Karlsruhe Institute of Technology

Flavour Tagging Calibration and Measurement of B_s Oscillations and CP Asymmetry

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Physik des
Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Phys. Jan Morlock

aus

Pforzheim

Tag der Prüfung:	18. Juni 2010
Referent:	Prof. Dr. M. Feindt
Korreferent:	Prof. Dr. G. Quast

Contents

List of Tables

List of Figures

1 Introduction

The field of particle physics explores elementary constituents of matter and interactions between them. During the last century remarkable achievements were made in this field - smaller and smaller objects could be resolved. In this sense, the research objects gradually shifted from molecules over atoms and nucleons to point-like particles, which are assumed to be elementary. Interactions between these constituents are mediated by four fundamental forces. Three of them are described by the Standard Model of Particle Physics. Judged by its ability to describe observations and its capability to make predictions, this model is extremely successful.

According to the standard model the constituents of matter are quarks and leptons. Both groups consist of six different species, so called flavours, which may appear as particle or anti-particle. Quarks never appear as free particles. Only bound states of quark and anti-quark, called mesons, and three quarks, called baryons, are observed. The proton, for example, is composed of two up (u) and one down (d) quark.

The focal point of this thesis is the B_s^0 meson. It is composed of an anti-bottom-quark and a strange-quark. It has a mass of 5366.3 MeV/c^2 and its mean lifetime amounts approximately 1.47 ps^{-1}. Besides, it possesses a very interesting feature: it continuously transforms itself into its own anti-particle. This phenomenon is referred to as flavour oscillation or quark mixing.

The frequency of this process Δm_s is an important physics parameter. So far, its most precise measurement was achieved in 2006 by the CDF collaboration [5] as

$$\Delta m_s = (17.77 \pm 0.10(\text{stat.}) \pm 0.07(\text{syst.}))\ \text{ps}^{-1}. \tag{1.1}$$

Apart from that, many properties of the B_s^0 are still unexplored. Therefore it is a subject of many studies. One of those studies is the examination of the B_s^0 system with respect to a fundamental symmetry called $\mathcal{CP}$. Previous studies already showed that this symmetry is violated in other meson systems. The measurement has many aspects and is performed as a common effort within the CDF-II collaboration. One of its features is that it uses the assessment of the flavour eigenstate of the B_s^0 meson at production time to increase its sensitivity to the parameters of interest. Procedures providing this information are called flavour taggers. Their output values are referred to as tagging decision and dilution. The former is an integer number, defined as +1 for a B meson, −1 for $\bar{B}$ and 0 if a decision could not be reached. The latter is related to the probability for the tagging decision to be correct.

It shall be understood that the event-by-event assignment of those probabilities is a non-trivial task. On the other hand analyses using this information depend on the correct assessment of these quantities. Therefore a calibration of the flavour tagger must be performed. This is one major topic of this thesis.

The calibration consists of a determination of the B_s mixing amplitude on measured data. This is a technical parameter with a range of values which is interpreted in the following way: an amplitude consistent with one means that the tagger assesses its decisions and thus its performance correctly. A value smaller than one indicates that a given tagger overestimates itself. According to that, a value greater than one implies an underestimation of the decision power. After the mixing amplitude is determined, all event-by-event dilution values are multiplied by it. In this way a highly reliable tagging information is achieved. The mixing frequency is determined at the same time. In the context of this calibration, it can be seen as a test quantity providing cross-checks.

The calibration is carried out in a similar way as the original measurement of the mixing frequency. This affords an interesting opportunity: in comparison with the last measurement of the mixing frequency, nearly five times the amount of raw data is now available. The second topic of this thesis is therefore to provide an update of the mixing frequency presented above.

In a nutshell, the thesis at hand has two aims: the first one is to calibrate a tagger using B_s^0 oscillations and the second one is to improve the result (equation 1.1) on the mixing frequency itself.

This thesis continues in chapter 2 with a theoretical introduction on mixing and $\mathcal{CP}$ violation. The data used for this analysis was measured at the Collider Detector at Fermilab located at the Tevatron. Both are subject of the third chapter. Chapter 4 describes the details of data taking, event selection, simulation and data preprocessing. Afterwards a model is derived in the fifth chapter which contains both physics and nuisance parameters. Chapter 6 describes how those nuisance parameters are determined step-by-step. It ends with an estimation on the mean lifetime of the B_s^0 meson. Chapter 7 continues at this point with the measurement of the mixing amplitude. It is one of the input parameters of the measurement of the $\mathcal{CP}$-violating phase β_s. A short outline of this measurement and up-to-date results are provided in chapter 8. The final measurement of the mixing frequency is performed with a different flavour tagger configuration. This is presented in chapter 9. A conclusion of this analysis is drawn in the last chapter.

2 Theoretical Overview

2.1 Standard Model

The Standard Model (SM) of Particle Physics [6] is a quantum field theory which was established in the 1970s. Until then four distinct forces were known, which could be characterised by their different ranges of interaction and their relative strengths[1] [7]: the strong force has a range of about 10^{-15} m. The weak force, known from radioactive decays, has a range of 10^{-17} m and is about 10^{-7} times weaker. The strength of the electromagnetic force is determined by the fine structure constant $\alpha \approx 10^{-2}$ and has an infinite range. Gravity, the fourth force, also has an infinite range. Its relative size, about 10^{-39} in comparison with the strong force, is, to date, too weak to be observable in laboratory experiments of particle physics.

The achievement of the standard model was the elaboration of a common description of the strong, weak and electromagnetic forces in the language of a quantum gauge field theory. In addition, it combines the weak and electromagnetic forces in a single electroweak force.

The standard model is based on elementary matter fields consisting of fermions with spin-1/2 and their anti-particles. They are classified into quarks and leptons. The individual properties of both can be found in table 2.1. The most striking feature is their mass which varies over a large scale. The particles are further arranged into three generations: the first one contains the up (u) and down (d) quark, which are the constituents of nucleons, pions and other mesons responsible for nuclear binding. It also contains the electron (e^-) which builds up atomic shells and the electron neutrino (ν_e) which is emitted with a positron (e^+) in the nuclear β^+-decay.

The quarks of the second and third generation are unstable: the strange (s), the charm (c) and the bottom quark (b) appear as constituents of heavier short-lived particles. The most massive quark, the top quark (t), does not even form a bound state with other elementary particles, it decays within a time scale in the order of 10^{-24} s. Also the charged leptons of these generations, the muon (μ) and the tau (τ), decay to particles of the first family.

---

[1] The comparison is drawn assuming two protons at rest which are just in contact with each other.

			Mass [c^{-2}]	**Charge [e]**	**Generation**
Quarks	u	up	1.5 to 3.3 MeV	+2/3	1
	d	down	3.5 to 6.0 MeV	-1/3	1
	s	strange	105^{+25}_{-35} MeV	+2/3	2
	c	charm	$1.27^{+0.07}_{-0.11}$ GeV	-1/3	2
	b	bottom	$4.20^{+0.17}_{-0.07}$ GeV	+2/3	3
	t	top	$171.3 \pm 1.1 \pm 1.2$ GeV	-1/3	3
Leptons	e	electron	$0.510998910 \pm (1.3 \cdot 10^{-8})$ MeV	-1	1
	ν_e	e-neutrino	< 2 eV	0	1
	μ	muon	$105.6583668 \pm (3.8 \cdot 10^{-6})$ MeV	-1	2
	ν_μ	μ-neutrino	< 0.19 MeV	0	2
	τ	tau	1776.84 ± 0.17 MeV	-1	3
	ν_τ	τ-neutrino	< 18.2 MeV	0	3

Table 2.1: Mass, charge and generation of the quarks and leptons. The masses were taken from Ref. [1]. Upper limits are provided for the neutrino masses, which are assumed to be zero within the standard model.

Spin-1 gauge bosons are responsible for all interactions described by the standard model, including decays or scatter processes, between the mass fields described above. A particle is affected by a given fundamental force if it possesses a certain property or quantum number the force couples to. This is further described in the following.

The theory of Quantum Chromodynamics (QCD) deals with the strong force and underlies a $SU(3)$ symmetry group. The strong force is mediated by the exchange of massless, electrically neutral bosons called gluons between particles that carry a quantum number which is referred to as colour. This theory requires eight different gluons, $g_1, g_2, \ldots, g_8$. Colour is not only carried by quarks but also by the gluons themselves. As a consequence, the potential energy between two coloured particles increases significantly with increasing distance. Thus, quarks and gluons cannot appear as free particles, but exist only inside composite objects, called hadrons, with no net colour charge. This property is called confinement. Three experimentally indistinguishable colours are existing which are by convention referred to as red, green and blue for particles. Accordingly, anti-particles carry anti-red, anti-blue and anti-green. In this way a variety of bound colour-neutral quark states is conceivable. However, only two combinations were observed: baryons (e.g. protons) are composed of three quarks of different colours, mesons (e.g. pions) contain quark and anti-quark.

The field theory dealing with the electromagnetic force is referred to as Quantum Electrodynamics (QED). It is mediated by the photon (γ), an electrically uncharged, massless particle. The photon couples to all charged particles such as quarks, charged leptons or the $W^{\pm}$-bosons introduced below. The fact that it is massless accounts for the unlimited range of electromagnetic interaction.

The weak force is mediated by the electrically charged $W^{\pm}$ and the neutral Z^0 bosons. They possess the following masses [1]:

$$m(W^{\pm}) = 80.398 \pm 0.025 \text{ GeV/c}^2, \tag{2.1}$$

$$m(Z^0) = 91.1876 \pm 0.0021 \text{ GeV/c}^2. \tag{2.2}$$

The magnitude of these values is the reason why the weak force, in contrast to the other fundamental forces, does not form bound states. The Z^0 couples to both leptons and quarks. The $W^\pm$ bosons couple only to fermions with negative chirality. This violation of parity is further exemplified in the next section.

Because the $W^\pm$ particles carry electric charge, they couple to γ. This observation already implies a gauge theory that unites the weak and electromagnetic interactions. It is commonly referred to as electroweak force and underlies a $SU(2) \otimes U(1)$ symmetry group. In detail, it consists of the fields W^1, W^2, W^3 and B. The original charged $W^\pm$ bosons are defined by

$$\begin{pmatrix} W^+ \\ W^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} W^1 \\ W^2 \end{pmatrix}. \tag{2.3}$$

By using the weak mixing angle $\sin^2(\theta_W) = 0.23119(14)$, the neutral bosons of the electromagnetic and weak interaction can be defined as

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix}. \tag{2.4}$$

The fact that the weak and the electromagnetic force manifest with different properties at low energies is referred to as electroweak symmetry breaking. In the standard model it is explained by the presence of the so-called Higgs field. It is introduced as a scalar field $\phi = |\phi| e^{i\theta}$ with a potential energy density function of

$$V(|\phi|^2) = -\mu^2|\phi|^2 + \lambda|\phi|^4. \tag{2.5}$$

The expectation value has its minimum at $\sqrt{2}\langle|\phi|\rangle = \mu/\sqrt{\lambda} = v$, different from zero and independent of θ. Quantum excitations of $|\phi|$ about its vacuum value are Higgs scalars with mass $m_H^2 = 2\mu^2 = 2\lambda v^2$.

The fundamental character of this field becomes clear in its interactions with the other particles: all particles which couple to the Higgs field cannot propagate at the speed of light - they acquire masses. Particles that do not interact, e.g. the photon or the gluons, remain massless. The origin of the Higgs field has not yet been determined, neither has there been a direct observation.

The standard model has been experimentally verified to a high degree of accuracy over a broad range of energy and processes. A fundamental contradiction to its predictions has not yet been observed. However, there are still many unresolved issues. Future analyses or new experiments, like the LHC, may shed light on them.

2.2 Symmetries

The examination of a model with respect to elementary symmetries plays a central role in physics. The reason for this is Noether's theorem which states that if a system has a symmetry property, then there are corresponding quantities whose values are conserved in time.

In the following three operations are defined which play an important role. They are applied on particles with momentum p which are represented by their wave-function $|\Psi(p)\rangle$.

- The $\mathcal{C}$ operator turns a particle into its anti-particle. The sign of the charge is reversed.

$$\mathcal{C}|\Psi(p)\rangle = |\bar{\Psi}(p)\rangle. \tag{2.6}$$

- The $\mathcal{P}$ operation performs a spatial inversion of coordinates

$$\mathcal{P}|\Psi(p)\rangle = |\Psi(-p)\rangle \tag{2.7}$$

- The $\mathcal{T}$ operation corresponds to a reversal of time. Momenta are hereby inverted

$$\mathcal{T}|\Psi(p)\rangle = |\Psi(-p)\rangle^*. \tag{2.8}$$

All three operations have the following properties: if one is applied twice, the effect is neutralised $\mathcal{C}^2 = \mathcal{P}^2 = \mathcal{T}^2 = 1$. Further all three operations permute.

The central question is the following: does a system, where one or more of the above operations were applied, behave exactly as the original one?

The answer is 'yes' for the concatenation of all three operations. This is known as the $\mathcal{CPT}$ theorem of Lüders-Pauli [8]. It is one of the most fundamental laws of particle physics and applies to every quantum field theory following the principles of locality and Lorentz invariance. It also implies that mass and lifetime are the same for particle and antiparticle. Further the magnetic moment has the same magnitude but opposite direction for both. locality, and

The answer is also 'yes' for all processes which are mediated by either the strong or the electromagnetic force. However, the weak force violates both the $\mathcal{P}$ and the $\mathcal{C}$ symmetry in a maximal way. In addition, the concatenation of both operations, $\mathcal{CP}$, is violated as well. The different kinds of $\mathcal{CP}$ violation are further explained in section 2.5.

The conservation of $\mathcal{CPT}$ and the violation of $\mathcal{CP}$ implies that the $\mathcal{T}$ symmetry must also be violated in exactly the same way. The first direct observation of this behaviour was provided by the CPLEAR collaboration [9].

From a theoretical point-of-view it is still a mystery why the weak force is the only source of $\mathcal{CP}$ violation in the standard model. To be precise, $\mathcal{CP}$ could be violated also by strong interactions. However, it is known from experiment that Θ_{QCD}, the non-perturbative parameter that determines the strength of this type of $\mathcal{CP}$ violation, is tiny, if not zero. It is therefore irrelevant to the discussions following.

2.3 The Cabibbo-Kobayashi-Maskawa Matrix

The charged current interactions of the weak force [1] for quarks are given by the Lagrangian

$$-\mathcal{L}_{W^\pm} = \frac{g}{\sqrt{2}} \bar{u_{Li}} \gamma^\mu (V_{CKM})_{ij} d_{Lj} W_\mu^+ + h.c. \tag{2.9}$$

Here $i, j = 1, 2, 3$ are the generation numbers of the quarks, which are ordered by their masses: $(u_1, u_2, u_3) \rightarrow (u, c, t)$ and $(d_1, d_2, d_3) \rightarrow (d, s, b)$. The gauge coupling constant is denoted as g. The above definition introduces the Cabibbo-Kobayashi-Maskawa matrix V_{CKM}. The most general notation of this 3×3 complex matrix is the following:

$$\mathbf{V_{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.10}$$

The CKM matrix connects the weak eigenstates to the flavour eigenstates. Expressed in simple terms, the probability for quark i $(i = u, c, t)$ to become quark j $(j = d, s, b)$ is proportional to the square of V_{ij}. In order to conserve probability it must be unitary:

$$\mathbf{V_{CKM}}\mathbf{V_{CKM}}^{\dagger} = \mathbf{V_{CKM}}^{\dagger}\mathbf{V_{CKM}} = 1. \tag{2.11}$$

This implies the following relations:

$$\sum_{j=1}^{3} |V_{ij}|^2 = 1 \quad \text{for} \quad i = 1, 2, 3, \tag{2.12}$$

$$\sum_{k=1}^{3} V_{ik}^{*} V_{jk} = 0 \quad \text{for} \quad i \neq j. \tag{2.13}$$

In total, nine constraints are acquired for the 18 real- and imaginary parts of the matrix. In this way the number of free parameters is reduced to nine. They can be interpreted as three angles and six phases. Furthermore, the physics of the standard model is invariant under a phase transformation of the fields:

$$d_j \longrightarrow e^{i\phi_{d,j}} d_j, \qquad u_i \longrightarrow e^{i\phi_{u,i}} u_i. \tag{2.14}$$

This implies that the corresponding CKM element may also be rotated by a given phase:

$$V_{ij} \longrightarrow V_{ij} e^{i(\phi_{d,j} - \phi_{u,i})} \tag{2.15}$$

The above invariance can be used to eliminate as many phases as possible: six phases imply five independent phase differences $\phi_{d,j} - \phi_{u,i}$. This means that five of the six phases in V_{CKM} are unobservable and can be rotated away. The physics one is referred to as Kobayashi-Maskawa phase δ. It is responsible for all $\mathcal{CP}$ violating phenomena in the standard model. So the fact that one can express V_{CKM} by three angles and one phase can be made manifest by choosing an explicit parametrisation. One possibility is

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.16}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ [1].

