Analysis of B_s flavor oscillations at CDF

Nuno T. Leonardo

Doctoral Dissertation

September 2006

Analysis of B_s flavor oscillations at CDF

by

Nuno Teotónio Viegas Guerreiro Leonardo

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

Doctor of Phylosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2006

Author . . . .

Department of Physics
September 21, 2006

Certified by . . . .

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

Accepted by . . . .

Thomas J. Greytak
Associate Department Head for Education, Professor

Analysis of B_s flavor oscillations at CDF

by

Nuno Teotónio Viegas Guerreiro Leonardo

Submitted to the Department of Physics
on September 21, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Phylosophy

Abstract

The search for and the study of flavor oscillations in the neutral $B_s\bar{B}_s$ meson system is an experimentally challenging endeavor. This constitutes a flagship analysis of the Tevatron proton–antiproton collider physics program. In this dissertation we develop such an analysis of the time-dependent B_s flavor oscillations using data collected with the CDF detector. The data samples are formed of both fully and partially reconstructed B meson decays, $B_s \to D_s\pi(\pi\pi)$ and $B_s \to D_s l\nu$.

A likelihood fitting framework is implemented and appropriate models and techniques developed for describing the mass, proper decay time, and flavor tagging characteristics of the data samples. The analysis is extended to samples of B^+ and B^0 mesons, which are further used for algorithm calibration and method validation. The B mesons lifetimes are extracted. The measurement of the B^0 oscillation frequency yields $\Delta m_d = 0.522 \pm 0.017$ ps^{-1}.

The search for B_s oscillations is performed using an amplitude method based on a frequency scanning procedure. Applying a combination of lepton and jet charge flavor tagging algorithms, with a total tagging power $\epsilon\mathcal{D}^2$ of 1.6%, to a data sample of 355 pb^{-1}, a sensitivity of 13.0 ps^{-1} is achieved.

In a preliminary study, we develop a same side kaon tagging algorithm, which is found to provide a superior tagging power, of about 4.0% for the B_s meson species. An evaluation of the dilution systematic uncertainties is not reported but after including the algorithm as is in the analysis the sensistivity is significantly increased to about 18 ps^{-1}. An indication of a B_s oscillation signal is seen at a frequency of about 17.5 ps^{-1}. We show that the extension of the analysis to the increasing data samples with the inclusion of the same side tagging algorithm is capable of providing an observation of B_s mixing beyond the standard model expectation. We show also that the improved knowledge of Δm_s has a considerable impact on constraining the CKM matrix elements.

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

Contents

Em memória de meus avós
Jorge e Francisca

Preamble

To discover what the universe is made of and how it works, probing the deepest structure of space-time, is the challenge of particle physics. It is the aim to understand the elementary constituents of matter and their interactions. Such constituents, despite their dynamics being governed by quantum mechanics along with special relativity, are referred to as elementary *particles*. Theoretical models are developed to describe the particles and their interactions, which are based on and, according to the principles of modern science, ought to be confronted with experiment.

From signals in intricate detector systems to grand theories

Most elementary particles do not exist under normal circumstances in nature, but can be created and detected during energetic collisions of other particles. This is achieved in the laboratory with particle accelerators and detectors. The higher the energy of the collision, the deeper the structure of matter can be probed. Fermilab's Tevatron accelerates protons and antiprotons in an underground ring of 1 kilometer radius, and smashes them head-on to produce the world's highest energy collisions currently obtained in the laboratory.

The remnants of such collisions are recorded as they traverse the material of detector systems surrounding the collision point. The traces of their passage, registered in the form of electronic signals, are used to reconstruct the particles' trajectories and ultimately infer their properties and interactions. The CDF detector, used in this dissertation, is an apparatus of about 12 meters in all three dimensions. It is curious that for studying the smallest objects in nature the largest research apparatuses are required. CDF has been built and is operated by an international collaboration of about 700 physicists.

CDF is a general purpose detector, with a broad physics program aimed at scrutinizing the existing theoretical models and searching for exotic, unexpected phenomena. Despite the remarkable success of the current standard theoretical formulation in accounting for the experimentally observed phenomena, it is believed to be an incomplete description of nature, and that a more fundamental theory awaits discovery. The overarching goal, which is pursued in several distinct ways, is to find and understand what physics may lie beyond. It is the beauty and sophistication of experiments, the mathematical elegance and intricacy of theoretical conjectures, and all that lies in between that makes particle physics a most exciting human endeavor.

The fluctuating identity of neutral B mesons

B mesons are particles with very interesting properties. They are unstable, decaying to more stable, lighter products shortly after being produced in the collision. Nevertheless, they still live for a relatively long time, compared to other decaying sub-atomic particles, about 1.5 ps, flying a few millimeters in the center of the detector before decaying. One can then hope that they do something interesting while they are "alive". Indeed, the neutral B mesons have the sui generis property of undergoing spontaneous transitions to their antiparticles, and back again. These transitions between matter and antimatter are extremely rapid. This is especially true for the B_s (pronounced "B sub s") mesons, for which such *flavor oscillations* are expected to occur at a rate of several trillion times per second. As it can be expected, it is extremely difficult to measure oscillations this rapid, and their observation has in fact evaded the many experimental attempts performed in the past.

The relevance of the study of B_s flavor oscillations goes well beyond the determination of a peculiar property of this meson system. The measurement is used to constrain fundamental parameters of the underlying flavor model. It can constitute a probe for new physics, or even shed light on the mechanisms responsible for the observed matter–antimatter asymmetry in the universe. It is a major flagship measurement of the Tevatron physics program and, arguably, the most complex analysis ever attempted at a hadron collider.

The analyses

The search for and the study of B_s flavor oscillations constitute the objective of the analyses developed in this dissertation. After having collected large samples of B_s mesons their properties need to be extracted. The mesons' flight distance in the detector needs to be precisely measured, within tens of micrometers. Their momenta must also be determined from its decay products as accurately as possible. These quantities provide a determination of the time between production and decay of the B_s meson, in its own rest-frame. Finally, one must determine its flavor, *i.e.* whether it was in a particle or antiparticle state, both when it decayed and at the time it was produced. The latter task is achieved with techniques named *flavor tagging*. Oscillations are searched for as time dependent flavor asymmetries in the data samples. In order to pin down the oscillation frequency most precisely, a sophisticated, multivariate mathematical fit of the data is developed and performed. The samples and all the developed tools and techniques must be thoroughly understood and calibrated. This leads to the use of yet additional samples, containing higher yields of lighter B meson species, namely B^0 and B^+, for method validation and tool calibration, which adds further to the complexity of the procedure.

This document is organized as follows. An overview of the theoretical foundations of neutral B meson mixing, and of the experimental methods and strategy employed is presented in Chapter 1. Chapter 2 gives a brief description of the accelerator complex and of the detector apparatus. The latter is employed to collect the data samples used in the dissertation, which themselves are presented and characterized in Chapter 3. In Chapter 4 the fitting framework is first introduced, and likelihood techniques developed for befittingly describing the data samples at hand, leading to the measurement of the B mesons' lifetimes. The flavor tagging methods are introduced in Chapter 5, while a novel algorithm is further developed in Chapter 8. The flavor tagging information is incorporated into the likelihood description and applied to samples of the lighter B species in Chapter 6. As a result, a measurement of the oscillation frequency in the B^0 system is accomplished therein, along with a simultaneous calibration of the tagging techniques themselves. The calibrated flavor tagging algorithms are finally applied to the samples of B_s mesons in Chapter 7, where a search for their rapid oscillations is carried out employing a suitable frequency scanning method. The obtained results are combined with other existing measurements in Chapter 9 for constraining the flavor model parameters.

Chapter 1

Introduction

The phenomenon of B–$\bar{B}$ flavor oscillations, or mixing, refers to particle-antiparticle transitions in the neutral B meson systems, where B can be either B^0 or B_s, formed respectively of a bottom antiquark and either a down or a strange quark. The rate at which these transformations occur is characterized by the mass difference between the two quantum states of the meson, denoted Δm. Neutral B meson mixing is currently the subject of very intense research in high energy physics. As it will be seen shortly, mixing measurements address many outstanding questions relating to the standard model (SM) of particle physics. In particular it directly relates to fundamental parameters of the theory, shedding light on the processes of CP violation, mass scales, and generation structure.

Experiments in the past few decades have verified the SM gauge structure of elementary particle interactions, in a comprehensive and very precise way. The flavor sector had not been explored to the same extent, and it constitutes one of the highest priorities of contemporary high energy physics. The SM serves as the current foundation for discussing flavor physics. The quark masses, their mixing and CP violation have there a common origin. Namely, they arise from Yukawa interactions with the (yet to be observed) Higgs field. However, all mass scales and their hierarchies are left unexplained, *i.e.* they are not predicted and need to be fixed by empirical inputs. This constitutes a shortcoming of the model.

There are several additional powerful experimental and phenomenological reasons to expect new physics (NP). The SM is thus often taken, a perspective we share, as a (remarkable) *effective* theory describing nature at the energy scales probed so far. Many extensions of the SM have been sought, ranging from compositeness, low or high scale supersymmetry (SUSY), grand unification (GUT), to extra dimensions and deconstruction, and string/M theories. Each of them overcomes some unsatisfactory aspects of the SM while adding considerably more unknowns.

There are theoretical hints that NP should be found at accessible energy scales. The available experimental data exclude, however, the possibility of sizable contributions to tree-level (*i.e.* lowest order perturbation theory) SM processes. Large new contributions are most likely to be present in loop-mediated processes. This renders meson mixing a fertile ground to reveal the influence of new interactions.

The measurement of B_s oscillations is thus quite interesting from a phenomenological perspective. If on the one hand the oscillation frequency is found within the expected SM region, it becomes a precision measurement providing strong constraints on the model parameters. In that case it will add to the remarkable success of the SM and further corroborate it as an effective theory at the energy scales so far probed. At the same time, it places further bounds on extensions of the SM, emphasizing that if there is NP at the TeV scale it must have a special flavor structure to satisfy the observed constraints. On the other hand, if the measurement is found incompatible with the existing SM constraints, it will constitute a direct indication of new physics. In either case, it is bound to be a most exciting result. From the experimental point of view, it is an extremely challenging endeavor and, most definitely, no less exciting.

Below we present the foundations of flavor oscillations within the standard flavor sector description and corresponding parameter constraining, followed by a survey of previous mixing measurements, and conclude with a synopsis and strategy of the data analyses to be performed in the dissertation.

1.1 Phenomenology of flavor oscillations

The quarks, along with the leptons (electron e, muon μ, tau τ, and associated neutrinos), constitute the known building blocks of matter. They appear as six *quark-flavors*: the so-called up-type (u, c, t) and down-type (d, s, b) quarks have electric charges $+\frac{2}{3}\mathsf{e}$ and $-\frac{1}{3}\mathsf{e}$, respectively (e being the magnitude of the electron's charge); the associated antiparticles have symmetric charges. The quark masses [1] (computed in a common renormalization scheme) present the following observed hierarchy

u (up)	c (charm)	t (top)
1.5-4 MeV/c^2	1.15-1.35 GeV/c^2	(174.3±5.1) GeV/c^2
d (down)	s (strange)	b (bottom)
4-8 MeV/c^2	80-130 MeV/c^2	4.1-4.4 GeV/c^2

which is not understood, lacking an experimentally established, fundamental theory of masses.

Transitions take place between the up-type and down-type quarks, with relative strengths which may be represented in a matrix of the form

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

Within the SM, such quark-flavor changing transitions are accomplished by the weak interactions, and the quark mixing matrix above is *unitary* and referred to as the CKM matrix. Flavor changing neutral currents (FCNC) corresponding to transitions among the up-type quarks or the down-type quarks are, at lowest level in perturbation theory, absent. The CKM elements correspond to fundamental parameters not predicted by the theory; they must be extracted from data.

Unlike leptons, quarks are not observed as physical particles, rather they are confined inside hadrons. Among these, the neutral B meson systems are of central importance for the phenomena we set out to study. These meson states are formed of a bottom (anti-) quark together with either a down or a strange quark,

$$B^0(\bar{b}d)\,,\bar{B}^0(b\bar{d}) \qquad \text{and} \qquad B_s(\bar{b}s)\,,\bar{B}_s(b\bar{s})\,.$$

The aforementioned quark-flavor changing interactions induce transitions among the neutral B and $\bar{B}$ states, for both the B^0 and B_s meson systems. By means of those (weak) interactions, the effective Hamiltonian describing the two-state system acquires off-diagonal contributions, such that the system's mass eigenstates differ from those states of definite quark-flavor. As a result, time-dependent transitions develop between the B and $\bar{B}$ states. These transitions are referred to as B meson *mixing* or *flavor oscillations*, and occur with a frequency determined by the difference Δm between the two mass eigenvalues of the system. The quantum mechanical description of the associated particle-antiparticle system is detailed in Appendix A.

By performing measurements of processes governed by the above mentioned weak transitions one is in principle able to obtain information about the underlying parameters of the theory, V_{ij}. In particular, the measurement of the oscillation frequencies of the B^0 and B_s systems, denoted respectively Δm_d and Δm_s, provides a determination of the moduli of the matrix elements V_{td} and V_{ts}. Such determinations in general suffer, nevertheless, from complications stemming from the fact that the quarks are confined in hadrons, and associated theoretical uncertainties enter into the interpretation of the measurements. These uncertainties in the case of the B^0 and B_s mixing systems are considerably large. As may be expected, however, since these mesons differ only by their light quarks, a significant uncertainty reduction is obtained for the ratio of the oscillation frequencies of the two systems. As a result, a precise determination of the ratio of the CKM elements $|V_{td}/V_{ts}|$ is possible.

Various other processes are sensitive to the different matrix elements. These involve weak decays and asymmetries, not necessarily restricted to B meson systems. Such measurements along with B mixing translate into constraints on the elements of the CKM matrix, and together are used to test its unitarity and the standard flavor description. Furthermore, B mixing, which occurs in the SM only at loop level, is expected to be particularly sensitive to new physics sources.

Next we further discuss the CKM ansatz, the origin and parameterization of the quark mixing matrix, and how measurements of flavor oscillations may be employed to constrain the underlying flavor model.

1.1.1 Quark masses and flavor changing interactions

Quark masses

The mass term in the Lagrangian density for a Dirac spinor field $\psi(x)$ is of the form

$$-m\bar{\psi}\psi \tag{1.1}$$

where m is the fermion mass, and $\bar{\psi}$ is the adjoint field ($\bar{\psi} \equiv \psi^\dagger \gamma^0$; γ^μ will refer to a consistent set of Dirac matrices [2]). Expanded in terms of the left and right chirality projections ($\psi_{R,L} \equiv \frac{1 \pm \gamma^5}{2} \psi$), (1.1) may be written as: $-m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)$.

The mass terms for the quark fields may, correspondingly, be in general expressed as

$$\mathcal{L}_M \;=\; -\bar{\mathbf{u}}_R^{\circ T}\, \mathbf{m}_{\rm u}\, \mathbf{u}_L^\circ - \bar{\mathbf{d}}_R^{\circ T}\, \mathbf{m}_{\rm d}\, \mathbf{d}_L^\circ \;+\; \text{h.c.} \;. \tag{1.2}$$

The up-type and down-type quark fields are contained in the column vectors $\mathbf{u}^\circ$ and $\mathbf{d}^\circ$, respectively (with dimension given by the number of families); "h.c." stands for the hermitian conjugate counterparts of the terms shown. The mass matrices $\mathbf{m}_{\rm u}$ and $\mathbf{m}_{\rm d}$ are, for generality, complex-valued and non-diagonal. The Lagrangian (1.2) should be regarded in a more general context as effectively obtained from a fundamental theory of masses.

Fermionic mass terms, as in (1.1) and (1.2), are not invariant under gauge transformations, where the right- and left-handed fields may belong to different representations of the gauge group. They cannot thus appear in the bare Lagrangian of such gauge invariant theories, and are instead effectively generated in general through spontaneous symmetry breaking. In the SM electroweak theory, the $SU(2) \otimes U(1)$ local symmetry spontaneous breakdown is achieved by introducing a scalar field with non-vanishing vacuum-expectation-value (*vev*). The scalar field is realized by the Higgs boson, and the mentioned mechanism thereby gives mass to gauge bosons, charged leptons, and quarks. Specifically, the fermion masses arise from Yukawa terms, coupling the Higgs and the fermion fields,

$$-Y \bar{\psi} \psi \phi \quad \xrightarrow[\text{symmetry breaking}]{\text{spontaneous}} \quad -Y \bar{\psi} \psi \,(v + \phi') \;, \tag{1.3}$$

where Y is the Yukawa coupling with the Higgs field $\phi(x)$, and v and $\phi'(x)$ are respectively its *vev* and excitation above background. The quark mass terms in (1.2) are then given by

$$\mathbf{m}_{\rm f} \;=\; v \mathbf{Y}_{\rm f} \qquad ({\rm f} = {\rm u}, {\rm d}) \;, \tag{1.4}$$

with $(\mathbf{Y}_{\rm f})_{ij}$ representing the associated Yukawa couplings. The origin of these couplings and their apparent hierarchy are not understood. The Higgs scalar field in the SM consists of a single doublet of the gauge group; and ϕ above in (1.3) corresponds more precisely to the scalar component not eaten up by gauge symmetry. The quark mass matrices however could potentially arise from a more elaborate Higgs sector, leaving behind a more complicated set of quark-Higgs couplings, with possibly multiple *vev*'s. The general algebraic construction (of the CKM ansatz) that follows would hold equally well in such scenarios.

The mass matrices, $\mathbf{m}_{\rm u}$ and $\mathbf{m}_{\rm d}$, may be diagonalized with the help of two unitary matrices each – L_u, R_u and L_d, R_d – as

$$L_u \mathbf{m}_{\rm u} R_u^\dagger = \hat{\mathbf{m}}_{\rm u} \;, \qquad L_d \mathbf{m}_{\rm d} R_d^\dagger = \hat{\mathbf{m}}_{\rm d} \;, \tag{1.5}$$

where $\hat{\mathbf{m}}_{\rm u}$ and $\hat{\mathbf{m}}_{\rm d}$ are diagonal, with real, positive eigenvalues, corresponding to the individual quark masses,

$$\hat{\mathbf{m}}_{\rm u(d)} = \text{diag}\left(m_{u(d)}, m_{c(s)}, m_{t(b)}\right) \ . \tag{1.6}$$

The quark fields in (1.2) are accordingly translated into their mass eigenstates, denoted $\mathbf{u}$ and $\mathbf{d}$, through the unitary transformations specified by (1.5),

$$\mathbf{u}_L = L_u\, \mathbf{u}_L^\circ\,, \quad \mathbf{u}_R = R_u\, \mathbf{u}_R^\circ\,, \quad \mathbf{d}_L = L_d\, \mathbf{d}_L^\circ\,, \quad \mathbf{d}_R = R_d\, \mathbf{d}_R^\circ\,. \tag{1.7}$$

Although the underlying (electroweak) theory is first written down in terms of the gauge basis states, actual calculations which confront theory with experiment are performed using the mass basis states.

Flavor changing interactions

The SM weak interactions induce flavor changing transitions through charged currents coupling to the $W^\pm$ bosons. For the quark fields, represented by the original field arrays $\mathbf{u}^\circ$ and $\mathbf{d}^\circ$, these flavor changing interactions are described by

$$\mathcal{L}_W \;=\; \frac{g}{\sqrt{2}}\, \bar{\mathbf{u}}_L^{\circ T}\, \gamma^\mu\, \bar{\mathbf{d}}_L^\circ\, W_\mu^+ \;+\; \text{h.c.}\ , \tag{1.8}$$

where g denotes the $SU(2)$ gauge coupling constant. Once the quark fields are expressed in the mass eigenstates basis, via the transformations (1.7), the above expression becomes

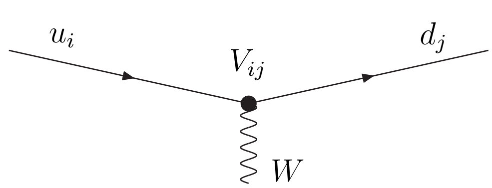
$$\mathcal{L}_W \;=\; \frac{g}{2\sqrt{2}}\, \bar{\mathbf{u}}^T\, \gamma^\mu (1-\gamma^5)\, \mathbf{V}\, \mathbf{d}\, W_\mu^+ \;+\; \text{h.c.}\ . \tag{1.9}$$

with

$$\mathbf{V} \equiv L_u L_d^\dagger\ . \tag{1.10}$$

The unitary quark-mixing matrix $\mathbf{V}$ is referred to as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3, 4].

The charged current interactions of (1.9) produce transitions between up- and down-type quarks, which occur across the various quark families, with amplitudes determined by the corresponding CKM matrix elements. This may be represented by the diagram



.

The b quark decays to a c or u quark and a virtual W boson, with couplings given by the V_{cb} or V_{ub} matrix elements, respectively. The dominant, tree-level diagrams

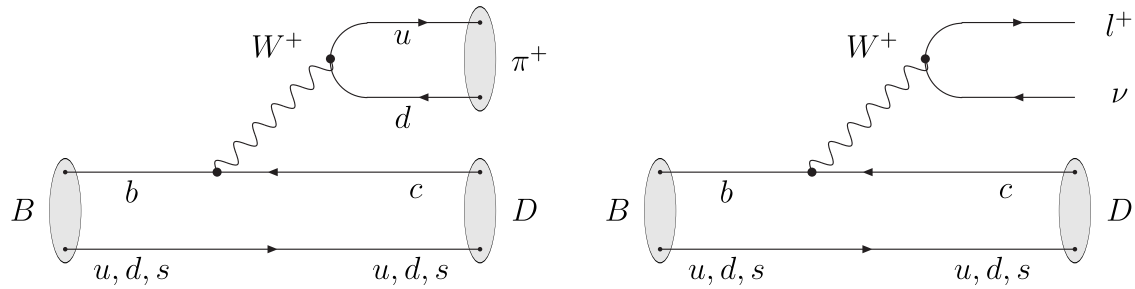


Figure 1-1: Tree-level Feynman diagrams describing B meson decays: $B \to D\pi^+$ (*hadronic*) and $B \to Dl^+\nu$ (*semileptonic*).

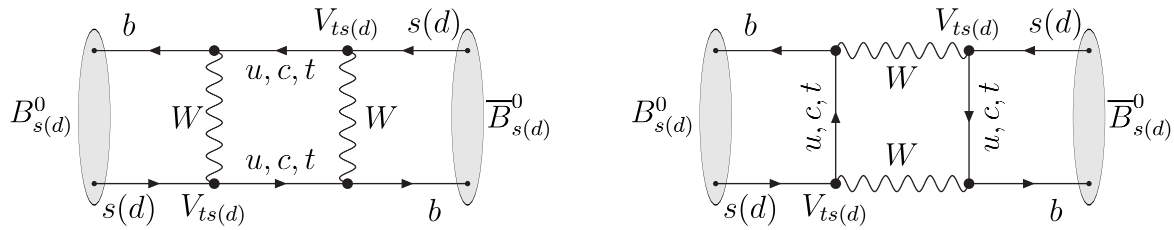


Figure 1-2: Leading Feynman diagrams contributing to B_s and B^0 flavor oscillations.

describing the decays of B mesons are correspondingly represented in Figure 1-1. In case the W boson exchange occurs with a $(u_i, \bar{d}_j)$ quark pair or a (l^+, ν) lepton pair, the B decay is in general also named *hadronic* or *semileptonic*, respectively.

The weak charged current interactions of (1.9) are further capable, at loop level, of inducing B to $\bar{B}$ transitions in the neutral B meson systems. These transitions are represented in Figure 1-2, and will be addressed below.

1.1.2 CKM ansatz and the unitarity triangle

We have seen above that the quark mass eigenstates, $\mathbf{u}$ and $\mathbf{d}$, differ in general from the quark field combinations which belong to the gauge group representations of the underlying theory. The states which participate in the transitions of the charged weak current are linear combinations of the mass eigenstates. By convention, the mixing is assigned to the down-type quarks, being expressed by the CKM matrix (1.10) as

$$\mathbf{d}' = \mathbf{V}\,\mathbf{d}\,, \tag{1.11}$$

or, in an expanded form,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.12}$$

The SM does not predict the content of the CKM matrix $\mathbf{V}$, rather, its matrix elements must be phenomenologically extracted from data.

We note in passing that the experimentally constrained observables, the CKM matrix and the mass matrices eigenvalues, are insufficient for reconstructing the full mass matrices (1.5). The latter requires knowledge of both left and right handed rotation matrices, $L_{u(d)}$ and $R_{u(d)}$. This is the starting difficulty one encounters in trying to construct a theory of masses. Such models proceed through theoretically motivated hypotheses which ought to be tested via predicted relations among the observables.

CKM matrix parameterization

The quark mixing matrix $\mathbf{V}$, being the product of unitary matrices (1.10), is itself unitary. A general 3×3 unitary matrix has 9 parameters. Among these, three are rotation angles; this is the number of parameters of a $O(3)$ rotation, *e.g.* the Euler angles. The remaining 6 parameters are phases. Some of these can be absorbed by making phase rotations of the quark fields; noticing that an overall phase, common to all six quark fields is redundant, we deduce that only 5 of these phases can be removed. This may be written explicitly as

$$\mathbf{V} \quad \mapsto \quad P(\sigma_1, \sigma_2, \sigma_3)\, \mathbf{V}\, P(0, \sigma_4, \sigma_5) \;, \tag{1.13}$$

with $P(\sigma_i, \sigma_j, \sigma_k) \equiv \mathrm{diag}(e^{i\sigma_i}, e^{i\sigma_j}, e^{i\sigma_k})$. Once the phases σ_i are absorbed into the re-definition of the quark fields, $\mathbf{V}$ finally contains 4 real parameters: 3 angles and 1 phase.

There are many ways to parameterize the CKM matrix in terms of four independent parameters. The numerical value of the physical phase, for instance, is not unique due to the arbitrariness in the choice of the unphysical phases, thus varying with the adopted parameterization. The standard CKM parameterization [5] is constructed as the product of three rotation matrices R_{ij} characterized by the Euler angles θ_{12}, θ_{23}, θ_{13} along with a phase δ,

$$\begin{aligned}
\mathbf{V} &= R_{23} P(-\delta,0,0) R_{13} P(\delta,0,0) R_{12} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} c_{13}c_{12} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - s_{23}s_{13}c_{12}e^{i\delta} & c_{23}c_{12} - s_{23}s_{13}s_{12}e^{i\delta} & s_{23}c_{13} \\ s_{23}s_{12} - s_{13}c_{23}c_{12}e^{i\delta} & -s_{23}c_{12} - s_{13}s_{12}c_{23}e^{i\delta} & c_{23}c_{13} \end{pmatrix} .
\end{aligned} \tag{1.14}$$

The mixing angles θ_{ij} (for generation labels $i, j = 1, 2, 3$) can all be chosen to lie in the first quadrant, such that $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$ are positive quantities, while the phase δ may vary between 0 and 2π.

Following the observation of a hierarchy between the mixing angles, $s_{13} \ll s_{23} \ll s_{12} \ll 1$, an expansion was proposed by Wolfenstein [6] in terms of powers of $\lambda = s_{12}$

($i.e.$ the sine of θ_{12}, which is identified with the Cabibbo angle), along with parameters A, ρ and η intended to be of order unity. We adopt here a generalization [7] where the real parameters A, λ, ρ and η are *defined* (to all orders) through the following relations,

$$s_{12} = \lambda\,, \quad s_{23} = A\lambda^2\,, \quad s_{13}e^{i\delta_{13}} = A\lambda^3(\rho + i\eta)\,. \tag{1.15}$$

Making this change of variables in the standard parameterization (1.14), we obtain the CKM matrix as a function of (A, λ, ρ, η) which satisfies unitarity exactly. We may next perform an expansion in powers of λ, to obtain

$$\mathbf{V} \;=\; \begin{pmatrix} 1-\frac{\lambda^2}{2}-\frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + A^2\lambda^5\left(\frac{1}{2}-\rho-i\eta\right) & 1-\frac{\lambda^2}{2}-\frac{1}{8}\lambda^4(1+4A^2) & A\lambda^2 \\ A\lambda^3\left(1-(\rho+i\eta)(1-\frac{\lambda^2}{2})\right) & -A\lambda^2 + A\lambda^4\left(\frac{1}{2}-\rho-i\eta\right) & 1-\frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6)\,. \tag{1.16}$$

This explicitly illustrates the observed hierarchy in the magnitude of the CKM elements: the on-diagonal elements are large, of order unity, while the off-diagonal entries are smaller, indicating the relative suppression of cross-generation transitions. For example, $b \to c$ transitions are suppressed by λ^2 and $b \to u$ transitions are suppressed by λ^3. Truncated expansions in λ are instructive, as for such illustrative purposes; however we find otherwise no need here to make such approximations.

Physically meaningful quantities are, of course, independent of the adopted parameterization. Phase-convention invariant quantities include the moduli, $|V_{ij}|$, of the matrix elements. The parameters $\lambda = |V_{us}|/\sqrt{|V_{ud}|^2 + |V_{us}|^2}$ and $A = |V_{cb}/V_{us}|/\lambda$ are thus phase invariant, while ρ and η, with $\rho + i\eta = V_{ub}^*/\left(A\lambda^3\right)$, are not.

CP violation

Irreducible phases in the quark mixing matrix, which imply the presence of weak complex couplings, lead to violation of the charge-parity CP symmetry [9, 8]. If the Hamiltonian is complex then the theory is not invariant under time reversal, $THT^{-1} \neq H$, due to the complex conjugation produced by the T operation. The combined operation CPT is a basic symmetry of any (local Poincaré invariant) quantum field theory; it follows that non-conservation of T implies that CP must be violated as well.

As we have seen, in the case of three quark generations there is one such single phase. The generalization to the case of n generations contains $(n-1)(n-2)/2$ phases along with $n(n-1)/2$ angles. For less than 3 generations no physical phase would remain, which makes the third quark generation a requirement for the existence of CP violation in the quark sector. This also means that the phase of the CKM matrix can have physical consequences only in processes involving all three generations, which typically corresponds to processes containing weak interaction loop contributions.

A phase-convention independent measure of CP violation in the SM quark sector is given by

$$\mathrm{I}m\ \det\left([\mathbf{m}_{\mathrm{u}}\mathbf{m}_{\mathrm{u}}^{\dagger},\mathbf{m}_{\mathrm{d}}\mathbf{m}_{\mathrm{d}}^{\dagger}]\right) \;=\; 2\,J\,(m_t^2-m_c^2)(m_t^2-m_u^2)(m_c^2-m_u^2) \times\ (m_b^2-m_s^2)(m_b^2-m_d^2)(m_s^2-m_d^2)\ . \tag{1.17}$$

The Jarlskog invariant [10], J, contains the dependence on the CKM elements, and has the general representation

$$\mathrm{I}m\ \left(V_{ij}V_{kl}V_{il}^*V_{kj}^*\right) = J\sum_{m,n=1}^{3}\epsilon_{ikm}\epsilon_{jln} \tag{1.18}$$

which in terms of the standard (1.14) and generalized Wolfenstein (1.16) parameterizations is given by

$$J = c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13}\sin\delta = A^2\lambda^6\eta\left(1-\lambda^2/2\right)+\mathcal{O}(\lambda^{10})\ . \tag{1.19}$$

The requirements for CP violation include therefore the non-degeneracy of the up-type and down-type quark masses, and the non-vanishing of the CKM phase, hence $J\neq 0$. Both conditions are experimentally verified.

Unitarity triangle

The unitarity of the CKM matrix leads to various relations among its elements, which may be expressed in terms of geometric representations. The unitarity conditions are summarized as

$$\sum_i V_{ij}V_{ik}^* \;=\; \delta_{jk} \;=\; \sum_i V_{ji}V_{ki}^*\ . \tag{1.20}$$

For example, from the orthogonality of the first and third columns, one has

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\ . \tag{1.21}$$

This relation requires the sum of three complex quantities to vanish, and can thus be represented in the complex plane as a triangle [11]. Phase transformations lead to rotations of the triangle in the complex plane. The angles and the sides of the triangle, which are given by the moduli of the matrix elements, are phase-convention independent and constitute physical observables. Overall, (1.20) defines six triangles corresponding to orthogonality relations. All such triangles have a common area, identical to $|J|/2$.

The geometric representation of the unitarity relation (1.21) is named the *unitarity triangle* (UT). The UT is customarily re-scaled such that one of the sides has unit length and is aligned with the real axis. This is represented in Figure 1-3. The re-scaling factor is chosen to be the inverse of $V_{cd}V_{cb}^*$, emphasizing the less well determined parameters (ρ,η). The apex of the UT is given by the following phase-convention

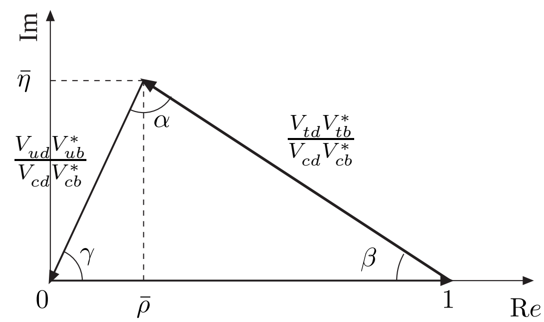


Figure 1-3: The unitarity triangle in the complex plane.

independent definition

$$\bar{\rho} + i\bar{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,. \tag{1.22}$$

Expressed in terms of the parameterization (1.16), one finds

$$\rho + i\eta = \frac{\sqrt{1 - A^2\lambda^4}\,(\bar{\rho} + i\bar{\eta})}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4\,(\bar{\rho} + i\bar{\eta})]}\,, \tag{1.23}$$

which is valid to all orders in λ; to leading order, $\bar{\rho}$ and $\bar{\eta}$ are given by

$$\bar{\rho} = \rho\left(1 - \lambda^2/2\right) + \mathcal{O}\left(\lambda^4\right)\,, \qquad \bar{\eta} = \eta\left(1 - \lambda^2/2\right) + \mathcal{O}\left(\lambda^4\right)\,. \tag{1.24}$$

The UT has the characteristic that all sides are of comparable size, of order λ^3 in the parameterization (1.16). This reveals convenient from an experimental point of view, if all angles and sides are to be precisely constrained. The sides of the UT (besides that along the real axis which is normalized to unity) read to all orders

$$R_u \equiv \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\bar{\rho}^2 + \bar{\eta}^2}\,, \tag{1.25}$$

$$R_t \equiv \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\bar{\rho})^2 + \bar{\eta}^2}\,. \tag{1.26}$$

The angles are defined as

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right)\,, \qquad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)\,, \qquad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)\,, \tag{1.27}$$

and may be expressed, again to all orders in λ, as

$$\begin{aligned} \cos\gamma &= \bar{\rho}/R_u\,, \quad \sin\gamma = \bar{\eta}/R_u\,, \\ \cos\beta &= (1-\bar{\rho})/R_t\,, \quad \sin\beta = \bar{\eta}/R_t\,, \\ \alpha &= \pi - \beta - \gamma\,. \end{aligned} \tag{1.28}$$

The CKM phase in the parameterization (1.14) reads $\delta = \gamma + A^2\lambda^4\eta + \mathcal{O}(\lambda^6)$.

The sides and the angles of the UT may all be experimentally determined. By doing so, one can over-constrain the shape of the triangle, testing the unitarity of the CKM matrix. As we shall see next, the measurement of flavor oscillations in both neutral B meson systems will provide a stringent constraint in one side, R_t, of the UT.

We also like to mention that in the future the copious production of B_s mesons — at the Tevatron, the LHC and $\Upsilon(5S)$ resonance machines — will allow further exploration of the unitarity triangle ($\mathrm{UT_s}$) which results from the orthogonality of the second and third columns of the CKM matrix. In this other triangle not all sides are of similar sizes. The smallest side corresponds to $V_{us}V_{ub}^*/V_{cs}V_{cb}^*$, and has a relative magnitude of the order of λ^2; the small angle opposing this side is given by

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) . \tag{1.29}$$

The determination of this phase, accessible for example through the study of the CP asymmetry in B_s decays into final CP eigenstates such as $J/\psi\phi$, is complementary to the Δm_s measurement which itself corresponds to the modulo of the mixing process amplitude (A.20).

1.1.3 Model constraining

Provided that the CKM matrix elements govern the flavor changing processes, one may be able to use experimental inputs such as measurements of decay rates, asymmetries, and mixing to constrain the matrix parameters. In particular, the unitarity of the CKM matrix is in this way tested, validating the three-generation SM, while any discrepancies found will provide insight on sources of new physics.

The connection between the properties of the b-flavored hadrons, which are what is experimentally detected, and the underlying quark dynamics is achieved employing effective field theory techniques. These proceed by separating the different energy scales involved, such that the high scale phenomena associated to the flavor structure may be treated separately from the complications of non-perturbative hadronic physics. The interactions at higher scales, basically, give rise to local operators at lower scales. An effective weak Hamiltonian describing the flavor changing processes is obtained through an operator product expansion [2]. The hadronic matrix elements are tackled by techniques such as heavy quark effective theory and lattice QCD. The most notable application of lattice QCD, in the context of CKM parameter determination, is in effect to mixing related quantities; uncertainties in those quantities are currently dominant in CKM fits.

The leading, lowest order diagrams that contribute to the $B_q^0\bar{B}_q^0$ ($q = d, s$) transitions are shown in Figure 1-2. They correspond to four-vertex "box" graphs containing two W and two up-type quark internal lines. The dominant fermion contribution in the loop is provided by the t quark; the contributions from the lighter quarks are suppressed by $(m_{u,c}/m_W)^2$. The effective coupling therefore becomes proportional to $(V_{tb}^*V_{td})^2$ and $(V_{tb}^*V_{ts})^2$, respectively, for the B^0 and B_s systems. Correspondingly, the

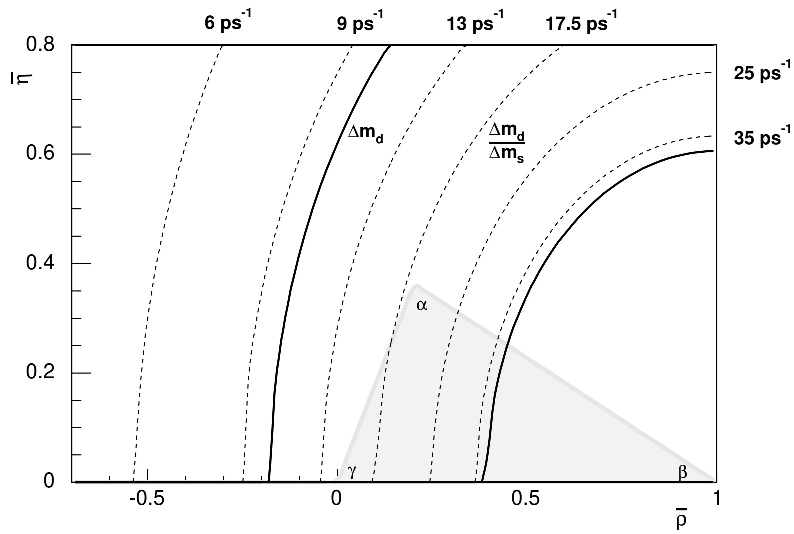


Figure 1-4: Constraints on the apex of the unitarity triangle from B mixing measurements; the continuous lines indicate the Δm_d allowed region (95% CL), while the dashed lines project a set of Δm_s values.

following relation may be written for the oscillation frequencies,

$$\Delta m_q \quad = \quad C_q \left| V_{tb}^* V_{tq} \right|^2 \ , \qquad (q = d, s) \ . \tag{1.30}$$

The coefficients C_q are evaluated in the framework of the effective theory, the full expression being given in (9.15); the derivation is outlined in Appendix A. In principle, and further noting that $V_{tb} \approx 1$, the relations (1.30) may be used to constrain the individual CKM matrix elements $|V_{td}|$ and $|V_{ts}|$ from the measurements of Δm_d and Δm_s. The power of these constraints nevertheless is hampered by the systematic uncertainties which characterize the lattice calculation of the hadronic matrix elements, which are contained in C_q. Several of these uncertainties cancel out once the ratio of the oscillation frequencies of the two systems is formed,

$$\frac{\Delta m_d}{\Delta m_s} \quad = \quad \frac{C_d}{C_s} \frac{|V_{td}|^2}{|V_{ts}|^2} \quad = \quad \frac{m_{B^0}}{m_{B_s}} \xi_\Delta^{-2} \frac{|V_{td}|^2}{|V_{ts}|^2} \ . \tag{1.31}$$

Here the parameter ξ_Δ quantifies the SU(3) flavor symmetry breaking corrections to the matrix elements, and can be calculated more accurately in lattice QCD than the matrix elements themselves. (Currently, while uncertainties on the computations of C_q are of the order of 15%, a precision of better than 4% is attained in the computation of ξ_Δ [12].)

The constraints on the CKM matrix elements are translated into probability regions in the space of its parameters (1.16), and in particular may be used to constrict the apex of the UT. For the B^0 system, (1.30) approximately describes a circle around

$(1,0)$ in the $(\bar{\rho},\bar{\eta})$ plane,

$$\Delta m_d \;=\; C_d \lambda^6 A^2 [(1-\bar{\rho})^2 + \bar{\eta}^2] \;, \tag{1.32}$$

to which a distortion appears at $\mathcal{O}(\lambda^{10})$. Despite Δm_d being currently known with very good precision, the aforementioned theoretical uncertainties limit its effectiveness as a constraint. This is represented graphically in Figure 1-4, which shows the wide allowed region representing the Δm_d constraint. Forming the ratio of the oscillation frequencies as in (1.31) a more powerful constraint may be obtained, which is expressed to leading order as

$$\Delta m_s \;=\; \Delta m_d \, \xi_\Delta^2 \, \frac{m_{B_s}}{m_{B^0}} \left(\frac{1-\frac{1}{2}\lambda^2}{\lambda} \right)^2 \frac{1}{(1-\bar{\rho})^2 + \bar{\eta}^2} \;. \tag{1.33}$$

Note that while Δm_s has only a weak direct dependence on the parameters $(\bar{\rho},\bar{\eta})$, neglected in (1.33), it effectively provides a determination of the non-perturbative parameters contained in C_q which enter (1.32). The UT's side R_t (1.26), which corresponds to the circle's radius, becomes proportional to the inverse of Δm_s. A set of Δm_s values is projected onto the unitarity plane in Figure 1-4. Lower exclusion limits for Δm_s can carry already very useful information, as they impose upper bounds on the magnitude of R_t. A full double-sided measurement of Δm_s will result in a particularly stringent constraint. The power of the latter will also benefit from improvements in the (lattice) calculation of the relevant matrix elements. Conversely, the measurement of Δm_s will effectively provide an "experimental" determination of those hadronic quantities, within the SM framework, which can then be compared against the theoretical predictions.

Other constraints

Several other measurements are used, along with neutral B meson mixing, to over-constrain the UT in a global CKM fit. We briefly mention next some of the most relevant.

The length of the side R_u (1.25) of the UT is constrained from the ratio $|V_{ub}/V_{cb}|$. This is determined from the relative rates of $b \to ul\bar{\nu}$ and $b \to cl\bar{\nu}$ decays, and corresponds to a circle centered at the origin in the $(\bar{\rho},\bar{\eta})$ plane,

$$\bar{\rho}^2 + \bar{\eta}^2 = C_1 \;. \tag{1.34}$$

The quantities C_i contain both results of experimental measurements and related theoretical computations. From the neutral kaon system, the measurement of indirect CP violation, which is quantified by the asymmetry parameter $|\epsilon_k|$, becomes also useful as the relevant matrix elements can be obtained with accountable, moderate systematic uncertainties. This translates roughly into a hyperbolical constraint of the form

$$\bar{\eta}[1 + C_3(1-\bar{\rho})] = C_2 \;. \tag{1.35}$$

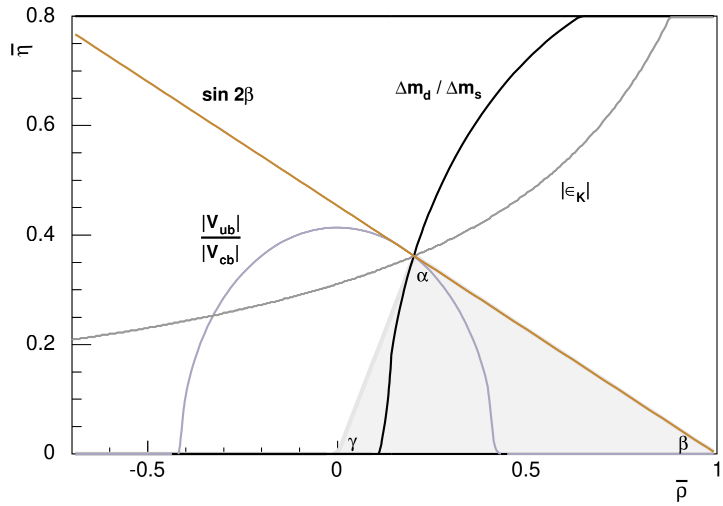


Figure 1-5: Constraints on the unitarity plane, where the compatibility among the observables has been enforced.

The UT angles are all accessible from the B sector, albeit with different sensitivity and purity. The measurements of β and γ are performed through B decays in charmonium and open-charm, respectively, and are theoretically clean, whereas the measurement of α in charmless B decays relies on theoretical assumptions. For the angle β, the leading experimental observable is $\sin 2\beta$, measured from time-dependent CP violation parameters in $b \to c\bar{c}s$ decays. The measurement of α and γ involve interference with transitions governed by the small CKM matrix element V_{ub}, and require relatively larger data samples. The form of the UT angles' constraints is encoded in (1.28).

Relative to $\bar{\rho}$ and $\bar{\eta}$, the parameters λ and A are currently measured with a considerably higher precision. The former, λ, is obtained from the magnitude of V_{us}, which is traditionally extracted from semileptonic kaon decays. The latter parameter, A, is determined from $|V_{cb}|$, being most accurately obtained from exclusive and inclusive semileptonic B decays to charm.

A graphical representation of the most relevant constraints, without uncertainties, is exhibited in Figure 1-5; some indicative values for the involved quantities are used while the compatibility between the constraints has been enforced for illustration.

Beyond the standard model

New physics (NP) may potentially bring sizable alterations to the standard flavor description. While additional tree-level contributions are in general anticipated to be suppressed in view of existing experimental constraints, large NP contributions may be present in loop-mediated processes where new interactions may play a significant role. This holds in particular for B meson mixing [13].

The presence of NP may manifest itself in several ways, including incompatibilities of different measurements related to a same SM quantity, significant enhancements of decay rates and asymmetries that are expected to vanish or be very small within the SM, inconsistencies among the values of the angles and/or the sides of the UT, or mixing may be found to differ significantly from SM predictions.

General NP models may introduce a large number of new parameters: flavor changing couplings, CP violating phases, short distance coefficients and matrix elements of new local operators. Nevertheless, a mixing process is described by a single amplitude, and can be parameterized, without loss of generality as in (A.39), in terms of two parameters which quantify the difference of the complex amplitude with respect to that of the SM. In the presence of NP, the neutral B_q^0 meson oscillation frequency may be parameterized accordingly as

$$\Delta m_q \;=\; \Delta m_q^{\rm SM}\,|1 + h_q\, e^{2i\sigma_q}|\;, \tag{1.36}$$

where $\Delta m_q^{\rm SM}$ is the SM contribution, and h_q and σ_q denote the relative magnitude and phase of the NP contribution. Inconsistencies with the SM expectations may then be quantified and represented in the $(h_q,\, \sigma_q)$ parameter space.

In the case of the B^0 system, the Δm_d and $\sin 2\beta$ constraints may be used to determine h_d and σ_d. The magnitude h_s and the phase σ_s in the B_s system may be correspondingly constrained from measurements of Δm_s and of the angle β_s (1.29). The observable $\sin 2\beta_s$ can be determined from the time-dependent analysis of the CP asymmetry in B_s decays such as $B_s \to J/\psi\,\phi$. The effectiveness of such a general, model independent approach will depend on the precision of the experimental measurements, including Δm_s, and on the accuracy of the SM theoretical prediction. The latter has currently sizable uncertainties arising from imprecisely determined CKM factors and hadronic matrix elements.

The ratio of oscillation frequencies $\Delta m_d/\Delta m_s$, as discussed above, is expected to provide a precise determination of the CKM ratio $|V_{td}/V_{ts}|$ within the SM. Remarkably, this remains true in many NP scenarios. In such classes of models, the virtual exchange of new particles in the box diagrams (Figure 1-2) induce modifications to the coefficients C_q in (1.30), namely of the type of (A.37), which cancel in the ratio. This is the case of various classes of NP scenarios, including for instance two-Higgs doublet models and minimal supersymmetry with flavor conservation.

There are large classes of models in which the biggest effects of NP occur in transitions involving the second and the third generations. These leave the flavor-changing transitions between the other families unaffected, thus respecting, in particular, the bounds from kaon physics. The B_s sector is clearly well suited to test such scenarios. The latter comprise certain SUSY GUT models, for example, which predict an enhancement of Δm_s compared to its SM prediction. Another popular scenario, with large effects in flavor physics, involves Higgs mediated FCNCs in the large $\tan\beta_{\rm SUSY}$ region ($\tan\beta_{\rm SUSY}$ is the ratio of the *vevs* of the two Higgs doublets). Here suppression of Δm_s proportional to $\tan^4\beta_{\rm SUSY}$ is predicted [14]. Models with an extra $U(1)'$ gauge boson Z' can induce FCNCs at tree-level if its coupling to physical fermions is non-diagonal. Such Z' bosons often occur, for instance, in the context of GUTs, superstring theories and theories with large extra dimensions.

Some NP scenarios retain the CKM structure. In the so-called minimal flavor violation (MFV) models, the dynamics of flavor violation is completely determined by the structure of the SM Yukawa couplings. In this case, all FCNC and CP violating phenomena are expressed in terms of the CKM matrix and the dominating top quark Yukawa couplings. Several MFV implementations exist in different contexts, including two-Higgs doublet models, supersymmetry, extra dimensions, *etc.* Other models allow noticeably larger modifications. The structure of the CKM matrix is changed, for example, with the addition of a fourth generation, extra singlet quarks, or in Left-Right symmetric models. Sizable Yukawa couplings to the light fermions are favored for instance in scenarios involving leptoquarks and Higgs models with flavor changing couplings.

In general, B_s mixing may receive significant, either positive or negative, new physics contributions. Depending on the magnitude of those contributions, and the precision of relevant theoretical computations, various models may be ruled out, or more or less severely constrained, by the experimental determination of Δm_s along with further and improved relevant measurements in the flavor sector.

1.2 Mixing history and status

Particle–antiparticle oscillations have been established in various neutral meson systems. The oscillation frequencies are extremely well measured in the K^0 and B^0 systems, while for the B_s the oscillations are very rapid and their frequency has yet to be observed.

Oscillations were first proposed for the kaon system in 1955 [15]. This lead to predicting the existence of a long-lived strange particle, the K_L^0, which was subsequently observed in 1956 [16]. The first CKM element, the Cabibbo quark mixing angle identified with θ_{12} in (1.14), was introduced in 1963 [3] to explain the small decay rates for particles carrying strangeness. This further lead in 1970 to the introduction of the unitarity quark mixing ansatz used by GIM [17] to postulate the existence of a fourth quark (charm). In 1973 Kobayashi and Maskawa [4] added to the model a third generation of quarks, motivated by the earlier discovery of CP violation in 1964 in K^0 decays [18]. The bottom and top quarks were subsequently discovered in 1977 [19] and 1995 [20], respectively, at Fermilab.

Evidence of neutral B meson mixing was first reported in 1987 by UA1 [21]. This consisted of a time-integrated analysis, based on the measurement of the ratio of like-sign muon pairs to unlike-sign muon pairs. The first observation of B^0 oscillations was published by ARGUS [22] in that same year, and included the unambiguous identification of a B^0B^0 pair. Confirmation of those results was provided by CLEO [23] in 1989. Various other time-integrated mixing analyses followed, at PEP, PETRA, and LEP experiments, as well as at CDF. Time-dependent measurements of B^0 mixing have since then yielded a precise determination of the oscillation frequency Δm_d. A compilation is presented in Figure 1-6. The world average value for Δm_d is 0.507 ± 0.005 ps^{-1} [24], being dominated by BaBar and Belle measurements [25].

That B_s mesons undergo oscillations was demonstrated early on. The above mentioned measurements by ARGUS and CLEO, which operate at e^+e^- colliders tuned to the $\Upsilon(4S)$ resonance, provided the time integrated probability that a B^0 oscillates. When these were compared with inclusive measurements performed at LEP, for instance, where in addition to B^0 also B_s mesons are produced, a significant contribution to the mixing signal from the latter was deduced. Subsequent time-dependent measurements, performed at LEP and SLD operating at the Z-pole, and at CDF in the previous data taking stage of the Tevatron operating at an energy $\sqrt{s} = 1.8$ TeV, established lower exclusion bounds on the oscillation frequency of the B_s system.

For the purpose of combining exclusion regions from the various experiments in a convenient and consistent fashion, the results on B_s oscillations are presented in terms of amplitude measurements for a spectrum of probe frequencies. Such a procedure [31] essentially entails searching for a peak in the power spectrum of the data as a function of frequency. A list of amplitude measurements is shown in Figure 1-7 for the specified probe frequency. The combined results are presented in Figure 1-8, giving a world average exclusion limit of $\Delta m_s > 14.4$ ps^{-1} at 95% C.L. [24], including contributions from ALEPH, DELPHI, OPAL, SLD and CDF. The corresponding combined sensitivity, given by the measured amplitude uncertainties, is 18.2 ps^{-1}.

The aforementioned CDF contribution was achieved during the former Tevatron data taking period. It consisted of a search for flavor oscillations using inclusively reconstructed B_s mesons in $\phi l X$ final states. An exclusion limit was placed at 5.8 ps^{-1} at 95% C.L. [29]. We emphasize that the results and averages presented in this introductory section do not include results obtained during the ongoing Tevatron data taking period (called Run II), in general, and in the presently documented analysis in particular. We also point out that througout this document, unless otherwise specified, Δm will be treated primarily as an oscillation frequency, as this is what is directly extracted from the data, and it is thus expressed in units of inverse-(pico)second. In Table 1.1 we also summarize the values of Δm in units of mass, for the kaon and bottom neutral systems.

Currently, the Tevatron is the only collider producing significant samples of B_s mesons. The CDF and DØ collaborations stand therefore in unique positions to explore the B_s system. A first and precise determination of its oscillation frequency, or the exclusion of the SM favored region, is a stated priority of the Tevatron Run II physics program. That is indeed the subject and objective of the analysis developed in this dissertation, using the B_s data samples collected with the CDF detector. In the process of carrying out such a time-dependent analysis, the meson's lifetime is extracted, as a by-product, from the mixing data samples. Other measurements involving the B_s meson, along with other b-hadron states, are also pursued at CDF; these include, among others, mass, branching ratios, CP asymmetries, and width differences.

The B_s meson is too heavy to be produced in decays of the $\Upsilon(4S)$ resonance, and cannot thus be studied at the B factories which operate at that energy. The kinematical threshold is satisfied though by the $\Upsilon(5S)$ resonance, which may be explored. For instance, the asymmetric e^+e^- KEKB collider was recently operated at this resonance in an engineering run, which allowed the Belle experiment to take first B_s data [30]. The excitement for studying the B_s meson is expected to be transferred in the future to the LHC experiments at CERN, namely the LHCb as well as the general purpose experiments, ATLAS and CMS. Not only with it be possible to corroborate the Tevatron results, and specifically those on Δm_s, to a higher precision, but novel flavor measurements will be achievable. The exploration of the B_s system being undertaken at the Tevatron is thus expected to be complemented and deepened at the LHC.

	$K^0\bar{K}^0$	$B^0\bar{B}^0$	$B_s\bar{B}_s$
mass [MeV/c^2]	497.648 ± 0.022	5279.4 ± 0.5	5369.6 ± 2.4
lifetime [ps]	89.53 ± 0.06, 51800 ± 400	1.536 ± 0.014	1.461 ± 0.057
Δm [$(\hbar/\mathrm{c}^2)\mathrm{s}^{-1}$]	$(0.5292 \pm 0.0010) \cdot 10^{10}$	$(0.507 \pm 0.005) \cdot 10^{12}$	$> 14.4 \cdot 10^{12}$
Δm [eV/c^2]	$(3.483 \pm 0.006) \cdot 10^{-6}$	$(3.337 \pm 0.033) \cdot 10^{-4}$	$> 94.8 \cdot 10^{-4}$

Table 1.1: Parameters of the oscillating kaon and bottom meson systems.

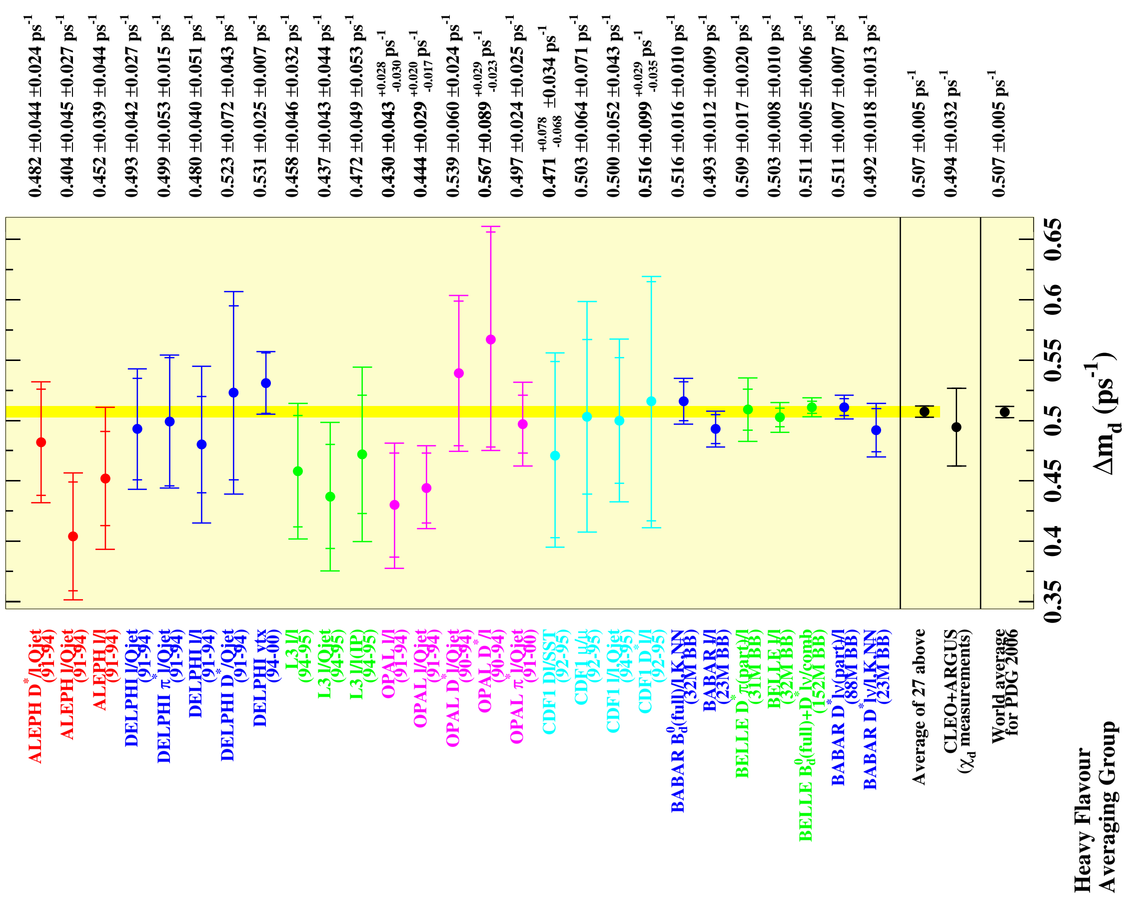


Figure 1-6: Experimental measurements of Δm_d.

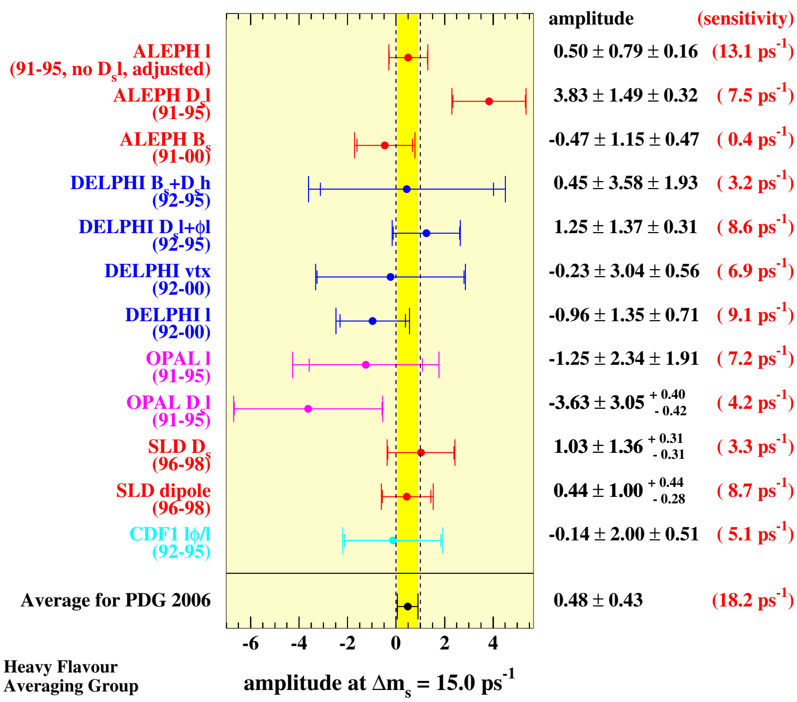


Figure 1-7: Experimental amplitudes and sensitivities at Δm_s probe frequency 15 ps^{-1}.

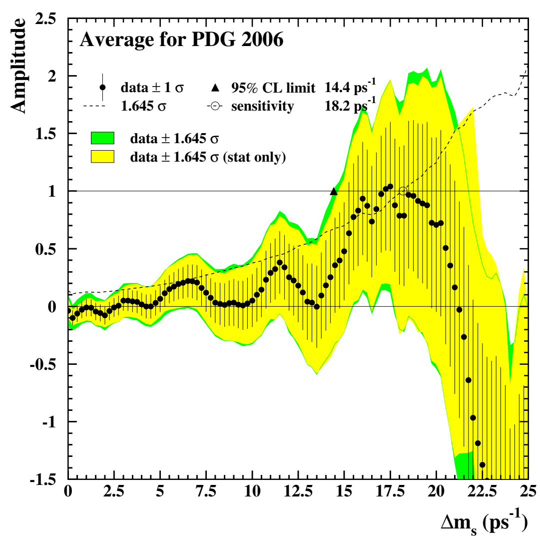


Figure 1-8: World combined results on Δm_s amplitude measurements.

1.3 Analysis synopsis and roadmap

The large expected value of the particle–antiparticle mixing frequency makes it challenging to observe and measure the time-dependence of B_s flavor oscillations. The amplitude of the oscillations is attenuated by effects resulting from background contamination of the samples, experimental resolutions on the decay proper time reconstruction, and performance limitations of the flavor tagging methods.

In the absence of such attenuation effects and further experimental complications, the probability density for an originally produced B_s meson to decay at a later time t as $\bar{B}_s$, or conversely, follows from basic quantum mechanics (A.24) and is given simply by

$$\mathcal{P}_{B_s \to \bar{B}_s} \sim \frac{1}{2\tau} e^{-t/\tau} \left[1 - \cos\left(\Delta m_s \, t\right)\right] , \tag{1.37}$$

where τ is the meson's lifetime and Δm_s is the frequency of oscillation we aim to determine.

A schematic representation of a typical B event is shown in Figure 1-9. The main ingredients required by the measurement may be summarized as follows:

- *data samples:* B events produced in the $p\bar{p}$ collisions need to be identified and collected, and the signal and background composition of the resulting data samples needs to be characterized,

- *decay time:* the proper time of the decay, t, is constructed from the flight distance of the meson in the detector, between the primary $p\bar{p}$ collision point and the meson's decay position, and the determination of the B momentum from its decay products,

- *flavor at production and decay:* it is necessary to find what the flavor (*i.e.* B or $\bar{B}$) of the meson was, both when it was produced and when it decayed, in order to determine whether it decayed as mixed (flavor of production different from flavor at the time of decay) or unmixed (flavor of production the same as flavor at the time of decay).

The above ingredients must be thoroughly characterized, with all involved quantities determined as accurately as possible and the probabilities of potential misdeterminations fully quantified.

In order to take the best advantage of the collected data samples and their characteristics a robust fitting framework is developed, which serves as the cornerstone of the data analysis. The optimal fit parameters are found using the maximum likelihood method based on the individual events input. The likelihood formulation is developed to accurately describe the characteristics of the data samples at hand, optimized for computational speed and accuracy, and implemented to efficiently accommodate the various reconstructed B decay modes employed.

The data analysis is extended to the lighter B^+ and B^0 meson species, for which larger data samples are available. This allows a determination of the properties of

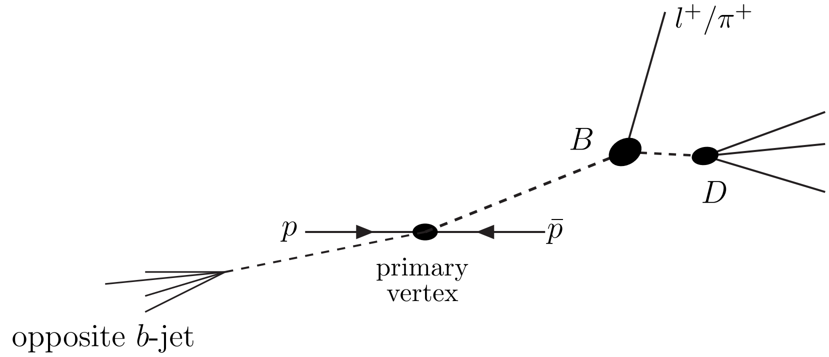


Figure 1-9: Schematic of a typical B event.

those mesons, and simultaneously serves the purposes of method validation, tool calibration, and framework testing and development.

Data samples

B mesons originate from the hadronization of b quarks produced in the proton–antiproton collisions. The b production cross-section at the Tevatron is about 0.1 mb, while the total $p\bar{p}$ cross-section, including elastic, diffractive and inelastic scattering, is of the order of 75 mb. The bottom production processes predominantly involve the creation of $b\bar{b}$ pairs. Once they have been produced, the b and $\bar{b}$ quarks both undergo fragmentation forming b hadrons, which include B^+, B^0 and B_s mesons, as well as heavier states such as B_c, Λ_b, Ξ_b, *etc.* It should be added also that a produced $b\bar{b}$ pair in a given beam-crossing is typically accompanied by additional background interactions which are omitted in Figure 1-9.

The signature which allows one to identify b hadrons from the other, more common collision products is their distinctive, long lifetimes. The B^+, B^0 and B_s mesons — which constitute our aimed signal samples and which we shall commonly denote by B — each have a lifetime τ of approximately 1.5 ps, or $c\tau \sim 450$ μm (c denoting the speed of light). When the momentum spectrum of the B mesons is folded in, the corresponding boost translates into average decay flights of the order of a few millimeters at the center of the detector, or more exactly inside the beampipe. This characteristic decay length signature is used to distinguish B signals from the myriad of background processes with their much larger cross-sections. Such characteristic secondary vertices are explored, as a matter of fact, already at the level of real-time event selection, or *trigger*ing. The daughter particles which originate from B decays usually have significant impact parameter with respect to the primary vertex; this is illustrated in Figure 1-10. The reconstructed trajectory of a long-lived (*e.g.* pion, kaon) charged particle, dubbed *track*, is said to be "detached" if the impact parameter, divided by its uncertainty, is large. CDF possesses a trigger processor which performs fast and precise track reconstruction based on information from the Silicon detectors, which is capable of looking for the presence of such detached tracks. This capability is central to the trigger strategy employed for gathering the majority of the samples used

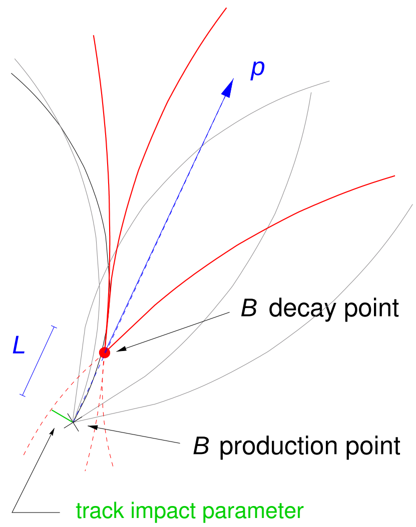


Figure 1-10: Sketch of a B decay.

in this dissertation. The detachment requisite is complemented by further topological and kinematical event information. More traditional trigger criteria based on the presence of leptons with moderate to high transverse momenta are also concurrently explored.

The B mesons are reconstructed in a variety of final states. These may be classified as corresponding to fully- or partially-reconstructed B decays. In the former case, one attempts to identify all particles participating in the decay, as for instance in $B \to D\pi$. The latter class includes for example decays of the type $B \to Dl\nu$, where one does not attempt to reconstruct the neutrino which escapes detection. The designations of *hadronic* and *semileptonic* decays will correspondingly be also employed to refer to the two classes. The leading Feynman diagrams associated with the mentioned example decays are shown in Figure 1-1. In general the semileptonic modes will have considerably larger ($\sim$10-fold) yields, while the samples' composition will be also more complex and difficult to assess. The hadronic samples despite having smaller yields provide a more complete information about the B meson, most notably its momentum.

The gathered samples of reconstructed B candidates will contain, in addition to the aimed signals, background events from various sources. The most harmful backgrounds are those which peak in the signal region (of a discriminating variable, for instance the candidate's mass) and which can more easily fake signals. The background sources and their characterization are specific to the individual decay modes. In the case of the semileptonic samples, not only is the signal and background separation complicated by the fact that the B momentum, and thus its mass, is not fully accessible, but furthermore a same lepton-D final state may in fact be achieved as product of various B mesons' decay chains which are not differentiated. The signal composition and more generally the characterization of physics-like background contributions from processes other than the nominally reconstructed decay mode benefit from Monte Carlo simulation of the relevant processes.

An overview of the CDF detector and its trigger system is provided in Chapter 2. The reconstruction of the B mesons' candidates and the characterization of the corresponding samples are explained in Chapter 3. The full collection of decay modes studied is summarized in Table 3.1.

Proper decay time

The proper decay time of a B meson candidate is calculated from the decay distance L between the production and decay points, the momentum p, and the meson's nominal mass M^B, as $t = L/\gamma\beta c = LM^B/p$. In fact, the projections of both the distance and the momentum on the plane transverse to the beamline are used. The proper time probability density function (PDF) is determined by the B meson lifetime (A.1),

$$\mathcal{P}(t) \sim \frac{1}{\tau} e^{-t/\tau} \ . \tag{1.38}$$

The above expression becomes modified once various necessary correcting and resolution effects are accounted for. These include the limited precision with which the B decay length can be measured. The incomplete momentum reconstruction in the case of the semileptonic modes induces further smearing of the PDF; this smearing effectively translates into an additional contribution to the proper time uncertainty (B.4). The detachment requirements employed for event selection and reconstruction also introduce a modification in the shape of the t-PDF which needs to be appropriately described.

The proper decay time likelihood description for the various samples and the calibration of the corresponding uncertainty are contained in Chapter 4. The lifetimes of the B mesons are measured; this also constitutes a final validation of the description of the samples before the introduction of flavor tagging information.

Flavor tagging

Determining the flavor of a B meson is equivalent to determining its b or $\bar{b}$ quark content. In general, flavor tagging exploits a correlation between the beauty flavor of the b-hadron and a *charge* in the event. The B mesons are reconstructed in flavor-specific final states. It follows that the flavor at the time of decay is given by the electric charges of the decay products. For example, a B_s decays to $D_s^-\pi^+$ or $D_s^- l^+ X$, while a $\bar{B}_s$ decays into $D_s^+\pi^-$ or $D_s^+ l^- X$.

The flavor at the time of production is more difficult to ascertain, rendering its determination a more complex task. Several different flavor tagging techniques are used for that purpose, being associated to two general strategies. The *opposite side* tagging (OST) methods exploit the fact mentioned above that bottom quarks are predominantly produced in $b\bar{b}$ pairs. These techniques thus aim at inferring the B meson production flavor by identifying the flavor of the second, accompanying b quark. Specifically, these techniques correspond to the *soft lepton* taggers, which attempt to identify electrons or muons from semileptonic b-decays, and the *jet charge* taggers which explore properties of clusters of tracks to statistically infer (the sign of) the b charge.

The *same side* tagging (SST) technique is based on correlations between particles produced in the fragmentation of b quarks to the B mesons. The SST algorithms, unlike the OST counterparts, are based on information carried by tracks found in the vicinity to the B meson being tagged. The SST performance, also unlike the OST case, is expected to vary among B species. In the case of B^+ and B^0, the mentioned charge-flavor correlations are enhanced by decays of P-wave B mesons (B^{**}). For the B_s mesons, the leading fragmentation track expected to be correlated to the B flavor is a kaon, and the algorithm's performance is thus enhanced by making use of information allowing the separation of kaons from the more abundant pions.

A given flavor tagging algorithm does not always provide a correct decision about the B meson flavor. A method's performance is conveniently quantified in terms of its tagging efficiency ϵ and dilution $\mathcal{D}$. The efficiency is the fraction of signal candidates with a flavor tag. The dilution is defined as $\mathcal{D} \equiv 1 - 2w$, where w is the mistag probability; in this way, a perfect tag would have unit dilution, while a random tag would have zero dilution. Besides the tag decision also the probability of its correctness, or equivalently the dilution, must be evaluated for the individual events. The figure of merit of a tagging algorithm is given by $\epsilon\mathcal{D}^2$.

The tagging methods are presented in detail in Chapter 5, while in Chapter 8 a novel SST method is further developed for suitably tagging the flavor of B_s mesons.

Mixing

The proper decay time PDF describing flavor oscillations may be expressed as

$$\mathcal{P}(t) \sim \frac{1}{2\tau} e^{-t/\tau} \left[1 \pm \mathcal{A}\mathcal{D} \cos\left(\Delta m\, t\right)\right] . \tag{1.39}$$

This is simply a slight elaboration of (1.37). The signs "+" and "−" refer to B candidates identified as unmixed and mixed, respectively. This identification is achieved with the flavor tagging methods, whose dilution $\mathcal{D}$ is provided to the fit, for each event, and is found to attenuate the oscillating term given by the cosine. The parameter $\mathcal{A}$ is introduced as describing the amplitude of the oscillation and is, for the moment, taken as unity.

Two mathematically equivalent methods are employed to extract from the data information about the oscillation frequency Δm. The first method involves performing a fit to the parameter Δm directly. In this way, the oscillation frequency is treated as a standard fit parameter, whose best estimate is found by maximizing the likelihood function for the data sample at hand, and whose uncertainty is determined by the variation of the likelihood function around the found maximum. This constitutes our chosen method for the measurement of the oscillation frequency in the B^0 system performed in Chapter 6. In fact, the fit is performed simultaneously for Δm and for the parameter $\mathcal{A}$, where the later constitutes in this case a calibrating factor of the dilution of the flavor tagging methods when applied to the actual mixing samples.

The second method for extracting information about the oscillation frequency is

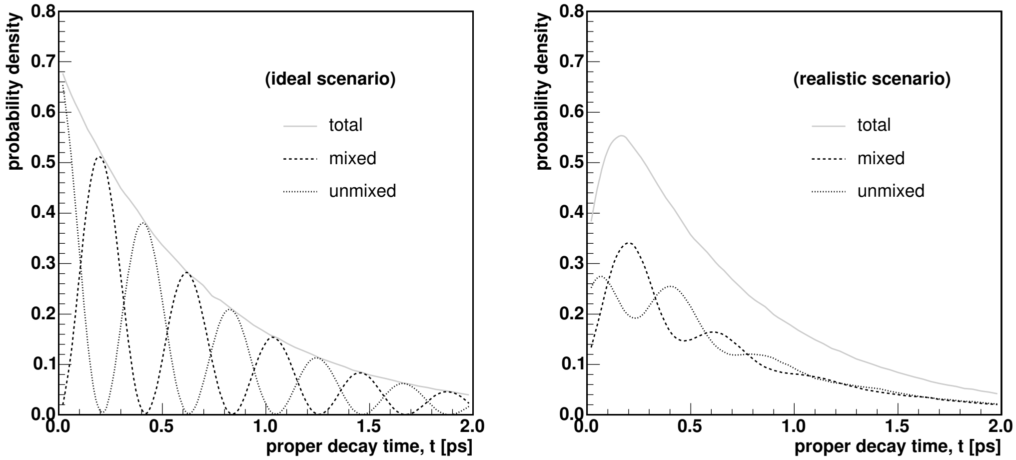


Figure 1-11: Realistic effects on the oscillation signal: *i*) ideal scenario, for an oscillation frequency of $15\,\mathrm{ps}^{-1}$ (left), and *ii*) the resulting signal after selection bias, decay length and momentum resolutions (right).

to determine the amplitude $\mathcal{A}$ as a function of the frequency which itself is fixed, at each step, to a different probe value. The method is appropriate for searching for fast oscillations, making it convenient for evaluating exclusion conditions on individual frequency values. A frequency point is excluded to a given confidence level (C.L.) if the hypothesis $\mathcal{A} = 1$ is excluded in a one-sided Gaussian test. Specifically, all values of Δm for which the combined variable $\mathcal{A} + 1.645\sigma_{\mathcal{A}}$ is smaller than 1 are excluded at 95% C.L., where $\sigma_{\mathcal{A}}$ is the total uncertainty on $\mathcal{A}$. The exclusion limit is defined as the largest frequency value below which all frequencies are excluded. The sensitivity for 95% C.L. is given by the range where $1.645\ \sigma_{\mathcal{A}} < 1$, *i.e.* it corresponds to the expected limit if an average observed amplitude $\mathcal{A} = 0$ (expected in the absence of oscillation) would be obtained. This is the method used in Chapter 7 for scanning for B_s oscillations. It allows as well a more straightforward combination of results from different analyses.

The necessary modifications stemming from resolution and bias effects mentioned above regarding the lifetime PDF (1.38) hold equally for the mixing PDF (1.39). The latter is represented in Figure 1-11 before and after resolution, bias and mistagging effects. These effectively cause the dampening of the oscillations.

The statistical significance of an oscillating signal may be approximated (B.17) by

$$\underset{\text{significance}}{\text{mixing}} \sim \sqrt{\frac{\epsilon\mathcal{D}^2\, S}{2}}\, \sqrt{\frac{S}{S+B}}\, e^{-\frac{\sigma_t^2 w^2}{2}} \;, \tag{1.40}$$

where S and B are the number of signal and background events in the selected sample and σ_t stands for the resolution on the proper decay time. The effective signal yield of the sample is seen to be scaled down by the tagging power $\epsilon\mathcal{D}^2$ which quantifies the limited performance of the flavor tagging algorithms employed. The effect of the proper decay time resolution is observed to be determining as well, particularly so in the case of higher oscillation frequencies. This allows one to anticipate the complementarity of the hadronic and semileptonic samples. The relatively large yields of the latter are expected to provide the dominant contributions to the lower range of the probed frequency spectrum, while the precise resolutions which characterize the former samples are expected to contribute the most discerning power at higher frequencies. The expression (1.40) can be used to estimate the sensitivity of the sample given its relevant characteristics. In Section 8.3 this is used to quantify expected sensitivity increases in view of improvements such as in flavor tagging and the increasing samples' size.

Constraining the CKM parameters

The obtained oscillation frequency information may be combined with other pieces of experimental and theoretical information relevant to infer the parameters of the unitarity triangle within the CKM framework. This is explored in Chapter 9. The various constraints are combined to form an inference framework based on Bayesian probability. The basic idea is that beliefs (*i.e.* probabilities) on the value of each input quantity are propagated into beliefs about the output quantities, such as the $\bar{\rho}$ and $\bar{\eta}$ CKM parameters. Such *posterior* probabilities can also be obtained for the input quantities themselves. The inference procedure may be carried out both including and excluding selected inputs, such as Δm_s, thus emphasizing its constraining effect.

A layout of the analyses and of this document is depicted in Figure 1-12.

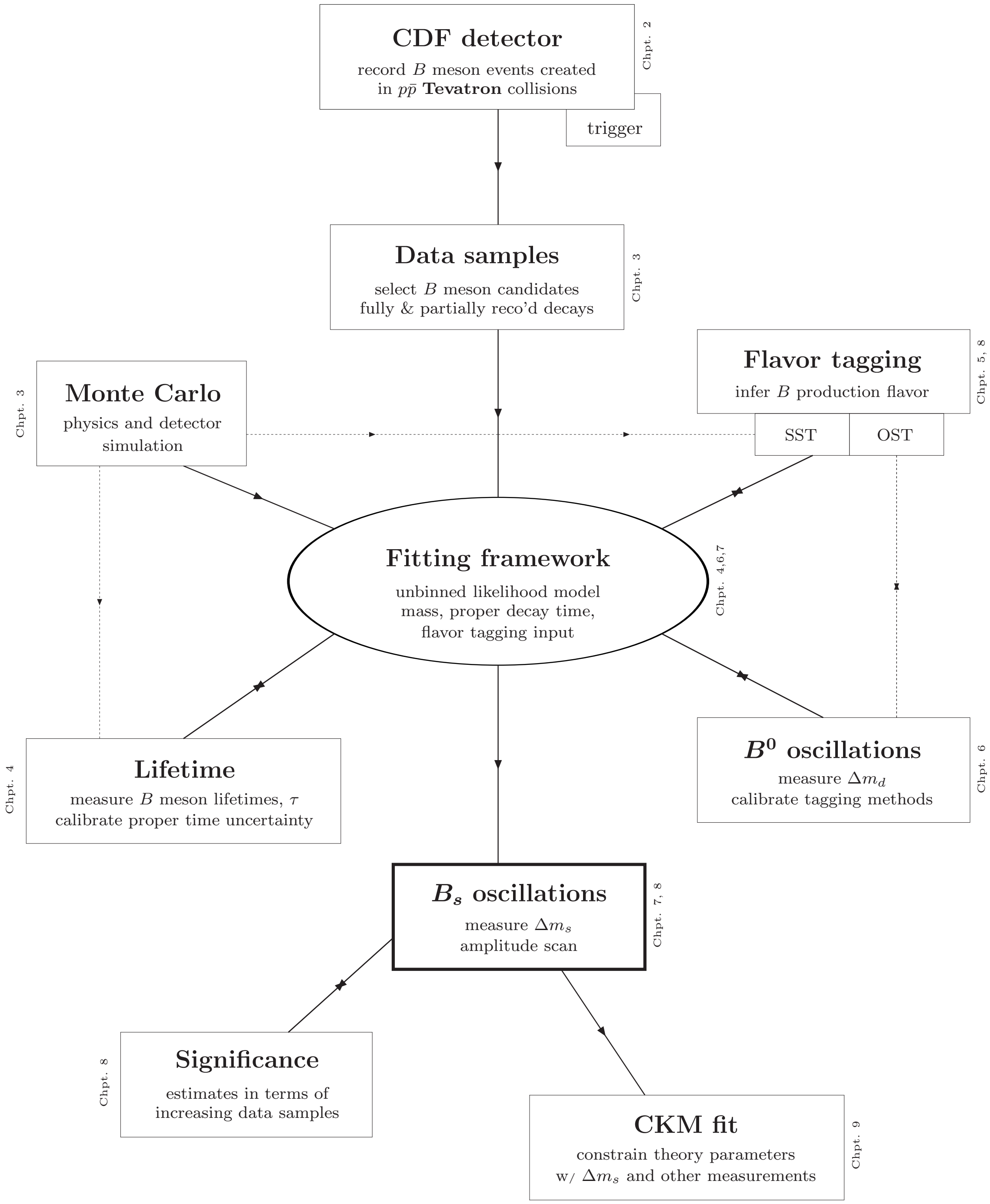


Figure 1-12: The big picture — analysis layout.

Chapter 2

Experimental apparatus

The Tevatron is currently the world's highest energy accelerator. It is lodged in an underground ring with radius of about 1 km located at the Fermi National Accelerator Laboratory (Fermilab), Illinois, USA. Protons and antiprotons ($p\bar{p}$) are brought into collision with a center of mass energy of approximately 1.96 TeV. The collisions take place at two points separated by about 120° along the ring. The CDF detector, which collects the data analyzed in this dissertation, surrounds one such collision point. The other collision point lays at the center of the DØ detector. Both the accelerator and the collider detectors underwent major upgrades between 1997 and 2001, mainly aimed at achieving, and to cope with, higher luminosities. The upgraded machine accelerates 36 bunches of protons and antiprotons, resulting in a time between bunch crossings of 396 ns. The period in which the current data is taken is referred to as the Run II of Tevatron and the detector is referred as CDF II. In the following sections we first outline the Fermilab accelerator complex, and describe the components of the CDF II detector, data taking structure and trigger systems used to collect and measure the properties of the particles produced in the $p\bar{p}$ collisions.

2.1 The Tevatron collider

In order to create the world's most powerful particle beams, Fermilab uses a series of accelerators. The diagram in Figure 2-1 shows the paths taken by protons and antiprotons from initial acceleration to collision in the Tevatron.

Proton source

The process leading to the production of $p\bar{p}$ colliding beams is initiated with a Cockcroft-Walton [32] pre-accelerator. Negatively charged ions H^- are created from the ionization of hydrogen gas and accelerated to a kinetic energy of 750 keV. The H^- ions are delivered to a 150 m long linear accelerator (Linac) [33]. The Linac uses a radio-frequency (RF) field, running at about 800 MHz, to further accelerate the H^- ions to an energy of 400 MeV. Before entering the next stage, a carbon foil removes the electrons from the H^- ions, leaving only the protons. The 400 MeV protons are

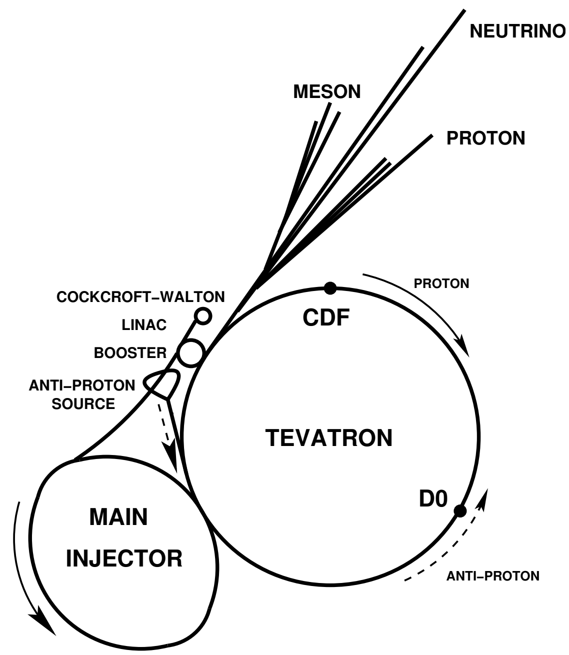


Figure 2-1: Layout of the Fermilab accelerator complex.

then injected into the Booster, a 74.5 m diameter circular synchrotron [33]. The intensity of the proton beam is increased by injecting new protons into the same orbit as the circulating ones. The protons travel around the Booster about 200,000 times to a final energy of 8 GeV. Protons emerge from the Booster in 84 bunches, with about 6×10^{10} protons each, spaced by about 19 ns.

Main Injector

Protons are extracted from the Booster into the Main Injector [34], a synchrotron of about 3 km in circumference which operates at 53 MHz. The Main Injector accelerates both protons and antiprotons from 8 GeV to 150 GeV. The bunches are also coalesced into 36 bunches per beam, before being finally injected into the Tevatron, where they undergo the final acceleration stage.

The Main Injector accomplishes further duties, which include the acceleration of protons to 120 GeV to be delivered to fixed target areas, the NuMI beamline for neutrino production, and the antiproton production target.

Antiproton production

One major advantage of a $p\bar{p}$ collider is that the proton and antiproton beams can circulate in opposite directions sharing the same magnet and vacuum systems. One major disadvantage of using $\bar{p}$ beams, nevertheless, is the associated cost of production. In the absence of technical problems with the accelerator such as system-to-system transfer inefficiencies, antiproton supply constitutes the most limiting factor for attaining higher luminosities.

The antiprotons are produced using a proton pulse of 120 GeV from the Main Injector, which is directed onto a nickel target. In the collisions, about 20 antiprotons are produced per one million incident protons, with a mean kinetic energy of 8 GeV. The produced antiprotons are collected and focused by a lithium lens and separated from the other by-products of the proton-nickel scattering using a bending magnet.

Exiting the collecting lens, the antiprotons have a large spread of momentum in longitudinal and transverse directions. Before being accelerated and further prepared for collisions, the particles must be confined to a smaller phase space volume. As this condition thermodynamically corresponds to a lower temperature, the process is referred to as beam *cooling*. Antiprotons are initially cooled in the Debuncher collected and further cooled using stochastic cooling [35] in the Accumulator. Stochastic cooling is a feedback-based method, where the particles' motion is sampled with pickup sensors placed around the beam, and their trajectories corrected with kicker electrodes and magnets. Such distinctive efforts are needed for obtaining cold antiproton beams without significant antiproton loss; protons, on the other hand are more readily available, and consequently cold beams may in this case be more simply produced by removing particles found outside the desired phase space volume. The *stacking* of antiprotons in the Accumulator, depending on the desired beam intensity, may take up to a day. The antiprotons have at this stage a characteristic energy of 8 GeV, with reduced beam size and momentum spread.

The stacked antiprotons, in a bunched beam configuration identical to that of the protons, are delivered back into the Main Injector. There they are accelerated to 150 GeV to be transferred next to the Tevatron for final acceleration and collision. Not all antiprotons are used during a *store* (time period of stable circulation that the colliding $p\bar{p}$ beams are retained in the Tevatron). The Recycler ring makes the recovery of these antiprotons (which amount to about 75% of the original injection quantity) possible for use in a later store, thus significantly reducing the stacking time. The Recycler is installed in the Main Injector enclosure, and functions as an antiproton storage ring. By storing antiprotons in the Recycler high antiproton intensities can be achieved as the stacking rate in the Accumulator is reduced at high antiproton intensities.

Tevatron

The final stage of Fermilab's accelerator chain is the Tevatron. It receives 150 GeV protons and antiprotons from the Main Injector and accelerates them to 980 GeV. The protons and antiprotons circle the Tevatron in opposite directions. The beams are brought to collision at the center of the two detectors, CDF II and DØ.

The antiprotons are loaded after the protons have been injected. Before antiproton injection, a set of electrostatic separators are used to create a pair of non-intersecting closed helical orbits with the protons circulating on one strand of the helix and the antiprotons on the other. This provides transverse separation of the proton and antiproton bunches as they pass each other at crossing points other than CDF and DØ.

number of bunches (N_B)	36
bunch rms [m]	0.37
bunch spacing [ns]	396
protons/bunch (N_p)	2.7×10^{11}
antiprotons/bunch ($N_{\bar{p}}$)	3.0×10^{10}
total antiprotons	1.1×10^{12}
β^* [cm]	35
interactions/crossing	2.3
peak luminosity [cm^{-2}s^{-1}]	1.2×10^{32}

Table 2.1: Tevatron Run II parameters.

Once loading is complete beams are accelerated to the maximum energy and collisions are initiated. Both protons and antiprotons circulate in three trains of 12 bunches, separated by about 2.6 μs, with bunches in each group spaced by 396 ns. After acceleration is complete, the beams are further focused in the CDF and the DØ interaction regions, the transverse size being reduced from about 1 mm to about 25 μm. This results in an increase of the chance of a proton and antiproton colliding.

The *instantaneous luminosity*, to which the rate of collisions is proportional, in the absence of a crossing angle or offset in beams position, is approximately given by

$$L = \frac{f N_B N_p N_{\bar{p}}}{2\pi(\delta_p^2 + \delta_{\bar{p}}^2)} F\left(\frac{\delta_l}{\beta^*}\right) , \tag{2.1}$$

where f is the bunch revolution frequency, N_B is the number of bunches, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch, and $\delta_{p(\bar{p})}$ is the protons (antiprotons) *rms* beam size at the interaction point. F is a form factor which corrects for the bunch shape and depends on the ratio of the bunch length δ_l to the beta function β^* at the interaction point (the beta function is a measure of the beam width, and it is proportional to the beam's x and y extent in phase space). The Run II design parameters are summarized Table 2.1. The *integrated luminosity*, defined as $\mathcal{L} = \int L\, dt$, is more relevant to physics analyses. The probability of occurrence for interactions is directly proportional to the cross section of the process σ [cm^2] and to $\mathcal{L}$ [cm^{-2}]. For cross sections observed at high energy collisions the preferred basic unit is the *barn* ($b = 10^{-24}$cm^2). The integrated luminosity delivered by the accelerator along with that recorded by CDF are illustrated in Figure 2-2.

2.2 The CDF II detector

The Collider Detector at Fermilab (CDF) is a general-purpose, azimuthally and forward-backward symmetric apparatus designed to study $p\bar{p}$ collisions at the Tevatron. Its design, accordingly, rather than being tailored toward a specific class of physics measurement, is optimized toward extracting different properties of all particle species emanating from the $p\bar{p}$ collisions. A diagram of the CDF II detector is

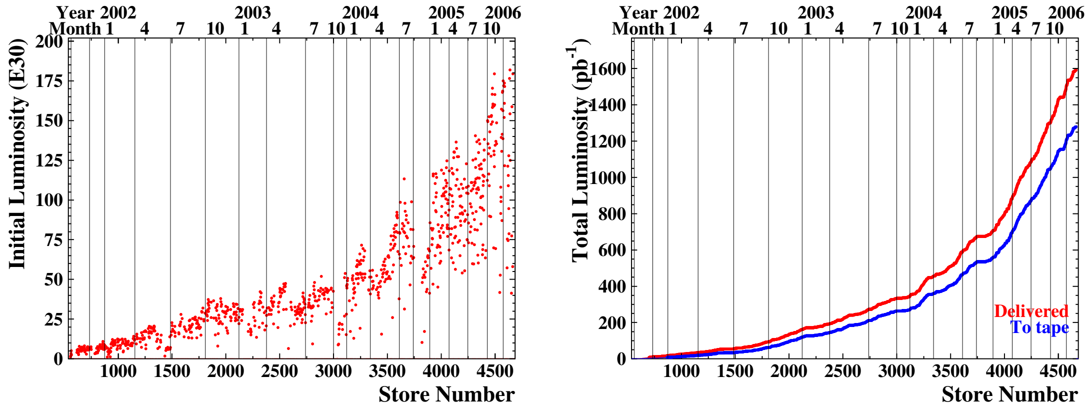


Figure 2-2: Initial luminosity [cm^{-2}s^{-1}] per Tevatron store (left), and integrated luminosity [pb^{-1}] delivered and recorded (right).

shown in Figure 2-3.

The innermost detector system is the integrated tracking and vertexing system. It consists of a Silicon microstrip detector and a multi-wire drift chamber, which have a cylindrical geometry concentric with the beam. It is designed to detect charged particles, measure their trajectories and momenta. Reconstructed particle trajectories are referred to as *tracks*. Multiple track reconstruction allows the identification of vertices, where either the $p\bar{p}$ interaction occurred (primary vertex) or the decay of a long-lived particle took place (secondary or displaced vertex).

The tracking system is surrounded by the Time of Flight (TOF) system, designed to provide particle identification for low-momentum charged particles. Both the tracking and Time of Flight systems are placed inside a superconducting coil, which generates a 1.4 Tesla solenoidal magnetic field parallel to the beam axis. The coil is surrounded by electromagnetic and hadronic calorimeters, which measure the energy of particles that shower when interacting with matter. The calorimetry systems are surrounded by the muon detector system. Muons interact with matter primarily through ionization, and act as minimally ionizing particles depositing only small amounts of ionization energy. Therefore, they are able to penetrate through the tracking, TOF, solenoid and calorimeter systems, with minimal interaction with the detector material. The muon chambers, therefore, are located radially outside the calorimeters.

All of the detector parts mentioned above provide information which is used in the analyses presented in this dissertation, and will be described in the following sections. A more complete and detailed description of the CDF II detector can be found in the Technical Design Reports [36].

