magnitude as the leading order ones. They could introduce large strong phases, thus accommodating the large direct CP-violating asymmetries observed in data and not predicted by QCDF predictions.

Comments on factorization

A key utility of factorization theorems for non-leptonic decays is that the expansions adopted are *systematic*, thereby providing a method to estimate, not only the central values of quantities, but also their uncertainties, for comparison with experimental data. Predictions from factorization can be compared with data in several ways, depending on the scale at which perturbation theory is used, or if light-cone sum rules, models, or other data are used to determine the hadronic parameters. QCDF and PQCD use perturbation theory at the scale $E_m\Lambda$ and light-cone sum-rules or simple estimates are used for numerical evaluation of most hadronic parameters. In this regime, all non-leptonic observables can be predicted and compared with experimental data. The comparison lead to identification of a subset of power corrections, empirically parameterized in terms of additional unknowns to be included in the numerical analysis. These power corrections are crucial to reach reasonable agreement with the data. But, in some cases, the magnitude of power-suppressed contributions competes with leading order diagrams. Then, a systematic modification of the power counting that promotes these effects to leading order is necessary. Thus, a limitation of these approaches is that no clear distinction is available between results that are model-independent consequences of the QCD in the heavy-quark limit, and results that depend on the expansion chosen, non-perturbative input parameters, etc. Moreover, a convincing proof that the power series converge is still lacking. For instance, owing to the intermediate mass of the charmed quark, the identification of a convergent expansion for the long-distance $c\bar{c}$ penguin processes remains controversial [28]. In the QCDF model, contributions from charming penguins are factorized and small. But it was argued that their effect could be larger than expected. Part of these inconveniences, at least for heavy-to-light decays, is overcome by the SCET approach.

1.4.2 Flavor symmetries

Use of *flavor symmetries* for the light quarks (q) is the most relevant application of the methods based on relationships between amplitudes of different channels, to reduce the number of hadronic parameters.

The strategies based on flavor symmetries rely substantially on data for extracting the unknown amplitudes from those measured experimentally. This is both their main advantage and limitation. The advantage is that all information on hadronic parameters is extracted from data, including all contributions such as final state interactions, annihilation contributions, etc., which are typically difficult to calculate in a model-independent way. But the uncertainties on the predictions are usually large, their magnitude being dependent on experimental knowledge. And, when different experimental measurements of the same quantity are controversial (e. g., in occasion of the Belle and *BABAR* disagreement on $A_{\mathsf{CP}}^{\mathrm{dir}}$ in the $B^0 \to \pi^+\pi^-$ decay [29]) the predictive power of such methods is severely affected.

Another limitation consists in the difficult assessment of the magnitude of symmetry-violations. This can be partially done using QCD-based numerical or analytical calculation for the factorizable portion of the violation. But the non-factorizable part has to be extracted from data, and this may require large samples of extremely rare b-meson decays.

In addition, amplitude decompositions based on flavor symmetries are useful phenomenologically to extract CKM-related information, but provide marginal information on the details of quark-gluon dynamics in non-leptonic b-meson decays. Convenient amplitude relations for $B^0_{(s)} \to h^+h^{'-}$ decays useful are obtained using SU(2)- and SU(3)-based flavor symmetries.

SU(2) isospin symmetry

Use of isospin relations is based on expansions in $m_{u,d}/\Lambda \sim 0.03 \ll 1$. Several methods for determining α or γ are based on the isospin symmetry, which is conserved to a few percent accuracy. However, in $B^0_{(s)} \to h^+h^{'-}$ decays, the effectiveness of this approach is currently limited by the number of unknown isospin parameters. And, even with more complete experimental information, it would still be important to complement this strategy with information on amplitudes from SU(3) or factorization. Otherwise, the possible presence of new physics effects may remain undetected in full fits of SU(2) amplitudes [30].

SU(3) symmetry

SU(3)-based amplitude relations are obtained combining the $m_s/\Lambda \sim 0.3 \ll 1$ expansion parameter with those of the SU(2) symmetry. Several Standard Model-based strategies based on SU(3) symmetry have been proposed. Still, the limited number of currently precise measurements available makes it necessary to introduce additional "dynamical assumptions" to further reduce the number of hadronic parameters. These assumptions usually rely on the additional knowledge of the strong matrix elements from the factorization approach. For their relevance in the present analysis, implications and limits of the application of a subgroup of the SU(3) symmetry to the phenomenology of $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decays are discussed in detail as follows.

1.5 $B^0_{(s)} \to h^+h^{'-}$ decays

Two-body non-leptonic charmless decays are the most widely used processes to study flavor physics in the b-meson sector. The large mass of the b-meson allows for a plethora of open channels, which provide multiple ways for testing the consistency of the Standard Model interpretation of CP violation. For each channel, observables include the CP-averaged branching fraction, the direct CP-violating asymmetries and, for certain decays of neutral mesons, the mixing-induced CP-violating asymmetry.

These decay modes are considered interesting since the beginning of theoretical studies in the $B^0_{(s)}$ meson sector. The initial interest was motivated by the opportunity of a theoretically solid determination of the CKM angle α. In absence of second-order QCD penguin contribution ($\bar{b} \to \bar{s}(\bar{d})$), a measurement of the time-dependent asymmetry in the $B^0 \to \pi^+\pi^-$ decay directly provides the unitarity-constrained CKM phase $\alpha = 180° - \beta - \gamma$, in analogy with the $B^0 \to J/\psi K^0_S$ channel for

the β phase. However, a significant penguin contribution invalidates this direct approach. Several methods have been proposed to bound the magnitude of the penguin effects in this determination. In 1990, Gronau and London proposed a technique based on isospin symmetries to determine the gluonic penguin effects neglecting electroweak penguin amplitudes [3]. The challenge of this approach resides in the experimental difficulty of reconstructing large samples of all neutral and charged $B \to \pi\pi$ decays, including the color-suppressed $B^0 \to \pi^0\pi^0$ decays. Subsequently, Silva and Wolfenstein extended the idea to SU(3) flavor symmetry, combining also the $K\pi$ modes [4], followed by several authors [31].

In 1998–2000, the CLEO experiment measured $\mathcal{B}(B^0 \to K^+\pi^-)/\mathcal{B}(B^0 \to \pi^+\pi^-) \simeq 4$ whereas if only tree processes were to contribute, $\mathcal{B}(B^0 \to K^+\pi^-)/\mathcal{B}(B^0 \to \pi^+\pi^-) \propto |V_{us}|^2/|V_{ud}|^2 \approx \mathcal{O}(\lambda^2) \approx 0.05$ was expected. This confirmed the relevant role of penguin amplitudes.

Today, it is common to think of penguin amplitudes as an opportunity to be exploited rather than as a limitation, because the increased complexity implies also an enriched phenomenology: higher-order penguin processes may provide access for new physics phases, induced by virtual contributions of new particles in the loops. The $B^0_{(s)} \to h^+h^{'-}$ decays can be used for determining the values of the CKM-related quantities, which may differ from the ones extracted from other tree-level dominated processes, possibly indicating non-Standard Model CP-violating phases. In addition, $B^0_{(s)} \to h^+h^{'-}$ decays can provide valuable information on low-energy strong dynamics in $B^0_{(s)}$ meson decays.

Amplitudes of $B^0_{(s)} \to h^+h^{'-}$ decays are dominated by $\bar{b} \to \bar{u}$ (tree-type) and $\bar{b} \to \bar{s}(\bar{d})$ (penguin-type) quark transitions (see figs. 1.4–1.7). The observed decay-rates are $\mathcal{O}(10^{-5})$ and smaller because the former processes involve leading-order diagrams that are CKM suppressed ($|V_{ub}| \ll |V_{cb}|$), while the latter involves higher-order diagrams. Several strategies based on flavor symmetries were

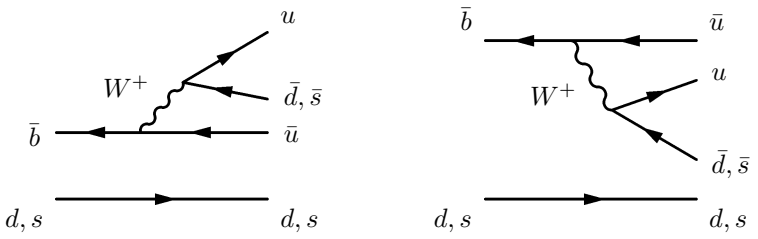

Figure 1.4: Color-allowed (left panel) and color-suppressed (right panel) tree (T) diagram contributing to $B^0_{(s)} \to h^+h^{'-}$ decays.

proposed to control the hadronic uncertainties in the predictions of $B^0_{(s)} \to h^+h^{'-}$ amplitudes. In the following, we focus the discussion on the U-spin flavor-symmetry because it has a specific interest in the context of this analysis.

1.5.1 Amplitude relations from U-spin flavor symmetry

The U-spin symmetry is a sub-group of the SU(3) flavor symmetry under which d quarks transform into s quarks. The $B^0 \to \pi^+\pi^-$ and the $B^0_s \to K^+K^-$ decays are an example of completely U-spin-symmetric channels. Relations (1.68)–(1.71) show an extended set of U-spin-symmetric decay modes

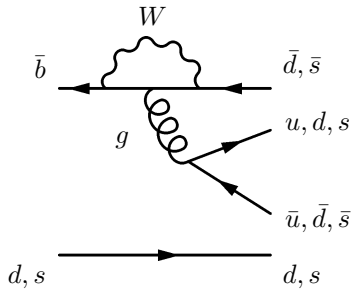

Figure 1.5: QCD-penguin (P) diagram contributing to $B^0_{(s)} \to h^+ h'^-$ decays.

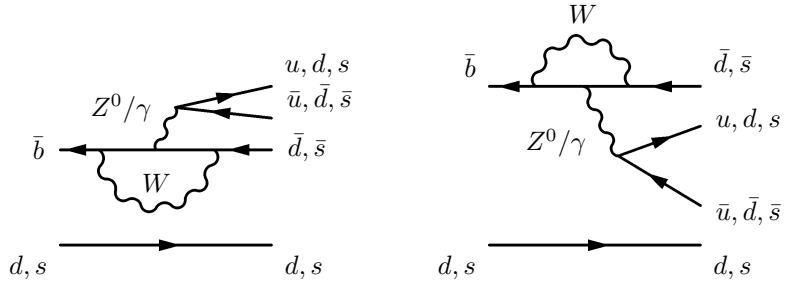

Figure 1.6: Color-allowed ($\mathrm{P_{EW}}$, left panel) and color suppressed ($\mathrm{P^C_{EW}}$, right panel) electroweak penguin diagram contributing to $B^0_{(s)} \to h^+ h'^-$ decays.

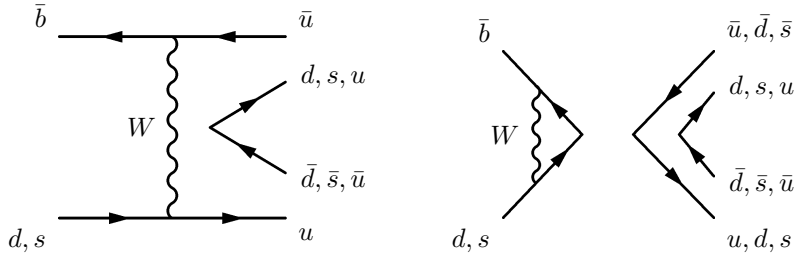

Figure 1.7: W-exchange (E, left panel) and penguin-annihilation (PA, right panel) diagram contributing to $B^0_{(s)} \to h^+ h'^-$ decays.

along with their amplitudes (indicated following the classification of figs. 1.4–1.7)

$$\underbrace{B^0 \to \pi^+\pi^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}+PA+E}} \qquad d \longleftrightarrow s \qquad \underbrace{B^0_s \to K^+K^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}+PA+E}} \tag{1.68}$$

$$\underbrace{B^0 \to K^+\pi^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}}} \qquad d \longleftrightarrow s \qquad \underbrace{B^0_s \to K^-\pi^+}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}}} \tag{1.69}$$

$$\underbrace{B^0 \to K^+\pi^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}}} \qquad d \overset{\text{spect.}}{\longleftrightarrow} s \qquad \underbrace{B^0_s \to K^+K^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}+PA+E}} \tag{1.70}$$

$$\underbrace{B^0 \to \pi^+\pi^-}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}+PA+E}} \qquad d \overset{\text{spect.}}{\longleftrightarrow} s \qquad \underbrace{B^0_s \to K^-\pi^+}_{\mathrm{T+P+\frac{2}{3}P^C_{EW}}} \tag{1.71}$$

The "spect." superscript labels relations in which U-spin-symmetry is applied only to the "spectator" quark, i. e., the valence quark of the $B^0_{(s)}$ meson that does *not* participate to the weak, quark-level process governing the decay. In these cases, the effect of annihilation and exchange diagrams (see fig. 1.7), in which *both* quarks of the $B^0_{(s)}$ meson participate to the weak transition, is assumed negligible. This assumption can be verified experimentally by measuring the rates of other $B^0_{(s)} \to h^+h'^-$ decays (see sec. 1.5.5).

U-spin symmetry is not exactly conserved in the Standard Model. The magnitude of its violation, due to both factorizable and non-factorizable hadronic matrix elements, is not precisely known but several authors expect it being a $\mathcal{O}(10\%)$ effect. The factorizable part of the violation can be calculated from QCD methods, such as light-cone sum-rules [32]; the full violation, that includes non-factorizable components, can only be inferred from data, by comparing amplitudes of U-spin-related modes.

Use of U-spin-symmetry has been proposed to limit the adverse effect of unknown penguin amplitudes and to extract the weak-phase γ from measurements on $B^0_{(s)} \to h^+h'^-$ decays. In 1993, Dunietz made the first proposal of relating the amplitudes of U-spin-symmetric $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ modes to extract the CKM phase α [33]. In 1999 Pirjol suggested that combining the branching-fractions of these modes constrains the penguin-induced uncertainty on the CKM phase α more effectively than the SU(3)- and SU(2)-based methods, in which the penguin contribution can not be distinguished from the rescattering one [34]. The definitive and most complete formalization of a strategy exploiting U-spin-symmetries applied to $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decays is due to Fleischer [5]. Further contributions to the Fleischer methods were given by Matias, London et al. [35, 36, 37]. Since the Fleischer proposal is the main theoretical motivation for studying $B^0_{(s)} \to h^+h'^-$ decays at CDF, it is discussed in more detail.

The $B^0 \to \pi^+\pi^-$ decay is governed by a $\bar{b} \to \bar{u}u\bar{d}$ transition, whose amplitude may be rewritten as

$$\mathcal{A}(B^0 \to \pi^+\pi^-) = V^*_{ub}V_{ud}\ (\mathcal{A}^u_{\mathcal{T}} + \mathcal{A}^u_{\mathcal{P}}) + V^*_{cb}V_{cd}\ \mathcal{A}^c_{\mathcal{P}} + V^*_{tb}V_{td}\mathcal{A}^t_{\mathcal{P}}, \tag{1.72}$$

where $\mathcal{A}^q_{\mathcal{T}}$ is the CP-conserving tree amplitude that includes the charged-current process and the exchange process ($\mathcal{T} = \mathrm{T} + \mathrm{E}$), and $\mathcal{A}^u_{\mathcal{P}}$, $\mathcal{A}^c_{\mathcal{P}}$, and, $\mathcal{A}^t_{\mathcal{P}}$ describe the QCD, and EW penguin and the penguin-annihilation topologies ($\mathcal{P} = \mathrm{P} + \mathrm{PA} + \frac{2}{3}\mathrm{P^C_{EW}}$) with internal quarks u, c, and t, respectively. Using eq. (1.33) to eliminate the CKM factor $V_{td}V^*_{tb} = -V_{ud}V^*_{ub} - V_{cd}V^*_{cb}$ and applying the Wolfenstein parameterization, we obtain

$$\mathcal{A}(B^0 \to \pi^+\pi^-) = \mathcal{C}\left(e^{i\gamma} - de^{i\theta}\right), \tag{1.73}$$

where

$$\mathcal{C} \equiv |V^*_{ub}V_{ud}|R_b\left(\mathcal{A}^u_{\mathcal{T}} + \mathcal{A}^u_{\mathcal{P}} - \mathcal{A}^t_{\mathcal{P}}\right), \quad \text{and} \quad de^{i\theta} \equiv \frac{1}{R_b}\left(\frac{\mathcal{A}^c_{\mathcal{P}} - \mathcal{A}^t_{\mathcal{P}}}{\mathcal{A}^u_{\mathcal{T}} + \mathcal{A}^u_{\mathcal{P}} - \mathcal{A}^t_{\mathcal{P}}}\right), \tag{1.74}$$

are hadronic parameters, and R_b is one side of the UT (see eq. (1.23)). The equations above are completely general parameterizations of CP-violating observables relying only on the unitarity of the CKM matrix. The large hadronic uncertainties affecting the $de^{i\theta}$ factor limit the possibility of a straightforward determination of the quantity $\sin(2\beta + 2\gamma)$ which would be directly accessible through

the measurement of $A_{\mathsf{CP}}^{\rm mix}(B^0 \to \pi^+\pi^-)$ if $d = 0$. However, even with $d \neq 0$, the following general expressions hold:

$$\begin{aligned} A_{\mathsf{CP}}^{\rm dir}(B^0 \to \pi^+\pi^-) &= -\frac{2d\sin(\theta)\sin(\gamma)}{1-2d\cos(\theta)\cos(\gamma)+d^2} \\ A_{\mathsf{CP}}^{\rm mix}(B^0 \to \pi^+\pi^-) &= \frac{\sin(\phi_d+2\gamma)-2d\cos(\theta)\sin(\phi_d+\gamma)+d^2\sin(\phi_d)}{1-2d\cos(\theta)\cos(\gamma)+d^2} \end{aligned} \tag{1.75}$$

where $\phi_d = 2\beta$ is known, up to a two-fold ambiguity, from the B-Factories measurements of the mixing-induced CP violation in $B^0 \to J/\psi K^0_S$ decays.[5]

Relation (1.75) show that the decay $B^0 \to \pi^+\pi^-$ may be used to probe γ, provided sufficient information on the hadronic parameters $\mathcal{C}$, d, and θ is available. The U-spin symmetry may contribute this information, through the combination of observables related to $B^0_s \to K^+K^-$ and $B^0 \to \pi^+\pi^-$decays.

The decay $B^0_s \to K^+K^-$ is a $\bar{b} \to \bar{s}u\bar{s}$ transition, thus involving tree and penguin diagrams as the $B^0 \to \pi^+\pi^-$ mode. Its CKM structure makes the penguin topologies dominant. In analogy to relation (1.73), the amplitude reads as

$$\mathcal{A}(B^0_s \to K^+K^-) = \left|\frac{V_{us}}{V_{ud}}\right| \mathcal{C}' \left(e^{i\gamma} - \left|\frac{V_{cs}V_{ud}}{V_{us}V_{cd}}\right| d' e^{i\theta'} \right), \tag{1.76}$$

where

$$\mathcal{C}' \equiv |V^*_{ub}V_{ud}| R_b \left(\mathcal{A}'^u_{\mathcal{T}} + \mathcal{A}'^u_{\mathcal{P}} - \mathcal{A}'^t_{\mathcal{P}}\right), \quad \text{and} \quad d' e^{i\theta'} \equiv \frac{1}{R_b}\left(\frac{\mathcal{A}'^c_{\mathcal{P}} - \mathcal{A}'^t_{\mathcal{P}}}{\mathcal{A}'^u_{\mathcal{T}} + \mathcal{A}'^u_{\mathcal{P}} - \mathcal{A}'^t_{\mathcal{P}}}\right), \tag{1.77}$$

Except for the interchange of a d quark by an s quark, the hadronic parameters $|\mathcal{C}'|$, d', and θ' have the same functional dependence on the penguin amplitudes as the $|\mathcal{C}|$, d, and θ parameters. In the U-spin limit, the following relations hold:

$$\mathcal{C}' = \mathcal{C}, \qquad d' = d, \qquad \text{and} \qquad \theta' = \theta. \tag{1.78}$$

If the U-spin violation is not maximal, or if it is measurable from data, this similarity in the functional dependence can be used to reduce the number of hadronic unknowns in $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decays.

1.5.2 Time-integrated $B^0_s \to K^+K^-$ decay rate

A first opportunity is provided by time integrated measurements of decay rate of the (still unobserved) $B^0_s \to K^+K^-$ channel. Fleischer and Matias indicate that the combination of $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ branching fractions is already sufficient to provide constraining information on γ or, alternatively, on the presence of non-Standard Model CP-violating phases [5, 35]. The ratio of $B^0_s \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ branching fractions has the following expression in terms of hadronic and

[5]Strictly speaking, the B-Factories measurements probe the quantity $\phi_d + \phi_K$, where ϕ_K is related to the weak K^0-$\overline{K}^0$ mixing phase; however, this is negligibly small in the Standard Model.

CKM-related parameters:

$$R_d^s \equiv \frac{\mathcal{B}(B_s^0 \to K^+K^-)}{\mathcal{B}(B^0 \to \pi^+\pi^-)} = \frac{1}{\epsilon}\left|\frac{\mathcal{C}'}{\mathcal{C}}\right| \frac{\epsilon^2 + 2\epsilon d' \cos(\theta')\cos(\gamma) + d'^2}{1 - 2d\cos(\theta)\cos(\gamma) + d^2} \times \frac{m_{B^0}}{m_{B_s^0}} \times \frac{\tau_{B^0}}{\tau_{B_s^0}} \times \frac{\sqrt{1 - \left(\frac{m_K}{m_{B_s^0}}\right)^2}}{\sqrt{1 - \left(\frac{m_\pi}{m_{B^0}}\right)^2}} \quad (1.79)$$

where

$$\epsilon \equiv \frac{\lambda^2}{1 - \lambda^2} \approx 0.053 \quad (1.80)$$

accounts for CKM factors,

$$\left|\frac{\mathcal{C}'}{\mathcal{C}}\right| \quad \text{and} \quad \frac{\epsilon^2 + 2\epsilon d' \cos(\theta')\cos(\gamma) + d'^2}{1 - 2d\cos(\theta)\cos(\gamma + d^2)} \quad (1.81)$$

account for U-spin-violation, and

$$\frac{\sqrt{1 - 2\left(m_K/m_{B_s^0}\right)^2}}{\sqrt{1 - 2\left(m_\pi/m_{B^0}\right)^2}} \approx 0.92 \quad (1.82)$$

is a kinematic factor. In the SU(3) limit, the three observables $A_{\mathsf{CP}}^{\mathrm{dir}}(B^0 \to \pi^+\pi^-)$, $A_{\mathsf{CP}}^{\mathrm{mix}}(B^0 \to \pi^+\pi^-)$, and R_d^s depend on the four theoretical parameters d, θ, ϕ_d, and γ. The phase ϕ_d is precisely known from the *BABAR* and Belle experiments. Assuming the Standard Model, the value of the phase γ can be used from independent measurements or from global CKM fits; thus, In this case, knowledge of two observables allows prediction of the third. This prediction can be compared with the direct experimental measurements to check the consistency of the approach.

More interestingly, one can resolve for d, γ, and θ using the three observables to check whether the result on γ is consistent with independent measurements. Of particular relevance is the comparison with independent determinations of the phase γ from decays not involving penguin amplitudes, such as $B^0_{(s)} \to D^-_{(s)}K^+$: a significant discrepancy in this case, would indicate the contribution of virtual non-Standard Model particles in the loops of the $B_s^0 \to K^+K^-$ penguin amplitudes.

However, it is known that the U-spin symmetry is not exactly conserved in the Standard Model(see sec. 1.5.3), but still, useful information is available through correlations in the $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ observable spaces: in presence of U-spin violation, eq. (1.79) is approximately valid yielding predictions affected by additional uncertainties. Nevertheless, Fleischer et al. show that even in case of significant U-spin violations, constraining information on the phase γ, or on the possible contribution of new physics can be extracted by comparing the experimental data with the Standard Model-allowed regions in the space of observables [35, 36].

This strategy, based on time-integrated $B_s^0 \to K^+K^-$ decay-rates is particularly suited for an application to the first available samples of the (yet unobserved) $B_s^0 \to K^+K^-$ decays: the measurement of $\mathcal{B}(B_s^0 \to K^+K^-)$ can be compared with the prediction extracted from the already available measurement of time-dependent $B^0 \to \pi^+\pi^-$ asymmetries from the B-Factories to check for indications of inconsistent values of γ or new, CP-violating physics-phases.

1.5.3 Time-dependent partial $B^0_s \to K^+K^-$ rate asymmetries

When large samples of $B^0_s \to K^+K^-$ will be available, the second step of the Fleischer method can be applied. Using eq. (1.76), the partial-rate decay asymmetries of the $B^0_s \to K^+K^-$ decay are written, in analogy with eq. (1.75), as

$$\begin{aligned} A^{\rm dir}_{\mathsf{CP}}(B^0_s \to K^+K^-) &= \frac{2(d'/\epsilon)\sin(\theta')\sin(\gamma)}{1-2(d'/\epsilon)\cos(\theta')\cos(\gamma)+(d'/\epsilon)^2} \\ A^{\rm mix}_{\mathsf{CP}}(B^0_s \to K^+K^-) &= \frac{\sin(\phi_s+2\gamma)-2(d'/\epsilon)\cos(\theta')\sin(\phi_s+\gamma)+(d'/\epsilon)^2\sin(\phi_s)}{1-2(d'/\epsilon)\cos(\theta')\cos(\gamma)+(d'/\epsilon)^2}, \end{aligned} \tag{1.83}$$

The phase $\phi_s = \mathcal{O}(1^\circ)$ is negligibly small in the Standard Model; it can be experimentally probed using the time-dependent rate asymmetries of $B^0_s \to J/\psi\phi$. Thus, the measurements of asymmetries in $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decays can be converted into theoretically reliable contours in the (γ, d) and (γ, d') planes.

Further progress is possible exploiting the U-spin flavor symmetry. In the approximation of conserved symmetry, the angle γ would be the only unknown in the system of four equations shown in relation (1.75) and relation (1.83). However, theoretical and experimental difficulties challenge the extraction of the phase γ through this strategy:

Magnitude of U-spin violation – a theoretical limitation to the validity of the above method arises from the unknown magnitude of the violation of U-spin symmetry in the Standard Model. No method has been devised yet to fully determine this magnitude. Constraining bounds on the factorizable portion of the violation can be extracted from QCD calculations [32], but little information is presently available for the non-factorizable portion. However, as the experimental measurements of U-spin-related $B^0_{(s)} \to h^+h'^-$ decays increase, additional constraining information will become available.

Measurement of $B^0_s \to K^+K^-$ asymmetries – experimentally, the measurement of time-dependent $B^0_s \to K^+K^-$ decay asymmetries is challenging. While the CP-violating asymmetries in $B^0 \to \pi^+\pi^-$ decays are known with increasing precision from the measurements at the Belle and *BABAR* experiments, their $B^0_s \to K^+K^-$ counterparts are yet to be measured. Currently, the $B^0_s \to K^+K^-$ decays can be produced only at the Tevatron. A measurement of CP-violating time-dependent asymmetries requires fitting the time-evolution (see eq. (1.59)) of a sample of decays in which the B^0_s flavor at the time of production is known (i. e., flavor-tagged decays). This requires a large initial sample of $B^0_s \to K^+K^-$ decays, since an unavoidable factor of ≈ 20 reduction in statistics is due to the limitations of the production-flavor tagging-algorithms. Additional dilution, due to the time-dependent treatment, further affects the final statistical uncertainties on the asymmetries. Furthermore, high resolution in reconstructing the decay vertices is required to resolve the fast oscillation of B^0_s mesons. Further discussion is in sec. 10.7.

1.5.4 Role of the $B^0_s \to K^-\pi^+$ decay

The $B^0_s \to K^-\pi^+$ decay, which is the (still unobserved) U-spin partner of the $B^0 \to K^+\pi^-$ decay, has been indicated by several authors as a source of rich opportunities of investigation.

In 1998 Gronau and Rosner proposed an SU(3)-based strategy to measure the CKM phase γ [38]. Experimentally, their proposal is accessible since it involves only time-integrated measurements of partial rates of untagged $B^0_{(s)}$ and B^+ decays into kaons and pions. But, from the theoretical stand-point, it needs some unknown corrections, due to rescattering effects and $\mathrm{P}^{\mathrm{C}}_{\mathrm{EW}}$ contributions. Gronau and Rosner indicated the $B^0_s \to K^-\pi^+$ mode as the key to reduce one of the major sources of theoretical uncertainty, i. e., the parameter associated to the ratio between tree and penguin amplitudes (P/T) [39]. Using the spectator-U-spin symmetry, which transforms the U-spin partners decays $B^0_s \to K^-\pi^+$ and $B^0 \to K^+\pi^-$ into another pair of U-spin partners, $B^0_s \to K^+K^-$ and $B^0 \to \pi^+\pi^-$, the authors devise a method to precisely measure P/T, provided separate measurement of $B^0_s \to K^-\pi^+$ and $\overline{B}^0_s \to K^+\pi^-$ decay rates are available.

More recently, a paper by Lipkin drew additional attention on the $B^0_s \to K^-\pi^+$ decay [40]. The observation of $\mathcal{O}(10\%)$ direct CP violation in the $B^0 \to K^+\pi^-$ decay [2] is not supported by an expected similar effect in the $B^+ \to K^+\pi^0$ decay, which differs only by the spectator quark. This raised interpretations about this disagreement being a possible indication of new physics. Lipkin suggests that partial decay-rates in $B^0_s \to K^-\pi^+$ provide a stringent test that necessarily confirms or rules out the presence of new physics involved in $B^0 \to K^+\pi^-$ amplitudes. This approach, inspired to an earlier work by Gronau [41], exploits an accidental cancellation between CKM factors involved in $B^0 \to K^+\pi^-$ and $B^0_s \to K^-\pi^+$ decays, which yields *equal* expected direct CP asymmetries in these decays. In fact, penguin (tree) amplitudes dominate the $B^0 \to K^+\pi^-$ ($B^0_s \to K^-\pi^+$) process; in the B^0_s decay, the penguin amplitude is suppressed, with respect to its B^0 counterpart, by exactly the same factor that enhances the amplitude of the interfering tree amplitude. This accidental correspondence is valid only within the framework of the Standard Model, thereby any significant disagreement between the measured partial rate asymmetries of strange and non-strange b-meson $K\pi$ decays is strong convincing indication of new physics.

The relevance of this test resides in its nearly complete independence of models. No flavor symmetry are required, but only charge-conjugation invariance for all final state rescattering. In addition, the large asymmetries expected in the $B^0_s \to K^-\pi^+$ decay make this test relatively easy, from an experimental point of view, when $B^0_s \to K^-\pi^+$ decays will be observed.

1.5.5 Role of penguin-annihilation and exchange amplitudes

Amplitudes of penguin-annihilation and exchange diagrams, in which all initial-state quarks undergo a transition, are difficult to predict with current phenomenological models. Their contribution needs to be neglected, for instance, when assuming spectator-U-spin relations (e. g., $B^0 \to K^+\pi^- \longleftrightarrow B^0_s \to K^-\pi^+$). And QCDF predictions are also affected by extremely large uncertainties due to the empirical corrections necessary to control soft end-point divergences.

But, in general, PA and E topologies may carry different CP-violating and CP-conserving phases with respect to the leading processes, thereby influencing the determination of CKM-related parameters from widely-used decays as $B^0 \to \pi^+\pi^-$. Experimental information on their magnitude is, therefore, particularly desirable.

In the Standard Model, the decays $B^0 \to K^+K^-$ and $B^0_s \to \pi^+\pi^-$ proceed only through PA and E diagrams. A measurement of their decay rates (or improved constraints on them) would thereby

provide valuable estimates of the magnitude of PA and E contributions. Theoretical predictions estimate the branching fractions of $B^0 \to K^+K^-$ and $B^0_s \to \pi^+\pi^-$ modes in the 10^{-7}–10^{-8} range [42, 43, 44]. While the $B^0 \to K^+K^-$ mode is reconstructable also the B-Factories, currently only CDF has access to the $B^0_s \to \pi^+\pi^-$ decay mode.

1.6 Experimental considerations

The measurements involving b-hadrons produced in hadron-hadron collisions had a role since the beginning of the experimental study of flavor physics in the b-meson sector, as confirmed by the discovery of the b-quark, the pioneering results of the UA1 Collaboration [45], and the CDF measurements obtained from the 1992–96 data-taking period (Run I) [46]. However, the largest experimental contribution to the understanding of the physics of the b-quark derives from e^+e^- machines operating at the $\Upsilon(4S)$ resonance (just above the open beauty threshold) or at the Z^0 pole. In the following, we discuss the peculiar features of the hadronic environment and compare them with their e^+e^- counterparts (see tab. 1.1). This provides a better understanding of the experimental difficulties encountered in the measurements presented in this thesis, and of the strategies adopted to overcome them.

In mid 1977, an excess in the 9.5 GeV/c^2 region of the invariant $\mu^+\mu^-$-mass distribution from products of collisions of 400 GeV/c protons on nuclear targets was interpreted as a bound quark-antiquark state containing a new fundamental particle, the b-quark [47]. The original enhancement was soon resolved into two resonances, Υ' and Υ'', and rapidly confirmed by the PLUTO and Double Arm Spectrometer II experiments at the DOppel RIng-Speicher e^+e^- storage ring (DORIS), which accurately determined the properties of the new resonances in 1978 [48]. Since 1980, further refinements of these measurements and discovery of new bound $b\bar{b}$ states were provided by the CLEO and the Columbia University-Stony Brook experiments at the Cornell Electron Storage Ring (CESR), and by the Crystal Ball detector at the upgraded DORIS ring [49]. In the 1980s, the experimental knowledge on the phenomenology of the b-quark increased considerably. The CLEO experiment observed enhancements in the single electron and muon inclusive cross-sections in e^+e^- collisions, providing evidence of open b-meson production followed by semileptonic decays. Subsequently, the first samples of exclusive and semileptonic b-meson decays were reconstructed. The Mark II Collaboration at the PEP ring measured a B^0 meson lifetime of about 1 ps, confirmed by several other experiments [50]. The long lifetime of b-hadrons compared with lifetimes of light-flavors is a key feature that makes their experimental study so rich. If exploited with detectors with sufficient spatial resolution, it constitutes a highly-discriminating quantity in rejecting background, and allows studying the flavor properties of neutral b-mesons through measurements of their time evolution. In fact, after some indications from the CLEO and Underground Area 1 (UA1) experiments [51], convincing evidence of B^0-$\overline{B}^0$ mixing was found by the ARGUS Collaboration [52], which detected like-sign semileptonic decays of $B\overline{B}$ pairs. This also provided the first indication of high-frequency flavor oscillations of B^0_s mesons, combined with the UA1 measurement [51].[6] Measurements on b-physics in the 1990s were dominated by the CLEO experiment and the LEP experiments (especially ALEPH — Apparatus for LEP PHysics), which observed several new decay-modes of b-mesons.

[6]The UA1 result also indicated a large value of the *top*-quark mass, $m_t \gtrsim 50$ GeV/c^2.

1.6.1 The B-Factories

Following the original 1987 proposal by Oddone [53], the latest evolution of accelerators for the study of b-meson decays are the B-Factories: high-luminosity e^+e^- colliders with asymmetric beam energies that produce $\Upsilon(4S)$ resonances with 0.4–0.6 Lorentz boost. The $\Upsilon(4S)$ meson decays 96% of the times into $B\overline{B}$ pairs ($B = B^0$ or B^+) which in turn decay in vertices typically spaced apart by 200–300 μm. This allows determination of the time-interval between the two decays with sufficient precision for measuring time-dependent CP-violating asymmetries, given the availability of silicon detectors with good enough spatial resolution. Operating the colliders at an energy corresponding to the $\Upsilon(4S)$ resonance, just above the open beauty threshold, reduces backgrounds because of tight kinematic constraints and absence of fragmentation products.

The experiments installed at the B-Factories (*BABAR* [54] and Belle [55]) are large-acceptance asymmetric spectrometers with excellent performance in reconstructing charged-particle trajectories (silicon micro-vertex detectors plus drift chambers with helium-based gas admixtures in 1.5 T magnetic fields), that allows discrimination of heavy-flavor decay vertices and measurement of decay-time differences. Redundant information from silica quartz (*BABAR*) or aerogel (Belle) Cherenkov counters, time-of-flight detectors, thallium-doped cesium iodide electromagnetic calorimeters, and an outermost layers of muon detectors allow identification of muons, electrons, and hadrons.

1.6.2 The hadron collider environment

The primary advantage of high-energy hadron colliders is the large cross-section for b-quark production. At the Fermilab Tevatron $p\bar{p}$ collider, the dominant production process is non-resonant inclusive b-quark pair-production of the type $p\bar{p} \to b\bar{b}X$.[7] The corresponding cross section, multiplied by the efficiency that at least one resulting b-hadron is within the detector acceptance (i. e., its decay products have approximately $p_T > 300$ GeV/c and $|\theta| < 40^\circ$ with respect to the beam axis) is $\sigma(p\bar{p} \to b\bar{b}+X) \approx 50$ μb [56]. This, corresponding to roughly 5,000 b-hadrons within CDF acceptance per second (at typical current luminosities), is large with respect to production cross-sections at e^+e^- machines: 1 nb at the $\Upsilon(4S)$ resonance and 7 nb at the Z^0 pole. As a consequence, experiments at the B-Factories require $\mathcal{O}(10^3)$ higher luminosities for collection of samples comparable in size to the Tevatron samples.

A second advantage of hadron collisions is the available center-of-mass energy. At the Tevatron ($\sqrt{s} = 1.96$ TeV) this is sufficient to produce all species of b-hadrons: not only B^0 and B^+ mesons, produced in large quantities in $\Upsilon(4S)$ decays, but also B_s^0 and B_c^+ mesons and b-baryons. In addition, the typical relativistic factors (i. e., Lorentz boost or $\beta\gamma$) of b-hadrons produced in $p\bar{p}$ collisions are larger with respect to the B-Factories. This results in larger decay-lengths, which allow probing shorter scales in the time-evolution of heavy-flavors.

However, the hadron collider environment has its disadvantages. The large $b\bar{b}$ cross-section is still about three order of magnitudes smaller than the $p\bar{p}$ inelastic cross-section, resulting in $\mathcal{O}(10^{-9})$ signal-to-background ratios at production, for typical branching-fractions of interesting processes. This poses a challenging task for the trigger systems. In this respect, the component of momentum of b-hadrons

[7] The Tevatron collider is described in sec. 2.1.

perpendicular to the beam line (transverse momentum), which scales approximately as the b-quark mass ($\langle p_{\mathrm{T}}(B)\rangle \approx 5$ GeV/c), provides a first discrimination from light-quark background, generally distributed at lower values. But the transverse momentum distribution of b-hadrons is a rapidly falling function: most b-hadrons have low transverse momenta and decay into particles often having $p_{\mathrm{T}} < 1$ GeV/c. The need to select low-momentum particles conflicts with the accept-rate limitations of the data acquisition systems. Furthermore, since the longitudinal component of b-hadron momenta is frequently large, their decay products tend to be boosted along the beam direction, thus escaping the detector acceptance. If one b-hadron is within CDF acceptance, the other one is within acceptance only $\mathcal{O}(20\%)$ of the time.

Reconstruction of b-hadron decays in a hadron collider is also plagued by the more complex event structure. Different sources contribute particles entering the detector acceptance when protons collide against antiprotons.

Hard interaction – in most hard $\mathrm{p\bar{p}}$ interactions, only one proton constituent (a valence — u, d — or a "sea" quark, or a gluon) undergoes hard-scattering against one antiproton constituent. This two-to-two parton scattering is the leading interaction that may produce the $b\bar{b}$ pair and additional particles through initial and final state radiation.

Underlying event – when a hard collision occurs, the remnants of proton and antiproton rearrange in color-neutral hadrons which may have a momentum perpendicular to the beam sufficient to enter the detector acceptance. The underlying event also includes the products of multiple parton interactions, i. e., multiple hard-scattering between different constituents of the same $\mathrm{p\bar{p}}$ pair.

b-hadron fragmentation process – quarks are never observed as independent entities, but are confined inside hadrons with a partner antiquark (in mesons) or diquark (in baryons). Fragmentation (or hadronization) is the long-distance, non-perturbative QCD process of transition from a single, final-state quark to an observable color-singlet hadron. In this process, a number of accompanying hadrons is produced in a local region around the hadronizing quark, as a result of the fragmentation of color lines of strong force. The fragmentation of the $b\bar{b}$ pair is exploited in the present analysis to reject background (see sec. 3.7.2). The fragmentation of all other quarks and gluons produced in the event, including those generated in the final- and initial-state radiation, in the underlying event etc. is generally source of nuisance in the analysis.

Pile-up event – when a beam of protons crosses a beam of antiprotons, multiple hard interactions may occur between different proton-antiproton pairs. Each hard interaction contribute with an associated fragmentation process and underlying event.

Occasionally, products of interactions of the beams with residual gas contaminating the vacuum of the beam-pipe add to the above sources. As a consequence, in a $\mathrm{p\bar{p}}$ collider as the Tevatron, high particle multiplicities per event are observed (30–40 charged particles with $p_{\mathrm{T}} \gtrsim 50$ MeV), with multiple production vertices, and event shapes lacking well-defined structures (see fig. 1.8). This results in greater detector occupancies and in increased difficulty in reconstructing b-hadron decay topologies. The $\mathrm{e^+e^-}$ colliders, instead, benefit from low particle multiplicities and well defined spherical (back-to-back) event shapes at the B-Factories (LEP) (see fig. 1.9).

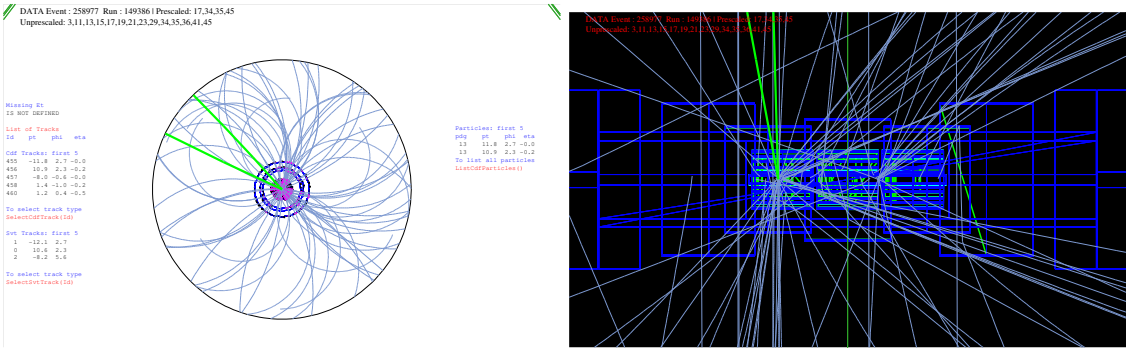

Figure 1.8: Computerized reconstruction of an event containing a $B^0_{(s)} \to h^+h^{'-}$ candidate (green tracks) in the CDF II detector. Left panel shows an enlargement of the drift chamber (see sec. 2.3.5) projected in the plane transverse to the beam. Right panel shows an enlargement of the inner silicon micro-vertex detector (see sec. 2.3.3) projected in the plane that contains the beam. Two distinct $p\bar{p}$ interactions are clearly distinguishable.

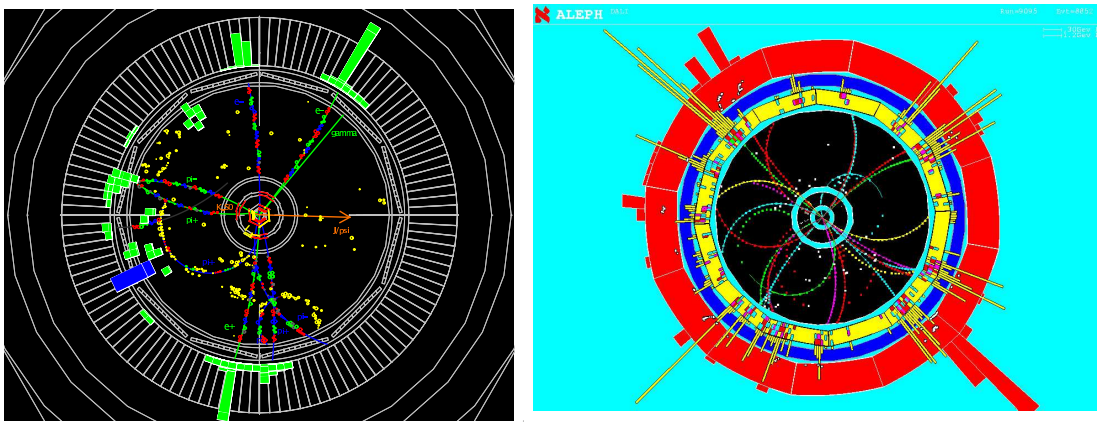

Figure 1.9: Computerized reconstruction of an event containing a $B^0 \to J/\psi K^0_S$ candidate in the *BABAR* detector (left panel) and of a typical event in the ALEPH detector (right panel).

In hadron-hadron collisions, therefore, extraction of heavy-flavor signals from the large background is the first experimental challenge. In particular, the role of the trigger system is crucial for efficiently collecting heavy-flavor decays and reject overwhelming backgrounds in the high collision-rate environment typical of hadron colliders. At the Tevatron, selecting final states containing single or dileptons of low to moderate transverse momenta allows isolating large samples of b-hadron decays, because semileptonic ($B \to \ell^- \bar{\nu}_\ell X$, where $\ell = e,\ \mu$) and *charmonium* ($B \to J/\psi X \to [\ell^+\ell^-]X$) decays account for approximately 20% of total b-hadron widths and have recognizable experimental signatures. In order to exploit them, lepton identification capability and accurate reconstruction of charged-particle trajectories are required. But the most powerful discriminant for heavy flavor decays against background is their relatively long ($\approx$ 1.5 ps, for $B^0_{(s)}$ mesons) lifetime. Heavy flavors

produced at the Tevatron have momenta whose component in the plane transverse to the beam-line is often large enough to allow measuring the displacement between their production and decay positions (≈ 500 μm), provided a vertex detector of sufficient resolution. This feature allows rejecting a large fraction of light-quark background while keeping significant fractions of signal. High-precision measurement of particle momenta and accurate reconstruction of decay vertices are mandatory for an effective discrimination based on decay lengths.

Before the current period of Tevatron operation, final-state leptons were used at the trigger level for a first pre-selection of samples enriched in heavy-flavor decays, while the vertex displacements criteria were applied only in the data analysis, to further enhance the signal purities. This choice, dictated by instrumental limitations, excluded the non-leptonic heavy flavor decays from the reach of experiments at hadron colliders.

Since March 2001, for the first time at a hadron collider, the CDF experiment exploits high-resolution identification of charged particles produced in decays displaced from beam line at trigger level (see sec. 2.5.2). This unique capability allows collection of samples of hadronic heavy flavor decays (e. g., $B^0_{(s)} \to h^+h'^-$, $B^0_{(s)} \to D^-_{(s)}\pi^+$, etc.) of unprecedented size and purity, disclosing a broad program of measurements, such the one described in this thesis, previously inaccessible.

A next generation of experiments dedicated to studying b-physics in hadron collisions will soon follow the CDF experience. The "Large Hadron Collider b" experiment (LHCb) is currently being set up to study pp collisions at $\sqrt{s} = 14$ TeV [57]. The LHCb detector is a single-arm spectrometer designed to exploit the large production cross-section and the high spatial correlation (i. e., large acceptance for both b-hadrons) of $b\bar{b}$ pairs emitted at small angles with respect to the LHC beams. Starting in 2008, a $b\bar{b}$ production rate of 0.1–1 MHz is expected during five years of operations, in which approximately 2 fb^{-1} per year will be collected. The $\mathcal{O}(10^2$ GeV$/c)$ momentum of b-hadrons combined with an hodoscope of silicon micro-strip detectors will ensure precise reconstruction of decay vertices with minimal multiple scattering. A dipole magnet will produce an integrated field $\int Bdl = 4$ T$\cdot$m to bend the trajectories of charged particles, whose momentum will be measured using a combined system of silicon micro-strip detectors and straw-tube drift cells. Two ring-imaging Cherenkov detectors will provide charged hadron identification, crucial for separation of kinematically similar signal modes and for identifying the flavor of the b-hadron at production. Sampling calorimeters using scintillating tiles and lead or steel absorbers will measure the energy of electrons (for reconstructing semileptonic and charmonium decays), photons (final states with neutral pions and eta mesons, radiative decays), and hadrons. Multi-wire proportional chambers and gas-electron multipliers will detect penetrating charged particles for muon identification (semileptonic, charmonium, and rare decays).

1.7 Introduction to the analysis

1.7.1 Current experimental status on $B^0_{(s)} \to h^+h'^-$ decays

In 1993, the first evidence of decays of a b-meson into two charmless pseudo-scalar mesons was claimed by the CLEO experiment, which observed $13.6^{+4.7}_{-3.9}$ events of type $B^0 \to h^+h'^-$ in 1.47×10^6 $B\overline{B}$ pairs produced in $\int \mathcal{L}dt = 1.37$ fb^{-1} of e$^+$e$^-$ collisions at the Υ(4S) resonance [58]. Later, in 1998, they quoted the first measurement of an individual branching fraction, $\mathcal{B}(B^0 \to K^+\pi^-) = [15^{+5}_{-4}$ $(stat.) \pm$

	$e^+e^- \to \Upsilon(4S) \to \overline{B}B$	$e^+e^- \to Z^0 \to b\bar{b}$	$p\bar{p} \to b\bar{b}X$
Accelerator	CESR, PEP-II, KEKB	LEP	Tevatron
Detector	CLEO, *BABAR*, Belle	ALEPH, DELPHI, L3, OPAL,	CDF II, DØ
$\sigma(b\bar{b})$	1 nb	7 nb	50 μb
$\sigma(b\bar{b})/\sigma(\text{bck})$	0.26	0.22	0.001
Typical $b\bar{b}$ rate	10 Hz	0.1 Hz	5–10 kHz
Flavors	B^0 (50%), B^+ (50%)	B^0 (40%), B^+ (40%) B_s^0 (10%), B_c^+ ($< 0.1\%$), b-baryons (10%)	B^0 (40%), B^+ (40%) B_s^0 (10%), B_c^+ ($< 0.1\%$), b-baryons (10%)
Boost $\langle\beta\gamma\rangle$	0.06–0.6	6	1–4
Pile-up events	0	0	1–5
Track multiplicity	~ 5	~ 10	~ 30
Trigger	Inclusive	Inclusive	displaced tracks, ℓ

Table 1.1: Schematic comparison of some relevant parameters for b-physics measurements in different experimental environments. "DELPHI" stands for "DEtector with Lepton, Photon, and Hadron Identification", "OPAL" for "Omni-Purpose Apparatus at LEP". All numerical values are approximate.

$1.4\ (syst.)] \times 10^{-6}$, using 3.3×10^6 $B\overline{B}$ pairs, corresponding to $\int \mathcal{L}dt = 3.14\ \text{fb}^{-1}$ [59]. This large decay-rate provided the first experimental indication that $b \to sg$ penguin amplitudes give an important contribution to two-body charmless b-meson decays. A confirmation of this hierarchy was obtained in 2000, when the $B^0 \to \pi^+\pi^-$ decay was observed and its branching fraction measured to be $\mathcal{B}(B^0 \to \pi^+\pi^-) = [4.6^{+1.6}_{-1.4} \pm 0.5] \times 10^{-6}$, using a larger sample of 9.66×10^6 $B\overline{B}$ pairs ($\int \mathcal{L}dt = 9.13\ \text{fb}^{-1}$) [60]. From the same sample, the CLEO Collaboration quoted also the first measurement of partial rate asymmetry in $B^0 \to K^+\pi^-$ decays: $A_{\mathsf{CP}}(B^0 \to K^+\pi^-) = -0.04 \pm 0.16\ (stat.) \pm 0.02\ (syst.)$ [61].

More recently, $B\overline{B}$ samples $\mathcal{O}(50)$ times larger, available at the *BABAR* and Belle experiments, allowed measurements of time-integrated decay-rates of impressive precision, and the first measurements of time-dependent decay-rates (see tab. 1.2).

		BABAR	Belle
$A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$		$-0.108 \pm 0.024 \pm 0.008$	$-0.093 \pm 0.018 \pm 0.008$
$\mathcal{B}(B^0 \to K^+\pi^-)$	$[10^{-6}]$	$19.7 \pm 0.6 \pm 0.6$	$20.0 \pm 0.4 \pm 0.8$
$\mathcal{B}(B^0 \to \pi^+\pi^-)$	$[10^{-6}]$	$5.8 \pm 0.4 \pm 0.3$	$5.1 \pm 0.2 \pm 0.2$
$\mathcal{B}(B^0 \to K^+K^-)$	$[10^{-6}]$	< 0.40 @ 90% CL	< 0.25 @ 90% CL

Table 1.2: Time-integrated decay-rates for the modes of interest as measured by the *BABAR* [62, 63] and the Belle [64, 65, 66] experiments, as of Summer 2006.

While charmless hadronic decays of B^0 (and B^+) mesons have been extensively explored experimentally in the last decade, significantly less information is available about the B_s^0 meson. No charmless decays of a B_s^0 meson into kaons or pions has yet been observed. The only information

currently available are the following upper limits (at the 90% CL) on branching fractions, obtained by the ALEPH Collaboration in a sample of four million hadronic Z^0 decays produced in e^+e^- collisions at the Z^0 resonance: $\mathcal{B}(B^0_s \to K^+K^-) < 59 \times 10^{-6}$, $\mathcal{B}(B^0_s \to K^-\pi^+) < 210 \times 10^{-6}$, and $\mathcal{B}(B^0_s \to \pi^+\pi^-) < 170 \times 10^{-6}$ [67].

CDF has, therefore, the first and unique opportunity to improve our understanding of b-meson decays by complementing the rich experimental data available for charmless hadronic decays of B^0 and B^+ mesons with measurements of similar modes for the B^0_s meson.[8]

Joint study of B^0 and B^0_s decays was originally envisaged also at the Hadron Electron Ring Accelerator facility (HERA-B) single-magnet forward spectrometer at the Deutsches Elektronen-Synchrotron [68]. They expected to reconstruct approximately 800 $B^0 \to \pi^+\pi^-$ decays per year in collision of 820 GeV/c protons on a fixed wire target. However, detector-related problems delayed the commissioning of the experiment and no result on $B^0_{(s)} \to h^+h^{'-}$ decays has been claimed yet by the HERA-B Collaboration.

1.7.2 Analysis overview

All branching fractions measured in the present work are relative (i. e., ratios of branching fractions) rather than absolute branching fractions.

The primary motivation for a relative measurement is that ratios of branching fractions of $B^0_{(s)} \to h^+h^{'-}$ decays have often more physics relevance than absolute measurements. They are usually easier to predict and to interpret theoretically than absolute quantities, especially if amplitude relations based on Standard Model symmetries allow (partial) cancellation of hadronic unknowns in the ratios, as discussed in sec. 1.4 and 1.5.

The second, important motivation for a relative measurement is the cancellation of several systematic effects. Consider the expression for the measurement of an absolute branching-fraction (at the Tevatron) of, say, the $B^0_s \to K^+K^-$ decay:

$$\mathcal{B}(B^0_s \to K^+K^-) = \frac{N(B^0_s \to K^+K^-)}{2\varepsilon \times \sigma(\mathrm{p\bar{p}} \to b\bar{b} + X) \times f_s \times \int \mathcal{L}dt}, \tag{1.84}$$

where $N(B^0_s \to K^+K^-)$ is the yield of reconstructed candidates, ε indicates the total trigger, reconstruction, and analysis efficiencies, $\sigma(\mathrm{p\bar{p}} \to b\bar{b} + X)$ is the cross-section for hadroproduction of b-quarks, f_s is the probability that a b-quark produced in $\mathrm{p\bar{p}}$ collisions hadronizes into a B^0_s meson, and $\int \mathcal{L}dt$ is the integrated luminosity corresponding to the size of the data sample. All these factors are affected by significant uncertainties. These can have large effect on absolute measurements, whereas partially or completely cancel in relative measurements:

efficiency – the knowledge of the total efficiency for each mode, which is $\mathcal{O}(‰)$, is affected by important uncertainties. These are partly due to the frequent changes (improvements) in detector and trigger configuration which took place during the collection of the first CDF data, on which this analysis is based. On the other hand, even in stable trigger and detector configurations,

[8] The future role of the Belle experiment when KEKB will be operating at a center of mass energy of the $\Upsilon(5S)$ resonance is discussed in sec. 10.7.

there is an intrinsic difficulty in measuring with the accuracy needed for an absolute measurement the trigger and reconstruction efficiencies for (relatively) low momentum particles in the complicate environment of hadron collisions. The need to reject backgrounds several orders of magnitude larger than the signals, results in triggers designed to optimize purity at the expense of efficiency. Trigger and reconstruction efficiencies are difficult to measure with high accuracy. Common reconstruction efficiencies and systematic uncertainties cancel in relative measurements involving kinematically similar modes. This yields reduced residual uncertainties that are easier to determine and less sensitive to changes of the trigger and detector configurations.

Production cross-section – the ($p_{\mathrm{T}}(B)$-dependent) value of the production cross-section is affected by a $\approx 15\%$ uncertainty [56]. This reflects into a systematic uncertainty in absolute branching-fraction measurements, and cancels completely in a relative measurement.

Fragmentation fraction – the 2.5%–10% uncertainty on the probability for a b-quark to hadronize to the desired hadron [18] becomes a systematic uncertainty in an absolute measurement, whereas it cancels completely (partially) in relative measurements between $B^0_{(s)}$ mesons of same (different) flavor.

Luminosity – the irreducible 6% uncertainty in the CDF integrated luminosity (see sec. 2.4.4), would directly propagate as a systematic uncertainty on the measurement of absolute branching fractions, whereas it cancels completely in a relative measurement.

To ensure control of remaining systematic uncertainties, we adopted a data-driven approach. Whenever possible, we relied on collision data as opposed to computer simulation using pseudo-random sampling (i. e., "Monte Carlo"). This choice is rewarding in terms of reliability of results, at the price of the additional difficulty of reconstructing several control samples. In fact, with current Monte Carlo programs, it is extremely difficult to reliably simulate b-hadron production and associated backgrounds in a hadron collision environment at the level of accuracy required by this kind of measurements. Results from simulation are either strongly model-dependent and plagued by large systematic effects, or require extensive tunings based on data, thereby losing the benefits of simulation. We limited the use of Monte Carlo simulation to the strictly necessary, namely the modeling of the *kinematic of signal* decay-modes and the evaluation the *relative acceptances* between different signal modes. This choice provides reliable results because the decay kinematics of signal and the detector geometry are accurately reproduced by the standard Monte Carlo simulations. Modeling of background distributions and evaluation of other efficiencies was based on data.

The first challenge of the analysis is the extraction of the $B^0_{(s)} \to h^+h^{'-}$ signal, starting from a $\mathcal{O}(10^{-9})$ signal-to-background ratio at production: the combination of a dedicated trigger (sec. 2.5), with an analysis selection optimized in an unbiased way, allowed the first reconstruction ever of a $B^0_{(s)} \to h^+h^{'-}$ signal at a hadron collider. But this signal contains contributions from several, kinematically similar, B^0 and B^0_s meson decay-modes ($B^0 \to K^+\pi^-$, $B^0 \to \pi^+\pi^-$, $B^0_s \to K^+K^-$, ...), overlapping with unknown proportions in a single mass peak. For the desired measurement of decay rates, we needed to discriminate among the different contributions. This second challenge required the development of a sophisticated multivariate likelihood fit because the available resolution in mass and in particle identification is insufficient for an event-by-event separation of the decays.

We exploited the kinematic differences between modes and the measurement of specific ionization of charged particles (dE/dx) in the drift chamber, to determine the composition of signal and background. This required a detailed calibration of the dE/dx with large samples of $D^{*+} \to D^0\pi^+ \to [K^-\pi^+]\pi^+$ decays. In order to infer the desired measurements of branching fractions from the observed numbers of signal events, we had to correct for the relative efficiency associated to the reconstruction of different decay modes.

After the discussion of the experimental apparatus in the next chapter, the description of the data analysis occupies the rest of this thesis.

Chapter 2

The upgraded Collider Detector at the Fermilab Tevatron

The data used for this work were collected by the upgraded Collider Detector at the Fermilab Tevatron collider (CDF II). This chapter provides a concise description of the complex infrastructure, accelerator and detector, involved in producing the data sample. A more detailed description of the tracking and the trigger systems is given, for the crucial role they have in the present analysis.

2.1 The Fermilab Tevatron collider

The Tevatron collider is currently the world highest energy accelerator. It provides collisions of antiprotons with protons at a center-of-mass energy of 1.96 TeV. The Tevatron is an underground circular proton synchrotron, 1 km in radius, the last stage of a system of accelerators, storage rings, and transfer lines located at the Fermi National Accelerator Laboratory (FNAL or Fermilab), about 50 km west from Chicago, Illinois, United States. While the machine operates in collider mode, "bunches" of protons, circulating clockwise as seen from above and spaced by 396 ns, collide against a similar beam of antiprotons accelerated counter-clockwise, both at energies of 980 GeV. A bunch is a collection of particles contained within one radio-frequency "bucket" (defined below).

The Energy Doubler, now known as the Tevatron, was commissioned in 1983 as the first large-scale superconducting synchrotron in the world. The first $p\bar{p}$ collisions occurred in 1985 and, since then, various periods of collider operations alternated with fixed-target operations or shut-down periods for upgrades of the machine. Each period of Tevatron collider operations is conventionally identified as a *Run*.[1] Table 2.1 contains a summary of the Tevatron operations and performance since its construction. The present analysis uses the first physics-quality data collected in Run II.

[1] The Run is not to be confused with the *run*, defined in CDF as a continuous period of data-taking in approximately constant detector and beam conditions.

Date		$\sqrt{s}$ [TeV]	$\mathcal{L}$ [cm^{-2}s^{-1}]	$\int \mathcal{L}dt$ [pb^{-1}]
Mar 1983	End of the construction	–	–	–
Jul 1983	Proton energy 512 GeV	–	–	–
Oct 1983	Fixed-target program began	–	–	–
Feb 1984	Proton energy 800 GeV	–	–	–
Oct 1985	CDF observed first p$\bar{\text{p}}$ collisions	1.6	few 10^{24}	–
Oct 1986	Proton energy 900 GeV	–	–	–
Jun 1988–May 1989	Run 0: first physics collider Run	1.8	2×10^{30}	$\simeq 4.5$
Aug 1992–Feb 1996	Run I	1.8/0.63	28×10^{30}	$\simeq 180$
Aug 2000	Beam energy 980 GeV	–	–	–
Mar 2001	Run II began	1.96	5×10^{30}	–
Sep 2006	Record luminosity	1.96	229×10^{30}	$\simeq 1800$

Table 2.1: Chronological overview of the Tevatron operation and performance. The third column reports the peak luminosity. The fourth column reports the delivered integrated luminosity. The last row shows the performance as of this writing.

The performance of the Tevatron collider is evaluated in terms of two key parameters: the available center-of-mass energy, $\sqrt{s}$, and the instantaneous luminosity, $\mathcal{L}$. The former defines the accessible phase-space for the production of resonances in the final states. The latter is the coefficient of proportionality between the rate of a given process and its cross-section σ:

$$\text{rate [events s}^{-1}] = \mathcal{L} \text{ [cm}^{-2}\text{s}^{-1}] \times \sigma \text{ [cm}^2].$$

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

$$n(T) = \int_0^T \mathcal{L}\sigma dt. \tag{2.1}$$

Assuming an ideal head-on p$\bar{\text{p}}$ collision with no crossing angle between the beams, the instantaneous luminosity is defined as

$$\mathcal{L} = 10^{-5} \frac{N_{\text{p}} N_{\bar{\text{p}}} B f \beta\gamma}{2\pi\beta^* \sqrt{(\epsilon_{\text{p}} + \epsilon_{\bar{\text{p}}})_x (\epsilon_{\text{p}} + \epsilon_{\bar{\text{p}}})_y}} \mathcal{H}(\sigma_z/\beta^*) \qquad [10^{30}\text{cm}^{-2}\text{s}^{-1}]. \tag{2.2}$$

It depends on the following Tevatron parameters: the number of circulating bunches in the ring ($B = 36$), the revolution frequency ($f = 47.713$ kHz), the Lorentz relativistic factor (boost, $\beta\gamma = 1045.8$ at 980 GeV), the average numbers of protons ($N_{\text{p}} \approx 250 \times 10^9$) and antiprotons ($N_{\bar{\text{p}}} \approx 25 \times 10^9$) in a bunch, an empiric "hourglass" factor ($\mathcal{H} = 0.6$–0.7), which is a function of the ratio between the longitudinal r.m.s. width of the bunch ($\sigma_z \approx 60$ cm) and the "beta function" calculated at the interaction point ($\beta^* \approx 31$ cm), and the 95% normalized emittances of the beams ($\epsilon_{\text{p}} \approx 18\pi$ mm mrad and $\epsilon_{\bar{\text{p}}} \approx 13\pi$ mm mrad after injection).[2] The dominant limiting factor of the luminosity is the availability of monochromatic antiprotons that can be efficiently transferred through the accelerator

[2]The hourglass factor is a parameterization of the longitudinal profile of the beams in the collision region, which assumes the shape of an horizontal hourglass centered in the interaction region. The beta function is a parameter

chain for final collisions. The current Tevatron performance benefits from an increase of a factor of 2.5 in this critical parameter, with respect to the beginning of Run II.

The Tevatron is an approximately circular synchrotron employing 772 dipole, 2 half-dipole, and 204 quadrupole superconducting magnets. Each is approximately 6 m long, 4 tons in mass, and is made of NbTi alloy filaments embedded in copper, kept at 4.3 K temperature by a large cryogenic system. A 4400 A current flows through each magnet to produce the 4.2 T magnetic field necessary to keep the particles on their orbit, while they are accelerated by eight radio-frequency cavities (RF) driven at approximately 53.105 Hz. Motions or friction by the approximately 4000 N/cm of outward pressure are avoided by epoxy-covered steel collars bound around the magnets.

The particles are accelerated through the RF buckets. A bucket is one interval of the longitudinal restoring force provided by the RF cavities that results in a stable phase-space where a bunch may be captured and accelerated. The Tevatron provides beams for experiments in different modes (fixed-target, collider, etc.). For the purpose of the present analysis, we describe the procedure for obtaining a continuous period of collider operation using the same collection of protons and antiprotons, called a *store*. Further details can be found in Ref. [69].

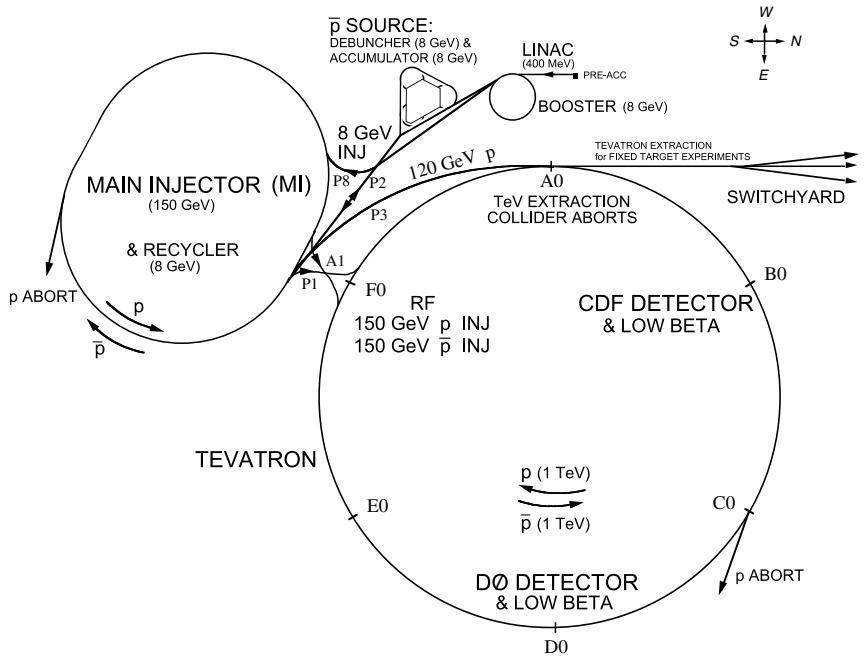

Figure 2.1: Illustration of the Fermilab Tevatron collider.

convenient for solving the equation of motion of a particle through an arbitrary beam transport system. The emittance ϵ measures the phase-space occupied by the particles of the beam. Three independent two-dimensional emittances are defined. The quantity $\sqrt{\beta\epsilon}$ is proportional to the r.m.s. width of the beam in the corresponding phase plane. On-line measurements of the transverse emittances are performed at the Tevatron with various methods, including flying through the beam a 7 μm wire and by measuring the cascade of losses, which is proportional to the beam intensity, or detecting the synchrotron light radiated by the particles at the edge of a dipole magnet.

2.1.1 Proton production

Hot hydrogen gas is passed through a *magnetron*, which extracts a 50–55 mA current of 15–22 keV H^- ions, subsequently accelerated every 66 milliseconds to 750 keV by a three-staged diode-capacitor voltage multiplier (Cockroft-Walton) accelerator. The proton beam, segmented into bunches, is then injected into a two-staged 150 m long linear accelerator (Linac, see fig. 2.1). First, a drift tube accelerator resonating at 201.249 MHz accelerates bunches of protons up to 116 MeV; then, a side-coupled cavity accelerator at 804.996 MHz increases their energy to 401.5 MeV before injection into the Booster.[3] The Booster (see fig. 2.1) is an alternating gradient synchrotron (orbit radius of 75.5 m) that accelerates protons to 8 GeV in 33 milliseconds, sweeping from ≈ 38 to 53.105 MHz. At injection, a thin carbon foil is used to strip the electrons from the H^- ions to obtain protons. Injecting H^- ions rather than protons into the Booster allows the injection to proceed over multiple revolutions of the beam around the Booster Ring (usually 10–12). If protons were instead injected, the magnetic field used to inject new protons onto orbit in the Booster would also deflect the already revolving protons out of orbit. There are two basic modes during collider operations: antiproton accumulation and injection.

2.1.2 Antiproton production and accumulation

In accumulation mode, one set of 84 proton bunches (approximately 8×10^{12} p in total) is extracted from the Booster at 8 GeV and injected into the Main Injector every 2.2 seconds.[4] The Main Injector (see fig. 2.1) is a 53.105 MHz circular synchrotron (528.5 m in radius), with 18 accelerating cavities and conventional magnets. The protons are accelerated to 120 GeV, and then extracted and directed to the antiproton production station, a rotating 7 cm-thick target made of nickel alloys containing chromium, iron and other metals. The particles produced in the interaction are spatially wide-spread. They are collected and focused with a cylindrical lithium lens (760 T/m).[5] Eight GeV/c negatively-charged secondaries are momentum-selected by a 1.5 T pulsed dipole magnet. Typically, 21 antiprotons are collected for each 10^6 protons on target, resulting in a stacking rate of approximately 10–20 mA/h. The emerging antiprotons have a bunch structure similar to the one of the incident protons and are delivered to the Debuncher storage ring (see fig. 2.1). This rounded triangular synchrotron, 90 m of mean radius, transforms the antiproton pulses in a continuous beam of monochromatic antiprotons. Stochastic cooling and bunch rotation are applied during many cycles.[6] From the Debuncher, 8 ± 0.018 GeV antiprotons are transferred with 60%–70% efficiency into the Accumulator, a concentric storage ring 75 m in mean radius (see fig. 2.1), where they are stacked and cooled

[3] The side-coupled technique avoids the need of increasing the drift-tube lengths by producing a traveling wave that moves along with the ions, keeping them into the accelerating phase.

[4] The novel "slip-stacking" technique, implemented in 2005, allows grouping two (instead of one) sets of protons bunches, for a greater antiproton production-rate.

[5] Lithium is used to minimize beam loss from multiple-scattering.

[6] Stochastic cooling is a technique used to reduce the transverse and energy spread of a particle beam without any accompanying beam-loss. This is achieved by applying iteratively a feedback mechanism that senses with extreme sensitivity the beam deviation from the ideal orbit with electrostatic plates, processes and amplifies it, and transmits an adequately-sized synchronized correction pulse to another set of plates downstream [70]. Bunch rotation is an RF manipulation technique that, using adequate phasing, transforms a beam with a large time spread and a small energy spread in a beam with a large energy spread and a small time spread, or viceversa.

with a variety of systems until the maximum antiproton intensity is reached. Since 2004, optimized antiproton accumulation is achieved using the Recycler Ring (see fig. 2.1). This is a constant 8 GeV-energy storage-ring placed in the Main Injector enclosure, that uses permanent magnets (magnetized strontium ferrite, the same as in refrigerators). It is used to gather antiprotons that are periodically transferred from the Accumulator (95% transfer efficiency) thus maintaining it at its optimum intensity regime. Recently, relativistic electron cooling was successfully implemented in the Recycler, further enhancing the Tevatron performance [71].[7]

2.1.3 Injection and collisions

Every 10–20 h, antiproton accumulation is stopped in preparation for injection. A set of seven proton bunches is extracted from the Booster, injected into the Main Injector, accelerated to 150 GeV, coalesced with $\approx$ 90% efficiency into a single bunch of $\approx 300 \times 10^9$ p, and then injected into the Tevatron.[8] This process is repeated every 12.5 seconds, until 36 proton bunches, separated by 396 ns, are loaded into the Tevatron central orbit. Typically, 65% of the protons in the Main Injector are successfully transferred to the Tevatron. The electrostatics separators (about 30 pairs of metal plates) are then activated in the Tevatron, in preparation for antiproton injection.

Four sets of 7–11 $\bar{\text{p}}$ bunches are extracted from the Accumulator (or from the Recycler) to the Main Injector, accelerated to 150 GeV, coalesced with $\approx$ 80% efficiency into four 30×10^9 $\bar{\text{p}}$ bunches separated by 396 ns, and then injected into the Tevatron, where protons are counter-rotating. Protons and antiprotons circulate in the same enclosure, sharing magnet and vacuum systems. The separators minimize the beam-beam interactions, by keeping the proton and the antiproton beams, each about half a millimeter thick, into two non-intersecting closed helical orbits separated by approximately five millimeters (3σ–5σ) as they revolve in opposite directions. This allows controlling each beam nearly independently. The injection process is repeated nine times until 36 antiproton bunches circulate in the Tevatron.

Sweeping the Tevatron RF by $\approx$ 1 kHz, the beam is then accelerated in about a minute from 150 to 980 GeV, at which energy one particle completes the full revolution of the Tevatron circumference in 21 μs at $0.9999996c$. The beams are finally brought into collision at the two instrumented interaction-points located along two straight sections of the Tevatron: DØ and BØ, where the DØ and CDF II detectors, respectively, are located. Although the power produced in the collision is only 1–2 W, the stored energy of the beam is about 1.7 MJ, corresponding, approximately, to the kinetic energy of a 4.5 ton truck moving at 100 km/h. Special high-power quadrupole magnets ("low-β squeezers"), installed on the beam pipe at either side of the detectors, reduce the transverse spatial spread of the beams to maximize the collision rate in the interaction regions. The resulting transverse spatial distribution of the luminous region is approximately a two-dimensional Gaussian, with $\sigma_{\text{T}} \approx 30$ μm. The typical longitudinal dimension of a bunch is 60–70 cm. The interaction regions have a roughly Gaussian distribution along the beam direction, with r.m.s. width $\sigma_z \approx 28$ cm.[9] The center of the

[7] Electron cooling is a method of damping through the interaction between the antiproton beam and an electron beam propagating together at the same average velocity.

[8] Coalescing is the process of compacting into one dense bunch many smaller bunches.

[9] Whereas one may expect a bunch length $\sigma_z \approx 60$ cm to distribute $\text{p}\bar{\text{p}}$ interactions over a length of $60/\sqrt{2}$ cm, this length is in fact less than 30 cm owing to the variation of the transverse beam profile along the beam (z) axis.

luminous region is shifted toward the nominal interaction point by fine tuning of the squeezers. The 36 bunches of protons (antiprotons) are distributed among the 1113 buckets in three equispaced "trains" of 12 bunches each. The inter-bunch spacing is 396 ns (21 buckets) within a train, while a 2.6 μs spacing (139 buckets, "abort gap") is kept between trains. The need for the abort gap is two-fold: it allows antiprotons injection (in coincidence with the proton abort gap) without perturbing the already revolving protons with the injecting magnet. Furthermore, when beam abortion is needed, the abort gap allows ramping-up the deflecting magnets without interfering with the beam during the transient, possibly damaging the detectors. As a consequence of this bunch distribution, the average bunch-crossing rate is 1.7 MHz, resulting from a 2.53 MHz rate, when the proton and antiproton trains are crossing, and zero rate in correspondence of the abort gaps.

The transverse profile of the beam is shaped to its optimized configuration to avoid detector damage from the tails of the p ($\bar{\mathrm{p}}$) distributions interacting with the beam pipe: retractable collimators (iron plates) are moved perpendicularly toward the beam and trim-off the residual halo. When the beam profile is narrow enough and the conditions are safely stable, the detector is powered and the data-taking starts.

The number of overlapping inelastic interactions N for each bunch crossing is a Poisson-distributed variable that depends on the instantaneous luminosity. The observed distribution of the multiplicity of interaction vertices yields $\bar{N} \approx 0.2$, 1.0, 2.0, and 6.0 for respectively, $\mathcal{L} \approx 1 \times 10^{31}$, 5×10^{31}, 10×10^{31}, and 30×10^{31} luminosities [72]. The luminosity decreases as a function of time because of the interactions of the beam with residual molecules of gas that escaped the vacuum of the beam pipe, beam-halo interactions, and $\bar{\mathrm{p}}$ depletion due to the collisions. The luminosity profile can be empirically modeled by the sum of two exponential functions:

$$\mathcal{L}(t) = \mathcal{L}_0 \left(c_S e^{-t/t_S} + c_L e^{-t/t_L} \right), \tag{2.3}$$

where the c_i, and t_i parameters depend on the initial luminosity ($\mathcal{L}_0$). Their values are typically $c_S \approx 0.4$ and $t_S \approx 4$ h, for the short-lived component, and $c_L \approx 0.6$ and $t_L \approx 23$ h, for the long-lived one. During the 10–20 h of a store, the luminosity decreases by a factor of 2.5–5, the majority of data being collected at $\mathcal{L} \approx \mathcal{L}_0/2$. Just after the final injection, a new antiproton accumulation cycle is started. When the antiproton stack is sufficiently large and the colliding beams are degraded, the detector high-voltages are switched-off and the store is dumped. The beam is extracted via a switch-yard and sent to an absorption zone.

Beam abortion can occur also accidentally when a superconducting magnet rises its temperature above the critical value (i. e., the magnet "quenches"), destroying the orbit of the beams.[10] The time between the end of a store and the beginning of collisions of the next one is typically 2 h, during which time calibrations of the subdetectors and cosmic rays tests are usually performed.

2.1.4 Tevatron performance

Since the beginning of Run II the Tevatron performance has been steadily increasing. Currently (as of 2006), the Tevatron is running at a center-of-mass energy of 1.96 TeV with an inter bunch-crossing

[10] During a quench, one cubic liter of liquid helium expands to 700 cubic liters of gaseous helium within a quarter of a second. More than 1500 custom-made relief valves allow the Tevatron cryogenic system to manage the rapid expansion.

time of 396 ns. The original plan of shortening the inter bunch-crossing to 132 ns, to reduce pile-up events, has been discarded. The Tevatron set the world record of highest peak luminosity for a hadron collider of 229.0×10^{30} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (September, 20^{st} 2006). As of this writing, a single store lasts for about 20–30 h yielding ≈ 6 pb^{-1} integrated luminosity. Antiprotons are accumulated at $10^{11}\bar{\mathrm{p}}$/h rates and transferred through the accelerator chain with average 75% overall efficiency for the next store. The Tevatron delivers typically collisions corresponding to 30 pb^{-1} per week. As of October 2006, physics quality data corresponding to 1.6 fb^{-1} are stored on permanent memories.

2.2 The CDF II detector

The CDF II detector is a large multi-purpose solenoidal magnetic spectrometer surrounded by 4π fast, projective calorimeters and fine-grained muon detectors. It is installed at the BØ interaction point of the Tevatron (see fig. 2.2) to determine energy, momentum and, whenever possible, the identity of a broad range of particles produced in 1.96 TeV $\mathrm{p\bar{p}}$ collisions. It was designed, built, and operated by a team of physicists, technicians, and engineers that, as of this writing, spans 60 institutions of 13 countries, and includes approximately 618 official members. Several upgrades modified the design of the original facility commissioned in 1985.[11] The most extensive upgrade started in 1995 and led to the current detector whose operation is generally referred to as Run II.

2.2.1 Coordinates and notation

CDF II employs a right-handed Cartesian coordinate system with origin in the BØ interaction point, assumed coincident with the center of the drift chamber (see sec. 2.3.5). The positive z-axis lies along the nominal beam-line pointing toward the proton direction (east). The (x, y) plane is therefore perpendicular to either beams, with positive y-axis pointing vertically upward and positive x-axis in the horizontal plane of the Tevatron, pointing radially outward with respect to the center of the ring.

Since the colliding beams of the Tevatron are unpolarized, the resulting physical observations are invariant under rotations around the beam line axis. Thus, a cylindrical (r, ϕ, z) coordinate system is particularly convenient to describe the detector geometry. Throughout this thesis, *longitudinal* means parallel to the proton beam direction (i. e., to the z-axis), and *transverse* means perpendicular to the proton beam direction, i. e., in the $(x, y) \equiv (r, \phi)$ plane.

Since protons and antiprotons are composite particles, the actual interaction occurs between individuals partons (valence or sea quarks and gluons) contained within them. Each parton carries a varying fraction of the (anti)proton momentum, not known on an event-by-event basis. As a consequence of the possible imbalance in the longitudinal components of the momenta of interacting partons, possible large velocities along $\hat{z}$ for the center-of-mass of the parton-level interaction may occur. In hadron-collisions environments, it is customary to use a variable invariant under $\hat{z}$ boosts as an unit of relativistic phase-space, instead of the polar angle θ. This variable is the *rapidity* defined as

$$Y = \frac{1}{2} \ln \left[\frac{E + p\cos(\theta)}{E - p\cos(\theta)} \right], \tag{2.4}$$

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

where $(E, \vec{p})$ is the energy-momentum four-vector of the particle.[12] Under a $\hat{z}$ boost to an inertial frame with velocity β, the rapidity of a particle transforms linearly, according to $Y \rightarrow Y' \equiv Y + \tanh^{-1}(\beta)$, therefore Y is invariant since $dY \equiv dY'$. However, a measurement of rapidity still requires a detector with accurate particle identification capabilities because of the mass term entering E. Thus, for practical reasons, it is often preferred to substitute Y with its approximate expression η in the ultra-relativistic limit, usually valid for products of high-energy collisions:

$$Y \overset{p \gg m}{\longrightarrow} \eta + \mathcal{O}(m^2/p^2), \tag{2.5}$$

where the *pseudo-rapidity* $\eta \equiv -\ln[\tan(\theta/2)]$ is only function of the momenta. As the event-by-event longitudinal position of the actual interaction is distributed around the nominal interaction point with 30 cm r.m.s. width, it is useful to distinguish *detector pseudo-rapidity*, $\eta_{\rm det}$, measured with respect to the (0, 0, 0) nominal interaction point, from *particle pseudo-rapidity*, η, which is measured with respect to the z_0 position of the real vertex where the particle originated.[13]

Mapping the solid angle in terms of (pseudo)-rapidity and azimuthal angle is also convenient because the density of final-state particles in energetic hadronic collisions is approximately flat in the (Y, ϕ) space. The (pseudo)-rapidity dependence was observed experimentally, the azimuthal dependence derives from the unpolarized beams.[14] Other convenient variables used are the transverse component of the momentum with respect to the beam axis ($p_{\rm T}$), the "transverse energy" ($E_{\rm T}$), and the approximately Lorentz-invariant angular distance ΔR, defined as

$$\vec{p}_{\rm T} \equiv (p_x, p_y) \rightarrow p_{\rm T} \equiv p\sin(\theta), \quad E_{\rm T} \equiv E\sin(\theta), \quad \text{and} \quad \Delta R \equiv \sqrt{\eta^2 + \phi^2}. \tag{2.6}$$

Throughout this thesis, the magnitude of the vector $\vec{p}_{\rm T}$ (and of any vector $\vec{v}$) is indicated as $p_{\rm T}$ (v) , instead of $|\vec{p}_{\rm T}|$ $(|\vec{v}|)$, for a simpler notation.