The absolute values of the on-diagonal entries, which represent the coupling between quarks of the same generation, is close to one. In comparison, the couplings between

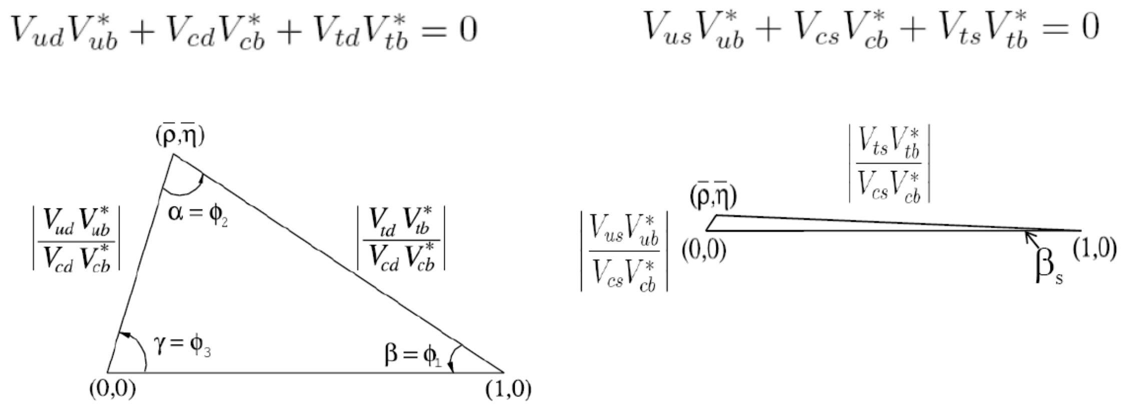


Figure 2.1: Two of the six unitarity triangles are shown with their corresponding relation. The length of the sides are rescaled by dividing each side by the well-known values $V_{cd}V_{cb}^*$ (left-hand side) and $V_{cs}V_{cb}^*$ (right-hand side).

the first and the second generation are suppressed by a factor of $\lambda = \sin\theta_{Cabibbo} = 0.22$. This continues with a factor of λ^2 between the second and third and finally with a factor of λ^3 between the first and third generation. The hierarchy explained is considered in the following approximation:

$$\mathbf{V_{CKM}} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.17}$$

It was presented by Lincoln Wolfenstein in 1983 [10] and contains four independent, real-valued parameters A, λ, ρ and η.

Since the entries of the CKM matrix are fundamental parameters of the standard model, their determination is crucial. It can be achieved by measuring processes which are sensitive to the elements or to relations between them. One important set was already given in expression 2.13, which stands for six equations. Each one can be represented by a triangle in the complex plane. The triangles which play a role in the thesis at hand can be found in figure 2.1 and are further described in sections 2.6 and 2.7.

2.4 Time Evolution

There are several bases of eigenstates which are used to describe the neutral B_s^0 meson system [11]. The flavour eigenstates are defined by the quark content. They are

$$|B_s^0\rangle = |\bar{b}s\rangle \qquad \text{and} \qquad |\bar{B}_s^0\rangle = |b\bar{s}\rangle. \tag{2.18}$$

Their time evolution is given by the Schrödinger equation,

$$i\frac{\partial}{\partial t}\begin{pmatrix} |B_s^0\rangle \\ |\bar{B}_s^0\rangle \end{pmatrix} = \mathbf{H}\begin{pmatrix} |B_s^0\rangle \\ |\bar{B}_s^0\rangle \end{pmatrix}. \tag{2.19}$$

This equation contains the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$, which are both hermitian and 2×2 in size:

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix}. \tag{2.20}$$

$\mathcal{CPT}$ invariance (see section 2.2) requires $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. The latter is referred to as decay width and can be related to the mean lifetime as $\tau = 1/\Gamma$. The eigenvalues of $\mathbf{H}$ are

$$\lambda_L = (M - \frac{i}{2}\Gamma) + \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}), \tag{2.21}$$

$$\lambda_H = (M - \frac{i}{2}\Gamma) - \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}). \tag{2.22}$$

The indices H and L refer to the heavy and light mass eigenstates which are defined below. The substitution q/p introduced above is defined as

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.23}$$

This definition is helpful when it comes to the eigenvectors of the effective Hamiltonian, which are

$$|B_s^L\rangle = p\ |B_s^0\rangle + q\ |\bar{B}_s^0\rangle, \tag{2.24}$$

$$|B_s^H\rangle = p\ |B_s^0\rangle - q\ |\bar{B}_s^0\rangle. \tag{2.25}$$

In order to preserve normalisation, the complex parameters p and q must fulfil the requirement $|p|^2 + |q|^2 = 1$.

The masses of both eigenstates, m_L and m_H, are the real parts of their respective eigenvalues. Similarly, the decay widths, Γ_L and Γ_H, are the negative half of the imaginary parts:

$$\lambda_L = m_L - \frac{i}{2}\Gamma_L, \tag{2.26}$$

$$\lambda_H = m_H - \frac{i}{2}\Gamma_H. \tag{2.27}$$

Mass and decay width difference between the heavy and light mass eigenstates are defined as

$$\Delta m = m_H - m_L \qquad \text{and} \qquad \Delta\Gamma = \Gamma_L - \Gamma_H. \tag{2.28}$$

By taking the difference between the complex-valued equations 2.27 and 2.26, two real-valued equations can be derived which connect the mass and decay width differences to the off-diagonal elements of the effective Hamiltonian:

$$\Delta m^2 - \frac{1}{4}\Delta\Gamma^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2, \tag{2.29}$$

$$\Delta m \cdot \Delta\Gamma = -\,\mathrm{Re}(M_{12}\Gamma_{12}^*). \tag{2.30}$$

Using the definitions above, the time evolution of the mass eigenstates is

$$|B_s^L(t)\rangle = |B_s^L\rangle \; e^{-im_L t - \frac{1}{2}\Gamma_L t}, \tag{2.31}$$

$$|B_s^H(t)\rangle = |B_s^H\rangle \; e^{-im^H t - \frac{1}{2}\Gamma_H t}. \tag{2.32}$$

Both formulae can be translated in order to express the time evolution of the flavour eigenstates:

$$|B_s(t)\rangle = g_+(t)|B_s\rangle + \frac{q}{p} g_-(t)|\bar{B}_s\rangle, \tag{2.33}$$

$$|\bar{B}_s(t)\rangle = \frac{p}{q} g_-(t)|B_s\rangle + g_+(t)|\bar{B}_s\rangle. \tag{2.34}$$

The expression $g_\pm(t) = \frac{1}{2}(e^{-i\lambda_L t} \pm e^{-i\lambda_H t})$ was hereby used as substitution. The following two relations between these time evolution functions are later helpful for providing expectation values:

$$|g_\pm(t)|^2 = \frac{1}{2}\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) \pm \cos(\Delta m t)\right] e^{-\Gamma t}, \tag{2.35}$$

$$g_+^*(t) g_-(t) = \frac{1}{2}\left[-\sinh\left(\frac{\Delta\Gamma}{2}t\right) + i\sin(\Delta m t)\right] e^{-\Gamma t}. \tag{2.36}$$

With this it is possible to determine the decay rate $\Gamma(B_s^0(t) \to f)$ [11] of a meson produced as B_s^0 into some final state f. It is defined as

$$\Gamma(B_s^0(t) \to f) = \frac{1}{N_B}\frac{dN(B_s^0(t) \to f)}{dt}. \tag{2.37}$$

Here, N_B is the total number of B_s^0 mesons produced at time $t = 0$. $dN(B_s^0(t) \to f)$ denotes the number of decays of initial B_s^0 mesons into the final state f occurring within the time interval between t and $t + dt$. An analogous definition holds for $\Gamma(\bar{B}_s^0(t) \to f)$. Both can be expressed in notation of quantum mechanics as

$$\Gamma(B_s^0(t) \to f) = N_f |\langle f|B_s^0(t)\rangle|^2, \tag{2.38}$$

$$\Gamma(\bar{B}_s^0(t) \to f) = N_f |\langle f|\bar{B}_s^0(t)\rangle|^2. \tag{2.39}$$

Similar equations can be given for the $\mathcal{CP}$-conjugate final state of $|\bar{f}\rangle = \mathcal{CP}|f\rangle$:

$$\Gamma(B_s^0(t) \to \bar{f}) = N_f |\langle \bar{f}|B_s^0(t)\rangle|^2, \tag{2.40}$$

$$\Gamma(\bar{B}_s^0(t) \to \bar{f}) = N_f |\langle \bar{f}|\bar{B}_s^0(t)\rangle|^2. \tag{2.41}$$

N_f is a time-independent normalisation factor which has the same value in all four equations. Before providing a general expression, it is helpful to define the decay amplitudes

$$A_f = \langle f|B_s^0\rangle,\; A_{\bar{f}} = \langle \bar{f}|B_s^0\rangle,\; \bar{A}_f = \langle f|\bar{B}_s^0\rangle \text{ and } \bar{A}_{\bar{f}} = \langle \bar{f}|\bar{B}_s^0\rangle. \tag{2.42}$$

Further the following ratios are often used to classify the different kinds of $\mathcal{CP}$-violation:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \qquad \text{and} \qquad \lambda_{\bar{f}}^{-1} = \frac{p}{q}\frac{A_{\bar{f}}}{\bar{A}_{\bar{f}}}. \tag{2.43}$$

	$\boldsymbol{K^0/\bar{K}^0}$	$\boldsymbol{D^0/\bar{D}^0}$	$\boldsymbol{B^0/\bar{B}^0}$	$\boldsymbol{B^0_s/\bar{B}^0_s}$
$\boldsymbol{\tau}$ $[\boldsymbol{ps}]$	89.53 ± 0.05	0.4101 ± 0.0015	1.525 ± 0.009	$1.472^{+0.024}_{-0.026}$
	51160 ± 200			
$\boldsymbol{\Gamma}$ $[\boldsymbol{ps^{-1}}]$	$5.594 \cdot 10^{-3}$	2.438	0.656	0.679
$\lvert\boldsymbol{y}\rvert$	0.9965	$0.0078^{+0.0018}_{-0.0019}$	$\leq 0.01^*$	0.005 ± 0.019
$\boldsymbol{\Delta m}$ $[\boldsymbol{ps^{-1}}]$	$(5.290 \pm 0.015) \cdot 10^{-3}$	$0.0237^{+0.0066}_{-0.0071}$	0.507 ± 0.005	17.77 ± 0.12
$\boldsymbol{x}$	0.9461	0.00972	0.776 ± 0.008	26.1 ± 0.5

Table 2.2: Parameters of the four neutral oscillating meson pairs. Theoretical expectation values are denoted by an asterisk.

With this it is finally possible to provide general expressions for the four decay rates:

$$\begin{aligned}\Gamma(B^0_s(t) \to f) = N_f|A_f|^2 e^{-\Gamma t}\biggl[&\frac{1+|\lambda_f|^2}{2}\cosh\frac{\Delta\Gamma t}{2} + \frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) \\ &- \mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma t}{2} - \mathrm{Im}(\lambda_f)\sin(\Delta m t)\biggr],\end{aligned} \tag{2.44}$$

$$\begin{aligned}\Gamma(\bar{B}^0_s(t) \to f) = N_f|A_f|^2\left|\frac{p}{q}\right|^2 e^{-\Gamma t}\biggl[&\frac{1+|\lambda_f|^2}{2}\cosh\frac{\Delta\Gamma t}{2} - \frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) \\ &- \mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma t}{2} + \mathrm{Im}(\lambda_f)\sin(\Delta m t)\biggr],\end{aligned} \tag{2.45}$$

$$\begin{aligned}\Gamma(B^0_s(t) \to \bar{f}) = N_f|\bar{A}_{\bar{f}}|^2\left|\frac{q}{p}\right|^2 e^{-\Gamma t}\biggl[&\frac{1+|\lambda_{\bar{f}}^{-1}|^2}{2}\cosh\frac{\Delta\Gamma t}{2} - \frac{1-|\lambda_{\bar{f}}^{-1}|^2}{2}\cos(\Delta m t) \\ &- \mathrm{Re}(\lambda_{\bar{f}}^{-1})\sinh\frac{\Delta\Gamma t}{2} + \mathrm{Im}(\lambda_{\bar{f}}^{-1})\sin(\Delta m t)\biggr],\end{aligned} \tag{2.46}$$

$$\begin{aligned}\Gamma(\bar{B}^0_s(t) \to \bar{f}) = N_f|\bar{A}_{\bar{f}}|^2 e^{-\Gamma t}\biggl[&\frac{1+|\lambda_{\bar{f}}^{-1}|^2}{2}\cosh\frac{\Delta\Gamma t}{2} + \frac{1-|\lambda_{\bar{f}}^{-1}|^2}{2}\cos(\Delta m t) \\ &- \mathrm{Re}(\lambda_{\bar{f}}^{-1})\sinh\frac{\Delta\Gamma t}{2} - \mathrm{Im}(\lambda_{\bar{f}}^{-1})\sin(\Delta m t)\biggr].\end{aligned} \tag{2.47}$$

The above equations were derived without using approximations. They can therefore be applied on any neutral meson where mass eigenstates differ from flavour eigenstates. Within the standard model those are, besides of the $|B^0_s\rangle$, also $|B^0\rangle = |\bar{b}d\rangle$, $|D^0\rangle = |c\bar{u}\rangle$ and $|K^0\rangle = |\bar{s}d\rangle$.

However, in some of these meson systems approximations lead to drastic simplifications of the above formulae. It is therefore enlightening to compare the measured values Γ, Δm and $\Delta\Gamma$ for the four mentioned mesons. This is done in table 2.2. The following dimensionless quantities are hereby used for comparison:

$$x = \frac{\Delta m}{\Gamma} \qquad \text{and} \qquad y = \frac{\Delta\Gamma}{2\Gamma}. \tag{2.48}$$

2.5 Classifying $\mathcal{CP}$-Violation

The complex parameter λ_f which was introduced in the previous section as

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \tag{2.49}$$

plays a pivotal role in classification of $\mathcal{CP}$ violation: if it deviates from the real value of ±1, it is proved that the given system violates $\mathcal{CP}$ symmetry.

If the reason for this are different decay amplitudes $|\bar{A}_f| \neq |A_f|$, it is called $\mathcal{CP}$ violation in decay or direct $\mathcal{CP}$ violation. A good example for this class is the decay of B^0 into $K^+\pi^-$.

If λ_f is different from one because of $|q| \neq |p|$, this type is referred to as $\mathcal{CP}$ violation in mixing or indirect $\mathcal{CP}$ violation. It implies that the transition rates, particle → antiparticle and antiparticle → particle, are different for a given flavour-oscillating meson. The decay $K_L \to \pi^\pm \ell^\mp \nu_\ell$ can be chosen as example for this class.

Finally, a deviation may also arise if the absolute value of λ_f is indeed equal or close to 1 but the imaginary part assumes a non-zero value, $\mathrm{Im}(\lambda_f) \neq 0$. This class is known as mixing-induced $\mathcal{CP}$ violation or $\mathcal{CP}$ violation in the interference between decay and mixing. The primary example is $B^0 \to J/\psi K_S$ which is used to determine the CKM angle

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right). \tag{2.50}$$

It is worth mentioning that a given system may also exhibit more than one kind of $\mathcal{CP}$ violation.

2.6 Mixing Frequency

The measurement of the mixing frequency of the B_s^0 plays a central role for the thesis at hand. It is therefore further explained in this section.

The initial hint for B_s^0 oscillations was obtained by combining observations of the UA1 experiment [12] which was located at the SPS collider with observations of the ARGUS experiment [13] located at the electron-positron collider ring DORIS II at DESY. Afterwards, limits were set by the LEP experiments ALEPH [14], DELPHI [15, 16] and OPAL [17] and the SLD [18] experiment at SLC. The first measurements were performed in 2006 by both Tevatron experiments. To date, the most significant result was achieved by the CDF-II collaboration [5], which measures

$$\Delta m_s = (17.77 \pm 0.10(\mathrm{stat.}) \pm 0.07(\mathrm{syst.}))\ \mathrm{ps}^{-1}. \tag{2.51}$$

The mixing frequency of the B^0 is very well measured by Belle and Babar. The current value is

$$\Delta m_d = (0.507 \pm 0.005)\ \mathrm{ps}^{-1}. \tag{2.52}$$

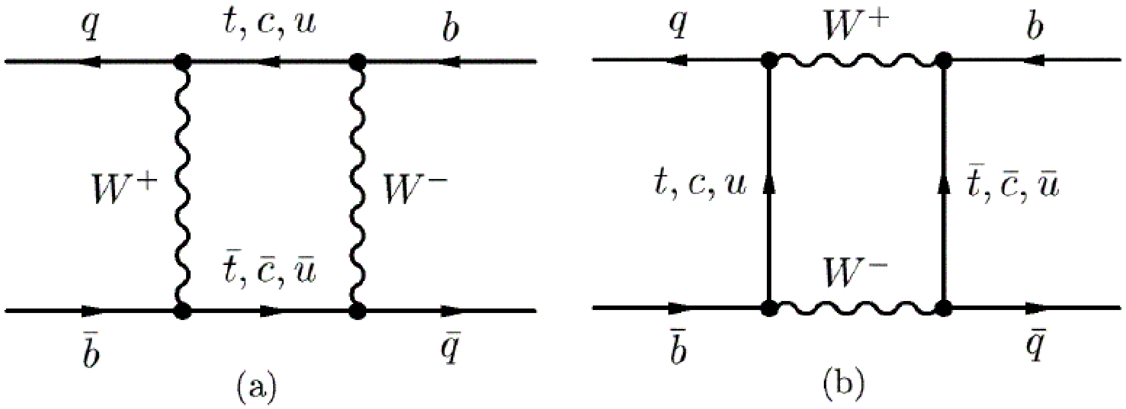


Figure 2.2: Box diagrams for mixing between B_q^0 and $\bar{B}_q^0$ with $q = d, s$.