Unlike in e^+e^- collisions, in $p\bar{p}$ collisions not all of the center of mass energy of the $p\bar{p}$ system is absorbed by the participants in the fundamental interaction. It is the partons inside the proton and antiproton (valence or sea quarks, or gluons) that are involved in the fundamental interactions, and these carry only a fraction of the proton and antiproton momenta. Furthermore, as a result of a possible imbalance in

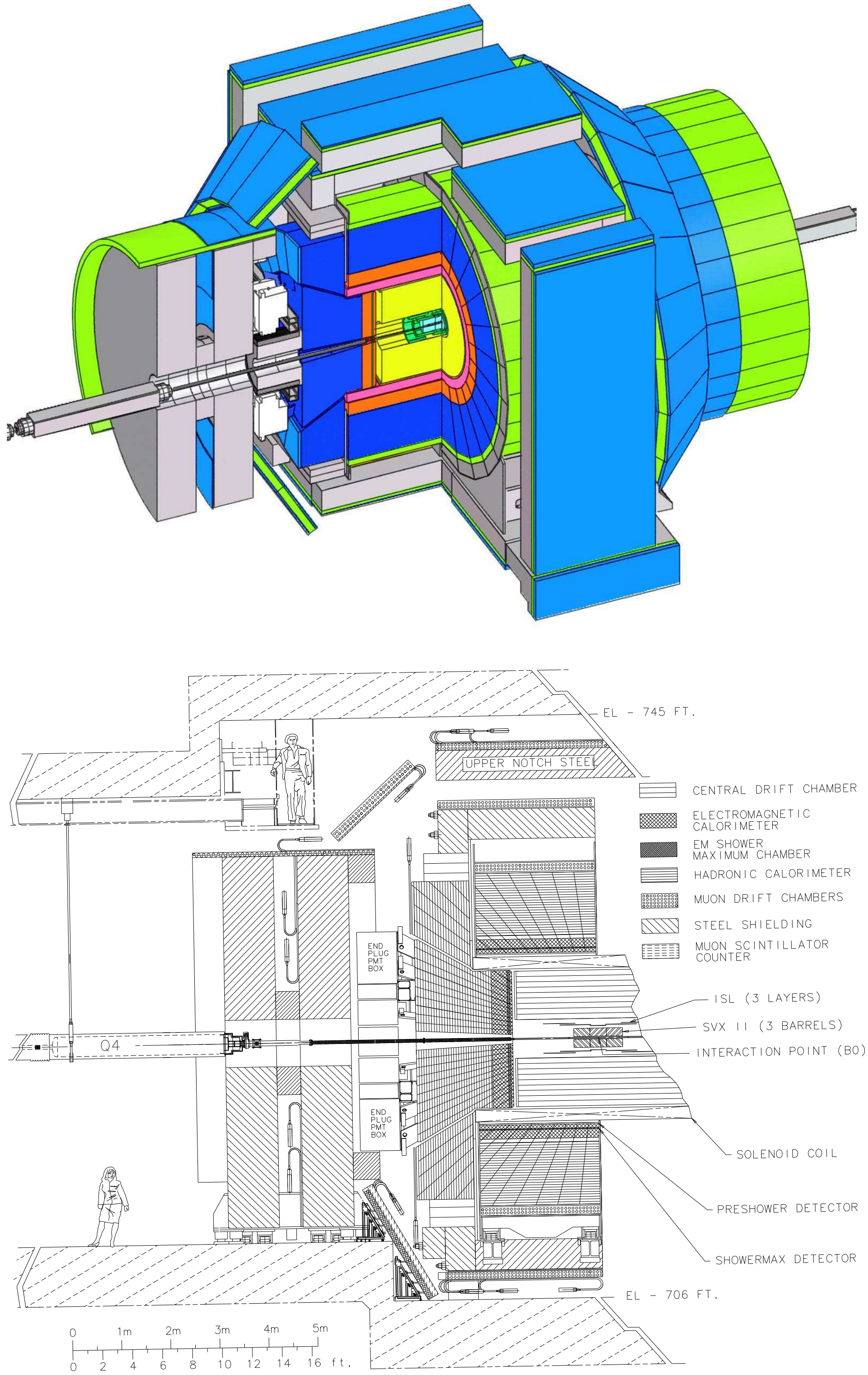


Figure 2-3: The CDF II detector with a quadrant cut (top) and elevation view (bottom) exposing the different sub-detectors.

the momenta of the two interacting partons, the observed physics interactions often have large boosts along the beam direction (the "longitudinal" direction). That is not the case on the plane perpendicular ("transverse") to the beam, given the negligible transverse motion of the colliding particles. It follows, for instance, that the vectorial sum of all particles's transverse momenta in a collision vanishes.

The geometrical center of the detector, roughly coincident with the beam crossing region, serves as the origin for the CDF coordinate systems. In the Cartesian system, the x axis is in the (horizontal) plane of the accelerator, pointing radially outward, and the y axis points upwards. The beams travel through the detector approximately parallel to the the z axis, with protons moving in the positive z direction. The detector's barrel-like shape makes it convenient to introduce cylindrical (r, ϕ, z) or polar (ρ, ϕ, θ) coordinates. The r and ρ coordinates indicate the associated radial distances from the origin, and ϕ is the azimuthal angle. The r-ϕ (or x-y) plane is referred as the *transverse plane*, as it is perpendicular to the beam line. The polar angle θ is defined relative to the z axis. It is convenient to introduce yet another variable, as an alternative to the polar angle; namely, the pseudorapidity η is defined as

$$\eta \equiv -\ln \tan \left(\frac{\theta}{2}\right) . \tag{2.2}$$

This quantity coincides with the ultrarelativitic, massless limit of a particle's rapidity

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} . \tag{2.3}$$

Differences in y, and in the mentioned limit therefore also in η, are invariant under Lorentz boosts along the z direction. The angular variable η is often used for specifying the geometrical coverage of the detector sub-systems. The variable

$$\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.4}$$

is commonly used for specifying angular distances.

2.2.1 Tracking and vertexing

Charged particles cause ionization as they pass through matter, leaving trails of charge and energy clusters, denoted *hits*, along their paths. Once detected hits can be used to reconstruct the particles' trajectories, in the process known as *tracking*. From the intersection of multiple reconstructed tracks, *vertices* may be obtained.

CDF's innermost tracking device is a silicon microstrip vertex detector, which consists of three concentric systems. A layer of silicon sensors, called Layer 00 (L00) [37], is installed directly onto the beryllium vacuum beam pipe, with the sensors at radii 1.35 and 1.62 cm from the beam. The layer of silicon on the beam pipe is followed by five concentric layers of silicon sensors, denoted the Silicon Vertex Detector SVX [38], located at radii between 2.45 and 10.6 cm. The Intermediate Silicon Layers, ISL [39], are the outermost silicon sub-detector systems, consisting of one layer at a radius of 22 cm in the central region and two layers at radii 20 and 28 cm in the forward regions. L00 provides $r-\phi$ information, while SVX and ISL provide both $r-\phi$ and z measurements.

Surrounding the silicon detector is the Central Outer Tracker (COT) [40], a 3.1 m-long cylindrical open-cell drift chamber covering radii from 40 to 137 cm. The layout of the systems is shown in Figure 2-4.

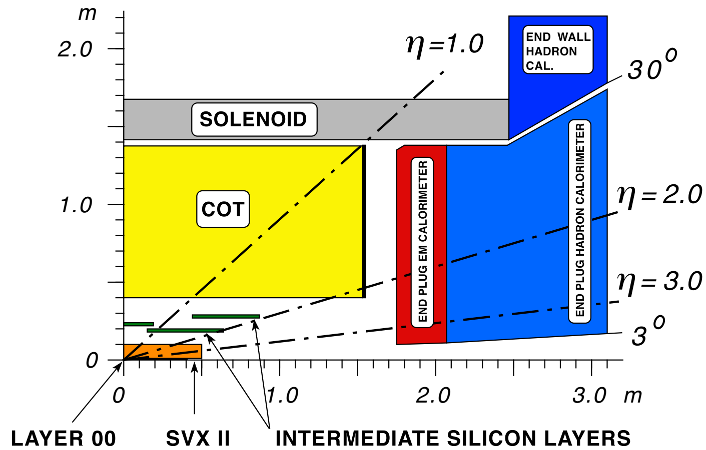


Figure 2-4: The CDF II tracker layout showing the different sub-detector systems.

Silicon vertex detectors

A silicon tracking detector is fundamentally a reverse-biased p-n junction. The passage of a charged particle leaves a trail of electron-hole pairs induced by ionization, which drift in opposite directions in the existing electric field. The reverse-biased voltage increases the gap between the conduction and valence bands across the p-n junction and reduces the current from thermal excitation. By segmenting the p or n side of the junction into finely spaced *strips* and reading out the charge deposition separately on every strip, a well-localized signal is obtained. The typical distance between two strips is about 60 μm. Charge deposition from a single particle passing through the silicon sensor will usually be detected on a *cluster* of strips rather than just one. The hit position is extracted by weighting the strip positions by the amount of charge collected.

There are two types of microstrip detectors: single and double-sided. In single-sided detectors only one (p) side of the junction is segmented into strips. Double-sided detectors have both sides of the junction segmented into strips. The benefit of double-sided detectors is that while one (p) side has strips parallel to the z direction, providing ϕ position measurements, the other (n) side can have strips at an angle (stereo angle) with respect to the z direction, which will give z position information. The innermost layer, L00, is made of single-sided silicon sensors which only provide $r-\phi$ measurements. The SVX and ISL are made of double-sided silicon sensors. As

	layer 0	layer 1	layer 2	layer 3	layer 4
number of ϕ strips	256	384	640	768	869
number of z strips	256	576	640	512	869
stereo angle	90°	90°	+1.2°	90°	-1.2°
ϕ strip pitch [μm]	60	62	60	60	65
z strip pitch [μm]	141	125.5	60	141	65
active width [mm]	15.30	23.75	38.34	46.02	58.18
active length [mm]	72.43	72.43	72.38	72.43	72.43

Table 2.2: Relevant parameters for the layout of the sensors of the SVX layers.

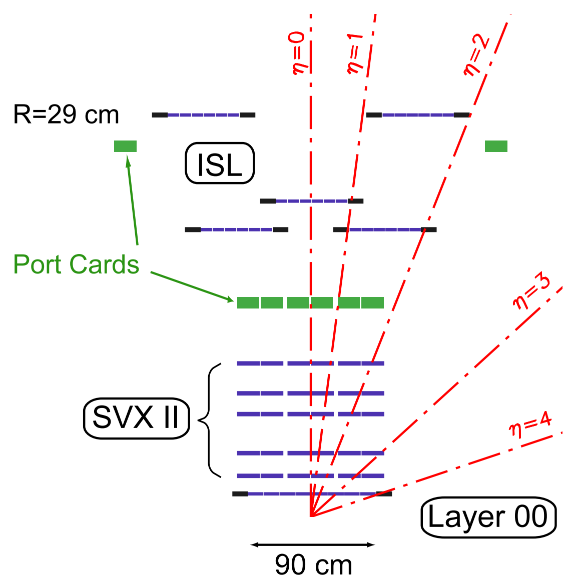


Figure 2-5: Coverage of the different silicon sub-detectors projected into the $r-z$ plane; the r and z axes have different scales.

shown in Table 2.2, the SVX layers have different stereo angles. Two layers have a 1.2° stereo angle and three have a 90° stereo angle. The ISL detector provides small angle (1.2°) stereo information.

The SVX silicon sensors are supported by carbon fiber rails in assemblies called *ladders*. Each ladder hosts four pieces of silicon sensors, whose readout electronics are mounted directly to the surface of the silicon sensor at each end of the ladder. Each of the 5 SVX concentric layers is formed of 12 ladders in ϕ, with a small overlap at the edges for improved coverage. The ladders are further arranged in three, 29 cm long barrels, mounted end-to-end along the z axis. Each barrel contains a total of 60 ladders, mounted between two beryllium bulkheads which provide mechanical support and which also carry the water cooling lines for the readout electronics.

The coverage of the silicon detector sub-systems is shown in Figure 2-5.

Drift chamber

The Central Outer Tracker, COT [40], is a cylindrical multi-wire open-cell drift chamber, with a coverage of $|\eta| < 1$. The system is segmented into 8 concentric *superlayers*, as represented in Figure 2-6. A superlayer is divided into *cells* in ϕ. Each cell contains 12 sense wires and 17 potential (for field shaping) wires along with adjacent ground field (cathode) sheets on either side. The potential wires alternate with the sense wires at a pitch of 0.3556 cm. The distance between the wires and the field sheets is 0.88 cm. Both the sense and the potential wires are 40 μm diameter gold plated Tungsten. The field sheets are 6.35 μm thick Mylar with vapor-deposited gold on both sides, and are shared among adjacent cells. The cells layout is shown in Figure 2-7.

The maximum drift distance is approximately the same for all superlayers. Thus, the number of cells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30,240 sense wires. The wires in superlayers 2, 4, 6 and 8 run along the z direction ("axial"), while the wires in the other superlayers are strung at a small angle (2°) with respect to the z direction ("stereo").

The COT chamber is filled with an Argon-Ethane gas mixture and Isopropyl alcohol (49.5:49.5:1). The voltages on the sense wires are 2600-3000 volts and 1000-2000 volts on the potential wires; the ground sheets are grounded. Charged particles that pass through the gas mixture leave a trail of ionization electrons. The electric field created by the cathode field sheets and the potential wires cause the electrons to drift toward the sense wires; the drift velocity is $\sim 50\mu$/ns. Due to the magnetic field in which the COT is immersed, electrons drift at a Lorentz angle of $\sim 35°$. Therefore, the cell is tilted by 35° with respect to the radial direction to compensate for this effect. After the tilt the electrons drift approximately perpendicularly to the wire plane. The maximum electron drift time is less than 200 ns, much shorter than the bunch crossing frequency of 396 ns.

When the ionization electrons drift close to the wires surface the local $\frac{1}{r}$ electric field vigorously accelerates them. This further causes secondary ionization thus initiating a limited avalanche, which produces a signal (hit) on the wire large enough to be read out by the electronics attached to the end of the wire. Signals on the sense wires are processed by the ASDQ (amplifier, shaper, discriminator with charge encoding) chip, which provides input protection, amplification, pulse shaping, baseline restoration, discrimination and charge measurement [41]. The leading edge gives the arrival time information and the pulse width encodes the amount of charge collected by the wire. Upon calibration the digital width is thus used in the measurement of the ionization energy loss along the trail of the particle through the COT, dE/dx, for particle identification. The pulse is sent through ~ 105 cm of micro-coaxial cable, via repeater cards to time to digital converter (TDC) boards in the collision hall. Hit times are later processed by pattern recognition software to form helical tracks. The hit resolution of the COT is about 140 μm. The transverse momentum (p_T) resolution is $\sigma_{p_T}/p_T \sim 0.15\% \cdot p_T$.

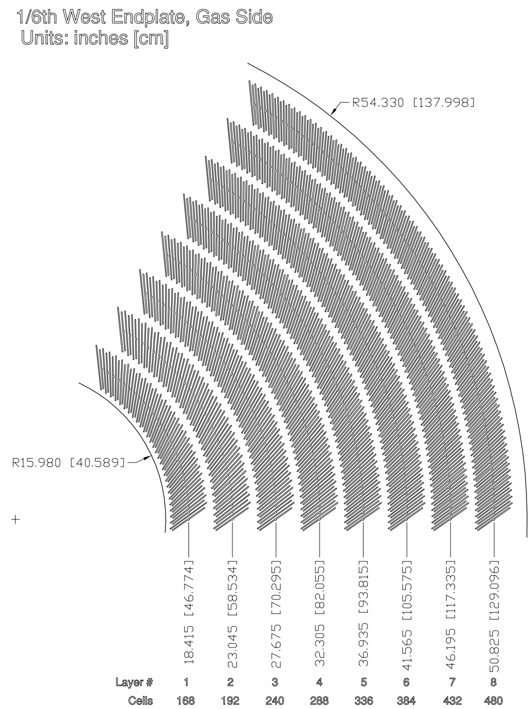


Figure 2-6: Layout of wire planes on a COT endplate.

Track reconstruction

Charged particles moving through a uniform solenoidal magnetic field, such as that which permeates the CDF tracking volume, follow helical trajectories. To uniquely describe such a trajectory in three dimensions five parameters are needed. Three of those parameters are chosen to describe a position and two more are associated to the momentum vector at that position. Specifically, the set of parameters is: C, $\cot\theta$, d_0, ϕ_0, z_0. The latter three correspond to the r, ϕ and z cylindrical coordinates of the point of closest approach of the helix to the z axis, C is the helix curvature and θ is the angle between the z axis and the particle's momentum vector.

The projection of the helix on the transverse $r{-}\phi$ plane is a circle. The curvature C is a signed quantity given by $C = \frac{q}{2R}$, where R is the radius of the circle and the signal of the particle's charge q is given by its bending direction in the magnetic field. The transverse momentum p_T can be calculated from the helix curvature and magnetic field (B_{solenoid}) as

$$p_T = \frac{cB_{solenoid}}{2|C|}\,. \tag{2.5}$$

The longitudinal momentum p_z component can then be also calculated as $p_z = p_T \cdot \cot\theta$. The direction of the transverse momentum vector $\vec{p}_T$ is implicitly given by ϕ_0, noting that at any given point of the helix the track momentum is tangent to the

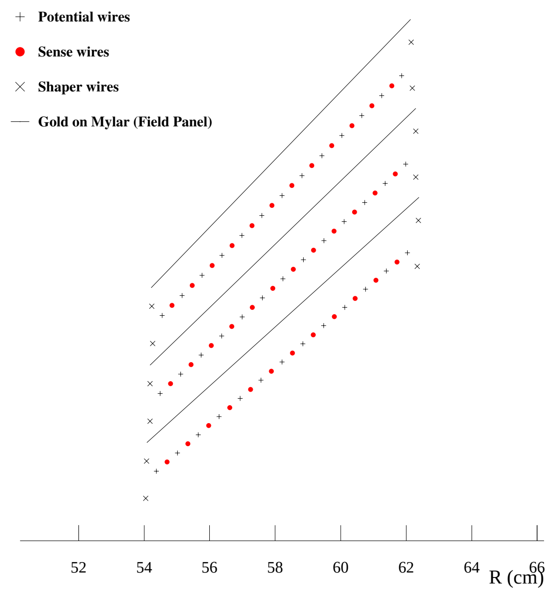


Figure 2-7: Transverse view of three adjacent COT cells.

helix.

The *impact parameter* d_0 of a track is a signed variable, whose absolute value corresponds to the distance of closest approach of the track to the beamline. It is defined as

$$d_0 \quad = \quad q \cdot \left(\sqrt{x_c^2 + y_c^2} - R \right) \ , \tag{2.6}$$

where (x_c, y_c) is the center of the helix as projected onto the x-y plane.

Pattern recognition

The process of track reconstruction begins using only COT information. The COT electronics report hit time and integrated charge (codified as the leading edge and width of the digital pulse, respectively, from the TDC) for every sense wire. The hit locations are reconstructed from the time difference between when the ionization occurs (the collision time plus the particle's time of flight) and when the signal is picked up by the wire. Corrections from the global time offset of the wires and cables, electronic channel pedestals, charged-based time slewing and non-uniform drift velocities are accounted for.

The first step in pattern recognition is to look for circular paths in the axial superlayers of the COT. Sets of 4 or more hits are searched for in each axial superlayer to form a straight line, or *segment*. Once segments are found, there are two approaches to track finding. One approach is to link together segments for which the measurements of curvature and ϕ_0 are consistent with lying tangent to a common circle. The other approach is to constrain the circular fit to the beamline starting from the outermost superlayer, and then add hits which are consistent with this path. Once a circular path is found in the $r-\phi$ plane, segments and hits in the stereo superlayers are added depending on their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not.

Once a track is reconstructed in the COT, it is extrapolated into the SVX. A three-dimensional "road" is formed around the extrapolated track, based on the estimated errors on the track parameters. Starting from the outermost layer, and working inward, silicon clusters found inside the road are added to the track. As a cluster is added, a new track fit is performed, which modifies the error matrix for the track parameters and produces a narrower road. In the first pass of this algorithm, $r-\phi$ clusters are added. In the second pass, clusters with stereo information are added to the track. If there are multiple tracks with differing combinations of SVX hits associated with the same COT track, the track with the largest number of SVX hits is chosen.

The track reconstruction efficiency in the COT is about 95% for tracks which pass through all 8 superlayers ($p_T \geq 400$ MeV/c), and 98% for tracks with $p_T \geq 10$ GeV/c. It depends mostly on the number of tracks present, as hits from one track can shadow hits from the other track. The track reconstruction efficiency for the addition of SVX information is about 93% for tracks with at least 3 SVX $r-\phi$ hits.

Vertexing

A particle's production point cannot be inferred from a reconstructed helix alone. In general its production vertex is found from the intersection with one or more helices associated with particles assumed to have originated from the same space point. The process of finding the spatial coordinates, along with corresponding uncertainties, of the tracks' intersection point is referred to as *vertexing*. Algorithms employed at CDF for this purpose are described in [42].

For decaying particles, we define the displacement L_{xy} as

$$L_{xy} = \vec{d} \cdot \hat{p}_T \,, \tag{2.7}$$

where $\vec{d}$ is the displacement vector of the decay vertex in the transverse plane, and $\hat{p}_T$ is the unit vector in the direction of $\vec{p}_T$.

2.2.2 Calorimetry

Surrounding the CDF tracking volume, and outside the solenoid coil, are found the electromagnetic (EM) calorimeters followed by the hadronic (HA) calorimeters. Comparison of the energy deposition in the electromagnetic and the hadronic calorimeters provides separation between electrons and photons (whose energy is mostly absorbed in the electromagnetic calorimeter) and hadrons (which deposit a large fraction of their energy in the hadronic calorimeter).

	η coverage	thickness	energy resolution [%]
CEM	$\lvert\eta\rvert < 1.1$	$19X_0$	$13.5/\sqrt{E\sin\theta} \oplus 2$
PEM	$1.1 < \lvert\eta\rvert < 3.6$	$21X_0$	$16/\sqrt{E\sin\theta} \oplus 1$
CHA	$\lvert\eta\rvert < 1.1$	$4.5\lambda_0$	$75/\sqrt{E\sin\theta} \oplus 3$
WHA	$0.7 < \lvert\eta\rvert < 1.3$	$4.5\lambda_0$	$75/\sqrt{E\sin\theta} \oplus 3$
PHA	$1.3 < \lvert\eta\rvert < 3.6$	$7\lambda_0$	$80/\sqrt{E\sin\theta} \oplus 5$

Table 2.3: Pseudorapidity coverage, thickness, and energy resolution (with E in GeV) for the different calorimeter sub-detectors of CDF II; λ_0 signifies interaction length and X_0 radiation length.

The calorimeter detectors are segmented in η and ϕ towers that point to the interaction region, in an arrangement referred to as projective tower geometry. The calorimeter system is divided into *central*, *wall* and *plug* regions. They are denoted central electromagnetic (CEM), central hadronic (CHA), wall hadronic (WHA), plug electromagnetic (PEM) and plug hadronic (PHA) sub-systems. The central calorimeters cover the pseudorapidity region of about $|\eta| < 1$. Each tower spans 15° in azimuth and about 0.1 in pseudorapidity. Plug calorimeters extend the pseudorapidity coverage to $|\eta|$=3.6, and use variable segmentation: 7.5° to 15° in ϕ, and 0.1 to 0.6 in $|\eta|$.

The CEM and PEM towers consist of lead sheets interspersed with scintillator as the active detector medium. The light signal is wavelength-shifted and carried by light guides to the photomultiplier tubes (PMTs), which measure the number of scintillation photons produced in the EM shower that is formed during the particles' passage through the detector. Both calorimeters are equipped with shower maximum detectors, CES and PES for central and plug regions respectively. These are proportional chambers of wires and strips, embedded at the depth of about 6 radiation lengths (X_0), which measure the coordinates of showers produced by electrons and photons with 1-3 mm accuracy. Shower maximum detectors are used for identification of electrons and photons by matching the position of electromagnetic showers with incident tracks. They also provide measurements of the transverse shower profile which are used to separate photons from neutral pions, and the pulse height helps to identify electromagnetic showers. Another set of multi-wire proportional chambers is located between the solenoid coil and the CEM at a radius of about 168 cm. These constitute the preshower detector (CPR), which samples the electromagnetic showers that begin in the solenoid material (1.08 X_0), providing enhanced soft electron identification and separation from lesser ionizing particles such as muons and pions.

The hadronic calorimeters are located immediately behind the electromagnetic calorimeters, with matching segmentation. The CHA and WHA towers are formed of alternating layers of iron and scintillator. The pseudorapidity coverage, thickness and resolutions for the different electromagnetic and hadronic calorimeters are given in Table 2.3. A detailed description can be found in [36].

2.2.3 Muon chambers

The radially outermost component of CDF II is the muon system, which comprises sets of drift chambers and scintillators. While electrons and hadrons deposit most of their energy in the calorimeter material, muons which are minimum ionizing penetrate much more material than any other charged particles, and given enough momentum reach the muon chambers. There they leave a track segment, called a muon *stub*. If a stub matches a track measured in the COT then the two are combined to form a muon candidate.

The CDF II detector contains four muon systems: the central muon (CMU), the central muon upgrade (CMP), the central muon extension (CMX), and the intermediate muon (IMU) detectors [43]. The coverage of each sub-system is represented in Figure 2-8.

The CMU detector is located around the outside of the central hadronic calorimeter at a radial distance of 347 cm. It is segmented in ϕ into 24 wedges, with a gap of $2.4°$ between wedges which limit the coverage in ϕ to about 84%. The CMU is also divided into East (positive z) and West (negative z) halves with about a 18.4 cm gap at $\eta = 0$ between the two. Each wedge is further segmented azimuthally into three $4.2°$ modules; there is a total of 144 modules. A module consists of 4 layers of 4 rectangular drift cells. A cell has at its middle a 226 cm long, 50 μm stainless steel sense wire, parallel to the z axis, and is filled with an identical (Argon-Ethane-alcohol) gas mixture as the COT.

The CMP consists of drift chambers located behind a 60 cm iron shield. The CMP chambers are comprised of rectangular ($2.5\,\mathrm{cm} \times 15\,\mathrm{cm} \times 640\,\mathrm{cm}$), single-wire drift tubes configured in four layers with alternate half-cell staggering. In combination with the CMU, the CMP improves further the purity of muon identification. Muon candidates with stubs in both CMU and CMP are referred to as CMUP muons.

The CMX detector provides an extension to the central muon systems to cover $0.6 < |\eta| < 1.0$. It is a conical arrangement of drift tubes similar to those of CMP cells and a sandwich of scintillators. It has a $30°$ gap (top, East) which is filled with the instrumentation for the solenoid cryogenic system. The CMX is azimuthally segmented in $15°$ wedges, each made up of 8 layers of drift tubes. The tubes are 180 cm long, have a rectangular cross-section, and are staggered.

The IMU consists of a barrel of drift chambers and scintillator counters mounted on the outer radius of two (non-energized) steel toroids with additional counters between the toroids, in a pseudorapidity region of $1.0 < |\eta| < 1.5$. The chambers and counters are similar to those in the CMP. Each chamber covers $12.5°$ in azimuth, and is about 363 cm long.

Shielding serves the desired duty of preventing hadrons from reaching the muon chambers. The effective hadronic shielding provided mostly by calorimeters, solenoid coil, and additional steel absorbers in the path to each muon sub-system is quantified in Table 2.4 in terms of pion interaction lengths. While an increasing amount of absorber improves muon purity, it also causes muons themselves to loose energy. As a result, muons below certain momentum thresholds do not reach the muon detectors; the rangeout thresholds for the individual systems are summarized in Table 2.4.

Another issue is Coulomb scattering, the effect of which becomes amplified with the introduction of additional material. Multiple Coulomb scattering randomly deflects a particle's path through the material, inducing deviations of the detected trajectory in the muon system from that expected by extrapolation from the tracking volume.

It remains possible, however, that a small fraction of hadrons that interact late in the calorimeters will produce secondary particles that still reach the muon chambers; these are referred to as *punch-throughs*. Additional background may be due to real muons from decays of pions and kaons, referred to as *decays-in-flight*, or simply to fake stubs produced by electronic noise.

Track-stub matching

Muon candidates are reconstructed by matching the stub position in the muon detector to a track measured in the COT and extrapolated to the muon chambers. Using the timing information from the drift cells of the muon systems, hit positions are found. A muon stub is obtained by fitting sets of hits using a least-square method to a segment line. A stub must have at least 3 hits associated to it. Stubs need then to be matched to COT tracks. For stubs reconstructed in the CMU, CMP, CMX and IMU, the set of tracks with transverse momentum above 1.4 GeV/c, 2.2 GeV/c, 1.5 GeV/c and 2.5 GeV/c, respectively, are considered. The tracks are extrapolated using a simplified geometry model of the muon candidate's motion in the non-uniform magnetic field in the calorimeter.

The matching procedure compares the position of the stub in a given muon chamber with the extrapolated position of the track. Differences in position and direction, along with stub χ^2 and track covariance matrix are used. The evaluated match distance in the $r{-}\phi$ plane between the projected track and the stub, denoted ΔX, is required to be smaller than 30 cm, 60 cm, 50 cm and 90 cm, respectively, for CMU, CMP, CMX and IMU muons. The mentioned matching and minimum momentum requirements are chosen in order to maximize the muon reconstruction efficiency while maintaining high muon purity.

Additional matching variables are further defined and evaluated, which serve as discriminating variables for muon identification. These include, for all muon subsystems, the opening angle $\Delta\Phi$ between the stub and the direction of the track extrapolation, projected on to the $r{-}\phi$ plane. The distance in the r-z plane, denoted ΔZ, is evaluated for all muon types except CMP (the CMP chambers measure the location of the stub in the $r{-}\phi$ plane only). Additional information such as the number of hits used to reconstruct the stub, and the χ^2 fit of the stub are also provided.

2.2.4 Time of Flight

Outside the tracking system, but still inside the superconducting magnetic coil, CDF II has a Time of Flight (TOF) [44] system. It is designed to distinguish low momentum pions, kaons and protons by measuring the time it takes these particles to travel from the primary vertex of the $p\bar{p}$ collision to the TOF system. The main

	η coverage	pion interaction length	minimum muon p_T
CMU	$\lvert\eta\rvert < 0.6$	5.5	1.4 GeV/c
CMP	$\lvert\eta\rvert < 0.6$	7.8	2.2 GeV/c
CMX	$0.6 < \lvert\eta\rvert < 1.0$	6.2	1.4 GeV/c
IMU	$1.0 < \lvert\eta\rvert < 1.5$	6.2-20	1.4-2.0 GeV/c

Table 2.4: Pseudorapidity coverage, pion interaction length, and minimum detectable p_T for the different muon sub-detectors of CDF II.

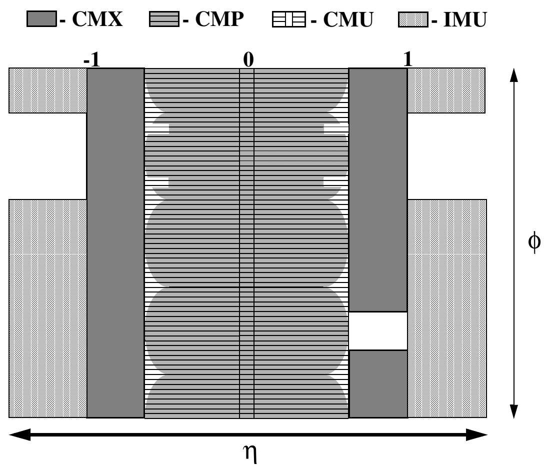


Figure 2-8: Coverage in η-ϕ of each muon system.

physics motivation for the introduction of this detector sub-system is to enhance particle identification for flavor tagging.

The TOF system consists of 216 scintillator bars, approximately 280 cm long and with a cross-section of 4 cm $\times$ 4 cm. The bars are arranged into a barrel at a radius of about 138 cm, and in the 4.7 cm of radial space between the COT cylinder and the cryostat of the superconducting solenoid. The pseudorapidity coverage is roughly $|\eta| < 1$. The scintillator bars are read out at both ends by fine-mesh photomultiplier tubes (PMT; Hamamatsu R7761), which provide a gain of $\sim 30,000$ inside the 1.4 Tesla magnetic field of the CDF solenoid.

Particles passing through the scintillating material of the bars deposit energy causing small flashes of visible light detected by the PMT. The signal from the photomultiplier tube is processed by a pre-amplifier circuit mounted directly onto the tube, and sent to the readout electronics in the collision hall. The readout electronics perform both time and amplitude digitization of the signal. The TDC information is a digitization of the time when the pulse leading edge reaches a fixed discriminator threshold. This time depends on the amplitude of the pulse, as larger pulse heights fire the discriminator at an earlier time than smaller pulse heights (*time-walk*). Larger pulses also give better timing resolution, as light attenuates while travelling through the scintillator material. Therefore, particles passing through the bar near the photomultiplier tube have better timing resolution than those which are farther away.

Particle identification

The TOF system expands CDF's particle identification (PID) capability. Along with the measurement of specific ionization dE/dx in the COT, the measurement of the particle's flight time provides complementary particle discrimination power in the low p_T region.

PID with TOF is performed by measuring the time of arrival t_{flight} of a particle at the scintillator with respect to the collision time, t_0. The particle's mass m is then determined as

$$m = \frac{p}{c}\sqrt{\frac{c^2\, t^2_{\text{flight}}}{L^2} - 1} \; , \tag{2.8}$$

where the track's momentum p and path length L between the beam collision point and the scintillator are measured in the tracking system.

The achieved timing resolution relies on the calibration of the response of each channel. The time t_i at which the discriminator of channel i is fired, for a track hitting the scintillator at a position z along the length of the bar, is described by

$$t_i = \alpha_i + t_0 + t_{\text{flight}} + \frac{1}{c_s}\left(\frac{l}{2} \pm z\right) - S_i(Q_i) \; , \tag{2.9}$$

where the constant offset α_i accounts for cable and electronics delays, l is the length of the scintillator bar, c_s is the effective speed of light propagation in the scintillator, and the last term describes the effect of time-walk introduced by the use of leading edge pulse discriminators, which depends on the integrated charge Q_i of the PMT pulse. The positive (negative) sign multiplying the z hit coordinate corresponds to the case where a PMT is on the positive z (negative z) end of the bar.

The TOF resolution may be estimated from the difference between the time of flight measured for each track t_{flight}, and the expected time of flight assuming the track is a pion. The calculation is improved by computing the expected flight time assuming various particle hypotheses with specific probabilities (f_π, f_k and f_p for pions, kaons and protons, respectively). Additionally, the resolution for any given channel degrades with the hit displacement from the photomultiplier tube; a linear parameterization in z suitably accounts for the effect. The timing resolution at the face of the PMT is about 110 ps.

2.3 Data acquisition and trigger systems

The Tevatron proton and antiproton bunches cross every 396 ns at the center of the CDF detector. Such a high collision rate, of about 2.5 million interactions per second, makes it not practical to store information about every single $p\bar{p}$ collision.

The readout of the entire detector produces an event size of the order of 200 kB, which would thus require a data recording rate of 0.5 GB/s and would amount to a few hundred tera-bytes of data in a single week of running. Furthermore, the time it takes to read the entire detector out, about 2 ms, would be long enough for another 5,000 or so interactions to take place. However, most processes of interest have cross sections at least three orders of magnitude smaller than the total $p\bar{p}$ cross section (at $\sqrt{s}$ = 1.96 TeV). The total hadronic cross section (including elastic, inelastic, and diffractive processes) is about 75 mb, while for instance the $b\bar{b}$ production cross section is only about 0.1 mb. This makes it possible to circumvent the above mentioned difficulties by reading out and storing only events with signatures found of interest. The process of real-time event selection is referred to as *trigger*.

The data acquisition (DAQ) and trigger systems are intrinsically coupled providing a common infrastructure for data taking. The design of the pipelined DAQ system and three-level staged deadtimeless trigger is represented in Figure 2-10. An event is passed on to the following trigger level if it has been accepted by the previous stage, being otherwise discarded. The input event rate is reduced at each level, providing increasing time for more complex and accurate reconstruction tasks.

The first level of the trigger, Level 1, rejects the vast majority of the events. A period of 5.5 μs is allowed for Level 1 to reach its decision. The necessary delay is attained through pipeline storage of the readout data available at the front-end electronics. The buffers are 14 (396 ns) bunch crossings deep, and for every Tevatron clock cycle, the event is moved up one slot in the pipeline. By the time it reaches the end of the pipeline the trigger will have reached a decision whether to accept or reject the event. The Level 1 decision is based on only a subset of the detector parts along with quick pattern recognition and filtering algorithms. The original 2.5 MHz event rate is reduced at Level 1 to less than 50 KHz. Events accepted by Level 1 trigger are moved to one of four on-board Level 2 buffers. The second trigger level performs a limited event reconstruction through a more careful analysis of the readout information. It also takes longer, about 20 – 30 μs, to reach a decision. To ensure maximal speed, both Level 1 and Level 2 triggering mechanisms are implemented with custom designed electronics. The DAQ system allows the Level 2 trigger to accept as many as 300 events per second. Following a Level 2 accept, the various detector data fragments are assembled together by the Event Builder (EVB) system to be transferred to the last trigger stage. The third level of the trigger, Level 3, is implemented as a PC computing farm. Parallel event processing by many nodes allows for nearly 1 s to be allocated for the trigger decision, permitting a fairly thorough event reconstruction. An event rate below 75 Hz is achieved at the end of the final trigger level.

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a *trigger path.* The CDF trigger system implements about an hundred trigger paths, which compete for DAQ bandwidth. An event will be accepted if and only if it passes the requirements of any one of these paths. Events accepted by the trigger get written to a mass storage device.

All data manipulations performed from then on are referred to as data handling. These include offline reconstruction of physics objects, such as tracks, vertices, electrons, muons, jets, *etc*, suitable for data analysis. This stage, denominated *production*,

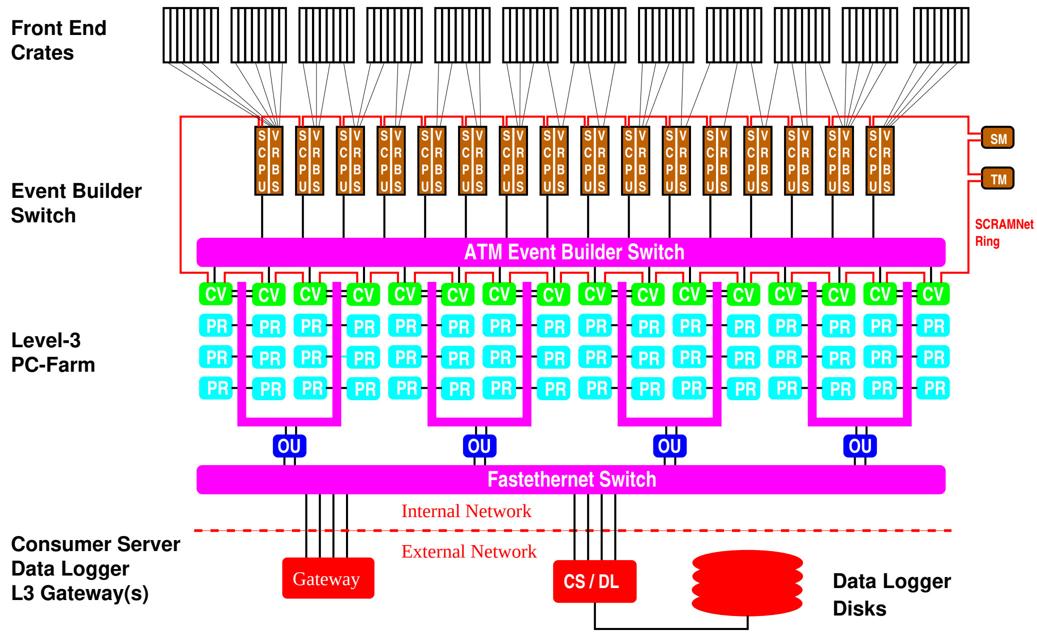


Figure 2-9: Representatio of the DAQ system layout: event building and Level 3 filtering.

involves similar operations to those performed at Level 3, only with more elaborate algorithms, and using the most accurate and up-to-date detector calibrations and alignments.

2.3.1 Data flow

The DAQ system is responsible for collecting the data fragments from the front-end electronics systems, for events satisfying the trigger criteria, up to the final stage in which complete event records are formed and sent to mass storage.

The front-end and trigger electronics are packaged as VME modules that reside in about 120 crates in the system. These modules process detector signals and make the data available to the DAQ system through the VME bus. Each front-end crate contains at least one processor board for hardware initialization and event readout, running under the VxWorks operating system. The crates contain also a *controller* module which distributes, through the VME backplane, received timing synchronization signals. These timing signals are dictated by the Tevatron clock, coherently with the bunch crossings, and ensure global synchronization of readout electronics and trigger.

The controller module further provides the interface of the VME modules to the Trigger System Interface (TSI), as well as the data interface to the EVB. The TSI is responsible for receiving the decisions from the hardware triggers (Level 1 and Level 2), communicating them to the front-end crates, and supervising the data flow up until it is transferred to the EVB.

Event building and processing farm

Upon Level 2 accept, the data from the various front-end crates are collected and transmitted to the Level 3 trigger by the event builder system. The EVB and the Level 3 form a crucial and intricate system of fast data transmission and control networks, and processing units, whose structure is outlined in Figure 2-9.

The data are first received by the VME Readout Boards (VRB). Each VBR is fiber linked to a group of front-end crates. The VRB are themselves part of 15 EVB crates, each of which is controlled by a single board processing unit referred to as "Scanner CPU" (SCPU) running VxWorks. The 15 EVB crates are connected to 16 *converter* node PCs of the Level 3 farm through an Asynchronous Transfer Mode (ATM) network switch for data transfer. Data flow between the SCPUs and converter nodes is controlled by the Scanner Manager (SM). The SM is a task running in a processing unit located at an additional EVB crate, and constitutes the EVB interface with the TSI. Communication among the SCPUs, converter nodes and TSI is performed over a serial-ring reflective-memory control network denominated SCRAMNet ("shared common random access memory network").

Event building and data transfer from the front-end electronics to the Level 3 involve the following stages. Front-end processors complete data readout and, via the controller module, load data to the VRBs. The TSI passes the Level 2 accept message to the SM. The SM instructs the SCPUs to read in and combine the event fragments from all VRBs in their local crates. SCPUs acknowledge loading completion to the SM. Meanwhile, converter nodes if ready make their status known to the SM. The SM selects a single converter node among those which reported themselves available. SCPUs are directed by the SM to transfer event fragments, through the ATM switch, to the selected converter node. The full *event record* thus assembled at the converter node constitutes from this point on the single and only piece of information about a particular event.

The Level 3 farm is formed of roughly three hundred dual-processor PCs running the Linux operating system. The structure of the farm is also represented in Figure 2-9. It is arranged in 16 sub-farms, each containing a converter node and several (12 to 16) *processor* nodes; there are also 8 *output* nodes, each shared by two sub-farms. The event assembled at a given converter node is sent via a fast Ethernet connection to one of the processor nodes of the associated sub-farm. There the event passes through a trigger "filter" executable to achieve near-final quality reconstruction on which the Level 3 decision is based. If the executable decides to accept an event, it is then passed to the output nodes of the farm. These nodes send the event onward to the Consumer Server / Data Logger (CS/DL) system for storage first on disk, and later on tape.

A detailed description of the implementation and operation of the EVB and Level 3 systems, along with the integration in the full DAQ system, is provided in [45].

2.3.2 Real-time event selection

The structure of the trigger system is illustrated in Figure 2-10.

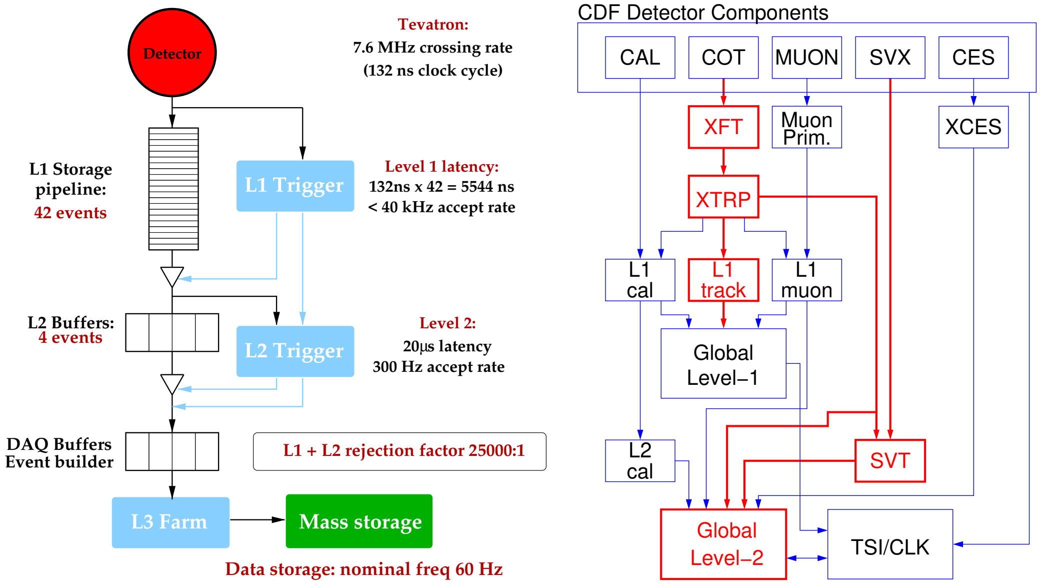


Figure 2-10: Block diagram of the CDF DAQ and trigger systems.

Level 1

The Level 1 trigger is a synchronous system, providing an accept / reject decision every bunch crossing, using a fully pipelined architecture. The Level 1 decision is based on a limited amount of event information from the COT, calorimeters and muon chambers, which is used to form coarse versions of physics objects (such as tracks, electrons and muons) referred to as trigger *primitives*.

Track primitives are identified by the Extremely Fast Tracker (XFT) [46]. A rough measurement of the transverse momentum and azimuthal direction of a track is formed by using the hit information of the 4 axial superlayers of the COT. The identification process involves three stages: hit classification, segment finding and segment linking. Each wire hit is classified as either "prompt", if the drift time is lower than 66 ns, or as "delayed", if the drift time ranges from 66 ns to 220 ns. Track segments are searched for in each axial superlayer by grouping together adjacent COT cells in groups of 4 and by comparing the hit data to a list of pre-loaded patterns. The patterns vary depending on the combination of prompt and delayed hits and the track angle through the cell or track p_T. Each found segment is characterized by its mean ϕ position and slope information. Once track segments have been found, sets of four, each belonging to a different superlayer, are compared against a list of templates in a $\Delta\phi$ window of 1.25°. Matches are searched for which correspond to valid tracks with $p_T \geq 1.5$ GeV/c and constrained to the beamline. If no track is found, the search is performed on segments in the innermost three axial superlayers.

Finally, the XFT reports the tracks p_T, ϕ_6 (ϕ at the superlayer 6) and charge. The achieved XFT track resolution is $\sigma(p_T) = 1.7\%/\text{GeV/c}$ and $\sigma_{\phi_0} = 5$ mrad, only about a factor of 10 coarser than the offline reconstruction. If more than 6 tracks are found an automatic Level 1 accept is generated; otherwise the decision is dependent on the specific trigger requirements on p_T and ϕ accordingly coded in look-up tables.

The XFT tracks are extrapolated to the calorimeter and muon detector systems via the Extrapolator unit (XRTP) for matching with electron towers and muon stubs. The calorimeter trigger is based both on object primitives (electrons, photons and jets) and global event variables (total transverse energy $\sum E_T$ and missing transverse energy $\not{E}_T$). Primitives are formed by applying thresholds to individual calorimeter trigger towers. Transverse energy (E_T) is found by summing the digitized calorimeter data into trigger towers weighted by $\sin\theta$. Muon and dimuon primitives are derived from hits in the muon chambers or coincidences of hits with the scintillators. Track matching is performed at a granularity of about 1.5°.

Level 2

The Level 2 trigger is an asynchronous system which processes events accepted by Level 1 based on better precision and additional event information. Level 2 is implemented in two stages, with a latency of about $10 - 20\ \mu s$ each, respectively achieved by primitive building hardware and by programmable processors (Alpha) which operate on the primitive inputs to perform further event topology characterization and impose tighter requirements.

The Silicon Vertex Trigger (SVT) [47] processor uses the $r{-}\phi$ hits in the silicon detectors to extend the XFT track primitives inside the SVX volume. It provides a measurement of the track impact parameter d_0, while also improving the determination of p_T and ϕ_0. It thus enables the trigger to distinguish primary and secondary particles, and hence to collect, in particular, large samples rich in heavy flavor decays.

The SVT structure reflects the SVX detector's 12-fold azimuthal symmetry and 3 barrel segmentation, and does tracking separately for each ϕ sector and barrel. An SVT track candidate requires coincidence of an XFT track and hits in four (out of the five available) silicon layers. The XFT tracks are swum into the SVX detector region, forming "roads" within which clusters of charge in the silicon layers have to be found. The SVT uses a list of pre-loaded patterns for finding the best coincidences. The information about the XFT outer-track and the four silicon hits is finally fed into a linearized fitter which returns the measurements of p_T, ϕ_0 and d_0 for the track.

The average SVT processing time is about 19 μs as indicated in Figure 2-11. The SVT impact parameter resolution for tracks with $p_T > 2$ GeV/c is about 35 μm, comparable to that of offline tracks which do not use Layer 00 information (which is not yet available in the SVT). This is shown in Figure 2-11; the width of the Gaussian fit shown for the tracks impact parameter distribution is 47 μm, which is the combination of the intrinsic SVT impact parameter resolution and the transverse size of the beam profile (about 30 μm).

The calorimeter trigger is also improved at Level 2. Clustering of trigger towers is performed by applying seed and shoulder thresholds, which is used to form jets and reduce corresponding rates. The shower maximum strip chambers (CES) are further used to refine the track matching, and reduce the fake electron and photon rates.

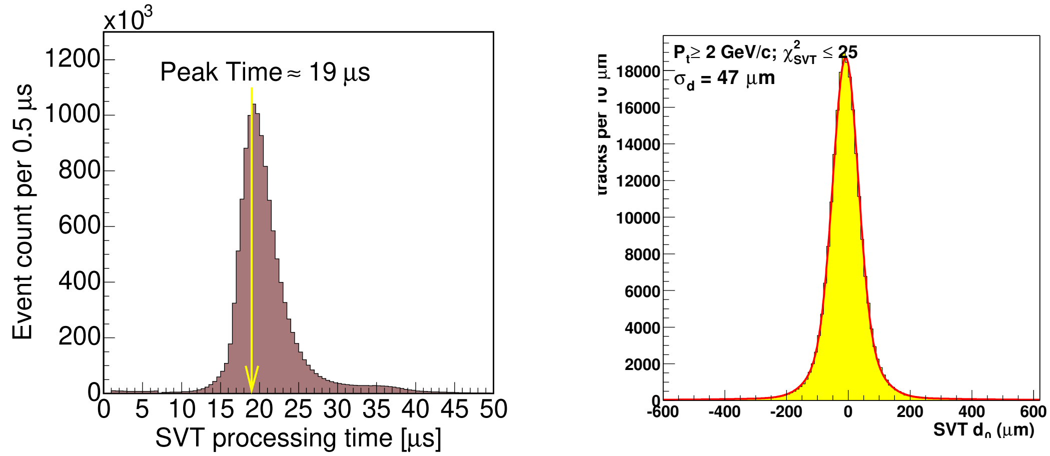


Figure 2-11: SVT processing time (left) and track impact parameter resolution (right).

Level 3

At Level 3 nearly offline quality of event reconstruction is achieved, taking advantage of the complete event record and improved resolutions not available to the lower levels. Tight matching of tracks to calorimeter and muon system information is implemented. Three dimensional track reconstruction becomes available for the first time. Vertices may be explicitly reconstructed and L_{xy} extracted. Invariant masses of associated particle candidates, for instance, may also be used as trigger criteria. Results from the previous levels may be used to drive the algorithms.

Chapter 3

Data samples selection and composition

The data samples used in this dissertation were recorded with the CDF detector in the period from February 2002 to August 2004. These correspond to an integrated luminosity of approximately 355 pb^{-1} after imposed quality requirements, *i.e.* taken with the full detector systems properly functioning.

The following B meson decay modes are reconstructed,

- $B^+ \to J/\psi K^+$,
  $B^+ \to \bar{D}^0\pi^+$, $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$,
  $B^+ \to \bar{D}^0 l^+\nu$,
- $B^0 \to J/\psi K^{*0}$,
  $B^0 \to D^-\pi^+$, $B^0 \to D^-\pi^+\pi^-\pi^+$, $B^0 \to D^{*-}\pi^+$, $B^0 \to D^{*-}\pi^+\pi^-\pi^+$,
  $B^0 \to D^- l^+\nu$, $B^0 \to D^{*-} l^+\nu$,
- $B_s \to D_s^-\pi^+$, $B_s \to D_s^-\pi^+\pi^-\pi^+$,
  $B_s \to D_s^- l^+\nu$,

where the lepton l stands for electron or muon, and the involved charm and strange mesons are reconstructed in the following channels,

- $D^- \to K^+\pi^-\pi^-$,
- $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$,
- $D^{*-} \to \bar{D}^0\pi^-$,
- $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to \pi^+\pi^-\pi^-$,
- $\phi \to K^+K^-$,
- $K^{*0} \to K^+\pi^-$,
- $J/\psi \to \mu^+\mu^-$.

Charge-conjugated states are implicitly assumed throughout the text. The full list of decay chains is shown explicitly in Table 3.1.

The B meson decays studied fall into two categories: fully reconstructed and partially reconstructed. The former category corresponds to the $B \to J/\psi K$ and $B \to D\pi(\pi\pi)$ modes. The latter class is realized by the $B \to Dl\nu$ decays, where the neutrino and possibly other decay products are missed. The resulting samples will be sometimes also referred to as *hadronic* and as *semileptonic*, respectively.

The semileptonic modes provide relatively larger samples of B mesons. The full reconstruction which characterizes the hadronic modes, however, is translated in a more accurate determination of the candidates' proper decay time, which will reveal an important effect especially in the analysis of B_s oscillations.

In this chapter we give an outline of the trigger strategies used to collect the data samples, and describe the criteria employed to select the signal candidates, corresponding to the above mentioned decays. The accurate understanding of the selected samples will benefit from Monte Carlo simulation of the involved signals, including detector and trigger selection effects. Such Monte Carlo samples will be employed in the process of signal optimization, as well as to assess the samples composition.

3.1 Trigger requirements

The data are acquired through trigger requirements, imposed *online*, *i.e.* in real-time, by the data acquisition system.

Muons allow for a clean identification, and are used to collect samples of J/ψ mesons, partially fed by B decays. The possibility of triggering on events displaced from the interaction point, by exploring the long lifetime of b-hadrons, is made possible for the first time at a hadron detector by the CDF Silicon Vertex Trigger (SVT) processor. Such criteria for selecting events formed of displaced tracks are used in the other two of the three trigger paths employed, which are described next.

Di-muon trigger

The B decays involving a J/ψ meson in the final state are selected by identifying a pair of muon candidates, with opposite charges, and a mass, $M_{\mu\mu}$, close to the J/ψ nominal mass. The specific requirements are summarized by the following:

A. Level 1

- two XFT tracks with opposite charge,
- each track is matched with two muon stubs,
- each CMU(CMX) muon has $p_T^{\mathrm{XFT}} > 1.5\,(2.2)$ GeV/c,
- $\Delta\phi_6(\mathrm{CMU}, \mathrm{CMU}) < 135°$, no cut in $\Delta\phi_6(\mathrm{CMU}, \mathrm{CMX})$,

B. Level 3

- $2.7 < M_{\mu\mu} < 4$ GeV/c^2,

$B^+ \to J/\psi K^+$,	$J/\psi \to \mu^+\mu^-$
$B^0 \to J/\psi K^{*0}$,	$J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$
$B^+ \to \bar{D}^0\pi^+$,	$\bar{D}^0 \to K^+\pi^-$
$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$,	$\bar{D}^0 \to K^+\pi^-$
$B^0 \to D^-\pi^+$,	$D^- \to K^+\pi^-\pi^-$
$B^0 \to D^-\pi^+\pi^-\pi^+$,	$D^- \to K^+\pi^-\pi^-$
$B^0 \to D^{*-}\pi^+$,	$D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-$
$B^0 \to D^{*-}\pi^+$,	$D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$,	$D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-$
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$,	$D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$
$B_s \to D_s^-\pi^+$,	$D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$
$B_s \to D_s^-\pi^+$,	$D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$
$B_s \to D_s^-\pi^+$,	$D_s^- \to \pi^-\pi^+\pi^-$
$B_s \to D_s^-\pi^+\pi^-\pi^+$,	$D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$
$B_s \to D_s^-\pi^+\pi^-\pi^+$,	$D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$
$B^{+,0} \to \bar{D}^0 l^+ X$,	$\bar{D}^0 \to K^+\pi^-$
$B^{+,0} \to D^- l^+ X$,	$D^- \to K^+\pi^-\pi^-$
$B^{+,0} \to D^{*-} l^+ X$,	$D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-$
$B_s \to D_s^- l^+ X$,	$D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$
$B_s \to D_s^- l^+ X$,	$D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$
$B_s \to D_s^- l^+ X$,	$D_s^- \to \pi^-\pi^+\pi^-$

Table 3.1: The full list of decay channels analyzed.

where p_T^{XFT} is the track transverse momentum as measured by XFT, and $\Delta\phi_6$ is the opening angle at the COT superlayer 6.

Two displaced-tracks trigger

The trigger criteria include requiring two displaced tracks forming a displaced vertex. The specific requirements are summarized as follows:

A. Level 1

- two XFT tracks with opposite charges,
- each track has $p_T^{\mathrm{XFT}} > 2.04$ GeV/c,
- $\sum p_T^{\mathrm{XFT}} > 5.5$ GeV/c,
- $0° < \Delta\phi_6 < 135°$,

B. Level 2

- for each track, 100 μm $\leq |d_0^{\mathrm{SVT}}| \leq 1$ mm,
- each track has $p_T^{\mathrm{SVT}} > 2$ GeV/c,
- $\sum p_T^{\mathrm{SVT}} > 5.5$ GeV/c,
- $2° < \Delta\phi_0 < 90°$,

C. Level 3

- $|\Delta z_0| < 5$ cm,
- two-track vertex $L_{xy} > 200$ μm,

where d_0^{SVT} is the impact parameter as measured in the SVT, Δz_0 is the distance between the two tracks along the beam axis, and L_{xy} is the distance in the transverse plane of the two-track vertex with respect to the primary vertex.

Lepton and displaced-track trigger

The strategy is to identify semileptonic B decays by requiring the presence of a lepton with large transverse momentum together with a displaced track. The specific requirements are summarized as follows:

A. Level 1

for the lepton:

- track with $p_T^{\mathrm{XFT}} > 4$ GeV/c,
- $E_T > 4$ GeV, $E_{\mathrm{HAD}}/E_{\mathrm{EM}} < 0.125$, or
  track with stubs in both CMU and CMP,

for the displaced-track:

- $p_T^{\mathrm{XFT}} > 4.09$ GeV/c,
- $0° < \Delta\phi_6 < 100°$,

B. Level 2

for the displaced-track:

- 120 μm $\leq |d_0^{\mathrm{SVT}}| \leq 1$ mm,
- $p_T^{\mathrm{SVT}} > 2$ GeV/c,

C. Level 3

for the lepton displaced-track pair:

- $2° < \Delta\phi < 90°$,
- mass < 5 GeV/c^2 .

These trigger samples contain large numbers of semileptonic B meson decays. These are primarily used in an inclusive fashion for the study and optimization of opposite-side flavor tagging methods. Semi-exclusive semileptonic signals may additionally be extracted, although this is not done in the context of this dissertation.

3.2 Event reconstruction

The data samples collected according to the specified trigger criteria undergo further manipulations, which lead to precise and complete reconstruction of physics objects suitable for data analysis. During the production stage more accurate detector calibration and alignment information is available than was used during data taking. Tracks, which are reconstructed at CDF using several algorithms [48], form the basis for the reconstruction of B meson decay candidates.

3.2.1 Track selection and preparation

Tracks are used to reconstruct decaying particles' vertices, the resolution of which is a determining factor for the measurements to be performed. This renders the use of well-measured tracks further imperative. It is therefore important to ensure that tracks which are used for reconstructing candidates meet standards for good quality in each of the detector sub-systems. Furthermore, all tracks are refit in order to achieve improved determination of track parameters and associated uncertainties.

Mis-alignments together with hits caused by noise in the tracking system produce fake and mis-measured tracks in the events. The number of such occurrences are decreased by requiring for the track a minimum number of hits in the drift chamber and silicon detector. Specifically this track quality criteria involve requiring at least 10 axial and 10 stereo COT hits, 3 SVX hits, and a minimum track momentum of 350 MeV/c.

Selected tracks are required to have a COT parent, with a physical covariance matrix, and a helix fit attached. The ascribed uncertainties to the COT hit positions, which do not take into account the effect of multiple scattering in the material of the COT system, determine an underestimation of the track parameters returned by the track fit. In order to correct for COT resolution effects which are not well understood, the covariance matrix of the COT track is rescaled with the following empirical scaling factors [49] (with p_T in GeV/c):

$$\begin{aligned}
s(\lambda) &= \sqrt{1+p_\lambda(1+\lambda^2)^{1.5}/p_T^2} && \text{with } p_\lambda = 0.58;\\
s(C) &= \sqrt{1+p_C/p_T^2} && \text{with } p_C = 5.33;\\
s(z_0) &= \sqrt{1+p_{z_0}(1+\lambda^2)^{1.5}/p_T^2} && \text{with } p_{z_0} = 0.653;\\
s(d_0) &= \sqrt{1+p_{d_0}/p_T^2} && \text{with } p_{d_0} = 3.01;\\
s(\phi_0) &= \sqrt{1+p_{\phi_0}/p_T^2} && \text{with } p_{\phi_0} = 3.7.
\end{aligned}$$

Here $\lambda = \cot\theta$, with θ being the polar angle of the track; the signed curvature of the track is referred to as C, with the magnitude given by the inverse of the trajectory diameter; the parameters z_0, ϕ_0, and d_0 are the coordinates of the point of closest approach to the beamline. The factors are applied to the covariance matrix as $c_{ij}^{\text{new}} = s_i s_j c_{ij}^{\text{old}}$, where the indices i and j represent the various helix parameters.

The rescaled track is finally refit with the addition of the silicon hits. For properly positioning the silicon hits an alignement table is used. When the data gets officially reconstructed the final alignement table is usually not yet available. Therefore we use here the most up-to-date version, for optimal resolutions. If available, hits on the Layer 00 silicon subsystem are also included in the track refit procedure. The inclusion of the additional constraints provided by this innermost silicon detector system on the particle tracks results in improvements in vertexing resolution. This translates eventually in $\sim 10\%$ improvements in decay length resolution. The refitting of the track takes into account corrections to energy loss in the detector material for kaon, pion, electron or muon hypotheses according to the requested species. The description of the passive material settings is based on [50].

The location of the primary interaction vertex of the event, where the b hadron was produced, needs to be determined. A good approximation of its location is the beam spot. We use an improved determination which is provided by an iteractive procedure [51]. A set of loosely selected tracks is combined to form a vertex, and at each fit iteration the track with worst χ^2 is eliminated. The procedure is repeated until no track has a χ^2 larger than a maximum allowed value. The tracks belonging to the reconstructed B meson decay candidate are excluded from the procedure.

3.2.2 Pre-selection and BStntuple

Trigger confirmation

The confirmation of the trigger requirements is important given that there are differences between the online and offline reconstruction tracks. For example, the SVT fit does not use the exact same algorithm and information which characterizes the tracks obtained from production. Connecting these production tracks with the online trigger information avoids considering so-called volunteer candidates. These are such candidates for which not all trigger tracks correspond to B decay products. These may arise, for example, when an SVT trigger track is provided by a fake XFT track, which got several accidental hits assigned and has thus high probability to have a large impact parameter; or when the lepton or a SVT track belong in fact to the other B hadron in the event.

In order to ensure that we have a displaced SVT track a matching is performed between the production and SVT tracks. The SVT matching algorithm compares the curvature and the ϕ of the two tracks, and requires for the quadratic sum of the normalized distance in the two parameters between the tracks a maximum default value. Differences in ϕ, for example, may also to some extent have origin in the fact that different silicon alignments are used for SVT and offline reconstruction. Only SVT tracks which have an SVT χ^2 larger than 25 are considered. To promote a production level track to being matched to an online SVT trigger track we further require that the SVT measurements of p_T and d_0 comply with $p_T > 2$ GeV/c and 0.012 cm $< |d_0| < 0.100$ cm.

General pre-selection

Candidates for each particle decay tree are constructed from the bottom up. For example, in the case of $B \to l^+ D_s^- X (D_s^- \to \phi\pi^-, \phi \to K^+K^-)$ we start with a $\phi \to K^+K^-$ decay, which in the next step is used to make candidates for the $D_s^- \to \phi\pi^-$ decay and finally those are used to form $B \to l^+ D_s^- X$ decays. In each step the candidates are subjected to a set of selection requirements which is chosen to quickly reject uninteresting candidates. Charge correlations among the tracks are applied first. Candidates with duplicate tracks are removed. The trigger is confirmed when applicable: for $D_{(s)}$ candidates one matched SVT trigger track is required in the case of semileptonic modes; for the hadronic B decays, two of the tracks forming the B candidate are required to be matched to an SVT trigger track pair, and these may belong to any child of the B meson: D_s^-, D^-, D^0, D^{*-}, $\pi(\pi\pi)$ or any combination thereof.

The raw mass is calculated based on the fully corrected track momentum and candidates are only accepted within loose mass windows. A vertex fit [42] is performed to the decay topology, and the resulting χ^2 in the r-ϕ plane, $\chi^2_{r\phi}$, the fitted mass, and the two-dimensional decay length, L_{xy}, are used to reject candidates. Cuts are only applied if tighter cuts are used later in the analysis.

BStntuple

We have implemented a framework, here referred to as the `BStntuple` [52], for efficiently storing and accessing the selected data information which forms the B mesons candidates. It shares the basic structure of the standard `stntuple` [53], which itself constitutes, in practice, a more sophisticated `ntuple` together with a set of convenient tools; its format shares features of both micro-DST and `Root` files. It contains structures to hold the reconstructed candidates information (stable and decaying objects) as well as taggers' information (decision, raw dilution), and particle identification information (TOF, dE/dx, muon and electron quantities). The actual data blocks correspond to instances of these classes for specific decay modes, which are implemented by cloning prototype modules appropriate for the decay topology. This has revealed to be an efficient and uniforming framework, in terms of both CPU usage time and procedure sharing, which has facilitated candidate reconstruction and the ntupling process for the various decays and data samples employed in the analysis.

3.2.3 Selection optimization

In deciding on the requirements for selecting B candidates which will form our samples the goal is to optimize the samples' effective statistical significance in view of the measurements which are to be performed.

The significance of a sample for a mixing analysis is estimated via (B.17), and in principle this full quantity should be considered as optimization criteria. This includes a quantification of the signal yield and purity, through the variable

$$\frac{N_S}{\sqrt{N_S + N_B}}\,, \tag{3.1}$$

where N_S and N_B are the number of selected signal and background events, respectively. This is the main criteria for sample selection optimization which has been employed.

In addition, it is possible to appropriately weight signal events according to their proper time resolution, by including the following multiplicative term

$$\frac{1}{N_S}\sum_{i=1}^{N_S} e^{-(w\sigma_i)^2}\,, \tag{3.2}$$

where σ_i is the event resolution, and w stands for a value in the vicinity of the mixing frequency being probed. These additional effects are relevant mostly for the study of flavor oscillations in samples of B_s mesons. Such a correction to the optimization criteria has in fact been implemented for the semileptonic B_s samples, while no appreciable requirement modifications have been induced. Contributions from flavor tagging, which would depend on algorithm specifics, to the effective significance (B.17), are not considered in the selection process.

The optimization procedure is performed after a set of variables with appropriate discriminating power between signal and background is found. For each set of cut values of the variables being optimized, the quantity to be maximized, which we have addressed above, is evaluated. The effect of the selection requirements on the background is evaluated using the data sample itself, while for signal Monte Carlo simulated events are employed, in order to ensure that cuts are not biased. The number of background events is estimated by applying the selection requirements to a sample of the data where no signal events are present ("sideband"). The mass of the sideband events is fit with a linear ($J/\psi K$ and Dl modes) or an exponential ($D\pi(\pi\pi)$ modes) model, which is extrapolated to and integrated in the signal region. The number of signal events is determined from the Monte Carlo sample, after this is rescaled to the number of events found in the mass peak in data before optimization.

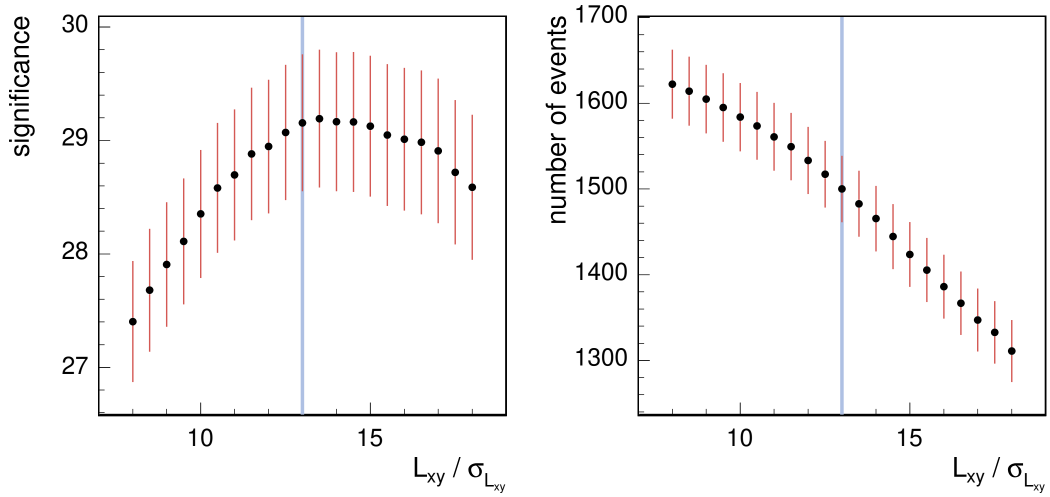


Figure 3-1: Illustration of selection optimization procedure applied to $L_{xy}/\sigma_{L_{xy}}$: significance (left) and efficiency (right).

The procedure is iterative. At a given iteration, a set of best values is found for the set of variables being optimized – each variable is varied at a time while all others are kept fixed, and that value which maximizes the significance is found. The procedure is repeated until convergence is observed. The optimization process is illustrated in Figure 3-1 for the cut $L_{xy}/\sigma_{L_{xy}}$. Besides the significance being maximized, shown on the left, the other variable which is monitored in the process is the analysis efficiency, shown on the right hand side figure. When a set of points is found with similar significance, the point with highest efficiency is chosen.

3.2.4 Reconstruction of B decays

Variables

Several variables have been identified which are useful for reducing the combinatorial background levels of our samples. A selection based on these variables is optimized, as described above, for each of the decay channels reconstructed, in order to achieve maximal signal significance.

B mesons are relatively long-lived particles thus producing measurable decay distances in the detector. That is, the B decay vertex tends to be significantly displaced from the primary interaction point. Additionally, D mesons have themselves a finite (though somewhat smaller) lifetime as well, making their decay vertex displaced with respect to that of the parent B. The selection criteria attempt therefore to select the resulting decay topologies, which are furthermore kinematically consistent with the particle hypotheses forming the studied decay trees.

Various variables are available for implementing selection criteria which allow to statistically identify the signal decays with such characteristics from backgrounds produced by other, random track combinations. A list of some of the most common follow. The quality of fits to tracks forming vertices of B and D meson candidates is given by the $\chi^2_{r\phi}$ in the transverse r-ϕ plane or the three dimensional vertex fit probability. The transverse distance, L_{xy}, of B and D meson decay vertices relative to the primary interaction vertex, or the transverse decay distance of the D meson relative to the B meson, $L_{xy}(B \leftarrow D)$, as well as the transverse decay distance uncertainty, $\sigma_{L_{xy}}$, are used to identify displaced objects. The transverse momenta, p_T, of particle candidates are widely useful. So are the masses associated to vertices, which are provided by the vertex fit, and to track systems such as that of a D-lepton pairs m_{Dl}, and mass differences such as $m(D^*) - m(D^0)$, which also provide powerfull criteria.

Building blocks

The $J/\psi \to \mu^+\mu^-$ decay candidates are formed from pairs of trigger muons and are required to have a mass within ± 80 MeV/c^2 of the world average J/ψ mass.

The $K^{*0} \to K^+\pi^-$ candidates are reconstructed from track pair combinations that fit into a common vertex. They are required to fall into a $\pm(50-60)$ MeV/c^2 mass window about the K^{*0} world average mass. Due to the broadness of the K^{*0} resonance and the absence of particle identification in the selection analysis, it is not uncommon for a track pair to have both possible mass assignments $K^+\pi^-$ and π^+K^- found in the allowed mass region for the K^{*0}. No attempt is made to discriminate between the true and the swapped K^{*0} candidates, and all combinations are accepted. While this ambiguity is removed in the case of the $D_s^- \to K^{*0}K^-$ decay, through charge correlation, the fraction of signal $B^0 \to J/\psi K^{*0}$ candidates formed of the swapped mass assignment is estimated in Section 3.3 from Monte Carlo predictions.

The ϕ meson candidates are reconstructed from a pair of oppositely charged kaons, which is required to have a mass consistent with the ϕ nominal world average mass value, in a window of ± 12 MeV/c^2.

The decays of the charmed mesons $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ are reconstructed by combining two, four and three tracks, respectively, and applying the standard vertex fit. A loose mass cut is applied to the D^0 and D^+ candidates by requiring that they be in the range of $(1.77, 1.97)$ GeV/c^2. The cuts for the D vertex quality vary depending on the decay channel of the parent B. In the semileptonic samples the D^0 meson is reconstructed in the lower multiplicity mode only.

The $D^{*-} \to \bar{D}^0\pi^-$ meson candidates are reconstructed by combining the D^0 with appropriately charged pion candidate tracks. Both $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^-\pi^+$ collections are used for the D^{*-} reconstruction in hadronic samples, while only the former is employed in the semileptonic samples. The reconstructed mass of the D^{*-} is required to be within ± 80 MeV/c^2 of the PDG value. Subsequently, a cut on $\Delta M_{D^*} \equiv m_{D^*} - m_{D^0}$ is applied. In the mass difference calculation, the fitted masses m_{D^*} and m_{D^0} are used; the track 4-momenta of the D^0 daughters are taken from the stand-alone D^0 mass fit without the D^0 mass constraint. The momentum of the slow pion from D^* is taken from the fit of the D^*; for D^*l decays, this momentum should have a magnitude of at least 0.4 GeV/c.

In any triple-pion decay mode such as $B^0 \to D^-\pi^+\pi^-\pi^+$ or $B^0 \to D^{*-}\pi^+\pi^-\pi^+$, there are three pion tracks which come directly from the B decay vertex. Triple pion combinations are formed from all eligible tracks in an event and are fitted to have a common origin using the standard vertex fit. Vertex quality requirements are applied on these vertices. In addition, an invariant mass cut allows for significant reduction of otherwise overwhelming combinatorics from random tracks in the event. For a triple pion combination to come from a B meson in $B \to D\pi\pi\pi$ decay, it has to have a mass no larger than m_B-m_D. The chosen cuts on the $\pi\pi\pi$ vertex quality and mass are shown in Table 3.3.

The D_s meson candidates are reconstructed in three final states: $\phi\pi$, K^*K, and $K\pi\pi$. The requirement on the invariant mass corresponds to the range of $(1.90, 2.04)$ GeV/c^2. The possibility of having background coming from reflections under this D_s mass region has been investigated, by modifying the mass assignment of the particles and checking the different mass distributions. The $D_s \to K^*K$ case has a substantial reflection from $D \to K^*\pi$, $K^* \to K\pi$ and $D \to K\pi\pi$ (non-resonant) decays. The requirement $|m(K\pi\pi) - m_{PDG}^{D^+}| < 20 - 24$ MeV/c^2 while having a 60% efficiency for signal reduces that background to negligible levels. In the $D_s \to \pi\pi\pi$ channel there is a large contribution coming from $D^* \to D^0\pi$, $D^0 \to K\pi\pi$ decays; requiring $m(K\pi\pi) - m(K\pi) > 150$ MeV/c^2 was shown to reduce to very low levels the D^* background, with essentially no loss of signal candidates. In the case of $D_s \to \phi\pi$, due to the narrow ϕ mass window, no substantial reflections are observed. The cuts used for reconstructing D_s in the semileptonic samples are shown in Table 3.5. The helicity angle ψ is calculated in the ϕ or K^{*0} rest frame, and is defined as the angle between the momentum of the D_s candidate and the direction of the two decay products, e.g. $\phi \to K^+K^-$.

For selecting the lepton tracks from B decays we use lepton identification likelihood techniques, which combine related variables from various detector components. These coincide with those developed for the soft lepton taggers [54, 55]. The transverse momentum threshold for the lepton identification has been set to $p_T \geq 1.5$ GeV/c. For electrons, the likelihood is required to be larger than 0.9. In case of muon identification the cut value depends on the muon system: 0.05 (CMUP), 0.50 (CMU,CMP,CMX), and 0.70 (IMU).

B meson reconstruction

B meson candidates are formed from the building blocks constructed above. Vertex fits are applied to the resulting combinations, according to the reconstructed decay channels. The masses of the J/ψ, D^0, D^-, D_s^- meson candidates are constrained in these fits to the corresponding world average mass values. No pointing constraints are imposed for either J/ψ, D or B mesons.

The decay candidates $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ are formed by pairing the J/ψ candidates with charged tracks and with K^{*0}. The vertex fit is performed on the three and four track combinations for the B^+ and B^0 candidates, respectively. The kinematic and quality cuts applied to the resulting candidates are listed in Table 3.2.

decays	$B^+ \to J/\psi K^+$	$B^0 \to J/\psi K^{*0}$
B vertex probability $>$	10^{-3}	10^{-4}
$p_T(B)$ [GeV/c] $>$	5.5	6.0
$p_T(K^{*0,+})$ [GeV/c] $>$	1.6	2.6
$\sigma_{\rm ct}$ [μm] $<$	150	150

Table 3.2: Summary of reconstruction requirements for $J/\psi K$ samples.

The list of $B \to D^{(*)}\pi(\pi\pi)$ modes is found in Table 3.1. For each newly formed candidate, the full vertex fit is performed. A number of (optimized) kinematic and vertex quality cuts are then applied to reduce the combinatorial background; these are specified in Table 3.3.

		B^+ modes		B^0 modes			
		$\bar{D}^0\pi^+$	$\bar{D}^0\pi^+\pi^-\pi^+$	$D^-\pi^+$	$D^-\pi^+\pi^-\pi^+$	$D^{*-}\pi^+$	$D^{*-}\pi^+\pi^-\pi^+$
$\chi^2_{r\phi}(B)$	$<$	15	8	15	11	17	20
$\chi^2_{r\phi}(D)$	$<$	15	4	15	15	16	16
$\chi^2_{r\phi}(\pi\pi\pi)$	$<$	—	6	—	—	—	—
$\sigma_{L_{xy}}(B)$ [μm]	$<$	—	—	—	—	200	200
$\frac{L_{xy}}{\sigma}(B)$	$>$	8	15	11	14	4.5	8
$\frac{L_{xy}}{\sigma}(D)$	$>$	—	10	—	14	3	6
$L_{xy}(B \to D)$ [μm]	$>$	-150	0	-300	0	—	—
$\lvert d_0(B)\rvert$ [μm]	$<$	80	55	110	70	180	180
$p_T(B)$ [GeV/c^2]	$>$	5.5	8.5	5.5	7.0	4.0	9.0
$p_T(\pi_B)$ [GeV/c^2]	$>$	1.0	—	1.2	—	1.5	0.4
$p_T(\pi_{soft})$ [GeV/c^2]	$>$	—	—	—	—	0.4	0.4
$m_{3\pi}$, GeV/c^2	$<$	—	1.75	—	—	—	2.3
$\lvert\Delta M_{D^*} - 145.4\rvert$ [MeV/c^2]	$<$	—	—	—	—	2.5	2.5
$\Delta R(D, \pi_B)$	$<$	2.0	—	1.5	—	—	—

Table 3.3: Summary of reconstruction requirements for $D^{(*)}\pi(\pi\pi)$ samples; D^* decays with both $D^0 \to K\pi$ and $D^0 \to K\pi\pi\pi$ are included.

The reconstructed hadronic B_s channels include $B_s \to D_s^-\pi^+$ and $B_s \to D_s^-\pi^+\pi^-\pi^+$ with D_s decaying through three different channels: $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. The mode $B_s \to D_s^-\pi^+\pi^-\pi^+$, $D_s^- \to \pi^-\pi^+\pi^-$ has not been considered due to the large amount of combinatorial background. The full list of selection cuts is shown in Table 3.4.

The collection of partially reconstructed semileptonic B^+, B^0, and B_s decays is listed in Table 3.1. Each candidate is obtained by combining a lepton and a D meson, on which a full vertex fit is performed. The optimization of kinematic and vertex quality cuts has been performed for the six Dl nominal modes individually. We point out that the requirement on the invariant mass of the lepton and the D meson, m_{Dl}, has revealed to be rather powerful in rejecting background, especially the denominated *fakes* background which is of non-combinatorial type as will be analyzed

$B_s \to \ldots$		$D_s^- \pi^+$			$D_s^- \pi^+ \pi^- \pi^+$	
$D_s \to \ldots$		$\phi\pi^-$	$K^{*0}K^-$	3π	$\phi\pi^-$	$K^{*0}K^-$
$\chi^2_{r\phi}(B)$	$<$	15	8	6	15	8
$\chi^2_{r\phi}(D)$	$<$	14	15	15	15	10
$\sigma_{L_{xy}}(B)$ [μm]	$<$	400	400	400	400	400
$\frac{L_{xy}}{\sigma}(B)$	$>$	7	9	13	11	18
$\frac{L_{xy}}{\sigma}(D)$	$>$	2	2	2	11	15
$L_{xy}(B \leftarrow D)$ [μm]	$>$	-200	-200	-100	-50	100
$\lvert d_0(B)\rvert$ [μm]	$<$	60	60	60	70	50
$p_T(B)$ [GeV/c^2]	$>$	5.5	5.5	6.0	6.0	5.0
$p_T(\pi_B)$ [GeV/c^2]	$>$	1.2	1.3	1.5	—	—
$p_T(K^\pm, \pi^\pm)$ [GeV/c^2]	$>$	0.35	0.40	0.40	0.45	0.35
m_{KK}, [GeV/c^2]		[1.013,1.28]	—	—	[1.010,1.031]	—
$\lvert m_{K\pi} - m_{K^*}\rvert$ [MeV/c^2]	$<$	—	55	—	—	50
$\lvert m_{KK\pi} - m_{D_s}\rvert$ [MeV/c^2]	$<$	—	25	—	—	—
$\lvert m_{K\pi\pi} - m_{D^-}\rvert$ [MeV/c^2]	$>$	—	24	—	—	24
$\lvert m_{K\pi\pi} - m_{K\pi}\rvert$ [MeV/c^2]	$>$	—	—	160	—	—
$\Delta R(D, \pi_B)$	$<$	—	1.5	1.5	—	—

Table 3.4: Summary of reconstruction requirements for $D_s\pi(\pi\pi)$ samples.

.

in the following section. The selection criteria are summarized in Table 3.5 for all Dl samples.

	lD^0	lD^+	$lD_s[\phi\pi^-]$	$lD_s[K^{*0}K^-]$	$lD_s[\pi^+\pi^-\pi^-]$
B vertex probability $>$	10^{-6}	10^{-5}	10^{-7}	10^{-5}	10^{-5}
$\chi^2_{xy}(D)$ $<$	20	20	20	20	20
$L_{xy}/\sigma_{L_{xy}}(D \to PV)$ $>$	6	11	5	8	11
$L_{xy}/\sigma_{L_{xy}}(D \to PV)$ $>$	2	2	—	2	2
$\lvert cos\,\psi\rvert$ $>$	—	—	0.3	0.3	—
$p_T(K^\pm, \pi^\pm)$ [GeV/c] $>$	0.5	0.7	0.7	0.7	0.7
m_{lD} [GeV/c^2]	2.9 - 5.3	2.9 - 5.3	2.9 - 5.3	2.9 - 5.3	2.9 - 5.3

Table 3.5: Summary of reconstruction requirements for Dl samples; the requirements for the D^*l sample are identical to those of the D^0l sample.

3.3 Samples composition

Samples of both fully and partially reconstructed B meson candidates are studied in this dissertation. In this section we assess the yields and study the various sources and type of candidates which contribute to their composition.

3.3.1 Sub-samples classification

The signal yields of the fully reconstructed decay samples are summarized in Table 3.6.

As it was mentioned in Section 3.2.2, it is required that the production tracks matched to the two trigger tracks be among the B daughters. In the case of semileptonic samples, if the lepton corresponds to one such track the event is said to be of the "*B trigger*" type, otherwise both trigger tracks are matched to the D daughters and the event is thus denoted of the "*D trigger*" type. For each of the Dl decay samples – D^*l, D^0l, D^+l, and D_s^-l – the following event sub-categories are considered: μ and e events, B and D trigger events. As these sub-samples have relatively different characteristics, such as background fractions and decay-length distributions, the analysis takes into account this classification with the aim of gaining in overall signal significance. The initial six semileptonic decay samples become therefore arranged in a total of 24 sub-samples, according to this further classification. The yields for these sub-samples are provided in Table 3.7.

sample	N_S
$B^+ \to J/\psi K^+$	4,950
$B^+ \to \bar{D}^0\pi^+$	9,600
$B^+ \to \bar{D}^0\pi^\pm$	1,560
$B^0 \to J/\psi K^{*0}$	1,790
$B^0 \to D^-\pi^+$	8,420
$B^0 \to D^-\pi^+\pi^-\pi^+$	4,610
$B^0 \to D^{*-}\pi^+$ ($\bar{D}^0 \to K^+\pi^-$)	1,380
$B^0 \to D^{*-}\pi^+$ ($\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$)	1,010
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-$)	1,090
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$)	820
$B_s \to D_s^-\pi^+$, $D_s \to \phi\pi$	550
$B_s \to D_s^-\pi^+$, $D_s \to K^*K$	240
$B_s \to D_s^-\pi^+$, $D_s \to 3\pi$	110
$B_s \to D_s^-\pi^+\pi^-\pi^+$, $D_s \to \phi\pi$	160
$B_s \to D_s^-\pi^+\pi^-\pi^+$, $D_s \to K^*K$	60

Table 3.6: Summary of the estimated signal event yields for the fully reconstructed samples.

The following ct regions are allowed for the Dl samples,

$D^{(*)}$:	$ct \in [50, \infty)$ μm ,
$lD_s^-[\phi\pi^-]$, B trigger:	$ct \in [50, \infty)$ μm ,
$lD_s^-[\phi\pi^-]$, D trigger:	$ct \in [100, \infty)$ μm ,
$lD_s^-[K^{*0}K^-,\ \pi^+\pi^-\pi^-]$, B and D triggers:	$ct \in [100, \infty)$ μm .

3.3.2 Fully reconstructed B samples

The reconstructed nominal signal decays in the $J/\psi K$ samples are $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$. In order to identify misreconstructed candidates in the samples, the

sub-sample	trigger type	N_S	sub-sample	trigger type	N_S
μD^0	B	103,300	$\mu D_s[\phi\pi^-]$	B	5,230
μD^0	D	18,100	$\mu D_s[\phi\pi^-]$	D	1,150
μD^+	B	11,300	$\mu D_s[K^{*0}K^-]$	B	1,750
μD^+	D	3,200	$\mu D_s[K^{*0}K^-]$	D	650
μD^*	B	47,800	$\mu D_s[\pi^+\pi^-\pi^-]$	B	1,530
μD^*	D	13,400	$\mu D_s[\pi^+\pi^-\pi^-]$	D	750
eD^0	B	53,800	$eD_s[\phi\pi^-]$	B	2,910
eD^0	D	10,200	$eD_s[\phi\pi^-]$	D	720
eD^+	B	6,400	$eD_s[K^{*0}K^-]$	B	970
eD^+	D	1,900	$eD_s[K^{*0}K^-]$	D	360
eD^*	B	26,400	$eD_s[\pi^+\pi^-\pi^-]$	B	850
eD^*	D	7,600	$eD_s[\pi^+\pi^-\pi^-]$	D	340

Table 3.7: Summary of the signal event yields per Dl sub-sample.

signal selection criteria is applied to Monte Carlo samples of b-hadron $\to J/\psi X$ decays [56, 57]. Figure 3-2 shows the mass distributions obtained from various contributing channels in such inclusive Monte Carlo samples, reconstructed as $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$.

The fitting mass region for these samples is $(5.17, 5.39)$ GeV/c^2. This was chosen to avoid sources of partially reconstructed decays which are present for masses below about 5.17 GeV/c^2. Two predominant contributions from misreconstructed decays remain, arising from incorrect identification of kaons and pions.

The Cabibbo-suppressed $B^+ \to J/\psi\pi^+$ decays, with the pion misidentified as a kaon, have a reconstructed mass slightly above the nominal B^+ mass. This contribution amounts to a fraction of candidates relative to the nominal decay which is reported in Table 3.8.

In the case of the $B^0 \to J/\psi K^{*0}$ mode, a self-reflection arises from $K^{*0} \to K^+\pi^-$ decays which are reconstructed with wrong (*swapped*) pion and kaon mass assignments to the two forming tracks. The relative contribution of these swapped candidates is extracted from a sample of signal Monte Carlo to which the selection criteria are applied. We observe that the mass distributions for both the correctly reconstructed candidates and those formed of swapped K^{*0} have a general Gaussian shape, as shown in Figure 3-3. These are characterized by widths of 10.5 and 25.5 MeV/c^2, with that of the swapped component being broader; the relative fraction is reported in Table 3.8.

The list of fully reconstructed $B \to D\pi(\pi\pi)$ decays is included in Table 3.1. The reconstructed mass spectra for these modes show various satellite structures. Their source is investigated by identifying the contributions in Monte Carlo simulation from the involved decays when reconstructed as the nominal signal modes [58]. Figures 3-4 and 3-5 illustrate these for B^+, B^0, and B_s decays, in a wide mass range.

A narrow fitting mass range is chosen which, as it was the case for the $J/\psi K$ samples, excludes, or substantially reduces, several of the structures appearing in the

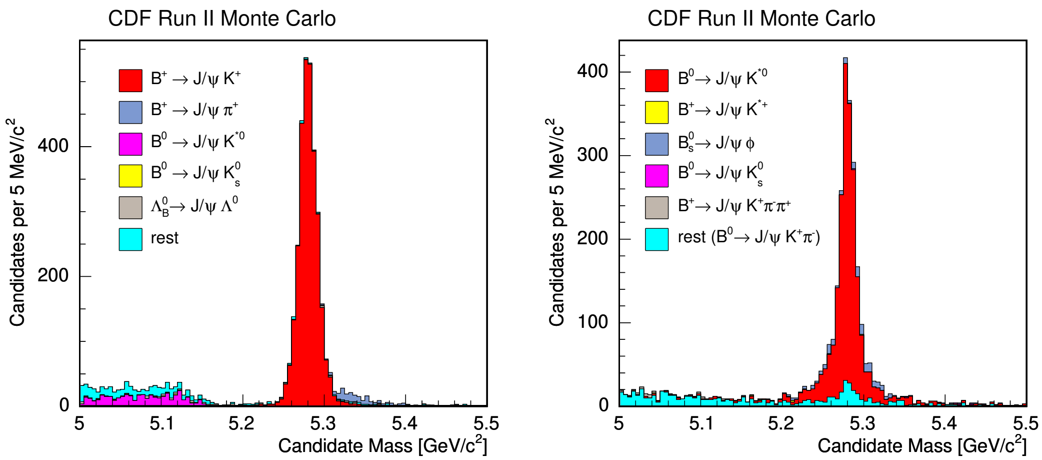


Figure 3-2: Monte Carlo mass distributions of b-hadron $\rightarrow J/\psi X$ decays, reconstructed as $B^+ \rightarrow J/\psi K^+$ (left), and $B^0 \rightarrow J/\psi K^{*0}(K^+\pi^-)$ (right).

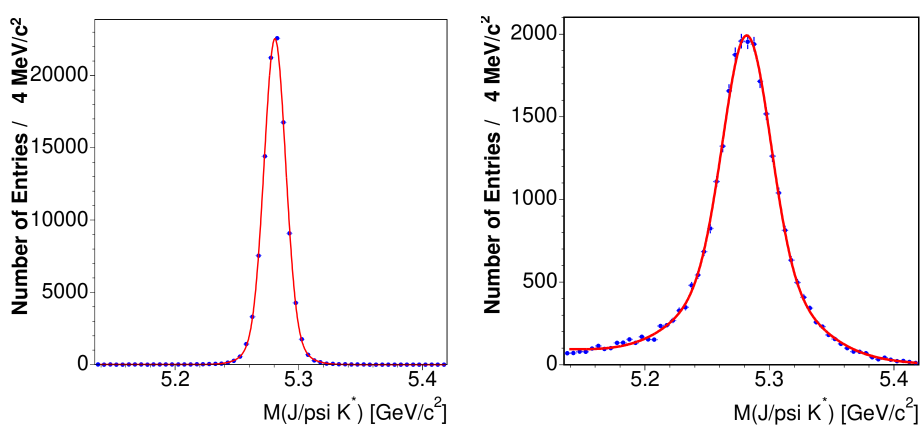


Figure 3-3: Mass distribution for $B^0 \rightarrow J/\psi K^{*0}$ candidates reconstructed in the signal MC sample when the K^* is reconstructed correctly (left), and when it is misreconstructed with swapped K and π assumptions (right).

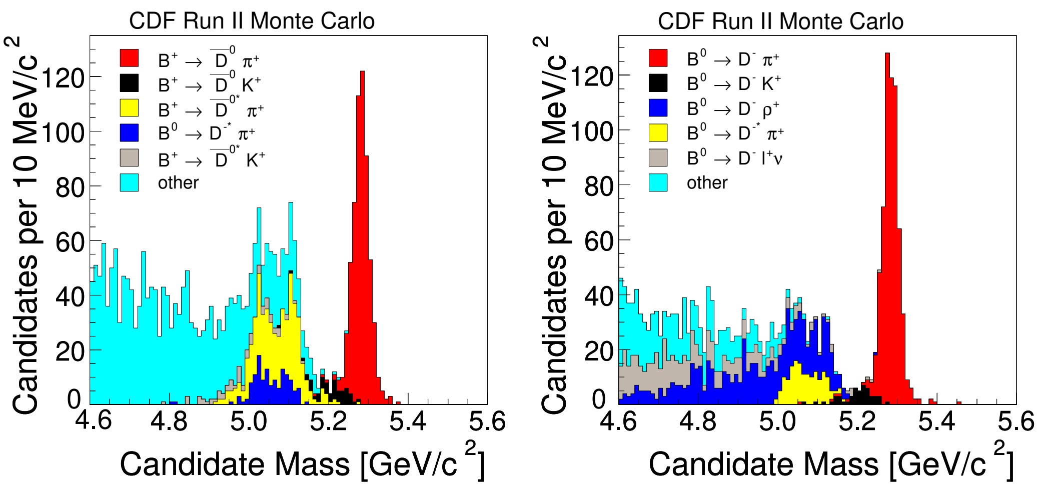


Figure 3-4: Monte Carlo mass distributions of b-hadron $\to DX$ decays, reconstructed as $B^+ \to \bar{D}^0\pi^+$ ($\bar{D}^0 \to K^+\pi^-$) (left), and $B^0 \to D^-\pi^+$ ($D^- \to K^+\pi^-\pi^-$) (right).

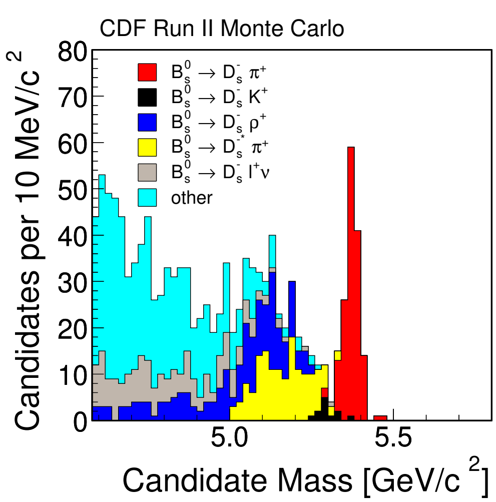


Figure 3-5: Monte Carlo mass distributions of b-hadron $\to D_sX$ decays, reconstructed as $B_s \to D_s^-\pi^+$ ($D_s^- \to \phi\pi^-$).