2.2.2 Overview

CDF II (see fig. 2.2) is a 5000-ton assembly of subdetectors, $\approx$ 16 m in length by $\approx$ 12 m in diameter, which can be moved from its garaged position, in the CDF assembly building, to its operation position on the Tevatron beam line. The 31.4 m move takes one day. The CDF II detector was designed and constructed with an approximately cylindrically symmetric layout both in the azimuthal plane and in the "forward" ($z > 0$, east) "backward" ($z < 0$, west) directions with spatial segmentation of its subcomponents roughly uniform in pseudo-rapidity and azimuth. CDF II is composed of several specialized subsystems, each one designed to perform a different task, arranged in a standard layout for multipurpose detectors; starting from the interaction point, particles emitted within the acceptance region encounter in sequence: a thin wall beryllium vacuum chamber, a high-precision tracking system

[12] The rapidity can be derived from the Lorentz-invariant cross-section: $E\frac{d^3\sigma}{(dp)^3} = E\frac{d^2\sigma}{\pi p_{\rm T} dp_{\rm T} dp_z}$. Observing that only E and p_z change under z boosts, we can replace them by a variable Y such as $E\frac{dY}{dp_z} = 1$. Solving for Y we get eq. (2.4).

[13] An idea of the difference is given by considering that $\eta_{\rm det} \approx \eta \pm 0.2$ if the particle was produced at $z = 60$ cm from the nominal interaction point.

[14] The charged particle density distribution as a function of rapidity is invariant under Lorentz transformation, and was found to be approximately constant (within a wide rapidity range) at about four charged particles per unit of rapidity at the Tevatron.

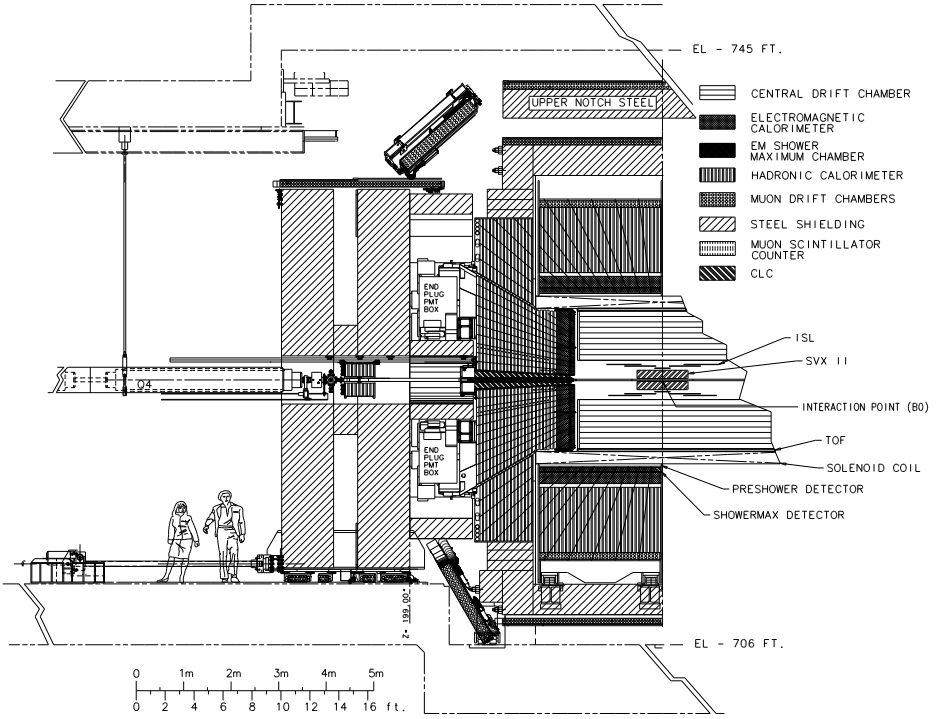

Figure 2.2: Elevation view of one half of the CDF II detector.

composed by an inner silicon system and an outer drift-chamber, a time-of-flight detector, a solenoidal magnet and its return steel yoke, finely segmented sampling calorimeters, and muon detectors.

Its main features are an excellent tracking performance, which provides high mass resolution and precisely reconstructed decay-vertices, good electron and muon identification capabilities combined with charged-hadron identification, and an advanced trigger system that fully exploits the high event-rates. A detailed description of the CDF II detector can be found elsewhere [73] and in specific references cited for each subdetector. In the following, we emphasize the tracking and the trigger systems, which are the aspects of the detector more specific to this analysis.

2.3 The tracking system

Three-dimensional charged particle tracking is achieved through an integrated system consisting of three silicon inner subdetectors and a large outer drift-chamber, all contained in a superconduct-

ing solenoid. The 1.41 T magnetic field and the 130 cm total lever arm provide excellent tracking performances (see fig. 2.3).

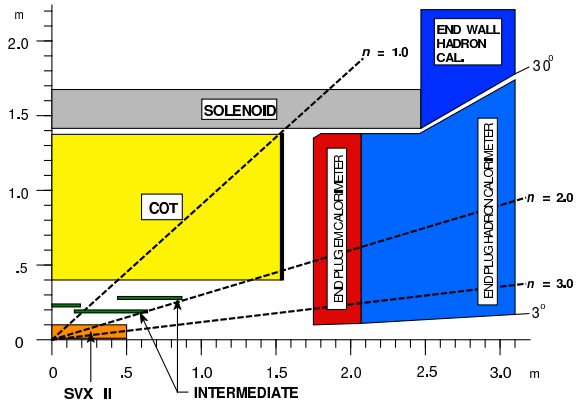

Figure 2.3: Elevation view of one quadrant of the inner portion of the CDF II detector showing the tracking volume surrounded by the solenoid and the forward calorimeters.

In the central region ($|\eta_{\rm det}| \lesssim 1$), 7 silicon samplings (one in the (r,ϕ) view plus six in the (r,ϕ,z) view), and 96 chamber samplings (48 (r,ϕ) plus 48 (r,z)) are available between 1.6 and 132 cm. In the forward and backward regions ($1 \lesssim |\eta_{\rm det}| \lesssim 2$), 8 silicon samplings (one in the (r,ϕ) view plus seven in the (r,ϕ,z) view) are available between 1.6 and 29 cm, along with partial information from the chamber.

The high number of samplings over the 88 cm lever arm of the chamber ensure precise determination of curvature, azimuth, and pseudo-rapidity of the tracks in the central region. The chamber provides also track seeds for pattern-recognition in silicon.

The core of the silicon detector is the Silicon VerteX detector (SVXII). It provides five three-dimensional measurements that extend the lever arm by 41.5 cm toward the beam thus allowing more precise determination of the trajectories and identification of decay-vertices displaced from the beam-line. The SVXII has an outer and an inner extension.

The outer extension, i. e., the Intermediate Silicon Layers (ISL), provides a single (double) three-dimensional silicon measurement in the central (forward-backward) region, at intermediate radial distance from the chamber. The ISL allows efficient linking between tracks reconstructed in the chamber and hits detected in the SVXII, and extends the track finding at pseudo-rapidities $1 \lesssim |\eta_{\rm det}| \lesssim 2$, where the chamber coverage is marginal.

The inner extension, i. e., the Layer ØØ (LØØ), is a light-weight silicon layer placed on the beam-pipe. It recovers the degradation in resolution of the reconstructed vertex position due to multiple scattering on the SVXII read-out electronics and cooling system, installed within the tracking volume. In addition, the LØØ, being made of state-of-the-art radiation-tolerant sensors, will extend the lifetime of the whole system when the effects of radiation damage will degrade the performance of the inner SVXII layers. The integrated design of the tracking system allowed commonality of components among subdetectors (read-out chip, support structures, etc.) thus simplifying the construction and

the operation.

All 722,432 channels from the $\approx 7.0\text{m}^2$ silicon active-surface employ 5644 radiation-tolerant, custom integrated read-out chips of the same type. This chip allows independent cycles of digitization of data and analog processing of subsequent data. It optimizes charge collection using sparsification, neighbor logic, and common-mode noise suppression.[15] The discriminated differential pulse from each channel is preamplified, digitized and propagated to the downstream data-acquisition. The ISL and the SVXII, whose mass is approximately 128 kg, share also the carbon-fiber supporting structure.

The total amount of material in the silicon system, averaged over ϕ and z, varies roughly as $\frac{0.1X_0}{\sin(\theta)}$ in the $|\eta_{\text{det}}| \lesssim 1$ region, and roughly doubles in $1 \lesssim |\eta_{\text{det}}| \lesssim 2$ because of the presence of cables, cooling bulk-heads, and portions of the support frame.[16] The average amount of energy loss for a charged particle is roughly 9 MeV. The total heat load of the silicon system is approximately 4 kW. To prevent thermal expansion, relative detector motion, increased leakage-current, and chip failure due to thermal heating, the silicon detectors and the associated front-end electronics are held at roughly constant temperature ranging from $-6°$C to $-10°$C for LØØ and SVXII, and around $10°$C for ISL, by an under-pressurized water and ethylene-glycol coolant flowing in aluminum pipes integrated in the supporting structures.[17]

2.3.1 The magnet

A 1.4116 T solenoidal magnetic field is maintained in the region $r \lesssim 150$ cm $|z| \lesssim 250$ cm by circulating a 4650 A current (current density 1150 A/m) through 1164 turns of an aluminum-stabilized NbTi/Cu super-conducting coil. The field is oriented along the positive $\hat{z}$ (proton) direction and is uniform to within 0.1% in the $|z| \lesssim 150$ cm volume, where tracking measurements are made (see fig. 2.3). The tiny non-uniformities, mapped out during detector construction, are treated as small perturbations in the track-fitting algorithms. During data-taking, the field is continuously monitored by nuclear magnetic resonance probes with 0.01% accuracy. Any deviation from the mapped values is applied as a correction to the measured track parameters. The threshold to radially escape the magnetic field for a particle is $p_{\text{T}} \gtrsim 0.3$ GeV/c while the trajectory of a particle with $p_{\text{T}} = 30$ GeV/c deviates only 1.6 cm from a straight path of 150 cm. The solenoid is 4.8 m in length, 1.5 m in radius, $0.85X_0$ in radial thickness (for normally incident particles.), and is cooled by forced flow of two-phase helium. Outside the coil, the return of the field flux is a box-shaped steel yoke, 9.4 m high by 7.6 m wide by 7.3 m long. It avoids interference between the field and the proper operations of the photo-multiplier tubes (PMT) used in the calorimeters.

[15] Neighbor-logic consists in a programmable threshold for each chip that allows read-out of channels over threshold and of their nearest-neighbors only, speeding-up the read-out and minimizing the data size. Noise suppression subtracts from the signal of all channels a pedestal calculated event-by-event for each chip.

[16] The symbol X_0 indicates the radiation length.

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

2.3.2 Layer 00

The LØØ is the innermost layer of the micro-vertex silicon detector [74]. It consist of a single, castellated layer of single-sided, AC-coupled silicon sensors mounted directly on the beam pipe at radii, alternating in ϕ, of 1.35 or 1.62 cm from the beam. It provides full azimuthal and $|z| \lesssim$ 47 cm longitudinal coverage. Longitudinally adjacent sensors (0.84 (or 1.46) cm $\times$ 7.84 cm) are ganged in modules of 15.7 cm active-length arranged into twelve partially-overlapping ϕ sectors, and six longitudinal barrels. These radiation-tolerant sensors are biased to $\mathcal{O}$(500 V), which allows full depletion after $\mathcal{O}$(5 Mrad) integrated radiation doses. The strips are parallel to the beam axis allowing sampling of tracks in the (r, ϕ) plane. The inter-strip implant pitch of 25 μm with floating alternate strips results in 50 μm read-out pitch. The analog signals of the 13,824 channels are fed via fine-pitch cables, up $\approx$ 50 cm long, to the front-end electronics outside the tracking volume. These cables unexpectedly pick-up ambient noise, causing a significant fraction of channels to have non-uniform and event-by-event changing pedestals that can not be controlled with dynamic pedestal subtraction. As a consequence, all channels have to be read-out, and pedestals are fit to 6th-order Chebyshev polynomial functions and subtracted off-line on an event-by-event basis. The fit procedure is iterative and excludes channels with isolated, positive signals (hits).

2.3.3 Silicon VerteX detector II

The SVXII is a fine resolution silicon micro-strip vertex detector which provides five three-dimensional samplings of tracks at 2.45 (3.0), 4.1 (4.6), 6.5 (7.0), 8.2 (8.7), and 10.1 (10.6) cm of radial distance from the beam (see fig. 2.3) with full pseudo-rapidity coverage in the $|\eta_{\rm det}| \lesssim 2$ region (see fig. 2.4(a)) [75]. This corresponds to a length of $|z| \lesssim 96$ cm along the beam-line, sufficient to cover the $\sigma_z \approx 28$ cm longitudinal spread of the luminous region. The SVXII has a cylindrical geometry coaxial with the beam, and its mechanical layout is segmented in three 32 cm axial sections ("barrels") $\times$ twelve 30° azimuthal sectors ("wedges") $\times$ five equally-spaced radial layers. A small overlap between the edges of adjacent azimuthal sectors helps wedge-to-wedge alignment (see fig. 2.4(b)). Sensors in a

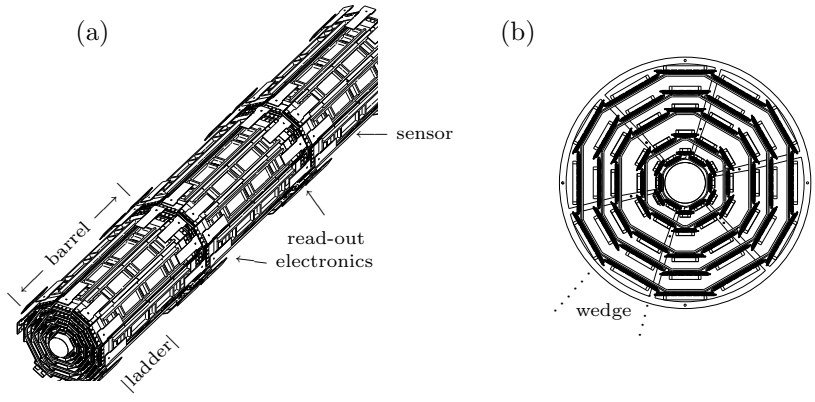

Figure 2.4: Schematic illustration of the three instrumented mechanical barrels of SVXII (a) and of the cross-section of a SVXII barrel in the (r, ϕ) plane (b).

single layer are arranged into independent longitudinal read-out units, called "ladders". Each ladder comprises two, double-sided sensors and a multi-layer electronic board, all glued on a carbon-fiber support. Front-end electronics, biasing circuits, and fan-out are located on the board that serves the pair of sensors whose strips are wire-bonded together resulting in a 15 cm active length. At a given radial layer and azimuth, each barrel contains pairs of ladders stacked length-wise head-to-head to keep the read-out electronic at the two outside extremities of the barrel (see fig. 2.4(a)). The active surface consists of double-sided, AC-coupled, 7.5 cm $\times$ 1.5–5.8 cm silicon sensors with micro-strips implanted on a 300 μm thick, high resistivity bulk. Bias is applied through integrated poly-silicon resistors. On one side, all sensors have axial strips (i. e., parallel to the beam direction) spaced by approximately 60–65 μm, for a precise reconstruction of the ϕ coordinate. On the reverse side, the following combination of read-out pitch (strip orientations with respect to the beam) is used: 141 μm (90°), 125.5 μm (90°), 60 μm (−1.2°), 141 μm (90°), 65 μm (1.2°), from the innermost to the outermost layer for reconstructing the z coordinate. A total of 405,504 electronics channels are used for SVXII.

2.3.4 Intermediate Silicon Layers

The ISL is a silicon tracker placed at intermediate radial distance between the SVXII and the drift chamber (see fig. 2.3), and covering the $|\eta_{\text{det}}| \lesssim 2$ pseudo-rapidity range for a total length of 174 cm along z [76]. At $|\eta_{\text{det}}| \lesssim 1$ a single layer of silicon sensors is mounted on a cylindrical barrel at radius of 22.6 (or 23.1 cm). At $1 \lesssim |\eta_{\text{det}}| \lesssim 2$ two layers of silicon sensors are arranged into two pairs of concentric barrels (inner and outer). In the inner (outer) barrel, staggered ladders alternate at radii of 19.7 and 20.2 cm (28.6 and 29.0 cm). One pair of barrels is installed in the forward region, the other one in the backward region. Each barrel is azimuthally divided into a 30° structure matching the SVXII segmentation. The basic read-out unit consists of an electronic board and three sensors ganged together resulting in a total active length of 25 cm. ISL employs 888 5.7 cm $\times$ 7.5 (or 6.7) cm double-sided, AC-coupled, 300 μm-thick sensors. Each sensor has axial strips spaced by 112 μm on one side, and 1.2°-angled strips spaced 112–146 μm on the reverse, for 303,104 total channels.

2.3.5 Central Outer Tracker

A multi-wire, open-cell drift chamber provides charged particle tracking at large radii in the central pseudo-rapidity region ($|\eta_{\text{det}}| \lesssim 1$, see fig. 2.3) [77]. The Central Outer Tracker (COT) has an hollow-cylindrical geometry, its active volume spans from 43.4 to 132.3 cm in radius and $|z| \lesssim 155$ cm in the axial direction. Arranged radially into eight "super-layers", it contains 96 planes of wires that run the length of the chamber between two end-plates (see fig. 2.5(a)). Each super-layer is divided into ϕ cells; within a cell, the trajectory of a charged particle is sampled at 12 radii (spaced 0.583 cm apart) where sense wires (anodes) are strung. Four super-layers employ sense-wires parallel to the beam axis, for the measurement of the hit coordinates in the (r, ϕ) plane. These are radially interleaved with four stereo super-layers whose wires are alternately canted at angles of $+2°$ and $-2°$ with respect to the beam-line. Combined read-out of stereo and axial super-layers allows the measurement of the (r, z) hit coordinates. Each super-layer is azimuthally segmented into open drift-cells. The drift cell layout (see fig. 2.5(b)) consists of a wire plane closed azimuthally by cathode sheets spaced approximately

2 cm apart. The wire plane contains sense wires alternating with field-shaping wires, which control the gain on the sense wires optimizing the electric field intensity. The cathode is a 6.35 μm-thick Mylar sheet with vapor-deposited gold shared with the neighboring cell.[18] Innermost and outermost radial extremities of a cell (i. e., the boundaries between super-layers) are closed both mechanically and electrostatically by Mylar strips with an additional field-shaping wire attached, the shaper wire.

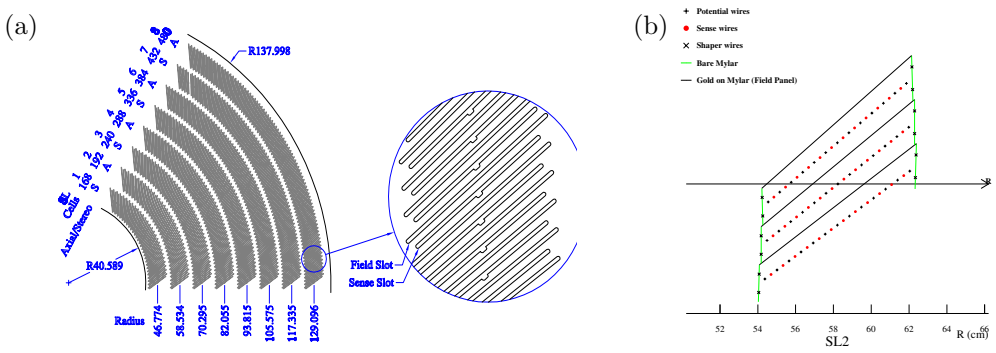

Figure 2.5: A 1/6 section of the COT end-plate (a). For each super-layer is given the total number of cells, the wire orientation (axial or stereo), and the average radius [cm]. The enlargement shows in details the slot were wire planes (sense) and field sheet (field) are installed. Sketch of an axial cross-section of three cells in super-layer 2, (b). The arrow shows the radial direction.

Both the field sheet and wire plane have a center ($z \approx 0.0$ cm) support rod that limits motion due to electrostatic forces. Each wire plane contains 12 sense, 13 field-shaping, and 4 shaper wires, all made of 40 μm-diameter gold-plated tungsten. Wire planes are not aligned with the chamber radius: a $\zeta = 35°$ azimuthal tilt (see fig. 2.5(b)) partially compensates for the Lorentz angle of the drifting electrons in the magnetic field.[19] The tilted-cell geometry helps in the drift-velocity calibration, since every high-p_{T} (radial) track samples the full range of drift distances within each super-layer. Further benefit of the tilt is that the left-right ambiguity is resolved for particles coming from the z-axis since the ghost track in each super-layer appears azimuthally rotated by $\arctan[2\tan(\zeta)] \approx 54°$, simplifying the pattern recognition problem.[20]

A 50:50 gas admixture of argon and ethane bubbled through isopropyl alcohol (1.7%) flows at 9.45 l/min in the active volume of the chamber with its pressure being continuously monitored by four probes. Since 2003, the flux has been increased by a factor of ten to contrast the adverse effect of wire aging (see sec. 4.3). High voltage is applied to the sense and field-shaping wires to generate a 1.9

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

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

[20] Each pulse on a given wire has a two fold ambiguity corresponding to the two incoming azimuthal drift trajectories. The signals from a group of nearby radially wires will satisfy the configuration for two tracks, one from the actual particle trajectory and another "ghost track" originated by the ambiguity.

kV/cm drift electric-field. This value, combined with the drift gas, results in a maximum drift-time of about 177 ns along a maximum drift-distance of 0.88 cm, allowing for read-out and processing of the COT data between two consecutive bunch-crossings. The average 180 kV/cm field present at the surface of the sense wire produces typical gains of 2×10^4. The 30,240 sense wires are read-out by the front-end chip, which provides input protection, amplification, shaping, baseline restoration, discrimination, and charge measurement. The input-charge information is encoded (logarithmically) in the signal width for dE/dx sampling, and is fed to a time-to-digital converter that records leading and trailing-edge times of signal in 1 ns bins. Hit times are later processed by the pattern recognition software to reconstruct trajectories. Owing to the relevance of the dE/dx measurement in this analysis, the extraction of the dE/dx information in the COT is detailed in appendix A. The material of the COT corresponds to an average $0.017X_0$ for tracks at normal incidence.

2.3.6 Tracking performance

The only physics objects used in this analysis are the tracks. Within an uniform axial magnetic field in vacuum, the trajectory of a charged particle produced with non-zero initial velocity in the bending plane of the magnet is described by an helix. The arc of an helix described by a charged particle in the magnetic volume of CDF is parameterized using three transverse, and two longitudinal parameters:

- $\boldsymbol{C}$ – signed helix (half)-curvature, defined as $C \equiv \frac{q}{2R}$, where R is the radius of the helix. This is directly related to the transverse momentum: $p_{\mathrm{T}} = \frac{cB}{2|C|}$;
- $\boldsymbol{\varphi_0}$ – ϕ direction of the particle at the point of closest approach to the z-axis;
- $\boldsymbol{d_0}$ – signed impact parameter, i. e., the distance of closest approach to the z-axis, defined as $d_0 \equiv q(\sqrt{x_c^2 + y_c^2} - R)$, where (x_c, y_c) are the coordinates of the center-guide;
- $\boldsymbol{\lambda}$ – the helix pitch, i. e., $\cot(\theta)$, where θ is the polar direction of the particle at the point of its closest approach to the z-axis. This is directly related to the longitudinal component of the momentum: $p_z = p_{\mathrm{T}} \cot(\theta)$;
- $\boldsymbol{z_0}$ – the z coordinate of the point of closest approach to the z-axis.

The trajectory of a charged particle satisfies the following equations [78]:

$$x = r \sin(\varphi) - (r + d_0) \sin(\varphi_0) \tag{2.7}$$

$$y = -r \cos(\varphi) + (r + d_0) \cos(\varphi_0) \tag{2.8}$$

$$z = z_0 + s\lambda, \tag{2.9}$$

where s is the projected length along the track, $r = 1/2C$, and $\varphi = 2Cs + \varphi_0$. The reconstruction of a charged-particle trajectory consists in determining the above parameters through an helical fit of a set of spatial measurements ("hits") reconstructed in the tracking detectors by clustering and pattern-recognition algorithms. The helical fit takes into account field non-uniformities and scattering in the detector material.

For this analysis, only COT-seeded silicon tracks were used, because the pattern recognition algorithms that use stand-alone silicon information would have given marginal contribution for two reasons. First, the impact of silicon stand-alone tracking becomes important in the region $1 \lesssim |\eta| \lesssim 2$ where the COT coverage is incomplete. However, this region of acceptance is already excluded in our analysis, since the trigger that collects $B^0_{(s)} \to h^+ h^{'-}$ decays uses the COT information (see sec. 3.2). Secondly, the algorithms for silicon stand-alone tracking were not yet optimized as of this analysis.

All tracks were first fit in the COT and then extrapolated inward to the silicon. This approach guarantees fast and efficient tracking with high track purities. The greater radial distance of the COT with respect to the silicon tracker results in a lower track density and consequent fewer accidental combination of hits in the track reconstruction. A concise overview of the tracking algorithms is given in the following, see Ref. [79] for more details.

Tracking in the COT

In each event and for each sense wire above threshold, the COT front-end provides the information on the integrated charge and on the arrival time of the avalanche ("hit time") with a few nanoseconds resolution. The hit time depends on several factors including the time at which the collision occurred with respect to the Tevatron clock, the time spent by the charged particle produced in the collision to reach the proximity of the wire, the drift time of the ionized charge in the gas, the profile of the charge pulse (different amplitudes trigger the discriminator at different times), non-uniform electronic signal propagation among channels, etc. The hit time, after calibration and correction for all these effects, is interpreted as drift time and used for pattern recognition.

First, in each super-layer a segment seed is constructed by fitting to a straight line a triplet of hits in adjacent wire-planes [80]. The angle between the segment and the radial direction is required to be consistent with the one of a tangent to a circular trajectory of a charged particle with $p_{\mathrm{T}} \gtrsim 0.3$ GeV/c originated near the z-axis. This removes ghost tracks due the left-right azimuthal ambiguity with respect to a wire. Further hits in the super-layer are iteratively added to the seed when lying within 1 mm (i. e., 20 ns at 50 μm/ns velocity).

Then, axial pattern recognition takes place by reconstructing circles (i. e., two-dimensional tracks) in the four axial COT super-layers. Two algorithms are used in parallel for a higher efficiency; eventual track duplicates are removed at the end of the process. The first algorithm applies a χ^2-fit to all hits belonging to matching segments among different super-layers. The second algorithm defines a circle using one segment and the beam position. Then it searches for all hits within 1 cm in radius from the circle and fills a 200 μm-binned histogram with the radius of each hit found. If the higher statistics bin contains at least 10 hits, these are used to fit a track. If additional hits are found within 750 μm from the track, they are included and the track is refit.

Stereo information is added from the outer to the inner super-layers. A three-dimensional track is refit after progressive addition of stereo segments reconstructed in cells adjacent to those where axial segments were found (segment-linking). Once this step is completed, the three-dimensional tracks are used to find the z coordinates of the vertices in the event. Association of stereo hits that match the found z vertices is attempted for those axial tracks that failed stereo segment-linking. Final tracks are refit using detailed magnetic field maps and drift model.

All channels of the COT are properly working. Its efficiency for tracks is typically 99%. The single-hit resolution is 140 μm, including a 75 μm contribution from the ≈ 0.5 ns spread in the measurement of the time of the interaction. Internal alignments of the COT cells are maintained within 10 μm using cosmic rays. Curvatures effects from gravitational and electrostatic sagging are under control within 0.5% by equalizing the difference of E/p between electrons and positrons as a function of $\cot(\theta)$. The typical resolutions on track parameters are the following: $\sigma_{p_T}/p_T^2 \approx 0.0015$ $(\mathrm{GeV}/c)^{-1}$, $\sigma_{\varphi_0} \approx 0.035°$, $\sigma_{d_0} \approx 250$ μm, $\sigma_\theta \approx 0.17°$, and $\sigma_{z_0} \approx 0.3$ cm for tracks fit with no silicon information or beam constraint. Since it directly affected this analysis, aging of the chamber is discussed in sec. 4.3.

Tracking in the silicon detector

When a silicon strip is above threshold, its neighboring strips are also read-out. This results in strings of contiguous strips above threshold with local minima and maxima of signal. A clustering algorithm split these strings in correspondence of the local minima. If any local maximum exist, the closest local minima are associated to it, forming a cluster. If the minimum is equidistant from two maxima, its charge is split between the two clusters. The precise position of each cluster, i. e., the silicon hit, is lastly calculated with a charge-weighed mean over the strips of the cluster.

COT tracks are used to seed the reconstruction of tracks in silicon with a progressive algorithm that attaches silicon hits from the outermost layers moving inward. First axial and then stereo silicon hits are associated to the COT track with an iterative process. Silicon hits are searched in a 4σ-wide extrapolation of the COT track in the silicon layers, where σ are the uncertainties on the estimated track parameters, which are updated at each iteration to account for the multiple scattering effects. At each iteration (i. e., in each layer), several silicon hits may be found in the search region. For each of them, a candidate track is considered and refit, but only the two tracks with highest hit multiplicity are considered for the next layer. At the end of the process, only the track with highest hit multiplicity and better χ^2 is retained.

As of this writing, 96% of the LØØ, 84% of the SVXII, and 82% of the ISL modules are recording physics-quality data (i. e., with less than 1% digital error rate). Two unexpected phenomena, occurred at the beginning of Run II, were the dominant causes of the current inefficiency [81]. The first contribution is related to two beam incidents: in March 2002, the failure of multiple Tevatron RF cavities debunched the beam, causing high losses and consequent quenching of the magnets. The beam was aborted with an uncontrolled deflection that exposed the CDF II detector to a flux of more than 10^7 minimum ionizing particles per cm^2 in less than 150 ns; in November 2002, a failure on the deflecting magnets induced an incidental Tevatron abort in which some bunches were deflected into CDF II. Addition of faster interlock systems and more collimators prevented further occurrence of such incidents.

The second phenomenon was the break-up of wire-bonds oriented orthogonally to the magnetic field, due to resonant Lorentz forces occurring in read-out tests at 16 kHz frequency. This was addressed with an optimized configuration of the read-out parameters, and inhibiting further read-out if a resonance is detected. A temporary inefficiency was induced by 35% of ISL cooling-lines being blocked by epoxy, after installation. This prevented a large portion of ISL from being active, until all the lines were cleared with a laser (in January 2003).

The signal-to-noise ratio ranges from 14:1 for the (r, ϕ) layers of the SVXII to 10:1 for the LØØ. The best (r, ϕ) position resolution achieved is 9 μm, using two-strip clusters in SVXII. The z_0 resolution is typically 70 μm. An active real-time optical survey keeps the SVXII axis parallel to the beam within 20 μm along the SVXII length. Tight assembling tolerances (10 μm in ϕ and 40 μm in r) combined with a set of off-line algorithms provide internal and global LØØ, SVXII, and ISL alignment accurate within 20 μm, and constantly monitored in time. The excellent overall accuracy of the silicon alignment is confirmed by the fluctuations of the measured impact parameters of prompt particles as a function of z and ϕ, which do not exceed 2 μm. This can be compared with a typical impact parameter resolution of a few tenths of microns. The average offline tracking efficiency is 94%.[21] In the $1 \lesssim |\eta| \lesssim 2$ region, where no COT coverage is present, seeding the silicon-only track with calorimeter information provides efficiencies over 70%, with minimal fake rates.

Aging effects due to long-term radiation are the primary concern for the lifetime of the detector. Depletion voltages are regularly monitored and, by extrapolating the current data, we expect type inversion of the inner layer of SVXII after 3–5 fb^{-1} integrated luminosity. The silicon information improves the impact parameter resolution of tracks which, depending on the number (and radial distance) of the silicon hits, may reach $\sigma_{d_0} \approx 20$ μm (not including the transverse beam size). This value, combined with the $\sigma_{\mathrm{T}} \approx 30$ μm transverse beam size, is sufficiently small with respect to the typical transverse decay-lengths of heavy flavors (a few hundred microns) to allow separation of their decay-vertices from production vertices. The silicon tracker improves also the stereo resolutions up to $\sigma_\theta \approx 0.06°$, and $\sigma_{z_0} \approx 70$ μm, while the transverse momentum and the azimuthal resolutions remain approximately the same of COT-only tracks.

Fit of the beam-line and reconstruction of vertices

For the data used in the present analysis, the position of the primary p$\bar{\mathrm{p}}$ interaction is found by combining the information on the transverse plane from the beam-line position, with the information on the longitudinal position from the z_0 of tracks.

beam position – the average beam position through the SVXII is determined on-line and off-line for each run, using the correlation (eq. (2.13)) between impact parameter and azimuthal coordinate of tracks (see sec. 2.5.2). An additional off-line algorithm determines the beam position using only COT information. It uses an iterative vertex-fit procedure that loops on the tracks and performs multiple refitting of the vertex, after excluding progressively the tracks that mostly contribute to the χ^2. The comparison between the SVXII and COT beam-lines is used to check the relative alignment of the two subdetectors.

primary vertex – the primary vertex position is reconstructed by clustering the z_0 of tracks whose z_0 are within 5 cm, and that are consistent with passing through the beam-position. In bunch crossings with pile-up events, the 5 cm limit minimizes the contribution of tracks originated from distinct primary vertices, whose observed Δz_0 distribution has a roughly Gaussian shape with about 1 cm r.m.s. width.

[21] This efficiency is defined as the probability to attach silicon hits to an isolated COT track with $p_{\mathrm{T}} \gtrsim 1.5$ GeV/c.

secondary vertex – the CTVMFT package [78] reconstructs decay vertices displaced from the production vertices. Equations (2.7)–(2.9), solved for s, d_0, and z_0, yield

$$s = \frac{1}{2C} \arcsin\left[(2C)x\cos(\varphi_0) + (2C)y\sin(\varphi_0)\right] \tag{2.10}$$

$$d_0 = y\cos(\varphi_0) - x\sin(\varphi_0) - \frac{1}{C}\sin^2(Cs) \tag{2.11}$$

$$z_0 = z - \lambda s. \tag{2.12}$$

Given a set of N tracks, CTVMFT determines the coordinates of their decay vertex, i. e., the set (x_v, y_v, z_v) that satisfy above equations, varying the track parameters within their uncertainties to minimize the quantity $\chi^2 \equiv \sum_{i=1}^{N} \vec{v}_i^T \mathbf{G}^{-1} \vec{v}_i$. The symbol $\vec{v}_i$ indicates the vector of the distance between the track constrained through eqs. (2.10)–(2.12) and the original unconstrained one in the five-dimensional space of track parameters; $\mathbf{G}$ is the covariance matrix of the track parameters evaluated at the distance of closest approach to the z axis. In case of success, the parameters of the tracks originated from the secondary vertex are refit using the additional information on its position.

2.4 Other detectors

In this section are briefly discussed the subdetectors not used in this analysis.

2.4.1 Time-of-flight detector

The CDF II capability of identifying charged hadrons is expanded at low momenta by the Time of Flight detector (TOF) [82]. By measuring the arrival time (t) of a charged particle with respect to the bunch-crossing time, the TOF infers the mass of the particle according to the relation $m = \frac{p}{c}\sqrt{\frac{c^2t^2}{L^2} - 1}$, where the momentum p and the path length L are precisely measured by the tracking system. A cylindrical array of 216 scintillating bars, oriented along the beam axis, is installed in the 4.7 cm radial space between the outer surface of the COT and the cryostat of the superconducting solenoid at an average radius of 140 cm, which corresponds to 5 ns flight-time for the fastest particles. The light pulse induced by the passage of a charged particle is collected at both ends of each bar, 279 cm in length ($|\eta_{\text{det}}| \lesssim 1$) and 4×4 cm^2 in cross-section ($\Delta\phi = 1.7°$), into 432 fine-mesh, 19-stage photomultipliers able to maintain adequate gains even in the 1.4 T magnetic field. The preamplified PMT signal follows two parallel paths: the timing signal is discriminated and digitized, while the charge signal is digitized for use at trigger level and for subsequent extraction of off-line corrections. Long attenuation-length and fast rise-time scintillator, along with accurate calibrations, ensure a measured resolution (i. e., the uncertainty on the relative timing between the collision and the TOF hits) of $\sigma_t \approx 110$ ps. This guarantees a separation between charged pions and kaons with $p_{\text{T}} \lesssim 1.6$ GeV/c equivalent to 2σ, assuming Gaussian distributions. Unfortunately, in high ($\mathcal{L} \gtrsim 5 \times 10^{31}$ cm^{-2}s^{-1}) luminosity conditions, the occupancy of the single bars determines a degradation in efficiency, which is about 60% per track.

2.4.2 Calorimeters

Outside the solenoid, calorimeters covering the region $|\eta_{\text{det}}| \lesssim 3.6$ measure the energy flow from hadrons, electrons, or photons, using "shower" sampling based on layers of high-Z passive absorber interspersed with layers of plastic scintillator, and identify neutrinos via transverse energy imbalance.[22] The calorimeters are finely segmented in solid angle around the nominal collision point and coarsely segmented outward from the collision point (in-depth segmentation). Angular segmentation is organized in projective *towers*. Each tower is an independent read-out unit which subtends a portion of the solid angle, namely a rectangular cell in the $(\eta_{\text{det}} - \phi)$ space, with respect to the nominal interaction region. In-depth segmentation of each tower consists of two independent compartments: the inner one samples the electromagnetic component of the shower, while the outer one samples the hadronic fraction of the deposited energy. Different fractions of energy release in the two compartments distinguish photons and electrons from hadrons.

The initial calibration of the central calorimeter was done using cosmic rays and 50 GeV electrons and pions in a test beam. The test-beam calibration of the forward electromagnetic calorimeters employed 5–260 GeV positrons, pions, and muons. Long term variations of the response are monitored using radioactive ^{137}Cs (^{60}Co) sources in the central (forward) calorimeter. Short term (once per store) gain variations are monitored using an automated system of light-emitting diodes and xenon flashers for the electromagnetic portion, and a nitrogen laser for the hadronic portion. *In situ* calibrations are also exploited by matching the trajectories of isolated charged particles with the calorimeter towers, and comparing calorimeter energy with the momentum measured in the COT.

Central calorimeter

In the $|\eta_{\text{det}}| \lesssim 1.1$ region, the electromagnetic calorimeter appears as an hollow cylinder occupying the radial region between 173 and 208 cm [83]. It is made of four arches, each subtending 180° and divided into 12 azimuthal 15°-sectors (see fig. 2.6(a)). Each sector consists of 31 layers of 5 mm thick polystyrene scintillator radially interleaved with 30 aluminum-clad lead sheets, 3.2 mm-thick. Each sector is divided into ten η_{det} towers ($\Delta\eta_{\text{det}} \times \Delta\phi \approx 0.11 \times 15°$ per tower). To maintain a constant thickness in X_0, compensating the $\sin(\theta)$ variation between towers, some lead layers are replaced with increasing amounts of acrylic as a function of η_{det}.[23] The blue light from each tower is collected, wave-length shifted into green light by sheets of acrylic material placed at both azimuthal tower boundaries, and guided to two photo-tubes per tower. The outer two towers in one wedge are missing to allow accessing the solenoid. The total number of instrumented towers is 478. At a radial depth of 5.9 radiation lengths (X_0), where the peak of shower development is typically located, an array of multi-wire proportional chambers measures the transverse shower-shape with 2.0 mm resolution (for 50 GeV electrons). A further set of multi-wire proportional chambers was located in the gap between the outer surface of the solenoid and the first layer of the calorimeter to monitor photon conversions started in the tracker material or in the solenoid. During the fall 2004 shut-down, this system was substituted with a finely segmented layer of scintillator tiles [84].

The hadronic compartment comprises two subsystems: the central ($|\eta_{\text{det}}| \lesssim 0.9$) and the end-wall

[22] See pag. 48 for a definition of transverse energy.

[23] The number of lead layers varies from 30 in the innermost ($|\eta_{\text{det}}| \approx 0.06$) tower to 20 in the outermost ($|\eta_{\text{det}}| \approx 1.0$).

($0.7 \lesssim |\eta_{\rm det}| \lesssim 1.3$) section [85]. Both consist of four "C"-shaped arches for a total of 48 azimuthal sectors. Each central wedge is segmented into nine $\eta_{\rm det}$ towers matching in size and position the electromagnetic towers, for 384 towers in total. The end-wall section has six additional towers, three of which matching the central hadronic towers (see fig. 2.2) for a total number of 288 towers. A central hadronic tower is constructed of 32 layers of steel absorber, 2.5 cm thick, alternating radially with 1.0 cm-thick acrylic scintillator. The wall towers are similar but contain only 15 layers of 5.1 cm-thick absorber.

The total thickness of the electromagnetic section corresponds to approximately $19X_0$ ($1\lambda_{\rm int}$, where $\lambda_{\rm int}$ is the pion nuclear absorption length in units of g cm^{-2}), for a relative energy resolution $\sigma_E/E = 13.5\%/\sqrt{E\sin(\theta)} \oplus 2\%$.[24] The total thickness of the hadronic section corresponds to approximately $4.5\lambda_{\rm int}$, for an energy resolution of $\sigma_E/E = 50\%/\sqrt{E\sin(\theta)} \oplus 3\%$ for the central, and $\sigma_E/E = 75\%/\sqrt{E\sin(\theta)} \oplus 4\%$ for the end-wall.

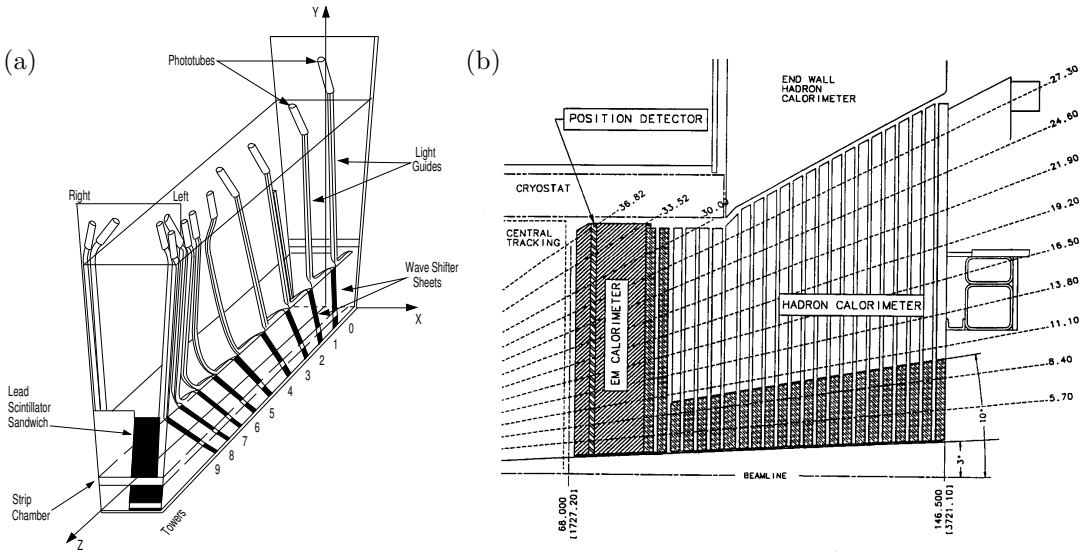

Figure 2.6: Schematic illustration of an azimuthal sector of the central electromagnetic calorimeter (a). Elevation view of one quarter of the plug calorimeter (b).

Forward calorimeter

The electromagnetic (hadronic) coverage is extended in the region 1.10 (1.30) $\lesssim |\eta_{\rm det}| \lesssim 3.64$ by a scintillating tile calorimeter (see fig. 2.6(b)) [86]. It is composed of two independent and identical devices, installed at longitudinally symmetric positions with respect to the interaction point (east and west). The absorber of the electromagnetic part consists of 23 annular lead plates, 2.77 m outer diameter, with a central hole to house the beam pipe. Each one is made out of 4.5 mm-thick calcium-tin-lead sandwiched between two 0.5 mm-thick stainless-steel sheets. In between the absorber plates,

[24] The first term is called the "stochastic" term and derives from the intrinsic fluctuations of the shower sampling process and of the PMT photo-electron yield. The second term, added in quadrature, depends on the calorimeter non-uniformities and on the uncertainty on the calibrations. All energies are in GeV.

4 mm-thick scintillator tiles are organized azimuthally in 15° triangular sectors. In each tile, the signal is collected independently by embedded wave-length shifter fibers connected to the photo-tubes. A thicker (10 mm) amount of scintillator, independently read-out and installed in the first layer, acts as preshower detector. Shower-maximum sampling is available at a radial depth of $\approx 6X_0$ using two tilted layers of scintillator strips with 5 mm pitch. The hadronic section consists of 23 annular layers of 5 cm-thick iron absorber alternated with 6 mm scintillator layers. The outer radius of each module increases at increasing $|z|$ giving the characteristic "plug" shape to the calorimeter. Each module is segmented into 12 azimuthal sectors, each subtending 30°. Within each sampling layer, the scintillator is arranged in tiles similar to those used in the electromagnetic compartment.

The total thickness of the electromagnetic section corresponds to approximately $21X_0$ ($1\lambda_{\rm int}$), for an energy resolution of $\sigma_E/E = 16\%/\sqrt{E\sin(\theta)} \oplus 1\%$.[25] The total thickness of the hadronic section corresponds to approximately $7\lambda_{\rm int}$, for an energy resolution of $\sigma_E/E = 74\%/\sqrt{E\sin(\theta)} \oplus 4\%$.

2.4.3 Muon detectors

The tracker, the magnet return yoke, the calorimeter, and additional steel shielding act as muon filters: they absorb the majority of charged particles, thus allowing detection of the more penetrating ones in multiple layers of drift chambers placed in the outermost shell of the detector. Four independent systems detect penetrating charged particles in the $|\eta_{\rm det}| \lesssim 1.5$ pseudo-rapidity range, employing similar combinations of drift tubes, scintillation counters, and absorbers with different azimuthal coverages [87]. Table 2.2 shows the most relevant design parameters of these detectors. Single wire, rectangular drift chambers are arranged in arrays with various azimuthal segmentations and coupled with scintillator counters. The chambers, filled with a 50:50 admixture of argon and ethane, have sense wires parallel to the beam axis operating in proportional regime. In each azimuthal sector, stacks of up to eight layers of chambers are overlaid along the radial direction to allow coincidences among hits. The chambers are staggered in various patterns of alternating layers, for azimuthal ambiguity resolution. The difference of the drift electrons arrival-times between neighboring cells provides up to 250 μm hit-position resolution in the (r, ϕ) view. Division of the charge collected at the opposite ends of sense wires allows a measurement of the z coordinate of the hit with up to 1.2 mm resolution. Resolutions were measured using cosmic rays. The charge from each of the 7,316 channels is preamplified, shaped, discriminated, and digitized. The arrival time is digitized also. Scintillators provide timing information to suppress backgrounds due to secondary interactions in the beam pipe material and to cosmic rays. Timing also allows association of cell hits to the appropriate bunch-crossing, since the maximum drift time in the chambers (see tab. 2.2) exceeds the inter bunch-crossing time. When a short track-segment (*stub*), resulting from three matching radial layers, corresponds to the outward extrapolation of a COT track, a muon candidate is identified and associated to the corresponding momentum measured in the tracker.

The Central MUon detector (CMU) is located around the outside of the central hadron calorimeter at a radius of 347 cm from the beam axis and covers the region $0.03 \lesssim |\eta_{\rm det}| \lesssim 0.63$. Each array covers 12.6° in ϕ, and a 2.4° gap between arrays limits the azimuthal coverage to 84% of the full angle. Each array is further segmented azimuthally into three 4.2° modules.

[25] See footnote at pag. 61 for an explanation of terms.

The Central Muon uPgrade (CMP) is a second set of drift chambers located behind an additional 60 cm of steel and arranged to enclose the central detector inside an approximately rectangular box. Its function is to cover the ϕ gaps of the CMU, and to enhance rejection of penetrating high energy hadrons, which are limited to a measured fraction of 1% of the total pions and 2–4% of the total kaons. Owing to the common CMU and CMP coverage, only one set of scintillators is used for both, and the z coordinate is measured only in the CMU.

At a radial distance of 400–600 cm from the beam axis, an arch-shaped arrangement of drift-cells and scintillation counters, the Central Muon eXtension (CMX), extends the muon coverage to the $0.6 \lesssim |\eta_{\rm det}| \lesssim 1$ region.

Coverage in the region $1 \lesssim |\eta_{\rm det}| \lesssim 1.5$ is completed by the Intermediate MUon system (IMU). Each cell-stack spans 1.25° in ϕ although its azimuthal coverage is limited by the presence of support structures (see tab. 2.2).

Parameter	CMU	CMP	CMX	IMU	Units
Polar coverage	$\lvert\eta_{\rm det}\rvert \lesssim 0.6$	$\lvert\eta_{\rm det}\rvert \lesssim 0.6$	$0.6 \lesssim \lvert\eta_{\rm det}\rvert \lesssim 1.0$	$1 \lesssim \lvert\eta_{\rm det}\rvert \lesssim 1.5$	–
Azimuthal coverage	302°	360°	360°	270°	Degrees
Maximum drift time	800	1,400	1,400	800	ns
Number of channels	2,304	1,076	2,208	1,728	–
Pion interaction-length	5.5	7.8	6.2	6.2–20.0	$\lambda_{\rm int}$
Minimum $p_{\rm T}(\mu)$	1.4	2.2	1.4	1.4–2.0	GeV/c

Table 2.2: Design parameters of the muon detectors. Pion interaction length is quoted for a reference axial angle $\theta = 90°$ in CMU and CMP, and $\theta = 55°$ in CMX.

2.4.4 Cherenkov Luminosity Counters

The luminosity ($\mathcal{L}$) is inferred from the average number of inelastic interactions per bunch crossing ($\overline{N}$) according to $\overline{N} \times f_{\rm b.c.} = \sigma_{\rm p\bar{p}-in.} \times \varepsilon \times \mathcal{L}$, where the bunch-crossing frequency ($f_{\rm b.c.}$) is precisely known from the Tevatron RF, $\sigma_{\rm p\bar{p}-in.} = 59.3 \pm 2.3$ mb is the inelastic p$\bar{\rm p}$ cross-section resulting from the averaged CDF and E811 luminosity-independent measurements at $\sqrt{s} = 1.8$ TeV [88], and extrapolated to $\sqrt{s} = 1.96$ TeV, and ε is the efficiency for detecting an inelastic scattering.

The Cherenkov Luminosity Counters (CLC) are two separate modules, covering the $3.7 \lesssim |\eta_{\rm det}| \lesssim 4.7$ range symmetrically in the forward and backward regions [89]. Each module consists of 48 thin, 110–180 cm long, conical, isobutane-filled Cherenkov counters. They are arranged around the beam-pipe in three concentric layers and point to the nominal interaction region. The base of each cone, 6–8 cm in diameter and located at the furthest extremity from the interaction region, contains a conical mirror that collects the light into a PMT, partially shielded from the solenoidal magnetic field. Isobutane guarantees high refraction index and good transparency for ultraviolet photons. With a Cherenkov angle $\theta_C = 3.4°$, the momentum thresholds for light emission are 9.3 MeV/c for electrons and 2.6 GeV/c for charged pions. Prompt charged particles from the p$\bar{\rm p}$ interaction are likely to traverse the full counter length, thus generating large signals and allowing discrimination from the smaller signals of angled particles due to the beam halo or to secondary interactions. In addition,

the signal amplitude distribution shows distinct peaks for different particle multiplicities entering the counters. This allows a measurement of $\overline{N}$ with 4.4% relative uncertainty in the luminosity range $10^{31} \lesssim \mathcal{L} \lesssim 10^{32}$ cm^{-2}s^{-1}. This accuracy, combined with the 4% relative uncertainty on the inelastic p$\bar{\text{p}}$ cross-section, results in an instantaneous luminosity measured with 5.9% relative uncertainty. This uncertainty does not affect the results of this analysis since ratios of branching fractions, instead of absolute branching fractions, are measured.

2.4.5 Forward detectors and beam monitoring

Six arrays of scintillator paddles, read-out by acrylic light-guides are wrapped around the beam line at longitudinal distances of ±6.6, ±23.2, and ±31.6 m from the interaction point at 3.8–10.8 cm radii. An additional array is placed at 56.4 m from the interaction point toward the west side. These counters are used for triggering diffractive events and for monitoring beam losses.

Two small, cylindrical calorimeters occupy the radial range $6 \lesssim r \lesssim 33$ cm in the region $580 \lesssim |z| \lesssim 640$ cm. They employ lead plates immersed in radiation-resistant liquid scintillator with a tower-less, homogeneous geometry suited for calorimetric tracking for diffractive physics measurements [90].

The furthermost ($z \approx -57$ m) components from the interaction point are scintillating fibers placed in the Tevatron vacuum inside retractable "Roman Pots", which are used as spectrometer for leading antiprotons in measurements of diffractive physics.

Further arrays of scintillation counters and ionization chambers are placed along the beam line at varying distances from the interaction point to monitor the beam halo and losses [91].

2.5 Trigger and Data Acquisition System

From the rule of thumb 1 μb = 1 Hz at $\mathcal{L} = 10^{30}$ cm^{-2}s^{-1}, we obtain that, at a typical Tevatron instantaneous luminosity $\mathcal{L} \approx 4 \times 10^{31}$ cm^{-2}s^{-1}, and with an inelastic p$\bar{\text{p}}$ cross-section of $\sigma_{\text{p}\bar{\text{p}}-\text{in.}} \approx 60$ mb, approximately 2.5×10^6 inelastic collisions per second occur, corresponding to one inelastic p$\bar{\text{p}}$ interaction per bunch crossing on average.[26] Since the read-out of the entire detector needs about 2 ms on average, after the acquisition of one event, another approximately 5,000 interactions would occur and remain unrecorded. The percentage of events which are rejected solely because the trigger is busy processing previous events is referred to as trigger *deadtime*.

On the other hand, the average size of the information associated to each event from the $\mathcal{O}(10^6)$ total CDF II channels is 140 kbyte. Even in case of deadtime-less read-out of the detector, in order to record all events, an approximate throughput and storage rate of 350 Gbyte/s would be needed, largely beyond the possibilities of currently available technology.

If one would adapt the 2.53 MHz interaction-rate to the 50–100 Hz storage rate attainable at CDF II by choosing randomly the 2 storable events out of the 100,000 occurred, only four $B^0_{(s)} \to h^+h'^-$ decays would be expected in our sample of $\int \mathcal{L}dt = 180$ pb^{-1}. However, since the cross-sections of most interesting processes are 10^3–10^{12} times smaller than the inelastic p$\bar{\text{p}}$ cross-section, the above problems may be overcome with an on-line preselection of the most interesting events. This is the task

[26] Abort gaps can be neglected for this estimate.

of the trigger system, which evaluates the partial information provided by the detector and discards the uninteresting events on-line.

The CDF II trigger is a three-level system that selectively reduces the acquisition rate, with virtually no deadtime, i. e., keeping each event in the trigger memory a time sufficient to allow for a trigger decision without inhibiting acquisition of the following events (see fig. 2.7). Each level receives the accepted event from the previous one and, provided with detector information of increasing complexity and with more time for processing, applies a logical "OR" of several set of programmable selection criteria to make its decision.

Prior to any trigger level, the bunched structure of the beams is exploited to reject cosmic-ray events by gating the front-end electronics of all subdetectors in correspondence of the bunch crossing. The front-end electronics of each subdetector, packaged in Versa Module Eurocard modules hosted in about 120 crates, has a 42-cells deep pipeline synchronized with the Tevatron clock-cycle (i. e., 132 ns). The Tevatron clock picks-up a timing marker from the synchrotron RF and forwards this bunch-crossing signal to the trigger and to the front-end electronics. Since the inter-bunch time is 396 ns (instead of the 132 ns originally planned), the pipeline collects data corresponding to a maximum of $42 \times 132/396 = 14$ bunch crossings, automatically rejecting 2/3 of cycles corresponding to the crossing of empty buckets. For each crossing, data enter the pipeline for read-out and eventual use at Level-2, and a Level-1 decision on a preceding crossing is made before the corresponding data reach the end of the pipeline. The Level-1 has 132 ns $\times$ 42 $\simeq$ 5.5 μs to make its decision before the contents of the buffer is deleted. On a Level-1 accept, the data from the Level-1 buffer are passed to the four-cell Level-2 buffer integrated in the front-end electronics of each subdetector, and the event is queued for a Level-2 decision. While data in a Level-2 buffer are being processed, they cannot be overwritten by incoming data corresponding to a subsequent Level-1 accept. If a Level-1 accept occurs while all four Level-2 buffers are occupied, trigger deadtime is incurred. The 5.5 μs $\times$ 4 $\simeq$ 20 μs latency of the Level-2 decision is less than approximately 80% of the average time between Level-1 accepts, to minimize deadtime. On a Level-2 accept, the entire detector is read-out, thereby emptying a cell in all detector buffers for the next event; the event is queued for read-out in Level-3 and for eventual storage to permanent memory.

The following description emphasizes the aspects of the trigger specific to this analysis: particular detail is devoted to the devices dedicated to the identification of tracks produced in decays displaced from the hard $p\bar{p}$ interaction vertex. These tracks populate events enriched in long-lived heavy-flavor decays, including the $B^0_{(s)} \to h^+ h^{'-}$ decays we wish to reconstruct.

2.5.1 Level-1

At Level-1, a synchronous system of custom-designed hardware process a simplified subset of data in three parallel streams to reconstruct coarse information from the calorimeters (total energy and presence of single towers over threshold), the COT (two-dimensional tracks in the transverse plane), and the muon system (muon stubs in the CMU, CMX, and CMP chambers). A decision stage combines the information from these low-resolution physics objects, called "primitives", into more sophisticated objects, e. g., track primitives are matched with muon stubs, or tower primitives, to form muon,

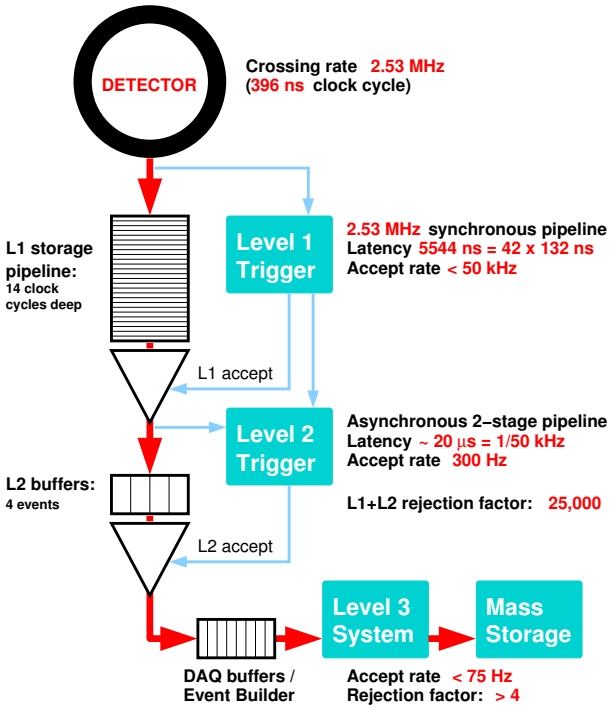

Figure 2.7: Functional block diagram of the CDF II trigger and data acquisition system.

electron, or jet objects, which are subjected to basic selections.[27]

Drift chamber track-processor

The eXtremely Fast Tracker (XFT) is a custom processor that identifies two-dimensional tracks in the (r, ϕ) view of the COT (transverse plane) in time with the Level-1 decision. It uses pattern matching to first identify short segments of tracks and then to link them into full-length tracks [92]. After classifying the hits of the four axial COT super-layers in "prompt" (0–66 ns) or "delayed" hits (67–220 ns), depending upon the observed drift-time within the cell, track segments are reconstructed in each axial super-layer. A pattern-matching algorithm searches for coincidences between the observed combinations of hits in each super-layer — a minimum of 11 (out of 12) hits is required — and a set of predetermined patterns. If a coincidence between segments crossing four super-layers is found, two-dimensional XFT-tracks are reconstructed by linking the segments. The segments are compared with a set of about 2,400 predetermined patterns corresponding to all tracks with $p_{\mathrm{T}} \gtrsim 1.5$ GeV/c originating from the beam line. The comparison proceeds in parallel in each of the 288 azimuthal 1.25°-sectors in which XFT logically divides the chamber. If no track is found using all four super-layers, then the best track found in the innermost three super-layers is output. The track-finding efficiency and

[27] A particle jet is a flow of observable secondary particles produced in a spatially collimated form, as a consequence of the hadronization of partons produced in the hard collision.

the fake-rate with respect to the off-line tracks depend on the instantaneous luminosity, and were measured to be $\varepsilon \approx 96\%$, and 3%, respectively, for tracks with $p_{\mathrm{T}} \gtrsim 1.5$ GeV/c at $\mathcal{L} \simeq 10^{31}$ cm^{-2}s^{-1}. The observed momentum resolution is $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}^2 \approx 0.017(\mathrm{GeV}/c)^{-1}$, and the azimuthal resolution is $\sigma_{\varphi_6} \approx 0.3^\circ$, where φ_6 is the azimuthal angle of the track measured at the sixth COT super-layer, located at 106 cm radius from the beam line.

2.5.2 Level-2

At Level-2, an asynchronous system of custom-designed hardware processes the time-ordered events accepted by the Level-1. Additional information from the shower-maximum strip chambers in the central calorimeter and the axial hits in the SVXII is combined with Level-1 primitives to produce Level-2 primitives. A crude energy-clustering is done in the calorimeters by merging the energies in adjacent towers to the energy of a seed tower above threshold. Level-1 track primitives matched with consistent shower-maximum clusters provide refined electron candidates whose azimuthal position is known with 2° accuracy. Information from the (r, ϕ) sides of the SVXII is combined with Level-1 tracks primitives to form two-dimensional tracks with resolution similar to the off-line one. Finally, an array of programmable processors makes the trigger decision, while the Level-2 objects relative to the following event accepted at Level-1 are already being reconstructed.

Silicon Vertex Trigger

Reconstructing decay vertices on-line is technically challenging and requires constrained geometrical fitting of (previously reconstructed) high-resolution tracks at high event-rates. The Silicon Vertex Trigger (SVT) detects instead impact parameters of the charged particles, which is faster than fully reconstructing their decay vertices, but still provides information on the lifetime of the decaying particle [93]. The full spatial resolution of silicon detectors is needed to discriminate $\mathcal{O}(100\ \mu\mathrm{m})$ impact parameters from the $\mathcal{O}(10\ \mu\mathrm{m})$ beam spot. Thus the SVT requires the coincidence of hits in four axial SVXII layers with a XFT track. Since the silicon signals are digitized only after the Level-1 accept decision, the SVT is used at Level-2, whose average latency is around 20 μs. Within this time, the SVT reconstructs two-dimensional tracks in the bending plane of the spectrometer with off-line resolution, a task that typically needs thousands of milliseconds to be accomplished by the off-line CPUs. SVT speed is largely due to a highly-parallelized structure whose segmentation matches the SVXII 12-fold azimuthal symmetry: each 30° azimuthal sector of each of the six longitudinal half-barrels is processed by its own asynchronous, data-driven pipeline.

The SVT receives in input the XFT tracks and the digitized pulse-heights from four SVXII layers. It first finds charge clusters in silicon, by converting a sparsified list of channel numbers and pulse heights into charge-weighed hit centroids. At this point the pattern recognition is separated in two stages. First, a low-resolution stage is implemented by grouping together adjacent detector channels into "super-bins". Their width in the azimuthal direction is programmable, with 250–700 μm typical values. A set containing about 95% of all super-bin combinations compatible with the trajectory of a charged particle with $p_{\mathrm{T}} \gtrsim 2$ GeV/c originated from the beam line ("patterns") is calculated in advance from simulation and stored in a large memory. For each azimuthal sector, only the 32,768 most probable patterns are stored. On-line, an algorithm detects low-resolution candidate tracks called

"roads" by matching super-bins containing hits with the stored patterns. A road is a combination four excited super-bins in different SVXII layers plus the XFT track parameters, which are logically treated as additional hits (see fig. 2.8(a)). Maximum parallelism is exploited at this step to speed-up

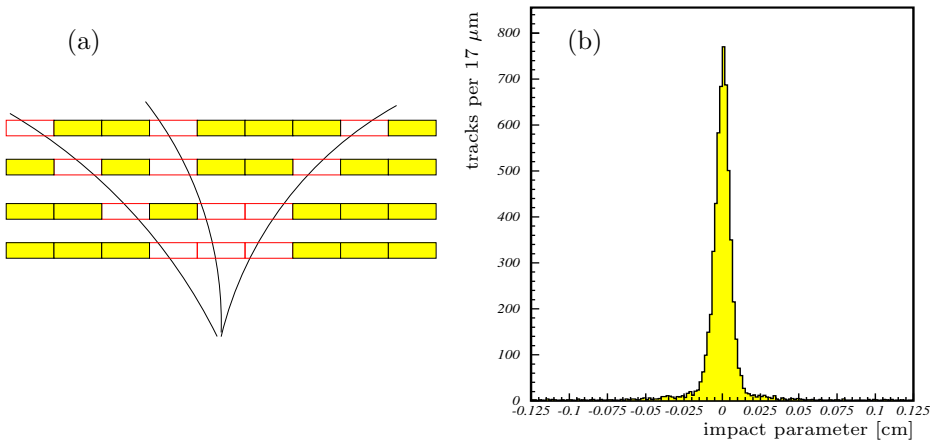

Figure 2.8: Schematic illustration of combinations of super-bins (in the transverse plane) corresponding to the passage of charged particles in four radial silicon layers (a). Impact parameter distribution as measured by the SVT (b).

the processing, using a working principle similar to the one of the bingo game: while the silicon hits are being read out, each "player" marks the matching super-bins on his "score-card"; each "bingo" corresponds to a road and is retained for further processing. A maximum of 64 roads per event, each one having a maximum of 8 hits per super-bin, is output. At this stage, pattern recognition is done during detector read-out with no additional processing time. The resolution is coarse enough to reduce the fraction of accidental combinations, but fine enough to separate most tracks. Once a track is confined to a road, most of the pattern recognition is done, leaving the remaining ambiguities, as multiple hits in the same super-bin, to the stage of track fitting.

In principle, no exact linear relation exists between the transverse parameters C, φ_0, and d_0 of a track in a solenoidal field, and the coordinates at which the track intersects a radial set of flat detector planes. But for $p_{\mathrm{T}} \gtrsim 2$ GeV/c, $|d_0| \lesssim 1$ mm and $|\Delta\varphi_0| \lesssim 15°$, a linear fit biases the reconstructed d_0 by at most a few percent. The track-fitting process exploits this feature by expanding the non-linear constraints and the parameters of the real track to first order with respect to the reference track associated to each road. A linear expansion in the hit positions of both the track parameters and the χ^2 is used. The fit process is thus reduced to computing a few scalar products, which is done within 250 ns per track. The needed constants, which depend on detector geometry and alignments, are evaluated in advance and stored in an internal memory. The output of the SVT are the reconstructed parameters of the two-dimensional track in the transverse plane: φ_0, p_{T}, and the impact parameter, d_0. The list of parameters for all found tracks is sent to Level-2 for trigger decision.

The SVT measures the impact parameter with $\sigma_{d_0,\mathrm{SVT}} \approx 35$ μm r.m.s. width, with an average

latency of 24 μs, 9 μs of which being spent waiting for the start of the read-out of silicon data. This resolution is comparable with the off-line one for tracks not using LØØ hits, and yields a distribution of impact parameter of prompt tracks with respect to the z axis with $\sigma_{d_0} \approx 47$ μm (see fig. 2.8(b)) when combined with the transverse beam-spot size.[28] The SVT efficiency is higher than 85%. This efficiency is defined as the ratio between the number of tracks reconstructed by SVT and all XFT-matched off-line silicon tracks that are of physics analysis quality.

The impact parameter is a relative quantity measured with respect to the beam. If the actual beam position in the transverse plane is shifted by an amount d_{fake} with respect to the origin of the SVT reference frame, all prompt tracks appear to SVT as having $\mathcal{O}(d_{\text{fake}})$ impact parameters. This is relevant since the beam is usually displaced from its nominal (0, 0, z) position. Between Tevatron stores, $\mathcal{O}(500$ μm) displacements in the transverse plane and $\mathcal{O}(100$ μrad) slopes with respect to the detector axis may occur. In addition, the beam can drift by $\mathcal{O}(10$ μm) in the transverse plane even during a single store. However, a simple geometric relation prescribes that the impact parameter of a track (d_0), calculated with respect to a point displaced from its production vertex, is a sinusoidal function of its azimuthal coordinate (φ_0):

$$d_0 = y_v \cos(\varphi_0) - x_v \sin(\varphi_0), \tag{2.13}$$

where $\vec{x}_v = (x_v, y_v)$ are the coordinates of the production vertex (see fig. 2.9). Using eq. (2.13), the

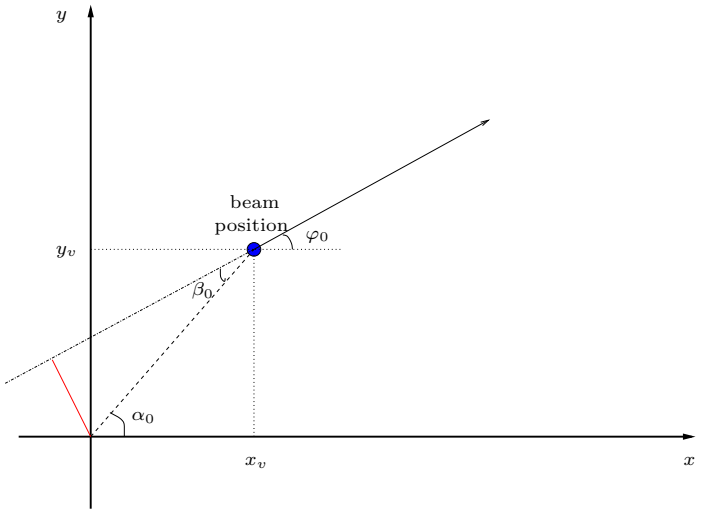

Figure 2.9: Schematic illustration, in the plane perpendicular to the beam, of a track emerging from a primary vertex displaced from the origin of the coordinate system. Equation (2.13) is easily verified by noticing that $\beta_0 = \alpha_0 - \varphi_0$.

SVT measures the actual coordinates of the beam position with respect to the detector system and subtracts them from the measured impact parameters, in order to provide physical impact parameters. Using about 10^5 tracks every 30 seconds, six transverse beam positions (one for each SVXII semi-

[28] Prompt tracks are those associated to particles produced in the hard $p\bar{p}$ interaction.

barrel) are determined on-line. The six samplings along the $\hat{z}$ direction provide a measurement of the slope of the beam with respect to the nominal z-axis.

For the proper measurement of impact parameters, the beam slope is more harmful than the transverse drift, because it breaks the cylindric symmetry of the system. The SVT does not have access to the z_0 coordinate of tracks. For each track, only the longitudinal coordinate of the SVXII half-barrel that detected the track is known. But half-barrels are too long (16 cm) for allowing a reliable correction of the beam slope. When significant slopes are observed, the Tevatron beam division is alerted, and they apply a corrective action on the magnets.

Beam mis-alignments affect also the SVT efficiency. Owing to its modular structure and to the limited size of the pattern bank, the SVT can not identify charged particles that cross adjacent SVXII wedges. In normal conditions, these are only a small fraction of $p_{\mathrm{T}} > 2$ GeV/c particles, typically due to the bending trajectory and of the finite beam-spot size. However, in presence of beam offset from the nominal position, this fraction significantly increases, thus inducing SVT inefficiency.

2.5.3 Level-3

The digitized output relative to the Level-2-accepted event arrives fragmented from all subdetectors via optical fibers. It is collected by a custom hardware switch that arranges it in the proper order and transfers it to 292 (as of this writing) commercial computers, running LINUX and organized in a modular and parallelized structure of 16 subsystems [94]. The ordered fragments are assembled in the *event record*, a block of data that univocally corresponds to a bunch crossing and is ready for the analysis of the Level-3 software. The event reconstruction benefits from full detector information and improved resolution with respect to the preceding trigger levels, including three-dimensional track reconstruction, tight matching between tracks and calorimeter or muon information, and calibration information. If an event satisfies the Level-3 requirements, the corresponding event record is transferred to mass storage at a maximum rate of 20 Mbyte/s. A fraction of the output is monitored in real time to search for detector malfunctions, to derive calibrations constants and to graphically display events. The Level-3 decision is made after the full reconstruction of the event is completed and the integrity of its data is checked, a process that takes a few milliseconds.

2.5.4 Trigger performance

A maximum of 64, 128, and 256 different combinations of selection requirements can be implemented at Level-1, Level-2 and Level-3, respectively. An unique sequence of these requirements is called a trigger logic-path. The set of all active trigger-paths, along with the corresponding thresholds and calibration constants, are downloaded to the data acquisition (DAQ) and to the trigger at the beginning of each run, during which they remain unchanged. The trigger performance was continuously optimized while the data used in this analysis were collected (February 2002–September 2003).

Level-1 – the number of different combinations of requirements increased from 40 to 56. The available output rate passed from 12 kHz to 18 kHz, of which roughly 90% were occupied by tracking-based triggers for b-physics, 5% by lepton triggers for high-p_{T} physics, and 5% by jet and photon triggers for QCD and new physics searches.

Level-2 – the number of different combinations of requirements increased from 68 to 116. The available output rate passed from 250 Hz to 300 Hz (50% tracking, 30% jet and photon, and 20% lepton).

Level-3 – the number of different combinations of requirements increased from 125 to 185. The available output rate passed from 50 Hz to 75 Hz (40% tracking, 30% jet and photon, and 30% lepton).

The trigger deadtime never exceeded 5%, and its integrated value in the data used for this analysis was below 1%.

2.6 Operations and data quality

The proper operation of the detector and the quality of the on-line data-taking is continuously ensured by "crews" of five collaborators plus one technician which alternate on duty with eight-hours shifts, plus several subdetector experts available on request. The on-line crew, in communication with the Tevatron crew, ensures smooth data-acquisition, monitors the crucial parameters of all subdetector, and intervenes in case of malfunctions. The average data-taking efficiency is 85%. The inefficiency is approximately equally shared in a 5% arising at the beginning of the store, when the detector is not powered while waiting for stable beam conditions, a 5% due to trigger deadtime, and a 5% due to unexpected detector or DAQ problems.

When no beam is present, cosmic-rays runs are taken, or calibrations of the subdetector are done. During the Tevatron shut-down periods, the crew coordinates and helps the work of experts that directly access the detector.

Each time that at least one of the trigger paths fires, an "event" is labeled with a progressive number. Events are grouped into runs, i. e., periods of continuous data-taking in constant configurations of trigger table, set of active subdetectors and so forth.[29] Several parameters of the operations (e. g., beam-line position and slope, set of calibrations, etc.) are stored in the database on a run-averaged format.

All data manipulations occurring some time after the data are written to permanent memories are referred to as *off-line* processes, as opposed to the on-line operations that take place in real time, during the data-taking. The most important off-line operation is the processing with a centralized *production* analysis that generates collections of high-level physics objects suitable for analysis, such as tracks, vertices, muons, electrons, jets, etc. from low-level information such as hits in the tracking subdetectors, muon stubs, fired calorimeter towers, etc. [95]. During the production, more precise information about the detector conditions (e. g., calibrations, beam-line positions, alignment constants, masks of malfunctioning detector-channels, etc.) and more sophisticated algorithms are used than those ones available at the Level-3 of the trigger. The production may be repeated when improved detector information or reconstruction algorithms become available: this typically occurs once or twice every year. The reprocessing uses large farms of commercial processors that reconstruct

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

approximately 10^7 events per day employing approximately 2–5 s per event with 1 GHz CPU.[30] The added information increases the event size by typically 20% after production.

To ensure homogeneous data-taking conditions, each run undergoes a quality inspection. On-line shift operators, off-line production operators, and subdetector experts certify in what fraction of data the running conditions for all relevant subdetectors are compliant to physics-quality standards.

When detectable problems of the detector occur, the data-taking is quickly stopped, so very short runs are likely to contain corrupted data. Runs with fewer than 10^8 live Tevatron clock-cycles, or fewer than 10^4 (10^3) Level-1 (Level-2) accepts, or containing data corresponding to an integrated luminosity $\int \mathcal{L}dt < 1$ nb^{-1} are excluded from physics analysis. On-line shift operators further exclude the runs in which temporary or test trigger tables were used.[31] Runs whose data underwent problems or software crashes during the production are excluded off-line.

Accurate integrated luminosity measurements are ensured in physics-quality data by requiring the CLC to be operative during the data-taking and by verifying that a set of luminosity and beam-monitor probe quantities are within the expected ranges. Shift operators ensure that Level-1 and Level-2 trigger operate correctly and that the rate of SVXII data corruption errors is smaller than 1%.[32] SVT experts verify that the on-line fit and subtraction of the beam position is done correctly and that the SVT occupancy is within the expected limits. In addition, higher level quantities, such as event yields of $J/\psi \to \mu^+\mu^-$, $D^0 \to K^-\pi^+$, and $D^{*+} \to D^0\pi^+$ decays are monitored on-line and are required to be within the expected ranges. For analyses that use COT information, the minimum integrated luminosity required is 10 nb^{-1} and the fraction of noisy COT channels is required to be smaller than 1%.

2.7 Monte Carlo simulation of detector and trigger

Estimation of the fraction of events of a certain type that escape the detector acceptance, or detailed studies of the expected response of the detector to the passage of particles is a common need in many analyses. Usually, complex detector geometries and the numerous effects that need to be accounted for in predicting their response make it the analytical derivation of the relevant distributions impractical or impossible. Monte Carlo techniques are an useful and widely-used complement for this problem. Although we chose a data-driven approach for this analysis, use of simulation was unavoidable for some purposes. We provide here a short overview of the standard CDF II simulation. Further details can be found in Ref. [96].

In the standard CDF II simulation, the detector geometry and material are modeled using the version 3 of the GEANT package [97] tuned to test-beam and collision data. GEANT receives in input the positions, the four-momenta, and the identities of all particles produced by the simulated collisions that have long enough lifetimes to exit the beam pipe.[33] It simulates their passage in the detector, modeling

[30] The event size, and the processing-time increase roughly linearly with the instantaneous luminosity.

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

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

[33] The generation of simulated samples is described in sec. 3.6.

their interactions (*bremsstrahlung*, multiple scattering, nuclear interactions, photon conversions, etc.) and the consequent generation of signals on a single channel basis. Specific packages substitute GEANT for some subdetectors: the calorimeter response is simulated with GFLASH, a faster parametric shower-simulator [98] tuned for single-particle response and shower-shape using test-beam data (8–230 GeV electrons and charged pions) and collision data (0.5–40 GeV/c single isolated tracks); the drift-time within the COT is simulated using the GARFIELD standard package [99] further tuned on data; the charge-deposition model in the silicon uses a parametric model, tuned on data, which accounts for restricted Landau distributions, production of δ-rays, capacitive charge-sharing between neighboring strips, and noise [100].[34] Furthermore, the actual trigger logic is simulated. The output of the simulated data mimics the structure of collision data, allowing their analysis with the same reconstruction programs used for collision data.

The detector and trigger configuration undergo variations during data-taking. Minor variations may occur between runs, while larger variations occur, for instance, after major hardware improvements, or Tevatron shut-down periods. For a more detailed simulation of the actual experimental conditions, the simulation has been interfaced with the off-line database that reports, on a run-by-run basis, all known changes in configuration (position and slope of the beam line, relative mis-alignments between subdetectors, trigger-table used, set of SVT parameters) and local or temporary inefficiencies in the silicon tracker (active coverage, noisy channels, etc.). This permits simulating the detailed configuration of any set of real runs to use it, after proper luminosity reweighing, for modeling the realistic detector response in any given subset of data.

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

Chapter 3

Sample selection and signal extraction

In order to extract the desired information from the $B^0_{(s)} \to h^+h'^-$ decays, the events of interest (referred to as "signal") need to be extracted from the multitude of other uninteresting events ("background"). This chapter describes the three main stages of this process. The first step is the on-line trigger selection, which collects events most likely containing $B^0_{(s)} \to h^+h'^-$ decays. The baseline selection is a next step that uses a simple off-line improvement of the trigger selection to extract a visible $B^0_{(s)} \to h^+h'^-$ signal. The optimized selection is the last step aimed at obtaining the sample such that the statistical uncertainties on the quantities one wishes to measure are minimized.

3.1 $B^0_{(s)} \to h^+h'^-$ decays at CDF

The topology of a $B^0_{(s)} \to h^+h'^-$ decay is simple. Two charged, pseudo-scalar mesons ($\pi^+\pi^-$, $K^+\pi^-$, $K^-\pi^+$, and K^+K^-) from the decay of a long-lived $B^0_{(s)}$ meson produce, in the detector, two oppositely-curved tracks intersecting in a space-point a few hundred microns away from the location of the $p\bar{p}$ interaction, where the $B^0_{(s)}$ meson was produced.

Such a simple topology poses a first experimental challenge at CDF: rare processes ($\mathcal{B} \approx 10^{-5}$), immersed in a background $\mathcal{O}(10^9)$ times larger, need to be selected on-line, relying uniquely on tracking resources. In fact, the signature of $B^0_{(s)} \to h^+h'^-$ decays lacks the most used discriminating features; for instance, no leptons are present to exploit the good CDF muon and electron identification capability, nor narrow intermediate resonances (e. g., a ϕ meson) can be used to provide additional kinematic constraints.

The $B^0_{(s)} \to h^+h'^-$ final states contain only kaons and pions which, unfortunately, are also the most common particles present in the background. However, if the momentum of the $B^0_{(s)}$ meson has a sufficiently large component in the plane transverse to the beam-line, the displacement between production and decay positions of the $B^0_{(s)}$ meson can be measured with the silicon tracker. This provides a highly discriminating quantity for background rejection.

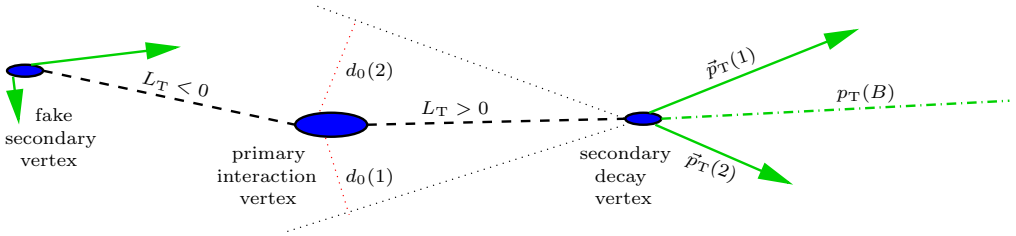

Figure 3.1: Illustration of a $p\bar{p}$ event containing a $B^0_{(s)} \to h^+h'^-$ decay, projected into the transverse plane. Ellipses indicate vertices, arrows indicate the transverse momenta (i. e., the direction) of charged particles. Nothing is to scale.

Before discussing the details of trigger and off-line selection, it is useful to define some relevant quantities used in the analysis. All quantities are calculated in the laboratory frame, and are illustrated in fig. 3.1. For the present description, we neglect the curvature of $\mathcal{O}(100\ \mu\text{m})$ samplings of trajectories of charged particles with $\mathcal{O}(\text{GeV}/c)$ momenta. In proximity of the beam line, tracks in the bending plane of the solenoid are approximated with straight lines.

Transverse plane – the plane perpendicular to the proton beam direction, in which the profile of the interaction region is approximately Gaussian with r.m.s. $\sigma_T \approx 30\ \mu$m.