The importance of the mixing frequencies of both mesons lies in the connection to the CKM elements. As a matter of fact, $|V_{td}|$ and $|V_{ts}|$ cannot be measured from tree-level decays of the top quark [1]. Therefore, the oscillations of the neutral B mesons are used. To be precise, the determination of the ratio of both mixing frequencies gives a clean and useful constraint on the CKM matrix elements mentioned above. Figure 2.2 shows the dominating Feynman diagrams of these processes. The contribution of the top quark is apparent as virtual particle inside the boxes.

The standard model prediction [4] for both mixing frequencies can be expressed by

$$\Delta m_d = \frac{G_F^2}{6\pi^2} \eta_B m_{B_d} f_{B_d}^2 B_d m_W^2 S(x_t) |V_{td} V_{tb}^*|^2, \tag{2.53}$$

$$\Delta m_s = \frac{G_F^2}{6\pi^2} \eta_B m_{B_s} f_{B_s}^2 B_s m_W^2 S(x_t) |V_{ts} V_{tb}^*|^2. \tag{2.54}$$

G_F is here the Fermi coupling constant. The matrix elements $f_{B_q}\sqrt{B_q}$ $(q = d, s)$ are taken from Lattice QCD. η_B denotes a perturbative QCD correction factor to the Inami-Lim function $S(x_t)$ [19] with $x_t := m_t^2/m_W^2$. The masses of the mesons and the $W^\pm$-boson are given by m_{B_d}, m_{B_s} and m_W.

By using the mixing frequency of the B^0 meson, which is precisely known from the experiments Belle and Babar, a constraint can be created on the CKM triangle shown in figure 2.1 on the left side: the factor $|V_{td} V_{tb}^*|$ occurring in equation 2.53 and its standard deviation describe an annulus around $(1, 0)$, which is shown in yellow in figure 2.3 [4]. Additional constraints stemming from other measurements are also plotted.

The mass difference between B_s^H and B_s^L has only a weak relative dependence on the Wolfenstein parameters ρ and λ. Nevertheless, a measurement of Δm_s is useful in the CKM fit since it indirectly leads to an improvement of the constraint from the Δm_d measurement on $|V_{td} V_{tb}^*|^2$. This can be derived by taking the ratio of equation 2.54 and equation 2.53. In this way common factors cancel out and

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s} f_{B_s}^2 B_s}{m_{B_d} f_{B_d}^2 B_d} \left| \frac{V_{ts}}{V_{td}} \right|^2 = \frac{m_{B_s^0}}{m_{B^0}} \xi^2 \left| \frac{V_{ts}}{V_{td}} \right|^2 \tag{2.55}$$

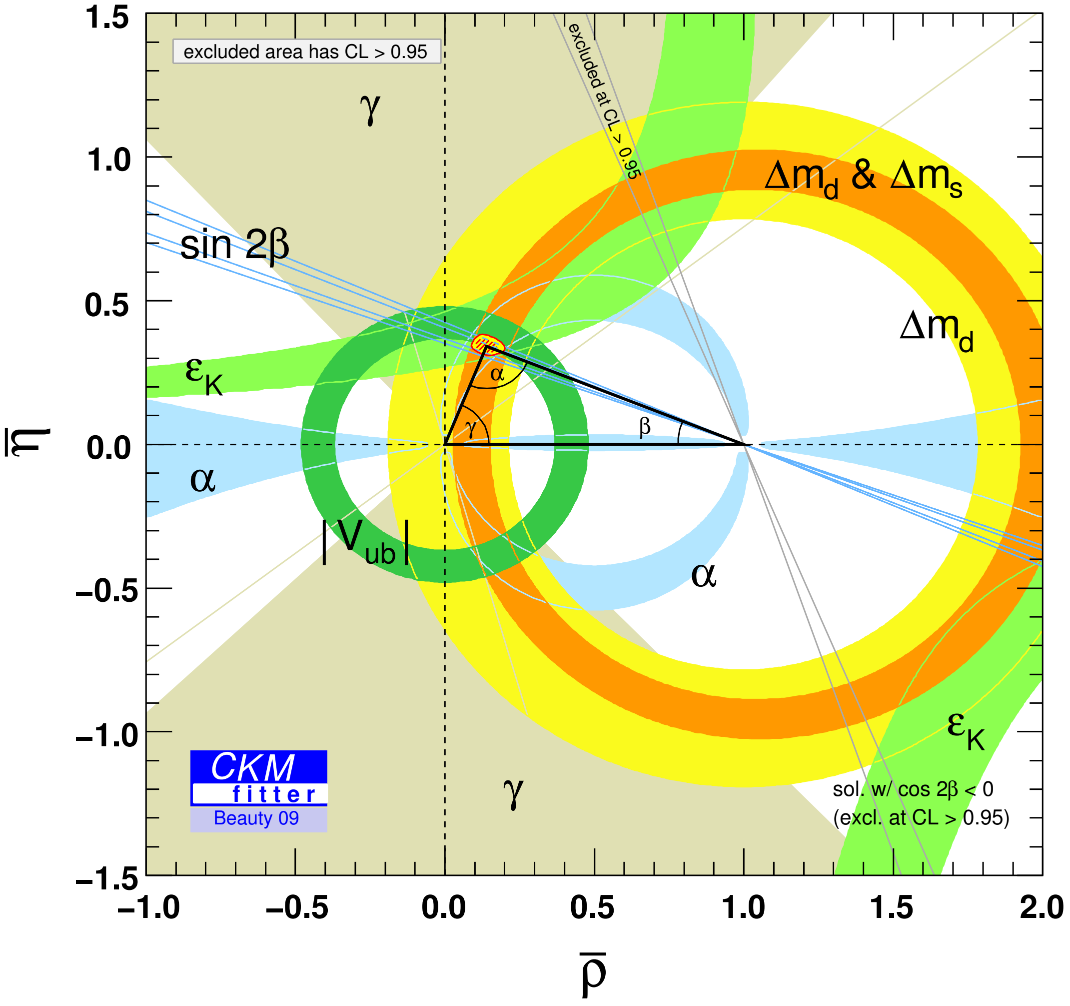


Figure 2.3: Constraints in the $(\bar{\rho},\bar{\eta})$ plane including the most recent inputs in the global CKM fit [4]. The quantities on the axis can be defined via $\bar{\rho}+i\bar{\eta} = -(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$.

is obtained. The parameter $\xi = (f_{B_s}\sqrt{\hat{B}_{B_s}})/(f_{B_d}\sqrt{\hat{B}_{B_d}})$ can be calculated more accurately in Lattice QCD than the matrix elements themselves. Therefore, a measurement of Δm_s improves the knowledge of $f_{B_d}\sqrt{B_d}$ and refines the constraint in the triangle described above. By using the above result on Δm_s and $\xi = 1.21^{+0.047}_{-0.035}$ [5] a ratio of

$$\left|\frac{V_{td}}{V_{ts}}\right| = 0.2060 \pm 0.0007(\text{stat.+syst.})^{+0.0081}_{-0.0060}(\text{theory}) \tag{2.56}$$

is obtained. The corresponding contour is shown in figure 2.3 in orange colour.

Flavour specific final states and their $\mathcal{CP}$ conjugate are appropriate decay channels for the measurement of the mixing frequency. For example the final state $f = D_s^-\pi^+$ can be reached by a B_s^0 but not by a $\bar{B}_s^0$ meson. This implies that the decay flavour of the meson can flawlessly be determined by its final state - usually the pion(s) from the secondary vertex are used.

Coming back to equations 2.44 to 2.47 the decay amplitudes $\bar{A}_f$ and $A_{\bar{f}}$ are zero. In this way λ_f and $\lambda_{\bar{f}}^{-1}$ become also zero and the following equations are obtained:

$$\Gamma(B_s^0(t) \to f) = N_f|A_f|^2 e^{-\Gamma t}\frac{1}{2}\left[\cosh\frac{\Delta\Gamma t}{2} + \cos(\Delta m t)\right], \tag{2.57}$$

$$\Gamma(\bar{B}_s^0(t) \to f) = N_f|A_f|^2\left|\frac{p}{q}\right|^2 e^{-\Gamma t}\frac{1}{2}\left[\cosh\frac{\Delta\Gamma t}{2} - \cos(\Delta m t)\right], \tag{2.58}$$

$$\Gamma(B_s^0(t) \to \bar{f}) = N_f|\bar{A}_{\bar{f}}|^2\left|\frac{q}{p}\right|^2 e^{-\Gamma t}\frac{1}{2}\left[\cosh\frac{\Delta\Gamma t}{2} - \cos(\Delta m t)\right], \tag{2.59}$$

$$\Gamma(\bar{B}_s^0(t) \to \bar{f}) = N_f|\bar{A}_{\bar{f}}|^2 e^{-\Gamma t}\frac{1}{2}\left[\cosh\frac{\Delta\Gamma t}{2} + \cos(\Delta m t)\right]. \tag{2.60}$$

Assuming no direct $\mathcal{CP}$-violation $|A_f|$ and $|\bar{A}_{\bar{f}}|$ can be equated. Two empirical approximations can be made for both neutral B meson systems. Those are $\Delta m \gg \Delta\Gamma$ and $|M_{12}| \gg |\Gamma_{12}|$. Based on these assumptions, the decay width difference is neglected for the measurement of the mixing frequency. It is therefore set to zero and possible biases are evaluated later. This implies $|q/p| = |p/q| = 1$ and the above four equations simplify to two:

$$\Gamma(B_s^0(t) \to f) = \Gamma(\bar{B}_s^0(t) \to \bar{f}) = N_f|A_f|^2 e^{-\Gamma t}\frac{1}{2}\Big[1 + \cos(\Delta m t)\Big] \tag{2.61}$$

$$\Gamma(\bar{B}_s^0(t) \to f) = \Gamma(B_s^0(t) \to \bar{f}) = N_f|A_f|^2 e^{-\Gamma t}\frac{1}{2}\Big[1 - \cos(\Delta m t)\Big] \tag{2.62}$$

In this way only a sensitivity to the mean lifetime and to the mixing frequency remains. Both equations will be picked up again in chapter 5.

2.7 Mixing Phase

One common effort, which is currently made by the CDF-II collaboration, is to measure the $\mathcal{CP}$-violating phase β_s in the B_s^0 system. The thesis at hand makes a significant contribution to this measurement. The following section gives an overview to this topic and puts this work to its context.

The measurement is based on a decay channel where both the B_s^0 and the $\bar{B}_s^0$ are able to decay into the same final state f. In addition, the final state is required to be a $\mathcal{CP}$ eigenstate, in the following denoted as $f_{\mathcal{CP}}$. This means that out of the four decay amplitudes introduced in section 2.4 only two remain. Those are

$$A_{f_{\mathcal{CP}}} \qquad \text{and} \qquad \bar{A}_{f_{\mathcal{CP}}}. \tag{2.63}$$

As a consequence both λ-parameters become the same:

$$\lambda_f = \lambda_{\bar{f}} =: \lambda_{f_{\mathcal{CP}}} = \frac{q}{p}\frac{\bar{A}_{f_{\mathcal{CP}}}}{A_{f_{\mathcal{CP}}}}. \tag{2.64}$$

The decay channel, which is introduced below, is assumed to have only one dominant type of $\mathcal{CP}$ violation. It is interference between mixing and decay (see also section 2.5). As consequence of assuming no $\mathcal{CP}$ violation in decay, the absolute value of both decay amplitudes become the same value:

$$\left|\frac{\bar{A}_{f_{\mathcal{CP}}}}{A_{f_{\mathcal{CP}}}}\right| = 1 \tag{2.65}$$

Furthermore, assuming no $\mathcal{CP}$ violation in mixing implies

$$\left|\frac{q}{p}\right| = 1. \tag{2.66}$$

Combining both equations, it follows that the absolute value of $\lambda_{f_{\mathcal{CP}}}$ is also one:

$$|\lambda_{f_{\mathcal{CP}}}| = 1 \tag{2.67}$$

By applying these definitions on the general equations 2.44 to 2.47, they simplify to two expressions:

$$\Gamma(B_s^0(t) \to f_{\mathcal{CP}}) = N_f|A_{f_{\mathcal{CP}}}|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - \mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma t}{2} - \mathrm{Im}(\lambda_f)\sin(\Delta m t)\right], \tag{2.68}$$

$$\Gamma(\bar{B}_s^0(t) \to f_{\mathcal{CP}}) = N_f|A_{f_{\mathcal{CP}}}|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - \mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma t}{2} + \mathrm{Im}(\lambda_f)\sin(\Delta m t)\right]. \tag{2.69}$$

The B_s^0 decay channel [20] which features the above properties is

$$B_s^0 \to J/\psi\phi,\ J/\psi \to \mu^+\mu^-,\ \phi \to K^+K^-. \tag{2.70}$$

The described interference is illustrated in figure 2.4 using two Feynman diagrams. In both Feynman diagrams the quark sub-process $\bar{b} \to \bar{c}c\bar{s}$ is mediated by a W-tree diagram. This process is predominant and allows for drawing a connection between the CKM angle $\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right)$ and λ_f. The following expression is valid to a very good approximation:

$$\lambda_f = \exp(2i\beta_s). \tag{2.71}$$

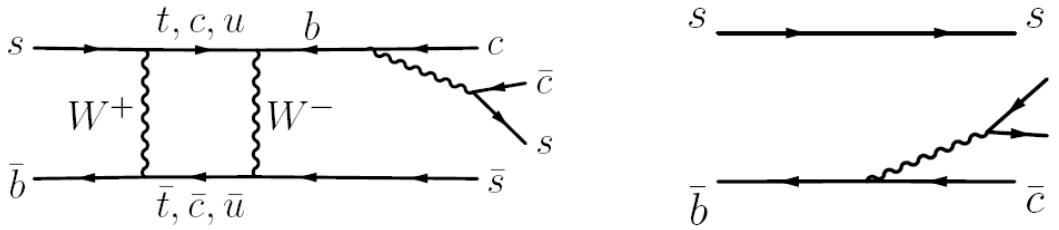


Figure 2.4: The interference between decays with and without mixing is illustrated for the decay channel $B_s^0 \to J/\psi\phi$.

Applied to equations 2.68 and 2.69, a sensitivity to β_s is finally established:

$$\Gamma(B_s^0(t) \to f_{\mathcal{CP}}) = N_f |A_{f_{\mathcal{CP}}}|^2 e^{-\Gamma t} \left[\cosh \frac{\Delta\Gamma t}{2} - \cos(2\beta_s) \sinh \frac{\Delta\Gamma t}{2} - \sin(2\beta_s) \sin(\Delta m t) \right] \tag{2.72}$$

$$\Gamma(\bar{B}_s^0(t) \to f_{\mathcal{CP}}) = N_f |A_{f_{\mathcal{CP}}}|^2 e^{-\Gamma t} \left[\cosh \frac{\Delta\Gamma t}{2} - \cos(2\beta_s) \sinh \frac{\Delta\Gamma t}{2} + \sin(2\beta_s) \sin(\Delta m t) \right] \tag{2.73}$$

Both equations contain the mixing frequency Δm_s, which would also allow for a measurement if $\sin(2\beta_s)$ differs from zero.

Coming back to equations 2.29 and 2.30, it is useful to introduce another phase. Based on the empirical approximations $\Delta m \gg \Delta\Gamma$ and $|M_{12}| \gg |\Gamma_{12}|$, both equations simplify as follows:

$$\Delta m = 2\ |M_{12}| \tag{2.74}$$

$$\Delta\Gamma = 2\ |\Gamma_{12}| \cos(\phi_s). \tag{2.75}$$

In the previous equation, the mixing phase ϕ_s was introduced. It is defined as

$$\phi_s = \arg\left(-\frac{M_{12}}{\Gamma_{12}} \right). \tag{2.76}$$

It can be shown that the existence of new physics (NP) would affect both phases introduced by the same quantity ϕ_s^{NP} [2] as

$$2\beta_s^{J/\psi\phi} = 2\beta_s^{SM} - \phi_s^{NP} \quad \text{and} \quad \phi_s = \phi_s^{SM} + \phi_s^{NP} \tag{2.77}$$

In case ϕ_s^{NP} dominates over the small values $2\beta_s^{SM} \approx 0.04$ and $\phi_s^{SM} \approx 0.0042$, the standard model phases can be neglected. In this way a simple relationship is obtained:

$$2\beta_s^{J/\psi\phi} = -\phi_s^{NP} = -\phi_s \tag{2.78}$$

The expectation values of all three quantities of interest can be found in table 2.3. The discussion of this topic will be continued in chapter 8.

Description	Quantity	Value
CKM angle	β_s	0.020 ± 0.005
Mixing phase	ϕ_s	0.0042 ± 0.0014
Decay width difference	$\Delta\Gamma$	$(0.096 \pm 0.039)\ \mathrm{ps}^{-1}$

Table 2.3: Standard model expectation values [2].