	$B^+ \to J/\psi K^+$	$B^0 \to J/\psi K^{*0}$
$B^+ \to J/\psi\pi^+$	2.50 ± 0.06 %	-
$K^{*0}(K \leftrightarrow \pi)$ swap	-	16.9±0.4 %

Table 3.8: Contributions from non-combinatorial components in the $J/\psi K$ samples relative to the nominal signals.

low mass region relative to that of the nominal signal. The upper limit of the mass interval is chosen at 5.6 GeV/c^2, while the lower limit is set at 5.2 GeV/c^2 for B^+ and B^0 decays, and at 5.3 GeV/c^2 for the B_s modes.

Despite having selected a restricted mass range, there is still some contamination which leaks into this narrower fitting mass region. The contribution of such partial reconstructed decays is extrapolated to the narrower mass region from a fit to the data performed in the extended mass spectrum. The numerical results for their contributions relative to the nominal signals are summarized in Table 3.9. These extended fits are in addition most relevant for determining the combinatorial background behavior. The details on the wide mass range fits are addressed further below.

partially-reconstructed fractions			
$B^+ \to \bar{D}^0\pi^+$	–	$B_s \to D_s^-[\phi\pi^-]\pi^+$	0.59 %
$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$	–	$B_s \to D_s^-[K^{*0}K^-]\pi^+$	2.67 %
$B^0 \to D^-\pi^+$	–	$B_s \to D_s^-[\pi^-\pi^+\pi^-]\pi^+$	0.66 %
$B^0 \to D^-\pi^+\pi^-\pi^+$	–	$B_s \to D_s^-[\phi\pi^-]\pi^+(\pi^-\pi^+)$	0.99 %
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$	–	$B_s \to D_s^-[K^{*0}K^-]\pi^+(\pi^-\pi^+)$	0.41 %

Table 3.9: Contributions from partially reconstructed backgrounds to the $B \to D\pi(\pi\pi)$ samples; the relative fractions are calculated in the narrow mass range, from Monte Carlo simulation.

The Cabibbo-suppressed decays $B \to D^{(*)}K(\pi\pi)$, where the kaon track is misassigned the pion mass, appear as small contributions underneath the main mass peaks in the hadronic samples. The shape and the relative amount of these components can be predicted from Monte Carlo simulation, where the branching fraction of the Cabibbo-suppressed mode is set to the expected amount based on the Cabibbo angle relative to the Cabibbo-favored modes. These contributions are of the level of about 6% relative to the nominal signals. The fractions evaluated for each mode are specified in Table 3.10, and the mass shape is illustrated for the main hadronic B_s mode in Figure 3-6.

There is some amount of cross-talk among the various hadronic modes. For example, the decay channels $B_s \to D_s^-\pi^+$ ($D_s^- \to K^+K^-\pi^-$) may be reconstructed as $B^0 \to D^-\pi^+$ ($D^- \to K^+\pi^-\pi^-$) in case the kaon is misidentified as a pion. The inverse also occurs, when $B^0 \to D^-\pi^+$ ($D^- \to K^+\pi^-\pi^-$) decays are misreconstructed in one of the $B_s \to D_s^-\pi^+$ modes. There are additional background sources coming from misreconstructed $\Lambda_b^0 \to \Lambda_c^+\pi^-$ ($\Lambda_c^+ \to p^+K^-\pi^+$) decays, where the proton is

cabibbo-suppressed fractions			
$B^+ \to \bar{D}^0\pi^+$	6.16 %	$B_s \to D_s^-[\phi\pi^-]\pi^+$	6.09 %
$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$	7.87 %	$B_s \to D_s^-[K^{*0}K^-]\pi^+$	5.66 %
$B^0 \to D^-\pi^+$	5.80 %	$B_s \to D_s^-[\pi^-\pi^+\pi^-]\pi^+$	5.95 %
$B^0 \to D^-\pi^+\pi^-\pi^+$	6.68 %	$B_s \to D_s^-[\phi\pi^-]\pi^+(\pi^-\pi^+)$	5.00 %
$B^0 \to D^{*-}\pi^+$	6.34 %	$B_s \to D_s^-[K^{*0}K^-]\pi^+(\pi^-\pi^+)$	5.00 %
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$	–		

Table 3.10: Contributions from Cabibbo-suppressed components to the $B \to D\pi(\pi\pi)$ samples; the relative fractions are calculated in the narrow mass range, from Monte Carlo simulation.

misidentified with a kaon or a pion. The contributions of such decays relative to the nominal signals are obtained from the ratio of the relevant branching fractions, and trigger and reconstruction efficiencies evaluated from Monte Carlo simulation. These fractions are summarized for all B_s channels in Tables 3.11 and 3.12.

$B_s \to \ldots$	$D_s^-[\phi\pi^-]\pi^+$	$D_s^-[K^{*0}K^-]\pi^+$	$D_s^-[\pi^-\pi^+\pi^-]\pi^+$
$B_s \to D_s^-[\phi\pi^-]\pi^+$	1	$0.249 \pm 0.075\%$	–
$B_s \to D_s^-[K^{*0}K^-]\pi^+$	$0.233 \pm 0.064\%$	1	–
$B_s \to D_s^-[\pi^-\pi^+\pi^-]\pi^+$	–	$0.0073 \pm 0.0063\%$	1
$B^0 \to D^-[K^+\pi^-\pi^-]\pi^+$	$1.26 \pm 0.40\%$	$1.88 \pm 0.29\%$	$0.0587 \pm 0.0092\%$
$\Lambda_b \to \Lambda_c^+[pK^-\pi^+]\pi^-$	$0.598 \pm 0.21\%$	$11.6 \pm 1.9\%$	–
$\Lambda_b \to \Lambda_c^+[p\pi^-\pi^+]\pi^-$	$0.0180 \pm 0.0039\%$	$0.236 \pm 0.039\%$	$5.16 \pm 0.85\%$

Table 3.11: Contributions from fully reconstructed backgrounds to the $B_s \to D_s^-\pi^+$ samples; the fractions are calculated in the narrow mass range, from Monte Carlo simulation.

Wide range mass fits

In order to gain understanding of background components in the hadronic $D\pi(\pi\pi)$ samples, we perform a fit to the mass distribution of the data in an extended mass range. Specifically, the selected wider mass intervals are $(4.4, 7.6)$ and $(4.6, 5.6)$ GeV/c^2 for the B^+ and B^0, and the B_s samples, respectively.

The region to the left of the nominal signal peak in the mass distribution contains various structures. The corresponding sources are identified from Monte Carlo simulation, as illustrated in Figures 3-4 and 3-5. These include partially reconstructed $B_{(s)} \to D_{(s)}\pi X$ and $B_{(s)} \to D_{(s)}l\nu$ decays. In the latter case the lepton is misreconstructed as a pion, while the former class of candidates includes $B_{(s)} \to D_{(s)}\rho$ decays, and $B_{(s)} \to D^*_{(s)}\pi$ decays, with the $D^*_{(s)}$ decaying to $D_{(s)}\gamma$ and to $D_{(s)}\pi^0$ and the γ and π^0 not being detected.

$B_s \to \ldots$	$D_s^-[\phi\pi^-]\pi^+\pi^-\pi^+$	$D_s^-[K^{*0}K^-]\pi^+\pi^-\pi^+$
$B_s \to D_s^-[\phi\pi^-]\pi^+(\pi^-\pi^+)$	1	0.55 ± 0.21 %
$B_s \to D_s^-[K^{*0}K^-]\pi^+(\pi^-\pi^+)$	0.371 ± 0.13 %	1
$B^0 \to D^-[K^+\pi^-\pi^-]\pi^+(\pi^-\pi^+)$	1.91 ± 0.87 %	6.3 ± 2.8 %
$\Lambda_b \to \Lambda_c^+[pK^-\pi^+]\pi^-(\pi^+\pi^-)$	1.16 ± 0.19 %	14.2 ± 2.4 %
$\Lambda_b \to \Lambda_c^+[p\pi^-\pi^+]\pi^-(\pi^+\pi^-)$	0.076 ± 0.013 %	0.43 ± 0.071 %

Table 3.12: Contributions from fully reconstructed backgrounds to the $B_s \to D_s^-\pi^+\pi^-\pi^+$ samples; the fractions are calculated in the narrow mass range, from Monte Carlo simulation.

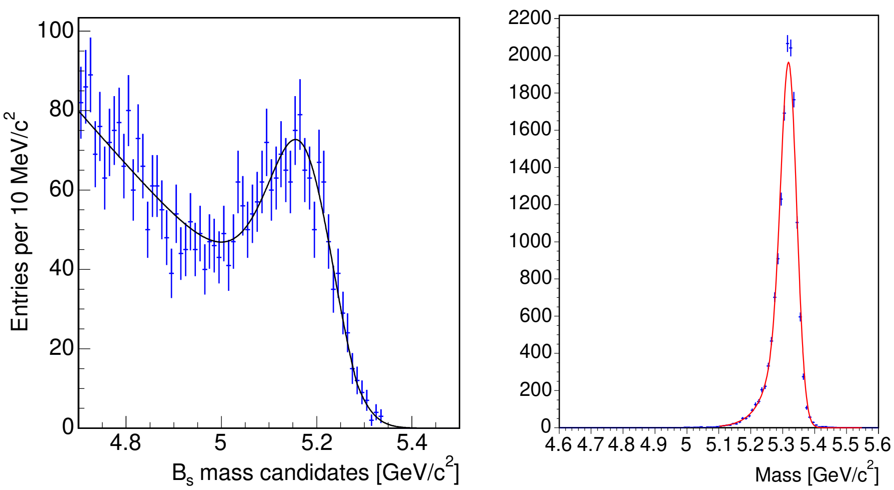


Figure 3-6: Mass templates for partially reconstructed (left) and Cabibbo-suppressed (right) background decays, illustrated with the $B_s \to D_s^-\pi^+$ ($D_s^- \to \phi\pi^-$) channel.

The shape of the templates describing the background constituted by partially reconstructed decays is modeled using inclusive B Monte Carlo samples, $b \to D_{(s)}X$. The $B_{(s)} \to D_{(s)}\rho, D^*_{(s)}\pi$ decays, and those which include leptons in the final state or in which more than a pion is lost are treated by separate templates. These correspond to a smeared Gaussian model, and a linear shape with a cut-off, respectively. In Figure 3-6 the combined templates are illustrated for the main B_s decay mode. The relative normalizations between the two template models is allowed to float in the fit.

The fully reconstructed backgrounds have already been addressed. These tend to have shapes of bumps not much wider than the signal peak, and place themselves in the vicinity of it. As it was observed before, the description of these components needed to be inferred from Monte Carlo simulation, which is imposed in the fit.

The contribution of the Cabibbo-suppressed decays, as it was also addressed, is predicted from Monte Carlo simulation. The relative fractions are summarized in Table 3.10. In the fit to the data the shape of this component is also fixed from simulation; its width is dominated by kinematic smearing, rather than detector resolution,

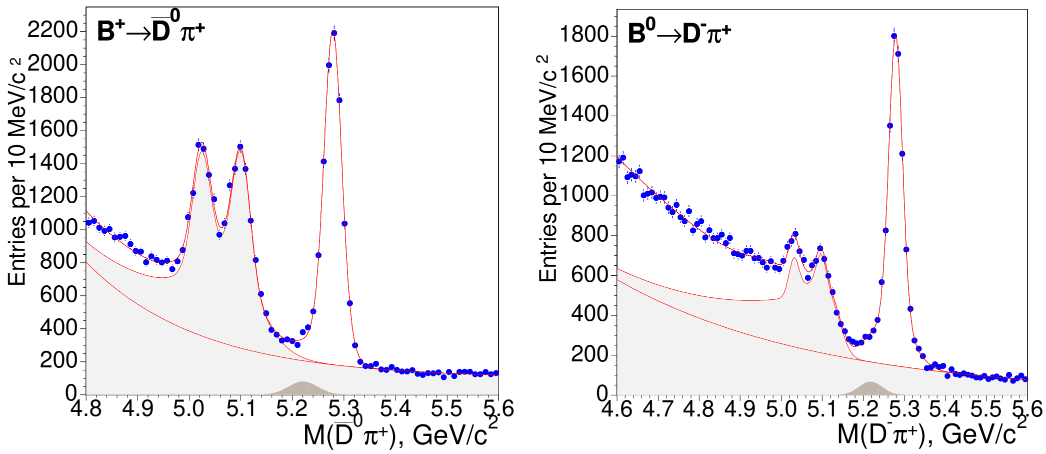


Figure 3-7: Wide range mass fits for $B^+ \to \bar{D}^0\pi^+$ (left) and $B^0 \to D^-\pi^+$ (right) candidates.

due to mass mis-assignment. The template is illustrated in Figure 3-6.

The wide range mass fit projections are shown in Figures 3-7 and 3-8 for the main $\bar{D}^0\pi^+$, $D^-\pi^+$, $D_s^-\pi^+$, and $D_s^-\pi^+\pi^-\pi^+$ channels. The shape parameters and normalization of the combinatorial background component float free in the fits.

3.3.3 Partially reconstructed B samples

The Dl candidates are formed of a combination of a lepton and D meson pair consistent with having originated from a B meson decay. The lepton l is identified as either an electron or a muon, and D stands for either of D^0, D^+, D^{*-}, and D_s^-, which are fully reconstructed.

None of the final states D^0l, D^+l, $D^{*+}l$ or D_sl is formed exclusively from a single, specific B decay mode. The reconstruction of the B meson is *partial*, with the neutrino and possibly other decay products being missed. Furthermore, the D^0l, D^+l, $D^{*+}l$ final states originate from a mixture of B^+ and B^0 decays through excited charm states. These inclusive modes are represented as follows,

- $B^{+,0} \to \bar{D}^0 l^+ X$, $\bar{D}^0 \to K^+\pi^-$,
- $B^{+,0} \to D^- l^+ X$, $D^- \to K^+\pi^-\pi^-$,
- $B^{+,0} \to D^{*-} l^+ X$, $D^{*-} \to \bar{D}^0\pi^-$, $\bar{D}^0 \to K^+\pi^-$,
- $B_s \to D_s^- l^+ X$, $D_s^- \to \phi\pi^-$,
- $B_s \to D_s^- l^+ X$, $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$,
- $B_s \to D_s^- l^+ X$, $D_s^- \to \pi^+\pi^-\pi^-$.

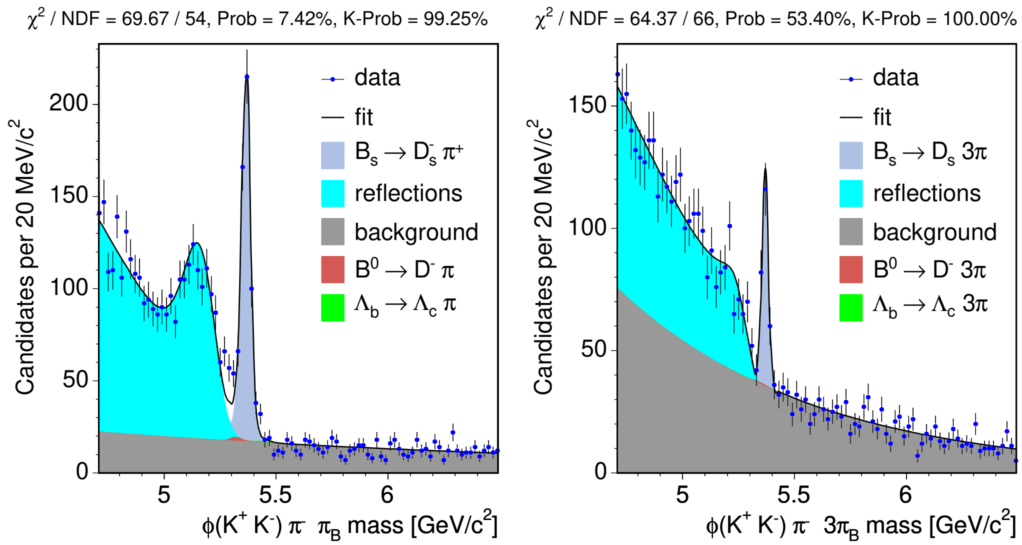


Figure 3-8: Wide range mass fits for $B_s \to D_s^-[\phi\pi^-]\pi^+$ (left) and $B_s \to D_s^-[\phi\pi^-]\pi^+(\pi^-\pi^+)$ (right) candidates.

The $B \to Dl$ transition may not be direct, as the indicated final state may be achieved through a chain of intermediate states. X stands for the neutrino from the B semileptonic decay, and other unreconstructed decay products from intermediate channels.

This dissertation deals with the study of B properties, which are specific to the meson species: B^+, B^0, and B_s. In the inclusive samples, accordingly, we ought also to provide independent description for the three species. In particular, the relative fractions of B^+ and B^0 which contribute to each of the D^0l, D^+l, $D^{*+}l$ samples must be determined. The determination of these fractions, and corresponding uncertainties, is done based on the values of the so-called sample composition parameters, and the trigger and reconstruction efficiencies. This issue is addressed further below.

In the semileptonic decays, the mass of the complete B system cannot be determined due to the reconstruction being partial. Instead, the reconstructed mass of the involved D mesons is employed to describe the samples. There are several categories of background events which may mimic our semileptonic signals by providing both lepton and D candidates with the correct charge correlation. Certain physics processes involving B hadron decays other than the considered signal channels may contribute to the Dl signature. These are denoted by *physics* type background, and its contribution is evaluated from Monte Carlo simulation. Additionally, our Dl candidates may have contaminations resulting from incorrect reconstruction of either the D meson or the B-daughter lepton. Combinatorial background formed of events of the former class presents itself as a linear contribution in the D mass distribution. Events of the latter class are characterized using a sample obtained by selecting fake leptons and are thus denoted by *fakes* background.

Physics background

This class of background is realized by B hadron decays other than the semileptonic B signals which also feed into our Dl decay signatures. More specifically, this may correspond for instance to $B \to DDX$ decays where one of the charm mesons decays semileptonically. B decays to tau leptons with these decaying semileptonically are also part of this background class except in the case of the $D_s l$ samples where $B_s \to \tau D_s X, \tau \to l\nu\nu$ decays are chosen and treated as a part of the signal. The relative overall contribution from these background components typically amounts up to 16% of the $D_s l$ modes and to about 6% of the remaining Dl modes; such fractions are shown in Table 3.13 for the considered background components.

The contributions are evaluated, for each of the 24 semileptonic decay sub-samples, from a corresponding simulated inclusive physics background sample. These inclusive samples are obtained by generating Monte Carlo samples for each of the contributing background modes and weighting them according to the relevant branching fractions [1] and reconstruction efficiencies. The fractions are summarized in Table 3.14.

decay		$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]
$B_s \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to \mu^+ X$	—	—	—
$B_s \to D_s^{(*+)} D_s^{(*-)} X,$	$D_s^{(*)} \to \mu^+ X$	3.0 ± 0.9	3.0 ± 0.9	2.9 ± 0.9
$B^+ \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to \mu^+ X$	6.2 ± 2.5	6.2 ± 2.5	6.0 ± 2.4
$B^0 \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to \mu^+ X$	6.3 ± 2.5	6.3 ± 2.5	6.0 ± 2.4
total		15.5 ± 3.6	15.5 ± 3.6	14.9 ± 3.5
$B_s \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to e^+ X$	—	—	—
$B_s \to D_s^{(*+)} D_s^{(*-)} X,$	$D_s^{(*)} \to e^+ X$	2.1 ± 0.6	2.3 ± 0.7	1.7 ± 0.5
$B^+ \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to e^+ X$	5.4 ± 2.2	5.1 ± 2.0	5.2 ± 2.1
$B^0 \to D_s^{(*)} D^{(*)} X,$	$D^{(*)} \to e^+ X$	5.3 ± 2.1	5.0 ± 2.0	5.1 ± 2.0
total		12.8 ± 3.1	12.4 ± 3.0	12.0 ± 2.9

Table 3.13: Expected fraction of each physics background component, relative to the sum of signal and physics background, for each selected D_s decay mode, and for muons and electrons separately (averaged over B and D trigger types).

Fakes background and m_{Dl} analysis

This class of background arises from Dl candidates formed of a real D meson together with a track which may not be a real lepton or which did not really originate from a B decay. The lepton is said to be a *fake* in the former case, and in the latter it may have originated from a $c\bar{c}$ event.

The characterization of background events of this type may be attempted in few ways. One would be to explore the sample of wrongly correlated candidates, *e.g.* D^+l^+, given that no specific charge correlation between the D and l candidates is expected for events of the type mentioned above. This has in fact been explored

physics background					
sub-sample	trigger type	fraction [%]	sub-sample	trigger type	fraction [%]
μD^0	B	14.8 ± 5.0	$\mu D_s[\phi\pi^-]$	B	14.8 ± 5.0
μD^0	D	18.1 ± 6.2	$\mu D_s[\phi\pi^-]$	D	18.1 ± 6.2
μD^+	B	14.7 ± 5.0	$\mu D_s[K^{*0}K^-]$	B	14.7 ± 5.0
μD^+	D	17.8 ± 6.2	$\mu D_s[K^{*0}K^-]$	D	17.8 ± 6.2
μD^*	B	14.2 ± 4.9	$\mu D_s[\pi^+\pi^-\pi^-]$	B	14.2 ± 4.9
μD^*	D	16.3 ± 5.7	$\mu D_s[\pi^+\pi^-\pi^-]$	D	16.3 ± 5.7
eD^0	B	10.2 ± 3.6	$eD_s[\phi\pi^-]$	B	10.2 ± 3.6
eD^0	D	17.7 ± 6.1	$eD_s[\phi\pi^-]$	D	17.7 ± 6.1
eD^+	B	8.7 ± 3.0	$eD_s[K^{*0}K^-]$	B	8.7 ± 3.0
eD^+	D	10.4 ± 3.6	$eD_s[K^{*0}K^-]$	D	10.4 ± 3.6
eD^*	B	11.9 ± 4.1	$eD_s[\pi^+\pi^-\pi^-]$	B	11.9 ± 4.1
eD^*	D	9.4 ± 3.2	$eD_s[\pi^+\pi^-\pi^-]$	D	9.4 ± 3.2

Table 3.14: Fractions of physics background, for all 24 semileptonic decay subsamples, with respect to the sum of signal and physics background.

in an earlier stage of the analysis. The method which we employ is based on a sample of fake leptons and real D events, which is obtained by applying lepton *anti*-selection; *i.e.* a sample in which the lepton track candidates are required to have low likelihood (specifically a lepton likelihood cut $\leq$0.05 is applied). The corresponding reconstructed D mass and Dl proper decay time distributions are illustrated in Figure 4-6.

The fraction of fakes background events in our samples is determined from a fit to the invariant mass distribution of the Dl pair candidates, m_{Dl}. Firstly, the combinatorial background component is subtracted, using the m_{Dl} distribution for D mass-sideband candidates. This distribution is then fitted with the model

$$f_{\text{fakes}} L_{m_{Dl}}^{\text{fakes}} + (1 - f_{\text{fakes}}) \left[f_{\text{phys.}} L_{m_{Dl}}^{\text{phys.}} + (1 - f_{\text{phys}}) L_{m_{Dl}}^{\text{signal}} \right], \tag{3.3}$$

where the $L_{m_{Dl}}$ and f denote the m_{Dl} templates and fractions of the indicated components. For the signal and the physics background these are determined from Monte Carlo simulation. The m_{Dl} template for the fakes background $L_{m_{Dl}}^{\text{fakes}}$ is obtained from the fake lepton sample after D mass-sideband subtraction. These various m_{Dl} distributions are illustrated in Figure 3-9. The background events have on average a significantly lower m_{Dl} than the signal events. The fraction f_{fakes} of the fakes background component which we set out to determine is the only free parameter in the fit, while the other two fraction are either fixed or constrained to the Monte Carlo values.

The fit is performed in the m_{Dl} range of $(2.05, 5.35)$ GeV/c^2 for the $D_s l$ samples, and of $(2.05, 5.50)$ GeV/c^2 for the other semileptonic samples. The input fractions and templates mentioned above are thus determined for this looser m_{Dl} selection criteria.

The fit result for the f_{fakes} parameter is finally propagated into the nominal selection range $(2.9, 5.3)$ GeV/c^2 specified in Table 3.5. The final fakes background fractions are summarized in Table 3.15 for all Dl decay sub-samples.

fakes background					
sub-sample	trigger type	fraction [%]	sub-sample	trigger type	fraction [%]
μD^0	B	2.4 ± 0.5	$\mu D_s[\phi\pi^-]$	B	2.7 ± 0.6
μD^0	D	6.8 ± 1.4	$\mu D_s[\phi\pi^-]$	D	7.7 ± 1.8
μD^+	B	2.5 ± 0.5	$\mu D_s[K^{*0}K^-]$	B	3.7 ± 0.8
μD^+	D	6.3 ± 1.3	$\mu D_s[K^{*0}K^-]$	D	8.7 ± 1.9
μD^*	B	2.0 ± 0.4	$\mu D_s[\pi^+\pi^-\pi^-]$	B	4.4 ± 0.9
μD^*	D	5.3 ± 1.1	$\mu D_s[\pi^+\pi^-\pi^-]$	D	7.7 ± 1.7
eD^0	B	1.9 ± 0.5	$eD_s[\phi\pi^-]$	B	4.0 ± 0.9
eD^0	D	7.0 ± 1.4	$eD_s[\phi\pi^-]$	D	8.8 ± 1.9
eD^+	B	2.1 ± 0.4	$eD_s[K^{*0}K^-]$	B	2.1 ± 0.5
eD^+	D	8.9 ± 1.8	$eD_s[K^{*0}K^-]$	D	8.9 ± 2.0
eD^*	B	2.0 ± 0.4	$eD_s[\pi^+\pi^-\pi^-]$	B	3.5 ± 0.8
eD^*	D	5.1 ± 1.0	$eD_s[\pi^+\pi^-\pi^-]$	D	7.3 ± 1.7

Table 3.15: Fractions of fakes background, for all 24 semileptonic decay sub-samples, with respect to the sum of signal with fakes and physics backgrounds.

Description of B^+ and B^0 content of $D^{(*)}l$ samples

The samples D^0l, D^+l, $D^{*+}l$ are formed of contributions from 24 decay chains of B^+ and B^0 mesons. These are illustrated schematically in Figure 3-10. The fraction of events with a signature k arising, say, from B^0 decays is obtained by performing a sum, over all B^0 decay chains h that contribute to that decay signature (k), of the product of the involved branching ratios with the selection efficiency; conceptually,

$$S_k^0 = \sum_{B^0 \to h \in k} \mathcal{B}(B \to lD^{(**)}\ldots)\times\mathcal{B}(D^{(**)} \to D^{(*)}\ldots)\times\ldots\times\mathcal{B}(D^+ \to K\pi\pi)\times\epsilon(B^0 \to h).$$

A similar expression S_k^+ may be obtained for B^+ decays. These relations as written depend upon absolute branching ratios and efficiencies, but by removing an overall normalization only relative fractions matter. Indeed, the contributing fraction from B^0 decays to a sample k is given by

$$f_k^0 = \frac{S_k^0}{S_k^0 + S_k^+}\,, \tag{3.4}$$

where efficiencies are expressed relative to that of the direct decay chain (*i.e.* $B \to Dl\nu$), and, for the most part, only ratios of branching ratios are needed.

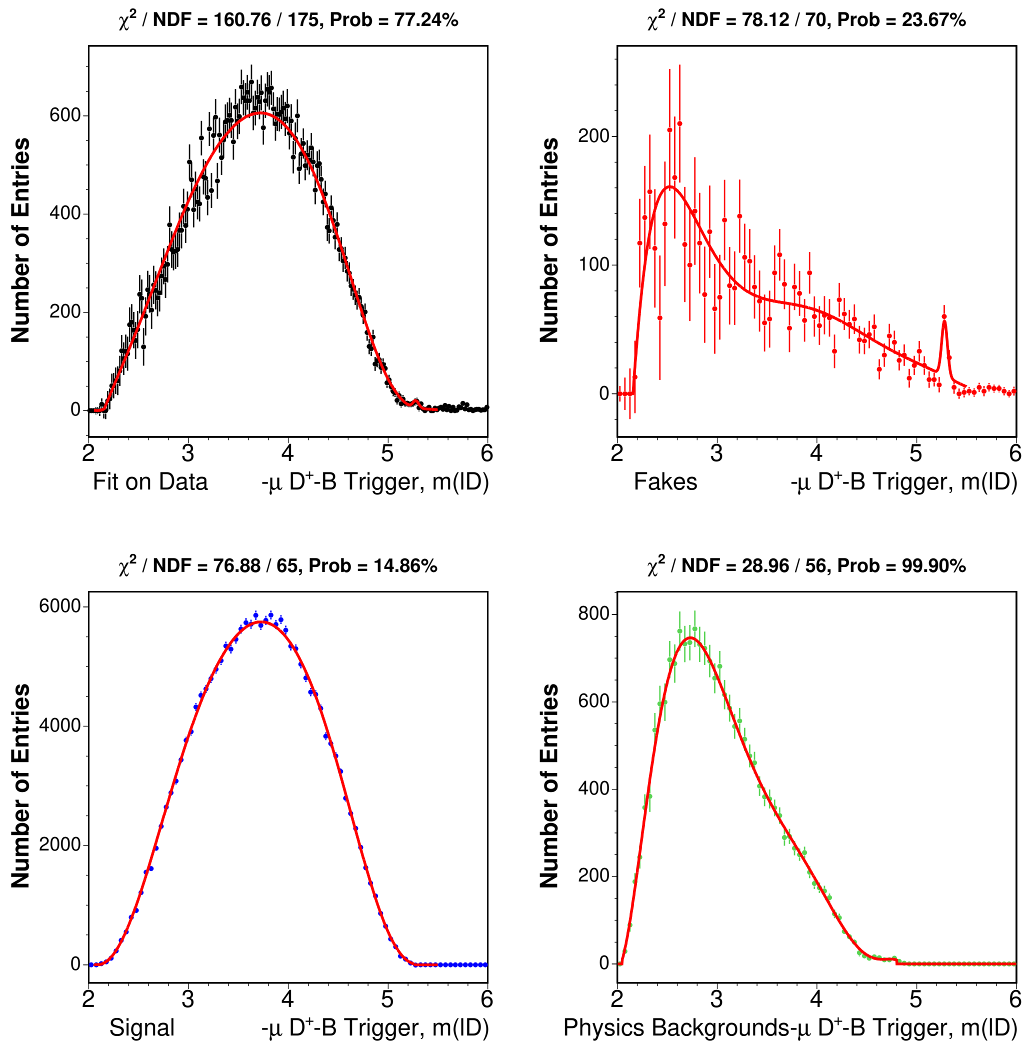


Figure 3-9: Fit to m_{Dl} [GeV/c^2] distribution on data (top left), together with templates for fake background (top right, from fake lepton sample), and signal and physics background (bottom left and bottom right, respectively, from Monte Carlo simulation). This is illustrated for the μD^+ B trigger sub-sample.

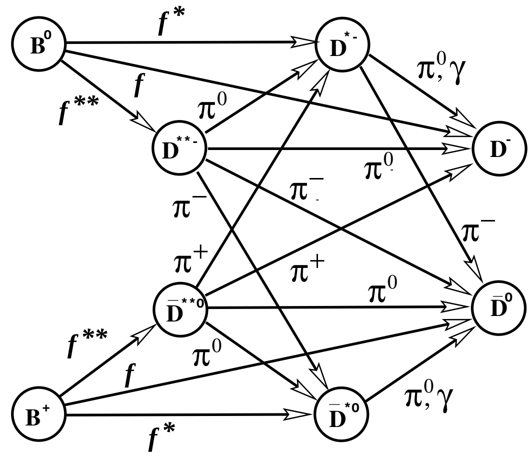


Figure 3-10: State diagram of $B^{+,0} \to DlX$ transitions.

The signal composition of the sample may for our purposes be characterized as follows. We denote by f, f^*, and f^{**} the relative fractions for B decaying into a D (*i.e.* D^0 or D^+), D^* (*i.e.* D^{*-} or D^{*0}), and D^{**} (*i.e.* an excited D state other than D^*), with $f + f^* + f^{**} = 1$. The rate that D^{**} states decay into D^* states relative to the rate that D^{**} states decay directly into D mesons is denoted by P_V. The efficiency for reconstructing the decay $D^{*-} \to \bar{D}^0\pi^-$ is represented by ϵ^*. The relative fractions of the individual decay chains are then given in Table 3.17, in terms of the set of independent sample composition parameters shown in Table 3.16 – f^{**}, $R_f(\equiv f^*/f)$, P_V, ϵ^* – and the relevant branching fractions [1].

R_f	2.14 ± 0.14
f^{**}	0.31 ± 0.05
P_V	0.627 ± 0.26

Table 3.16: Semileptonic $B^{+,0}$ signal composition parameters.

A needed correction for the sample composition, as referred to above, is the general effect of trigger and reconstruction efficiencies. The different decay chains are characterized by potentially different kinematic behavior, which results in different efficiencies for being selected at trigger or reconstruction levels, thus altering the decay chains relative contribution to a given decay signature. Since only relative efficiencies are involved as noted above, many uncertainties cancel in the calculation. These efficiencies are determined from 21 signal Monte Carlo samples, and given in Table 3.18, for each lD sample, relative to the corresponding direct decay mode.

	decay signatures		
	$D^{*-}l$	$D^{-}l$	$D^{0}l$
B^{+} decay chains			
$\to \overline{D}^{**0}l^{+}\nu$			
$\overline{D}^{**0} \to D^{*-}\pi^{+}_{**}$			
$D^{*-} \to \overline{D}^{0}\pi^{-}_{*}$	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})\epsilon^{*}$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})(1-\epsilon^{*})$
$D^{*-} \to D^{-}\pi^{0}_{*}$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(D^{-}\pi^{0}_{*})$	—
$D^{*-} \to \overline{D}^{0}\gamma$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(D^{-}\gamma)$	—
$\overline{D}^{**0} \to \overline{D}^{*0}\pi^{0}_{**}$			
$\overline{D}^{*0} \to \overline{D}^{0}\pi^{0}_{*}$	—	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(\overline{D}^{0}\pi^{0}_{*})$
$\overline{D}^{*0} \to \overline{D}^{0}\gamma$	—	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(\overline{D}^{0}\gamma)$
$\overline{D}^{**0} \to D^{-}\pi^{+}_{**}$	—	$f^{**}(1-P_V)\frac{2}{3}$	—
$\overline{D}^{**0} \to \overline{D}^{0}\pi^{0}_{**}$	—	—	$f^{**}(1-P_V)\frac{1}{3}$
$\to \overline{D}^{*0}l^{+}\nu$			
$\overline{D}^{*0} \to \overline{D}^{0}\pi^{0}_{*}$	—	—	$f^{*}\mathcal{B}^{*}(\overline{D}^{0}\pi^{0}_{*})$
$\overline{D}^{*0} \to \overline{D}^{0}\gamma$	—	—	$f^{*}\mathcal{B}^{*}(\overline{D}^{0}\gamma)$
$\to \overline{D}^{0}l^{+}\nu$	—	—	f
B^{0} decay chains			
$\to D^{**-}l^{+}\nu$			
$D^{**-} \to \overline{D}^{*0}\pi^{-}_{**}$			
$\overline{D}^{*0} \to \overline{D}^{0}\pi^{0}_{*}$	—	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(D^{-}\pi^{0}_{*})$
$\overline{D}^{*0} \to \overline{D}^{0}\gamma$	—	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^{*}(\overline{D}^{0}\gamma)$
$D^{**-} \to D^{*-}\pi^{0}_{**}$			
$D^{*-} \to \overline{D}^{0}\pi^{-}_{*}$	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})\epsilon^{*}$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})(1-\epsilon^{*})$
$D^{*-} \to D^{-}\pi^{0}_{*}$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(D^{-}\pi^{0}_{*})$	—
$D^{*-} \to D^{-}\gamma$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^{*}(D^{-}\gamma)$	—
$D^{**-} \to \overline{D}^{0}\pi^{-}_{**}$	—	—	$f^{**}(1-P_V)\frac{2}{3}$
$D^{**-} \to D^{-}\pi^{0}_{**}$	—	$f^{**}(1-P_V)\frac{1}{3}$	—
$\to D^{*-}l^{+}\nu$			
$D^{*-} \to \overline{D}^{0}\pi^{-}_{*}$	$f^{*}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})\epsilon^{*}$	—	$f^{*}\mathcal{B}^{*}(\overline{D}^{0}\pi^{-}_{*})(1-\epsilon^{*})$
$D^{*-} \to D^{-}\pi^{0}_{*}$	—	$f^{*}\mathcal{B}^{*}(D^{-}\pi^{0}_{*})$	—
$D^{*-} \to D^{-}\gamma$	—	$f^{*}\mathcal{B}^{*}(D^{-}\gamma)$	—
$\to D^{-}l^{+}\nu$	—	f	—

Table 3.17: The various B^{+} and B^{0} decay chains and their contributions to the composition of the three general categories of decay signatures ($D^{*-}l$, $D^{-}l$, and $\bar{D}^{0}l$); the branching ratio of $D^{*} \to X$ is denoted by "$\mathcal{B}^{*}(X)$".

	decay chain					relative efficiencies: trigger	relative efficiencies: reconstruction
$D^0 l$	$B^0 \rightarrow$		$D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	1.22 ± 0.02	0.90 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^0 D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	0.85 ± 0.01	0.92 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	π_{**}^-	$\overline{D}^0 \rightarrow$	$K\pi$	0.95 ± 0.01	0.91 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^- D^{*0} \rightarrow$	$\gamma \overline{D}^0 \rightarrow$	$K\pi$	0.83 ± 0.01	0.92 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^- D^{*0} \rightarrow$	$\pi_*^0 \overline{D}^0 \rightarrow$	$K\pi$	0.84 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow$			$\overline{D}^0 \rightarrow$	$K\pi$	1	1
	$B^+ \rightarrow$		$D^{*0} \rightarrow$	$\gamma \overline{D}^0 \rightarrow$	$K\pi$	1.32 ± 0.02	0.93 ± 0.01
	$B^+ \rightarrow$		$D^{*0} \rightarrow$	$\pi_*^0 \overline{D}^0 \rightarrow$	$K\pi$	1.32 ± 0.02	0.93 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	π_{**}^0	$\overline{D}^0 \rightarrow$	$K\pi$	0.83 ± 0.01	0.94 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^0 \overline{D}^{*0} \rightarrow$	$\gamma \overline{D}^0 \rightarrow$	$K\pi$	0.94 ± 0.01	0.95 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^0 \overline{D}^{*0} \rightarrow$	$\pi_*^0 \overline{D}^0 \rightarrow$	$K\pi$	0.84 ± 0.01	0.93 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^+ D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	0.83 ± 0.01	0.91 ± 0.01
$D^- l$	$B^0 \rightarrow$			$D^- \rightarrow$	$K\pi\pi$	1	1
	$B^0 \rightarrow$		$D^{*-} \rightarrow$	$\gamma D^- \rightarrow$	$K\pi\pi$	1.20 ± 0.02	0.96 ± 0.01
	$B^0 \rightarrow$		$D^{*-} \rightarrow$	$\pi_*^0 D^- \rightarrow$	$K\pi\pi$	1.19 ± 0.02	0.96 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	π_{**}^0	$D^- \rightarrow$	$K\pi\pi$	0.85 ± 0.01	0.95 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^0 D^{*-} \rightarrow$	$\gamma D^- \rightarrow$	$K\pi\pi$	0.78 ± 0.01	0.97 ± 0.01
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^0 D^{*-} \rightarrow$	$\pi_*^0 D^- \rightarrow$	$K\pi\pi$	0.78 ± 0.01	0.95 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	π_{**}^+	$D^- \rightarrow$	$K\pi\pi$	1.06 ± 0.02	0.77 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^+ D^{*-} \rightarrow$	$\gamma D^- \rightarrow$	$K\pi\pi$	0.91 ± 0.01	0.82 ± 0.01
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^+ D^{*-} \rightarrow$	$\pi_*^0 D^- \rightarrow$	$K\pi\pi$	0.91 ± 0.01	0.81 ± 0.01
$D^{*-} l$	$B^0 \rightarrow$		$D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	1	1
	$B^0 \rightarrow$	$D^{**-} \rightarrow$	$\pi_{**}^0 D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	0.68 ± 0.01	1.05 ± 0.02
	$B^+ \rightarrow$	$\overline{D}^{**0} \rightarrow$	$\pi_{**}^+ D^{*-} \rightarrow$	$\pi_*^- \overline{D}^0 \rightarrow$	$K\pi$	0.66 ± 0.01	0.99 ± 0.02

Table 3.18: Relative trigger and reconstruction efficiencies of the 21 B^+ and B^0 signal decay chains.

3.4 Monte Carlo samples

Samples of Monte Carlo simulated b-hadron decays are used at several stages along this dissertation. Their applications include the following: The estimation of the signal yields in the selection optimization procedure. The determination of data samples composition, as well as of characteristics of signal and physics type background components, such as mass and proper decay time distributions. Reconstruction efficiencies, together with treatment of trigger, selection, and partial reconstruction effects on proper decay time distributions. Testing of fitting framework.

A Monte Carlo sample is generated for each nominal B signal mode contributing to our data samples. Samples of additional b-hadron decays are also generated to estimate physics type backgrounds. These include exclusive samples where a single B meson decay chain is enforced, as well as inclusive samples where various B decays contributing to a common final state signature are included. As examples of the latter case we have: $B \rightarrow J/\psi X$ with $J/\psi \rightarrow \mu^+\mu^-$; $B \rightarrow \bar{D}^0 X$ with $\bar{D}^0 \rightarrow K^+\pi^-$; or $B^{+,0} \rightarrow D^- l\nu X$ with $D^- \rightarrow K^+\pi^-\pi^-$. In order to produce such generic inclusive samples in an efficient fashion only the desired decays are generated. This follows from the selection of the channels leading to a specified final state (such as the J/ψ or D meson modes mentioned) from a default global decay table, in which the relative branching ratios need to be re-evaluated [59].

Generation and decay of b-hadrons

For the generation and the fragmentation of b-quarks we employ the `Bgenerator` [60] Monte Carlo program. This is based on Next to Leading Order QCD calculations [61] and uses the Peterson fragmentation function [62]. The generation process involves a single b-quark, which then fragments to a b-hadron. No additional fragmentation products or proton remnants are present.

For simulating b-hadron decays we use the `EvtGen` [63] program, which has been extensively tuned by the experiments operating at the $\Upsilon(4S)$ resonance.

Realistic simulation

The detector response to the generated events is provided by the `Geant` [64] simulation framework, which models the detector geometry and the behavior of its active components at hit level. Time dependent detector inefficiencies, which arise for example from temporary disablement of subdetector components, are taken into account in the simulation, and the data taking period is simulated run-by-run with the event numbers per run calculated on the basis of the integrated luminosity of the run.

The data acquisition and trigger systems response is also simulated and the output of the simulation mimics the real raw data structure. Events are then passed through the production stage during which physics objects (*e.g.* tracks, leptons, jets, *etc*) are created. Finally, the simulated samples are stored in the form of `BStntuples` [53], and the analysis selection and reconstruction criteria are applied.

The agreement of relevant kinematic distributions obtained from the data and the simulated samples was validated. Where applicable, Monte Carlo reweighting was performed in order to account for run-by-run prescaling of the trigger paths.

3.5 Résumé

The selection and reconstruction criteria employed for collecting and forming the various data samples used in this dissertation have been presented.

The raw data selected by the online trigger system is sent for permanent storage, being then passed to the production stage where the most detailed calibrations and reconstruction algorithms are applied. Such datasets undergo further preparation, where the event tracks and vertices are refit and constructed under specific particle hypotheses. The collections of particle candidates – such as J/ψ, D, and B mesons – are pre-selected and stored in `BStntuples` which contain the candidates information. Optimized selection criteria requirements are then applied to form the final collection of B meson candidates. The relevant information, for the next stages of the analysis, pertaining to each individual B candidate is retrieved and appended to an `ascii` file; these selected data include mass, momentum, decay length, as well as flavor tagging information. For each decay sub-sample such a data file is constructed which contains the required input to be provided to the unbinned likelihood framework where the fits of the data are performed — which constitutes the subject of the following chapters. The sequence of data formats and steps may be summarily represented as

$$\texttt{Datasets} \rightarrow \texttt{Pre-selection} \rightarrow \texttt{BStntuple} \rightarrow \texttt{Selection} \rightarrow \texttt{Fit input file}.$$

The approximate B signal yields for the various sample categories, reconstructed from the ≈ 355 pb^{-1} of data analyzed, are summarized as

	B^+	B^0	B_s
$J/\psi K$	5,000	1,800	–
$D\pi(\pi\pi)$	11,200	17,300	1,100
Dl	130,000	130,000	13,000

The composition of the data samples has been studied, and the various signal and background components were assessed and characterized. This was achieved in part through the generation and use of Monte Carlo samples, which were described. These samples were used as well in the selection optimization procedure, and further provide necessary input to be employed in subsequent chapters.

Chapter 4

Measurement of B mesons lifetimes

4.1 Fitting technique

We develop a fitting framework based on the maximum likelihood estimation method, which we use to extract the parameters of interest from the B data samples. It is implemented in an *unbinned* fashion, where the input information provided directly to the fitter is that pertaining to the individual B meson candidates (these will be referred to as *events*). The unbinned likelihood description allows to explore more thoroughly the information contained in the data samples.

4.1.1 Maximum likelihood

The likelihood function $\mathcal{L}$ is a measure for the probability of observing the data set at hand, characterized by measurements $\mathbf{t} = \{t_i\}$, given the parameters $\boldsymbol{\tau}= \{\tau_j\}$. It is defined here as the joint probability density for the independent measurements $\mathbf{t}$ as a function of the parameters $\boldsymbol{\tau}$,

$$\mathcal{L}(\mathbf{t}|\boldsymbol{\tau}) = \prod_{\{t_i\}} \mathcal{P}(t_i|\boldsymbol{\tau})\,. \tag{4.1}$$

The model parameters $\boldsymbol{\tau}$ are varied, and their best estimates, denoted by *fitted* parameter values, correspond to the set which maximizes $\mathcal{L}$. This is implemented using the minimization program `Minuit` [?, 66], provided with the score function $-2\ \ln(\mathcal{L})$.

In case a fit to a common set of parameters is to be performed simultaneously to several B samples $\{s\}$, the product of the likelihoods $\mathcal{L}_s$ of individual samples is formed, and the quantity to be minimized then becomes

$$-2 \sum_{\{s\}} \ \ln\ (\mathcal{L}_s)\,. \tag{4.2}$$

The necessary normalization of the underlying probability density function, $\int \mathcal{P} d\mathbf{t} =$

1, for each event, is implemented analytically, resulting in an optimization of the fitter speed and accuracy.

A given candidate has certain probabilities of belonging to the different components which form the sample being fit. The event likelihood is thus formed of terms describing these various components. In general, these contain the *signal* and *background* classes. Signal refers to those sample components the model description of which involve directly the parameters of interest. For the background components it is sufficient in general to provide an empirical description of the corresponding distributions. Denoting by f_B the fraction of background, and by $(1 - f_B)$ the fraction of signal, the likelihood becomes

$$\mathcal{L} = \prod_i \left[(1 - f_B) \cdot \mathcal{P}_i^S + f_B \cdot \mathcal{P}_i^B\right] \tag{4.3}$$

where $\mathcal{P}_i^S$ and $\mathcal{P}_i^B$ denote the signal and background likelihood components, evaluated for the i-th candidate.

More generally, the signal and background classes may be both formed of distinct components, which benefit from separate treatment in the likelihood model. The candidates forming such components may belong to one of the following categories:

- The nominal signal, where all tracks forming the candidate were correctly reconstructed as the B daughter particles of the decay channel being reconstructed.
- Partially reconstructed b-hadrons, where some tracks have not been reconstructed, or the decay products contain neutral particles.
- Misidentified tracks, where a particle has been wrongly reconstructed due to mis-identification of a track with a lepton, pion, kaon, or proton.
- Misreconstructed B-decays, where, although the tracks have been assigned the correct masses, the resulting particles did not originate from the B decay being reconstructed.
- The combinatorial background, corresponding to other track combinations with an invariant mass lying in the mass fitting region.

The various specific components which contribute to each of the samples being fit have been identified in Section 3.3.

Each sample component is assigned corresponding PDFs in the spaces of the input variables. The likelihood is then formed of the joint PDFs, $\mathcal{P}^\alpha$, obtained for each component,

$$\mathcal{L} = \prod_i \sum_\alpha f_\alpha \mathcal{P}_i^\alpha \,, \tag{4.4}$$

where the indices i, α run over the number of events and number of sample components, respectively, and f_α denote the component fractions, with $\sum_\alpha f_\alpha = 1$.

While nominal signal and combinatorial background are in general the dominant contributions, in the following paragraphs we identify the specific realizations also of the other categories in our data samples.

4.1.2 Fit input variables

The quantities which serve as input to the fit include the mass, the proper decay time, and the proper decay time resolution of the reconstructed B candidates. These variables are introduced in the following, and will be denoted by the lower-case symbols m, t and σ_t, respectively.

Proper decay time

The principal measurements performed in this dissertation correspond to parameters which determine the distribution of the reconstructed proper decay time of B meson candidates. This quantity is related to the observed B candidate decay distance, and is defined as

$$t \equiv L_{xy}\frac{M^B}{p_T}. \tag{4.5}$$

Here p_T is the transverse momentum of the reconstructed system. L_{xy} is the projection on the direction of the transverse momentum p_T of the displacement vector in the transverse plane, $\vec{d}$, from the B meson production vertex to the its decay vertex,

$$L_{xy} = \vec{d}\cdot\frac{\vec{p}_T}{|p_T|}. \tag{4.6}$$

The world average value for the mass of the involved B meson, M^B, is used in the definition of t. For the signal PDF the fit estimate of the B mass is not used because it would introduce unnecessary uncertainties in the determination of t.

In case of fully reconstructed signal candidates, t is identified with the proper decay time of the B meson, $t = t^B$,

$$t^B = \frac{L^B}{\beta\gamma} = L^B\frac{M^B}{p^B} = L^B_{xy}\frac{M^B}{p^B_T}, \tag{4.7}$$

where L, β, and γ are the flight distance, velocity, and Lorentz boost factor of the B candidate.

For partially reconstructed decays, the B transverse momentum p^B_T, and thereby the B proper decay time t^B, cannot be measured. For example, the neutrino from the decay $B \to Dl\nu$ is not detected. In these cases t is sometimes also referred to as *pseudo*-proper decay time of the B candidate, and differs from t^B by a kinematic correction factor relating the B and Dl systems, $t^B = t\cdot\kappa$, defined by

$$\kappa = \frac{L^B_{xy}}{L^{Dl}_{xy}}\frac{p^{Dl}_T}{p^B_T}. \tag{4.8}$$

An average distribution, $\mathcal{F}(\kappa)$, for the κ-factor is obtained from Monte Carlo simulation, and this is a necessary ingredient when forming the signal PDF.

We will generally refer to the measured quantity t as proper decay time, while the prefix *pseudo* should be implicit for decays which are not completely reconstructed.

Although t does not have a specific meaning for background components, it remains a well defined quantity, and is employed in both signal and background PDFs.

Proper decay time resolution

The detector resolution and track fitting methods imply a finite precision for the measurement of the B decay-length, $\sigma_{L_{xy}}$, and therefore in the determination of the proper decay time t,

$$\sigma_t = \sigma_{L_{xy}} \cdot \frac{M^B}{p_T}. \tag{4.9}$$

Contributions to the proper time resolution associated to uncertainties on the reconstructed transverse momentum and on the world average B masses are negligible. This information ought to be included in the relevant PDFs for the signal. Indeed, the proper decay time resolution is a determining factor in the B_s mixing analysis when attempting to resolve rapid, time dependent oscillations.

An additional complication arises from the fact that the uncertainty returned by the vertex fitter is in general underestimated, arising from the complexity of describing the track hit uncertainties. This requires a correcting *scale factor* to be applied,

$$\sigma_t \quad \mapsto \quad S_t \cdot \sigma_t \,. \tag{4.10}$$

This scale factor S_t is extracted directly from the data samples, as a fit parameter adjusting the Gaussian width describing a prompt component in the proper decay time distribution. However, such a prompt contribution is only present in $J/\psi K$ modes, and is suppressed by the trigger selection in the other B samples. In those t biased samples of B^+ and B^0 decays a common average value for S_t is used. A more detailed study is performed in Section 4.7, in anticipation to the study of B_s oscillations, where the proper time uncertainties becomes a critical issue. In samples of B_s meson decays the scale factor S_t has an event dependent value, computed as described therein.

Mass

The inclusion of the reconstructed candidate mass, m, is crucial for separating signal from background components, and corresponding PDFs are included in the likelihood description.

In fully reconstructed decays it is the mass of the B candidate which is provided as input to the fit. For partially reconstructed decays, one uses instead the mass of the D meson candidate, as the B momentum is not completely reconstructed.

Although the B, or D, meson masses, M, will be parameters of the fit, the corresponding fit results for M should not be taken as definite measurements, the relevant systematic effects not being evaluated. We do not include the uncertainty on the candidate mass determination in the likelihood description.

4.1.3 Likelihood factors

The distributions of the fit input variables presented in the previous section – $m,\ t,\ \sigma_t$ – are described by likelihood factors, which for each sample component (e.g. signal, backgrounds) are given by

$$\mathcal{P} = L_m\, L_t\, L_{\sigma_t}. \tag{4.11}$$

The mass and proper time spaces are disjoint, and the corresponding probabilities multiply. The probability density for the proper time, t, depends on the smearing effect determined by the resolution σ_t. The factorization (4.11) is then

$$\mathcal{P}(m,t,\sigma_t) = \mathcal{P}(m)\cdot\mathcal{P}(t,\sigma_t) = \mathcal{P}(m)\cdot\mathcal{P}(t|\sigma_t)\cdot\mathcal{P}(\sigma_t). \tag{4.12}$$

These factors are the probability density functions of the corresponding argument variables, and are thus unit normalized

$$\int L_m\, dm = \int L_t\, dt = \int L_{\sigma_t}\, d\sigma_t = 1\ .$$

The forms of the PDFs for mass and proper time, L_m and L_t, are given in the sections below in terms of the fitting parameters. The σ_t factor is obtained directly from data as follows: for the background, $L^B_{\sigma_t}$ is given by the σ_t distribution for events in the mass side-band region; for signal events, $L^S_{\sigma_t}$ corresponds to the normalized σ_t distribution obtained for candidates in the mass signal region, after mass side-band subtraction. These distributions are illustrated in Figure 4-1 for the indicated modes.

The components contributing to each sample are identified in Section 3.3 and summarized in Table 4.1. Separate likelihood terms as in (4.11) have to be provided for each component.

	signal	background
$J/\psi K$	B^+, B^0	combinatorial, Cabibbo-suppressed, K^* swap
$D\pi(\pi\pi)$	B^+, B^0, B_s	combinatorial, Cabibbo-suppressed, fully and partially rec.
Dl	B^+/B^0, B_s	combinatorial, physics, fakes

Table 4.1: Summary of data sample component categories, from Section 3.3.

4.2 Proper time likelihood formalism for signal

The proper decay time distribution for the signal is formed of an exponential decay characterized by the B meson lifetime τ, smeared by the detector resolution, and corrected by possible trigger, selection, and partial reconstruction effects.

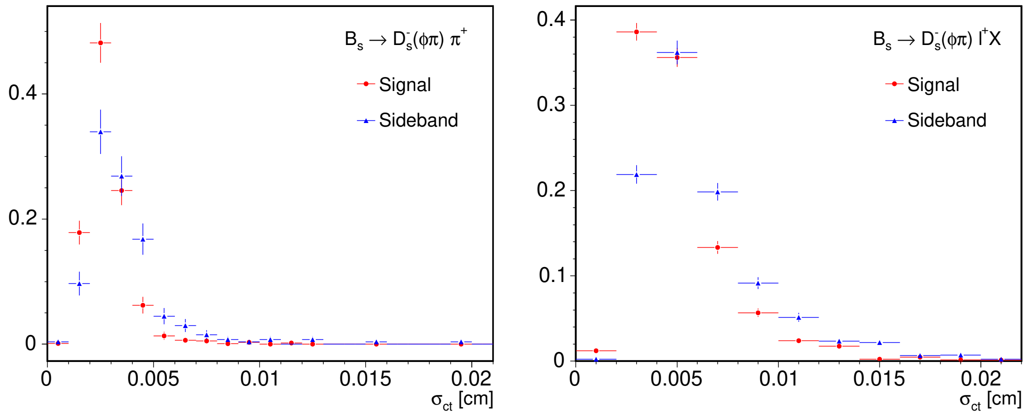


Figure 4-1: Distributions of calibrated proper decay time resolution σ_t, for signal and background events, for the main hadronic (left) and semileptonic (right) B_s decay modes.

4.2.1 Fully reconstructed and unbiased proper time

The expected true proper decay time follows an exponential decay form

$$E(t;\tau) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \theta(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{\tau} e^{-\frac{t}{\tau}}, & t \geq 0 \end{cases} \tag{4.13}$$

determined by the B meson lifetime, τ. The semicolon is used here to separate the main function argument from its parameters. The finite vertex resolution, described by a Gaussian function, induces the smearing of the expected proper time. The width of the resolution function is given by the measured proper time uncertainty for the individual event,

$$G(t;\sigma_t) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{t^2}{2\sigma_t^2}}. \tag{4.14}$$

The observable t is given by the sum of the true proper decay time (4.13) with the effect (jitter) of the detector resolution (4.14). The resulting PDF for t is accordingly given by the convolution of the two distributions,

$$\begin{aligned} L(t|\sigma_t, \tau) &= E(t;\tau) \otimes G(t;\sigma_t) \\ &= \int_{-\infty}^{\infty} \frac{1}{\tau} e^{-t'/\tau} \theta(t') \cdot \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t')^2}{2\sigma_t^2}}\ dt' \\ &= \frac{1}{2\tau}\ e^{-\frac{1}{\tau}(t-\frac{\sigma_t^2}{2\tau})} \cdot \mathrm{Erfc}\left(\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right) \end{aligned} \tag{4.15}$$

The complementary error function, $\mathrm{Erfc}(z) \equiv \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-u^2} du$, has been employed. That (4.15) has unit normalization follows from the fact that the normalization of the convolution of a function with a Gaussian coincides with the normalization of the function itself,

$$\int_{-\infty}^{+\infty} f(t) \otimes G(t) dt = \int_{-\infty}^{+\infty} f(t') \int_{-\infty}^{+\infty} G(t-t') dt\ dt' = \int_{-\infty}^{+\infty} f(t') dt' . \tag{4.16}$$

4.2.2 Fully reconstructed and biased proper time

Certain selection cuts, applied at either event triggering or reconstruction levels, may affect the expected proper time distribution, causing deviations from the shape given by (4.15).

For example, suppose a selection level threshold is applied directly to the proper time, $t > t_0$. Or, as it is most often the case, that such a cut is applied to the B meson decay-length or its significance, $L_{xy}/\sigma_{L_{xy}} > \alpha$. This translates again to applying a selection directly on the proper time, $t > t_\alpha$, where for each event the threshold is given by $t_\alpha = \alpha \cdot \left(\sigma_{L_{xy}} \frac{m_B}{p_T}\right)$. We note that such selection criteria are based on reconstructed, thus smeared, observables. The PDF becomes

$$L(t|\sigma_t, \tau, t_\alpha) \;=\; \frac{1}{\mathcal{N}} \cdot \frac{1}{2\tau}\, e^{-\frac{1}{\tau}(t - \frac{\sigma_t^2}{2\tau})} \cdot \mathrm{Erfc}\left(\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t \tau}\right) \cdot \theta(t - t_\alpha)\,, \tag{4.17}$$

where the step function, $\theta(t)$, directly implements the selection $t > t_\alpha$. A normalization factor, $\mathcal{N}$, needs now to be evaluated, and this computation is to be performed for the individual events. For the case at hand, an analytical integration of (4.17) gives

$$\mathcal{N}(\sigma_t, \tau, t_\alpha) \;=\; \frac{1}{2}\left[e^{-\frac{1}{\tau}(t_\alpha - \frac{\sigma_t^2}{2\tau})} \cdot \mathrm{Erfc}\left(\frac{\sigma_t^2 - t_\alpha\tau}{\sqrt{2}\sigma_t \tau}\right) + \mathrm{Erfc}\left(\frac{t_\alpha}{\sqrt{2}\sigma_t}\right)\right] . \tag{4.18}$$

We have thus seen that in case a selection cut is imposed on an observable directly related to the proper decay time t, the effect on the PDF can be implemented through multiplication with an acceptance function, $\mathcal{E}(t)$, which in those cases is simply a step function, $\mathcal{E}(t) = \theta(t)$, centered at the corresponding cut value for each event. More generally, however, there are cuts which are not expressed as simple threshold conditions on t. The track impact parameter criteria is an example of such a cut. In these more complex cases, the sculpting effect on the proper time PDF can be described in a similar fashion,

$$\begin{aligned} L(t|\sigma_t, \tau) &= \frac{1}{\mathcal{N}} \cdot [\ E(t;\tau) \otimes G(t;\sigma_t)\ ] \cdot \mathcal{E}(t) \\ &= \frac{1}{\mathcal{N}} \cdot \frac{1}{2\tau}\, e^{-\frac{1}{\tau}(t - \frac{\sigma_t^2}{2\tau})} \cdot \mathrm{Erfc}\left(\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t \tau}\right) \cdot \mathcal{E}(t) \end{aligned} \tag{4.19}$$

via an efficiency function, $\mathcal{E}(t)$. The normalization factor

$$\mathcal{N}(\sigma_t,\tau) \;=\; \int_{-\infty}^{\infty} \frac{1}{2\tau}\; e^{-\frac{1}{\tau}(t-\frac{\sigma_t^2}{2\tau})} \cdot \mathrm{Erfc}\left(\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right) \cdot \mathcal{E}(t)\; dt \tag{4.20}$$

needs as before to be evaluated for each single event.

4.2.3 Partially reconstructed and biased proper time

For cases where the B meson is not fully reconstructed, the measured input observable t is related to the B meson proper decay time through a kinematical factor, obtained from Monte Carlo simulation. For semileptonic decays, such κ-factor has been defined in (4.8).

In this case, the original lifetime exponential of (4.13) takes the form

$$E'(t;\tau) = E(t;\tau) \otimes_\kappa \mathcal{F}(\kappa) = \int \frac{\kappa}{\tau} e^{-\frac{\kappa t}{\tau}} \theta(\kappa t) \cdot \mathcal{F}(\kappa)\; d\kappa \tag{4.21}$$

where $\mathcal{F}(\kappa)$ is the normalized κ-factor distribution, and the κ-factor smearing operator $\otimes_\kappa$ is defined by the second equality. In a similar fashion, the final PDF becomes

$$\begin{aligned} L(t|\sigma_t,\tau) \;&=\; \frac{1}{\mathcal{N}} \cdot [\; E(t;\tau) \otimes_\kappa \mathcal{F}(\kappa) \otimes G(t;\sigma_t)\;] \cdot \mathcal{E}(t) \qquad (4.22)\\ &=\; \frac{1}{\mathcal{N}} \cdot \int \frac{\kappa}{2\tau}\; \exp\left(-\frac{\kappa}{\tau}(t - \frac{\kappa\sigma_t^2}{2\tau})\right) \cdot \mathrm{Erfc}\left(\frac{\kappa\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right) \cdot \mathcal{F}(\kappa)\; d\kappa \cdot \mathcal{E}(t)\,. \end{aligned}$$

The construction of the efficiency function, $\mathcal{E}(t)$, is specified in a later section, as is the evaluation of the normalization factor

$$\mathcal{N}(\sigma_t,\tau) \;=\; \int \left[\int_{-\infty}^{+\infty} e^{-\frac{\kappa}{\tau}t}\; \mathrm{Erfc}\left(\frac{\kappa\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right)\; \mathcal{E}(t)\; dt\right] \cdot \frac{\kappa}{2\tau}\; \exp\left(\frac{\kappa^2\sigma_t^2}{2\tau^2}\right) \cdot \mathcal{F}(\kappa)\; d\kappa\,. \tag{4.23}$$

which needs to be performed for each event.

4.3 Describing trigger and reconstruction effects on proper time

The shape of the proper decay time distribution of a sample is potentially affected by trigger and reconstruction selection criteria. Such effects need to be incorporated into the relevant PDFs, along with the physics and resolution contributions. This is achieved by introducing an appropriate efficiency function, $\mathcal{E}(t)$, which multiplies the proper time signal PDF. Here we show how such functions are constructed, based on Monte Carlo simulation of the involved B decays, in which full detector effects, trigger, and offline selection are incorporated. The construction of the κ-factor distribution, necessary for partially reconstructed modes, is also presented.

4.3.1 The t-efficiency function

The t-efficiency function is defined as a function of the *reconstructed* proper decay time. This is appropriate as the selection criteria whose effect is to be described are themselves applied to reconstructed observables. An efficiency function defined in terms of un-smeared quantities is not compatible with a straightforward implementation of the t-resolution in the PDF [67].

The efficiency function $\mathcal{E}(t)$ relates the proper time distribution of a sample of events obtained with the given selection criteria, to the corresponding proper time distribution of the sample without the bias introduced by the selection. Indeed, it is constructed as the ratio of such two distributions. Specifically, the biased distribution is obtained directly from Monte Carlo simulation, where the description of the full biasing selection is included. And the denominator, describing the corresponding unbiased distribution, is obtained in an analytical fashion: the lifetime exponential smeared with the t-resolution, also obtained from the same Monte Carlo sample. That is, the t-efficiency function is given, for fully reconstructed decays, by

$$\mathcal{E}(t) \quad = \quad \frac{t\text{−distribution after selection}}{\sum_{\{\sigma_t\}} \frac{1}{\tau} e^{-t/\tau} \otimes G(t;\sigma_t)}. \tag{4.24}$$

The parameter τ denotes the B^+, B^0 or B_s, as appropriate, lifetime value, identical to that used in the simulation from which the numerator is constructed [1]. The sum is weighted over the reconstructed proper time uncertainty distribution of the Monte Carlo simulated events. Figure 4-2 illustrates the t-efficiency distributions obtained for two selected hadronic modes.

For partially reconstructed decays a similar definition as in (4.24) is employed

$$\mathcal{E}(t) \quad = \quad \frac{t\text{−distribution after selection}}{\sum_{\{\sigma_t\}} \frac{1}{\tau} e^{-t/\tau} \otimes G(t;\sigma_t) \otimes_\kappa \mathcal{F}(\kappa)} \tag{4.25}$$

where the smearing with the κ-factor distribution $\mathcal{F}(\kappa)$ has been included in the construction of the unbiased distribution.

A t-efficiency function is constructed for each of the B mesons whose semileptonic decay channels contribute to a given sub-sample. This is done based on an inclusive Monte Carlo simulation containing all relevant channels. Figure 4-3 illustrates the t-efficiency distributions obtained for selected decay Dl sub-samples. These reflect the distinct pseudo proper time distributions which characterize the samples of B and D trigger types. For candidates for which the B-daughter lepton is not a trigger track (D trigger type) the t distribution is accordingly not suppressed as much by the trigger at low pseudo proper times.

We note that an overall scaling $\sigma_t \mapsto r \cdot \sigma_t$ is applied, in the definitions (4.24) and (4.25) above, to the reconstructed proper time uncertainties of the Monte Carlo events. The factor r is the Gaussian width of the distribution $(t^{\text{reconstructed}} - t^{\text{true}})/\sigma_t$, obtained from the Monte Carlo sample, and its numerical value is about 1.1.

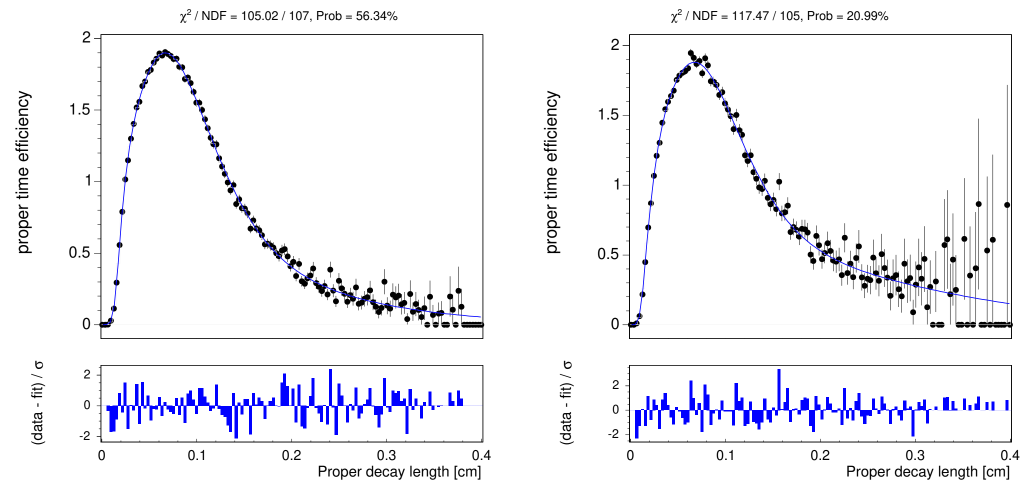


Figure 4-2: t-efficiency distribution and fitted function $\mathcal{E}(t)$for the fully reconstructed $B^+ \to \bar{D}^0\pi^+$ (left) and $B_s \to D_s^-[\phi\pi^-]\pi^+$ (right) decay channels

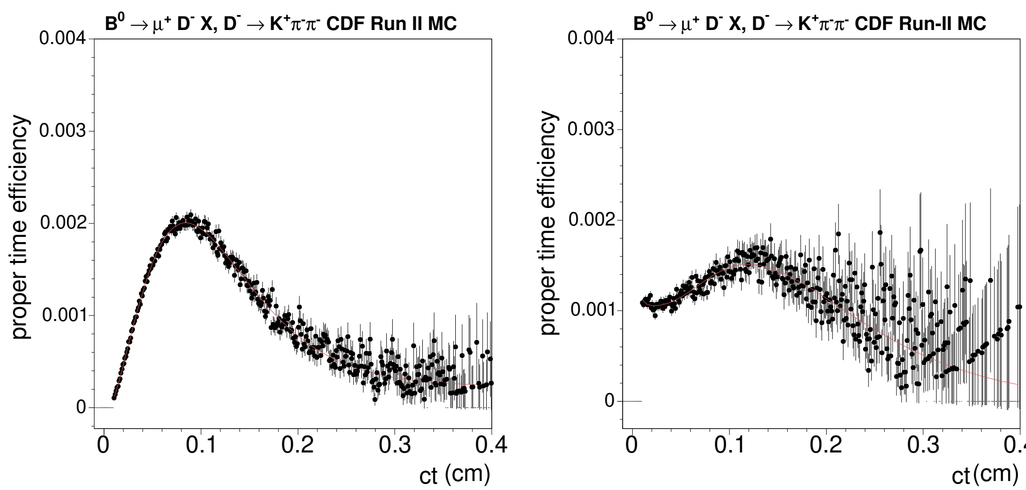


Figure 4-3: t-efficiency distribution and fitted function $\mathcal{E}(t)$for the partially reconstructed $B^0 \to D^-\mu^+X$ channels, for B trigger (left) and D trigger (right) type candidates.

4.3.2 The κ-factor distribution

The distribution $\mathcal{F}(\kappa)$ of κ-factors (4.8) is constructed for each B meson type present in each of the 24 Dl decay sub-samples from an inclusive Monte Carlo sample.

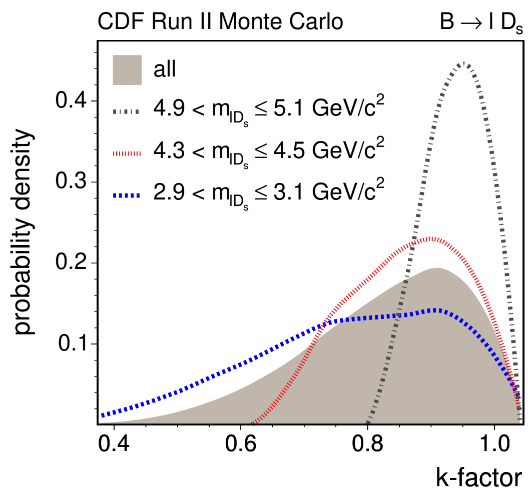


Figure 4-4: Distribution $\mathcal{F}(\kappa)$ of κ-factors, also shown for distinct ranges of the $D_s l$ mass.

The κ-factor distribution is correlated to the mass of the Dl system, m_{Dl}. Indeed, the fraction of the B momentum which is reconstructed, quantified by the κ-factor, tends to be larger for candidates with larger m_{Dl}. That is, for candidates with m_{Dl} close to the B mass, not much momentum could have been carried away. Most importantly, such events tend to contribute a reduced spread for $\mathcal{F}(\kappa)$ as the phase space available for the missing neutrino, and additional non-reconstructed particles, is also reduced. This is shown explicitly in Figure 4-4. We exploit these dependences in the case of the $D_s l$ samples, by assigning separate κ-factor distributions depending on the candidate's $m_{D_s l}$. Specifically the following ranges are used:

$$m_{D_s l} \in [1.0, 3.0],\ [3.0, 3.5],\ [3.5, 4.0],\ [4.0, 4.5],\ [4.5, 5.3]\ \mathrm{GeV/c^2}\ .$$

This classification is performed in view of the expected increased significance resulting from the better characterization of the κ-factor resolution of the candidates.

4.3.3 PDF normalization factor

We now address how the evaluation of the necessary PDF normalization factors, which have been defined in (4.20) and (4.23), is performed. We recall that these depend not only on fitting parameters, such as τ, but also on properties of the event, such as σ_t. Accordingly, the computation of these quantities need to be performed at each iteration step of the fit, for each single event. The time required by the fit convergence process becomes a serious issue that needs to be addressed. We derive a convenient parameterization of the t-efficiency function which will allow us to perform an analytical integration in t of the PDFs in (4.19) and (4.22).

A given shape can in general be approximavely described by a number of templates. For example, in the rational approximation a shape is described by the ratio of two polynomials of certain degrees. Here however we are interested in finding simpler parameterizations, which will render the normalization integral feasible, by adding the least complexity to the remaining part of the PDF (to be multiplied by the t-efficiency itself). This reasoning leads to considering the class of functions represented by the basis $\{t^n e^{-t/\tau_n}\}$, resulting in a parameterization of the t-efficiency function involving terms of the following type

$$\mathcal{E}(t) \sim \sum_n a_n \, t^n \, e^{-\frac{t}{\tau_n}} \, \theta(t-\zeta_n). \tag{4.26}$$

We perform a fit to the t-efficiency curve defined for the t-biased modes, using a combination of templates of the form given in (4.26) up to second order, to determine the involved parameters $\{a_n, \tau_n, \zeta_n\}$. The following explicit forms are found convenient for the $D\pi(\pi\pi)$ and Dl modes, respectively:

$$\mathcal{E}(t) = \begin{cases} \sum_{j=0}^{2} a_j (t-\zeta_j)^2 e^{-\frac{t}{\tau_j}} \theta(t-\zeta_j) & \text{(hadronic)} \\ (a_0 + a_1\, t + a_2\, t^2) \left(f\, e^{-\frac{t}{\tau_1}} + (1-f)\, e^{-\frac{t}{\tau_2}} \right) \theta(t-\zeta) & \text{(semileptonic)} \end{cases} .$$

The fits to the t-efficiency curves are illustrated in Figure 4-2 for selected hadronic channels, and in Figure 4-3 for semileptonic modes where it can be seen that the latter expression accommodates well both B and D trigger type sub-samples.

The template expressions parameterizing the t-efficiency curves involve only terms of the form of (4.26). These then render the integrations in t expressed in (4.20) and (4.23) analytically feasible. The result of such calculations is explicitly included in Appendix D.1. The additional integration involving the κ-factor distribution necessary for the semileptonic samples, both in the evaluation of the PDF (4.22) and its normalization (4.23), is performed numerically, by discretizing the $\mathcal{F}(\kappa)$ distribution and summing over the corresponding bins.

4.4 Mass PDF

The likelihood description of the mass subspace provides important separation power among the various signal and background components present in our samples.

For the nominal signal component, the width of the mass distribution is dominated by the detector resolution. The corresponding PDF will be given by a Gaussian function, unit-normalized in the fitting range $(M_{\min}, M_{\max})$,

$$G(m; M, \sigma_m, M_{\min}, M_{\max}) = \frac{\frac{1}{\sqrt{2\pi}\sigma_m} e^{-\frac{1}{2}\left(\frac{m-M}{\sigma_m}\right)^2}}{\frac{1}{2}\left[\mathrm{Erf}\left(\frac{M_{\max}-M}{\sqrt{2}\sigma_m}\right) + \mathrm{Erf}\left(\frac{M-M_{\min}}{\sqrt{2}\sigma_m}\right)\right]}, \tag{4.27}$$

with mean and width given by the fit parameters M and σ_m. The error function, $\mathrm{Erf}(z) \equiv \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du$, is used to express the normalization factor.

The combinatorial background is modeled through an empirical parameterization, as determined from the mass side-band region. For the $J/\psi K$ and Dl samples, a linear model $am + b$ is employed. Upon unit-normalization in the mass fitting region $(M_{\rm min}, M_{\rm max})$, it is given by

$$A(m; a, M_{\rm min}, M_{\rm max}) \quad = \quad a\,m + \frac{1}{M_{\rm max} - M_{\rm min}} \left[1 - \frac{a}{2}\left(M_{\rm max}^2 - M_{\rm min}^2\right)\right]. \tag{4.28}$$

For the $D\pi$ modes an exponential shape is also found useful to model the combinatorial shape,

$$B(m; b, M_{\rm min}, M_{\rm max}) = \frac{be^{-bm}}{e^{-bM_{\rm min}} - e^{-bM_{\rm max}}}. \tag{4.29}$$

In order to allow for a better determination of the background shape parameters, a preliminary fit is performed in an extended mass range.

$B^+ \to J/\psi K^+$

The nominal signal and combinatorial background are described by a Gaussian (4.27) and a linear (4.28) model, respectively,

$$\begin{aligned} L_m^{J/\psi K}(m|M, \sigma_m) &= G(m; M, \sigma_m, M_{\rm min}, M_{\rm max})\,, \\ L_m^{\rm comb}(m) &= A(m|a)\,. \end{aligned}$$

The contribution of the Cabibbo-suppressed $B^+ \to J/\psi\pi^+$ mode is fixed to 2.5% of the total B^+ signal, and its mass shape is taken from Monte Carlo simulation, being approximated by a Landau distribution. The template is positioned relative to the mass fit parameter, M, and is otherwise forced to those values found from simulation,

$$L_m^{J/\psi\pi}(m|M) \quad = \quad \frac{\left(m - M + \delta_{J/\psi\pi}\right)^6 e^{\frac{M-m-\delta_{J/\psi\pi}}{b}}}{a_6 b^7 - e^{\frac{M-M_{\rm max}-\delta_{J/\psi\pi}}{b}}\left[a_6 b^7 + \sum_{i=1}^{i=6} a_i b^i \left(M_{\rm max} - M + \delta_{J/\psi\pi}\right)^{7-i}\right]}\,,$$

with $\vec{a} \equiv (1, 5, 30, 120, 360, 720)$, $\delta_{J/\psi\pi} = -3$ MeV/c^2, $b = 9$ MeV/c^2.

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

The description of the nominal signal and combinatorial background components is identical to that of the previous decay.

Candidates formed with an incorrect (*swapped*) mass assignment to the kaon-pion pair forming the K^* meson amount to a fraction of 16.9% relative to the nominal signal. The shape of its mass distribution is modeled by a Gaussian function centered at the nominal signal mass, as determined from Monte Carlo simulation,

$$L_m^{\rm swap}(m|M) \quad = \quad G(m; M, \sigma_{\rm swap}, M_{\rm min}, M_{\rm max}) \qquad \text{with} \quad \sigma_{\rm swap} = 25 \text{ MeV/c}^2\,.$$

$B^+ \to \bar{D}^0\pi^+$

The mass distribution for the nominal signal is better described by a pair of Gaussian functions (4.27),

$$L_m^{D\pi}(m|M,\sigma_m) \;=\; (1-f)\,G(m;M,\sigma_m,M_{\min},M_{\max}) + f\,G(m;M,r\sigma_m,M_{\min},M_{\max})\,;$$

the factors f, r are found from a preliminary fit to the data performed in the wide mass range.

The combinatorial background is described by the combination of a linear (4.28) and an exponential (4.29) models,

$$L_m^{\text{comb}}(m|b) \;=\; (1-f)\,A(m|a) + f\,B(m;b,M_{\min},M_{\max})\,;$$

a flat $a = 0$ linear model is found to suffice. The relative fraction f of flat and exponential components as well as the exponential decay parameter b are determined from the wide mass range fit. Only the overall combinatorial background fraction is allowed to float in the narrow mass range fit.

The mass template for the misreconstructed Cabibbo-suppressed $B^+ \to \bar{D}^0K^+$ mode is determined from Monte Carlo simulation. It is given by a Gaussian model whose width, relative fraction and displacement with respect to the nominal signal peak are fixed from simulation,

$$L_m^{DK}(m|M) \;=\; G(m;M-\delta_{DK},\sigma_{m_{DK}},M_{\min},M_{\max})\,,$$

with $\delta_{DK} = 57$ MeV/c^2 and $\sigma_{m_{DK}} = 27$ MeV/c^2.

$B^0 \to D^-\pi^+$

The nominal signal is described by a double-Gaussian model as the previous mode, and the combinatorial background by an exponential (4.29) model. A Gaussian model is used for the Cabibbo-suppressed $B^0 \to D^-K^+$ decay, identically to what was done for the previous decay mode.

The mass template $L_m^{B_s}(m)$ for the misreconstructed $B_s \to D_s^-\pi^+$ ($D_s^- \to K^+K^-\pi^-$) background is obtained from Monte Carlo simulation, and is given by two Gaussian functions centered at 5.315 GeV/c^2, of widths 45 MeV/c^2 and 22 MeV/c^2 with fractions 81% and 19%, respectively.

The misreconstructed $\Lambda_b^0 \to \Lambda_c^+\pi^-$ ($\Lambda_c^+ \to p^+K^-\pi^+$) background decay is parameterized, $L_m^{\Lambda_b^0}(m)$, as well from Monte Carlo simulation by an exponential of decay constant 76 MeV/c^2 convoluted with a Gaussian of width 36 MeV/c^2 and centered at 5.439 GeV/c^2.

$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$, $B^0 \to D^-\pi^+\pi^-\pi^+$

The nominal signal and combinatorial background are described by single Gaussian (4.27) and falling exponential (4.29) models, respectively. The Cabibbo-suppressed $B \to DK\pi\pi$ decays are described by a Gaussian template whose width, displacement and fraction relative to the nominal signal are all determined, and fixed, from Monte Carlo simulation.

$B^0 \to D^{*-}\pi^+$

The mass of the nominal signal in $D^*\pi$ is modeled with a single Gaussian (4.27) component, while the Cabibbo-suppressed contribution from D^*K is handled by another Gaussian defined by parameters which are derived from signal Monte Carlo. The sole source of background events is combinatorics, which are well described in mass by a single linear (4.28) component.

$B^0 \to D^{*-}\pi^+\pi^-\pi^+$

The mass model is yet further simplified for the two modes with the $D^{*-}(\pi\pi\pi)^+$ topology as there is no Cabibbo-suppressed contribution to the B^0 signal. The mass subspace is entirely defined by a single Gaussian (4.27) peak on a linear (4.28) combinatorial background.

$B_s \to D_s^-\pi^+$, $B_s \to D_s^-\pi^+\pi^-\pi^+$

The nominal signals are described by a Gaussian function (4.27). The combinatorial background is modeled via the combination of a flat (4.28) and an exponential (4.29) distribution.

The Cabibbo-suppressed $B_s \to D_s^- K^+$ and $B_s \to D_s^- K^+\pi^-\pi^+$ decays are treated in an identical fashion to the previous modes, via a Gaussian model fixed from Monte Carlo. The shapes of the fully reconstructed Λ_b^0 and B^0 mass peaks are similarly fixed to those found from Monte Carlo. The same applies to the mass distributions from the partially reconstructed $B \to D_s^-\pi X$ background decays which are modeled with a Gaussian (4.27) and a linear (4.28) function.

$B^{+,0} \to DlX$

To model the D mass distribution associated with the signals, and the *fakes* and *physics* background components we use a common double Gaussian distribution,

$$L_m^{Dl}(m|M,\sigma_{m,1},\sigma_{m,2}) = (1-f)\,G(m;M,\sigma_{m,1},M_{\min},M_{\max}) + f\,G(m;M,\sigma_{m,2},M_{\min},M_{\max})\,.$$

We assume a simple linear shape (4.28) for the mass distribution of the combinatorial background.

$B_s \to D_s^- l^+ X$

The description in mass space of all involved components – signal as well as combinatorial, fakes, and physics backgrounds – is identical to that presented for the previous Dl samples.

4.5 Proper decay time PDF

The description of the proper decay time subspace involves the physics parameters which we are primarily interested in extracting from the data. The B mesons lifetimes do appear in the formalism developed in Section 4.2 for the signal components. The extraction of such parameters requires additionally an appropriate description of the other contributing components in the samples.

A background source common to all samples is that of the combinatorial type. Its description in the proper decay time space is found from the mass-sideband regions. A common *empirical* model is employed, which is found to provide an effective description across the data samples. It is first presented in a more general form, while its specific realizations for the individual samples are addressed afterwards. It involves a linear combination of Gaussian and smeared exponential terms. Terms of the latter form have been evaluated in (4.15). These are more generally written as

$$G_\alpha(t) \equiv G(t;\sigma_\alpha,\Delta_\alpha) = \delta(t-\Delta_\alpha)\otimes G(t;\sigma_\alpha)\,, \tag{4.30}$$

$$H_\alpha(t) \equiv H(t;\lambda_\alpha,\sigma_\alpha,\Delta_\alpha) = E(t-\Delta_\alpha;\lambda_\alpha)\otimes G(t;\sigma_\alpha)\,. \tag{4.31}$$

The model in its general form involves: one Gaussian term, one negatively and one positively short-lived exponential component, and one positively long-lived exponential component; let us refer to these by the subscripts $'o'$, $'-'$, $'+'$, $'++'$, respectively ($0<\lambda_-,\lambda_+\ll\lambda_{++}$). The resulting PDF may then be expressed as

$$\begin{aligned} &L_t^{\rm comb}(t|\sigma_o,\Delta_o,f_o;\lambda_-,\sigma_-,\Delta_-,f_-;\lambda_+,\sigma_+,\Delta_+,f_+;\lambda_{++},\sigma_{++},\Delta_{++},f_{++}) \\ &= \quad f_oG_o(t)+f_-H_-(-t)+f_+H_+(t)+f_{++}H_{++}(t)\,, \end{aligned} \tag{4.32}$$

where $f_{\alpha\in\{o,-,+,++\}}$ denote the relative fractions, with $\sum_\alpha f_\alpha=1$. Next we specify the model realizations (and simplifications) which are applied to the individual samples, and simultaneously attempt to motivate the above parameterization itself.

The description of the combinatorial background components in samples characterized by an *unbiased* proper decay time distribution is of special importance. The $J/\psi K$ modes provide such class of unbiased samples, for which the proper time distribution is dominated by a peak centered at $t=0$. This is due to candidates constructed from tracks originating from the primary vertex, which for the current case mostly involve the combination of a prompt J/ψ with prompt track(s). As the candidates forming this prompt component are characterized by zero proper time, its description is provided mathematically by a delta function, $\delta(t)$, convoluted with the resolution function (4.14); *i.e.* it is of the form of (4.30). The remaining sources of combinatorial background include: candidates involving tracks with erroneous hits, or belonging in reality to different displaced vertices (short-lived); physics backgrounds not explicitly accounted for, fake J/ψ from sequential semileptonic $b\to c\mu\nu\to s\mu\mu\nu\nu$ decays, or true displaced J/ψ paired with a random track (long-lived). The description of these other sources is provided, in an *effective* fashion, by the short and long-lived components of the form of (4.31). The combinatorial proper time PDF for the unbiased samples may be summarized as

$$\begin{aligned} & L_t^{\text{comb}}(t|\sigma_t, S_t; \lambda_-, f_{++}; \lambda_+, f_+; \lambda_{++}, f_{++}) \\ = \quad & f_o G(t; S_t\sigma_t) + f_- H_-(-t) + f_+ H_+(t) + f_{++} H_{++}(t) \,. \end{aligned} \tag{4.33}$$

The displacement parameters are here taken to be zero, $\Delta_\alpha = 0$, and instead of the fitting parameters σ_α a common smearing resolution is used, which is determined for each event by the corresponding measured t-uncertainty, $\sigma_\alpha = S_t\sigma_t$. The scaling factor S_t is a fit parameter. As it was mentioned above, the prompt component formed of candidates with zero proper time should be described by the resolution function (4.14). In fact, deviations observed in the distribution of the data from this form must be due to an incorrect estimation of the t resolution, σ_t. The model of (4.33) thus allows to determine the necessary proper time uncertainty scale factor from the fit to the data. Such correcting factor is propagated to the full likelihood model, and in particular to the signal proper time PDFs derived in Section 4.2.

For the biased samples, provided by the hadronic and semileptonic modes, the model of (4.32) is suitable with the following modifications. The trigger and candidate selection criteria remove events with lower proper times. This bias therefore substantially removes the negatively lived, H_-, and prompt, G_o, components. In effect, for the $D_{(s)}\pi(\pi\pi)$ modes a model without both such components is used, as it is found to describe data well. In the combinatorial PDF used for the $D_{(s)}l$ modes a Gaussian component displaced from the origin ($\Delta_o > 0$) is found useful to describe the data, and the negatively lived term is again not employed ($f_- = 0$).

We re-emphasize that the prompt component is substantially removed in the case of biased samples. This implies that the method for calibrating the proper time uncertainties which was described for the $J/\psi K$ modes cannot be applied to these samples in the same fashion. For the lifetime fits of the biased samples the t-uncertainties, which are inputs to the fit, are scaled with the predetermined factors S_t. This issue is addressed further in Section 4.7, where we explain the details involved in determining and transferring these necessary correcting factors which are particularly relevant for the study of B_s oscillations performed in Chapter 7.

For the hadronic modes, the associated Cabibbo-suppressed channels share the same proper time description as the nominal signal components. Generally, associated exclusive backgrounds follow a model description provided by Monte Carlo simulation.

In the remainder of this section we describe the proper time PDFs of the additional background components which pertain to each individual sample.

$B^+ \to J/\psi K^+$

The proper decay time distribution for the nominal signal is that obtained in (4.15) for fully reconstructed, unbiased signal samples,

$$L(t|\sigma_t, S_t, \tau) \quad = \quad E(t;\tau) \otimes G(t; S_t\sigma_t)\,; \tag{4.34}$$

the scale factor S_t, as mentioned above, is introduced as a fit parameter. The Cabibbo-suppressed $B^+ \to J/\psi\pi^+$ component is described by a model identical to that of the signal, with a lifetime forced to $1.01 \times \tau$ as predicted from simulation. The PDF for the combinatorial background is given in (4.33).

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

The model description of the nominal signal and combinatorial background is identical to that of the above charged mode. The proper time distribution of the K^*-swapped candidates is negligibly different from that of correctly reconstructed candidates, and share thus in the fit the same model as the nominal signal.

$B^+ \to \bar{D}^0\pi^+$

The proper time PDF for the nominal signal has been obtained in (4.19) for the fully reconstructed, biased samples. The Cabibbo-suppressed modes share the identical description. That for the combinatorial background is realized by two positively lived components of the form of (4.31), with a common displacement Δ_{bg} and smearing resolution parameter σ_{bg}.

$B^0 \to D^-\pi^+$

In addition to that referred to the previous mode, the Λ_b^0 and B_s^0 backgrounds are modeled from and fixed to Monte Carlo simulation via terms of the form of (4.31).

$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$, $B^0 \to D^-\pi^+\pi^-\pi^+$

The description employed for the nominal signal, as well as for the Cabibbo-suppressed modes, is that given in (4.19) for the fully reconstructed, biased samples. The combinatorial background, in the case of the $B^0 \to D^-\pi^+\pi^-\pi^+$ sample, is described by two independent, smeared exponential terms – of the form of (4.31) –, while for the $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ sample one single such term is found adequate.

$B^0 \to D^{*-}\pi^+$

The nominal signal component in the t subspace is the standard expression derived for the hadronic modes (4.19), which is shared as well by the Cabibbo-suppressed D^*K component. The combinatorial background is fitted with two positively lived exponential tails. These tails share a common offset from zero, Δ_{bg}, and are smeared by a common Gaussian resolution whose width, σ_{bg}, is also a fit parameter.

$B^0 \to D^{*-}\pi^+\pi^-\pi^+$

The proper time distribution is modeled exactly as for $B^0 \to D^{*-}\pi^+$.

$B_s \rightarrow D_s^- \pi^+$, $B_s \rightarrow D_s^- \pi^+ \pi^- \pi^+$

The proper time PDF for the nominal signal is that obtained in (4.19) for the fully reconstructed, biased samples. Such PDF is shared by the Cabibbo-suppressed modes. The same is also true for the partially reconstructed B_s candidates entering in the fit region. A more detailed analysis induces a κ-factor correction, analogous to that implemented for semileptonic modes, which for events in the considered mass range is very close to unit, and is thus neglected.

The shape of the combinatorial background is described by a single positively lived exponential component of the form of (4.31).

$B^{+,0} \rightarrow DlX$

The reconstructed B signals in each of the three Dl samples involve both B^+ and B^0 meson decays. Accordingly, the proper decay time PDF must be formed of the combination

$$L_t^{Dl}(t|\sigma_t, \tau_u, \tau_d) \quad = \quad f_u \cdot L_t^u\,(t|\sigma_t, \tau_u) + f_d \cdot L_t^d\,(t|\sigma_t, \tau_d)\,. \tag{4.35}$$

Here τ_u and τ_d are the fit parameters, representing the B^+ and the B^0 meson lifetimes. The PDF components L^u and L^d are of the form of (4.22) as derived for the partially reconstructed, biased samples. The involved t-efficiency functions and κ-factor distributions are derived independently for the combination of B^+ and of B^0 channels. The relative fractions f_u and f_d ($f_u + f_d = 1$) for the two components are computed from an inclusive Monte Carlo sample, assuming the world average values [1] for the sample composition parameters, a summary of which is provided in Table 3.16. The relative signal fraction f_d of B^0 mesons in the $D^0 l$, $D^+ l$ and $D^* l$ samples is about 29%, 86%, and 90%. The difference between the electron and muon samples is below 0.5%.

The combinatorial background is described by the general model expressed by (4.32) with absent negatively lived component ($f_- = 0$). A similar model is used for the *prompt* component, which is composed of one Gaussian (4.30) and one smeared exponential (4.31) term.

The physics backgrounds are described in an inclusive way by a Gaussian smeared exponential term of the form of (4.31).

$B_s \rightarrow D_s^- l^+ X$

The likelihood model for describing the t distribution of the nominal B_s signal component has been derived for the partially reconstructed modes and is expressed by (4.22).

Each physics background component is described by an independent t-PDF term of the form of that used for the signal. The corresponding PDF (4.22) is formed using a κ-factor distribution and a t-efficiency function which are derived from Monte Carlo simulation of the associated B signal decay. The nominal values of the B^+ and B^0 lifetimes [1] replace the B_s lifetime fit parameter when appropriate. The (pseudo) proper time t itself along with its uncertainty are determined, according to (4.5) and (4.9), using the nominal mass of the involved B meson.

4.6 Fitting procedure and results

In this section we present in further detail the procedure for performing the simultaneous mass and proper decay time fits to the data samples based on the models described previously, together with the extracted results.

The likelihood factors modeling the mass L_m and proper decay time L_t for the individual samples components have been presented in the Sections 4.4 and 4.5 above. Additionally, the proper decay time uncertainty likelihood factor L_{σ_t} is constructed. As discussed in Section 4.1.3, two such distributions are obtained for each mode, from the mass-sideband and signal (mass-sideband subtracted) regions, which are respectively associated to the combinatorial and to the remaining sample components.