Transverse momentum ($\vec{p}_T$) – the projection of the momentum vector onto the transverse plane. This quantity is the simplest discriminant between heavy-flavor signals and background as discussed in sec. 1.6.2 Raising the thresholds on transverse momentum is also an effective way to control the trigger accept-rates. Another useful quantity used in the selection is the scalar sum of the transverse momenta of the two particles, $\sum \boldsymbol{p_T} \equiv \boldsymbol{p_T(1)} + \boldsymbol{p_T(2)}$.

Primary vertex – the space-point of the reconstructed primary $p\bar{p}$ interaction, where a b-quark pair, once produced, quickly hadronizes to a b-hadron pair. The primary-vertex reconstruction has been detailed in sec. 2.3.6. An event may contain multiple primary vertices due to multiple hard $p\bar{p}$ collisions occurring in the same bunch-crossing ("pile-up" event, see sec. 1.6.2).

Secondary vertex – the space-point in which the decay of a long-lived particle occurs. The components of its displacement with respect to the primary vertex in the transverse plane are indicated by the vector $\vec{\boldsymbol{x}}_{\boldsymbol{v}} = \vec{\beta}_T \gamma ct = (\vec{p}_T/m)ct$, for a particle of mass m and momentum p that decays at a time t after its production. The secondary-vertex reconstruction has been detailed in sec. 2.3.6. Multiple secondary vertices may be present in the same event. They can be due to the intersection of tracks from various sources, including the decay of the *other* heavy-flavor produced in the event, the decay of additional heavy-flavors produced in a pile-up event, fake (i. e., due to accidental combinations of noise hits) or mis-measured (i. e., partially contaminated by noise hits) tracks.[1]

Transverse decay-length ($\boldsymbol{L_T}$) – the displacement of the secondary vertex with respect to the primary one, projected onto the transverse momentum vector of the decaying particle ($\vec{p}_T(B)$).

[1] In the dominant $p\bar{p} \to b\bar{b}\ (c\bar{c}) + X$ production process, two heavy-flavors are produced in the event.

The transverse displacement of the secondary vertex ($\vec{x}_v$) may be not collinear with $\vec{p}_{\rm T}(B)$ because of measurement uncertainties. Thus, the transverse decay-length,

$$L_{\rm T} \equiv \frac{\vec{p}_{\rm T} \cdot \vec{x}_v}{p_{\rm T}}, \tag{3.1}$$

is usually preferred to $\vec{x}_v$ as an estimator of transverse-length travelled before decay. This quantity is typically positive for true long-lived decays, while it is negative or positive with almost equal probability for decays from a fake secondary vertex or for combinations of prompt tracks, although in the latter case its value is comparable with its resolution.

Impact parameter (d_0) – the component of the distance of closest approach between a track and the primary vertex in the transverse plane. This is a signed quantity defined as

$$d_0 \equiv \frac{\hat{z} \cdot (\vec{p}_{\rm T} \wedge \vec{x}_v)}{p_{\rm T}}, \tag{3.2}$$

where the scalar product with the unit vector pointing toward the proton direction ($\hat{z}$) determines its sign and the symbol $\wedge$ indicates the vector product. The impact parameter is typically different from zero for products of long-lived decays, while it is comparable with the convolution of its resolution and the transverse size of the beam for particles produced in the vicinity of the primary vertex (prompt background).

Azimuthal opening angle ($\Delta\varphi_0$) – the opening angle between the two outgoing particles projected on the transverse plane. The distribution of this quantity in $B^0_{(s)} \to h^+h'^-$ decays depends on the distributions in impact parameter and transverse momentum. However, it has generally a slowly-varying shape for signal candidates, while it shows two enhancements around 0° and 180° for background candidates (see fig. 3.2). Pairs of quasi-collinear tracks are found in hadronic jets, due to light-quark fragmentation, or in highly occupied regions of the detector, due to combinations of fake tracks; pairs of azimuthally-opposed tracks are found in back-to-back jets of generic QCD background.

3.2 The B_PIPI trigger path

The first proposal for a trigger dedicated to $B^0_{(s)} \to h^+h'^-$ decays at CDF was presented at the Snowmass workshop on *"B physics at hadron accelerators"* in 1993 [102]. The B_PIPI trigger path, i. e., the sequence of Level-1, Level-2, and Level-3 requirements used to collect the data for this analysis, is a selection of the Displaced-Tracks Trigger consisting in an improvement of that original idea.

In the following, we outline just the most common trigger configuration, while the improvements that had relevant effects in the analysis are summarized in sec. 3.2.4.

3.2.1 Level-1 requirements

At Level-1, the reconstruction of charged-particles trajectories relies on the tracks reconstructed by the XFT in the transverse plane. Their parameters are $p_{\rm T}$, φ_6, and q. The azimuthal angle is measured at

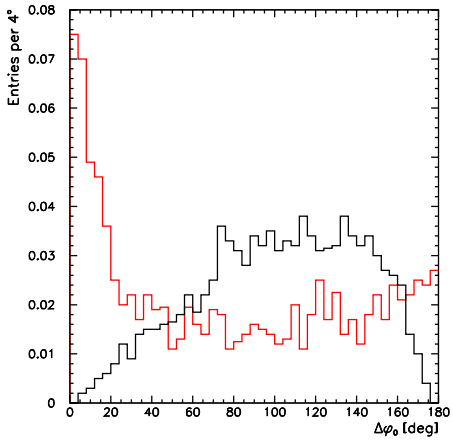

Figure 3.2: Distribution of azimuthal opening angle between tracks with $p_{\mathrm{T}} > 2$ GeV/c and $d_0 > 100$ μm from simulated $B^0 \to \pi^+\pi^-$ decays (black plot). Distribution of azimuthal opening angle for pairs of tracks with $p_{\mathrm{T}} > 2$ GeV/c from background candidates in data (red plot). Background candidates are reconstructed from the "minimum bias" sample collected by CDF Run I. The minimum bias trigger collects random $\mathrm{p\bar{p}}$ collisions thus providing a sample dominated by events with QCD light-quark production. (Plot taken from Ref. [101].)

a radial distance corresponding to the super-layer 6 of the COT and it is labeled φ_6 rather than φ_0. A pair of oppositely-curved XFT-tracks is required, consistently with the decay of a neutral particle into two charged particles. A $p_{\mathrm{T}} > 2.04$ GeV/c requirement on the transverse momentum of each XFT-track is imposed by hardware constraints to keep a sustainable trigger accept-rate. In addition, track pairs due to light-quark fragmentation in back-to-back jets are rejected with the $0^\circ < \Delta\varphi_6 < 135^\circ$ requirement on the azimuthal opening angle between tracks. To further reduce the Level-1 accept-rate, a requirement on the scalar sum of the transverse momenta of the pair is applied: $\sum p_{\mathrm{T}} > 5.5$ GeV/c.

3.2.2 Level-2 requirements

The trigger requires a pair of oppositely-curved SVT tracks that satisfy a minimal linearized-fit quality requirement: $\chi^2_{\mathrm{SVT}} < 25$ [103].[2] Transverse momentum and $\sum p_{\mathrm{T}}$ requirements from Level-1 are confirmed on SVT tracks. In addition, a $100 < |d_0| < 1000$ μm requirement is applied to the impact parameter of each track. The lower threshold rejects a large fraction of background tracks from prompt light-flavor decays; the upper threshold is dictated by hardware constraints (see sec. 2.5.2), but also reduces the contamination from strange-hadron decays, mis-measured tracks due to silicon noisy strips, and charged products of nuclear interactions in the detector material.[3]

[2] The efficiency of the $\chi^2_{\mathrm{SVT}} < 25$ requirement on unbiased samples is approximately 97%.

[3] At typical Lorentz boosts of the outgoing particles from Tevatron collisions, s-hadrons have decay length of 2–1500 cm. In addition, a low, nearly flat background of fake tracks is known to be uniformly distributed over the whole d_0 spectrum.

The impact parameter requirement is the single most selective one of the whole trigger selection. It reduces the trigger accept-rate by a factor $\mathcal{O}(100)$, while still keeping $\mathcal{O}(10\%)$ efficiency on signal. The spatial resolution of SVT in identifying secondary vertices is further exploited: positive decay-length of the $B^0_{(s)}$ candidate is required, $L_{\rm T}(B) > 200$ μm, along with a $|d_0(B)| < 140$ μm requirement on its impact parameter. The latter imposes that the candidate originates from the primary vertex, rejecting sequential $B \to DX \to h^+h^{'-}X$ decays. Events with $L_{\rm T}(B) < -200$ μm are retained as well for studies of background (see sec. 3.8.1). The azimuthal opening-angle requirement is further tightened with respect to the previous trigger stage to $20^\circ < \Delta\varphi_0 < 135^\circ$, to reduce the fraction of events with light-quark background.

3.2.3 Level-3 requirements

The Level-1 and Level-2 criteria are reapplied on Level-3 tracks. In addition, a requirement on the longitudinal separation between the two tracks at the point of their minimum distance from the beam is applied: $|\Delta z_0| < 5$ cm. This significantly reduces the fraction of combinations of two tracks descending from particles produced in distinct primary vertices (pile-up events). A $|\eta| < 1.2$ requirement on tracks excludes events with particles outside the XFT fiducial acceptance. The Level-3 mass-resolution, comparable with the off-line one, allows a loose $4.0 < m_{\pi^+\pi^-} < 6.0$ GeV/c^2 requirement on the reconstructed invariant $\pi^+\pi^-$-mass of the particle pair. This is adequate for reducing the Level-3 accept-rate, while keeping events populating a sufficiently wide mass-spectrum for signal and background studies.

3.2.4 Changes in the trigger configuration

In the first portion of data, the XFT required a coincidence of 10 out of 12 hits in each axial super-layer to form a track segment (see sec. 2.5.1). An optimized configuration requiring a coincidence of 11 out of 12 hits was applied after October 2002 (run 152646). This allowed a $\mathcal{O}(30\%)$ reduction of the Level-1 trigger rate with just $\mathcal{O}(5\%)$ efficiency loss on signal.

Another crucial improvement was implemented at Level-2: the SVT requires the coincidence of four strip clusters in four radial SVXII layers (see sec. 2.5.2), resulting in an overall efficiency proportional to the fourth power of the single hit efficiency. Until June 2003, the clusters were searched in the four innermost SVXII layers. After then, a majority logic was introduced, that searches the coincidence of four clusters in any combination of four (out of five) SVXII layers. The looser requirement yields a 18% increase in the single track efficiency, and provides additional benefits as an almost-doubled event yield for two-body decays, like $D^0 \to K^-\pi^+$, with negligible worsening of the signal purity.[4]

In the first portion of data, corresponding to about $\int \mathcal{L}dt \simeq 120$ pb^{-1}, the code for reconstructing silicon tracks was too slow to be executed within the time allowed by the Level-3 latency. The spatial resolution of the silicon detectors was therefore exploited by looking for a matching between SVT

[4]The increase in the single-track efficiency is $4(1-\varepsilon_{\rm clust})$ where $\varepsilon_{\rm clust} \simeq 95\%$ is the clustering efficiency in a single SVXII layer. The resulting 20% increase reduces to about 18% because only 90% of the SVXII wedges had five properly working layers. Further benefit is provided by the opportunity to reconstruct an additional 15% of tracks: those that cross two different SVXII electrical barrels within a wedge. In conclusion, the SVT trigger efficiency for a two-body decay increased by a factor of about $(1+0.18+0.15)^2 - 1 \simeq 80\%$.

and COT tracks, i. e., by requiring proximity in curvature ($|C^{\rm COT} - C^{\rm SVT}| < 1.5 \times 10^{-4}$ cm^{-1}) and azimuth ($|\varphi_0^{\rm COT} - \varphi_0^{\rm SVT}| < 0.86^\circ$) to form Level-3 tracks. The matched tracks had their d_0 assigned from the corresponding SVT track, whereas other parameters were assigned from the corresponding COT track. In May 2003, a faster silicon-tracking algorithm was implemented at the Level-3 trigger. Since then, Level-3 tracks are required to be associated to silicon hits in at least three different axial SVXII layers. This improvement allowed more than a factor of two reduction in Displaced-Tracks Trigger rate, with just a $\mathcal{O}(5\%)$ inefficiency for heavy-flavor decays.

3.3 Other Displaced-Tracks Triggers

This is a concise description of the `B_CHARM` path, which was implemented to collect efficiently multi-body heavy-flavor decays, such as $B_s^0 \to D_s^- \pi^+$, that are used as control samples for the present measurement. At Level-1 the `B_CHARM` and `B_PIPI` paths share common requirements; this explains why, in the `B_PIPI` path, the $\Delta\varphi_0 > 20^\circ$ requirement was applied only at Level-2, although the information needed for this requirement was available since Level-1. At subsequent levels, the `B_CHARM` selection is similar to the `B_PIPI` one but it has different thresholds on the impact parameter of SVT tracks ($|d_0| > 120$ μm), on the azimuthal opening angle ($2^\circ < \Delta\varphi_0 < 90^\circ$), and does not contain any requirement on the impact parameter of the B candidate.

During the 20 hours of a typical Tevatron store, the instantaneous luminosity decreases by a factor of typically 2.5–5 with respect to the initial luminosity. Since the set of trigger selections is optimized for an average reference luminosity, during the store, a fixed selection is either too loose, suffering high trigger accept-rates and inducing dead-time in the trigger decision, or is too tight, leaving a fraction of trigger bandwidth unused. CDF uses simultaneously trigger selections that have different $p_{\rm T}$-thresholds and self-adjusting prescale factors ("dynamic prescale") that change on the milliseconds scale to optimally exploit the available bandwidth.[5]

At luminosities $\mathcal{L} \gtrsim 4.5 \times 10^{31}$ cm^{-2}s^{-1}, the Level-1 of the Displaced-Tracks Trigger selection causes a trigger dead-time greater than the upper limit of 5%. Thus, dynamic prescaling and introduction of additional trigger paths, the `B_PIPI_HIGHPT` and `B_CHARM_HIGHPT` trigger, was required. These variants are similar to the `B_PIPI` and `B_CHARM` paths respectively, but have tighter requirements on the transverse momenta of XFT tracks ($p_{\rm T} > 2.46$ GeV/c), and on their scalar sum ($\sum p_{\rm T} > 6.5$ GeV/c) at Level-1. These additional requirements, confirmed at subsequent trigger stages, allow sustainable trigger accept-rates at high luminosities, at the price of some reduction of the signal yields.

3.4 `B_PIPI` trigger performance

Table 3.1 contains the most relevant requirements applied at each level of the `B_PIPI` trigger, along with the typical `B_PIPI` rates and the available trigger bandwidths. A semi-quantitative illustration of the impact of the trigger selection on the signal-to-background ratio for $B_{(s)}^0 \to h^+ h^{'-}$ decays is shown in

[5] A trigger prescaled of a factor N (acceptance fraction) will only accept every randomly-chosen Nth event that satisfies the trigger requirements. This reduces the trigger accept-rate by a factor N, leaving unchanged the signal-to-background ratio in the sample.

Quantity	Units	Level-1	Level-2	Level-3
$p_{\mathrm{T}}(1),\ p_{\mathrm{T}}(2)$	GeV/c	> 2.04	> 2.0	> 2.0
$\Delta\varphi_0$	Degrees	$[0°, 135°]$	$[20°, 135°]$	$[20°, 135°]$
$\sum p_{\mathrm{T}}$	GeV/c	> 5.5	> 5.5	> 5.5
$\lvert d_0(1)\rvert,\ \lvert d_0(2)\rvert$	μm	–	$[100,\ 1000]$	$[100,\ 1000]$
$\lvert d_0(B)\rvert$	μm	–	< 140	< 140
$\lvert L_{\mathrm{T}}(B)\rvert$	μm	–	> 200	> 200
$\lvert\eta(1)\rvert,\ \lvert\eta(2)\rvert$	–	–	–	< 1.2
$m_{\pi^+\pi^-}$	GeV/c^2	–	–	$[4.0, 6.0]$
Accept-rate	Hz	10000	30	0.5
Trigger bandwidth	Hz	16000	250	50

Table 3.1: Upper part: summary of the most relevant trigger requirements. Lower part: average B_PIPI accept-rates at $\mathcal{L} = 4 \times 10^{31}$ cm^{-2}s^{-1}, a typical peak luminosity in 2002–2003, when the data of this analysis were collected.

fig. 3.3, where we assumed a production cross-section times branching fraction $\sigma_{\mathrm{p\bar{p}}\to B^0_{(s)}+X} \times \mathcal{B}(B^0_{(s)} \to h^+h'^-) \approx 30\ \mu\mathrm{b} \times 3 \times 10^{-5} \approx 1$ nb for the signal, and an inelastic p$\bar{\mathrm{p}}$ cross-section $\sigma_{\mathrm{p\bar{p}}} = 60$ mb at $\sqrt{s} = 1.96$ TeV.

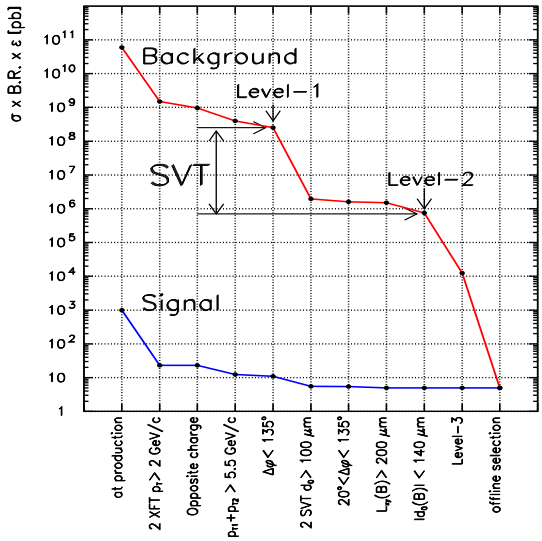

Figure 3.3: Cross-section times branching-fraction times detection efficiency for $B^0_{(s)} \to h^+h'^-$ decays compared with the cross-section times efficiency for generic inelastic p$\bar{\mathrm{p}}$ background, as a function of the selection requirements.

3.5 Extraction of the $B^0_{(s)} \to h^+h'^-$ signal

The first step of the off-line analysis consists in applying a *baseline* selection to the sample collected by the B_PIPI trigger between February 9, 2002, (run 138815) and September 2, 2003, (run 168640). After application of the standard CDF data-quality requirements (see sec. 2.6) the sample size correspond to an integrated luminosity of $\int \mathcal{L}dt = 180 \pm 11$ pb^{-1}. In the baseline selection, trigger requirements are reapplied using high-resolution off-line quantities to remove $B^0_{(s)} \to h^+h'^-$ candidates not satisfying the trigger selection, and two-particle secondary-vertices are fitted in the transverse plane.

3.5.1 Track preparation

This analysis is based on tracks. Tracks were reconstructed by the production executable (see sec. 2.3.6) using the SVXII and COT hits, the detailed magnetic map of the tracking volume, and taking into account the measured angular and translational mis-alignments among the SVXII, the COT, and the beam-line. Silicon hits in the LØØ and ISL detectors were not used because the alignment corrections for these subdetectors were not available at the time of this analysis.

For the sake of an accurate determination of the decay lengths of heavy-flavor decays, it is necessary to use events in which both primary and secondary vertices are reconstructed with good spatial resolution. We therefore used only tracks whose reconstruction included silicon hits. Mis-alignments and noise hits in the silicon detectors and in the COT cause a contamination of fake or mis-measured tracks in the events. The fraction of such undesirable tracks was reduced by applying a set of standard criteria: we selected tracks reconstructed using at least 42 COT hits, of which a minimum of 20 were found in the axial super-layers, and a minimum of 20 in the stereo super-layers. Each track was also required to be associated to hits in at least three $r-\phi$ SVXII layers and to result from a converged helix-fit with positive-definite error matrix.

To increase computation speed in the off-line production, the error matrix of the track fit in the COT is estimated disregarding the effect of multiple scattering in the detector material. According to the standard CDF prescription, we compensated for this approximation by refitting the tracks, after rescaling the covariance matrix of the COT track with an appropriate set of empirical scale factors. The rescaled COT track is used to seed the refit of the combined COT+SVXII tracks. The refitting uses an algorithm based on Kalman filtering [104] and includes energy-loss corrections for kaons and pions, according to the chosen mass assignment for each particle. The refitting procedure, the tracking alignments, and the GEANT description of the detector material have been carefully studied at CDF by independent analyses for the measurement of b-hadron masses [105].

3.5.2 Trigger confirmation

Since SVT tracks are determined with a different fitting algorithm with respect to the off-line tracks, the sample may contain candidates that did not satisfy the trigger selection ("volunteers"). Common sources of volunteers are, for instance, track pairs in which a track from a b-hadron decay is combined with a fake track, or with a track from the decay of the other b-hadron of the event. Volunteers are undesirable in decay-rate measurements like ours, because we need to use Monte Carlo simulation

to determine the reconstruction efficiencies for each mode. Since the Monte Carlo does not simulate volunteers (see sec. 3.6), we need to exclude them also from data. Volunteers were excluded by requiring matching between the offline-track pair forming the $B^0_{(s)} \to h^+h'^-$ candidate and two SVT tracks in each event;[6] then the complete set of trigger requirements was applied to SVT quantities of the matched tracks, thus repeating the real trigger decision in the off-line analysis.

3.5.3 Reconstruction of $B^0_{(s)} \to h^+h'^-$ candidates

Decays of b-mesons are expected to appear as peaking structures in the otherwise smooth invariant mass spectrum. The off-line reconstruction of $B^0_{(s)} \to h^+h'^-$ candidates was solely based on tracking, disregarding any form of particle identification. In each event, the two-particle invariant mass was computed for all possible pairs of oppositely-curved tracks satisfying the criteria described in sec. 3.5.1 and 3.5.2. We used the measured momenta and we arbitrarily assigned the charged-pion mass to both tracks. The two tracks were constrained by the CTVMFT algorithm (see sec. 2.3.6) to originate from a common vertex in the transverse plane. The vertex fit was restricted to the transverse plane because the alignment constants for the (r, z) layers of the SVXII, needed for a three-dimensional vertex fit, were not available at the time of this analysis. In case of a converged vertex-fit with satisfactory quality, the pair was promoted to be a $B^0_{(s)} \to h^+h'^-$ candidate, and retained for further processing, if its invariant $\pi^+\pi^-$-mass lie in the $4.2 < m_{\pi^+\pi^-} < 6.0$ GeV/c^2 range. During the reconstruction, we applied a baseline selection, consistent in reapplying the trigger selection on off-line quantities (see tab. 3.1), to quickly reject uninteresting candidates. In addition, we rejected tracks reconstructed outside the SVT fiducial acceptance ($|\eta| \lesssim 1.0$), and pairs with positive product of impact parameters.[7]

The invariant $\pi^+\pi^-$-mass distribution of the resulting sample is shown in fig. 3.4. A bump appears at masses around the nominal $B^0_{(s)}$ meson masses. A simple, binned χ^2-fit of the distribution to a Gaussian function for the signal, over a negative exponential plus a constant function for the background, provides an estimate of about 1200 signal events. The signal is centered at about 5.25 GeV/c^2, with about 30 MeV/c^2 r.m.s. deviation, and ≈ 0.2 signal-to-background ratio at the peak.

The distribution in fig. 3.4 is an important achievement: for the first time in a hadron collider, $B^0_{(s)} \to h^+h'^-$ decays, with $\mathcal{O}(10^{-5})$ branching-fractions, are reconstructed, and this is made possible even at the level of trigger selection by the SVT.

3.6 Monte Carlo simulation of $B^0_{(s)} \to h^+h'^-$ decays

This section describes the simulated Monte Carlo samples used in several parts of the analysis.

We used the BGENERATOR package which allows generating large samples of b-hadron decays [107]. BGENERATOR simulates production and decay of b-hadrons only: no fragmentation products, collision

[6] The algorithm required proximity in curvature and azimuthal opening angle: $\sqrt{\frac{(C^{\rm off}-C^{\rm SVT}-\overline{C})^2}{\sigma_C^2} + \frac{(\varphi_0^{\rm off}-\varphi_0^{\rm SVT}-\overline{\varphi}_0)^2}{\sigma_{\varphi_0}^2}} < 95$, where "off" labels off-line quantities, "SVT" labels SVT quantities, the mean of the differences between off-line and SVT quantities, measured in $J/\psi \to \mu^+\mu^-$ data, are over-lined, and σ are the corresponding standard deviations.

[7] An independent analysis of the similar $D^0 \to h^+h'^-$ decays [106] proved the $d_0(1) \times d_0(2) < 0$ μm^2 requirement to be effective in increasing the signal purity with negligible inefficiency.

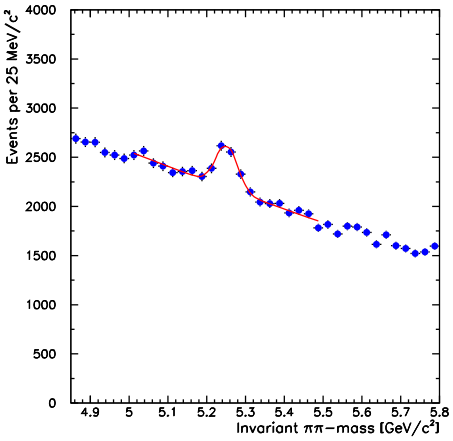

Figure 3.4: Invariant $\pi^+\pi^-$-mass distribution of the events passing the trigger selection and reconstructed off-line. The result of a Gaussian (signal) over a negative exponential and a constant (background) fit is overlaid in red.

remnants, or pile-up events are present in the simulated data. This package provides less complex events, particularly suited for a fast processing with the detector simulation, but no information about QCD backgrounds or fragmentation phenomenology can be extracted from the simulated samples. This information is available when using full $p\bar{p}$ interaction generators, like PYTHIA, but still they would require extensive and dedicated tuning to reproduce the phenomenology of events with the accuracy required by b-physics analyses. Furthermore, a huge amount of computing power would be needed to generate background samples of adequate size for analyses like ours, with $\mathcal{O}(10^9)$ rejection factors for background. Hence we chose the simpler, and more reliable, approach of using Monte Carlo simulation for the signal, while extracting the information on background from collision data.

For each of the following modes: $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B^0 \to K^+K^-$, $B_s^0 \to K^+K^-$, $B_s^0 \to K^-\pi^+$, $B_s^0 \to \pi^+\pi^-$, $\Lambda_b^0 \to p\pi^-$, and $\Lambda_b^0 \to pK^-$, we generated 10^6 decays for a resulting sample approximately 20 times larger in size than the data sample after the trigger and analysis requirements. Single $B_{(s)}^0$ mesons were directly generated using a two-dimensional distribution in transverse momentum and rapidity of the $B_{(s)}^0$ meson as external input (see fig. 3.5). This distribution combines the spectrum in transverse momentum measured in $B_{(s)}^0 \to J/\psi X$ decays by CDF [56], and a constant rapidity-distribution in the range $|Y| < 1.3$. The time evolution and decay of b-hadrons were simulated according to the most recent experimental and theoretical knowledge by the EVTGEN package [108], which properly accounts for phase space, helicity, and angular distributions of the decay products.[8] We used EVTGEN with the option of no flavor oscillations ($\Delta m_d = \Delta m_s = 0$).

Since the analysis relies on selective requirements on lifetime-sensitive quantities (e. g., d_0 and L_T), a proper treatment of $B_{(s)}^0$ meson lifetimes in the simulation is crucial for an accurate description of

[8] EVTGEN was originally developed and extensively tested at the *BABAR* and CLEO experiments.

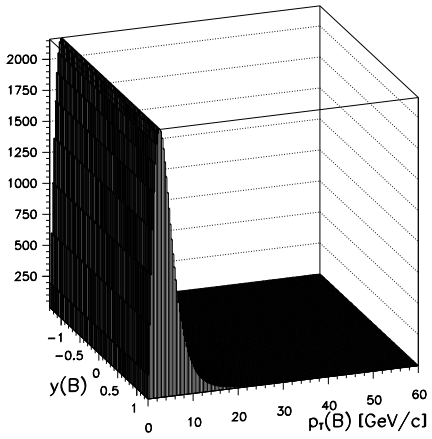

Figure 3.5: Two-dimensional distribution in rapidity and transverse momentum of the b-hadron used in the generation of simulated $B^0_{(s)} \to h^+h^{'-}$ decays.

the data. For the pseudo-proper decay-length of B^0 and Λ^0_b decays we used the world-average values:[9] $c\tau(B^0) = 460 \pm 4$ μm and $c\tau(\Lambda^0_b) = 368 \pm 24$ μm [109]. For the $B^0_s \to K^-\pi^+$ mode, we chose the world-average value measured in flavor-specific decays: $c\tau(B^0_s \to K^-\pi^+) = 438 \pm 17$ μm [109]. Special care was needed for the $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+\pi^-$ decays. Their time evolution depends on the CP contents of these modes and on the relative width-difference between the B^0_s mass-eigenstates $(\Delta\Gamma_s/\Gamma_s)$.[10] At first order, it is written as:

$$\Gamma(B^0_s(t) \to h^+h^-) + \Gamma(\overline{B}^0_s(t) \to h^+h^-) \propto R_H e^{-\Gamma_{s,H}t} + R_L e^{-\Gamma_{s,L}t}, \tag{3.3}$$

where R_H (R_L) is the relative contribution of the heavy (light) mass eigenstate to the given mode and $\Gamma_{s,H}$ ($\Gamma_{s,L}$) is the corresponding decay-width. Neither the CP contents of $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+\pi^-$ decays, nor the value of $\Delta\Gamma_s/\Gamma_s$ are known. However, a reasonable assumption can be made, based on Standard Model predictions. Since the K^+K^- and $\pi^+\pi^-$ states are in S-wave, they are CP-even states. For these decays, under standard assumptions, the penguin amplitude is dominant over the tree amplitude. Since the penguin amplitude has zero CP-phase, only the CP-even eigenstate of the B^0_s meson decays to K^+K^- and $\pi^+\pi^-$ final states. Thus, 95% of these decays is contributed by the CP-even (short) component within the Standard Model. In case of new physics, this is no longer true, since a CP-phase can arise from the new-physics-induced penguin amplitude. We therefore calculated the values for $c\tau(B^0_s \to h^+h^-)$ under the following assumptions:

- the $B^0_s \to h^+h^-$ modes are 100% pure decays of the short B^0_s eigenstate;

[9] Pseudo-proper decay-length is intended as the product of the mean life with the speed of light.

[10] We adopt the convention $\Delta\Gamma_s = \Gamma_{s,L} - \Gamma_{s,H}$ where H and L refer to the heavy and light mass eigenstates of the B^0_s meson. According to the Standard Model, the following relations hold: $\Delta\Gamma_s/\Gamma_s > 0$, $B^0_{s,L} \equiv B^0_{s,\mathrm{even}} \equiv B^0_{s,\mathrm{short}}$, and $B^0_{s,H} \equiv B^0_{s,\mathrm{odd}} \equiv B^0_{s,\mathrm{long}}$, since mass and CP-eigenstates coincide.

- the value of the width difference is $\Delta\Gamma_s/\Gamma_s = 0.12 \pm 0.06$, according to the Standard Model predictions [110];

- the widths for flavor-specific B^0 and B_s^0 decays are equal, according to the Standard Model predictions: $\Gamma_s = \Gamma_d$.

This results in

$$\begin{aligned} c\tau(B_s^0 \to h^+h^-) &= \frac{c}{\Gamma_{s,\text{short}}} = \frac{c}{\Gamma_s + \Delta\Gamma_s/2} = \frac{c}{\Gamma_d + 0.12 \cdot \Gamma_d/2} \\ &= [434 \pm 4(\Gamma_d) \pm 12(\Delta\Gamma_s)]\ \mu\text{m}, \end{aligned} \tag{3.4}$$

where we separated the uncertainty into the 1% contribution from the uncertainty on Γ_d, and the contribution from the 0.06 uncertainty on $\Delta\Gamma_s/\Gamma_s$.

Lastly, all needed information on final-state particles is passed to the CDF II detector and trigger simulation that produces the simulated event in the standard format (see sec. 2.7).

Figure 3.6 shows the two-particle invariant mass distributions for some simulated samples. Each decay mode is reconstructed assigning the correct mass to each particle. The central part of each distribution was fit with a simple, binned χ^2-fit to a Gaussian distribution. All distributions were correctly centered at the nominal B^0 or B_s^0 meson masses, with about 22 MeV/c^2 estimated r.m.s. width. A dependence of the observed width on the $B_{(s)}^0$ meson transverse momentum was found in the simulation. The width increased by about 1 MeV/c^2 in increasing the $B_{(s)}^0$ meson transverse momentum from 4 to 20 GeV/c. Similar features were observed for the other simulated modes, not reported in fig. 3.6 for brevity.

We verified the simulation by comparing the distributions of several quantities used in the analysis with collision data. In data, a significant contribution from background events contaminates the signal region (see fig. 3.4). Thus, for extracting unbiased distributions of the signal quantities, we did a background subtraction: for each quantity to be compared, we subtracted the distributions for *background* candidates from the distributions for *signal plus background* candidates. The signal plus background candidates are defined as those found in the invariant-mass range $5.125 < m_{\pi^+\pi^-} < 5.400$ GeV/c^2 of fig. 3.4. For the background candidates, we assumed that their contribution underneath the signal peak is dominated by pairs of random tracks satisfying the selection requirements (combinatorial background). We sampled this component using candidates at masses higher with respect to the signal peak since candidates at lower masses include partially reconstructed $B_{(s)}^0$ meson decays, such as $B^0 \to \rho^+\pi^- \to [\pi^+\gamma]\pi^-$, which are kinematically excluded from the signal region;[11] we therefore used, as background candidates, those in the mass range $5.37 < m_{\pi^+\pi^-} < 5.55$ GeV/c^2. Before the subtraction, the distribution of background candidates was appropriately rescaled to the number of background events expected underneath the signal.

We compared the distributions of several variables of the two outgoing particles (transverse momentum, pseudo-rapidity, and impact parameter), and of the candidate (transverse momentum, pseudo-rapidity, sum of transverse momenta of outgoing particles, azimuthal opening angle, transverse decay-length, impact parameter). The agreement between simulation and data is satisfactory for all variables,

[11] See sec. 3.8.1 for further details on the background composition.

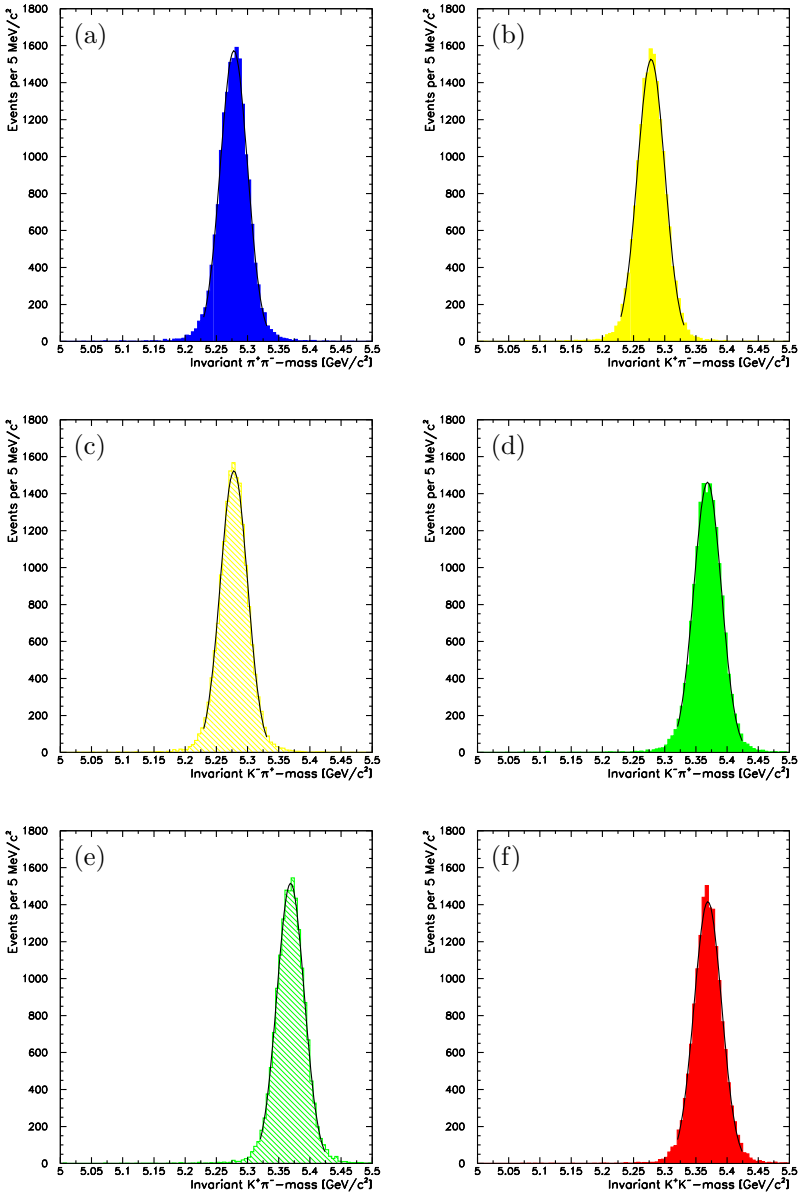

Figure 3.6: Two-particle invariant mass distributions in simulated samples of (a) $B^0 \to \pi^+\pi^-$ and $\overline{B}^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $\overline{B}^0 \to K^-\pi^+$, (d) $B^0_s \to K^-\pi^+$, (e) $\overline{B}^0_s \to K^+\pi^-$, (f) $B^0_s \to K^+K^-$ and $\overline{B}^0_s \to K^+K^-$ decays. The result of a simple Gaussian fit is overlaid (black, solid line).

indicating that the simulated sample describes the data with enough accuracy for our purpose. Figure 3.7 shows an example of such agreement.

Since the simulation available when the analysis was done did not reproduce well the *resolutions* on the tracking quantities (e. g., $\sigma_{p_{\rm T}}$, $\sigma_{L_{\rm T}}$ and so forth), we did not use discriminating variables based

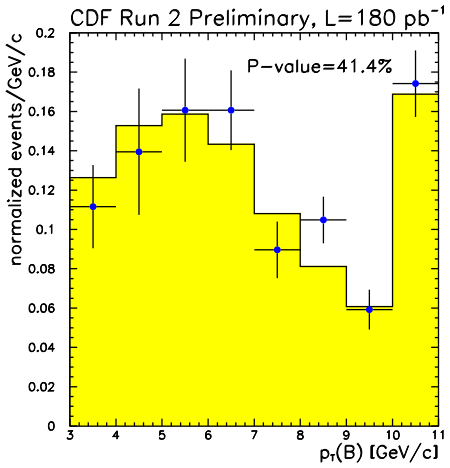

Figure 3.7: Background-subtracted $p_{\mathrm{T}}(B)$-distribution in $B^0_{(s)} \to h^+h^{'-}$ decays. Data (points with error bars) are compared with Monte Carlo simulation (filled histogram). The rightmost bin contains overflows.

on these resolutions. The only relevant consequence of this limitation is an underestimation of the mass widths in the simulated data.[12] Based on the comparison between the observed widths of several two-body decays ($D^0 \to h^+h^{'-}$, $J/\psi \to \mu^+\mu^-$, and $\Upsilon \to \mu^+\mu^-$) and the predictions of the simulation, we determined an expected r.m.s. width $\sigma_m = 28 \pm 0.3$ MeV/c^2 for each single $B^0_{(s)} \to h^+h^{'-}$ decay mode, instead of the $\sigma_m = 22 \pm 0.2$ MeV/c^2 estimated from simulation in fig. 3.6.

3.7 Improved extraction of the $B^0_{(s)} \to h^+h^{'-}$ signal

For the most efficient use of the statistical information in our data, we optimized the selection criteria, starting from the sample shown in fig. 3.4.

3.7.1 Unbiased selection optimization

For any given sample, an optimization of the selection is a procedure that selects a subsample that provides the smallest (expected) statistical uncertainty on the quantity one wishes to measure. It is often desirable that the optimization be conducted in an unbiased way, i. e., free from subjective inputs or arbitrary tuning, to ensure the reproducibility of results in independent samples. Biases may affect the optimization when this is performed on the same data set used for the measurement. This may lead to artificial enhancements of the precision of the measurement, due not to a better use of the statistical information, but to sample-specific random fluctuations. This bias can be avoided, for instance, by using simulated data samples for the optimization, or by excluding from the measurement

[12] For $D^0 \to h^+h^{'-}$ decays, for instance, the observed mass-width in data was around 10.0 MeV/c^2, to be compared with approximately 7.9 MeV/c^2 in the simulation.

the subsample of data used for the optimization.

In principle, to define an unbiased optimization for our measurement of ratios of branching fractions, one should try all possible configurations of selection requirements, repeat the measurement in each of the resulting subsamples, and then apply the optimal selection. In practice, the number of the possible selections is often so large that repeating the full measurement many times may become difficult. However, a fast and reliable method to evaluate the resolution expected from a measurement before actually carrying it out is provided by the Minimum Variance Bound (MVB) [111] (see description in appendix B for details). Given the data, the MVB analytically provides an upper bound to the precision that can be achieved on a parameter, whatever the estimation procedure used. The problem is therefore reduced to finding the analytical expression of the MVB for measuring the branching ratio of a signal over a background. In the simplified case of a counting experiment to determine the number S of signal events within a total number of $S+B$ events, the expected statistical resolution on the signal yield σ_S, estimated with the MVB, obeys the following expression:

$$\sigma_S \propto \frac{1}{\mathcal{S}} = \frac{\sqrt{S+B}}{S}. \tag{3.5}$$

An optimization procedure would imply, in this case, to evaluate the quantity $\mathcal{S}$ for all combinations of selection requirements. The optimal selection would be one that maximizes $\mathcal{S}$. The expression of eq. (3.5), which is rigorously valid for a counting experiment, is still sufficiently accurate in the case of a likelihood fit of a continuous distribution. We therefore based our optimization on the quantity $\mathcal{S}$, which was maximized as a function of different configurations of the selection requirements.

3.7.2 Isolation of the b-meson

Before proceeding with the details of the optimization, it is convenient to introduce an additional variable, useful in the off-line selection: the isolation.

Given their hard fragmentation, b-hadrons tend to carry a larger fraction of the transverse momentum of the particles produced in the fragmentation, with respect to lighter hadrons [112]. Following several Run I analyses [113], we constructed the variable "isolation of the B candidate". The isolation is an estimator of the fraction of momentum, available from the b-quark fragmentation, carried by the b-meson:

$$I|_{R=1}(B) = \frac{p_{\mathrm{T}}(B)}{p_{\mathrm{T}}(B) + \sum\limits_{i \neq j : B \to j}^{R} p_{\mathrm{T}}(i)}, \tag{3.6}$$

where the sum in the right-hand term of the denominator runs over all fragmentation tracks, identified as tracks (other than those of the B candidate decay-chain) satisfying standard track-quality requirements and found in a local region around the flight direction of the B candidate.[13] Such region is parameterized as a cone in the $(\eta - \phi)$ space, unitary in radius ($R = \sqrt{\phi^2 + \eta^2} = 1$), whose apex is the primary vertex and the axis collinear with $\vec{p}_{\mathrm{T}}(B)$ (see fig. 3.8).

[13] For the computation of the isolation variable, we loosened the requirements of sec. 3.5.1: we used all tracks with a converged helix-fit, with no requirement on the minimum number of associated COT or silicon hits.

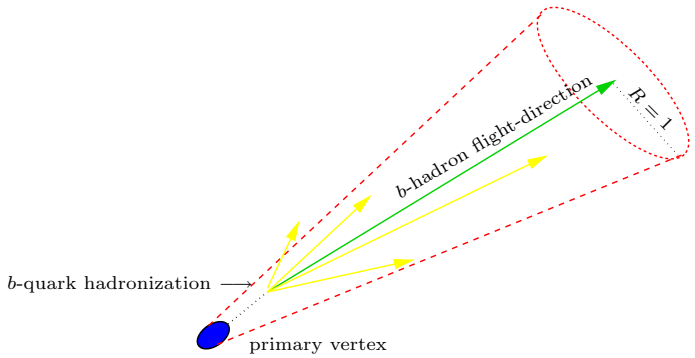

Figure 3.8: Illustration of the fragmentation of a b-quark into a b-hadron. The isolation cone is shown in red, the directions of the charged particles produced in the fragmentation are shown in yellow. Nothing is to scale.

When the decay products of the b-meson are contained in the cone, $I(B)$ is just the fraction of transverse momentum within the cone carried by the b-meson.[14] Candidates with high values of isolation are more likely to be b-mesons than candidates with low isolation.

The introduction of the isolation, however, adds further complexity in the analysis: its distribution depends on the mechanism of hadronization of the b-quark, which is not described by the signal-only simulation discussed in sec. 3.6. We therefore had to use collision data to characterize this new observable in the analysis. Section 8.5 is devoted to the extraction of the efficiency of the isolation requirement.

3.7.3 Optimized selection

In the optimization, we varied the thresholds on the following discriminating quantities: d_0 of the two tracks descending from the candidate, $L_{\rm T}(B)$, $|d_0(B)|$, $\sum p_{\rm T}$, and $I(B)$. We avoided using signal events from data to keep the optimization unbiased; for each jth configuration of the selection requirements we evaluated the quantity

$$\mathcal{S}_j = \frac{S_j}{\sqrt{S_j + B_j}}, \tag{3.7}$$

function of the expected number of signal (S_j) and background (B_j) events, defined as follows:

S_j – the number of simulated $B^0_{(s)} \to h^+h^{'-}$ decays that passed the jth configuration of the selection requirements, normalized to the 1200 $B^0_{(s)} \to h^+h^{'-}$ decays observed in data after the baseline selection. This guarantees that, in each step of the optimization, signal and background yields are correctly normalized. The $B^0_{(s)} \to h^+h^{'-}$ signal is simulated as an admixture of the expected dominant modes in the proportions resulting from the theoretical and experimental knowledge at the time of the analysis: $B^0 \to \pi^+\pi^-$ (15%), $B^0 \to K^+\pi^-$ (60%), $B^0_s \to K^+K^-$ (20%), and $B^0_s \to K^-\pi^+$(5%).

B_j – the number of background events passing the jth configuration of the selection requirements.

[14] Since we use $R = 1$ all through this thesis, we henceforth assume $I(B) \equiv I|_{R=1}(B)$.

We define background events as those found in the $5.37 < m_{\pi^+\pi^-} < 5.55$ GeV/c^2 invariant $\pi^+\pi^-$-mass region ("mass sideband") in the data.[15] We adequately normalized the amount of the events found in the sideband to what was expected underneath the peak.

Starting from the baseline selection described in sec. 3.5.3, each requirement was tightened independently of the others, scanning an adequate range of values for its threshold. Table 3.2 shows the chosen ranges and step widths for each requirement, which resulted in 4.8×10^3 total selections.

The effectiveness of the procedure relies on the capability to reproduce the real efficiency, for background and for signal, corresponding to each selection. Reliable efficiencies for background were ensured by our choice of using collision $B^0_{(s)} \to h^+h^{'-}$ data. For the signal, the validation of sec. 3.6 guarantees that the simulation reproduces with sufficient accuracy the efficiencies of all quantities used in the optimization, except for the isolation. The efficiency of the isolation requirement was extracted separately and assumed factorizable with the other efficiencies. We extracted it from control samples of $B^0_{(s)}$ decays fully reconstructed in data with a method described in sec. 8.5.

The optimal selection is reported in the last column of tab. 3.2.

Quantity	Units	Lower edge	Higher edge	Step width	*Optimum*
Minimum $L_{\rm T}$	μm	200	650	50	> 300
Minimum $\sum p_{\rm T}$	GeV/c	5.5	7.5	0.5	> 5.5
Min($\lvert d_0(1)\rvert$, $\lvert d_0(2)\rvert$)	μm	100	250	10	> 150
Maximum $\lvert d_0(B)\rvert$	μm	140	50	10	< 80
Minimum $I(B)$	–	0.4	1.0	0.1	> 0.5

Table 3.2: Quantities used in the optimization of the selection and ranges in which they were varied. The results of the optimization are shown in last column.

The dependence of $\mathcal{S}$ on each scanned quantity is shown in fig. 3.9.

3.8 The final sample

Table 3.3 summarizes the final requirements used to select the $B^0_{(s)} \to h^+h^{'-}$ sample from which we measured the decay rates. Figure 3.10 shows the $\pi^+\pi^-$-invariant mass distribution of the sample resulting from the optimized selection. A clear $B^0_{(s)} \to h^+h^{'-}$ signal peak emerges from a smooth background. A satisfactory (reduced $\chi^2 \simeq 1.1$) binned fit of the distribution to a Gaussian function for the signal, over a negative exponential plus a constant function for the background, provides an estimate of 893 ± 47 signal events, centered at 5.253 ± 0.002 GeV/c^2, with 38 ± 2 MeV/c^2 r.m.s. width, and an approximate 2.1 signal-to-background ratio at the peak. The impact of the optimized selection is an improvement of a factor 20 in background rejection, with just a 25% inefficiency on signal with respect to the baseline selection.

[15] This mass sideband was chosen according to the arguments mentioned in sec. 3.6, and detailed in sec. 3.8.1.

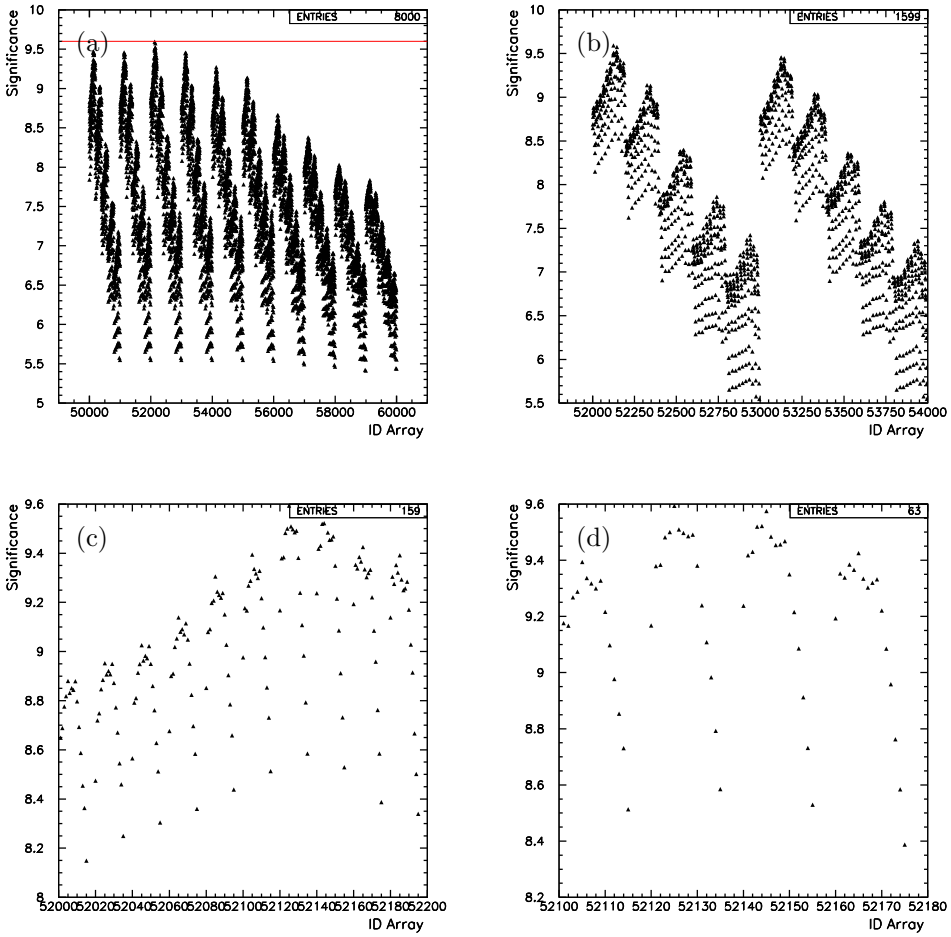

Figure 3.9: Figure of merit $\mathcal{S}$ as a function of the combination of selection requirements; each combination is uniquely identified by an ID index on the abscissa. The requirement on minimum isolation is $I(B) > 0.5$ in all plots. Plot (a) spans all combinations; the red line intersects the point corresponding to the optimal selection. Magnification of the region corresponding to $L_{\mathrm{T}}(B) > 300$ μm (left structure) and $L_{\mathrm{T}}(B) > 350$ μm (right structure) requirements (b). Magnification of the region corresponding to the requirements $L_{\mathrm{T}}(B) > 300$ μm and $\sum p_{\mathrm{T}} > 5.5$ GeV/c (c). Magnification of the region corresponding to the requirements $L_{\mathrm{T}}(B) > 300$ μm, $\sum p_{\mathrm{T}} > 5.5$ GeV/c, and 90 μm $> |d_0(B)| > 60$ μm (d).

Quantity of the track	Units	Requirement
Axial Si hits	–	≥ 3
Axial COT hits	–	≥ 20
Stereo COT hits	–	≥ 20
Total COT hits	–	≥ 42
p_{T}	GeV/c	> 2.0
$\lvert\eta\rvert$	–	< 1.0
$\lvert d_0\rvert$	μm	[150, 1000]
Quantity of the candidate		
$q(1) \times q(2)$	e^2	-1
$d_0(1) \times d_0(2)$	$\mu\mathrm{m}^2$	< 0
L_{T}	μm	> 300
$\sum p_{\mathrm{T}}$	GeV/c	> 5.5
$\lvert d_0\rvert$	μm	< 80
I	–	> 0.5
$\lvert\eta\rvert$	–	< 1.0
$\Delta\varphi_0$	Degrees	[20°, 135°]
$m_{\pi^+\pi^-}$	GeV/c	[4.2, 6.0]

Table 3.3: Summary of the optimized selection for the extraction of the final $B^0_{(s)} \to h^+h'^-$ sample.

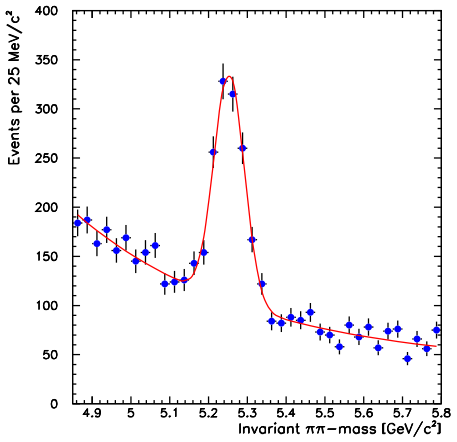

Figure 3.10: Invariant $\pi^+\pi^-$-mass distribution of the events passing the optimal selection. The result of a Gaussian (signal) over a negative exponential and a constant (background) fit is overlaid in red.

3.8.1 Background composition

Owing to the extremely selective criteria used to isolate the final sample, a detailed understanding of the background is challenging. Contributions to the background include an unknown admixture

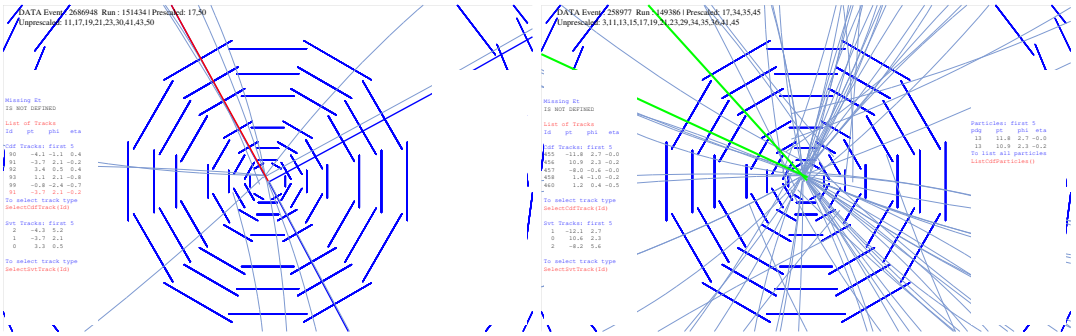

Figure 3.11: Computerized reconstruction of two events containing $B^0_{(s)} \to h^+h'^-$ candidates as they appear in the transverse plane. The LØØ and SVXII cross sections are visible in the center (blue segments).

of rare events from heavy-flavors, light-quarks, resolution tails and so forth. However, the invariant $\pi^+\pi^-$-mass distribution already provides a first insight on the background composition. This becomes evident if the distributions for the events passing the final selection (black plot in fig. 3.12, labeled as "signal sample") and the corresponding distribution for the events with "negative-$L_{\rm T}$" are compared. The negative-$L_{\rm T}$ sample contains events that pass the selection of tab. 3.3 except for the requirement on the transverse decay-length, which is inverted to $L_{\rm T}(B) < -300$ μm (see fig. 3.12, red plot). The negative-$L_{\rm T}$ sample is composed of "unphysical" decays, i. e., decays in which the reconstructed direction of the transverse momentum vector of the candidate appears opposite to the reconstructed direction of its displacement with respect to the primary vertex (see sec. 3.1). The invariant-mass shape of these random pairs of tracks is independent of their $L_{\rm T}(B)$ to a good approximation; thus, the negative-$L_{\rm T}$ sample provides an useful model of this background component in the signal sample. From the comparison of the two distributions shown in fig. 3.12, we infer the background composition as follows

Combinatorial background – it is mostly composed of random pairs of charged particles, displaced from the beam-line ($|d_0| > 150$ μm), accidentally satisfying the selection requirements. Its dominant sources include generic QCD background of light-quark decays, lepton pairs from Drell-Yan processes [114], pairs of mis-measured tracks, combinations of a mis-measured track with a track from an heavy-flavor decay, or combinations of two tracks originated from two independent heavy-flavor decays of the event ($b\bar{b}$ and $c\bar{c}$ production).[16] This is consistent with the smooth, slowly decreasing invariant $\pi^+\pi^-$-mass distribution of the negative-$L_{\rm T}$ sample that populates the whole mass range. This distribution is almost coincident with the corresponding shape in the signal sample for masses higher than 5.4 GeV/c. In this region, therefore, as well as in the signal region ($5.125 < m_{\pi^+\pi^-} < 5.4$ GeV/c), the combinatorial component is the prominent contribution to the background.

Partially-reconstructed heavy-flavor decays – a change in the slope of the mass distribution of

[16] Mis-measured tracks, i. e., tracks (partially) based on noise hits, are known to have an approximately constant distribution in impact parameter extended up to large values.

the signal sample, at masses just smaller than the signal mass, indicates an additional background source having an $L_{\rm T}$ distribution biased toward positive values. This contribution is readily interpreted as mis-reconstructed b-hadron decays. These are multi-body b-hadron decays (e. g., $B^0_{(s)} \to \rho^-\pi^+$, $B^0 \to \rho^- K^+$, $B^0_s \to K^-\rho^+$ and many others), in which only two tracks were reconstructed, resulting in the typical "shoulder"-shape, that is suppressed around 5.1 GeV/c, because their contribution is limited to the $m_{\pi^+\pi^-} < m_{B^0_{(s)}}$ region for kinematic reasons.

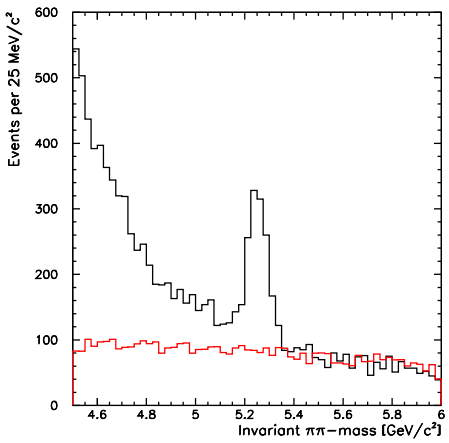

Figure 3.12: Invariant $\pi^+\pi^-$-mass distribution of the events passing the optimized selection of tab. 3.3 (black). The corresponding distribution for events with $L_{\rm T}(B) < -300$ μm is overlaid (red).

3.8.2 Signal composition

The observed width of the $B^0_{(s)} \to h^+h^{'-}$ signal is approximately 38 MeV/c^2, larger than what expected from the simulation for a single (e. g., $B^0 \to \pi^+\pi^-$ or $B^0 \to K^+\pi^-$) decay. This indicates that the $B^0_{(s)} \to h^+h^{'-}$ signal is the overlap of signals from different B^0 and B^0_s decay-modes with unknown proportions. Theoretical [24] and experimental [109] knowledge at the time of this analysis predicts sizable contributions from two known B^0 modes ($B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$) and two, still unobserved, B^0_s modes ($B^0_s \to K^+K^-$ and $B^0_s \to K^-\pi^+$), but, in principle, many other may contribute. In order to obtain the desired branching-fraction measurements, it is necessary to separate the contributions of the different components of the signal.

The measurement of the composition of the $B^0_{(s)} \to h^+h^{'-}$ signal is the scope of this analysis, and it is described in the next four chapters. In chap. 4, we see that a separation on an event-by-event basis is not feasible, and we introduce the statistical approach that combines information from PID and kinematics into an unbinned multivariate maximum likelihood (ML) fit to determine the contribution of each mode. It includes also choice of the kinematic observables and the achieved PID performance. In chap. 5 we detail the dedicated calibration of the dE/dx measurement, while the likelihood fit and its results are described in chap. 6 and chap. 7.

Chapter 4

Separation of individual $B^0_{(s)} \to h^+h'^-$ modes

The $B^0_{(s)} \to h^+h'^-$ signal reconstructed in the previous chapter is expected to contain several B^0 and B^0_s meson decay-modes in unknown proportions. We need to discriminate the individual contributions for extracting the desired measurements of branching fractions. The available resolution in mass and in particle identification is insufficient for an event-by-event separation of the decays. This chapter is devoted to the description of how the information from kinematics and particle identification was used to achieve a statistical discrimination among the individual modes.

4.1 Introduction

Figure 4.1 shows the invariant $\pi^+\pi^-$-mass distribution of the different expected decay-modes, resulting from the Monte Carlo simulation. Besides the established $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ modes, we included contributions from $B^0_s \to K^+K^-$, $B^0_s \to K^-\pi^+$, $B^0_s \to \pi^+\pi^-$, $B^0 \to K^+K^-$, $\Lambda^0_b \to \mathrm{p}\pi^-$, and $\Lambda^0_b \to \mathrm{p}K^-$ decays. We used branching fractions based on current experimental knowledge and theoretical predictions. We used the world-average measurements of the production fractions of B^0 mesons, B^0_s mesons, and b-baryons from fragmentation of a b-quark: $f_d = 0.397$, $f_s = 0.107$, and $f_{\mathrm{baryon}} = 0.107$ [109]. For the branching fractions of the established modes, we used the most recent results from the B-Factories: $\mathcal{B}(B^0 \to K^+\pi^-) = 18.5 \times 10^{-6}$ and $\mathcal{B}(B^0 \to \pi^+\pi^-) = 4.8 \times 10^{-6}$ [109]. For the other modes, we used typical theoretical predictions: $\mathcal{B}(B^0 \to K^+K^-) = 0.1 \times 10^{-6}$, $\mathcal{B}(B^0_s \to K^-\pi^+) = 8.0 \times 10^{-6}$, and $\mathcal{B}(B^0_s \to \pi^+\pi^-) = 0.1 \times 10^{-6}$ from Ref. [24], and $\mathcal{B}(\Lambda^0_b \to \mathrm{p}\pi^-) = 1.0 \times 10^{-6}$ and $\mathcal{B}(\Lambda^0_b \to \mathrm{p}K^-) = 1.5 \times 10^{-6}$ from Ref. [115].

Despite the excellent CDF II mass-resolution, the different decay-modes are too closely spaced in mass to be resolved;[1] they appear overlapping in a single peak, broader than the expected mass resolution for an individual decay. In addition, PID resolution is insufficient for an event-by-event

[1] See, for instance, some observed mass-widths: $\sigma_m \approx 14$ MeV/c^2 in $J/\psi \to \mu^+\mu^-$ decays, or $\sigma_m \approx 9$ MeV/c^2 in $D^0 \to h^+h'^-$ decays.

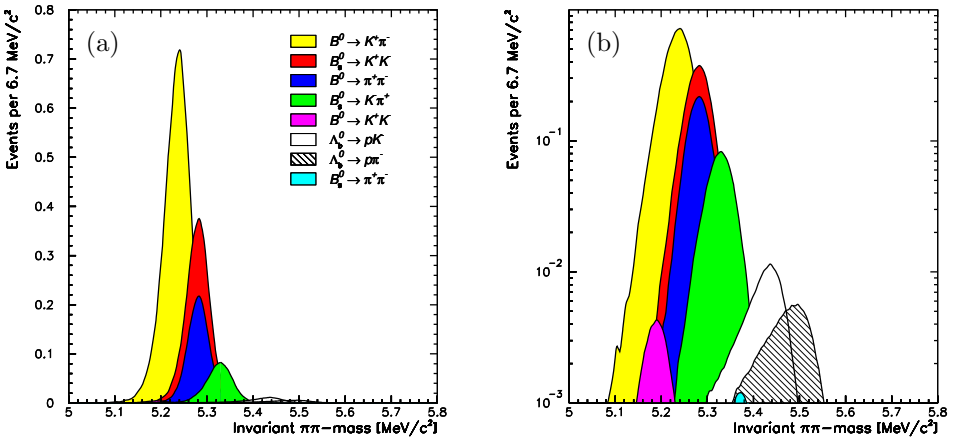

Figure 4.1: Invariant $\pi^+\pi^-$-mass distribution of the simulated decay-modes contributing to the $B^0_{(s)} \to h^+h^{'-}$ signal.

identification of particles in the final states. Only the momenta of charged particles are measured, while their masses need to be arbitrarily assigned when the invariant mass of the decay is computed. Unfortunately, whatever mass assignment we choose (a single one for all modes), the invariant-mass distributions of modes with mis-assigned masses are inevitably broadened: even with an infinitely precise mass resolution, their reconstructed invariant-mass would vary as a function of the momenta of the outgoing particles. For the distribution of fig. 4.1 we chose the charged pion mass for both outgoing particles. The mis-assigned invariant $\pi^+\pi^-$-mass of the $B^0 \to K^+\pi^-$ mode peaks at a value about 45 MeV/c^2 lower than the nominal B^0 meson mass, while the invariant $\pi^+\pi^-$-mass of the $B^0_s \to K^-\pi^+$ mode peaks at a value some 45 MeV/c^2 higher, although 45 MeV/c^2 lower than the nominal B^0_s meson mass. Any contribution from mis-reconstructed $B^0_s \to K^+K^-$ decays centers at the B^0 meson mass, because the B^0-B^0_s mass difference approximately compensates the effect of mis-identifying both kaons. While the mass r.m.s. width is approximately 25 MeV/c^2 for the properly reconstructed $B^0 \to \pi^+\pi^-$ mode, the widths of other modes appear to be larger (about 30 MeV/c^2), as a consequence of incorrect mass assignments.

However, even though an event-by-event separation of the individual contribution of each involved process is beyond CDF possibilities, a statistical separation can be exploited for extracting the desired measurements of decay rates. This was based on the combination of kinematic information of the decays with final-states PID, and required a dedicated calibration of the PID response to exploit the maximum separation.

4.2 Kinematic separation

To discriminate among decay modes, we exploit kinematic differences. These differences are small, since the $\approx$ 90 MeV/c^2 difference between B^0_s and B^0 masses, and the $\approx$ 350 MeV/c^2 difference

between kaon and pion masses are small compared with the typical energy available for each outgoing particle in the decay rest-frame ($E \approx 2.5$ GeV).

Evaluating, for each event, the different invariant masses resulting from all possible mass assignments to the particle pair (i. e., $\pi^+\pi^-$, $K^+\pi^-$, π^+K^-, and K^+K^-) induces unavoidable and large correlations between the likelihood terms, with complications for their description and use in the fitting procedure. In addition, this choice would considerably increase the number of observables needed in the likelihood function. If contributions from other modes (e. g., $\Lambda_b^0 \to \mathrm{p}\pi^-$) are considered, additional observables add up to the set of fit observables.

On the other hand, choosing a single, privileged mass-assignment capable of discriminating the modes better than other assignments is not convenient. The resulting separation power remains limited whatever assignment is chosen, because not all the available kinematic information is efficiently exploited.

We defined two variables that summarize the information carried by the values of invariant-mass resulting from all possible mass assignments to the outgoing particles. In a decay of a particle into two bodies of momenta $\vec{p}_1$ and $\vec{p}_2$ and masses $\overline{m}_1$ and $\overline{m}_2$, the invariant mass of the decaying particle satisfies the following relation:

$$m^2_{\overline{m}_1\overline{m}_2} = \left(\sqrt{\overline{m}_1^2 + p_1^2} + \sqrt{\overline{m}_2^2 + p_2^2}\right)^2 - (\vec{p}_1 + \vec{p}_2)^2. \tag{4.1}$$

Similarly, the invariant mass of the pair resulting from a different mass assignment to the outgoing particles is

$$m^2_{m_1m_2} = \left(\sqrt{m_1^2 + p_1^2} + \sqrt{m_2^2 + p_2^2}\right)^2 - (\vec{p}_1 + \vec{p}_2)^2, \tag{4.2}$$

where the incorrect mass m_1 (m_2) is assigned to the particle with momentum $\vec{p}_1$ ($\vec{p}_2$) with, in general, $m_1 \neq m_2 \neq \overline{m}_1 \neq \overline{m}_2$. The $B^0_{(s)} \to h^+h'^-$ data are acquired with a trigger requirement of particles with transverse momenta larger than 2 GeV/c. In this regime, the limit of relativistic decay-products is sufficiently accurate, $m^2_{1,2}, \overline{m}^2_{1,2} \ll p^2_{1,2}$, thus the difference between the mis-reconstructed and the true mass of the pair can be expanded in the following Taylor series up to $\mathcal{O}(m^2/p^2)$:

$$m^2_{m_1m_2} - m^2_{\overline{m}_1\overline{m}_2} \approx \left(1 + \frac{p_1}{p_2}\right)\left(m_2^2 - \overline{m}_2^2\right) + \left(1 + \frac{p_2}{p_1}\right)\left(m_1^2 - \overline{m}_1^2\right). \tag{4.3}$$

The relation (4.3) allows writing the invariant mass of the decay, $m^2_{m_1m_2}$, for any mass assignment to the outgoing particles (m_1, m_2), as a function of just two, loosely correlated variables: an invariant mass $m^2_{\overline{m}_1\overline{m}_2}$ obtained from a single arbitrary choice of mass assignment, and the ratio of momenta p_1/p_2 (momentum imbalance). After choosing a single mass-assignment for all events, thus obtaining the observable $m^2_{m_1m_2}$, of which the p.d.f. is function, the likelihood term corresponding to a given decay-mode is easily written in terms of the nominal B^0 or B^0_s mass (i. e., the value of $m_{\overline{m}_1\overline{m}_2}$ obtained with the correct mass assignment for the given mode), and of the momentum imbalance. Our choice of the charged-pion mass (m_π) for the two particles yields

$$m^2_{\pi^+\pi^-} \approx m^2_{\overline{m}_1\overline{m}_2} + \left(1 + \frac{p_1}{p_2}\right)\left(m_\pi^2 - \overline{m}_2^2\right) + \left(1 + \frac{p_2}{p_1}\right)\left(m_\pi^2 - \overline{m}_1^2\right). \tag{4.4}$$

For each event (i. e., for each set of observed $m_{m_1m_2} = m_{\pi^+\pi^-}$ and p_1/p_2), eq. (4.4) takes on different forms, depending on the mode one is referring to. For the $B^0 \to K^+\pi^-$ mode, for instance, $\overline{m}_1 = m_{K^+}$, $\overline{m}_2 = m_{\pi^-}$ or viceversa by construction, hence $m^2_{\overline{m}_1\overline{m}_2} = m_{B^0}$.

The advantage of this approach consists in the fact that all kinematic information is summarized in just two, loosely correlated, variables, while a larger number of strongly-correlated variables would have been needed if all mass assignments had been used in the likelihood.

Charge-kinematics flavor tagging

By combining charge information with kinematic information, one can gain discrimination between $K^+\pi^-$ and $K^-\pi^+$ final states. In $K\pi$ decays the heavier kaon tends to carry higher momentum than the pion; hence, decays in which the positively (negatively) charged particle has higher momentum than the other one are preferably $K^+\pi^-$ ($K^-\pi^+$) events. If one neglects a $\mathcal{O}(‰)$ fraction of doubly-Cabibbo-suppressed decays, this provides some separation between $B^0_{(s)}$ and $\overline{B}^0_{(s)}$ decays in $K\pi$ final states, because $K^+\pi^-$ states are produced only by B^0 and $\overline{B}^0_s$ decays, while $K^-\pi^+$ states are produced only by $\overline{B}^0$ and B^0_s decays.

We labeled the outgoing particles according to the increasing magnitude of their momenta, index "1" labels the charge (q_1), mass (m_1), and momentum ($\vec{p}_1$) of the lower momentum particle in the decay, index "2" labels the corresponding quantities of the higher momentum particle. Then we defined a "signed momentum-imbalance" as

$$\alpha = \left(1 - \frac{p_1}{p_2}\right) \times q_1, \tag{4.5}$$

which takes values in the finite interval $[-1, 1]$. Equation (4.4) can be rewritten in terms of α as

$$m^2_{\pi^+\pi^-} \approx m^2_{\overline{m}_1\overline{m}_2} + (2 - |\alpha|)(m^2_\pi - \overline{m}^2_2) + \left(1 - \frac{1}{|\alpha| - 1}\right)(m^2_\pi - \overline{m}^2_1), \tag{4.6}$$

where $\overline{m}_1$ ($\overline{m}_2$) is the mass of the lower (higher) momentum particle.

Table 4.1 reports the analytic expressions for the average invariant $\pi^+\pi^-$-mass of the decay, as a function of the signed momentum imbalance, for each specific $B^0_{(s)} \to h^+h^{'-}$ mode. The formulas are obtained by solving eq. (4.6) for $m_{\pi^+\pi^-}$, and replacing all masses with the nominal values appropriate to each decay mode. For each mode, and for every value of α, the observed $\pi^+\pi^-$-invariant mass is distributed as a Gaussian function, owing to the detector resolution smearing of the natural width of the $B^0_{(s)}$ meson. The mean of the distribution is labeled as $\mathcal{M}(\alpha) \equiv E[m^2_{\pi^+\pi^-}(\alpha)]$, where $E[x]$ indicates the expectation value of the random variable x (see tab. 4.1).

Figure 4.2 shows a graphical representation of the relation between the invariant $\pi^+\pi^-$-mass and the signed momentum-imbalance, in which the analytical formula of the $B^0 \to K^+\pi^-$ mode in tab. 4.1 is smeared with the assumed 25 MeV/c^2 mass resolution, and the distribution in signed momentum-imbalance is assumed flat. The mean of the invariant-mass distribution, $\mathcal{M}(\alpha)$, shifts as a function of the signed momentum imbalance. The gain in information due to this mass-momentum correlation over the simpler invariant mass information (represented by the projection of the plot in fig. 4.2 onto the $m_{\pi^+\pi^-}$ axis) comes from the differences among shifting patterns in different modes.

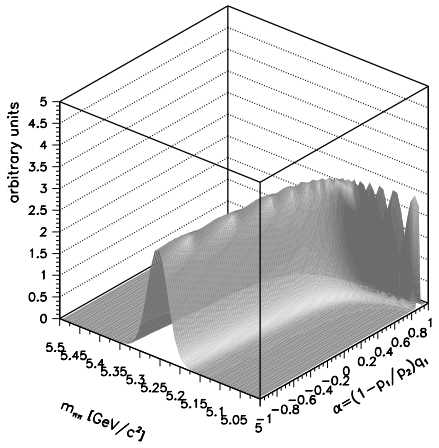

Figure 4.2: Graphical representation of the two-dimensional distribution of invariant $\pi^+\pi^-$-mass as a function of the signed momentum-imbalance for the $B^0 \to K^+\pi^-$ mode. The invariant mass distribution is assumed to be Gaussian with constant 25 MeV/c^2 standard deviation. The α-distribution is assumed to be a constant.

mode	$\mathcal{M}^2(\alpha) = m^2_{\pi^+\pi^-}(\alpha < 0)$	$\mathcal{M}^2(\alpha) = m^2_{\pi^+\pi^-}(\alpha > 0)$
$B^0, \overline{B}^0 \to \pi^+\pi^-$	$m^2_{B^0}$	$m^2_{B^0}$
$B^0 \to K^+\pi^-$	$m^2_{B^0} + (2+\alpha)(m^2_\pi - m^2_K)$	$m^2_{B^0} + \left(1 + \frac{1}{1-\alpha}\right)(m^2_\pi - m^2_K)$
$\overline{B}^0 \to K^-\pi^+$	$m^2_{B^0} + \left(1 + \frac{1}{1+\alpha}\right)(m^2_\pi - m^2_K)$	$m^2_{B^0} + (2-\alpha)(m^2_\pi - m^2_K)$
$B^0, \overline{B}^0 \to K^+K^-$	$m^2_{B^0} + \left(3 + \alpha + \frac{1}{1+\alpha}\right)(m^2_\pi - m^2_K)$	$m^2_{B^0} + \left(3 - \alpha + \frac{1}{1-\alpha}\right)(m^2_\pi - m^2_K)$
$B^0_s \to K^-\pi^+$	$m^2_{B^0_s} + \left(1 + \frac{1}{1+\alpha}\right)(m^2_\pi - m^2_K)$	$m^2_{B^0_s} + (2-\alpha)(m^2_\pi - m^2_K)$
$\overline{B}^0_s \to K^+\pi^-$	$m^2_{B^0_s} + (2+\alpha)(m^2_\pi - m^2_K)$	$m^2_{B^0_s} + \left(1 + \frac{1}{1-\alpha}\right)(m^2_\pi - m^2_K)$
$B^0_s, \overline{B}^0_s \to K^+K^-$	$m^2_{B^0_s} + \left(3 + \alpha + \frac{1}{1+\alpha}\right)(m^2_\pi - m^2_K)$	$m^2_{B^0_s} + \left(3 - \alpha + \frac{1}{1-\alpha}\right)(m^2_\pi - m^2_K)$
$B^0_s, \overline{B}^0_s \to \pi^+\pi^-$	$m^2_{B^0_s}$	$m^2_{B^0_s}$

Table 4.1: Analytic expressions of the squared invariant-mass of the $B^0_{(s)}$ meson candidate with pion mass assignment to both particles ($m_{\pi^+\pi^-}$) as a function of the signed momentum-imbalance (α) for each $B^0_{(s)} \to h^+h^{'-}$ mode. Left column: values for $\alpha < 0$, i. e., the negatively-charged particle carries smaller momentum. Right column: values for $\alpha > 0$, i. e., the positively-charged particle carries larger momentum.

The distributions of invariant $\pi^+\pi^-$-mass as a function of the signed momentum-imbalance for simulated $B^0_{(s)} \to h^+h^{'-}$ decays are shown in fig. 4.3 (see sec. 3.6 for details on the simulated samples). In spite of the smearing effect of the mass resolution, different trends for the different modes are visible. Differences between $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ decays, and between $B^0_s \to K^-\pi^+$ and $\overline{B}^0_s \to K^+\pi^-$ decays are also evident. The latter are used to measure the CP-violating decay-rate

asymmetry in these modes. The shape in the $B^0 \to \pi^+\pi^-$ mode is obviously a straight line centered at the B^0 meson mass, since the chosen mass assignment is correct for this mode. Small, if any, kinematic separation is expected between $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ modes, because the $B^0_s \to K^+K^-$ curve is also approximately constant, and overlaps the $B^0 \to \pi^+\pi^-$curve. However, an increased separation between these two modes is provided by the PID information, since both particles in final states are different.

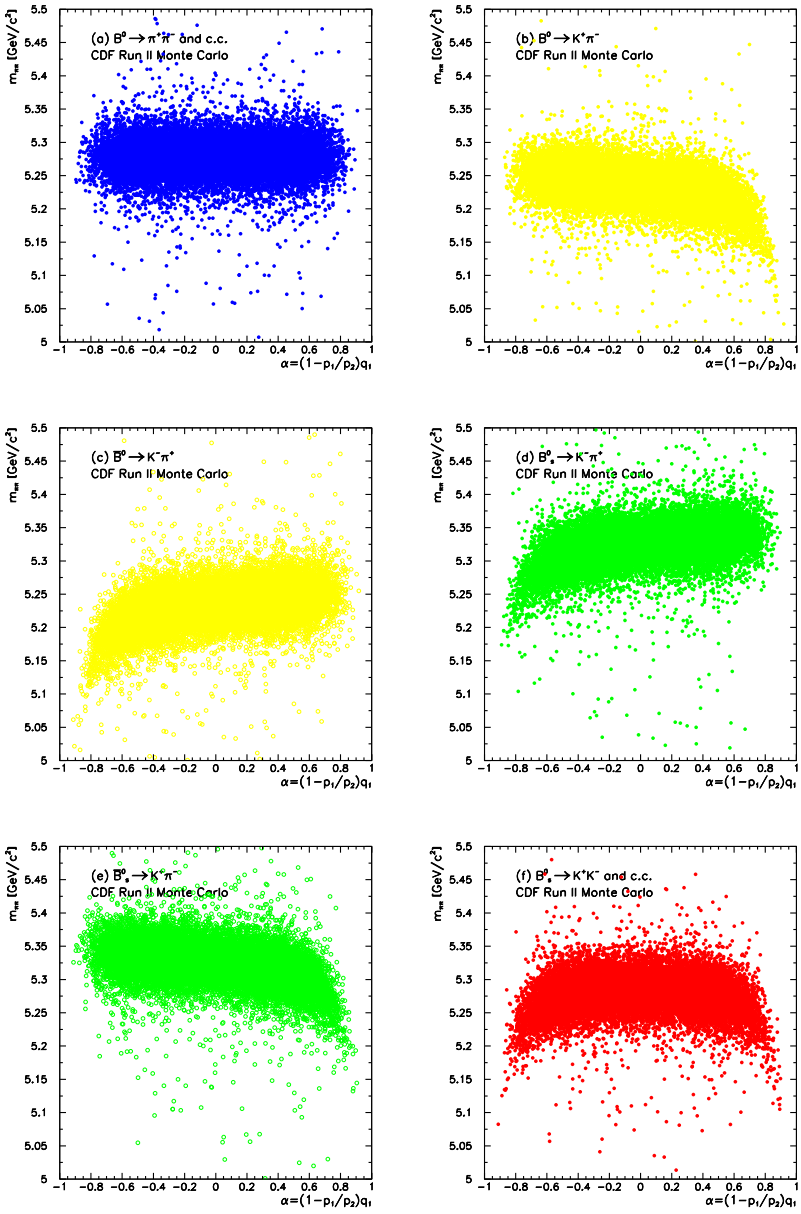

Figure 4.3: Distribution of the invariant $\pi^+\pi^-$-mass as a function of the signed momentum imbalance for the following simulated modes: (a) $B^0 \to \pi^+\pi^-$ and $\overline{B}^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $\overline{B}^0 \to K^-\pi^+$, (d) $B_s^0 \to K^-\pi^+$, (e) $\overline{B}_s^0 \to K^+\pi^-$, (f) $B_s^0 \to K^+K^-$ and $\overline{B}_s^0 \to K^+K^-$.