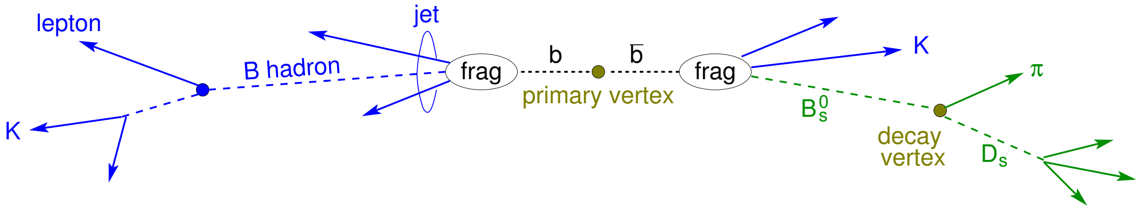


Figure 2.5: Draft of a typical B_s^0 event.

2.8 Tagging and Calibration

The decay rates which were presented in the previous two sections rely on the knowledge of the meson's flavour at production time. Without this information it would not be possible to measure the mixing frequency Δm_s at all. A measurement of the $\mathcal{CP}$ violating phase β_s would still be possible however with a decreased sensitivity [21, 22].

Procedures providing information about the production flavour are called flavour taggers. Their output values are referred to as tagging decision ξ and dilution D. The former is an integer number, defined as $+1$ for a B_s^0 meson, -1 for $\bar{B}_s^0$ and 0 if a decision could not be reached. The latter is related to the probability P for the tagging decision to be correct:

$$D = 2 \cdot P - 1. \tag{2.79}$$

While the production of single b quarks is possible at the Tevatron by electroweak processes, the production of a $b\bar{b}$ quark pair, which is mediated by quantum chromodynamics, has a much higher cross section [23]. The presence of a $b\bar{b}$ quark pair is therefore assumed in the following. Figure 2.5 shows a schematic view of a typical B_s^0 event.

It starts with the production of a $\bar{b}b$ quark pair at the primary vertex. Both quarks are decoupled from each other. In the events of interest, at least one quark hadronises to a B_s^0 or $\bar{B}_s^0$. The corresponding hemisphere is called same side. The hemisphere of the remaining b quark is called opposite side. According to this definition two approaches are used at the Tevatron: opposite side tagging and same side tagging.

Opposite side tagging is based on the fact that the quark at the opposite side also has opposite charge. The decay products are therefore used for drawing conclusions about the opposite side quark and thus about the signal quark. Leptons and kaons are hereby clean evidences. In addition, the weighted sum of a jet's tracks and, if available, the vertex charge of the opposite side B hadron can be used.

As already indicated above, the signal B_s^0 is obtained by pulling a $s\bar{s}$ quark pair from vacuum. The probability that the remaining $\bar{s}$ quark will end up as a kaon is large. Tagging approaches which are based on this process are referred to as same side tagging.

From the above description it already becomes evident that the event-by-event assignment of tagging decisions and probabilities is a non-trivial task. On the other hand, analyses like the measurement of the $\mathcal{CP}$ violation phase, which use this information, depend on the correct assessment of these quantities. At this point the measurement of the B_s^0's mixing frequency comes into play. A methodological trick is hereby used: a global parameter is introduced as a scale factor for the event-by-event dilution. It is referred to as mixing amplitude. The range of values of this technical parameter is interpreted in the following way: an amplitude consistent with one means that the tagger assesses its decisions and thus its performance correctly. A value smaller than one indicates that a given tagger overestimates itself. According to that, a value greater than one implies an underestimation of the decision power. Altogether, this is the reason why the presented measurement possesses an outstanding role: the performance of a flavour tagger is evaluated on measured data.

To date, the same side kaon tagger (SSKT) [24], which is used inside this analysis, is the most powerful stand-alone tagger at the CDF-II experiment. It has undergone several stages of development. For the following study, the SSKT version is used which combines kinematic information with particle identification. It is essentially the same configuration as was used for the observation of B_s^0 mixing [5]. The main difference is that information about the specific ionisation loss dE/dx is not exploited for tracks with transverse momentum below 2 GeV/c.

3 Experimental Setup

The data used within this analysis was taken with the CDF-II detector located at the Tevatron proton-antiproton collider. The following chapter gives an overview of this experiment.

3.1 Fermilab

The Fermi National Accelerator Laboratory (abbreviation: FNAL or Fermilab) is a research facility near Batavia, Illinois, USA. Committed to particle physics it was founded in 1967. It is a host to both theoretical and experimental research groups and accommodates many experiments.

Two major components of the Standard Model of Particles Physics were discovered at Fermilab: the bottom quark in 1977 and the top quark in 1995 [25].

3.2 Tevatron

The Tevatron is the proton-antiproton collider with the highest centre-of-mass energy worldwide. Built at Fermilab, it was completed in 1983.

The first period of data taking, called Collider Run I, was performed with a centre-of-mass energy of 1.8 TeV. Afterwards, both the Tevatron and its experiments underwent major upgrades. Data taking with a centre-of-mass energy of 1.96 TeV began in 2001. It is referred to as Collider Run II and will be maintained at least until end of 2011.

The preparation of protons and antiprotons and the different stages of acceleration are subject of the following section. A schematic view is shown in figure 3.1. Further information can be found in Refs. [26, 27].

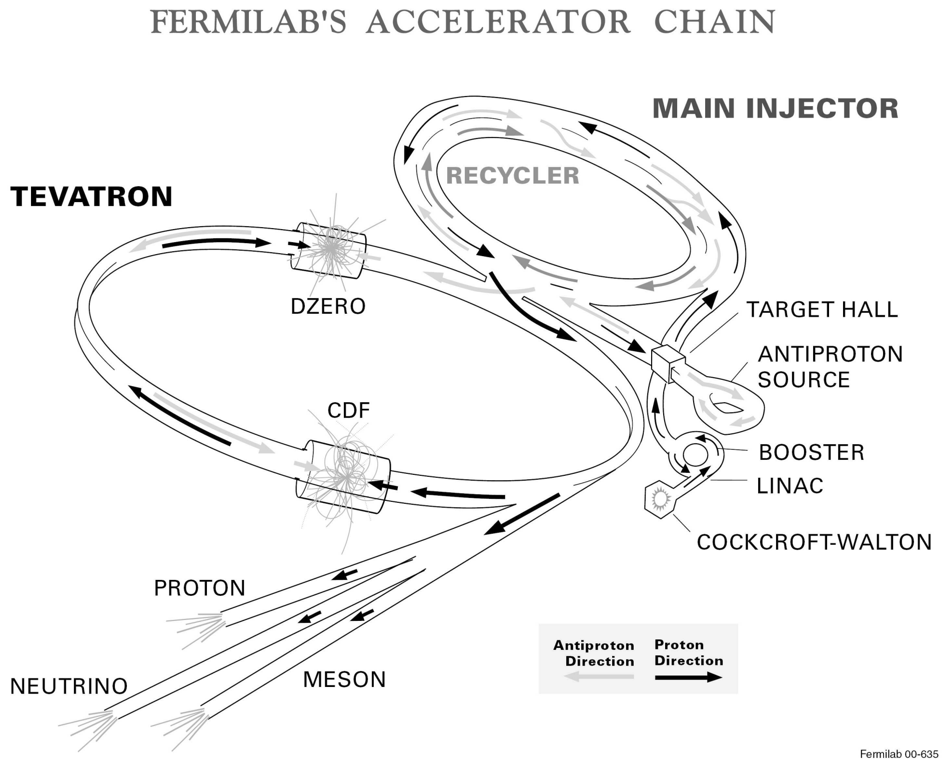


Figure 3.1: Schematic view of the Fermilab acceleration chain.

3.2.1 Cockcroft-Walton Pre-Accelerator

The Cockcroft-Walton Pre-Accelerator provides the first stage of acceleration. Inside this device hydrogen gas is ionised. Subsequently, the negatively charged particles are accelerated in a series of positive voltages to an energy of 750 keV. Afterwards they enter the Linear Accelerator.

3.2.2 Linear Accelerator

The Linear Accelerator (LINAC) is approximately 130 m long. It consists of a series of drift tubes which are arranged in a row. The length of these tubes and the gaps between them increase in beam direction. Oscillating electric fields are used inside the gaps to accelerate the hydrogen ions. The polarisation has hereby the same direction for all fields.

Particles which are located inside the gaps during reversion of polarity are decelerated. Inside the drift tubes the ions are shielded from the electric field. By using this technique bunches are formed consisting of about $1.5 \cdot 10^9$ particles. The time interval between the bunches amounts 5 ns. At the end of the Linear Accelerator the ions have an energy of 400 MeV. They pass through a carbon foil which strips the electrons leaving only protons [28].

3.2.3 Booster

The third stage of acceleration is a synchrotron called Booster which is 475 meters in circumference. Here the protons are bended by magnetic fields to a circle which is passed up to 20,000 times. With each revolution the protons are accelerated by electric fields. Finally they leave the Booster with an energy of 8 GeV and enter the Main Injector [28].

3.2.4 Main Injector

The Main Injector is a synchrotron with a circumference of 3 km. It performs multiple tasks: its first function is the acceleration of the protons coming from the Booster up to 150 GeV. Its second function is the production and acceleration of antiprotons. In order to do so, protons with a mean energy of 120 GeV are taken out of the Main Injector and conducted into the Antiproton Source. Here they hit a nickel target producing a variety of particles (mostly protons, pions and neutrons) including antiprotons. In order to separate these antiprotons from the background particles a pulsed magnet is used. They are afterwards focussed by a lithium lens and stored in the Accumulator. After producing a sufficient number of antiprotons, they are directed into the Recycler. Here stochastic cooling is applied in order to provide a monochromatic beam. Afterwards they are accelerated to 150 GeV. Finally, both protons and antiprotons are injected into the Tevatron.

3.2.5 Tevatron

The Tevatron tunnel has a radius of 1 km and is buried approximately 7.5 m below grade, underneath an earthen berm. It is divided into six sectors of equal size. The boundaries of these sections are referred to as A0, B0, C0, D0, E0 and F0. Beam injection and acceleration is done near A0. The CDF-II experiment is located at B0. As the name already implies, the DØ experiment is located at D0.

After the particles are injected into the Tevatron, their energy is increased from 150 up to 980 GeV. In a top view, the protons travel in clockwise direction while the anti-protons travel in counter-clockwise direction. The beams inside the Tevatron are arranged into 36 proton and 36 anti-proton bunches. They are conducted inside vacuum pipes and more than 1,000 super-conducting dipole magnets keep them on a circular path. In addition, quadrupole and correction magnets are used to focus the beams and to fine-tune their orbits. Collisions at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV are arranged at the centre of each experiment. The time interval between two bunch crossings amounts 396 ns. Apart from interaction regions the beams are kept clearly separated from each other.

The instantaneous luminosity L connects the total cross section σ of a given interaction with the event rate dN/dt:

$$\frac{dN}{dt} = L \cdot \sigma \tag{3.1}$$

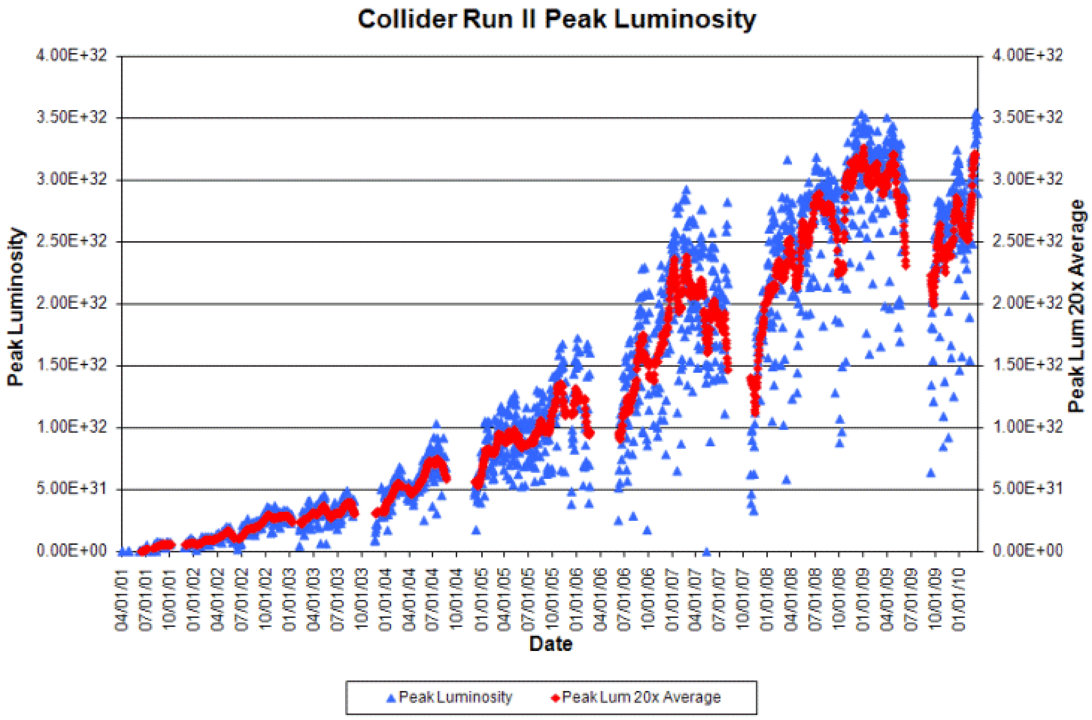


Figure 3.2: Instantaneous luminosity in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ plotted as a function of time.

It is a property of the experiment's mode of operation and can be expressed by technical parameters:

$$L = f \cdot n \cdot \frac{N_1 N_2}{A} \tag{3.2}$$

n is hereby the number of bunches per beam and f is the revolution frequency. The number of particles in each bunch is denoted by N_i and A is defined as the cross section of the beam.

Over the years, the Tevatron operations group continuously improved this figure of merit. The last record value was $376.6 \cdot 10^{30}\ \mathrm{s}^{-1}\mathrm{cm}^{-2}$ on March 18th, 2010. Figure 3.2 shows the evolution of the instantaneous luminosity over the last years.

If the instantaneous luminosity is integrated over time, the integrated luminosity is acquired. Under optimal circumstances, it is proportional to the amount of data available for a detectable decay channel. Figure 3.3 shows the integrated luminosity as a function of time.

3.3 The CDF-II Detector

The CDF-II experiment is a multi-purpose detector with azimuthal symmetry. It has a length of 12 meters and a radius of about 6 meters. It consists of several layers, most of them arranged coaxially around the beam line. An elevation view can be found in figure 3.4. The innermost layers are tracking detectors, followed by

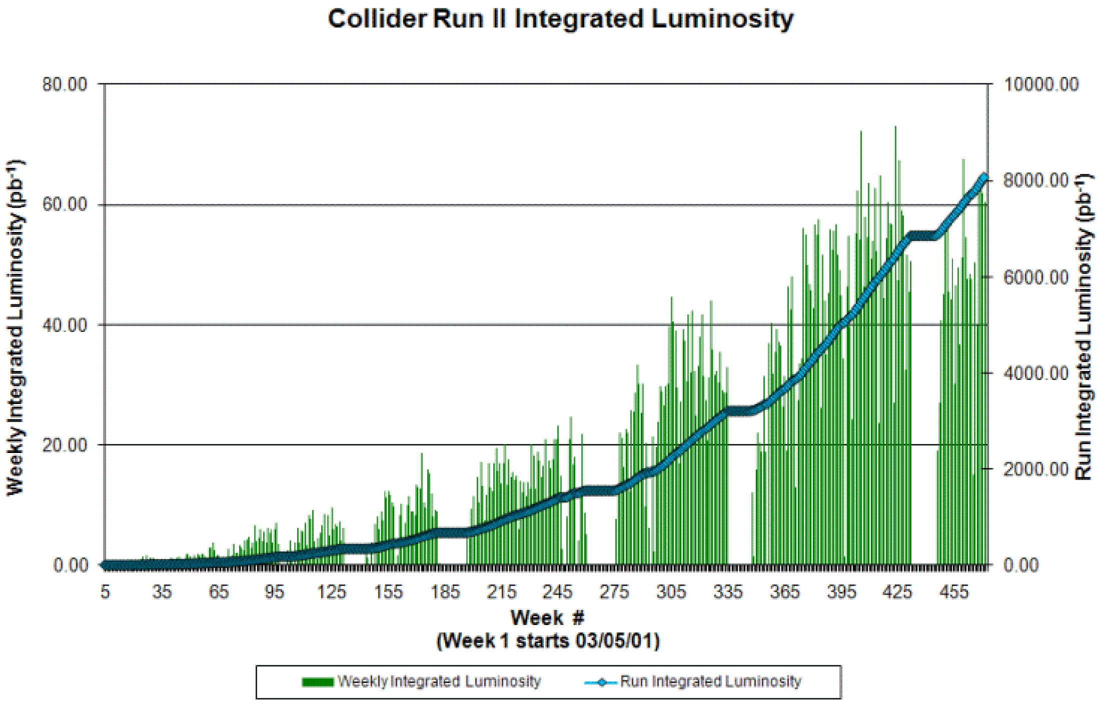


Figure 3.3: Integrated luminosity in pb^{-1} plotted as a function of time.

the time-of-flight detector, the electromagnetic and hadronic calorimeters and finally the muon tracking chambers.

Directions and positions are defined using the following axis: the point of collision, which is located at the centre of the detector, is taken as origin. Starting from that point the z-axis runs parallel to the beamline in direction of the protons. The x-y plane is located perpendicular to that axis: the x-axis points radially outward from the Tevatron and the y-axis points upward to the sky.