Before fitting the data samples, the necessary Monte Carlo quantities and templates are computed and implemented in the fitting framework for each individual mode. These may involve t-efficiency functions (Section 4.3.1), κ-factor distributions (Section 4.3.2), and certain sample composition parameters and templates (Section 3.3).

In general, the fit to the data proceeds in several steps. The results of the fitted parameters obtained at each step are used to set the initial values, constrain or fix the parameter values for the following step. Combinatorial background parameters are inferred first. This is usually accomplished by fitting candidates which belong to the mass-sideband regions, in which case solely the combinatorial background model is employed. For the $D\pi(\pi\pi)$ samples, due to the structures that populate the lower mass-sideband region to the left of the signal peak, a preliminary fit is performed in an extended mass region, as in Section 3.3.2. In the Dl samples, information about further background sources is extracted from additional data samples, formed of selected fake candidates, as described in Section 3.3.3. In the final stages of the procedure, the fit is performed in the full signal region. The description of backgrounds may be relaxed as appropriate. The main parameters of interest, *i.e.* the B mesons lifetimes, are then finally extracted.

The correction to the proper time uncertainty is obtained directly from the fit in the case of the $J/\psi K$ decay modes. It corresponds to the parameter S_t introduced in (4.33). As mentioned before, these samples have a prominent prompt background component whose width readily allows for the determination of this parameter. That is not the case for the remaining samples, where such prompt components are highly suppressed by trigger and reconstruction requirements. For the B^+ and B^0 modes reconstructed in all $D\pi(\pi\pi)$ and Dl samples, a common correcting factor (4.10) is used. A value of 1.4 is employed, which is based on the results obtained from the mentioned $J/\psi K$ fit and according to average results also obtained in Section 4.7. While for the case of the B^+ and B^0 mesons the usage of a single numerical value is sufficient, for the B_s case a more detailed scale factor evaluation is necessary. For the B_s samples a scale factor is inferred for each candidate, according to the procedure explained in Section 4.7. The proper time uncertainties are thus scaled before being provided as input to the fit.

4.6.1 Fully reconstructed decays

All parameters of the fit model for the $J/\psi K$ modes are determined directly and at once from the fit to the data except for the amount and the shape of the Cabibbo-suppressed and K^{*0} swapped contributions. The fit results for these samples are shown in Table 4.4, and the fit projections are found in Figures 4-7 and 4-8.

In the case of the $D\pi(\pi\pi)$ modes, combinatorial background mass parameters are pre-determined in a wide-range mass-only fit (Section 3.3.2). Among these only the overall combinatorial fraction is left floating in the nominal, narrow mass-range fit. The shape of the double-Gaussian model used in mass space to describe the signal peak in the $D\pi$ modes is also determined in such wide-range mass fit. No such preliminary fit is necessary in the case of the $D^*\pi(\pi\pi)$ modes for adjusting the simpler flat combinatorial mass model. The associated t-space parameters are then found from the fit to the candidates in the upper mass-sideband region. The relative fractions of the Cabibbo-suppressed and partially reconstructed decays are shown in Tables 3.10 and 3.9. These fractions along with the corresponding templates are determined from Monte Carlo simulation (Section 3.3). The fit projections are shown in Figures 4-9–4-21, and the numerical results are listed in Tables 4.5–4.8.

The B meson lifetimes as measured in the fully reconstructed modes are summarized in Table 4.2.

4.6.2 Partially reconstructed decays

For the semileptonic samples, the mass and the proper decay time models for the combinatorial background are adjusted by fitting candidates in the D mass-sideband regions. This is illustrated in Figure 4-5 for a selected sub-sample. The corresponding fraction is determined from the fit to the full signal mass region.

The relative fractions of the physics background are determined from Monte Carlo simulation and are summarized in Table 3.14 for all 24 semileptonic decay sub-samples. They are fixed parameters in the fit. Except in the case of the $D_s l$ samples, for which a signal-like treatment was adopted, a fixed, inclusive t template derived from Monte Carlo simulation is used to describe the proper decay time for the physics background component of each sub-sample. This is illustrated in Figure 4-5 for a selected sub-sample.

The description of the fakes background component is obtained from a preliminary fit to the fake lepton sample, as addressed in Section 3.3.3. This sample is sufficiently large to allow for the determination of the proper decay time distribution of the fakes background. Fit projections are shown in Figure 4-6. The relative fraction of this component is extracted from yet another preliminary fit, namely to the mass distribution of the Dl candidates, m_{Dl}. The procedure is thoroughly explained in Section 3.3.3. The numerical values of the fractions are summarized in Table 3.15. The characterization of the fakes background is fixed in the nominal lifetime fit. We

sample	$c\tau$ [μm]
$B^+ \to J/\psi K^+$	499.4±8.0
$B^+ \to \bar{D}^0\pi^+$	494.5±8.1
$B^+ \to \bar{D}^0\pi^\pm$	470.4±18.1
(combined B^+ fit	495.0±5.2)
$B^0 \to J/\psi K^{*0}$	458.4±11.0
$B^0 \to D^-\pi^+$	475.7±8.1
$B^0 \to D^-\pi^+\pi^-\pi^+$	458.4±11.5
$B^0 \to D^{*-}\pi^+$ ($\bar{D}^0 \to K^+\pi^-$)	495.3±22.4
$B^0 \to D^{*-}\pi^+$ ($\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$)	449.7±19.6
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-$)	449.3±21.0
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-\pi^+\pi^-$)	489.5±24.0
(combined B^0 fit	468.1±4.8)
$B_s \to D_s^-\pi^+$, $D_s \to \phi\pi$	464.76±27.81
$B_s \to D_s^-\pi^+$, $D_s \to K^*K$	464.91±41.41
$B_s \to D_s^-\pi^+$, $D_s \to 3\pi$	417.89±58.15
$B_s \to D_s^-\pi^+\pi^-\pi^+$, $D_s \to \phi\pi$	590.38±89.30
$B_s \to D_s^-\pi^+\pi^-\pi^+$, $D_s \to K^*K$	514.07±106.39
(combined B_s fit	472.73±21.08)

Table 4.2: Summary of measured B lifetimes in fully reconstructed modes.

note that the ability for background characterization carried by the Dl mass distributions, which was here used through a separate fit, could in principle be optimally exploited in a simultaneous fit. That is, by including m_{Dl} of the candidates as an additional input together with the corresponding likelihood factor.

For the D^0l, D^+l, D^*l final states, all 12 decay sub-samples are fitted simultaneously. A different set of parameters is allocated for each of the sub-samples being fit, with the B^+ and the B^0 lifetimes as the only common parameters. The fit likelihood projections in the mass and pseudo proper decay time spaces are shown in Figures 4-22–4-33. The 12 decay D_sl sub-samples are fitted individually. The fit projections are displayed in Figures 4-34–4-45.

The B meson lifetimes as measured in the partially reconstructed modes are summarized in Table 4.3.

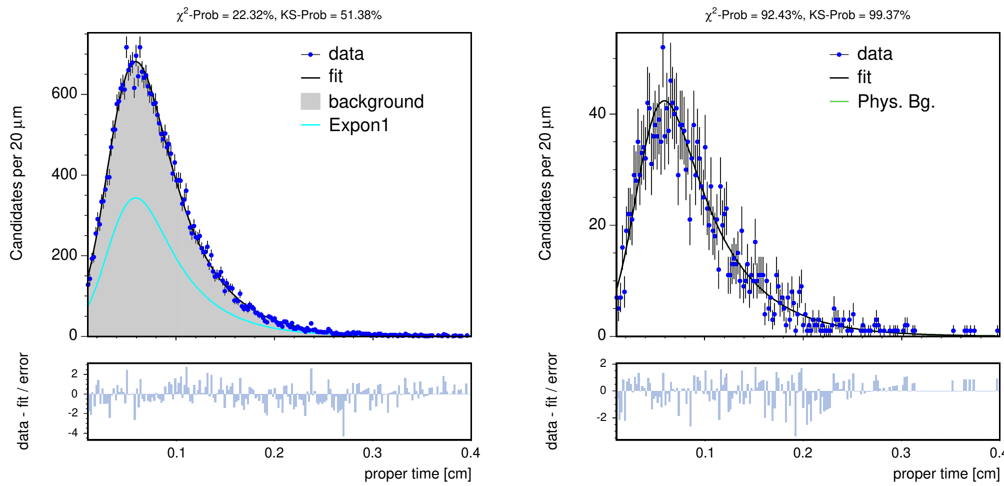


Figure 4-5: Proper decay time distribution of combinatorial (left) and physics (right) backgrounds, for the μD^+ B trigger sub-sample.

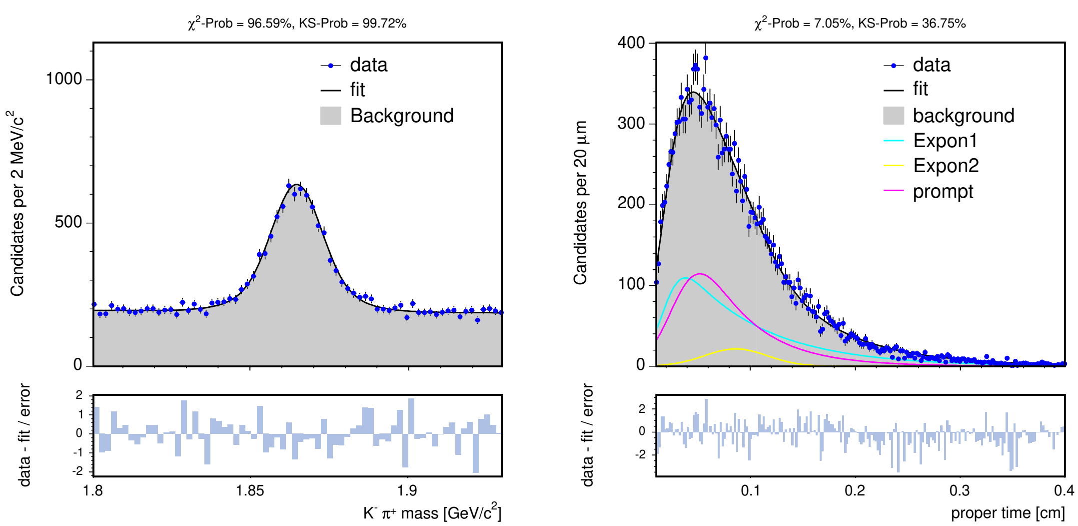


Figure 4-6: Fake lepton sample: mass and proper decay time distributions, for the μD^0 B trigger sub-sample.

Sub-sample	Trigger type	$c\tau_u$ [μm]	$c\tau_d$ [μm]	Sub-sample	Trigger type	$c\tau_s$ [μm]
μD^0	B	498.25±3.76	–	$\mu D_s(\phi\pi^-)$	B	440.8 ± 10.9
eD^0	B	489.55±4.66	–	$eD_s(\phi\pi^-)$	B	460.6 ± 14.7
μD^0	D	504.46±7.42	–	$\mu D_s(\phi\pi^-)$	D	471.2 ± 21.3
eD^0	D	502.63±9.62	–	$eD_s(\phi\pi^-)$	D	422.3 ± 25.8
μD^+	B	–	465.81± 4.36	$\mu D_s(K^{*0}K^-)$	B	437.9 ± 23.6
eD^+	B	–	462.82± 5.34	$eD_s(K^{*0}K^-)$	B	402.6 ± 27.9
μD^+	D	–	475.31± 7.38	$\mu D_s(K^{*0}K^-)$	D	445.6 ± 32.3
eD^+	D	–	460.31± 9.26	$eD_s(K^{*0}K^-)$	D	444.9 ± 45.7
μD^*	B	–	476.97± 7.25	$\mu D_s(\pi^+\pi^-\pi^-)$	B	392.2 ± 28.1
eD^*	B	–	468.35±10.41	$eD_s(\pi^+\pi^-\pi^-)$	B	409.4 ± 35.8
μD^*	D	–	459.31±12.62	$\mu D_s(\pi^+\pi^-\pi^-)$	D	480.0 ± 43.1
eD^*	D	–	448.24±16.41	$eD_s(\pi^+\pi^-\pi^-)$	D	420.6 ± 45.4
average		496.60±2.62	466.68± 2.53	average		441.14 ± 6.41

Table 4.3: Summary of measured B lifetimes in partially reconstructed modes; in the fits to the individual sub-samples the lifetime of a single B species is fitted, while those of the other present B mesons are fixed to their world average values [1].

parameter	$B^+ \to J/\psi K^+$	$B^0 \to J/\psi K^{*0}$
M [MeV/c^2]	5279.0 ± 0.2	5279.4 ± 0.3
σ [MeV/c^2]	12.8 ± 0.2	10.7 ± 0.3
f_{bg}	0.894 ± 0.002	0.863 ± 0.003
a[MeV/c$^{2^{-1}}$]	-1.09 ± 0.35	-1.78 ± 0.61
$c\tau$ [μm]	499.4 ± 8.0	458.4 ± 11.1
S_t	1.334 ± 0.007	1.376 ± 0.013
f_+	0.061 ± 0.004	0.069 ± 0.007
f_{++}	0.021 ± 0.002	0.036 ± 0.006
f_-	0.012 ± 0.001	0.013 ± 0.002
$c\tau_+$ [μm]	79.0 ± 7.5	80.0 ± 12.6
$c\tau_{++}$ [μm]	448.5 ± 36.6	415.2 ± 40.7
$c\tau_-$ [μm]	138.4 ± 10.3	130.0 ± 16.0

Table 4.4: Mass and proper decay time fit results for $J/\psi K$ modes.

parameter	$B^+ \to \bar{D}^0\pi^+$	$B^0 \to D^-\pi^+$
M[MeV/c^2]	5278.3 ± 0.2	5278.6 ± 0.3
σ [MeV/c^2]	16.5 ± 0.2	16.2 ± 0.3
f_{bg}	0.390 ± 0.005	0.324 ± 0.007
$c\tau$ [μm]	494.5 ± 8.1	475.7 ± 8.1
f_{++}	0.800 ± 0.083	0.226 ± 0.613
$c\tau_+$ [μm]	477.2 ± 63.3	227.8 ± 39.8
$c\tau_{++}$ [μm]	208.5 ± 15.7	384.1 ± 149.6
Δ_{bg} [μm]	178.2 ± 3.0	205.9 ± 4.1
σ_{bg} [μm]	50.4 ± 2.1	55.5 ± 3.0

Table 4.5: Mass and proper decay time fit results for $D\pi$ modes.

parameter	$B^0 \to D^-\pi^+\pi^-\pi^+$	$B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$
M [MeV/c^2]	5278.5 ± 0.4	5277.9 ± 0.5
σ [MeV/c^2]	14.2 ± 0.4	13.9 ± 0.5
f_{bg}	0.861 ± 0.004	0.527 ± 0.014
b^{-1}[GeV/c^2]	1.330 ± 0.061	2.856 ± 0.261
$c\tau$ [μm]	458.4 ± 11.5	470.4 ± 18.1
$c\tau_+$ [μm]	436.2 ± 21.4	505.3 ± 15.2
Δ_+ [μm]	203.7 ± 7.8	268.5 ± 7.9
σ_+ [μm]	37.7 ± 4.3	75.2 ± 6.6
f_{++}	0.622 ± 0.089	—
$c\tau_{++}$ [μm]	281.8 ± 11.8	—
Δ_{++} [μm]	295.3 ± 9.3	—
σ_{++} [μm]	68.0 ± 3.1	—

Table 4.6: Mass and proper decay time fit results for $D\pi\pi\pi$ modes.

parameter	$D^0 \to K\pi$	$D^0 \to K\pi\pi\pi$
M [MeV/c^2]	5278.1 ± 0.6	5279.3 ± 0.7
σ [MeV/c^2]	18.2 ± 0.6	17.8 ± 0.6
f_{bg}	0.387 ± 0.016	0.299 ± 0.018
a[GeV/c$^{2^{-1}}$]	-2.85 ± 0.96	-4.51 ± 1.33
$c\tau$ [μm]	495.3 ± 22.4	449.7 ± 19.6
f_{++}	0.047 ± 0.040	0.052 ± 0.051
$c\tau_+$ [μm]	221.9 ± 17.5	241.2 ± 25.7
$c\tau_{++}$ [μm]	970.6 ± 386.0	1104.1 ± 509.9
Δ_{bg} [μm]	100.3 ± 4.3	81.1 ± 5.3
σ_{bg} [μm]	24.7 ± 3.5	22.2 ± 4.3

Table 4.7: Mass and proper decay time fit results for $B^0 \to D^{*-}\pi^+$ modes, fitted separately by D^0 decay channel.

parameter	$D^0 \to K\pi$	$D^0 \to K\pi\pi\pi$
M [MeV/c^2]	5277.9 ± 0.7	5277.4 ± 0.7
σ [MeV/c^2]	15.5 ± 0.7	14.6 ± 0.7
f_{bg}	0.726 ± 0.011	0.676 ± 0.014
a[GeV/c$^{2^{-1}}$]	-3.29 ± 0.44	-3.36 ± 0.56
$c\tau$ [μm]	449.3 ± 21.0	489.5 ± 24.0
f_{++}	0.088 ± 0.073	0.468 ± 0.274
$c\tau_+$ [μm]	206.3 ± 13.5	181.5 ± 46.5
$c\tau_{++}$ [μm]	496.3 ± 134.8	359.1 ± 63.8
Δ_{bg} [μm]	152.0 ± 2.9	141.3 ± 3.7
σ_{bg} [μm]	37.3 ± 2.3	29.9 ± 2.8

Table 4.8: Mass and proper decay time fit results for $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ modes, fitted separately by D^0 decay channel.

parameter	$B_s \to D_s^-\pi^+$ channels		
	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^-\pi^+\pi^-$
M [GeV/c^2]	5.365272 ± 0.001206	5.369043 ± 0.001358	5.367352 ± 0.000037
σ [GeV/c^2]	0.021013 ± 0.001195	0.016042 ± 0.001234	0.021484 ± 0.002567
f_{bg}	0.436024 ± 0.019620	0.517976 ± 0.029545	0.728162 ± 0.029775
b^{-1}[GeV/c^2]	0.436343 ± 0.118268	2.439732 ± 0.183415	1.468729 ± 0.145326
$c\tau$ [cm]	0.046476 ± 0.002781	0.046491 ± 0.004141	0.041789 ± 0.005815
$c\tau_+$ [cm]	0.019139 ± 0.001303	0.022452 ± 0.001757	0.019421 ± 0.001425
Δ_+ [cm]	0.012299 ± 0.000738	0.014345 ± 0.000837	0.019063 ± 0.000819
σ_+ [cm]	0.003431 ± 0.000644	0.003639 ± 0.000706	0.004524 ± 0.000664

Table 4.9: Mass and proper decay time fit results for $B_s \to D_s^-\pi^+$ modes.

parameter	$B_s \to D_s^-\pi^+\pi^-\pi^+$ channels	
	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$
M [GeV/c^2]	5.368652 ± 0.000002	5.368372 ± 0.001627
σ [GeV/c^2]	0.013856 ± 0.001497	—
f_{bg}	0.844878 ± 0.015780	0.857557 ± 0.026283
b^{-1}[GeV/c^2]	1.138907 ± 0.085058	1.359105 ± 0.124292
$c\tau$ [cm]	0.059038 ± 0.008930	0.049950 ± 0.002930
$c\tau_+$ [cm]	0.046073 ± 0.001930	0.051407 ± 0.010639
Δ_+ [cm]	0.026917 ± 0.001065	0.033521 ± 0.001842
σ_+ [cm]	0.007232 ± 0.000944	0.010109 ± 0.001550

Table 4.10: Mass and proper decay time fit results for $B_s \to D_s^-\pi^+\pi^-\pi^+$ modes.

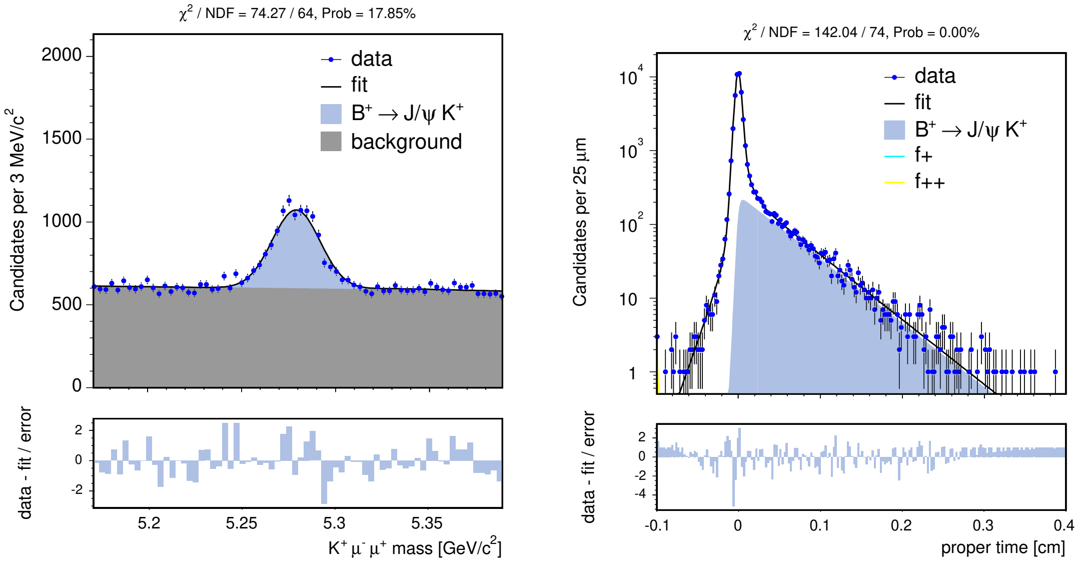


Figure 4-7: Mass and proper decay time fit projections for $B^+ \to J/\psi K^+$ decay.

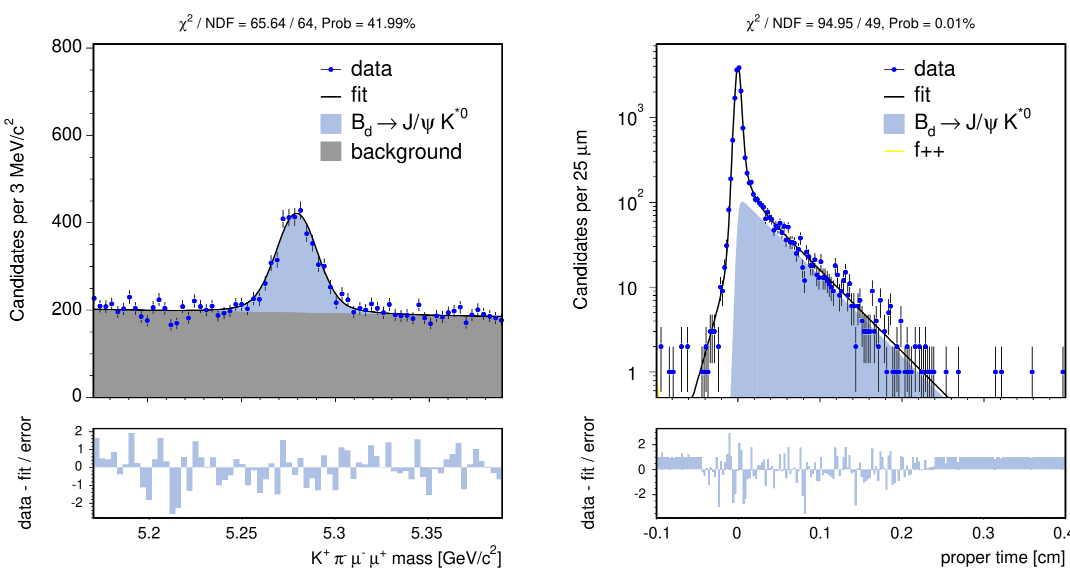


Figure 4-8: Mass and proper decay time fit projections for $B^0 \to J/\psi K^{*0}$ decay.

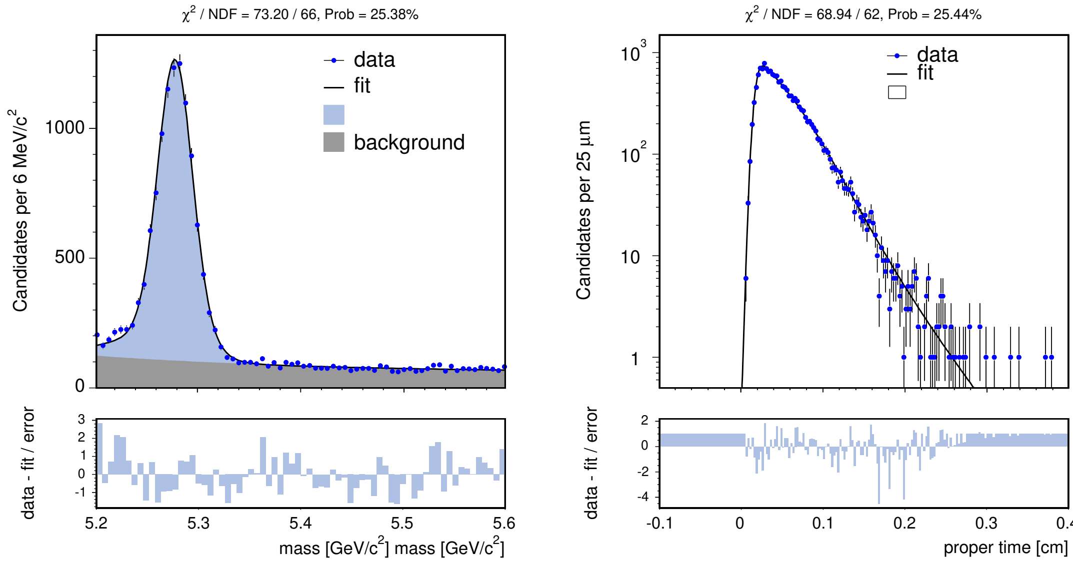


Figure 4-9: Mass and proper decay time fit projections for $B^+ \to \bar{D}^0\pi^+$ decay.

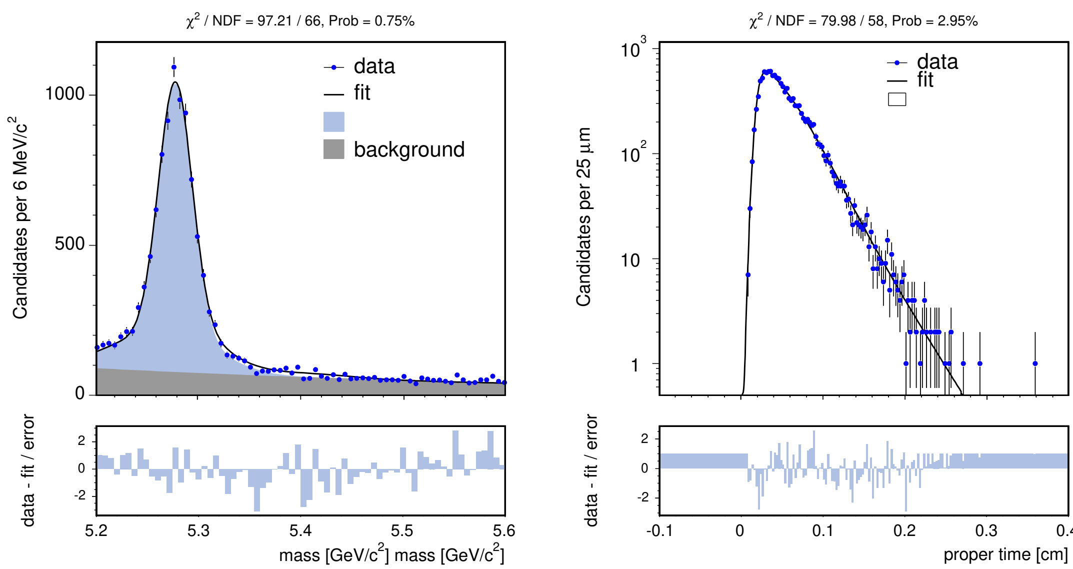


Figure 4-10: Mass and proper decay time fit projections for $B^0 \to D^-\pi^+$ decay.

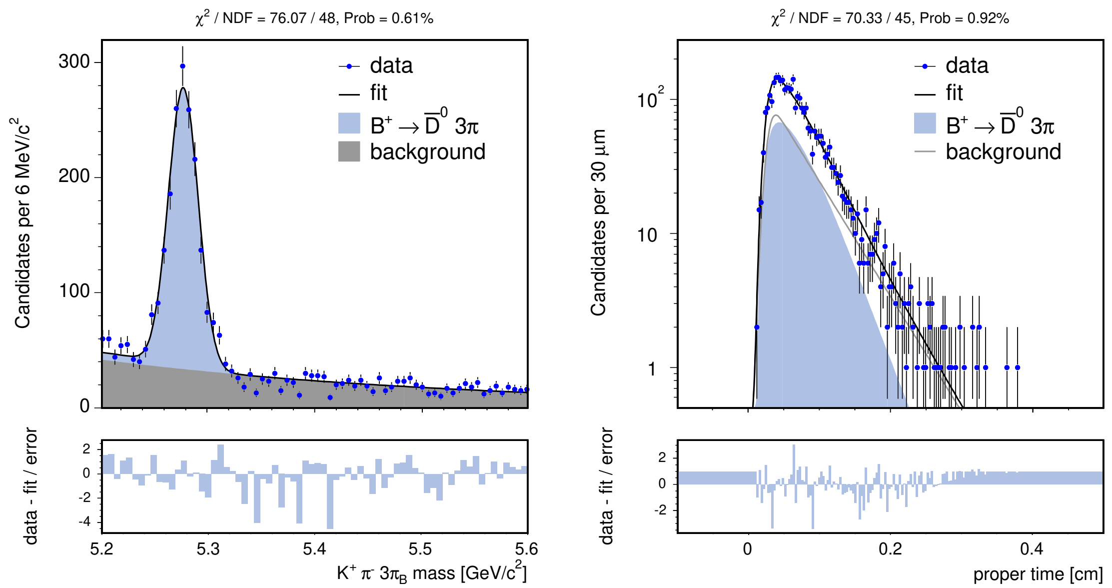


Figure 4-11: Mass and proper decay time fit projections for $B^+ \to \bar{D}^0\pi^+\pi^-\pi^+$ decay.

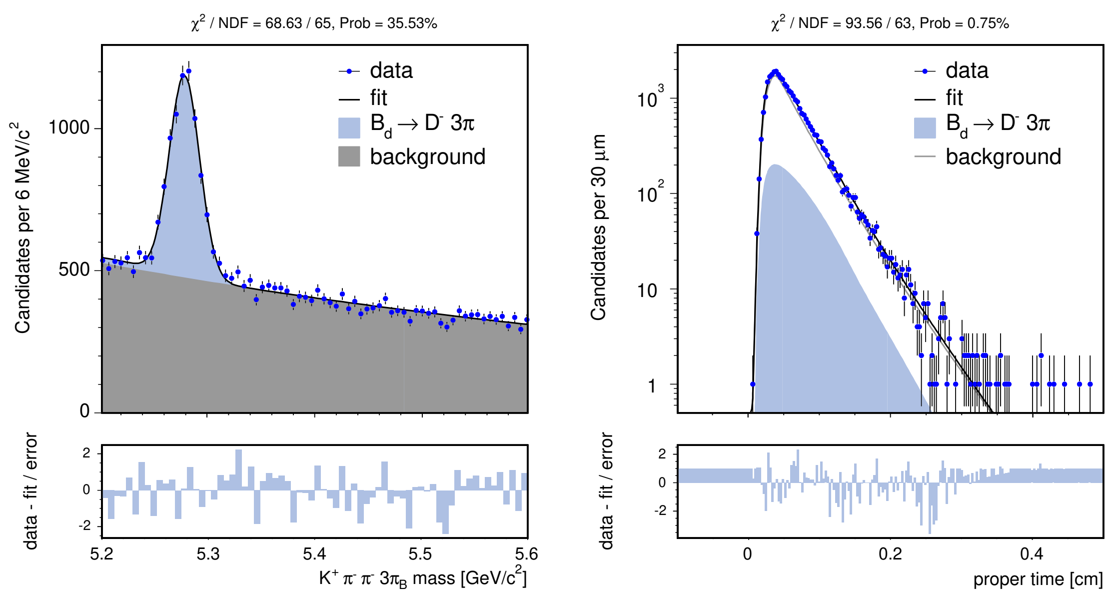


Figure 4-12: Mass and proper decay time fit projections for $B^0 \to D^-\pi^+\pi^-\pi^+$ decay.

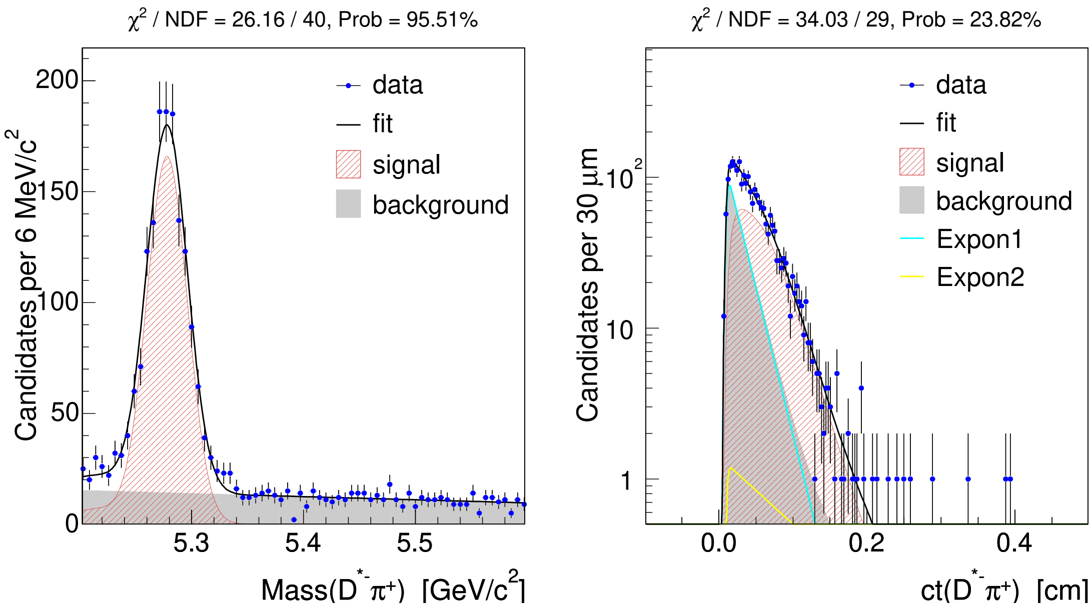


Figure 4-13: Mass and proper decay time fit projections for $B^0 \to D^{*-}\pi^+$, with $D^0 \to K\pi$.

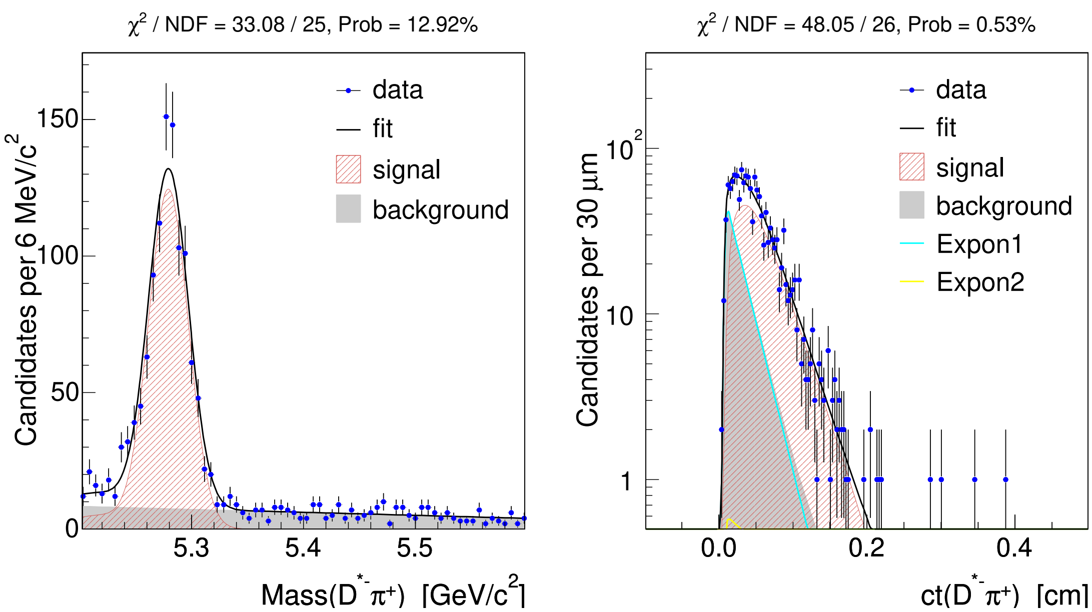


Figure 4-14: Mass and proper decay time fit projections for $B^0 \to D^{*-}\pi^+$, with $D^0 \to K\pi\pi\pi$.

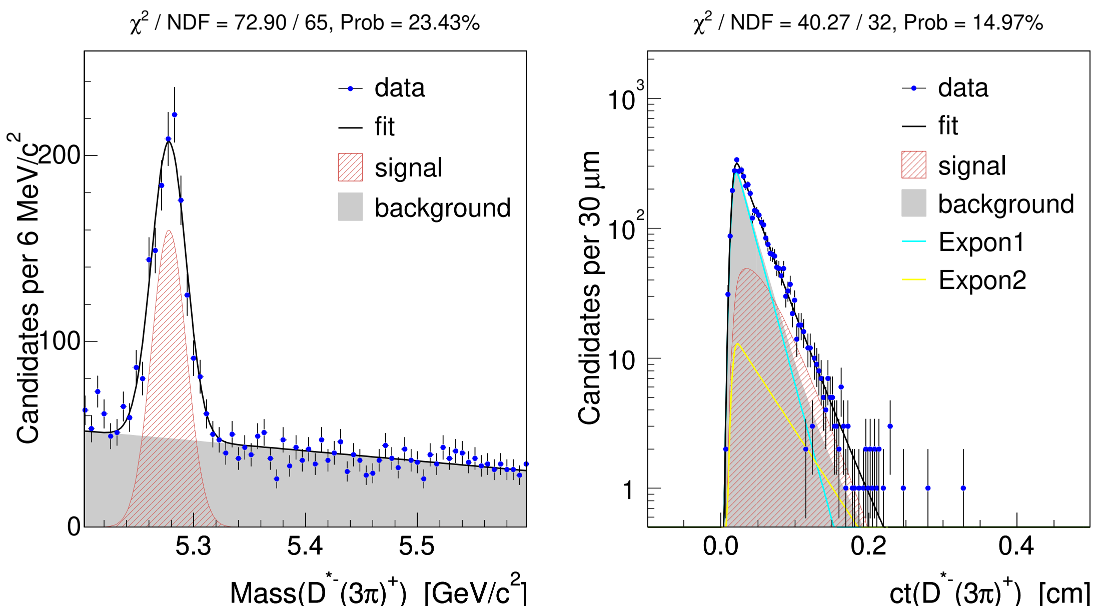


Figure 4-15: Mass and proper decay time fit projections for $B^0 \to D^{*-}(\pi\pi\pi)^+$, with $D^0 \to K\pi$.

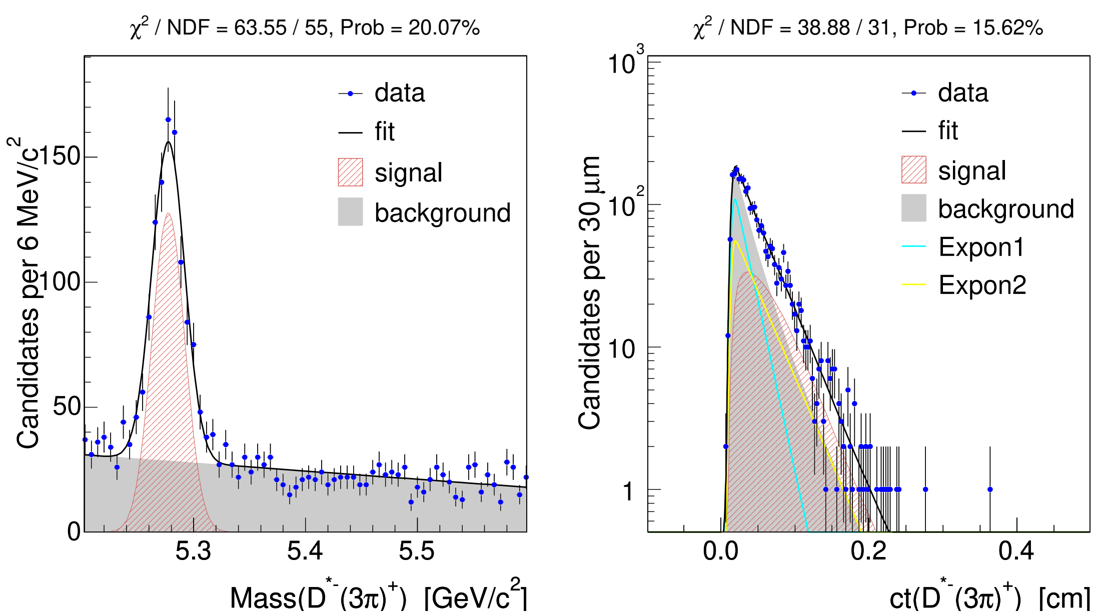


Figure 4-16: Mass and proper decay time fit projections for $B^0 \to D^{*-}(\pi\pi\pi)^+$, with $D^0 \to K\pi\pi\pi$.

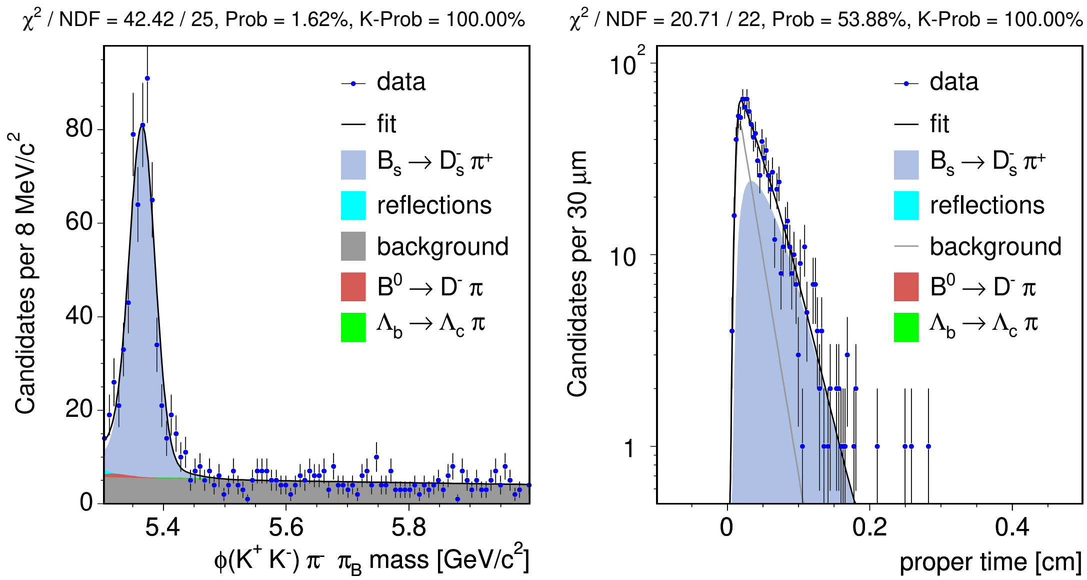


Figure 4-17: Mass and proper decay time fit projections for $B_s \to D_s^- \pi^+$ with $D_s^- \to \phi\pi^-$.

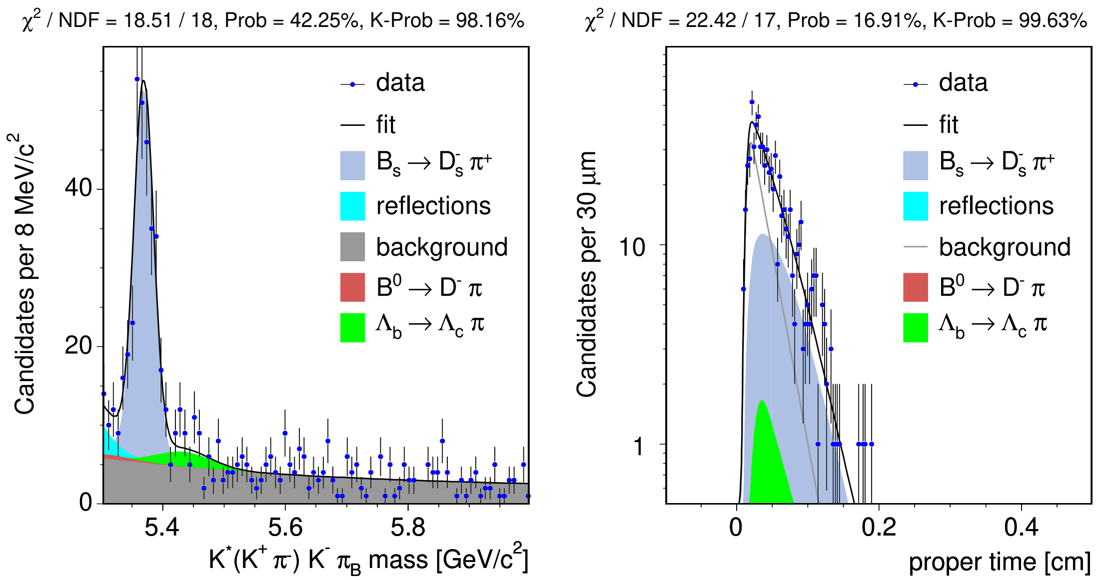


Figure 4-18: Mass and proper decay time fit projections for $B_s \to D_s^- \pi^+$ with $D_s^- \to K^{*0} K^-$.

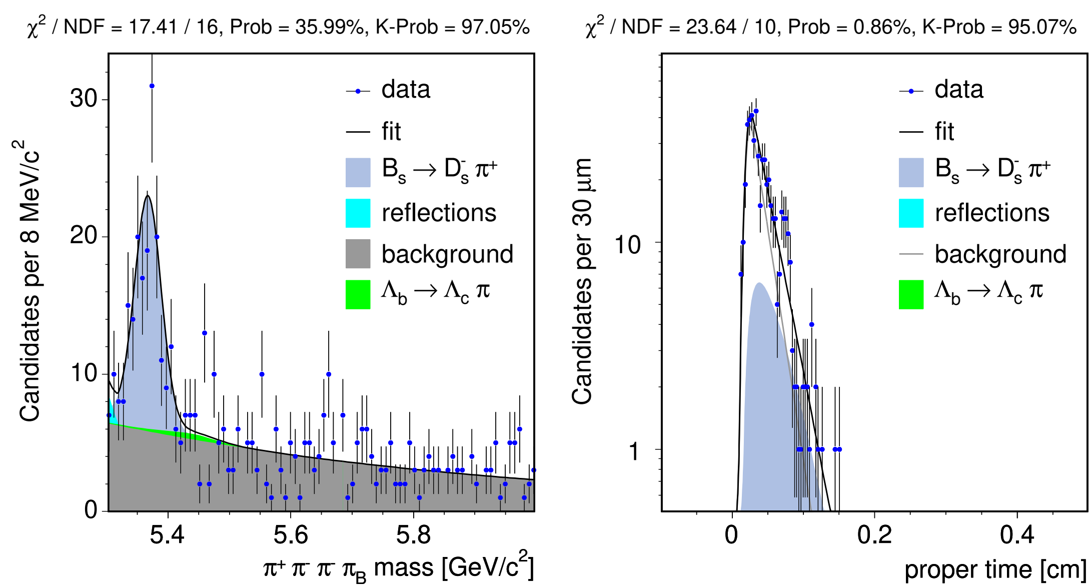


Figure 4-19: Mass and proper decay time fit projections for $B_s \to D_s^- \pi^+$ with $D_s^- \to \pi^- \pi^+ \pi^-$.

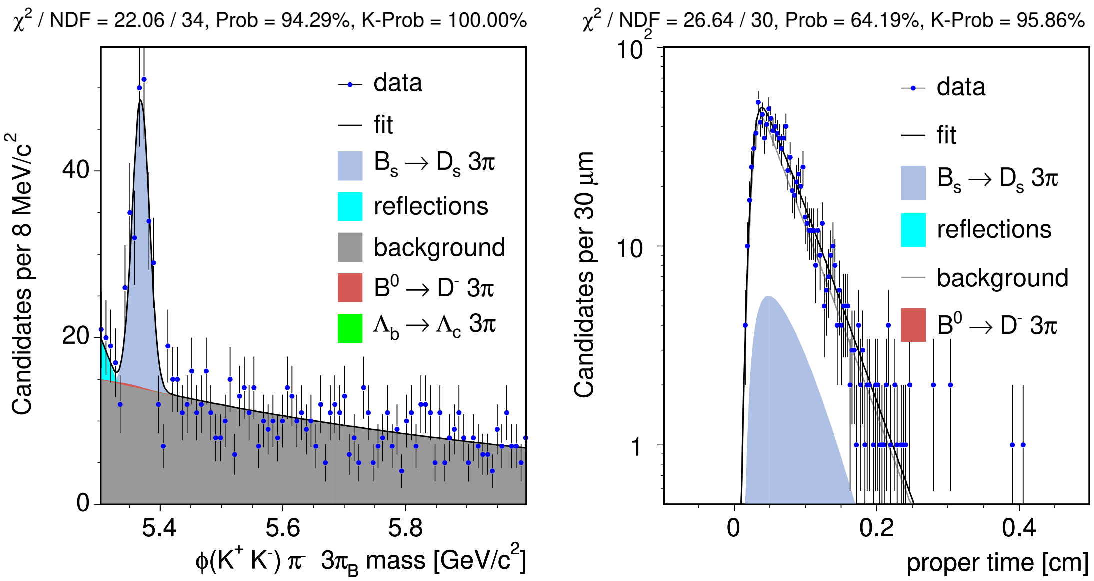


Figure 4-20: Mass and proper decay time fit projections for $B_s \to D_s^- \pi^+ \pi^- \pi^+$ with $D_s^- \to \phi \pi^-$.

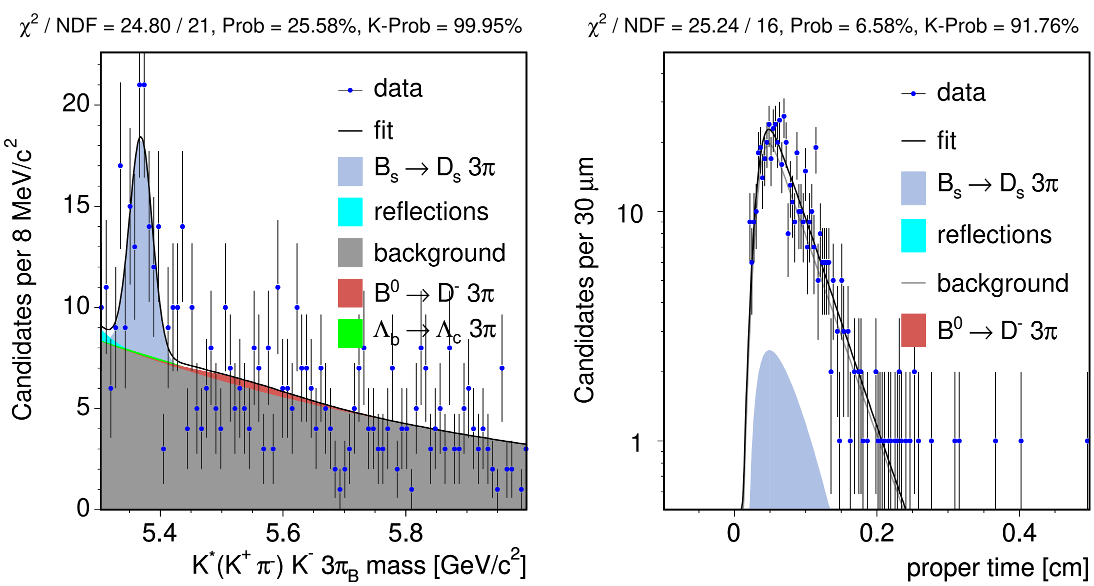


Figure 4-21: Mass and proper decay time fit projections for $B_s \to D_s^- \pi^+ \pi^- \pi^+$ with $D_s^- \to K^{*0} K^-$.

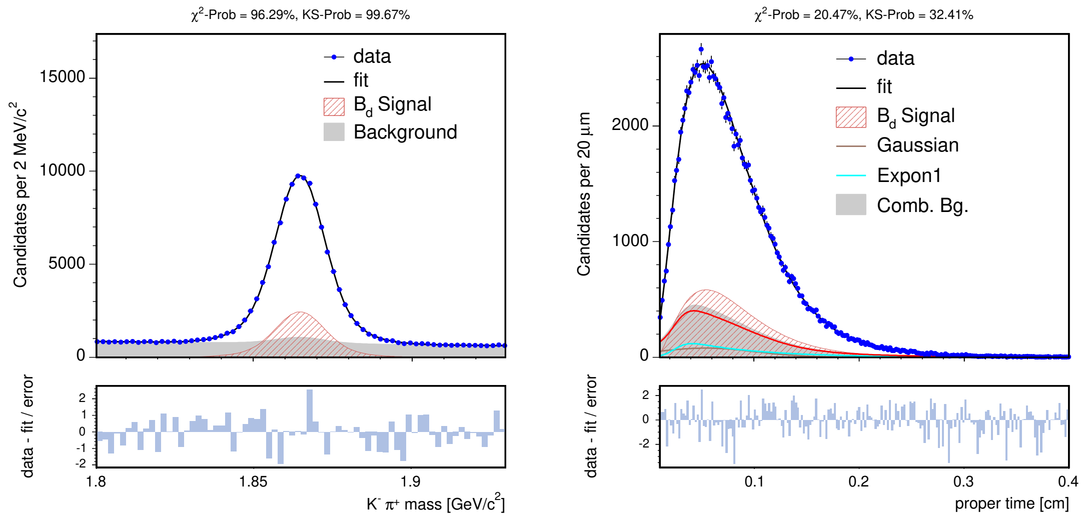


Figure 4-22: Mass and proper decay time fit projections for μD^0 B trigger.

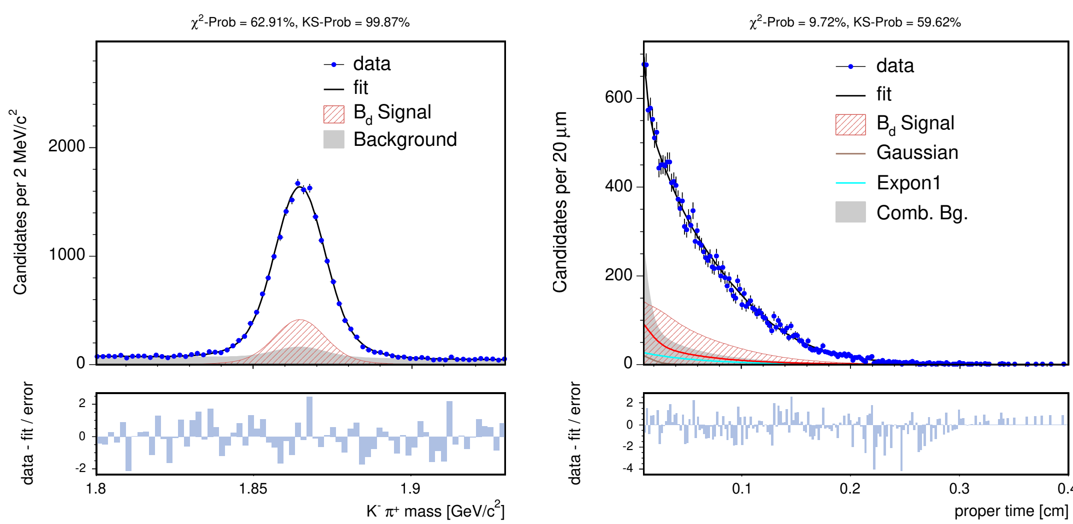


Figure 4-23: Mass and proper decay time fit projections for μD^0 D trigger.

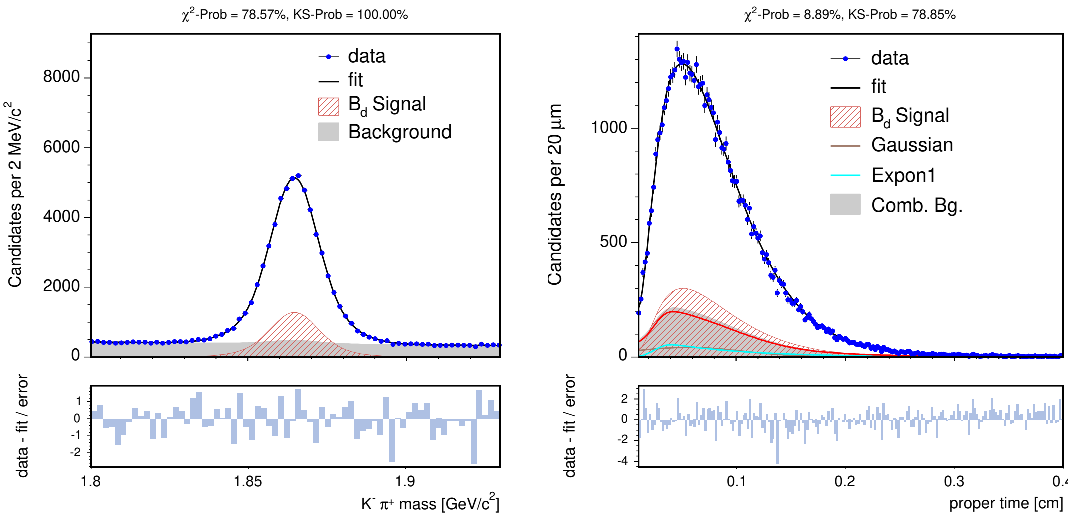


Figure 4-24: Mass and proper decay time fit projections for eD^0 B trigger.

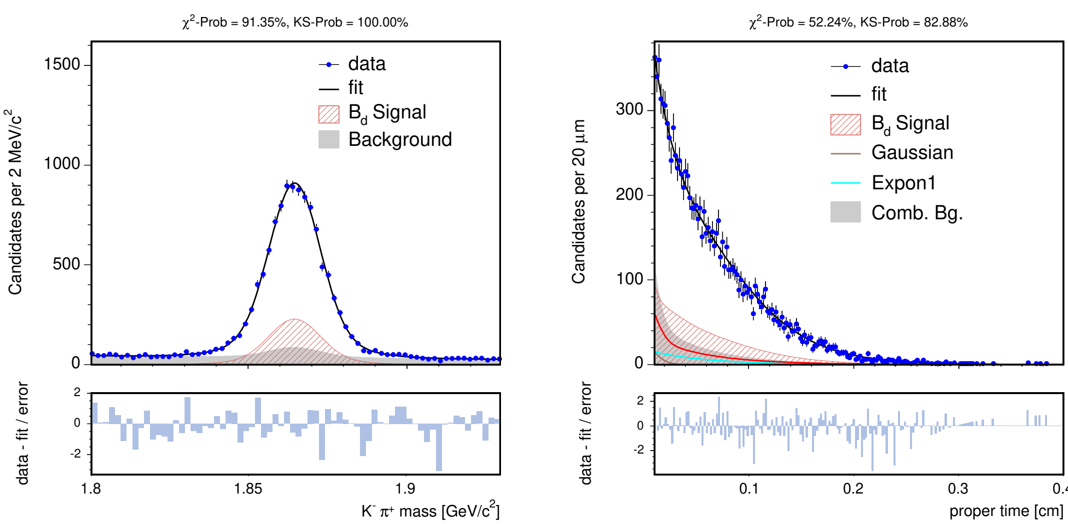


Figure 4-25: Mass and proper decay time fit projections for eD^0 D trigger.

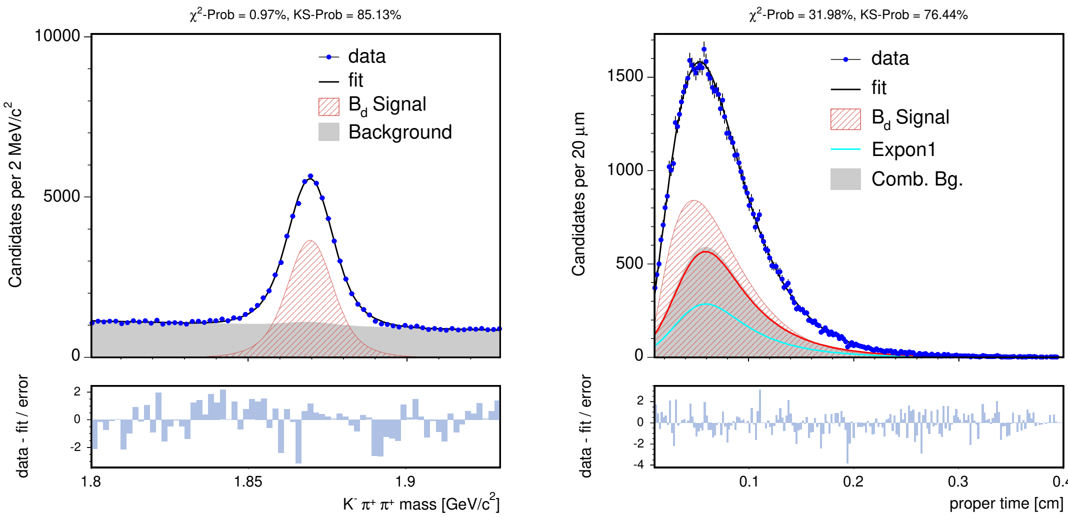


Figure 4-26: Mass and proper decay time fit projections for μD^+ B trigger.

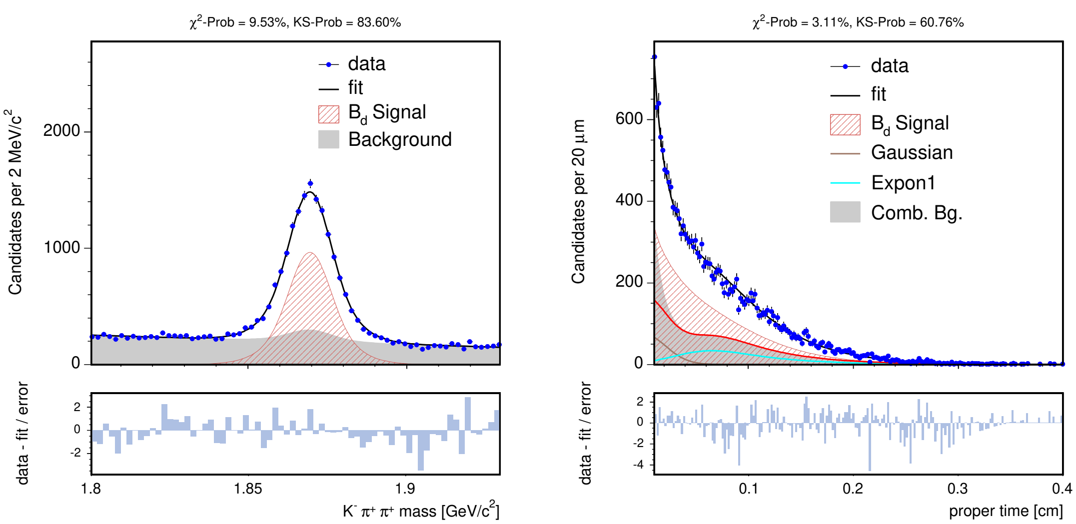


Figure 4-27: Mass and proper decay time fit projections for μD^+ D trigger.

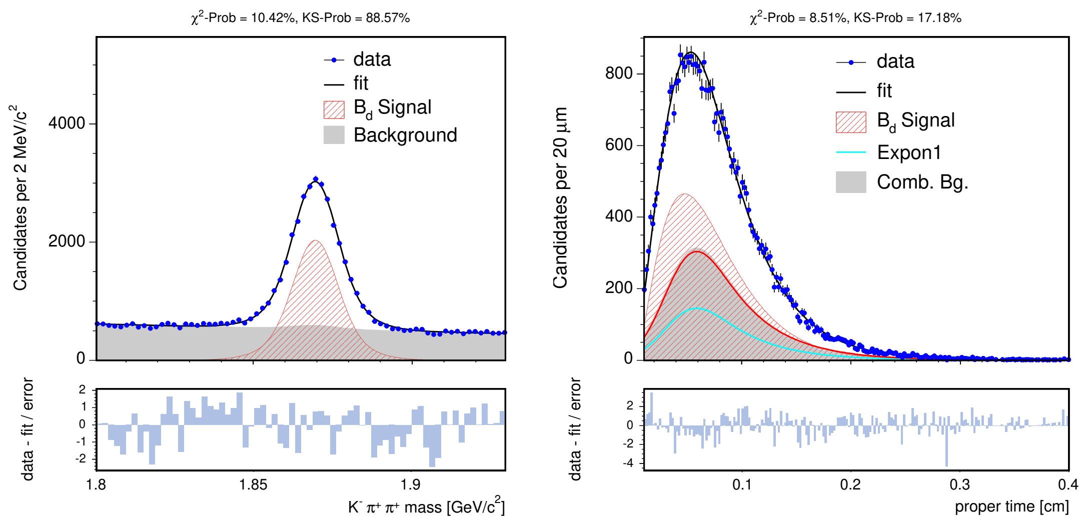


Figure 4-28: Mass and proper decay time fit projections for eD^+ B trigger.

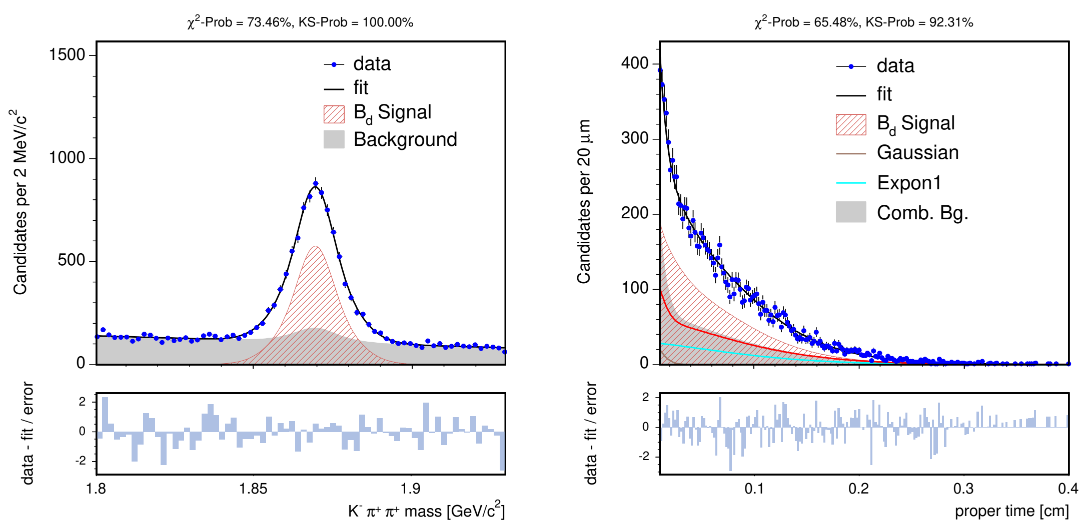


Figure 4-29: Mass and proper decay time fit projections for eD^+ D trigger.

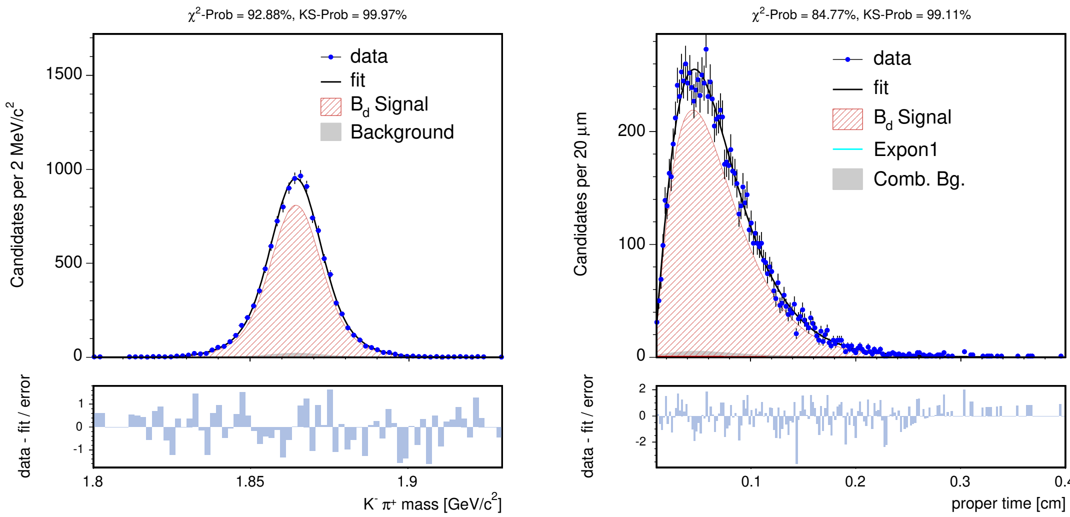


Figure 4-30: Mass and proper decay time fit projections for μD^{*} B trigger.

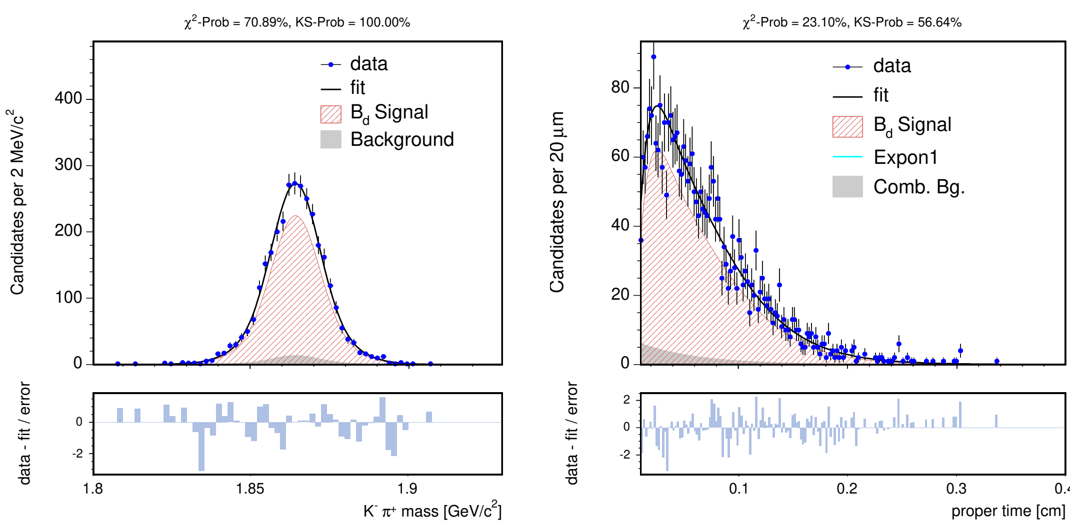


Figure 4-31: Mass and proper decay time fit projections for μD^{*} D trigger.

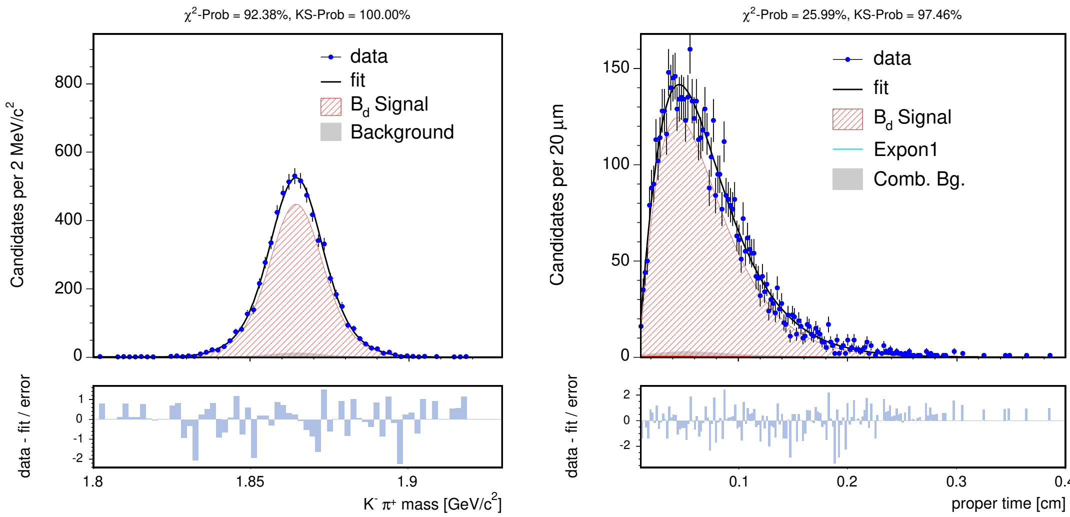


Figure 4-32: Mass and proper decay time fit projections for eD^* B trigger.

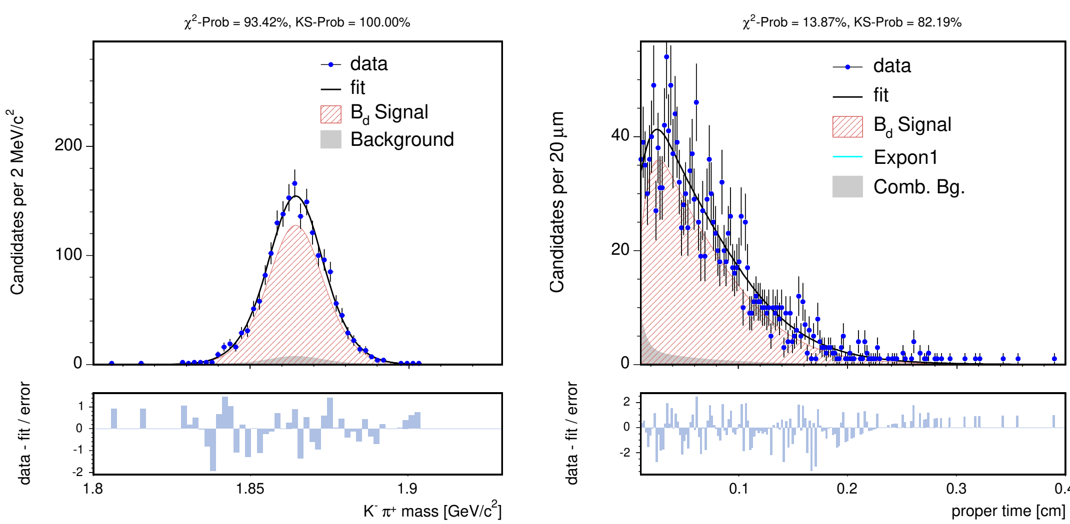


Figure 4-33: Mass and proper decay time fit projections for eD^* D trigger.

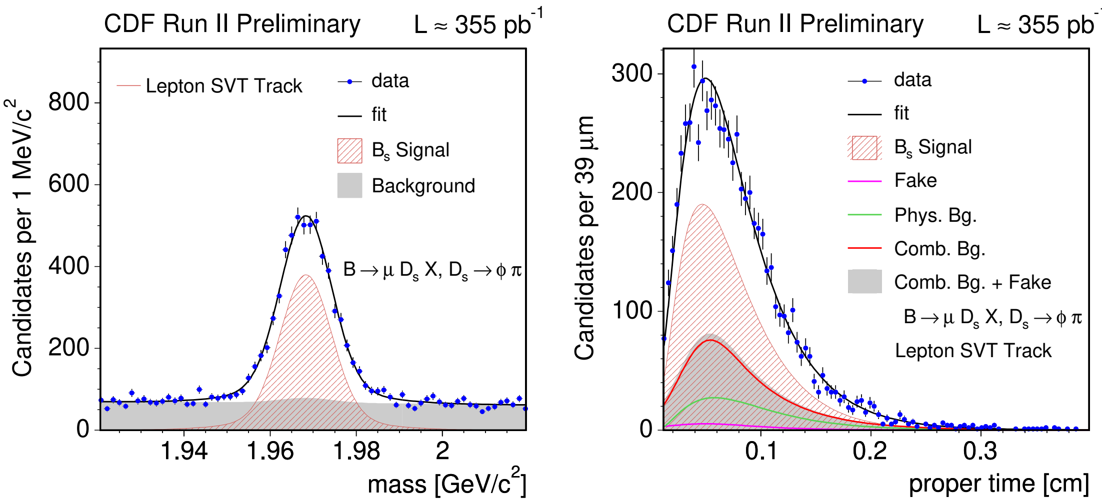


Figure 4-34: Mass and proper decay time fit projections for $\mu^+ D_s^-$ B trigger with $D_s^- \to \phi\pi^-$.

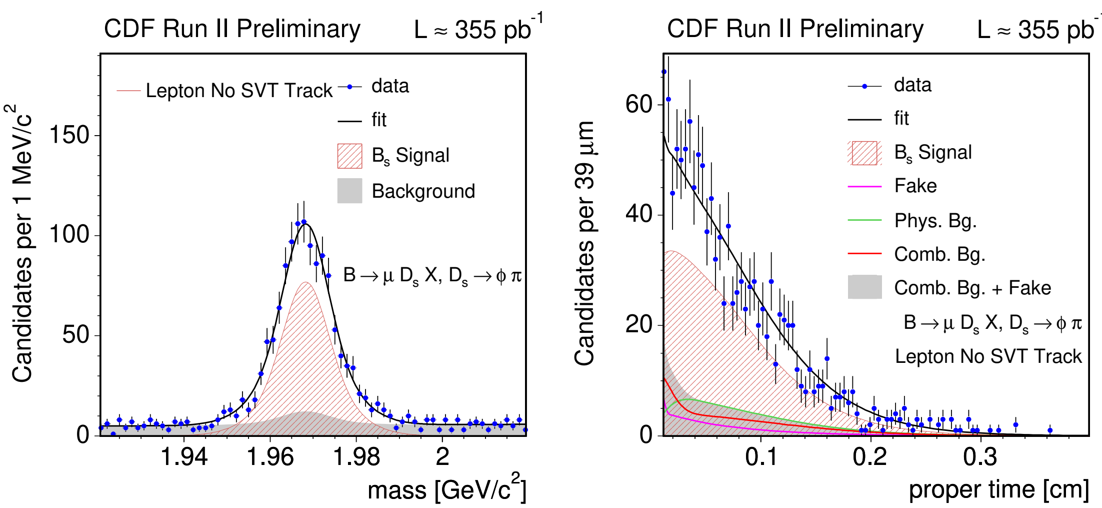


Figure 4-35: Mass and proper decay time fit projections for $\mu^+ D_s^-$ D trigger with $D_s^- \to \phi\pi^-$.

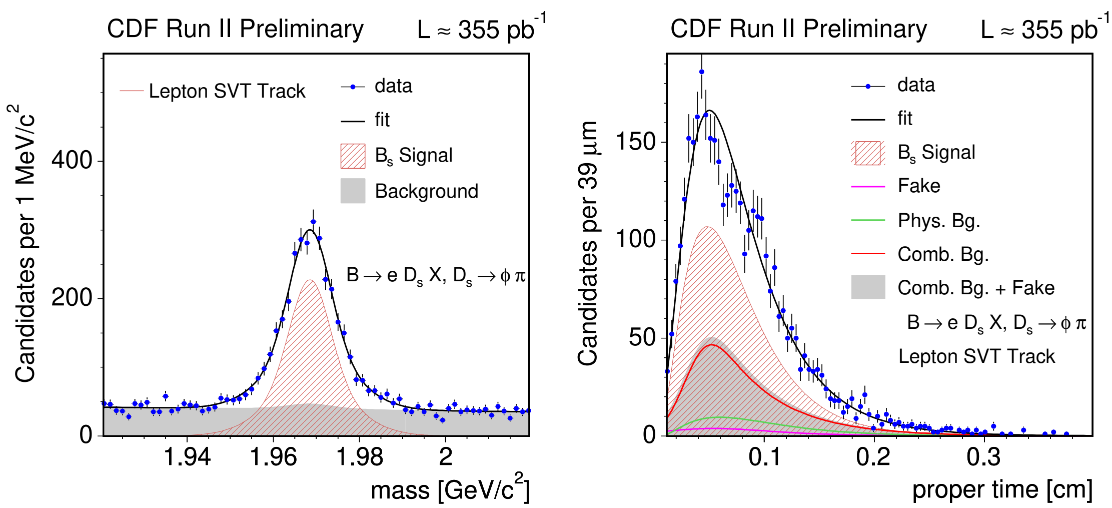


Figure 4-36: Mass and proper decay time fit projections for $e^+D_s^-$ B trigger with $D_s^- \to \phi\pi^-$.

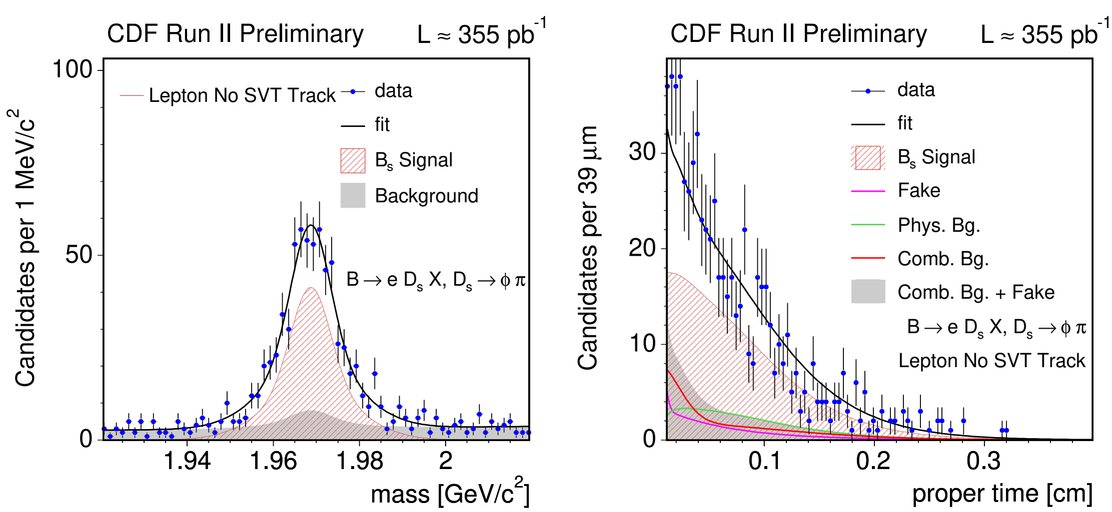


Figure 4-37: Mass and proper decay time fit projections for $e^+D_s^-$ D trigger with $D_s^- \to \phi\pi^-$.

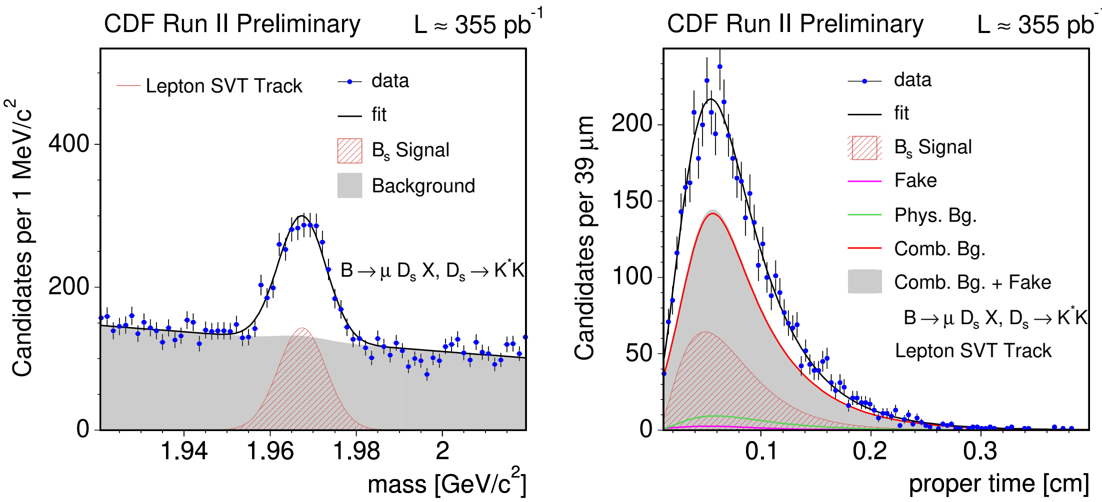


Figure 4-38: Mass and proper decay time fit projections for $\mu^+ D_s^-$ B trigger with $D_s^- \to K^{*0} K^-$.

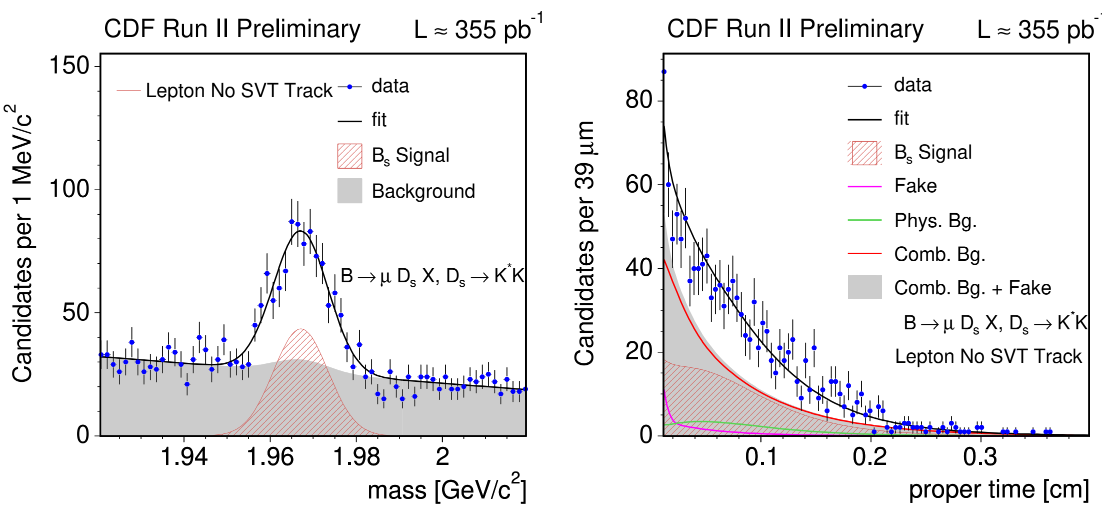


Figure 4-39: Mass and proper decay time fit projections for $\mu^+ D_s^-$ D trigger with $D_s^- \to K^{*0} K^-$.

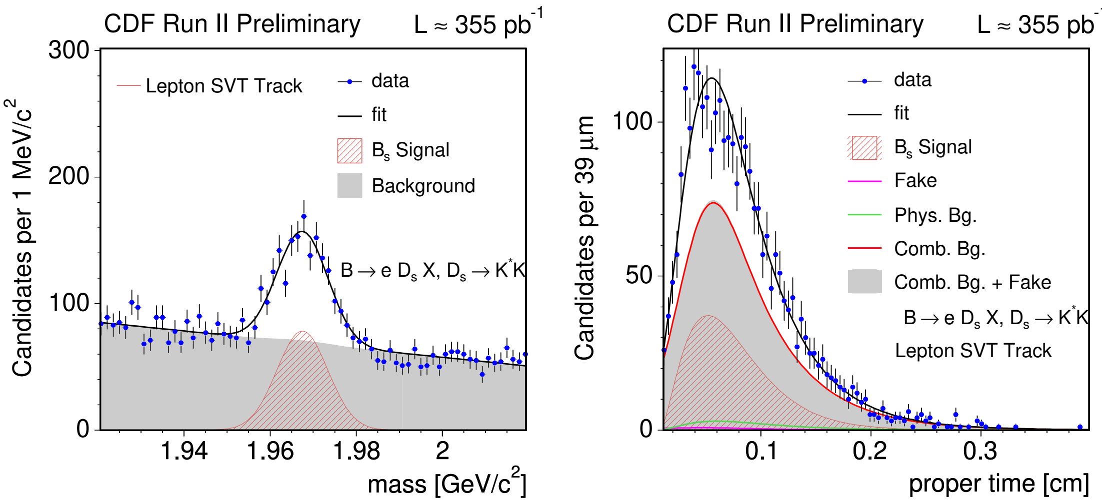


Figure 4-40: Mass and proper decay time fit projections for $e^+D_s^-$ B trigger with $D_s^- \to K^{*0}K^-$.

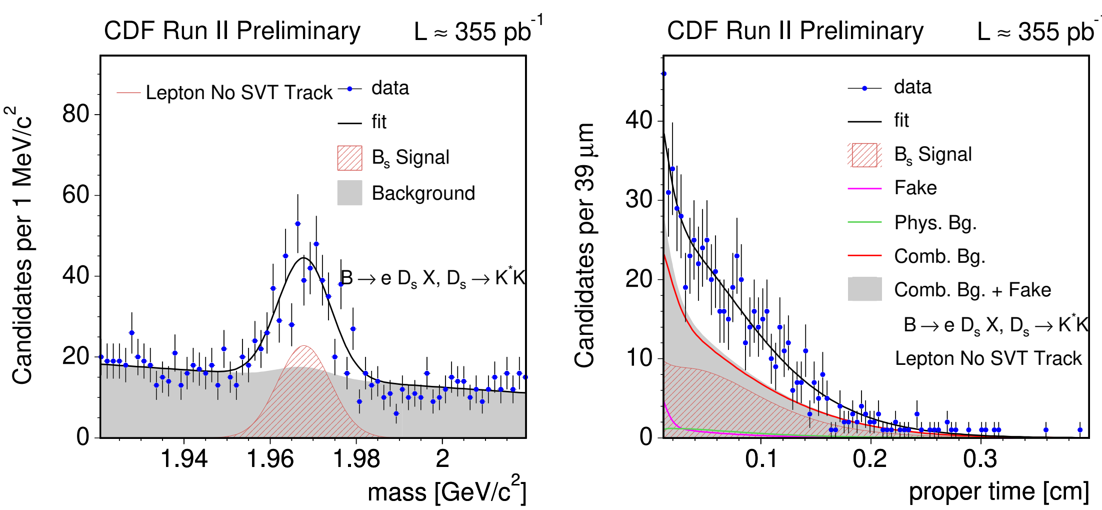


Figure 4-41: Mass and proper decay time fit projections for $e^+D_s^-$ D trigger with $D_s^- \to K^{*0}K^-$.

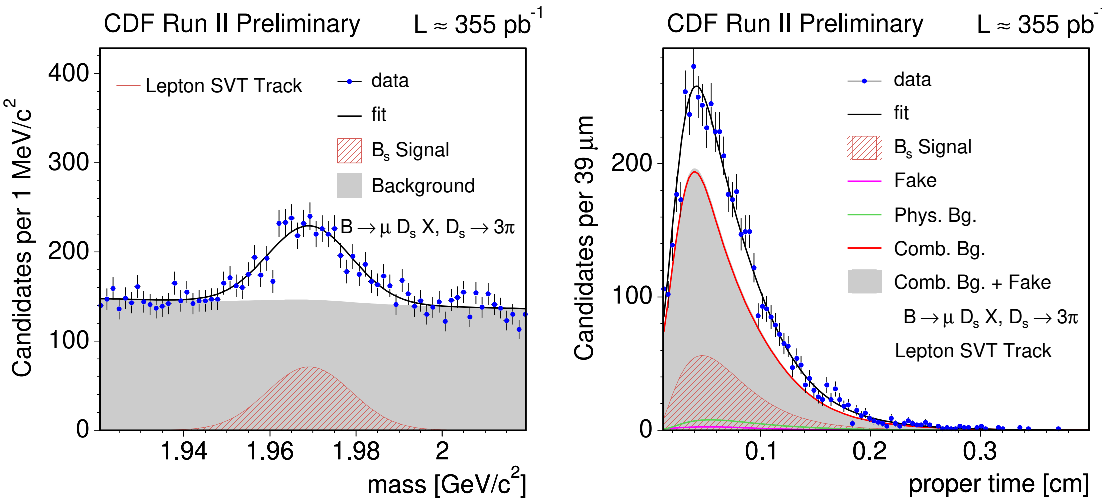


Figure 4-42: Mass and proper decay time fit projections for $\mu^+ D_s^-$ B trigger with $D_s^- \to \pi^+\pi^-\pi^-$.

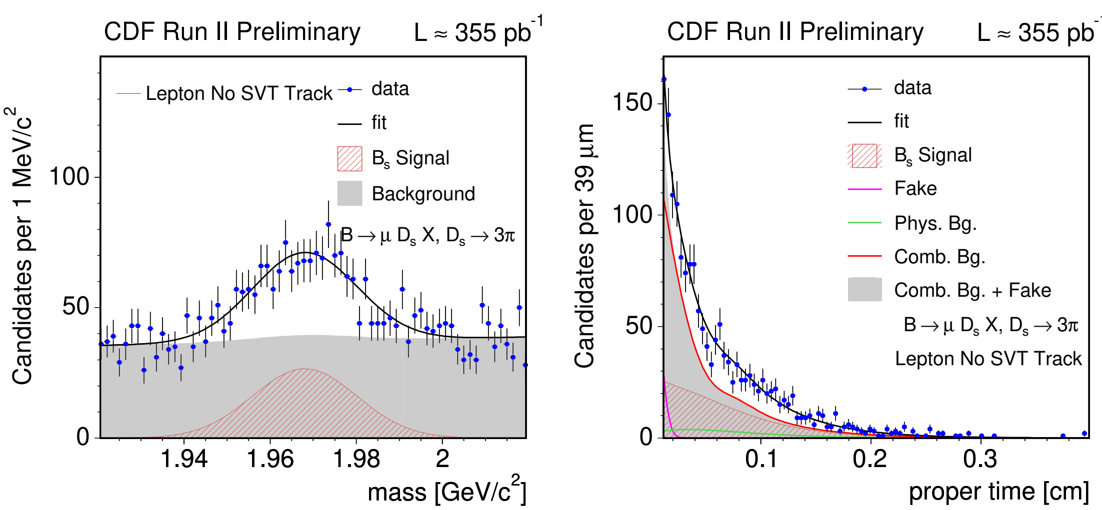


Figure 4-43: Mass and proper decay time fit projections for $\mu^+ D_s^-$ D trigger with $D_s^- \to \pi^+\pi^-\pi^-$.

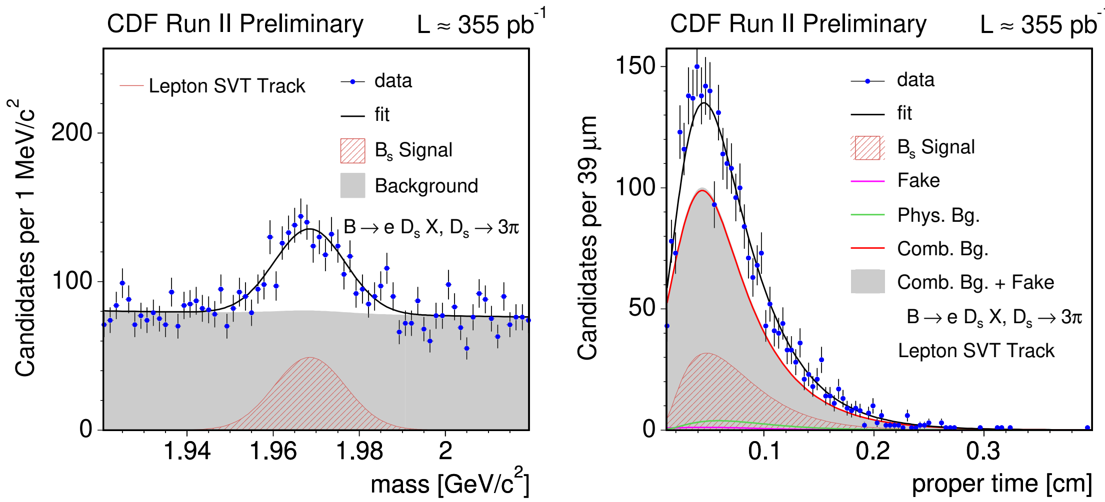


Figure 4-44: Mass and proper decay time fit projections for $e^+D_s^-$ B trigger with $D_s^- \to \pi^+\pi^-\pi^-$.