The same features are more evident in fig. 4.4, in which the distributions of fig. 4.3 are $m_{\pi^+\pi^-}$-averaged within each bin (profile plots) to visually remove the effect of mass-resolution smearing. The precise correspondence between the analytical functions of tab. 4.1 (solid, black lines overlaid)

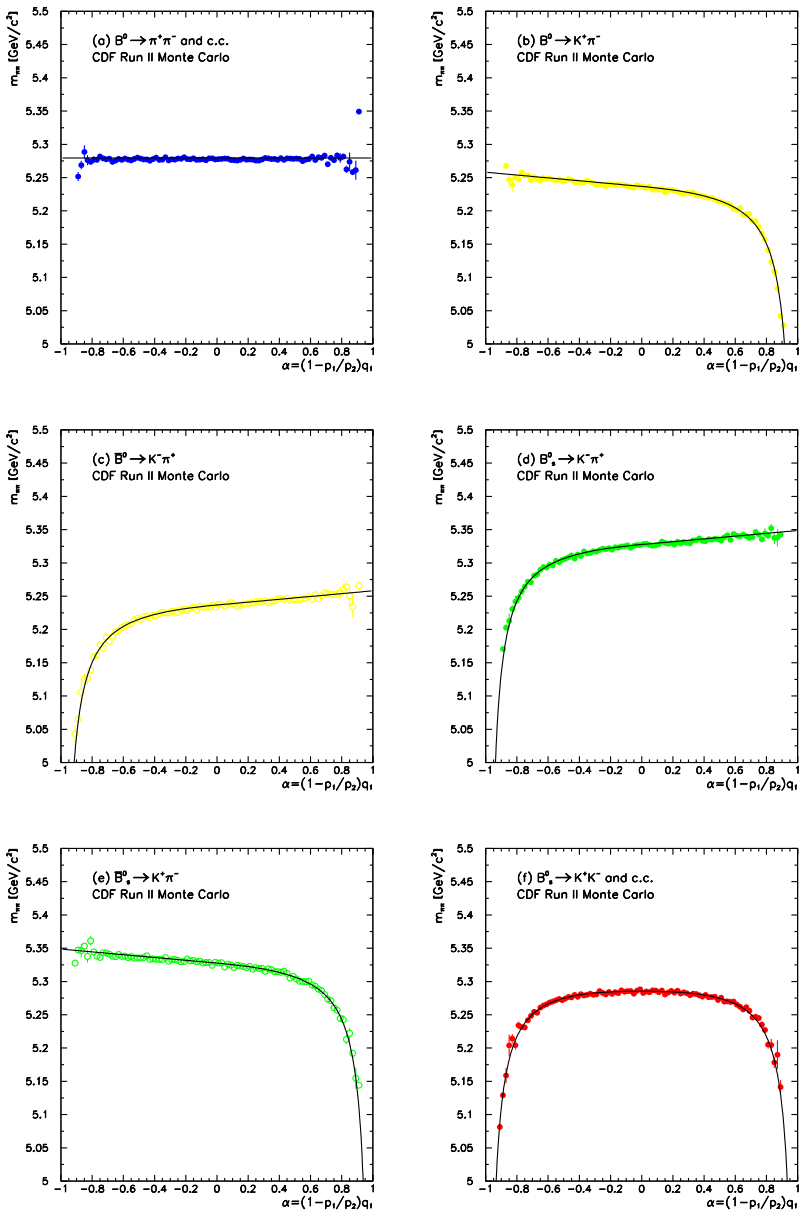

Figure 4.4: Profile plots of the invariant $\pi^+\pi^-$-mass as a function of signed momentum imbalance for the following simulated modes: (a) $B^0 \to \pi^+\pi^-$ and $\overline{B}^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $\overline{B}^0 \to K^-\pi^+$, (d) $B^0_s \to K^-\pi^+$, (e) $\overline{B}^0_s \to K^+\pi^-$, (f) $B^0_s \to K^+K^-$ and $\overline{B}^0_s \to K^+K^-$. The corresponding analytical expressions (from tab. 4.1) are overlaid (black, solid line).

and the simulated data indicate that the approximation used to extract eq. (4.3) is very accurate in the kinematic range of interest. We effectively summarized the kinematic information in just two

variables, with virtually no loss of information. A general feature of these plots is that slopes are enhanced at the boundaries of the α domain ($|\alpha| \approx 1$), suggesting that the kinematic separation is more effective in those decays where the momenta of the final particles are strongly unbalanced.

4.3 Charged particle identification

Identification of hadrons in CDF is difficult, since the detector was mostly designed for high-p_{T} physics measurements [116] and not optimized for these capabilities. The TOF is the only subdetector entirely devoted to this function, but its performances are marginal for particles with momenta greater than 2.0 GeV/c. Similarly, specific ionization from the silicon tracker is of little help, because its identification power is only effective for particles with $p_{\mathrm{T}} \lesssim 800$ MeV/c.[2]

Conversely, for charged particles of $p_{\mathrm{T}} \gtrsim 2$ GeV/c, the identification from the rate of energy loss through ionization (dE/dx) in the gas that fills the active volume of the drift chamber is reasonably effective. To a good approximation, the most probable dE/dx value for a charged particle is a function of its velocity; thus, if the momentum of the particle is measured, the mass can also be determined. In the COT, the signal induced on each sense-wire depends on the amount of ionization charge produced by the passage of the charged particle near the wire. It is measured in nanoseconds because it is encoded as the digital pulse-width between the leading and the trailing-edge times of the hit. Multiple samplings along the trajectory of the charged particle allow a more reliable estimation of dE/dx, which has usually a broad distribution (see appendix A). The COT samples a maximum of 96 dE/dx measurements per track, from which a 80% truncated mean is calculated to avoid the adverse effect of long positive tails in the estimation of the average dE/dx. A more detailed description of the dE/dx measurement in the COT is given in appendix A and in references therein.

Off-line dE/dx information could not be used directly in our analysis. In the first two years of data-taking, an unexpected wire-aging was observed, and the effect was severe enough to significantly degrade the performance of the chamber. Gain losses of up to 50% were observed as a function of time and location in the chamber. The aging was attributed to polymer build-up in the avalanche, which grew in a strong radiation environment and deposited on sense wires. The importance of the effect depended on local temperature and gas-flow patterns. The wire coating was identified as being mostly composed of carbon, hydrogen and a small amount of oxygen. Eventually, it was found that operating the chamber with about 100 ppm of O_2 added to the gas admixture (compared with a standard level of less than 12 ppm), reversed two years of gain loss in less than ten days [117].

Although the data collected when the COT performance was maximally compromised were excluded from physics analyses, smaller gain variations, due to the first, undetected effects of degradation, were observed throughout all data. Therefore, an accurate calibration of the uniformity of the dE/dx response in time and over the chamber volume was required. This is particularly important for analyses like ours, that rely on statistical separation, where even modest gain-variations may significantly reduce the available separation power.

The calibration of the dE/dx has been a crucial ingredient for the present measurement. This

[2] A separation equivalent to the one between two equal Gaussian distributions spaced by one standard deviation apart is obtained for kaons and pions at this momentum.

analysis was the first one using dE/dx in Run II. After our calibration, several other CDF measurements took advantage from the improved dE/dx performance [118, 119]. The details of the calibration procedure are described in the next chapter, which can be skipped at a first reading, if one prefers to continue directly toward the discussion of the measurement. In this case, we anticipate the relevant results of the calibration as follows.

The goal of the calibration was to remove any spurious dE/dx dependence on local properties of the chamber or on features of the particles other than velocity. This had the effect of optimizing the PID performance in terms of *separation power* and *correlations* in the response.

Separation power – the capability of discriminating among different classes of charged particles is crucial because the larger the separation, the smaller the statistical uncertainty on measurements of sample composition. Before the calibration, the observed difference between dE/dx distributions of kaons and pions was equivalent to an approximately 1.10σ separation between two Gaussian distributions, σ being the r.m.s. width (see fig. 4.5(a)). The improved performance due to calibration provided a separation approximately equivalent to 1.4σ (see fig. 4.5(b)) nearly constant in the momentum range of interest. The observed 1.4σ separation is sufficient to statistically distinguish two samples obtaining an uncertainty only a factor 0.6 worse than the one achievable by means of an "ideal" PID, i. e., the one with PID-observables having non-overlapping distributions. The most relevant separation capability needed in our analysis is between $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ contributions, which are almost indistinguishable kinematically (see fig. 4.1). Since both final state particles are different, in this case we benefit from approximately $1.4 \times 1.4 \simeq 2\sigma$ separation, corresponding to 75% of the ideal separation.

In addition to the increase in separation (see sec. 5.2.2–5.2.4), we obtained the curves of the expected dE/dx for kaons and pions (see sec. 5.3), and we accurately modeled the dE/dx distributions for a proper inclusion in the fit of the sample composition (see sec. 5.8).

Correlations – Even more critical, although often overlooked, are the correlations between dE/dx measurements. Residual time-dependent gain variations cause non-zero correlations between observed dE/dx of particles in the same event, yielding biased estimates of the sample composition, if the dE/dx is applied to multiple particles per event.

As a result of our calibration, we halved the initial track-to-track correlation (see sec. 5.7) remaining with a residual 11% correlation, which was modeled for a proper inclusion in the fit (see sec. 5.8).

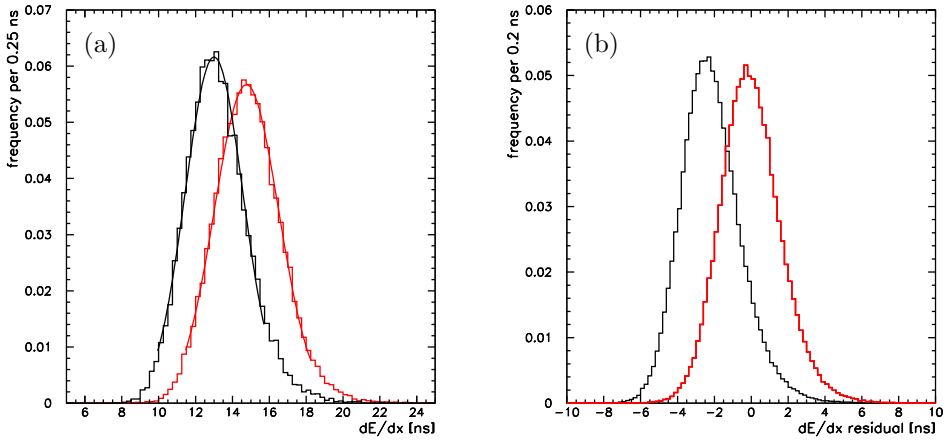

Figure 4.5: Distribution of the off-line dE/dx (before the calibration) for negatively-charged kaons (black histogram) and pions (red histogram) with $2.5 < p < 3.0$ GeV/c (a). Distribution of the calibrated dE/dx around the average pion response for negatively-charged kaons (black histogram) and pions (red histogram) (b). The curves of the average dE/dx response as a function of momentum were not available before the calibration. We therefore restricted the distributions in plot (a) to a narrow momentum-range to minimize the spread due to the dE/dx dependence on momenta.

Chapter 5

Calibration of the measurement of specific ionization in the drift chamber

This chapter describes the calibration of the specific ionization signal produced by the passage of charged particles in the drift chamber. Data were used to equalize the response as a function of the trajectory of the incident particle in the chamber and of time. Special emphasis is given to the study of correlations between the observed dE/dx of particles in the same event. The calibration provided a significantly improved separation between classes of particles and reduced correlations. The chapter also discusses the extraction of the modeling of the dE/dx distributions to be used in the fit of the composition of the $B^0_{(s)} \to h^+h^{'-}$ sample.

5.1 Calibration sample

Final states of $B^0_{(s)} \to h^+h^{'-}$ modes and of the dominant backgrounds contain charged kaons and pions. For calibration purposes, we therefore used a copious and pure sample of charged kaons and pions from D^0 decays in the chain

$$D^*(2010)^+ \to D^0\pi^+ \to [K^-\pi^+]\pi^+. \tag{5.1}$$

The $\mathcal{O}(5\ \mu\text{b})$ cross-section for production of central D^{*+} mesons with $p_{\text{T}} > 6$ GeV/c in $\text{p}\bar{\text{p}}$ collisions [120], combined with a 2.6% branching fraction for the process (5.1) [109], are efficiently exploited by the Displaced-Tracks Trigger, which collects extremely abundant signal samples. The strong D^{*+} decay unambiguously identifies the flavor of the D^0 meson, which is selected in its dominant, Cabibbo-favored decay $D^0 \to K^-\pi^+$.[1] Thus, the final states always contain two like-sign charged pions and one oppositely-charged kaon, allowing collection of high-purity samples of kaons and pions from the

[1] The $\mathcal{O}(10^{-3})$ contamination of the doubly-Cabibbo-suppressed mode $D^0 \to K^+\pi^-$ is negligible for our purposes [121].

D^0 decay. The pion originated from the D^{*+} decay is denoted as "soft pion", since it has a much lower transverse momentum than all other particles in the decay.

The sample was collected by the `B_CHARM` trigger (see sec. 3.3), a path that belongs, along with the `B_PIPI` path, to the Displaced-Tracks Trigger. A large fraction of trigger requirements is common to these two paths. Thus, most trigger-dependent effects on the dE/dx of $B^0_{(s)} \to h^+h^{'-}$ final states are automatically accounted for in the calibration. The sample contains the same physics-quality runs used for the $B^0_{(s)} \to h^+h^{'-}$ analysis (see sec. 2.6).

Following Ref. [106], the reconstruction of signal (see fig. 5.1) was solely based on tracking and on the information on the identity of D^0 decay-products provided by the charge of the soft pion. The quality requirements described in sec. 3.5.1 were applied to all tracks of the candidate decay. One

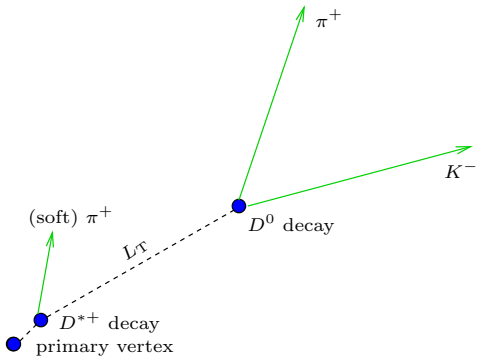

Figure 5.1: Schematic sketch of the $D^{*+} \to D^0\pi^+ \to [K^-\pi^+]\pi^+$ decay topology in the plane transverse to the proton beam direction.

$D^0 \to K^-\pi^+$ and one $\overline{D}^0 \to K^+\pi^-$ candidate were formed for each pair of oppositely-curved tracks found in the XFT fiducial region ($|\eta| < 1$). Further requirements on the product of impact parameters of tracks ($d_0(K) \times d_0(\pi) < 0$ cm^2), on candidate transverse momentum ($p_T(D^0) > 5.5$ GeV/c), on its transverse decay-length ($L_T(D^0) > 300$ μm), and on its impact parameter ($|d_0(D^0)| < 140$ μm) were applied to reject $\mathcal{O}(10\%)$ contributions [120] from non-prompt D^{*+} decays. Candidates with reconstructed invariant mass within 200 MeV/c^2 from the world average D^0 mass [109] were combined with a third charged particle with $p_T > 0.4$ GeV/c (soft pion) to form a $D^{*+} \to D^0\pi^+$ candidate. The charged pion mass is assigned to the like-sign pair of particles. The difference between the reconstructed D^{*+} and D^0 masses was required to be within 1.5 MeV/c^2 from the nominal value of 0.1454 GeV/c^2 to reduce backgrounds (combinations of true D^0 decays with random tracks, random three-track combinations that satisfy the selection requirements, etc.).

The selection results in about 3.3×10^5 signal decays, shown in the invariant $K\pi$-mass distribution of fig. 5.2(a). A D^0 signal with about 10 MeV/c^2 r.m.s. width is visible over a low background composed of random track combinations that accidentally meet the selection requirements (combinatorial component) and of mis-reconstructed D^0 decays. The combinatorial component populates uniformly the whole mass range. $D^0 \to K^+K^-$ decays ($\mathcal{B} \simeq 3.9 \times 10^{-3}$) in which a kaon is mis-assigned the pion mass, and $D^0 \to K^-\pi^+\pi^0$ ($\mathcal{B} \simeq 13\%$) decays in which a π^0 is not reconstructed, contaminate the lower mass-side of the signal. The higher mass-side contains, among others, $D^0 \to \pi^+\pi^-$ decays

($\mathcal{B} \simeq 1.4 \times 10^{-3}$), in which a pion is mis-assigned the kaon mass [106]. The final calibration sample is

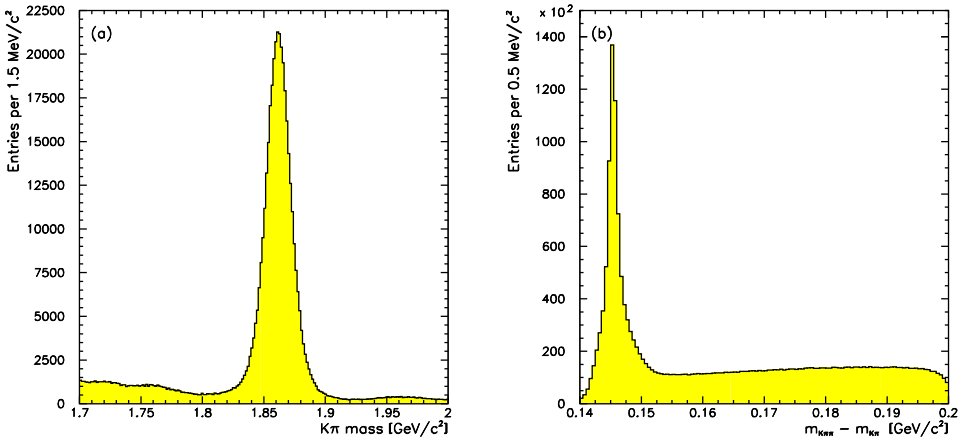

Figure 5.2: Invariant $K\pi$-mass distribution (a) and distribution of the mass difference between D^{*+} and D^0 candidates, $m_{D^{*+}} - m_{D^0}$ (b).

restricted to the D^0 candidates found within 40 MeV/c^2 from the world average D^0 mass [109], and resulted in 184,368 $D^{*-} \to \overline{D}^0\pi^- \to [K^+\pi^-]\pi^-$ and 178,258 $D^{*+} \to D^0\pi^+ \to [K^-\pi^+]\pi^+$ candidates. At this level we did not subtract the background, whose contamination is already reduced to $\mathcal{O}(5\%)$.

Kaons (or pions) with positive and negative charges were studied separately to properly account for the effects due to the asymmetry of the COT layout, which makes the efficiency of charge-collection dependent on the sign of the track curvature (see sec. 2.3.5). In the following, for brevity, we show the distributions for particles of only one charge, when the effects observed for the other one were similar.

5.2 Calibration procedure

We studied the observed dE/dx response as a function of a set of "macroscopic" parameters such as time, track azimuthal angle, pseudo-rapidity, etc. We did not check for anomalous effects associated to lower level COT structures such as super-layers, cells, or hit widths, because these were already accounted for in the standard CDF low-level corrections described in appendix A. Our calibration procedure consists of the following steps:

1. extensive search for possible dE/dx gain variations as functions of several global and track-related quantities;

2. identification of the subset of variables $\{x_i\}$ showing the most relevant effects;

3. removal of the dE/dx dependence on each x_j variable; we first corrected for the most severe effects, proceeding toward the smallest ones, while taking care of minimizing the cross-correlation between the variables;

4. extraction of the ionization curve, i. e., the function that describes the expected average dE/dx for a charged particle as a function of its Lorentz boost ($\beta\gamma$);

5. determination of the dE/dx resolution;

6. evaluation of the separation between kaons and pions;

7. extraction of the dE/dx correlations.

The crucial point is the correction (third step). In the simple case of a gain variation as a function of just one detector-related or track-related variable, e. g., x, the corrected dE/dx is simply obtained by fitting the dependence on that variable and subtracting the result from the uncorrected dE/dx. When the dE/dx depends on multiple quantities, e. g., x_1 and x_2, two cases are possible. If the dependencies are factorizable, i. e., $\frac{dE}{dx}(x_1, x_2) = f(x_1) \times g(x_2)$, with f and g being generic functions, then the corrected dE/dx is easily determined in two stages. First, the dependence on one variable, say x_2, is corrected in a subsample restricted to a limited domain in x_1, such that dE/dx is approximately constant as a function of x_1 in that domain and the subsample contains sufficient statistics. Then, the resulting correction is applied to the whole sample, in which, finally, the one-dimensional dependence on x_1 can be corrected. The worst case is when multiple dependencies can not be factorized, i. e., $\frac{dE}{dx}(x_1, x_2) = h(x_1, x_2) \neq f(x_1) \times g(x_2)$. In principle, a multivariate fit of dE/dx as a function of all quantities provides a multi-dimensional function that corrects simultaneously all effects. In practice, the size of the sample needed to obtain reliable corrections increases significantly with the number of effects to be corrected for, and the parameterization of the correction becomes more and more complex. Therefore, this approach is often combined with iterative corrections in a compromise between the conflicting needs for a simultaneous correction of correlated dependencies, and for samples with adequate statistics.

In the present case, the observed dE/dx was affected by multiple anomalous dependencies correlated in a non-factorizable way and complicated by unavoidable, intrinsic correlations between the geometric and kinematic quantities. For instance, a gain variation as a function of pseudo-rapidity is induced by the natural dependence of ionization on the Lorentz boost ($\beta\gamma$) combined with the $\beta\gamma - \eta$ correlation caused by the selection requirements and the kinematics of the decay.

We applied multiplicative, rather than additive, corrections because most dE/dx dependencies are expected to derive from geometric or gas-related effects of the chamber. They are therefore expected to influence the gain, rather than the offset, of the response.

5.2.1 Search for gain variations

An extensive scan of dE/dx distributions as a function of a number of variables evidenced four major anomalous effects (see fig. 5.3): a dependence on the time (run number) correlated with effects dependent on three track-related quantities, the azimuthal angle (φ_0), the pseudo-rapidity (η), and the multiplicity of associated COT-hits. Similar effects were observed also in control samples of muons from J/ψ decays and electrons from conversions. No further gain variations were found as a function of chamber pressure, instantaneous luminosity, or z_0 of tracks.

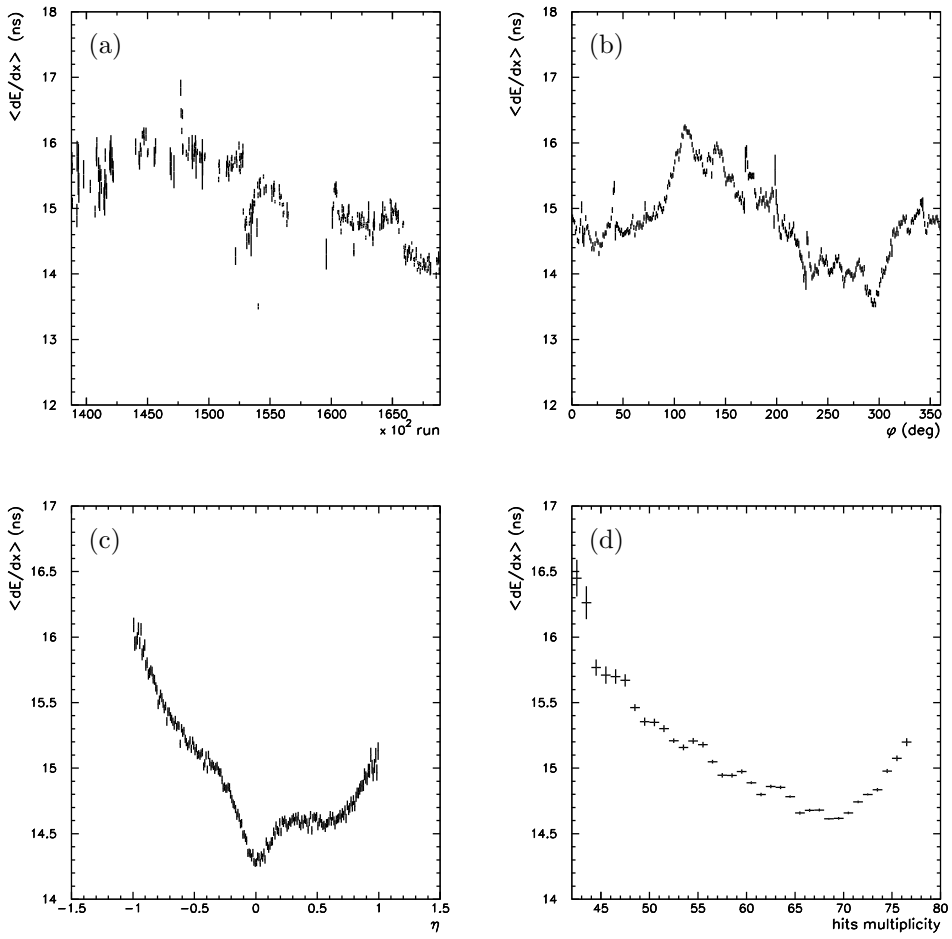

Figure 5.3: Profile plots of dE/dx as a function of run number (a), and of the azimuthal angle (b), pseudo-rapidity (c), and hit multiplicity (d) of the track. All distributions include contributions from positively and negatively-charged kaons and pions. The average ionization $\langle dE/dx \rangle$ is the dE/dx value averaged along the corresponding variable within each bin.

The φ_0 dependence is the most severe effect (see fig. 5.3). A "triangle"-shaped periodic pattern with ≈ 2.5 ns shifts between the measured dE/dx values is observed. Large dE/dx variations (≈ 2 ns) are observed also as a function of pseudo-rapidity. An uniform dE/dx degradation of about 1.5 ns, on average, is observed as a function of time, qualitatively confirming the hypothesis of a progressive gain-loss from coating of the wires (see sec. 4.3). An effect of about 1.5 ns is also observed as a function of the hit multiplicity.

The magnitude of these effects can be compared with the typical separation between the distributions of dE/dx for kaons and pions of fig. 5.4. A simple Gaussian χ^2-fit of the central part of each distribution yields a kaon response distributed around 13.0 ns with 1.6 ns r.m.s. spread, and a

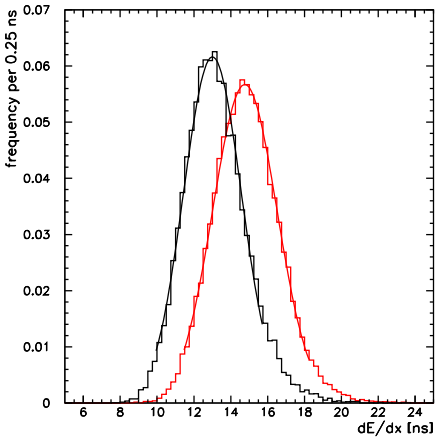

Figure 5.4: Distribution of the off-line dE/dx for negatively-charged kaons (black plot) and pions (red plot) with $2.5 < p < 3.0$ GeV/c.

pion response distributed around 14.8 ns with 1.7 ns r.m.s. spread. Despite the partial broadening due to the 500 MeV/c range in momentum, the separation is comparable in size to the observed gain variations.

Thus, the need for dE/dx calibration is two-fold: on one hand, this would increase the separation power; on the other one, the calibration is required to avoid biases (as a function of time, or of the position of the track in the chamber, for instance) in the determination of the composition of the $B^0_{(s)} \to h^+h^{'-}$ sample. This aspect is the most critical, and was the original and major motivation for carrying out the calibration before proceeding with our measurement.

The corrections were applied in the following order:

1. time-dependent correction for the azimuth dependence;
2. time-dependent correction for the pseudo-rapidity dependence;
3. correction for the hit multiplicity.

At each step, the correction was applied to the dE/dx value resulting from the previous step. We partitioned the data set in eleven independent subsamples, contiguous in time, to partially decouple geometric from time-dependent effects. The goal was to obtain samples with adequate statistics and limited dE/dx variations as a function of time. We studied the geometric dependencies separately within each subsample. The full set of run ranges is shown in fig. 5.5. The major known hardware and working-conditions changes were taken into account in partitioning the sample, including the implementation of hardware corrections for the gas pressure (since run 142559, see appendix A), the introduction of the XFT one-miss logic configuration (since run 152646, see sec. 3.2.4), and the introduction of the 4/5 SVT majority-logic (since run 164304, see sec. 3.2.4). An example of the impact of such changes in dE/dx-related quantities is given in fig. 5.5(b): since run 167830, a significant

fraction of Displaced-Tracks Trigger data were collected by the `B_CHARM_HIGHPT` (see sec. 3.3) trigger path, which applies higher thresholds on transverse momenta. A separate treatment is needed for these data to take into account the higher average boost of particles (see fig. 5.5(b)).

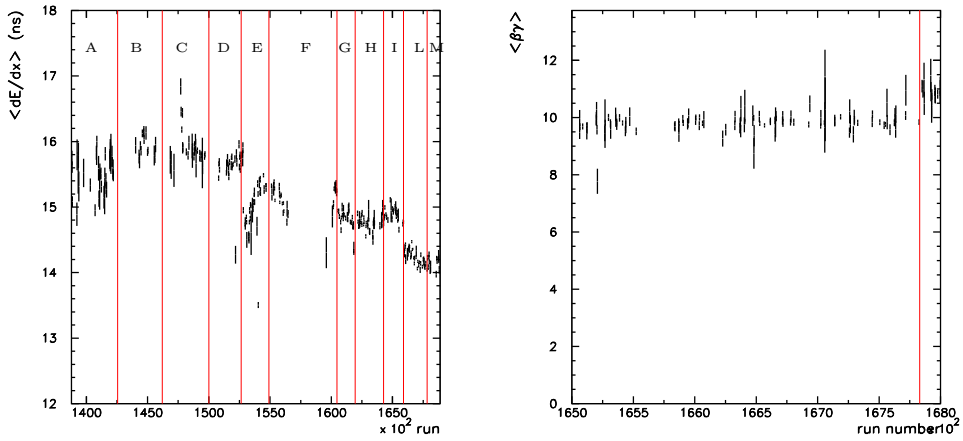

Figure 5.5: Profile plot of dE/dx as a function of run number for kaons and pions (left panel). Red lines indicate the separations between run ranges, which are labeled in alphabetical order. Profile plot of the Lorentz boost as a function of run number for kaons (right panel). The red line corresponds to run 167830.

5.2.2 Time-dependent correction for gain variation with azimuthal position

First we corrected for the azimuthal dependence, which is the largest effect observed. We found that this effect is independent of the pseudo-rapidity and multiplicity of associated COT hits of the track, but it degrades with time. Figure 5.6 depicts the average measured dE/dx as a function of φ_0 in 9 (out of 11) run ranges for positively-charged kaons and pions. Time and φ_0 dependencies of the dE/dx are not factorizable, since the magnitude of azimuthal dependence changes (worsens) with time. This is consistent with the hypothesis of a time-dependent degradation of the gain due to wire aging. The interpretation of the azimuthal pattern of the gain loss is complicated by the details of gas-flow patterns within the COT. However, fig. 5.6 shows that the region with the largest gain-reduction corresponds to the coldest portion of the chamber, i. e., the sector that lies directly under the silicon detector, which is kept at a temperature below 0°C. Convection currents in the gas could also play a role. This suggested a temperature-dependence of wire coating, consistently with what was observed in the H1 chamber at HERA which used a similar gas [122]. Besides aging, gravitational and electrostatic sagging of the wires were investigated as possible causes of this effect. The observed azimuthal pattern is inconsistent with this hypothesis.

The effect was removed with a time-dependent φ_0 correction, i. e., an independent set of φ_0

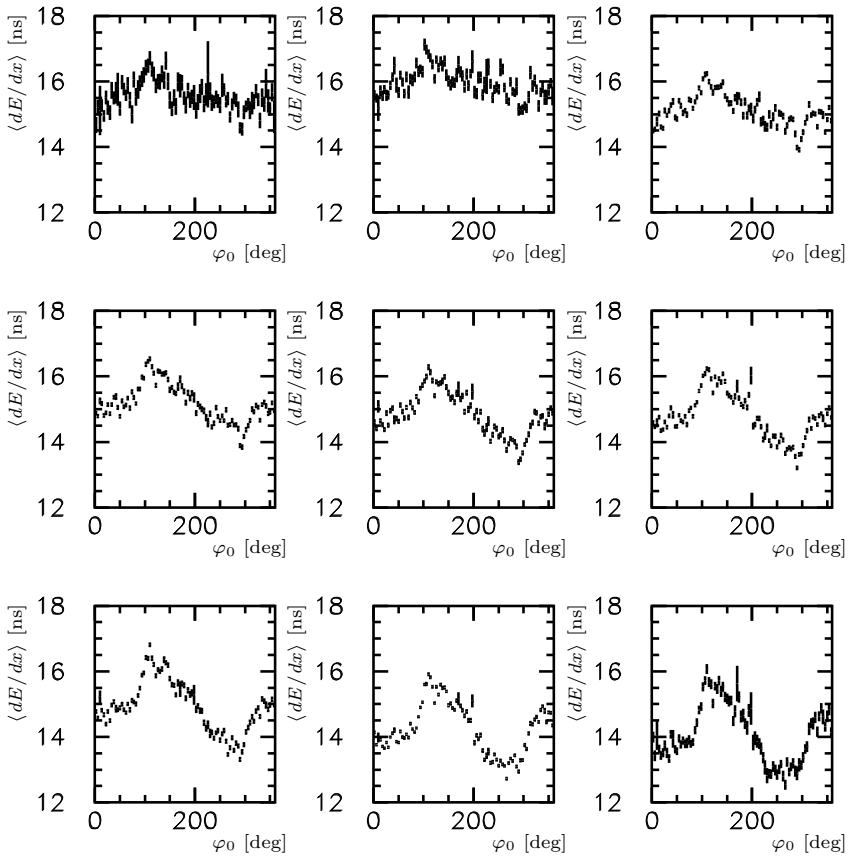

Figure 5.6: Profile plot of dE/dx as a function of track φ_0 for positively-charged kaons and pions. Each plot corresponds to a different period of data taking (see fig. 5.5).

corrections was determined for each of the run ranges of fig. 5.5. Since the sharp peaks of fig. 5.6 are poorly modeled with smooth fitting functions, we applied a per-bin multiplicative correction that rescales the average dE/dx in each φ_0-bin, and in each run range, to an arbitrarily chosen value of 15 ns. The result of the correction is reported in fig. 5.7.

5.2.3 Time-dependent correction for gain variation with axial position

The second largest effect is a function of pseudo-rapidity. The combined effect of kinematics, selection biases, and cell geometry cause a coupling of the pseudo-rapidity, hit multiplicity, and boost distributions of particles in our sample. In order to study the η dependence independently of such correlations, we restricted the calibration sample to the subset of particles with boost in the range $7 < \beta\gamma < 9$ and with $69 \leq \mathscr{N} \leq 73$ associated COT hits. Such $\beta\gamma$-range is small enough to make dE/dx variations

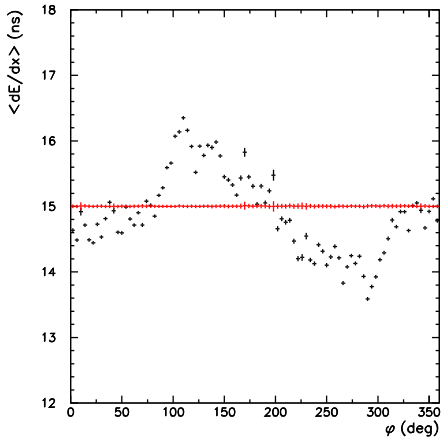

Figure 5.7: Profile plot of dE/dx as a function of track φ_0 for positively-charged kaons and pions. Black (red) points refer to the dE/dx evaluated before (after) the correction.

within the subsample negligible for our purpose, and it has the highest density of particles (kaons only) in our sample.[2] Similarly, the chosen $\mathscr{N}$-domain is the only one in which the dE/dx appears fairly constant, as shown in fig. 5.3(d). As for the azimuthal dependence, the gain degradation as a function of pseudo-rapidity is time-dependent in a non-factorizable way (see fig. 5.8). In earlier data, figs. 5.8(a)–5.8(c), the pattern appears symmetric between negative and positive pseudo-rapidities, as one would expect from a consequence of the sample-specific correlations between pseudo-rapidity and Lorentz boost. Later on, the degradation increases and the gain loss worsens toward the exhaust end of the chamber ($\eta > 0$ side). Such asymmetric pattern is consistent with the observed aging due to polymer build-up, since the net gas flow in the chamber is directed toward positive $\hat{z}$. Although the steep, "V"-shaped feature appearing around $|\eta| \lesssim 0.2$ was not further investigated, past Run I experience suggested a possible detector-induced gain saturation affecting signals from vertical tracks [123].

We applied a time-dependent, multiplicative η-correction by fitting to 11th-degree polynomials the observed patterns in each run range. This removed the dependence and adjusted the average dE/dx value to the conventional value of 15 ns. The correction determined from the subsample described above was then applied to the whole sample (see fig. 5.9).

5.2.4 Correction for gain variation with sampling multiplicity

The observed dE/dx also shows a non-trivial dependence on the number $\mathscr{N}$ of COT hits associated to the track (see fig. 5.10(a)). The dE/dx decreases in going from $\mathscr{N} = 42$ to $\mathscr{N} \approx 70$, after which value it shows an almost constant trend followed by a moderate increase. Multiple and inter-

[2] Owing to the 2 GeV/c trigger threshold on the transverse momentum, the minimum $\beta\gamma$ for pions in Displaced-Tracks Trigger data is $\beta\gamma = p/m \approx 2/0.139 \approx 14$.

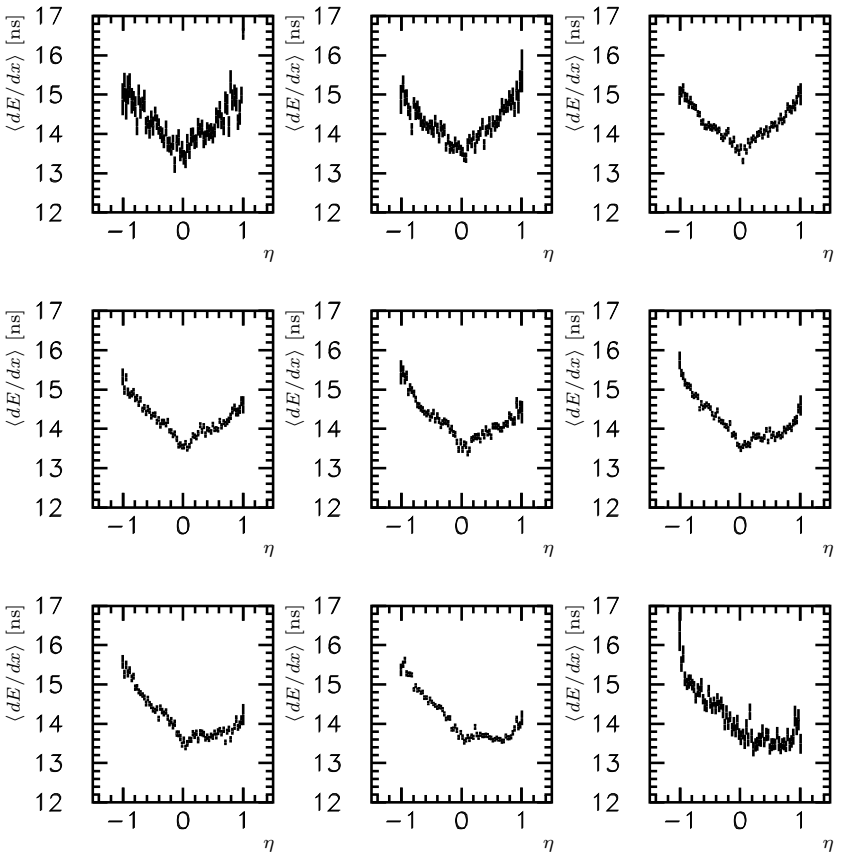

Figure 5.8: Profile plots of dE/dx (already corrected for the φ_0 dependence) as a function of track pseudo-rapidity in different run ranges for a subsample of positively-charged kaons and pions with specified Lorentz boost and number of associated COT hits (see text).

correlated factors affect hit-multiplicity distributions. Lorentz boost plays a role, for instance, because faster charged-particles produce larger ionization, whose effect on hit multiplicity is two-fold: on one hand a larger fraction of wires collects signals above discriminator threshold, on the other hand, large ionization causes larger pulse-widths, which have higher probability to be merged, during read-out, with subsequent hits in high occupancy events, hence decreasing hit multiplicities. The former contribution dominates, causing an average excess of two hits for pions with respect to kaons. An interpretation of the observed $\mathcal{N}$-dependence in terms of the above issues is not simple. However, a pattern of this kind is expected from the truncated-mean estimation of an asymmetric distribution, as in the case of the off-line dE/dx. We studied this estimator as a function of the sampling multiplicity (see fig. 5.10(b)) in an ensemble of pseudo-experiments based on positive Cauchy distributions whose mean values were estimated with an 80% truncation. An effect similar to the one observed in data is

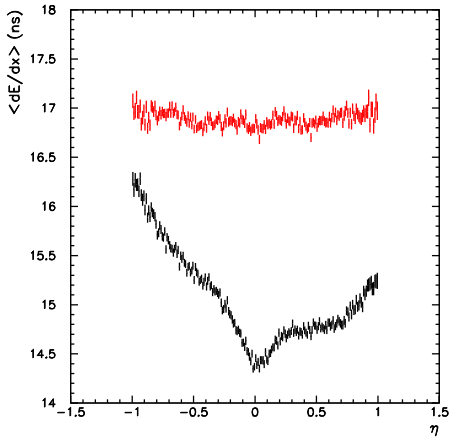

Figure 5.9: Average dE/dx response (already corrected for the φ_0 dependence) as a function of track pseudo-rapidity for positively-charged kaons and pions. Black (red) points refer to the dE/dx evaluated before (after) the correction.

evident, with comparable size (which depends on the skewness of the distribution).

The gain variation as a function of the number $\mathscr{N}$ of associated COT hits does not vary with time, allowing for a time-integrated correction. Lorentz-boost correlation was minimized by restricting the sample to the domain of particles (only kaons) with $7 < \beta\gamma < 8$. We corrected the effect on hit multiplicity by fitting it to a 4th-degree polynomial function that removed the slope, and set the average dE/dx to the conventional value of 15 ns. The correction is first obtained from the chosen $\beta\gamma$-restricted subsample, and it is subsequently extended to the whole sample (see fig. 5.10(a)). Although the effect is considerably reduced, a residual slope is still visible. This indicates that the $\beta\gamma$ and the $\mathscr{N}$ dependencies do not factorize exactly, as assumed in our correction procedure. The residual dependence is accounted for in the treatment of dE/dx correlations (see sec. 5.7).

5.3 Extraction of the expected ionization curve

Once the dominant effects were corrected, we extracted the curve of the expected dE/dx as a function of the Lorentz boost. We first divided the sample according to the Lorentz boost of the particles ($\beta\gamma$-bins); in each bin we computed the dE/dx sample-mean (denoted as "average dE/dx" or $\langle dE/dx \rangle$ in the following);[3] then we fit the average dE/dx as a function of the Lorentz boost to the following modified Bethe-Bloch curve (see appendix A):

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2}\left[c_1 \ln\left(\frac{\beta\gamma}{b+\beta\gamma}\right) + c_0\right] + a_1(\beta-1) + a_2(\beta-1)^2 + C, \tag{5.2}$$

[3]Sample mean was preferred to any other estimator because it minimizes the variance.

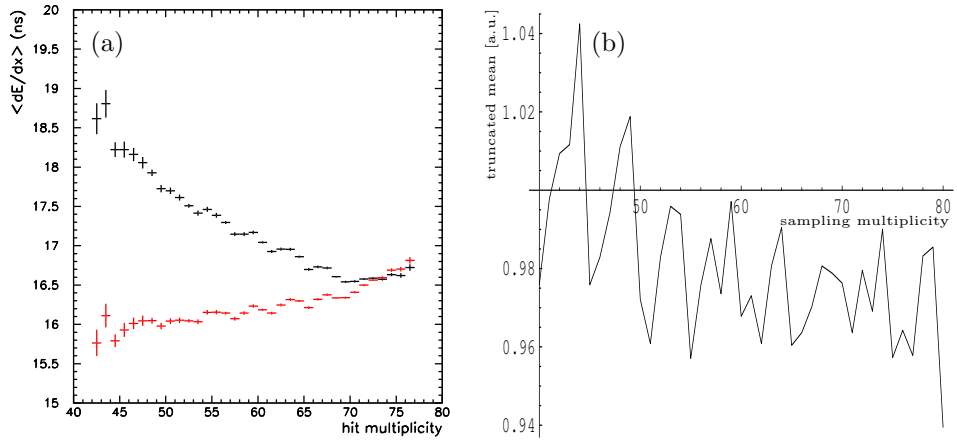

Figure 5.10: Profile plot of dE/dx (already corrected for the azimuthal and axial effects) as a function of hit multiplicity for positively-charged kaons and pions (a). Black (red) points refer to the dE/dx evaluated before (after) the correction. Average truncated mean of a set of Cauchy-distributed variables, as a function of the number of variables (b).

with a_i, b, c_j, and C free parameters in the fit. The resulting function is usually referred to as the "ionization curve".

Since contamination from background may lead to a sample-dependent curve and consequent biases in estimates of the sample composition, the purity of the sample is of paramount importance in extracting the curve. We both increased the sample purity in the region of the D^0 signal and applied a background subtraction to the dE/dx distributions. First, we defined a *signal plus background* sample, tightening the requirements of sec. 5.1: we used D^0 candidates with observed mass within 20 MeV/c^2 (compared with 40 MeV/c^2, used for calibration) from the world average [109]. Then we defined a *background* sample that contained D^0 candidates with reconstructed mass within 40 to 60 MeV/c^2 from the nominal one. Finally, we subtracted the $\langle dE/dx \rangle$ distribution of background events from the $\langle dE/dx \rangle$ distribution of signal plus background events. For each ith bin in $\beta\gamma$ we calculated the sideband-subtracted expected dE/dx as

$$\left\langle \frac{dE}{dx} \right\rangle_i = \frac{\langle S_i \rangle n_i - \langle B_i \rangle m_i}{n_i - m_i}, \tag{5.3}$$

where S_i is the dE/dx sample-mean (over the n_i entries in the bin) of particles in the signal-plus-background sample, and B_i is the sample-mean (over the m_i entries in the same bin) of particles in the background sample. In each bin, n_i includes a number m_i of expected background events. The expression for the standard deviation of the expected background-subtracted dE/dx is not trivial:

$$\begin{aligned} \sigma_{\langle dE/dx \rangle_i} &= \frac{1}{(n_i - m_i)^4} \times \left\{ n_i^4 \sigma_{S_i} + m_i^5 \sigma_{B_i}^2 + n_i m_i^2 \left[n_i \sigma_{S_i} - 2 m_i^2 \sigma_{B_i}^2 + (\langle S_i \rangle - \langle B_i \rangle)^2 \right] \right. \\ &+ \left. n + i^2 m_i \left[m_i^2 \sigma_{B_i}^2 - 2 n_i \sigma_{S_i} + (\langle S_i \rangle - \langle B_i \rangle)^2 \right] \right\}, \end{aligned} \tag{5.4}$$

where σ_{S_i} and σ_{B_i} are the uncertainties on the number of signal-plus-background and background events in the ith bin.

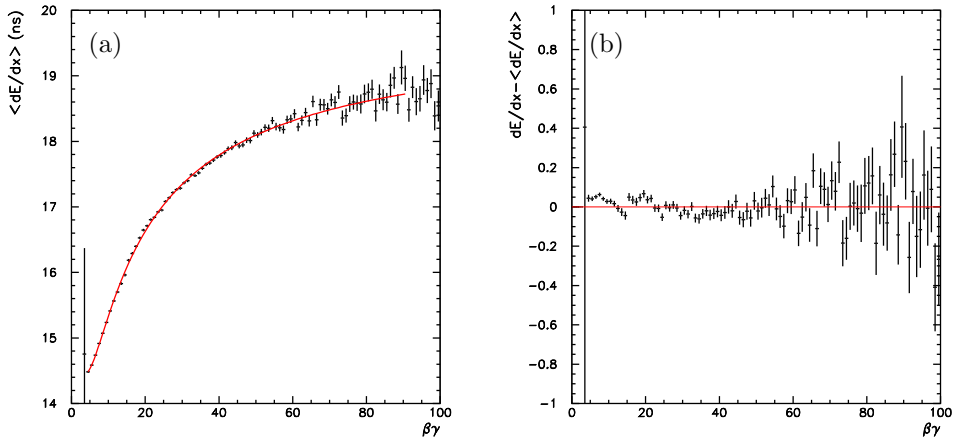

Figure 5.11: Average dE/dx (after all corrections) as a function of $\beta\gamma$ of positively-charged kaons and pions (a). A fit to the modified Bethe-Bloch curve (see eq. (5.2)) is overlaid in red. Average difference between data and fit function as a function of $\beta\gamma$ (b). The red line marks the reference zero-value.

Figure 5.11(a) and fig. 5.12(a) depict the average dE/dx after all corrections as a function of the Lorentz boost for kaons (in the range $\beta\gamma \gtrsim 4$) and pions (in the range $\beta\gamma \gtrsim 14$) with the result of a binned ML fit to the function of eq. (5.2) overlaid in red. The fit does not describe accurately the data (see fig. 5.11(b) and fig. 5.12(b)): positive and negative $\mathcal{O}(0.1$ ns) discrepancies, and the quality of matching between pion and kaon data at the pion turn-on ($\beta\gamma \gtrsim 14$) is poor. The corresponding plots for negatively-charged pions and kaons, figs. 5.12(a) and (b), exhibit similar features.

These effects suggest for us to use separate models for the expected dE/dx of pions and kaons, while keeping the original separation in charge, for an additional insight in the observed mis-match (see fig. 5.13). For a given boost, higher ionization is detected from positively-charged kaons than from pions while for negatively-charged particles the opposite behavior is observed. The different average ionization between kaons and pions, therefore, seems to be an effect depending on the track curvature (i. e., mass and charge), rather than just on the mass of particles. This is consistent with being caused by the asymmetry of the COT cells. A more detailed investigation would require a complete revision of the CDF low-level standard calibration (see appendix A). However, since the $\mathcal{O}(0.1$ ns) bias is small compared with 1.5 ns typical dE/dx resolutions, we chose the simpler solution of fitting separately each particle type and charge to extract particle-dependent expected dE/dx values.

Figure 5.14 shows the fit residuals (difference between data and fit functions). The residual for positively-charged particles is still not perfectly flat at zero, but the size of previous biases is significantly reduced and a smooth matching between kaons and pions is obtained.

In presence of a small systematic effect in curvature, we therefore chose to use particle-dependent curves for the expected dE/dx throughout the analysis. In the need for the maximum separation

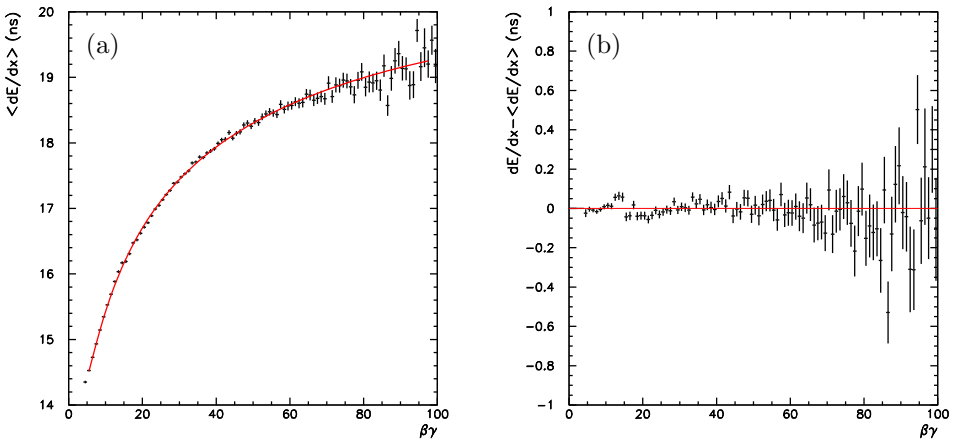

Figure 5.12: Average dE/dx (after all corrections) as a function of $\beta\gamma$ of negatively-charged kaons and pions (a). A fit to the modified Bethe-Bloch curve is overlaid in red. Average difference between data and fit function as a function of $\beta\gamma$ (b). The red line marks the reference zero-value.

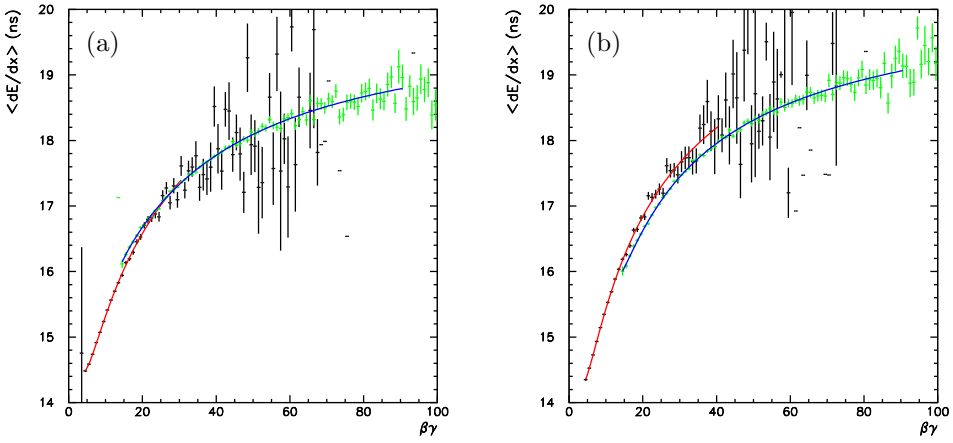

Figure 5.13: Average dE/dx (after all corrections) as a function of $\beta\gamma$ of positively (a) and negatively-charged (b) particles. In each plot, two distributions are superimposed, one corresponding to kaons (black points with fit function overlaid in red), the other one corresponding to pions (green points with fit function overlaid in blue).

power, this allows exploiting a more accurate parameterization of dE/dx.

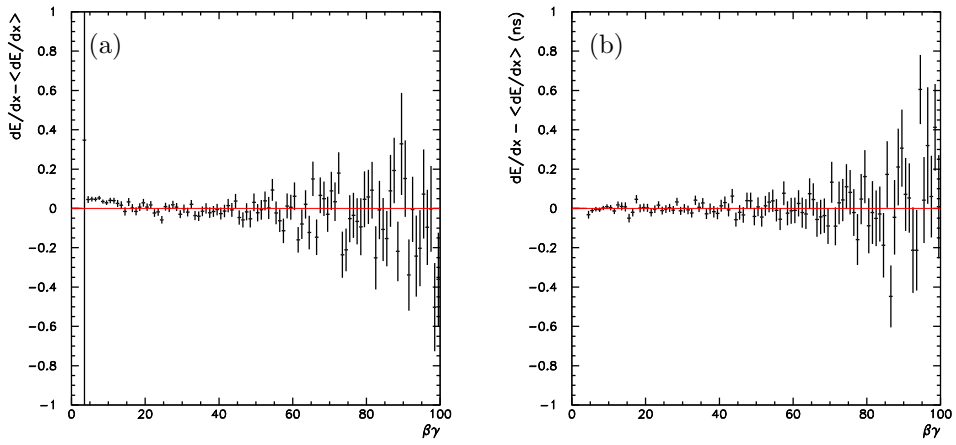

Figure 5.14: Average difference between data and fit function shown in fig. 5.13 as a function of $\beta\gamma$ of positively (a) and negatively-charged (b) kaons and pions.

5.4 Checking the corrections

We verified the improvements provided by the dE/dx tuning by studying the distributions of "dE/dx residuals". The dE/dx residual (in m_A mass hypothesis) of a charged particle with momentum p and observed specific energy-loss $dE/dx_{\rm obs}$, is defined as

$$\delta_A = \frac{dE}{dx}_{\rm obs} - \frac{dE}{dx}_A, \tag{5.5}$$

in which dE/dx_A is the expected dE/dx determined from the function eq. (5.2) evaluated at $\beta\gamma = p/m_A$.

The residual of an ideal dE/dx would show no dependencies. Comparing the residuals before and after tuning allows estimating the benefits and the reliability of the tuning (see figs. 5.15 and 5.16). The calibration greatly improved the uniformity and stability of dE/dx response: the ≈ 2 ns variation in gain as a function of time is reduced to a $\lesssim 1$ ns effect; the ≈ 2.5 ns azimuthal dependence was reduced down to the ≈ 0.5 ns level; The 1–1.5 ns axial dependence was reduced to $\lesssim 0.5$ ns effect as well as the effect on hit multiplicity. Similar features are present in the corresponding distributions for negatively-charged particles. However, the dependencies are not completely removed: second-order residual effects are still visible. Sections 5.7 and 6.2.1 discuss on how they have been taken into account in the analysis.

5.5 Resolution

Additional, and potentially useful, information for separation resides in the resolution of the dE/dx measurement, available on a per-track basis. The dE/dx resolution is defined as the r.m.s. width of a Gaussian distribution that fits the bulk of the distribution of dE/dx residual. In general, the observed

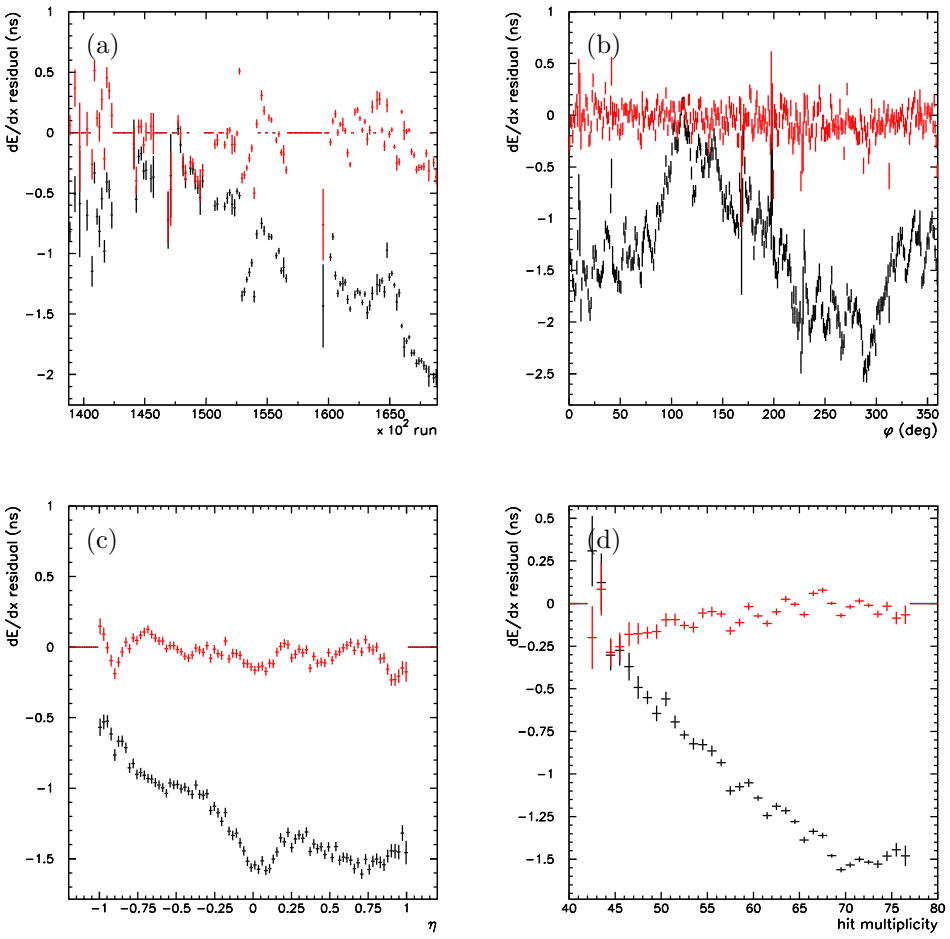

Figure 5.15: Distribution of residuals (with kaon hypothesis) for positively-charged kaons as a function of run number (a), and of azimuth (b), pseudo-rapidity (c), and hit-multiplicity (d) of the track. Black (red) points are obtained using the uncorrected (corrected) dE/dx.

deviations of resolution from its typical 1.5 ns value are modest, thus providing limited additional information. The dependence on hit multiplicity is the dominant effect involving the resolution (see fig. 5.17). It derives from the sampling process associated to the dE/dx measurement. The observed $\sigma_{dE/dx} \propto a - b \cdot \mathcal{N}$ dependence is different from the expected Poisson fluctuation $\sigma_{dE/dx} \propto 1/\sqrt{\mathcal{N}}$. This might be due to non-linearities from low-level gain corrections (which vary among super-layers) and from the truncated mean algorithm.

We chose not to use dE/dx resolution as a discriminating variable in the analysis, after checking that its contribution to the separation power is marginal.

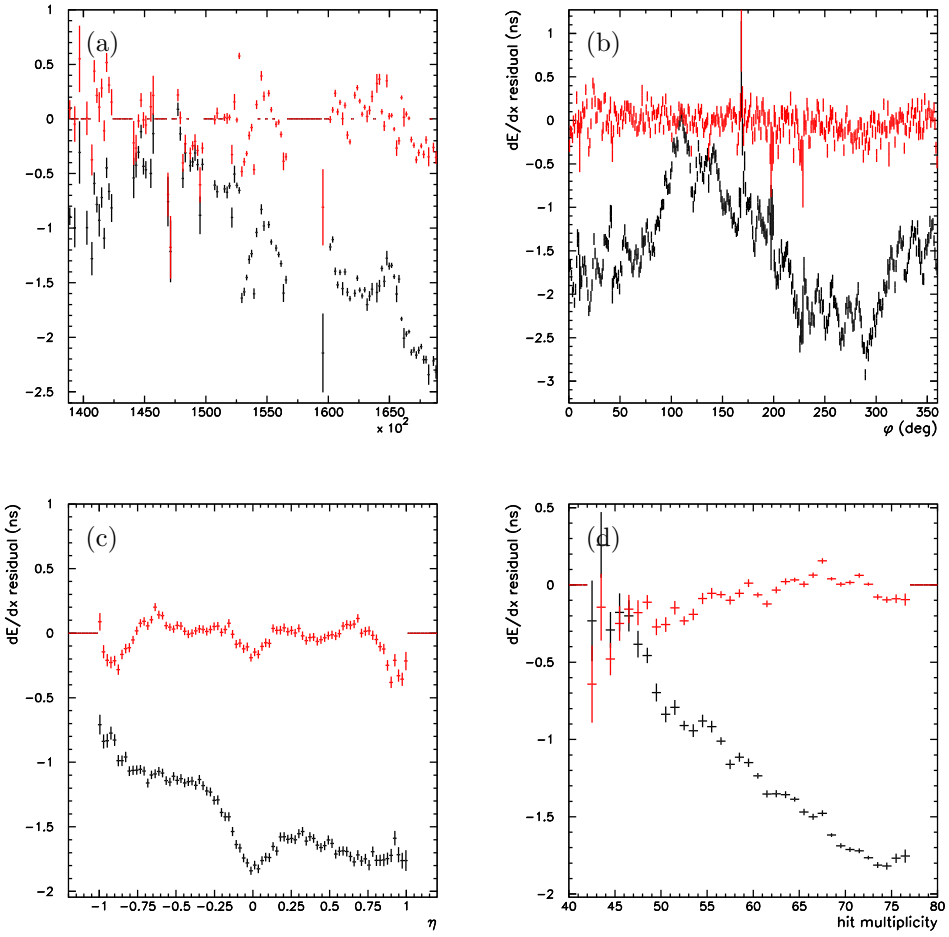

Figure 5.16: Distribution of residuals (with pion hypothesis) for positively-charged pions as a function of run number (a), and of azimuth (b), pseudo-rapidity (c), and hit multiplicity (d) of the track. Black (red) points are obtained using the uncorrected (corrected) dE/dx.

5.6 Separation

Any improvement in particle identification capability quantifies the impact of the dE/dx corrections. Given a PID-related observable, (dE/dx residual, for instance) the identification performance relies on the difference in the distributions of the chosen observable between the classes of events to be identified. Such difference is generally expressed in terms of a *separation* between those distributions.

The conventional way to quote the separation is to provide an estimation of the distance between the centers of the distributions in units of their standard deviations (σ). The reliability of this estimation degrades increasingly as the distributions deviate from the Gaussian shape. In presence of long tails or strongly asymmetric distributions, the separation estimated with this approach is not

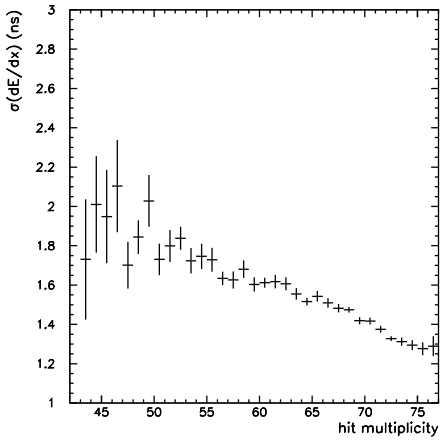

Figure 5.17: Distribution of dE/dx resolution as a function of the number of associated COT hits for positively-charged kaons and pions.

reliable. In addition, this choice has limited applicability: if the PID information is given in form of a multi-dimensional observable, no straightforward and unambiguous way to quote the separation is provided.

We therefore chose a different approach [124]. Suppose that one uses PID information in a sample of N events, of which a fraction f are pions and a fraction $1-f$ kaons (see fig. 5.18). An unambiguous estimator of the separation power is the precision $(1/\sigma_f)$ in estimating the fraction f, where σ_f is the statistical uncertainty on the fraction. The precision is bounded from above by the value

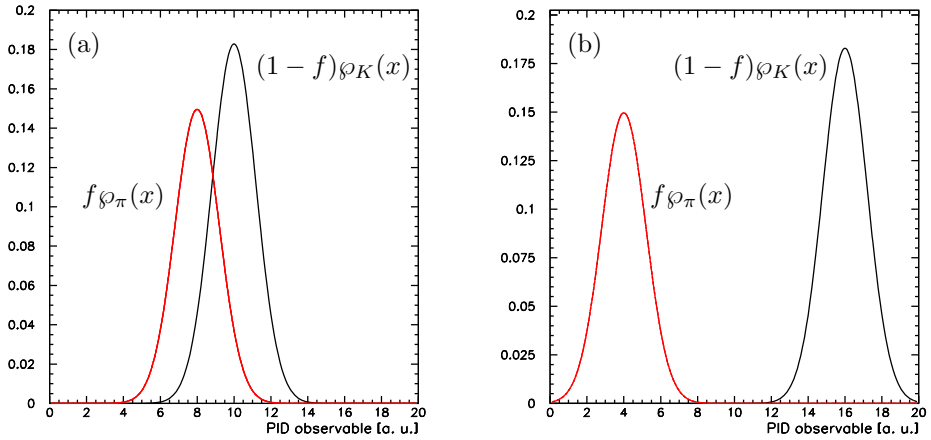

Figure 5.18: Illustration of the distribution of a simulated PID-observable (x, in the text) for pions (red curve) and kaons (black curve) in case of partial (a) and total (b) event separation.

$1/\sigma_f^{\text{best}} = \sqrt{N/f(1-f)}$ obtained in the ideal case of classes of events being totally separated, i. e., the distributions of the chosen PID observable having zero overlap (see fig. 5.18(b)). In this limiting case, the only uncertainty in assigning an event to one of the two classes comes from the Binomial fluctuation due to the finite sample-size.

We quote the separation between kaons and pions as the relative precision with respect to the ideal case $s = \sigma_f^{\text{best}}/\sigma_f$. The resolution σ_f can in principle be determined by repeating a maximum likelihood fit of the fraction f on a sufficient number of pseudo-experiments that simulate the experimental data, and evaluating the spread of results around the input values. We used, instead, the Minimum Variance Bound (see appendix B). Choosing a variable x to separate the classes of events, in the simple case of two classes, the MVB is written as

$$\sigma_f^2 = \frac{1}{N}\left[\int \frac{(\wp_\pi(x) - \wp_K(x))^2}{f\wp_\pi(x) + (1-f)\wp_K(x)}dx\right]^{-1}, \tag{5.6}$$

where $\wp_\pi(x)$ and $\wp_K(x)$ are the probability distributions of x, normalized to unit area, for pions and kaons respectively (see fig. 5.18).[4] Following our approach, the separation power of the variable x in the given sample is determined by evaluating

$$s = \sigma_f^{\text{best}}/\sigma_f = \sqrt{f(1-f)\int \frac{(\wp_\pi(x) - \wp_K(x))^2}{f\wp_\pi(x) + (1-f)\wp_K(x)}dx}. \tag{5.7}$$

This quantity is independent of the sample size, but depends on the true values of fractions as generally happens for resolutions. This is also intuitive, since an easier separation is expected among populations similar in size. The quantity s ranges from zero, i. e., no separation corresponding to completely overlapping distributions, to one, i. e., the maximum achievable separation in the given sample. The quantity s is well-defined whatever the shape and the dimensionality of the PID distributions involved. In addition, any s value can be analytically converted in an "equivalent $n\sigma$ separation" to quote, in a more conventional way, the separation in σ units one would have observed if the distributions of the chosen PID variable were Gaussian.

The achieved separation power depends on the variable (i. e., x) used to measure it, and a wise choice of the variable (or set of variables) may enhance the actual separation. To find the optimal variable, we checked a few combinations of observed dE/dx, expected dE/dx, and dE/dx resolutions.[5] We chose to use dE/dx residuals with pion mass hypothesis, δ_π, since no significant enhancements in separation were found by using other variables.

We evaluated the separation using an approximately 97%-pure sample of kaons and pions from candidates with mass within 20 MeV/c^2 from the nominal D^0 mass (see fig. 5.19). Following relation (5.7), the resulting separation is $s^+ \simeq 59.53\%$ for positively-charged particles and $s^- \simeq 58.60\%$ for negatively-charged particles, approximately constant in the momentum range of interest $2 \lesssim p \lesssim 20$ GeV/c. These values correspond to approximately 1.43σ and 1.39σ Gaussian-equivalent

[4] The symbol x may stand for a set of many variables, discrete or continuous, and the integral extends over the whole x domain.

[5] We studied dE/dx residuals, residuals normalized by the dE/dx uncertainty, and the Neyman-Pearson Test (NPT) variable. The NPT is the ratio of the likelihood functions corresponding to the two classes ($\wp_\pi/\wp_K$) and was tried because it is the most statistically powerful variable when a cut is used to separate events, i. e., cutting on the NPT variable provides the least background at fixed efficiency.

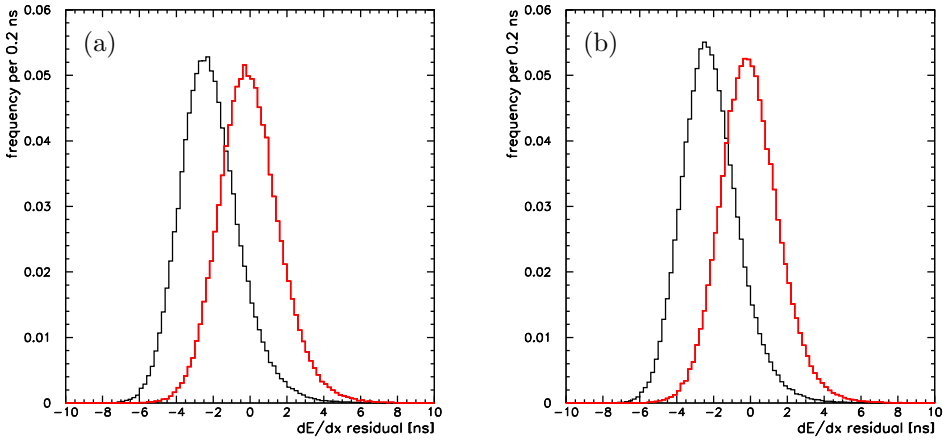

Figure 5.19: Distribution of dE/dx around the average pion response for negatively (a) and positively-charged (b) pions (red line) and kaons (black line).

separations, respectively. Since s (see eq. (5.7)) depends on the specific proportions among classes of events present in the sample, the above values hold only for samples with approximately equal contributions from pions and kaons. This condition is (accidentally) satisfied in the $B^0_{(s)} \to h^+h'^-$ sample (see chap. 7).

	Separation s [%]	Equivalent separation [Gaussian σ]
Uncorrected	$\simeq 48.78$	$\simeq 1.10$
$\varphi_0 - t$	$\simeq 54.90$	$\simeq 1.28$
$\eta - t$	$\simeq 56.43$	$\simeq 1.33$
Hit multiplicity	$\simeq 59.70$	$\simeq 1.43$

Table 5.1: Observed separation between positively-charged kaons and pions as a function of the corrections applied.

Table 5.1 contains the values of the separation evaluated after each step of the calibration. A 22% improvement is achieved in passing from the uncorrected dE/dx to the dE/dx after the corrections. The values refer to positively-charged kaons and pions; similar results are found for negatively-charged particles.

5.7 Correlations

Figure 5.20 shows the distribution of the residual for pions (with pion hypothesis) as a function of the residual for kaons (with kaon hypothesis) after all dE/dx corrections have been applied. A non-zero, positive correlation is apparent from the shape of the distribution, corresponding to a correlation

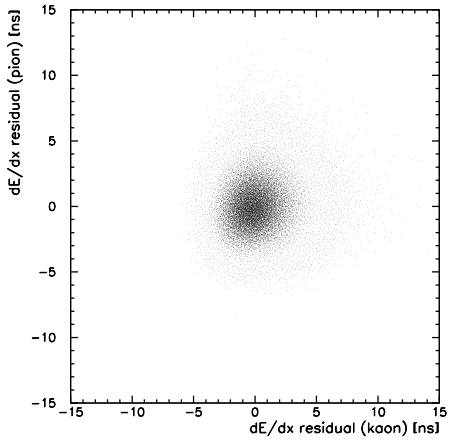

Figure 5.20: Residual for pions (with pion hypothesis) as a function of the residual for kaons (with kaon hypothesis).

coefficient $\rho \simeq 11\%$.[6]

5.7.1 Source of the correlation

With an ideal PID detector, no correlation is expected between independent measurements. Non vanishing correlation indicates the presence of residual dE/dx gain variations. An uncorrected gain variation, in fact, would induce a correlation between the observed ionizations of distinct particles, through the inevitable correlations present in the calibration sample.

The mechanism through which the correlations of the calibration sample induce correlations in the dE/dx response is easily explained in mathematical terms. Suppose to use dE/dx information from two particles per event. One wants to properly write the distributions of dE/dx for the two particles, in general $\wp(dE'/dx', dE''/dx'')$. If the dE/dx shows a gain variation as a function of, say φ_0, the p.d.f. becomes conditional at given pair of azimuthal angles:

$$\wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\right) \quad \longrightarrow \quad \wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\middle|\varphi_0', \varphi_0''\right) \times \wp\left(\varphi_0', \varphi_0''\right). \tag{5.8}$$

Hence, the desired distribution is obtained as

$$\begin{aligned}\wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\right) &= \int \wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\middle|\varphi_0', \varphi_0''\right) \times \wp\left(\varphi_0', \varphi_0''\right) d\varphi_0' d\varphi_0'' \\ &= \int \wp\left(\frac{dE'}{dx'}\middle|\varphi_0'\right) \wp\left(\frac{dE''}{dx''}\middle|\varphi_0''\right) \times \wp\left(\varphi_0', \varphi_0''\right) d\varphi_0' d\varphi_0''. \end{aligned} \tag{5.9}$$

If the azimuthal angles are independent observables,

$$\wp(\varphi_0', \varphi_0'') = \wp(\varphi_0') \times \wp(\varphi_0''), \tag{5.10}$$

[6] The correlation coefficient in this case is $\rho = \frac{E[\delta_\pi \times \delta_K] - E[\delta_\pi] \times E[\delta_K]}{\sigma_{\delta_\pi} \times \sigma_{\delta_K}}$, in which $E[x]$ indicates the expected value of x, and σ are sample standard-deviations.

and eq. (5.9) can be factorized as:

$$\begin{aligned} \wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\right) &= \int \wp\left(\frac{dE'}{dx'}\middle|\varphi_0'\right) \wp\left(\varphi_0'\right) d\varphi_0' \int \wp\left(\frac{dE''}{dx''}\middle|\varphi_0''\right) \wp\left(\varphi_0''\right) d\varphi_0'' \\ &= \wp\left(\frac{dE'}{dx'}\right) \wp\left(\frac{dE''}{dx''}\right). \end{aligned} \tag{5.11}$$

However, eq. (5.10) does not usually hold for particles in a same event. In particular, if the particles originate from the same decay, their kinematic quantities (φ_0, η, ...) will be strongly dependent, preventing the factorization of the dE/dx distribution and yielding correlated values of the observed dE/dx. The sources of correlation can be divided into two types:

Global effects – these are all effects not-related to kinematics. Suppose that the dE/dx shows gain variations as a function of the instantaneous luminosity: $\frac{dE}{dx} = \frac{dE}{dx}(\beta\gamma, \mathcal{L})$. Then, since the kaon and the pion from a D^0 decay are reconstructed in the same event (i. e., in the same conditions of luminosity), their observed dE/dx would appear correlated by the common dependence on luminosity. This may apply to a variety of global variables, such as time, pressure or temperature of the gas, and so forth.

Local effects – these are all effects related to kinematics. Suppose that the dE/dx shows gain variations as a function of the azimuthal angle of emission of the particle: $\frac{dE}{dx} = \frac{dE}{dx}(\beta\gamma, \varphi_0)$. Then, since the azimuthal angle of a kaon and a pion from a D^0 decay are correlated by the kinematic of the decay and the selection biases, their observed dE/dx would become correlated. This may apply to a variety of local variables, such as η, z_0, hit multiplicity, etc.

5.7.2 Impact of the correlations

In presence of correlations between observables of a likelihood fit, not only the statistical precision is degraded, but biased estimates may occur. The correlation between observed values of dE/dx plays a role when dE/dx information of multiple particles per event is used, as in our case. If the values dE'/dx', dE''/dx'',..., dE^n/dx^n are observed for n particles in an event, the presence of correlation imply that these random variables are not independently distributed, i. e., the probability of observing a set of dE/dx values is not factorizable in the single probabilities:

$$\wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}, \ldots, \frac{dE^n}{dx^n}\right) \neq \wp'\left(\frac{dE'}{dx'}\right) \times \cdots \times \wp^n\left(\frac{dE^n}{dx^n}\right). \tag{5.12}$$

The joint p.d.f. in the left-hand side term of above relation needs to be used, instead of the factorized form in the right-hand side. Otherwise, when the dE/dx is used as a discriminating observable in a fit, the multi-particle probability distribution for a dE/dx observable would be incorrect. Similarly, when an event-by-event dE/dx-based selection is applied (i. e., when "cutting" on the dE/dx), the expected overall efficiency corresponding to the chosen requirements would be wrong. Both cases may yield biased estimates of the sample composition.

Relation (5.12) may hold also in absence of correlation, since *dependence* between variates, a weaker condition than *correlation*, is the harmful feature.[7] However, following a common colloquial

[7] As an example, the variates x and $y = x^2$ with flat x-distribution in $[0, 1]$ are fully dependent but show zero correlation.

convention, in this thesis we refer to this dependence as correlation, because in our specific case (common-mode gain-variations, see next section) correlation and dependence coincide.

5.7.3 Study of the correlations

We investigated the combined effect of any possible residual gain variation by allowing for a generic, time-dependent common-mode fluctuation $c(t)$ that affects and correlates the observed dE/dx values in the event. In particular, we extracted the variance (σ_c^2) of the distribution of the common mode, as an estimator of the size of the correlation. We denote the probability distributions of dE/dx residual for pions (with pion mass-hypothesis) as $\wp_\pi(\delta_\pi)$, with standard deviation σ_π. An analogous notation is adopted for kaons. If δ_π and δ_K were independent variables, the probability distribution of their sum $(\delta_K + \delta_\pi)$ would satisfy

$$\wp(\delta_\pi + \delta_K) = \wp_\pi(\delta_\pi) * \wp_K(\delta_K), \tag{5.13}$$

in which the symbol $*$ indicates the Fourier convolution product.[8] Similarly, their difference $\delta_\pi - \delta_K$ would be distributed as

$$\wp(\delta_\pi - \delta_K) = \wp_\pi(\delta_\pi) * \wp_{-K}(-\delta_K), \tag{5.14}$$

where $\wp_{-K}(-\delta_K)$ is the distribution of the negative residual for kaons ($\frac{dE}{dx}_K - \frac{dE}{dx}_{\rm obs}$), whose variance satisfies the condition $\sigma_K^2 = \sigma_{-K}^2$. Since the variance of a convolution product is the sum of variances of the convoluted distributions, the standard deviations of the distributions of sum and difference are equal:

$$\sigma_{\pi+K} = \sigma_{\pi-K} = \sqrt{\sigma_\pi^2 + \sigma_K^2}. \tag{5.15}$$

On the other hand, if the two residuals are correlated by a common-mode fluctuation, the observed residual ($\delta^{\rm obs}$) is written as the sum of the intrinsic, uncorrelated residual with the common-mode shift:

$$\delta_\pi^{\rm obs} = \delta_\pi + c \quad \text{and} \quad \delta_K^{\rm obs} = \delta_K + c. \tag{5.16}$$

Therefore, the sum of observed residuals, $\delta_\pi^{\rm obs} + \delta_K^{\rm obs} = \delta_\pi + \delta_K + 2c$, is distributed as

$$\wp(\delta_\pi^{\rm obs} + \delta_K^{\rm obs}) = \wp_\pi(\delta_\pi) * \wp_K(\delta_K) * \wp_c(2c), \tag{5.17}$$

whereas their difference, $\delta_\pi^{\rm obs} - \delta_K^{\rm obs} = \delta_\pi + c - \delta_K - c = \delta_\pi - \delta_K$, is distributed as

$$\wp(\delta_\pi^{\rm obs} - \delta_K^{\rm obs}) = \wp_\pi(\delta_\pi) * \wp_{-K}(-\delta_K). \tag{5.18}$$

Equations (5.17) and (5.18) show that, in presence of a common mode, the sum of residuals has greater variance than their difference, $\sigma_{\pi+K}^2 > \sigma_{K-\pi}^2$. The standard deviation of the correlation is easily obtained:

$$\sigma_c = \frac{1}{2}\sqrt{\sigma_{\pi+K}^2 - \sigma_{\pi-K}^2}. \tag{5.19}$$

Following eq. (5.19), we used the distributions of sum and difference of the observed residual to estimate the magnitude of time-dependent common modes. We compared the variances of the correlation extracted from two samples:

[8] Henceforth, "convolution" always denote the Fourier convolution product.

physical sample – the sample used in the calibration: each event contains a kaon and a pion produced in the decay of a single D^0 meson. In this sample, the common mode receives contributions from gain variations due to *global* effects (pressure, time, luminosity,...) and *local* effects (φ_0, η, ...) since kaons and pions are kinematically correlated.

mixed-decay sample – an artificial sample containing events in which the kaon and the pion originate from decays of two distinct D^0 mesons, reconstructed in two consecutive events. In this sample, we expect the magnitude of the common mode to be smaller than in the physical sample. The contribution to the gain variations due to *local* effects are reduced, since pion and kaon are kinematically uncorrelated. In addition, the larger the time scale of the gain variation with respect to the typical time interval between the collection of two D^0 decays, the smaller becomes the contribution of *global* effects.

This allowed studying what fraction of the correlation (σ_c) is enhanced by the kinematics of the decay and what is due to global effects (see fig. 5.21). The correlation is expected to reduce after dE/dx

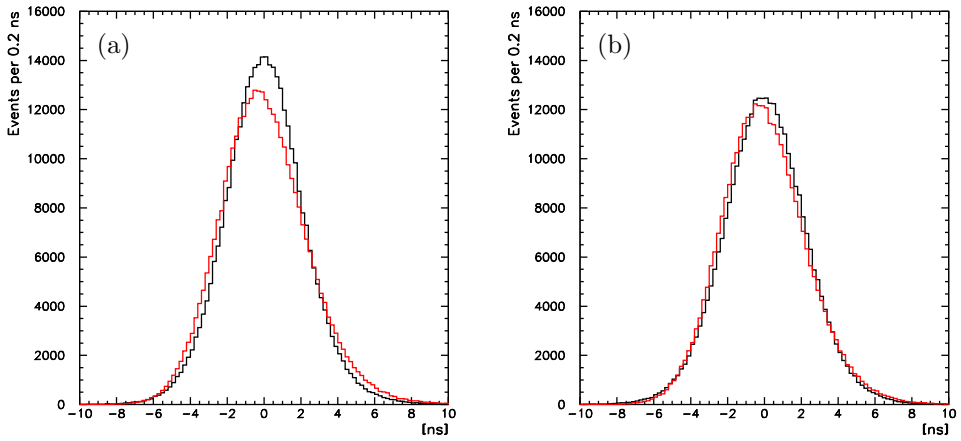

Figure 5.21: Distribution of the sum (black line) and difference (red line) of residuals for a kaon (in kaon hypothesis) and a pion (in pion hypothesis) from a D^0 decay (a). Same distribution of plot (a) obtained by combining kaons and pions from decays of D^0 candidates in adjacent events (b).

tuning. tab. 5.2 shows σ_c evaluated after each step of the dE/dx calibration. A factor of two reduction in the size of the correlation confirms that our corrections were effective. Nevertheless, a 11% residual correlation (corresponding to $\sigma_c \simeq 0.52$ ns) still affects the calibration sample. This has been modeled and properly included in the likelihood fit of the $B^0_{(s)} \to h^+h'^-$ sample (see next section).