The coordinate system which is used in the following is geared to the symmetry of the detector: r is the distance to the z-axis, ϕ is the angle to the x-axis in the x-y-plane and θ the angle to the z-axis in the y-z-plane. Instead of θ, the pseudorapidity η is used more often. It is Lorentz invariant and defined by

$$\eta = -\log(\tan(\theta/2)). \tag{3.3}$$

3.3.1 Tracking System

The tracking system consists of a silicon detector and a drift chamber. Both are contained inside a 1.4 T solenoidal magnetic field. Their geometric proportions are summarised in table 3.1. The purpose of the tracking system is to resolve the tracks of charged particles as precise as possible. Both devices are described in the following, an outline can be found in figure 3.5.

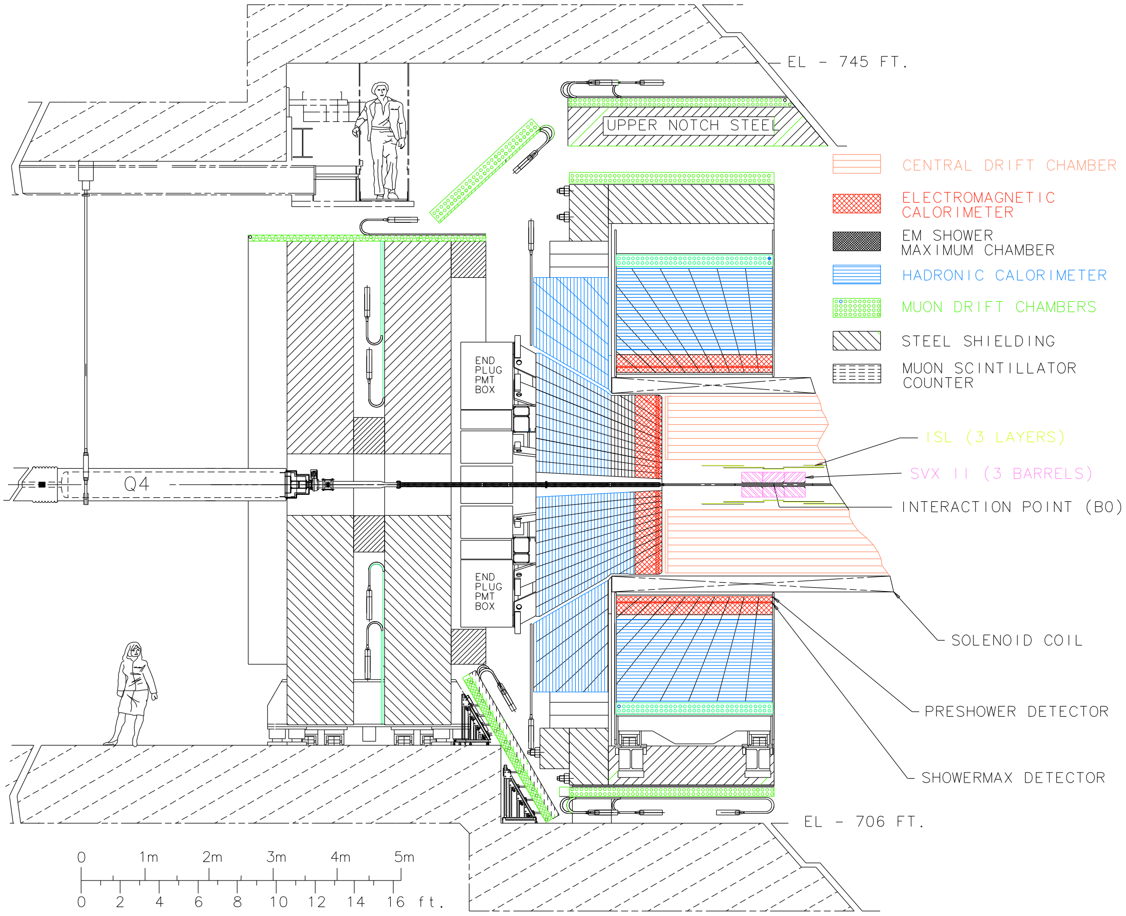


Figure 3.4: Elevation view of the CDF detector.

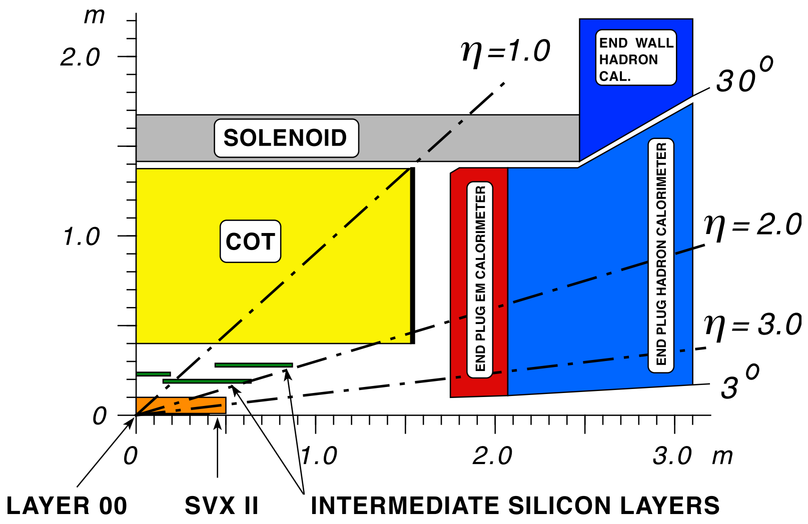


Figure 3.5: The CDF tracking system consists of the silicon detector and the central outer tracker. Both are enclosed by a solenoid providing a magnetic field. Three of the total five calorimeters (see section 3.3.3) are also shown.

	Radial coverage / [cm]	**Pseudorapidity coverage**
Layer 00	1.35 to 1.65	$\|\eta\| \leq 4.0$
SVX II	2.4 to 10.7	$\|\eta\| \leq 2.0$
ISL	20 to 28	$\|\eta\| \leq 1.9$
COT	44 to 132	$\|\eta\| \leq 1.0$

Table 3.1: Geometric properties of the silicon detector and the central outer tracker.

3.3.1.1 Silicon Vertex Detector

The CDF silicon detector [29, 30] is closest to the interaction point and consists of 12 identical wedges in ϕ. It is composed of three separate sub-detectors: the SVX-II [31], the ISL [32] and the L00 [33].

Each SVX-II wedge contains 5 layers of double-sided silicon, oriented parallel to the beam-pipe at radii from 2.5 to 10.6 cm. Three of the layers possess 90° stereo strips on the other side, the remaining two have 1.2° stereo strips. The SVX-II can be read out quickly in parallel and is designed as a high resolution vertex detector. In this respect, it provides input for a displaced vertex trigger, which is described in section 4.1.

The ISL (Intermediate Silicon Layers) is located between the central outer tracker (see subsection 3.3.1.2) and the SVX-II. It consists of one central barrel in $|\eta| < 1$ with one layer and two forward barrels consisting of two layers. Each layer is made out of 1.2° stereo double-sided silicon. The ISL has two roles: it provides an additional tracking layer to link tracks between SVX-II and COT and increases the forward tracking coverage in the region between $1 < |\eta| < 2$.

At a distance of 1.35 cm, the L00 (Layer 00) is directly mounted on the beam pipe. It provides another high resolution tracking point and compensates for a potential failure of the inner layers of the SVX-II due to radiation damage.

3.3.1.2 Central Outer Tracker

The central outer tracker (COT) [34, 35] is a cylindrical open-cell drift chamber spanning from 44 to 132 cm in radii. It is segmented into 8 super-layers alternating stereo and axial wires. The stereo angle between both amounts ±2°. Each super-layer contains 12 sense wires alternated with 13 potential wires which provide the field shaping within the cell, yielding a total of 96 measurement layers. High voltage and readout electronics are mounted on aluminium end-plates, providing timing and charge deposit information. The chamber is filled with argon and ethane in equal shares. In addition, small amounts of maintenance gases are usually added.

The main function of the central outer tracker is to find charged tracks in the region $|\eta| < 1.0$ with transverse momenta as low as 400 MeV/c. Apart from that, the charge deposit information can also be used for particle identification which is described in section 3.3.2.

3.3.1.3 Track Reconstruction

Track reconstruction [36] begins in the Central Outer Tracker. The first step is the formation of line segments from hits in each super-layer. Consistent line segments from the axial layers are linked together to form a track candidate and the hit positions are fitted to a circle. Line segments in stereo layers are then linked to the 2-dimensional track and a helix fit is performed. The transverse momentum resolution of the COT is measured using cosmic ray events to be

$$\frac{\sigma(p_T)}{p_T^2} = 0.0017\ [\mathrm{GeV/c}]^{-1}. \tag{3.4}$$

The next step is to extrapolate each COT track into the SVX-II and add hits that are consistent with the track. A window around the track is established based on the errors on the COT track parameters. If a hit in the outer SVX-II layer lies within the window, it is added to the track. A new track fit is then performed, resulting in a new error matrix and a new window. This window is then used to add hits from the next SVX-II layer and the procedure is repeated over all layers. If no hit is found within the search window, the algorithm proceeds to the next layer. There may be multiple track candidates with different combinations of SVX-II hits associated with one COT track. In this case the track with the largest number of SVX-II layers with hits is chosen.

3.3.2 Particle Identification

The CDF detector does not feature a direct measurement of a stable particle's mass. Instead two measured quantities are capable of at least providing estimations on the identity of stable charged particles: the time-of-flight and the specific ionisation loss. The first is measured by the time-of-flight detector, the latter by the drift chamber. The separation power of both systems is a function of the transverse momentum p_T and can be found in figure 3.6. It is evident that both systems complement each other. Therefore their information is usually combined.

3.3.2.1 Energy Loss through Ionisation

In addition to its tracking capabilities, the central outer tracker features also a measurement of the rate of energy loss through ionisation (dE/dx) [37] in the gas that fills the active volume of the drift chamber. The most probable dE/dx value for a charged particle (other than electron) is a function of its velocity. Thus, if the momentum of the particle is measured, the mass can also be determined.

In the COT, the signal induced on each sense-wire depends on the amount of ionisation charge produced by the passage of the particle near the wire. Multiple samplings along the trajectory of the charged particle allow a reliable estimation of dE/dx which has usually a broad distribution. The COT samples a maximum of 96 dE/dx measurements per track from which a 80 % truncated mean is calculated to avoid the adverse effect of long positive tails in the estimation of the average dE/dx.

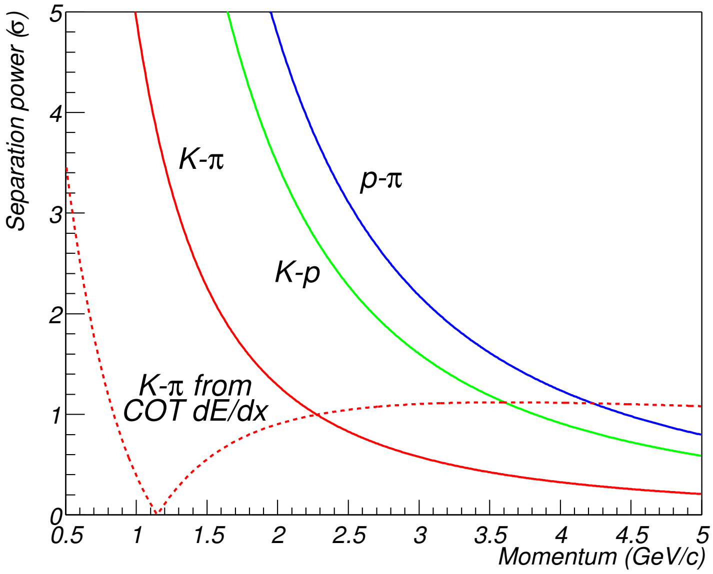


Figure 3.6: The separation power of tof and dE/dx is plotted as a function of transverse momentum p_T. The discrimination between different particle species is shown for tof in solid lines. The separation between kaons and pions using dE/dx is plotted in dashed line.

3.3.2.2 Time-of-Flight

The CDF time-of-flight detector (TOF) [38, 39] is located between the central outer tracker and the cryostat of the super-conducting solenoid at a mean distance of 140 cm from the beam line. This detector consists of 216 bars of Bicron BC-408 scintillator material. Each bar possesses a length of 279 cm and a footprint of 4 x 4 cm. In this way each one covers a pseudo-rapidity interval of roughly $|\eta| < 1$ and approximately $1.7°$ in ϕ. Special photomultipliers are attached to both ends of each bar.

The time-of-flight (tof) of a particle is the difference of the arrival-time (t_a) at the TOF detector and its production time which is called event-t_0 or vertex-t_0:

$$tof = t_a - t_0 \tag{3.5}$$

While the t_a is measured by the detector, the vertex-t_0 has to be computed per vertex out of the t_a measurements of the other tracks with associated tof information coming from the same production vertex. This estimation considers the particle hypotheses of all event tracks.

For its calculation the expected time-of-flight is used:

$$tof_{exp} = \frac{s}{c}\sqrt{1 + \frac{m^2}{p^2}} \tag{3.6}$$

This quantity is calculated using the 3D arc-length from the vertex to the entry-point in the time-of-flight detector (s), the momentum of the particle (p) and a mass hypothesis (m). The first two of those values are determined from the tracking

	Pseudorapidity coverage
CEM	$\|\eta\| \leq 1.1$
CHA	$\|\eta\| \leq 0.9$
WHA	$0.8 \leq \|\eta\| \leq 1.2$
PEM	$1.1 \leq \|\eta\| \leq 3.64$
PHA	$1.2 \leq \|\eta\| \leq 3.64$

Table 3.2: Pseudorapidity coverage of the different calorimeters.

system which intersects with the detector. In formula 3.6, c denotes the speed of light. Usually four particle assumptions, pion, kaon, proton and muon, are used.

By comparing the measured time-of-flight with the expected one for a given particle hypothesis ($tof_{exp}(\pi)$, $tof_{exp}(K)$, $tof_{exp}(p)$ or $tof_{exp}(\mu)$), the probability for each species can be determined.

3.3.3 Calorimeters

The calorimeters of the CDF-II detector [40] were designed for particles with high transverse momentum ($p_T > 20$ GeV/c^2). Therefore they are usually not used inside b-physics analyses. The only exception is the information provided by the electromagnetic calorimeters because it can be used to identify electrons. For reasons of completeness, the following subsection gives an introduction on the CDF calorimeters.

The purpose of the calorimetry system is to measure the energy of all particles except for muons and neutrinos by full absorption. In order to achieve this, the CDF detector possesses in total five calorimeters. All of them are scintillator-based sampling calorimeters employing fiber-based techniques for light extraction.

The central calorimeters are located outside of the solenoid. The region of $|\eta| < 1.1$ is covered by the Central Electromagnetic Calorimeter (CEM) [41] and, on the outside, by the Central Hadron Calorimeter (CHA) [42]. The basic structural units of both are 15° wedges which are subdivided in 10 projective towers each subtending 0.1 units of pseudorapidity. Light from all active scintillator elements of each tower is collected on each side of the wedge by wavelength-shifter and transported to the photomultipliers by means of Plexiglas guides which are routed through the clearances between wedges. Lead is used as absorber material inside the CEM, while iron is used in the CHA.

The CHA is complemented by the Wall Hadron Calorimeter (WHA) [42] which extends its range to $|\eta| < 1.2$. It is composed of iron-scintillator but the sampling fraction is smaller than for the CHA because higher forward energies require more absorber and its resolution is correspondingly worse.

The forward region is covered by the End-Plug Electromagnetic Calorimeter (PEM) [43], a lead-scintillator device, and the End-Plug Hadron Calorimeter (PHA), an iron-scintillator device. These calorimeters span the range $1.1 < |\eta| < 3.64$. Both devices can also be found in figure 3.5.

Table 3.2 summaries the coverage of the calorimeters in pseudo-rapidity.