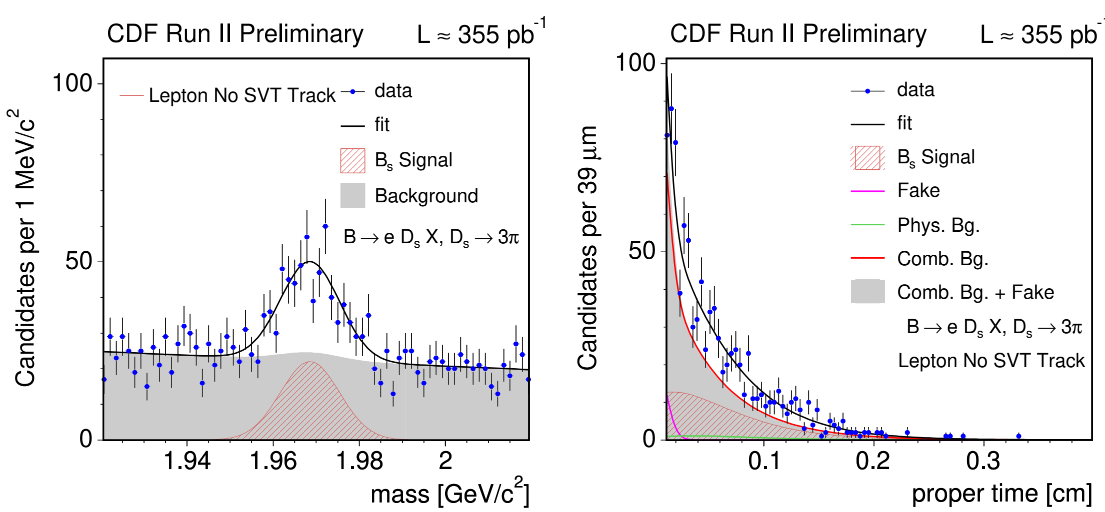


Figure 4-45: Mass and proper decay time fit projections for $e^+D_s^-$ D trigger with $D_s^- \to \pi^+\pi^-\pi^-$.

4.7 Calibration of proper time resolution

The accurate estimation of the t resolution is relevant for ensuring a precise description of the signal proper decay time distribution. It directly enters the signal PDFs. This effect becomes more conspicuously mattering in the study of t-dependent oscillations, in particular when probing high oscillation frequencies and for establishing reliable exclusion regions.

Corrections to the proper decay time uncertainty are parameterized through a *scale factor* S_t as introduced in Section 4.1.2. This factor is extracted from the data directly, namely by adjusting the width of a prompt Gaussian component describing the t distribution of combinatorial background events. This was in effect the procedure employed for the $J/\psi K$ samples where the parameter S_t is part of the fitting procedure, as expressed in (4.33), being determined predominantly from the mass-sidebands which contain a large fraction of prompt events. For the other samples, however, which are affected by the SVT trigger requirements, such prompt component is greatly reduced, and the σ_t scale factor cannot thus be determined concurrently with the fit in the exact same fashion described for the t-unbiased samples.

An alternative, more detailed approach is put forth for estimating corrections to the proper time uncertainties for the t-biased samples. It is based on a calibration sample designed to be predominantly prompt in decay time and also sharing topological and kinematic attributes with those B decay signal samples. With such a sample at hand, the same fitting framework and techniques which were described earlier in this chapter are applied to extract the scale factor values from the fit to the prompt peak. The additional background components of the sample are also described in the fit.

As already mentioned, the effect of these corrections is more determining for the study of B_s oscillations. To further facilitate the transfer of the scale factor from the calibration sample to the B_s samples its dependence on various topological and kinematical quantities is parameterized. For the B_s case therefore per-event corrections are derived based on those parameterizations. For simplicity, for the t-biased B^+ and B^0 samples a common value is derived. The proper time uncertainties σ_t, which are part of the fit input, are then re-scaled as described in (4.10) before the data is passed on to the fitter.

Calibration sample

The main intended characteristic of the calibration sample is that it be *prompt* – that is, it should contain B-like vertex candidates which are expected to coincide with the primary vertex of the $p\bar{p}$ interaction. This is accomplished by removing any requirements on impact parameter or transverse decay length which would bias or sculpt the distribution in proper decay time. Additionally, it should mimic as closely as possible the kinematics and vertex topology of the t-biased signal samples – $B \to D\pi(\pi\pi)$ and $B \to DlX$.

The same two displaced tracks trigger dataset, from which the signal samples are extracted, is used. The method of reconstruction for the vertex candidates is to pair a real D meson with a prompt track (or three prompt tracks, depending on the channel topology) at the primary vertex. Both trigger tracks are required to be from the D meson and *not* the prompt track(s) with which it is paired. Cuts similar to those used for selecting the B samples are used, in order to have similar sample kinematics from the start. Tight cuts are also imposed on the D candidates to further maximize their signal content. Selection requirements are compiled in Table 4.11.

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

Table 4.11: Selection requirements for σ_t calibration sample.

The resulting calibration sample is larger than the corresponding hadronic B_s samples by a factor of ~ 500 for the $D\pi$ topology and ~ 100 for the $D\pi\pi\pi$ topology.

Global fit

A fit is performed to the proper decay time distribution of the sample. No distinguishing structures extist in mass space between the components of the sample, and the likelihood is given by the t-PDF only. The latter corresponds to a model similar to that used for parameterizing the combinatorial background in t-unbiased samples given by (4.33). The dominant component is, as designed, given by a prompt Gaussian. It is formed of events containing a real, prompt D meson and one (three) prompt track(s). In addition, there is a smaller contribution [68] from secondary D mesons originating from B decays, mis-reconstructed D mesons, and tracks not originating from the primary vertex. These are described by negatively and positively lived exponential tails of the form of (4.31). More specifically, and in addition to the prompt component, the $D\pi$ topologies are found to be best fitted with one long lived and one short lived exponential tail pair symmetric about zero, while for the $D\pi\pi\pi$ topologies a single asymmetric exponential tail pair is found to form an adequate model. Representative t distributions and associated fit projections for the $D\pi$ and $D\pi\pi\pi$ topologies are shown in Figure 4-46.

The components just described which make up the PDF correspond to a delta function (4.30) and exponential (4.31) functions, convoluted with a Gaussian resolution function, $G(t; S_t\sigma_t)$, whose width is given by the candidate's measured t uncertainty multiplied by the scale factor, S_t. The latter, which is common to all PDF terms, is the fit parameter of interest. Table 4.12 lists the scale factor fit results for the total sample. These corroborate the value of 1.4 which was chosen as the common scale factor value to be used for all B^+ and B^0 t-biased samples.

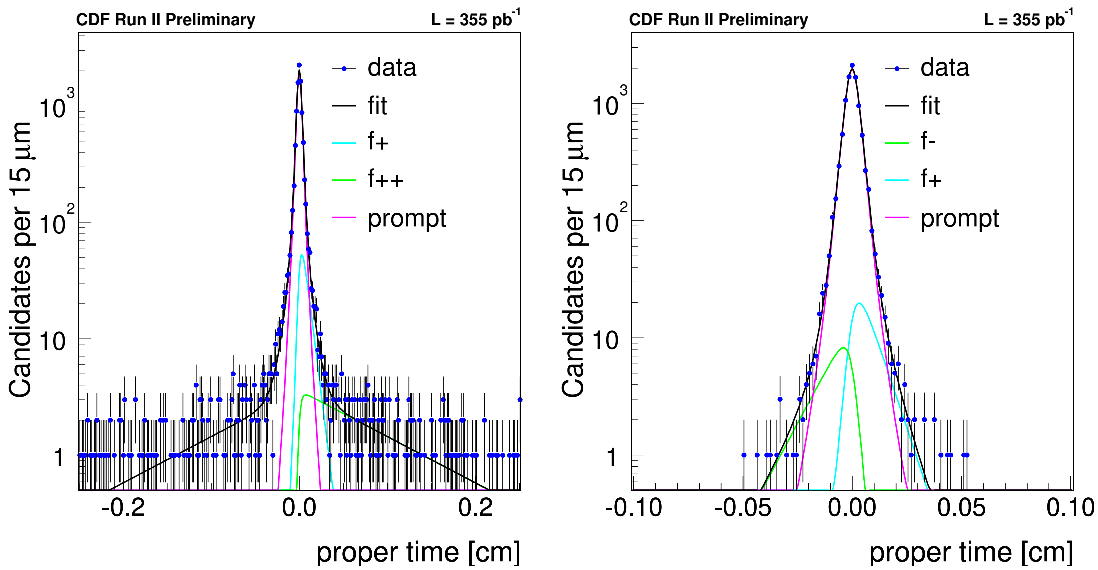


Figure 4-46: Representative t distributions and fit projections for the $D\pi$ (left) and $D\pi\pi\pi$ (right) topologies.

topology	N	S_t
$D\pi$	338000	1.406 ± 0.003
$D\pi\pi\pi$	35500	1.492 ± 0.008

Table 4.12: Total sample size N and fitted scale factor S_t for the calibration sample topologies; uncertainties are statistical only.

Scale factor dependencies and tuning

We now pursue a parameterization of the scale factor based on a set of variables on which it is observed to have dependence. These dependences are then used to estimate the scale factors for the signal samples on a per-event basis. The motivation for doing so is two-fold. It allows for a more systematic and proper transfer between the calibration and the signal samples, by effectively taking into account relative *kinematical* differences. Additionally, by scaling the measured proper time resolution by different values, candidates with good resolution are identified and used to the fullest extent, while candidates with poorer resolution are de-weighted in the smearing (4.14) of the signal PDF.

Due to its large size, the calibration sample can be binned in different variables and examined for the dependences of S_t. The following set of variables characterizing the $D\pi(\pi\pi)$ system is employed in the parameterization:

- angular distance (between the D and the pion(s)), ΔR,
- isolation, $I = p_T(D\pi(\pi\pi))/\sum_i p_T^i(\Delta R < 0.7)$,

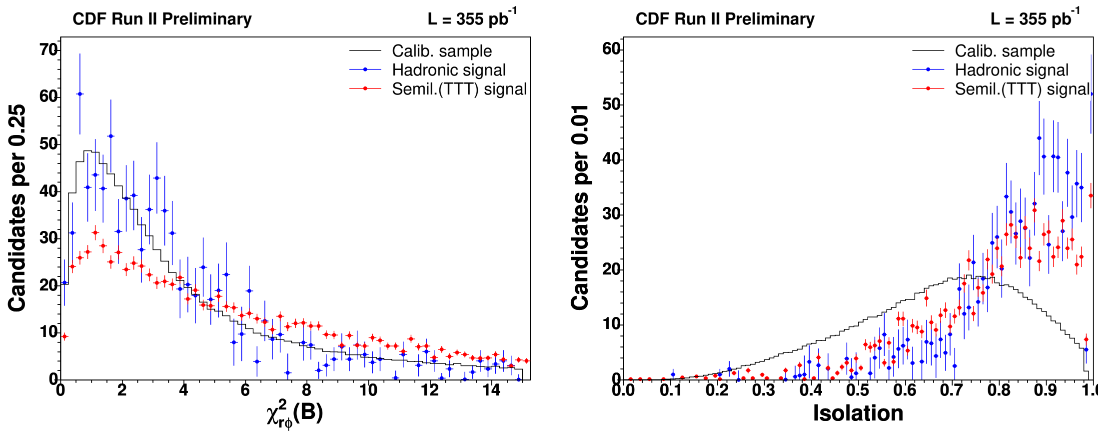


Figure 4-47: Comparison of the calibration sample (black line) and sideband-subtracted $B_s \rightarrow D_s\pi/l$ distributions of variables from hadronic (blue) and semileptonic (red) samples, for the variables $\chi^2(B)$ (left) and isolation (right). Distributions are scaled to facilitate comparison.

- pseudo rapidity, η,
- beamline axis position, z,
- chi-square of the vertex fit in the transverse plane, χ^2.

Figure 4-47 overlays the distributions of the calibration sample and the sideband-subtracted signal samples $B_s \rightarrow D\pi/l$, for the selected variables. The disagreements in these distributions are not of concern, precisely because the respective dependences of S_t are parameterized (Figure 4-48). Rather, the important consideration here is that the calibration sample covers the phase space of the signal samples.

The scale factor parameterization is implemented by factorizing the dependence on the various variables, that is

$$S_t(\Delta R, I, \eta, z, \chi^2) = f_{\Delta R} \cdot f_I \cdot f_\eta \cdot f_z \cdot f_{\chi^2} \ , \tag{4.36}$$

where the various f denote polynomial functions (in fact, parabolas) obtained by fitting the dependence on the individual quantities as illustrated in Figure 4-48. The polynomial parameters are adjusted in an iterative fashion. That is, the scale factor variation with respect to one quantity is first fitted and corrected for, and then the correction for a succeeding variable is extracted and applied. The pattern is continued until the parameterized scale factor produces a reasonably constant behavior close to unit. This is illustrated in Figure 4-48 where the original dependence patent on the leftmost plot is absorbed once the parameterization is applied as it is seen in the rightmost plot.

It is the tuned parameterization achieved in (4.36) that is used to scale the proper time uncertainties (4.10), according to the individual candidates' characteristics, in

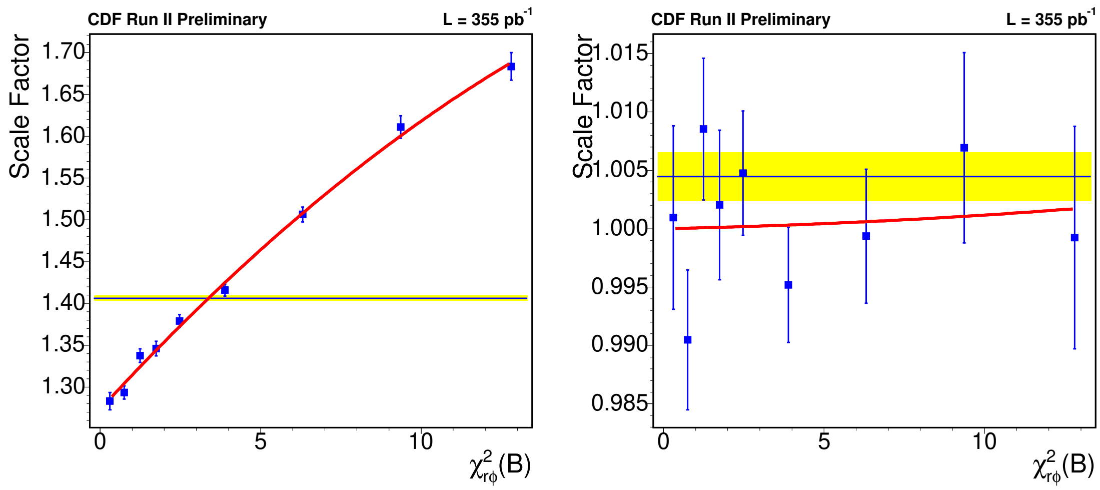


Figure 4-48: Scale factor dependence before (left) and after (right) parameterization, illustrated for the χ^2 vertex fit in the $D\pi$ topology.

the case of the B_s signal samples. In Table 4.13 we summarize the average of the corrections to σ_t implemented in the various hadronic B_s modes.

$B_s \to D_s^- \pi^+$, $D_s \to \phi\pi$	1.44137
$B_s \to D_s^- \pi^+$, $D_s \to K^* K$	1.39889
$B_s \to D_s^- \pi^+$, $D_s \to \pi\pi\pi$	1.37909
$B_s \to D_s^- \pi^+ \pi^- \pi^+$, $D_s \to \phi\pi$	1.39659
$B_s \to D_s^- \pi^+ \pi^- \pi^+$, $D_s \to K^* K$	1.39722

Table 4.13: Average correction factors to σ_{ct} for different B_s modes.

Systematic uncertainties

Systematic aspects associated to the method employed for deriving the corrections to the proper time uncertainties are now addressed. Foremost, we intend to ensure that the achieved corrections are not incorrectly favorable when establishing exclusion conditions for the oscillation frequency. Hence only an upper bound on the scale factor is measured.

The following systematic sources are considered.

***D* background content:** The likelihood model used for fitting the calibration sample employs several simplifications. A main assumption in this respect is that the D candidates are purely signal, in view of the tight D selection requirements, and no modeling of combinatorial D background is thus included. To evaluate the effect of this model simplification we increase, by approximately two times, the level of background acceptance in the sample by doubling the width of the D mass cut window. We then determine the shift in the fitted average scale factor values on this background-enriched sample.

***D* impact parameter requirement:** The requirement on the D impact parameter is used to select candidates whose momentum points to the primary vertex. This may bias the vertex calibration by preferentially rejecting badly measured events which could have large scale factors. To evaluate this effect the impact parameter selection requirement is varied by 20 μm around the nominal 100 μm cut, and the modified samples are fitted for the average residual scale factors.

Transferred parameterization: The scale factor parameterization is to be applied to the various B_s signal modes. While it is derived in a calibration sample involving only D^- candidates, those target B_s samples are based on distinct D_s topologies. Other calibration topologies were designed, including $D_s\pi$, $D^0\pi$, and additional $D\pi\pi\pi$ modes. The systematic uncertainty associated with the transfer of scale factors between different D decay topologies is evaluated by fitting for the average scale factors in the various samples with interchanged parameterizations; the largest observed deviation from unit is taken as the uncertainty.

Tuning residuals: The tuning procedure employs a limited number of variables while there may be others on which the scale factor also depends. The assumption was also made that the dependences completely factorize. The effect of such incompleteness of the tuning procedure are small and limited to the level of the residual deviations from the average fitted scale factor. These are taken as the dominant relative systematic uncertainty source. The relative systematic uncertainties evaluated for each source are shown in Table 4.14.

systematic source	relative uncertainty [%]
D background	< 1.0
D impact parameter	≤ 1.0
transferred parameterization	3.5
tuning residuals	4.0
total assigned	6.0

Table 4.14: Contributions to the σ_t scale factor systematic uncertainty.

4.8 Résumé

The unbinned fitting framework and associated likelihood description of the data samples in mass and proper decay time spaces in the absence of flavor tagging information have been presented.

The quantities of central interest for this dissertation are associated directly to the description of the proper decay time of the reconstructed B meson signals in the data samples. The discrimination of these signals from the various other components which constitute the samples chiefly benefits from the characterization of the involved mass distributions. The fits of the data are accordingly performed based on input information from both proper time and mass spaces in a concurring fashion. This input data is moreover provided for the individual B mesons candidates, thus following an unbinned approach, which maximizes the usage of the available information extracted from the data samples.

The components which contribute to the data samples were identified in Chapter 3. Each of these components is modelled in the spaces of the input quantities to the likelihood function. The characterization of background contributions of the combinatorial type is inferred from mass-sideband regions, while that of backgrounds of the physics type, originating from non-signal b-hadron decays, readily benefits from expected templates found from Monte Carlo simulation of those processes. The B signals are recognized in mass space as Gaussian peaks whose widths are determined by the detector resolution.

The B mesons lifetimes, here the main parameters of interest, are part of the proper decay time probability density functions for signal components. Analysis methods are developed in order to give necessary account of effects coming from trigger and selection requirements, and partial reconstruction which characterize the data samples. Analytical techniques are further examined and implemented that result in an optimization of speed and accuracy of the fitting process.

The measured B meson lifetimes are summarized (uncertainties are statistical only):

	lifetime, τ [ps]
B^+	1.654±0.008
B^0	1.557±0.007
B_s	1.479±0.020

These are found to be in good agreement with the world average values [1] — 1.670 ± 0.017 ps (B^+), 1.535 ± 0.014 ps (B^0), 1.460 ± 0.057 ps (B_s). The statistical precision achieved is competitive with those averages. Preliminary evaluation of the systematic uncertainties performed in previous iterations of the analysis estimate less than 4% relative systematic uncertainty contributions.

The issue of proper time uncertainty calibration, of particular relevance for the study of B_s oscillations, is also thoroughly addressed.

The fitting framework here described further provides the underlying structure for the study of neutral B meson mixing once treatment of flavor information is incorporated. This information is acquired by applying to the data samples flavor tagging techniques. Such are the matters to be explored in the ensuing chapters.

Chapter 5

Flavor tagging

The flavor of B mesons refers throughout this dissertation to their particle B or anti-particle $\bar{B}$ state. It is, equivalently, given by the corresponding $\bar{b}$ or b quark content: a meson state B_q, where q denotes a u, d or s quark, contains a $\bar{b}$ quark; a $\bar{B}_q$ state is formed instead of a b quark. The B meson candidates are reconstructed in flavor specific final states, listed in Table 3.1. For example, in $B \to D\pi$ and $B \to DlX$ decays, the positively (negatively) charged pion or lepton is associated to a B ($\bar{B}$) state. The flavor at decay time is therefore determined from the decay products of the B meson candidate. For neutral B meson systems, B^0 or B_s, mixing occurs between the two flavor states, B-$\bar{B}$, as these do not coincide with the corresponding mass eigenstates (Appendix A). The study of flavor oscillations which occur in these systems therefore requires, in addition to the determination of the final state meson flavor, the identification of the B-flavor at production time. Several flavor tagging techniques are used in this task.

In the Tevatron's $p\bar{p}$ collisions b-hadrons are produced from the hadronization of b-quarks originating from $b\bar{b}$ pairs produced in the hard parton interactions. The B meson whose decay products satisfy the trigger requirements is referred to as the "trigger B", while the other b-hadron in the event is called the "opposite-side B". The flavor tagging algorithms may also be broken down into two classes: *same-side tagging* (SST) and *opposite-side tagging* (OST). The former explores flavor-charge correlations between the trigger B and tracks nearby in phase space, while the latter is based on the identification of some property of the opposite-side B to determine its b quantum number, from which the production flavor of the trigger B can be inferred.

A distinguishing aspect between the same-side and opposite-side methods is that the performance of the former depends on the species of the B meson candidate being tagged, while that of the latter does not. This is the case as the hadronization of the b quark and of the $\bar{b}$ quark take place in an incoherent fashion. It follows, in particular, that the properties of the OST algorithms may be measured for instance in samples of B^+ and B^0 mesons and transferred directly to samples of B_s mesons. This possibility is also explored in the coming chapters. The same strategy is not applicable in the case of the SST method.

The SST and OST techniques are explored respectively within and outside an isolation domain which, for a given event, corresponds to a cone ΔR of 0.7 around the reconstructed B candidate direction. The two classes of tagging algorithms make in this way use of uncorrelated information.

In general, for realistic algorithms and depending also on the characteristics of each event, a flavor tagging method does not always give a correct answer and sometimes fails alltogether to make a decision. Let, for a given sample, N_r denote the number of candidates correctly tagged, N_w denote the number of incorrect tags, and N_n the number of candidates for which the tagger failed to provide a decision. The following quantities are thus conveniently defined. The tagging *efficiency*, ϵ, is given by the fraction of events which are non-trivially tagged, that is for which the algorithm provides a definite decision,

$$\epsilon \quad = \quad \frac{N_r + N_w}{N_r + N_w + N_n} \; . \tag{5.1}$$

The tagging *dilution*, $\mathcal{D}$, is a measure of the quality of the tagging decision,

$$\mathcal{D} \quad = \quad \frac{N_r - N_w}{N_r + N_w} \; , \tag{5.2}$$

being related to the mistagging probability, $w = N_w/(N_r + N_w)$, by $\mathcal{D} = 1 - 2w$.

The tagging *effectiveness*, $\epsilon\mathcal{D}^2$, provides the figure of merit of a tagging algorithm. It effectively scales the size of a mixing sample: the statistical power of a sample of size N tagged by the algorithm is equivalent to that of a perfectly tagged sample of size $N\,\epsilon\mathcal{D}^2$. Accordingly, the expected statistical uncertainty, σ, in a mixing measurement becomes (B.17)

$$\sigma \quad \propto \quad \frac{1}{\sqrt{N \cdot \epsilon\mathcal{D}^2}} \; . \tag{5.3}$$

In developing and optimizing a tagging algorithm it is therefore the tagging power $\epsilon\mathcal{D}^2$, rather than ϵ and $\mathcal{D}$ separately, which is to be maximized.

5.1 Opposite side tagging

A highest quality opposite-side tagging method would involve the full reconstruction of the opposite-side B, and discovering it to be, say, a charged B meson which does not undergo flavor mixing. From this one would be assured of the b or $\bar{b}$ quark content of the opposite-side B, and thus also of the production flavor of the trigger B which would be the opposite. Although the dilution would be maximal (*i.e.* unity) for such an algorithm, in practice the efficiency for reconstructing completely the opposite-side b hadron is much too small to be useful. Instead of the full reconstruction, thus, inclusive properties of the opposite-side B indicative of its flavor are employed, this way achieving an appropriate compromise between dilution and efficiency.

The tagging methods specifically explore the sign of the charge of leptons originating from semileptonic decays of the opposite-side B, as well as that of the opposite-side b-jets.

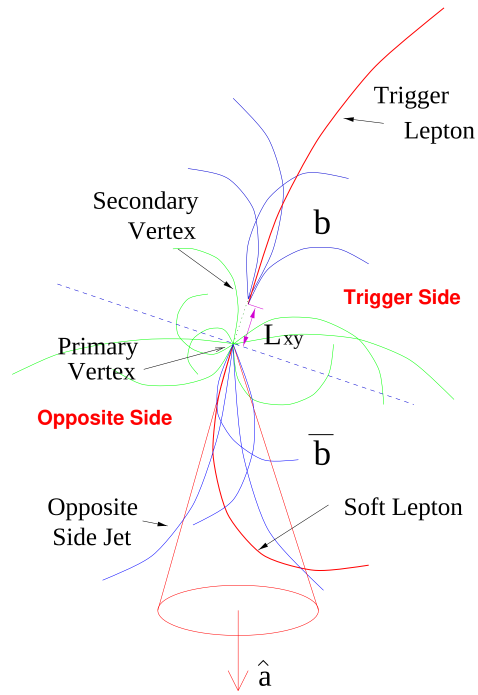


Figure 5-1: Illustration of a $b\bar{b}$ event and opposite-side tagging.

The performance of the algorithms is established in a high statistics inclusive B semileptonic trigger sample. The trigger and opposite sides are illustrated in Figure 5-1. These will be further described below. The performance is ascertained by comparing the charge of the trigger side lepton with that of the opposite-side lepton or jet. Each event is classified accordingly into one of the following three mutually exclusive categories: the charge of the trigger lepton coincides with (i) or is opposite to (ii) the charge of the opposite-side lepton or jet, or, otherwise, (iii) the latter cannot be inferred for the event. The number of events in each class is used as an estimate of the number of incorrectly tagged (N_w), correctly tagged (N_r), and non-tagged (N_n) events, respectively. These estimates are in turn used to compute the efficiency (5.1) and dilution (5.2) of the algorithm, after having taken into account effects having to do with the sample composition and mixing.

5.1.1 Soft lepton taggers

Lepton tagging exploits the charge sign of the lepton in the decays $b \to Xl$. In case the lepton originates from the semileptonic decay of the opposite-side B,

$$b \to c\, l^- \bar{\nu}_l X\ , \qquad\qquad \bar{b} \to \bar{c}\, l^+ \nu_l X\ ,$$

and in case the latter did not undergo flavor oscillation, the lepton charge determines the production b-flavor.

The accuracy of this determination is affected in principle by several factors. In an ensemble of tags, the opposite-side b hadron can be seen as a mixture of B^+, B^0, B_s and other meson and baryon states, which are not identified. While the B^+ and b baryons fully contribute, the neutral B mesons however undergo flavor mixing: a fraction of the B^0 mesons will have oscillated into a state different from the production state before decaying, thus giving an incorrect tag, and the B_s system, being fully mixed, does not provide any tagging power. It may happen, additionally, that the lepton does not come from the direct b semileptonic decay, but instead from the sequential transition $b \to c \to lX$. The charge of the lepton is in this case opposite to that of the direct decay, and results in an incorrect tag. It may also happen that the lepton candidate is misidentified, *i.e.* that it is fake. The overall correctness of the flavor determination is statistically determined, and expressed by the tagging dilution.

The method is explored and optimized for muon and electron samples and the corresponding algorithms are respectively named *soft muon tagger* (SMT) [54] and *soft electron tagger* (SET) [55]. In each case, a multivariate lepton likelihood $\mathcal{L}^l$ ($l = \mu, e$) is constructed for discriminating real from fake leptons. It combines information from the various calorimeter and muon detector subsystems, along with track matching and additional lepton identification variables. The lepton likelihood estimator is constructed as

$$\mathcal{L}^l = \frac{S^l}{S^l + B^l} \tag{5.4}$$

where the signal S^l and background B^l likelihood terms

$$S^l = \prod_i S_i^l, \qquad B^l = \prod_i B_i^l, \tag{5.5}$$

are given as the product of the corresponding probability distributions of the employed discriminating variables. The set of used variables is specific to each algorithm, and is given below. The likelihood construction (5.4) implies that it is expected to approach unity for real leptons and zero for fake leptons.

Dependences of the methods' dilutions are further explored, in terms of the lepton likelihood, the transverse momentum of the lepton relative to the axis of the jet to which it belongs, and, for muon candidates, the involved muon detector subsystems.

Muon identification

Muons are identified as charged particles which traverse the tracking and calorimeter systems of the detector, reaching the muons chambers. A muon candidate thus consists of a stub in a muon chamber associated with a COT track that extrapolates in the vicinity of the stub. The matching between the muon stub and the extrapolated track is quantified through the variables ΔX, $\Delta\Phi$, and ΔZ. These matching variables depend on the muon transverse momentum. For muons of lower momenta the effects of multiple scattering are larger, and the matching is less accurate. Momentum independent matching variables are obtained after rescaling with the corresponding p_T dependent widths, $\Delta x = \Delta X/\sigma_{\Delta X}$, $\Delta\phi = \Delta\Phi/\sigma_{\Delta\Phi}$, and $\Delta z = \Delta Z/\sigma_{\Delta Z}$.

In addition to the track–stub matching variables, the energy depositions by the muon candidate in the electromagnetic and hadronic calorimeters, denoted $E_{\rm em}$ and $E_{\rm had}$, are considered as well. Dependences of $E_{\rm had}$ on the transverse momentum of the muon candidate are considered, by parameterizing it in three p_T ranges. Momentum dependences are neglected for $E_{\rm em}$. The isolation, I, of a track j is defined as the ratio between the track transverse momentum, p_T^j, and the sum of the transverse momenta of all tracks within a cone $\Delta R < 0.4$ about the track j,

$$I = \frac{p_T^j}{\sum_i p_T^i}\,, \qquad \Delta R(i,j) = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4\,. \tag{5.6}$$

To account for the dependence of the electromagnetic energy distribution on isolation, different $E_{\rm em}$ templates are used for isolated ($I > 0.5$) and non-isolated ($I < 0.5$) muons. No noticeable dependence on isolation is observed for $E_{\rm had}$.

The characteristic behavior of the discriminating variables for real muons is extracted from samples of $J/\psi \to \mu^+\mu^-$ decays. Fake muon candidates may be produced by misidentified pions, kaons and protons, that reach the muon chambers by punching-through the calorimeters. These are studied in samples of $K_s^0 \to \pi^+\pi^-$ (pions), $D^0 \to K^-\pi^+$ (kaons) and $\Lambda^0 \to p^+\pi^-$ (protons) decays. The obtained distributions of the discriminating variables – Δx, $\Delta\phi$, Δz, $E_{\rm em}$, $E_{\rm had}$ – for both real and fake muons are taken as templates for forming the likelihood factors S^μ and B^μ according to (5.5). The most discriminating power is provided by the hadronic energy distributions, and, especially for high p_T muon candidates, by the track-stub matching quantities.

Electron identification

Electrons are identified as charged particles whose energy is deposited mostly in the electromagnetic calorimeter. Electron candidates consist of COT tracks that extrapolate to the calorimeter, satisfying the thresholds of 1.0 GeV/c in transverse momentum and of 0.8 GeV/c^2 in transverse electromagnetic energy. A 2-tower cluster is formed by adding to the extrapolated seed tower its nearest neighboring tower in z if the transverse energy deposited therein exceeds 100 MeV/c^2. The electromagnetic energy $E_{\rm em}$ is defined as the combined deposited energy in the 2-tower. The hadronic energy $E_{\rm had}$ is evaluated in a similar fashion. The local isolation I of a track is defined, similarly to (5.6), as the ratio of its transverse momentum to the sum of transverse momenta of the tracks in a $\Delta R < 0.7$ cone which extrapolate to the 2-tower cluster. Locally isolated candidates ($I = 1$) are required to satisfy $E_{\rm had}/E_{\rm em} < 0.125$, while for locally non-isolated candidates ($I < 1$) $E_{\rm had}/E_{\rm em} < 0.5$ is imposed. The ratio of the electromagnetic energy to the track momentum, $E_{\rm em}/p$, is a strong discriminator against pions particularly for isolated electron candidates.

Further discriminating information is provided by the maximum shower (CES) and the central pre-radiator (CPR) detector sub-systems. The χ^2 matching variables in the CES, χ^2_x and χ^2_z, provide a comparison of the shower profiles in the CES wire and strip views with the same profiles extracted from electron test beams. The variable $E_{\rm ces}/p^*$ corresponds to the corrected wire cluster pulse height in the CES, and provides very good separation between electrons and pions even at low p_T. The CES variables ΔX and ΔZ correspond to the distance in the transverse and r-Z planes, between the track extrapolated to the CES radius and the actual cluster position measured in the CES. These are stabilized against p_T variations by scaling with the corresponding p_T dependent widths, $\Delta x = \Delta X/\sigma_{\Delta X}$ and $\Delta z = \Delta Z/\sigma_{\Delta Z}$. Δx is further signed with the track charge. Distributions of these variables are formed for locally isolated and non-isolated candidates, and their discriminating power increases with p_T. The pulse height in the CPR corrected for its $\sin(\theta)$ dependence, $Q_{\rm cpr}$, is a rather strong discriminator between electrons and pions. The same is also true for the energy loss measured in the COT, dE/dx, whose usefulness is enhanced by the fact that, unlike the CES and CPR based variables, its separation power increases with the p_T of the track candidate.

A pure electron sample is obtained by reconstructing electrons from photon pair-conversions, $\gamma \to e^+e^-$. Hadrons such as pions, kaons, and protons can also fake electron candidates. Properties of fake candidates are obtained from samples of $K^0_s \to \pi^+\pi^-$ decays. These samples are used to extract parameterized distributions of the discriminating variables – $E_{\rm had}/E_{\rm em}$, $E_{\rm em}/p$, χ^2_x , χ^2_z , $q\Delta x/\sigma_x$, $\Delta z/\sigma_z$, $E_{\rm ces}/p^*$, $Q_{\rm cpr}$, dE/dx – for both real and fake electrons, to form the associated likelihood factors S^e and B^e as in (5.5).

Tag selection

When the algorithms are employed to tag the production flavor of a B meson candidate, certain criteria are applied to the potential soft lepton candidates and, if those criteria are satisfied, one of them is finally elected to provide the tag decision.

The soft lepton candidates are required not to coincide with the reconstructed daughter tracks of the trigger-side B meson. The transverse momentum is required to exceed 1.5 GeV/c and 2.0 GeV/c for muon and electron candidates, respectively. An impact parameter requirement of $|d_0| < 0.3$ cm is used to help rejecting hadrons which decay into muons. Track quality requirements of at least 10 axial and 10 stereo COT hits along with at least 2 SVX r-ϕ hits are imposed on electron candidates.

Jets present in the event are found based on a track cone clustering algorithm [69]. If multiple candidates satisfy the mentioned criteria, a globally isolated lepton (no other tracks found within $\Delta R = 0.7$) is chosen if found. Otherwise, the appointed candidate is that with the highest relative transverse momentum p_T^{rel} with respected to the jet axis to which it belongs.

Few additional criteria are further imposed bearing in mind the concurrent application of the various tagging methods to common mixing samples. A minimum lepton likelihood of 0.05 is required for both electron and muon candidates; events with such poor quality soft leptons are expected to be more accurately tagged with the jet charge method.

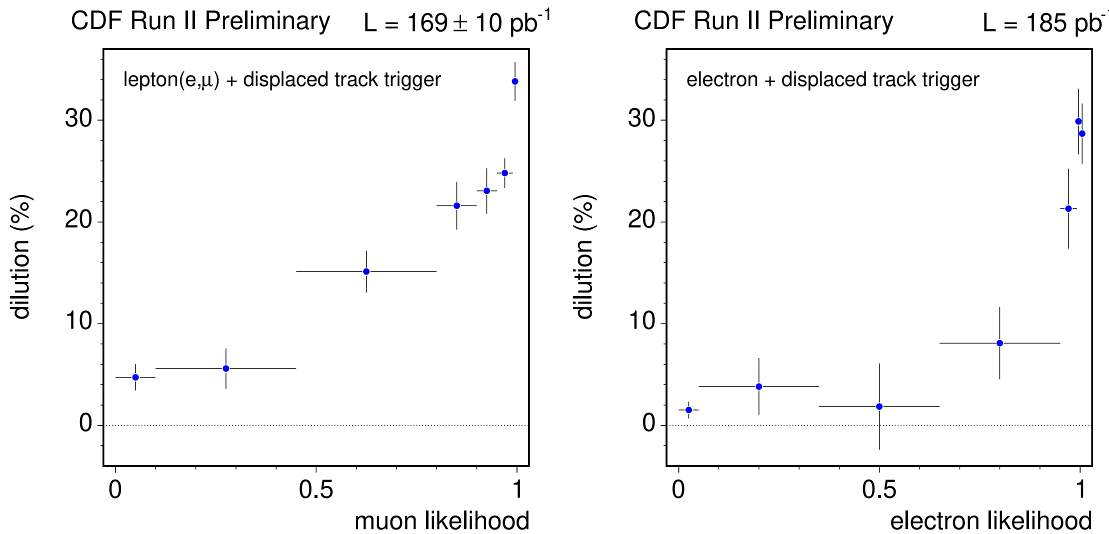


Figure 5-2: Dilution dependencies of the soft lepton taggers as function of the muon and electron likelihoods; the various muon types have been combined on the left plot.

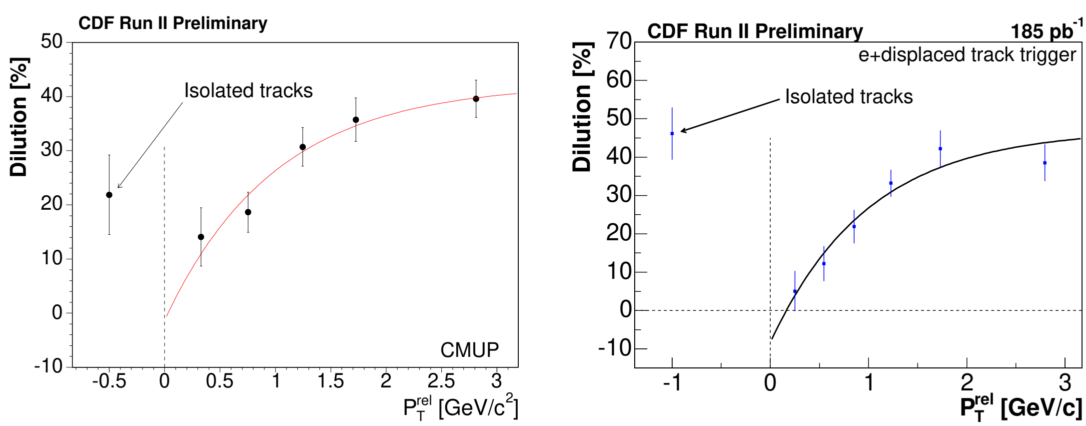


Figure 5-3: Dilution dependencies of the soft lepton taggers as function of the transverse momentum relative to the jet axis: (left) SMT for CMUP muon type, (right) SET for $\mathcal{L}^e > 0.85$.

Dilution dependences

The performance of the tagging algorithms shows expected dependences that are explored. Namely, once quantities on which the performance depends are identified, rather than imposing hard thresholds for selecting only the best quality tags, the dilution is instead parameterized in terms of such variables. The tagging efficiency is not decreased, and the more reliable tag gets a higher weight thus enhancing the overall tagging power.

One such variable is the lepton likelihood, $\mathcal{L}^l$. The dilution increase with increasing values of the lepton likelihood is accounted for, and shown in Figure 5-2. This behavior is expected based on that a more pure lepton sample should reflect the b flavor more accurately.

The dilution is verified also to increase with the transverse momentum p_T^{rel} of the lepton with respect to the axis of the jet in which the lepton is found. An adequate empirical description of this dependence is provided by

$$\mathcal{D}(p_T^{rel}) \;=\; A_\ell \left(1 - e^{-p_T^{rel}+\delta_\ell}\right) \; . \tag{5.7}$$

and is shown in Figure 5-3. This qualitative behavior originates from the fact that leptons coming from b quark decays tend to be more widely spread in the plane transverse to the b direction than in jets originating from lighter quarks, due to the larger phase space available.

For the SMT algorithm, a separate implementation is derived according to the detector sub-system which identifies the muon stub: CMU, CMP, CMUP (muons with stubs in both CMU and CMP), CMX, and IMU.

5.1.2 Jet charge taggers

Jet charge tagging exploits the charge of jets on the opposite-side. If the jet which originates from the opposite-side b-quark is identified, the charge sign of the b, and thus its flavor, is expected to be given on average by the sign of the weighted sum of the charges of tracks forming the jet.

The implementation [70] of the method takes as input various track and jet properties. The way the various discriminating quantities are extracted and characterized, as well as combined, is distinct from the strategy employed in the case of the SLTs described in Section 5.1.1. Samples of Monte Carlo events are used to study $b\bar{b}$ properties. Good agreement needs thus to be achieved between the relevant quantities extracted from these samples and those observed in the data, which is accomplished by taking into account in the simulation additional $b\bar{b}$ pair creation processes beyond the leading order $b\bar{b}$ production mechanism. The b-jet properties are extracted from these samples which are finally used for jet selection in the data. The available quantities are then combined in an artificial neural network (NN). An improved combination relative to that which could be obtained via a likelihood ratio is achieved, as correlations between the various input quantities are intrinsically taken into account.

Probabilities that individual tracks belong to B decay products are estimated;

these are used to define jet variables, which are combined to provide b-jet probabilities. Once the highest probability jet is selected, the flavor decision provided by the method is given by the sign of the charge computed for the tagging jet as

$$Q_{\mathrm{jet}} \equiv \frac{\sum_i q^i \, p_T^i \, (1+T_P^i)}{\sum_i p_T^i \, (1+T_P^i)} \tag{5.8}$$

where the index i runs over all tracks in the jet, q^i is the charge of the i-th track, p_T^i its transverse momentum, and T_P^i is the track probability indicating how likely it is to be a b daughter. The method is optimized for three mutually exclusive categories of jet quality:

Class 1 (JVX): jets containing an identified secondary vertex, with decay length significance $L_{xy}/\sigma_{L_{xy}}$ greater than 3,

Class 2 (JJP): jets not in the above class and with at least one track with probability T_P greater then 50%,

Class 3 (JPT): jets not in the above classes.

These classes provide different tagging performances and are characterized separately. Dilution dependencies on the jet charge Q_{jet} and NN output for jet probability are further explored and parameterized.

Jet reconstruction

All opposite-side tracks are first identified. These are defined, for each event, to be those found outside the isolation cone $\Delta R > 0.7$ relative to the momentum direction of the trigger B, with tracks used in its reconstruction being explicitly excluded. Tracks need then to satisfy requirements on impact parameter ($|d_0| < 0.15$ cm), on z_0 ($|z_0 - z_B| < 1$ cm), and on transverse momentum ($p_T > 0.4$ GeV/c) to be considered as candidate jet constituents.

Jets in an event are found using a standard CDF software package [69] which implements a track based cone clustering algorithm. Those tracks with $p_T > 1.0$ GeV/c are considered as jet seeds. Pairs of such tracks within a cone ΔR of 1.5 are merged, thus forming new seeds with momentum given by the sum of the momenta of the two original tracks. The procedure is repeated until no track pairs satisfy the merging criteria. Finally, jets are formed by adding to these final seeds the remaining, lower p_T tracks within a cone ΔR of 1.5.

Track probability

Tracks are assigned probabilities for having originated from a B decay chain. This is achieved with a neural network, which is optimized on Monte Carlo events. The NN output is evaluated during the learning process on the basis of track matching to the generator level information.

The input quantities to the NN are selected based on the power to discriminate between whether the track was created at the primary or a secondary vertex, in conjunction with the corresponding level of agreement between data and Monte Carlo. Among the most powerful input variables are the track impact parameter d_0, along with its impact parameter significance δ_0/σ_{d_0} signed with respect to the jet momentum $\vec{p}_{\rm jet}$, with

$$\delta_0 = |d_0|\,\mathrm{sign}(\vec{d_0}\cdot\vec{p}_{\rm jet})\,.$$

The transverse momenta and ΔR, along with the track rapidity with respect to the jet axis are also provided.

The structure employed for the track probability NN consists of three neuron layers. Each input variable is associated to one node in the input layer, the intermediate layer contains six nodes, and the output layer if formed of a single node which provides the network decisions.

Jet probability

The jets reconstructed on the opposite side may have varying characteristics, depending on several factors, which for instance include the presence of the opposite-side B in the detector acceptance, its momentum and decay length. Probabilities are assigned for the candidate jets which indicate the likelihood for being a b-jet. This is achieved employing a neural network combining several jet variables; a three layer structure similar to that used for the track NN is chosen. Several of these variables are constructed using the track probability, denoted T_P, information previously obtained. The weighted number of tracks in the jet is given by $\sum_i T_P^i$. The jet probability J_P is computed as

$$J_P \;=\; P_N \sum_{j=0}^{N-1} \frac{(-\ln P_N)^j}{j!} \qquad \text{with} \quad P_N = \prod_{i=1}^{N} T_P^i\,.$$

Input variables related to properties of secondary vertices, such as the corresponding χ^2 probability, the momentum fraction and the number of tracks involved are considered. Other variables are related to the kinematics and the shape of the jet, such as p_T, invariant mass, and jet spread.

Dilution dependences

Once the tagging jet is identified as the highest probability jet, the method's decision is provided by the sign of the effective charge of the jet in (5.8).

The correctness of this decision is reflected in the tagging dilution. The latter is observed to depend on the following two independent variables: the absolute value of the jet charge $|Q_{\rm jet}|$, as computed in (5.8), and the jet probability variable $P_{\rm nn}$, returned by the b-jet probability NN. The dilution increases with the two variables as expected, and its dependency is observed to be linear. In order to take advantage of such dependencies, each of the samples of tagging jet classes – JVX, JJP, and JPT

– is further divided in 10 bins of the variable $|Q_{\rm jet}| \cdot P_{\rm nn}$, along with an extra bin for jets containing a single track. These dependencies are illustrated in Figure 5-4.

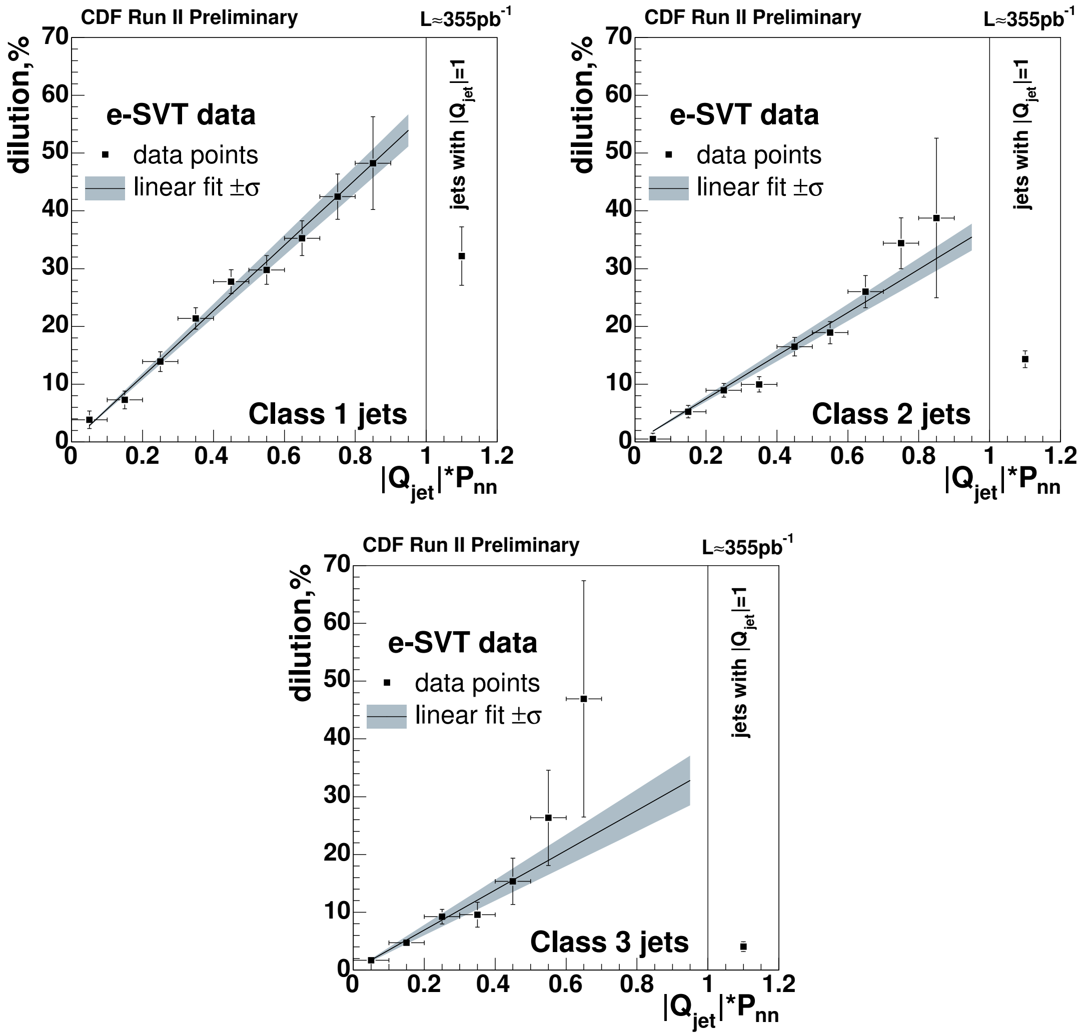


Figure 5-4: Dilution dependency as function of $|Q_{\rm jet}| \cdot P_{\rm nn}$, for each of the jet charge classes: (1) JVX, (2) JJP and (3) JPT.

5.1.3 Tagging performance

Data sample

The opposite-side tagging algorithms are established in an inclusive sample rich in B mesons, collected with the *lepton and displaced track trigger* described in Section 3.1. The large sample size allows for the exploration of dilution dependencies and reasonably accurate parameterizations in terms of relevant characteristics of the events. The sample sizes used for the soft lepton and jet charge taggers correspond to integrated luminosities of approximately 180 and 355 pb^{-1}, respectively.

The employed trigger sample is a not a pure sample of B decays. In addition to events from semileptonic B decays it also contains semileptonic charm decays, hadrons that fake the trigger lepton, and other backgrounds. Several of these background contributions are removed by imposing an event selection criteria which restricts the invariant mass of the lepton and displaced track system to the range of $(2.0, 4.0)$ GeV/c^2. To remove the effect of remaining background events, a subtraction procedure is applied to the distributions of interest. The procedure is based on the assumption that the background sources are symmetric in the signed impact parameter of the displaced track defined as

$$\delta_0^{\mathrm{SVT}} \equiv |d_0|\,\mathrm{sign}(\vec{d}_0 \cdot \vec{p}_{l,\,\mathrm{SVT}})\,,$$

where $\vec{d}_0$ points from the primary vertex to the point of closest approach of the displaced track, and $\vec{p}_{l,\,\mathrm{SVT}}$ is the the combined momentum of the trigger lepton and the displaced track. Figure 5-5 shows the mass distributions of the lepton and displaced track pair for positive and negative signed impact parameter δ_0^{SVT}. Distributions characteristic of pure signal events are obtained by subtracting the distribution with negative δ_0^{SVT} from the corresponding distribution with positive signed impact parameter [71]. In particular, this subtraction procedure is applied in the determination of the number of signal events belonging to each of the tagging classes used in the computation of the two quantities of interest: the efficiency and the dilution of the tagging algorithms.

Given that the trigger-side B meson is only inclusively reconstructed, a few additional track selection requirements are imposed for reinforcing that no trigger-side B daughters are considered as tag candidates. Both the trigger lepton and the displaced trigger track are explicitly excluded from the set of soft lepton candidates and from the set of tracks forming the jet candidates. The mass of the system formed by an eligible opposite-side track along with the trigger lepton and the displaced trigger track is required to be greater than about 5 GeV/c^2. Soft lepton track candidates that belong to jets containing the trigger lepton are rejected, and so are track candidates for tagging jets which lie within a cone ΔR of 1.6 relative to the trigger lepton and displaced track system.

Trigger side dilution correction

The OST dilution is calculated from the asymmetry in the number of events belonging to the classes of agreeing and disagreeing charges, between the trigger lepton and the opposite-side lepton or jet. The observed asymmetry however is decreased, *i.e. diluted*, due to effects related not solely to the opposite-side but also to the trigger-side. Denoting the corresponding asymmetry contributions by $\mathcal{D}$ and $\mathcal{D}_{trig}$, and to the extent that these are uncorrelated, the observed asymmetry $\mathcal{D}_{raw}$ is given by their product. That is,

$$\mathcal{D} \;=\; \frac{\mathcal{D}_{raw}}{\mathcal{D}_{trig}}\,, \tag{5.9}$$

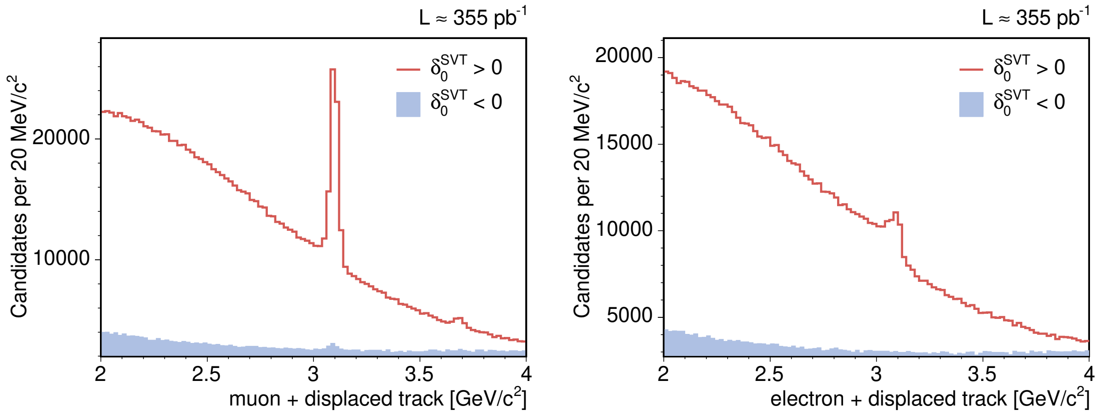


Figure 5-5: Mass distribution of the lepton and displaced track system, for the muon (left) and the electron (right) samples; the peaks at $\sim$ 3.1 GeV/c^2 correspond to $J/\psi \to l^+l^-$ decays.

which implies that the OST dilution is greater than the measured, raw asymmetry.

The trigger B may have undergone flavor oscillation by the time it decayed; also, the trigger lepton may have been produced in a sequential $b \to c \to l$ transition. Such trigger-side dilution effects are evaluated from Monte Carlo samples of semileptonic B decays including detector and trigger simulations. The dilution correction [71] is given by

$$\mathcal{D}_{trig} \quad = \quad 0.641 \pm 0.002\ (stat.) \pm 0.024\ (syst.)\ .$$

The OST dilution $\mathcal{D}$ is obtained, following (5.9), from the measured raw asymmetry and the above trigger-side correction.

Tagging power

We finally present the tagging performance of the opposite-side algorithms which have been described, as measured in the lepton and displaced track trigger sample.

The tagging efficiency is given by the fraction of events for which an eligible opposite-side object – soft lepton or jet – is identified. For the dilution, however, the achieved parameterizations imply that each tagged event is to be assigned a specific value, depending on the event properties accounted for through the dependencies implemented in each algorithm. The expected dilution and efficiency in each bin, $\mathcal{D}_k$ and ϵ_k, considering the discretized dependences, are combined to form the partial tagging power in the bin, $\epsilon_k \cdot \mathcal{D}_k^2$. The overall tagging power and efficiency are given by the sum over all bins,

$$\epsilon\mathcal{D}^2 = \sum_k \epsilon_k \cdot \mathcal{D}_k^2 ,$$
$$\epsilon = \sum_k \epsilon_k .$$

An *effective dilution* $\mathcal{D}_{\text{eff}}$ may then be defined as

$$\mathcal{D}_{\text{eff}} = \sqrt{\frac{\epsilon\mathcal{D}^2}{\epsilon}} .$$

The efficiency, effective dilution and tagging power for the soft lepton and jet charge taggers are summarized in Tables 5.1 and 5.2. The tagging performance is evaluated for the various algorithms separately. When applied to a given sample a fraction of events will be multiply tagged. For this reason, it also follows that the combined tagging power does not correspond simply to the sum of the partial $\epsilon\mathcal{D}^2$ contributions from the different algorithms.

soft lepton	ϵ [%]	$\mathcal{D}_{\text{eff}}$ [%]	$\epsilon\mathcal{D}^2$ [%]
muon, SMT	6.08 ± 0.04	33.9 ± 1.0	0.698 ± 0.042
electron, SET	14.22 ± 0.06	16.0 ± 0.7	0.366 ± 0.031

Table 5.1: Soft lepton tagging performance, for muon and electron algorithms.

jet type	ϵ [%]	$\mathcal{D}_{\text{eff}}$ [%]	$\epsilon\mathcal{D}^2$ [%]
class 1, JVX	10.53 ± 0.03	18.9 ± 0.3	0.376 ± 0.011
class 2, JJP	28.52 ± 0.05	11.9 ± 0.2	0.404 ± 0.014
class 3, JPT	56.56 ± 0.07	5.0 ± 0.2	0.140 ± 0.011

Table 5.2: Jet charge tagging performance, for the three exclusive jet categories.

5.2 Same side tagging

Same side tagging exploits the flavor-charge correlation between the B meson and associated particles produced nearby in phase space [72]. For each B meson species, the B and $\bar{B}$ states are expected to be more likely accompanied by leading fragmentation tracks of distinct charges. This is illustrated in Figure 5-6.

One can think of the hadronization process, in a simplified manner, as pulling light quark pairs from the vacuum and forming hadrons from nearby quarks. Consider for instance the case of a $B^0(\bar{b}d)$ meson. In order for it to be formed, the light quark pair which is nearest in the fragmentation chain to the initial heavy quark $\bar{b}$ must

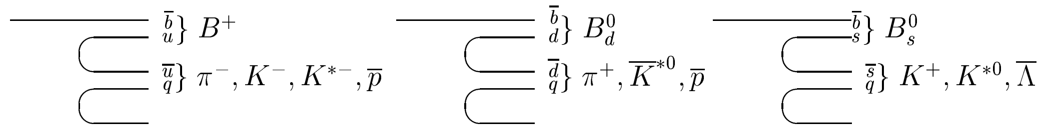


Figure 5-6: Charge correlation of B mesons with leading fragmentation tracks.

have been a $d\bar{d}$ pair. This leaves a $\bar{d}$ quark at the dangling end of the fragmentation chain. If the second nearest light quark pair is $u\bar{u}$, then the nearest meson in the fragmentation chain will be a π^+, which can be used to tag the flavor of the initial $\bar{b}$. If the second nearest light quark pair is a $d\bar{d}$ pair, then the nearest meson is a π^0, a neutral particle that therefore has no tagging power. However, the dangling end of the fragmentation chain remains a $\bar{d}$. If the third nearest light quark pair is a $u\bar{u}$ pair, then the second nearest meson will be a π^+, which can be used as a flavor tag.

Identical phenomenological descriptions as just offered for the B^0 mesons naturally hold for the other B meson species as well, the B^+ and the B_s. However the various possible hadrons which may be produced in the hadronization process associated to each B meson species lead to different overall likelihoods that a specific track charge is produced.

Additionally, also as pointed out in [72] the decay products of orbitally-excited B mesons usually referred to as B^{**} induce flavor-charge correlations which are coincident with those arising from fragmentation. Such B^{**} states correspond to P-wave levels of a b quark and a light antiquark, which may decay to $B\pi$. For instance, a B^{**+} can decay to $B^0\pi^+$ and not to $\bar{B}^0\pi^+$. These contributions are larger for B^+ than for B^0, and absent in the case of B_s mesons.

In summary, therefore, the following correlations are expected, on a statistical basis, between the production B meson flavor and a leading charged track:

- $B^+, \bar{B}^0, \bar{B}_s$ are more likely accompanied by a *negatively* charged track,
- B^-, B^0, B_s are more likely accompanied by a *positively* charged track.

The task of a SST algorithm is to identify such leading track in the event.

The same high statistics inclusive sample of a lepton and displaced track which was used to establish the OST methods in Section 5.1.3 cannot be used in the case of the SST method as its performance is expected not to be common for distinct B meson species. Instead, samples of fully reconstructed B decays will be employed to establish the algorithm performance from observed flavor asymmetries. These will be appropriate for the B^+ and B^0 mesons. The performance in the B_s case cannot however be extracted directly from corresponding data samples, and its evaluation is not addressed in the current chapter, being postponed until Chapter 8.

5.2.1 Algorithm

Candidate tracks are selected within an isolation cone of $\Delta R = 0.7$ relative to the B candidate. Tracks outside this region are used by the OST methods. The tags

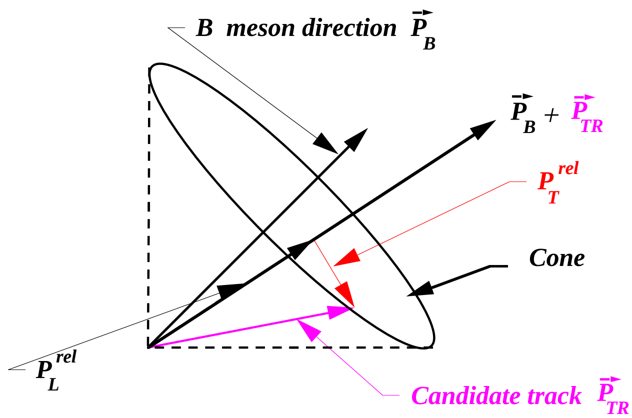


Figure 5-7: Illustration of the construction of the quantity p_T^{rel}.

are expected to originate from the primary vertex, and the track candidates should thus be consistent with it: the z_0 is required to be within 2 cm, and the impact parameter d_0 is required to be within 10 standard deviations from zero. A minimum transverse momentum p_T of 450 MeV is imposed to avoid significant effects from charge asymmetries inherent to the detector performance for low momentum tracks.

Once the above selection criteria are imposed, there may be none, single or multiple tracks which are selected. The fraction of events for which the former occurs determines the efficiency. In case a single track is selected it becomes the tagging track. Otherwise, if there are multiple tag candidates, distinct track properties may be explored (as studied in Chapter 8) to select the tagging track. In the present algorithm implementation, the chosen track is that with the smallest momentum transverse to the direction given by the combined B and track momenta, p_T^{rel}, as indicated in Figure 5-7. String fragmentation models indicate that particles produced in the b quark hadronization chain have small momenta transverse to the direction of the b quark momentum.

5.2.2 Flavor asymmetry analysis

The evaluation of the SST algorithm performance is achieved in samples of fully reconstructed B decays, corresponding to an integrated luminosity of about 270 pb^{-1}. Decays of charged and neutral B mesons are used. The dilution for the B^+ mesons is obtained from the overall flavor asymmetry, while for the B^0 mesons a time dependent asymmetry analysis is implemented.

Data samples

The charged and neutral B mesons are reconstructed in the following decay modes:

- $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow \bar{D}^0\pi^+$,

	ϵ [%]	$\mathcal{D}$ [%]	$\epsilon\mathcal{D}^2$ [%]
B^+	61.0 ± 0.5	19.5 ± 1.4	2.33 ± 0.34
B^0	63.7 ± 0.9	12.8 ± 2.2	1.00 ± 0.35

Table 5.3: Same side tagging algorithm performance for B^+ and B^0 mesons.

- $B^0 \to J/\psi K^{*0}$, $B^0 \to D^{(*)-}\pi^+(\pi^+\pi^-)$.

The signal selection follows the criteria described in Chapter 3. The samples composition and corresponding mass distributions are also addressed therein.

Asymmetry

In order to measure the flavor asymmetries, each data sample is subdivided into three disjoint subsamples according to flavor tagging information. One such subsample contains the events for which no tagging track was found. The other two of these subsamples are obtained by comparing the charge of the selected tagging track with the B decay products, to determine whether or not the tagged production flavor coincides with the decay flavor. The mass distributions for each subsample are formed, and the obtained histograms are fitted to determine the corresponding number of signal events. Figure 5-8 shows the mass distributions for the flavor tagged B^+ candidates in the $J/\psi K$ subsamples.

In the case of charged B mesons, the measured number of signal events in each subsample are taken as estimates of the number of correctly tagged N_r, incorrectly tagged N_w, and non-tagged N_n events. These are in turn used to compute the efficiency (5.1) and the dilution (5.2) of the tagging method.

For samples of neutral B mesons, the measured flavor asymmetry corresponds not only to the tagging dilution but contains also the effects of flavor mixing. The strategy in this case is to measure the asymmetry as above, but in given ranges of proper decay time t. Finally the asymmetry dependency on proper time is fitted with the model

$$A(t_i) \quad = \quad \frac{\{e^{-t_i/\tau}\,\mathcal{D}\,\cos(\Delta m_d\, t_i)\} \otimes G(t_i;\sigma_t)}{e^{-t_i/\tau} \otimes G(t_i;\sigma_t)\ .} \tag{5.10}$$

where τ and Δm_d stand for the lifetime and oscillation frequency of the B^0 system, and σ_t is the proper decay time resolution; t_i denotes the mean value of the respective proper decay time bin. The measured asymmetries for each proper time bin are illustrated in Figure 5-9 for the $B^0 \to D^-\pi^+$ decay sample, along with the fitted time-dependent asymmetry (5.10) projection.

The combined results for the SST algorithm performance obtained from fits of the charged and neutral B meson samples are summarized in Table 5.3.

The value of the oscillation frequency of the B^0 system, which appears in (5.10), is simultaneously extracted. From the combined fit of the studied neutral modes we obtain

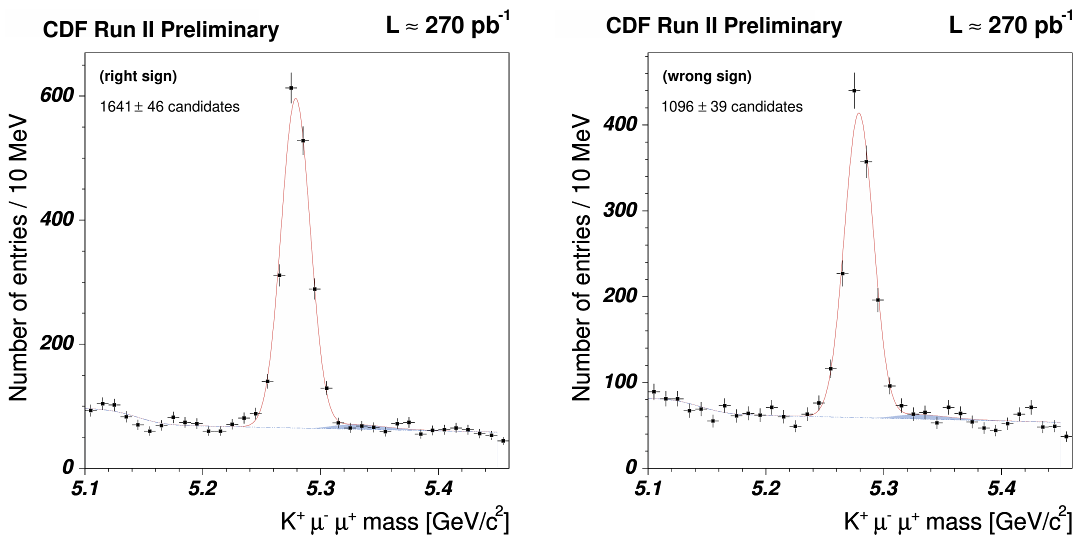


Figure 5-8: Mass distributions of $B^+ \to J/\psi K^+$ SST tagged candidates, for which the charges of the kaon and the tagging track disagree (left) and agree (right); the signal yields in the two samples are used to determine the tagging dilution.

$$\Delta m_d = 0.526 \pm 0.056 \text{ (stat.) } \pm 0.005 \text{ (syst.) ps}^{-1} .$$

The quoted systematic uncertainty was evaluated by varying the range and parameters in the individual mass fits, the effect of swapped candidates in $J/\psi K^*$, as well as the proper time resolution (varied from 0 to 90 μm), which provide the leading contributions. This constitutes the first mixing measurement performed at Run II. It is also the first time such a measurement is performed at a hadron collider using fully reconstructed decays.

5.3 Résumé

The determination of the b flavor at production time is a crucial task in the study of flavor oscillations. It is accomplished by the so-called flavor tagging methods which have been presented, and which will be employed in the following chapters.

B mesons are produced from the hadronization of b and $\bar{b}$ quarks which are originally produced in $b\bar{b}$ pairs. The determination of whether a B meson resulted from the hadronization of a b or a $\bar{b}$ quark can accordingly be achieved in two fashions. It can be inferred from information carried by tracks in the vicinity of the B candidate, or by the fragmentation or decay products of the other, accompanying b-hadron in the event. The corresponding classes of flavor tagging algorithms are called *same side* and *opposite side* methods, respectively, and denoted in short by SST and OST.

The aim of the SST is to detect the production flavor of the trigger B meson by identifying the leading accompanying fragmentation track. The objective of the

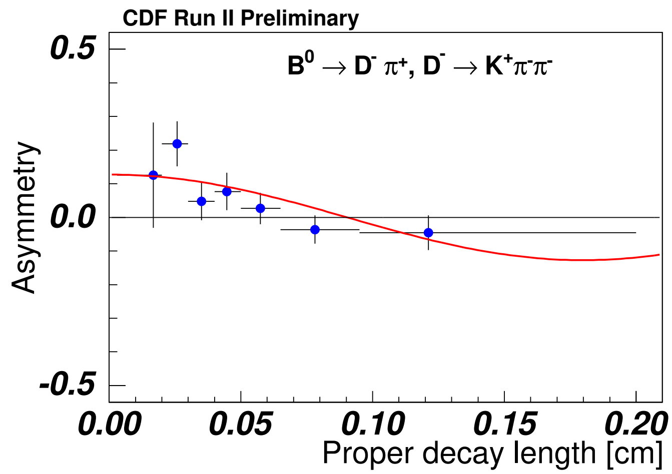


Figure 5-9: Time-dependent flavor asymmetry for $B^0 \to D^-\pi^+$ candidates tagged by the SST algorithm; the oscillation pattern is determined by the mixing frequency Δm_d and the corresponding amplitude is given by the tagging dilution.

OST is to tag the flavor of the b-hadron that did not fire the trigger to determine the flavor of the trigger B meson when it was produced. Two distinct OST strategies are explored. One, called soft lepton tagger, aims at identifying the leptons which originate from semileptonic decays of the opposite-side b-hadron. The other OST method, called jet charge tagger, explores the charges of the tracks surrounding the b-hadron direction whose combination is correlated to the charge of the b-quark itself.

The tagging power of an algorithm is given by $\epsilon\mathcal{D}^2$, where the efficiency ϵ corresponds to the fraction of events which are tagged, and the dilution $\mathcal{D}$ is related to the probability for the algorithm to provide a correct flavor determination. The OSTs are established in an inclusive sample of semileptonic decays. The high statistics of this sample further allows for the algorithms' performance to be parameterized which translates into increased power. The performance of the SST is dependent on the species of B meson candidates being tagged. The algorithm is thus preferably established in samples of fully reconstructed B meson decays. The tagging performance is summarized as:

method	tagging power, $\epsilon\mathcal{D}^2$ [%]
opposite-side (sum)	1.98 ± 0.26
same-side, B^+	2.33 ± 0.34
same-side, B^0	1.00 ± 0.35

When applied to samples of B^0 decays, a preliminary measurement of the oscillation frequency Δm_d is achieved, establishing the feasibility of mixing analyses at CDF II. The SST performance for B_s mesons cannot be extracted directly from data; this issue will be further addressed in Chapter 8.

Chapter 6

Measurement of B^0 oscillations

6.1 Fitting technique

The analysis of B flavor oscillations is performed using the unbinned maximum likelihood estimation method. The fitting framework and technique are built upon those implemented in Chapter 4 for the analysis of mass and proper decay time in the absence of flavor tagging information. The latter constitutes the ingredient which needs now to be incorporated.

The likelihood function is as in (4.4)

$$\mathcal{L} = \prod_i \sum_\alpha f_\alpha \mathcal{P}_i^\alpha \tag{6.1}$$

given by the product over the sample candidates (denoted by the index i) of the combination of the likelihood terms describing each of the sample components (denoted by the index α).

Besides the fit input quantities that have already been introduced in Section 4.1.2 – mass m, proper decay time t, and uncertainty σ_t – the novel information to be provided to the fit is the *decision* ξ and the expected *dilution* $\mathcal{D}$ of the tagging methods evaluated for the individual B meson candidates. The tagging decision constitutes the comparison of the flavor states of the B candidate at production and decay times performed by the tagging methods, while the tagging dilution is related to the correctness of the decision.

The likelihood terms describing each sample component, $\mathcal{P}^\alpha$, are given by the joint PDFs of the fit input variables,

$$\mathcal{P} \;\; = \;\; L_m \, L_t \, L_\xi \, L_\mathcal{D} \, L_{\sigma_t}. \tag{6.2}$$

The factorization given in (6.2) along with the meaning of the involved likelihood factors can be derived similarly to what was done in (4.12) using the rules of conditional probability, and relative dependences among the variables, as follows

$$\begin{aligned}\mathcal{P}(m, t, \sigma_t, \xi, \mathcal{D}) &= \mathcal{P}(m) \cdot \mathcal{P}(t, \xi|\mathcal{D}, \sigma_t) \cdot \mathcal{P}(\mathcal{D}) \cdot \mathcal{P}(\sigma_t) \\ &= \mathcal{P}(m) \cdot \mathcal{P}(t|\xi, \mathcal{D}, \sigma_t) \cdot \mathcal{P}(\xi|\mathcal{D}, \sigma_t) \cdot \mathcal{P}(\mathcal{D}) \cdot \mathcal{P}(\sigma_t) \; . \end{aligned} \quad (6.3)$$

The mass is decoupled from the functional dependence of probabilities on the other observables. The description in proper decay time space is going to be determined by the candidate lifetime along with mixing, tagging and resolution effects. In general, it will be distinct for events belonging to different tag classes, as specified by the tagging decision, and will depend on the candidate dilution in addition to its proper decay time resolution.

The likelihood factors L_{σ_t} and $L_{\mathcal{D}}$ do not involve likelihood parameters and are given instead by distributions obtained directly from the data samples. For each of the involved input quantities, namely the uncertainty σ_t and the dilutions $\mathcal{D}$ for the various taggers, two distributions are constructed: one from mass-sideband candidates and the other from the mass signal region after mass-sideband subtraction. These are respectively associated to the combinatorial background and the remaining sample components. The actual distribution used for $L_{\mathcal{D}}$ is that associated to the tagging method which provides the tagging decision ξ for the candidate at hand.

The description of the mass probability distributions L_m for the various sample components is addressed in Section 4.4. The likelihood models for proper decay time and flavor tagging contain the parameters of interest to be extracted from the fit to the data samples, and are thoroughly derived below.

6.2 Flavor tagging information

Several b-flavor tagging methods are applied to the B meson candidates. These were presented in Section 5. The class of algorithms denoted opposite-side taggers, in brief OST, is applied to all samples, and the same-side tagger, SST, method is applied solely to the fully reconstructed $B^{+,0}$ candidates.

The tagging information is contained in the following quantities: decision ξ, dilution $\mathcal{D}$, and efficiency ϵ. The tagging decision takes on the discrete values $\xi \in \{+1, -1, 0\}$. A positive (negative) decision indicates that the tagged b-flavor at the time the B meson was produced coincides with (is opposite to) the b-flavor at the time of decay. A null decision is assigned whenever the tagger fails to identify the b flavor at production time. The fraction of events for which a non-trivial decision ($\xi \neq 0$) is achieved defines the tagging efficiency, while the probability that such a decision is correct is determined by the tagging dilution,

$$\mathcal{P}(\text{correct decision}) = \frac{1 + \mathcal{D}}{2} \, .$$

The OSTs expected dilution is evaluated for each candidate, and provided as input data to the fit along with the SST and OST decisions.

The various OST methods are combined in an exclusive way. Specifically, the taggers are hierarchically ranked based on expected performance; whenever non-trivial decisions $\xi = \pm 1$ are achieved for a given candidate by a set of methods, only that provided by the foremost is used. This is further elucidated in Table 6.1.

opposite-side method	abbreviation	decision hierarchy
soft muon	SMT	evaluated first
soft electron	SET	if SMT failed
jet charge, class 1	JVX	if SMT and SET failed
jet charge, class 2	JJP	if SMT, SET and JVX failed
jet charge, class 3	JPT	if SMT, SET, JVX and JJP failed

Table 6.1: List of opposite-side taggers and decision hierarchy.

6.3 Likelihood formalism for flavor tagged samples

The likelihood descriptions of proper decay time and flavor tagging spaces are in general interconnected, and are more conveniently summarized by the joint PDF

$$L_{t,\xi} \;=\; \mathcal{P}(t, \xi | \mathcal{D}, \sigma_t)\,. \tag{6.4}$$

As it will be shown, in the case of sample components which do not undergo flavor oscillations, the proper time PDF may be defined in a decoupled fashion from the tagging decision information. It will then coincide with the t-PDFs which were derived in Chapter 4, in the absence of flavor tagging information. The probability of observing a given flavor tagging decision will in that case be determined exclusively from quantities which characterize the flavor tagging methods when applied to the data samples – the tagging efficiency and dilution.

For those components describing neutral B meson candidates a cosine term is included for characterizing the time-dependent flavor oscillations. Correspondingly, the probability for a given tagging decision to be observed will depend on that for mixing to have occurred, and will thus have a t dependence.

We proceed by first considering the simpler case of a single tagging algorithm, before extending the derived formalism to the actual situation of several taggers.

6.3.1 Non-oscillating components

We first address those sample components for which mixing effects do not occur. This is the model used for describing charged B meson signals, as well as backgrounds such as those of combinatorial type. In this case proper time and tagging probabilities may be treated in a decoupled fashion,

$$L_{t,\xi} \;=\; \mathcal{P}(\xi | \mathcal{D}) \cdot \mathcal{P}(t | \sigma_t) = L_\xi \, L_t\,. \tag{6.5}$$

The proper decay time PDF, L_t, coincides in this case with the modeling derived in Chapter 4. The tagging likelihood term L_ξ is new and needs now to be addressed.

The probability that an event, as associated to a given sample component, is tagged is expressed by the tagging efficiency ϵ for that component, which is a parameter of the fit. The non-tagging ($\xi = 0$) probability is then given by $1 - \epsilon$. This is summarized by the factor p_ϵ defined as

$$p_\epsilon(\xi) = \mathcal{P}(\xi|\epsilon) = \epsilon|\xi| + (1-\epsilon)(1-|\xi|) = \begin{cases} 1-\epsilon & \text{for } \xi = \ \ 0\,, \\ \epsilon & \text{for } \xi = \pm 1\,. \end{cases} \quad (6.6)$$

In case a non-trivial tagging decision is achieved, the probability that it is correct is determined by the tagging dilution evaluated for the candidate, and is given by $\frac{1+\mathcal{D}}{2}$. Provided no oscillations take place, as we are assuming to be the case, the indication of flavor change ($\xi = -1$) must be due to mistagging, whose probability is $\frac{1-\mathcal{D}}{2}$. This can be summarized as

$$L_\xi(\xi|\mathcal{D},\epsilon) = p_\epsilon(\xi) \cdot \frac{1+\xi\mathcal{D}}{1+|\xi|} = \begin{cases} 1-\epsilon & \text{for } \xi = \ \ 0\,, \\ \epsilon\ \frac{1\pm\mathcal{D}}{2} & \text{for } \xi = \pm 1\,; \end{cases} \quad (6.7)$$

where we note that the necessary normalization $\sum_{\xi\in\{0,+1,-1\}} L(\xi) = 1$ is verified.

6.3.2 Oscillating signal

When mixing is present, specific proper decay time distributions characterize the samples associated to the different tagging decisions. The probability for a flavor change to be ($\xi = -1$) or not be ($\xi = +1$) verified, at true proper time t, may be expressed as

$$\mathcal{P}(\xi|t,w) = \frac{1+\xi\cos wt}{2}\,, \quad (6.8)$$

where $w = \Delta m$ is the oscillation frequency of the system. The effect of mistagging is readily evaluated as

$$\mathcal{P}(\xi|t,\mathcal{D},w) = \frac{1+\xi\cos wt}{2}\frac{1+\mathcal{D}}{2} + \frac{1-\xi\cos wt}{2}\frac{1-\mathcal{D}}{2} = \frac{1+\xi\mathcal{D}\cos wt}{2}\,, \quad (6.9)$$

and additionally, using (6.6), so is that of tagging inefficiency

$$\mathcal{P}(\xi|t,\mathcal{D},\epsilon,w) = p_\epsilon(\xi) \cdot \frac{1+\xi\mathcal{D}\cos wt}{1+|\xi|} = \begin{cases} 1-\epsilon & \text{for } \xi = 0\,, \\ \epsilon\ \frac{1\pm\mathcal{D}\cos wt}{2} & \text{for } \xi = \pm 1\,. \end{cases} \quad (6.10)$$

The joint PDF for true proper decay time and tagging decision is given by $\mathcal{P}(t,\xi|\mathcal{D},\epsilon,w) = \mathcal{P}(\xi|t,\mathcal{D},\epsilon,w)\ \mathcal{P}(t)$, where the proper time probability decoupled from tagging information (that is integrated over the tagging decisions) has the lifetime exponential decay form of (4.13),

$$\begin{aligned}\mathcal{P}(t,\xi|\mathcal{D},\epsilon,w,\tau) &= p_\epsilon(\xi)\cdot\frac{1+\xi\mathcal{D}\cos wt}{1+|\xi|}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t) \\ &= p_\epsilon(\xi)\cdot\frac{E(t;\tau)+\xi\mathcal{D}\,C(t;\tau,w)}{1+|\xi|}\,, \end{aligned} \tag{6.11}$$

where we have for convenience introduced the functions E and C, whose definitions are implicitly given by the second equality (as in (4.13)).

For partially reconstructed modes, the pseudo proper decay time needs to be corrected by a kinematical factor, κ, defined in (4.8). This correction is achieved through the smearing of (6.11) with the κ-factor distribution $\mathcal{F}(\kappa)$ (Section 4.2.3). Specifically, the functions E and C become modified as follows

$$\begin{aligned}E'(t;\tau)+\xi\mathcal{D}\,C'(t;\tau,w) &= [E(t;\tau)+\xi\mathcal{D}\,C(t;\tau,w)]\otimes_\kappa\mathcal{F}(\kappa) \\ &= \int(1+\xi\mathcal{D}\cos(w\kappa t))\,\frac{\kappa}{\tau}e^{-\frac{\kappa t}{\tau}}\theta(\kappa t)\cdot\mathcal{F}(\kappa)\;d\kappa\,,\end{aligned} \tag{6.12}$$

where the κ-smearing operator $\otimes_\kappa$, originally defined in (4.21), has here been extended.

The effects of detector resolution, trigger and selection requirements are addressed as well in Section 4.2. With the description of such effects included, the joint PDF can accordingly be expressed as

$$L_{t,\xi}(t,\xi|\mathcal{D},\epsilon,w,\tau,\sigma_t) = p_\epsilon(\xi)\frac{1}{\mathcal{N}}\frac{E(t;\tau)+\xi\mathcal{D}\;C(t;\tau,w)}{1+|\xi|}\otimes_\kappa\mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\,, \tag{6.13}$$

where $G(t;\sigma_t)$ is the proper decay time resolution function (4.14), and $\mathcal{E}(t)$ is the t-efficiency function defined in Section 4.3.1. The normalization constant $\mathcal{N}$ is given in (4.20) and (4.23). The PDF normalization condition $\sum_\xi\int L(t,\xi)dt = 1$ is verified.

To conclude we mention that (6.13) may be cast in a form which explicitly verifies the factorization expressed in (6.2),

$$\begin{aligned}L_{t,\xi}(t,\xi) &= L_t(t|\xi)\cdot L_\xi(\xi)\,, \\ L_t(t|\xi) &= \frac{1}{N_\xi}\,\frac{E(t;\tau)+\xi\mathcal{D}\;C(t;\tau,w)}{1+|\xi|}\otimes_\kappa\mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\,, \\ L_\xi(\xi) &= f_\xi\,.\end{aligned} \tag{6.14}$$

The factors N_ξ are found by evaluating the normalization condition $\int L_t(t|\xi)dt = 1$ for each decision. The factors f_ξ, to be determined, must similarly ensure the corresponding normalization condition $\sum_{\xi\in\{0,-1,+1\}} L_\xi(\xi) = 1$. Furthermore, they correspond to the relative fractions of the various tagging decisions. The requirements $f_0 = 1-\epsilon$, $f_+ + f_- = \epsilon$, and $f_+/f_- = N_+/N_-$ readily imply

$$f_\xi = [\epsilon|\xi| + (1-\epsilon)(1-|\xi|)]\cdot\frac{N_\xi}{\mathcal{N}}\,, \tag{6.15}$$

and (6.13) is recovered. We in particular emphasize that the evaluation of the factors N_ξ which would involve the computation of integrals of cosine terms are not necessary. Solely the normalization factor $\mathcal{N}$which was originally addressed in the context of the lifetime analysis is required.

6.3.3 Multiply tagged candidates

We address now the case in which several decisions, provided by different tagging methods, are available for a single event. In case these decisions are correlated – as between those provided by different OSTs – one of the algorithms is selected, while for uncorrelated taggers (as between OST and SST) the information provided by the algorithms is fully combined.

For candidates concurrently tagged by several OST algorithms, correspondingly characterized by multiple decisions $\{\xi_i\}$ and dilutions $\{\mathcal{D}_i\}$, only the information provided by the selected one among these, as it was described in Section 6.2, is employed in the fit. Each algorithm is in this way effectively employed to tag the flavor of a fraction of candidates. These fractions are denoted by the parameters $\{\epsilon_i\}$. An event being not tagged would imply a trivial decision $\xi_i = 0$ among all algorithms, which occurs for a fraction $(1 - \sum_i \epsilon_i)$ of the candidates. Let the selected algorithm be identified by the index j. The factor of (6.6) becomes then

$$p_{\{\epsilon_i\}}(\xi_j) = \epsilon_j|\xi_j| + (1 - \sum_i \epsilon_i)(1 - |\xi_j|) = \begin{cases} 1 - \sum_i \epsilon_i & \text{for} \quad \forall_i\, \xi_i = \ \ 0\ , \\ \epsilon_j & \text{for} \quad \ \ \xi_j = \pm 1\ . \end{cases} \tag{6.16}$$

The PDF (6.13) becomes accordingly generalized to

$$\begin{aligned} &L_{t,\xi}(t, \xi_j|\mathcal{D}_j, \{\epsilon_i\}, w, \tau, \sigma_t) \\ &\quad = p_{\{\epsilon_i\}}(\xi_j) \cdot \frac{1}{\mathcal{N}} \frac{E(t;\tau) + \xi_j \mathcal{D}_j\, C(t;\tau,w)}{1 + |\xi_j|} \otimes_\kappa \mathcal{F}(\kappa) \otimes G(t;\sigma_t) \cdot \mathcal{E}(t)\ . \end{aligned} \tag{6.17}$$

When information provided by both the SST algorithm and a selected OST method is to be used in the fit, the combination of such information needs to be implemented in the likelihood model. The issue of combining independent taggers is explored in Appendix C. The parameters ξ', $\mathcal{D}'$, and ϵ' are used to characterize the SST, while unprimed symbols continue to be used for the OSTs. The efficiency factors are given simply by the multiplication

$$p_{\{\epsilon_i\}}(\xi_j) \cdot p_{\epsilon'}(\xi') \ = \ \begin{cases} (1 - \sum_i \epsilon_i) \cdot (1 - \epsilon') & \text{for} \quad \forall_i\, \xi_i = \ \ 0, \quad \xi' = \ \ 0\ , \\ (1 - \sum_i \epsilon_i) \cdot \epsilon' & \text{for} \quad \forall_i\, \xi_i = \ \ 0, \quad \xi' = \pm 1\ , \\ \epsilon_j \cdot (1 - \epsilon') & \text{for} \quad \ \ \xi_j = \pm 1, \quad \xi' = \ \ 0\ , \\ \epsilon_j \cdot \epsilon' & \text{for} \quad \ \ \xi_j = \pm 1, \quad \xi' = \pm 1\ . \end{cases} \tag{6.18}$$

The (t,ξ)-PDF takes in this case the following form

$$
\begin{aligned}
L_{t,\xi}(t,\xi_j,\xi'|\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon',w,\tau,\sigma_t) \;=\; & p_{\{\epsilon_i\}}(\xi_j)\cdot p_{\epsilon'}(\xi') \qquad (6.19)\\
\times\frac{1}{\mathcal{N}}\frac{(1+\xi_j\xi'\mathcal{D}_j\mathcal{D}')\,E(t;\tau)+(\xi_j\mathcal{D}_j+\xi'\mathcal{D}')\,C(t;\tau,w)}{(1+|\xi_j|)(1+|\xi'|)} & \otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\;.
\end{aligned}
$$

The joint proper decay time and tagging PDF of (6.19) is written in a most general form for the signal likelihood components. It straightforwardly reduces to the cases of lesser taggers derived before. It is suitable directly for the partially-reconstructed, t-biased samples. Its form for describing the fully reconstructed samples is recovered by imposing $\mathcal{F}(\kappa)=\delta(\kappa-1)$, and, for the t-unbiased samples, $\mathcal{E}(t)=1$. Incidentally, it may also be seen to hold for both neutral and charged B meson decay modes, with the parameter identification $w=\Delta m$ and by fixing $w=0$, respectively.

For components where mixing does not participate, or it is not to be explicitly parameterized, the descriptions of tagging and proper time is performed in a decoupled fashion, (6.5). The t-PDF coincides in those cases with that given in Chapter 4 for the sample component at hand. In adddition, so does the description of each of the independent flavor taggers. The joint (t,ξ)-PDF may be expressed as follows

$$
\begin{aligned}
L_{t,\xi}(t,\xi_j,\xi'|\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon') &= L_\xi(\xi_j|\mathcal{D}_j,\{\epsilon_i\})\cdot L_\xi(\xi'|\mathcal{D}',\epsilon')\cdot L_t(t)\\
&= p_{\{\epsilon_i\}}(\xi_j)p_{\epsilon'}(\xi')\cdot\frac{1+\xi_j\mathcal{D}_j}{1+|\xi_j|}\frac{1+\xi'\mathcal{D}'}{1+|\xi'|}\cdot L_t(t)\;. \qquad (6.20)
\end{aligned}
$$

6.4 Mixing and dilution calibration

In the fits of the B^0 and B^+ data samples described in the current chapter, we are interested in measuring the oscillation frequency Δm_d along with the taggers performance. The former appears as argument of the cosine function describing flavor oscillations as part of the signal PDF for neutral modes, while the taggers dilutions are determined, simultaneously, from both charged and neutral signal components.

The dilution of the OSTs is predicted, and assigned to each candidate as fit input, based on the parameterizations from Chapter 5. These parameterizations are achieved in high statistics samples with high B signal purity, and provide an optimal tagging power by distinguishing among candidates with poorer and higher dilutions. Accordingly, in place of extracting from the fit an average dilution for each OST, we use the referred per-event predicted dilution values, and allow instead for overall calibration factors. A single scale factor S_i is employed for each individual opposite side tagger, as a fit parameter multiplying the corresponding predicted dilution in the likelihood model.

The introduction of such dilution scale factors allows for the quantification of differences between the predicted and actual dilutions of the OST algorithms applied to our signal samples. If the dilution parameterizations are adequate and directly applicable to the signal samples to be fitted, the scale factors are then expected to be consistent with unit.

Furthermore, the behavior of the OSTs is expected to be identical when applied to samples of different B meson species. The calibration of the OSTs dilution through determination of the dilution scale factors is thus among the most interesting results to be presently obtained. These calibrations obtained based on the fully and partially reconstructed samples will be transferred directly to the kinematically and topologically similar B_s samples, where flavor oscillations will be studied in Chapter 7. Then, the dilution scale factors are provided as necessary input information to the likelihood model.

The joint proper time and tagging decision PDFs for the signal components derived in the previous section are modified slightly to include the dilution scale factor parameters $\{S_i\}$. For the partially reconstructed B^+ and B^0 signals in the DlX samples, following (6.17), it takes the form

$$\begin{aligned} L_{t,\xi}(t,\xi_j|\,S_j\,,\mathcal{D}_j,\{\epsilon_i\},\Delta m_d,\tau_u,\tau_d,\sigma_t) \;&=\; p_{\{\epsilon_i\}}(\xi_j)\,\frac{1}{1+|\xi_j|} \\ \times\,\{\; f_u\cdot\frac{1}{\mathcal{N}_u}\,[1+\xi_j\,S_j\,\mathcal{D}_j]\,E(t;\tau_u)\otimes_\kappa \mathcal{F}_u(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}_u(t) \\ +f_d\cdot\frac{1}{\mathcal{N}_d}\,[E(t;\tau_d)+\xi_j\,S_j\,\mathcal{D}_j\;C(t;\tau_d,\Delta m_d)]\otimes_\kappa \mathcal{F}_d(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}_d(t)\;\}\;, \end{aligned} \tag{6.21}$$

where the $u(d)$-indexed quantities refer to the $B^+(B^0)$ components. In this expression the expected common OST behavior for different B species is manifest: the same dilution and scale factor parameter are used in the B^+ and B^0 PDF terms.

For the fully reconstructed samples, in addition to the OST algorithms the SST method is also used. For the SST, no dilution information is provided as input to the fit. Instead, an average dilution $\mathcal{D}'$ is included in the likelihood model as a fit parameter. The joint tagging and proper time PDF for the fully reconstructed signal components is given by (6.19) with the inclusion of the OST dilution scale factors,

$$\begin{aligned} L_{t,\xi}(t,\xi_j,\xi'|S_j,\mathcal{D}_j,\{\epsilon_i\},\mathcal{D}',\epsilon',w,\tau,\sigma_t) \;&=\; p_{\{\epsilon_i\}}(\xi_j)\frac{1}{1+|\xi_j|}\cdot p_{\epsilon'}(\xi')\frac{1}{1+|\xi'|} \\ \times\frac{1}{\mathcal{N}}\,\{\,(1+\xi_j\xi' S_j\mathcal{D}_j\mathcal{D}')\,E(t;\tau)+(\xi_j S_j\mathcal{D}_j+\xi'\mathcal{D}')\,C(t;\tau,w)\,\}\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\;. \end{aligned} \tag{6.22}$$

We note that, contrarily to the OST case, the SST properties are expected to be distinct for different B species; *e.g.* the value of the dilution parameter $\mathcal{D}'$ is anticipated to substantially differ when fitting B^+ and B^0 meson samples.

Sample components describing B decays other than signal are in general assigned the same tagging parameters as signal. In particular, the Cabibbo-suppressed decays share the same tagging and proper decay time description as the signal.

For non-physics background components the predicted OST dilution is meaningless, and it is thus not used to describe the tagging or proper time PDFs. In order to account for possible background flavor asymmetries, the tagging decision(s) are used in the background likelihood modeling together with dilution-like parameters, $\mathcal{D}^B$. These will be referred to as background dilution parameters, and one such parameter is associated to each of the (OST and SST) tagging methods, for each sample. No time-dependent asymmetries are assumed, nor expected, and the model of (6.20) is employed, for both combinatorial and fakes backgrounds.

The proper decay time models L_t for the sample components other than nominal signals where mixing is present (*i.e.* B^0 modes) coincide with those found in Chapter 4. The likelihood models in mass space L_m for all the individual components of the various data samples are also those described therein. The remaining likelihood factors, L_{σ_t} and $L_{\mathcal{D}}$, do not explicitly involve fit parameters and are given by distributions obtained as described in Section 6.1. Some further considerations related to the treatment of flavor tagging information that are specific to individual samples are provided below.

$B^{+,0} \to J/\psi K$

The Cabibbo-suppressed $B^+ \to J/\psi\pi^+$ decays share the same proper time and tagging model with the nominal signal.