5.8 Model of the dE/dx distributions

Use of dE/dx information in a likelihood fit requires modeling the distributions of the desired observables. It is convenient to stress the difference between *observed* dE/dx quantities, i. e., those affected

	σ_c [ns]	$\sigma_c^{\rm mix}$ [ns]
Uncorrected	$\simeq 1.045$	$\simeq 0.54$
$\varphi_0 - t$	$\simeq 0.73$	$\simeq 0.21$
$\eta - t$	$\simeq 0.72$	$\simeq 0.25$
Hit multiplicity	$\simeq 0.52$	$\simeq 0.24$

Table 5.2: Width of the distribution of correlation for physical and mixed-decay samples as a function of the correction applied.

by common-mode fluctuations, and *intrinsic* quantities, that would have been observed in absence of correlations. Since the intrinsic residual and the correlation are, by construction, independent variables (see eq. (5.16)), the (known) distribution of the observed residuals is the convolution of their unknown distributions:

$$\wp(\delta^{\rm obs}) = \wp(\delta + c) = \wp(\delta) * \wp(c). \tag{5.20}$$

The models for the intrinsic residuals, $\wp(\delta)$, and for the correlations, $\wp(c)$, were extracted from the distributions of observed residuals, $\wp(\delta^{\rm obs})$, of pions and kaons from D^0 decays. We expanded each term of the right-hand side of eq. (5.20) in sum of Gaussian distributions, and we fit the distributions of observed residuals to extract the unknown parameters. In practice, the first two terms of the expansion were sufficient to model accurately intrinsic residuals and correlations:

$$\wp_K(\delta_K) = q \cdot \mathscr{G}_{K'}(\delta_K) + (1-q) \cdot \mathscr{G}_{K''}(\delta_K) \tag{5.21}$$

$$\wp_\pi(\delta_\pi) = p \cdot \mathscr{G}_{\pi'}(\delta_\pi) + (1-p) \cdot \mathscr{G}_{\pi''}(\delta_\pi) \tag{5.22}$$

$$\wp_c(c) = r \cdot \mathscr{G}_{c'}(c) + (1-r) \cdot \mathscr{G}_{c''}(c) \tag{5.23}$$

where we used the following notation for the Gaussian distribution: $\mathscr{G}_s(x) = \mathscr{G}(x;\mu_s,\sigma_s) = \frac{1}{\sigma_s\sqrt{2\pi}} e^{-\frac{(x-\mu_s)^2}{2\sigma_s^2}}$. Independent parameterizations were assumed for the distributions of intrinsic residuals for positively and negatively-charged particles. Mean (μ), variance (σ^2) and fraction of each Gaussian were determined with a simultaneous, binned ML fit of the following combinations of observed residuals:

$$\wp_K(\delta_K^{\rm obs}) = \wp(\delta_K) * \wp(c) = (\mathscr{G}_{K'} + \mathscr{G}_{K''}) * (\mathscr{G}_{c'} + \mathscr{G}_{c''}) \tag{5.24}$$

$$\wp_\pi(\delta_\pi^{\rm obs}) = \wp(\delta_\pi) * \wp(c) = (\mathscr{G}_{\pi'} + \mathscr{G}_{\pi''}) * (\mathscr{G}_{c'} + \mathscr{G}_{c''}) \tag{5.25}$$

$$\wp(\delta_\pi^{\rm obs} + \delta_K^{\rm obs}) = (\mathscr{G}_{\pi'} + \mathscr{G}_{\pi''}) * (\mathscr{G}_{K'} + \mathscr{G}_{K''}) * (\mathscr{G}_{2c'} + \mathscr{G}_{2c''}) \tag{5.26}$$

$$\wp(\delta_\pi^{\rm obs} - \delta_K^{\rm obs}) = (\mathscr{G}_{\pi'} + \mathscr{G}_{\pi''}) * (\mathscr{G}_{-K'} + \mathscr{G}_{-K''}), \tag{5.27}$$

where the relative normalization factors (p, q, r) were included in the fit, but omitted above for a clearer notation, and where $\mathscr{G}_{2s}(x) = \mathscr{G}(x; 2\mu_s, \sqrt{2}\sigma_s)$ and $\mathscr{G}_{-s}(x) = \mathscr{G}(x; -\mu_s, \sigma_s)$.

Figure 5.22 shows a satisfactory agreement between the chosen model and the distributions of intrinsic residuals and correlations. The only, slight discrepancies between data and model appear in the positive side of the distribution, with effects smaller than 1%. Figure 5.23(a) shows the resulting models of the intrinsic residuals. Although we allowed for independent residual distributions

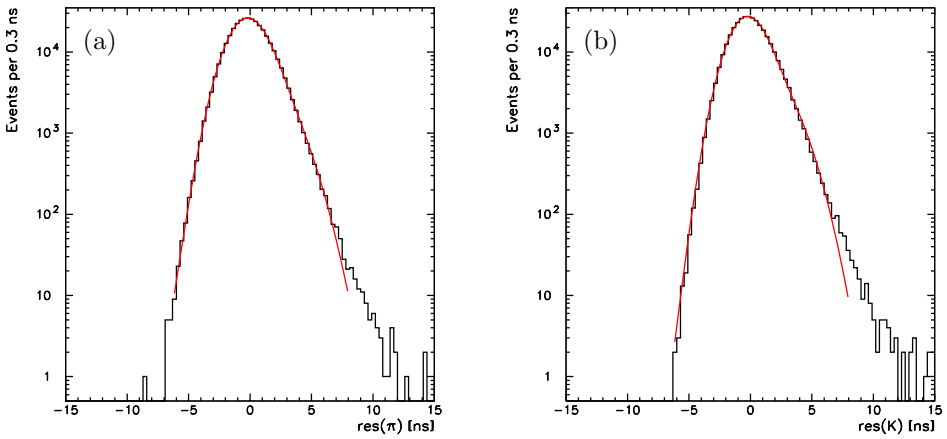

Figure 5.22: Distribution of observed dE/dx residual for pions (with pion mass hypothesis) (a), and for kaons (with kaon mass hypothesis) (b). Results of the fit to the functions in eq. (5.25) and eq. (5.24) are overlaid (red, solid line).

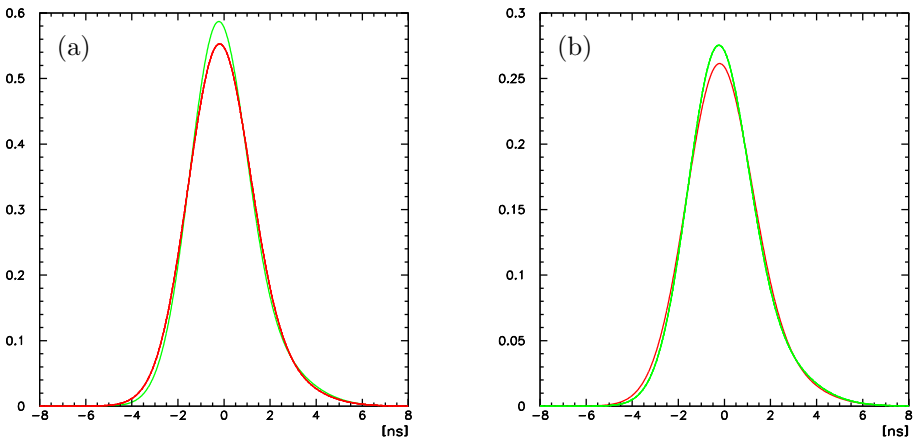

Figure 5.23: Probability density function of the intrinsic (a) and observed (b) residual for pions (red) and kaons (green). The p.d.f. correspond to positively-charged particles.

between kaons and pions, the extracted shapes are similar (see fig. 5.23), both showing non-Gaussian positive tails. Figure 5.24 shows that the differences between residuals of positively and negatively-charged particles are tiny. These small differences between kaons and pions and between positively and negatively-charged particles have been ascribed to a systematic dependence of the dE/dx response on track curvature, caused by the geometric and electrostatic asymmetry of the COT drift-cells. For given Lorentz boost, trajectories of charged particles with different masses or charge have different

curvatures, being sensitive to the systematic effects in the efficiency of charge-collection by the COT sense-wires. Figure 5.25 shows the extracted correlation functions. Following sec. 5.7.3, we extracted

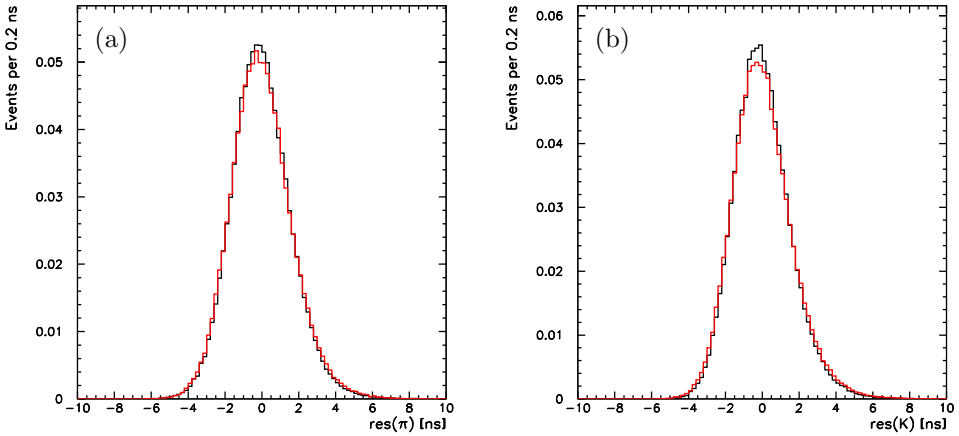

Figure 5.24: Distributions of the observed residuals for positively (black) and negatively-charged (red) pions (a) and kaons (b).

two models for the correlation function: one from the physical sample and the other one from the mixed-decay sample. Both models show a narrow core centered at zero, with a small, long positive tail. The width of the correlation between pion and kaon from the same decay is larger than the width in the kinematically independent sample. This effect is expected from the kinematic enhancement of dE/dx correlations discussed in sec. 5.7.

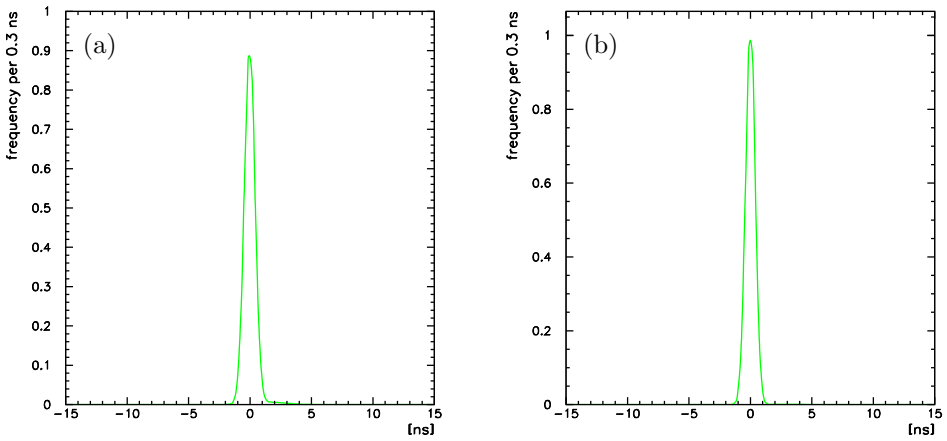

Figure 5.25: Probability density function of the correlation between dE/dx of pions and kaons from the same D^0 decays (a), and from decays of two D^0 candidates reconstructed in consecutive events (b).

Chapter 6

Statistical determination of the sample composition

This chapter describes how the kinematic and PID information discussed in previous chapters was combined in a maximum likelihood fit to statistically determine the composition of the $B^0_{(s)} \to h^+h'^-$ sample. We chose a strictly data-driven approach, using the data to model all likelihood terms, except for the kinematic distributions for signal events. Special emphasis is also devoted to discussing the tests of the fit on simulated samples and the form of the likelihood function, which was carefully chosen to avoid any possible bias.

6.1 The fit

The composition of the $B^0_{(s)} \to h^+h'^-$ sample, including the relative yields of $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ decays, was statistically determined with a multivariate maximum likelihood fit [125] that combined kinematic and PID information. We wrote an unbinned likelihood $\mathscr{L}(\vec{\theta}|\vec{x})$, function of the n parameters of the model $\vec{\theta} = (\theta_1, \ldots, \theta_n)$. For any specific set of values of the unknown parameters $\vec{\theta}$, $\mathscr{L}$ is the joint probability density for obtaining the particular set of values of the discriminating variables observed in the sample, $\vec{x}$. In the fitting process, we maximized the likelihood function with respect to the unknown parameters, thus determining their ML estimate. Technical details of the numerical maximization of the likelihood by means of the MINUIT package [126] are given in appendix B.

An appropriate choice of the discriminating observables is crucial to fully exploit the available information while keeping a simple analytic expression of the likelihood. The goal is to use the maximum available information through the minimum number of observables. In addition, the more independent the chosen observables are, the simpler is factorizing and modeling the joint probability density.

6.1.1 Kinematic observables

In the kinematic portion, we exploited the correlation between the invariant mass of the particle pair with pion mass assignments ($m_{\pi^+\pi^-}$) and the signed momentum imbalance, $\alpha = (1 - p_1/p_2)q_1$, accordingly to what is shown in sec. 4.2. The arbitrary assignment of the pion mass to both particles is convenient with respect to the $K\pi$ assignment because, being symmetric, it prevents complications from interchanged $K^+\pi^- \leftrightarrow K^-\pi^+$ assignments.

Additional information from momenta, complementary to the information from their ratio, is exploited by using also the scalar sum of momenta of the particles, $p_{\rm tot} = p_1 + p_2$. This quantity is required because, in spite of its limited separation power (it hardly discriminates all signals from background), it allows a proper factorization of the likelihood function, as it is shown in sec. 6.2.3. Using the variables $(\alpha, p_{\rm tot})$ in place of (p_1, p_2) to access momentum information is convenient, because it allows writing the probability distributions in terms of lesser correlated quantities.

6.1.2 PID observables

Particle identification information was summarized in a single observable for each charged particle, the "kaonness" κ, defined as

$$\kappa = \frac{dE/dx_{\rm obs} - dE/dx_{\pi}}{dE/dx_{K} - dE/dx_{\pi}}. \tag{6.1}$$

The average of this quantity is, by construction, zero for pions and one for kaons, with almost momentum-independent distribution for both types of particles. This is particularly convenient in our case, since all $B^0_{(s)} \to h^+h'^-$ modes have only pions and kaons in their final states, and also the background composition is expected to be dominated by these particles.

In summary, the likelihood was calculated by evaluating the joint probability density corresponding to the values of the following, five discriminating observables for each event:

1. $\boldsymbol{m_{\pi^+\pi^-}}$ – the invariant mass of the pair of final state particles with pion mass assignments;
2. $\boldsymbol{\alpha}$ – the signed momentum imbalance between the two particles;
3. $\boldsymbol{p_{\rm tot}}$ – the scalar sum of particle momenta;
4. $\boldsymbol{\kappa_1}$ – the kaonness (function of the dE/dx) of the lower-momentum particle;
5. $\boldsymbol{\kappa_2}$ – the kaonness of the higher-momentum particle.

6.2 Likelihood function

The likelihood function $\mathscr{L}$ was written as a product, over the N events in the sample, of the single-event likelihood $\mathscr{L}_i$:

$$\mathscr{L}(\vec{\theta}|\vec{x}) \equiv \mathscr{L}(\vec{\theta}) = \prod_{i=1}^{N} \mathscr{L}_i(\vec{\theta}). \tag{6.2}$$

Once the final selection was applied (see tab. 3.3), each event contained only one $B^0_{(s)} \to h^+h^{'-}$ candidate. Thus, additional correlation terms in the likelihood, needed in case of multiple candidates per event, could be neglected. The single-event likelihood is written as

$$\mathscr{L}_i(\vec{\theta}) = (1-b)\mathscr{L}^{\rm sig} + b\mathscr{L}^{\rm bck} \tag{6.3}$$

In eq. (6.3), $\mathscr{L}^{\rm sig}$ is the term that describes the probability distributions of the observables for signal events and $\mathscr{L}^{\rm bck}$ is the corresponding term for background events; b is the fraction of background events, to be determined by the fit.

6.2.1 Signal likelihood

The likelihood that describes the signal is a sum of factorized probability density functions:

$$\begin{aligned}
\mathscr{L}^{\rm sig} &= \sum_{j=1}^{s} f_j(\wp^m_j\wp^p_j\wp^{\rm PID}_j) = \\
&\quad f_{B^0\to\pi^+\pi^-}(\wp^m_{B^0\to\pi^+\pi^-}\wp^p_{B^0\to\pi^+\pi^-}\wp^{\rm PID}_{B^0\to\pi^+\pi^-}) \\
&+ \quad f_{B^0\to K^+\pi^-}(\wp^m_{B^0\to K^+\pi^-}\wp^p_{B^0\to K^+\pi^-}\wp^{\rm PID}_{B^0\to K^+\pi^-}) \\
&+ \quad f_{\overline{B}^0\to K^-\pi^+}(\wp^m_{\overline{B}^0\to K^-\pi^+}\wp^p_{\overline{B}^0\to K^-\pi^+}\wp^{\rm PID}_{\overline{B}^0\to K^-\pi^+}) \\
&+ \quad f_{B^0_s\to K^+K^-}(\wp^m_{B^0_s\to K^+K^-}\wp^p_{B^0_s\to K^+K^-}\wp^{\rm PID}_{B^0_s\to K^+K^-}) \\
&+ \quad f_{B^0_s\to K^-\pi^+}(\wp^m_{B^0_s\to K^-\pi^+}\wp^p_{B^0_s\to K^-\pi^+}\wp^{\rm PID}_{B^0_s\to K^-\pi^+}) \\
&+ \quad f_{\overline{B}^0_s\to K^+\pi^-}(\wp^m_{\overline{B}^0_s\to K^+\pi^-}\wp^p_{\overline{B}^0_s\to K^+\pi^-}\wp^{\rm PID}_{\overline{B}^0_s\to K^+\pi^-}) \\
&+ \quad f_{...}(\wp^m_{...}\wp^p_{...}\wp^{\rm PID}_{...}),
\end{aligned} \tag{6.4}$$

in which the index j runs over all expected $B^0_{(s)} \to h^+h^{'-}$ modes, that is, $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $\overline{B}^0 \to K^-\pi^+$, $B^0_s \to K^-\pi^+$, $\overline{B}^0_s \to K^+\pi^-$, $B^0_s \to K^+K^-$, $B^0 \to K^+K^-$, $B^0_s \to \pi^+\pi^-$, etc.[1] The parameters f_j are their fractions (with respect to the total signal) to be determined by the fit. From the $(s-1)$ independent fractions resulting by the normalization condition,

$$f_s = 1 - \sum_{j=1}^{s-1} f_j, \tag{6.5}$$

we determined the yield of each mode. We conventionally labeled as $\wp^m$ the term that describes the invariant-mass distributions ("mass term"), as $\wp^p$ the term that describes the momentum distributions ("momentum term"), and as $\wp^{\rm PID}$ the term that models the dE/dx density ("PID term"). We show in sec. 6.2.3 that this factorization is not trivial since the three terms of the p.d.f. are inter-related by the dependencies between mass, momentum, and dE/dx observables.

The model of the kinematic terms of the likelihood, $\wp^p$ and $\wp^m$, was extracted from Monte Carlo simulation (see sec. 3.6), while the model of the PID term was determined from the samples of $D^0 \to h^+h^{'-}$ decays described in sec. 5.1.

[1] C-conjugate modes are considered distinct for decays in $K\pi$ final states.

Probability density function of the mass term

The invariant $\pi^+\pi^-$-mass distribution of each decay mode is dominated by two contributions: one, present only in the terms of the non-$\pi^+\pi^-$ modes, is the mass shift, function of the momentum imbalance of the decay. It is due to mis-assigned masses of the outgoing particles, as discussed in sec. 4.2; the other one is the effect of the finite momentum resolution of the detector. We accounted for both effects by writing a mass p.d.f. which is conditional for given momentum imbalance: for each observed value of invariant $\pi^+\pi^-$-mass and of momentum imbalance, we assumed the mass distributed with Gaussian spread around the specific average-value expected at that momentum imbalance, $\mathcal{M}(\alpha)$. An event, whose candidate has an invariant $\pi^+\pi^-$-mass $m_{\pi^+\pi^-}$ and a momentum imbalance α, contributes to the mass p.d.f. of the jth signal mode with the following term:

$$\wp_j^m(m_{\pi^+\pi^-}|\alpha;\vec{\theta}) = \mathscr{G}(m_{\pi^+\pi^-}|\alpha;\vec{\theta}) = \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m_{\pi^+\pi^-}-\mathcal{M}_j(\alpha)}{\sigma_m}\right)^2}, \tag{6.6}$$

in which the dispersion σ_m accounts for the detector resolution, and the mean value $\mathcal{M}_j(\alpha)$ depends on the mode (j) and on the observed signed momentum imbalance (α). For each mode, we included the analytic expression of $\mathcal{M}_j(\alpha)$ shown in tab. 4.1, which depends on the masses m_{B^0}, $m_{B_s^0}$, m_K, and m_π as external inputs. For the masses of charged pion and kaon, we used the world average values [109]. For the $B^0_{(s)}$ mesons, we used the masses measured by CDF in Run II: $m_{B^0} = 5279.63 \pm 0.53$ (*stat.*) ± 0.33 (*syst.*) MeV/c^2 and $m_{B_s^0} = 5366.01 \pm 0.73$ (*stat.*) ± 0.33 (*syst.*) MeV/c^2 [105]. Using the masses measured with the same apparatus and reconstruction code as for $B^0_{(s)} \to h^+h^{'-}$ data allows cancellation of common systematic uncertainties. Therefore, only the statistical uncertainty in the measurement of b-hadron masses of Ref. [105] contributes to the systematic uncertainty in our measurement (see sec. 9.2.2).

We used a single mass resolution σ_m for all $B^0_{(s)} \to h^+h^{'-}$ modes, as predicted by the simulation in sec. 3.6. The value of σ_m was obtained rescaling the mass-width observed in simulated $B^0_{(s)} \to h^+h^{'-}$ decays to correct for the discrepancy observed between data and simulation (see sec. 3.6). We extracted the scale factor by comparing the observed widths of several two-body decays ($D^0 \to h^+h^{'-}$, $J/\psi \to \mu^+\mu^-$, and $\Upsilon \to \mu^+\mu^-$) with the predictions of the simulation. This yielded a 1.26 scale factor to be applied to the simulation, resulting in an expected r.m.s. width of $\sigma_m = 28 \pm 0.3$ MeV/c^2 for each individual $B^0_{(s)} \to h^+h^{'-}$ mode.

The effect of the small variation of the mass resolution as a function of the transverse momentum of the $B^0_{(s)}$ meson (about 1.5 MeV/c^2, spanning the range $2 \lesssim p_{\rm T}(B) \lesssim 15$ GeV/c, see sec. 3.6) was included in the final checks of the measurement (see sec. 9.2.1).

Probability density function of the momentum term

The momentum p.d.f. is extracted from the simulated signal-samples described in sec. 3.6. Since the signed momentum-imbalance (α) and the scalar sum of momenta ($p_{\rm tot}$) are not independent observables, we used a joint p.d.f., $\wp_j^p(\alpha, p_{\rm tot};\vec{\theta})$, to model their distributions. Selection requirements and kinematic correlations between the outgoing particles cause domain and shape of the α distribution to vary as a function of $p_{\rm tot}$. In fact, both momenta (p_1 and p_2) are necessarily larger than 2 GeV/c, and their sum is larger than 5.5 GeV/c, because of the trigger requirements on their transverse

components. This translates into the following conditions on α and $p_{\rm tot}$:

$$p_1 = p_{\rm tot}\left(\frac{1-|\alpha|}{2-|\alpha|}\right) > 2 \text{ GeV}/c \quad \text{and} \quad p_{\rm tot} > 5.5 \text{ GeV}/c. \tag{6.7}$$

The other constraint on momentum, $p_2 = p_{\rm tot}/(2-|\alpha|) > 2$ GeV/c, is automatically satisfied if eq. (6.7) holds. The domain of the joint p.d.f. is defined by the requirements of relation (6.7) and it is shown in fig. 6.1. The joint p.d.f. is written as the probability density of the scalar sum of momenta,

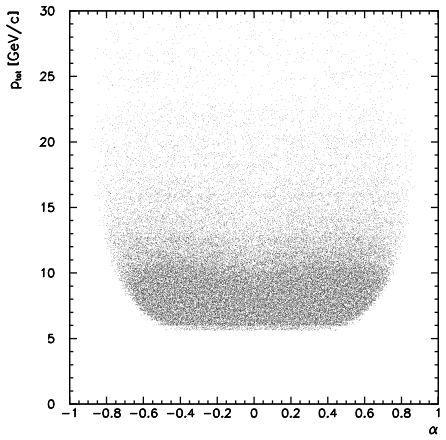

Figure 6.1: Allowed region in the ($p_{\rm tot}$, α) domain. Distribution of the scalar sum of momenta as a function of the momentum imbalance in simulated $B^0_{(s)} \to h^+h^{'-}$ decays.

$\wp(p_{\rm tot};\vec{\theta})$, times the conditional probability density of the signed momentum-imbalance at given $p_{\rm tot}$, $\wp(\alpha|p_{\rm tot};\vec{\theta})$. We empirically chose the shapes of both densities, whose parameters were determined by means of a two-dimensional, binned, ML fit of the simulated distributions. An event, whose candidate has an observed scalar sum of momenta $p_{\rm tot}$ and an observed momentum-imbalance α, contributes to the momentum p.d.f. of the jth signal mode with the following term:

$$\begin{aligned}\wp^p_j(\alpha, p_{\rm tot};\vec{\theta}) &= \wp_j(p_{\rm tot};\vec{\theta}) \times \wp_j(\alpha|p_{\rm tot};\vec{\theta}) \\ &= \frac{1}{K_j}\left(e^{c_j p_{\rm tot}}\sum_{l=0}^{4} a_{l,j}\, p^l_{\rm tot}\right) \times \left[\sum_{m=0}^{6} b_{m,j}\,\alpha^m \left(\frac{p_{\rm tot}-2}{p_{\rm tot}-4}\right)^m\right],\end{aligned} \tag{6.8}$$

in which the $p_{\rm tot}$ density is the product of an exponential function times a 4th-degree polynomial, whereas the conditional p.d.f. of α is a 6th-degree polynomial in the α variable scaled by a factor $(p_{\rm tot}-2)/(p_{\rm tot}-4)$ deriving from the constraint on the domain of α given by relation (6.7). The index j of the a, b, and c parameters denotes their dependence on the decay mode. The odd b terms are set to zero for decays into $\pi^+\pi^-$ and K^+K^- final states, where α distributions are symmetric because the two outgoing particles have same mass. The normalization factor K_j for each mode is calculated with numerical two-dimensional integration of the p.d.f. in the appropriate domain of α and $p_{\rm tot}$.

Figure 6.2 shows the distributions of the scalar sum of momenta as a function of the momentum imbalance for the simulated modes. These are fit to the functions of eq. (6.8) to obtain the momentum

templates shown in fig. 6.3. We checked the agreement between model and data by overlaying the templates to the α-distributions of simulated data, sampled in different $p_{\rm tot}$ ranges. An example of this check is shown in fig. 6.4 for the $B^0 \to \pi^+\pi^-$ mode. A similar agreement was found for all simulated modes.

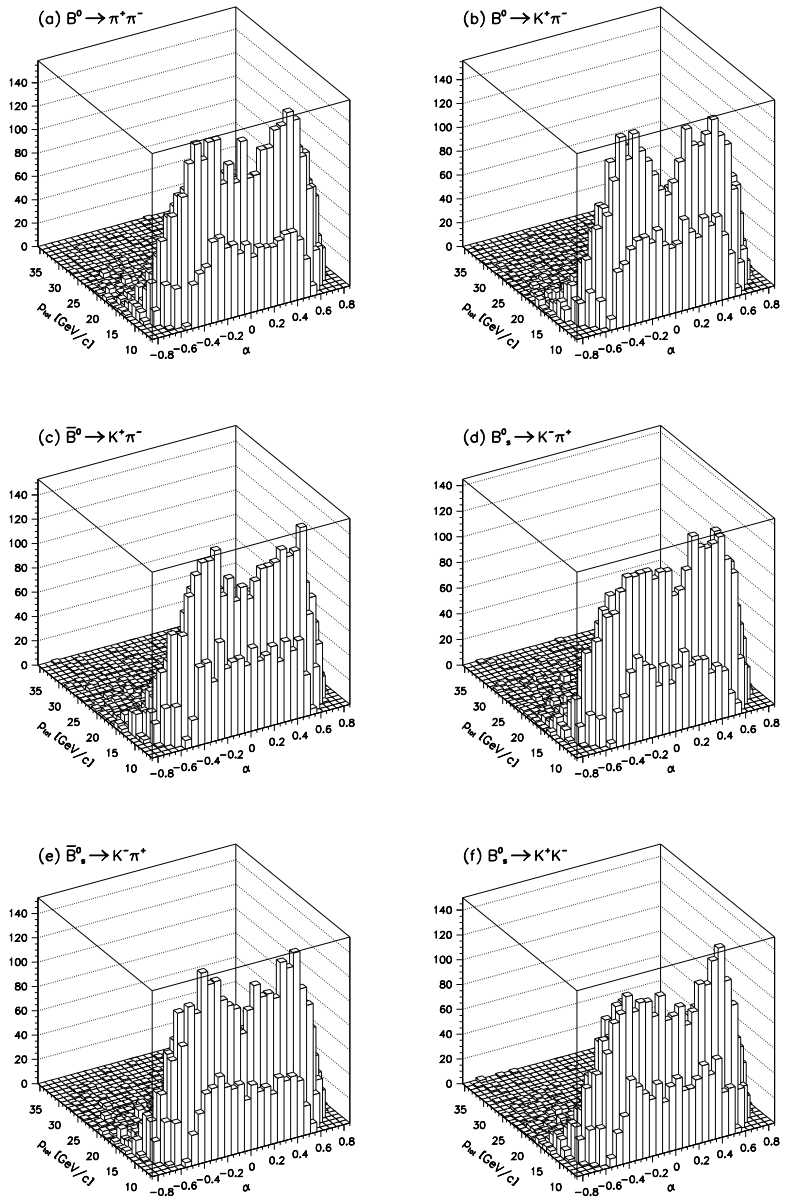

Figure 6.2: Distribution of the scalar sum of momenta as a function of the signed momentum imbalance in the following simulated decays: (a) $B^0 \to \pi^+\pi^-$ and $\overline{B}^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $\overline{B}^0 \to K^-\pi^+$, (d) $B_s^0 \to K^-\pi^+$, (e) $\overline{B}_s^0 \to K^+\pi^-$, and (f) $B_s^0 \to K^+K^-$ and $\overline{B}_s^0 \to K^+K^-$.

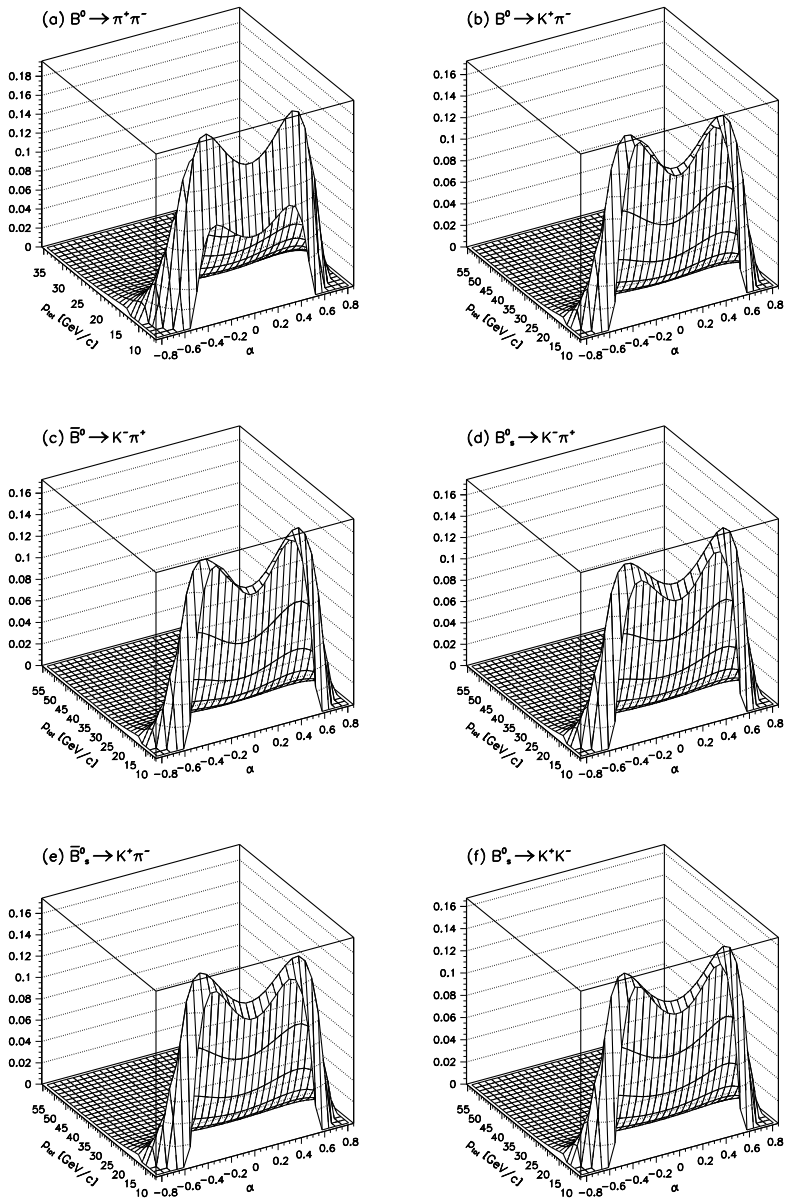

Figure 6.3: Joint $(\alpha, p_{\rm tot})$ templates as determined from simulated signal samples for the following modes: (a) $B^0 \to \pi^+\pi^-$and $\overline{B}^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $\overline{B}^0 \to K^-\pi^+$, (d) $B_s^0 \to K^-\pi^+$, (e) $\overline{B}_s^0 \to K^+\pi^-$, and (f) $B_s^0 \to K^+K^-$ and $\overline{B}_s^0 \to K^+K^-$.

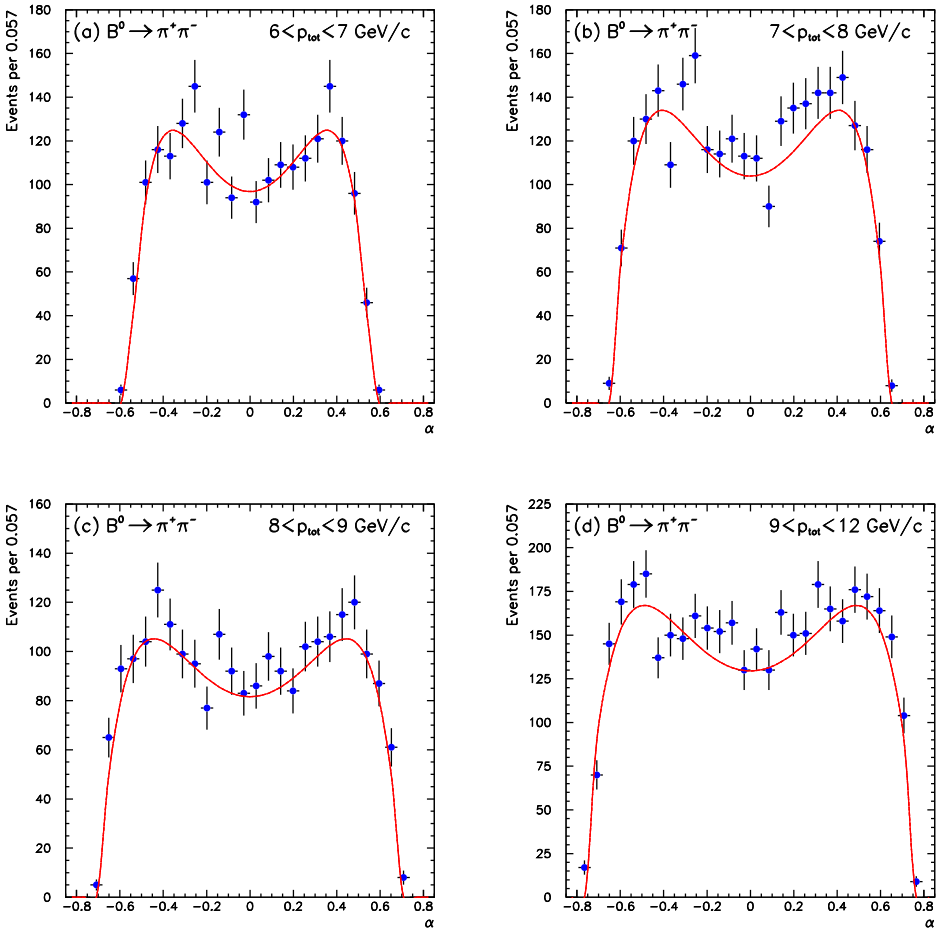

Figure 6.4: Distributions (points with error bars) of the signed momentum imbalance for simulated $B^0 \to \pi^+\pi^-$ and $\overline{B}^0 \to \pi^+\pi^-$ decays sampled in four distinct ranges of scalar sum of momenta: $6 < p_{\rm tot} < 7$ GeV/c (a), $7 < p_{\rm tot} < 8$ GeV/c (b), $8 < p_{\rm tot} < 9$ GeV/c (c), and $9 < p_{\rm tot} < 12$ GeV/c (d). The corresponding templates are overlaid (solid line).

Probability density function of the PID term

The p.d.f. of the PID information can not be factorized in the probability densities of the two particles, because of the 11% correlation between the observed dE/dx values in an event (see sec. 5.7). We therefore wrote a two-particle, joint p.d.f. that incorporates the probability densities for the intrinsic dE/dx observables of each particle, and for the correlation function.

We used the shapes of intrinsic residuals and correlation determined in sec. 5.8 to write the p.d.f., which results from a convolution integral plus a transformation of variables. The convolution combines the intrinsic dE/dx residuals of both particles (δ_1 and δ_2) through the p.d.f. of correlation, $\wp_c(c)$,

yielding the following p.d.f. for the jth mode:

$$\wp_j(\delta_1^{\rm obs},\delta_2^{\rm obs};\vec{\theta}) = [\wp_l(\delta_1)\times\wp_m(\delta_2)] * \wp_c(c) = \int_{-\infty}^{+\infty} \wp_l(\delta_{1,l}-c)\wp_m(\delta_{2,m}-c)\wp_c(c)dc, \tag{6.9}$$

which, at this stage, is independent of momenta. The indices $l,m = K$, or π depend on the identity of particles in the final states of the jth mode, and determine the choice of the shapes for intrinsic residuals within the integral, which are slightly different between pions and kaons (see fig. 5.23).

The p.d.f. in eq. (6.9) is function of a different set of observables for each mode, e. g., δ_π for the $B^0\to\pi^+\pi^-$ case, δ_π and δ_K for the $B^0\to K^+\pi^-$ case, and so on. To avoid biases related to using different sets of observables in different terms of the likelihood function [127], we rewrote the likelihood in terms of a single observable, the kaonness κ. However, changing variable from residual to kaonness induces an additional dependence on momentum in the PID term. In fact, for each particle type, the following relation holds:

$$\delta = (\kappa - \langle\kappa\rangle)\left(\frac{dE}{dx}_K - \frac{dE}{dx}_\pi\right) \equiv (\kappa-\langle\kappa\rangle)\,\Delta \tag{6.10}$$

in which Δ indicates the difference between expected dE/dx values evaluated in kaon and pion mass-hypothesis, which is function of momentum. The joint p.d.f. in terms of κ is

$$\wp(\kappa_1^{\rm obs},\kappa_2^{\rm obs};\vec{\theta}) = \int_{-\infty}^{+\infty} \wp(\delta_1-c)\wp(\delta_2-c)\left|\frac{\partial(\delta_1,\delta_2)}{\partial(\kappa_1,\kappa_2)}\right|\wp_c(c)dc, \tag{6.11}$$

in which $\delta\equiv\delta(\kappa)$. After writing out the Jacobian, the above equation becomes

$$\wp(\kappa_1^{\rm obs},\kappa_2^{\rm obs};\vec{\theta}) = \int_{-\infty}^{+\infty} \wp(\delta_1-c)\wp(\delta_2-c)\Delta_1\Delta_2\wp_c(c)dc. \tag{6.12}$$

The transformation from residual to kaonness brings the momenta into the probability density through the differences of expected dE/dx values, $\Delta_1(\alpha,p_{\rm tot})$ and $\Delta_2(\alpha,p_{\rm tot})$. Hence, the correct expression of the joint p.d.f. function of κ becomes a conditional probability density at given momenta: $\wp(\kappa_1^{\rm obs},\kappa_2^{\rm obs};\vec{\theta}) \longrightarrow \wp(\kappa_1^{\rm obs},\kappa_2^{\rm obs}|\alpha,p_{\rm tot};\vec{\theta})$.

An event in which δ_1 and δ_2 are the observed dE/dx residuals of the particle pair, contributes to the jth signal mode of the PID p.d.f. with the following term:

$$\wp_j^{\rm PID}(\kappa_1^{\rm obs},\kappa_2^{\rm obs}|\alpha,p_{\rm tot};\vec{\theta}) = \wp_{l,m}^{\rm PID}(\kappa_1^{\rm obs},\kappa_2^{\rm obs}|\alpha,p_{\rm tot};\vec{\theta}) = \frac{1}{K_j}\int_{-\infty}^{+\infty} \wp_l(\delta_{1,l}-c)\wp_m(\delta_{2,m}-c)\Delta_1\Delta_2\wp_c(c)dc, \tag{6.13}$$

which includes $K_j = 1/\int_{-\infty}^{+\infty}\wp_j(\kappa_1^{\rm obs},\kappa_2^{\rm obs}|\alpha,p_{\rm tot};\vec{\theta})d\kappa_1^{\rm obs}d\kappa_2^{\rm obs}$ as a normalization factor. The explicit expression of the p.d.f. was evaluated analytically using MATHEMATICA$^{\rm TM}$, a commercial software-package for computation [128]; we omit here its explicit expression because of its length and lack of visually intelligible features.

6.2.2 Background likelihood

The single-event likelihood for background is factorized in three terms:

$$\mathscr{L}^{\rm bck} = \wp_{\rm bck}^m\wp_{\rm bck}^p\wp_{\rm bck}^{\rm PID}. \tag{6.14}$$

The model of mass and momentum terms was extracted in $B^0_{(s)} \to h^+h^{'-}$ data from events distributed in the invariant mass regions at lower ($4.850 < m_{\pi^+\pi^-} < 5.125$ GeV/c^2) or higher ($5.4 < m_{\pi^+\pi^-} < 5.8$ GeV/c^2) masses with respect to the signal. These regions are henceforth identified as "mass sidebands". As for the signal, the model of the PID term for background was determined from the samples of $D^0 \to h^+h^{'-}$ decays described in sec. 5.1.

Probability density function of the mass term

The different contributions to the background (combinatorial and partially reconstructed heavy-flavors, sec. 3.8.1) are not included in the simulation. Rather than searching *a priori* a physics-motivated model capable to reproduce the mass shapes of the various contributions, we empirically found that an exponential function summed with a constant is adequate in the invariant-mass range $4.850 < m_{\pi^+\pi^-} < 5.800$ GeV/c^2. Each candidate, with invariant $\pi^+\pi^-$-mass $m_{\pi^+\pi^-}$ and momentum imbalance α, contributes to the mass probability density of the background with the following factor:

$$\wp^m_{\rm bck}(m_{\pi^+\pi^-}|\alpha;\vec{\theta}) = \frac{1}{K}\left(e^{c_1 m_{\pi^+\pi^-}} + c_2\right), \tag{6.15}$$

in which K is the normalization constant $K = \int_{4.85}^{5.80}\left(e^{c_1 m_{\pi^+\pi^-}} + c_2\right) dm_{\pi^+\pi^-}$, and c_i are free parameters in the fit. Figure 6.5 shows that the functional form chosen for the background is appropriate.

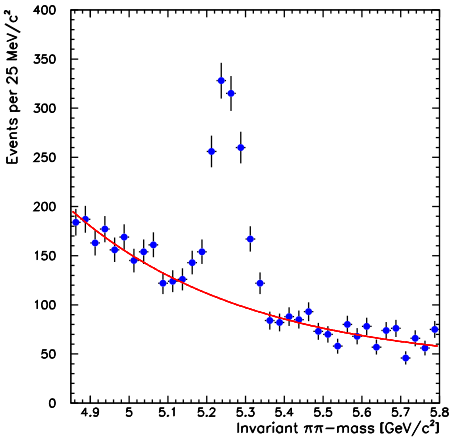

Figure 6.5: Invariant $\pi^+\pi^-$-mass distribution of $B^0_{(s)} \to h^+h^{'-}$ candidates passing the selection of sec. 3.8 (collision data). The function of eq. (6.15) is overlaid (red, solid line).

Probability density function of the momentum term

As for the signal, we used a joint p.d.f. that parameterizes simultaneously the distribution of momentum imbalance (α) and of scalar sum of momenta ($p_{\rm tot}$). The p.d.f. is further split in the density

of the scalar sum, $\wp(p_{\rm tot};\vec{\theta})$, times the conditional density of the momentum imbalance, $\wp(\alpha|p_{\rm tot};\vec{\theta})$. We empirically chose the shapes of both densities, whose parameters were determined in data with a binned, two-dimensional ML fit of the distributions for candidates from mass sidebands. An event with observed scalar sum of momenta $p_{\rm tot}$ and observed momentum-imbalance α contributes to the momentum p.d.f. of the background with the following term:

$$\begin{aligned}\wp^{p}_{\rm bck}(\alpha,p_{\rm tot};\vec{\theta}) &= \wp(p_{\rm tot};\vec{\theta})\times\wp(\alpha|p_{\rm tot};\vec{\theta}) = \\ &\frac{1}{K}\left[1+\left(\frac{p_{\rm tot}-\lambda}{a}\right)^2\right]^{-s} e^{\left[-\nu\ \tan^{-1}\left(\frac{p_{\rm tot}-\lambda}{a}\right)\right]}\times\sum_{m=0}^{6} b_m\alpha^m\left(\frac{p_{\rm tot}-2}{p_{\rm tot}-4}\right)^m \quad (6.16)\end{aligned}$$

where the $p_{\rm tot}$ density is a Pearson type-IV distribution [129], while the α conditional density is a 6th-degree polynomial in the α variable scaled by the usual factor $(p_{\rm tot}-2)/(p_{\rm tot}-4)$.[2] Since the momentum distribution of particles from background is not expected to vary between positively and negatively-charged particles, the shape of the α distribution was forced to be symmetric by setting the odd b terms to zero. The normalization factor K is calculated through a numerical two-dimensional

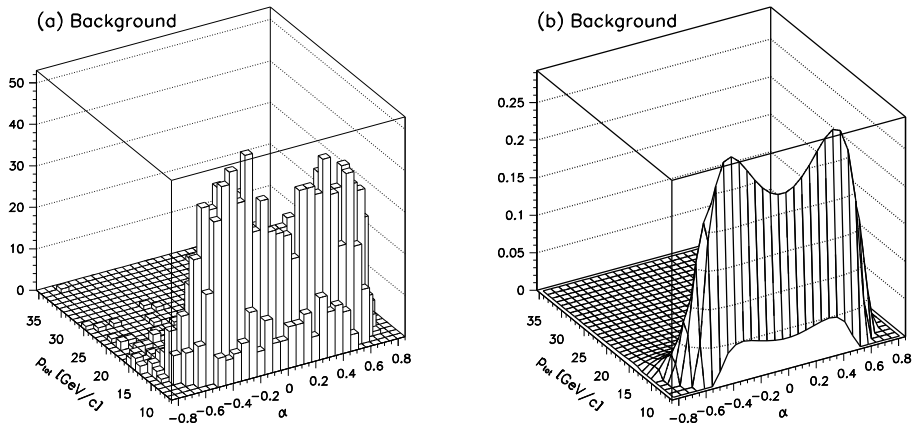

Figure 6.6: Distribution of the scalar sum of momenta as a function of the momentum imbalance in background events (a). Joint $(\alpha, p_{\rm tot})$ template as determined in background events (b).

integration of the p.d.f. in the appropriate domain of α and $p_{\rm tot}$. Figure 6.6(a) shows the distribution of the scalar sum of momenta as a function of the momentum imbalance in background events. A fit to the function of eq. (6.16) yields the momentum template shown in fig. 6.6(b). We checked the agreement between model and data by overlaying the template to the α-distribution of background events, sampled in different $p_{\rm tot}$ ranges (see fig. 6.7).

[2]The Pearson type-IV distribution is useful to approximate a broad class of asymmetric distributions with extensive tails.

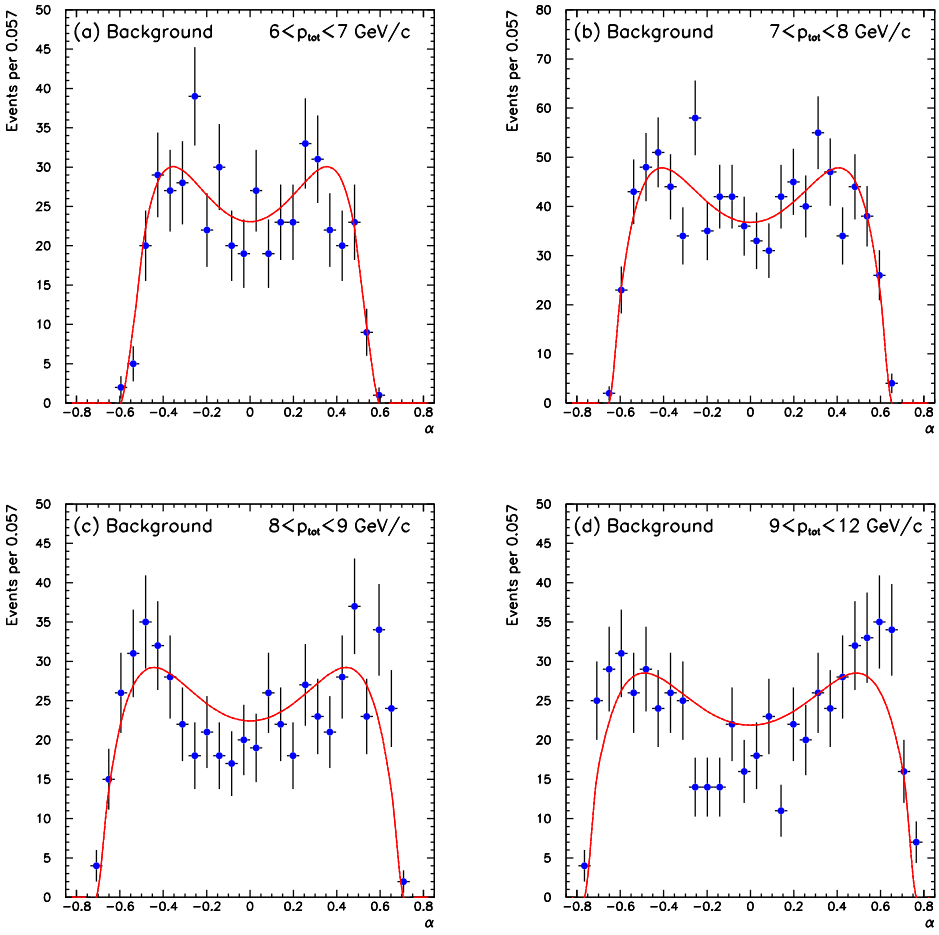

Figure 6.7: Distributions (points with error bars) of the momentum imbalance for background events sampled in four distinct ranges of scalar sum of momenta: $6 < p_{\rm tot} < 7$ GeV/c (a), $7 < p_{\rm tot} < 8$ GeV/c (b), $8 < p_{\rm tot} < 9$ GeV/c (c), and $9 < p_{\rm tot} < 12$ GeV/c (d). The corresponding templates are overlaid (solid line).

Probability density function of the PID term

The PID p.d.f. for background uses the same joint two-particle p.d.f. as for the signal. The term corresponding to each possible pair of particle types (l, m) in background, is weighed by a factor $w_l w_m$. Each weight w_l is proportional to the fractional contribution of particles of type l to the background, and it is a free parameter in the fit. We allowed for independent kaon, proton, electron and pion (or muon) contributions. Muons and pions were not differentiated since their contributions are indistinguishable; the ≈ 1.5 ns dE/dx resolution is insufficient to resolve the difference between their ionization rates, which is inappreciable because of the small difference in mass. This does not affect the signal composition, since muon contamination in the signal peak is negligible, if any, due

to the small rates expected for $B^0_{(s)}$ meson decays in muon pairs [130], and muons from semileptonic heavy-flavor decays don not have a peaking distribution in mass. We used the model of residuals and correlation for kaons and pions described in sec. 5.8, but we lacked the corresponding model for protons and electrons. We assumed that the distribution of the residual for protons (electrons) with proton (electron) mass hypothesis were the same as of pions: $\wp_e(\delta_e) = \wp_{\rm p}(\delta_{\rm p}) = \wp_\pi(\delta_\pi)$. This assumption is valid to a good extent since, in principle, the shapes of residuals (with the correct mass hypothesis) are independent of the particular type of particle. The small differences observed between kaon and pion residuals in fig. 5.23, suggest that we need to associate a systematic uncertainty due to this assumption (see sec. 9.3.2).

A candidate decaying to particles with $\kappa_1^{\rm obs}$ and $\kappa_2^{\rm obs}$ observed "kaonnesses", sum of observed scalar momenta $p_{\rm tot}$, and observed momentum imbalance α, contributes to the PID term of the likelihood of background with the following probability density function:

$$\wp^{\rm PID}_{\rm bck} = \sum_{l,m=e,\pi,K,{\rm p}} w_l w_m [\wp_{l,m}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}|\alpha, p_{\rm tot}; \vec{\theta}) + \wp_{m,l}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}|\alpha, p_{\rm tot}; \vec{\theta})]. \tag{6.17}$$

The explicit expression of $\wp_{l,m}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}|\alpha, p_{\rm tot}; \vec{\theta})$ is shown in eq. (6.13). While in the signal case the l, m indexes run over kaons and pions only, in the above equation they include also protons and electrons. We measured independent fractions of particle species in each of the invariant $\pi^+\pi^-$-mass ranges listed in tab. 6.1. This allows for *a priori* different background compositions in these regions, that account for the distinct sources (combinatorial and mis-reconstructed b-meson decays) outlined in sec. 3.8.1.

Legend	Mass range [GeV/c^2]
Low-mass region	$4.850 < m_{\pi^+\pi^-} < 5.125$
Signal region	$5.125 < m_{\pi^+\pi^-} < 5.400$
High-mass region	$5.400 < m_{\pi^+\pi^-} < 5.800$

Table 6.1: Mass regions with independent background compositions.

6.2.3 The likelihood revisited

Special care is needed in writing a multivariate likelihood, since potentially large fit-biases may arise when the densities of dependent observables are factorized improperly. In our likelihood function, the kinematic and PID terms are coupled by their common dependence on momentum, resulting in the factorization of the p.d.f. schematically shown below:

$$\mathscr{L}_i(\vec{\theta}) \sim \wp^m(m_{\pi^+\pi^-}|\alpha; \vec{\theta}) \times \wp^p(\alpha, p_{\rm tot}; \vec{\theta}) \times \wp^{\rm PID}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}|\alpha, p_{\rm tot}; \vec{\theta}). \tag{6.18}$$

The momentum dependence of the mass term is obvious from sec. 4.2, since the invariant $\pi^+\pi^-$-mass of modes with final states other than two pions varies with the momentum imbalance. As a consequence, the mass term is a conditional density at given α, multiplied by the momentum density. But the momentum affects also the κ observable via the presence of expected dE/dx values in the expression of κ.

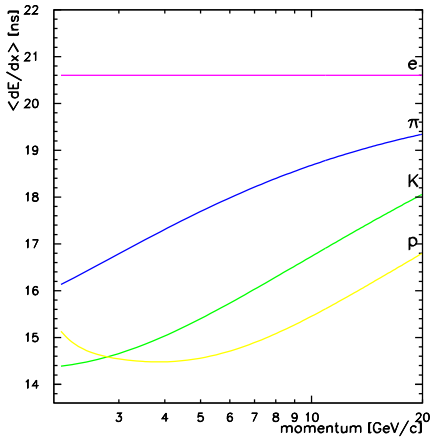

Figure 6.8: Average dE/dx as a function of particle momentum (in logarithmic scale).

In general, the dE/dx information is used to distinguish at least two classes of particles (type a and type b); for each type, the κ observable can be written as a function of the *proper* dE/dx residual (the one evaluated with "true" mass hypothesis):

$$\kappa(a) = \frac{dE/dx_{\text{obs}} - dE/dx_{\pi}}{dE/dx_{K} - dE/dx_{\pi}} = \frac{dE/dx_{\text{obs}} - dE/dx_{a}}{dE/dx_{K} - dE/dx_{\pi}} + \frac{dE/dx_{a} - dE/dx_{\pi}}{dE/dx_{K} - dE/dx_{\pi}} \tag{6.19}$$

$$\kappa(b) = \frac{dE/dx_{\text{obs}} - dE/dx_{\pi}}{dE/dx_{K} - dE/dx_{\pi}} = \frac{dE/dx_{\text{obs}} - dE/dx_{b}}{dE/dx_{K} - dE/dx_{\pi}} + \frac{dE/dx_{b} - dE/dx_{\pi}}{dE/dx_{K} - dE/dx_{\pi}}. \tag{6.20}$$

Under the assumption that distributions of the proper residuals have the same shape, which is true with good approximation, the first terms in the right-hand sides of the above equations are equal for any given momentum. Hence, $\kappa(b)$ is obtained by shifting $\kappa(a)$ by a factor $\frac{dE/dx_{b} - dE/dx_{a}}{dE/dx_{K} - dE/dx_{\pi}}$ which depends only on differences between expected dE/dx values with different mass hypotheses. If these differences were not a function of momentum, i. e., if all curves in fig. 6.8 were parallel, then the κ distribution for any type of particle could be obtained from other shapes with a simple shift. However, by allowing (in background) for contributions of electrons and protons, whose curves are not parallel to the ones relative to pions and kaons, we introduced a momentum dependence in the κ distribution. We therefore wrote the PID density as a conditional probability at given momenta. The components of the likelihood and the external inputs needed to model them are summarized in tab. 6.2.

6.3 Cross-checks of the fit

The fitting code has been extensively tested on ensembles of simulated pseudo-experiments. In this section we report a compilation of the most significant tests.

Term		Template	Source
Signal Mass	$\wp_j^m(m_{\pi^+\pi^-}\|\alpha)$	Analytic	Two-body kinematics
		Data	$m_{B^0}, m_{B_s^0}, J/\psi, \Upsilon \to \mu^+\mu^-$, $D^0 \to h^+h^{'-}$
		PDG	m_K, m_π
Sign. momentum	$\wp_j^p(\alpha, p_{\rm tot})$	MC	
Sign. PID	$\wp_j^{\rm PID}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}\|\alpha, p_{\rm tot})$	Data	$D^0 \to h^+h^{'-}$
Background mass	$\wp_{\rm bck}^m(m_{\pi^+\pi^-}\|\alpha)$	Data	$B^0_{(s)} \to h^+h^{'-}$ sidebands
Bckg. momentum	$\wp_{\rm bck}^p(\alpha, p_{\rm tot})$	Data	$B^0_{(s)} \to h^+h^{'-}$ sidebands
Bckg. PID	$\wp_{\rm bck}^{\rm PID}(\kappa_1^{\rm obs}, \kappa_2^{\rm obs}\|\alpha, p_{\rm tot})$	Data	$D^0 \to h^+h^{'-}$

Table 6.2: Schematic summary of the likelihood components.

6.3.1 Statistical resolution

We studied the features of the fit under several configurations of the most critical parameters, such as background contamination, uncertainty on the absolute scale of masses, and PID performance. We used the simulated signal decays described in sec. 3.6 to reproduce the experimental circumstances of the $B^0_{(s)} \to h^+h^{'-}$ fit. The kinematics of background was simulated by extracting pseudo-random numbers according to the distributions of background events in data. Since the CDF II simulation does not reproduce the dE/dx measurement, this was simulated with a simplified model, which included only pions and kaons with Gaussian dE/dx responses and a constant 1σ separation between them. Half of the pseudo-experiments contained 245 $B^0_{(s)} \to h^+h^{'-}$ decays (low-statistics sample), half contained 2450 decays (high-statistics sample). Input (true) values for the fraction of each decay-mode with respect to the total signal were chosen according to the following realistic values:

- $f_{B^0\to\pi^+\pi^-} = f_{\overline{B}^0\to\pi^+\pi^-} = 7.5\%$;
- $f_{B^0\to K^+\pi^-} = f_{\overline{B}^0\to K^-\pi^+} = 30\%$;
- $f_{B_s^0\to K^-\pi^+} = f_{\overline{B}_s^0\to K^+\pi^-} = 2.5\%$;
- $f_{B_s^0\to K^+K^-} = f_{\overline{B}_s^0\to K^+K^-} = 10\%$.

We fit the simulated data with two configurations: in the *kinematic and PID fit* we used the full likelihood function; in the *kinematic fit* we used only the kinematic portion of the likelihood. This allowed understanding the relative contribution of kinematics and PID information to the separation power. Part of the simulated samples contained 50% fraction of background in the signal region, part contained no background; this allowed studying what fraction of the statistical resolution was dominated by the purity of the signal, and what by the close spacing in mass between signals. We also investigated two configurations of the absolute scale of masses: one with scale free to vary in the fit, and the other one in which the scale was constrained to a Gaussian distribution with 3 MeV/c^2 standard deviation, chosen as a conservative upper bound on the understanding of the global mass-scale of the CDF II detector.

Kinematic fit

In the kinematic fit, only the $m_{\pi^+\pi^-}$, α, and $p_{\rm tot}$ observables were used and the densities $\wp^{\rm PID}$ in eqs. (6.5) and (6.14) were set to constants. Tables 6.3 and 6.4 report the results of the kinematic fits applied to the pseudo-experiments. The symbol A_{CP} indicates the following asymmetry of decay-rates: $\frac{f_{\overline{B}^0\to K^-\pi^+}-f_{B^0\to K^+\pi^-}}{f_{\overline{B}^0\to K^-\pi^+}+f_{B^0\to K^+\pi^-}}$. Each row corresponds to a different fit, applied to samples with no background (first two rows), and with background (last two rows).

In general, the kinematic information alone appears insufficient to distinguish the $B^0\to\pi^+\pi^-$ from the $B^0_s\to K^+K^-$ decays. This effect was expected, since their contributions almost perfectly overlap in the invariant $\pi^+\pi^-$-mass spectrum (see fig. 4.1). However, the estimate of the summed $B^0\to\pi^+\pi^-$ and $B^0_s\to K^+K^-$ yields ($\hat{f}_{B^0_{(s)}\to h^+h^-}$) is stable around the true value in all the fit configurations, and well separated from the other contributions even in presence of a realistic background. The statistical resolution in the low-statistics samples is significantly degraded when the mass scale is floating (see, in particular, the resolution on the estimate of A_{CP}). In the high-statistics sample, the available information is sufficient to determine the absolute mass-scale with about 2 MeV/c^2 uncertainty. Adding the background affects the statistical resolution in both samples, with a degradation ranging from 20% to more than 100%, depending on the chosen parameter; larger worsening affects lower-yield modes, as the $B^0_s\to K^-\pi^+$, in which a small fluctuation of the number of events assigned to the background may result in significant changes of the small signal yield. A row-to-row comparison between tabs. 6.3 and 6.4 shows the expected $1/\sqrt{N}$ scaling of the resolutions with statistics.

Kinematic and PID fit

The full fit uses all five discriminating observables of sec. 6.1 and the complete likelihood shown in eqs. (6.5) and (6.14). Tables 6.5 and 6.6 show the results of the full fit applied to samples of pseudo-experiments with low and high statistics using different configurations of background and absolute mass-scale. Adding the dE/dx information significantly improves the separation between $B^0\to\pi^+\pi^-$ and $B^0_s\to K^+K^-$ modes: the statistical uncertainties on the estimate of these fractions are reduced by factors of three to four with respect to the kinematic fit. The uncertainties on other parameters also show 10%–40% improvements; the only exception is the $B^0_s\to K^-\pi^+$ fraction, which shows little, if any, improvement because its small value is dominated by background fluctuations. Conversely, we found that the resolution on A_{CP} is less affected by the background than other parameters, especially at high statistics. Our previous considerations on the effects of the mass scale and of the background hold with the full fit as well, even though with reduced magnitudes.

We also verified that the full fit is able to properly estimate global shift of masses of up to 30 MeV/c^2, without introducing any bias, or degrading the estimates of the fractions. We also checked the ideal performance with a perfect PID: a factor of two improvement in resolution of each signal fraction would be obtained if the kaon-pion separation were equivalent to six Gaussian standard deviations.

m constr. [MeV/c^2]	$\hat{b}$	$\hat{f}_{B^0\to K^+\pi^-}$	$\hat{f}_{B^0\to\pi^+\pi^-}$	$\hat{f}_{B^0_s\to K^+K^-}$	$\hat{f}_{B^0_{(s)}\to h^+h^-}$	$\hat{f}_{B^0_s\to K^-\pi^+}$	$\hat{A}_{\mathsf{CP}}$	m scale [MeV/c^2]
3	–	0.578 ± 0.063	0.379 ± 0.184	-0.049 ± 0.172	0.330 ± 0.063	0.093 ± 0.026	-0.052 ± 0.192	-2.8 ± 2.5
Free	–	0.462 ± 0.086	0.433 ± 0.186	-0.013 ± 0.190	0.421 ± 0.074	0.117 ± 0.035	-0.022 ± 0.233	-10.1 ± 4.9
3	0.515 ± 0.053	0.620 ± 0.091	0.210 ± 0.196	0.108 ± 0.181	0.318 ± 0.102	0.062 ± 0.055	-0.319 ± 0.279	0 ± 3.4
Free	0.519 ± 0.054	0.323 ± 0.105	0.409 ± 0.197	0.176 ± 0.186	0.584 ± 0.108	0.093 ± 0.064	-0.418 ± 0.541	-18.1 ± 6.1

Table 6.3: Results of the kinematic fit on a sample of 254 simulated signal events. In this fit no PID information is used. The fraction of background in the signal region is zero (fixed in the fit), or 50% (floating in the fit) and the scale of masses is either constrained to a Gaussian distribution with 3.0 MeV/c^2 r.m.s. or free. Input values of the fractions are the following: $f_{B^0\to\pi^+\pi^-} = 15\%$, $f_{B^0\to K^+\pi^-} = f_{\overline{B}^0\to K^-\pi^+} = 30\%$, $f_{B^0_s\to K^+K^-} = 20\%$, and $f_{B^0_s\to K^-\pi^+} = f_{\overline{B}^0_s\to K^+\pi^-} = 2.5\%$.

m constr. [MeV/c^2]	$\hat{b}$	$\hat{f}_{B^0\to K^+\pi^-}$	$\hat{f}_{B^0\to\pi^+\pi^-}$	$\hat{f}_{B^0_s\to K^+K^-}$	$\hat{f}_{B^0_{(s)}\to h^+h^-}$	$\hat{f}_{B^0_s\to K^-\pi^+}$	$\hat{A}_{\mathsf{CP}}$	m scale [MeV/c^2]
3	–	0.592 ± 0.026	0.245 ± 0.050	0.101 ± 0.043	0.347 ± 0.026	0.062 ± 0.008	0.060 ± 0.068	-0.4 ± 1.3
Free	–	0.590 ± 0.029	0.246 ± 0.051	0.101 ± 0.043	0.347 ± 0.026	0.062 ± 0.009	0.061 ± 0.068	-0.4 ± 1.4
3	0.512 ± 0.017	0.645 ± 0.038	0.198 ± 0.066	0.116 ± 0.055	0.314 ± 0.036	0.041 ± 0.018	0.042 ± 0.086	3.2 ± 2.0
Free	0.518 ± 0.018	0.680 ± 0.042	0.162 ± 0.068	0.123 ± 0.054	0.286 ± 0.039	0.034 ± 0.019	0.038 ± 0.083	5.4 ± 2.3

Table 6.4: Results of the kinematic fit on a sample of 2540 simulated signal events. In this fit no PID information is used. The fraction of background in the signal region is zero (fixed in the fit), or 50% (floating in the fit) and the resolution on the scale of masses is either constrained to a Gaussian distribution with 3.0 MeV/c^2 r.m.s. or free. Input values of the fractions are the following: $f_{B^0\to\pi^+\pi^-} = 15\%$, $f_{B^0\to K^+\pi^-} = f_{\overline{B}^0\to K^-\pi^+} = 30\%$, $f_{B^0_s\to K^+K^-} = 20\%$, and $f_{B^0_s\to K^-\pi^+} = f_{\overline{B}^0_s\to K^+\pi^-} = 2.5\%$.

m constr. [MeV/c^2]	$\hat{b}$	$\hat{f}_{B^0\to K^+\pi^-}$	$\hat{f}_{B^0\to\pi^+\pi^-}$	$\hat{f}_{B^0_s\to K^+K^-}$	$\hat{f}_{B^0_{(s)}\to h^+h^-}$	$\hat{f}_{B^0_s\to K^-\pi^+}$	$\hat{A}_{CP}$	m scale [MeV/c^2]
3	–	0.570 ± 0.057	0.193 ± 0.050	0.152 ± 0.047	0.345 ± 0.058	0.086 ± 0.025	0.046 ± 0.133	-3.1 ± 2.3
Free	–	0.494 ± 0.071	0.218 ± 0.054	0.187 ± 0.052	0.405 ± 0.065	0.101 ± 0.029	0.075 ± 0.147	-8.2 ± 3.8
3	0.535 ± 0.049	0.577 ± 0.081	0.193 ± 0.072	0.198 ± 0.066	0.391 ± 0.090	0.032 ± 0.053	-0.162 ± 0.178	-1.5 ± 3.0
Free	0.529 ± 0.050	0.374 ± 0.105	0.287 ± 0.080	0.281 ± 0.075	0.569 ± 0.106	0.057 ± 0.059	-0.163 ± 0.247	-14.2 ± 5.9

Table 6.5: Results of the full fit on a sample of 254 simulated signal events. The fraction of background in the signal region is zero (fixed in the fit), or 50% (floating in the fit) and the scale of masses is either constrained to a Gaussian distribution with 3.0 MeV/c^2 r.m.s. or free. Input values of the fractions are the following: $f_{B^0\to\pi^+\pi^-} = 15\%$, $f_{B^0\to K^+\pi^-} = f_{\overline{B}^0\to K^-\pi^+} = 30\%$, $f_{B^0_s\to K^+K^-} = 20\%$, and $f_{B^0_s\to K^-\pi^+} = f_{\overline{B}^0_s\to K^+\pi^-} = 2.5\%$.

m constr. [MeV/c^2]	$\hat{b}$	$\hat{f}_{B^0\to K^+\pi^-}$	$\hat{f}_{B^0\to\pi^+\pi^-}$	$\hat{f}_{B^0_s\to K^+K^-}$	$\hat{f}_{B^0_{(s)}\to h^+h^-}$	$\hat{f}_{B^0_s\to K^-\pi^+}$	$\hat{A}_{CP}$	m scale [MeV/c^2]
3	–	0.604 ± 0.020	0.166 ± 0.016	0.170 ± 0.015	0.336 ± 0.020	0.060 ± 0.008	0.001 ± 0.042	0.2 ± 1.0
Free	–	0.605 ± 0.021	0.165 ± 0.016	0.170 ± 0.015	0.336 ± 0.020	0.060 ± 0.008	0.001 ± 0.042	0.2 ± 1.0
3	0.521 ± 0.015	0.644 ± 0.028	0.150 ± 0.022	0.177 ± 0.019	0.327 ± 0.027	0.029 ± 0.016	-0.032 ± 0.050	3.2 ± 1.4
Free	0.523 ± 0.016	0.657 ± 0.029	0.143 ± 0.022	0.174 ± 0.019	0.317 ± 0.028	0.026 ± 0.017	-0.032 ± 0.049	4.1 ± 1.6

Table 6.6: Results of the full fit on a sample of 2540 simulated signal events. The fraction of background in the signal region is zero (fixed in the fit), or 50% (floating in the fit) and the resolution on the scale of masses is either constrained to a Gaussian distribution with 3.0 MeV/c^2 r.m.s. or free. Input values of the fractions are the following: $f_{B^0\to\pi^+\pi^-} = 15\%$, $f_{B^0\to K^+\pi^-} = f_{\overline{B}^0\to K^-\pi^+} = 30\%$, $f_{B^0_s\to K^+K^-} = 20\%$, and $f_{B^0_s\to K^-\pi^+} = f_{\overline{B}^0_s\to K^+\pi^-} = 2.5\%$.

6.3.2 Bias and variance of the ML estimate

Maximum likelihood estimates based on samples of finite statistics often suffer from statistical biases. To investigate these effects and the stability of the minimization code, which might experience difficulties in finding the proper minimum independently of the initial values of parameters when dealing with a fairly complicate p.d.f., we studied the distributions of the pull values of the fit. The pull value of each fit parameter θ_i estimated is defined as follows:

$$\mathcal{P}(\theta_i) = \frac{\hat{\theta}_i - \theta_i}{\hat{\sigma}_{\hat{\theta}_i}}, \tag{6.21}$$

in which $\hat{\theta}_i$ is the estimate of the parameter, and $\hat{\sigma}_{\hat{\theta}_i}$ is the estimate of its uncertainty. We studied the pull values using an ensemble of 280 pseudo-experiments that simulated the experimental circumstance of the fit on $B^0_{(s)} \to h^+h^{'-}$ data. Each pseudo-experiment consisted of the simulated distributions of the five discriminating observables (mass, momentum imbalance, scalar sum of momenta, and dE/dx of both tracks) corresponding to 4900 total events. The distributions for each signal mode and for background were generated according to the corresponding likelihood term using a pseudo-random number generator.[3] The fractions of each signal mode and of background fluctuated from sample to sample according to a multinomial distribution whose mean value was the set of true parameters $\vec{\theta}$. This choice properly accounts for the statistical fluctuations of signal and background fractions among samples, while keeping constrained the total number of events. We fit the composition of all pseudo-experiments using the same likelihood as for the data; we then derived the pull-value distributions of the relevant physics quantities from the estimated parameters and uncertainties (see fig. 6.9).

The pull-values are Gaussian-distributed with approximately unit variance and negligible bias (see tab. 6.7). The distribution of $\mathscr{L}$ in the proximity of its minimum is, therefore, approximately Gaussian for each estimated parameter, which is not obvious with finite samples and complicated probability densities. This guarantees the estimated uncertainty of each parameter $\hat{\sigma}_{\hat{\theta}_i}$ to be such that the range $[\hat{\theta}_i - \hat{\sigma}_{\hat{\theta}_i}, \hat{\theta}_i + \hat{\sigma}_{\hat{\theta}_i}]$ is the shortest that contains the true value θ_i with about 68% probability. Note that the pull values corresponding to ratios of fractions are not necessarily Gaussian-distributed because such ratios are not primary fit parameters, but they are ratios of fit parameters (the fit determines the absolute fractions). Since the distribution of the ratio of two Gaussian-distributed quantities has infinite variance, the Gaussian assumption for the distribution of pull values of ratios is just a first-order approximation.

Fit parameter	Pull mean	Pull standard deviation	χ^2/d.o.f.
$f_{B^0\to\pi^+\pi^-}/f_{B^0\to K^+\pi^-}$	-0.04 ± 0.06	0.94 ± 0.05	16.4/25
A_{CP}	-0.01 ± 0.06	0.92 ± 0.06	25.1/25
$f_{B^0_s\to K^+K^-}/f_{B^0\to K^+\pi^-}$	0.01 ± 0.06	0.99 ± 0.05	15.3/25
$f_{B^0_s\to K^-\pi^+}/f_{B^0\to K^+\pi^-}$	-0.07 ± 0.07	0.93 ± 0.05	25.8/25

Table 6.7: Results of the Gaussian fit of the pull-value distributions.

[3] We used the RANLUX generator, from the CERN (Conseil Europeén pour la Recherche Nucleaire) software libraries, with "luxury-level" three, interfaced with a C code compiled with GCC on a commercial LINUX PC.

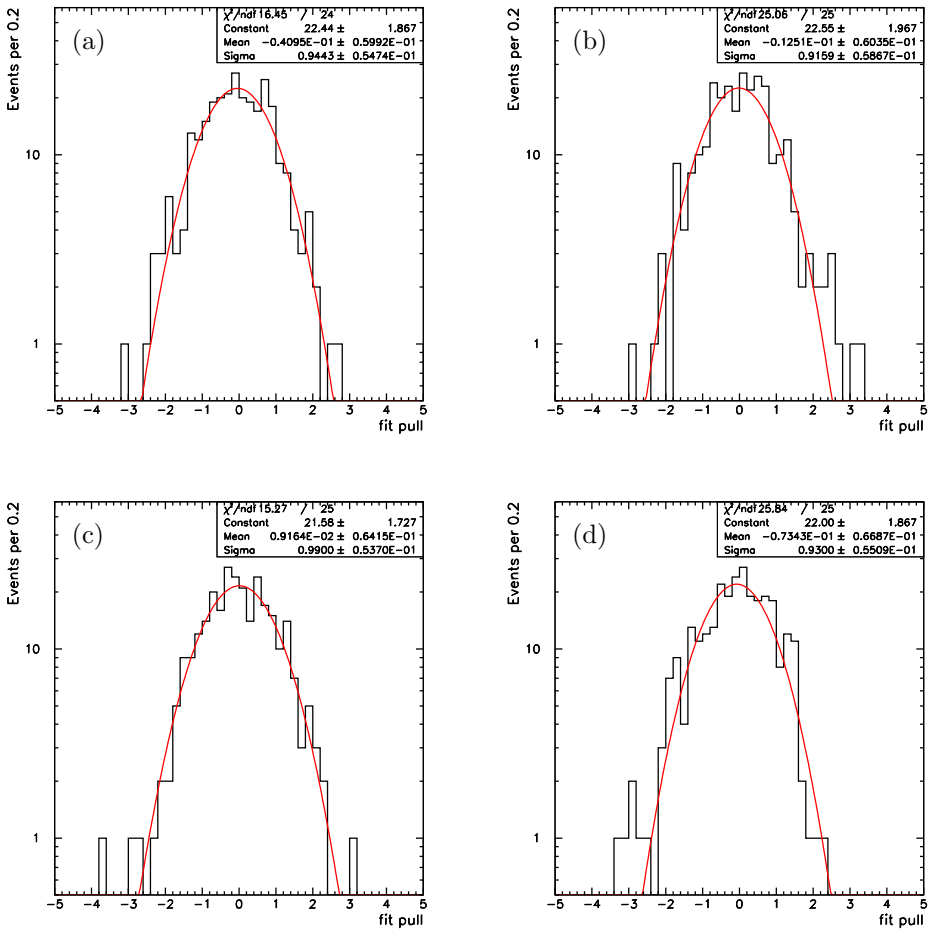

Figure 6.9: Distribution of the pull values for the relevant fit parameters: (a) $f_{B^0\to\pi^+\pi^-}/f_{B^0\to K^+\pi^-}$, (b) $A_{\mathsf{CP}}(B^0\to K^+\pi^-)$, (c) $f_{B_s^0\to K^+K^-}/f_{B^0\to K^+\pi^-}$, and (d) $f_{B_s^0\to K^-\pi^+}/f_{B^0\to K^+\pi^-}$. Results of χ^2-fits to Gaussian functions are overlaid in red.

Chapter 7

Result of the sample composition measurement

After checking the ML fit in simulated samples, we applied it to data for estimating the composition of the $B^0_{(s)} \to h^+h^{'-}$ sample. This chapter discusses the results of this measurement.

7.1 Fit results

After all checks of the likelihood, the fit was applied to the $B^0_{(s)} \to h^+h^{'-}$ sample passing the final selection (see sec. 3.8). We considered only candidates whose discriminating observables satisfied the following conditions.

- $4.850 < m_{\pi^+\pi^-} < 5.800$ GeV/c^2: this mass range contains the whole $B^0_{(s)} \to h^+h^{'-}$ signal, and allows a proper extrapolation of the background shape underneath the signal peak;
- $-0.85 < \alpha < 0.85$: this imbalance range excludes boundary regions where the model of the corresponding p.d.f. becomes inaccurate and may lead to undesired systematic uncertainties. The requirement removed only a few events, all lying outside the signal region;
- $p_{\rm tot} > 5.75$ GeV/c: we chose to exclude a few candidates with lower $p_{\rm tot}$ values, because the likelihood function became negative for these candidates, due to the steep turn-on of the corresponding $p_{\rm tot}$-density;
- $-10 < \kappa_1, \kappa_2 < 10$: candidates with extremely unlikely values of the observed dE/dx are excluded to reduce the probability of contamination from fake tracks or tracks with corrupted dE/dx information. In this case there are fewer than ten excluded events.

The total number of events is $N = 4897$. Given the production fractions and branching fractions involved (see sec. 4.1), we expect sizable contributions from two known modes, $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, and two not yet observed modes, $B^0_s \to K^+K^-$ and $B^0_s \to K^-\pi^+$. Their fractions with respect to the total signal were left free to vary in the fit. Contributions of the rarer $B^0 \to K^+K^-$

and $B_s^0 \to \pi^+\pi^-$ modes were not considered, because the existing upper limits (at the 90% confidence level) on their branching fractions, $\mathcal{B}(B^0 \to K^+K^-) < 3.7 \times 10^{-7}$ and $\mathcal{B}(B_s^0 \to \pi^+\pi^-) < 2.1 \times 10^{-4}$ [18], along with typical theoretical predictions, $\mathcal{B}(B^0 \to K^+K^-) = (0.01\text{–}0.21) \times 10^{-6}$ and $\mathcal{B}(B_s^0 \to \pi^+\pi^-) = (0.027\text{–}0.155) \times 10^{-6}$ [24], suggest negligible contributions in our sample. Similar considerations apply to $\Lambda_b^0 \to \mathrm{p}\pi^-$ and $\Lambda_b^0 \to \mathrm{p}K^-$ modes, whose contribution to our sample is excluded by an independent analysis [131]. The fractions of the $B_s^0 \to \pi^+\pi^-$ and $B^0 \to K^+K^-$ modes were fixed at zero and were determined in a separate fit (see sec. 7.2).

We numerically minimized the quantity $-2\ln(\mathscr{L})$ using the MINUIT package of the CERN software libraries [126]. Technical details of the numerical minimization are reported in appendix B. The minimization converges in about 15 minutes on a commercial LINUX computer, with 2.0 GHz processor and 1 GB RAM, yielding the fit results, the estimated minimum value of $-2\ln(\mathscr{L})$, and the positively-defined covariance matrix. The result of the fit is shown in tab. 7.1, and the corresponding correlation matrix is shown in sec. 7.1.3. Table 7.2 summarizes only the relevant physics quantities determined by the fit.

7.1.1 Signal composition

The fit determined a total $B_{(s)}^0 \to h^+h^{'-}$ signal yield of about 900 events, in agreement with the simple estimate from the invariant $\pi^+\pi^-$-mass fit shown in sec. 3.8. Significant signals are seen for $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ modes, which have already been established at the B-Factories [29], but so far were never observed at a hadron collider. In addition, we found about 240 decays of the previously unobserved $B_s^0 \to K^+K^-$ mode. This is the first observation of a B_s^0 meson decaying into two pseudo-scalar mesons. No evidence is obtained for the rarer $B_s^0 \to K^-\pi^+$ mode, nor for an asymmetry in $B^0 \to K^+\pi^-$ decays, which is consistent with zero within its 8% estimated statistical uncertainty. The estimated statistical uncertainties on the fractions are around 3% for each decay mode.