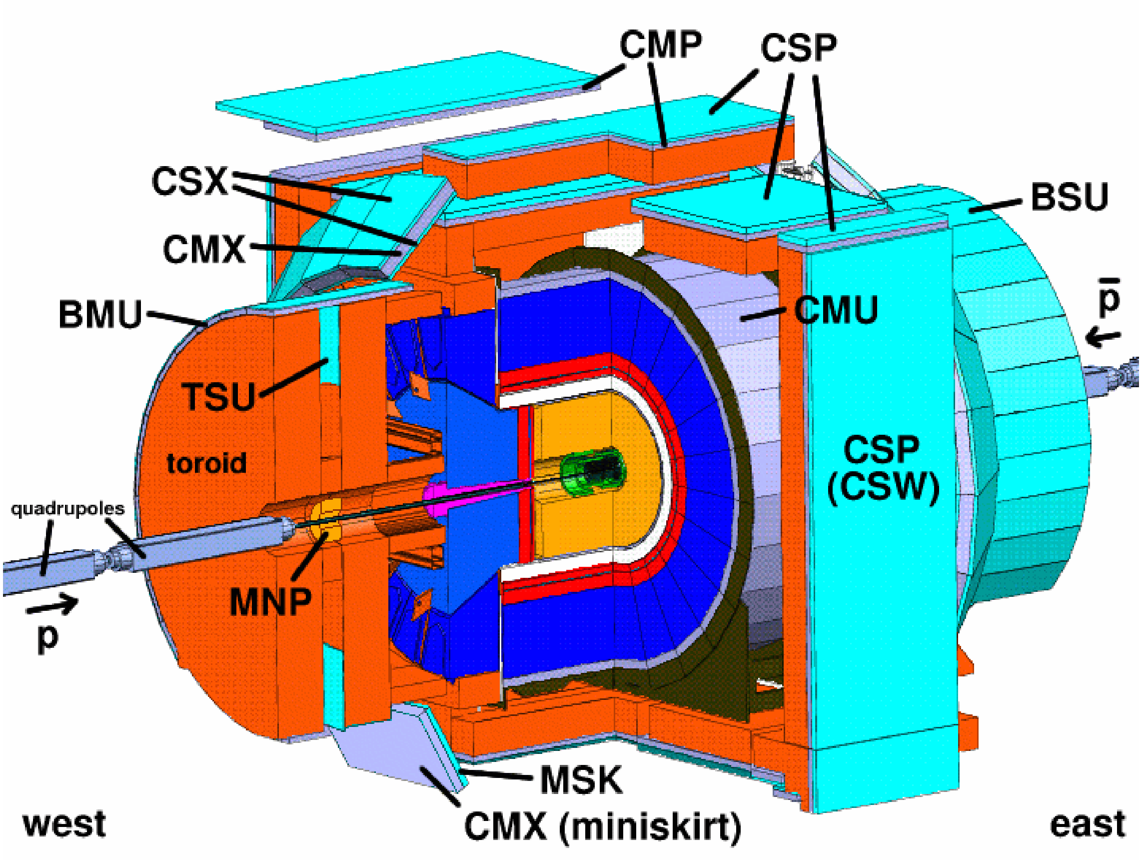


Figure 3.7: Cutaway view of the CDF-II detector with muon chambers labeled.

Both the CEM and the PEM feature devices allowing for position measurements at shower maximum. They are called Central Electromagnetic ShowerMax chamber (CES) and Plug Electromagnetic ShowerMax detector (PES) respectively. Their information can also be used to improve particle identification.

In addition, a Central Pre-Radiate chamber (CPR) and a Plug Pre-Radiate detector (PPR) are used to discriminate between electrons and pions and to improve photon measurements.

3.3.4 Muon Detectors

The muon momenta are measured by the tracking system. However, a key signature is obtained from specialised muon chambers and counters [44, 45] which are deployed behind of the calorimeters so as to exploit the muons' ability to traverse thick absorbers. Figure 3.7 shows a view of the detector with muon chambers sketched in.

The central region is covered by the the Central Muon Detector (CMU) [46], which is mounted behind steel absorbers to reduce beam-splash backgrounds. It is complemented by the Central Muon Extension chambers (CMX) and scintillators (CSX)

	Pseudorapidity coverage
CMU	$\|\eta\| \leq 0.6$
CMP/CSP	$\|\eta\| \leq 0.6$
CMX/CSX	$0.6 \leq \|\eta\| \leq 1.0$
BMU	$1.0 \leq \|\eta\| \leq 1.5$
/BSU-front	$1.0 \leq \|\eta\| \leq 1.25$
/BSU-rear	$1.25 \leq \|\eta\| \leq 1.5$
/TSU	$1.3 \leq \|\eta\| \leq 1.5$

Table 3.3: Properties of the different muon chambers.

and the Central Muon Upgrade chambers (CMP) and scintillators (CSP)[1] [47]. All devices described cover the full azimuthal range.

The forward region is handled by the Intermediate Muon System (IMU). It consists of a barrel-shaped array of muon chambers (BMU) and scintillators (BSU), mounted parallel to the beamline, and one ring-shaped array of scintillators (TSU) mounted perpendicular to the beamline. They complete the continuous muon coverage up to $|\eta| < 1.5$. The BMU consists of single-wire rectangular drift tubes mounted in four half-cell-staggered stacks each covering $1.25°$ of azimuth. One fourth of the azimuthal range is uninstrumented due to interference with the support structure and moving mechanism.

The coverage of the muon chambers is summarised in figure 3.3.

---

[1] The parts of CSP on the north and south walls are sometimes referred to as CSP Wall Scintillators (CSW)

4 Event Selection

The quality of the measured data is the result of different stages of data processing: the data is selected by hard- and software triggers. Afterwards the signal to background ratio is increased by the application of precuts and artificial neural networks. Still, different kinds of signal and non-signal contributions are present inside the data samples. In order to handle them, simulated data, which are referred to as Monte Carlo events, are employed. A detailed description is given in this chapter.

4.1 Trigger System

With a collision rate of 1.7 MHz and an event size of the order of 100 kB, the raw data flow at CDF-II is beyond current storage capabilities. The data taking is therefore governed by a trigger system [48, 49] consisting of three levels. It is described in the following with emphasis on b-physics.

The first level is implemented in hardware and receives data at the bunch crossing rate. The maximum allowed latency amounts 5.5 μs while the maximum allowed output rate is 30 kHz. One of the key components is the Extremely Fast Tracker (XFT) [50] processor which reconstructs tracks inside the drift chamber and. If appropriate, it matches them with lepton information coming from the electromagnetic calorimeters or from muon chambers.

The second trigger level is implemented in hard- and software and has a typical latency of 20 μs and a maximum output rate of 1000 Hz. For b-physics, the key component is the Silicon Vertex Trigger (SVT) processor [51], which exploits the lifetime of b- and c-quark hadrons. The decay vertices of those particles are often located a few millimetres away from the primary interaction vertex. Therefore, the SVT uses both the level-1 XFT tracks and SVX information as input. It applies tracking algorithms and provides 2D reconstructed parameters (p_T, ϕ, d) as output for all tracks. Typical impact parameter resolutions are of the order of 50 μm including the beam size.

The third trigger level is a dedicated computing farm that performs full event reconstruction with a speed-optimised version of the offline code. In most of the cases, the level-3 reconfirms decisions of the previous levels with higher precision. The final output rate to tape is ∼ 100 Hz.

For b-physics, the most important trigger paths are the di-muon trigger, the displaced track plus lepton trigger and the two-track trigger. Each one was designed for a certain group of decay signatures.

The data used for this analysis were taken with the two-track trigger. It requires two displaced tracks with $p_T > 2$ GeV/c respectively and an opening angle between 2° and 135°. The impact parameter d_0 of each track is required to be between 0.1 mm and 1 mm. In addition, the decay length in the transverse plane L_{xy} must be greater than 0.12 mm. Further one of the following scenarios, also called sub-paths, must be fulfilled:

LOW_PT: The scalar sum of the transverse momenta of both tracks must be greater than 4 GeV/c.

CHARM: The scalar sum of the transverse momenta of both tracks must be greater than 5.5 GeV/c. Further both tracks must possess opposite charge.

HIGH_PT: The scalar sum of the transverse momenta of both tracks must be greater than 6.5 GeV/c. Further both tracks must possess opposite charge.

4.2 Event Selection

As already mentioned in chapter 2, the following decay channels are used for the measurement of the mixing frequency Δm_s:

$$\begin{aligned}
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-,\\
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-,\ K^* \to K^+ \pi^-,\\
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+ \pi^- \pi^-,\\
&B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-.
\end{aligned}$$

First of all, events are only considered if they were taken during runs which were marked as good by the data quality group. Further, the following hit requirements must be met by the event tracks: the sum of stereo and axial hits in the silicon detector must be at least three. For the central outer tracker, at least 10 stereo and at least 10 axial hits are required. The lower limit for the transverse momenta of the stable particles is set to 0.35 GeV/c. The event itself must meet one of the level 2 trigger scenarios which are described above. For the reconstructed B_s^0 candidate a generous mass window is chosen from 4.8 GeV/c^2 to 6.0 GeV/c^2.

The tracks of each event are fitted with the mass hypotheses of the respective decay. Afterwards, it is reconstructed from bottom to top. In order to increase the signal to background ratio, neural networks are applied. They were optimised for the purpose of accepting as much signal events as possible while rejecting combinatorial background as much as possible. The figure of merit used for this optimisation is the yield significance, defined by $S/\sqrt{S+B}$. This expression is also a part of the amplitude significance, which is defined as

$$\mathcal{S} = \frac{S}{\sqrt{S+B}} \cdot \sqrt{\frac{\varepsilon D^2}{2}} \cdot e^{-\sigma_{ct}^2 \Delta m_s^2/2}. \tag{4.1}$$

For each decay channel one neural network is applied. Afterwards, events are only considered if their signal probability exceeds a certain value. This cut was determined by consideration of the yield significance for each decay channel separately. Further details on training and optimisation of the applied neural networks can be found in Ref. [52, 53, 54, 55].

The candidate events were measured between February 2002 and June 2009. This corresponds to data taking periods 0 through 25. Corresponding run ranges, online time and datasets can be found in table 4.1 on page 44. The data amount corresponds to an integrated luminosity of approximately 5.2 fb^{-1}.

4.3 Data Composition

After applying the reconstruction and selection algorithms, various kinds of signal and non-signal contributions are contained inside the data sample. The reason for this is obvious - they meet all requirements applied. The following section describes the different classes which the selected events are expected to belong to.

One way to classify these samples is to ask whether there is a real B_s^0 meson in the initial state. This is the case for the following contributions:

- **Signal:** The decay channel which is referred to as signal is $B_s^0 \to D_s^- \pi^+$ or $B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-$
- **Signal with final state radiation:** The charged particles in the decay channel at hand may emit bremsstrahlung. This affects the kinematics of the decay. In the following it will be denoted by the suffix $(n\gamma)$, e.g. $B_s^0 \to D_s^- \pi^+ (n\gamma)$.
- **Cabibbo suppressed signal:** This term refers to decays similar to the signal only with the pion from the B_s meson replaced by a kaon. In the selection and reconstruction applied this phenomenon is not considered explicitly. In comparison to the signal the size of this contribution is expected to be small (see table 4.2 on page 45).
- **Partially reconstructed decays:** This means that at least one particle was not considered in the reconstruction. Possible reasons are that it was not found or that it is electrically neutral. The remaining charged particles of the selected decay match the signal signature. Two partially reconstructed decay channels possess a distinct shape in mass: the first one is $B_s^0 \to D_s^- \rho^+$, $\rho^+ \to \pi^+ \pi^0$, the second one is $B_s^0 \to D_s^{*-} \pi^+$, $D_s^{*-} \to D_s^- \gamma$ and $D_s^{*-} \to D_s^- \pi^0$. Miscellaneous partially reconstructed decays are denoted by $B \to D_s^- X^+$.

The contributions which do not originate from real B_s^0 mesons can be divided into two classes. The first one consists of real physics events, where a wrong particle hypothesis was assigned to one of the decay products. The reflections expected are described in the following:

- **Background coming from B^0-Mesons:** If a kaon hypothesis is assigned to one of the pions from the D^- in the real decay $B^0 \to D^- \pi^+,\ D^- \to K^+ \pi^- \pi^-$, a D_s^- is misreconstructed. The subsequent combination of this particle with a pion leads to the reconstruction of a fake B_s^0 instead of a real B^0. This kind of background is only expected in $B_s^0 \to D_s^- \pi^+, D_s^- \to K^* K^-,\ K^* \to K^+ \pi^-$.

- **Background coming from Λ_b-Baryons:** Wrong mass hypotheses for the stable particles in the decay $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\Lambda_c^+ \to pK^-\pi^+$ may result in the reconstruction of a fake D_s^+ meson instead of a real Λ_c^+ baryon. Again the subsequent combination of this particle with a pion leads to the reconstruction of a fake $\bar{B}_s^0$ instead of a real Λ_b^0. This kind of reflection is expected in $B_s^0 \to D_s^-\pi^+, D_s^- \to K^*K^-,\ K^* \to K^+\pi^-$ and $B_s^0 \to D_s^-\pi^+,\ D_s^- \to \pi^+\pi^-\pi^-$.

The second class of the non-B_s^0 meson contribution is called **combinatorial background**. This means that an appropriate set of at least partially independent tracks match the selection criteria.

Table 4.2 (page 45) summarises the different contributions together with their expected fractions. All values are given with respect to the proper signal. As an example, A_{Cb} is defined as

$$A_{Cb} := \frac{N(B_s^0 \to D_s^- K^+)}{N(B_s^0 \to D_s^-\pi^+)} \tag{4.2}$$

in the single pion modes. Gaussian constraints are later applied for these parameters. Initial values for the physics reflections can be found in Ref. [54]. The fraction of signal events with final state radiation was assumed to be 4.0 % for both decays. This is the default value of the PHOTOS module [56] which was used for the generation of the simulated events for the signal (see also subsection 4.4).

4.4 Simulated Events

As discussed in previous section, several contributions are expected in the different data samples. Except for the combinatorial background, one simulated dataset is created for each such contribution:

For each decay channel a simulated data sample is produced for the signal using PYTHIA [57]. This data sample also contains final state radiation which is simulated by the PHOTOS module [56]. The mean B_s^0 lifetime used for the creation amounts $c\tau = 0.0459$cm.

The Monte Carlo generator BGenerator II [58] is used for generating templates for the physics backgrounds including the Cabibbo reflection. In addition, one simulated dataset containing partially reconstructed decays is created for each B_s^0 decay channel.

The same neural networks which were trained for the selection of B_s^0 mesons on measured data are applied on all simulated samples created. The same cut levels were hereby chosen.

The PYTHIA samples which are used to simulate the signal part of the respective decay were produced using detector parameters taken during periods 0 through 4. Since then, the operations of the Tevatron has changed: the initial luminosity in each store was gradually increased. With this the instantaneous luminosity tended to increase as well. Associated with this change is an alteration of the trigger subpath composition. The origin of this is the finiteness of the trigger bandwidth on level 1 and level 2. In order to use it in an optimal way with consideration of interests of

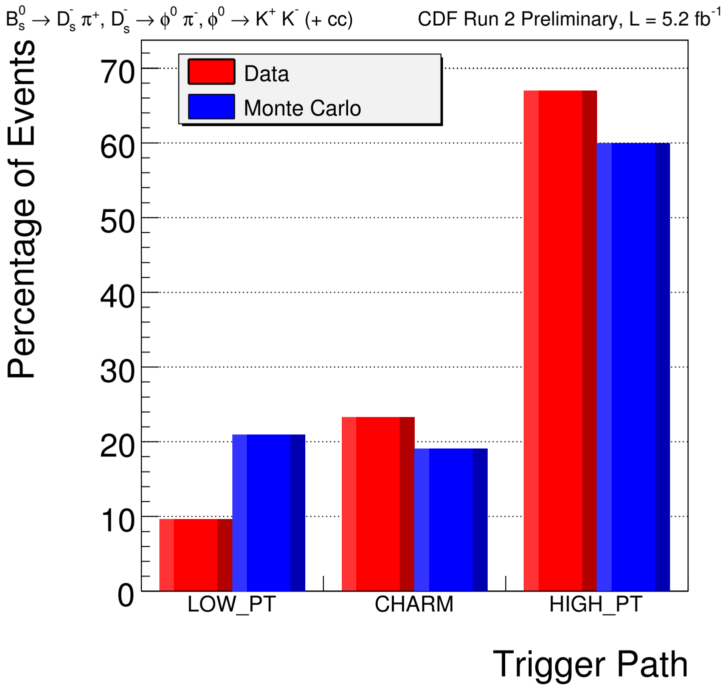


Figure 4.1: The composition of the level 2 trigger sub-paths are compared between simulated (blue) and measured data (red).

the different analysis groups at the same time, a mechanism called trigger prescaling is employed. It artificially suppresses certain trigger paths if others seem to be more adequate at a given time. In addition, interdependencies between triggers are existing. Further details on the B Physics triggers can be found in Ref. [59] and Ref. [60].

For the analysis at hand, this is the reason why the composition of the level 2 trigger sub-paths LOW_PT, CHARM and HIGH_PT has changed if simulation is compared to the entire dataset. Figure 4.1 illustrates this discrepancy. The technique which was used to obtain this histogram on measured data, a sideband subtraction, is described in section 6.

With HIGH_PT being a subset of CHARM and CHARM being a subset of the LOW_PT scenario, the three variables have the following meaning:

- **LOW_PT** means that only the LOW_PT scenario is fulfilled
- **CHARM** means that the LOW_PT and the CHARM requirements are met. The HIGH_PT scenario is not fulfilled
- **HIGH_PT** means that all criteria are fulfilled.

In order to deal with this disagreement, weights are derived. They are obtained by dividing the fraction of measured data events by the fraction of simulated events for each trigger sub-path, respectively. This procedure is performed for all decay channels separately. A comparison between them, based on measured events, can be found in figure 4.2 as complementing information. It is clearly recognisable that they all possess the same tendency.