The K^*-swap component deserves special treatment, because the decay flavor for these candidates is in fact opposite to that indicated by the reconstruction. The B flavor at decay time is inferred from the charges of the $K^{*0} \to K^+\pi^-$ daughter particles, and the misassignement of kaon and pion hypotheses leads to an incorrect inference of the actual B meson decay flavor. For tagged events, therefore, the taggers decisions are flipped for the swap component, which otherwise shares the same proper decay time and tagging descriptions as those used for signal.

$B^{+,0} \to D^{(*)}\pi(\pi\pi)$

The proper time and tagging joint PDF for the nominal signal component is given by (6.22), which is shared by the associated Cabibbo-suppressed $B \to D^{(*)}K(\pi\pi)$ decay component.

In the sample of $B^0 \to D^-\pi^+$ candidates, the B_s and Λ_b^0 background components are described through models of the form of (6.20), where the proper decay time is modeled in identically the same fashion as was used in the absence of flavor tagging information addressed in Chapter 4. For the misreconstructed B_s decays, no global flavor asymmetry is considered, following from the rapid oscillations. For the Λ_b^0 component the tagging parameters are common to those used for the signal, except for the SST dilution which has not been measured and to which a value of 16% is assigned. A systematic variations of this value will be considered as part of the uncertainty.

For the $D^*\pi(\pi\pi)$ samples, the dilution template distributions $L_{\mathcal{D}}$ are derived commonly for all four modes as these are affected by limited statistics.

$B^{+,0} \to DlX$

The signal components of the semileptonic samples are described by the model derived in (6.21). The remaining components are modeled in the proper decay time space as in Chapter 4, while the tagging parameters are contained in the likelihood factor L_ξ given by the general prototype of (6.7).

For the physics backgrounds, the OST dilution is given for each candidate by its predicted value, while the tagging parameters $\{S_i\}$ and $\{\epsilon_i\}$ coincide with those used for describing the signal components.

The description of the combinatorial and fakes background components does not make use of the predicted OST dilution input values, and are each described by an independent set of flavor asymmetry $\{\mathcal{D}_i\}$ and efficiency $\{\epsilon_i\}$ parameters.

6.5 Fitting procedure and results

The model parameters describing the mass and proper time space for our samples have been obtained in the fits to the data performed in Chapter 4 in the absence of flavor tagging information. At present, the aim is to determine those parameters introduced in the likelihood models that are associated to tagging and mixing. These correspond to:

1. the taggers efficiencies, $\{\epsilon_i\}$ and ϵ', for backgrounds and signal components,
2. the flavor asymmetries of non-physics backgrounds for each tagger,
3. the OST dilution scale factors $\{S_i\}$ and (for the fully reconstructed modes) the SST dilutions, for signal and physics backgrounds components,
4. the mixing frequency Δm_d, for the B^0 modes.

A characterization of the non-physics background components is achieved by fitting the individual data samples separately. The taggers efficiencies and flavor asymmetries for the combinatorial background are determined from fits to candidates in the mass-sideband region. The efficiencies for the fakes background component in the semileptonic modes are constrained to those of the signal, and its flavor asymmetry parameters are determined from the full fit to the individual samples.

Finally, combined fits are performed for the fully and the partially reconstructed B samples separately to determine the tagging parameters which describe both signal and physics backgrounds, as well as the B^0 mixing frequency. In such combined fits the likelihood maximization is performed on the product of the likelihoods of the individual samples according to (4.2), which are functions of those common parameters.

In the fit to the combination of partially reconstructed B samples, a simultaneous determination of the taggers signal efficiencies ϵ_i, the OSTs dilution scale factors S_i, and the mixing parameter Δm_d are achieved. In the fit of the fully reconstructed modes the SST dilutions for B^+ and B^0 are in addition determined.

The fitted values of these parameters are presented in Table 6.2. For the main parameters of interest – OST dilution scale factors, SST dilutions, and oscillation frequency – the systematic uncertainties are also shown, which are evaluated in Section 6.6.

Fit projections are shown for the high statistics, semileptonic samples in Figure 6-1. These are obtained for each individual OST method, as the tagged flavor, time-dependent asymmetry for signal. Denoting by $N_\xi(t)$ the pseudo proper decay time distribution of events for which the tagging decision is ξ, the shown asymmetry is constructed as

$$\frac{N_+(t) - N_-(t)}{N_+(t) + N_-(t)} \, . \tag{6.23}$$

For the likelihood projections themselves, $N_\xi(t)$ are based on (6.21). The figures also show these asymmetries obtained for the B^+ and B^0 terms separately, characterized respectively by a flat and an oscillating pattern. For the points representing the data in those figures, the proper time distributions $N_\xi(t)$ are obtained from the sub-samples of candidates characterized by the different tagging decisions, after subtracting background contributions using the associated fit models, and evaluating the corresponding asymmetries in the introduced t-bins.

It is also interesting to evaluate the overall tagging effectiveness, $\epsilon\mathcal{D}^2$. For the SST this quantity is readily computed from the combined fit results, while for the OSTs it is estimated by

$$\epsilon S^2 \left\langle \mathcal{D}^2 \right\rangle$$

where the average over per-event predicted dilutions is evaluated, and further mass-sideband subtracted. The results are found in Table 6.3.

parameter	hadronic fit	semileptonic fit
ϵ_{SMT}	0.0458 ± 0.0013	0.04808 ± 0.00044
ϵ_{SET}	0.0261 ± 0.0009	0.02959 ± 0.00036
ϵ_{JVX}	0.0821 ± 0.0016	0.08208 ± 0.00057
ϵ_{JJP}	0.2837 ± 0.0027	0.27578 ± 0.00093
ϵ_{JPT}	0.5146 ± 0.0030	0.51626 ± 0.00104
ϵ'_{SST}	0.6339 ± 0.0028	–
S_{SMT}	$0.95 \pm 0.08 \pm 0.02$	$0.936 \pm 0.038 \pm 0.013$
S_{SET}	$1.06 \pm 0.11 \pm 0.05$	$1.072 \pm 0.045 \pm 0.017$
S_{JVX}	$0.88 \pm 0.13 \pm 0.03$	$0.917 \pm 0.058 \pm 0.013$
S_{JJP}	$0.98 \pm 0.12 \pm 0.03$	$1.001 \pm 0.060 \pm 0.028$
S_{JPT}	$1.20 \pm 0.22 \pm 0.09$	$0.838 \pm 0.106 \pm 0.043$
$\mathcal{D}'_{SST}(B^+)$	$0.209 \pm 0.011 \pm 0.001$	–
$\mathcal{D}'_{SST}(B^0)$	$0.130 \pm 0.018 \pm 0.003$	–
Δm_d [ps^{-1}]	$0.536 \pm 0.028 \pm 0.006$	$0.511 \pm 0.020 \pm 0.014$

Table 6.2: Tagging and mixing results from the combined fits in the fully and the partially reconstructed B^+ and B^0 samples; the first uncertainty is statistical and the second systematic.

tagger	$\epsilon\mathcal{D}^2$ [%]	
	hadronic	semileptonic
SMT	$0.559 \pm 0.094 \pm 0.027$	$0.551 \pm 0.048 \pm 0.020$
SET	$0.264 \pm 0.054 \pm 0.022$	$0.308 \pm 0.030 \pm 0.008$
JVX	$0.230 \pm 0.068 \pm 0.017$	$0.247 \pm 0.033 \pm 0.010$
JJP	$0.347 \pm 0.084 \pm 0.020$	$0.366 \pm 0.045 \pm 0.021$
JPT	$0.152 \pm 0.055 \pm 0.024$	$0.076 \pm 0.019 \pm 0.009$
total OST	$1.553 \pm 0.163 \pm 0.050$	$1.550 \pm 0.083 \pm 0.029$
SST, B^0	$1.074 \pm 0.302 \pm 0.050$	—
SST, B^+	$2.773 \pm 0.296 \pm 0.027$	—

Table 6.3: Tagging effectiveness $\epsilon\mathcal{D}^2$ for each tagging method.

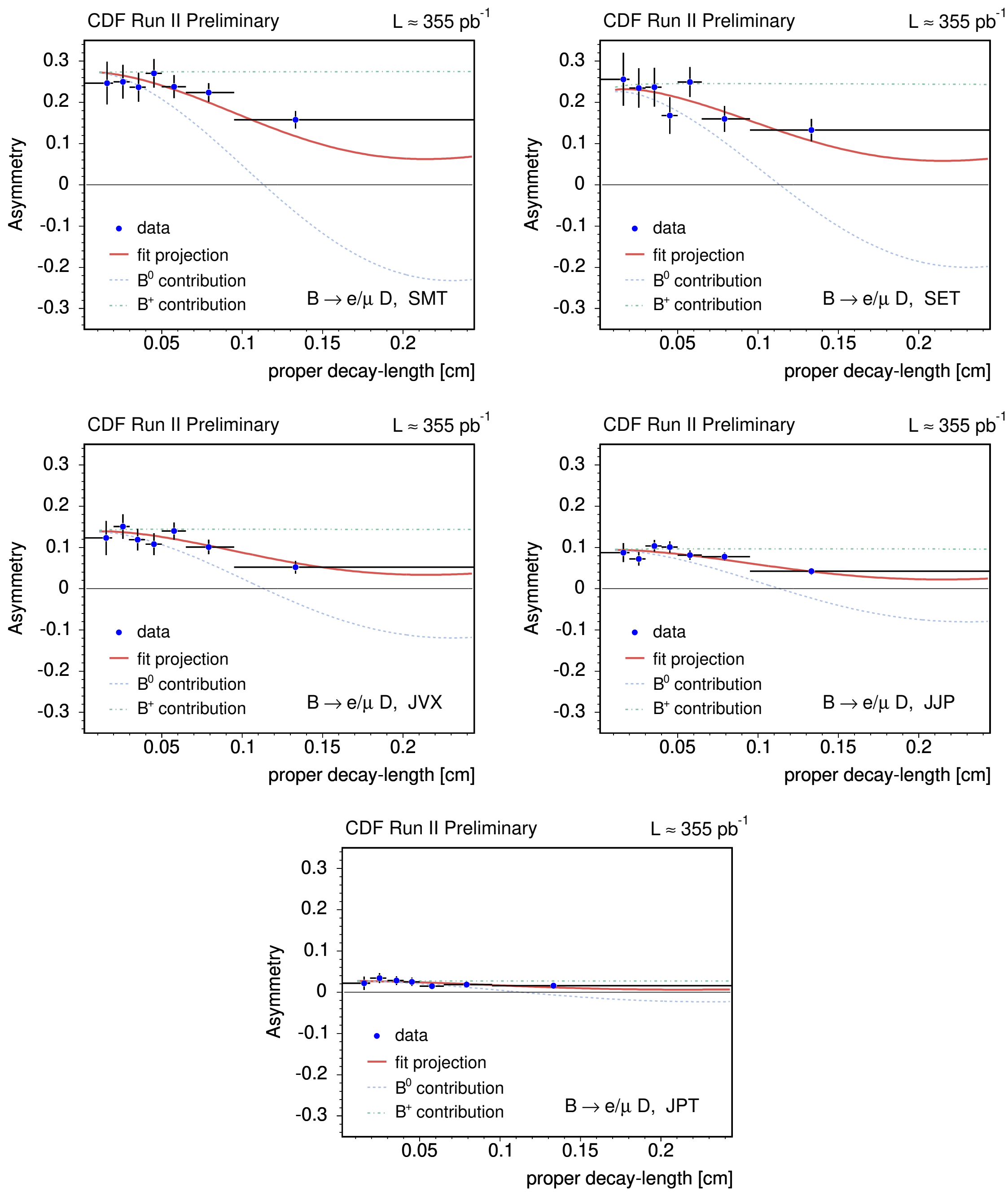


Figure 6-1: Asymmetry projections of the combined semileptonic fit, for the individual OST methods.

6.6 Systematic uncertainties

Systematic uncertainties are evaluated for the oscillation frequency Δm_d, as well as for the OST dilutions scale factors and the SST dilutions, by repeating the fits under modified conditions dictated by the considered sources enumerated below.

Combinatorial background $\boldsymbol{m}$ and $\boldsymbol{t}$ template parameters: The description of several background components is pre-determined and fixed in the final combined fits. Modifications of these parameters are propagated into systematic uncertainties for our primary parameters of interest. For the $D\pi$ modes, mass parameters for the combinatorial background component are pre-determined from a wide-range mass fit. They have uncertainties which are translated into systematic uncertainties by varying their values within $\pm 1\sigma$ and re-running the combined fit. For the semileptonic modes, such parameter variations are instead achieved by performing the fits in mass-sideband regions extended by ± 50 MeV/c^2. For the fully reconstructed samples, additionally, those parameters which are fixed to the values found in the fits of the individual modes are smeared according to the corresponding uncertainties. Specifically, such uncertainties are used as the widths of Gaussian-distributed corrections, which are added for each such parameter to its nominally fixed value. The combined fit is repeated many (about 150) times, and the resulting distributions of the parameters of interest are crudely Gaussian, the widths of these distributions being taken as estimates of systematic uncertainty.

Physics background levels: Various background component contributions are evaluated from Monte Carlo simulation and further specific branching ratios information. The fractions of Cabibbo-suppressed components in $D\pi\pi\pi$ samples are reset to 0-12% of the associated signal component contribution. For the $B^0 \to J/\psi K^{*0}$ mode, the fraction of events with swapped $K\pi$ mass assignment in K^{*0} candidates relative to that of correctly reconstructed signal candidates is re-estimated from an independent Monte Carlo sample. The modified relative fraction value thus obtained is 19%. In the $B^0 \to D^-\pi^+$ sample, the amounts of Λ_b^0 and B_s background decays are varied by 30%. For the semileptonic samples various branching ratios are involved in the determination of the physics background fraction which are poorly measured. Fractional variations of $\pm 25\%$ relative uncertainty are considered.

Semileptonic signal composition: Each of the three reconstructed Dl signal final states involve multiple decay chains as discussed in Section 3.3. The sample composition parameters for the B^+ and the B^0 signals are varied within the associated uncertainties. For the parameter P_V, for which no direct measurement yet exists, we use 0.627$\pm$0.26, otherwise variations of one standard deviation are considered. These variations imply not only modifications of the B^+ and B^0 relative fractions, but also of the κ-factor distributions and t-efficiency functions, which are considered in the estimation of the systematic uncertainty.

Fakes background in semileptonic samples: Two systematic sources are addressed regarding the description of the fakes background: its fraction and the shape of its proper time template. The latter is obtained from a fit to the fakes lepton sample, while the former is determined from a fit to the m_{Dl} distribution. Variations of both the fraction and the shape of this background component within the uncertainties obtained in these fits are propagated into contributions to the systematic uncertainty.

Proper time uncertainty calibration, S_t: The raw values of the proper decay time uncertainty σ_t are in general underestimated, and a scaling correction needs to be applied (4.10). Whereas the scale factors S_t for the $J/\psi K$ samples are determined as an integral part of the fit (4.33), for the $D\pi(\pi\pi)$ and DlX samples such direct determination is not attainable, because there is no prompt peak in proper time in these samples. For the t-biased hadronic and semileptonic samples we thus apply a fixed, common scale factor of 1.40 to the t resolution. This value is motivated by the $J/\psi K$ samples scale factor results of ~ 1.35, and by those obtained in Section 4.7. The fits are repeated by shifting this value, and the variations considered are $[1.0, 1.8]$ and $[1.3, 1.5]$ for the hadronic and semileptonic samples, respectively. The largest variations observed in the fit parameters (relative to the results with nominal fit conditions) are taken as the associated systematic uncertainty.

Construction of t-efficiency curves, $\mathcal{E}(t)$: The trigger selection and reconstruction requirements induce biasing effects in the proper decay time distribution which are accounted for by the t-efficiency function. This is constructed for each nominal signal mode from Monte Carlo samples which include trigger simulation and to which the sample selection is applied. These samples are modified at a time by different criteria. The B mesons lifetime input values are shifted by ± 1 standard deviation [1]. An observed difference between SVX and SVT d_0 residuals in Monte Carlo and data motivates the introduction of an extra smearing, of $\sim 12\mu$m, of the impact parameter d_0 before SVT-trigger confirmation. Tuning of resolution and efficiency for hits in the innermost silicon layer is applied to the simulation, which was demonstrated to further improve the agreement with data of L_{xy} and $\sigma_{L_{xy}}$ distributions. Corresponding t-efficiency curves are derived at a time and the combined fits repeated.

Discretized κ-factor distributions, $\mathcal{F}(\kappa)$: The proper time model for signal components in semileptonic samples involves the smearing with Monte Carlo-based κ-factor distributions, which are used to describe the effects of partial B reconstruction. About 500 variations of those histograms are produced by randomly modifying the entry of each bin according to its statistical uncertainty. The fits are repeated using such modified κ-factor distributions, and the associated systematic uncertainty contribution is taken as the width of the distribution of fit values for each parameter.

Dilution templates, $L_{\mathcal{D}}$: The likelihood factors $L_{\mathcal{D}}$, which estimate for each OST algorithm the overall probability for a signal or background event to have a given predicted dilution, are realized by mass-sideband (subtracted) distributions. The bins of these template histograms are not always highly populated due to relatively low sample sizes and/or low tagging efficiencies. Sets of about 300 modified distributions each are generated, for every original template, by Gaussian-smearing the contents of each bin according to its statistical uncertainty. The associated systematic uncertainty is taken as the Gaussian-width of the distribution of fitted parameter results.

Additional effects evaluated for exclusive samples: The following sources are further evaluated for the exclusively reconstructed samples. In the high statistics exclusive $D\pi$ samples, the signal component is described in mass space by a double

Gaussian, as it is observed to provide a better fit probability than the simpler single Gaussian model. As a systematic check, an alternative mass model involving a Breit-Wigner and Gaussian shape is used. For the Λ_b^0 component, appearing as part of the $B^0 \to D^-\pi^+$ sample, the SST dilution is fixed to 16% in the nominal fit. Although this quantity has not yet been measured, we have surmised that it is similar to the signal dilutions of B^+ and B^0. An associated systematic uncertainty is thus evaluated by repeating the primary fit with the Λ_b^0 SST dilution set to 12% and 20%. To cover the possibility that flavor tagging is slightly different among the modes, we compare the nominal fit with an alternative where all signal tagging efficiencies (SST, OSTs) and SST dilutions are fixed to the values found in individual fits of the modes. The background tagging efficiency parameters are fixed in the combined fit to the values found in individual fits for each of the analyzed decay modes. The statistical parameter uncertainties observed in these fits are used to generate sets of corrections to the nominal parameter values. The combined fit is repeated about 250 times, once for each such set, and the Gaussian-width of the distributions of fit results of the parameters of interest are taken as corresponding systematic uncertainty. In order to take into account correlation effects arising from the usage of the SST in the determination of the OST scale factors, the combined fit is repeated excluding SST information. We examine the shift in dilution scale factors between the two fits. Clearly, these variations are undefined for the SST dilutions, since they are not included in the varied fit. These are not applicable to Δm_d either, because its value is expected to change with the removal of SST information. However, for the dilution scale factor parameters, we opt to include the mentioned shifts as rather conservative estimates of possible correlation effects.

For some of the contributing systematic uncertainty sources above described, the estimation method adopted is based on an inherent statistical procedure. This may deliberately result on a potential inflation of those particular systematic contributions. Regardless, we emphasize that the final measurement of the main parameters, summarized in Table 6.2, is dominated by statistical uncertainty.

The contributions to the systematic uncertainties are compiled in Tables 6.4 and 6.5, for the fully and the partially reconstructed combined modes. Among the prominent sources are the dilution templates, particularly for lower efficiency taggers and smaller size samples. This provides in effect the dominant contributions for the fully reconstructed samples. We mention in passing that the size of these contributions are, however, expected to decrease with the increasing data samples. For the partially reconstructed modes, the dominant source of systematic uncertainties for the dilution scale factors is the fraction of the fakes background, especially for the high statistics, low purity taggers. For the Δm_d parameter, the predominant systematic uncertainty comes from the Dl signal composition. Although the present Δm_d measurement is dominated by statistical uncertainty, sample composition effects are ineluctably expected to limit the precision of future measurements in semileptonic decays.

source	relative uncertainty [%]							
	S_{SMT}	S_{SET}	S_{JVX}	S_{JJP}	S_{JPT}	$\mathcal{D}_{SST}^{B^+}$	$\mathcal{D}_{SST}^{B^0}$	Δm_d
combinatorial bckg.	0.3	0.6	0.7	0.7	0.7	0.3	1.6	0.6
physics bckg. level								
K^{*0} swap in $J/\psi K^{*0}$	< 0.1	< 0.1	0.3	0.4	0.2	< 0.1	0.1	< 0.1
Λ_b^0 and B_s in $D^-\pi^+$	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.4	0.1
Cabibbo-suppressed	0.4	0.3	0.8	0.1	0.6	< 0.1	0.6	0.2
signal shape for $D\pi$	< 0.1	0.4	0.5	< 0.1	< 0.1	< 0.1	0.2	0.2
scale factor on σ_t	< 0.1	< 0.1	0.2	< 0.1	< 0.1	< 0.1	0.2	0.2
t-efficiency function	< 0.1	< 0.1	0.6	0.3	1.0	< 0.1	0.4	0.1
Dilution templates	2.2	3.9	3.3	1.0	2.5	0.5	1.5	0.7
Tagging efficiencies	0.5	0.4	0.3	< 0.1	0.4	0.1	0.6	0.5
Λ_b^0 SST dilution in $D\pi$	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	—	< 0.1	< 0.1
SST removal	0.6	1.4	1.1	2.6	7.2	—	—	—
total	2.4	4.2	3.7	2.9	7.8	0.6	2.4	1.1

Table 6.4: Summary of systematic uncertainties in fully reconstructed modes.

source	relative uncertainty [%]					
	S_{SMT}	S_{SET}	S_{JVX}	S_{JJP}	S_{JPT}	Δm_d
combinatorial background	0.1	0.1	0.1	0.2	0.3	0.2
physics background	0.1	0.1	0.1	< 0.1	0.1	0.8
fakes background fraction	0.1	0.6	1.2	2.7	5.0	0.1
fakes background shape	0.1	0.1	0.3	0.3	0.4	0.1
signal composition	0.3	0.2	0.3	0.5	0.1	2.6
scale factor on σ_t	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1
t-efficiency function	0.2	0.1	0.2	0.2	0.4	0.2
κ-factor binning	0.2	0.1	0.1	0.1	0.3	0.1
dilution templates	1.3	1.4	0.7	0.5	0.4	0.5
total	1.4	1.6	1.4	2.8	5.1	2.8

Table 6.5: Summary of systematic uncertainties in partially reconstructed modes.

6.7 Résumé

The likelihood description of the data samples has been extended to include b-flavor information, and has been applied to the samples of B^0 and B^+ meson candidates in flavor oscillation and tagging measurements.

The opposite-side and same-side tagging methods – abbreviated OST and SST – are described in Chapter 5. We apply these techniques to compare the flavor, *i.e.* B or $\bar{B}$ state, of our $B^{+,0}$ candidates at production and decay times. The OST is used for tagging the semileptonic samples, while both OST and SST are applied to the fully reconstructed modes. The OST algorithms are combined in an exclusive fashion, such that for a given candidate only the decision provided by one selected algorithm is used. The combination of OST and SST information is derived, and also implemented in the fitting model.

The samples description in mass and proper decay time spaces achieved in Chapter 4 is augmented to incorporate the candidates flavor tagging information. For neutral B candidates, the corresponding likelihood signal component contains an additional time-dependent oscillating term which describes mixing. The B^0 oscillation frequency is determined directly as a fit parameter. The combined fit result is

$$\Delta m_d = 0.522 \pm 0.017 \text{ ps}^{-1} .$$

The measurement is in good agreement with the world average value, and the uncertainty is still dominated by the statistical contributions.

Predicted OST dilution values are assigned to the individual candidates based on properties of the event, and an overall dilution calibration factor is introduced as a fit parameter for each algorithm. For the SST, the average dilutions are measured directly as parameters of the fit. The average effectiveness of the two classes of tagging methods in the combined samples is:

method	tagging power, $\epsilon\mathcal{D}^2$ [%]
opposite-side	1.55 ± 0.08
same-side, B^+	2.77 ± 0.30
same-side, B^0	1.07 ± 0.30

The distinct behavior of the SST for different B meson species is emphasized. The combined OST performance is that obtained with the exclusive algorithm combination employed.

The treatment of flavor tagging and mixing here described in the context of the $B^{+,0}$ samples further serves as the basis for the study of flavor oscillations in the B_s system to be carried out in the following chapter. The achieved taggers' calibration further provides a crucial ingredient for inferring exclusion conditions for the B_s oscillation frequency.

Chapter 7

Search for B_s oscillations

7.1 Fitting technique

The unbinned maximum likelihood fitting framework is developed and described in previous chapters. The likelihood model presented in the context of the measurement of B^0 flavor oscillations in Chapter 6 serves as the basis also for the study of mixing in the B_s system, in the partially $D_s l$ and the fully $D_s\pi(\pi\pi)$ reconstructed modes.

The fit input quantities are, for the individual B_s candidates, the following: mass m, proper decay time t, proper decay time uncertainty σ_t, tagging decision ξ, and tagging dilution $\mathcal{D}$. Each sample component α is modeled in the spaces of these input variables through corresponding likelihood factors $\mathcal{P}$, which are evaluated for each candidate i and combined to form the likelihood function $\mathcal{L}$ which is to be maximized in the fitting process,

$$\mathcal{L} = \prod_i \sum_\alpha f_\alpha \mathcal{P}_i^\alpha \qquad \text{with} \qquad \mathcal{P} = L_m\, L_t\, L_\xi\, L_{\mathcal{D}}\, L_{\sigma_t}.$$

The likelihood factors L_{σ_t} and $L_{\mathcal{D}}$ do not themselves contain fit parameters and are realized by distributions obtained from the data sample being fit (Section 6.1). The mass PDF L_m was described in Section 4.4 for the various B_s sample components. The likelihood models for proper decay time L_t and flavor tagging L_ξ follow a general description consonant with that given in Chapter 6 for the $B^{+,0}$ samples. However, a few outstanding differences arise, which are addressed in the following sections. The introduction of the amplitude, which is the primary fitting parameter of interest, is the most prominent difference. It appears as part of the description of the signal components, and arises in the implementation of the fitting method used for the study of the rapid time-dependent flavor oscillations which characterize the B_s system.

7.2 Flavor tagging and input calibration

The leading factors determining the significance of a B_s oscillating signal, besides the data samples size and purity, are the flavor tagging performance and the proper decay time resolution. The calibration of these fit input quantities, accomplished in previous chapters, is thus crucial.

The b-flavor information is provided by the opposite-side tagging methods. An exclusive combination of those methods is used, as explained in Section 6.2, such that for a B_s candidate which is non-trivially tagged by several algorithms the elected tagging decision is that provided by the algorithm with the highest average dilution. The tagging algorithm raw dilutions are evaluated for each candidate according to the dependencies on event properties presented in Chapter 5. Further calibration of such dilution parameterizations is achieved in the mixing and tagging studies undertaken in Chapter 6. We implement therefore the re-scaling of the dilutions for each of the tagging algorithms,

$$\mathcal{D}_j \quad \mapsto \quad S_j \cdot \mathcal{D}_j \ , \tag{7.1}$$

according to the overall dilution scale factors $\{S_i\}$ shown in Table 6.2. Specifically, the factors applied to the $D_s\pi(\pi\pi)$ and $D_s l$ modes are those found for the kinematically similar hadronic and semileptonic $B^{+,0}$ samples, respectively.

The raw proper decay time uncertainty σ_t returned by the vertex fitter is in general underestimated as described in Chapter 4. A detailed calibration procedure is implemented in Section 4.7. The re-scaling of this fit input quantity is thus achieved,

$$\sigma_t \quad \mapsto \quad S_t \cdot \sigma_t \ , \tag{7.2}$$

on a per-event basis, through the scale factor parameterizations (4.36) explored in Section 4.7 which take into account dependencies on topological and kinematical vertex characteristics.

7.3 Amplitude method

The analysis method which we employ for the study of flavor oscillations in the B_s system is distinct from that used in the B^0 system in that the oscillation frequency Δm_s is not directly determined as a parameter of the fit.

Such an alternative method is motivated by the B_s mixing frequency being expected to be at least ~ 30 times larger that Δm_d, and that therefore the corresponding time-dependent oscillation patterns may not be adequately discernible with our present sample size and effective resolution for a direct observation to be feasible.

The amplitude method [31] used for probing flavor oscillations in the B_s system is based on the introduction of a Fourier-like coefficient, the amplitude $\mathcal{A}$, multiplying the cosine modulation term in the signal model,

$$1 \pm \cos(wt) \quad \mapsto \quad 1 \pm \mathcal{A} \cdot \cos(wt) \ . \tag{7.3}$$

By fixing the oscillation frequency to a given test value, the fit result for the parameter $\mathcal{A}$ is expected to be unit in case the probed frequency coincides with the true oscillation frequency of the system, $w = \Delta m_s$, and be zero otherwise. The execution of this method involves performing a scan in w and a measurement of the amplitude at each value. The output of the procedure is accordingly a list of fitted values $\{\mathcal{A}, \sigma_{\mathcal{A}}\}$ obtained for each probed frequency.

An analysis is said to have sensitivity in a given frequency range if the expected uncertainty on the measured amplitudes is small enough compared to unity, so that the two values $\mathcal{A} = 1$ and $\mathcal{A} = 0$ may be distinguished. The sensitivity of the analysis is here defined as the value of the frequency w for which a measured null amplitude value $\mathcal{A} = 0$ would imply the exclusion of $\mathcal{A} = 1$ at the desired confidence level. The degree of exclusion of a given frequency in the scan, for which the measured amplitude and associated uncertainty are $\mathcal{A}$ and $\sigma_{\mathcal{A}}$, is given by

$$\frac{1}{\sqrt{2\pi}\,\sigma_{\mathcal{A}}} \int_{-\infty}^{1} e^{-\frac{(x-\mathcal{A})^2}{2\,\sigma_{\mathcal{A}}^2}}\, dx \,. \tag{7.4}$$

Specifically, for a confidence level of 95%, which is nominally used, the exclusion and sensitivity conditions are expressed accordingly as follows:

$$\begin{aligned} \mathcal{A} + 1.645 \cdot \sigma_{\mathcal{A}} &< 1 \qquad && 95\%\,\text{C.L. exclusion condition}\,, \\ 1.645 \cdot \sigma_{\mathcal{A}} &= 1 \qquad && 95\%\,\text{C.L. sensitivity condition}\,. \end{aligned}$$

The exclusion *limit* is defined as the largest frequency value below which all frequencies are excluded.

A notable advantage of the indirect probe for oscillations offered by the method stems from the fact that the dependence on $\mathcal{A}$ is linear (7.3). The measurement of $\mathcal{A}$ is hence Gaussian (7.4), and the issue of merging different experimental measurements is straightforward. In effect, the amplitude results obtained at a given frequency point by two experiments (labeled 1 and 2) may be combined [31],

$$\mathcal{A} = \left(\frac{\mathcal{A}_1}{\sigma_{\mathcal{A},1}^2} + \frac{\mathcal{A}_2}{\sigma_{\mathcal{A},2}^2} \right) \cdot \sigma_{\mathcal{A}}^2 \qquad \text{with} \qquad \frac{1}{\sigma_{\mathcal{A}}^2} = \frac{1}{\sigma_{\mathcal{A},1}^2} + \frac{1}{\sigma_{\mathcal{A},2}^2}\,, \tag{7.5}$$

as independent measurements of a same physics quantity.

7.4 Likelihood description

We review in this section the likelihood description for the various sample components. The characterization of the samples accomplished in Chapter 4 is independent of flavor information. The treatment of the latter is achieved in Chapter 6. In the following we thus concentrate mostly on tagging and mixing related aspects which are specific to the B_s samples.

The mixing PDF for signal components, taking into consideration the description of tagging, proper time resolution, t-biasing effects, and partial reconstruction, has been previously derived (6.17). With the introduction of the amplitude parameter (7.3), the PDF takes the following form,

$$
\begin{aligned}
&L_{t,\xi}(t,\xi_j|\mathcal{A};\mathcal{D}_j,\{\epsilon_i\},w,\tau,\sigma_t) \qquad (7.6)\\
&= \begin{cases}
\left(1-\sum_i \epsilon_i\right) \frac{1}{\mathcal{N}}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t)\otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t) & \text{for} \quad \forall_i\,\xi_i = \ \ 0\ , \\
\epsilon_j \frac{1-\mathcal{D}_j\mathcal{A}\cos(wt)}{2} \frac{1}{\mathcal{N}}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t)\otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t) & \text{for} \quad \xi_j = -1\ , \\
\epsilon_j \frac{1+\mathcal{D}_j\mathcal{A}\cos(wt)}{2} \frac{1}{\mathcal{N}}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t)\otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t) & \text{for} \quad \xi_j = +1\ , \\
\end{cases}\\
&= \ p_{\{\epsilon_i\}}(\xi_j)\cdot\frac{1}{\mathcal{N}}\,\frac{E(t;\tau)+\xi_j\mathcal{D}_j\ \mathcal{A}\ C(t;\tau,w)}{1+|\xi_j|}\otimes_\kappa \mathcal{F}(\kappa)\otimes G(t;\sigma_t)\cdot\mathcal{E}(t)\ .
\end{aligned}
$$

The parameter w takes on the value of the oscillation frequency being probed. The j-index is used to indicate the algorithm employed for tagging the event, the efficiency factor $p_{\{\epsilon_i\}}$ being given by (6.16). The proper decay time resolution function $G(t;\sigma_t)$ is defined in (4.14), and the t-efficiency function $\mathcal{E}(t)$ in Section 4.3.1. For the semileptonic modes, the effects of partial reconstruction are described through smearing with the κ-factor distribution $\mathcal{F}(\kappa)$ addressed in Section 4.3.2. Such effects are not present in the fully reconstructed modes, which formally corresponds to imposing $\mathcal{F}(\kappa)=\delta(\kappa-1)$. The definitions of the functions E and C are tacitly implied by the second equality in the equation. The normalization constant $\mathcal{N}$ is given in (4.20) and (4.23).

From (7.6), the importance of the taggers calibration expressed in (7.1) becomes apparent. It readily shows that the introduction of the amplitude parameter forbids the possibility of simultaneously extracting the tagging dilution which it directly multiplies. In fact, the determination of the main parameter of interest, $\mathcal{A}$, requires thus the accurate knowledge of the dilution, which justifies the necessity of previously accomplishing its calibration.

For non-physics background components, the tagging predicted dilutions $\mathcal{D}_j$ have no particular meaning, thus not being used in the corresponding PDFs. The latter contain instead terms which account for possible background flavor asymmetries which are globally described by dilution-like fit *parameters*, D_j. For such components, where mixing effects are not considered, the descriptions of proper decay time and tagging are decoupled. Namely, the following model is used

$$
L_{t,\xi}(t,\xi|D_j,\{\epsilon_i\}) \ \ = \ \ p_{\{\epsilon_i\}}(\xi_j)\,\frac{1+\xi_j D_j}{1+|\xi_j|}\ \cdot\ L_t(t)\ . \qquad (7.7)
$$

The j-index denotes the tagging algorithm providing the decision for the event of which the likelihood is being evaluated. A specific set of fit parameters $\{\epsilon_i, D_i\}$ accounts for tagging related effects. The proper decay time PDF L_t coincides with that presented in Section 4.2 for the relevant background components. Specifically, the model expressed by (7.7) is employed for the combinatorial background in all samples, as well as the fakes background in the semileptonic samples.

In general, an identical model as used for signal (7.6) is employed for the physics backgrounds as well. Specifically, for the $D_s\pi(\pi\pi)$, the Cabibbo-suppressed, the exclusively and partially reconstructed B_s background modes all share the same model and parameters employed for the signal component.

A full signal-like treatment is devoted to the partially reconstructed physics backgrounds in the $D_s l$ samples. That is, a model identical to that of the nominal signal (7.6) is employed to describe the contributions coming from the involved additional B_α (with $\alpha = u, d, s$) decays,

$$\begin{aligned} &L^{\alpha}_{t,\xi}(t, \xi_j|\lambda_\alpha; \mathcal{D}_j, \{\epsilon_i\}, w_\alpha, \tau_\alpha, \sigma_t) \\ &= \; p_{\{\epsilon_i\}}(\xi_j) \cdot \frac{1}{\mathcal{N}_\alpha} \frac{E(t;\tau_\alpha) + \xi_j \mathcal{D}_j \; \lambda_\alpha \; C(t;\tau_\alpha, w_\alpha)}{1 + |\xi_j|} \otimes_\kappa \mathcal{F}_\alpha(\kappa) \otimes G(t;\sigma_t) \cdot \mathcal{E}_\alpha(t) \; , \end{aligned} \tag{7.8}$$

where $\lambda_u = \lambda_d = 1$, and $\lambda_s = \mathcal{A}$; $w_u = 0$, $w_d = \Delta m_d$, and w_s stands for the Δm_s hypothesis. The tagging efficiency parameters $\{\epsilon_i\}$ are identified with those associated to the signal component in (7.6). The parameters τ_u, τ_d, and τ_u correspond to the B^+, B^0, and B_s mesons lifetimes, respectively. Specific κ-factor distributions $\mathcal{F}_\alpha(\kappa)$, and t-efficiency functions $\mathcal{E}_\alpha(t)$ are derived from Monte Carlo simulation of the contributing decays to each component.

7.5 Fits of the blinded data

The fits to the data samples are performed first with the true tagging information provisionally *hidden*. A *blinding* strategy in data analysis, generally speaking, implies avoiding knowing the final answer of the measurement until the procedure is fully specified and carried out. We adopt such a strategy which allows nevertheless for the full characterization of the samples to be achieved, except for the main quantity of interest, $\mathcal{A}$, along with the background flavor asymmetries. The analysis is fully performed in this fashion. Next, in Section 7.6, the systematic uncertainties are evaluated and, finally, the amplitude fits are repeated with the *unblinded* tagging input, in Section 7.7.

The tagging decision at input to the fitter is multiplied by the factor $(-1)^n$, with n standing for an integer specifying the event's ordering and thus being completely unconnected to tagging information. This scrambling of the tagging decisions thus hides the effects of a potential oscillation signal in the probed frequency region, and forbids that any interpretation be made from the central values of the fitted amplitude. We note, nevertheless, that the tagging status $\xi = 0$ and $\xi \neq 0$ are not confused by the scrambling criteria, and that thus the fractions of each such class of candidates for each tagging method remains unaffected. This way the taggers' efficiencies may be determined from the fits. The tagging dilutions are provided to the fit as input with no modification. The procedure allows accordingly for the determination of the uncertainty on the amplitude $\sigma_{\mathcal{A}}$, for each probed frequency, and therefore the sensitivity of the fitted samples is inferred.

The determination of the main parameter of interest, the amplitude, is performed once all other fit parameters have been found and thereafter fixed. The model parameters describing the mass and proper decay time spaces for the B_s samples are obtained in the fits to the data performed in Chapter 4 in the absence of flavor information. The parameters which remain to be determined depend on the flavor taggers' information, and correspond to:

1. the taggers' efficiencies $\{\epsilon_i\}$ for backgrounds and signal components,
2. the flavor asymmetries $\{D_i\}$ of non-physics backgrounds for each tagger,
3. the amplitude, $\mathcal{A}$, for each fixed value of the oscillation frequency, w for Δm_s.

Fits are first performed to the individual samples for the tagging efficiency and the flavor asymmetry parameters of the non-physics backgrounds. Those describing the combinatorial backgrounds are found from fits to the mass-sideband candidates. For the fakes background in the semileptonic modes, tagging asymmetries are imposed to be null, and the efficiency parameters are commonly identified with those to be determined for the signal components. A combined fit is then performed to the hadronic samples, as well as to the semileptonic ones. The efficiency parameters are thus found, commonly for either the hadronic or the semileptonic samples, which are associated to the signal and physics-background sample components.

The final stage of the fitting procedure are the amplitude fits. Here the amplitude parameter is the only free parameter floating, to be adjusted in the likelihood maximization, all other parameters being fixed to the values formerly determined. The amplitude is a common parameter in the fits performed simultaneously to all subsamples of the fully, and of the partially reconstructed modes. The frequency region $(0, 20)$ [ps^{-1}] is discretized in steps of 0.25 ps^{-1}, and the amplitude fit is repeated at each such fixed frequency value. The scan results correspond thus to the sequence of measured pair values $(\mathcal{A}, \sigma_{\mathcal{A}})$, of the fitted amplitude and associated uncertainty, for each Δm_s hypothesis.

The results of the blinded amplitude scans are shown in Figures 7-1 and 7-2 for the combined $D_s\pi(\pi\pi)$ and $D_s l$ samples, respectively. We note that the amplitude uncertainties shown are statistical only. Technically, in fact, while asymmetric uncertainties are provided by the likelihood minimization in fitting, only the upper uncertainties are used.

Although exclusion conditions, which do depend on the central fitted values, cannot be inferred from these blinded results, several assertions can be drawn. Similar statistical sensitivities are observed for the hadronic and semileptonic samples, of approximately 11 ps^{-1}. However such value alone hides in fact the rather distinct observed behavior in the two scans at lower and higher frequencies, which are predominantly determined by the samples' statistics and resolutions, respectively. As expected, thus, the amplitude uncertainties in the lower region of the spectrum are considerably smaller in the semileptonic scan than in the hadronic, while the reverse is true for the upper frequency region. This relative complementarity of the two classes of B_s samples will result in a considerable improved behavior in the combined amplitude scan.

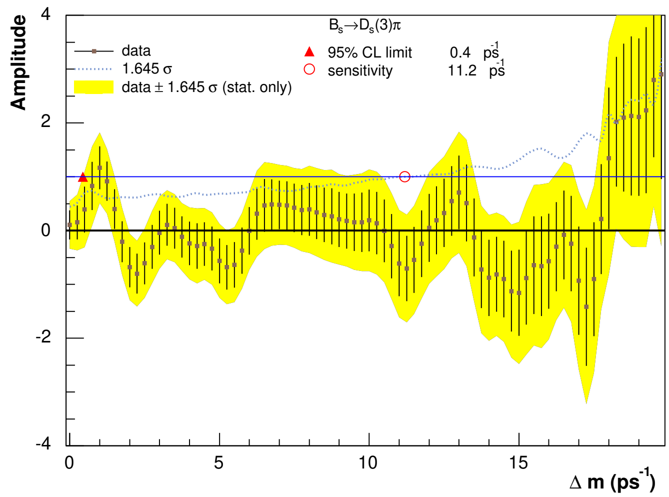


Figure 7-1: Blinded amplitude scan for $D_s\pi(\pi\pi)$ samples; uncertainties are statistical only.

7.6 Systematic uncertainties

Method of evaluation

A few issues arise regarding the estimation of systematic uncertainties on the amplitude parameter. For the purpose of establishing exclusion conditions for a given Δm_s hypothesis, both the corresponding fitted amplitude value and its uncertainty are directly employed. The point to emphasize is that systematic induced changes in the amplitude and in the statistical uncertainty on the amplitude, at a given frequency point in the scan, are correlated. The adequate evaluation of systematic uncertainties is derived in [31] and is expressed as

$$\sigma_{\mathcal{A}}^{\mathrm{syst}} \quad = \quad \Delta\mathcal{A} + (1-\mathcal{A})\,\frac{\Delta\,\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}}\ , \tag{7.9}$$

where $\Delta\mathcal{A}$ and $\Delta\,\sigma_{\mathcal{A}}$ are the observed variations in the fitted values for the amplitude and its statistical uncertainty induced by the systematic effect, relative to the values obtained in the nominal fit configuration, $\mathcal{A}$ and $\sigma_{\mathcal{A}}$.

The evaluation of the systematic uncertainties is fully based on toy Monte Carlo simulation. The latter provides a reliable parametric description of the fit input data, consistent with the likelihood models employed for each sample component. The parameters are set to those values previously found in fits to the data in Section 7.5. The toy Monte Carlo is generated with the value of the mixing frequency Δm_s for which the systematic uncertainty is to be evaluated. It is in effect this latter requirement of correspondence between the point of the amplitude scan and the true Δm_s value that has lead to the use of the toy Monte Carlo exclusively based approach to systematics evaluation. The toy Monte Carlo samples contain comparable statistics to the

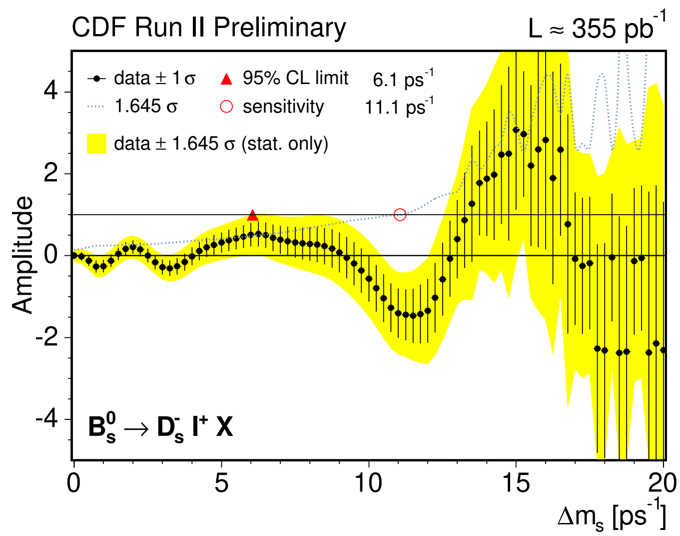


Figure 7-2: Blinded amplitude scan for $D_s l$ samples; uncertainties are statistical only.

data samples, in order to get distributions of returned amplitudes and uncertainties representative of those which would be expected in the data.

For each source of systematic uncertainty considered, a large set of about $1,000$ toy Monte Carlo samples is generated at each Δm_s value. Each of these samples is then fit under the nominal conditions, giving fit results $(\mathcal{A}^0, \sigma^0_{\mathcal{A}})$, and again with the variation on the systematic effect incorporated, with results $(\mathcal{A}^1, \sigma^1_{\mathcal{A}})$. The following quantity, analogous to (7.9), is then formed for each sample (i),

$$(\sigma^{\rm syst}_{\mathcal{A}})^i = (\mathcal{A}^1 - \mathcal{A}^0) + (1 - \mathcal{A}^0)\,\frac{\sigma^1_{\mathcal{A}} - \sigma^0_{\mathcal{A}}}{\sigma^0_{\mathcal{A}}}\ , \tag{7.10}$$

and its distribution is prepared for each samples' ensemble. The uncertainty contribution from each systematics source, at a given frequency point, is obtained as the mean of the above distribution, in case the systematic variation corresponds to turning an effect on or off, or to step-wise parameter variations. For a continuous variation, when parameter values are modified according to a distribution, the corresponding systematic contribution corresponds to the width of the distribution obtained from (7.10).

The systematic uncertainties evaluation procedure is repeated at a number of Δm_s values sampled in the probed frequency range. The contributions from the various systematics sources obtained at each such frequency point are added in quadrature. The outcome is then a profile $\sigma_{\mathcal{A}}(\Delta m_s)$ which is interpolated to give the systematic uncertainties for all frequency points in the scan. These are combined in quadrature to the statistical uncertainty on the amplitude obtained from the nominal fit. The exclusion condition, derived in Section 7.3, for a given probed frequency point $w =$

Δm_s becomes

$$\mathcal{A}(w) + 1.645\sqrt{\sigma_{\mathcal{A}}(w)^2 + \sigma_{\mathcal{A}}^{\mathrm{syst}}(w)^2} < 1 \ .$$

Uncertainty sources

Systematic uncertainties on the amplitude parameter are evaluated for the following sources.

Physics background levels: The amounts of physics backgrounds are known only to limited precision. In general, variations of the corresponding fraction parameters are implemented within uncertainties. For the Cabibbo-suppressed $B_s \to D_s K(\pi\pi)$ modes relative fraction uncertainties of 50% are assigned. The contribution from the partially reconstructed decays in the hadronic samples is nominally fixed to the Monte Carlo estimation. Variations are implemented where mass template parameters are allowed to float.

Semileptonic signal composition: The branching fractions of inclusive B_s decays to $D_s l$ and $D_s^{*(*)} l$ are not well known. The ratio of branching fractions for these contributions is varied by 20%. The κ -factor distributions and t-efficiency functions are modified accordingly as well.

Non-physics backgrounds in semileptonic samples: The fractions and shapes of the fakes background contributions are obtained respectively from fits to the fakes lepton samples and to the $m_{D_s l}$ distributions. Variations within the statistic uncertainties obtained from those fits are imposed as systematic effects. For the combinatorial background parameter variations are induced by extending the D_s mass-sideband regions by ± 50 MeV/c^2.

Dilution of fakes background in semileptonic samples: A null flavor asymmetry is assigned to the fakes background in the nominal fit. Variations of these dilution background parameters are taken within the dilution input values for signal.

Dilution and proper time uncertainty input calibration: Variations of the scale factors used for dilution (7.1) and proper decay time (7.2) calibration are taken in accordance with the values quoted in Tables 6.2 and 4.14, respectively.

Dilution and proper time uncertainty templates: The likelihood factors $L_{\mathcal{D}}$ and L_{σ_t} normally correspond to distributions obtained from sideband and sideband-subtracted data. The limited sample sizes and tagging efficiencies imply that some of these distributions may not be adequately populated. Variations of the dilution templates are constructed by continuously modifying the bin contents, through Gaussian-smearing, according to their statistical uncertainties. For the $D_s\pi(\pi\pi)$ samples, uniform templates are employed for the L_{σ_t} factors in the nominal fits. The resulting systematic effects are estimated by including non-trivial templates in the fits to the toy Monte Carlo.

Construction of t-efficiency curves, $\mathcal{E}(t)$: The t-efficiency functions are constructed from realistic Monte Carlo simulation, and small modifications may thus be induced by simulation related effects. The procedure used for evaluation of corresponding systematic shifts is similar to that described in Section 6.6. These correspond to relatively small contributions, which are neglected for the hadronic modes.

Unaccounted mixing effects in physics backgrounds: A full signal-like description (7.8) of the physics backgrounds is implemented for the semileptonic samples. For the hadronic samples, however, mixing effects are not considered in the models employed for the partially reconstructed B_s and misreconstructed B^0 contributions. No flavor asymmetry is assigned to these. Variations are estimated by incorporating oscillations, at the corresponding frequencies, in the toy Monte Carlo simulation.

Non-negligible $\Delta\Gamma/\Gamma$: The derivation of the proper decay time model in mixing used in the nominal fits are made in the assumption of zero lifetime difference between the two B_s mass eigenstates. In the case of a non-negligible lifetime difference contributions corresponding to the two $\Gamma = 1/\tau$ values are introduced (A.18) in the signal model. Its effect is evaluated by modifying the Monte Carlo generation model. A value of $\Delta\Gamma/\Gamma = 0.2$ is used in the simulation.

Resolution model: The detector resolution of the proper decay time is nominally modeled through a single Gaussian function (4.14). The effect of more elaborate resolution models is evaluated by fitting Monte Carlo samples generated with models involving an additional Gaussian and an exponential functions. Specifically, the former is characterized by a with of $2.5\,\sigma_t$ and a fraction of 17%, and the latter by a decay constant of about 100 μm and 1% fraction.

Summary of uncertainty results

The evaluated contributions from systematic sources discussed above for the fully and the partially reconstructed combined samples are compiled in Tables 7.1 and 7.2, respectively, for selected frequency points in the amplitude scan. The dominant contributions are the dilution scale factors, for the hadronic modes, and the level of physics backgrounds in the semileptonic samples. A graphical representation of the variation of each such contribution with Δm_s, obtained through polynomial interpolation, is presented in Figures 7-3 and 7-4. The uppermost curve corresponds to the combined systematic uncertainties for each case.

We point out that the evaluated systematic uncertainties on the amplitude are considerably smaller than the corresponding statistical uncertainties. We note also that with increasing sample sizes the dominant systematic sources will be also better controlled, and are thus not expected to impose precision limitations.

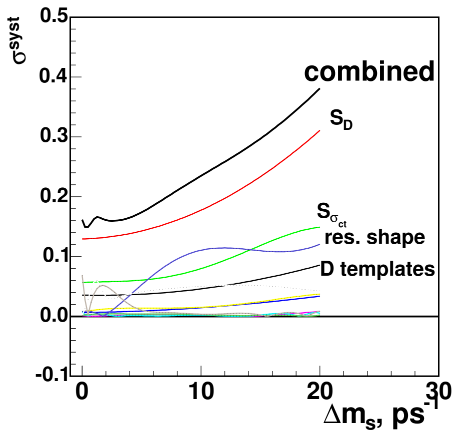


Figure 7-3: Summary graph of systematic uncertainties in the hadronic scan.

source	uncertainty at selected frequency point				
	0 ps^{-1}	5 ps^{-1}	10 ps^{-1}	15 ps^{-1}	20 ps^{-1}
physics background level	0.069	0.020	0.002	0.034	0.005
σ_{ct} scale factor	0.057	0.061	0.080	0.118	0.149
σ_{ct} templates	0.001	0.003	0.003	0.001	0.008
dilution scale factors	0.129	0.143	0.177	0.233	0.311
dilution templates	0.036	0.037	0.046	0.062	0.086
mixing in physics background	0.008	0.004	0.004	0.003	0.006
non-negligible $\Delta\Gamma/\Gamma$	0.002	0.041	0.050	0.051	0.043
resolution model	0.008	0.053	0.110	0.110	0.120
total systematic uncertainty	0.162	0.174	0.235	0.297	0.381
statistical uncertainty	0.251	0.400	0.567	0.846	1.177

Table 7.1: Summary of the uncertainties on the amplitude at selected frequency points in the hadronic scan.

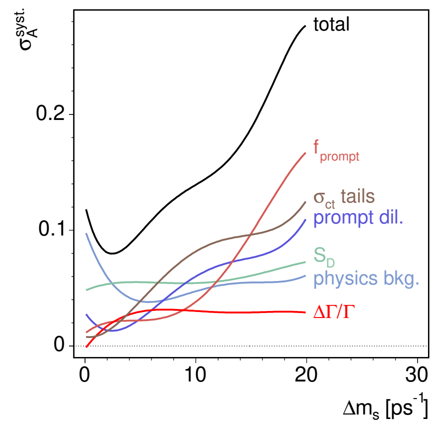


Figure 7-4: Summary graph of systematic uncertainties in the semileptonic scan.

source	uncertainty at selected frequency point				
	0 ps^{-1}	5 ps^{-1}	10 ps^{-1}	15 ps^{-1}	20 ps^{-1}
physics background level	0.107	0.041	0.044	0.054	0.060
fakes background level	0.012	0.024	0.042	0.102	0.177
combinatorial background level	0.011	0.021	0.029	0.036	0.042
signal composition	0.002	0.004	0.005	0.006	0.010
t-efficiency function	0.001	0.006	0.015	0.021	0.035
σ_t scale factor	0.001	0.019	0.024	0.030	0.070
dilution scale factors	0.047	0.054	0.054	0.061	0.073
fakes background dilution	0.036	0.027	0.057	0.075	0.109
non-negligible $\Delta\Gamma/\Gamma$	0.001	0.029	0.031	0.029	0.029
resolution model	0.006	0.031	0.080	0.094	0.125
total systematic uncertainty	0.123	0.093	0.138	0.187	0.277
statistical	0.088	0.252	0.565	0.758	1.850

Table 7.2: Summary of the uncertainties on the amplitude at selected frequency points in the semileptonic scan.

7.7 Results

The final step of the analysis, after having fully established the systematic uncertainties and the sensitivity of the measurements on the blinded samples, is to repeat the amplitude scans by removing the randomization of the tagging decision. The results of these unblinded amplitude scans are shown in Figures 7-5 and 7-6 for the fully and partially reconstructed B_s samples, respectively. The 95% exclusion limits obtained separately with the hadronic and semileptonic modes are 9.8 ps^{-1} and 10.4 ps^{-1}, respectively.

The combined analysis results are presented in Figure 7-7 which shows the amplitude scan obtained by combining, using (7.5), the hadronic and the semileptonic amplitude measurements. The uncertainty on the amplitude remains smaller than unity up until 17 ps^{-1}. The 95% sensitivity condition is verified exactly at 13.0 ps^{-1}, and approximately until 17 ps^{-1}, as indicated by the significance curve represented in the scan by the dashed line. The region defined by the set of probed points which are excluded at 95% C.L. is the following,

$$\text{excluded region}: \quad (0.00, 8.50) \cup (10.50, 11.50) \cup (12.50, 16.75)\ \text{ps}^{-1}\ .$$

The 95% exclusion limit achieved is 8.6 ps^{-1}.

The amplitude results obtained in previous experiments are summarized in Table 1-7. The achieved Δm_s sensitivity with our current data samples is competitive with the best single experiment. The combined amplitude scan based on previously published results is shown in Figure 1-8. The 95% C.L. exclusion limit and sensitivity are 14.4 ps^{-1} and 18.2 ps^{-1}, respectively. The modified world average amplitude scan with the inclusion of the amplitude measurements represented in Figure 7-7 is shown in Figure 7-8. The improved 95% C.L. exclusion limit is 16.6 ps^{-1}, and the sensitivity is pushed to 19.6 ps^{-1}. Under the assumption that Δm_s lies within the probed spectrum, all frequency values are excluded at 95% C.L. except for the following double-sided interval

$$\text{allowed region}: \quad (16.6, 20.8)\ \text{ps}^{-1}\ .$$

The impact of the reported results obtained with the current data samples on the world average knowledge of Δm_s is already considerable.

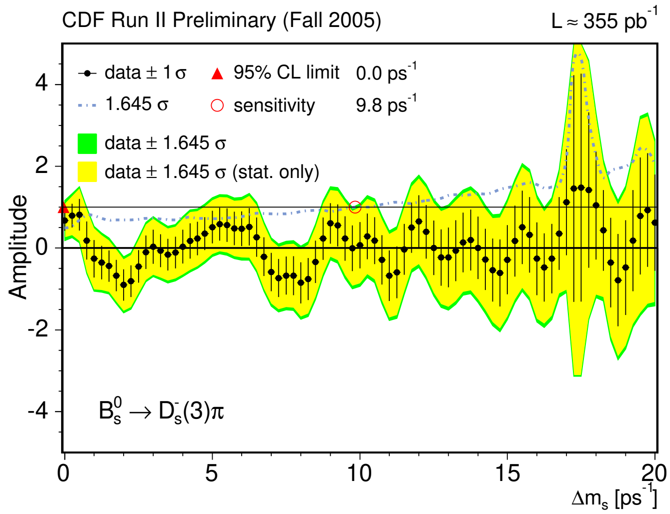


Figure 7-5: Amplitude scan in unblinded data for the hadronic samples.

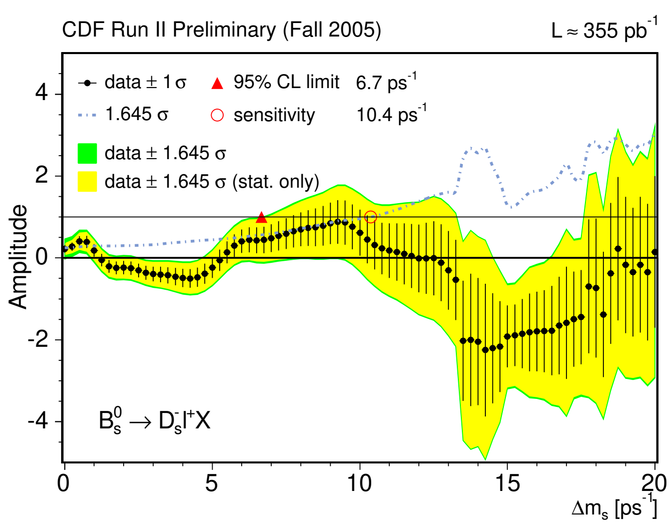


Figure 7-6: Amplitude scan in unblinded data for the semileptonic samples.

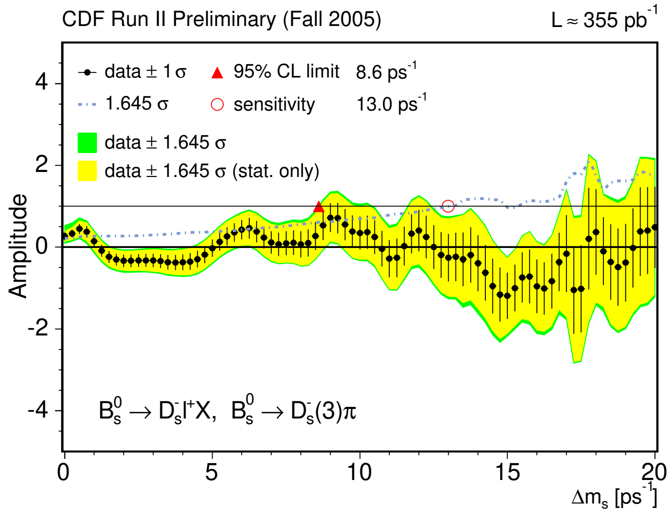


Figure 7-7: Combined amplitude scan in unblinded data for the hadronic and semileptonic samples.

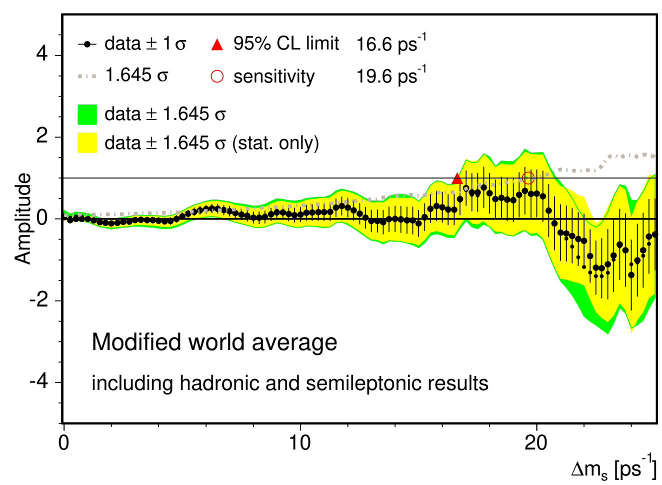


Figure 7-8: Modified world average amplitude scan with hadronic and semileptonic results.

7.8 Résumé

The study of B_s flavor oscillations in fully and partially reconstructed data samples was accomplished, based on the unbinned fitting framework, likelihood model, flavor tagging and input calibration previously presented.

The opposite-side tagging methods presented in Chapter 5 and further calibrated in Chapter 6 are applied to the hadronic $B_s \to D_s\pi(\pi\pi)$ and the semileptonic $B_s \to D_s lX$ decay samples. The underlying samples' description in mass and proper decay time spaces was achieved in Chapter 4, where the calibration of proper decay time uncertainties was also implemented.

In anticipation to the rapid oscillation frequency characterizing the B_s system, an alternative method to the direct determination, as applied for Δm_d in Chapter 6, is employed. The method consists in performing a scan in frequency and measuring at each such frequency value the amplitude of the probed oscillation. This so-called amplitude method is appropriate for setting exclusion conditions, and allows for a straightforward combination of existing experimental measurements.

A blinding technique is adopted, which consists of provisionally randomizing the tagging decisions. The full analysis is first performed in this fashion. The systematic uncertainties on the amplitude are evaluated taking into consideration correlated variations of the amplitude value and its statistical uncertainty, using toy Monte Carlo simulation. The dominant contribution to the combined uncertainty is manifestly statistical.

The quantities which are typically defined to characterize the results of the amplitude scan are the 95% C.L. sensitivity and exclusion limit. From the combined semileptonic and hadronic scans, these are given by:

$$\begin{aligned} \text{exclusion limit}: &\quad 8.6\ \text{ps}^{-1} \quad (95\%\ \text{C.L.}) \\ \text{sensitivity}: &\quad 13.0\ \text{ps}^{-1} \quad (95\%\ \text{C.L.}) \end{aligned}$$

These two quantities provide a short, incomplete summary of the results. The amplitude measurements performed at each probe frequency are combined with previously published results, having already a considerable impact on the accumulated knowledge of Δm_s. From such combination, assuming that the B_s oscillation frequency is smaller than 25 ps^{-1}, the interval of frequencies not excluded at 95% C.L. is given by

$$(16.6, 20.8)\ \text{ps}^{-1}\ .$$

The increasing size of our data samples along with other improvements are bound to further contribute significantly for our knowledge of the Δm_s parameter. In Chapter 8 we present one outstanding improvement in flavor tagging. The degree of exclusion achieved along the probed frequency spectrum is contained in the full amplitude scan. This complete set of experimental information about Δm_s is what is employed for all practical purposes, including for constraining the CKM matrix elements which we address in Chapter 9.

Chapter 8

Same-side kaon tagging and projections

8.1 Algorithm development

The basic ideas behind the method of same-side tagging have been exposed in Chapter 5. The specific algorithm presented therein was applied, along with the opposite-side tagging methods, to samples of fully reconstructed B^+ and B^0 decays, in Chapter 6. The algorithm's dilution was measured, for the charged modes, as the observed overall flavor asymmetry and, in the case of the neutral modes, from the analysis of the time-dependent flavor asymmetry. A similar evaluation of the method's performance in samples of B_s mesons was not provided then, as it is forbidden by the rapid mixing oscillations which characterize the system, for the available data samples.

The goal presently is to develop an optimized same-side tagging method, and estimate its performance so it can be applied in the analysis of B_s flavor oscillations. The strategy in algorithm development is to validate the Monte Carlo description of the relevant processes and observables, and subsequently extract the algorithm's performance from Monte Carlo samples. The cornerstone of such validation is to be provided by the successful performance description, verified against data, for the B^+ and B^0 mesons.

The generation of large Monte Carlo samples further offers the possibility of identifying and parameterizing dilution dependencies on relevant quantities. This will result in increased tagging power. For the B_s mesons, as illustrated in Figure 5-6, the leading fragmentation particle is expected to be a kaon. The performance of the method can thus be enhanced through the use of particle identification techniques for identifying kaons among the track candidates. The use of such information when available will accordingly be explored, and the corresponding algorithm is sometimes referred to as same-side *kaon* tagger.

Unlike the study of the tagging methods presented in Chapter 5, where a *binned* approach was employed throughout, the proposed SST algorithms' study benefits presently from the more thorough *unbinned* likelihood technique and description of the samples developed in previous chapters.

8.1.1 Data and Monte Carlo samples

Samples of fully reconstructed decays of the B^+, B^0 and B_s meson species are studied.

Monte Carlo generation

The Monte Carlo samples ought to contain a complete description of the hadronization process, and of the properties of the resulting fragmentation particles accompanying the B mesons.

Events are generated with `Pythia` [73] (version 6.216); the program's parameter `msel` is set to 1. All of the following QCD high-$p_\perp$ processes are generated: $f_i f_j \to f_i f_j$, $f_i \bar{f}_i \to f_k \bar{f}_k$, $f_i \bar{f}_i \to gg$, $f_i g \to f_i g$, $gg \to f_k \bar{f}_k$ and $gg \to gg$, where $f_{i,j,k}$ stand for fermions and g for gluons. The following $b\bar{b}$ production mechanisms [74] are correspondingly included. The leading process, denoted *flavor creation*, corresponds to the production of a $b\bar{b}$ pair by gluon fusion or by the annihilation of light quarks via the 2-to-2 parton subprocesses $q\bar{q} \to b\bar{b}$ and $gg \to b\bar{b}$. Another production process, referred to as *flavor excitation*, occurs when a virtual heavy quark from the parton distribution of a beam particle is put on mass shell through scattering by a light quark or gluon, via the subprocesses $gb \to gb$ and $qb \to qb$. A third source, denoted *gluon splitting*, comes from reactions with only gluons and light quarks participating in the 2-to-2 hard parton scattering subprocess, and where the $b\bar{b}$ pair is produced in a gluon shower either by a final or initial state gluon. All such processes need to be included in Monte Carlo generation for obtaining an accurate description of data.

The default `Pythia` Lund fragmentation model was used to describe the production and the properties of particles formed in the hadronization process. The energy taken away from the b string by the B meson is described by a symmetric Lund function, whose parameterization has been taken from fits to LEP data.

The default values for B^{**} masses and widths were updated with recent measurements [1]. A fraction of 20% of the B^+ and B^0 mesons was set to have originated from B^{**} decays.

In order to efficiently produce large Monte Carlo samples, the underlying processes from an original sample are kept, while the heavy flavor particles are re-decayed, using the program `EvtGen` [63]. No neutral B meson mixing is included in the simulation.

Selection and fitting model

The criteria for signal reconstruction follows the selection presented in Chapter 3. The unbinned likelihood based fitting model developed in Chapters 4 and 6 is employed.

For selecting B decays in the $J/\psi K$ samples so far criteria have been employed to ensure that no bias in proper decay time is introduced. Presently, however, a minimum decay distance significance threshold ($L_{xy}/\sigma_{L_{xy}} \gtrsim 4.5$) is imposed on the B candidates in order to further suppress the background contamination. This is relevant for obtaining meaningful comparisons of signal distributions in data to those obtained from Monte Carlo, where mass-sideband subtraction is performed for the former. The proper decay time bias introduced by this selection condition is exactly described in the likelihood model, as derived in Section 4.2 and expressed in (4.17).

Figures 8-1 and 8-2 show the likelihood projection in mass and proper time spaces for $J/\psi K$ sample decays. The decrease in the combinatorial background level, when comparing to Figure 4-7, is apparent from the mass distributions, as it is from the suppression of the prompt peak in proper decay time.

The general level of agreement of Monte Carlo and data is illustrated in Figure 8-3 for selected B meson candidates' distributions.

8.1.2 Pre-selection of track candidates

For being considered as potential tagging candidates tracks must fulfill certain criteria. These have been motivated and outlined in Section 5.2. The specific requirements imposed on the track candidates are the following:

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

Figure 8-4 shows data and Monte Carlo distributions compared for some of the involved quantities, where all cuts but that on the corresponding observable have been applied. In general, a satisfactory level of agreement is observed.

The set of tagging candidates is identified with those tracks which satisfy the above pre-selection criteria. The transverse momentum p_T of the selected candidates is shown in Figure 8-6 for data and Monte Carlo simulation. The partial distributions for the generator level identified particles are also represented. The origin of the pion and kaon track candidates is found from generator level information and is shown in Figure 8-5; right (wrong) sign indicate whether the track has the correct (incorrect) charge correlation with the B flavor expected by the SST.

The set of tagging candidates may contain none, single or multiple elements. Figure 8-7 shows the tagging candidates multiplicity for the selected decays. When one or multiple tagging candidates have been accepted the corresponding event is assigned to one of the following classes:

Agreeing class: if only one single track has been selected, or the charges of the selected tracks coincide;

Disagreeing class: if not all of the selected tracks have a common charge sign.

Rather roughly, about 60% of the time there's at least one track candidate, and approximately 40% of the time the track candidates have a common charge. For events belonging to the former class, the SST decision is unambiguously given by the identified track(s) charge. For the latter class, a decision needs to be made about which track should be selected as the tagging track. Several algorithm implementations are explored in the coming sections for accomplishing such purpose.

8.1.3 Candidate algorithms

The aim of the SST method is to identify the charge signal of the leading track, expected to be found *nearest* in phase space to the B meson. When several track candidates are available for a given event, a specific definition of *nearest* track needs to be adopted for selecting the tag. The following implementation possibilities are explored:

1. p_T^{rel}, selects the track with the smallest relative transverse momentum to the direction given by the sum of the track and B momenta,

2. p_L^{rel}, selects the track with the largest relative longitudinal momentum to the direction given by the sum of the track and B momenta,

3. p_T, selects the track with the largest transverse momentum,

4. $M_{B\pi}$, selects the track with the smallest invariant mass with the B,

5. ΔR, selects the track with the smallest ΔR relative to the B,

6. $p_L^{rel}Q$, the decision is given by the sum of the charges of the tag candidate tracks, or, in case this gives a null decision, by that of the p_L^{rel} algorithm above.

The first criteria listed corresponds to that employed in Section 5.2. The definition of the quantities p_T^{rel} and p_L^{rel} is illustrated in Figure 5-7. The corresponding distributions are shown in Figure 8-8.

The various listed criteria, which are more or less strongly inter-correlated, each offer in principle some discriminating power. In the future, these may therefore be combined along with additional discriminating information through an adequate multivariate mechanism, such as an artificial neural network. Currently nevertheless we are interested in adopting a simpler selection criteria which will be used in data and in Monte Carlo simulation to further explore the tagging related processes associated to the various B meson species.

Figure 8-9 provides a comparison of the tagging performance of the listed candidate algorithms. These are applied solely to events that contain multiple track candidates with differing charges. For these cases, a decision is always achieved, implying that all algorithms deliver identical tagging efficiencies.

The p_L^{rel} algorithm shows a relatively good performance, and it is the one we decide to elect among the list of candidates above. In particular, we conclude that it performs better than the p_T^{rel} algorithm which has formerly been adopted.

8.1.4 Parameterized dilution

The parameterization of potential dilution dependences offers several advantages over using an overall, average dilution value. In general and foremost it results in a relative increase of the algorithm's tagging performance. In addition it accounts for variations of the parameterized quantity which may exist between samples, and, as it is here most important, between data and Monte Carlo simulation. If for example less high momentum events would be found in data than in Monte Carlo (for instance due to a prescaling of the trigger not implemented in the simulation) predicted dilutions depending on that quantity would automatically assign to the data sample a smaller (as it would be the case) overall dilution relative to the Monte Carlo sample, while the shape of the predicted (*i.e.* parameterized) dilution wouldn't be affected.

The classification of tagged events into the classes of agreeing and disagreeing charges was introduced anticipating differences in expected characteristic dilutions. Accordingly, dilutions are evaluated and assigned separately to each of those classes. As expected a better performance is found for the agreeing case compared to the disagreeing case.

The SST dilution reveals an expected increase with the transverse momentum p_T of the tagging track. This dependency is shown in Figure 8-10 for the adopted p_L^{rel} algorithm, for the agreeing and disagreeing classes. It is parameterized as

$$\mathcal{D}(p_T) = \alpha_0 - \alpha_1 \cdot e^{-\alpha_2 \cdot p_T} . \tag{8.1}$$

This shape is in general suitable, as the projections in Figure 8-10 indicate. However, a deviant behavior is observed for the B^0 mesons in the class of disagreeing tracks. This is understood from the generator level information displayed in Figure 8-11. In effect, at higher momentum kaons tend to be more often selected as the tagging track, and these contribute negatively to the dilution in the case of the B^0 mesons. We mention in passing that the referred deviant behavior is observed for the p_L^{rel} and p_T algorithms, but not for others, such as the p_T^{rel} implementation. For the disagreeing class in B^0 modes only an average dilution is used, as the available sub-samples size is too small for allowing an accurate description of the resulting shape.

8.1.5 Kaon identification

The tagging tracks contribute distinctively to the SST dilution depending on the identity of the associated particles. This can already be inferred from the schematic in Figure 5-7, and it is demonstrated by the generator level information displayed in Figures 8-5 and 8-11. In particular, in the case of B_s mesons the correct SST charge correlation is expected to be provided by kaons, while no contribution is expected on average from either pions or protons. The performance of the SST method is in this way expected to benefit considerably from the use of particle identification information which may be available for the tracks.

Particle identification (PID) information for charged tracks is provided at CDF by two means: energy loss dE/dx in the tracking chamber COT, and time of flight measured by the TOF detector. The two measurements are complementary in that the latter is more effective for low momentum tracks and the former is more effective for intermediate momenta. An optimized separation power is obtained by combining the corresponding information,

$$L(\text{particle}) = P_{TOF}(\text{particle}) \cdot P_{dE/dx}(\text{particle}) \qquad \text{with} \quad \text{particle} = K, p, \pi \ ,$$

where P_{TOF} and $P_{dE/dx}$ are the probability density functions which have been determined on pure particle samples. We use the following combined likelihood ratio, defined for the kaon hypothesis,

$$\log(LH(PID)) = \log\left(\frac{P_{TOF}(K) \cdot P_{dE/dx}(K)}{f_p \cdot P_{TOF}(p) \cdot P_{dE/dx}(p) + f_\pi \cdot P_{TOF}(\pi) \cdot P_{dE/dx}(\pi)}\right) \ , \quad (8.2)$$

where $f_p = 0.1$ and $f_\pi = 0.9$ are the prior probabilities for background composition. In case no TOF or no dE/dx information is available, the corresponding probabilities are taken as 1. The dE/dx, TOF, and combined likelihoods are illustrated in Figures 8-12 and 8-13.

With particle identification information available for the track candidates, an alternative criteria for selecting the tagging track among multiple candidates is proposed: select that with the highest probability for being a kaon. This will be referred to as the PID algorithm, and it is proposed having specifically in mind the B_s case. Although it is clearly not optimal for the B^+ and B^0 cases, we extend its study to these high statistics B meson samples to further test the Monte Carlo description.

The algorithms performance is studied as a function of the value of the PID variable for the tagging track. The observed dilution dependences are shown in Figure 8-14 for the classes of agreeing and disagreeing track candidates. For the B^+ we see high dilution both for more pion like and more kaon like tracks. For the B^0, the pions are likely to provide the expected SST charge-flavor correlation (positive dilution), while kaons are more likely anti-correlated (negative dilution). The pattern is somewhat inverted for the B_s, where kaons are potentially good tagging tracks; pions do not carry in this case any information on the B production flavor (zero dilution). The decrease of predicted dilution at very hight kaon probability values is related to the fact that both particle ID systems (dE/dx and TOF) give best kaon-pion separation for low momentum tracks. Thus the high kaon probability tracks are most likely true low momentum kaons which are coming from the underlying event. However leading fragmentation tracks are more likely to have higher momenta.

8.1.6 Tagging performance

The tagging performance is evaluated using the unbinned likelihood framework developed in previous chapters, applied to the exclusive decay samples indicated above.

The results are summarized in Tables 8.1 and 8.2 for the p_L^{rel} and PID algorithms, respectively. The tagging efficiency is denoted by ϵ, and the average dilution by $\mathcal{D}$, which are floating parameters of the fit.

The evaluation of the tagging performance with the parameterized dilution proceeds as follows. The value of the dilution is predicted for each event according to the classification and dependencies which were derived in the previous sections and obtained from the corresponding Monte Carlo samples. The predicted dilution values are provided as input to the fit. Correspondingly, dilution templates for signal and background are derived and used in the likelihood model. The procedure is identical to that used in Chapter 6 for the purpose of OST calibration. A dilution scale factor $S_{\mathcal{D}}$ is introduced as a floating parameter of the fit, which directly multiplies the dilution in the proper decay time PDF for flavor tagged events. Finally, the effective dilution $\mathcal{D}_{\text{eff}}$ is obtained

$$\mathcal{D}_{\text{eff}} = S_{\mathcal{D}} \sqrt{\langle \mathcal{D}^2 \rangle} \,, \tag{8.3}$$

from the fitted scale factor value, and the dilution squared average over signal events.

p_L^{rel} algorithm [%]		B^+		B^0		B_s	
		$J/\psi K^+$	$\bar{D}^0\pi^+$	$J/\psi K^{*0}$	$D^-\pi^+$	$J/\psi\phi$	$D_s^-\pi^+$
MC	ϵ	55.2±0.1	55.9±0.1	54.4±0.1	56.6±0.1	49.5±0.3	52.1±0.3
	$\mathcal{D}$	25.3±0.2	25.7±0.3	14.9±0.3	15.1±0.4	16.2±0.8	19.0±0.8
	$S_{\mathcal{D}}$	100.3±0.7	100.3±0.9	101.1±1.8	98.0±2.1	99.7±4.4	100.5±3.3
	$\mathcal{D}_{\text{eff}}$	28.5±0.2	29.0±0.3	16.8±0.3	17.2±0.4	18.6±0.8	22.8±0.7
	$\epsilon\mathcal{D}^2_{\text{eff}}$	4.5±0.1	4.7±0.1	1.5±0.1	1.7±0.1	1.7±0.2	2.7±0.2
data	ϵ	60.0±0.7	58.4±0.5	57.3±1.0	57.2±0.6	48.1±2.9	49.3±2.3
	$\mathcal{D}$	22.3±1.9	25.9±1.4	10.7±4.6	13.3±2.9	—	—
	$S_{\mathcal{D}}$	91.8±6.3	103.6±4.7	85.9±29.1	110.0±16.0	—	—
	$\mathcal{D}_{\text{eff}}$	26.4±2.1	30.4±1.3	13.6±5.4	19.0±2.5	18.6±1.0	23.7±1.0
	$\epsilon\mathcal{D}^2_{\text{eff}}$	4.2±0.6	5.4±0.5	1.0±0.7	2.1±0.6	1.7±0.2	2.8±0.3

Table 8.1: Tagging performance of the p_L^{rel} algorithm in Monte Carlo and data; the quoted uncertainties are statistical only.

A relatively better performance is in general obtained with the PID algorithm for the B_s mesons, and with the p_L^{rel} algorithm for the $B^{+,0}$ samples. It should also be clear that their combination, along with the use of additional tracks' information, will potentially result in further improvements. The introduction of dilution parameterizations induce an absolute gain of about 2-5% in dilution. The observation that the dilution scale factors are close to unity further indicates that those parameterizations provide adequate descriptions in both Monte Carlo and data.

The tagging dilution is not *measured* for the B_s *data* samples. The reason, to restate it, being that the rapid flavor oscillations do not allow for a direct measurement of the unknown oscillation frequency, as it is done for the B^0 system, given current sample sizes, resolutions and flavor taggers. The strategy therefore is to use the Monte Carlo predictions as estimates of the tagger performance in data. In Tables 8.1 and 8.2, the dilution and tagging power quoted for the B_s samples use obtained accordingly employing the predicted dilution and scale factor from simulation. For the hadronic B_s mixing data sample the tagging power of the same side kaon (PID)

PID algorithm [%]		B^+ $J/\psi K^+$	B^+ $\bar{D}^0\pi^+$	B^0 $J/\psi K^{*0}$	B^0 $D^-\pi^+$	B_s $J/\psi\phi$	B_s $D_s^-\pi^+$
MC	ϵ	55.2±0.1	55.9±0.1	54.4±0.1	56.6±0.1	49.5±0.3	52.1±0.3
	$\mathcal{D}$	24.2±0.3	24.5±0.3	12.7±0.3	12.9±0.4	18.9±0.8	21.8±0.8
	$S_{\mathcal{D}}$	100.0±0.7	100.0±0.9	98.0±1.7	98.5±2.1	98.8±3.6	98.8±2.5
	$\mathcal{D}_{\text{eff}}$	27.0±0.2	27.5±0.2	17.9±0.3	17.4±0.4	22.3±0.8	28.5±0.7
	$\epsilon\mathcal{D}^2_{\text{eff}}$	4.0±0.1	4.2±0.1	1.7±0.1	1.7±0.1	2.5±0.2	4.2±0.2
data	ϵ	60.0±0.7	58.4±0.5	57.3±1.0	57.2±0.6	48.1±2.9	49.3±2.3
	$\mathcal{D}$	21.7±1.9	25.4±1.4	10.7±4.6	14.2±2.9	—	—
	$S_{\mathcal{D}}$	87.6±6.7	95.1±4.7	78.8±25.1	104.7±14.4	—	—
	$\mathcal{D}_{\text{eff}}$	23.9±2.2	25.7±1.3	13.7±5.5	17.6±2.3	22.1±1.0	28.3±1.1
	$\epsilon\mathcal{D}^2_{\text{eff}}$	3.4±0.5	3.9±0.4	1.1±0.7	1.8±0.5	2.4±0.3	4.0±0.4

Table 8.2: Tagging performance of the PID algorithm in Monte Carlo and data; the quoted uncertainties are statistical only.

algorithm implemented is

$$\epsilon\mathcal{D}^2 = 4.0 \pm 0.4 \ .$$

The evaluated uncertainty corresponds to statistical effects. Systematic variations associated to fragmentation and simulation processes in the Monte Carlo are necessary for assessing the full dilution uncertainty. The latter is relevant in the absence of a signal in the mixing analysis, for the purpose of evaluating frequency exclusion conditions.

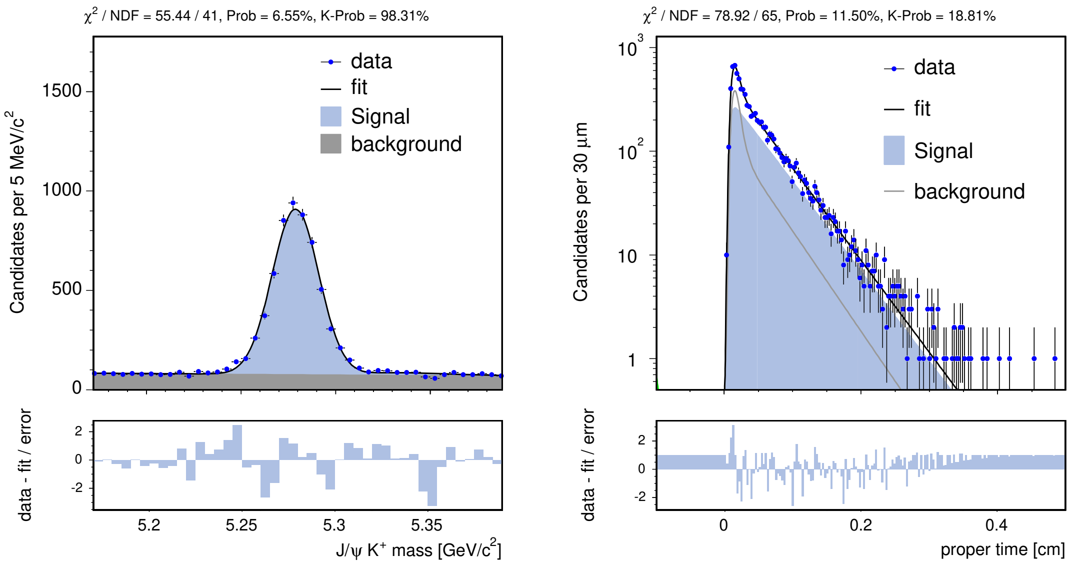


Figure 8-1: Mass and proper decay time fit projections for $B^+ \to J/\psi K^+$ decay, with $L_{xy}/\sigma_{L_{xy}}$ selection requirement.

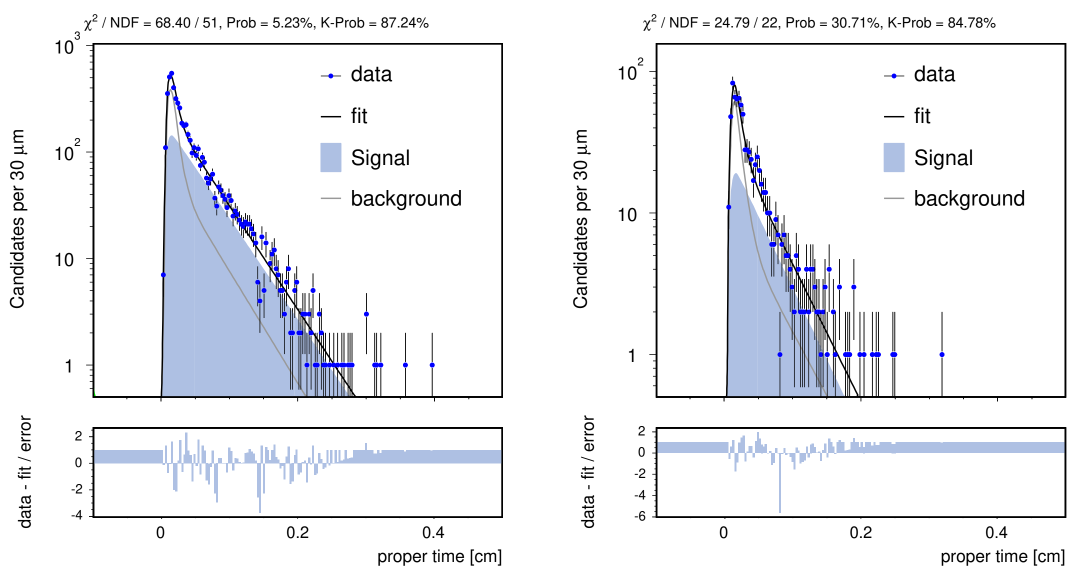


Figure 8-2: Proper decay time fit projections for $B^0 \to J/\psi K^{*0}$ and $B_s \to J/\psi\phi$ decays.

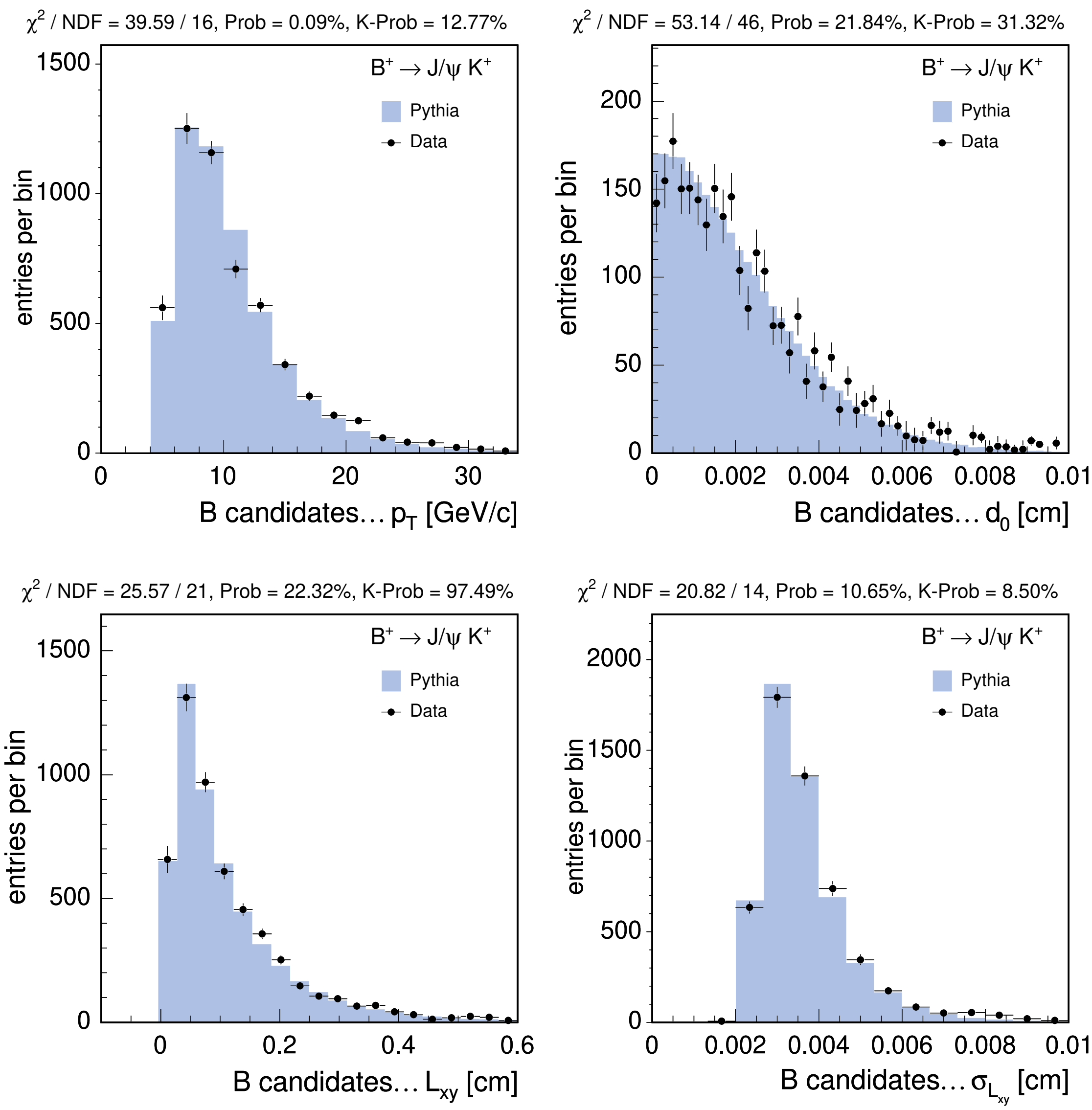


Figure 8-3: Comparison of distributions between data and Monte Carlo: transverse momentum p_T, impact parameter d_0, flight distance in the transverse plane L_{xy}, and error in L_{xy}, for B candidates reconstructed in $J/\psi\, K^+$ decays.

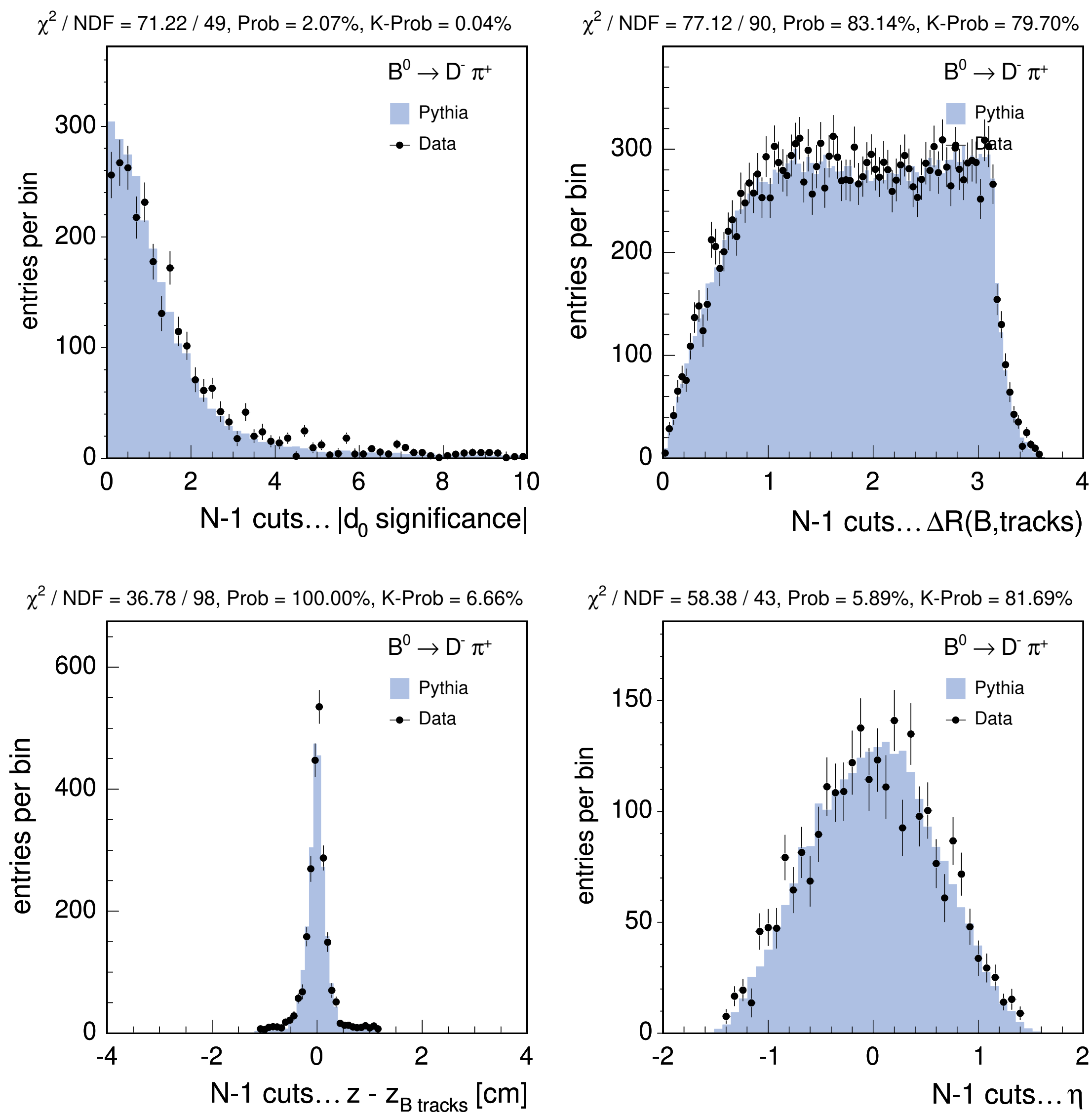


Figure 8-4: Comparison of track variables used for the pre-selection of tagging track candidates, illustrated for $B^0 \rightarrow D^-\pi^+$ decays.