Estimated ratios of yields

Although the fit determines the individual fraction of each mode, we combined these fractions in the ratios of yields shown in the upper portion of tab. 7.1 to extract the final measurements of ratios of branching fractions. We aimed at the best statistical resolution on the relative branching fractions by normalizing the yields of the $B^0 \to \pi^+\pi^-$, $B_s^0 \to K^+K^-$, and $B_s^0 \to K^-\pi^+$ modes to the yield of the $B^0 \to K^+\pi^-$ reference mode. This reference mode was chosen because it has the largest yield (i. e., introduces the smallest relative statistical uncertainty in the ratios), and because its branching fraction is measured with high precision at the B-Factories [29]. We quote also the $\frac{B^0 \to \pi^+\pi^-}{B_s^0 \to K^+K^-}$ ratio of yields, although it is just a different expression for the $\frac{B_s^0 \to K^+K^-}{B^0 \to K^+\pi^-}$ ratio, because it allows a convenient theoretical interpretation (see sec. 10.1).

7.1.2 Background composition

The estimate of the background composition and of its mass shape are shown in the lower portion of tab. 7.1. Evidence for pion (and muon), kaon, proton, and electron contributions is found along the

Quantity	Fit estimate	MINUIT code
$\hat{f}_{B^0\to\pi^+\pi^-}$	0.134 ± 0.030	1
$\hat{f}_{B^0\to K^+\pi^-}$	0.600 ± 0.034	2
$\left(\hat{f}_{\overline{B}^0\to K^-\pi^+} - \hat{f}_{B^0\to K^+\pi^-}\right) / \left(\hat{f}_{\overline{B}^0\to K^-\pi^+} + \hat{f}_{B^0\to K^+\pi^-}\right)$	-0.022 ± 0.078	3
$\hat{f}_{B_s^0\to K^-\pi^+}$	0.003 ± 0.028	4
$\hat{f}_{B_s^0\to K^+K^-}$	0.262 ± 0.035	–
$\hat{f}_{B_s^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.437 ± 0.074	–
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.224 ± 0.057	–
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B_s^0\to K^+K^-}$	0.513 ± 0.145	–
$\hat{f}_{B_s^0\to K^-\pi^+}/\hat{f}_{B^0\to K^+\pi^-}$	0.005 ± 0.046	–
$\hat{b}$	0.816 ± 0.009	7
$\hat{w}_{\pi,\mu}$ (low-mass region)	0.550 ± 0.017	8
$\hat{w}_{\pi,\mu}$ (signal region)	0.474 ± 0.027	9
$\hat{w}_{\pi,\mu}$ (high-mass region)	0.463 ± 0.021	10
$\hat{w}_K$ (low-mass region)	0.303 ± 0.036	–
$\hat{w}_K$ (signal region)	0.350 ± 0.057	–
$\hat{w}_K$ (high-mass region)	0.264 ± 0.051	–
$\hat{w}_e$ (low-mass region)	0.015 ± 0.005	11
$\hat{w}_e$ (signal region)	0.019 ± 0.006	12
$\hat{w}_e$ (high-mass region)	0.027 ± 0.007	13
$\hat{w}_{\mathrm{p}}$ (low-mass region)	0.132 ± 0.028	14
$\hat{w}_{\mathrm{p}}$ (signal region)	0.157 ± 0.044	15
$\hat{w}_{\mathrm{p}}$ (high-mass region)	0.246 ± 0.043	16
$\hat{c}_1$ (background slope)	-2.296 ± 0.415	18
$\hat{c}_2$ (background constant)	12.397 ± 21.967	19

Table 7.1: Fit results. Signal (background) related quantities are reported in the upper (lower) section. The last column reports the legend to convert the MINUIT coding of the fit parameters into physics quantities for interpreting the correlation matrix shown at pag. 160; the missing codes refer to parameters not being part of the set of primary fit parameters ($\vec{\theta}$). C-conjugate modes are implied except for the parameter in the third row.

whole mass range. As expected, the pions (contaminated by the muons) dominate the background composition. Their 50% fraction is estimated with $\approx 2\%$ statistical uncertainty. A part of the $\approx 30\%$ fraction of kaons is presumably due to two physics sources: displaced tracks from two uncorrelated heavy-flavor decays (i. e., one track from each one of the two heavy-flavors in the event), and tracks from the same, partially-reconstructed multi-body b-hadron decay. While the former contributes to the whole mass-spectrum, the latter is enhanced in the signal region and in its lower-mass side, as confirmed by the corresponding larger estimated kaon fractions. The $\approx$ 5% accuracy on the kaon fraction is worse than the accuracy for pions, because of the correlation between kaon and proton components, due to limited dE/dx separation between them. It is beyond the scope of this measurement to fully understand the source of the large fraction of protons found in the background.

However, decays of baryons, which are produced both in the $p\bar{p}$ interaction and in secondary nuclear interactions in the detector material, along with residual protons from beam-gas interactions, have the kinematic features to significantly contribute.[1] In addition, a small ($\approx 2\%$), but statistically significant (3σ–4σ) electron component is present in the whole mass range, presumably due to semi-leptonic decays of heavy flavors.

Mode	Estimated fraction [%]	Estimated yield
$B^0 \to \pi^+\pi^- + \overline{B}^0 \to \pi^+\pi^-$	13.4 ± 3.0	121 ± 27
$B_s^0 \to K^-\pi^+ + \overline{B}_s^0 \to K^+\pi^-$	0.0 ± 2.8	3 ± 25
$B_s^0 \to K^+K^- + \overline{B}_s^0 \to K^+K^-$	26.2 ± 3.5	237 ± 32
$B^0 \to K^+\pi^- + \overline{B}^0 \to K^-\pi^+$	60.0 ± 3.4	542 ± 31
$B^0 \to K^+\pi^-$	30.7 ± 2.9	277 ± 28
$\overline{B}^0 \to K^-\pi^+$	29.4 ± 2.9	265 ± 28

Table 7.2: Physics results from the fit. C-conjugate modes are not implied.

7.1.3 Correlations

We studied in detail the correlation matrix (shown below) to search for possible large correlations that may have suggested a better choice of the fit parameters. The matrix reports the correlation coefficients, defined as $\rho_{ij} = \mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j)/\hat{\sigma}_{\hat{\theta}_i}\hat{\sigma}_{\hat{\theta}_j}$, in which $\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j)$ is the off-diagonal element of the estimated covariance matrix of the fit. The conversion from the MINUIT coding of the fit parameters to the physical quantities is shown in last column of tab. 7.1.

```
 PARAMETER  CORRELATION COEFFICIENTS
 NO.       1     2     3     4     7     8     9    10    11    12    13    14    15    16    18    19
  1    1.000
  2   -0.411 1.000
  3    0.010 0.004 1.000
  4   -0.344-0.096 0.004 1.000
  7    0.064 0.067-0.023-0.387 1.000
  8    0.000 0.000 0.000 0.000 0.000 1.000
  9   -0.301-0.016 0.007-0.026 0.020 0.000 1.000
 10    0.000 0.000 0.000 0.000 0.000 0.000 0.000 1.000
 11    0.000 0.000 0.000 0.000 0.000-0.397 0.000 0.000 1.000
 12   -0.035 0.010 0.000-0.001-0.037 0.000-0.297 0.000 0.000 1.000
 13    0.000 0.000 0.000 0.000 0.000 0.000 0.000-0.445 0.000 0.000 1.000
 14    0.000 0.000 0.000 0.000 0.000 0.283 0.000 0.000-0.077 0.000 0.000 1.000
 15    0.027-0.007 0.020-0.008-0.064 0.000 0.269 0.000 0.000-0.068 0.000 0.000 1.000
 16    0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.277 0.000 0.000-0.087 0.000 0.000 1.000
 18    0.034 0.047-0.013-0.232 0.358 0.000 0.013 0.000 0.000-0.023 0.000 0.000-0.039 0.000 1.000
 19    0.034 0.050-0.013-0.236 0.359 0.000 0.013 0.000 0.000-0.023 0.000 0.000-0.039 0.000 0.998 1.000
```

Concerning the physical fit parameters (i. e., pars. 1–4), the largest correlation, about 40%, is observed between the estimated fractions of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ decays (pars. 1 and 2). Although the distributions in $\pi^+\pi^-$-mass of these modes are not completely overlapping, the

[1] Previous studies from CDF Run I showed that beam-gas interactions produce a sizable excess of protons with large impact parameters, which are prominent in the lower (0.5 GeV/c) part of the transverse-momentum spectrum [132].

estimated $B^0 \to \pi^+\pi^-$ fraction is about 20% of the $B^0 \to K^+\pi^-$ fraction. The fit has limited power in separating the bulk of the $B^0 \to \pi^+\pi^-$ from the higher-mass tail of the $B^0 \to K^+\pi^-$ mode, also because the available dE/dx separation is limited, in this case, to just one track per candidate. A similar reason explains the $\approx 35\%$ correlation between $B_s^0 \to K^-\pi^+$ and $B^0 \to K^+\pi^-$ fractions (pars. 2 and 4), in which a zero observed fraction of $B_s^0 \to K^-\pi^+$ decays is contaminated by the tails of the dominant $B^0 \to K^+\pi^-$ mode, with no separation by dE/dx. The $\approx 30\%$ correlation observed between the $B^0 \to \pi^+\pi^-$ fraction and the fraction of background pions underneath the signal (pars. 1 and 9) occurs because the fit tends to reabsorb into the $B^0 \to \pi^+\pi^-$ fraction any fluctuation of the pion fraction in background. The observed trend to small, negative correlations between fractions of signal modes is due to their constraint to have unit sum (see eq. (6.5)).

Concerning the other fit parameters, the largest correlation ($\approx 100\%$) is observed between the slope of the exponential (par. 18) and the constant term (par. 19) of the invariant $\pi^+\pi^-$-mass shape of background. Such a large value, together with an estimated constant term compatible with zero, indicates that our choice for these parameters is not optimal; however, this has no effect on the measurement of the fractions. A $\approx 45\%$ correlation is observed between the estimated fractions of electrons and pions in the background at the high-mass side of the signal (pars. 10 and 13). The fit barely distinguishes a 2%–3% fraction of electrons from the dominant ($\approx 50\%$) pion component which, at higher momenta, have similar dE/dx response. However, not only this effect is limited in size, but it does not affect any of the physics measurements. All other correlations are minor.

7.1.4 Distribution of errors

The distributions of the estimated uncertainties and their correlations were investigated by scanning the functional form of the likelihood in the proximity of its maximum; only the results are briefly reported below, while a description of the method is given in appendix B. This approximate, graphic method exploits the fundamental property for the variance of an estimator, the Minimum Variance Bound. In case of sufficiently regular likelihood functions and large data samples, the logarithm of the likelihood function admits the following, second-order Taylor-expansion in the vicinity of its maximum $\vec{\theta} \approx \vec{\hat{\theta}}$:

$$\ln[\mathscr{L}(\vec{\theta})] \simeq \ln[\mathscr{L}(\vec{\hat{\theta}})] - \frac{1}{2}\left[(\vec{\theta})^T \mathbf{V^{-1}} \vec{\theta}\right], \tag{7.1}$$

in which $(\vec{\theta})^T$ is the transposed vector of fit parameters and $\mathbf{V^{-1}}$ is the inverse of the covariance matrix. This implies that $\ln(\mathscr{L})$, at least locally around its maximum, is approximated by an n-dimensional ellipsoid, where n is the multiplicity of free parameters in the fit. The likelihood shape therefore resembles a multi-Gaussian function. In the simple, two-dimensional case,

$$\ln[\mathscr{L}(\vec{\theta})] \simeq \ln[\mathscr{L}(\vec{\hat{\theta}})] - \frac{1}{2(1-\rho)}\left[\left(\frac{\theta_1 - \hat{\theta}_1}{\hat{\sigma}_{\hat{\theta}_1}}\right)^2 + \left(\frac{\theta_2 - \hat{\theta}_2}{\hat{\sigma}_{\hat{\theta}_2}}\right)^2 - 2\rho\left(\frac{\theta_1 - \hat{\theta}_1}{\hat{\sigma}_{\hat{\theta}_1}}\right)\left(\frac{\theta_2 - \hat{\theta}_2}{\hat{\sigma}_{\hat{\theta}_2}}\right)\right], \tag{7.2}$$

where the contour $\ln[\mathscr{L}(\vec{\theta})] \simeq \ln[\mathscr{L}(\vec{\hat{\theta}})] - s$ defines an ellipse centered at the estimates, and in which the ratio between the main axes depends on the estimated correlation coefficient. We studied two-dimensional contours of constant likelihood around the maximum to detect any deviation of the estimated uncertainties from the Gaussian behavior (see fig. 7.1). Each contour is obtained by varying

a pair of highly-correlated parameters, while keeping all other parameters fixed at the values that maximize the likelihood. Samplings of the contours are obtained with the MNCONTOUR option of the MINUIT program (see appendix B). No significant deviation from the expected ellipsoid shape is

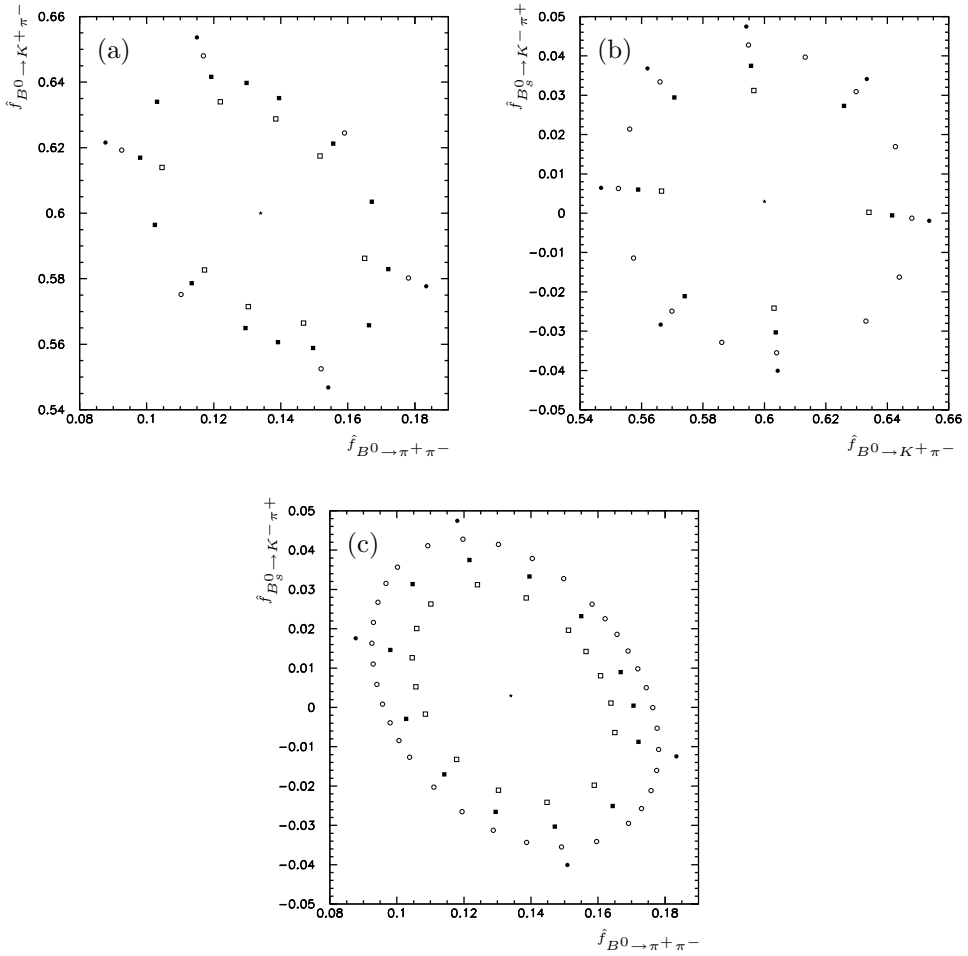

Figure 7.1: Samplings of the contour of constant likelihood $\ln(\mathscr{L}) = \ln[\mathscr{L}(\vec{\hat{\theta}})] - s$ in the proximity of the ML estimate (indicated with a star) in the following two-dimensional planes of parameters: $(f_{B^0\to\pi^+\pi^-}, f_{B^0\to K^+\pi^-})$ (a), $(f_{B^0\to K^+\pi^-}, f_{B^0_s\to K^-\pi^+})$ (b), and $(f_{B^0_s\to K^-\pi^+}, f_{B^0\to\pi^+\pi^-})$ (c). The contours correspond to deviations of the parameters by 1σ (open squares), 1.5σ (filled squares), 2σ (open circles), and 2.5σ (filled circles) from their ML estimates.

found.

7.1.5 Fit quality

The principle of maximum likelihood does not directly provide a method for testing the goodness-of-fit. Different approaches have been proposed in literature for testing the quality of unbinned likelihood

fits, but none has been proved rigorously correct [133]. However, a standard, qualitative method to obtain a visual test of the fit quality consists in comparing the distributions in data with the joint p.d.f. that corresponds to the likelihood function evaluated with the maximizing set of parameters $\hat{\vec{\theta}}$. In practice, for each discriminating observable x, the functional form of the p.d.f. (evaluated with $\vec{\theta} = \hat{\vec{\theta}}$) is integrated along the observables other than x and, after appropriate normalization, it is overlaid ("projected") to the binned distribution of x in data. This allows detecting possible macroscopic discrepancies between the observed distributions and the model. Distributions of the discriminating observables with fit projections overlaid are shown in figs. 7.2–7.5.

The fit reproduces accurately all the observed distributions. Figures 7.5(a) and (b) reports the distributions of κ for tracks from candidates in the signal region only ($5.125 < m_{\pi^+\pi^-} < 5.400$ GeV/c^2) for a further check on whether the fit properly determines the sample composition underneath the signal peak. The fit reproduces accurately the data both in the central part of the distribution, where the signal is present, and in the tails at lower (higher) values of κ, where electrons (protons) from background contribute.

The fit projections show clearly the different impact of each observable in separating the single modes. The mass and momentum observables contribute mostly in distinguishing the $B^0 \to K^+\pi^-$ mode from the others, whereas poor separation is obtained between $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$. The dE/dx observables, instead, separate $B^0 \to \pi^+\pi^-$ from $B_s^0 \to K^+K^-$ events effectively, thus complementing the kinematic information.

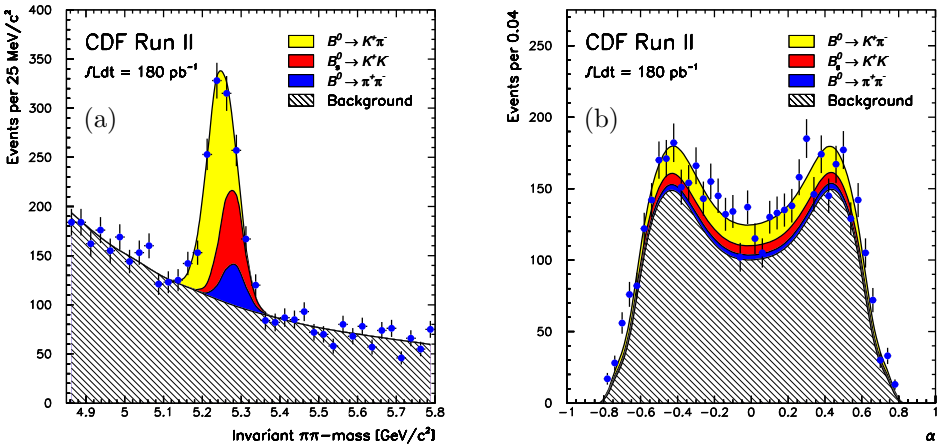

Figure 7.2: Invariant $\pi^+\pi^-$-mass distribution (points with error bars) with fit projection overlaid (shaded functions) (a). Signed momentum imbalance distribution (points with error bars) with fit projection overlaid (shaded functions) (b). In all fit projections of figs. 7.2–7.5 the p.d.f. of all signals (shaded in color) are summed with each other and with the p.d.f. of background (hatched in black).

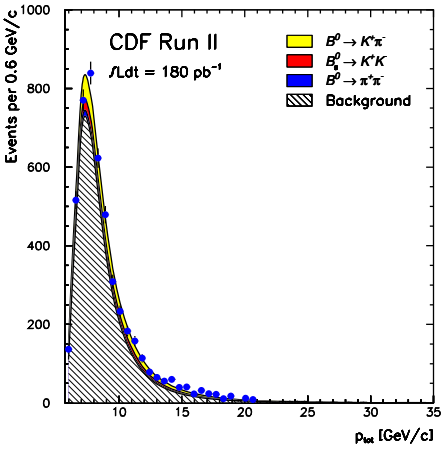

Figure 7.3: Distribution of the scalar sum of momenta (points with error bars) with fit projection overlaid (shaded functions).

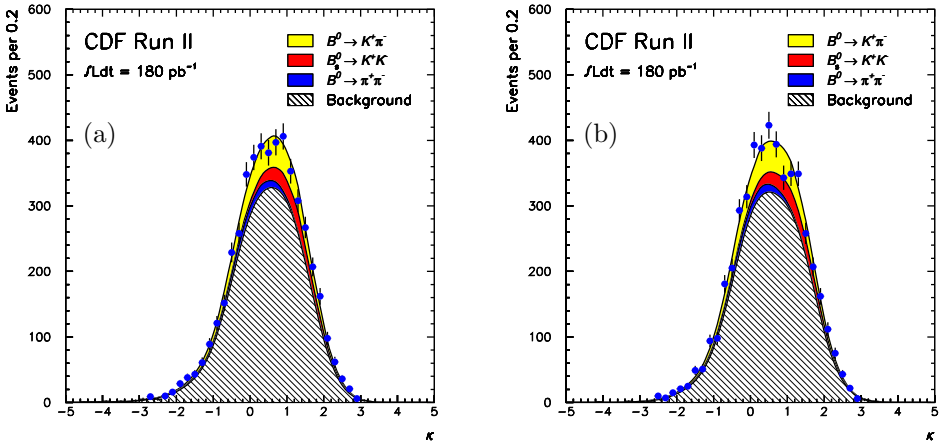

Figure 7.4: Distribution of the κ observable (points with error bars) with fit projection overlaid (shaded functions) for the smaller-momentum particle (a) and for the larger-momentum particle (b).

7.2 Search for rare $B_s^0 \to \pi^+\pi^-$ and $B^0 \to \pi^+\pi^-$ modes

Contributions of the rare $B^0 \to K^+K^-$ and $B_s^0 \to \pi^+\pi^-$ modes were not considered in the fit listed in tab. 7.1. We searched for these modes with two dedicated fits, by adding the fraction of $B^0 \to K^+K^-$ (or $B_s^0 \to \pi^+\pi^-$) decays as a free parameter in the likelihood. Table 7.3 reports the results of the fit for the $B^0 \to K^+K^-$ mode (second column), and for the $B_s^0 \to \pi^+\pi^-$ mode (third column).

No evidence for the presence of $B^0 \to K^+K^-$ and $B_s^0 \to \pi^+\pi^-$ decays was found in our sample.

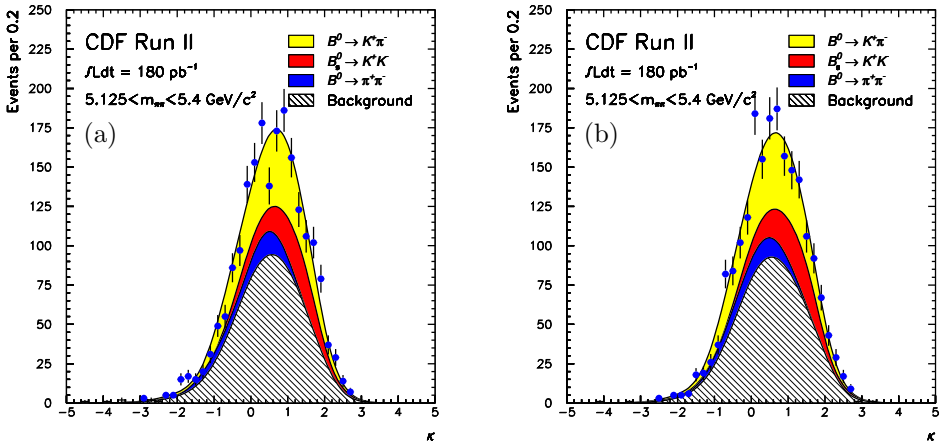

Figure 7.5: Distribution of the κ observable from candidates in the signal region (points with error bars) with fit projection overlaid (shaded functions) for the smaller-momentum particle (a) and for the larger-momentum particle (b).

Quantity	Fit estimate	Fit estimate
$\hat{f}_{B^0\to\pi^+\pi^-}$	0.133 ± 0.030	0.133 ± 0.030
$\hat{f}_{B^0\to K^+\pi^-}$	0.589 ± 0.042	0.605 ± 0.035
$\frac{\hat{f}_{\overline{B}^0\to K^-\pi^+} - \hat{f}_{B^0\to K^+\pi^-}}{\hat{f}_{\overline{B}^0\to K^-\pi^+} + \hat{f}_{B^0\to K^+\pi^-}}$	-0.023 ± 0.079	-0.022 ± 0.078
$\hat{f}_{B_s^0\to K^-\pi^+}$	0.004 ± 0.027	0.009 ± 0.029
$\hat{f}_{B_s^0\to K^+K^-}$	0.263 ± 0.035	0.265 ± 0.035
$\hat{f}_{B^0\to K^+K^-}$	0.011 ± 0.026	–
$\hat{f}_{B_s^0\to\pi^+\pi^-}$	–	-0.011 ± 0.017
$\hat{f}_{B_s^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.447 ± 0.079	0.438 ± 0.075
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.225 ± 0.058	0.220 ± 0.057
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B_s^0\to K^+K^-}$	0.505 ± 0.144	0.501 ± 0.145
$\hat{f}_{B^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.018 ± 0.044	–
$\hat{f}_{B_s^0\to\pi^+\pi^-}/\hat{f}_{B_s^0\to K^+K^-}$	–	-0.043 ± 0.064

Table 7.3: Results of the fit for rare $B^0 \to K^+K^-$ (second column) and $B_s^0 \to \pi^+\pi^-$ (third column) modes. C-conjugate modes are implied except for the parameter in the third row.

Allowing for additional signal modes did not bias the estimation of fit parameters, which are consistent with those of tab. 7.1. This holds for the signal parameters (shown in tab. 7.3) and for the background parameters (omitted for brevity). We studied the pull-values of the $\frac{f_{B^0\to K^+K^-}}{f_{B^0\to K^+\pi^-}}$ and $\frac{f_{B_s^0\to\pi^+\pi^-}}{f_{B_s^0\to K^+K^-}}$ ratios of fractions (see fig. 7.6), according to what described in sec. 6.3.2. The pull-values show the expected Gaussian shape, centered in zero and with unit variance (see tab. 7.4). We therefore used the results of tab. 7.3 to set frequentist upper limits for the branching fractions of $B^0 \to K^+K^-$ and $B_s^0 \to \pi^+\pi^-$

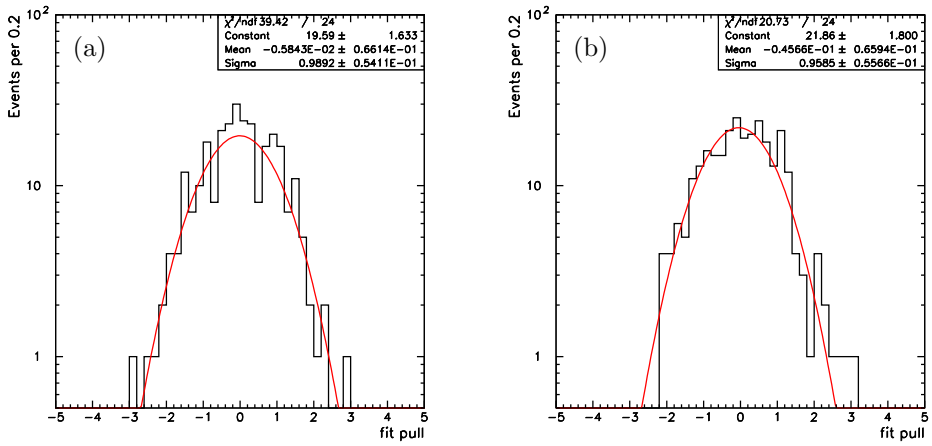

Figure 7.6: Distribution of the pull values for the following fit parameters: (a) $f_{B^0\to K^+K^-}/f_{B^0\to K^+\pi^-}$, (b) $f_{B^0_s\to\pi^+\pi^-}/f_{B^0_s\to K^+K^-}$. Results of χ^2-fits to Gaussian functions are overlaid in red.

Fit parameter	Pull mean	Pull standard deviation	χ^2/d.o.f.
$f_{B^0\to K^+K^-}/f_{B^0\to K^+\pi^-}$	-0.01 ± 0.07	0.99 ± 0.05	39.4/24
$f_{B^0_s\to\pi^+\pi^-}/f_{B^0_s\to K^+K^-}$	-0.04 ± 0.07	0.96 ± 0.06	20.7/24

Table 7.4: Results of the Gaussian fit of the pull value distributions.

modes, based on Gaussian distributions of fit pulls and likelihood-ratio ordering [134]. This procedure is detailed in sec. 10.4–10.6, after including the systematic uncertainties in the extraction of limits.

Chapter 8

Efficiency corrections

In order to infer the desired measurements of branching fractions from the observed numbers of signal events, it is essential to know the total efficiency associated to the reconstruction of each decay-mode. Contribute to this efficiency include efficiency of the on-line trigger requirements and of the off-line data reduction procedure. This chapter details the evaluation of these efficiencies.

8.1 Overview

Table 8.1 contains the ratios of fractions estimated by the fit, along with the ratios of decay rates we wish to extract from them. The symbols f_d and f_s indicate the production fractions of B^0 and B_s^0 mesons from the fragmentation of a b-quark in $p\bar{p}$ collisions.

Measured quantity	Estimated value		Physics quantity
$\frac{\hat{f}_{\overline{B}^0\to K^-\pi^+} - \hat{f}_{B^0\to K^+\pi^-}}{\hat{f}_{\overline{B}^0\to K^-\pi^+} + \hat{f}_{B^0\to K^+\pi^-}}$	$(-2.2 \pm 7.8)\%$	$\Longrightarrow$	$\frac{\mathcal{B}(\overline{B}^0\to K^-\pi^+) - \mathcal{B}(B^0\to K^+\pi^-)}{\mathcal{B}(\overline{B}^0\to K^-\pi^+) + \mathcal{B}(B^0\to K^+\pi^-)}$
$\hat{f}_{B_s^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.44 ± 0.07	$\Longrightarrow$	$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0\to K^+K^-)}{\mathcal{B}(B^0\to K^+\pi^-)}$
$\hat{f}_{B_s^0\to K^-\pi^+}/\hat{f}_{B^0\to K^+\pi^-}$	0.00 ± 0.04	$\Longrightarrow$	$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0\to K^-\pi^+)}{\mathcal{B}(B^0\to K^+\pi^-)}$
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.22 ± 0.06	$\Longrightarrow$	$\frac{\mathcal{B}(B^0\to\pi^+\pi^-)}{\mathcal{B}(B^0\to K^+\pi^-)}$
$\hat{f}_{B^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.02 ± 0.04	$\Longrightarrow$	$\frac{\mathcal{B}(B^0\to K^+K^-)}{\mathcal{B}(B^0\to K^+\pi^-)}$
$\hat{f}_{B_s^0\to\pi^+\pi^-}/\hat{f}_{B_s^0\to K^+K^-}$	-0.04 ± 0.06	$\Longrightarrow$	$\frac{\mathcal{B}(B_s^0\to\pi^+\pi^-)}{\mathcal{B}(B_s^0\to K^+K^-)}$

Table 8.1: Results from the fit and related measurements. Except for the first row, C-conjugate modes are implied.

To transform the ratios of event yields into ratios of branching fractions, we need to correct each yield for its reconstruction efficiency, according to the following equations (in eq. (8.1) C-conjugate modes are *not* implied):

$$A_{CP}(B^0 \to K^+\pi^-) = \frac{\hat{f}_{\overline{B}^0\to K^-\pi^+} \times \frac{\varepsilon_{\rm kin}(B^0\to K^+\pi^-)}{\varepsilon_{\rm kin}(\overline{B}^0\to K^-\pi^+)} - \hat{f}_{B^0\to K^+\pi^-}}{\hat{f}_{\overline{B}^0\to K^-\pi^+} \times \frac{\varepsilon_{\rm kin}(B^0\to K^+\pi^-)}{\varepsilon_{\rm kin}(\overline{B}^0\to K^-\pi^+)} + \hat{f}_{B^0\to K^+\pi^-}}, \tag{8.1}$$

$$\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = \frac{\hat{f}_{B^0\to\pi^+\pi^-}}{\hat{f}_{B^0\to K^+\pi^-}} \times \frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0 \to \pi^+\pi^-)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0 \to \pi^+\pi^-)}, \tag{8.2}$$

$$\frac{\mathcal{B}(B^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = \frac{\hat{f}_{B^0\to K^+K^-}}{\hat{f}_{B^0\to K^+\pi^-}} \times \frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0 \to K^+K^-)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0 \to K^+K^-)}, \tag{8.3}$$

$$\frac{\mathcal{B}(B^0_s \to \pi^+\pi^-)}{\mathcal{B}(B^0_s \to K^+K^-)} = \frac{\hat{f}_{B^0_s\to\pi^+\pi^-}}{\hat{f}_{B^0_s\to K^+K^-}} \times \frac{\varepsilon_{\rm kin}(B^0_s \to K^+K^-)}{\varepsilon_{\rm kin}(B^0_s \to \pi^+\pi^-)} \times \frac{c_{\rm XFT}(B^0_s \to K^+K^-)}{c_{\rm XFT}(B^0_s \to \pi^+\pi^-)}, \tag{8.4}$$

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = \frac{\hat{f}_{B^0_s\to K^+K^-}}{\hat{f}_{B^0\to K^+\pi^-}} \times \frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0_s \to K^+K^-)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0_s \to K^+K^-)} \times \frac{\varepsilon_{\rm iso}(B^0)}{\varepsilon_{\rm iso}(B^0_s)}, \tag{8.5}$$

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s \to K^-\pi^+)}{\mathcal{B}(B^0 \to K^+\pi^-)} = \frac{\hat{f}_{B^0_s\to K^-\pi^+}}{\hat{f}_{B^0\to K^+\pi^-}} \times \frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0_s \to K^-\pi^+)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0_s \to K^-\pi^+)} \times \frac{\varepsilon_{\rm iso}(B^0)}{\varepsilon_{\rm iso}(B^0_s)}. \tag{8.6}$$

In eqs. (8.1)–(8.6) the total efficiency is factorized in three terms,

$$\varepsilon = \varepsilon_{\rm kin} \times c_{\rm XFT} \times \varepsilon_{\rm iso}, \tag{8.7}$$

detailed as follows.

Kinematic efficiency ($\varepsilon_{\rm kin}$) – this term mostly accounts for acceptance effects. It includes the trigger efficiency and the efficiency of the off-line reconstruction and selection. The trigger efficiency accounts for any signal decay produced in $p\bar{p}$ collisions outside the geometric acceptance of the detector, or lost for other trigger-related inefficiencies. Examples are decays in which the two particles are within the COT acceptance but are not reconstructed by the XFT or by the SVT, or properly-reconstructed decays that fail the trigger requirements (e. g., opening angle, or impact parameter criteria, etc.). The off-line efficiency accounts for all triggered decays that fail the off-line reconstruction and selection, e. g., events excluded because involving subdetectors resulted partially or temporarily unreliable, or events failing the selection of sec. 3.8.[1]

The kinematic efficiency was extracted from simulated Monte Carlo samples. For this reason, this term does not include the contribution of the isolation requirement which is treated separately.

Track-trigger efficiency ($c_{\rm XFT}$) – this correction takes into account the discrepancy between simulation and data in the efficiency of the XFT trigger for charged kaons and pions. It was extracted from collision data.

Isolation efficiency ($\varepsilon_{\rm iso}$) – this is the efficiency for a signal event to satisfy the requirement on isolation. It was extracted from collision data.

[1] Occasionally, further and more accurate inspection of the results of routine detector-calibrations shows hardware problems not acknowledged during the data-taking.

8.2 Kinematic efficiencies

The kinematic efficiencies were extracted by using simulated samples of $B^0_{(s)} \to h^+h^{'-}$ decays (see sec. 3.6). The efficiency for each mode was obtained as follows:

$$\varepsilon_{\rm kin} = \frac{N^*}{N}, \tag{8.8}$$

where N is the number of generated decays for each mode, and N^* is the number of generated decays that passed the simulation of trigger, the reconstruction, and the off-line analysis requirements. Use of simulated samples is reliable because the effects contributing to the kinematic efficiency are either properly reproduced by the CDF II simulation, or common to all modes, thus vanishing in the ratios.

Any geometric acceptance effect is properly taken into account because the simulation reproduces the relevant kinematic distributions of the decays and includes an accurate description of the CDF II detector geometry (see sec. 2.7). Additional prominent contributions arise from the different probability of interaction with matter among positively-charged kaons, negatively-charged kaons, and pions, and from the different decay-in-flight probability between pions and kaons. Several CDF measurement confirm that the GEANT package reproduces these effects at the level of accuracy required for our purpose. All other residual contributions to the efficiency (e. g., exclusion of decays involving temporarily-damaged parts of the detector, etc.) can be disregarded: these effects introduce the same inefficiency in all modes; thus, the inefficiency is canceled out in the ratios. The opportunity to exploit such cancellations was one of the motivation for a measurement of relative, rather than absolute, branching fractions.

8.2.1 Kinematic correction for the branching fraction measurements

We applied the trigger simulation, the event reconstruction, and the off-line analysis selection to the sample of 10^6 generated decays per mode described in sec. 3.6. Table 8.2 reports, in second column, the yields of reconstructed decays after applying the trigger simulation, the event reconstruction, and the off-line selection requirements (except for the isolation). The resulting kinematic efficiencies are $\mathcal{O}(1\%)$. From the event yields of second column of tab. 8.2 we obtain the following ratios of kinematic efficiencies:

$$\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0 \to \pi^+\pi^-)} = 0.97 \pm 0.01, \tag{8.9}$$

$$\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0_s \to K^+K^-)} = 1.08 \pm 0.01, \tag{8.10}$$

$$\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0_s \to K^-\pi^+)} = 1.02 \pm 0.01, \tag{8.11}$$

$$\frac{\varepsilon_{\rm kin}(B^0_s \to K^+K^-)}{\varepsilon_{\rm kin}(B^0_s \to \pi^+\pi^-)} = 0.89 \pm 0.01, \tag{8.12}$$

and

$$\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0 \to K^+K^-)} = 1.02 \pm 0.01. \tag{8.13}$$

Mode	$N\varepsilon_{\rm kin}$	$N\varepsilon_{\rm kin}c_{\rm XFT}$
$B^0 \to \pi^+\pi^-$	13928 ± 118	13255 ± 115
$B^0 \to K^+\pi^-$	13567 ± 117	12308 ± 111
$B^0 \to K^+K^-$	13260 ± 115	11407 ± 107
$B_s^0 \to K^+K^-$	12595 ± 112	10900 ± 104
$B_s^0 \to K^-\pi^+$	13340 ± 115	12109 ± 110
$B_s^0 \to \pi^+\pi^-$	14116 ± 119	13455 ± 116

Table 8.2: Number of candidates in simulated events after trigger, reconstruction, and analysis selection, excluding the isolation cut (second column). Same quantities after correction for the XFT bias (third column).

8.2.2 Kinematic correction for the CP asymmetry measurement

The only acceptance effect that matters in the measurement of $A_{\rm CP}(B^0 \to K^+\pi^-)$ is the detector-induced charge asymmetry between positively and negatively charged kaons, due to their different probability of strong interaction in the tracker material. All other inefficiencies, such as the $\mathcal{O}(1\%)$ asymmetry between positively and negatively-charged particles induced by the COT cell-geometry, cancel out in the ratio involved in the expression of the asymmetry.

Following Ref. [106], we used large samples of simulated single kaons with momentum distributions similar to those of $B^0_{(s)} \to h^+h^{'-}$ decays to extract the probabilities of missing the reconstruction of a kaon due to hadronic interactions in the tracker. The results were $p_{\rm T}(B)$-averaged to yield a $(2.8 \pm 0.1)\%$ probability for positively-charged and $(4.5 \pm 0.1)\%$ for negatively-charged kaons, in agreement with the published cross-sections on deuterium [109]. The resulting correction factors are

$$\varepsilon_{\rm had}(K^+) = 0.972 \pm 0.001 \quad \text{and} \quad \varepsilon_{\rm had}(K^-) = 0.955 \pm 0.001, \tag{8.14}$$

where $\varepsilon_{\rm had}(K^+) \equiv \varepsilon_{\rm kin}(B^0 \to K^+\pi^-)$ and $\varepsilon_{\rm had}(K^-) \equiv \varepsilon_{\rm kin}(\overline{B}^0 \to K^-\pi^+)$. Our level of understanding of the charge asymmetries is confirmed by the measurement of $D^0 \to h^+h^{'-}$ partial widths of Ref. [106], where these effects are found to be under control at the 0.5% level.

8.3 Track-trigger efficiency

In the typical momentum range of the $B^0_{(s)} \to h^+h^{'-}$ decay-products, pions (with typical Lorentz boost $\beta\gamma \approx 15$) ionize more than kaons ($\beta\gamma \approx 4$). Thus, pions produce larger average pulse-widths in the COT sense-wires with respect to kaons. This results in typically higher numbers of hits above threshold associated to pions. The XFT requirement on the minimum number of axial COT-hits above threshold per track (see sec. 2.5.1) is therefore less efficient for kaons than for pions, thus biasing the observed ratios of yields between modes with different multiplicities of kaons and pions in their final states. An additional correction is needed to account for this effect, which is not reproduced by the simulation because of incorrect energy-loss modeling.

A time-dependent study on data made for the CDF measurement of $D^0 \to h^+h^{'-}$ partial widths [106] finds an average 6% (see fig. 8.1) higher efficiency for pions with respect to kaons. This was

measured using the XFT-unbiased prong in three-body $D^+_{(s)}$ meson decays ($D^+ \to K^-\pi^+\pi^+$, $D^+ \to K^+K^-\pi^+$and $D^+_s \to K^+K^-\pi^+$) triggered on two tracks. Using the XFT information associated to the non-triggered kaon (or pion) in the event, the fraction of kaons (pions) reconstructed off-line ($N_{\rm off}$) that would have passed the XFT requirements was measured as a function of the transverse momentum:

$$\varepsilon_{\rm XFT} = \frac{N_{\rm XFT}}{N_{\rm off}}. \tag{8.15}$$

For each charged particle of transverse momentum $p_{\rm T}$, the ratio of XFT efficiencies between data

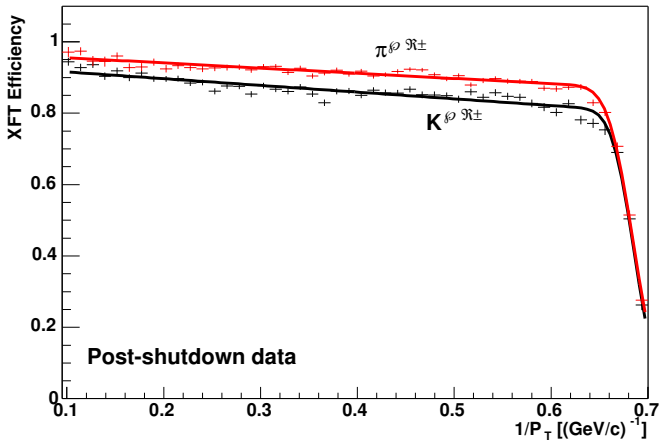

Figure 8.1: XFT efficiency for positively-charged kaons (black points) and pions (red points) as a function of $1/p_{\rm T}$, as measured in $D^+_{(s)}$ meson decays [106]. Fit functions are overlaid (solid lines).

and Monte Carlo simulation ($c_{\rm XFT}$) was fit to a straight-line function of $1/p_{\rm T}$:

$$c_{\rm XFT}(p_{\rm T}) = \frac{\varepsilon^{\rm data}_{\rm XFT}}{\varepsilon^{\rm MC}_{\rm XFT}} = k_0 + \frac{k_1}{p_{\rm T}}, \tag{8.16}$$

in which $\varepsilon^{\rm data}_{\rm XFT}$ is the XFT-efficiency measured on collision data, and $\varepsilon^{\rm MC}_{\rm XFT}$ is the XFT efficiency extracted from the simulation. The parameters of the above formula, determined from binned χ^2-fits as a function of time (in terms of runs), are shown in tab. 8.3.[2] Following Ref. [106], we applied the

Param.	Units	138809–152636	152637–156487	156488–164274	164275–168889
$k_0(\pi)$	-	1.024 ± 0.003	1.008 ± 0.006	0.985 ± 0.006	1.002 ± 0.005
$k_0(K)$	-	1.017 ± 0.003	0.959 ± 0.006	0.952 ± 0.004	0.957 ± 0.004
$k_1(\pi)$	GeV/c	0.0	-0.10 ± 0.02	-0.08 ± 0.02	-0.12 ± 0.02
$k_1(K)$	GeV/c	0.0	-0.11 ± 0.02	-0.13 ± 0.01	-0.14 ± 0.01

Table 8.3: Estimated parameters of the model of the XFT bias.

[2]The chosen run sectioning accounts for major trigger variations, already described in sec. 3.2.4.

correction to the simulated samples: the contribution to the final yield of each simulated $B^0_{(s)} \to h^+h^{'-}$ event was reweighed by the product $c_{\rm XFT}[p_{\rm T}(h)] \times c_{\rm XFT}[p_{\rm T}(h^{'})]$. The two factors deriving from eq. (8.16) were evaluated using the proper set of k_i parameters from tab. 8.3 and the specific value of transverse momentum of the particle. Table 8.2 shows the resulting reweighed yields (third column). The XFT-corrections with their statistical uncertainties are

$$\frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0 \to \pi^+\pi^-)} = 0.95 \pm 0.02, \tag{8.17}$$

$$\frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0_s \to K^+K^-)} = 1.05 \pm 0.02, \tag{8.18}$$

$$\frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0_s \to K^-\pi^+)} = 1.00 \pm 0.02, \tag{8.19}$$

$$\frac{c_{\rm XFT}(B^0_s \to K^+K^-)}{c_{\rm XFT}(B^0_s \to \pi^+\pi^-)} = 0.91 \pm 0.02, \tag{8.20}$$

and

$$\frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0 \to K^+K^-)} = 1.05 \pm 0.02. \tag{8.21}$$

8.4 Results (not involving isolation efficiency)

At this point we have the corrections for those ratios of branching fractions not needing the isolation efficiency (eqs. (8.1)–(8.4)) and we can extract the corresponding measurements. All associated uncertainties account for the statistical fluctuations due to finite samples only. The corrections for the remaining measurements (eqs. (8.5)–(8.6)) are treated in the rest of this chapter.

8.4.1 Partial rate asymmetry in the $B^0 \to K^+\pi^-$ decay

Following eq. (8.1), we corrected the fit results for the kaon charge-asymmetry factor, $\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(\overline{B}^0 \to K^-\pi^+)} \simeq 1.018$, to extract the direct asymmetry:

$$\frac{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) - \mathcal{B}(B^0 \to K^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) + \mathcal{B}(B^0 \to K^+\pi^-)} = -0.013 \pm 0.078\ (stat.). \tag{8.22}$$

8.4.2 Ratio $\mathcal{B}(B^0 \to \pi^+\pi^-)/\mathcal{B}(B^0 \to K^+\pi^-)$

Using the kinematic and XFT efficiencies $\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0 \to \pi^+\pi^-)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B^0 \to \pi^+\pi^-)} \simeq 0.93$ (see eq. (8.2)), we corrected the fit results to extract the following measurement of ratio of branching fractions:

$$\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.21 \pm 0.05\ (stat.). \tag{8.23}$$

8.4.3 Ratio $\mathcal{B}(B^0 \to K^+K^-)/\mathcal{B}(B^0 \to K^+\pi^-)$

Using the kinematic and XFT efficiencies, $\frac{\varepsilon_{\rm kin}(B^0\to K^+\pi^-)}{\varepsilon_{\rm kin}(B^0\to K^+K^-)} \times \frac{c_{\rm XFT}(B^0\to K^+\pi^-)}{c_{\rm XFT}(B^0\to K^+K^-)} \simeq 1.08$ (see eq. (8.3)), we corrected the fit results to extract the following measurement of ratios of branching fractions:

$$\frac{\mathcal{B}(B^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.02 \pm 0.05\ (stat.). \tag{8.24}$$

8.4.4 Ratio $\mathcal{B}(B_s^0 \to \pi^+\pi^-)/\mathcal{B}(B_s^0 \to K^+K^-)$

Using the kinematic and XFT efficiencies, $\frac{\varepsilon_{\rm kin}(B_s^0\to K^+K^-)}{\varepsilon_{\rm kin}(B_s^0\to \pi^+\pi^-)} \times \frac{c_{\rm XFT}(B_s^0\to K^+K^-)}{c_{\rm XFT}(B_s^0\to \pi^+\pi^-)} \simeq 0.81$ (see eq. (8.4)), we corrected the fit results to extract the following measurement of ratio of branching fractions:

$$\frac{\mathcal{B}(B_s^0 \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to K^+K^-)} = 0.03 \pm 0.05\ (stat.). \tag{8.25}$$

8.5 Isolation efficiency

The ratios between B^0 and B_s^0 yields need an additional correction for the different efficiency of the isolation requirement when applied to B^0 or to B_s^0 mesons. The distribution of isolation (see sec. 3.7.2) depends on several factors:

1. the multiplicity and momenta of the charged-particles not belonging to the b-meson decay-chain and produced in the b-quark fragmentation, or in the underlying event, or in possible pile-up events;

2. the transverse momentum of the b-meson, since the sharing of transverse momentum between the b-meson and its neighboring particles depends on the energy of the b-meson itself.[3]

This has some consequences. First, the isolation is independent of the decay mode of the b-meson, but it does depend on its production mechanism. We therefore expect different isolation distributions between B^0 and B_s^0 mesons, which are produced through different fragmentation processes. A second consequence is the need to study the isolation using collision data, since typical $\mathrm{p\bar{p}}$ collision simulators, as PYTHIA, cannot be trusted to reproduce reliably the details of the b-quark fragmentation. In addition, the dependence on the transverse momentum of the b-meson needs to be considered.

We measured, as a function of $p_{\rm T}(B)$, the isolation efficiencies $\varepsilon_{\rm iso}(B^0)$ and $\varepsilon_{\rm iso}(B_s^0)$, corresponding to the $I(B) > 0.5$ requirement, using several fully reconstructed $B^0_{(s)}$ meson decays available in data.

8.5.1 Choice of the control samples

We chose fully reconstructed decays since the determination of transverse momentum and direction of the $B^0_{(s)}$ meson is required. Since samples sharing the same trigger as of $B^0_{(s)} \to h^+h^{'-}$ decays, and with similar final states are desirable, we used the exclusive $B^0_{(s)} \to D^-_{(s)}\pi^+$ decays. However, common

[3] QCD measurements show that the fraction of energy carried by a charged particle in a jet does not depend linearly on the energy of the jet itself.

thresholds in transverse momentum between the B_PIPI (which collects $B^0_{(s)} \to h^+h'^-$ decays) and the B_CHARM (which collects $B^0_{(s)} \to D^-_{(s)}\pi^+$ decays) trigger paths cause a higher turn-on point in the $p_{\rm T}(B)$ distribution of the $B^0_{(s)} \to D^-_{(s)}\pi^+$ with respect to the $B^0_{(s)} \to h^+h'^-$ modes. A two-body decay-product carries a larger average-fraction of the momentum of the decaying particle than the product of a multi-body decay. Thus, b-mesons in the B_PIPI path had typically smaller transverse momenta than in the B_CHARM path (see fig. 8.2 and figs. 8.3(a) and (c)). While a 30%–40% fraction of $B^0_{(s)} \to h^+h'^-$ candidates have transverse momenta smaller than 6 GeV/c, $B^0_{(s)} \to D^-_{(s)}\pi^+$ candidates are removed, by the trigger bias, from this region of the spectrum.

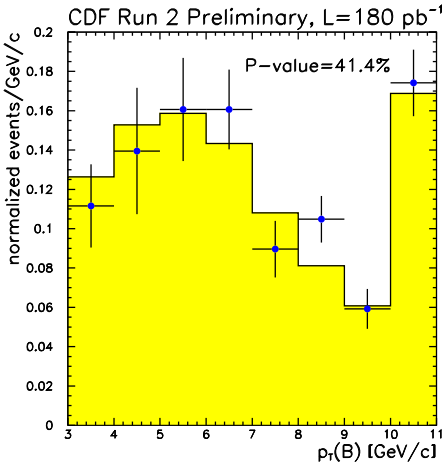

Figure 8.2: Background-subtracted $p_{\rm T}(B)$-distribution in $B^0_{(s)} \to h^+h'^-$ decays. Data (points with error bars) are compared with Monte Carlo simulation (filled histogram). Rightmost bin contains overflows. See sec. 3.6 for details on background subtraction.

We therefore complemented our sample with $B^0_{(s)}$ mesons selected in their $B^0_{(s)} \to J/\psi X$ decays by the di-muon trigger. This trigger selects events enriched in $J/\psi \to \mu^+\mu^-$ decays by requiring azimuthal matching between two oppositely-curved XFT-tracks with $p_{\rm T} > 1.5$ GeV/c and two non-adjacent track-segments in the CMU detectors. The lower threshold on the transverse momenta of muons provides signals with transverse momenta distributions extended down to $p_{\rm T}(B) \approx 4$ GeV/c (see figs. 8.3(b) and (d)). The decay chains of the samples used to measure the isolation efficiency are summarized below:

- $\boldsymbol{B^0_{(s)} \to D^-_{(s)}\pi^+}$ modes:
  - $\boldsymbol{B^0 \to D^-\pi^+ \to [K^+\pi^-\pi^-]\pi^+}$,
    with approximate total branching fraction 2.5×10^{-4};
  - $\boldsymbol{B^0_s \to D^-_s\pi^+ \to [\phi(1020)\pi^-]\pi^+ \to [[K^+K^-]\pi^-]\pi^+}$,
    with approximate total branching fraction 6.6×10^{-5} .
- $\boldsymbol{B^0_{(s)} \to J/\psi X}$ modes:

– $\boldsymbol{B^0 \to J/\psi K^*(892)^0 \to [\mu^+\mu^-][K^+\pi^-]}$,
with approximate total branching fraction 7.7×10^{-5};

– $\boldsymbol{B_s^0 \to J/\psi\phi(1020) \to [\mu^+\mu^-][K^+K^-]}$,
with approximate total branching fraction 2.6×10^{-5}.

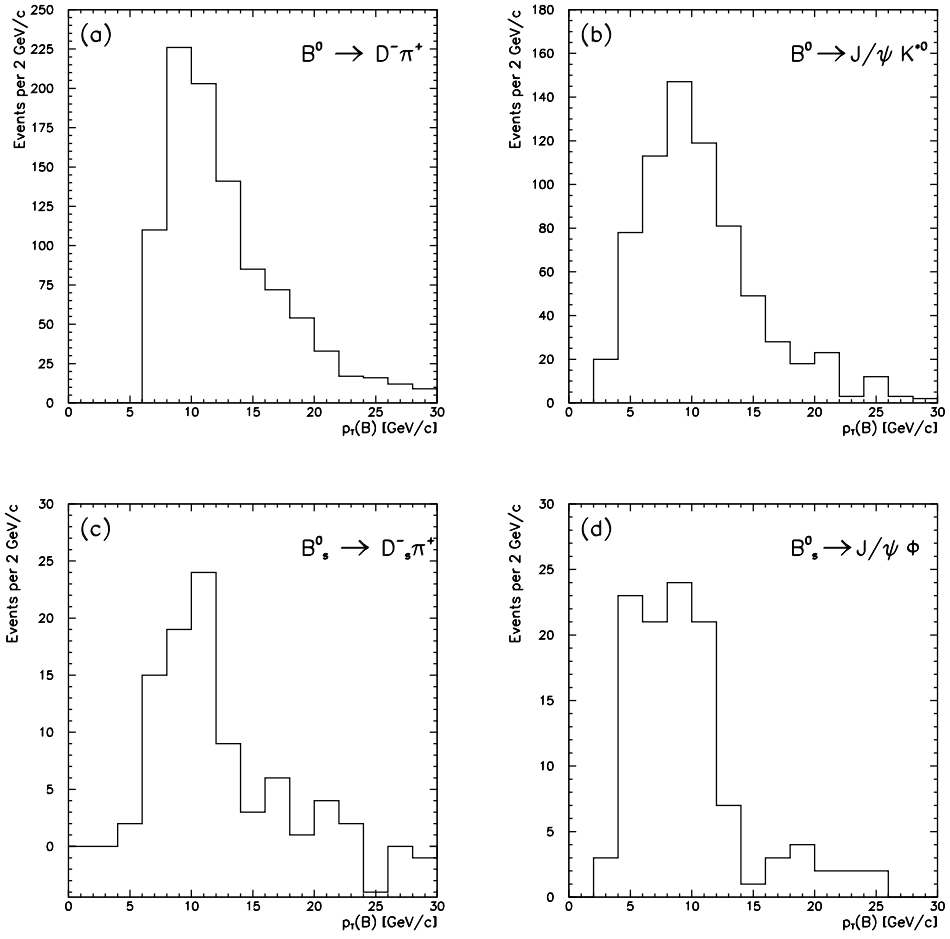

Figure 8.3: Background-subtracted $p_{\rm T}(B)$-distributions in $B^0 \to D^-\pi^+$ (a), $B^0 \to J/\psi K^{*0}$ (b), $B_s^0 \to D_s^-\pi^+$ (c), and $B_s^0 \to J/\psi\phi$ decays (d). A single sideband in the high-mass side of the $B_{(s)}^0 \to D_{(s)}^-\pi^+$ signals was used to extrapolate the background underneath the peak, avoiding partially-reconstructed $B_{(s)}^0$ decays present in the low-mass side of the signal. In the $B_{(s)}^0 \to J/\psi X$ case two symmetric mass-sidebands were used.

8.5.2 Data samples and event selection

All $B^0_{(s)} \to D^-_{(s)}\pi^+$ and $B^0_{(s)} \to J/\psi X$ decays were reconstructed in the same run set as of the $B^0_{(s)} \to h^+h'^-$ analysis and the same track-quality and fiducial-acceptance requirements were applied (see sec. 3.5.1). The reconstruction of all signals was solely based on tracking.

$B^0_{(s)} \to D^-_{(s)}\pi^+$ samples

The $B^0 \to D^-\pi^+$ and $B^0_s \to D^-_s\pi^+$ decays (see fig. 8.4) were reconstructed in data collected by the `B_CHARM` and `B_CHARM_HIGHPT` triggers (see sec. 3.3).

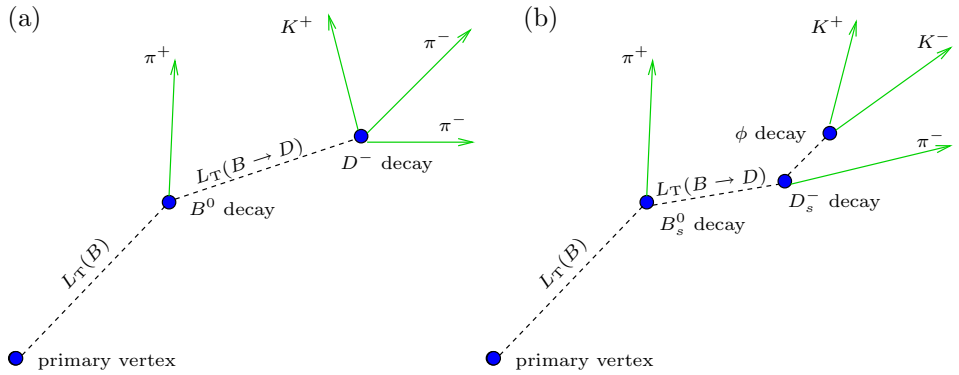

Figure 8.4: Illustration of the $B^0 \to D^-\pi^+ \to [K^+\pi^-\pi^-]\pi^+$ (a) and of the $B^0_s \to D^-_s\pi^+ \to [\phi\pi^-]\pi^+ \to [[K^+K^-]\pi^-]\pi^+$ (b) decay topologies in the transverse plane.

The selection for reconstructing pure $B^0_{(s)}$ meson signals without biasing their isolation distributions was derived following Ref. [135]. Table 8.4 contains the selection requirements, which are specified as follows. First, $D^-_{(s)}$ candidates were reconstructed, by attempting all combinations of three tracks consistent with a $D^- \to K^+\pi^-\pi^-$ (or $D^-_s \to K^+K^-\pi^-$) decay. Once three tracks originating from a common vertex were found, a fourth track with pion mass-assignment was added to reconstruct the $B^0_{(s)}$ candidate. The $D^-_{(s)}$ meson vertex momentum was required to intersect in space with the fourth track. While fitting the $B^0_{(s)}$ meson decay-vertex, the $D^-_{(s)}$ meson masses were constrained to their world-average values [109] to improve mass resolution.[4] In the $B^0_s \to D^-_s\pi^+$ case, further background rejection was achieved by requiring a ϕ in the D^-_s meson decay. We exploited the narrow width of the ϕ resonance to suppress the background by rejecting D^-_s candidates with invariant mass of the two oppositely-charged kaons differing by more than 10 MeV/c^2 from the world average ϕ mass [109]. The requirement on the impact parameter of the $B^0_{(s)}$ meson (see tab. 8.4) ensured that the candidate originated from the primary vertex. Candidates with large, positive decay-length were selected by imposing requirements on the impact parameters of their daughter particles, and on $L_T(B)$ and $L_T(B \to D)$.

[4] We verified that the mass-constrained fit did not bias the signal yield by repeating the fit after removing the constraint.

Variable	Units	$B^0 \to D^-\pi^+$	$B_s^0 \to D_s^-\pi^+$
$p_{\rm T}(D)$	GeV/c	> 5.0	> 3.0
$p_{\rm T}(B)$	GeV/c	> 6.0	> 5.5
$p_{\rm T}(\pi_B)$	GeV/c	> 1.6	> 1.9
$\lvert d_0(B)\rvert$	μm	< 80	< 80
$\lvert d_0(\text{track})\rvert$	cm	< 0.1	< 0.1
$\lvert d_0(\text{track})\rvert$	μm	> 120	> 120
$L_{\rm T}(B \to D)$	μm	> -150	> -150
$L_{\rm T}(B)$	μm	> 250	> 400
ϕ mass	GeV/c^2	–	$1.010 < m_{KK} < 1.030$
D mass	GeV/c^2	1.8694	1.9683
$\Delta R(D, \pi_B)$	–	< 1.5	< 1.5
$\chi^2_{\rm T}(D)$	–	< 10	< 10
$\chi^2_{\rm T}(B)$	–	< 15	< 15

Table 8.4: The $B^0_{(s)} \to D^-_{(s)}\pi^+$ selection. $\Delta R(D, \pi_B) \equiv \sqrt{(\Delta\eta)^2 + (\Delta\varphi_0)^2}$ where $\Delta\eta$ and $\Delta\varphi_0$ are pseudo-rapidity and azimuthal angle of the pion candidate originated from the $B^0_{(s)}$ meson decay-vertex, with respect to the direction of the $D^-_{(s)}$ meson. The quantities $\chi^2_{\rm T}(B)$ and $\chi^2_{\rm T}(D)$ are quality parameters of the fits to respectively the $B^0_{(s)}$ and $D^-_{(s)}$ candidate decay-vertices in the transverse plane.

Figures 8.5(a) and (c) show the invariant-mass distributions resulting from the selection of tab. 8.4. We reconstructed about 2.2×10^3 $B^0 \to D^-\pi^+$ decays with about 1.4 signal purity at the peak, and a small, but pure ($\approx$ 3.4 peak purity), sample of about 140 $B_s^0 \to D_s^-\pi^+$ decays.[5] Two distinct background components are visible: an exponential-like decreasing contribution, distributed over the whole mass-range — but more evident in the high-mass side of the signal — and an additional component present only in the low-mass side of the signal. The former is combinatorial background due to either a $D^-_{(s)}$ meson signal combined with a displaced track, or to a random four-track combination that satisfied the selection requirements. The latter is due to partially reconstructed $B^0_{(s)}$ meson decays such as $B^0_{(s)} \to D^-_{(s)}\rho^+$, and $B^0_{(s)} \to D^{*-}_{(s)}\pi^+$ decays, where the $D^-_{(s)}$ meson is properly reconstructed and the π^0 (or the photon) from the $\rho^+ \to \pi^+\pi^0$ ($D^{*-}_{(s)} \to D^-_{(s)}\pi^0/\gamma$) decay is undetected.

$B^0_{(s)} \to J/\psi X$ samples

$B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi\phi$ modes (see fig. 8.6) were reconstructed in data collected by the di-muon trigger. Following the standard CDF criteria for reconstructing $J/\psi X$ modes (see, for instance, Ref. [105]), we relaxed the requirements on the transverse momenta of $B^0_{(s)}$ mesons and of their daughter particles to gather enough signal events in the lower part of the $p_{\rm T}(B)$-spectrum. We partially compensated the worsening of the signal purity, induced by looser momentum-thresholds, by tightening the $L_{\rm T}(B)$ and $d_0(B)$ requirements.

[5] We indicate as signal purity N_s/N_b where N_s (N_b) are the signal (background) events in the invariant-mass bin corresponding to the peak.

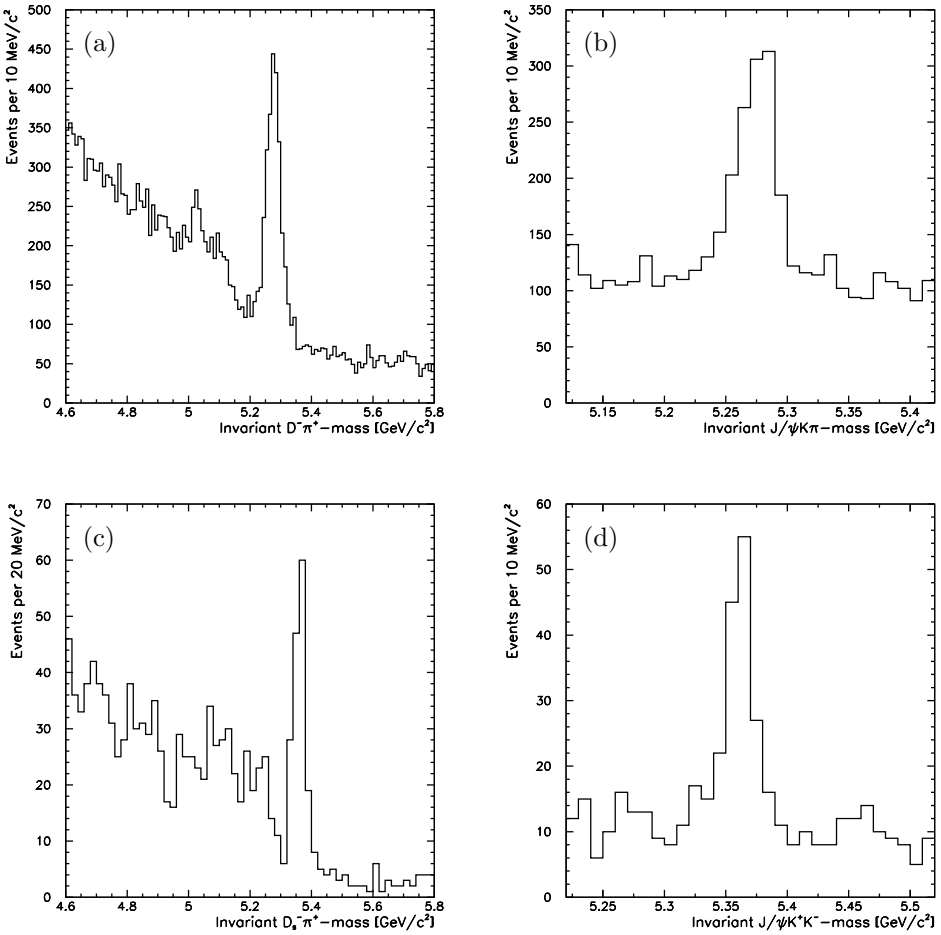

Figure 8.5: Invariant $D^-\pi^+$-mass distribution of $B^0 \to D^-\pi^+$ candidates (a); invariant $J/\psi K\pi$-mass distribution of $B^0 \to J/\psi K^{*0}$ candidates (b); invariant $D_s^-\pi^+$-mass distribution of $B_s^0 \to D_s^-\pi^+$ candidates (c); invariant $J/\psi K^+K^-$-mass distribution of $B_s^0 \to J/\psi\phi$ candidates (d).

Table 8.5 summarizes the selection criteria, which are discussed as follows. Two oppositely-charged particles, assigned the muon masses, were constrained to a common vertex. Candidates with a di-muon mass within 80 MeV/c^2 from the world-average J/ψ mass [109] were associated to a pair of tracks consistent with a $K^{*0} \to K^+\pi^-$ (or $\phi \to K^+K^-$) decay. Then, the four tracks were required to pass through a common vertex, while the invariant $\mu^+\mu^-$-mass was constrained to the world average J/ψ mass [109] to improve mass resolution. Of the two possible mass assignments in the K^{*0} reconstruction, $K^+\pi^-$ and $K^-\pi^+$, we chose the one yielding the invariant mass closest to the pole K^{*0} mass [109]. This choice, wrong in about 10% of the cases [136], has negligible effect in our measurement of efficiency (see sec. 8.5.4). The trigger threshold on transverse momenta of CMU muons was confirmed on off-line quantities (see tab. 8.5). The requirements on transverse decay-length and impact parameter select $B^0_{(s)}$ candidates originated from the primary vertex and with large, positive

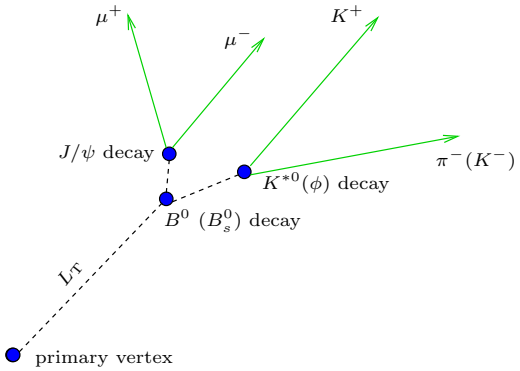

Figure 8.6: Illustration of the $B^0 \to J/\psi K^{*0} \to [\mu^+\mu^-][K^+\pi^-]$ ($B_s^0 \to J/\psi\phi \to [\mu^+\mu^-][K^+K^-]$) decay topology in the transverse plane.

decay-length.

Figures 8.5(b) and (d) show the invariant-mass distributions resulting from the selection of tab. 8.5. We reconstructed about 770 $B^0 \to J/\psi K^{*0}$ decays with approximate signal-to-background ratio 1.0 at the peak, and about 100 $B_s^0 \to J/\psi\phi$ decays with about 1.5 peak purity. The background, mostly uniform at both sides of the signals, is dominated by random four-track combinations that accidentally meet the selection requirements and random two-track combinations associated with a real J/ψ meson.

Variable	Units	$B^0 \to J/\psi K^{*0}$	$B_s^0 \to J/\psi\phi$
$p_{\rm T}(\mu)$	GeV/c	> 1.5	> 1.5
$p_{\rm T}$ of non-ψ daughter	GeV/c	$p_{\rm T}(K),\ p_T(\pi) > 1.0$	$p_{\rm T}(K),\ p_{\rm T}(K) > 0.5$
$p_{\rm T}$(non-ψ)	GeV/c	$p_{\rm T}(K^{*0}) > 2.0$	$p_{\rm T}(\phi) > 1.5$
$p_{\rm T}(B)$	GeV/c	> 3.0	> 3.0
$\lvert d_0(B)\rvert$	μm	< 100	< 100
$L_{\rm T}(B)$	μm	> 100	> 100
J/ψ mass	GeV/c^2	$3.017 < m_{\mu\mu} < 3.177$	$3.017 < m_{\mu\mu} < 3.177$
Non-ψ mass	GeV/c^2	$0.842 < m_{K\pi} < 0.942$	$1.010 < m_{KK} < 1.030$
$\chi^2_{\rm T}(B)$	–	< 30	< 30

Table 8.5: The $B^0_{(s)} \to J/\psi X$ selection. "Non-ψ" denotes the K^{*0} (ϕ) resonance in B^0 (B_s^0) decays.

8.5.3 Efficiency measurement

We define the efficiency of the $I(B) > 0.5$ requirement in a sample of N_s b-meson decays as being

$$\varepsilon_{\rm iso}(B) = \frac{N_s^{\rm iso}}{N_s}, \tag{8.26}$$

where $N_s^{\rm iso} \leq N_s$ is the number of decays in which the isolation of the b-meson is larger than 0.5. We extracted $\varepsilon_{\rm iso}(B)$ with a joint fit of the number of signal events passing the isolation requirement

($N_s^{\rm iso}$), and of the number of events failing the requirement ($N_s - N_s^{\rm iso}$). This was done for each decay-mode separately, and as a function of the transverse momentum of the b-meson. We first determined $\varepsilon_{\rm iso}(B^0 \to D^-\pi^+)$, $\varepsilon_{\rm iso}(B_s^0 \to D_s^-\pi^+)$, $\varepsilon_{\rm iso}(B^0 \to J/\psi K^{*0})$, etc.; then, the results for b-mesons of the same flavor, measured in distinct decay-modes, were combined to obtain the final efficiencies.

The yields were fit by maximizing a likelihood function that uses the information contained in two discriminating observables: the invariant mass of the B candidate (m_B, see figs. 8.7 and 8.8), and a boolean variable ("flag") that is true if the candidate satisfies the isolation requirement $I(B) > 0.5$. We numerically maximized the unbinned likelihood function $\mathscr{L}$ using the MINUIT package [126] (see appendix B). The likelihood $\mathscr{L}$ is function of the vector $\vec{\theta} = (\theta_1, \ldots, \theta_n)$ of the n unknown parameters, and is written as follows:

$$\mathscr{L}(\vec{\theta}) = e^{-(N_s+N_b)} \prod_{i=1}^{N} \begin{cases} N_s \varepsilon_{\rm iso}^s \mathscr{L}_i^{\rm sig} + N_b \varepsilon_{\rm iso}^b \mathscr{L}_i^{\rm bck-iso} & I(B) > 0.5; \\ N_s (1-\varepsilon_{\rm iso}^s) \mathscr{L}_i^{\rm sig} + N_b (1-\varepsilon_{\rm iso}^b) \mathscr{L}_i^{\rm bck-noiso} & I(B) \leq 0.5. \end{cases} \tag{8.27}$$

For each mode, the two-variate (m_B and flag) fit was implemented as a combination of two, simultaneous, mono-variate fits applied to two independent samples: the events passing and the events failing the $I(B) > 0.5$ requirement. For each event passing (failing) the requirement, the first (second) branch of the above likelihood function was used. In eq. (8.27), the exponential factor accounts for the Poisson fluctuations of the expected total number of events $N_s + N_b$, and the product runs over the observed number of events in our sample, N.[6] The expected numbers of signal (N_s) and background (N_b) events, and the isolation efficiencies on signal ($\varepsilon_{\rm iso}^s$) and on background ($\varepsilon_{\rm iso}^b$) were determined by the fit. The p.d.f. used in the likelihoods that describe signal ($\mathscr{L}_i^{\rm sig}$) and background ($\mathscr{L}_i^{\rm bck-iso}$ and $\mathscr{L}_i^{\rm bck-noiso}$) distributions are detailed as follows.

Probability density function for signal

The $B_{(s)}^0$ meson invariant-mass distribution of each mode is described with a Gaussian probability distribution:

$$\mathscr{L}^{\rm sig} = \wp(m_B; \vec{\theta}) = \mathscr{G}(m_B; \mu, \sigma) = \frac{1}{K} e^{-\frac{1}{2}(\frac{m_B-\mu}{\sigma})^2}, \tag{8.28}$$

where the mean, μ, and the standard deviation, σ are parameters to be determined by the fit, and the normalization constant, $K = \int_a^b e^{-\frac{1}{2}(\frac{m_B-\mu}{\sigma})^2} dm_B$, depends on the mass-range of the fit, $a < m_B < b$.

Probability density function for background

We described the background shapes with different empirical probability distributions depending on the decay-mode considered: an exponential function in $B_{(s)}^0 \to D_{(s)}^- \pi^+$ modes, and a linear function in $B_{(s)}^0 \to J/\psi X$ modes.

[6] When the number of observations in the sample is itself a Poisson random variable, the likelihood is often referred to as *extended*; it is the usual likelihood function with the feature that the sample size is part of the result of the estimation [137].

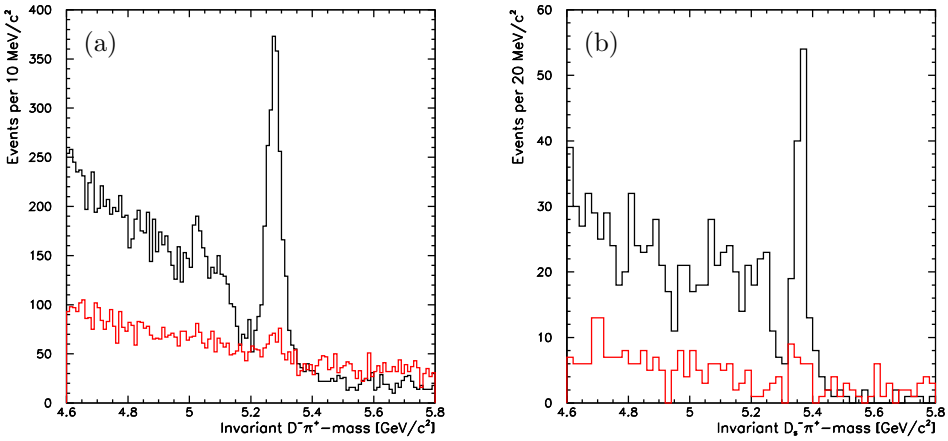

Figure 8.7: Invariant $D^-\pi^+$-mass distributions of $B^0 \to D^-\pi^+$ candidates passing (black histogram), and failing (red histogram), the $I(B) > 0.5$ requirement (a). Invariant $D_s^-\pi^+$-mass distributions of $B_s^0 \to D_s^-\pi^+$ candidates passing (black histogram), and failing (red histogram), the $I(B) > 0.5$ requirement (b).

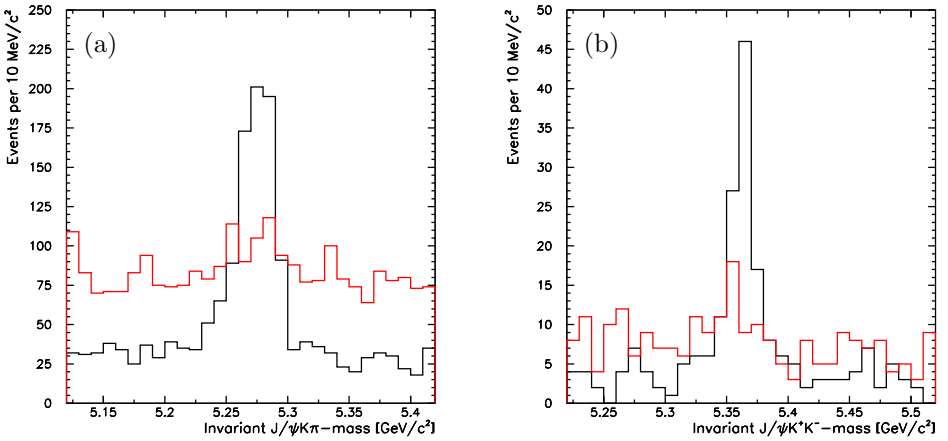

Figure 8.8: Invariant $J/\psi K\pi$-mass distribution of $B^0 \to J/\psi K^{*0}$ candidates passing (black histogram), and failing (red histogram), the $I(B) > 0.5$ requirement (a). Invariant $J/\psi K^+K^-$-mass distributions of $B_s^0 \to J/\psi\phi$ candidates passing (black histogram), and failing (red histogram) the $I(B) > 0.5$ requirement (b).

Since the invariant-mass distribution of background events in $B^0_{(s)} \to D^-_{(s)}\pi^+$ modes varies with the isolation (see fig. 8.7), we used two different probability functions:

$$\mathscr{L}^{\mathrm{bck-iso}} = \wp(m_B;\vec{\theta}) = \frac{1}{K_0}e^{c_0 m_B} \quad \text{and} \quad \mathscr{L}^{\mathrm{bck-noiso}} = \wp(m_B;\vec{\theta}) = \frac{1}{K_1}e^{c_1 m_B}, \tag{8.29}$$

where $\mathscr{L}^{\mathrm{bck-iso}}$ is used for candidates with $I(B) > 0.5$ and $\mathscr{L}^{\mathrm{bck-noiso}}$ for candidates with $I(B) < 0.5$. In the $B^0_{(s)} \to J/\psi X$ sample, instead, the mass shape of the background does not depend on the isolation (see fig. 8.8), allowing for an unique parameterization:

$$\mathscr{L}^{\mathrm{bck-iso}} = \mathscr{L}^{\mathrm{bck-noiso}} = \wp(m_B; \vec{\theta}) = \frac{1}{K_2}(a_0 + a_1 m_B) \tag{8.30}$$

for any value of $I(B)$.

Comments on the likelihood formulation

The vector of unknown parameters, $\vec{\theta}$, was estimated for each mode, independently, to determine

$\hat{\boldsymbol{N}}_{\boldsymbol{s}}$ – the estimate of the signal yield in the sample;

$\hat{\boldsymbol{N}}_{\boldsymbol{b}}$ – the estimate of the background yield in the sample;

$\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\sigma}}$ – the estimate of signal mean and standard deviation;

$\hat{\boldsymbol{\varepsilon}}^{\boldsymbol{s}}_{\mathbf{iso}}$ – the estimate of the efficiency of the $I(B) > 0.5$ requirement on signal events;

$\hat{\boldsymbol{\varepsilon}}^{\boldsymbol{b}}_{\mathbf{iso}}$ – the estimate of the efficiency of the $I(B) > 0.5$ requirement on background events;

$\hat{\boldsymbol{c}}_{\mathbf{0}}$ and $\hat{\boldsymbol{c}}_{\mathbf{1}}$ (or $\hat{\boldsymbol{a}}_{\mathbf{0}}$ and $\hat{\boldsymbol{a}}_{\mathbf{1}}$) – the estimate of the parameters of the mass distribution of background in $B^0_{(s)} \to D^-_{(s)}\pi^+$ (or $B^0_{(s)} \to J/\psi X$) modes.

Using the simultaneous fit of eq. (8.27) is convenient. Since several parameters (N_s, N_b, $\varepsilon^s_{\mathrm{iso}}$, $\varepsilon^b_{\mathrm{iso}}$, μ, σ, a_0, and a_1) are shared between the two branches of the likelihood, optimal use of the available information is achieved in determining these parameters jointly. In addition, the use of common parameters helps the convergence of the maximization of the likelihood in the low statistics, low purity subsamples of candidates that fail the isolation requirement (see figs. 8.9(b) and (d)). Lastly, several systematic effects, not dependent on isolation, cancel out in the ratio of signal yields between the two samples (see sec. 8.5.4).

Fit of the $B^0_{(s)} \to D^-_{(s)}\pi^+$ samples

Fitting the invariant-mass distributions of the hadronic samples is complicated by the sources of background described in sec. 8.5.2. Handling the contribution from partially reconstructed b-meson decays in the low-mass side of the signals was particularly difficult, especially for candidates failing the $I(B) > 0.5$ requirement. These samples contain a small amount of signal over a large amount of background (see figs. 8.7(a) and (b)), possibly inducing large systematic uncertainties.[7]

We simplified the problem by restricting the fit ranges to mass domains where the expected contribution of mis-reconstructed b-mesons is negligible, and the background has the simple shape exhibited by high-mass events. We shifted the lower bound of the fit range toward higher masses at the expense of excluding a small fraction of signal events. This fraction was reincorporated in the final estimate through a correction factor. In $B^0 \to D^-\pi^+$ decays, for instance, we chose the mass range

[7] We expected signal efficiencies $\varepsilon^s_{\mathrm{iso}} \simeq 80\%$–$90\%$ and background efficiencies $\varepsilon^b_{\mathrm{iso}} \simeq 30\%$–$40\%$.