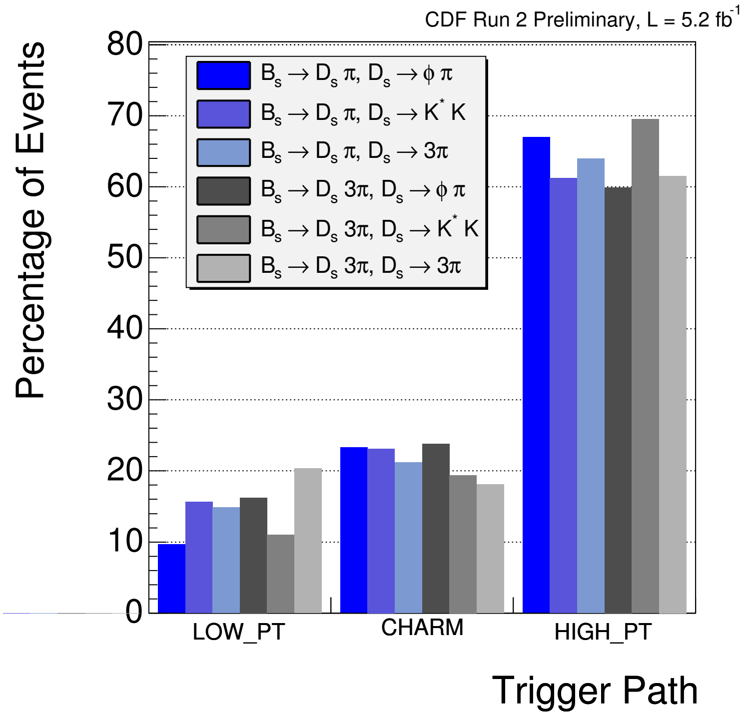


Figure 4.2: Composition of the three different level 2 triggers shown for different B_s^0 decay channels.

The derived weights for the decay channel $B_s^0 \rightarrow D_s^- \pi^+$, $D_s^- \rightarrow \phi^0 \pi^-$ are

$$w_{LOW\_PT} = 0.46, \tag{4.3}$$

$$w_{CHARM} = 1.22, \tag{4.4}$$

$$w_{HIGH\_PT} = 1.12. \tag{4.5}$$

They are used only on simulation. So each time a simulated event is used, it enters with the given number as weight.

In chapter 7 it is shown that, although the difference between simulated and measured data is quite striking, the effect of this reweighting procedure is negligible.

4.5 Variable Definitions

Several properties of the reconstructed B_s^0 meson are used inside the unbinned maximum likelihood. Each of those variables can be seen as a fitting dimension. The following list summarises them along with the reason why they are used:

- The reconstructed **invariant mass** m of the B_s^0, measured in GeV. The reason to use this quantity is that several contributions expected in our sample are clearly separated in mass. Therefore, it is the best discriminating variable used inside the fit.
- The **proper decay time** of the B_s^0 meson multiplied by the speed of light: ct. It is obtained by dividing the decay length of the B_s^0-meson in the x-y

plane L_{xy} (given in cm) by its transverse momentum p_T (given in GeV) and multiplying it by the reconstructed invariant mass m (given in GeV):

$$ct = \frac{L_{xy} m}{p_T}. \tag{4.6}$$

It has therefore the unit cm. Seen in general, mixing is a temporal phenomenon. Therefore, the proper decay time is an essential part of each measurement of the mixing frequency.

- The **resolution of the proper decay time** σ_{ct}, measured in cm. The expression for this variable can be derived by applying Gaussian error propagation on the above formula.

  $$\sigma_{ct} = \sqrt{\left(\frac{m}{p_T}\right)^2 \sigma^2_{L_{xy}} + \left(\frac{L_{xy}\, m}{p_T^2}\right)^2 \sigma^2_{p_T}}. \tag{4.7}$$

  This analysis only considers fully reconstructed decay channels. This implies that the resolution of the transverse momentum is expected to be negligible comparing to the decay length resolution. By doing so, the above equation simplifies to

  $$\sigma_{ct} = \frac{\sigma_{L_{xy}}\, m}{p_T}. \tag{4.8}$$

  By using this variable, events with a more precise decay time measurement are implicitly considered with a greater weight.

- The **decay flavour** of the B_s^0 meson ξ_D, a unitless quantity. This variable is obtained by the decay reconstruction. It is derived using the charge of the pion originating from the B_s^0 candidate. The decay flavour is therefore a binary quantity taking either -1 or +1 as value. It is an inherent property of the reconstruction that this assignment is always correct for a given signal decay.

- The **production flavour** of the B_s^0-candidate ξ_P, a unitless quantity. This variable is obtained by b-flavour tagging. The approaches used at CDF to achieve this are described in section 2.8.

  Most taggers do not possess the capability of providing a decision for each event. The range of values is therefore not only -1 or +1 (B_s^0 or $\bar{B}_s^0$) but also 0 (no decision). The fraction of events a tagger gives an estimation on is called efficiency:

  $$\varepsilon = \frac{N_+ + N_-}{N_+ + N_- + N_0}. \tag{4.9}$$

  N_+ (N_-) hereby stands for the number of events tagged with a positive (negative) production flavour. N_0 denotes events without tagging decision.

  An estimation of mixing frequency and amplitude can only be provided if both production flavour and decay flavour is available for a sufficient number of events.

- The **dilution** D is a unitless quantity which is also given by the flavour tagger algorithm. It is related to the probability for the tagging decision (see previous item) to be correct by the following equation:

  $$D = \frac{N_R - N_W}{N_R + N_W} = 2 \cdot P - 1, \tag{4.10}$$

  N_R (N_W) denotes hereby the number of correct (incorrect) decisions. It is noteworthy that the dilution does not apply if the tagging algorithm did not reach a decision.

  The reason for using this quantity is to assign higher weights for B_s^0 meson candidates which possess more reliable tagging decisions.

  By using the efficiency, the mixing amplitude and the mean value of D^2, the tagging power can be determined:

  $$\mathcal{T} = \varepsilon \mathcal{A}^2 D^2. \tag{4.11}$$

  It can be seen as a benchmark quantity for a given tagger. Typical values for the tagging powers at hadron colliders vary up to ~ 5 %.

4.6 Resolution Scaling

The mixing significance formula, which is presented above, also contains the uncertainty on the proper decay time. The correct estimation of that quantity is therefore of particular importance for the amplitude measurement of the fast oscillating B_s^0 meson. However, a known problem of the decay reconstruction used is a general underestimation of this uncertainty. Figure 4.3 illustrates this fact using simulated events. Here the difference between the true and the measured proper decay time divided by the estimated uncertainty is shown event-by-event. The distribution is fitted by a Gaussian which is assumed as resolution model. Three properties of this fit are discussed in the following: its width, its mean value and its quality.

The width of the fitted Gaussian is the factor by which the reconstructed proper decay time is underestimated. It is roughly 1.38. The usual way to deal with this discrepancy is to apply a scale factor. However, rather than choosing a constant number for scaling a one-dimensional function is employed. The pseudo-rapidity η is used as variable. So a plot of the standard deviation of the proper decay time pull distribution (difference between the measured decay time and the true decay time divided by the uncertainty) is drawn in bins of η. This function is fitted by a second order polynomial:

$$s(\eta) = a \cdot \eta^2 + b \cdot \eta + c. \tag{4.12}$$

The fit results (without unit) for the parameters a,b and c are

$$a = (7.92791 \pm 0.35811) \cdot 10^{-2}, \tag{4.13}$$

$$b = (-2.08479 \pm 2.61885) \cdot 10^{-3}, \tag{4.14}$$

$$c = 1.35375 \pm 0.00201. \tag{4.15}$$

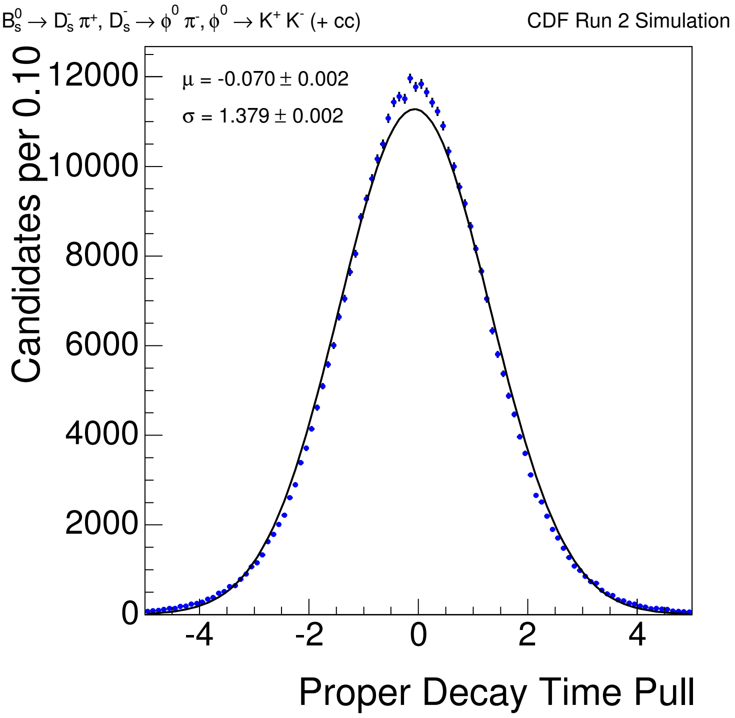


Figure 4.3: The difference between the true and the measured proper decay time divided by the resolution is shown event-by-event for the B_s^0 meson. A Gaussian is fitted to the distribution.

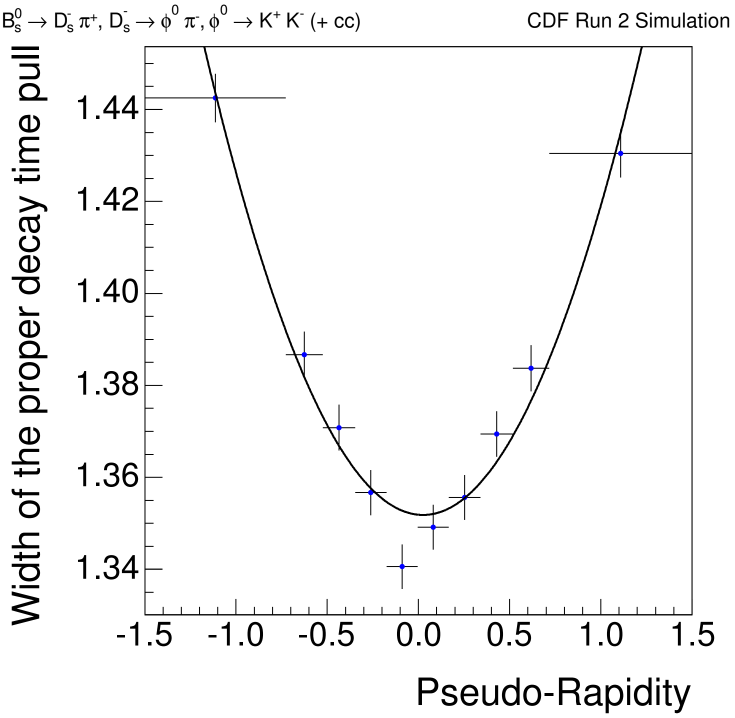


Figure 4.4: The width of the proper decay time pull distribution (difference between the simulated decay time and the measured decay time divided by the decay time resolution) is drawn over the pseudo-rapidity η.

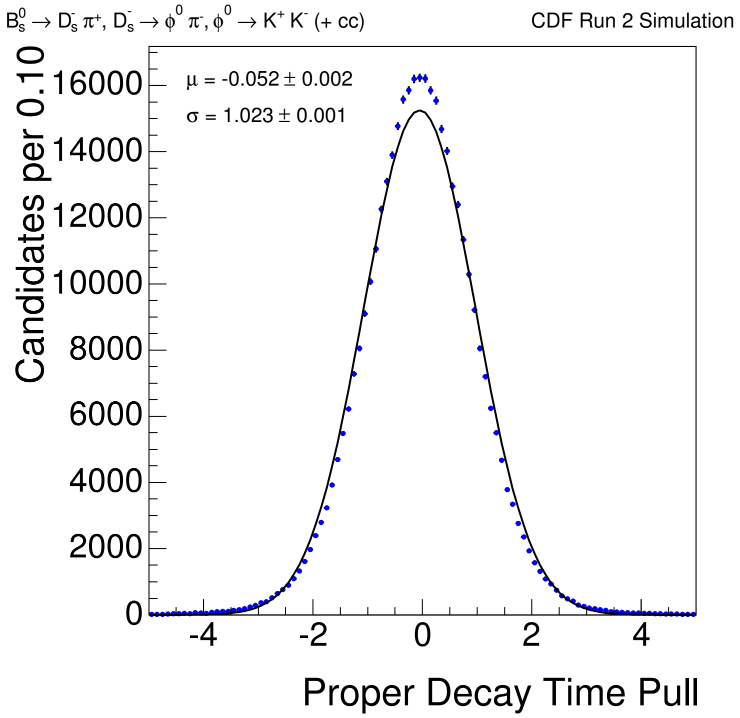


Figure 4.5: The difference between the true and the measured proper decay time divided by the scaled resolution is shown event-by-event for the B_s^0 meson. A Gaussian is fitted to the distribution.

The corresponding fit can be found in figure 4.4.

Each event of each data sample is multiplied by that function. This implies that the same function is used for all B_s^0 decay channels. A pull plot using the scaled proper decay time resolution can be found in figure 4.5. The width of the fitted Gaussian is now close to one. The mean value is slightly smaller in comparison to the previous fit.

In an earlier study [61], a similar fit was done using prompt decays taken from measured data. A comparison by eye shows that the corresponding projections resemble each other closely, which suggests that the simulated events are valid, at least to this point.

Concerning equation 4.12 the question arises whether additional scale functions would be necessary to handle this discrepancy. On the other hand a constant number, as used in many CDF studies, could also be sufficient for the analysis at hand. Therefore the effect of this scaling procedure on the parameters of interest is evaluated in chapter 7.

The mean value of the fitted Gaussian is noteworthy because it is not consistent with zero. By multiplying it by the mean proper decay time resolution of 0.0028 cm, the mean deviation

$$< ct_{measured} - ct_{true} >= 1.96 \cdot 10^{-4} \text{ cm} \tag{4.16}$$

is obtained. This value can be compared to the period length of the B_s^0 mixing frequency (assumption: $\Delta m_s = 17.77\ \mathrm{ps}^{-1}$):

$$cT = c \cdot \frac{2\pi}{\Delta m_s} \approx 1.06 \cdot 10^{-2}\ \mathrm{cm}. \tag{4.17}$$

This means that the expected period length is more than fifty times higher than the observed discrepancy. Based on this study, the observed shift will be neglected in the following.

Finally, the quality of the fits to both pull distributions show that the mixture of two Gaussians would give a better resolution model. This fact is confirmed by earlier studies [62] which were derived on measured data.

However, using two Gaussians would complicate the fit model and eventually double the computing time of the unbinned maximum likelihood fit (chapter 5). Besides it is arguable if the simulated data is reliable to this high degree. The resolution model is therefore taken as single Gaussian and systematic effects originating from this assumption are evaluated later in chapter 7.

Period	**Run-Range**	**Online Time**	**Dataset**
25	275873 - 277511	05 May 09 - 13 Jun 09	0m
24	274123 - 275848	22 Mar 09 - 04 May 09	0m
23	272470 - 274055	15 Feb 09 - 21 Mar 09	0m
22	271072 - 272214	02 Jan 09 - 10 Feb 09	0m
21	268155 - 271047	12 Oct 08 - 01 Jan 09	0m
20	266528 - 267718	24 Aug 08 - 04 Oct 08	0m
19	264101 - 266513	01 Jul 08 - 24 Aug 08	0m
18	261119 - 264071	18 Apr 08 - 01 Jul 08	0m
17	258880 - 261005	28 Feb 08 - 16 Apr 08	0k
16	256840 - 258787	27 Jan 08 - 27 Feb 08	0k
15	254800 - 256824	05 Dec 07 - 27 Jan 08	0k
14	252836 - 254683	28 Oct 07 - 03 Dec 07	0k
13	241665 - 246231	13 May 07 - 04 Aug 07	0j
12	237845 - 241664	01 Apr 07 - 13 May 07	0j
11	233133 - 237795	31 Jan 07 - 30 Mar 07	0j
10	228664 - 233111	24 Nov 06 - 31 Jan 07	0j
9	222529 - 228596	01 Sep 06 - 22 Nov 06	0i
8	217990 - 222426	09 Jun 06 - 01 Sep 06	0i
7	210012 - 212133	14 Jan 06 - 22 Feb 06	0i
6	206990 - 210011	10 Nov 05 - 14 Jan 06	0i
5	203819 - 206989	05 Sep 05 - 09 Nov 05	0i
4	201350 - 203799	20 Jul 05 - 04 Sep 05	0h
3	198380 - 201349	21 May 05 - 19 Jul 05	0h
2	195409 - 198379	19 Mar 05 - 20 May 05	0h
1	190697 - 195408	07 Dec 04 - 18 Mar 05	0h
0	138425 - 186598	04 Feb 02 - 22 Aug 04	0d

Table 4.1: Period, run ranges, online time and dataset for the events used in this analysis [3].