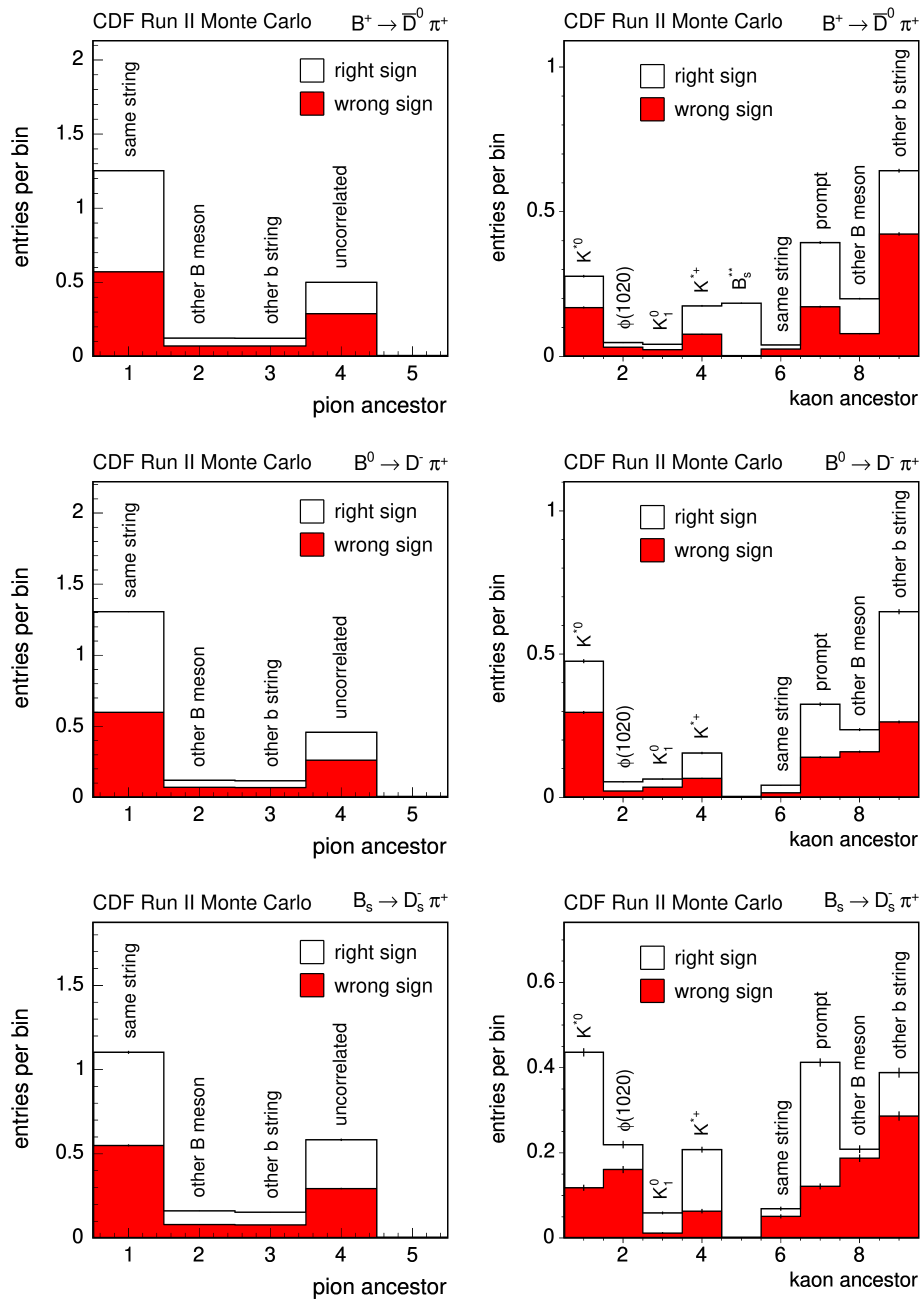


Figure 8-5: Generator level identification of the origin of the pion and kaon track candidates.

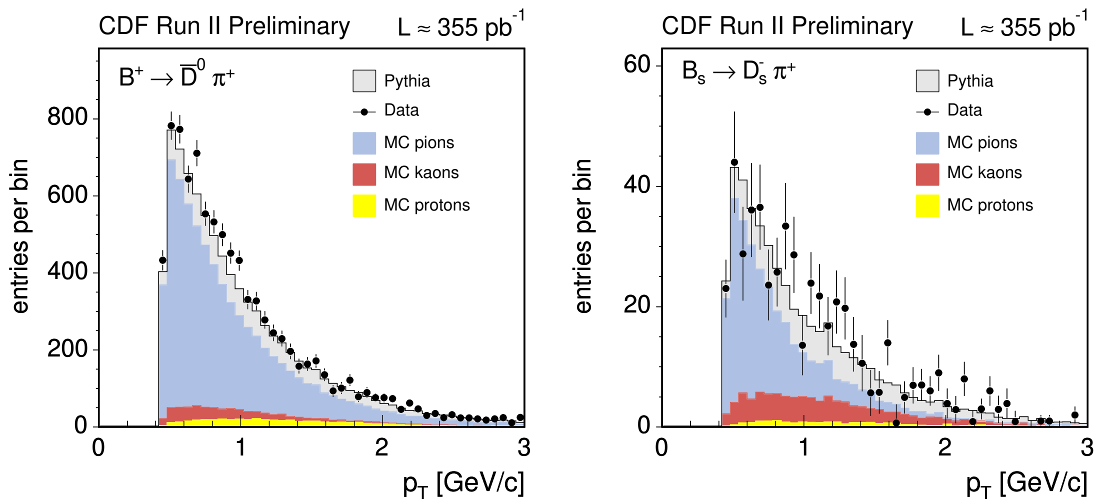


Figure 8-6: Tagging candidates' p_T comparison between data and Monte Carlo, illustrated for the $B^+ \to \bar{D}^0\pi^+$ and $B_s \to D_s^-\pi^+$ modes; the particle identities of the Monte Carlo tracks are also shown.

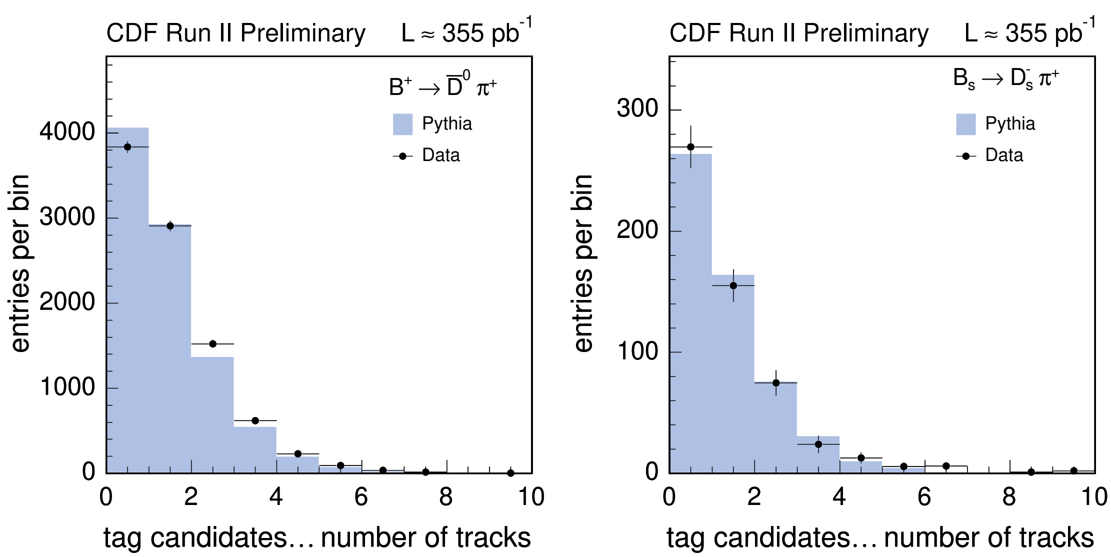


Figure 8-7: Tagging candidates's multiplicity comparison between data and Monte Carlo, illustrated for the $B^+ \to \bar{D}^0\pi^+$ and $B_s \to D_s^-\pi^+$ modes.

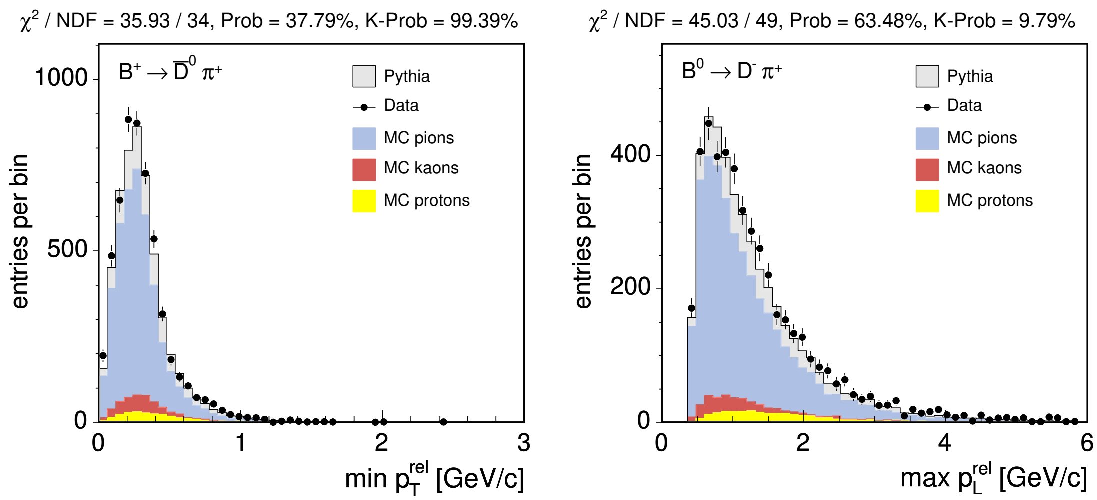


Figure 8-8: Tagging track minimum p_T^{rel} and maximum p_L^{rel}, for data and Monte Carlo.

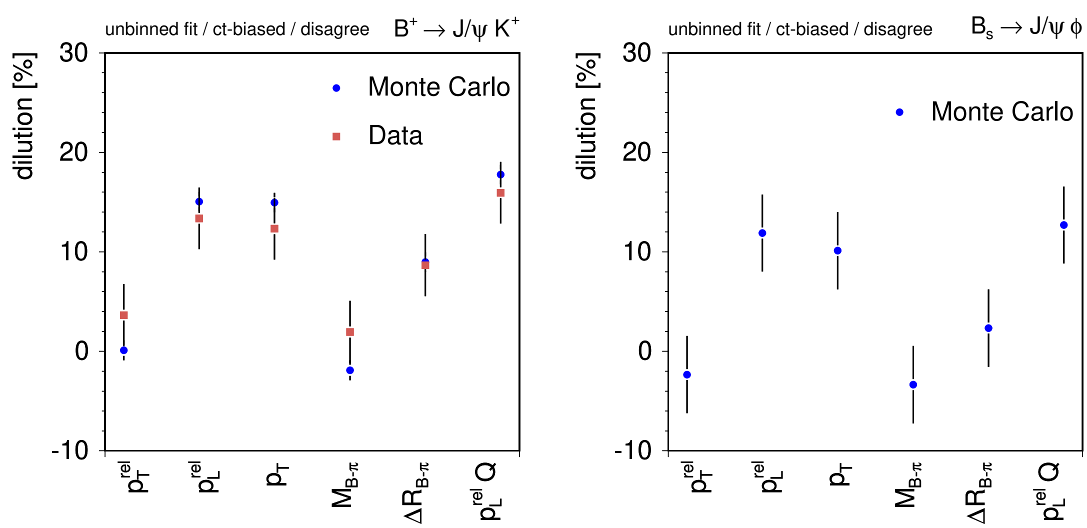


Figure 8-9: Relative performance of several candidate algorithms, applied to events with multiple tagging track candidates of disagreeing charges in B^+ and B_s samples.

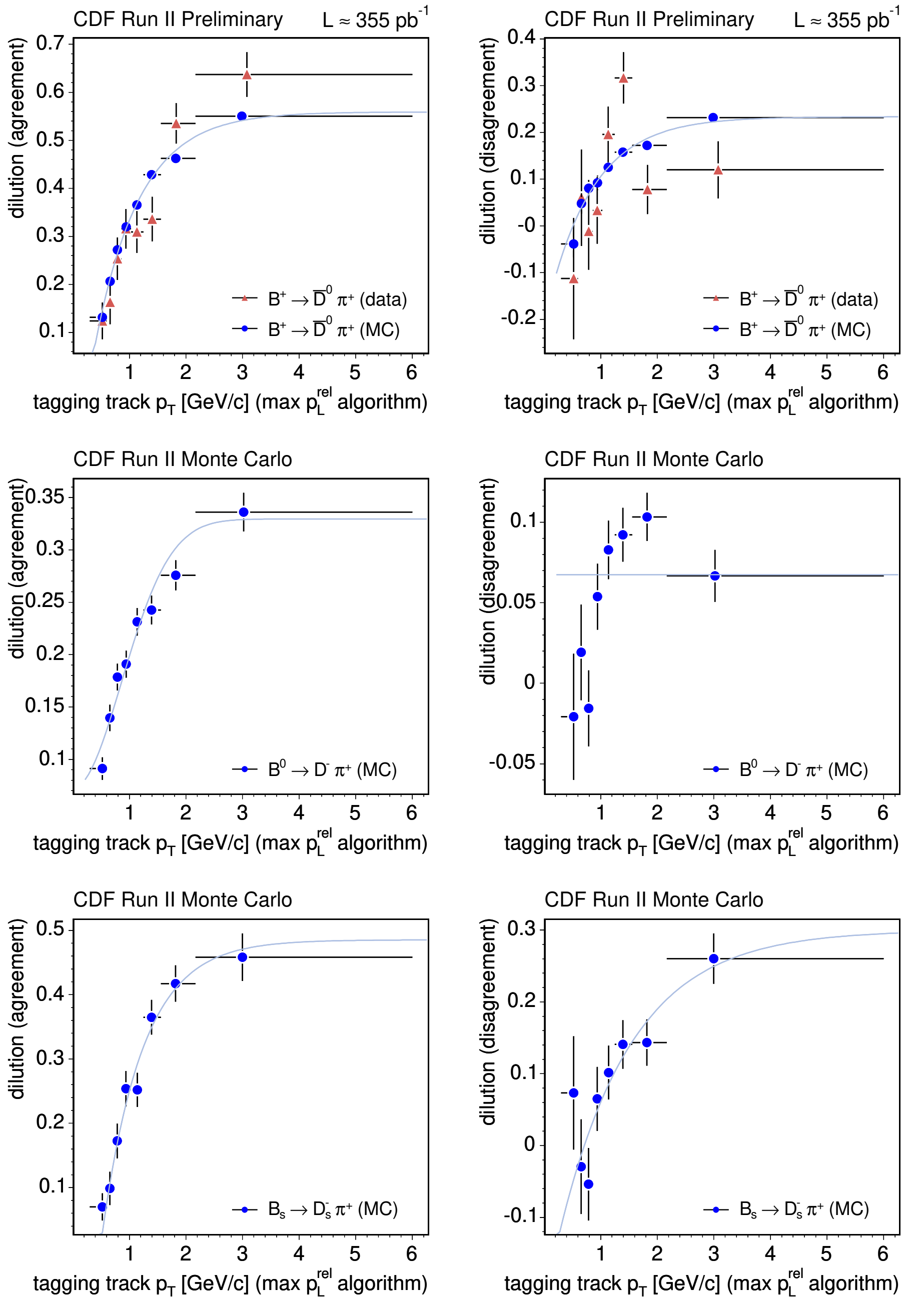


Figure 8-10: Dilution of the p_L^{rel} algorithm as a function of the tagging track p_T for classes of agreeing and disagreeing tagging candidate charges.

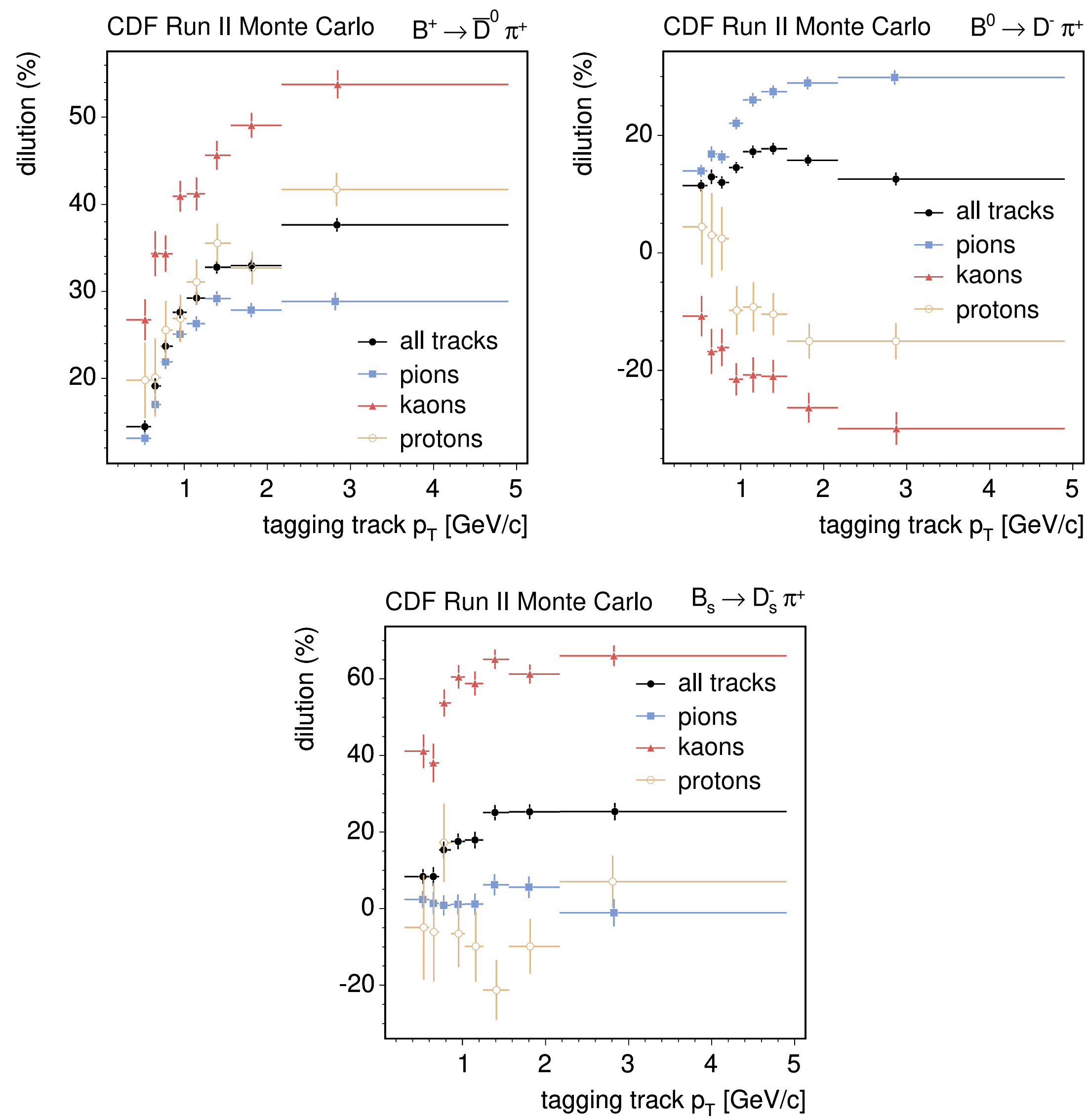


Figure 8-11: Dilution dependence of the p_L^{rel} algorithm on the tagging track p_T for different particle species, from Monte Carlo information.

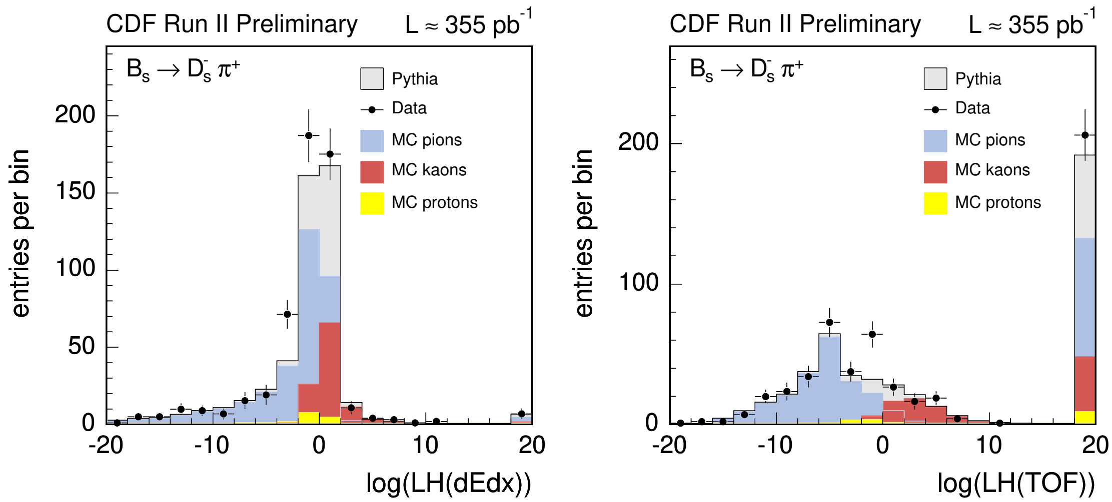


Figure 8-12: Likelihood distribution for particle identification with dE/dx (left) and TOF (right), shown for $B_s \to D_s^- \pi^+$ decays; the rightmost bin in the distributions refers to tracks for which the corresponding PID information is not available.

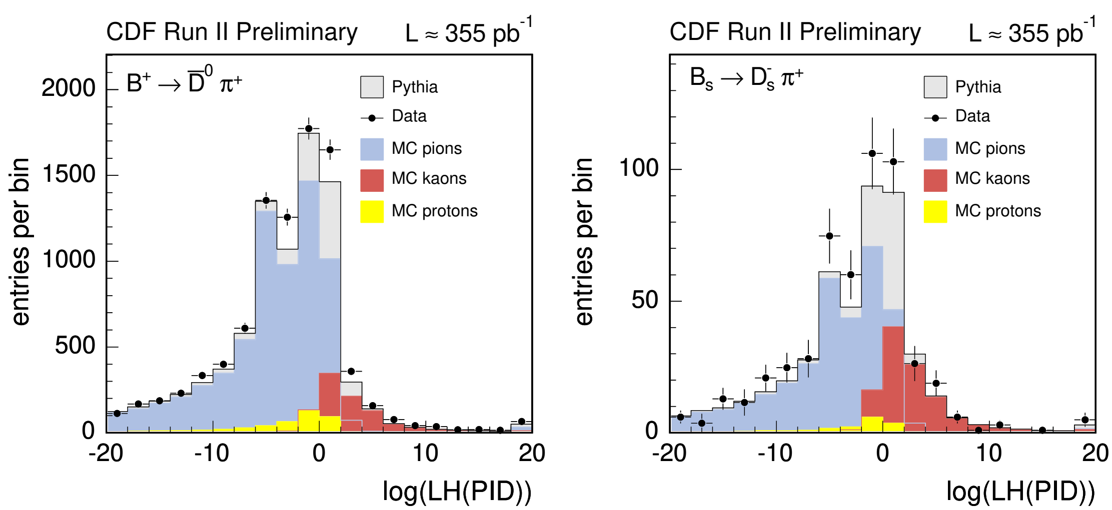


Figure 8-13: Combined likelihood distribution for particle identification, using both dE/dx and TOF information, shown for $B^+ \to \bar{D}^0 \pi^+$ and $B_s \to D_s^- \pi^+$ decays.

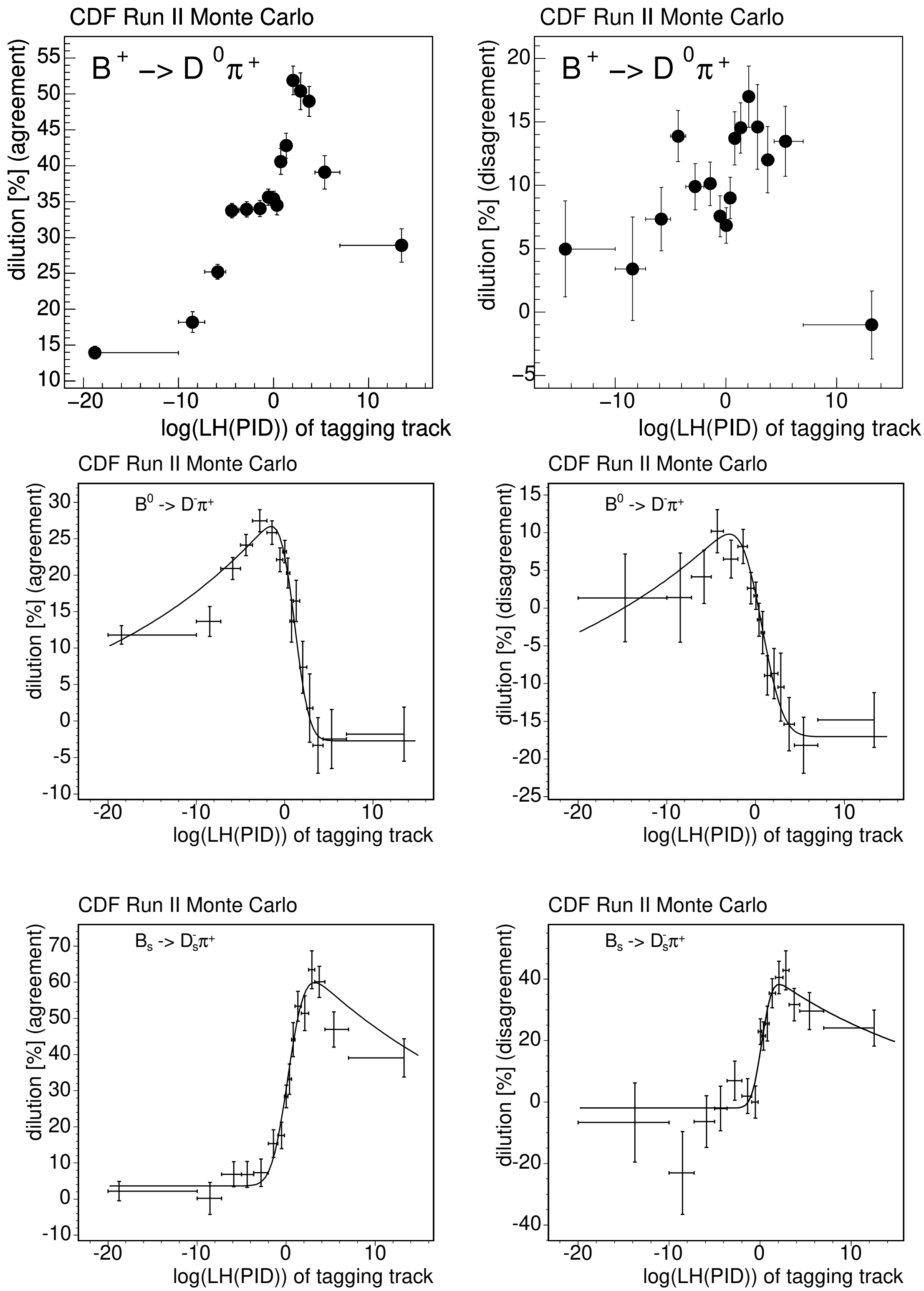


Figure 8-14: Dilution of the PID algorithm as a function of the tagging track PID value, for the classes of agreeing and disagreeing candidate charges.

8.2 Effect on amplitude scan

We apply the SST method to the fully reconstructed B_s modes. The parameterized PID algorithm presented in Section 8.1 is employed to provide per-event predicted dilutions, based on the Monte Carlo calibration achieved in the previous section. We point out that we do not report a full evaluation of the systematic uncertainties associated to the predicted dilution of this algorithm, and its usage here is in this sense preliminary. The analysis procedures are those presented in Chapter 7. The method for combining flavor tagging information is that presented in Section 6.3.3. The signal likelihood factor $L_{\mathcal{D}}$ for the SST is provided by dilution templates obtained from the Monte Carlo samples.

The hadronic scan, where only statistical uncertainties are represented, is shown in Figure 8-15. The increase in sensitivity is very large, as anticipated, following the considerable improvement in tagging power brought about by the inclusion of the same-side tagging method. A value of 18.2 ps^{-1} in sensitivity is achieved, while the lower exclusion limit is 13.9 ps^{-1}. The updated combined world average scan, shown in Figure 8-16, is dominated by these results. The 95% C.L. sensitivity and exclusion limit obtained with this combined scan become 16.6 ps^{-1} and 21.7 ps^{-1}, respectively.

In the probed frequency spectrum all values are practically excluded except in the range of (16.6, 18.2) ps^{-1}. This corresponds to a signature consistent with B_s oscillations at about 17.5 ps^{-1}, where the amplitude value is largest. At that frequency, the measured amplitude significantly deviates from the background hypothesis, $\mathcal{A} = 0$, by 3.2 standard deviations, without accounting for effects of systematic uncertainties.

It is important to throuroughly estimate relevant systematic effects, such as those asscociated with the SST algorithm dilution, and potentially additional sources such as those related to the absolute determination of the proper decay time scale. While the effects of the former are not expected to noticeably affect the ratio $\mathcal{A}/\sigma_{\mathcal{A}}$, the latter can introduce small shifts in Δm_s itself. A complete evaluation of the significance of the observed signature is needed, for accurately quantifying the probability that it be potentially due to a statistical fluctuation of the data.

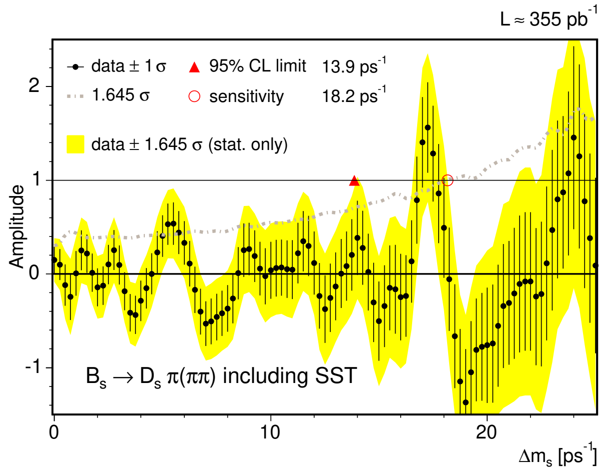


Figure 8-15: Updated amplitude scan for the hadronic modes, obtained with the addition of the parameterized PID SST algorithm; statistical uncertainties only.

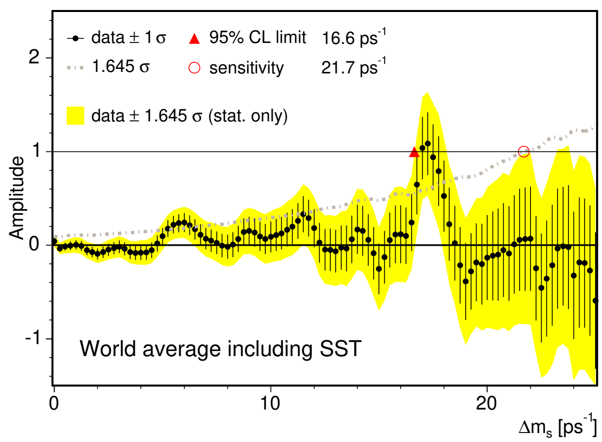


Figure 8-16: Combination of the hadronic scan of Figure 8-15 with the semileptonic and world average scans; statistical uncertainties only.

8.3 Sensitivity projection

The effect on the mixing analysis sensitivity of improvements, such as those related to the increases in the tagging power and in the size of the collected data samples, may be estimated in an analytical fashion. In effect, the measured values of the uncertainties on the amplitude can be approximately described analytically, as derived in Appendix B. Sensitivity estimates may thus be obtained through the re-scaling of the resulting significance curve.

The factors which determine the statistical uncertainty on the amplitude, $\sigma_{\mathcal{A}}$, and therefore the sensitivity of the data, include the following:

- the signal yield S and purity S/B; the dependence is $\sigma_{\mathcal{A}} \propto \frac{\sqrt{S+B}}{S}$. These factors depend on the delivered luminosity, the data acquisition and trigger systems efficiency, and the performance of signal extraction.

- the tagging power, $\epsilon\mathcal{D}^2$; the dependence is given by $\sigma_{\mathcal{A}} \propto \frac{1}{\sqrt{\epsilon\mathcal{D}^2}}$. It corresponds to the performance of the applied flavor tagging methods.

- the B proper decay time resolution, σ_{t_B}; the dependence is $\sigma_{\mathcal{A}} \propto e^{(\sigma_{t_B} w)^2/2}$. This corresponds to the accuracy of the measurement of the B proper decay time t_B, given by the decay length and momentum resolutions. It is determined by the experiment's tracking systems, the track and the vertex reconstruction algorithms, and, for the partially reconstructed modes, by how well the B momentum can be inferred. Due to the exponential dependency referred, this factor becomes increasingly more determining when larger oscillation frequencies w are probed.

The statistical uncertainty expected for the fitted amplitude is accordingly expressed by (B.17). In the case of fully reconstructed decays, this becomes

$$\sigma_{\mathcal{A}} \sim \sqrt{\frac{2}{\epsilon\mathcal{D}^2}} \cdot \frac{\sqrt{S+B}}{S} \cdot \exp\left(\frac{\sigma_t^2 w^2}{2}\right) \qquad \text{(exclusive modes)} ,$$

where the proper decay time resolution is determined by the vertex resolution. For partially reconstructed decays, an additional factor D_p, expressed in (B.7), needs to be considered, due to the uncertainty in the B momentum determination caused by the incomplete reconstruction,

$$\sigma_{\mathcal{A}} \sim \sqrt{\frac{2}{\epsilon\mathcal{D}^2}} \cdot \frac{\sqrt{S+B}}{S} \cdot \exp\left(\frac{\kappa^2\sigma_t^2 w^2}{2}\right) \cdot D_p^{-1}(w, \sigma_p) \qquad \text{(inclusive modes)} ,$$

where σ_p is estimated as the ratio of the root-mean-square deviation to the average of the κ-factor distribution $\mathcal{F}(\kappa)$ (Section 4.3.2), which are about 0.12 and 0.80, respectively.

Representative average values for the hadronic and semileptonic B_s samples are illustrated in Table 8.3. The evaluation of the significance based on the analytical expressions above, however, is performed individually for each decay mode. It takes

	hadronic	semileptonic
signal yield, S	1100	15000
signal-to-noise ratio, S/B	3.4	2.3
vertex resolution, σ_t [fs]	96	157
momentum resolution, σ_p	–	0.15

Table 8.3: Representative values of B_s sample parameters.

into account the distributions of proper decay time resolution σ_t. These distributions are obtained from mass sideband subtracted data, after calibration (Section 4.7), and are illustrated in Figure 4-1. The κ-factor distributions, illustrated in Figure 4-4, are also employed. The different contributions, as given by those equations, are added in quadrature, as it is done when combining the contributions from the various decay channels.

The amplitude uncertainties $\sigma_{\mathcal{A}}(w)$ measured in the hadronic and semileptonic data samples, and from the combined scans, shown in Figures 7-5 – 7-7 obtained with the OST in Chapter 7, are reproduced in Figure 8-17. The corresponding analytical significance curve projections are also represented, and are seen to provide an appropriate description. The analytical curves have an intrinsic smooth behavior, which replaces the statistical fluctuations that characterize the measured values, especially at higher probed frequencies. Also, they further provide an extension beyond the probed spectrum range. The curves predicted with the inclusion of the SST are also overlaid in Figure 8-17. The tagging power values used are those reproduced in Table 8.4. The estimated sensitivity values are given by the intersection of those curves with the horizontal unit line also displayed.

OST $\epsilon\mathcal{D}^2$	0.0155
SST $\epsilon\mathcal{D}^2$	0.04
combined $\epsilon\mathcal{D}^2$	0.0455

Table 8.4: Input tagging effectiveness.

The procedure employed for sensitivity estimation can be summarized as follows:

1. check the analytical description against measured $\sigma_{\mathcal{A}}$ values,
2. re-generate $\sigma_{\mathcal{A}}$ curves according to improvements considered,
3. identify corresponding sensitivity value as intersection point with line of the desired confidence level.

The projected sensitivities are displayed in the graphs of Figure 8-18, as a function of a scale to the yields of the samples which we have analyzed. Sensitivities are evaluated for 95% exclusion, as well as 3σ and 5σ observation. It shows that a potential exclusion of frequency values up to $30\,\mathrm{ps}^{-1}$ or the observation of a potential signal below $20\,\mathrm{ps}^{-1}$ may be achieved in a short time scale by extending the analysis to the increasing data samples being collected. The fully reconstructed decay samples are expected to provide the leading contribution, as anticipated for such large frequencies, while the semileptonic samples are expected to provide a secondary but meaningful contribution to the combined results.

Various further improvements are possible which were not explicitly considered. They can, nevertheless, also be translated into effective yield increases, via the relations expressed above. We mention, to conclude, a few such possible improvements. Several improvements in signal reconstruction are possible. The selection procedure may be implemented based on a multivariate criteria, adapted based on overall event characteristics rather than on rigid cuts on the various quantities. This could be achieved for example through neural network training, or other multivariate discriminating techniques. Additional event quantities such as particle identification information of the tracks may be employed. In this fashion both the signal yield and purity may be enhanced. Additional signal statistics may also be gathered by extracting signal events from other data samples, such as the lepton and displaced track trigger previously described, and others already obtained with current trigger strategies. These strategies can themselves be optimized having directly in mind maximal mixing sensitivity. The hadronic decay samples may be more efficiently explored in the following specific way. The mass region to the left of the signal peak in $B_s \to D_s^- \pi^+$ decays, in Figure 3-8, is populated with several partially reconstructed decays, as indicated in the Monte Carlo composition histogram of Figure 3-5. These contain in particular the signal $B_s \to D_s \rho$ and $B_s \to D_s^* \pi$ decays, where additional π and γ decay products are missed in the reconstruction. The inclusion in the analysis of these partially reconstructed B_s decays presupposes a treatment of the pseudo proper decay time analogous to that developed for the semileptonic modes. The associated κ-factor is however expected to be relatively better localized, close to unity, having a smaller spread than the analogous distributions in semileptonic modes, which follows from the more restrictive phase space available for the unreconstructed decay products. Considerable increases in sensitivity are thus expected. Also on the flavor tagging front improvements are to be expected. A more thorough exploration of the opposite-side is attainable, for example by combining the various lepton, jet and track-PID quantities in a concurring fashion. As already alluded to, further algorithm improvements for same-side tagging are implementable as well.

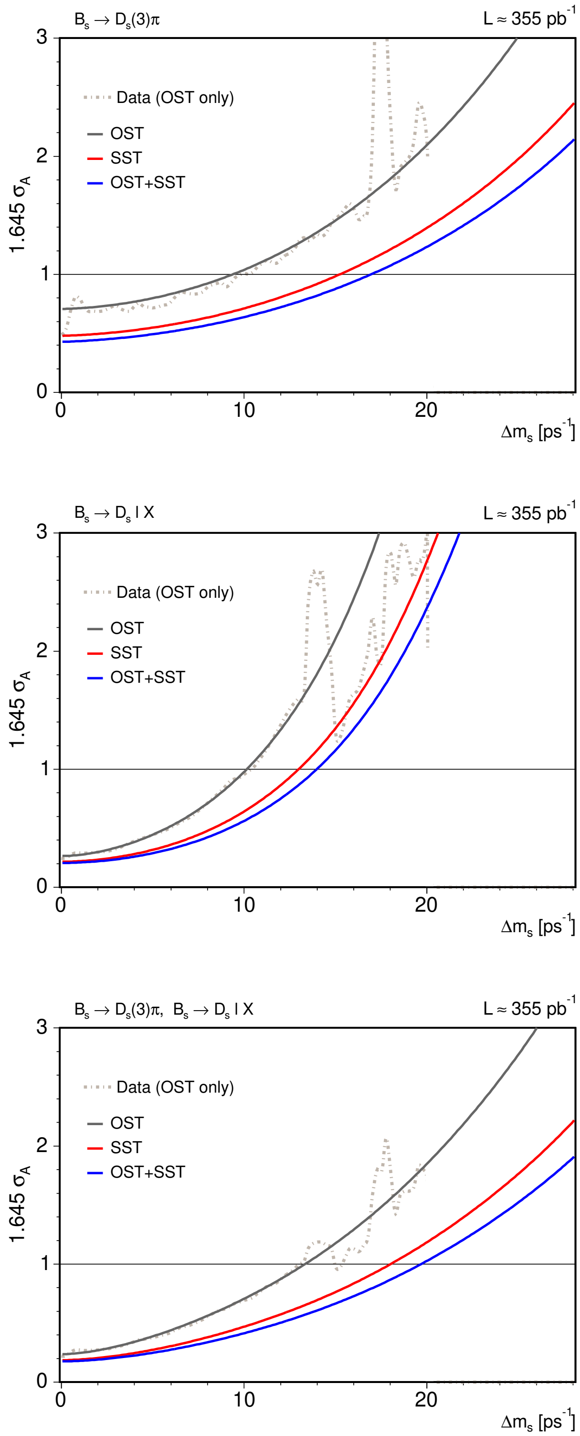


Figure 8-17: Significance curves for the hadronic, semileptonic and combined analyses, with estimated improvements lead by the inclusion of the SST.

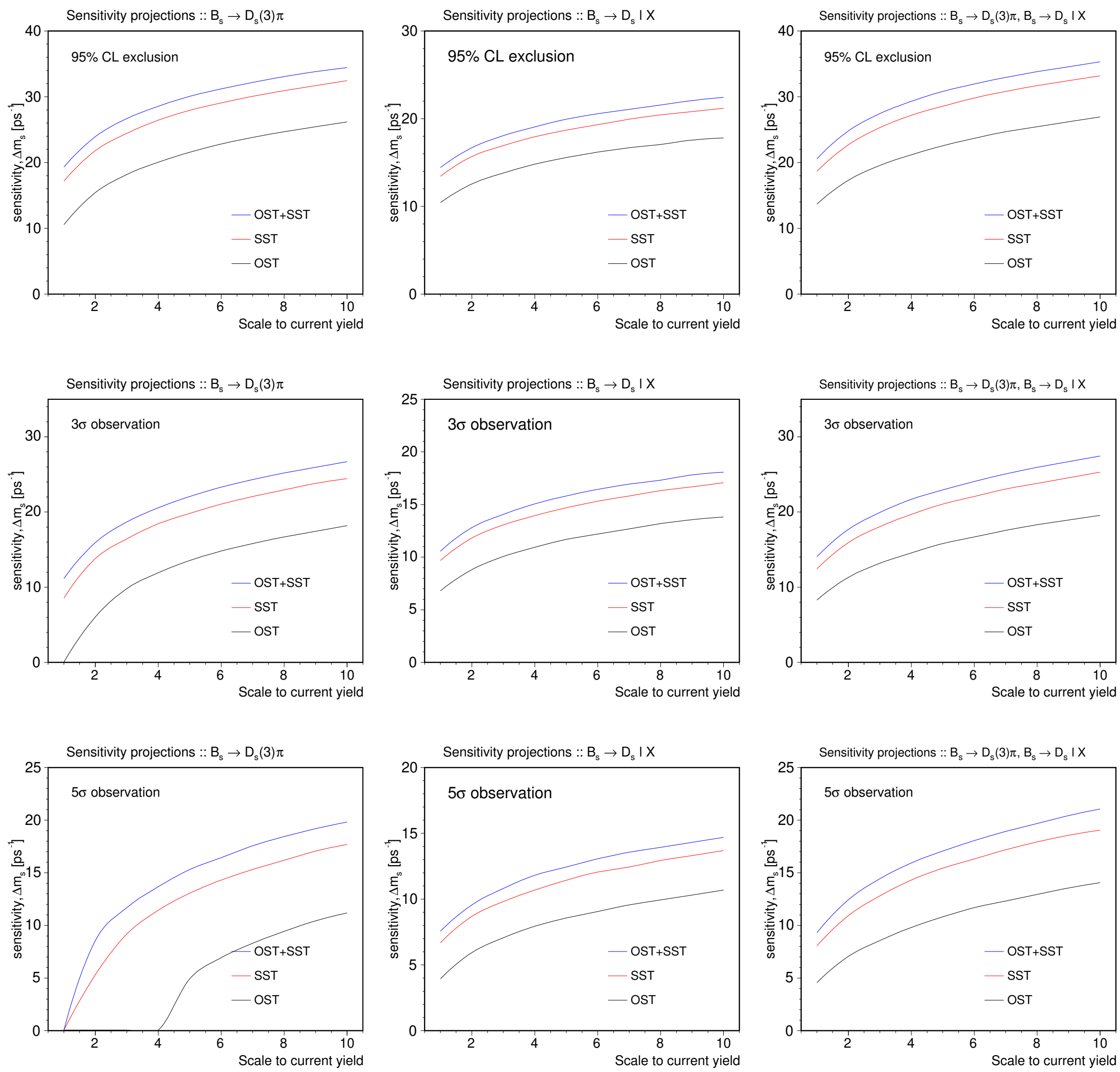


Figure 8-18: Sensitivity projections based on the hadronic, semileptonic and combined analyses; estimates (y-axis) are shown as a function of increasing signal yields, represented relatively to current yields (x-axis), and are provided for the confidence levels of 95% exclusion, 3σ and 5σ observation; the individual contributions from OST and SST are also displayed.

8.4 Résumé

The same-side (kaon) tagging method has been developed and optimized, and its performance established in Monte Carlo samples. The application to the B_s mixing analysis results in remarkable gains in sensitivity, estimates of which have been evaluated in terms of the increasing data samples.

The same-side tagging method is based on flavor charge correlations between the B candidate and tracks found in its vicinity. It further differs from the opposite-side tagging methods in that the tagging performance depends on the B meson species. Unlike those OST methods, the SST performance cannot thus be measured on data samples of B^+ and B^0 decays and transferred directly as input to the B_s analysis. The strategy adopted consists of performing a thorough performance estimation based on Monte Carlo samples. The procedure involves the validation of the Monte Carlo simulation, through comparison against data of various relevant distributions. Systematic variations of fragmentation and other simulation processes should be also implemented. The successful description of the tagging properties achieved in Monte Carlo is ultimately verified in data samples of B^+ and B^0 mesons.

Various algorithms are initially explored. Two such implementations are elected and their performance fully evaluated. One of these, denoted p_L^{rel} algorithm, is based on pure kinematical quantities, and is the chosen method for B^+ and B^0 meson samples. The other, denoted PID algorithm, further uses particle identification information for selecting kaon tracks expected to be associated, from fragmentation, with B_s mesons. The tagging performance is summarized as:

algorithm		tagging power, $\epsilon\mathcal{D}^2$ [%]
p_L^{rel}	B^+	5.4 ± 0.5
	B^0	2.1 ± 0.6
PID	B_s	4.0 ± 0.4

The SST performance achieved is considerably superior to that of the OST methods. A combined tagging power of above 5% for B_s is reached. This determines a dramatic increase in the sensitivity of the samples. A signature consistent with B_s oscillations is verified at 95% C.L. in the frequency range of (16.6, 18.2) ps^{-1}, which lays within the observed sensitivity region. A measured $\mathcal{A}/\sigma_{\mathcal{A}}$ significance of 3.2 is found, not including systematic effects. A thorough evaluation of the probability that the observed signature be potentially due to statistical fluctuation in the data should be performed.

The analysis sensitivity to B_s flavor oscillations is estimated in view of the increasing size of the data samples being collected. A parameterized analytical procedure is employed, and the analysis reach for exclusion and observation of Δm_s is estimated. The potential exclusion of frequency values below 30 ps^{-1} and the precise observation up to 20 ps^{-1} is found to be achievable within a short term by extending the analysis using additional hadronic and semileptonic data collected.

Chapter 9

Constraining the quark-mixing parameters

9.1 Likelihood technique

A global fit to the CKM matrix elements may be performed using several methods [75, 76, 77, 78]. These treat in different ways the available information on the experimental and theoretical uncertainties. We employ a Bayesian approach, following [75], to construct a global inference function, from which probability intervals for the relevant parameters may be derived. In implementing our code we follow the general method documentation [75]. The uncertainties are described in terms of PDFs which quantify the confidence on the values of the involved variables.

We first motivate the procedure for a single constraint; a more general description will follow. The oscillation frequencies of the neutral B meson systems can be related, within the SM, to the CKM parameters $\bar{\rho}$ and $\bar{\eta}$ through an equation of the type

$$(1-\bar{\rho})^2 + \bar{\eta}^2 = c\,, \tag{9.1}$$

where c is a quantity formed of the experimentally measured Δm ($c \propto \Delta m_d, \Delta m_s^{-1}$), and of other theoretically determined parameters.

In the ideal case where c would be perfectly known, the constraint expressed by (9.1) would result in a curve in the $(\bar{\rho}, \bar{\eta})$ plane, *i.e.* a circle of radius $\sqrt{c}$. The PDF describing our beliefs in the $\bar{\rho}$ and $\bar{\eta}$ values would be

$$f(\bar{\rho}, \bar{\eta}|c) = \delta((1-\bar{\rho})^2 + \bar{\eta}^2 - c)\,. \tag{9.2}$$

The points in the circumference would appear as likely. This would remain so in the absence of other experimental piece of information, or theoretical prejudice, which might exclude points outside a determined *physical region*, or in general lead to the assignment of different weights to the various points.

In a realistic case c is not known exactly, the available knowledge about its value being contained in a corresponding PDF, $f(c)$. This way, instead of a single circle, there is in reality an infinite collection of curves, each having a weight $f(c)$. The expected values for $\bar{\rho}$ and $\bar{\eta}$ are thus obtained from

$$f(\bar{\rho},\bar{\eta}) = \int f(\bar{\rho},\bar{\eta}|c)\; f(c)\; dc\,. \tag{9.3}$$

Supposing a best experimental estimate for c would be given by $\hat{c}$, with uncertainty σ_c, and assuming a Gaussian distribution, the previous equation would take the form

$$\begin{aligned} f(\bar{\rho},\bar{\eta}) &= \int \delta((1-\bar{\rho})^2+\bar{\eta}^2-c)\frac{1}{\sqrt{2\pi}\sigma_c}e^{-\frac{1}{2}\left(\frac{c-\hat{c}}{\sigma_c}\right)^2}\; dc \\ &= \frac{1}{\sqrt{2\pi}\sigma_c}e^{-\frac{1}{2}\left(\frac{(1-\bar{\rho})^2-\bar{\eta}^2-\hat{c}}{\sigma_c}\right)^2}\,. \end{aligned} \tag{9.4}$$

In a more general case, c may be formed from various input quantities $\{x_i\}$, denoted by $\mathbf{x}$, and generally described by a joint PDF $f(\mathbf{x})$; when the various x_i can be considered independent, the joint distribution simplifies to $f(\mathbf{x}) \sim \prod_i f(x_i)$. Denoting by $c(\mathbf{x})$ the dependency of c on the input quantities $\mathbf{x}$, $f(c)$ can generally be obtained as

$$f(c) = \int f(\mathbf{x})\delta(c-c(\mathbf{x}))\; d\mathbf{x}\,. \tag{9.5}$$

We describe now more generally the procedure employed in this analysis. It involves the construction of a global inference, $\mathcal{L}$, relating $\bar{\rho}$, $\bar{\eta}$, the constraints $\mathbf{c} = \{c_j\}_{j=1}^{M}$, and the parameters $\mathbf{x} = \{x_i\}_{i=1}^{N}$. The various constraints $\mathbf{c}$, standing for Δm_d, $\Delta m_s/\Delta m_d$, ϵ_K, $|V_{ub}/V_{cd}|$, $\sin 2\beta$, may be expressed as

$$c_j = c_j(\bar{\rho},\bar{\eta};\mathbf{x})\,, \tag{9.6}$$

where the parameters $\mathbf{x}$ denote here all experimentally measured and theoretically calculated quantities on which $\mathbf{c}$ depend. The set of measured constraint values is represented by $\hat{\mathbf{c}} = \{\hat{c}_j\}_{j=1}^{M}$.

Making use of Bayes's theorem,

$$\begin{aligned} \mathcal{L}(\bar{\rho},\bar{\eta},\mathbf{c},\mathbf{x}|\hat{\mathbf{c}}) &\propto f(\hat{\mathbf{c}}|\bar{\rho},\bar{\eta},\mathbf{c},\mathbf{x})\cdot f(\mathbf{c},\mathbf{x},\bar{\rho},\bar{\eta}) \\ &\propto f(\hat{\mathbf{c}}|\mathbf{c})\cdot f(\mathbf{c}|\mathbf{x},\bar{\rho},\bar{\eta})\cdot f(\mathbf{x},\bar{\rho},\bar{\eta}) \\ &\propto f(\hat{\mathbf{c}}|\mathbf{c})\cdot \delta(\mathbf{c}-\mathbf{c}(\mathbf{x},\bar{\rho},\bar{\eta}))\cdot f(\mathbf{x})\cdot f_0(\bar{\rho},\bar{\eta})\,. \end{aligned} \tag{9.7}$$

Here $f_0(\bar{\rho},\bar{\eta})$ is the *prior* distribution for $\bar{\rho},\bar{\eta}$, which we take as uniform; $f(\mathbf{x})$ denotes similarly the prior joint PDF for parameters $\mathbf{x}$. In the derivation we have noted that c_j are unequivocally determined, within the SM, from the values of $\bar{\rho}$, $\bar{\eta}$, and $\mathbf{x}$, and that $\hat{\mathbf{c}}$ depends on those parameters only through $\mathbf{c}$. Considering the independence of the various quantities, (9.7) becomes

$$\mathcal{L}(\bar{\rho},\bar{\eta},\mathbf{x}) \;\propto\; \prod_{j=1,M} f(\hat{c}_j|c_j(\bar{\rho},\bar{\eta},\mathbf{x}))\times\prod_{i=1,N} f_i(x_i)\,, \tag{9.8}$$

where the constraints imposed by the δ-functions in the previous expression are assumed, and the prior, constant f_0 distribution was also omitted.

The relation (9.8) constitutes our sought-after global inference. Within the framework of Bayes statistics, and upon normalization, the left-hand side of the relation is the *posterior* PDF for the argument parameters. The probability distribution for any of the involved parameters can be achieved by integration over the remaining, sometimes also called *nuisance*, quantities.

In a Bayesian approach the various uncertainties are treated in a similar fashion, such that there is no conceptual distinction between those due to random fluctuations in the measurements, those about the parameters of the theory, or those associated to systematics of parameters known but with limited accuracy. Indeed, a systematic uncertainty on a parameter on which the measured constraints depend may be handled by adding the parameter to the collection $\mathbf{x}$.

We consider two models for describing the uncertainties. A Gaussian model is chosen when the uncertainty is dominated by statistical effects, or there are many contributions to the systematic uncertainty, so that the central limit theorem applies. Otherwise, a uniform distribution is used for the uncertainty. When both Gaussian and flat uncertainty components are available for a parameter, the resulting PDF is obtained by convoluting the two distributions. That is, for an observable parameter x of true value $\bar{x}$, with Gaussian and uniform uncertainty components, σ_g, σ_u, one has for the parameter and its PDF, $f(x)$,

$$\begin{aligned} x &= \bar{x} + x_g + x_u \ , \\ f(x) &= \delta(x - \bar{x}) \otimes \mathrm{Gaus}(x|\sigma_g) \otimes \mathrm{Unif}(x|\sigma_f) \ . \end{aligned} \tag{9.9}$$

Besides the constraints themselves, we classify the involved parameters into two categories: *(i) varied*, for which we construct PDFs, and which are what we have been denoting by $\mathbf{x}$ (*e.g.* the top mass); and *(ii) fixed*, which are taken as constant (*e.g.* the W mass).

Joint PDF for $(\bar{\rho}, \bar{\eta})$ and other *posterior* probabilities

The combined probability distribution for $\bar{\rho}$ and $\bar{\eta}$ is obtained by integrating (9.8) over the (here *nuisance*) parameters $\mathbf{x}$,

$$\mathcal{L}(\bar{\rho}, \bar{\eta}) \ \propto \ \int \prod_{j=1,M} f(\hat{c}_j | c_j(\bar{\rho}, \bar{\eta}, \{x_i\})) \times \prod_{i=1,N} f_i(x_i) \ dx_i \ \times f_0(\bar{\rho}, \bar{\eta}) \ . \tag{9.10}$$

The integration is performed using Monte Carlo methods. The normalization can then be trivially performed, and all moments can also be easily computed. This expression shows explicitly that whereas *a priori* all values of $\bar{\rho}$ and $\bar{\eta}$ are equally likely by assumption, *i.e.* $f_0(\bar{\rho}, \bar{\eta}) = \text{const.}$, *a posteriori* the probability clusters in a region of maximal likelihood.

The probability regions in the $(\bar{\rho}, \bar{\eta})$ plane are constructed from the PDF obtained in (9.10). These are called *highest posterior density* regions, and are defined such that

$\mathcal{L}(\bar{\rho},\bar{\eta})$ is higher everywhere inside the region than outside,

$$P_w := \{z = (\bar{\rho},\bar{\eta}) \ : \ \int_{P_w} \mathcal{L}(z)dz = w; \ \mathcal{L}(z') < \min_{P_w}\mathcal{L}(z), \forall_{z' \notin P_w}\} \ . \qquad (9.11)$$

The single parameter PDF can also be obtained in the same fashion. For example, the PDF for $\bar{\rho}$ is obtained as

$$\mathcal{L}(\bar{\rho}) \ \propto \ \int \mathcal{L}(\bar{\rho},\bar{\eta}) \ d\bar{\eta} \ , \qquad (9.12)$$

from which its expected value can be calculated together with the corresponding highest posterior density intervals.

A similar procedure could be in principle used in order to obtain the PDF for other desired parameters. Alternatively, one may use the probability function for transformed variables; *i.e.*, that for $\mathbf{u}(\mathbf{x})$ one has $f(\mathbf{u}) = f(\mathbf{x})|\partial\mathbf{x}/\partial\mathbf{u}|$, where the last factor denotes the Jacobian. This way, the PDF for a parameter x is effectively obtained through a weighted integration over either $\bar{\rho}$ or $\bar{\eta}$; for instance,

$$\mathcal{L}(x) \ \propto \ \int \mathcal{L}(x,\bar{\eta}) \ d\bar{\eta} \ = \ \int \mathcal{L}(\bar{\rho},\bar{\eta}) \ \left|\frac{d\bar{\rho}}{dx}\right| \ d\bar{\eta} \ , \qquad (9.13)$$

where $\mathcal{L}(\bar{\rho},\bar{\eta})$ has been computed in (9.10) above. Besides the probability distribution for $\bar{\rho}$ and $\bar{\eta}$, we are also interested in obtaining the *posterior* distribution for the Δm_s observable itself. The latter may be obtained as

$$\begin{aligned} \mathcal{L}(\Delta m_s) \ &= \ \int \mathcal{L}(\bar{\rho},\bar{\eta}) \left|\frac{d\bar{\eta}}{d\,\Delta m_s}\right| \ d\bar{\rho} \\ &= \ \int \mathcal{L}(\bar{\rho},\bar{\eta}) \, \frac{1}{\Delta m_s} \frac{(1-\bar{\rho})^2 + \bar{\eta}^2}{2\bar{\eta}} \ d\bar{\rho} \ , \qquad (9.14) \end{aligned}$$

where the Jacobian has been calculated from the constraint (9.20).

9.2 Constraints

9.2.1 Neutral *B* meson mixing

Flavor oscillations in the neutral B meson systems are described within the SM by the electroweak box diagrams of Figure 1-2, which are dominated by t quark exchange. The evaluation of corresponding effective Hamiltonian matrix elements, following the formalism presented in Appendix A, leads to the following expression for the oscillation frequencies (with $q = s, d$)

$$\Delta m_q = \frac{G_F^2}{6\pi^2} m_W^2 \eta_{B_q} m_{B_q} B_{B_q} f_{B_q}^2 S_0(x_t) |V_{tq}^* V_{tb}|^2 \ . \qquad (9.15)$$

Here G_F is the Fermi constant; η_{B_q} is a QCD correction factor calculated in NLO; m_{B_q} and m_W are the B_q meson and W boson masses. The Inami-Lim function is given by

$$S_0(x_t) \;=\; x_t\left[\frac{1}{4}+\frac{9}{4}\frac{1}{1-x_t}-\frac{3}{2}\frac{1}{(1-x_t)^2}\right]-\frac{3}{2}\left[\frac{x_t}{1-x_t}\right]^3\ln x_t, \tag{9.16}$$

which describes the $|\Delta B| = 2$ transition amplitude in the absence of strong interaction, with $x_t \equiv m_t^2/m_W^2$ denoting the ratio of the t quark and W boson masses.

The dominant uncertainties in (9.15) come from the evaluation of the hadronic quantities: the B meson decay constant, f_{B_q}, and the bag factor, B_{B_q}, which parameterize the value of the hadronic matrix element (A.32). The ratio of these hadronic quantities for the two systems,

$$\xi_\Delta = \frac{f_{B_s}\sqrt{B_{B_s}}}{f_{B_d}\sqrt{B_{B_d}}}\,, \tag{9.17}$$

is more accurately obtained from lattice QCD calculations [12].

$\boldsymbol{\Delta m_d}$ constraint

Expressing (9.15) above, for the B^0 case, in terms of the Wolfenstein parameters, we obtain

$$\Delta m_d = \frac{G_F^2 m_W^2}{6\pi^2}A^2\lambda^6[(1-\bar\rho)^2+\bar\eta^2]m_{B_d}\frac{f_{B_s}^2 B_{B_s}}{\xi_\Delta^2}\eta_{B_d}S(x_t)\;. \tag{9.18}$$

Note that ξ_Δ and $f_{B_s}\sqrt{B_{B_s}}$ are used instead of $f_{B_d}\sqrt{B_{B_d}}$, which renders the constraint more effective [75]. in view of current parameter uncertainties.

The quantities with dominant uncertainties in (9.18) are $f_{B_s}\sqrt{B_{B_s}}$, ξ_Δ, A and λ, which are varied parameters of the fit. A Gaussian constraint is implemented in the global likelihood,

$$e^{-\frac{1}{2}\left(\frac{\Delta m_d-\widehat{\Delta m}_d}{\sigma_{\Delta m_d}}\right)^2}\,, \tag{9.19}$$

where Δm_d is provided by the r.h.s. of (9.18), while $\widehat{\Delta m}_d$ and $\sigma_{\Delta m_d}$ denote the experimentally measured values.

$\boldsymbol{\Delta m_s}$ constraint

A more powerful constraint of the side $|V_{td}|/(\lambda|V_{cb}|)$ of the unitarity triangle is obtained from the ratio of oscillation frequencies. The constraint associated to Δm_s is expressed, from (9.17) and (9.18), as

$$\Delta m_s = \Delta m_d\frac{m_{B_s}}{m_{B_d}}\xi_\Delta^2\frac{(1-\lambda^2/2)^2}{\lambda^2}\frac{1}{(1-\bar\rho)^2+\bar\eta^2}\,, \tag{9.20}$$

where Δm_d is here taken as an experimental input.

The parameters with dominant uncertainties in (9.20) are ξ_Δ, A and λ, which are varied parameters of the fit. The constraint is implemented via the likelihood ratio, $\mathcal{R}$ as given by (9.28), after accessing the experimentally measured amplitude point $(\mathcal{A}, \sigma_\mathcal{A})$ associated to the frequency value obtained by evaluating the r.h.s. of (9.20).

9.2.2 Other constraints and input

$|V_{ub}|/|V_{cb}|$ constraint

The CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ are measured from both inclusive and exclusive semileptonic B decays. $|V_{cb}|$ is more accurately obtained from exclusive $B \to D^{(*)} l \bar{\nu}_l$ and inclusive semileptonic b decays to charm. The extraction of $|V_{ub}|$ from charmless semileptonic inclusive decays, $B \to X_u l \bar{\nu}_l$, is performed but not without complications. In effect, experimental techniques aimed at suppressing the large $B \to X_c l \bar{\nu}_l$ background result in the introduction of additional theoretical uncertainties. Among the determinations from exclusive decays, such as $B \to \pi l \bar{\nu}_l$, $B \to \rho l \bar{\nu}_l$, $B \to w l \bar{\nu}_l$, and $B \to \eta l \bar{\nu}_l$, the former is the most advanced, as both experimental and lattice calculations are under best control.

In terms of the re-scaled Wolfenstein parameters, the ratio $|V_{ub}|/|V_{cb}|$ is expressed as the constraint

$$\left| \frac{V_{ub}}{V_{cb}} \right| = \frac{\lambda}{1 - \lambda^2/2} \sqrt{\bar{\rho}^2 + \bar{\eta}^2} \, . \tag{9.21}$$

Both Gaussian and flat uncertainties are computed, and a corresponding convoluted PDF is employed in the implementation of the constraint.

$|\epsilon_K|$ constraint

The kaon systems provide various possible constraints, from mixing, CP violation, and rare decays. Only indirect CP violation is used in the CKM fit, since the corresponding matrix elements can be obtained by lattice QCD with controlled systematic uncertainties. The measurement of indirect CP violation in the neutral K system is expressed through the parameter ϵ_K.

In terms of the Wolfenstein parameters, $|\epsilon_K|$ is given by

$$\begin{aligned} |\epsilon_K| &= \frac{G_F^2 f_K^2 m_K m_W^2}{6\sqrt{2}\pi^2 \Delta m_K} B_K A^2 \lambda^6 \bar{\eta} [-\eta_1 x_c + A^2 \lambda^4 \left(1 - \bar{\rho} - \left(\bar{\rho}^2 + \bar{\eta}^2\right) \lambda^2\right) \eta_2 S_0(x_t) \\ &+ \eta_3 S_0(x_c, x_t)] \, . \end{aligned} \tag{9.22}$$

The short distance QCD corrections are codified in the coefficients η_1, η_2 and η_3, and are functions of the charm and top quark masses, and of the QCD scale parameter Λ_{QCD}; the η_i have been calculated in next to leading order (NLO) QCD. The Inami-Lim functions, which describe the $|\Delta S| = 2$ transition amplitude in the absence of

fixed parameters	
G_F $[\mathrm{GeV}^{-2}\hbar^3\mathrm{c}^3]$	1.16637×10^{-5}
M_W $[\mathrm{GeV/c}^2]$	80.425 ± 0.038
f_K $[\mathrm{GeV}]$	0.1598 ± 0.0015
m_K $[\mathrm{GeV/c}^2]$	0.49765 ± 0.00002
Δm_K $[\mathrm{GeV/c}^2]$	$(3.483 \pm 0.006) \times 10^{-15}$
η_{tt}	0.574 ± 0.004
m_{B^0} $[\mathrm{GeV/c}^2]$	5.2794 ± 0.0005
η_B	0.55 ± 0.007
m_{B_s} $[\mathrm{GeV/c}^2]$	5.3696 ± 0.0024

Table 9.1: Input values of *fixed* parameters.

strong interactions, are given by

$$\begin{aligned} S_0(x_t) &= x_t \left[\frac{1}{4} + \frac{9}{4}\frac{1}{1-x_t} - \frac{3}{2}\frac{1}{(1-x_t)^2}\right] - \frac{3}{2}\left[\frac{x_t}{1-x_t}\right]^3 \ln x_t\,, \qquad (9.23) \\ S_0(x_c,x_t) &= -x_c \ln x_c + x_c \left[\frac{x_t^2 - 8x_t + 4}{4(1-x_t)^2}\ln x_t + \frac{3}{4}\frac{x_t}{x_t - 1}\right], \quad x_q \equiv \frac{m_q^2}{m_W^2}\,. \end{aligned}$$

The parameters with dominant uncertainties are B_K, η_1, η_3, m_c and m_t. A Gaussian constraint is implemented for $|\epsilon_K|$.

$\sin 2\beta$ constraint

A direct determination of the angles of the unitarity triangle can be achieved via measurements of CP asymmetries in various B decays. In $b \to c\bar{c}s$ transitions the time dependent CP violation parameters measured from the interference between decays with and without mixing coincide with $\sin 2\beta$ to a very good approximation. The theoretically cleanest cases are the $B^0 \to J/\psi K_{S,L}$ decays, while the world average uses additional measurements involving $\psi(2S)K_S$, $\chi_{c1}K_S$ and $J/\psi K^{*0}$ ($K^{*0} \to K_S\pi^0$) final states. In $b \to c\bar{c}d$ decays, such as $B^0 \to J/\psi\pi^0$ and $B^0 \to D^{(*)}D^{(*)}$, unknown contributions from penguin-type diagrams compromise the clean extraction of $\sin 2\beta$, and consequently are not taken into account in its average used for the CKM fit.

In terms of the re-scaled Wolfenstein parameters the constraint is expressed as

$$\sin 2\beta = \frac{2\bar{\eta}(1-\bar{\rho})}{\bar{\eta}^2 + (1-\bar{\rho})^2}\,, \qquad (9.24)$$

being free of hadronic uncertainties. A Gaussian constraint is implemented for $\sin 2\beta$.

The input values employed in the fit, for both fixed and varied parameters, are specified in Tables 9.1 and 9.2.

varied parameters	
$\lvert V_{cb}\rvert_{\text{incl.}}$	$(41.6 \pm 0.7) \times 10^{-3}$
$\lvert V_{cb}\rvert_{\text{excl.}}$	$(41.3 \pm 1.0 \pm 1.8) \times 10^{-3}$
$\lvert V_{ub}\rvert_{\text{incl.}}$	$(44.9 \pm 3.3) \times 10^{-4}$
$\lvert V_{ub}\rvert_{\text{excl.}}$	$(35.0 \pm 4.0) \times 10^{-4}$
$\lvert V_{us}\rvert$	0.2258 ± 0.0014
$\lvert\epsilon_K\rvert$	$(2.280 \pm 0.013) \times 10^{-3}$
B_K	$0.79 \pm 0.04 \pm 0.09$
$\overline{m}_c$ [GeV/c^2]	1.3 ± 0.1
$\overline{m}_t$ [GeV/c^2]	163.8 ± 3.2
η_{cc}	1.38 ± 0.53
η_{ct}	0.47 ± 0.04
$\sin 2\beta$	0.675 ± 0.026
Δm_d [$\hbar$/c^{2}ps^{-1}]	0.507 ± 0.005
$f_{B_s}\sqrt{B_{B_s}}$ [GeV]	0.262 ± 0.035
ξ_Δ	1.23 ± 0.06
Δm_s	amplitude scan

Table 9.2: Input values of *varied* parameters; the first and when available the second uncertainties are treated respectively as Gaussian and flat errors.

9.3 Making use of the Δm_s amplitude information

A generic constraint is implemented in the fit from the measured central value and uncertainty of the associated experimental observable. For implementing the Δm_s constraint however a separate treatment is motivated by the fact that a definite measurement has not been available. However, rather than using the exclusion region at a given confidence level, a more complete, continuous information about the degree of exclusion is used. Such information is contained in the full amplitude scan. The use of the amplitude measurements in the probed frequency spectrum effectively amounts to employing the Δm_s likelihood profile, as we shall see next. This provides more thorough information than just the frequency point which maximizes the likelihood, along with confidence bounds, and it is applicable regardless of whether or not a well defined, significant maximum has been found.

The measured values of the amplitude and its uncertainty, $\mathcal{A}$ and $\sigma_{\mathcal{A}}$, may be used to derive [31], in the Gaussian approximation, the log-likelihood function, $\Delta \ln \mathcal{L}^\infty(\Delta m_s)$, referenced to its value for an infinite oscillation frequency,

$$\Delta \ln \mathcal{L}^\infty(\Delta m_s) \;=\; \ln \mathcal{L}(\infty) - \ln \mathcal{L}(\Delta m_s) \;=\; \left(\frac{1}{2} - \mathcal{A}\right)\frac{1}{\sigma_{\mathcal{A}}^2}\,. \qquad (9.25)$$

The expected average log-likelihood value for the cases where the probe frequency corresponds to the true oscillation frequency of the system (*mixing* case) or is far from

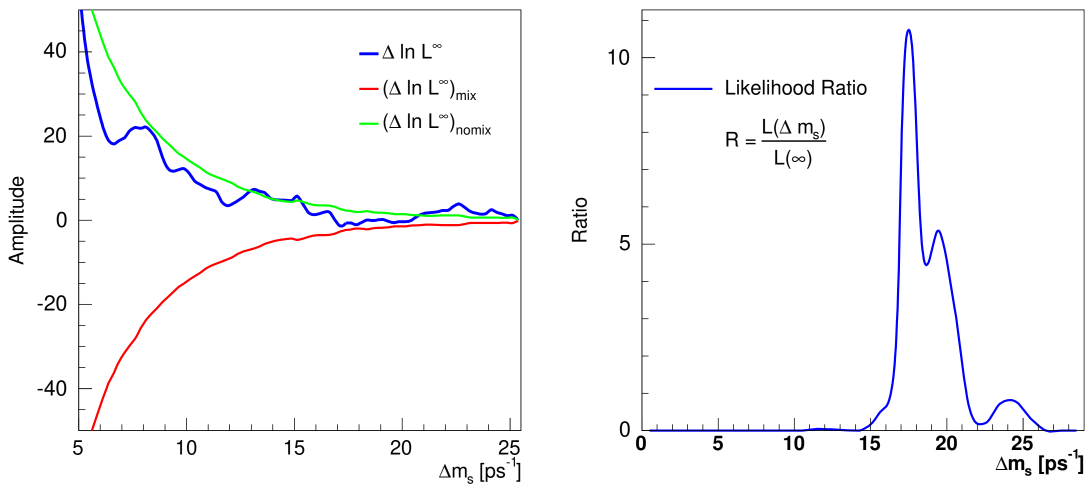


Figure 9-1: Delta log-likelihood and likelihood ratio distributions for Δm_s (obtained from amplitude scan of Figure 7-8).

it (*no-mixing* case), which are characterized respectively by unit and null expected amplitude values, are given by

$$\Delta \ln \mathcal{L}^{\infty}(\Delta m_s)_{\rm mix} \;=\; -\frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2}\,, \tag{9.26}$$

$$\Delta \ln \mathcal{L}^{\infty}(\Delta m_s)_{\rm nomix} \;=\; +\frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2}\,. \tag{9.27}$$

The log-likelihood difference, according to the central limit theorem of likelihood theory, is χ^2-distributed, $\Delta \ln \mathcal{L} = \frac{1}{2}\chi^2$. We therefore translate the amplitude scan into the likelihood ratio

$$\mathcal{R}(\Delta m_s) \;=\; e^{-\Delta \ln \mathcal{L}^{\infty}(\Delta m_s)} \;=\; \frac{\mathcal{L}(\Delta m_s)}{\mathcal{L}(\infty)} \;=\; e^{-\frac{\frac{1}{2}-\mathcal{A}(\Delta m_s)}{\sigma_{\mathcal{A}}^2(\Delta m_s)}}\,. \tag{9.28}$$

We re-state that the exponent in (9.28) corresponds to the χ^2, or log-likelihood, difference between the cases where an oscillation signal is present and absent, for which the expected amplitude value is 1 and 0, respectively. Hypotheses for Δm_s associated with larger $\mathcal{A}$ values and smaller uncertainties $\sigma_{\mathcal{A}}$ in the scan contribute a larger weight in the fit.

The likelihood ratio (9.28) constitutes the Δm_s input to the CKM fit. This is illustrated in Figure 9-1 as obtained from the amplitude scan of Figure 7-8. The corresponding graphs representing (9.26) and (9.27) are also shown.

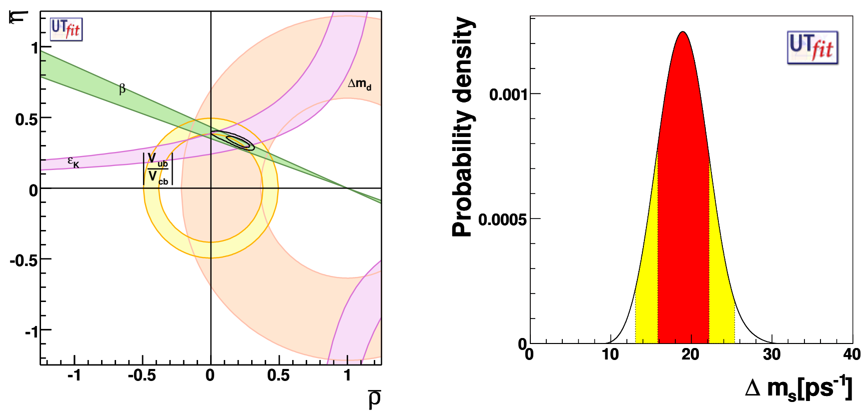


Figure 9-2: Posterior probability distributions in the $(\bar{\rho}, \bar{\eta})$ plane (left) and for Δm_s (right), without using the Δm_s constraint [79].

Extending the amplitude spectrum

The amplitude fits are commonly performed for Δm_s values lower than a given threshold, beyond which the fit behavior may become unstable. On the other hand, in the framework of the CKM fit, it is in principle desirable to have $\mathcal{R}$ defined for all positive frequency values, which in turn demands for a continuation of the measured amplitude spectrum.

The extrapolation of the value of $\sigma_{\mathcal{A}}$ may be achieved through an analytical description of the significance curve. Indeed, (B.17) provides a description of the expected significance dependency on Δm_s as determined by the sample's characteristics, such as signal yields and purity, tagging power, and effective proper time resolutions. For the combined amplitude scans from several measurements, which is the case of the world average, for cases of measurements for which detailed information about the samples are not available, the same general analytical dependence could be still used to extend the measured part of the spectrum by adjusting parameters.

The extrapolation of $\mathcal{A}$ itself in not possible. However, it is here sufficient to note that its expected values lie in the vicinity of either zero or unity as already stated. Therefore, for the unmeasured part of the spectrum, the exponent in (9.28) becomes small, and quickly approaches zero. We take this asymptotic limit for extending $\mathcal{R}$ beyond the experimentally probed frequency region.

9.4 Posterior probability distributions

The global likelihood function (9.10) involves the product of terms corresponding to the individual constraints, which include $|V_{ub}/V_{cb}|$, $|\epsilon_K|$, $\sin 2\beta$, Δm_d and Δm_s. The likelihood may then be projected onto desired subspaces. The constraint on the apex of the unitarity triangle is achieved by performing the fit projection on the $(\bar{\rho}, \bar{\eta})$ space (9.12). As mentioned earlier, the projection over a given parameter or

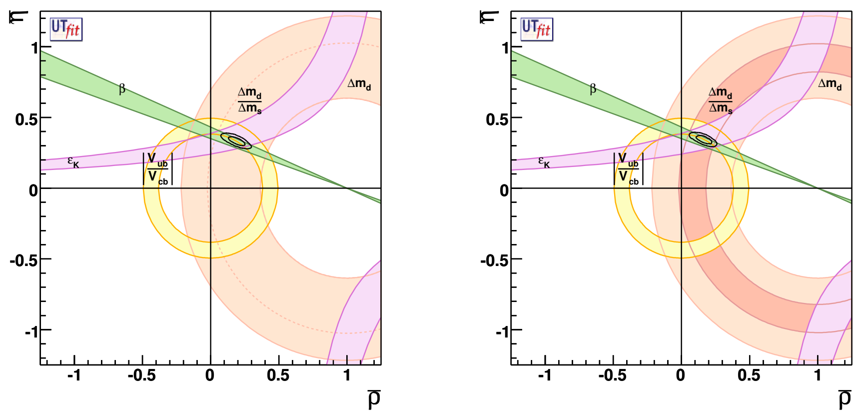


Figure 9-3: Posterior probability distributions in the $(\bar{\rho}, \bar{\eta})$ plane [79], using the Δm_s constraints from the amplitude scan in Figure 7-8 and the preliminary scan of Figure 8-16.

constraint is also achievable as (9.13) and (9.14) denote. The constraining power of a given constraint may be inspected by disabling it in the likelihood construction.

The posterior probability distributions, obtained with the Bayesian framework implementation of the UT fit group [79], are shown in Figures 9-2 and 9-3. The fit is performed without the Δm_s constraint in the former. In the latter, the Δm_s constraint is implemented using the following information: (i) the amplitude scan in Figure 7-8, and (ii) the preliminary amplitude scan of Figure 8-16 (which we emphasize does not contain systematic effects).

The frequency range favored by the data may be perceived from the amplitude scans directly. The likelihood ratio provides perhaps a better illustration of the information contained in the scan. In Figure 9-1, the values of about 17.5 ps^{-1} have the highest likelihood, but still with a considerable spread. The theory expectation, obtained independently of the experimental Δm_s data, is shown in Figure 9-2, and is consistent with the spectrum range favored by the data. The highest posterior density is better localized in the unitary plane once the Δm_s constraint is included in the fit. This is shown by the left plot of Figure 9-3; this is more pronounced in the right hand side plot, which however is preliminary as no systematic effects have been evaluated for the Δm_s input.

A study of the impact of the Δm_s results on constraining the SM and scenarios beyond the standard model is presented in [80].

9.5 Résumé

We have analyzed the constraining of the CKM matrix parameters achieved with the Δm_s amplitude scan information along with other existing measurements in an inference framework based on Bayesian statistics.

The magnitude of the right side of the UT is poorly known. Despite the high precision of the Δm_d measurement, its constraining power is limited by large theoretical uncertainties. A more powerful constraint is obtained with the inclusion of information about Δm_s, as the systematic uncertainty on the theoretical quantities involved in the ratio of the two oscillation frequencies is considerably reduced.

The SM expectation for Δm_s is inferred from a fit performed excluding the Δm_s input. With the currently available precision, the theory expectation agrees with the region favored by the data.

In the future, this consistency will be checked to a greater degree, employing the improved experimental determination of Δm_s. In order to take the best advantage of a coming precise determination of Δm_s, improvements on the lattice calculations will be also required.

Conclusions

In this dissertation we develop an analysis of the time-dependent B_s flavor oscillations using fully and partially reconstructed decay modes. It comprises various ingredients, techniques and calibrations. These lead to the analysis of additional, higher statistics samples of the lighter B meson species, B^+ and B^0. We perform measurements of the B mesons lifetimes and of the oscillation frequency Δm_d, which are found to be compatible with the corresponding world average results. They also serve as further validation of the developed analysis models and techniques.

We report results obtained with a dataset corresponding to an integrated luminosity of 355 pb^{-1}. A combination of opposite side tagging flavor algorithms, with a tagging power of about 1.6 %, is implemented. Using these methods alone, a sensitivity of 13.0 ps^{-1} is achieved, which is found to be competitive with or above any other previously reported single result.

We develop a preliminary same side kaon flavor tagging algorithm, which provides a tagging power of about 4% when applied to the B_s mixing samples. No systematic uncertainties of this algorithm are included here but applying it as is, the sensitivity of the dataset is significantly increased to about 18 ps^{-1}. A signature consistent with B_s oscillations is found within the sensitivity region at about 17.5 ps^{-1}. This is also compatible with the range of the probed frequency spectrum favored by the combination of previous measurements performed by various experiments, and with theory expectations.

The end purpose of the analysis techniques and framework which we have implemented and described in this dissertation is the precise measurement of the Δm_s parameter or the exclusion of the standard model favored region. [1] We show that such goal is within reach by iterating this analysis, including the same side kaon tagging algorithm, with the further accumulated data samples. We show also that the

[1] During the time this thesis was being written, we reported two updated results of the analysis obtained with a larger dataset of 1 fb^{-1}. First, the interval of $(16.96, 17.91)$ ps^{-1} is established at 95% C.L., while the probability that random fluctuations could produce a comparable signal is found to be 0.2% [81]. And finally, with the same 1 fb^{-1} of data and with further improvements in selection, in flavor tagging, and extended signal extraction, a first and definitive observation with a significance superior to 5 standard deviations is achieved [82, 83]. The measured result reported therein is $\Delta m_s = 17.77 \pm 0.10(\mathrm{stat}) \pm 0.07(\mathrm{syst})\ \mathrm{ps}^{-1}$.

increasing knowledge of Δm_s has a considerable impact on the determination of the CKM parameters and on the constraining of the unitarity triangle.

Appendix A

Quantum mechanics of particle-antiparticle oscillations

We present in this section the general quantum mechanical formalism for particle-antiparticle oscillations. It describes a situation where particle and antiparticle are distinguished by an internal quantum number, F, like beauty, strangeness, charm, lepton number, baryon number, *etc.* The charged mesons, for instance B^+/B^- and B_c^+/B_c^-, are not considered, as electric charge violation is not contemplated; unlike those other symmetries, electric charge conservation is protected by gauge, thus exact, symmetry. The π^0 mesons constitute their own antiparticles and are thus also excluded.

The formalism applies to the neutral B meson systems, $B_s\bar{B}_s$ and $B^0\bar{B}^0$, studied in this dissertation, and hold similarly for the $K^0\bar{K}^0$, $D^0\bar{D}^0$, neutron/antineutron and neutrino/antineutrino systems. Such a generic system will be here denoted by $P^0\bar{P}^0$, where P^0 and $\bar{P}^0$ are flavor eigenstates, distinguished merely by the internal quantum number F.

A.1 Effective Hamiltonian

An unstable particle can be described by an Hamiltonian, $\mathcal{H} = m - \frac{i}{2}\Gamma$, through the non-relativistic Schrödinger equation $i\partial_t\psi = \mathcal{H}\psi$. The solution

$$|\psi\rangle_t = e^{-imt}e^{-\frac{1}{2}\Gamma t}|\psi_0\rangle \tag{A.1}$$

reproduces the exponential law of radioactive decay, as $|\langle\psi_0|\psi\rangle_t|^2 = e^{-\Gamma t}$, with lifetime $\tau \equiv 1/\Gamma$. The Hamiltonian is not real (*i.e.* hermitian), since it describes the decay of a particle by its vanishing.

The $P^0\bar{P}^0$ pair can be described similarly as a decaying two-component quantum state. The effective Hamiltonian of the system will be formed of a component $\mathcal{H}_0$ which preserves the characteristic quantum number ($\Delta F = 0$) along with a component inducing $\Delta F \neq 0$ transitions; this can be written as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\Delta F}\ . \tag{A.2}$$

Particle-antiparticle transitions are $\Delta F = 2$ processes which are induced by $\mathcal{H}_{\Delta F}$.

An arbitrary state of the system is represented by a vector in the Hilbert space as

$$|\psi\rangle = a\,|P^0\rangle + b\,|\bar{P}^0\rangle\,. \tag{A.3}$$

Its dynamics, in the Weisskopf-Wigner approximation [84], is determined by the time-dependent Schrödinger equation

$$i\,\frac{d}{dt}\psi = \mathcal{H}\,\psi\,. \tag{A.4}$$

The Hamiltonian contains a dispersive part and an absorptive part, its matrix representation being decomposed accordingly as $\mathbf{H}$ =$\mathbf{M}$-$\frac{i}{2}\mathbf{\Gamma}$, with $\mathbf{M}$ and $\mathbf{\Gamma}$ complex hermitian matrices. Working in the flavor basis, (A.4) may be expressed as

$$i\,\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} m - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^{*} - \frac{i}{2}\Gamma_{12}^{*} & m - \frac{i}{2}\Gamma \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\,. \tag{A.5}$$

The diagonal Hamiltonian matrix elements describe the mass and decay width of the flavor eigenstates. CPT invariance, which is a basic feature of any local quantum field theory, guarantees equality of mass and lifetime of particles and antiparticles, leading to $M_{11} = M_{22} = m$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. The off-diagonal elements are responsible for $P^0\bar{P}^0$ transitions, where M_{12} represents virtual transitions and Γ_{12} represents real transitions through common decay modes.

The non-zero off-diagonal elements of the Hamiltonian matrix imply that the flavor eigenstates differ from the mass eigenstates. The latter will be referred to as *heavy (H)* and *light (L)* mass eigenstates, which are defined as

$$\begin{aligned} |P_L\rangle &= p\,|P^0\rangle + q\,|\bar{P}^0\rangle\,, \\ |P_H\rangle &= p\,|P^0\rangle - q\,|\bar{P}^0\rangle\,, \end{aligned} \tag{A.6}$$

with the complex coefficients p and q obeying the normalization condition $|p|^2 + |q|^2 = 1$. It should be noted that the states in (A.6) do not in general form an orthogonal set, as $\langle P_H|P_L\rangle = |p|^2 - |q|^2$ does not vanish if $|\frac{q}{p}| \neq 1$. In fact, this latter condition would mean that $P^0 \to \bar{P}^0$ and $\bar{P}^0 \to P^0$ transitions would occur at different rates (as it will become apparent below, see (A.18)), corresponding to CP violation in the mixing process. If CP is conserved, which occurs if $|\frac{q}{p}| = 1$ and $\arg\left(\frac{q}{p}\right) = 0$, in which case $q = p = \frac{1}{\sqrt{2}}$ up to some arbitrary phase convention, the mass and CP eigenstates coincide,

$$|P_{L,H}\rangle = \frac{|P^0\rangle \pm |\bar{P}^0\rangle}{\sqrt{2}} \qquad \text{with} \qquad \mathrm{CP}|P_{L,H}\rangle = \pm|P_{L,H}\rangle\,, \tag{A.7}$$

using a convention where $\mathrm{CP}\,|P^0\rangle = |\bar{P}^0\rangle$.

Solving the eigenvalue problem, $\det(\mathcal{H}-\lambda)=0$, we obtain

$$\begin{aligned} \lambda_{H,L} &= m - \frac{i}{2}\Gamma \pm Q\,, \quad \text{with} \qquad \text{(A.8)} \\ Q &\equiv \sqrt{\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)}\,. \end{aligned}$$

The eigenvalues can be further expressed as

$$\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L}\,, \quad \text{with} \qquad \text{(A.9)}$$

$$m_{H,L} = \mathrm{Re}\,(\lambda_{H,L})\,, \qquad \Gamma_{H,L} = -2\,\mathrm{Im}\,(\lambda_{H,L})\,. \qquad \text{(A.10)}$$

It is immediately seen that m and Γ are the average mass $\frac{1}{2}(m_H+m_L)$ and width $\frac{1}{2}(\Gamma_H+\Gamma_L)$. In order to obtain the explicit relationship between the matrix elements, M_{12} and Γ_{12}, and the observables

$$\Delta m \equiv m_H - m_L\,, \qquad \text{(A.11)}$$

$$\Delta\Gamma \equiv \Gamma_L - \Gamma_H\,, \qquad \text{(A.12)}$$

we can form the quantity $(\lambda_H-\lambda_L)^2=(2Q)^2$ to find, from (A.8) above,

$$\begin{aligned} (\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 &= 4|M_{12}|^2 - |\Gamma_{12}|^2\,, \qquad \text{(A.13)} \\ \Delta m\,\Delta\Gamma &= -4\,\mathrm{Re}\,(M_{12}\Gamma_{12}^*)\,, \\ \frac{q}{p} &= -\frac{\Delta m + i\,\Delta\Gamma/2}{2M_{12} - i\,\Gamma_{12}} = -\frac{2M_{12}^* - i\,\Gamma_{12}^*}{\Delta m + i\,\Delta\Gamma/2} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}\,. \end{aligned}$$

Mechanical analogon

The mechanical system formed of two coupled, identical pendula is also characterized by (A.5) [85]. In the absence of the coupling, they would be both described by an oscillation frequency m and a damping constant Γ. The two pendula correspond to the particle P^0 and antiparticle $\bar{P}^0$, in this case governed by $\mathcal{H}_0$ before the perturbation.

Once they are coupled by a spring characterized by elasticity proportional to M_{12} and damping constant Γ_{12} the solutions will comprise two eigenstates: *(i)* corresponding to a long-lived (*i.e.* low damping), light (*i.e.* low frequency) state, where the pendula oscillate in phase; and *(ii)* corresponding to a short-lived (*i.e.* high damping), heavy (*i.e.* high frequency) mode with a phase difference of 180°. The differences in frequency and damping for the two modes are given by $\Delta m = 2M_{12}$ and $\Delta\Gamma = 2\Gamma_{12}$.

As one pendulum is excited, it will transfer its energy to the other, and back, producing a beat with frequency $2\pi f_{12} = \Delta m$. This beat corresponds analogously to the oscillation between a particle P^0 and its antiparticle $\bar{P}^0$, where the mass difference Δm is actually observed as a frequency!

It should be noted, however, that in the mechanical system, unlike the case of the oscillating particles, M_{12} and Γ_{12} are strictly non-negative real numbers, and that the absence of non-trivial phases further prevents the system from simulating CP violation.

A.2 Time evolution

The solution to the system of coupled differential equations (A.5) is decomposed in single particle solutions (A.1) for the mass eigenstates. The latter evolve in time according to the corresponding eigenvalues found in (A.8),

$$|P_{L,H}\rangle_t \;=\; e^{-i\,\lambda_{L,H}\,t}\,|P_{L,H}\rangle \;=\; e^{-i\,m_{L,H}\,t-\frac{1}{2}\Gamma_{L,H}\,t}\,|P_{L,H}\rangle\;. \tag{A.14}$$

An arbitrary initial state of the system may be expressed as a linear combination of either flavor or mass eigenstates,

$$\begin{aligned}|\psi\rangle \;&=\; a_0\,|P^0\rangle + b_0\,|\bar{P}^0\rangle \;=\; \alpha_L\,|P_L\rangle + \alpha_H\,|P_H\rangle\,, \\ &\text{with}\qquad \alpha_{L,H} = \frac{1}{2}\left(\frac{a_0}{p} \pm \frac{b_0}{q}\right)\qquad \text{and} \\ &a_0 \;=\; p\,(\alpha_L + \alpha_H)\,,\qquad b_0 \;=\; q\,(\alpha_L - \alpha_H)\;.\end{aligned} \tag{A.15}$$

Its time evolution follows from (A.14), being given by

$$|\psi\rangle_t \;\equiv\; \mathcal{H}\,|\psi\rangle \;=\; \alpha_L\,|P_L\rangle_t + \alpha_H\,|P_H\rangle_t\;. \tag{A.16}$$

In particular, a state which is initially a pure flavor eigenstate (either $a_0 = 0$ or $b_0 = 0$) will evolve to a state of flavor admixture. Specifically, the time evolution of pure flavor eigenstates, as may be derived from the expressions above, is given by

$$\begin{aligned}|P^0\rangle_t \;&=\; \frac{1}{2p}\,(|P_L\rangle_t + |P_H\rangle_t) \;=\; g_+(t)\,|P^0\rangle + \frac{q}{p}\,g_-(t)\,|\bar{P}^0\rangle\,, \\ |\bar{P}^0\rangle_t \;&=\; \frac{1}{2q}\,(|P_L\rangle_t - |P_H\rangle_t) \;=\; \frac{p}{q}\,g_-(t)\,|P^0\rangle + g_+(t)\,|\bar{P}^0\rangle\,,\end{aligned} \tag{A.17}$$

where

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

The time dependent transition amplitudes squared for the initial states to evolve to a state of the same or the opposite flavor are correspondingly given by

$$\begin{aligned}
\left|\langle P^0|\mathcal{H}|\bar{P}^0\rangle\right|^2 &= \left|\frac{p}{q}\right|^2 |g_-(t)|^2 \ , \\
\left|\langle \bar{P}^0|\mathcal{H}|P^0\rangle\right|^2 &= \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \ , \\
\left|\langle P^0|\mathcal{H}|P^0\rangle\right|^2 &= \left|\langle \bar{P}^0|\mathcal{H}|\bar{P}^0\rangle\right|^2 = |g_+(t)|^2 \ ,
\end{aligned} \tag{A.18}$$

with

$$|g_\pm(t)|^2 = \frac{1}{2} e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma}{2}\, t\right) \pm \cos\left(\Delta m\, t\right)\right] \ .$$

A.3 Neutral $B\bar{B}$ meson systems

Here we address specific characteristics of the two neutral B meson systems – $B^0\bar{B}^0$ and $B_s\bar{B}_s$, which will be denoted $B_q^0\bar{B}_q^0$ ($q = d, s$), and point out appropriate formalism approximations. The symbol B will also be used for denoting the bottom quantum number.

The off-diagonal mass matrix elements are responsible for the $B_q^0\bar{B}_q^0$ transitions. M_{12} represents virtual transitions which provide the dominant contribution to the mixing amplitude. Γ_{12} represents the real transitions through common decay modes. The latter alone implies that $|\Gamma_{12}| \ll \Gamma$. These common decay modes are furthermore Cabibbo (*i.e.* CKM) suppressed. If Γ_{12} were to fully vanish, the relations (A.17) would yield $\Delta m = 2|M_{12}|$ and $\Delta\Gamma = 0$.