$5.25 < m_{D\pi} < 5.8$ GeV/c^2, which excludes some signal events, since the B^0 signal is centered at about 5.27 GeV/c^2 with about 25 MeV/c^2 r.m.s. width (see fig. 8.5(a)). The signal yield was determined by replacing, in the likelihood (eq. (8.27)), the parameter N_s with $N_s c$, c being the following correction:

$$c = \frac{\int_{5.25}^{5.80} e^{-1/2[(m_{D\pi}-\mu)/\sigma]^2} dm_{D\pi}}{\int_{-\infty}^{5.80} e^{-1/2[(m_{D\pi}-\mu)/\sigma]^2} dm_{D\pi}}. \tag{8.31}$$

The specific value of c depends on the fit estimates in each case. On average it is approximately 0.85, i. e., the fraction of signal events excluded from the fit region is typically around 15%.

Checks on pseudo-experiments showed that varying the fit ranges does not bias the estimated signal yields nor it degrades appreciably their statistical uncertainties.

$B^0_{(s)} \to D^-_{(s)}\pi^+$ results

Figures 8.9(a)–8.9(d) show the invariant-mass distributions for hadronic samples, with fit functions overlaid. To study the $p_{\rm T}(B)$ dependence, we divided the sample according to the transverse momentum of candidates in events with $6 < p_{\rm T}(B) < 10$ GeV/c and events with $p_{\rm T}(B) > 10$ GeV/c. Tables 8.6 and 8.7 summarize the corresponding fit results.

$p_{\rm T}$ range [GeV/c]	$\hat{\varepsilon}^s_{\rm iso}$ [%]	$\hat{\varepsilon}^b_{\rm iso}$ [%]	$\hat{c}$	$\hat{N}_s$	$\hat{N}_b$	$\hat{\mu}$, $\hat{\sigma}$ [MeV/c^2]
$6 < p_{\rm T}(B^0) < 10$	86.2 ± 2.9	31.2 ± 1.4	≈ 0.85	683 ± 55	1701 ± 49	5275 ± 2, 23 ± 2
$p_{\rm T}(B^0) > 10$	93.7 ± 1.3	49.9 ± 1.5	≈ 0.82	1518 ± 86	1649 ± 51	5275 ± 2, 27 ± 2
–	91.2 ± 1.3	40.6 ± 1.0	≈ 0.84	2177 ± 101	3358 ± 71	5275 ± 1, 25 ± 1

Table 8.6: Isolation efficiencies in the $B^0 \to D^-\pi^+$ sample. Last row refers to all events passing off-line selection requisites ($p_{\rm T}(B^0) > 6$ GeV/c, inclusive).

$p_{\rm T}$ range [GeV/c]	$\hat{\varepsilon}^s_{\rm iso}$ [%]	$\hat{\varepsilon}^b_{\rm iso}$ [%]	$\hat{c}$	$\hat{N}_s$	$\hat{N}_b$	$\hat{\mu}$, $\hat{\sigma}$ [MeV/c^2]
$6 < p_{\rm T}(B^0_s) < 10$	75.2 ± 9.6	39.8 ± 9.5	≈ 0.98	42 ± 8	48 ± 9	5359 ± 3, 15 ± 3
$p_{\rm T}(B^0_s) > 10$	91.5 ± 3.1	38.7 ± 11	≈ 0.91	102 ± 8	29 ± 2	5362 ± 3, 23 ± 3
–	84.4 ± 3.9	45.6 ± 8.2	≈ 0.94	136 ± 17	84 ± 8	5360 ± 2, 18 ± 3

Table 8.7: Isolation efficiencies in the $B^0_s \to D^-_s\pi^+$ sample. Last row refers to all events passing the off-line selection requisites ($p_{\rm T}(B^0_s) > 6$ GeV/c, inclusive).

Isolation efficiencies of typically 87% (43%) on signal (background) were measured. The efficiency on signal increases with the transverse momentum of the $B^0_{(s)}$ meson, which is intuitive, since the higher the transverse boost, the more collimated the fragmentation particles around the meson direction.

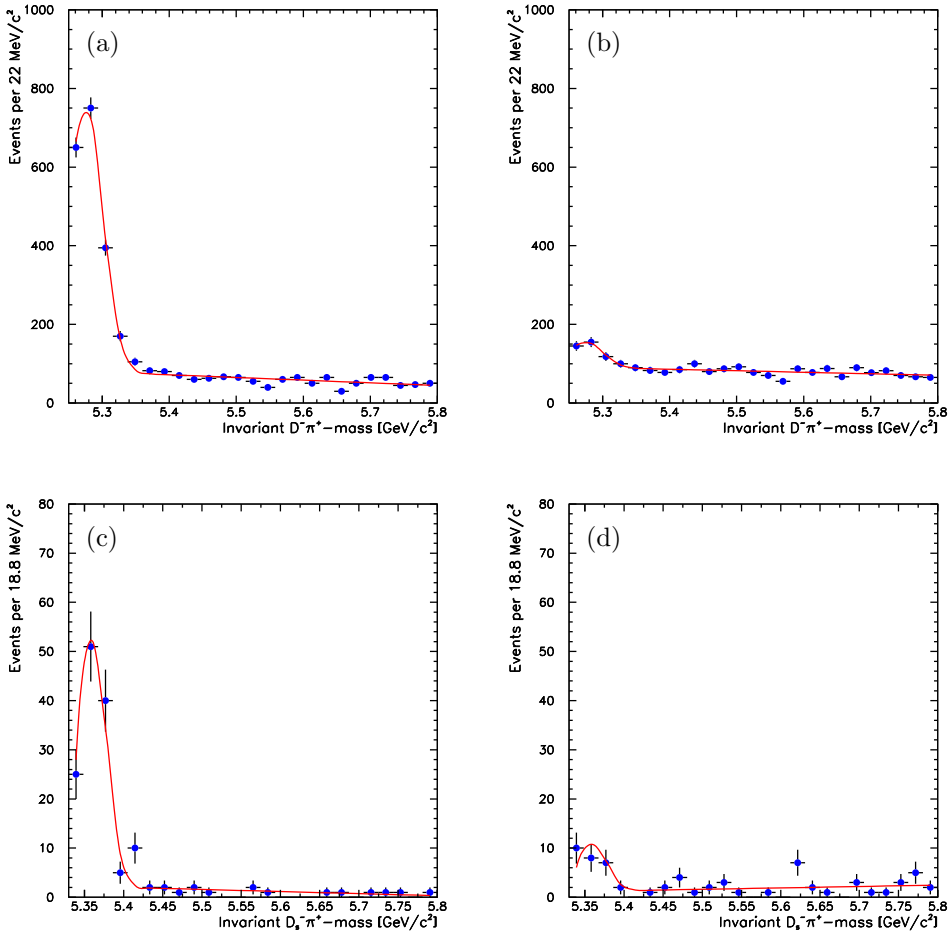

Figure 8.9: Invariant $D^-\pi^+$-mass distributions of $B^0 \to D^-\pi^+$ candidates passing (a), and failing (b), the isolation requirement. Invariant $D_s^-\pi^+$-mass distributions of $B_s^0 \to D_s^-\pi^+$ candidates passing (c), and failing (d), the isolation requirement. Fit functions are overlaid (red, solid line).

Fit of the $B^0_{(s)} \to J/\psi X$ samples

The fit range was not restricted in this case because of the uniform shape of background. Figures 8.10(a)–8.10(d) show the invariant-mass distributions for $B^0_{(s)} \to J/\psi X$ samples, with fit functions overlaid. Tables 8.8 and 8.9 contain the results of the fits applied to subsamples of different transverse momentum of the $B^0_{(s)}$ candidate. An example of such fits is reported in fig. 8.11. Isolation efficiencies

$p_{\rm T}$ range [GeV/c]	$\hat{\varepsilon}^s_{\rm iso}$ [%]	$\hat{\varepsilon}^b_{\rm iso}$ [%]	$\hat{N}_s$	$\hat{N}_b$	$\hat{\mu}$, $\hat{\sigma}$ [MeV/c^2]
$3 < p_{\rm T}(B^0) < 6$	57.5 ± 9.7	18.2 ± 1.3	110 ± 22	1115 ± 38	5263 ± 3, 17.0 ± 2.6
$6 < p_{\rm T}(B^0) < 10$	80.7 ± 4.6	28.5 ± 1.4	275 ± 24	1279 ± 40	5275 ± 1, 12.4 ± 1.0
$p_{\rm T}(B^0) > 10$	94.4 ± 3.0	38.0 ± 1.8	393 ± 26	930 ± 35	5276 ± 1, 14.8 ± 1.0
–	84.4 ± 2.9	27.9 ± 0.9	768 ± 42	3336 ± 66	5275 ± 1, 14.3 ± 0.7

Table 8.8: Isolation efficiencies in the $B^0 \to J/\psi K^{*0}$ sample. Last row refers to all events passing off-line selection requisites ($p_{\rm T}(B^0) > 3$ GeV/c, inclusive).

$p_{\rm T}$ range [GeV/c]	$\hat{\varepsilon}^s_{\rm iso}$ [%]	$\hat{\varepsilon}^b_{\rm iso}$ [%]	$\hat{N}_s$	$\hat{N}_b$	$\hat{\mu}$, $\hat{\sigma}$ [MeV/c^2]
$3 < p_{\rm T}(B^0_s) < 6$	70.1 ± 14.6	28.9 ± 4.1	22 ± 6	141 ± 13	5357 ± 2, 7.1 ± 2.0
$6 < p_{\rm T}(B^0_s) < 10$	90.0 ± 7.1	31.2 ± 4.8	42 ± 7	104 ± 11	5363 ± 1, 6.6 ± 1.1
$p_{\rm T}(B^0_s) > 10$	84.3 ± 7.3	40.7 ± 6.5	51 ± 9	74.4 ± 10	5365 ± 2, 13.6 ± 2.0
–	82.8 ± 5.7	32.8 ± 2.9	113 ± 14	321 ± 20	5362 ± 1, 10.2 ± 1.5

Table 8.9: Isolation efficiencies in the $B^0_s \to J/\psi\phi$ sample. Last row refers to all events passing off-line selection requisites ($p_{\rm T}(B^0_s) > 3$ GeV/c, inclusive).

of typically 83% (30%) on signal (background) were measured. Efficiencies on B^0 mesons was higher than in B^0_s mesons, and isolation efficiencies for signal increase when the b-meson momentum increases.

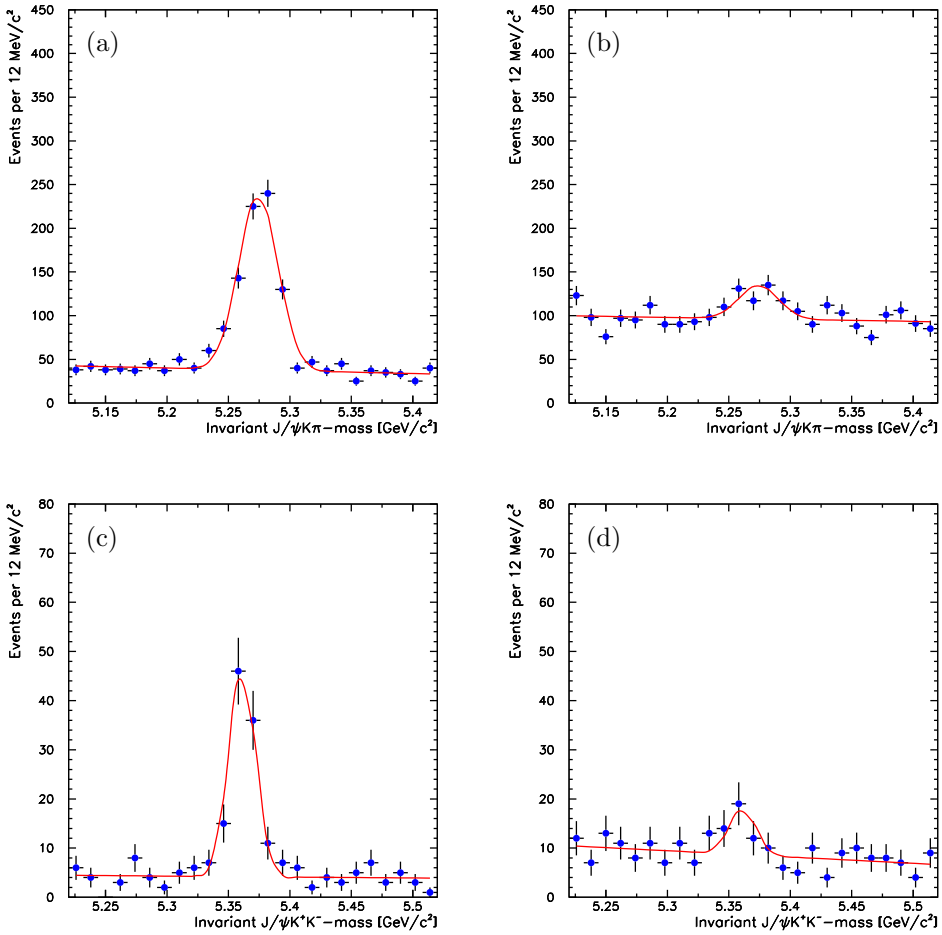

Figure 8.10: Invariant $J/\psi K\pi$-mass distributions of $B^0 \to J/\psi K^{*0}$ candidates passing (a), and failing (b), the isolation requirements. Invariant $J/\psi K^+K^-$-mass distributions of $B_s^0 \to J/\psi\phi$ candidates passing (c), and failing (d). Fit functions are overlaid (red, solid line).

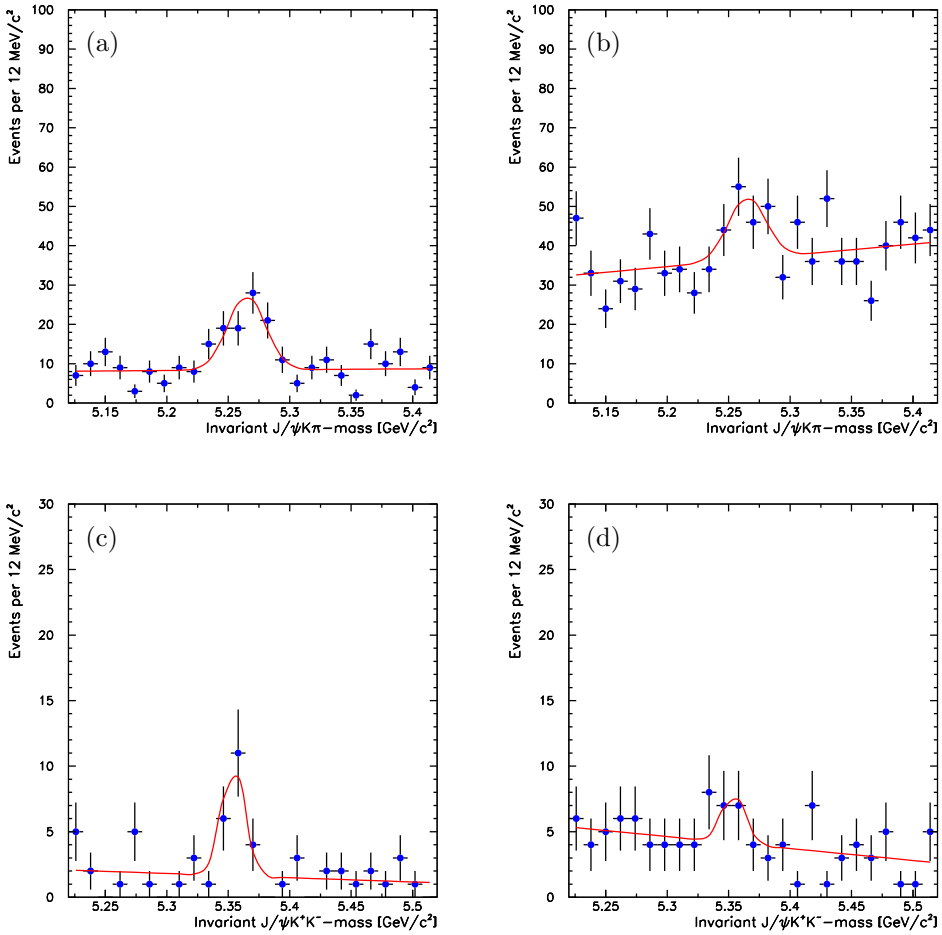

Figure 8.11: Invariant $J/\psi K\pi$-mass distributions of $B^0 \to J/\psi K^{*0}$ candidates with $p_{\mathrm{T}}(B) < 6$ GeV/c passing (a), and failing (b), the isolation requirement. Invariant $J/\psi K^+K^-$-mass distributions of $B_s^0 \to J/\psi\phi$ candidates with $p_{\mathrm{T}}(B) < 6$ GeV/c passing (c), and failing (d), the isolation requirement. Fit functions are overlaid (red, solid line).

8.5.4 Sample-averaged isolation efficiency

Tables 8.6–8.9 report the results of the measurement of the isolation efficiency. In the $3 < p_{\rm T}(B) < 6$ GeV/c subsample, the results are obtained from $J/\psi X$ decays only, otherwise , the two measurements of $\varepsilon^{s}_{\rm iso}$ from the $J/\psi X$ and the hadronic samples are combined. The total isolation efficiency on signal and its variance were obtained from a weighed mean,

$$\varepsilon^{s}_{\rm iso} = \frac{\varepsilon^{s-\psi}_{\rm iso}/\sigma^{2}_{\varepsilon^{s-\psi}_{\rm iso}} + \varepsilon^{s-{\rm had}}_{\rm iso}/\sigma^{2}_{\varepsilon^{s-{\rm had}}_{\rm iso}}}{1/\sigma^{2}_{\varepsilon^{s-\psi}_{\rm iso}} + 1/\sigma^{2}_{\varepsilon^{s-{\rm had}}_{\rm iso}}}, \qquad \sigma^{2}_{\varepsilon^{s}_{\rm iso}} = \frac{\sigma^{2}_{\varepsilon^{s-\psi}_{\rm iso}}\sigma^{2}_{\varepsilon^{s-{\rm had}}_{\rm iso}}}{\sigma^{2}_{\varepsilon^{s-\psi}_{\rm iso}} + \sigma^{2}_{\varepsilon^{s-{\rm had}}_{\rm iso}}}, \tag{8.32}$$

where $\varepsilon^{s-\psi}_{\rm iso}$ is the isolation efficiency on signal reconstructed in the $J/\psi X$ mode, $\varepsilon^{s-{\rm had}}_{\rm iso}$ is the efficiency from hadronic modes, and $\sigma_{\varepsilon^{s-\psi}_{\rm iso}}$ and $\sigma_{\varepsilon^{s-{\rm had}}_{\rm iso}}$ are the corresponding statistical uncertainties. Table 8.10 reports the final results of the averages. Although each efficiency depends individually on the

Sample	$p_{\rm T}$ range [GeV/c]	$\hat{\varepsilon}_{\rm iso}(B^0)$ [%]	$\hat{\varepsilon}_{\rm iso}(B^0_s)$ [%]	$\hat{\varepsilon}_{\rm iso}(B^0)/\hat{\varepsilon}_{\rm iso}(B^0_s)$
J/ψ	$3 < p_{\rm T}(B) < 6$	57.5 ± 9.7	70.1 ± 14.6	0.82 ± 0.22
J/ψ and hadronic	$6 < p_{\rm T}(B) < 10$	84.6 ± 2.4	84.8 ± 5.7	1.00 ± 0.08
J/ψ and hadronic	$p_{\rm T}(B) > 10$	93.8 ± 1.2	90.4 ± 2.8	1.04 ± 0.03

Table 8.10: Averages of the efficiencies measured with the $D^{-}_{(s)}\pi^{+}$ and $J/\psi X$ samples.

transverse momentum of the candidate, no evidence of a $p_{\rm T}(B)$ dependence of the ratios of efficiencies between B^0 and B^0_s mesons was found with the available statistics.

Systematic uncertainties

Systematic effects are negligible compared with statistical uncertainties. Any isolation independent systematic effect is canceled out in the ratio of signal yields, since it would contribute equally to numerator and denominator. This is the case of the $\approx 10\%$ contamination due to the broad self-reflection peak of the $B^0 \to J/\psi K^{*0}$ signal caused by interchanged $K \leftrightarrow \pi$ mass assignments. Any isolation-dependent systematic source should considered, in principle. This is the case, for example, of the known, asymmetric cross-feed between the reconstructed $B^0 \to J/\psi K^{*0}$ and $B^0_s \to J/\psi\phi$ samples, caused by $K-\pi$ mass mis-assignment, which occurs with a few percent incidence, according to simulations [105, 136]. However, these effects are small enough to be diluted in the 10% statistical uncertainties of the measurement.

8.5.5 Momentum-averaged isolation efficiency

The results of the fit of composition were corrected using an isolation efficiency, $\langle\varepsilon_{\rm iso}\rangle$, averaged over the transverse momentum of the $B^0_{(s)}$ meson. We extracted $\langle\varepsilon_{\rm iso}\rangle$ from the results of tab. 8.10, separately for B^0 and B^0_s meson decays, as follows: in each jth bin of transverse momentum, N^j_s is the yield of simulated signal after the full selection (except for the isolation requirement) is applied. The values of N^j_s are easily obtained from the yields in the third column of tab. 8.2 split in different bins of $p_{\rm T}(B)$, i. e., $\sum_j N^j_s = N\varepsilon_{\rm kin}c_{\rm XFT}$. Moreover, the isolation efficiency in each jth bin of transverse

momentum, $\hat{\varepsilon}^j_{\rm iso}$, is known (see tab. 8.10). We therefore obtain the simulated-signal yield expected in each $p_{\rm T}$-bin j after applying the isolation requirement, if this could have been done in the simulation as follows:[8] $N_s^{j-{\rm iso}} = N_s^j \hat{\varepsilon}^j_{\rm iso}$. This quantity, averaged along the $B^0_{(s)}$ meson momentum, becomes $N_s^{\rm iso} = N_s \langle \varepsilon_{\rm iso} \rangle$ that is $\sum_j N_s^{j-{\rm iso}} = \sum_j N_s^j \langle \varepsilon_{\rm iso} \rangle$. The averaged isolation efficiency is therefore obtained as $\langle \varepsilon_{\rm iso} \rangle = \sum_j N_s^{j-{\rm iso}} / \sum_j N_s^j = \sum_j \left(N_s^j \hat{\varepsilon}^j_{\rm iso} \right) / \sum_j N_s^j$. The resulting relative isolation efficiency with its statistical uncertainty is

$$\frac{\langle \varepsilon_{\rm iso}(B_s^0) \rangle}{\langle \varepsilon_{\rm iso}(B^0) \rangle} = 1.07 \pm 0.11. \tag{8.33}$$

8.6 Results (involving isolation efficiency)

The isolation efficiency, along with the kinematic and XFT efficiencies of sec. 8.4, completes the set of corrections we used to obtain the remaining ratios of branching fractions. All uncertainties account only for the statistical fluctuations due to finite samples.

8.6.1 Ratio $(f_s/f_d) \times \mathcal{B}(B_s^0 \to K^+K^-)/\mathcal{B}(B^0 \to K^+\pi^-)$

Following eq. (8.5), we corrected the fit results for the kinematic, the XFT, and the isolation efficiencies, $\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B_s^0 \to K^+K^-)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B_s^0 \to K^+K^-)} \times \frac{\langle \varepsilon_{\rm iso}(B^0) \rangle}{\langle \varepsilon_{\rm iso}(B_s^0) \rangle} \simeq 1.06$, to extract the following ratio of branching fractions:

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.46 \pm 0.08 \ (stat.), \tag{8.34}$$

8.6.2 Ratio $(f_s/f_d) \times \mathcal{B}(B_s^0 \to K^-\pi^+)/\mathcal{B}(B^0 \to K^+\pi^-)$

Following eq. (8.6), we corrected the fit results for the kinematic, the XFT, and the isolation efficiencies, $\frac{\varepsilon_{\rm kin}(B^0 \to K^+\pi^-)}{\varepsilon_{\rm kin}(B_s^0 \to K^-\pi^+)} \times \frac{c_{\rm XFT}(B^0 \to K^+\pi^-)}{c_{\rm XFT}(B_s^0 \to K^-\pi^+)} \times \frac{\langle \varepsilon_{\rm iso}(B^0) \rangle}{\langle \varepsilon_{\rm iso}(B_s^0) \rangle} \simeq 0.95$, to extract the following ratio of branching fractions:

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0 \to K^-\pi^+)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.01 \pm 0.04 \ (stat.), \tag{8.35}$$

8.7 Further considerations

Most corrections described in this chapter were obtained from data to comply with the data-driven approach chosen for the analysis. We used Monte Carlo simulation only when strictly necessary, in a context — the extraction of kinematic efficiencies — that ensures reliable results. In fact, this is the only part of the analysis where simulation enters directly the results. Using simulation to correct *ratios* of branching fractions is reliable since ratios of efficiencies are largely dominated by differences in geometric acceptance between $B^0_{(s)} \to h^+h^{'-}$ modes. These depend only on kinematics of the decays and on the detector geometry, two aspects that are accurately reproduced by the CDF II simulation. At first order, any possible more complex effect that might be poorly reproduced by

[8] We remind the reader that we use the BGENERATOR generator that does not simulate the $p\bar{p}$ interaction (see sec. 3.6).

the simulation cancels out in the ratios. The resulting input from simulation is modest: the largest kinematic correction amounts to a factor of about 12% (see eqs. (8.9)–(8.14)).

In the next chapter we evaluate the systematic uncertainties of the measurement, while the final results are discussed in chap. 10.

Chapter 9

Systematic uncertainties and checks

This chapter describes the evaluation of the uncertainties due to systematic effects relevant for this analysis.

9.1 General strategy

Whenever possible, we adopted the following general procedure to evaluate the influence of systematic effects on our analysis. We first identified the set of nuisance parameters considered as potential sources of systematic uncertainties. For each nuisance parameter s, we generated different sets of pseudo-experiments following the procedure discussed in sec. 6.3.2.[1] One set assumes, for the nuisance parameter, the nominal value s_0 used in the analysis on data ("standard" or "central" analysis). The other sets use alternative configurations, in which s is varied within a realistic range, e. g., $\pm 1\sigma_s$ when the source of systematic uncertainty is the statistical uncertainty σ_s on the parameter s. The resulting systematic uncertainty associated to s is the largest difference between the results of the analysis of the samples with alternative configurations, and the results of the sample with the nominal configuration.

The measurements described in this thesis focus on ratios of branching fractions of decays that are kinematically similar. We expect that most systematic effects that plague the individual modes, e. g., the uncertainty on the integrated luminosity of the sample, (mostly) cancel in the ratio, thus resulting in a smaller systematic uncertainty on the measured ratios. Only systematic effects that have a different impact on different modes were therefore considered. Furthermore, we ignored those systematic effects that induce uncertainties significantly, i. e., a factor $\mathcal{O}(10)$, smaller than their largest counterparts, because their contribution to the total systematic uncertainty is negligible anyhow.

Sections 9.2–9.8 contain the discussion on the dominant systematic uncertainties, while sec. 9.9 summarizes their magnitude for each measurement.

[1] The symbol s may indicate either a single parameter of a multidimensional parameter.

9.2 Systematic effects related to kinematics

Any uncertainty affecting the distributions of the kinematic observables of the fit ($m_{\pi^+\pi^-}$, α, $p_{\rm tot}$) is a potential source of systematic uncertainty for the final measurement.

9.2.1 Uncertainty on mass resolution

The simulation, used to model the signal distributions, underestimates the mass width of individual decay modes (see sec. 3.6). For the central fit, we scaled the $\sigma_m \simeq 22$ MeV/c^2 mass width observed in the simulation, to a value $\sigma_m = 28$ MeV/c^2, as derived from the comparison between data and simulation in $D^0 \to h^+h'^-$ decays and in di-muon decays of J/ψ and Υ(1S) mesons (see tab. 9.1). A

Decay mode	Q-value [MeV/c^2]	$\sigma_{m,\rm data}$ [MeV/c^2]	$\sigma_{m,\rm MC}$ [MeV/c^2]
$D^0 \to K^-\pi^+$	1231	$\simeq 10.0$	$\simeq 7.9$
$J/\psi \to \mu^+\mu^-$	2886	$\simeq 17.6$	$\simeq 13.2$
$\Upsilon(1S) \to \mu^+\mu^-$	9250	$\simeq 65.6$	$\simeq 50.0$

Table 9.1: Mass widths of two-body decays observed in collision data [105] and in simulation. The available kinetic energy of each process (Q-value) is reported.

discrepancy between the actual mass width in data and our assumed value may bias the estimated contribution of the individual modes to the peak, thus introducing a systematic deviation in the measurement of the signal composition.

We repeated the measurement on ensembles of pseudo-experiments generated with either $\sigma_m = 25$ MeV/c^2 or $\sigma_m = 31$ MeV/c^2, which correspond to variations by $\pm 50\%$ of the discrepancy between the data and the predictions of simulation. The maximum variation with respect to the central value was chosen as associated systematic uncertainty (see tab. 9.5).

Check for the $p_{\rm T}$-dependence of the mass width

The reconstructed mass-width of simulated decays depends moderately on the transverse momentum of the $B^0_{(s)}$ meson (see sec. 3.6). This effect was neglected in our default analysis. Thus, we checked its possible impact on the final results. We extracted from simulated data a straight-line parameterization of the $p_{\rm T}$-dependence of the mass width, $\sigma_m(p_{\rm T}) = 0.026 + 0.00029 \times p_{\rm T}$ GeV/c^2 with $p_{\rm T}$ in GeV/c, and we included it in the mass term of the likelihood (eq. (6.6)). The corresponding fit of composition in *data* show no statistically significant variation of the fit results with respect to our central fit, indicating a negligible systematic effect with respect to the statistical fluctuations.

9.2.2 Uncertainty on the nominal $B^0_{(s)}$ meson masses

The values of $B^0_{(s)}$ masses enter the likelihood through the analytic expressions of $\mathcal{M}(\alpha)$ for the mean of the invariant $\pi^+\pi^-$-mass distributions (see sec. 6.2.1). However, $B^0_{(s)}$ meson masses are associated to their experimental uncertainties, which introduce a systematic uncertainty in the present analysis.

Wrong values of the masses in the likelihood would result in a poor description of the $\pi^+\pi^-$-mass distributions of data, and in biases of the fit results. We included this effect considering the *statistical*, rather than the *total*, uncertainty on the input values of $B^0_{(s)}$ masses. In fact, we purposefully used in this analysis the masses measured in Ref. [105] with an analysis that used our own apparatus and reconstruction code, to cancel the common systematic uncertainties. The dominant systematic effects affecting the CDF II measurements of masses (tracker mis-alignments and incorrect absolute-scale of the momentum) are independent of the details of the considered decay-mode, being related only to the tracking performance. As a consequence, these uncertainties mostly cancel in CDF analyses that use mass information retrieved from other CDF measurements. Thus, only the statistical uncertainties on the input masses contribute to the systematic uncertainty in the present analysis.

We repeated the measurement on different samples of pseudo-experiments generated with different combinations of $B^0_{(s)}$ masses. The four possible combinations of B^0 and B^0_s masses were simulated by independently increasing (decreasing) by one statistical standard deviation the masses measured in Ref. [105]: $m_{B^0} = 5279.63 \pm 0.53$ (*stat.*) MeV/c^2 and $m_{B^0_s} = 5366.01 \pm 0.73$ (*stat.*) MeV/c^2.

The largest discrepancy between the results of fits in these samples was taken as systematic uncertainty (see tab. 9.5).

Check for a global mass scale

As a further check of our control on masses, we repeated the fit on *data* with an additional parameter to be determined by the fit: a global mass scale ξ. We added the mass scale in the $\mathcal{M}(\alpha)$ expression within the mass term of the signal p.d.f. (eq. (6.6)) as a global mass offset

$$m_{B^0} \to m_{B^0} + \xi \quad \text{and} \quad m_{B^0_s} \to m_{B^0_s} + \xi. \tag{9.1}$$

Other than for the addition of ξ, the likelihood was unchanged with respect to the central fit. Table 9.2 shows the fit results with ξ free to float (second column) compared with the results of the central fit (third column).

Quantity	Free mass-scale	Central fit
$\hat{f}_{B^0\to\pi^+\pi^-}$	0.116 ± 0.032	0.134 ± 0.030
$\hat{f}_{B^0\to K^+\pi^-}$	0.669 ± 0.054	0.600 ± 0.034
$\left(\hat{f}_{\overline{B}^0\to K^-\pi^+} - \hat{f}_{B^0\to K^+\pi^-}\right) / \left(\hat{f}_{\overline{B}^0\to K^-\pi^+} + \hat{f}_{B^0\to K^+\pi^-}\right)$	-0.034 ± 0.073	-0.022 ± 0.078
$\hat{f}_{B^0_s\to K^-\pi^+}$	-0.018 ± 0.030	0.003 ± 0.028
$\hat{f}_{B^0_s\to K^+K^-}$	0.233 ± 0.039	0.262 ± 0.035
$\hat{f}_{B^0_s\to K^+K^-} / \hat{f}_{B^0\to K^+\pi^-}$	0.348 ± 0.080	0.437 ± 0.074
$\hat{f}_{B^0\to\pi^+\pi^-} / \hat{f}_{B^0\to K^+\pi^-}$	0.174 ± 0.057	0.224 ± 0.057
$\hat{f}_{B^0\to\pi^+\pi^-} / \hat{f}_{B^0_s\to K^+K^-}$	0.498 ± 0.161	0.513 ± 0.145
$\hat{\xi}$ [GeV/c^2]	0.005 ± 0.003	0 constant

Table 9.2: Physical fit results obtained allowing an additive, global mass-scale to float (first column) compared with the results of the central fit, in which the mass scale is set to zero (second column).

The difference in the value of $-2\ln(\mathscr{L})$ obtained between the fit with floating mass scale and the

central fit is $-2\Delta\ln(\mathscr{L}) \simeq 2.39$, and the estimated value of ξ is just 1.6σ off zero. This indicates no statistical evidence for a shift of the mass scale in data.[2] The results with floating mass scale are consistent with those of the central fit with $\xi = 0$, which is the configuration chosen to exploit a better resolution on the physical parameters. This test shows also that the fit has sufficient statistical information to determine the global mass-scale in data with 2.8 MeV/c^2 accuracy, and with little deterioration in statistical resolution of other parameters.

9.2.3 Effect of final state electromagnetic radiation on masses and momenta

The unavoidable emission of low-energy photons from charged kaons and pions in the final states (final state radiation, FSR) may distort the distributions of kinematic variables ($m_{\pi^+\pi^-}$, α, and $p_{\rm tot}$) because of the unreconstructed photon-momentum. Since the resulting lower-mass tails of the mass distributions are not modeled in our fit, a systematic deviation in the observed ratios of yields may be present: the fit may mis-assign the contribution from decays that underwent radiation to other modes peaking at lower masses. The BGENERATOR simulator (see sec. 3.6) does not include FSR. Complementing it with available FSR-simulators, as PHOTOS [139], is of little help, because currently these packages are reliable in simulating genuine electroweak processes only, or processes fully calculable with perturbation theory within the Standard Model.

We therefore included FSR in our pseudo-experiments by using an analytical formula for leading radiation obtained directly from QED calculations [140, 141]. The expression is derived within scalar QED, under the approximation of a point-like effective weak vertex, and neglecting effects of non-*bremsstrahlung* emission. In Ref. [140] an expression for the cumulative distribution of the photon energy is derived for $K \to \pi\pi$ decays. We extrapolated the corresponding formula for $B^0_{(s)} \to \pi^+\pi^-$ and $B^0_{(s)} \to K^+K^-$ decays, following Ref. [141], and we conservatively quoted the deviation of the results from the fit without FSR as systematic uncertainty. No QED calculation was finalized yet for the $K^\pm\pi^\mp$ decays as of this analysis. The authors of Ref. [141] suggested that a reliable estimate of the effect can be obtained by averaging the expected FSR in a K^+K^- decay with the one in a $\pi^+\pi^-$ decay. We therefore simulated the radiation of $K^\pm\pi^\mp$ final states as if they were $\pi^+\pi^-$ or K^+K^- states, and we assessed a conservative systematic uncertainty as the largest deviation from the results without FSR (see tab. 9.5).

9.3 dE/dx-related systematic effects

When using the dE/dx information in fit of composition some assumptions were inevitable, yielding another class of systematic effects in our measurement.

[2] A fundamental property of the likelihood ratio test-statistics prescribes that, for sufficiently regular likelihoods and in the asymptotic limit, the quantity $-2\Delta\ln(\mathscr{L})$ between two ML estimators with a difference n in dimensionality, is distributed as a χ^2 with n degrees of freedom [138].

9.3.1 Uncertainty on the distribution of the common mode fluctuations

The observed dE/dx of kaons and pions from D^0 decays are correlated by common-mode fluctuations (see sec. 5.7), whose effect had to be included in the ML fit of composition. Since no independent way to determine the distribution of the common mode in the $B^0_{(s)} \to h^+h^{'-}$ decays is available, we used the distribution of the D^0 sample. However, the magnitude of the dE/dx correlation between particles in a decay partially depends on the kinematic correlations between the particles themselves, resulting in a sample-dependent distribution of the dE/dx correlations (see sec. 5.7). In particular, smaller dE/dx correlations are expected between $B^0_{(s)} \to h^+h^{'-}$ decay-products with respect to $D^0 \to K^-\pi^+$ products, due to the smaller kinematic correlations provided by the higher Q-value of the decay. Thus, by assuming no differences between the dE/dx correlation of two kinematically different samples as $B^0_{(s)} \to h^+h^{'-}$ and $D^0 \to K^-\pi^+$ decays we introduced a systematic effect.

In sec. 5.7 we determined the distribution of the common mode for kaons and pions from D^0 decays in two cases: one in which kaons and pions are produced in the decay of a single D^0 meson ("physical-sample"), the other one in which kaon and pion are produced in the decay of two distinct D^0 mesons, reconstructed in two consecutive events ("mixed-decay sample"). We used these independent models of the correlation function to assess a systematic uncertainty. We generated two ensembles of pseudo-experiments: one with common-mode distribution extracted from the same-decay sample, the other one with the distribution extracted from the mixed-decay sample. The former choice, which may overestimate the correlation in the $B^0_{(s)} \to h^+h^{'-}$ sample, is also the one adopted for the standard analysis on data. The latter represent a lower limit to the correlation expected in $B^0_{(s)} \to h^+h^{'-}$ decays. The variation between the results of the analysis performed in these different set of pseudo-experiments was chosen as the systematic uncertainty (see tab. 9.5).

9.3.2 Uncertainty on the distribution for electrons and protons

In sec. 5.8, we determined the dE/dx-residual distributions only for kaons and pions. Concerning electrons and protons, we included these (small) background components in the likelihood by assuming for them the pion distribution. In principle the distribution of the residual, i. e., of the observed dE/dx around its average value with correct mass assignment, is expected to be independent of the particle type. However, slight differences between the distributions of kaon and pion residuals were observed and ascribed to systematic dependence of the dE/dx on the track curvature (see sec. 5.8). One would expect similar differences between the distributions of pions, electrons, and protons. We generated four ensembles of pseudo-experiments, using in each one a different combination of pion and kaon shapes to be assigned to the proton and electron components. The largest deviation from the obtained results with respect to those obtained with the standard analysis was chosen as a conservative systematic uncertainty (see tab. 9.5).

9.3.3 Charge dependence of the distributions

Small differences between the residual distributions of positively and negatively-charged particles of the same type were observed (see fig. 5.24). In our measurement, we used the p.d.f. extracted from positively-charged particles only, possibly biasing the charge-sensitive fit parameters, as the direct

CP-asymmetry.

We generated four ensembles of pseudo-experiments; in each, a different combination of residual distributions for positively and negatively-charged particles was assigned to the outgoing particles of each kind. The largest deviation from the results obtained with the standard configuration was chosen as systematic uncertainty (see tab. 9.5).

9.4 Lifetime-related systematic effects

Lifetime-related discriminants (impact parameter and transverse decay-length) are crucial ingredients of the sample selection. The accuracy of the simulation in reproducing the distributions of these quantities is of paramount importance for a proper evaluation of the kinematic efficiencies, used to extract the ratios of branching ratios from the ratios of event yields. In case of poorly simulated lifetime-distributions of the $B^0_{(s)} \to h^+ h^{'-}$ modes (strictly speaking, the *ratios* of distributions are the relevant quantities), wrong efficiency corrections would be applied to the fit results, yielding systematically wrong measurements. Our choice of lifetimes is affected by some (modest) experimental uncertainties, and by some theoretical uncertainties coming from the unavoidable assumptions made (see sec. 3.6). Both induce a systematic uncertainty in the final ratios of branching fractions.

9.4.1 Uncertainty on the nominal $B^0_{(s)}$ meson lifetimes

In this section we assess the systematic uncertainty deriving from the experimental uncertainties on the lifetime of B^0 decays and of flavor-specific B^0_s decays only. The world-average values for these lifetimes are dominated by measurements based on large samples of semielectronic and semimuonic decays, thus having relatively small uncertainties: $\approx 4\%$ for B^0_s flavor-specific decays and $\approx 1.5\%$ for B^0 decays. We re-evaluated the kinematic efficiency for each mode (see sec. 8.2) after fluctuating by one standard deviation the input lifetime in the simulated samples. We assumed the worst cases, in which B^0 and B^0_s lifetimes fluctuated in opposite directions:

1. $c\tau(B^0) = 460 + 4 = 464$ μm and $c\tau(B^0_s \to K^-\pi^+) = 438 - 17 = 421$ μm;
2. $c\tau(B^0) = 460 - 4 = 456$ μm and $c\tau(B^0_s \to K^-\pi^+) = 438 + 17 = 455$ μm.

The resulting kinematic efficiencies were applied to the fit results.[3] The largest variation between these results and the central results was adopted as systematic uncertainty (see tab. 9.5).

9.4.2 Uncertainty on the width difference between B^0_s mass eigenstates

In addition to the above effect, the measurements involving $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+\pi^-$ modes suffers from an additional uncertainty due to the poor knowledge of the width difference between B^0_s mass eigenstates ($\Delta\Gamma_s/\Gamma_s$). The relative uncertainty on the Standard Model prediction for this

[3] Also the kinematic efficiencies involving $B^0_s \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decays are affected, since the value of $\Gamma_d = 1/\tau(B^0)$ enters in the extraction of their time-evolution (see eq. (3.4)).

quantity ($\Delta\Gamma_s/\Gamma_s = 0.12 \pm 0.06$, [110]) results in a 12 μm uncertainty on the pseudo-proper decay-lengths of $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+\pi^-$ modes. We re-evaluated the kinematic efficiency for $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+\pi^-$ modes (see sec. 8.2) after fluctuating by this quantity their input lifetimes in the simulation. The resulting kinematic efficiencies were applied to the fit results. The maximum variation between these results and the central value was chosen as systematic uncertainty (see tab. 9.5).

9.5 Systematic effects induced by charge asymmetries

An accurate understanding of all charge-related asymmetries is of critical importance for a correct estimation of the yield asymmetry between $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ modes that determines our measurement of direct CP-asymmetry. Any instrumental effect may fake an asymmetry, and the composition of background in terms of charge becomes crucial.

9.5.1 Uncertainty on the probability of interaction with matter of kaons

The $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ yields were corrected for the different probability of interaction with matter between positively and negatively-charged kaons (see sec. 8.2.2). Any uncertainty in the correction, which was extracted from Monte Carlo simulation and checked on real data, reflects itself as a systematic effect on the final $A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$ asymmetry. Although the GEANT package for detector simulation is known to reproduce accurately the different cross-sections for kaons, we allowed for a conservative systematic uncertainty. We applied to the fit result a new correction, consisting in the central one changed by the 25% of its value, and we quoted the difference with respect to the central result as a systematic uncertainty (see tab. 9.5).

9.5.2 Uncertainty on charge asymmetries in background

The form of the likelihood function implies that each species populates the background with equal proportions of positively and negatively-charged particles, i. e., the estimated fraction of each particle type ($\hat{f}_\pi$, $\hat{f}_K$, $\hat{f}_\mathrm{p}$, and $\hat{f}_e$) indicates the charge-averaged fraction of that type. If, instead, the background contains a different number of positively and negatively-charged particles of some types, possible fit biases may appear. Such a charge asymmetry of the background may mimic an asymmetry of the $B^0 \to K^+\pi^-$ signals, inducing a systematic effect on the measured $A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$ asymmetry. To check the hypothesis of measurable charge asymmetries in the background, we repeated the fit of composition on *data* allowing for independent fractions of positively and negatively-particles of each type in background. This resulted in nine additional parameters to be fit. The corresponding (physics related) fit results are shown in the first column of tab. 9.3. The results of the central fit are also reported for comparison. The difference between the value of $-2\ln(\mathscr{L})$ obtained with the central fit and the one from the fit with independent background fractions for positively and negatively-charged particles was $-2\Delta\ln(\mathscr{L}) \simeq 16.54$ (for nine additional degrees of freedom). In addition, all differences between fractions of oppositely-charged particles of the same type are largely consistent with zero within 1.5σ. Thus, no statistical evidence for charge asymmetries in the background is present in

Quantity	+/− bckg. split	Central fit
$\hat{f}_{B^0\to\pi^+\pi^-}$	0.135 ± 0.030	0.134 ± 0.030
$\hat{f}_{B^0\to K^+\pi^-}$	0.599 ± 0.034	0.600 ± 0.034
$\left(\hat{f}_{\overline{B}^0\to K^-\pi^+} - \hat{f}_{B^0\to K^+\pi^-}\right) / \left(\hat{f}_{\overline{B}^0\to K^-\pi^+} + \hat{f}_{B^0\to K^+\pi^-}\right)$	-0.073 ± 0.086	-0.022 ± 0.078
$\hat{f}_{B_s^0\to K^-\pi^+}$	0.004 ± 0.028	0.003 ± 0.028
$\hat{f}_{B_s^0\to K^+K^-}$	0.262 ± 0.035	0.262 ± 0.035
$\hat{f}_{B_s^0\to K^+K^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.437 ± 0.074	0.437 ± 0.074
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B^0\to K^+\pi^-}$	0.226 ± 0.057	0.224 ± 0.057
$\hat{f}_{B^0\to\pi^+\pi^-}/\hat{f}_{B_s^0\to K^+K^-}$	0.517 ± 0.146	0.513 ± 0.145

Table 9.3: Physical fit results obtained allowing separate fractions for positively and negatively-charged particles in background (first column) compared with the results of the central fit (second column).

our data.[4] Furthermore, the physics parameters show negligible variations between the two fits (see tab. 9.3): all differences amount to fractions of a standard deviation except for the differential rate between $B^0 \to K^+\pi^-$ and $\overline{B}^0 \to K^-\pi^+$ decay, which varied by about half of a standard deviation. However, the considerable degradation of its statistical resolution (7.8% → 8.6%) indicates that the shift on the estimate could be ascribed to statistical reasons.

Although the above results show no evidence of charge asymmetries in background, we still assigned a conservative systematic uncertainty on possible undetected charge-asymmetries of protons. Only protons and antiprotons were considered because of the possible contribution to the background of beam-gas interaction products, which contain more protons than antiprotons.[5] We therefore repeated *in data* a fit of composition allowing for separate fractions of protons and antiprotons. This determined the following $\mathrm{p\bar{p}}$ asymmetry in the signal region: $\hat{A}(\mathrm{p}) = (\hat{f}_{\mathrm{p}} - \hat{f}_{\bar{\mathrm{p}}})/(\hat{f}_{\mathrm{p}} + \hat{f}_{\bar{\mathrm{p}}}) = 0.06 \pm 0.05$. The results of this fit were used to generate an ensemble of pseudo-experiments to which the central analysis was applied. The difference between fit results obtained in samples with and without $\mathrm{p\bar{p}}$ asymmetry was used as systematic uncertainty (see tab. 9.5).

9.6 Likelihood-related systematic effects

An important class of systematic effects involves the expression of the likelihood function. If the dependencies among observable are improperly treated (i. e., the p.d.f. of two dependent observables are factorized), biases on the fit results may occur and invalidate the measurement. Even subtler effects arise when the fit uses an observable that has different distributions among different classes of events, which are not included in the p.d.f. [127].

[4] See footnote at pag. 194.
[5] See footnote at pag. 160

9.6.1 Effect of the dependence between the observables $m_{\pi^+\pi^-}$ and $p_{\rm tot}$

In the likelihood, we assumed the invariant $\pi^+\pi^-$-mass ($m_{\pi^+\pi^-}$) being independent of the scalar sum of momenta ($p_{\rm tot}$):

$$\mathscr{L}_i(\vec{\theta}) \sim \wp^m(m_{\pi^+\pi^-}|\alpha;\vec{\theta}) \times \wp^p(\alpha, p_{\rm tot};\vec{\theta}) \times \wp^{\rm PID}(\kappa_1^{\rm obs},\kappa_2^{\rm obs}|\alpha, p_{\rm tot};\vec{\theta}). \tag{9.2}$$

While this applies with good accuracy to the signal, it is an approximation for the background. Figure 9.1 shows the normalized invariant $\pi^+\pi^-$-mass distributions for background events with $p_{\rm tot} <$

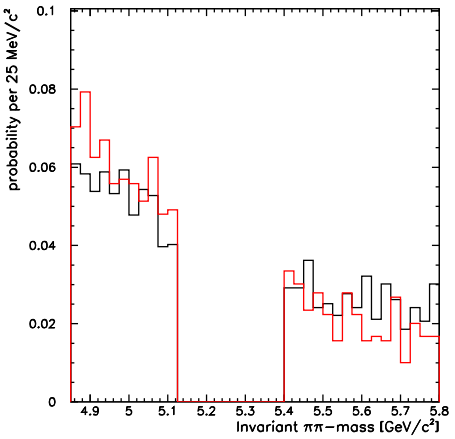

Figure 9.1: Normalized invariant $\pi^+\pi^-$-mass distribution of background events that pass the optimized selection and with $p_{\rm tot} < 9$ GeV/c (black plot) and $p_{\rm tot} > 9$ GeV/c (red plot).

9 GeV/c (black histogram) and with $p_{\rm tot} > 9$ GeV/c (red histogram). The shape slightly depends on $p_{\rm tot}$, thus, the factorized form of eq. (9.2) is not completely accurate, and may introduce undesired biases in the fit results. We determined the slopes of the exponential functions used to model the two distributions in fig. 9.1 with a binned χ^2-fit, and estimated a 32% relative difference between them. Then we used these shapes to generate two ensembles of pseudo-experiments, and we applied the standard fit of composition to these simulated samples. The maximum deviation between the observed fit results was taken as a systematic uncertainty (see tab. 9.5).

9.6.2 Uncertainty on the momentum distributions of background components

In sec. 6.2.2 we determined an inclusive momentum ($\alpha, p_{\rm tot}$) distribution for the background, using events in the mass sidebands of data. This provided an average p.d.f., consisting in the superimposition of the momentum distributions of an unknown admixture of protons, kaons, pions, muons, and electrons. As a general rule, whenever the model used in a multi-component fit depend on additional observables, one should always use the correct, complete likelihood expression, including the explicit distributions of all observables for all classes of events because, if they are different, significant biases

of the fit results may occur [127]. One of the discriminating observables used in the ML fit is the scalar sum of the momenta of the two outgoing particles ($p_{\rm tot}$), which is clearly sensitive to differences in momentum spectra between the different background components. Since in our central fit the likelihood function implicitly assumed the same momentum distribution for all background components, we need to investigate the possible biases induced by this approximation.

We studied the momentum distribution of each type of particle in the background of the $B^0_{(s)} \to h^+h'^-$ sample. To increase the number of background events, we released the requirement on isolation in the selection (see tab. 3.3). Then we defined a sample of background particles: i. e., the tracks from candidates with lower, $4.800 < m_{\pi^+\pi^-} < 5.125$ GeV/c^2, or higher, $5.4 < m_{\pi^+\pi^-} < 5.8$ GeV/c^2, $\pi^+\pi^-$-mass than the $B^0_{(s)} \to h^+h'^-$ signal mass. We further subdivided this sample in five subsamples according to the momentum of the outgoing particles. A measurement of the composition of each subsample may allow extracting useful information on the momentum spectra of each component. At fixed momentum, the distribution of the observed dE/dx around the average pion response is sensitive to the fraction of different species of charged particles of the sample. The available dE/dx resolution, comparable in size to the separation among the average responses for each particle type, along with the unavoidable momentum-spread of particles within each subsample, complicate the task. However, some useful information is still available, especially if the study is restricted to momentum ranges in which a reasonably high statistics is combined with well-separated dE/dx curves. We fit the composition of each subsample on a single-particle basis in terms of pion, kaon, proton, and electron components with a binned ML fit of the $\delta_\pi^{\rm obs}$ (i. e., residual with pion hypothesis) distribution. The shapes of the $\delta_\pi^{\rm obs}$ distributions for each particle type were fixed in the fit, only their relative normalization were allowed to float for determining the composition. Figure 9.2 shows an example of one fit. From the results on the momentum-dependent composition of background (see tab. 9.4), we

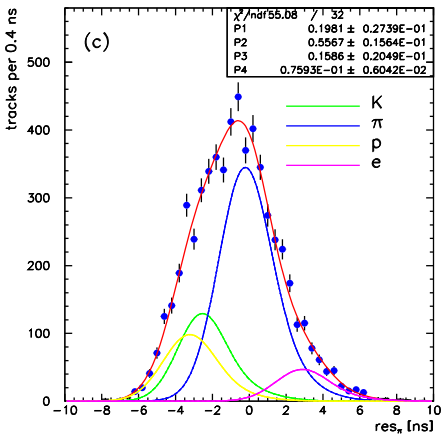

Figure 9.2: Distribution of the observed dE/dx around the average pion response for charged particles with $4 < p < 5$ GeV/c from background.

derived the momentum distributions of each particle type shown in fig. 9.3. We fit these distribution

p [GeV/c]	Entries	$\hat{f}_K$ [%]	$\hat{f}_\pi$ [%]	$\hat{f}_p$ [%]	$\hat{f}_e$ [%]
2.808–3.500	5807	16.8 ± 6.4	54.5 ± 1.5	23.3 ± 5.8	4.2 ± 0.4
3.500–4.000	4180	14.0 ± 4.6	55.8 ± 1.8	23.8 ± 3.8	5.2 ± 0.6
4.000–5.000	5923	19.7 ± 2.7	55.6 ± 1.6	15.9 ± 2.0	7.7 ± 0.6
5.000–5.800	2478	23.0 ± 3.4	57.1 ± 2.5	12.7 ± 2.3	5.7 ± 1.0
5.800–9.000	2506	29.0 ± 2.8	54.6 ± 2.7	5.5 ± 1.5	8.4 ± 1.2

Table 9.4: Momentum-dependent composition of background.

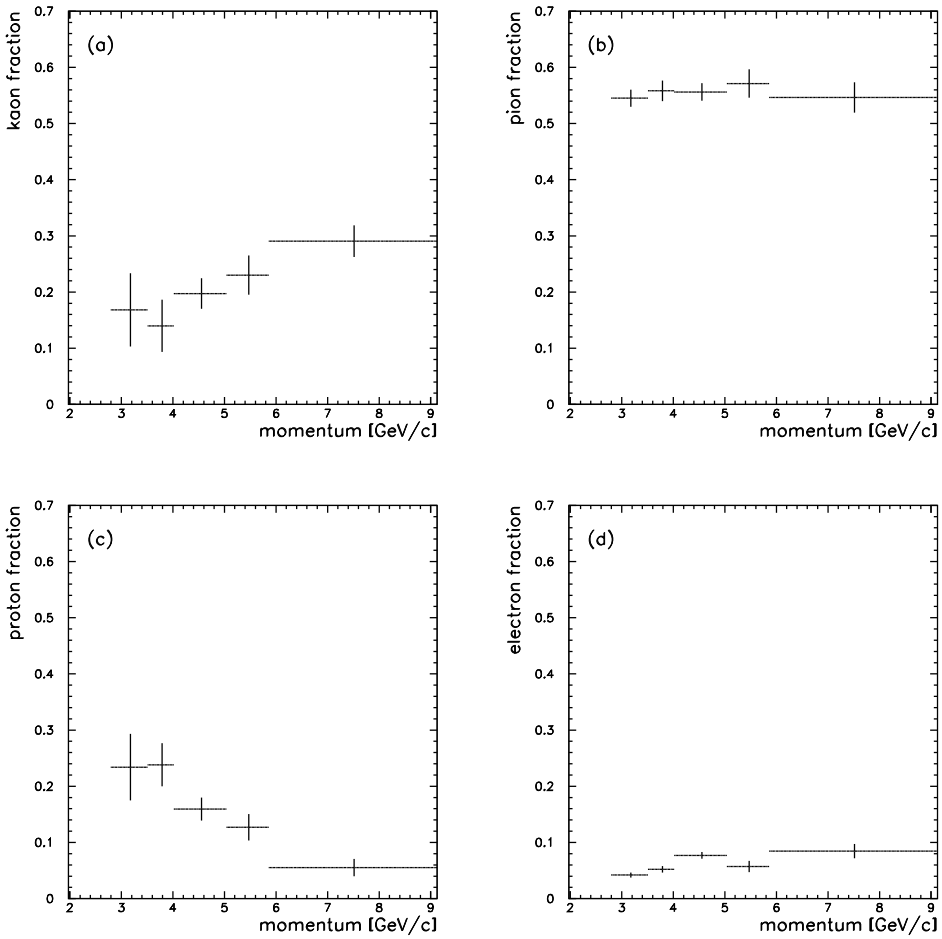

Figure 9.3: Fraction of kaons (a), pions (b), protons (c), and electrons (d) in background as a function of momenta.

with straight line functions, and we generated ensembles of pseudo-experiments in which the fraction of each particle species was reweighed as a function of the momentum, according to the fit results.

We quoted a conservative systematic uncertainty by using the difference between the fit results on pseudo-experiments reweighed with the momentum spectra of fig. 9.3, and pseudo-experiments in which the slopes of the reweighing functions were inverted. The resulting effect is small in comparison with other systematic effects (see tab. 9.5).

The results of tab. 9.4 have a physical interest in their own right. The capability of measuring fractions of individual species, and their momentum distributions in the background of such an aggressively selected sample may be exploited in a future, dedicate analysis for extracting insights on the different sources of background, or on additional, unexpected signals.

9.7 Efficiency-related systematic effects

In chap. 8, the ratios of observed decays were corrected for the relative efficiencies between the individual $B^0_{(s)} \to h^+h'^-$ modes. However, the relative efficiencies were determined in samples of finite statistics. The Poisson fluctuations of the events yields used to determine the efficiencies introduce a systematic effect in the final measurements of ratios of branching fractions.

9.7.1 Uncertainty on the isolation efficiency

In sec. 8.5, the relative isolation efficiency was determined with $\approx 10\%$ statistical uncertainty, which introduces a systematic uncertainty in measurements of ratios of branching fractions between B^0_s and B^0 decays. We re-evaluated the ratios of branching ratios after fluctuating the relative isolation efficiency by one standard deviation in either directions. The difference between the resulting branching fraction and the central result was quoted as systematic uncertainty (see tab. 9.5).

9.7.2 Effect of Poisson fluctuations in simulated samples

The relative kinematic efficiencies, used to convert the ratios of event yields in ratios of branching fractions, were determined within $\mathcal{O}(10\%)$ statistical uncertainties (see sec. 8.2). We re-evaluated each ratio of branching fractions by using acceptance corrections fluctuated by one standard deviation in either direction. The difference between the resulting branching fraction and the central result was used as systematic uncertainty (see tab. 9.5).

9.7.3 Uncertainty on the XFT-bias correction

In sec. 8.3, the ratios of observed decays were corrected for the different XFT triggering-efficiency for kaons and pions. The parameters of the functions used for this correction were known within statistical uncertainties, which induce a systematic uncertainty on our results. We re-evaluated the correction factors by independently fluctuating the coefficients of the polynomials of tab. 8.3, neglecting their correlations. We evaluated the efficiencies in the limiting (worst) cases in which the coefficients for pions and for kaons fluctuated in opposite directions by one standard deviation. This provides conservative variations of the efficiencies, which are upper limits to the ones expected including the

correlations between coefficients. The largest difference between the resulting branching fraction and the central result was quoted as a symmetric systematic uncertainty (see tab. 9.5).

9.8 Other systematic uncertainties

9.8.1 Uncertainty on the transverse momentum spectra of B_s^0 and B^0 mesons

The simulated samples were generated using a single $p_{\mathrm{T}}(B)$-spectrum for B^0 and B_s^0 mesons, extracted from an unknown admixture of b-hadrons reconstructed in CDF data (see sec. 3.6). Small differences between the transverse momentum distributions of B^0 and B_s^0 mesons are possible, because of the differences in their fragmentation processes. The effect of these differences in the present analysis might be two-fold: on one side, the modeling of the mass and momentum terms of the p.d.f. for the signal (see chap. 6) would be incorrect, possibly resulting in biases of the fit results; on the other hand, the ratios of kinematic efficiencies between B^0 and B_s^0 decay modes, extracted form the simulation, would be affected.

The currently available experimental data do not allow for an independent extraction of the B_s^0 fragmentation function. Some assumptions are needed for a semi-quantitative evaluation of this effect. It is reasonable to expect the fragmentation function of the B_s^0 meson to be softer compared with the one of the B^0 meson, since the b-quark produces an heavier quark from vacuum to hadronize in a B_s^0 meson. The mean value of the fraction of b-quark momentum inherited by the meson is expected to vary by approximately $\Delta z \approx \frac{m_{B_s^0} - m_{B^0}}{m_{B_s^0}/2 + m_{B^0}/2} \approx 2\%$ between B^0 and B_s^0 mesons [142].

We modified the $p_{\mathrm{T}}(B)$ spectrum in the Monte Carlo with a $\pm 1\%$ shift and we re-extracted the mass and momentum distributions for the signals. No significant difference within the statistical uncertainties was found with respect to the distributions of the standard analysis, thus we estimated this effect to be negligible. In addition, the simulated samples with adjusted spectra were used to re-evaluate the kinematic efficiencies for each mode, assuming the worst case in which the standard spectrum was shifted by 1% toward higher (lower) momenta for the B^0 (B_s^0) modes. The difference between the resulting ratios of branching fractions and the central results were quoted as systematic uncertainty (see tab. 9.5).

9.8.2 Uncertainty on the mass distribution of background

Since we empirically assumed an exponential model for the invariant $\pi^+\pi^-$-mass distribution of background candidates (see eq. (6.15)), we quoted a systematic uncertainty due to our limited knowledge of the real distribution. We first fit the invariant $\pi^+\pi^-$-mass distribution *in data*, trying various different empirical models: polynomial functions of various degrees, combinations of Gaussian functions, and so forth. We excluded all models that returned a poor quality of the fit (χ^2-test), retaining only the second and third degree polynomial functions. Then, we generated two ensembles of pseudo-experiments in which the invariant $\pi^+\pi^-$-mass of background events assumed these shapes, whose parameters had been previously extracted from data. Finally, we applied the full fit to these pseudo-experiments. The largest difference between the fit results obtained with alternative models and those obtained with the standard exponential model was quoted as systematic uncertainty (see tab. 9.5).

9.9 Systematic uncertainty summary

Table 9.5 contains the contribution of each effect to the systematic uncertainty associated to each measurement. Each systematic effect can be considered independent from the others. Thus, we evaluated the total systematic uncertainty for each measurement by adding in quadrature the individual contributions.

Source	$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)}$	$A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$	$\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)}$
Mass resolution	0.004	0.0015	0.002
Nominal $B^0_{(s)}$ masses	0.022	0.0033	0.010
Final state radiation	0.020	–	0.025
dE/dx: common-mode shape	0.026	0.0016	0.013
dE/dx: p and e residuals	0.001	0.0003	0.000
dE/dx: charge-dependent shapes	0.005	0.0052	0.001
Nominal $B^0_{(s)}$ lifetimes	0.004	–	–
Effect of $\Delta\Gamma_s/\Gamma_s$	0.006	–	–
Kaon charge-asymmetry	–	0.0009	–
$\mathrm{p\bar{p}}$ asymmetry	0.000	0.0097	0.000
$m_{\pi^+\pi^-} - p_{\mathrm{tot}}$ dependence	0.010	0.0018	0.002
Momenta in background	0.001	0.0005	0.000
Isolation efficiency	0.047	–	–
MC statistics	0.004	0.0004(∗)	0.003
XFT-bias correction	0.009	–	0.003
$p_{\mathrm{T}}(B)$ spectrum	0.006	–	–
Background model	0.011	0.0022	0.003
Total	0.07	0.012	0.03

Table 9.5: Summary of systematic uncertainties. The uncertainty on the asymmetry due to Poisson fluctuations of the Monte Carlo sample labelled with an asterisk (∗) is smaller that other uncertainties on the same row, because for this uncertainty we used a dedicated, high-statistics sample of simulated charged kaons (see sec. 9.5.1).

The dominant sources of systematic uncertainties vary among the measured quantities. The largest contributions are the uncertainty on the isolation efficiency, which is the dominant uncertainty for the $(f_s/f_d) \times \mathcal{B}(B_s^0 \to K^+K^-)/\mathcal{B}(B^0 \to K^+\pi^-)$ measurement, the effect of possible asymmetries between the fraction of protons and antiprotons in background, which dominates the $A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$ measurement, and the effect of final state radiation for the $\mathcal{B}(B^0 \to \pi^+\pi^-)/\mathcal{B}(B^0 \to K^+\pi^-)$ measurement. Further discussion of the systematic uncertainties is given in the next chapter.

Chapter 10

Results and their interpretation

We reconstructed, for the first time in hadron collisions, decays of b-mesons into two-body, charmless pseudo-scalar mesons, and we measured their branching fractions. We observed a new mode and we set improved upper limits on two unobserved modes. In all measurements, uncertainties are dominated by the statistical component. This chapter contains the final results, their discussion, and the perspectives of future $B^0_{(s)} \to h^+h'^-$ measurements at CDF.

10.1 Observation of the $B^0_s \to K^+K^-$ decay-mode

We report the first observation of the decay mode $B^0_s \to K^+K^-$, with a yield of 236 ± 32 events, corresponding to the following measurement of relative branching fraction:

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.46 \pm 0.08\ (stat.) \pm 0.07\ (syst.), \tag{10.1}$$

where f_s/f_d is the ratio of production fractions of B^0_s and B^0 mesons from the hadronization of a b-quark in $p\bar{p}$ collisions. This is the first observation of the decay of a B^0_s meson into a pair of pseudo-scalar mesons. Using the world-average values $\mathcal{B}(B^0 \to K^+\pi^-) = (18.9 \pm 0.7) \times 10^{-6}$ and assuming for f_s/f_d the world-average value from $p\bar{p}$ and e^+e^- collisions, $f_s/f_d = 0.261 \pm 0.038$ [29], we extract the following absolute branching fraction:

$$\mathcal{B}(B^0_s \to K^+K^-) = [33 \pm 6\ (stat.) \pm 7\ (syst.)] \times 10^{-6}, \tag{10.2}$$

which is compatible with the current world-best limit from the ALEPH Collaboration: $\mathcal{B}(B^0_s \to K^+K^-) < 59 \times 10^{-6}$ [67].

The contributions of the systematic and statistical components to the uncertainty are approximately equal. A sizable reduction of the systematic uncertainty is expected, along with the statistical one, as the data samples increase in size: the dominant sources of systematic uncertainty include the statistical uncertainty on the isolation efficiency, the statistical uncertainty on the nominal values of $B^0_{(s)}$ meson masses used in the likelihood, and the uncertainty on the effect of final state radiation (see fig. 10.1). The first two contributions are of statistical origin, thus expected to decrease as the

size of calibration samples used to determine them increases. The FSR-related uncertainty will be reduced in the next update of the analysis, since Isidori and Baracchini recently provided the QED corrections also for $B^0 \to K^+\pi^-$ modes [141]. This will allow incorporating the effect of FSR as a further efficiency correction to the central result. For comparison with theoretical predictions, it is

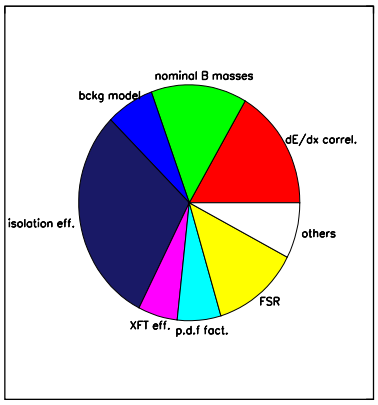

Figure 10.1: Circular chart of the systematic uncertainties (σ_i) contributing to the measurement of $(f_s/f_d) \times \mathcal{B}(B_s^0 \to K^+K^-)/\mathcal{B}(B^0 \to K^+\pi^-)$. The area of each sector is proportional to σ_i^2.

convenient to relate the $B_s^0 \to K^+K^-$ rate also to the rate of the U-spin conjugate mode $B^0 \to \pi^+\pi^-$:

$$\frac{f_d}{f_s} \times \frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to K^+K^-)} = 0.45 \pm 0.13\ (stat.) \pm 0.06\ (syst.). \tag{10.3}$$

10.1.1 A test of SU(3) symmetry breaking

Khodjamirian, Mannel and Melcher provide *quantitative* estimates of the flavor SU(3)-symmetry breaking in $B^0_{(s)} \to h^+h^{'-}$ decays [32]. This crucial information defines the limits of applicability of the SU(3) symmetry approach proposed to reduce and control hadronic uncertainties in the predictions of $B^0_{(s)} \to h^+h^{'-}$ amplitudes [5]. Our measurement provides a first experimental check of this prediction.

Amplitudes are written in terms of factorizable portions ($\mathcal{A}_{\rm fact}$, function of CKM factors, short distance operators, decay constants and form factors), multiplied by non-factorizable corrections (sums of terms suppressed by either α_s or inverse powers of the b-quark mass). Following an usual phenomenological approach, Khodjamirian et al. evaluate ratios of amplitudes, so that the magnitude of SU(3) violation becomes related to the ratio of the kaon and pion decay constants (f_K/f_π), to the ratio of $B \to K$ and $B \to \pi$ form factors, and to the ratio of non-factorizable terms. All these factors are calculated using the QCD sum-rules approach [25].

Neglecting the CKM factors, the authors calculate the ratio between the factorizable parts of

$B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ (or $B^0 \to K^+\pi^-$) amplitudes: $|\mathcal{C}'/\mathcal{C}|_{\text{fact.}}$ (see sec. 1.5.2).[1] They show that this is strongly affected by SU(3) violation, obtaining

$$\left|\frac{\mathcal{C}'}{\mathcal{C}}\right|_{\text{fact.}} = 1.52^{+0.18}_{-0.14}, \quad \text{and} \quad \frac{f_\pi}{f_K} \times \left|\frac{\mathcal{C}'}{\mathcal{C}}\right|_{\text{fact.}} = 1.25^{+0.14}_{-0.12}. \tag{10.4}$$

When comparing $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ amplitudes, U-spin symmetry predicts $|\mathcal{C}'/\mathcal{C}| = 1$ with no assumptions. However, extracting the parameter $|\mathcal{C}'/\mathcal{C}|$ from our measurement would require assumptions on the value of the hadronic parameters d and θ and on the CKM angle γ (see eq. (1.79)). These, combined with the experimental uncertainty of our result provide little information on the compatibility with the above predictions. Comparing $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ amplitudes, instead, allows a comparison without assumptions on the values of d, θ, and γ. The only assumption, in this case, is that exchange and penguin-annihilation topologies, which only affect the $B_s^0 \to K^+K^-$ amplitude, play a minor role. Our measurement of $\mathcal{B}(B_s^0 \to K^+K^-)$ allows the first check of eq. (10.4) with experimental data, yielding

$$\left(\frac{f_\pi}{f_K} \times \left|\frac{\mathcal{C}'}{\mathcal{C}}\right|\right)_{\text{CDF}} \simeq 1.8 \pm 0.4. \tag{10.5}$$

The experimental accuracy is still too poor for a statistically significant discrimination between the hypotheses of large U-spin violation or exact U-spin validity. However, the observed $B_s^0 \to K^+K^-$ to $B^0 \to K^+\pi^-$ amplitude ratio exceeds the predictions of Khodjamirian et al., suggesting large U-spin breaking in this process.

10.1.2 Comparison with theoretical predictions

The 27% uncertainty on our result is not sufficiently small to rule out any of the currently prominent theoretical predictions, which have comparable uncertainties.

Good consistency is found with the QCDF-based predictions of Beneke and Neubert, $\mathcal{B}(B_s^0 \to K^+K^-) = (28\text{–}36) \times 10^{-6}$ [24] and with a recent prediction based on the SCET, $\mathcal{B}(B_s^0 \to K^+K^-) = (18.2 \pm 6.8) \times 10^{-6}$ [143]. On the other hand, our result disfavors an alternative QCDF-based prediction by Sun, Zhu, and Du, that includes "chirally enhanced" power corrections and weak annihilation contributions: $\mathcal{B}(B_s^0 \to K^+K^-) = (7\text{–}17) \times 10^{-6}$ [144].

A satisfactory agreement is found also with the predictions based upon U-spin applied to $B^0 \to \pi^+\pi^-$ data by Buras, Fleischer, Recksiegel, and Schwab. They predict $\mathcal{B}(B_s^0 \to K^+K^-) = (35 \pm 7) \times 10^{-6}$ neglecting annihilation and exchange topologies, and partially including the effect of SU(3) violation: they use the value $\mathcal{C}'/\mathcal{C} = 1.76^{+0.15}_{-0.17}$ calculated by Khodjamirian ([32], see previous section), but neglect U-spin violation in the non-factorizable amplitudes [145]. Another recent, SU(3)-based prediction by Chiang and Zhou, $\mathcal{B}(B_s^0 \to K^+K^-) \approx (19 \pm 1\text{–}4) \times 10^{-6}$, is in reasonable agreement with our result [146].

Probably, the most interesting comparison is with the predictions of Descotes-Genon, Matias, and Virto that recently proposed a new approach to the calculation of $\mathcal{B}(B_s^0 \to K^+K^-)$ that complements the benefits of the QCD factorization expansion (QCDF) with those of the SU(3) symmetry relations

[1] The non-factorizable part involves complications from penguin and annihilation topologies and is neglected.

to control, for the first time in literature, the $1/m_b$ corrections [37]. The measured rate of the $B^0 \to K^0\bar{K}^0$ process [147] is connected to the $B_s^0 \to K^+K^-$ rate, using a combination of U-spin and isospin arguments and fully evaluating SU(3) breaking within QCDF. The resulting prediction for the branching fraction is $\mathcal{B}(B_s^0 \to K^+K^-) = (20 \pm 9) \times 10^{-6}$ within the Standard Model.[2] This method benefits from unprecedented theoretical precision, the uncertainty being dominated by the experimental uncertainty on $\mathcal{B}(B^0 \to K^0\bar{K}^0)$, which is expected to be further reduced with future measurements from the B-Factories. A further refinement by Baek, London, Matias, and Virto allows the prediction of $\mathcal{B}(B_s^0 \to K^+K^-) = (17 \pm 7) \times 10^{-6}$, valid under the assumption of positive direct CP-violating asymmetry in $B^0 \to K^0\bar{K}^0$ decays, $A_{\mathsf{CP}}^{\rm dir}(B^0 \to K^0\bar{K}^0) > 0$ [148]. Our result is largely consistent with both these predictions.

10.1.3 A probe for new physics and for the CKM phase γ

Recently, results of increasingly-high precision from the B-Factories suggest several mild discrepancies with the Standard Model predictions. Although not statistically significant, all these discrepancies hint at possible new physics arising in $\bar{b} \to \bar{s}$ transitions. Examples include the $\approx 2\sigma$ discrepancy between CP asymmetries in $\bar{b} \to \bar{s}q\bar{q}$ (where $q = u, d, s$) and $\bar{b} \to \bar{s}c\bar{c}$, expected to be approximately equal in the Standard Model, the "$B \to K\pi$ puzzle", etc. [3] The $B_s^0 \to K^+K^-$ decay, being governed by the $\bar{b} \to \bar{s}$ transition, is a natural candidate for investigating other possible discrepancies.

Following the strategy discussed in sec. 1.5.2, we investigated the impact of the measured $\mathcal{B}(B_s^0 \to K^+K^-)$ on the CKM picture by studying the compatibility of our measurement with allowed region in the space of $B_{(s)}^0 \to h^+h^{'-}$ observables. The comparison between allowed regions in the $[A_{\mathsf{CP}}^{\rm dir}(B^0 \to \pi^+\pi^-), R_d^s]$ space and the value of R_d^s *directly* measured in this thesis, provides an immediate test of the validity of Standard Model. A significant discrepancy between the values of R_d^s obtained with the two methods would indicate new physics. On the other hand, in case of consistency, a precise measurement of the observables would provide valuable information on the value of the CKM phase γ.

The derivation of allowed regions is detailed in sec. 1.5.2; in the following we list the values of experimental and theoretical input used for the comparison. The experimental inputs are the CP-violating asymmetries in the $B^0 \to \pi^+\pi^-$ decay measured by the Belle Collaboration, $A_{\mathsf{CP}}^{\rm dir} = -0.55 \pm 0.09$ and $A_{\mathsf{CP}}^{\rm mix} = -0.61 \pm 0.11$ [149], those measured by the *BABAR* Collaboration $A_{\mathsf{CP}}^{\rm dir} = -0.16 \pm 0.11$ and $A_{\mathsf{CP}}^{\rm mix} = -0.53 \pm 0.14$ [62], the world-average value of the phase $\phi_d = 2\beta = 43°$ as measured at the B-Factories using $B^0 \to J/\psi K_S^0$ decays [29], and the value of $R_d^s = 8.5 \pm 2.9$ directly measured in this thesis. The theoretical inputs are the range $58° \leq \gamma \leq 72°$ for the V_{ub} phase, and the following ranges for the SU(3)-violating parameters $1.49 \leq |\mathcal{C}'/\mathcal{C}| \leq 1.91$, $0.8 \leq d'/d \leq 1.2$, and $-40° \leq \theta' - \theta \leq 40°$.

Figure 10.2 shows the comparison between the value of R_d^s indirectly determined from the measurement of asymmetries at the B-Factories [149, 62], and the direct value measured in this thesis.

[2] The uncertainties quoted by Matias et al., are added in quadrature.

[3] In the Standard Model, the isospin-symmetry allows accurate predictions of ratios of branching fractions between different $B \to K\pi$ modes. The predictions are confirmed by the experimental data for the decays in which EW penguins may only be color suppressed, whereas (moderate) inconsistencies are found for modes in which EW penguin amplitudes are not suppressed. This phenomenon, frequently referred to as "$B \to K\pi$ puzzle" might be explained if non-Standard Modelphases would enhance the contribution of EW penguin amplitudes.

All inputs are allowed to vary within 1σ from their central values.

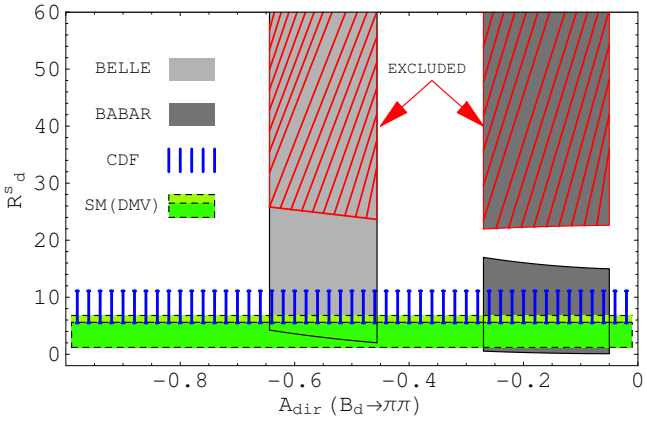

Figure 10.2: Allowed regions in the $[A_{\mathsf{CP}}^{\mathrm{dir}}(B^0 \to \pi^+\pi^-), R_d^s]$ space as resulting from Belle and *BABAR* measurements compared with the direct measurement of R_d^s. The dark gray zone corresponds to *BABAR* inputs, the light gray zone corresponds to the Belle inputs and the blue-hatched band corresponds to the direct measurement of this thesis. The light green band indicates the Standard Model value of R_d^s predicted in Ref. [37]. The improved prediction for R_d^s, obtained under the additional hypothesis that $A_{\mathsf{CP}}^{\mathrm{dir}}(B^0 \to K^0\bar{K}^0) > 0$ [148], is indicated by the dark green band. The red-hatched region is excluded using a conservative constraint on the penguin parameter d', extracted from the measurement of $\mathcal{B}(B^0 \to K^+\pi^-)$. *The plot is an updated version of the plot shown in fig. 3 of Ref. [36] and was kindly provided by Joaquim Matias for this thesis.*

The present measurement of $\mathcal{B}(B_s^0 \to K^+K^-)$ is fully compatible with the allowed regions obtained from the measurements at the B-Factories, excluding any indication of physics beyond the Standard Model at this level. The experimental and theoretical uncertainties are still too large for extraction of valuable information on the CKM phase γ. However, this method will benefit from increasingly-improved experimental and theoretical uncertainties allowing search of physics beyond the Standard Model in these decays, until the full power of measurements of CP-violating $B_s^0 \to K^+K^-$ decay asymmetries will be available.

10.2 Measurement of CP-asymmetry in the $B^0 \to K^+\pi^-$ decay

From the partial rate asymmetry between $\overline{B}^0 \to K^-\pi^+$ and $B^0 \to K^+\pi^-$ decays, we extracted the measurement of the direct CP-violating asymmetry.

Following the usual procedure, the flavor of the decaying meson was not identified. The asymmetry was derived assuming all $K^+\pi^-$ final states (from the decay of a particle with mass consistent with the B^0 meson mass) originated from B^0 decays, and all $K^-\pi^+$ states (from the decay of a particle with mass consistent with the B^0 meson mass) originated from $\overline{B}^0$ decays. We therefore neglected the contribution of doubly-Cabibbo-suppressed (DCS) decays ($B^0 \to K^-\pi^+$ and $\overline{B}^0 \to K^+\pi^-$), of asymmetries in flavor mixing ($B^0 \to \overline{B}^0 \to K^-\pi^+$ and $\overline{B}^0 \to B^0 \to K^+\pi^-$), and of their combined

effect. The contribution of DCS decays becomes relevant only in presence of direct CP asymmetry in the DCS modes. However, even in the extremely unlikely case of unit asymmetry, the size of the effect would be $\mathcal{O}(‰)$, i. e., not appreciable at the current level of experimental accuracy. The effect of flavor mixing is proportional to the size of the asymmetry in the mixing (i. e., $\propto \int_0^T N(B^0 \to \overline{B}^0)dt - \int_0^T N(\overline{B}^0 \to B^0)dt$), which is expected significantly smaller than our current experimental uncertainty. Even smaller is the effect of a possible asymmetry arising in the interference between decay and mixing due to $K^+\pi^-$ and $K^-\pi^+$ final states being accessible to both B^0 and $\overline{B}^0$ mesons through the DCS processes. The measured asymmetry is

$$A_{CP}(B^0 \to K^+\pi^-) \equiv \frac{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) - \mathcal{B}(B^0 \to K^+\pi^-)}{\mathcal{B}(\overline{B}^0 \to K^-\pi^+) + \mathcal{B}(B^0 \to K^+\pi^-)} = -0.013 \pm 0.078\ (stat.) \pm 0.012\ (syst.). \tag{10.6}$$

This result is compatible with the current results from the Belle, $A_{CP}(B^0 \to K^+\pi^-) = -0.093 \pm 0.018\ (stat.) \pm 0.008\ (syst.)$ [64], and the *BABAR*, $A_{CP}(B^0 \to K^+\pi^-) = -0.108 \pm 0.024\ (stat.) \pm 0.008\ (syst.)$ [62] experiments, although it is also compatible with zero, within its 8% uncertainty. Our measurement is not competitive with the recent B-Factories results, still it is the third best measurement, more precise than the (seven years old) CLEO measurement, $A_{CP}(B^0 \to K^+\pi^-) = -0.04 \pm 0.16\ (stat.) \pm 0.02\ (syst.)$ [61].

The systematic uncertainty is promising for future extension of this measurement to samples of increased size. The main contributions to this uncertainty are the effects of possible charge asymmetries in background and in the dE/dx response (see fig. 10.3). Although it may be difficult to strongly reduce these contribution in future updates of the measurement, their impact is already $\mathcal{O}(1\%)$, a value comparable with the uncertainties quoted in recent B-Factories measurements. On the other hand, the statistical uncertainty, which currently is the limiting factor in this measurement, is expected to reduce by a factor of at least three with the sample already available at CDF, and will further decrease as CDF keeps accumulating data.

10.3 Measurement of $\mathcal{B}(B^0 \to \pi^+\pi^-)/\mathcal{B}(B^0 \to K^+\pi^-)$

From the observed yields of 121 ± 27 $B^0 \to \pi^+\pi^-$ decays and 542 ± 30 $B^0 \to K^+\pi^-$ decays, we measured the following ratio of branching fractions:

$$\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.21 \pm 0.05\ (stat.) \pm 0.03\ (syst.). \tag{10.7}$$

This result is compatible with the measurements available at the B-Factories:[4] $\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.26 \pm 0.01\ (stat.) \pm 0.01\ (syst.)$ from the Belle Collaboration [65] and $\frac{\mathcal{B}(B^0 \to \pi^+\pi^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} = 0.29 \pm 0.02\ (stat.) \pm 0.02\ (syst.)$ measured at the *BABAR* experiment [63]. This agreement represents an important cross check of the validity of our analysis, that provides additional convincing evidence of the $B_s^0 \to K^+K^-$ observation. In $\pi^+\pi^-$-mass, the $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ distributions are completely overlapping (see figs. 4.1 and 7.2) and well separated from the $B^0 \to K^+\pi^-$ contribution. As a consequence,

[4] *BABAR* quotes absolute branching fractions. We calculated their ratio neglecting the correlations between the systematic uncertainties, which might be slightly over-estimated.

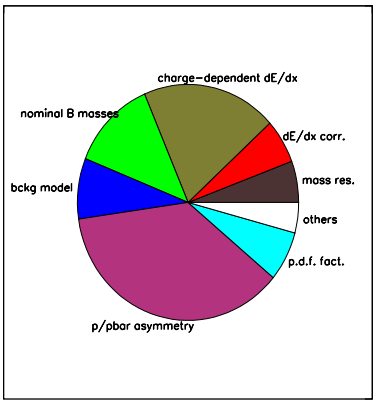

Figure 10.3: Circular chart of the systematic uncertainties (σ_i) contributing to the measurement of $A_{\mathsf{CP}}(B^0 \to K^+\pi^-)$. The area of each sector is proportional to σ_i^2.

the total number of joint $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ events is determined accurately even by fitting just the mass distribution. Any artificial enhancement of the $B_s^0 \to K^+K^-$ contribution due to a improper use of the dE/dx information would decrease the $B^0 \to \pi^+\pi^-$ contribution, thus resulting in a measured value of $\frac{\mathcal{B}(B^0\to\pi^+\pi^-)}{\mathcal{B}(B^0\to K^+\pi^-)}$ inconsistent with the B-Factories results.

Besides its role of cross-check, this results is promising for the future extension of the measurement to larger samples. With the sample already collected by the CDF experiment, we expect a reduction of more than a factor of three in the statistical uncertainty, further bound to decrease as CDF keeps gathering data. Also the systematic uncertainty, whose largest contribution derives from the effect of final state radiation (see fig. 10.4), will be strongly reduced. In the present measurement, this was included in the systematic uncertainties because the QED calculation of FSR in $B^0 \to K^+\pi^-$ decays were not available (see sec. 9.2.3). In the next update of the analysis the effect of FSR will be incorporated in the central result, because QED corrections have recently been published also for $B^0 \to K^+\pi^-$ modes [141].

10.4 Results on $B_s^0 \to K^-\pi^+$

From the observed yield of 3 ± 25 events, no evidence of the $B_s^0 \to K^-\pi^+$ decay was found. We thus set a frequentist upper limit on the corresponding branching fraction based on Gaussian distributions of fit pulls (see fig. 6.9(d)) and likelihood-ratio (LR) ordering, following Ref. [134]. Systematic uncertainties were added in quadrature to the statistical uncertainty for a proper inclusion of systematic effects in the extraction of the upper limit. We used the confidence intervals for the mean of a Gaussian distribution (i. e., the pull distribution) constrained to be non-negative (see tab. X of Ref. [134]). The resulting 90% confidence level (CL) upper limit on the ratio between the branching fractions of

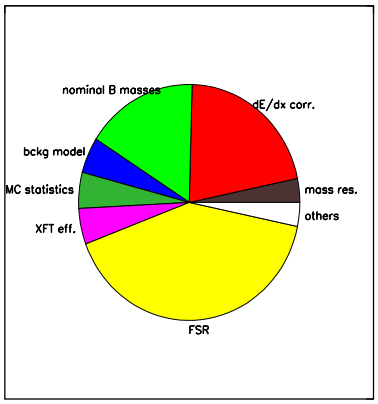

Figure 10.4: Circular chart of the systematic uncertainties (σ_i) contributing to the measurement of $\mathcal{B}(B^0 \to \pi^+\pi^-)/\mathcal{B}(B^0 \to K^+\pi^-)$. The area of each sector is proportional to σ_i^2.

$B_s^0 \to K^-\pi^+$ and $B^0 \to K^+\pi^-$ modes, multiplied by the ratio of production fractions, is

$$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B_s^0 \to K^-\pi^+)}{\mathcal{B}(B^0 \to K^+\pi^-)} < 0.08. \tag{10.8}$$

By normalizing the above result to the world-average value $\mathcal{B}(B^0 \to K^+\pi^-) = (18.9 \pm 0.7) \times 10^{-6}$ and by assuming for f_s/f_d the world-average value from $p\bar{p}$ and e^+e^- collisions, $f_s/f_d = 0.261 \pm 0.038$ [29], we obtain the following upper limit:

$$\mathcal{B}(B_s^0 \to K^-\pi^+) < 5.6 \times 10^{-6} \quad \text{at } 90\% \text{ CL}. \tag{10.9}$$

This result represents more than a factor of 30 improvement over the world-best limit available so far by the ALEPH Collaboration: $\mathcal{B}(B_s^0 \to K^-\pi^+) < 2.1 \times 10^{-4}$ at 90% CL [67].

Our result suggests a value for $\mathcal{B}(B_s^0 \to K^-\pi^+)$ at the lower end of current theoretical expectations. Beneke and Neubert predict $\mathcal{B}(B_s^0 \to K^-\pi^+) = (6.8\text{–}10.4) \times 10^{-6}$ basing their calculations on QCDF, in agreement with the results of Sun, Zhu, and Du, who improve the QCDF approach with chirally enhanced power corrections and contributions from electroweak annihilation [144]. Chiang and Zhou predict $\mathcal{B}(B_s^0 \to K^-\pi^+) = (5 \pm 1) \times 10^{-6}$, basing on SU(3) flavor symmetries [146]. Similar values are predicted also by calculations based on PQCD and on SCET. Yu, Li, and Lu expect $\mathcal{B}(B_s^0 \to K^-\pi^+) = (6.2\text{–}8.1) \times 10^{-6}$, assuming the $70° < \alpha < 130°$ range for the CKM phase α [150], while Williamson and Zupan obtain $\mathcal{B}(B_s^0 \to K^-\pi^+) = (4.9 \pm 1.8) \times 10^{-6}$ [143].

The comparison between the present upper limit and theoretical expectations suggest that valuable information on this decay will be available when this analysis will be extended to larger samples: either an observation of the $B_s^0 \to K^-\pi^+$ decay, or an upper limit on its branching fraction significantly lower than predictions will provide relevant input for tuning of phenomenological models on $B_{(s)}^0 \to h^+h'^-$ decays.

10.5 Results on $B_s^0 \to \pi^+\pi^-$

From the observed yield of -10 ± 15 events, no evidence of the pure-annihilation decay $B_s^0 \to \pi^+\pi^-$ was found. We thus set a frequentist upper limit on the corresponding branching fraction following the procedure detailed for the $B_s^0 \to K^-\pi^+$ limit (see sec. 10.4). The resulting 90% CL upper limit on the ratio between the branching fractions of $B_s^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ modes is

$$\frac{\mathcal{B}(B_s^0 \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to K^+K^-)} < 0.05. \tag{10.10}$$

By normalizing the above result to the $B_s^0 \to K^+K^-$ branching fraction shown in relation (10.1) and to the world-average value $\mathcal{B}(B^0 \to K^+\pi^-) = (18.9 \pm 0.7) \times 10^{-6}$, and by assuming for f_s/f_d the world-average value from $p\bar{p}$ and e^+e^- collisions, $f_s/f_d = 0.261 \pm 0.038$ [29], we obtain the following upper limit:

$$\mathcal{B}(B_s^0 \to \pi^+\pi^-) < 1.7 \times 10^{-6} \quad \text{at 90\% CL}. \tag{10.11}$$

This result represents an improvement of approximately two orders of magnitude over the world-best limit available so far by the ALEPH Collaboration: $\mathcal{B}(B_s^0 \to \pi^+\pi^-) < 1.7 \times 10^{-4}$ at 90% CL [67].

The sensitivity of our result approaches the expectations from recent calculations. Li, Lu, Xiao, and Yu calculate $\mathcal{B}(B_s^0 \to \pi^+\pi^-) = (4.2 \pm 0.6) \times 10^{-7}$ within the PQCD approach with Sudakov resummation, and including contributions from electroweak and QCD penguin amplitudes [43]. Beneke and Neubert, using the QCDF, predict smaller values, $\mathcal{B}(B_s^0 \to \pi^+\pi^-) = (0.24\text{–}1.55) \times 10^{-7}$ [24], in agreement with Yang et al. that also used QCDF, but with a different solution to avoid end-point divergences, that provides an expected rate of $\mathcal{B}(B_s^0 \to \pi^+\pi^-) = (1.24 \pm 0.28) \times 10^{-7}$ [44].

The CDF experiment is currently the only one capable to search for this decay mode. With the continuous increase of the collected data sample it has the unique opportunity to either obtain the first observation of a decay proceeding exclusively through annihilation topology or to improve the upper limits on its branching fraction below the theoretical expectations, thus providing constraining information.

10.6 Results on $B^0 \to K^+K^-$

From the observed yield of 10 ± 23 events, no evidence of the pure-annihilation $B^0 \to K^+K^-$ decay was found. We thus set a frequentist upper limit on the corresponding branching fraction following the procedure detailed for the $B_s^0 \to K^-\pi^+$ limit (see sec. 10.4). The resulting 90% CL upper limit on the ratio between the branching fractions of $B^0 \to K^+K^-$ and $B^0 \to K^+\pi^-$ modes is

$$\frac{\mathcal{B}(B^0 \to K^+K^-)}{\mathcal{B}(B^0 \to K^+\pi^-)} < 0.10. \tag{10.12}$$

By normalizing the above result to the world-average value $\mathcal{B}(B^0 \to K^+\pi^-) = (18.9 \pm 0.7) \times 10^{-6}$ [29], we obtain the following upper limit:

$$\mathcal{B}(B^0 \to K^+K^-) < 1.8 \times 10^{-6} \quad \text{at 90\% CL}. \tag{10.13}$$

This result is not yet competitive with the results available at the B-Factories, $\mathcal{B}(B^0 \to K^+K^-) < 0.25 \times 10^{-6}$ at 90% CL from the Belle Collaboration [66] and $\mathcal{B}(B^0 \to K^+K^-) < 0.40 \times 10^{-6}$ at 90% CL from the *BABAR* Collaboration [63]. However, competitive measurements are expected with the sample currently available at CDF. The expected sensitivities are still far from probing the full range of theoretical expectations, which, however, are affected by large uncertainties: Beneke and Neubert, for instance, predict $\mathcal{B}(B^0 \to K^+K^-) = (0.007\text{–}0.079) \times 10^{-6}$ [24].

10.7 Concluding remarks and future prospects

When the opportunity to trigger on $B^0_{(s)} \to h^+h^{'-}$ decays and reconstruct them efficiently in the CDF II detector was first proposed in 1993 [102], it was received with a lot of skepticism, motivated by the envisaged difficulties of reconstructing such rare signals in the challenging environment of $p\bar{p}$ collisions. The successful design and operation of the trigger on displaced tracks (SVT), along with the analysis shown in this thesis, confirms today the validity of that proposal.

In this thesis I describe how a signal of $B^0_{(s)} \to h^+h^{'-}$ decays was reconstructed for the first time at a hadron collider and was used to measure the relative branching fractions of individual decay modes and the CP-violating partial-rate asymmetry of $B^0 \to K^+\pi^-$ decays (see tab. 10.1).

Mode	Yield	Measured quantity	Derived $\mathcal{B}$ $[\times 10^{-6}]$
$B^0 \to K^+\pi^-$	542 ± 30	$A_{CP} = -0.013 \pm 0.078 \pm 0.012$	
$B^0 \to \pi^+\pi^-$	121 ± 27	$\frac{\mathcal{B}(B^0\to\pi^+\pi^-)}{\mathcal{B}(B^0\to K^+\pi^-)} = 0.21 \pm 0.05 \pm 0.03$	$3.9 \pm 1.0 \pm 0.6$
$B^0_s \to K^+K^-$	236 ± 32	$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s\to K^+K^-)}{\mathcal{B}(B^0\to K^+\pi^-)} = 0.46 \pm 0.08 \pm 0.07$	$33 \pm 6 \pm 7$
$B^0_s \to K^-\pi^+$	3 ± 25	$\frac{f_s}{f_d} \times \frac{\mathcal{B}(B^0_s\to K^-\pi^+)}{\mathcal{B}(B^0\to K^+\pi^-)} < 0.08$ @ 90% CL	< 5.6 @ 90% CL
$B^0_s \to \pi^+\pi^-$	-10 ± 15	$\frac{\mathcal{B}(B^0_s\to\pi^+\pi^-)}{\mathcal{B}(B^0_s\to K^+K^-)} < 0.05$ @ 90% CL	< 1.7 @ 90% CL
$B^0 \to K^+K^-$	10 ± 23	$\frac{\mathcal{B}(B^0\to K^+K^-)}{\mathcal{B}(B^0\to K^+\pi^-)} < 0.10$ @ 90% CL	< 1.8 @ 90% CL

Table 10.1: Final results. Absolute branching fractions are normalized to the world-average values $\mathcal{B}(B^0 \to K^+\pi^-) = (18.9 \pm 0.7) \times 10^{-6}$ and $f_s/f_d = 0.261 \pm 0.038$ [29]. The first quoted uncertainty is statistical, the second one systematic.

I report the first observation of the decay mode $B^0_s \to K^+K^-$, which is also the first observed decay of a B^0_s meson into two pseudo-scalar mesons. I report greatly improved upper limits on the branching fractions of $B^0_s \to K^-\pi^+$ and $B^0_s \to \pi^+\pi^-$ modes, and results in agreement with current world-averages for B^0 decay modes. These measurements have implications for both the knowledge of CKM sector of the Standard Model and the phenomenology of non-perturbative hadronic contributions in non-leptonic $B^0_{(s)}$ meson decays. The present results (see tab. 10.1) were published in Physical Review Letters [1].

10.7.1 $B^0_s \to K^+K^-$ observation

The observation of the $B^0_s \to K^+K^-$ decay and the measurement of its branching fraction have multiple relevance. It is the first measured amplitude of a charmless decay of a B^0_s into two pseudo-scalar

mesons. The comparison of the observed branching fraction with theoretical predictions provides valuable information for tuning the phenomenological models of hadronic $B^0_{(s)}$ meson decays (SU(3) symmetries, QCDF, SCET, QCD sum-rules, etc.) and for optimizing the choice of their input parameters. As an example of this impact, the measurement of $\mathcal{B}(B^0_s \to K^+K^-)$ provides the first experimental insight on the magnitude of the SU(3) breaking, a necessary test and ingredient for the U-spin-based method of extracting the CKM parameters.

In the flavor sector, the current resolution on $\mathcal{B}(B^0_s \to K^+K^-)$ is insufficient to derive strong constraints on the CKM parameters, as is frequently the case when dealing with measurements involving a single $B^0_{(s)} \to h^+h^{'-}$ decay-mode. However, combined with the full set of measurements of this thesis, for a more efficient use of information, it provides valuable experimental input. In particular, the measurement of $\mathcal{B}(B^0_s \to K^+K^-)$ provides the global CKM fits with the first (and currently unique) $B^0_{(s)} \to h^+h^{'-}$ observables from decays of B^0_s mesons.

A recent example of the possibilities of this approach was given in Ref. [151] by Malclès, who proposed a new data-driven technique in which relations among amplitudes of two-body charmless decays, including charged b-mesons and neutral final states, are derived in a model-independent way under the assumption of unitarity of the CKM matrix and at the limit of exact SU(3) symmetry. Electroweak penguin and annihilation-exchange topologies are *included* in the correlations between

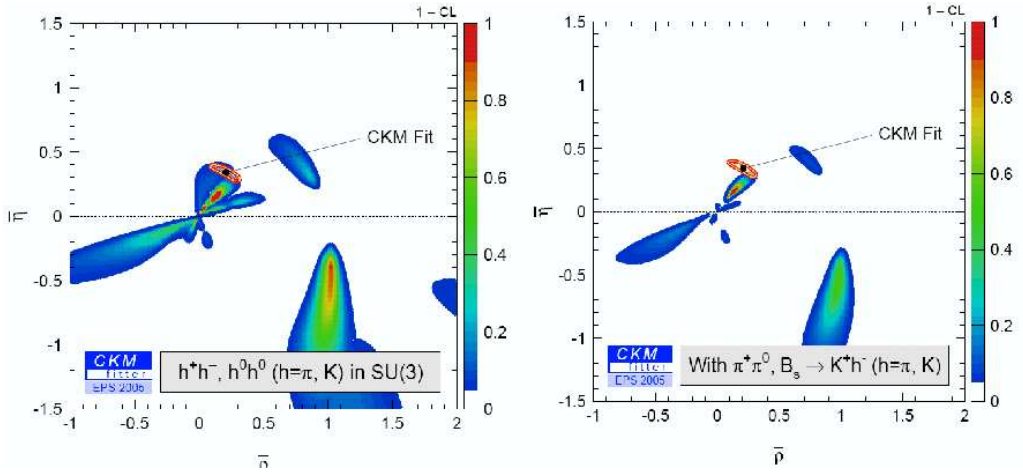

Figure 10.5: Allowed regions in the $(\bar{\rho}, \bar{\eta})$ plane of the CKM parameters as resulting from the global fits of Ref. [151]. Plot (a) uses experimental inputs from the B-Factories experiments. In plot (b) the measurements of this thesis are included. The color code indicates the probability (1-CL) of the χ^2-fit. For comparison, the results of the CKMfitter Group [152] are overlaid. The recently measured value of the frequency of B^0_s flavor oscillations is not included in these plots.

amplitudes and simplified SU(3)-breaking corrections are incorporated through π and K decay constants and conservative theoretical uncertainties.[5] Malclès shows that the combination of precise measurements from the B-Factories with the CDF results on branching fractions of B^0_s mesons in a

[5]The dominant fraction of the factorizable corrections is included, the non-factorizable corrections are assumed negligible.

global χ^2-fit provides non trivial constraints in the $(\bar{\rho}, \bar{\eta})$ plane of the CKM parameters. Figure 10.5 shows the impact of the results of this thesis within this approach: the area of allowed regions in the $(\bar{\rho}, \bar{\eta})$ plane is visibly reduced by the B_s^0 meson branching fractions measured at CDF.

The capability of collecting a copious $B_s^0 \to K^+K^-$ sample, demonstrated in this thesis, provides the opportunity to detect CP violation in the B_s^0 system for the first time. At given integrated luminosity, the $B_s^0 \to K^+K^-$ sample reconstructed in the CDF II detector is the second largest sample of fully reconstructed B_s^0 meson decays. This constitutes the first key ingredient for the measurement of CP-violating asymmetries in $B_s^0 \to K^+K^-$ decays, following the Fleischer proposal for a theoretically reliable measurement of the CKM phase γ [5]. This challenging measurement is the ultimate long-term goal of the analysis of $B_{(s)}^0 \to h^+h^{'-}$ decays at CDF, and will need the full Run II statistics. The recent measurement of B_s^0 flavor oscillations shows that all needed ingredients for this measurement are available: an excellent proper-time resolution of the CDF II tracker, which is necessary to resolve the sine and cosine terms of the $B_s^0 \to K^+K^-$ decay-rate asymmetry as a function of time (see eq. (1.59)); a good performance, $\epsilon D^2 \simeq 5\%$, in identifying the flavor of the b-mesonat production, despite the difficult environment of hadron collisions;[6] and a precise measurement of the mass-difference Δm_s, a necessary input parameter of the Fleischer strategy. With the full Run II statistics ($\approx$ 8 fb^{-1}) we expect a measurement of CP-violating asymmetries in $B_s^0 \to K^+K^-$ decays with $\mathcal{O}(10\%)$ resolution. This measurement is possible only at CDF and at the future LHCb experiment. The Lorentz boost of B_s^0 mesons that will be produced in the planned Υ(5S) run at the KEKB collider is insufficient for resolving the fast B_s^0 oscillations in the Belle vertex detector.[7]

The intermediate experimental goal between the currently available high-precision measurements of time-integrated rates and future time-dependent measurement of asymmetries, is the study of the time evolution of $B_s^0 \to K^+K^-$ decays. This provides crucial information on the lifetime difference in the B_s^0 meson system, $\Delta\Gamma_s/\Gamma_s$, as already demonstrated by a preliminary result obtained using the technique developed in this thesis to isolate the signal [155].

10.7.2 Other results

In the context of increasing the experimental information for tuning the phenomenological models of hadronic $B_{(s)}^0$ meson decays (SU(3) symmetries, QCDF, SCET, QCD sum-rules, etc.) the great improvement on upper limits on branching fractions of the $B_s^0 \to K^-\pi^+$ and $B_s^0 \to \pi^+\pi^-$ decays plays a key role. The result on the $B_s^0 \to \pi^+\pi^-$ decay-rate provides a tighter constraint on the still unknown magnitude of penguin-annihilation amplitudes, which are difficult to predict and introduce uncertainties in several processes (i. e., $B^0 \to \pi^+\pi^-$) used for extracting CKM phases.

[6] The performance of flavor-tagging methods is usually quoted as the product of a tagging efficiency (ϵ) by a squared "dilution" (D). The efficiency is the fraction of events in which the tagging-algorithm converges to a decision, the dilution is the asymmetry between correct and wrong tagging-decisions.

[7] The Belle experiment has access to B_s^0 decays when the KEKB collider operates at the energy of the Υ(5S) resonance. Following the CLEO success of 2003, when exclusive B_s^0 meson decays were observed in a $\int \mathcal{L}dt = 0.42$ fb^{-1} dataset collected at the Υ(5S) resonance [153], in 2005 the Belle experiment collected data during a three-days "engineering" run at the Υ(5S) energy and observed two reconstructed $B_s^0 \to K^+K^-$ candidates [154]. While a larger dataset corresponding to $\int \mathcal{L}dt = 21.7$ fb^{-1} is currently being analyzed, the Belle Collaboration expects the reconstruction of 30–40 $B_s^0 \to K^+K^-$ decays per $\int \mathcal{L}dt = 100$ fb^{-1} (corresponding approximately to three months of operation) in a future physics run at the Υ(5S) energy.

The upper limit on $\mathcal{B}(B_s^0 \to K^-\pi^+)$ approaches the lower end of theoretical expectations suggesting a possible observation in the near future. In this case, samples of $B_s^0 \to K^-\pi^+$ decays will be used to extract valuable information on the ratio between weak tree and penguin amplitudes to determine the CKM phase γ, following the proposal by Gronau and Rosner [39].

The present results on B^0 decays are not competitive with the B-Factories results. However, the small systematic uncertainty of the direct CP-violating asymmetry in the $B^0 \to K^+\pi^-$ decays, and the large $B^0 \to K^+\pi^-$ yield indicate that a competitive measurement will be available already with the sample available as of this writing. In addition, only CDF can combine measurements of asymmetries in the $B^0 \to K^+\pi^-$ and $B_s^0 \to K^-\pi^+$ modes to explore a theoretically robust, model-independent test for new physics proposed by Lipkin in a recent paper [40].

10.7.3 Outlook

Not only the present results are relevant in their own right, they are also extremely promising for future extensions of this analysis to larger samples. They are obtained from a sample, corresponding to $\int \mathcal{L}dt = 180$ pb^{-1}, which is a small fraction of the currently available data. In addition, being the first physics-quality data, they were collected when trigger and detector performances were still sub-optimal. Currently the Tevatron is running smoothly and its performance is steadily improving. Similarly, the operation of the CDF II detector and trigger has been optimized and improved through various upgrades. Typically, the Tevatron delivers collisions corresponding to 30 pb^{-1} per week, resulting in approximately 200 $B_{(s)}^0 \to h^+h^{'-}$ decays reconstructed in the CDF II detector, after the analysis. As of October 2006, the sample collected by the CDF experiment corresponds to $\int \mathcal{L}dt \approx$ 1.6 fb^{-1}.

The preliminary results obtained from a subsample corresponding to $\int \mathcal{L}dt \approx 1$ fb^{-1} show a signal yield of ≈ 6500 $B_{(s)}^0 \to h^+h^{'-}$ events with higher purity than the one obtained in this thesis [156]. Already with this sample, we expect results of same quality as current B-Factories results in measurements related to B^0 decays, and dramatically improved results on the B_s^0 sector, including the first observation of the $B_s^0 \to K^-\pi^+$ decay.

The Tevatron is planned to operate through September 2009 (at least), delivering a total integrated luminosity of 4.5–8.5 fb^{-1}. During this time, the CDF experiment will continue exploiting the physics opportunity of accessing jointly B^0 and B_s^0 decays into two-body charmless mesons. This prerogative is unique to CDF at least until the planned run of the KEKB collider at the center-of-mass energy of the Υ(5S) resonance or the LHCb data-taking (2008) will take place. This promotes the CDF experiment in a leading position for either improving our knowledge of the mechanism of CP-violation in the Standard Model or detecting the first indications of unexpected CP-violating phases.

Appendix A

Specific ionization measurement in the COT

A.1 Energy loss of relativistic charged particles in a medium

When a charged particle traverses a gaseous or condensed medium, an exchange of virtual photons with a broad spectral range, from the soft ultraviolet to the x-ray region, occurs. This is due to the (largely dominant) Coulomb interaction between the electromagnetic fields of the incoming particle and of the material. The photons interact with the atoms of the medium producing *excitation* and *ionization*. An indirect measurement of the total energy-loss is given by the number of ionization electrons emitted in the gas. The broad spectrum of photons and the thin samples usually employed lead to large fluctuations in the signal from a proportional counter. The typical ionization energy-loss distribution, at fixed velocity of the incident particle, has a peak with a long tail at higher values of energy loss ("Landau distribution"). The peak results from soft collisions in which the atom as a whole absorbs a virtual photon producing ionization. The tail results from hard collisions between the particle and quasi-free atomic electrons. This is due to the logarithmic divergence of the mean transferred energy per collision caused by the $1/E^2$ dependence of the cross-section for Rutherford scattering on one electron, and it is usually a nuisance in dE/dx measurements.

The electromagnetic field of the incoming particle propagates in the direction perpendicular to its motion as the velocity approaches the phase velocity of light in the medium, and the energy-loss cross-section increases as $\ln(\beta\gamma)$. Thus, the position of the peak of the energy-loss spectrum moves accordingly. The dependence of the ionization on velocity allows estimating the particle mass in the relativistic regime. In a medium of finite density, the polarization associated to the dielectric properties of the medium shields the transverse field of the propagating particle, causing a saturation of the energy loss (plateau). The average total energy-loss per unit length of a particle (heavier than the electron) of charge q traversing a gas volume with velocity $c\beta$ is approximated by the Bethe-Bloch formula [157]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi N e^4}{m_e c^2 \beta^2} q^2 \left[\ln\left(\frac{2 m_e c^2 \beta^2 \gamma^2}{I} \right) - \beta^2 - \frac{\delta(\beta)}{2} \right], \tag{A.1}$$

where N is the electron density in the medium, m_e (e) is the electron mass (charge), I is the mean excitation energy of the medium atoms, and $\delta(\beta)$ is the correction that accounts for the density effect at high velocities.[1]

The main features of the logarithmic rise of the energy-loss cross-section and of the density effect depend only on the virtual wave structure of the electromagnetic field of the incident particle and may be understood with the aid of a simple model [158]. Consider a scalar field ϕ coupled to a particle moving with constant velocity v along the x axis through a dispersive medium in which the phase velocity of the field is $u(\omega)$, a function of frequency. In the rest frame of the particle the field is static. In the frame of the medium, it is a wave packet that is neither dispersed nor attenuated and that moves along x with constant amplitude and phase with respect to the particle. Thus, the phase velocity along x of each frequency component must equal v, since it is a component of a static field as seen by an observer moving with the particle, $v = \omega/k_x$ (k is the wave number of the Fourier component of frequency ω). By definition of phase velocity, $u = \omega/|k|$. Therefore the space-time dependence of each frequency component is described in the two-dimensional case by the phase factor $e^{i(k_x x + k_y y - \omega t)} = e^{\frac{i\omega}{v}(x - vt + y\sqrt{v^2/u^2 - 1})}$. We denote as $\beta'(\omega)$ the dimensionless velocity $\frac{v}{u(\omega)}$. If $\beta' < 1$, the transverse dependence of the field is an evanescent wave of range $y_0(\omega) = \frac{v}{\omega\sqrt{1-\beta'^2}} = \beta'\gamma'\xi$, where $\gamma' = \frac{1}{\sqrt{1-\beta'^2}}$ and ξ is the free wavelength over 2π. For larger velocities, β' gets closer to unity and the range increases linearly with $\beta'\gamma'$. For the electromagnetic field of a charged particle moving in vacuum, $u = c$ for all frequencies, $\beta' = v/c = \beta$, and $\gamma' = \gamma$.

The propagation of the electromagnetic field with $\beta\gamma$ responsible for the relativistic rise is therefore seen as a general consequence of wave motion. In a medium, on the other hand, $u = c/\sqrt{\varepsilon}$ where $\varepsilon(\omega)$ is the dielectric constant. When $\varepsilon(\omega) > 1$ the range is $y_0(\omega) = \frac{v}{\omega\sqrt{1-\beta^2\varepsilon}}$ which saturates at $\frac{\xi}{\sqrt{1-\varepsilon}}$. Approximately 90% of this range is achieved when $\beta\gamma = \frac{2}{\sqrt{1-\varepsilon}}$, where the field saturates, becoming insensitive to any further increase of velocity. A simplified model of the dielectric constant yields $\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}$, where the plasma frequency is a function of the electron density N: $\omega_p^2 = \frac{4\pi N e^2}{m}$. Even in this approximation, the energy-loss cross-section saturates at $\beta\gamma = \frac{2\omega}{\omega_p} \propto \frac{1}{\sqrt{N}}$, a result close to the one observed experimentally and known as "density effect".

A.2 The measurement of energy loss in the COT

The COT was originally designed for accurate and efficient tracking, not being explicitly optimized for dE/dx measurements. The small drift cells, designed for fast response and high spatial-resolution measurements, reduce the number of charge clusters collected by each wire, thus increasing statistical fluctuations. In addition, the gas is kept at atmospheric pressure, while higher pressures are better for dE/dx measurements. However, a dedicated charge-integration mode can be enabled in the COT read-out chip for a measurement of dE/dx.

The COT samples the amount of ionization charge produced by a track by measuring the time-

[1] In the specific case of ionization of tracks in drift chambers the "restricted" energy-loss formula is more appropriate. It accounts for the fact that energy transfers high enough to knock out an electron from a gas atom result in a second track, whose energy will not contribute to the original track [158]. It differs from the Bethe-Bloch curve for an energy cut-off term in the logarithm.

over-threshold of the pulse on each wire associated to the track. Figure A.1 shows a pulse after each stage of the read-out chip. The average chamber pulse for a radial charged particle ionizing at minimum rate injects a charge of approximately 10 fC into the chip during its 8 ns peaking time. Series of resistors combined with diodes protect the chip from large negative and positive spikes. A pseudo-differential input in each channel is provided by two preamplifiers which have 1.5 mV/fC linear gain and 1.5 ns rise time. The shaper is fully differential with multi-pole shaping to cancel positive ion and preamplifier tails. Its gain is 25 mV/fC, and the associated undershoot of the pulse does not exceed 2%. The differential baseline-restoration circuit is AC-coupled, resulting in uniform channel-to-channel discriminator thresholds. It allows for high-rate performance with earlier retriggering in case of large charge depositions. The discriminator thresholds are set at ≈ 2 fC, close to the inherent noise level, to ensure high hit-efficiency and good timing-accuracy independently of Q fluctuations from particle to particle.

The measurement of charge Q can be optionally enabled (dE/dx-on), resulting in an output width of the discriminated signal proportional to $\ln(Q)$.[2] When the trailing-edge of the signal exceeds the threshold, a 28 ns shaping allows accurate integration of the charge. When the signal falls below threshold, integration is stopped. Off-line, an 80% truncated mean of the dE/dx samplings along

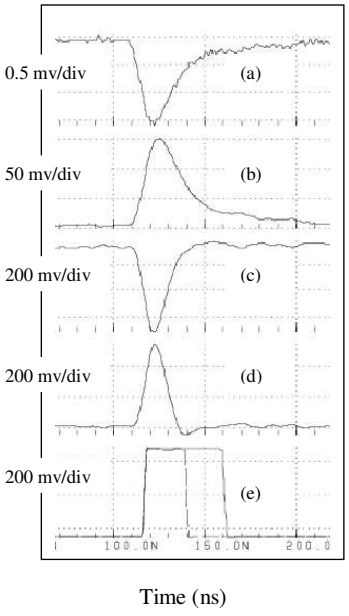

Figure A.1: An ≈ 80 fC pulse at the read-out chip: (a) preamplifier input, (b) preamplifier output, (c) shaper output, (d) baseline-restoration output, and (e) discriminator output with dE/dx on (longer pulse) or off (shorter pulse).

[2]In high-luminosity environment, with associated high occupancy, the dE/dx can be disabled to allow for a faster shaping and higher tracking efficiency.

a track is used to estimate the average ionization, in order to limit the adverse effect of the dE/dx positive tail due to the logarithmic divergence. The equation that better models the COT average energy-loss as a function of velocity is the following variant of the Bethe-Bloch curve:

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

with a_i, b, c_j, and C parameters extracted from data. The above function has all the features that are present in the Bethe-Bloch curve (eq. (A.1)). Parameters c_0 and c_1 represent the intensities of the $1/\beta^2$ fall and of the relativistic rise, respectively. Parameter b is associated with the COT gas properties, e. g., mean excitation energy of the gas atoms, etc. Parameters a_1 and a_2 provide further adjustment, especially in the low $\beta\gamma$ region.

The hit-widths, i. e., the individual charge collections output by the COT are subject to the following corrections, applied in the off-line production (sec. 2.6), to eliminate a number of detector related conditions [159]:

Hit merging – The distribution of the arrival times of clusters produced by a single track may have a large spread, especially if the incident particle passes close to a sense wire, thus crossing a high number of isochrones. This may cause the signal from one cluster to fall below discriminator threshold before the subsequent cluster (from the same track) reaches the wire. On each wire, adjacent hit-widths due to split pulses induced by the same track are merged to avoid this effect.

Electronic pedestal subtraction – it accounts for the offset between digital pulse-width and the corresponding collected charge.

Path-length correction – it removes the charge dependence on the local distance traversed by the particle near a wire. The polar dependence is due to the geometry of cells. The correction is proportional to $1/\sin(\theta)$.

High-voltage correction – it removes gain fluctuations due to different voltages of power-supplies among super-layers.

z **correction** – it removes the small dependence on the position of the hit along the sense-wire, caused by the attenuation of the signal while propagating along the sense wire to the read-out end of the wire.

Angle and drift distance corrections – the pulse shape depends on the drift distance and on the angle between the momentum of the incident particle and the drift field in a cell. The angle-distance dependence is corrected jointly.

Wire correction – wires at the end of the cell collect more charge than wires in the middle, because the physical boundary between super-layers constrains the diffusion of the charge. The collected charge among the 12 sense-wires in each cell is equalized.

Super-layer correction – it removes residual super-layer dependencies due to different cell-size between super-layers.

Pressure correction – it removes factors up to 5–6 of gain variation due to fluctuations of the gas pressure. Increases in pressure increase the amount of primary ionization, but reduce the mean free-path for electrons in the avalanche, producing an overall reduction in gain. An hardware pressure-feedback circuit and a specific off-line correction remove the effect.

The track-based calibration procedure described in chap. 5 is applied after all the above corrections have been applied.

Appendix B

Minimum Variance Bound and numerical minimization

B.1 Fisher information and Minimum Variance Bound

An useful figure of merit in data reduction is the "Fisher information". The amount of (Fisher) information given by an observation x about the parameter θ is defined by the following analytical quantity (if it exist):

$$I_x(\theta) = E\left[\left[\frac{\partial \ln(\mathscr{L})}{\partial\theta}\right]^2\right], \tag{B.1}$$

where x is the observable, $E[x]$ indicates the expectation value of x, and $\mathscr{L} \equiv \mathscr{L}(\theta)$ is the likelihood function of θ. $I_x(\theta)$ has some interesting properties: it increases as the number of observations increases, it is conditional to the estimation being done (data that are irrelevant to the parameters to be measured contain no information), and it is related to the statistical precision of the estimation. Thus, $I_x(\theta)$ provides an useful criterion for the efficient reduction of data: obtain the maximum data reduction consistent with minimum loss of information.

More generally, for a vector $\vec{\theta} = (\theta_1, \ldots, \theta_m)$ of parameters, the ij element of the information matrix is defined as

$$[I_x(\theta)]_{ij} = E\left[\frac{\partial \ln(\mathscr{L})}{\partial\theta_i}\frac{\partial \ln(\mathscr{L})}{\partial\theta_j}\right]. \tag{B.2}$$

Under general conditions of regularity of the likelihood (space of the observables independent of $\vec{\theta}$, $\mathscr{L}$ twice-differentiable, and commutation of $\int dx$ and $\partial/\partial\theta$), the following relation holds:

$$[I_x(\theta)]_{ij} = -E\left[\frac{\partial^2 \ln(\mathscr{L})}{\partial\theta_i\partial\theta_j}\right]. \tag{B.3}$$

The Fisher information is related to a fundamental bound of the maximum precision attainable in measuring a parameter $\vec{\theta}$ [111]. This is given by the Cramér-Rao inequality,

$$\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j) \geq \frac{1}{[I_x(\theta)]_{ij}} = \frac{1}{-E\left[\frac{\partial^2 \ln \mathscr{L}(\vec{\theta})}{\partial\theta_i\partial\theta_j}\right]}, \tag{B.4}$$

where $\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j)$ ($\equiv \sigma_i^2$ if $i = j$) is the element ij of the covariance matrix. The covariance matrix is a generalization to multiple dimensions of the variance of a scalar-valued random variable. It contains the covariances between elements of a vector of random variables. The covariance between the ith and jth elements of the vector is $\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j) = E[\theta_i\theta_j] - E[\theta_i]E[\theta_j]$. Relation (B.4) is easily generalizable to the case of biased estimators and applies to any estimator, not only to those derived from the ML principle. In the case of equality, the Minimum Variance Bound (MVB) is reached. Relation (B.4) is useful because, under fairly mild conditions of regularity of the problem, the maximum likelihood estimator meets the MVB. Thus, the MVB provides a convenient analytical calculation of the statistical power of an estimator, before carrying-out the measurement. Since relation (B.4) involves an expectation value, i. e., an integration over the observable space (x), its evaluation may be impractical. However, in case of sufficiently large data samples, the MVB expression is well approximated by

$$\mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j) \approx \frac{1}{-\frac{\partial^2 \ln \mathscr{L}(\vec{\theta})}{\partial\theta_i \partial\theta_j}\Big|_{\theta=\hat{\theta}}}, \tag{B.5}$$

where the second derivative is evaluated using the measured data at the ML estimates $\hat{\theta}$.

B.2 Numerical minimization with the MINUIT package

In this thesis, the maximum likelihood estimates ($\hat{\vec{\theta}}$) of the desired parameters ($\vec{\theta}$) were found by a numerical minimization of the function $-2\ln(\mathscr{L})$, where $\mathscr{L} \equiv \mathscr{L}(\vec{\theta})$ is the likelihood function expressed in terms of the probability density functions.

For the minimization, we used the MINUIT software package [126] which, varying the parameters $\vec{\theta}$, evaluates the function as a 64-bit floating point number. In addition, MINUIT analyzes the shape of the function in the neighborhood of the minimum to estimate the parameter uncertainties. Several options of varying reliability and computational speed are available for this task. For minimization, we used the MIGRAD option, which applies a stable variation of the Davidon-Fletcher-Powell algorithm [160] in a gradient search for the minimum value. This algorithm converges to the correct error matrix as is converges to the minimum of the function. At each iteration, it needs a "working approximation" of the covariance matrix and of the gradient vector at the current best point, to determine the current search direction. If the matrix is positive-definite, as expected for most physical functions in the vicinity of their minima, the search proceeds toward the minimum; if not, e. g., in case of numerical inaccuracies, an appropriate constant is added along the diagonal, as determined from the matrix eigenvalues.

For the evaluation of uncertainties, we compared the results of three different algorithms MIGRAD, HESSE, and MINOS:

1. MIGRAD – the uncertainties are calculated using relation (B.5): the matrix of second derivatives of $-2\ln(\mathscr{L})$ is numerically determined using finite differences, it is evaluated at the ML estimates, and inverted to find the covariance matrix. That is, MIGRAD uses the curvature of $-2\ln(\mathscr{L})$ at the minimum and assume the parabolic shape. The resulting uncertainties, therefore, account for parameter correlations, but not for non-linearities.

2. HESSE – in some cases, MIGRAD minimization may converge before a reliable estimate of the error matrix is available.[1] The HESSE option provides calculation of the covariance matrix after MIGRAD is converged. Again, uncertainties are approximate in case of non-linearities.

3. MINOS – this option allows inclusion of non-linearities (along with correlations) in the uncertainties. After MIGRAD minimization, each parameter θ_i is varied, each time minimizing $-2\ln(\mathscr{L})$ with respect to all other parameters. MINOS searches for the two values of θ_i for which the minimum of $-2\ln(\mathscr{L})$ takes on the values $-2\ln[\mathscr{L}(\hat{\vec{\theta}})]+\Delta$, where Δ is an user-specified positive quantity, typically corresponding to one standard deviation.

In addition to the evaluation of uncertainties, MINUIT allows graphical representation of two-dimensional contours of the $-2\ln(\mathscr{L})$ function in the proximity of its minimum. This is accomplished by the MNCONTOUR option, which uses the same algorithm of MINOS, varying two user-specified parameters at a time, and accounting for the correlations and non-linearities.

In the present analysis, the values of the event observables to be input to MINUIT were stored in an ASCII file and the function $-2\ln(\mathscr{L})$, written in C programming language, was externally called by MINUIT. The initial set of parameters $\vec{\theta}$ was also provided in another ASCII file. We verified the robustness of the fit estimates by checking their consistence against changing the set of initial parameters. For some values of the parameters, the function $\mathscr{L}$ may take on values equal or less than zero. In these cases "positive infinity" was added to $-2\ln(\mathscr{L})$ to avoid "pathologic" regions.

Occasionally, the domain of some parameters needed to be bounded to some specified values (e. g., $a \leq \theta_i \leq b$) for an easier convergence of the minimization. The resulting uncertainty may be incorrectly estimated in these cases, since MINUIT defines internally a non-linear transformation of the parameter space, that introduces further numerical inaccuracy. We dealt with this feature by limiting the domains of parameters just to find the minimizing set of parameters. Then, the set found was used as starting point for a new fit with free parameters and consequent proper evaluation of the uncertainties.

[1] Since for n parameters there are $n(n+1)/2$ elements of the error matrix, a number of MIGRAD iterations larger than n^2 is needed for a reliable estimate of the error matrix itself.

Bibliography

[1] A. Abulencia et al. (CDF Collaboration), *Observation of $B_s^0 \to K^+K^-$ and Measurements of Branching Fractions of Charmless Two-Body Decays of B^0 and B_s^0 Mesons in* $\mathrm{p\bar{p}}$ *Collisions at $\sqrt{s}$=1.96* TeV, Phys. Rev. Lett. **97**, 211802 (2006), [hep-ex/0607021].

[2] B. Aubert et al. (*BABAR* Collaboration), *Direct* CP *Violating Asymmetry in $B^0 \to K^+\pi^-$ Decays*, Phys. Rev. Lett. **93**, 131801 (2004), [hep-ex/0407057];
Y. Chao et al. (Belle Collaboration), *Evidence for direct* CP *Violation in $B^0 \to K^+\pi^-$ Decays*, Phys. Rev. Lett. **93**, 191802 (2004), [hep-ex/0408100].

[3] M. Gronau and D. London, *Isospin analysis of* CP *asymmetries in B decays*, Phys. Rev. Lett. **65**, 3381 (1990).

[4] J. P. Silva and L. Wolfenstein, *Determining the penguin effect on* CP *violation in $B^0 \to \pi^+\pi^-$*, Phys. Rev. D **49**, 1151 (1994), [hep-ph/9309283].

[5] R. Fleischer, *New Strategies to extract β and γ from $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$*, Phys. Lett. **B459**, 306 (1999), [hep-ph/9903456].

[6] J. H. Christenson et al., *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13**, 138 (1964);
A. Abashian et al., *Search for CP nonconservation in K_2^0 decays*, Phys. Rev. Lett. **13**, 243 (1964).

[7] A. Alavi-Harati et al. (KTeV Collaboration), *Observation of Direct* CP *Violation in $K_{S,L} \to \pi\pi$ Decays*, Phys. Rev. Lett. **83**, 22 (1999), [hep-ex/9905060];
V. Fanti et al. (NA48 Collaboration), *A new measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465**, 335 (1999), [hep-ex/9909022].

[8] L. Wolfenstein, *Violation of CP Invariance and the Possibility of Very Weak Interactions*, Phys. Rev. Lett. **13**, 562 (1964).

[9] A. B. Carter and A. I. Sanda, *CP Violation in Cascade Decays of B Mesons*, Phys. Rev. Lett. **45**, 952 (1980);
——— *CP Violation in B Meson Decays*, Phys. Rev. D **23**, 1957 (1981);
I. I. Y. Bigi and A. I. Sanda, *Notes on the Observability of CP Violations in B Decays*, Nucl. Phys. **B 193**, 85 (1981).

[10] B. Aubert et al. (*BABAR* Collaboration), *Observation of* CP *Violation in the B^0 Meson System*, Phys. Rev. Lett. **87**, 091801 (2001), [hep-ex/0107013];
K. Abe et al. (Belle Collaboration), *Observation of large* CP *Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87**, 091802 (2001), [hep-ex/0107061].

[11] A. D. Sakharov, *Violation Of CP Invariance, C Asymmetry, And Baryon Asymmetry Of The Universe*, Pisma Zh. Exp. Theor. Fiz. **5**, 32 (1967), English translation in JETP Lett. **5**, 24 (1967), reprinted in Sov. Phys. Usp. **34**, 392 (1991).

[12] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[13] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963);
M. Gell-Mann, *A Schematic Model for Baryons and Mesons*, Phys. Lett. **8**, 214 (1964);
G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and Its Breaking. 2.*, CERN-TH-412.

[14] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2**, 1285 (1970).

[15] G. Lüders, *Proof of the TCP theorem*, Ann. Phys. (N.Y.), **1** (1957);
W. Pauli, *Exclusion Principle, Lorentz Group, and reversal of space-time and charge*, in *Niels Bohr and the Development of Physics*, W. Pauli (ed.) New York: Pergamon (1955).

[16] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal* CP *Nonconservation*, Phys. Rev. Lett. **55**, 1039 (1985);
———, *A basis independent formulation of the connection between quark mass matrices,* CP *violation and experiment*, Z. Phys. C **29**, 491 (1985);
I. Dunietz, O. W. Greenberg, and D.-D. Wu, A priori *definition of maximal CP nonconservation*, Phys. Rev. Lett. **55**, 29351 (1985).

[17] L. Wolfenstein, *Parameterization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[18] W.-M. Yao et al. (Particle Data Group), *Review of particle physics*, J. Phys. **G** 33, 1 (2006).

[19] K. G. Wilson, *Non-Lagrangian Models of Current Algebra*, Phys. Rev. **179**, 1499 (1969).

[20] M. S. Sozzi and I. Mannelli, *Measurements of direct* CP *violation*, Riv. Nuovo Cim **26**, 1 (2003), [hep-ex/0312015].

[21] V. Weisskopf and E. Wigner, *Calculation of the natural brightness of spectral lines on the basis of Dirac's theory*, Z. Phys. **63**, 54 (1930), in German, and
———, *Over the natural line width in the radiation of the harmonius oscillator*, Z. Phys. **65**, 18 (1930), in German.

[22] A. Ali, G. Kramer, and C.-D. Lü, *Experimental tests of factorization in charmless nonleptonic two-body B decays*, Phys. Rev. D **58**, 094009 (1998), [hep-ph/9804363];
Y.-Y. Keum and H.-N. Li,, *Nonleptonic charmless B decays: Factorization versus perturbative QCD*, Phys. Rev. D **63**, 074006 (2001), [hep-ph/0006001];
M. Neubert,, *Application of QCD Factorization in Hadronic B decays*, in M. B. Voloshin (ed.), proceedings of the Fourth Workshop on Continuous Advances in QCD, Minneapolis, 256 (2000), [hep-ph/0008072].

[23] M. Wirbel, B. Stech, and M. Bauer, *Exclusive semileptonic decays of heavy mesons*, Z. Phys. C **29**, 637 (1985);
———, *Exclusive non-leptonic decays of D, D_s and B mesons*, Z. Phys. C **34**, 103 (1987).

[24] QCD Factorization was first proposed in,
M. Beneke et al., *QCD Factorization in $B \to \pi\pi$ Decays: Strong Phases and CP Violation in the Heavy Quark Limit*, Phys. Rev. Lett. **83**, 1914 (1999), [hep-ph/9905312] and
———, *QCD Factorization for exclusive non-leptonic B-meson Decays: General arguments and the case for heavy-light final states*, Nucl. Phys. **B 591**, 313 (2000), [hep-ph/0006124];
the most recent update is
M. Beneke and M. Neubert, *QCD factorization for $B \to PP$ and $B \to PV$ decays*, Nucl. Phys. **B 675**, 333 (2003), [hep-ph/0308039].

[25] M. A. Shifman, A. I. Vainshtein, and V. I. Zacharov, *QCD and resonance physics. Theoretical foundations*, Nucl. Phys. **B 147**, 385 (1979).

[26] Y.-Y. Keum, H.-N. Li, and A. I. Sanda, *Fat penguins and imaginary penguins in perturbative QCD*, Phys. Lett. **B504**, 6 (2001), [hep-ph/0004004];
———, *Penguin enhancement and $B \to K\pi$ decays in perturbative QCD*, Phys. Rev. D **63**, 054008 (2001), [hep-ph/0004173].

[27] C. W. Bauer, S. Fleming, and M. E. Luke, *Summing Sudakov logarithms in $B \to X_s\gamma$ in effective field theory*, Phys. Rev. D **63**, 014006 (2001), [hep-ph/0005275];
C. W. Bauer et al., *An effective field theory for collinear and soft gluons: Heavy to light decays*, Phys. Rev. D **63**, 114020 (2001), [hep-ph/0011336];
C. W. Bauer and I. W. Steward, *Invariant operators in collinear effective theory*, Phys. Lett. **B516**, 134 (2001), [hep-ph/0107001].

[28] M. Ciuchini et al., *Charming Penguins in B Decays*, Nucl. Phys. **B 501**, 271 (1997), [hep-ph/9703353].

[29] E. Barberio et al. (Heavy Flavor Averaging Group), *Averages of b-hadron Properties at the End of 2005*, [hep-ex/0603003].

[30] S. Baek et al., *Can one detect new physics in I=0 and/or I=2 contributions to the decays $B \to \pi\pi$?*, Phys. Rev. D **72**, 036004 (2005), [hep-ph/0506075].

[31] See, for instance,
Y. Grossman and H. R. Quinn, *Bounding the effect of penguin diagrams in $a_{CP}(B^0 \to \pi^+\pi^-)$*,

Phys. Rev. D **58**, 017504 (1998), [hep-ph/9712306];
J. Charles, *Taming the penguin contributions in the $B^0(t) \to \pi^+\pi^-$ CP asymmetry: Observables and minimal theoretical input*, Phys. Rev. D **59**, 054007 (1999), [hep-ph/9806468].

[32] A. Khodjamirian, T. Mannel, and M. Melcher, *Flavor SU(3) symmetry in Charmless B Decays*, Phys. Rev. D **68**, 114007 (2003), [hep-ph/0308297];
———, *Kaon distribution Amplitude from QCD Sum Rules*, Phys. Rev. D **70**, 094002 (2004), [hep-ph/0407226];
A. Khodjamirian, T. Mannel, and N. Offen, *Form Factors from Light-Cone Sum Rules with B-Meson Distribution Amplitudes*, [hep-ph/0611193].

[33] I. Dunietz, *Extracting CKM parameters from B decays*, in P. McBride and C.S. Mishra (ed.), Proceedings of the Workshop on B Physics at Hadron Accelerators, Snowmass, 83 (1993).

[34] D. Pirjol, *Bounding the penguin effect in the determination of α in $B^0(t) \to \pi^+\pi^-$*, Phys. Rev. D **60**, 054020 (1999), [hep-ph/9903447].

[35] R. Fleischer and J. Matias, *Searching for New Physics in Non-Leptonic B Decays*, Phys. Rev. D **61**, 074004 (2000), [hep-ph/9906274];
———, *Exploring CP Violation through Correlations in $B \to \pi K$, $B_d \to \pi^+\pi^-$, $B_s \to K^+K^-$ Observable Space*, Phys. Rev. D **66**, 054009 (2002), [hep-ph/0204101].

[36] D. London and J. Matias, *Testing the Standard Model with $B_s^0 \to K^+K^-$ Decays*, Phys. Rev. D **70**, 031502 (2004), [hep-ph/0404009].

[37] S. Descotes-Genon, J. Matias, and J. Virto, *Exploring $B_{d,s} \to KK$ decays through flavour-symmetries and QCD factorisation*, Phys. Rev. Lett. **97**, 061801 (2006), [hep-ph/0603239].

[38] M. Gronau and J. L. Rosner, *Weak phase γ from the ratio of $B \to K\pi$ rates*, Phys. Rev. D **57**, 6843 (1998), [hep-ph/9711246].

[39] M. Gronau and J. L. Rosner, *The role of $B_s^0 \to K^-\pi^+$ in determining the weak phase γ*, Phys. Lett. **B482**, 71 (2000), [hep-ph/0003119].

[40] H. J. Lipkin, *Is observed direct CP violation in $B_d \to K^+\pi^-$ due to new physics? Check standard model prediction of equal violation in $B_s \to K^-\pi^+$*, Phys. Lett. **B621**, 126 (2005), [hep-ph/0503022].

[41] M. Gronau, *U-spin symmetry in Charmless B Decays*, Phys. Lett. **B492**, 297 (2000), [hep-ph/0008292].

[42] F. Su et al., *Large Strong Phases and CP Violation in the Annihilation Processes $\overline{B}^0 \to K^+K^-$, $K^{*\pm}K^{\mp}$, $K^{*+}K^{*-}$*, [0604082].

[43] Y. Li et al., *Branching ratio and CP asymmetry of $B_s \to \pi^+\pi^-$ decays in the perturbative QCD approach*, Phys. Rev. D **70**, 034009 (2004), [hep-ph/0404028].

[44] Y.-D. Yang et al., *Revisiting the annihilation decay $\bar{B}_s \to \pi^+\pi^-$.*, Eur. Phys. J. **C44**, 243 (2005), [hep-ph/0507326].

[45] N. Ellis and A. Kernan, *Heavy quark production at the CERN* $\mathrm{p\bar{p}}$ *collider*, Phys. Rept. **195**, 23 (1990).

[46] See, for instance,
A. A. Affolder et al. (CDF Collaboration), *Measurement of* $\sin 2\beta$ *from* $B \to J/\psi K^0_S$ *with the CDF detector*, Phys. Rev. D **61**, 072005 (2000), [hep-ex/9909003];
F. Abe et al. (CDF Collaboration), *Measurement of the* $B^0\overline{B}^0$ *flavor oscillations frequency and study of same side flavor tagging of* B *mesons in* $\mathrm{p\bar{p}}$ *collisions*, Phys. Rev. D **59**, 032001 (1999), [hep-ex/9806026].

[47] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5* GeV *in 400-*GeV *Proton-Nucleus Collisions*, Phys. Rev. Lett. **39**, 252 (1977)

[48] Ch. Berger et al. (PLUTO Collaboration), *Observation of a narrow resonance formed in* $\mathrm{e^+e^-}$ *annihilation at 9.46* GeV, Phys. Lett. **B76**, 243 (1978);
C. W. Darden et al., *Observation of a Narrow Resonance at 9.46* GeV *in Electron-Positron Annihilations*, Phys. Lett. **B76**, 246 (1978), ibidem Phys. Lett. **B78**, 364 (1978).

[49] D. Andrews et al., *Observation of Three Upsilon States*, Phys. Rev. Lett. **44**, 1108 (1980);
T. Böringer et al., *Observation of* Υ*,* Υ'*, and* Υ'' *at the Cornell Electron Storage Ring*, Phys. Rev. Lett. **44**, 1111 (1980);
D. Andrews et al., *Observation of a Fourth Upsilon State in* $\mathrm{e^+e^-}$ *Annihilation*, Phys. Rev. Lett. **45**, 219 (1980);
G. Finocchiaro et al., *Observation of the* Υ''' *at the Cornell Electron Storage Ring*, Phys. Rev. Lett. **45**, 222 (1980).

[50] N. S. Lockyer et al., *Measurement of the Lifetime of Bottom Hadrons*, Phys. Rev. Lett. **51**, 1316 (1983);
E. Fernandez et al., *Lifetime of Particles Containing b Quarks*, Phys. Rev. Lett. **51**, 1022 (1980).

[51] C. Albajar et al. (UA1 Collaboration), *Search for* $B^0 - \overline{B}^0$ *oscillations at the CERN proton-antiproton collider*, Phys. Lett. **B186**, 247 (1987), erratum-ibidem Phys. Lett. **B197**, 565 (1987).

[52] H. Albrecht et al. (ARGUS Collaboration), *Observation of* $B^0 - \overline{B}^0$ *mixing*, Phys. Lett. **B192**, 245 (1987).

[53] P. Oddone, *Detector Considerations*, in D. H. Stork, proceedings of Workshop on Conceptual Design of a Test Linear Collider: Possibilities for a $B\overline{B}$ Factory, 423 (1987).

[54] B. Aubert et al. (*BABAR* Collaboration), *The BABAR detector*, Nucl. Instrum. Methods **A479**, 1 (2002), [hep-ex/0105044].

[55] M. T Chang et al. (Belle Collaboration), *A Study of CP Violation in B Meson Decays — Technical Design Report*, BELLE-TDR-3-95 (1995);
A. Abashian et al. (Belle Collaboration), *The Belle detector*, Nucl. Instrum. Methods **A476**, 117 (2002).

[56] D. Acosta et al. (CDF Collaboration), *Measurement of the J/ψ and b-Hadron Production Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1960 GeV*, Phys. Rev. D **71**, 032001 (2005), [hep-ex/0412071].

[57] S. Amato et al. (LHCb Collaboration), *LHCb Technical Proposal*, CERN-LHCC-98-04 (1998).

[58] M. Battle et al. (CLEO Collaboration), *Observation of B^0 Decay to Two Charmless Mesons*, Phys. Rev. Lett. **71**, 3922 (1993);
D. M. Asner et al. (CLEO Collaboration), *Search for exclusive charmless hadronic B decays*, Phys. Rev. D **53**, 1039 (1996), [hep-ex/9508004].

[59] R. Godang et al. (CLEO Collaboration), *Observation of exclusive Two-body B Decays to Kaons and Pions*, Phys. Rev. Lett. **80**, 3456 (1998), [hep-ex/9711010].

[60] D. Cronin-Hennessy et al. (CLEO Collaboration), *Study of Two-Body B Decays to Kaons and Pions: Observation of $B \to \pi^+\pi^-$, $B \to K^\pm\pi^0$, and $B \to K^0\pi^0$ Decays*, [hep-ex/0001010] (2000).

[61] S. Chen et al. (CLEO Collaboration), *Measurement of Charge Asymmetries in Charmless Hadronic B Meson Decays*, Phys. Rev. Lett. **85**, 525 (2000), [hep-ex/0001009].

[62] B. Aubert et al. (*BABAR* Collaboration), *Measurement of CP Asymmetries and Branching Fractions in $B \to \pi\pi$ and $B \to K\pi$ Decays*, to appear in the proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0607106].

[63] B. Aubert et al. (*BABAR* Collaboration), *Improved Measurements of the Branching Fractions for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, and a Search for $B^0 \to K^+K^-$*, [hep-ex/0608003], submitted to Phys. Rev. D.

[64] Y. Unno for the Belle Collaboration, presented at the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow.

[65] K. Abe et al. (Belle Collaboration), *Improved measurements of branching fractions for $B \to K\pi$ and $B \to \pi\pi$ decays*, to appear in the proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0609015].

[66] K. Abe et al. (Belle Collaboration), *Observation B Decays to Two Kaons*, to appear in proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0608049].

[67] D. Buskulic et al. (ALEPH Collaboration), *Observation of charmless hadronic B decays*, Phys. Lett. **B384**, 471 (1996).

[68] See, for instance
E. Artouni et al. (HERA-B Collaboration), *HERA-B. An Experiment to Study CP Violation in the B System Using an Internal Target at the HERA Proton Ring*, DESY-PRC-95-01 (1995).

[69] D. P. McGinnis, *Fermilab Tevatron operational status*, in C. Horak (ed.), proceedings of the Particle Accelerator Conference (PAC 05), (2005); in addition, detailed and updated information on the Tevatron is available in the following web-pages:
`http://www-bd.fnal.gov/runII/index.html`
`http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html`.

[70] D. Mohl et al., *Physics and Technique of Stochastic Cooling*, Phys. Rept. **58**, 73 (1980).

[71] S. Nagaitsev et al., *Experimental Demonstration of Relativistic Electron Cooling*, Phys. Rev. Lett. **96**, 044801 (2006).

[72] These numbers, based on a bunch-by-bunch study, come from a private communication with L. Ristori.

[73] R. Blair et al. (CDF II Collaboration), *The CDF II Detector, Technical Design Report*, FERMILAB-Pub-96/390-E CDF (1996).

[74] C. S. Hill, *Initial experience with the CDF layer 00 silicon detector*, Nucl. Instrum. Methods **A511**, 118 (2003).

[75] A. Sill, *CDF Run II silicon tracking projects*, Nucl. Instrum. Methods **A447**, 1 (2000).

[76] A. Affolder et al., *Status report of the intermediate silicon layers detector at CDF II*, Nucl. Instrum. Methods **A485**, 6 (2002).

[77] T. Affolder et al., *CDF Central Outer Tracker*, Nucl. Instrum. Methods **A526**, 249 (2004).

[78] J. Marriner, *Secondary vertex fit with mass and pointing constraints (CTVMFT)*, CDF Internal Note 1996 (1993), unpublished.

[79] P. Gatti, *Performance of the new tracking system at CDF II*, Tesi di Dottorato (Ph. D. thesis), University of Padova, FERMILAB-THESIS-2001-23 (2001);
S. Menzemer, *TrackingCal - A tracking and alignment software package for the CDF II silicon detector*, Ph. D. thesis, University of Karlsruhe and CDF Internal Note 5968 (2002), unpublished.

[80] C. Hays et al., *The COT Pattern Recognition Algorithm and Offline Code*, CDF Internal Note 6992 (2004), unpublished.

[81] C. S. Hill et al., *Operational experience and performance of the CDF II silicon detector*, Nucl. Instrum. Methods **A530**, 1 (2004).

[82] D. Acosta et al., *A Time-of-Flight detector in CDF-II*, Nucl. Instrum. Methods **A518**, 605 (2004).

[83] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nucl. Instrum. Methods **A267**, 272 (1988);
S. R. Hahn et al., *Calibration systems for the CDF central electromagnetic calorimeter*, Nucl. Instrum. Methods **A267**, 351 (1988).

[84] M. Gallinaro, *A New Scintillator Tile/Fiber Preshower Detector for the CDF Central Calorimeter*, IEEE Trans. Nucl. Sci. **52**, 879 (2005), [physics/0411056].

[85] S. Bertolucci et al., *The CDF central and endwall hadron calorimeter*, Nucl. Instrum. Methods **A267**, 301 (1988).

[86] M. Albrow et al., *The CDF plug upgrade electromagnetic calorimeter: test beam results*, Nucl. Instrum. Methods **A480**, 524 (2002);
M. Albrow et al., *A preshower detector for the CDF Plug Upgrade: test beam results*, Nucl. Instrum. Methods **A431**, 104 (1999);
G. Apollinari et al., *Shower maximum detector for the CDF plug upgrade calorimeter*, Nucl. Instrum. Methods **A412**, 515 (1998).

[87] G. Ascoli et al., *CDF central muon detector*, Nucl. Instrum. Methods **A268**, 33 (1988);
C. M. Ginsburg, *CDF Run 2 Muon System*, Eur. Phys. J. **33**, S1002 (2004).

[88] C. Avila et al. (E811 Collaboration), *A measurement of the proton-antiproton total cross-section at* $\sqrt{s}$*=1.8* TeV, Phys. Lett. **B445**, 419 (1999);
F. Abe et al. (CDF Collaboration), *Measurement of the anti-proton proton total cross-section at* $\sqrt{s}$*=546* GeV *and 1800* GeV, Phys. Rev. D **50**, 5550 (1994).

[89] D. Acosta et al., *The performance of the CDF luminosity monitor*, Nucl. Instrum. Methods **A494**, 57 (2002).

[90] K. Goulianos et al., *The CDF miniplug calorimeters at the Tevatron*, Nucl. Instrum. Methods **A518**, 42 (2004), [physics/0310156].

[91] M. Karagoz Unel and R. J. Tesarek, *Beam Halo monitoring at CDF*, Nucl. Instrum. Methods **A506**, 7 (2003), [hep-ex/0211055].

[92] E. J. Thomson et al., *Online Track Processor for the CDF Upgrade*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).

[93] For more details on the SVT, see the following conference proceedings (and references therein):
B. Ashmanskas et al., *The CDF Silicon Vertex Trigger*, Nucl. Instrum. Methods **A518**, 532 (2004), [physics/0306169];
M. Dell'Orso, *The CDF Silicon Vertex Trigger*, Nucl. Phys. **B** (Proc. Suppl.) **156**, 139 (2006).

[94] G. Gómez-Ceballos et al., *Event Builder and Level 3 at the CDF experiment*, Nucl. Instrum. Methods **A518**, 522 (2004).

[95] J. Antos et al., *Data processing model for the CDF experiment*, submitted to IEEE Trans. Nucl. Sci., [physics/0606042] (2006).

[96] E. Gerchtein and M. Paulini, *CDF detector simulation framework and performance*, eConf. **C0303241** TUMT005 (2003) [physics/0306031] (2003).

[97] R. Brun et al., *GEANT: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual*, CERN-DD-78-2-REV (1978).

[98] P. A. Movilla-Fernandez, *Performance of the CDF Calorimeter Simulation in Tevatron Run II.*, in R. Yoshida S. Magill (ed.), Proceedings of the XII International Conference on Calorimetry in High Energy Physics (CALOR 06), [physics/0608081] (2006).

[99] R. Veenhof, *GARFIELD, recent developments.*, Nucl. Instrum. Methods **A419**, 726 (1998).

[100] M. Gold, *Description of the Parameterized Charge Deposition Model*, CDF Internal Note 5871 (2002), unpublished.

[101] S. Donati and G. Punzi, *Two Track trigger with full SVT simulation*, CDF Internal Note 3780 (1998), unpublished.

[102] G. Punzi, S. Donati, and G. Gagliardi, *A $B^0 \to \pi^+\pi^-$ trigger for CDF*, in P. McBride and C.S. Mishra (ed.), Proceedings of the Workshop on B Physics at Hadron Accelerators, Snowmass, 227 (1993);
S. Donati, *Una strategia per la misura della asimmetria* CP *nel decadimento $B^0 \to \pi^+\pi^-$ a CDF*, Tesi di Dottorato (Ph. D. thesis), University of Pisa (1996), in Italian, unpublished.

[103] For an exact definition of the $\chi^2_{\rm SVT}$ quantity, see
S. Belforte et al., *Silicon Vertex Tracker Technical Design Report*, CDF Internal Note 3108 (1995), unpublished.

[104] For a simple introduction to the concepts of Kalman filtering, see
P. Maybeck, *Stochastic Models, Estimation, and Control*, Vol. 1, New York: Academic Press, 1979, and references therein.

[105] D. Acosta et al. (CDF Collaboration), *Measurement of Bottom-Quark Hadron Masses in Exclusive J/ψ Decays with the CDF Detector*, Phys. Rev. Lett. **96**, 202001 (2006), [hep-ex/0508022].

[106] D. Acosta et al. (CDF Collaboration), *Measurement of Partial Widths and Search for Direct CP Violation in D^0 Meson Decays to K^-K^+ and $\pi^-\pi^+$*, Phys. Rev. Lett. **94**, 122001 (2005), [hep-ex/0504006].

[107] P. Sphicas, *A $b\bar{b}$ Monte Carlo Generator*, CDF Internal Note 2655 (1994), unpublished;
K. Anikeev, P. Murat, and Ch. Paus, *Description of Bgenerator II*, CDF Internal Note 5092 (1999), unpublished.

[108] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods **A462**, 152 (2001);
W. Bell et al., *User Guide for EvtGen @ CDF*, CDF Internal Note 5618 (2003), unpublished.

[109] S. Eidelman et al. (Particle Data Group), *Review of particle physics*, Phys. Lett. **B592**, 1 (2004).

[110] M. Beneke et al., *Next-to-Leading Order QCD Corrections to the Lifetime Difference of B_s Mesons*, Phys. Lett. **B459**, 631 (1999), [hep-ph/9808385];
I. Dunietz, R. Fleischer, and U. Nierste, *In pursuit of new physics with B_s decays*, Phys. Rev. D **63**, 114015 (2001), [hep-ph/0012219].

[111] Although the inequality was first (implicitly) derived by Aitken and Silverstone, it is generally referred to as Cramér-Rao (or Fréchet) inequality, following the papers below:
C. R. Rao, *Information and accuracy attainable in the estimation of statistical parameters*, Bull. Calcutta Math. Soc., **37**, 81 (1945);
H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press (1946).

[112] J. D. Bjorken, *Properties of hadron distributions in reactions containing very heavy quarks*, Phys. Rev. D **17**, 171 (1978);
M. Suzuki, *Fifth quark and dimuon production by Neutrinos*, Phys. Lett. **B68**, 164 (1977).

[113] A few CDF Run I analyses used an isolation criterion as a discriminating observable:
D. Acosta et al. (CDF Collaboration), *Search for the Decay $B_s \to \mu^+\mu^-\phi$ in $\mathrm{p\bar{p}}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **65**, 111101 (2002);
———, *Search for radiative b-hadron decays in proton anti-proton collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. D **66**, 112002 (2002), [hep-ph/0208035].

[114] S. D. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High Energies*, Phys. Rev. Lett. **25**, 316 (1970), erratum-ibidem Phys. Rev. Lett. **25**, 902 (1970).

[115] See, for instance,
R. Mohanta, A. K. Giri, and M. P. Khanna, *Charmless two body hadronic decays of the Λ_b baryon*, Phys. Rev. D **63**, 074001 (2001), [hep-ph/0006109].

[116] D. Ayres et al., *Design Report Of the Fermilab Collider Detector Facility (Cdf), August 1981*, Batavia, USA, Fermilab (1981);
F. Abe et al. (CDF Collaboration), *The CDF Detector: An Overview*, Nucl. Instrum. Methods **A271**, 387 (1988).

[117] M. Binkley et al., *Aging in large CDF tracking Chambers*, Nucl. Instrum. Methods **A515**, 53 (2003);
D. Allspach et al., *Aging in the large CDF axial drift chamber*, IEEE Trans. Nucl. Sci. **52**, 2956 (2005).

[118] D. Acosta et al. (CDF Collaboration), *Evidence for $B_s^0 \to \phi\phi$ Decay and Measurements of Branching Ratio and A_{CP} for $B^+ \to \phi K^+$*, Phys. Rev. Lett. **95**, 031801 (2005), [hep-ex/0502044].

[119] dE/dx information plays an important role in the algorithms for inferring the flavor of the decaying B_s^0 meson in the following measurements:
A. Abulencia et al. (CDF Collaboration), *Measurement of the $B_s^0 - \overline{B}_s^0$ Oscillation Frequency*, Phys. Rev. Lett. **97**, 062003 (2006), [hep-ex/0606027] and
———, *Observation of $B_s^0 - \overline{B}_s^0$ Oscillations*, submitted to Phys. Rev. Lett., [hep-ex/0609040].

[120] D. Acosta et al. (CDF Collaboration), *Measurement of Prompt Charm Meson Production Cross Sections in* $p\bar{p}$ *Collisions at* $\sqrt{s} = 1.96$ TeV, Phys. Rev. D **91**, 241804 (2003), [hep-ex/0307080].

[121] A. Abulencia et al. (CDF Collaboration), *Measurement of the Ratio of Branching Fractions* $\mathcal{B}(D^0 \to K^-\pi^+)/\mathcal{B}(D^0 \to K^+\pi^-)$, Phys. Rev. D **74**, 031109(R) (2006), [hep-ex/0605027].

[122] C. Niebuhr, *Aging in the Central Jet Chamber of the H1 experiment*, Nucl. Instrum. Methods **A515**, 43 (2003).

[123] M. D. Peters, *A measurement of B Meson Oscillations using Inclusive Leptons in Proton-Antiproton Collisions at 1.8* TeV, Ph. D. thesis, University of California, Berkeley, FERMILAB-THESIS-1997-19 (1997).

[124] G. Punzi, *Useful formulas on statistical separation of classes of events*, [physics/0611219].

[125] The original papers in which the basic principles of the maximum likelihood estimation were developed are
R. A. Fisher, *On an absolute criterion for fitting frequency curves*, Messenger of Mathematics, **41**, 155 (1912);
———, *On the probable error of a coefficient of correlation deduced from a small sample*, Metron., **1**, 3 (1921);
———, *On the mathematical foundation of theoretical statistics*, Philos. Trans. Roy. Soc. London Ser. A **222**, 309 (1922).
For a modern description of the ML method and its properties, see pag. 46 of
A. Stuart, K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics - Vol. 2A, Classical Inference & the Linear Model*, London: Arnold, 6th edition (1999).

[126] F. James and M. Roos, *'MINUIT' A system for Function Minimization And Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10**, 343 (1975);
F. James, *MINUIT – Function Minimization and Error Analysis – Reference Manual*, CERN Program Library **D506** (1998).

[127] For a discussion on the bias related to the use of multiple variables in multi-component fits see G. Punzi, *Comments on Likelihood fits with variable resolution*, eConf. **C030908** WELT002 (2003), [physics/0401045].

[128] Detailed documentation on the MATHEMATICA™ software package is available at `http://www.wolfram.com`.

[129] K. Pearson, *Contributions to the Mathematical Theory of Evolution – II. Skew Variation in Homogeneous Material*, Philos. Trans. Roy. Soc. London A **186**, 343 (1895);
———, *Contributions to the Mathematical Theory of Evolution – X. Supplement to a Memoir on Skew Variation*, Philos. Trans. Roy. Soc. London A **197**, 443 (1901);
———, *Contributions to the Mathematical Theory of Evolution – XIX. Second Supplement to a Memoir on Skew Variation*, Philos. Trans. Roy. Soc. London A **216**, 429 (1916).

[130] A. Abulencia et al. (CDF Collaboration), *Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ Decays in $p\bar{p}$ Collisions with CDF II*, Phys. Rev. Lett. **95**, 221805 (2005), erratum-ibidem Phys. Rev. Lett. **95**, 249905 (2005), [hep-ex/0508036].

[131] D. Acosta et al. (CDF Collaboration), *Search for $\Lambda_b^0 \to p\pi^-$ and $\Lambda_b^0 \to pK^-$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **72**, 051104(R) (2005), [hep-ph/0507067].

[132] A. B. Wicklund and K. Byrum, *Calibration and Applications of the dE/dx from the CTC*, CDF Internal Note 2624 (1995), unpublished.

[133] J. Heinrich, *Pitfalls of Goodness-of-fit from Likelihood*, eConf. **C030908** MOcT001 (2003), [physics/0310167].

[134] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57**, 3873 (1998), [hep-ex/0605027].

[135] A. Abulencia et al. (CDF Collaboration), *Measurement of the Ratios of Branching Fractions $\mathcal{B}(B_s^0 \to D_s^-\pi^+)/\mathcal{B}(B^0 \to D^-\pi^+)$ and $\mathcal{B}(B^+ \to \overline{D}^0\pi^+)/\mathcal{B}(B^0 \to D^-\pi^+)$*, Phys. Rev. Lett. **96**, 191801 (2006), [hep-ex/0508014].

[136] D. Acosta et al. (CDF Collaboration), *Measurement of the Lifetime Difference between B_s Mass Eigenstates*, Phys. Rev. Lett. **94**, 101803 (2005), [hep-ex/0412057].

[137] Several introductory books on statistics for physicists treat the concept of extended likelihood, see, for instance,
G. Cowan, *Statistical Data Analysis*, Oxford: Oxford University Press, (1998); L. Lyons, *Statistics for nuclear and particle physicists*, Cambridge: Cambridge University Press, (1986); R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, Chichester: John Wiley & Sons, (1989).

[138] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Statist. **9**, 60 (1938).

[139] E. Barberio, B. Van Eijk, and Z. Was, *PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. **66**, 115 (1991);
E. Barberio and Z. Was, *PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0*, Comput. Phys. Commun. **79**, 291 (1994).

[140] V. Cirigliano, J. F. Donoghue, and E. Golowich, *$K \to \pi\pi$ phenomenology in the presence of electromagnetism*, Eur. Phys. J. **C**18, 83 (2000), [hep-ph/0008290].

[141] E. Baracchini and G. Isidori, *Electromagnetic corrections to non-leptonic two-body B and D decays*, Phys. Lett. **B633**, 309 (2006), [hep-ph/0508071].

[142] This estimate of the difference between transverse-momentum spectra of B_s^0 and B_s^0 meson was suggested in a private communication by M. L. Mangano and M. Cacciari.

[143] A. Williamson and J. Zupan, *Two body B decays with isosinglet final states in SCET*, Phys. Rev. D **74**, 014003 (2006) erratum-ibidem Phys. Rev. D **74**, 03901 (2006) [hep-ph/0601214].

[144] J. Sun, G. Zhu, and D. Du, *Phenomenological Analysis of Charmless Decays $B_s \to PP, PV$ with QCD Factorization*, Phys. Rev. D **68**, 054003 (2003), [hep-ph/0211154].

[145] A. Buras et al., *Anatomy of prominent B and K Decays and Signatures of CP-Violating New Physics in the Electroweak Penguin Sector*, Nucl. Phys. **B 697**, 133 (2004), [hep-ph/0402112].

[146] C.-W. Chiang and Y-F. Zhou, *Flavor SU(3) analysis of charmless B meson decays to two pseudoscalar mesons*, [hep-ph/0609128].

[147] B. Aubert et al. (*BABAR* Collaboration), *Evidence for $B^+ \to \bar{K}^0 K^+$ and $B^0 \to K^0 \bar{K}^0$, and Measurement of the Branching Fraction and Search for Direct CP Violation in $B^+ \to \bar{K}^0 K^+$*, Phys. Rev. Lett. **95**, 221801 (2005), [hep-ex/0507023].

[148] S. Baek et al., *$B_s^0 \to K^+K^-$ and $B^0 \to K^0\bar{K}^0$ Decays within Supersymmetry*, [hep-ph/0610109].

[149] K. Abe et al. (Belle Collaboration), *Observation of Direct CP-Violation in $B \to \pi\pi$ Decays with 535 Million $B\overline{B}$ pairs*, to appear in the proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0608035].

[150] X.-Q. Yu, Y. Li, and C.-D. Lü, *Branching ratio and CP asymmetry of $B_s \to \pi K$ decays in the perturbative QCD approach*, Phys. Rev. D **71**, 074026 (2005), erratum-ibidem Phys. Rev. D **72**, 119903 (2005), [hep-ph/0501152].

[151] J. Malclès, *Depuzzling $B \to K\pi$: constraints on the unitarity triangle from $B, B_s \to \pi\pi, K\pi, KK$ decays in the SU(3) limit*, to appear in the proceedings of the 41st Rencontres de Moriond on Electroweak Interactions and Unified Theories, [hep-ph/0606083].

[152] J. Charles et al. (CKMfitter Group), *CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories*, Eur. Phys. J. **C41**, 1 (2005), [hep-ph/0406184] and web-updates at `http://ckmfitter.in2p3.fr`.

[153] G. Bonvicini et al. (CLEO Collaboration), *Observation of B_s^0 Production at the $\Upsilon(5S)$ resonance*, Phys. Rev. Lett. **96**, 022002 (2006).

[154] A. Drutskoy et al. (Belle Collaboration), *Measurement of inclusive D_s, D^0 and J/ψ rates and determination of the $B_s^{(*)}\overline{B}_s^{(*)}$ production fraction in $b\bar{b}$ events at the $\Upsilon(5S)$ resonance*, to appear in proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0608015];
K. Abe et al. (Belle Collaboration), *Measurement of exclusive B_s^0 Decays at the $\Upsilon(5S)$.*, to appear in proceedings of the XXXIII International Conference on High Energy Physics (ICHEP2006), Moscow, [hep-ex/0610003].

[155] D. Tonelli for the CDF Collaboration, *CDF Hot Topics*, eConf. **C060409** (2006), [hep-ex/0605038].

[156] M. J. Morello for the CDF Collaboration, *Branching ratios and direct CP asymmetries of charmless decay modes at the Tevatron*, talk given at the 11th International Conference on B-Physics at Hadron Machines (BEAUTY2006).

[157] H. A. Bethe, *Theory of passage of fast corpuscular rays through matter*, Annalen Phys. **5**, 325 (1930), in German;
———, *Scattering of electrons*, Z. Phys. **76**, 293 (1932);
F. Bloch, *Stopping power of atoms with several electrons*, Z. Phys. **81**, 363 (1933).

[158] Detailed information on energy-loss processes are available in the following references:
R. Talman, *On the Statistics of Particle Identification Using Ionization*, Nucl. Instrum. Methods 159 (1979);
W. W. M. Allison and J. H. Cobb, *Relativistic charged particle identification by energy loss*, Ann. Rev. Nucl. Part. Sci. **30**, 253 (1980);
F. Sauli, *Principles of operations of multiwire proportional and drift chambers*, in *Experimental Techniques in High Energy Physics*, T. Ferbel (ed.), Menlo Park: Addison-Wesley (1987);
W. Blum and L. Rolandi, *Particle detection with drift chambers*, Berlin: Springer-Verlag (1993).

[159] S.-S. 'Eiko' Yu et al., *COT dE/dx Measurement and Corrections*, CDF Internal Note 6361 (2004), unpublished.

[160] W. C. Davidon, *Variance algorithm for minimization*, Comput. J. **10**, 406 (1968);
R. Fletcher and M. J. D. Powell, *A rapidly converging descent method for minimization*, Comput. J. **6**, 162 (1963).