The following inequalities hold empirically for both systems

$$|\Gamma_{12}| \ll |M_{12}|, \qquad \Delta\Gamma \ll \Delta m \ , \tag{A.19}$$

such that an expansion in the respective ratios results in a good approximation. An approximate solution to (A.17) is accordingly provided by the following expansions

$$\Delta m = 2|M_{12}|\left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \ , \tag{A.20}$$

$$\Delta\Gamma = 2\,|\Gamma_{12}|\cos\phi_{12}\left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right] \ , \tag{A.21}$$

$$\frac{q}{p} = -e^{-i\phi_M}\left[1 - \frac{1}{2}\left|\frac{\Gamma_{12}}{M_{12}}\right|\sin\phi_{12}\right] + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right) \ , \tag{A.22}$$

$$\text{with} \qquad \phi_M \equiv \arg(M_{12})\,, \quad \phi_{12} \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right) \ .$$

For the B^0 system, $\Delta m_d \approx 0.75\Gamma_d$, while for the B_s system existing experimental bounds give $\Delta m_s \gg \Gamma_s$. The existing experimental bounds on the fractional width differences place $\Delta\Gamma/\Gamma$ below 0.18 and 0.29 (95% C.L.) [1], respectively, for the B^0 and B_s systems, while predicted theory bounds are correspondingly at less than 1% and less than 20% [9].

For both systems, $|\frac{q}{p}| = 1$ holds to a very good approximation. In effect, the difference

$$1 - \left|\frac{q}{p}\right|^2 \approx \mathrm{I}m \left(\frac{\Gamma_{12}}{M_{12}}\right) \tag{A.23}$$

is estimated to be $\sim \mathcal{O}(10^{-3})$ for the B^0 and $\lesssim \mathcal{O}(10^{-4})$ for the B_s systems [1]. For these systems the $\Delta B = 2$ and the Cabibbo favored $\Delta B = 1$ effective operators are CP conserving.

The probability densities, denoted $\mathcal{P}$, for observing the initial (t=0) flavor eigentstates to decay at a later time t with the opposite or the same flavor, following (A.18) in the limit of $|\frac{q}{p}| = 1$ and negligible $\Delta\Gamma/\Gamma$, are given by

$$\begin{aligned}
\mathcal{P}_{B_q^0\to\bar{B}_q^0}(t) &= \mathcal{P}_{\bar{B}_q^0\to B_q^0}(t) = \frac{\Gamma}{2}e^{-\Gamma t}\left[1-\cos\left(\Delta m\, t\right)\right] , && \text{(A.24)}\\
\mathcal{P}_{B_q^0\to B_q^0}(t) &= \mathcal{P}_{\bar{B}_q^0\to\bar{B}_q^0}(t) = \frac{\Gamma}{2}e^{-\Gamma t}\left[1+\cos\left(\Delta m\, t\right)\right] . && \text{(A.25)}\\
& && \text{(A.26)}
\end{aligned}$$

The frequency of flavor transitions corresponds, as explicitly shown, to the mass difference between the two mass eigenvalues of the system — Δm constitutes therefore the target observable of time dependent flavor oscillation measurements.

Standard model predictions

While the description above holds rather generally, in the framework of a specific underlying interaction model the relevant matrix elements may be actually computed. In the Standard Model, $\mathcal{H}_0$ accounts for the strong and electromagnetic Hamiltonians, which have stable flavor eigenstates, while $\mathcal{H}_{\Delta F}$ corresponds to the weak interaction perturbation.

In the B systems the off-diagonal matrix elements are given by the leading term

$$M_{12}^q - \frac{i}{2}\Gamma_{12}^q = \langle B_q^0|\mathcal{H}_{\Delta B=2}|\bar{B}_q^0\rangle , \tag{A.27}$$

which corresponds to the box diagrams in Figure 1-2 containing internal W and up-type quark lines.

The dispersive (mass) part of the box graph is dominated by the t quark contribution. Its evaluation may be expanded as

$$M_{12}^q = \left(\frac{G_F}{4\pi}\right)^2 \left(V_{tb}V_{tq}^*\right)^2 C(\mu)\,\langle B_q^0|Q|\bar{B}_q^0\rangle_{(\mu)} , \tag{A.28}$$

where G_F is the Fermi constant, and V_{ij} denote the CKM matrix elements involved; the factors C and $\langle Q \rangle$ contain respectively the *short distance* (perturbative) and *long distance* (non-perturbative) contributions to the process amplitude, to be evaluated using a consistent renormalization scale μ and scheme.

The Wilson coefficient is given by

$$C(\mu) \quad = \quad m_W^2 \, S_0 \left(\frac{m_t^2}{m_W^2} \right) \eta_{B_q}(\mu) \ , \tag{A.29}$$

where m_W and m_t are the W boson and top quark masses, respectively, and η_{B_q} is a short distance QCD correction. The Inami-Lim function,

$$S_0(x) \quad = \quad x \left[\frac{1}{4} + \frac{9}{4} \frac{1}{1-x} - \frac{3}{2} \frac{1}{(1-x)^2} \right] - \frac{3}{2} \left[\frac{x}{1-x} \right]^3 \ln x \ , \tag{A.30}$$

from the loop [86] is approximated well by $0.784 \, x^{0.76}$. An evaluation of $S(m_t^2/m_W^2)$ and η_B within a consistent renormalization scheme yields $S_0 \approx 2.3$, $\eta_B \approx 0.55$.

Integrating out the internal (top quark and W boson) fields from the box diagram leads to a local operator of the form

$$Q \quad = \quad \bar{q}\gamma_\nu(1-\gamma^5)b \ \cdot \bar{q}\gamma^\nu(1-\gamma^5)b \ . \tag{A.31}$$

The corresponding hadronic matrix element is parameterized as

$$\langle B_q^0 | Q | \bar{B}_q^0 \rangle_{(\mu)} \quad = \quad -\frac{4}{3} m_{B_q} B_{B_q}(\mu) f_{B_q}^2(\mu) \ , \tag{A.32}$$

where m_{B_q} and f_{B_q} are the B_q^0 mass and decay constant, and B_{B_q} is the bag parameter arising from the vacuum insertion approximation. While the scale μ is arbitrary, the physical amplitude (A.28) is independent of both the renormalization scheme and scale.

The absorptive (lifetime) matrix element Γ_{12} involves weak decays common to both B_q^0 and $\bar{B}_q^0$. In contrast to the neutral kaon system, these are only a fraction of the B decays. Those states correspond to on-shell, energetically allowed transitions, and thus the top quark loops do not contribute — the leading internal quarks in the box graph are now charm and up quarks which are considerably lighter than the B mesons. It is the latter which here sets the scale, $\Delta\Gamma \propto m_{B_q}$, which, in view of $\Delta m \propto m_t$, supports the inequality $\Delta\Gamma \ll \Delta m$. The evaluation of the quark box diagram yields in this case

$$\begin{aligned} \Gamma_{12}^q \quad = \quad & -\frac{G_F^2}{8\pi} m_W^2 \eta'_{B_q} m_{B_q} B_{B_q} f_{B_q}^2 \, [\, (V_{tq}^* V_{tb})^2 \\ & + \quad V_{tq}^* V_{tb} V_{cq}^* V_{cb} \, \mathcal{O}\left(\frac{m_c^2}{m_b^2} \right) + \left(V_{cq}^* V_{cb} \right)^2 \, \mathcal{O}\left(\frac{m_c^4}{m_b^4} \right)] \ . \end{aligned} \tag{A.33}$$

The predicted relationship between the width and mass differences stem from the relations (A.28) and (A.33), along with (A.20) and (A.21). The ratio

$$\left|\frac{\Gamma_{12}}{M_{12}}\right| \;\simeq\; \frac{3\pi}{2}\frac{m_b^2}{m_W^2}\frac{1}{S_0(m_t^2/m_W^2)} \;\sim\; \mathcal{O}\left(\frac{m_b^2}{m_t^2}\right) \tag{A.34}$$

is approximately independent of CKM elements, and therefore the same for the B^0 and the B_s systems. The width difference may be estimated [9] from

$$\frac{\Delta\Gamma}{\Delta m} \;\simeq\; \frac{1}{2}\left|\frac{\Gamma_{12}}{M_{12}}\right| \;\simeq\; \mathcal{O}(10^{-3})\,. \tag{A.35}$$

A large Δm corresponds to a large lifetime difference between the heavy and light states. The fractional width difference for the B_s system is estimated to be $\Delta\Gamma_s/\Gamma_s \simeq (1 \sim 20)\%$, depending on the value of Δm_s.

Parameterization of new physics effects

The existence of new physics (NP) may modify the low-energy effective Hamiltonian in several ways. Additional local $\Delta B = 2$ operators may be generated, of the type

$$Q^{NP} \;\sim\; \sum \frac{c_{ij}}{\Lambda^2}\,\bar{q}\Gamma_i b\;\cdot\bar{q}\Gamma_j b\,, \tag{A.36}$$

where c_{ij} are the operators' strength and Λ is the NP scale. New contributions to the Wilson coefficients of the SM operators may arise, inducing modifications to the Inami-Lim functions,

$$S_0 \mapsto S_0 + \delta S^{NP}\,. \tag{A.37}$$

Rather generally, NP may introduce new amplitude contributions to the mixing process, which may be expressed as

$$M_{12}^q \;=\; M_{12}^{q,\,SM} + M_{12}^{q,\,NP} \qquad \text{with} \qquad M_{12}^{q,\,NP} \;=\; \langle B_q^0|\mathcal{H}_{\Delta B=2}^{NP}|\bar{B}_q^0\rangle\,. \tag{A.38}$$

It is usually convenient to parameterize the NP amplitude, in a model-independent fashion, relatively to the SM contribution, giving

$$M_{12}^q \;=\; M_{12}^{q,\,SM}\left(1 + h_q\,e^{2i\sigma_q}\right)\,, \tag{A.39}$$

where h_q (σ_q) quantify the NP contribution to the magnitude (phase) of the $B_q^0\bar{B}_q^0$ mixing process.

Appendix B

Mixing significance

In this section we explore an analytical characterization of the mixing analysis. A powerful mathematical technique for the analysis of periodic signals is the Fourier transform. The amplitude method employed for the search and the study of B_s oscillations is itself motivated and shares several advantages associated with the latter techniques, as well as with the thorough likelihood characterization of the data sample components. We follow and elaborate on the expositions in [31, 87, 88]. The significance of a mixing measurement is derived. Effects of resolution, partial reconstruction and biases on the proper time distribution of the sample are explored. Further and insightful properties can be obtained beyond our brief presentation.

B.1 Likelihood approach

The basic likelihood description in proper decay time space for flavor tagged candidates has the following form

$$f_\xi(t) \;=\; \frac{N}{2}\left(f_s\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t)[1+\xi\mathcal{D}\cos(wt)] + (1-f_s)(1+\xi\mathcal{D}_B)B(t)\right) \tag{B.1}$$

where N stands for the number of tagged events in the sample; f_s the fraction of signal events; $\mathcal{D}$ and $\xi \in \{-1,+1\}$ the flavor tagging dilution and decision; τ and w the lifetime and the oscillation frequency of the B system. The background model contains a possible small flavor tagging asymmetry, which is assumed to be time independent; it is accordingly described by the asymmetry parameter $\mathcal{D}_B$, along with an empirical t distribution, $B(t)$, characterized by a decay constant τ_B.

The flavor asymmetry may be obtained as the difference of the expressions for the two tagging decisions

$$a(t) \;=\; f_+(t) - f_-(t) = N\left(f_s\mathcal{D}\frac{1}{\tau}e^{-\frac{t}{\tau}}\theta(t)\cos(wt) + (1-f_s)\mathcal{D}_B B(t)\right) \,. \tag{B.2}$$

We shall focus in this section on the signal component. We assume the background may be effectively subtracted via additional likelihood factors describing discriminating quantities such as the candidates' mass along with others. Effects of detector resolution, trigger, selection and partial reconstruction are not included in the above expressions.

Detector resolution

The reconstructed proper decay time is given by

$$t = \frac{LM}{p}$$

where L and p are the (transverse) decay distance and momentum, and M the mass of the B meson. The uncertainty is generally given by

$$\sigma_t = \frac{L}{p}\sigma(M) \oplus \frac{M}{p}\sigma(L) \oplus \frac{\sigma(p)}{p}t$$

where σ denotes the partial uncertainties and the $\oplus$ symbol indicates that the terms combine in quadrature. The first, which is proportional to the uncertainty on the B hadron mass is negligible; in fact, the average world value for M is taken rather than the reconstructed candidate's mass. The uncertainty in the decay vertex position is to first order independent of the decay length itself, while the momentum resolution deteriorates with increasing momentum. We assume for the moment that the absolute uncertainty on the decay length and the relative uncertainty on the B momentum are Gaussian.

The smearing effect on the proper decay time distributions (B.1) may be expressed as

$$f_\xi(t) \mapsto \int_{-\infty}^{+\infty} f_\xi(t')\mathcal{R}(t',t)\,dt' \tag{B.3}$$

where $\mathcal{R}$ is the resolution function. The latter can be written as

$$\begin{aligned}\mathcal{R}(t',t) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma(p)} e^{-\frac{(p-p')^2}{2\sigma_p^2}} \frac{1}{\sqrt{2\pi}\sigma(L)} e^{-\frac{(pt-p't')^2}{2M^2\sigma(L)^2}} (\frac{p'}{M})\; dp \\ &\simeq \frac{1}{\sqrt{2\pi(\sigma_l^2+\sigma_p^2t^2)}} e^{-\frac{(t't')^2}{2(\sigma_L^2+\sigma_p^2t^2)}}\end{aligned}$$

where we have assumed $\sigma(p)/p' \ll 1$ and defined

$$\sigma_l \equiv \frac{M}{p'}\sigma(L) \simeq \frac{M}{p}\sigma(L) \qquad \text{and} \qquad \sigma_p \equiv \frac{\sigma(p)}{p'} \simeq \frac{\sigma(p)}{p}\;.$$

We may write

$$\sigma_t = \sigma_l \oplus \sigma_p\, t\;. \tag{B.4}$$

In addition to the resolution smearing, we may include the biasing effects in the proper decay time distribution induced by trigger and signal selection criteria. These are described by the t-efficiency function $\mathcal{E}(t)$, presented in Section 4.3.1. This function is defined in terms of the *reconstructed* proper time, and is thus applied after the detector resolution smearing.

The integration in (B.3), with the biasing effects incorporated, identically to (D.7) and (D.8), for the signal terms, is given by

$$\begin{aligned}
&\int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \left(1 + \xi \mathcal{D} \cos(wt)\right) \mathcal{R}(t', t)\, dt' \cdot \mathcal{E}(t) \\
&= \frac{1}{2\tau} e^{-\frac{1}{\tau}(t - \frac{\sigma_t^2}{2\tau})} e^{-\frac{\sigma_t^2 w^2}{2}} \cdot \mathrm{Re}\, \{e^{-iw(t - \frac{\sigma_t^2}{\tau})}\, \mathrm{Erfc}(\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t \tau} + i\frac{\sigma_t w}{\sqrt{2}})\} \cdot \mathcal{E}(t) \\
&\simeq \frac{1}{\tau} e^{-\frac{t}{\tau}} \left(1 + \xi e^{-\frac{\sigma_t^2 w^2}{2}} \mathcal{D} \cos(wt)\right) \cdot \mathcal{E}(t)\ .
\end{aligned}$$

The approximation holds in the limit that the exponential variation is negligible over the range of the resolution function, *i.e.* $\sigma_t \ll \tau$. A phase shift of the order of $w\sigma^2/\tau$ has also been neglected. The error function turn-on is taken to be superseded by that of the efficiency function. Accordingly, the resolution effects approximately reduce to a re-scaling of the cosine factor – that is to say, to a decrease of the oscillation amplitude, by

$$D_l(w, \sigma_t) \;\; = \;\; e^{-\frac{\sigma_t^2 w^2}{2}}\ . \tag{B.5}$$

In general, the decay distance uncertainty dominates over the momentum uncertainty, and the latter may then be safely neglected.

Partial reconstruction

In partial reconstructed B systems a correcting κ-factor to the proper decay time needs to be considered, as addressed in Section 4.2.3,

$$t \;\; \mapsto \;\; t \cdot \kappa\ .$$

The signal t PDF is modified accordingly,

$$\int \frac{\kappa}{\tau} e^{-\frac{\kappa t}{\tau}} \theta(t) \big(1 + \xi \mathcal{D} \cos(w\kappa t)\big)\, \mathcal{F}(\kappa)\, d\kappa\ .$$

where $\mathcal{F}(\kappa)$ is the the κ-distribution, defined in Section 4.3.2.

This results effectively in the smearing of the proper decay time, and its action may be estimated as part of the resolution function. For fully reconstructed decays $\mathcal{F}(\kappa) = \delta(\kappa - 1)$ and no effect is introduced. For partially reconstructed decays, we use for the current purpose a Gaussian approximation, based on the average $\bar{\kappa}$ and *rms* deviation σ_κ of the actual κ-distribution. For this latter case we have,

$$\begin{aligned} \mathcal{R}(t',t) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_\kappa} e^{-\frac{(\kappa-\bar{\kappa})^2}{2\sigma_{\bar{\kappa}}^2}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t\bar{\kappa}-t'\kappa)^2}{2\kappa^2\sigma_t^2}} \left(\frac{1}{\bar{\kappa}}\right) d\kappa \\ &\simeq \frac{1}{\sqrt{2\pi(\bar{\kappa}^2\sigma_t^2+\sigma_\kappa^2 t'^2)}} e^{-\frac{\bar{\kappa}^2(t-t')^2}{2(\bar{\kappa}^2\sigma_t^2+\sigma_\kappa^2 t'^2)}} \end{aligned} \tag{B.6}$$

Accordingly, the κ-factor effect is translated effectively as a time-dependent resolution, analogous to a direct momentum uncertainty contribution as addressed above. It results in a damping of the oscillation, approximated to the form of (B.5). Integrating over the time dependence we obtain

$$\begin{aligned} D_\kappa(w,\sigma_\kappa) &= \int_0^{+\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} e^{-\frac{w^2(\sigma_\kappa t)^2}{2}} dt \\ &= \sqrt{\pi} Y e^{Y^2} \mathrm{Erfc}(Y) \qquad \text{with} \quad Y \equiv \frac{1}{\sqrt{2}\sigma_\kappa w\tau} \ . \end{aligned} \tag{B.7}$$

B.2 Fourier transform approach

The Fourier Transform (FT) of a function $f(t)$ is defined as

$$\tilde{f}(\nu) = \int_{-\infty}^{+\infty} f(t)\, e^{-i\nu t}\, dt \ . \tag{B.8}$$

For practical numerical calculations the discrete Fourier transform is more suitable. Algorithms such as the Fast Fourier Transform are particularly computationally efficient (specifically, by reducing the number of needed arithmetical operations from $\mathcal{O}(N)$ to $\mathcal{O}(NlogN)$). For a set of measurements $\{t_i\}_{i=1}^N$ the discrete transformation is defined simply as

$$g(\nu) = \sum_{i=1}^{N} e^{-i\nu t_i} \ . \tag{B.9}$$

If the sample measurements $\{t_i\}$ correspond to a random extraction over a probability distribution function $f(t)$, then $g(\nu)$ is on average proportional to the continuous transform of $f(t)$,

$$\langle g(\nu) \rangle = \sum_{i=1}^{N} \int_{-\infty}^{+\infty} f(t_i)\, e^{-i\nu t_i}\, dt_i = N\, \tilde{f}(\nu) \ .$$

The Fourier transform of the mixing PDF expression (B.1) is given by

$$\tilde{f}_\xi(\nu) = \frac{N}{2}\left(f_s\left[\frac{1}{1+i\nu\tau} + \xi\mathcal{D}\frac{1+i\nu\tau}{w^2\tau^2+(1+i\nu\tau)^2}\right] + (1-f_s)(1+\xi\mathcal{D}_B)\widetilde{B}(\nu)\right) \tag{B.10}$$

and its real part by

$$\mathrm{Re}\left\{\tilde{f}_\xi(\nu)\right\} = \frac{N}{2}\big(f_s[\frac{1}{1+\nu^2\tau^2}+\xi\frac{\mathcal{D}}{2}\left(\frac{1}{1+(\nu-w)^2\tau^2}+\frac{1}{1+(\nu+w)^2\tau^2}\right)]$$
$$+(1-f_s)(1+\xi\mathcal{D}_B)\mathrm{Re}\left\{\widetilde{B}(\nu)\right\}\big) \ . \qquad \text{(B.11)}$$

Asymmetry

The difference between the expressions for events tagged as unmixed and mixed is given by

$$\tilde{f}_+(\nu)-\tilde{f}_-(\nu) = N\left(f_s\mathcal{D}\frac{1+i\nu\tau}{w^2\tau^2+(1+i\nu\tau)^2}+(1-f_s)\mathcal{D}_B\widetilde{B}(\nu)\right) \ .$$

Its real part evaluated at the signal oscillation frequency is

$$\Delta(w) \equiv \mathrm{Re}\left\{\tilde{f}_+(w)-\tilde{f}_-(w)\right\} = N\left(f_s\mathcal{D}\frac{1+2w^2\tau^2}{1+4w^2\tau^2}+(1-f_s)\mathcal{D}_B\widetilde{B}(w)\right)$$
$$\simeq \frac{1}{2}Nf_s\mathcal{D} \ . \qquad \text{(B.12)}$$

The approximation is very good for large oscillation frequencies, $w^2\tau^2 \gg 1$. For the background, the following considerations need to be made. In general the background flavor asymmetry is quite small, $|\mathcal{D}_B| \ll 1$. Furthermore, long lived background components have a negligible contribution, as long as $w^2\tau_B^2 \gg 1$. For less common situations, where sizable short-lived or prompt background components present a large asymmetry, the background cancellation in (B.12) becomes incomplete, and deviations are expected, especially for lower frequencies. Possible uncertainties on the description of such components require that corresponding systematic uncertainties be adequately evaluated.

Detector resolution

The effect of the proper decay time resolution was described by the convolution with the resolution function. Making use of the convolution theorem, $\widetilde{f \otimes g} = \tilde{f}\cdot\tilde{g}$, this is expressed as a multiplicative factor corresponding to the FT of the Gaussian resolution function,

$$D_l(\nu,\sigma_t) = e^{-\frac{\sigma_t^2\nu^2}{2}} \ .$$

Partial reconstruction

The effects of partial momentum reconstruction are translated into frequency smearing of the FT, for the signal component, via the κ-factor distribution,

$$\begin{aligned}
\tilde{f}_\xi(\nu) &= \frac{N}{2}\big(f_s \int_0^\infty [\frac{1}{1+i\nu\tau/\kappa} + \xi\mathcal{D}\frac{1+i\nu\tau/\kappa}{w^2\tau^2+(1+i\nu\tau/\kappa)^2}]\,\mathcal{F}(\kappa)\,d\kappa \\
&\quad + \ (1-f_s)(1+\xi\mathcal{D}_B)\widetilde{B}(\nu)\big)\ .
\end{aligned}$$

The asymmetry becomes

$$\Delta(w) \ = \ Nf_s\mathcal{D}\int_0^\infty \frac{1+w^2\tau^2(1+1/\kappa^2)}{1+2w^2\tau^2(1+1/\kappa^2)+w^4\tau^4(1-1/\kappa^2)^2}\,\mathcal{F}(\kappa)\,d\kappa$$

In case the κ-distribution is well localized around unity, $\mathcal{F}(\kappa)\sim\delta(\kappa-1)$, this expression reduces to (B.12). As its mean shifts to lower values, becoming broader, the term $(1-1/\kappa^2)$ stops being suppressed, and a damping of the signal peak is induced, more prominently so at larger oscillation frequencies.

The smearing effect caused by partial momentum reconstruction is more readily estimated by considering its effective contribution to the resolution of the proper decay time of the B meson. This contribution (B.6) has a time dependence, and the calculation cannot be simply performed using the convolution theorem as above (B.13). Instead one has

$$\begin{aligned}
\widetilde{a\otimes g_p}\,(\nu) &= \int_0^{+\infty}\frac{1}{\tau}e^{-\frac{t'}{\tau}}\cos(wt')\left(\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_p t'}e^{-\frac{1}{2}\left(\frac{t-t'}{\sigma_p t'}\right)^2}e^{-i\nu t}\,dt\right)dt' \\
&= \int_0^{+\infty}\frac{1}{\tau}e^{-\frac{t}{\tau}}\cos(wt)e^{-\frac{\nu^2\sigma_p^2t^2}{2}}e^{-i\nu t}\,dt\ .
\end{aligned}$$

This determines a reduction of the peak amplitude by a factor approximately equal to $D_p(w,\sigma_p)$ given in (B.7).

Proper time bias

The sculpting of the proper decay time distribution induced by possible trigger and selection criteria is accounted for by the t-efficiency function. The later is to be applied to the reconstructed proper time, thus after the resolution effects have been incorporated. This introduces complexity in the computation to be performed in frequency space. However, these are identical to those appearing in the context of PDF normalization in time space, as addressed in Appendix D. Alternatively, however, the FT can be evaluated using the inverse of the convolution theorem, which gives

$$\widetilde{f\mathcal{E}} = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{f}(\nu-\lambda)\,\widetilde{\mathcal{E}}(\lambda)\,d\lambda\ .$$

The FT of the indicated typical form of the t-efficiency function is given by

$$\mathcal{E}(t) = \sum_n a_n (t-\zeta_n)^n e^{-\frac{t}{\tau_n}} \theta(t-\zeta_n) \qquad \mapsto \qquad \widetilde{\mathcal{E}}(\nu) = \sum_n a_n \frac{n!\, \tau_n^{n+1}}{(1+i\nu\tau_n)^{n+1}} e^{-(1+i\nu\tau_n)\frac{\zeta_n}{\tau_n}} \ .$$

where the integral computation is reduced to the Gamma function.

The qualitative effect of a proper time bias can be obtained using the simplest form for the t-efficiency curve, namely a direct cut $t > \zeta$ as in (4.17). In the absence of resolution effects,

$$\begin{aligned} &\mathrm{Re}\,\{\int_0^{+\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} \cos(wt)\theta(t-\zeta) e^{-i\nu t}\} \simeq \mathrm{Re}\,\{e^{-\frac{\zeta}{\tau}} \frac{e^{-i(\nu-w)\zeta}}{1+i(\nu-w)\zeta}\} \\ &\simeq \frac{e^{-\frac{\zeta}{\tau}}}{1+(\nu-w)^2\tau^2} \left(\cos\left((\nu-w)\zeta\right) - \sin\left((\nu-w)\zeta\right)\cdot(\nu-w)\tau\right) \ . \end{aligned} \tag{B.13}$$

The amplitude of the signal peak decreases by a factor of $e^{-\frac{\zeta}{\tau}}$, as observed for $\nu = w$. A less anticipated effect, perhaps, is the modification induced in the FT profile. If no bias is present, *i.e.* $\zeta = 0$, the shape is that of the standard Breit-Wigner, $1/(1+(\nu - w)^2\tau^2)$. For non-negligible bias, $\zeta > 0$, however, a sinusoidal behavior is introduced. It results in particular, for realistic parameter values, in a small undershooting of the FT (or Amplitude) profile immediately before and after the signal peak. Note also that the sinusoidal terms also cause the narrowing of the signal peak. Further considerations about these effects are provided in Appendix F.

Statistical noise

Statistical fluctuations of the sample generate noise in the FT. While the peak size of a signal is given by (B.12), the determination of its statistical significance requires that an evaluation of the noise fluctuations be performed.

The real part of the discrete FT (B.9) is

$$\mathrm{Re}\,\{g(\nu)\} \ = \ \sum_j \cos(\nu t_j) \ .$$

Its variance is evaluated as

$$\begin{aligned}
\sigma^2\big(\mathrm{Re}\,\{g(\nu)\}\big) &= \langle \mathrm{Re}\,\{g(\nu)\}^2\rangle - \langle \mathrm{Re}\,\{g(\nu)\}\rangle^2 \\
&= \sum_{j,k\neq j}\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}\rangle \cdot \langle \mathrm{Re}\,\{e^{-i\nu t_k}\}\rangle + \sum_{j,k=j}\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}\cdot \mathrm{Re}\,\{e^{-i\nu t_k}\}\rangle \\
&\quad -\sum_{j,k}\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}\rangle \cdot \langle \mathrm{Re}\,\{e^{-i\nu t_k}\}\rangle \\
&= (N^2-N)\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}\rangle^2 + N\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}^2\rangle - N^2\langle \mathrm{Re}\,\{e^{-i\nu t_j}\}\rangle^2 \\
&= \frac{N}{2}\left(1-2\left(\frac{\mathrm{Re}\,\{\tilde{f}(\nu)\}}{\tilde{f}(0)}\right)^2 + \frac{\mathrm{Re}\,\{\tilde{f}(2\nu)\}}{\tilde{f}(0)}\right)\,, \qquad \text{(B.14)}
\end{aligned}$$

where we have used

$$\begin{aligned}
\langle \mathrm{Re}\,\{e^{-i\nu t}\}\rangle &= \frac{\mathrm{Re}\,\{\int e^{-i\nu t} f(t)\,dt\}}{\int f(t)\,dt} = \frac{\mathrm{Re}\,\{\tilde{f}(\nu)\}}{\tilde{f}(0)} \\
\langle \mathrm{Re}\,\{e^{-i\nu t}\}^2\rangle &= \frac{\int \cos^2(\nu t) f(t)\,dt}{\int f(t)\,dt} = \frac{\int (1+\cos(2\nu t))\,f(t)\,dt}{2\,\tilde{f}(0)} = \frac{1}{2}+\frac{1}{2}\frac{\mathrm{Re}\,\{\tilde{f}(2\nu)\}}{\tilde{f}(0)}\,.
\end{aligned}$$

We now apply the results of (B.14) to the PDF of (B.10). For frequencies in the vicinity of the true oscillation frequency w, the real part of the Fourier transform (B.11) and its normalization become

$$\begin{aligned}
\mathrm{Re}\,\{\tilde{f}_\xi(w)\} &= \frac{N}{2}\big(f_s[\frac{1}{1+w^2\tau^2} + \xi\mathcal{D}\frac{1+2w^2\tau^2}{1+4w^2\tau^2}] + (1-f_s)(1+\xi\mathcal{D}_B)\mathrm{Re}\,\{\widetilde{B}(w)\}\big) \\
&\simeq \frac{N}{4}f_s\xi\mathcal{D} \qquad (w^2\tau^2, w^2\tau_B^2 \gg 1) \\
\mathrm{Re}\,\{\tilde{f}_\xi(2w)\} &= \frac{N}{2}\big(f_s[\frac{1}{1+w^2\tau^2} + \xi\frac{\mathcal{D}}{2}(\frac{1}{1+w^2\tau^2}+\frac{1}{1+9w^2\tau^2})] \\
&\quad +(1-f_s)(1+\xi\mathcal{D}_B)\mathrm{Re}\,\{\widetilde{B}(w)\}\big) \simeq 0 \qquad (w^2\tau^2, w^2\tau_B^2 \gg 1) \\
\tilde{f}_\xi(0) &= \frac{N}{2}\big(f_s[1+\xi\mathcal{D}\frac{1}{1+w^2\tau^2}] + (1-f_s)(1+\xi\mathcal{D}_B)\big) \\
&\simeq \frac{N}{2} \qquad (w^2\tau^2 \gg 1, |\mathcal{D}_B| \ll 1)\,.
\end{aligned}$$

It can be seen that for large frequencies the ratio of the FT real part to the normalization becomes much smaller than unity. We have then for the uncertainty

$$\begin{aligned}
\sigma^2\big(\mathrm{Re}\,\{\tilde{f}_\xi(w)\}\big) &= \frac{N_\xi}{2}\left(1-2\left(\frac{\mathrm{Re}\,\{\tilde{f}_\xi(\nu)\}}{\tilde{f}_\xi(0)}\right)^2 + \frac{\mathrm{Re}\,\{\tilde{f}_\xi(2\nu)\}}{\tilde{f}_\xi(0)}\right) \\
&\simeq \frac{N_\xi}{2}\left(1-\frac{f_s^2\mathcal{D}^2}{2}\right) \qquad (w^2\tau^2 \gg 1)
\end{aligned}$$

where N_ξ denote the number of events with the corresponding tagging decision. It is given by

$$N_\xi \;=\; \tilde{f}_\xi(0) \qquad \text{with} \quad N_+ + N_- = \tilde{f}_+(0) + \tilde{f}_-(0) = N \;.$$

Finally, we arrive at

$$\begin{aligned} \sigma(\Delta) \;&=\; \sqrt{\sigma^2\big(\mathrm{Re}\,\{\tilde{f}_+(w)\}\big) + \sigma^2\big(\mathrm{Re}\,\{\tilde{f}_-(w)\}\big)} \\ &\simeq\; \sqrt{\frac{N}{2}}\sqrt{1 - \frac{f_s^2\mathcal{D}^2}{2}} \end{aligned} \qquad \text{(B.15)}$$

The result in (B.15) is obtained also in [31], through application of the Wiener-Kinchin theorem for the *rms* power spectrum. It is seen that the variance in the FT, from statistical fluctuations, has a behavior independent of the frequency. It is also important to note that all events in the sample, including background, contribute, and that this is to leading order irrespective of resolution and mistagging.

Significance

We have seen that the size of the FT resonance peak is given by

$$\Delta \;=\; \frac{1}{2} N f_s \,\mathsf{D}$$

where D stands for the combined dilution factors

$$\mathsf{D} \;=\; \mathcal{D}\, D_l(w,\sigma_t)\, D_p(w,\sigma_p)$$

from tagging $\mathcal{D}$, vertex resolution D_l, and effective momentum resolution D_p. The noise contribution (B.15) is

$$\sigma(\Delta) = \sqrt{\frac{N}{2}}\sqrt{1 - \frac{f_s^2\mathsf{D}^2}{2}} \simeq \sqrt{\frac{N}{2}} \;.$$

The statistical significance of the signal peak is finally

$$\mathcal{S} \;=\; \frac{\Delta}{\sigma(\Delta)} = \sqrt{\frac{N}{2}} f_s \,\mathsf{D} \;. \qquad \text{(B.16)}$$

We now denote the total number of signal and background events by S and B, respectively. We also notice that we've been considering in the above calculations only flavor tagged candidates. Denoting the tagging efficiency by ϵ, we have $N = \epsilon(S+B)$ We further note that the significance expression (B.16) constitutes an estimate of the uncertainty $\sigma_\mathcal{A}$ of the amplitude [31]. It may accordingly be re-written as

$$\frac{1}{\sigma_\mathcal{A}} = \sqrt{\frac{\epsilon\mathcal{D}^2\, S}{2}}\sqrt{\frac{S}{S+B}}\,\exp\left(-\frac{\sigma_t^2 w^2}{2}\right)\, D_p(w,\sigma_p) \;. \qquad \text{(B.17)}$$

Appendix C

Dilution factors

Measurements are in general affected by various experimental effects, stemming, for instance, from detector characteristics and resolutions, and sample backgrounds. In mixing analyses, the performance of the flavor tagging methods play additionally a prominent role. These sources effectively translate into dilution factors, which dampen the observable amplitude of the oscillations and decrease the sensitivity of the sample. In Appendix B we have described the contribution from various of these sources.

Flavor tagging deals with the determination of which state, B or $\bar{B}$, the meson is at a given stage. The tagging dilution $\mathcal{D}$ provides an estimate of the probability that such decision is correct. This probability for a flavor method α is given by $p_\alpha = \frac{1+\mathcal{D}_\alpha}{2}$, while the mistagging probability is the complement to unity, $1 - p_\alpha$.

In general the determination of the B flavor needs to be performed in different circumstances, employing distinct tagging approaches. For example, a comparison of the flavor at production and decay times is necessary to infer the mixed or unmixed classification of the event. The flavor of the meson at decay time may be obtained from its decay products, when flavor specific decays are employed. This determination becomes more unambiguous the more accurate and complete the reconstruction of the decay chain is. When the candidate reconstruction is pursued in a more inclusive fashion, more elaborate techniques may be used and their respective dilutions estimated.

We are considering systems characterized by two possible states or outcomes. That is, the meson is determined to be in either a particle or antiparticle state. Mixing is itself a two-state physics process: the meson, once detected, is either in the same or the opposite flavor state than when it was produced. In the latter case the associated dilution is given by the cosine term. For such systems it may be useful to define the matrix

$$O(\mathcal{D}_\alpha) \quad = \quad \begin{bmatrix} p_\alpha & 1 - p_\alpha \\ 1 - p_\alpha & p_\alpha \end{bmatrix} = \begin{bmatrix} \frac{1+\mathcal{D}_\alpha}{2} & \frac{1-\mathcal{D}_\alpha}{2} \\ \frac{1-\mathcal{D}_\alpha}{2} & \frac{1+\mathcal{D}_\alpha}{2} \end{bmatrix} ,$$

with

$$\mathcal{D}_\alpha = |O(\alpha)| = 2p_\alpha - 1 \ .$$

When two separate systems are involved, one may write

$$O(\mathcal{D}_\alpha) \cdot O(\mathcal{D}_\beta) \;\; = \;\; O(\mathcal{D}_\alpha \cdot \mathcal{D}_\beta) \; .$$

That is, the dilution factors multiply together.

When multiple tagging methods are employed to determine the flavor of the meson at a common stage, the flavor information may be combined, an issue which we address below.

C.1 Combining independent taggers

This section deals with the proper decay time likelihood description for events tagged by independent algorithms. The combination of flavor information obtained from multiple tagging algorithms has previously been addressed in [89].

Consider independent taggers $\{T_i\}$, each characterized by efficiency ϵ_i and dilution D_i. Denote the probability of decision by tagger T_i to be correct as $p_i^{rs} = \frac{1+D_i}{2}$ and incorrect as $p_i^{ws} = \frac{1-D_i}{2}$. The probability that the B meson has not mixed is $p_{um} = \frac{1+\cos(wt)}{2}$, and that it did mix is $p_m = \frac{1-\cos(wt)}{2}$. For neutral mesons, $w = \Delta m$ is the mixing frequency, while for charged mesons, $w = 0$. Let $p_i^{\pm}$ denote the probability for tagger T_i to give observable tagging decision $\pm$; here, the decision being"“+” (“-”) denotes that the production and decay flavors coincide (are opposite).

For the case of a single tagger, T_i, the probability for an event in the sample to be tagged is ϵ_i, and for it not to be tagged is $1-\epsilon_i$. For the case of two taggers, T_i and T_j, the probability for an event to be tagged by both algorithms is $\epsilon_i \cdot \epsilon_j$, for it to be tagged by T_i alone is $\epsilon_i \cdot (1-\epsilon_j)$, and by none of the algorithms is $(1-\epsilon_i)\cdot(1-\epsilon_j)$. Schematically,

$$p_\epsilon(\xi) = \begin{cases} (1-\epsilon_1)(1-\epsilon_2) & (\xi_1 = 0,\ \xi_2 = 0) \\ \epsilon_1(1-\epsilon_2) & (\xi_1 = \pm 1,\ \xi_2 = 0) \\ (1-\epsilon_1)\epsilon_2 & (\xi_1 = 0,\ \xi_2 = \pm 1) \\ \epsilon_1\epsilon_2 & (\xi_1 = \pm 1,\ \xi_2 = \pm 1) \end{cases} \; .$$

Assuming now that the event is tagged by a single algorithm, we are interested in expressing the probability for observing a particular decision. For example, a “+” decision can be arrived at in case the meson did not mix and the tagger gave the correct answer, or if an incorrect answer was given for a meson which did mix. This is expressed simply as follows,

$$\begin{aligned} p_i^+ &= p_i^{rs} \cdot p_{umix} + p_i^{ws} \cdot p_{mix} \\ &= \frac{1+D_i}{2} \cdot \frac{1+\cos(wt)}{2} + \frac{1-D_i}{2} \cdot \frac{1-\cos(wt)}{2} \\ &= \frac{1+D_i \cdot \cos(wt)}{2} \; . \end{aligned}$$

More generally, the probability for a non-trivial decision ξ is expressed as

$$p_i^{\xi} = \frac{1}{2}[1 + \xi D_i \cdot \cos(wt)] \qquad \text{(singly - tagged)}.$$

Next we extend this reasoning to find the corresponding expressions for events simultaneously tagged by independent algorithms. For example, we would like to evaluate the probability p^{++} that both taggers, T_1 and T_2, give "+" decisions. Despite the algorithms being independent, the equality $p^{++} = p_1^+ \cdot p_2^+$ does not in general hold; this is so due to physics-level correlations, *i.e.* mixing. Instead, and analogously to the single tagger case above, a positive decision for both taggers may be arrived at in case the meson did not mix and the taggers gave the correct answer, or if an incorrect answer was given and the meson did mix. This is expressed as follows,

$$\begin{aligned} p^{++} &= p_1^{rs} \cdot p_2^{rs} \cdot p_{um} + p_1^{ws} \cdot p_2^{ws} \cdot p_m \\ &= \frac{(1 + D_1 D_2) + (D_1 + D_2) \cdot \cos(wt)}{4} . \end{aligned}$$

More generally, the probability for non-trivial decisions ξ_1 and ξ_2 is expressed as

$$p^{\xi_1 \xi_2} = \frac{1}{4}[(1 + \xi_1 \xi_2 D_1 D_2) + (\xi_1 D_1 + \xi_2 D_2) \cdot \cos(wt)] \qquad \text{(doubly - tagged)} .$$

The treatment of the possible combinations of tagging decisions is specified in Table C.1. A general expression can be formally written as

$$P(\xi|t) = \frac{(1 + \xi_1 \xi_2 D_1 D_2) + (\xi_1 D_1 + \xi_2 D_2) \cos(wt)}{(1 + |\xi_1|)(1 + |\xi_2|)} .$$

ξ_1	ξ_2	efficiency factors	tagging decision probabilities	
0	0	$(1-\epsilon_1)\cdot(1-\epsilon_2)$	1	1
+	0	$\epsilon_1 \cdot (1-\epsilon_2)$	$p_1^{rs} \cdot p_{um} + p_1^{ws} \cdot p_m$	$\frac{1+D_1 \cdot \cos(wt)}{2}$
-	0	$\epsilon_1 \cdot (1-\epsilon_2)$	$p_1^{rs} \cdot p_m + p_1^{ws} \cdot p_{um}$	$\frac{1-D_1 \cdot \cos(wt)}{2}$
0	+	$(1-\epsilon_1) \cdot \epsilon_2$	$p_2^{rs} \cdot p_{um} + p_2^{ws} \cdot p_m$	$\frac{1+D_2 \cdot \cos(wt)}{2}$
0	-	$(1-\epsilon_1) \cdot \epsilon_2$	$p_2^{rs} \cdot p_m + p_2^{ws} \cdot p_{um}$	$\frac{1-D_2 \cdot \cos(wt)}{2}$
+	+	$\epsilon_1 \cdot \epsilon_2$	$p_1^{rs} \cdot p_2^{rs} \cdot p_{um} + p_1^{ws} \cdot p_2^{ws} \cdot p_m$	$\frac{(1+D_1 D_2)+(D_1+D_2)\cdot\cos(wt)}{4}$
-	-	$\epsilon_1 \cdot \epsilon_2$	$p_1^{rs} \cdot p_2^{rs} \cdot p_m + p_1^{ws} \cdot p_2^{ws} \cdot p_{um}$	$\frac{(1+D_1 D_2)-(D_1+D_2)\cdot\cos(wt)}{4}$
+	-	$\epsilon_1 \cdot \epsilon_2$	$p_1^{rs} \cdot p_2^{ws} \cdot p_{um} + p_1^{ws} \cdot p_2^{rs} \cdot p_m$	$\frac{(1-D_1 D_2)+(D_1-D_2)\cdot\cos(wt)}{4}$
-	+	$\epsilon_1 \cdot \epsilon_2$	$p_1^{rs} \cdot p_2^{ws} \cdot p_m + p_1^{ws} \cdot p_2^{rs} \cdot p_{um}$	$\frac{(1-D_1 D_2)+(D_2-D_1)\cdot\cos(wt)}{4}$

Table C.1: Probability factors for the possible decisions of two independent taggers.

Same side tagging in inclusive modes

In partially reconstructed modes, the tag-track chosen by the SST algorithm may belong to the B decay products. Indeed, in $B \to DlX$ modes, in a fraction (f_{**}) of the times a decay track, denoted π^{**}, is selected as the tag-track; this notation will be used in the following. Also, the first (second) tagger will be identified with the same (opposite) side algorithm: $T_1 = SST$ and $T_2 = OST$.

The charge of the lepton and that of the π^{**} are anti-correlated (Tables 3.17 and 3.18). As both the lepton and the π^{**} are final-state tracks, no information of the B initial flavor can be inferred, when the π^{**} is selected by the algorithm. Accordingly, the overall probability for a SST decision is given by

$$\begin{aligned} p_1^+ &= (1 - f_{**}) \cdot \frac{1 + D_1 \cdot \cos(wkt)}{2} \\ p_1^- &= (1 - f_{**}) \cdot \frac{1 - D_1 \cdot \cos(wkt)}{2} + f_{**} \end{aligned}$$

This may be summarized as

$$P(\xi|t) = (1 - f_{**}) \cdot \frac{1 + \xi D_1 \cdot \cos(w\kappa t)}{(1 + |\xi|)} + f_{**} \cdot \frac{1 - \xi}{2}|\xi| \; .$$

Here κ denotes the kinematical correction factor (4.8) required by the partial proper time reconstruction.

For studying the case when an event is tagged simultaneously by SST and OST, it is useful to consider first the hypotheses that a π^{**} is and is not selected by the SST. For the sake of concreteness, consider the evaluation of the case the SST yields decision "-" and the OST gives decision "+".

In the hypothesis that a π^{**} is not selected by the SST, the probability for the decisions is correlated to that of the B to have mixed. It is expressed as

$$\begin{aligned} p^{-+} &= p_1^{ws} \cdot p_2^{rs} \cdot p_{um} + p_1^{rs} \cdot p_2^{ws} \cdot p_m \\ &= \frac{(1 - D_1 D_2) + (D_2 - D_1) \cdot \cos(wkt)}{4} \end{aligned}$$

In the hypothesis that a π^{**} is selected by the SST, no physics-level correlations need to be considered (as the π^{**} does not carry initial flavor information), and the probability can be expressed as

$$p^{-+} = p_1^- \cdot p_2^+ = 1 \cdot \frac{1 + D_2 \cdot \cos(wkt)}{2}$$

The possible combinations of other non-trivial decisions can be evaluated in a similar fashion, and the result may be represented as

$$\begin{aligned} p^{\xi_1 \xi_2} &= (1 - f_{**}) \cdot \frac{1}{4}[(1 + \xi_1 \xi_2 D_1 D_2) + (\xi_1 D_1 + \xi_2 D_2) \cdot \cos(wkt)] \\ &+ \quad f_{**} \cdot \frac{1 - \xi_1}{2} \cdot \frac{1 + \xi_2 D_2 \cdot \cos(wkt)}{2} \; . \end{aligned}$$

The general expression summarizing the possible tagging decision combinations is accordingly written as

$$\begin{aligned}
P(\xi|t) &= (1-f_{**})\cdot\frac{(1+\xi_1\xi_2 D_1 D_2)+(\xi_1 D_1+\xi_2 D_2)\cdot\cos(wkt)}{(1+|\xi_1|)(1+|\xi_2|)} \\
&+ \quad f_{**}\cdot\frac{1-\xi_1}{2}|\xi_1|\cdot\frac{1+\xi_2 D_2\cdot\cos(wkt)}{1+|\xi_2|}\ .
\end{aligned}$$

Appendix D

Analytical likelihood evaluation

A considerable effort was put in developing and implementing PDF expressions allowing precise and efficient likelihood evaluations, specifically through analytical integration.

In general, the computation of probability integrals is less time consuming and more precise when these are solved analytically rather than numerically. At the very beginning one needs to evaluate convolution integrals with normal (resolution) distributions, and these are already complex enough, being solved only through the use of the error function. In cases where integrals of such functions need to be evaluated, not only for each iteration step of the fit, but also for each event input, the computation time, as well as the precision of the numerical method, are of concern. Other cases where multiple integrals of such types are involved require that at least part of the expressions be evaluated analytically in order to render the fits feasible in a reasonable amount of time.

D.1 Lifetime

The optimization of the computation speed is dependent on the possibility of evaluating analytically probability integrals, namely the PDF normalization to be performed for each input event. The issue is trivial for unbiased proper decay time distributions (4.15). In cases of explicit biases (4.17) the computation is also readily performed (4.18). The issue is not straightforward however for general biasing effects, induced both at trigger or reconstruction stages. There, the possibility of analytical PDF normalization relies upon convenient parameterizations of the proper decay time efficiency function, $\mathcal{E}(t)$, defined in (4.3).

The required computation is indicated in (4.20). In what follows the parameterization motivated in (4.26) is assumed. The normalization integrals may accordingly be expressed as

$$\begin{aligned}\mathcal{N}_n(\alpha) &= a_n\tau^n \cdot \frac{1}{2\tau}\, e^{\frac{\sigma_t^2}{2}\frac{\alpha^2}{\tau^2}} \cdot \int_{\zeta_n}^{+\infty} \frac{t^n}{\tau^n}\, e^{-(\alpha+\frac{\tau}{\tau_n})\frac{t}{\tau}}\ \mathrm{Erfc}\left(-\frac{\tau}{\sqrt{2}\sigma_t}\frac{t}{\tau} + \frac{\sigma_t}{\sqrt{2}}\frac{\alpha}{\tau}\right) dt \\ &= a_n\tau^n \cdot \frac{1}{2}\, e^{c^2} \cdot \int_{\zeta_n/\tau}^{+\infty} x^n\, e^{-bx}\ \mathrm{Erfc}\left(-ax+c\right)\ dx \end{aligned} \tag{D.1}$$

where a, b and c denote the following dimensionless positive parameters,

$$a = \frac{\tau}{\sqrt{2}\sigma_t}, \qquad b = \alpha + \frac{\tau}{\tau_n}, \qquad c = \frac{\sigma_t}{\sqrt{2}}\frac{\alpha}{\tau}\ .$$

While for fully reconstructed decays the parameter α is identical to unity, for partially reconstructed proper times it corresponds to the κ-factor, $\alpha = \kappa$. In the latter case, an additional integration over the k-factor parameter is to be performed, $\int_\kappa \ldots \kappa\mathcal{F}(\kappa)d\kappa$. While an analytical implementation of this integration is in principle achievable using identical techniques, it has not been found necessary.

The integration result, for the prototype parameterization

$$\mathcal{E}(t) \;=\; (a_0 + a_1 t + a_2 t^2)\cdot e^{-\frac{t}{\tau_0}} \cdot \theta(t-\zeta)\ , \qquad \alpha = 1 \tag{D.2}$$

is given by

$$\begin{aligned}\mathcal{N} \;=\; & \frac{1}{2b}\Big\{\left[a_0 + a_1\frac{\tau}{b}(1 - b\frac{\sigma_t^2}{\tau\tau_0}) + a_2\frac{\tau^2}{b^2}(2 + \frac{\sigma_t^2}{\tau^2}(1-\frac{\tau}{\tau_0})b + \frac{\sigma_t^4}{\tau^2\tau_0^2}b^2)\right] \\ & \cdot\left(2 - \mathrm{Erfc}\left(\frac{-1}{\sqrt{2}}(\frac{\zeta}{\sigma_t} + \frac{\sigma_t}{\tau_0})\right)\right) \cdot e^{\frac{\sigma_t^2}{2\tau_0^2}} \\ & + \left[a_0 + a_1\frac{\tau}{b}(1 + b\frac{\zeta}{\tau}) + a_2\frac{\tau^2}{b^2}\left(1 + (1 + b\frac{\zeta}{\tau})^2\right)\right] \\ & \cdot \mathrm{Erfc}\left(\frac{-1}{\sqrt{2}}(\frac{\zeta}{\sigma_t} - \frac{\sigma_t}{\tau})\right) \cdot e^{-b\frac{\zeta}{\tau} + \frac{\sigma_t^2}{2\tau^2}} \\ & + \left[a_1\tau + a_2\tau^2(\frac{\zeta}{\tau} + \frac{2}{b} - \frac{\sigma_t^2}{\tau\tau_0})\right] \cdot \frac{\sigma_t}{\tau}\sqrt{\frac{2}{\pi}} \cdot e^{-\frac{\zeta}{\tau}\left(\frac{\tau}{\tau_0} + \frac{1}{2}\frac{\zeta}{\tau}\frac{\tau^2}{\sigma_t^2}\right)}\Big\}\ . \end{aligned} \tag{D.3}$$

Probability integrals

The PDF integrals at hand involve the exponential function, power terms, and the error function, taking the following general form

$$I_n \;=\; \int t^n\, e^{-bt}\ \mathrm{Erfc}(-at + c)\ dt\ . \tag{D.4}$$

The complementary error function has the following integral representation on the real axis, which is taken as definition,

$$\mathrm{Erfc}(z) \equiv \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-u^2}\, du \,.$$

The following useful properties are also noted,

$$\mathrm{Erfc}(0) = 1, \qquad \mathrm{Erfc}(\infty) = 0, \qquad \mathrm{Erfc}(-\infty) = 2, \qquad \frac{\partial}{\partial z}\mathrm{Erfc}(z) = -\frac{2}{\sqrt{\pi}} e^{-z^2} \,,$$

along with the Taylor series and asymptotic expansions

$$\begin{aligned}
\mathrm{Erfc}(x) &= 1 - \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)n!} \qquad &(\text{small } x)\,, \\
\mathrm{Erfc}(x) &= \frac{e^{-x^2}}{x\sqrt{\pi}}\big(1 + \sum_{n=1}^{\infty} \frac{(-1)^n (2n)!}{n!(2x)^{2n}}\big) \qquad &(\text{large } x)\,.
\end{aligned}$$

Integrals of the type of (D.4) can be solved analytically. Calculations have been confirmed using the symbolic program `Mathematica` [90]. Tests of the analytical results against those obtained with numerical integration were performed. Results are expressed below in a form appropriate for *definite integration* of the expressions (D.4) on a domain of the real axis:

$$\begin{aligned}
I_0 &\simeq \frac{1}{b}\left(e^{\frac{b^2}{4a^2} - \frac{bc}{a}}\,\mathrm{Erfc}(-at + c - \frac{b}{2a}) - e^{-bt}\,\mathrm{Erfc}(-at+c) \right)\,, \\
I_1 &\simeq \left(\frac{1}{b^2} - \frac{1}{2a^2} + \frac{c}{ab} \right) e^{\frac{b^2}{4a^2} - \frac{bc}{a}}\,\mathrm{Erfc}(-at + c - \frac{b}{2a}) - \frac{1}{ab\sqrt{\pi}} e^{-(-at+c)^2 - bt} \\
&\quad - \frac{bt+1}{b^2} e^{-bt}\,\mathrm{Erfc}(-at+c)\,, \\
I_2 &\simeq \left(\frac{8a^4 - 2a^2b^2 + b^4}{4a^4b^3} + \frac{2c}{ab^2} - \frac{c}{a^3} + \frac{c^2}{a^2 b} \right) e^{\frac{b^2}{4a^2} - \frac{bc}{a}}\,\mathrm{Erfc}(-at + c - \frac{b}{2a}) \\
&\quad - \left(\frac{t}{ab\sqrt{\pi}} - \frac{b^2 - 4a^2 - 2abc}{2a^3b^2\sqrt{\pi}} \right) e^{-(-at+c)^2 - bt} - \frac{(bt+1)^2 + 1}{b^3} e^{-bt}\,\mathrm{Erfc}(-at+c)\,.
\end{aligned}$$

D.2 Mixing

The description of flavor oscillations involves the introduction of a cosine term in the proper decay time PDF. The latter (6.13) involves the following factors

$$\mathcal{P}_{\mathrm{exp}}(t;\kappa) = \left[e^{-\kappa\frac{t}{\tau}} \theta(t) \right] \otimes G(t;\sigma_t) \cdot \mathcal{E}(t)\,, \tag{D.5}$$

$$\mathcal{P}_{\mathrm{cos}}(t;\kappa) = \left[e^{-\kappa\frac{t}{\tau}} \theta(t) \cdot \cos(w\kappa t) \right] \otimes G(t;\sigma_t) \cdot \mathcal{E}(t)\,, \tag{D.6}$$

where w denotes the oscillation frequency, and the parameter κ is provisionally identified with unity (the κ-factor) for fully (partially) reconstructed modes.

The convolution integral in (D.5) has been already evaluated (4.15) in the context of the lifetime analysis

$$\mathcal{P}_{\rm exp}(t;\kappa) \;=\; \frac{1}{2} e^{-\frac{\kappa}{\tau}(t-\frac{\kappa\sigma_t^2}{2\tau})}\,{\rm Erfc}\left(\frac{\kappa\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right)\cdot\mathcal{E}(t)\ . \tag{D.7}$$

Regarding (D.6), it can be seen that it formally reduces to (D.5) upon extension to the complex plane. Indeed, by expanding the cosine in terms of exponential functions of imaginary phase, the following relation holds (with $i=\sqrt{-1}$)

$$\begin{aligned}\mathcal{P}_{\rm cos}(t;\kappa) \;&=\; {\rm Re}\,\{\mathcal{P}_{\rm exp}(t;\alpha)\} \qquad \text{with} \quad \alpha=\kappa(1+iw\tau)\ , \qquad (\text{D.8})\\ &=\; \frac{1}{2}e^{-\frac{\kappa}{\tau}(t-\frac{\kappa\sigma_t^2}{2\tau})}e^{-\frac{\kappa^2\sigma_t^2w^2}{2}}\cdot{\rm Re}\,\{e^{-iw\kappa(t-\frac{\kappa\sigma_t^2}{\tau})}\,{\rm Erfc}(\frac{\kappa\sigma_t^2-t\tau}{\sqrt{2}\sigma_t\tau}+i\frac{\kappa\sigma_t w}{\sqrt{2}})\}\cdot\mathcal{E}(t)\ .\end{aligned}$$

That is, the PDF computations arising in the framework of mixing analyses are accomplished as a complexification of those found for lifetime analyses.

A note on PDF normalization

In Section D.1 we have tackled the issue of analytically integrate expressions of the form of (D.5),

$$\mathcal{N}_{\rm exp}(t;\kappa) \;=\; \int_{-\infty}^{+\infty}\mathcal{P}_{\rm exp}(t;\kappa)\,dt \tag{D.9}$$

needed for proper decay time PDF normalization in the context of lifetime analyses. As it was addressed in Section 6.3.2, this is also the normalization which is needed in the context of mixing analyses.

While integration of expressions of the type (D.6),

$$\mathcal{N}_{\rm cos}(t;\kappa) \;=\; \int_{-\infty}^{+\infty}\mathcal{P}_{\rm cos}(t;\kappa)\,dt\ , \tag{D.10}$$

is in general not a requisite for the process of likelihood maximization, they may reveal useful in circumstances such as likelihood projections in tagged subspaces. Such expressions are thus evaluated and implemented in the fitting framework. For the unbiased cases the following is obtained,

$$\mathcal{N}_{\rm cos}(t) \;=\; \frac{\tau}{1+w^2\tau^2} \qquad\qquad \text{for} \quad \mathcal{E}(t)=1\ .$$

In general, the integration (D.10) may be obtained by extending the result of (D.9) to the complex plane, as

$$\mathcal{N}_{\rm cos}(t;\kappa) \;=\; {\rm Re}\,\{\mathcal{N}_{\rm exp}(t;\alpha)\} \qquad \text{with} \quad \alpha=\kappa(1+iw\tau)\ .$$

In particular, for proper decay time biases described by a generic t-efficiency function, parameterized as in (D.2), the normalization may be obtained correspondingly by evaluating in the complex plane the results obtained in Section D.1.

Extension to the complex plane

As pointed out, the likelihood computation for tagged events involves in general the task of evaluating expressions in the complex plane. This requires that a complex class be defined in the fitter framework, in order to handle the basic complex operations.

Additionally, one needs to evaluate the complementary error function of complex argument. For this purpose it is convenient to express the latter as

$$\begin{aligned} \mathrm{Erfc}(z) &= e^{-z^2} W(iz) \;, \\ W(z) &= e^{-z^2} \left[1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{u^2}\, du \right] \;, \end{aligned}$$

with z denoting a complex number, and $W(z)$ is the complex error function (also called Faddeeva function) which is evaluated with existing numerical algorithms. While this is the implementation adopted, we also mention in passing that the complementary error function may alternatively be expressed in terms of the incomplete gamma function,

$$\begin{aligned} \Gamma(a,x) &\equiv \int_x^\infty u^{a-1} e^{-u} du \\ \mathrm{Erfc}(z) &= \begin{cases} \frac{1}{\sqrt{\pi}}\Gamma(\frac{1}{2}, z^2), & \mathrm{Re}\,(z) > 0 \\ 2 - \frac{1}{\sqrt{\pi}}\Gamma(\frac{1}{2}, z^2), & \mathrm{Re}\,(z) < 0 \end{cases} \end{aligned}$$

for which convenient numerical methods are also available.

D.3 Log-likelihood expansion

Once the data samples have been characterized and the associated PDFs of the various input quantities established, the analysis of B_s oscillations contains a final additional step. This involves the introduction of an extra parameter, the amplitude $\mathcal{A}$, and a scanning procedure, which is presented in Section 7.3. The procedure requires that many fits be performed to the amplitude parameter, one for each probe frequency value. Such fits need to be executed considerably more times, in many samples of toy Monte Carlo events, for systematics evaluation.

The likelihood maximization procedure involved in the amplitude scanning can be made more time-effective in a few complementary ways. First of all advantage should be taken from the fact that the amplitude is the only floating parameter, and all remaining fit parameters are fixed. The computation of the various PDFs' terms not depending on the amplitude need to be evaluated once only for each event, and the results cached for use during the maximization process, which is to be repeated for the various probed frequencies.

Furthermore, the likelihood maximization may be achieved in a time-efficient fashion based on log-likelihood expansion and derivation. The event likelihood has a linear dependence on the amplitude parameter. The method can be more simply illustrated expressing (7.3) and its logarithmic series expansion as

$$\begin{aligned} 1+\upsilon\mathcal{A} &= 1+\mathcal{A}\cdot\xi\mathcal{D}\cos(wt) \ , \\ \ln(1+\upsilon\mathcal{A}) &= \upsilon\mathcal{A}-\frac{1}{2}\upsilon^2\mathcal{A}^2+\frac{1}{3}\upsilon^2\mathcal{A}^3-\frac{1}{4}\upsilon^4\mathcal{A}^4+\mathcal{O}(\upsilon^5) \qquad (|\upsilon|<1) \ . \end{aligned}$$

That is, the logarithm of the likelihood is expanded as a polynomial on the amplitude with constant and pre-determined coefficients. Likelihood maximization becomes then reduced to finding the polynomial roots; for lower polynomial degrees, the latter can be achieved analytically.

More generally, as introduced in Section 4.1 and Section 6.1, the likelihood has the following structure,

$$\mathcal{L}=\prod_i\sum_\alpha f_\alpha\mathcal{P}_i^\alpha \qquad \text{with} \qquad \mathcal{P}=L_m\,L_\mathcal{D}\,L_{\sigma_t}\,L_{t,\xi} \ ,$$

where the indices i and α run over the number of events and number of sample components, respectively, and f_α denote the component fractions. For the signal component, the proper decay time likelihood factor (7.6) has the following form

$$L_{t,\xi} = \frac{p_\epsilon}{1+|\xi|}\frac{1}{\mathcal{N}\tau}\left(\mathcal{P}_{\text{exp}}+\mathcal{A}\cdot\xi\mathcal{D}\mathcal{P}_{\cos}\right)$$

where $\mathcal{N}$ is evaluated in (D.3), and $\mathcal{P}_{\text{exp}}$ and $\mathcal{P}_{\cos}$ are given by (D.5) and (D.6), respectively. The likelihood logarithm may accordingly be cast in the form

$$\ln(\mathcal{L}(\mathcal{A})) = \sum_i \ln\left(\alpha_i+\mathcal{A}\cdot\beta_i(w)\right)=\sum_i\ln\left(1+\mathcal{A}\cdot\upsilon_i(w)\right)+\text{const.}$$

where α, β and υ correspond to combined likelihood factors with no dependence on the amplitude parameter. Making use of the logarithm series expansion, one has

$$\ln(\mathcal{L}(\mathcal{A})) = \sum_n\left(\sum_i\frac{1}{n}(-1)^{n-1}\upsilon_i^n\right)\cdot\mathcal{A}^n+\text{const.}$$

Alternatively to the standard `Minuit` minimization procedure based on (4.2), the amplitude values which maximize the likelihood may be found as the roots of the polynomial

$$\frac{d\ln(\mathcal{L}(\mathcal{A}))}{d\mathcal{A}} = \sum_n\left(\sum_i(-1)^{n-1}\upsilon_i^n\right)\cdot\mathcal{A}^{n-1}=0 \ .$$

The amplitude uncertainty estimates $\sigma_\mathcal{A}$ are obtained from the amplitude values for which the log–likelihood has varied by an amount of 0.5 relative to its found maximum.

Appendix E

Fitting framework validation

The unbinned maximum likelihood fit is of considerable complexity. In the development of the fitting framework, the starting point is provided by simplified models. Successive complications are then progressively introduced. The $J/\psi K$ samples were particularly useful at many stages in this respect. Their sample composition, mass and proper time models happen to be relatively simpler in general. When deriving the implementation of trigger bias in the proper decay time distribution, it was insightful again to work with the unbiased $J/\psi K$ samples. Biases were introduced as ad-hoc selection thresholds, such as simple, direct cuts on proper time. It became apparent, using actual data in this way, that certain approaches previously attempted do not suit. Once solutions for the simpler cases are found, further elaborations alongside similar lines are then pursued.

A technique that was consistently employed in the development of the fitting model and for checking implementation consistency was that of toy Monte Carlo. Events are sampled according to the likelihood model. If a large enough number of events is generated, its distribution will exactly coincide, by design, with the likelihood function. In these circumstances, a fit to the data should return precisely the parameter values that were employed for the generation. This is a useful consistency test. Differences between the expected value parameter and its value returned by the fit must be due to statistical fluctuations. The distribution of such differences, referred to as "pull", should be consistent with a unit Gaussian centered at zero. To achieve the proper statistical fluctuation, the number of events simulated for each sample is matched to that collected in the data. This allows a direct, quantitative determination of whether the values returned by the likelihood are unbiased estimators of the fit parameters.

A selection of pull distributions is shown in Figure E-1. The Gaussian characteristics of such distributions for selected parameters are shown in Tables E.1 – E.3.

parameter	mean	width	Gaussian fit prob.
M	0.025 ± 0.023	1.043 ± 0.017	0.937
σ_m	-0.036 ± 0.023	1.027 ± 0.016	0.786
f_{bg}	-0.012 ± 0.022	0.993 ± 0.016	0.787
$c\tau$	-0.025 ± 0.022	1.004 ± 0.016	0.231
backgr. $c\tau$	-0.049 ± 0.023	1.014 ± 0.016	0.228
backgr. offset	0.009 ± 0.022	0.999 ± 0.016	0.742
backgr. σ	-0.106 ± 0.023	1.026 ± 0.016	0.219

Table E.1: Characteristics of pull distributions for mass and lifetime fits; 1000 events per pseudo-experiment.

parameter	mean	width	Gaussian fit prob.
m	0.088 ± 0.031	0.987 ± 0.022	0.969
$c\tau$	-0.021 ± 0.033	1.029 ± 0.023	0.884
Δm_d	0.012 ± 0.033	1.047 ± 0.023	0.032
SMT ϵ	-0.064 ± 0.031	0.981 ± 0.022	0.780
JJP ϵ	-0.011 ± 0.031	0.988 ± 0.022	0.459
SMT $\mathcal{D}$	0.092 ± 0.031	0.994 ± 0.022	0.782
JJP $\mathcal{D}$	-0.002 ± 0.033	1.049 ± 0.023	0.985
SMT ϵ_{bg}	-0.029 ± 0.031	0.991 ± 0.022	0.950
JJP ϵ_{bg}	0.046 ± 0.031	0.970 ± 0.022	0.781
SMT $\mathcal{D}_{bg}$	-0.016 ± 0.032	1.017 ± 0.023	0.013
JJP $\mathcal{D}_{bg}$	0.018 ± 0.032	1.011 ± 0.023	0.554

Table E.2: Characteristics of pull distributions for mixing fits; 2000 events per pseudo-experiment, two taggers, $\Delta m = 0.5$ ps^{-1}.

parameter	mean	width	Gaussian fit prob.
m	0.034 ± 0.033	1.054 ± 0.024	0.610
$c\tau$	-0.055 ± 0.032	1.012 ± 0.023	0.389
$\mathcal{A}$	-0.007 ± 0.034	1.062 ± 0.024	0.662
SMT ϵ	-0.076 ± 0.033	1.042 ± 0.023	0.245
JJP ϵ	-0.052 ± 0.032	1.018 ± 0.023	0.235
SMT ϵ_{bg}	-0.028 ± 0.032	0.998 ± 0.022	0.568
JJP ϵ_{bg}	-0.005 ± 0.032	1.010 ± 0.023	0.827
SMT $\mathcal{D}_{bg}$	-0.034 ± 0.033	1.046 ± 0.023	0.697
JJP $\mathcal{D}_{bg}$	0.024 ± 0.032	1.022 ± 0.023	0.373

Table E.3: Characteristics of pull distributions for amplitude scan fits; 2000 events per pseudo-experiment, two taggers with fixed dilution scale factors, $\Delta m = 15$ps^{-1}.

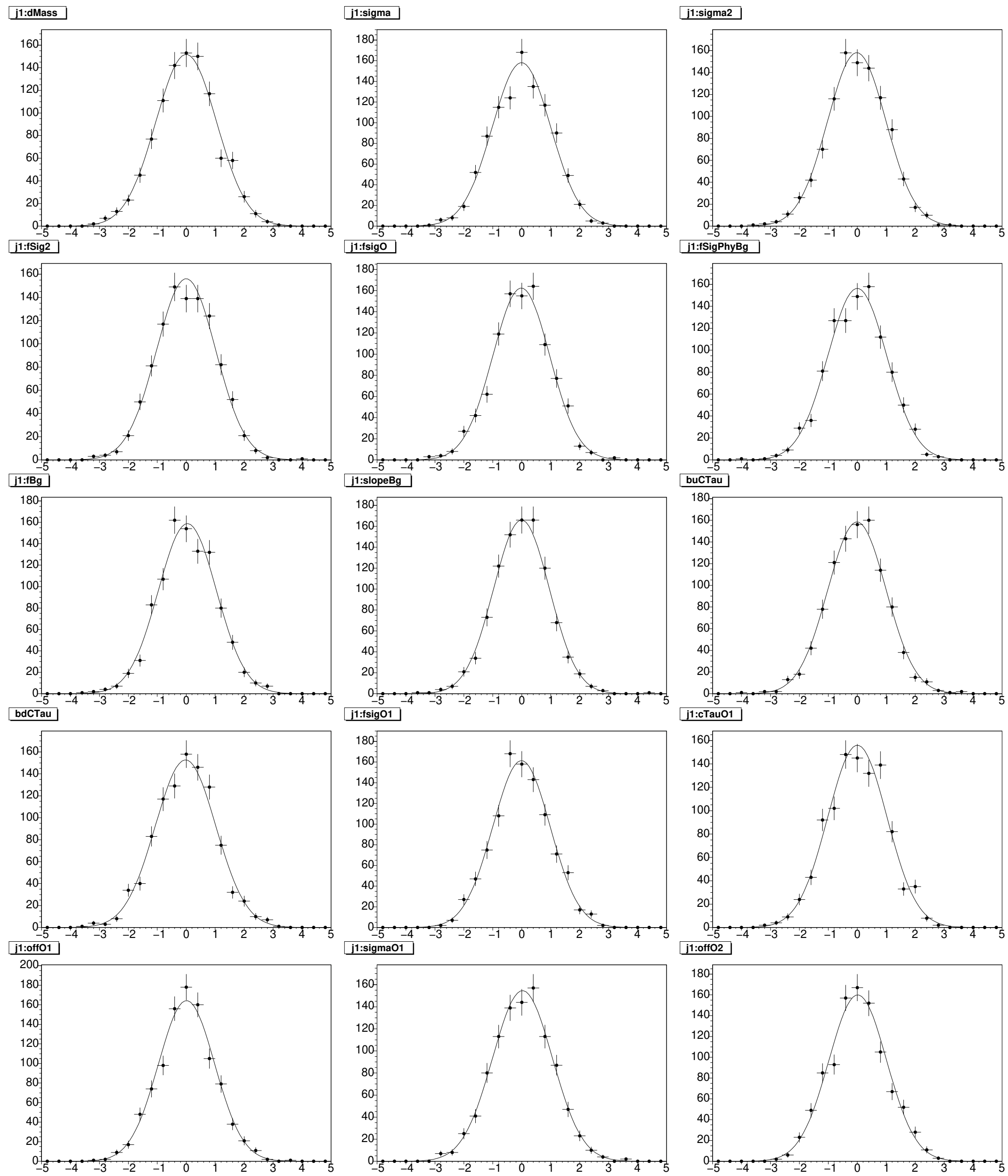


Figure E-1: Toy Monte Carlo pull distributions, for a selection of mass and proper time fit parameters.

Appendix F

Amplitude scanning

Shape of the amplitude scan profile

In the amplitude scan method, in case the probe frequency coincides with that of the oscillating system the amplitude is expected to be maximal and unity. As frequencies away from that true value are probed, the amplitude is expected to approach zero. The shape of the resonance peak is expected to be given, modulo slowly varying functions, by a Breit-Wigner function, whose width is determined by the lifetime of the system. This issue is investigated using Fourier analysis, as is done in Appendix B. In case, however, the proper decay time distribution is biased, deviating from the form of a smeared exponential, so-called *undershooting* are expected on both sides of the amplitude peak.

The appearance of such effects has been demonstrated for the case of a simple bias in (B.13). A graphical representation is provided in Figure F-1. This is a simple, convenient illustration of the origin behind the general effect. A more elaborate analytical treatment may be pursued to account for more realistic conditions. However, for such more complex, specific cases a simulation of the involved sample characteristics becomes appropriate. The scan obtained from a full fit of a toy Monte Carlo mixing sample is shown in Figure F-2, for a representative frequency of Δm_d. The generation of the sample is performed both including and excluding the characteristic proper time bias, using in the fit model the t-efficiency function and turning it off, as appropriate. The undershooting is verified in the former case and absent in the latter, as expected. The amplitude scan performed on an actual *data* sample of fully reconstructed B^0 decays is shown in Figure F-3.

We make now a worthy consideration about the width of the amplitude peak. As it is apparent from Figures F-1 and F-2, as well as from (B.13), the biasing of the proper time distribution induces a narrowing of the peak relative to the unbiased case. In this latter case, the width is determined essentially by the lifetime τ of the system. However, the same effect which, in the t biased case, is responsible for the undershooting, also induces the width narrowing. This translates in turn into a narrowing of the likelihood profile for Δm (which corresponds to the logarithm of the likelihood ratio in (9.28)) and ultimately in a decrease of the uncertainty in the

Δm determination. By preferentially selecting B events further displaced from the production point, samples with larger proper decay times are formed, which more precisely probe the frequency of the oscillations.

In Figure F-4 we show the Δm_s amplitude scan obtained after extending the analysis to a larger dataset of 1 pb^{-1} reported in [81]. The undershooting effects on both sides of the oscillation signal peak are visible. As demonstrated, they are just expected, and are a consequence of the characteristics of the proper time distributions, induced by the trigger and selection criteria employed to gather our data samples.

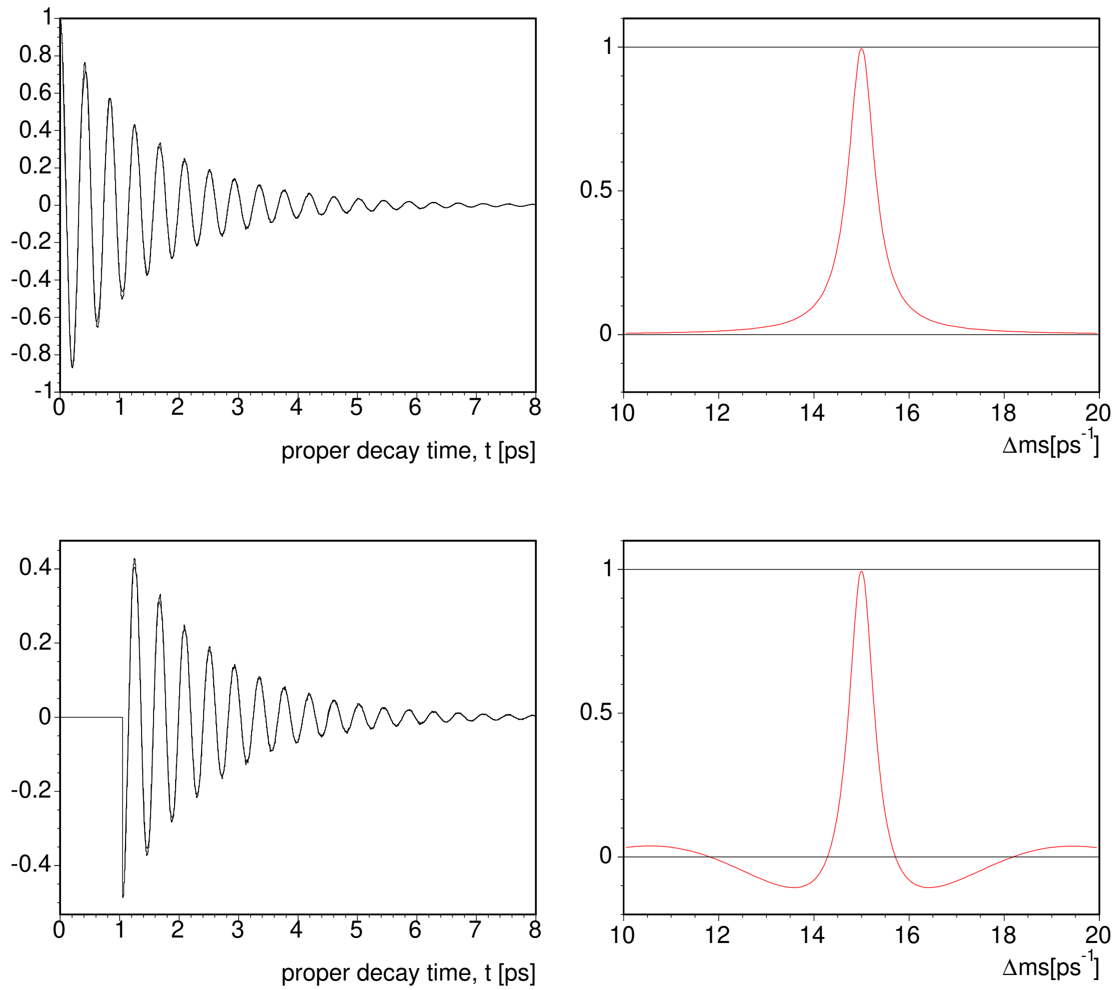


Figure F-1: Illustration of the amplitude scan shape for the case of unbiased proper decay time (top) and in the presence of a direct cut (bottom); the graphs on the left indicate the asymmetry distributions, while those on the right represent the corresponding Fourier transforms.

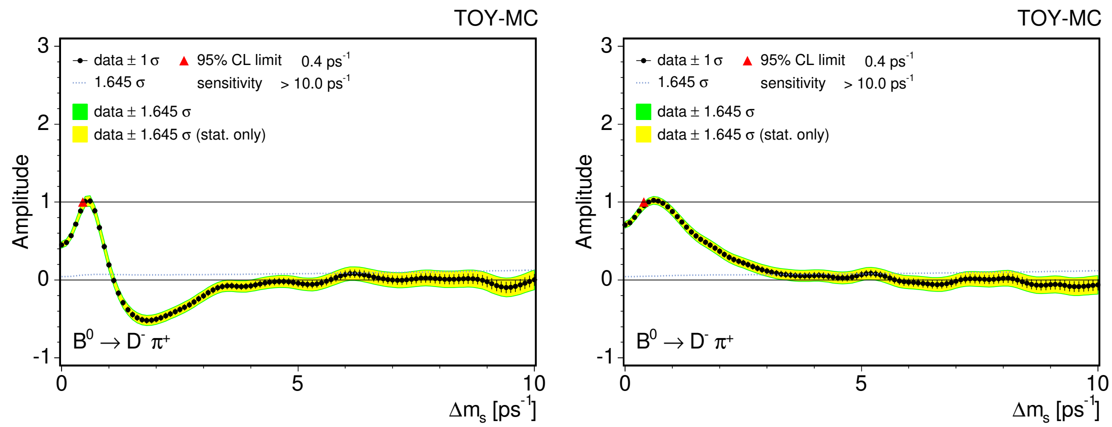


Figure F-2: Illustration of t-biasing effects on the amplitude scan, using realistic toy simulation for the Δm_d case: (left) with biased t distribution, and (right) for the unbiased case where the efficiency function was been disabled.

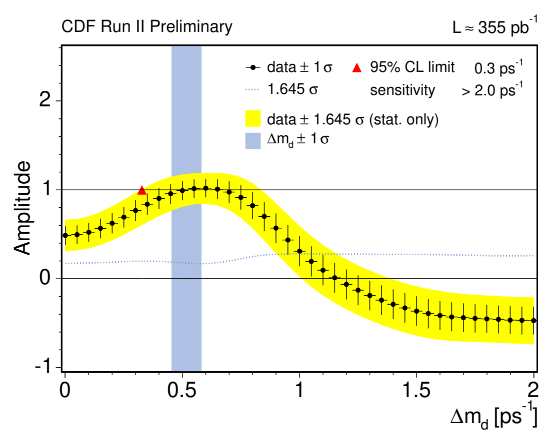


Figure F-3: Δm_d amplitude scan from the hadronic B^0 data samples.

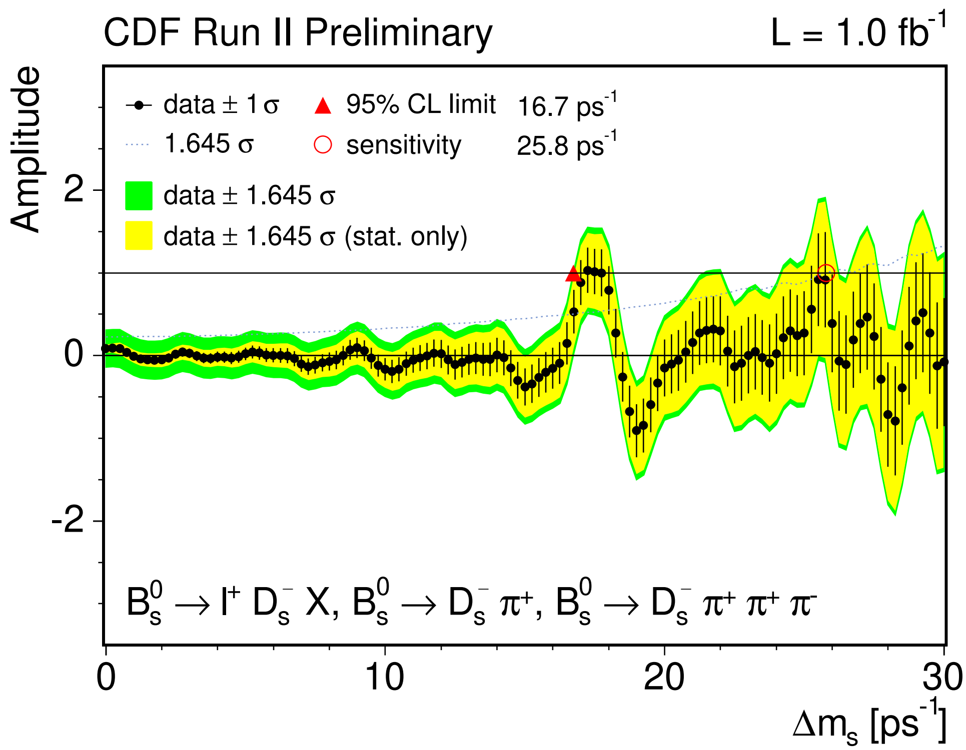


Figure F-4: Amplitude scan showing the first measurement of Δm_s .

List of Figures

List of Tables

Bibliography

[1] S. Eidelman *et al.* (Particle Data Group), *Review of particle physics*, Phys. Lett. B **592**, 1 (2004)

[2] M.E. Peskin, D.V. Schroeder, An introduction to quantum field theory, Perseus Books (1995); L. Ryder, Quantum field theory, Cambridge University Press (1996); S. Weinberg, The quantum theory of fields, Cambridge University Press (1995); there are many other excellent books on the central subject of quantum field theory, this being a list the author is best familar with

[3] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10**, 531 (1963)

[4] M. Kobayashi, T. Maskawa, *CP violation in the renormalizable theory of weak interactions*, Prog. Theor. Phys. **49**, 652 (1973)

[5] L. L. Chau, W. Y. Keung, Phys. Rev. Lett. **53** 1802, (1984)

[6] L. Wolfenstein, *Parameterization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** 1945, (1983)

[7] A. Buras, M. Lautenbacher, G. Ostermaier, Phys. Rev. D **50**, 3433 (1994)

[8] G.C. Branco, L. Lavoura, J.P. Silva, CP Violation, International series of monographs in physics, No. 103, Clarendon Press, Oxford, UK (1999)

[9] I.I. Bigi, A.I. Sanda, CP violation, Cambridge monographs on particle physics, nuclear physics, and cosmology, No. 9, Cambridge University Press (2000)

[10] C. Jarslog, Phys. Rev. Lett. **55**, 1039 (1985)

[11] The geometric interpretation of the unitarity triangle has been pointed out by Bjorken (1986)

[12] M Okamoto, *Full determination of the CKM matrix using recent results from lattice QCD*, Proc. Sci. LAT2005, 013 (2005)

[13] A. Höcker, Z. Ligeti, *CP violation and the CKM matrix*, hep-ph/0605217; R. Fleischer, *Flavour physics and CP violation*, hep-ph/0608010 (2006), references therein, and the various works which appeared following the recent Δm_s Tevatron results

[14] A. Buras, P. Chankowski, J. Rosiek, L. Slawianowska, Nucl. Phys. B **619**, 434 (2001)

[15] M. Gell-Mann, A. Pais, *Behavior of neutral particles under charge conjugation*, Phys. Rev. **97**(3), 1387 (1955)

[16] K. Lande *et al.*, *Observation of long-lived neutral V particles*, Phys. Rev. **103**, 1901 (1956)

[17] S. Glashow, J. Iliopoulos, L. Maiani, Phys. Rev. D **2**, 1285 (1970)

[18] J. Christenson, J. Cronin, V. Fitch, R. Turlay, Phys. Rev. Lett. **13**, 138 (1964)

[19] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977)

[20] F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995)

[21] C. Albajar *et al.* (UA1 Collaboration), Phys. Lett. B **186**, 247 (1987)

[22] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **192**, 245 (1987)

[23] M. Artuso *et al.* (CLEO Collaboration), Phys. Rev. Lett. **62**, 2233 (1989)

[24] Heavy Flavor Averaging Group, `www.slac.stanford.edu/xorg/hfag/osc/index.html`

[25] K. Abe *et al.* (Belle Collaboration), Phys. Rev. D **71**, 072003 (2005); **71**, 079903(E) (2005); N. C. Hastings *et al.* (Belle Collaboration), Phys. Rev. D **67**, 052004 (20053); B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **88**, 221803 (2002)

[26] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C. **35** (2004)

[27] K. Abe *et al.* (SLD Collaboration), Phys. Rev. D **67**, 012006 (2003)

[28] A. Heister *et al.* (ALEPH Collaboration), Eur. Phys. J. C. **29**, 143 (2003)

[29] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **82**, 3576 (1999).

[30] A. Drutskoy (for the Belle Collaboration), *Results from the $\Upsilon(5S)$ engineering run*, hep-ex/0605110 (2006)

[31] H.-G. Moser, A. Roussarie, *Mathematical methods for $B^0\bar{B}^0$ oscillation analysis*, Nucl. Instr. Methods A **384**, 491 (1997)

[32] G. Aubrecht *et al.*, Contemporary Physics Education Project (2003)

[33] C. Schmidt, *The Fermilab 400-MeV linac upgrade*, FERMILAB-CONF-93-111 (1993)

[34] Fermilab Beams Division, *Run II handbook*, `http://www-bd.fnal.gov/runII/index.html`

[35] J. Marriner, *Stochastic cooling overview*, FERMILAB-CONF-03-158 (2003)

[36] R. Blair *et al.* (CDF Collaboration), *The CDF II detector: Technical Design Report*, FERMILAB-PUB-96/390-E (1996); F. Abe *et al.* (CDF Collaboration), Nucl. Instr. and Meth. Phys. Res, **271A**, 387 (1988), FERMILAB-PUB-94/024-E (1994)

[37] T. Nelson *et al.*, FERMILAB-CONF-01/357-E (2001)

[38] A. Sill *et al.*, Nucl. Instr. Methods A **447**, 1–8 (2000)

[39] T. Affolder *et al.*, Nucl. Instr. Methods A **485**, 6–9 (2002)

[40] K. Pitts *et al.*, FERMILAB-CONF-96-443-E (1996)

[41] R. Wagner *et al.*, *CDF Central Outer Tracker*, CDF Note 6267(2003)

[42] J. Guimarães da Costa, *Beamline constraint in* `CTVMFT`, CDF Note 6439;J. Marriner, *Secondary Vertex Fit with Mass and Pointing Constraints* (`CTVMFT`), CDF Note 1996(1993)

[43] `http://huhepl.harvard.edu/~cmx/cdfnotes/`

[44] D. Acosta *et al.* (CDF Collaboration), *A time-of-flight detector in CDF II*, Nucl. Instr. Methods A **518**, 605 (2004)

[45] N. Leonardo *et al.*, *The event builder and level 3 manual*, CDF Note 6138(2002)

[46] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002)

[47] W. Ashmanskas *et al.*, FERMILAB-CONF-02/035-E; A. Bardi *et al.*, Nucl. Inst. Methods Phys. Res. A **485**, 6 (2002)

[48] C. Hays *et al.*, *The COT pattern recognition algorithm and offline code*, CDF Note 6992;C. Hays *et al.*, Nucl. Instr. Methods A **538**, 249 (2005); Y. Huang *et al.*, *Inside-out tracking*, CDF Note 6707(1993)

[49] A. Korn, R. Snyder, J. Boudreau, `cdfcodebrowser.fnal.gov/CdfCode/source/TrackingUserHL/`

[50] M. Feindt, S. Menzemer, K. Rinnert, `TrackingKal`–*a tracking and alignment software package for the CDF II silicon detector*, CDF Note 5968(2002)

[51] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D71, 052003 (2005); K. Burkett, J. Guimarães da Costa, D. Sherman, CDF Note 7859(2005)

[52] Ch. Paus *et al.*, *BottomMods*, `cdfcodebrowser.fnal.gov/CdfCode/source/BottomMods/`

[53] P. Murat *et al.*, *Stntuple*, `cdfcodebrowser.fnal.gov/CdfCode/source/Stntuple/`

[54] G. Giurgiu *et al.*, *Muon B flavor tagging - a likelihood approach*, CDF Note 7043, `www-cdf.fnal.gov/physics/new/bottom/072204.blessed-like-mu-tag/`

[55] V. Tiwari *et al.*, *Likelihood based electron tagging*, CDF Note 7121, `www-cdf.fnal.gov/physics/new/bottom/081204.blessed-like-el-tag/`

[56] A. Korn, *Measurement of the B hadron masses in exclusive J/ψ decay channels*, CDF Note 6963(2004)

[57] K. Anikeev, *Measurement of the lifetimes of B meson mass eigenstates*, CDF Note 7269(2004)

[58] I. Furić, *Measurement of the ratio of branching fractions* $\mathrm{BR}\,(B_s \to D_s^- \pi^+)\,/\,\mathrm{BR}\,(B^0 \to D^- \pi^+)$, CDF Note 7352(2004)

[59] N. Leonardo, *Redefining particle decays*, CDF Note 6425(2003), `cdfcodebrowser.fnal.gov/CdfCode/source/qq_i/decdef/`

[60] P. Sphicas (1994), *A $b\bar{b}$ Monte Carlo generator*, CDF Note 2655; K. Anikeev, P. Murat, Ch. Paus, *Description of Bgenerator II*, CDF Note 5092(1999)

[61] P. Nason, S. Dawson and R.K. Ellis, Nucl. Phys. **B303**, 607 (1988); Nucl. Phys. **B327** 49 (1989)

[62] C. Peterson *et al.*, Phys. Rev. **D27**, 105 (1983)

[63] W. Bell, J. Fernandez, L. Flores, F. Würthwein, R. Tesarek, *User guide for*`EvtGen` *CDF*, CDF Note 5618(2001); `EvtGen`–*a Monte Carlo generator for B physics*, `http://www.slac.stanford.edu/~lange/EvtGen/`

[64] R. Brun, K. Hakelberg, M. Hansroul, and J. Lasalle, `GEANT`–*detector description and simulation tool*, CERN-DD-78-2(-REV)

[65] F. James, `Minuit`–*function minimization and error analysis*, CERN Program Library Long Write-up D506, `http://consult.cern.ch/writeup/minuit/`

[66] R. Brun *et al.*, `Root`–*an object-oriented data analysis framework*, `http://root.cern.ch`

[67] N. Leonardo *et al.*, *Methods for likelihood computation in lifetime and mixing analyses*, CDF Note 7385(2004)

[68] C. Chen, *A measurement of the direct charm meson production cross section at CDF II*, Ph.D thesis, University of Pennsylvania (2003)

[69] I. Kravchenko, `cdfcodebrowser.fnal.gov/CdfCode/source/JetUserObjects/`

[70] C. Lecci, *A neural jet charge tagger for the measurement of the B_s-$\bar{B}_s$ oscillation frequency at CDF*, Ph.D thesis, University of Karlsruhe (2005)

[71] M. Jones *et al.*, *Sample composition of the l+SVT triggers*, CDF Note 6480(2003)

[72] A. Ali, F. Barreiro, Z. Phys. C **30**, 635 (1986); M. Gronau, A. Nippe, J L. Rosner, Phys. Rev. D **47**, 1988 (1993); M. Gronau, J L. Rosner, Phys. Rev. D **49**, 254 (1994)

[73] T. Sjöstrand *et al.*, `Pythia`, Computer Phys. Commun. **135**, 238 (2001)

[74] R. Field, *The sources of b quarks at the Tevatron and their correlations*, Phys. Rev. D **65**, 094006 (2002)

[75] M. Bona *et al.*, UTfit Collaboration, JHEP **0107**, 013 (2001); JHEP **0507**, 028 (2005); JHEP **0603**, 080 (2006); `http://utfit.roma1.infn.it/`

[76] J. Charles *et al.*, CKMfitter Collaboration, Eur. Phys. J. C**21**, 225 (2001); Eur. Phys. J. C**41**, 1 (2005); `http://ckmfitter.in2p3.fr/`

[77] A. Ali, D. London, Eur. Phys. J. C**18** (2001) 665; S. Mele, Phys. Rev. D **59** (1999) 113011; D. Atwood, A. Soni, Phys. Lett. B **508** (2001) 17.

[78] Y. Grossman, Y. Nir, S. Plaszczynski, M.-H. Schune, Nucl. Phys. B **511** (1998) 69; S. Plaszczynski, M.-H. Schune, hep-ph/9911280 (1999)

[79] Courtesy of UT fit Collaboration, many thanks to M. Bona, M. Pierini, A. Stocchi

[80] L. Velasco-Sevilla, N. Leonardo, *Impact of Δm_{B_s} on the determination of the unitary triangle and bounds on physics beyond the standard model*, in preparation (2006)

[81] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 062003 (2006)

[82] A. Abulencia *et al.* (CDF Collaboration), submitted to Phys. Rev. Lett., hep-ex/0609040 (2006)

[83] A. Abulencia *et al.* (CDF Collaboration), to be submitted to Phys. Rev. D (2006)

[84] V. Weisskopf, E. Wigner, Z. Phys. **63**, 54 (1930)

[85] R. Waldi, *Flavour oscillation and CP violation of B mesons*, M. Beyer (ed.), *CP violation in particle, nuclear, and astrophysics*, Springer (2002)

[86] T. Inami, C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981) [Erratum-ibid. **65**, 1772 (1981)].

[87] G. Boix, D. Abbaneo, *The B_s amplitude analysis*, JHEP **9908**, 004 (1999)

[88] F. Bedeschi, *B_s mixing in frequency space*, CDF Note 7871(2005); H.-G. Moser, Nucl. Instr. Methods A **295**, 435 (1990)

[89] P. Sphicas, *Combining flavor taggers*, CDF Note 3425(1995)

[90] Wolfram Research, `Mathematica` 5.0 for Mac OS X.