Identifier	**Type of Contribution**	**Decay**	**Fraction**
$\boldsymbol{B_s^0 \rightarrow D_s^-\pi^+,\ D_s^- \rightarrow \phi^0\pi^-}$			
	Signal	$B_s \rightarrow D_s\pi$	100 %
	Signal with final state radiation	$B_s \rightarrow D_s\pi(n\gamma)$	4.0 %
A_{Cb}	Cabibbo Suppressed Signal	$B_s \rightarrow D_sK$	5.5 %
$\boldsymbol{B_s^0 \rightarrow D_s^-\pi^+,\ D_s^- \rightarrow K^*K^-}$			
	Signal	$B_s \rightarrow D_s\pi$	100 %
	Signal with final state radiation	$B_s \rightarrow D_s\pi(n\gamma)$	4.0 %
A_{Cb}	Cabibbo Suppressed Signal	$B_s \rightarrow D_sK$	5.5 %
A_{B0}	B^0 contribution	$B^0 \rightarrow D\pi$	6.7 %
A_{Lm}	Λ_b contribution	$\Lambda_b \rightarrow \Lambda_c\pi$	9.2 %
$\boldsymbol{B_s^0 \rightarrow D_s^-\pi^+,\ D_s^- \rightarrow \pi^+\pi^-\pi^-}$			
	Signal	$B_s \rightarrow D_s\pi$	100 %
	Signal with final state radiation	$B_s \rightarrow D_s\pi(n\gamma)$	4.0 %
A_{Cb}	Cabibbo Suppressed Signal	$B_s \rightarrow D_sK$	5.5 %
A_{Lm}	Λ_b contribution	$\Lambda_b \rightarrow \Lambda_c\pi$	6.2 %
$\boldsymbol{B_s^0 \rightarrow D_s^-\pi^-\pi^+\pi^+,\ D_s^- \rightarrow \phi^0\pi^-}$			
	Signal	$B_s \rightarrow D_s3\pi$	100 %
	Signal with final state radiation	$B_s \rightarrow D_s3\pi(n\gamma)$	4.0 %
A_{Cb}	Cabibbo Suppressed Signal	$B_s \rightarrow D_sK\pi\pi$	5.5 %

Table 4.2: Expected backgrounds in the different data samples and their expected fraction with respect to the signal.

5 Maximum Likelihood Method

For this analysis, the most important technique is the maximum likelihood method. It is used to determine the quantities of interest. A description can be found in the this chapter.

5.1 Definition

A set of measured values x $(x_1, x_2, \ldots, x_n)$, where x_i is either a scalar or a vector, is given. Further a probability density function $f(x|a)$ is assumed to describe the measured values for a certain but unknown set of parameters a. It is required that $f(x|a)$ is normalised to one with respect to x for all possible values of a:

$$\int_\Omega f(x|a)dx = 1. \tag{5.1}$$

Based on this initial situation, the maximum likelihood method [63] provides the best estimation of the parameters. In order to do so the likelihood function

$$\mathcal{L}(a) = \prod_{i=1}^{n} f(x_i|a) \tag{5.2}$$

is built. It is a function of a and can be seen as the probability to observe the measured values for a given parameter set. On the other hand, it must be emphasised that $\mathcal{L}(a)$ is not a probability density function.

According to the maximum likelihood principle, the best estimation for a is the value for which $\mathcal{L}(a)$ becomes maximal. In other words, the probability for observing the measured values is optimised.

Technically, the negative log-likelihood function is used. It is obtained by applying, at first, the logarithm on equation 5.2 and afterwards multiplying the result by a negative value:

$$-2 \cdot \log(\mathcal{L}(a)) = -2 \cdot \sum_{i=1}^{n} \log(f(x_i|a)). \tag{5.3}$$

The logarithm is a monotonous function. Therefore, by using this transformation the position of the maximum is preserved. At the same time, the multiplication becomes a sum and a more convenient range of values is obtained for the overall result. The reason for multiplying the result by a negative number is based on the fact that algorithms, which are employed in order to find extrema are by convention minimisers. By using 2 as a prefactor the expression resembles the method of least squares.

In addition to the best estimation, the maximum likelihood method is also capable of providing the standard deviation for each parameter. This procedure is presented in the following. In order to simplify matters, the derivation is constraint to one free parameter only. The extension to multiple parameters works in an analogical way.

For a high enough number of measurements the likelihood function $\mathcal{L}(a)$ possesses a Gaussian shape. In this case, the log-likelihood function $F(a) := -2 \cdot \log(\mathcal{L}(a))$ can be expanded around the minimal value $\hat{a}$:

$$F(a) = F(\hat{a}) + \frac{1}{2} \cdot \frac{d^2F}{da^2}(a - \hat{a})^2 + \dots \tag{5.4}$$

The approximation does not include the first order because it is by definition zero for a minimum. Stopping at the second order, the likelihood function becomes

$$\mathcal{L}(a) = \text{const} \cdot \exp\left(- \frac{d^2F}{da^2}(a - \hat{a})^2 \right). \tag{5.5}$$

A comparison between this expression and a Gaussian function,

$$\text{const} \cdot \exp\left(- \frac{d^2F}{da^2}(a - \hat{a})^2 \right) \overset{!}{=} \text{const} \cdot \exp\left(- \frac{(a - \hat{a})^2}{2\sigma_{\hat{a}}^2} \right), \tag{5.6}$$

yields that the standard deviation of the parameter a can be estimated using the second derivative of the log-likelihood function:

$$\sigma_{\hat{a}} = \left(2 \cdot \left. \frac{d^2F}{da^2} \right|_{\hat{a}} \right)^{-1/2}. \tag{5.7}$$

Equation 5.4 shows that the log-likelihood function is expected to have a parabolic shape around the minimum $\hat{a}$. If this is not the case, the ideal way to deal with the situation is to find a non-linear transformation function $z = z(a)$, where $F(z)$ becomes parabolic. Afterwards, $\hat{z}$ and $\sigma_{\hat{z}}$ can be used in order to determine $\hat{a}$ and $\sigma_{\hat{a}}$. Another possibility is the determination of right and left standard deviations, σ_r and σ_l. This means that the algorithm actually follows the negative log-likelihood function out of the minimum. After a certain increment is reached, e.g. $\Delta F = 1$ for 1σ, the current distance to the minimum is taken as standard deviation. This is done on both sides resulting in two different values σ_l and σ_r.

The advantage of the maximum likelihood method is that it is consistent. This means that the best estimation of the parameter set $\hat{a}$ converges to the true value a_{true} as the number of observations increase:

$$\lim_{n \to \infty} \hat{a} = a_{true}. \tag{5.8}$$

The disadvantage of this method is that it is not necessarily unbiased if the estimation is performed on a finite number of observations. This means that the log-likelihood function around the minimum may deviate from a parabolic shape.

Throughout this analysis the maximum likelihood method is used in order to determine the parameters of interest. The function definition follows equation 5.3. Default minimisations are performed using Minuit [64]. If necessary, scans for the right and left standard deviations are done with Minos.

5.2 Combination

In order to extract common parameters from different data samples, the maximum likelihood method is extended. This is done by minimising the sum of the respective negative log-likelihood functions,

$$-2 \cdot \sum_{j} \ln \left(\mathcal{L}_j(\vec{a_j}) \right). \tag{5.9}$$

The index j denotes hereby the different data samples. The fit models are also allowed to differ from each other.

In the analysis at hand, the different decay channels are treated in this way. Common parameters are the mean B_s^0 lifetime, the mixing amplitude and the mixing frequency.

5.3 Fit Model

It was already mentioned in the previous sections that the maximum likelihood method relies on a model function which is assumed to describe the measured values for a certain choice of parameters. Some of those parameters are the quantities of interest (e.g. scale factor, mixing frequency or mean lifetime), the rest are nuisance parameters. The following section focusses therefore on the derivation of a data fit model. The consideration of correlations between the variables introduced in section 4.5 is hereby a crucial point.

An introduction is given in table 5.1, where the correlations between the fit variables are shown for signal and background. The ones for the signal were derived from simulated data. The coefficients for the combinatorial background were derived from measured data using the upper side band starting from 5.55 GeV/c^2 in invariant mass, where only this contribution is expected.

5.3.1 Classes

The unbinned maximum likelihood fit which is used to determine the parameters of interest is performed inside a narrow mass window ranging from 5.31 GeV/c^2 to 5.60 GeV/c^2. Inside this window the contributions of non-proper signal decays, like

Signal

	$\boldsymbol{m}$	$\boldsymbol{ct}$	$\boldsymbol{\sigma_{ct}}$	$\boldsymbol{D}$	$\boldsymbol{\xi_{Prd}}$	$\boldsymbol{\xi_{Dcy}}$
$\boldsymbol{m}$	100.0 %	0.4 %	0.8 %	0.4 %	0.3 %	-0.1 %
$\boldsymbol{ct}$	0.4 %	100.0 %	20.5 %	0.6 %	-0.1 %	0.2 %
$\boldsymbol{\sigma_{ct}}$	0.8 %	20.5 %	100.0 %	3.9 %	-0.0 %	0.0 %
$\boldsymbol{D}$	0.4 %	0.6 %	3.9 %	100.0 %	0.7 %	-0.1 %
$\boldsymbol{\xi_{Prd}}$	0.3 %	-0.1 %	-0.0 %	0.7 %	100.0 %	-0.5 %
$\boldsymbol{\xi_{Dcy}}$	-0.1 %	0.2 %	0.0 %	-0.1 %	-0.5 %	100.0 %

Combinatorial Background

	$\boldsymbol{m}$	$\boldsymbol{ct}$	$\boldsymbol{\sigma_{ct}}$	$\boldsymbol{D}$	$\boldsymbol{\xi_{Prd}}$	$\boldsymbol{\xi_{Dcy}}$
$\boldsymbol{m}$	100.0 %	3.5 %	6.6 %	-4.4 %	-1.0 %	-4.5 %
$\boldsymbol{ct}$	3.5 %	100.0 %	57.4 %	-0.5 %	-0.1 %	0.0 %
$\boldsymbol{\sigma_{ct}}$	6.6 %	57.4 %	100.0 %	3.3 %	3.9 %	-0.2 %
$\boldsymbol{D}$	-4.4 %	-0.5 %	3.3 %	100.0 %	0.2 %	-4.5 %
$\boldsymbol{\xi_{Prd}}$	-1.0 %	-0.1 %	3.9 %	0.2 %	100.0 %	-0.5 %
$\boldsymbol{\xi_{Dcy}}$	-4.5 %	0.0 %	-0.2 %	-4.5 %	-0.5 %	100.0 %

Table 5.1: Correlation matrix for simulated signal events and events taken from the combinatorial background.

final state radiation, Cabibbo suppressed signal and partially reconstructed decays, are expected to be small. In mass, the corresponding functions are therefore merged into a common function. For the other dimensions the same functions are used for all contributions. In the following, this ensemble will be referred to as signal (subscript *Sig*).

The other event classes which are considered in the probability density function are

- Combinatorial background (subscript *Comb*)
- B^0-Meson reflection (subscript B)
- Λ_b reflection (subscript Λ_b)

The overall probability density function can be written as

$$\begin{aligned} P(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) = {} & f_{Sig} \cdot P_{Sig}(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_{Comb} \cdot P_{Comb}(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_{\Lambda_b} \cdot P_{\Lambda_b}(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_B \cdot P_B(m, ct, \sigma_{ct}, \xi_D, \xi_P, D). \end{aligned} \tag{5.10}$$

The functions P_j inside this expression refer to probability density functions describing the corresponding class j. Internal parameters of the probability density functions are omitted. f_{Sig}, f_{Comb}, f_{Λ_b} and f_B are fit parameters. They can be seen as fractions describing the relative size of class j with respect to the total data sample at hand. The requirement of a normalisation equal to 1 leads to the following condition:

$$f_{Sig} + f_{Comb} + f_{\Lambda_b} + f_B = 1. \tag{5.11}$$

Subsection 6.2 describes how this constraint is realised technically.

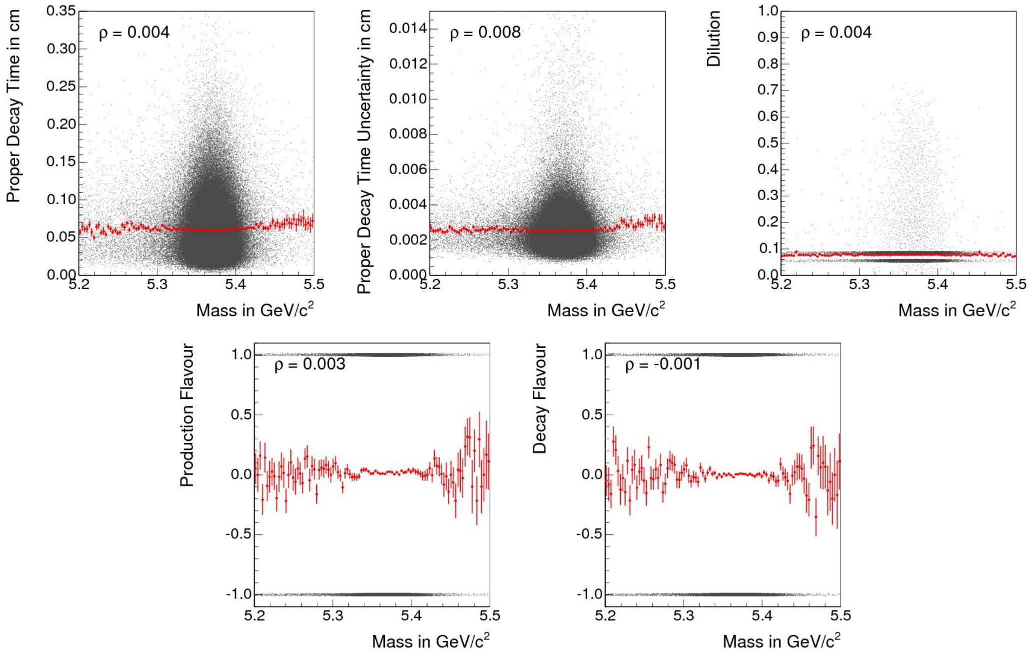


Figure 5.1: The correlation between mass and other variables is illustrated by combined scatter and profile plots. A data sample consisting of simulated B_s^0 mesons was used for their creation.

Each probability density function P_j can be factorised in a mass-dependent and a mass-independent term:

$$P_j(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) = P_j(m|ct, \sigma_{ct}, \xi_D, \xi_P, D) \cdot P_j(ct, \sigma_{ct}, \xi_D, \xi_P, D) \tag{5.12}$$

In order to give an impression of the small correlation between mass and any other variable used, combined scatter and profile plots can be seen in figure 5.1 and 5.2 for signal and combinatorial background, which are the two dominating classes.

The highest correlation coefficient between mass and any other variable is 6.6 %. A similar behaviour is expected from the physics reflections. Therefore, mass is assumed to be independent from the other variables. With this, the general equation 5.12 simplifies to

$$P_j(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) = P_j(m) \cdot P_j(ct, \sigma_{ct}, \xi_D, \xi_P, D) \tag{5.13}$$

and equation 5.10 becomes

$$\begin{aligned} P(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) = {} & f_{Sig} \cdot P_{Sig}(m) \cdot P_{Sig}(ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_{Comb} \cdot P_{Comb}(m) \cdot P_{Comb}(ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_{\Lambda_b} \cdot P_{\Lambda_b}(m) \cdot P_{\Lambda_b}(ct, \sigma_{ct}, \xi_D, \xi_P, D) \\ & + f_B \cdot P_B(m) \cdot P_B(ct, \sigma_{ct}, \xi_D, \xi_P, D). \end{aligned} \tag{5.14}$$

The mass terms are fitted by phenomenological functions. This is further described in chapter 6. The following subsections focus on the individual $P_j(ct, \sigma_{ct}, \xi_D, \xi_P, D)$ terms.

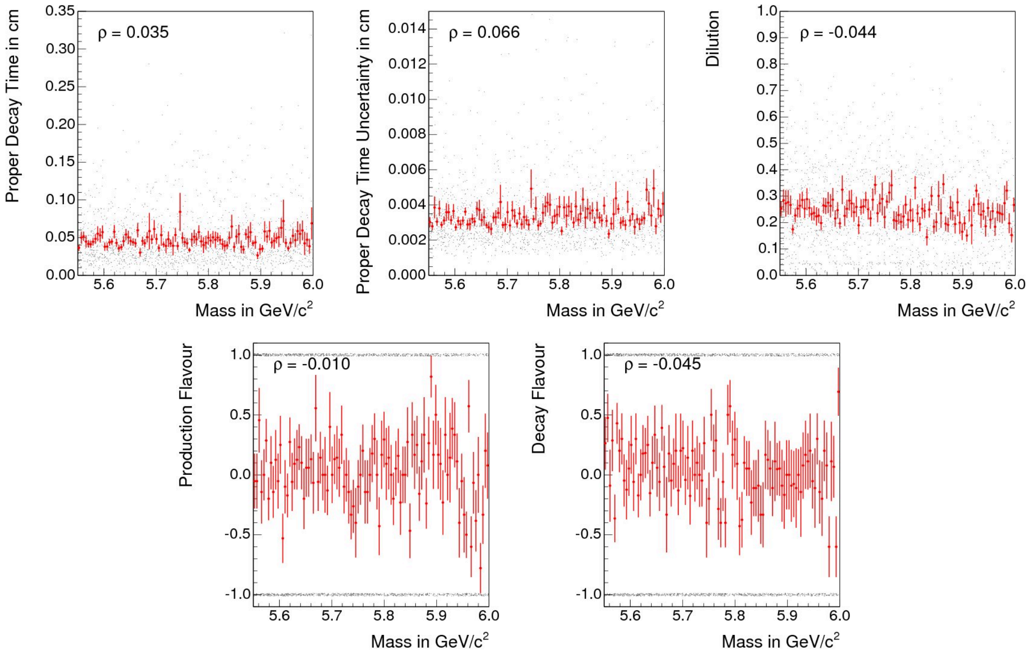


Figure 5.2: The correlation between mass and other variables is illustrated by combined scatter and profile plots. A data sample consisting of B_s^0 candidate events taken from the combinatorial background was used for their creation.

5.3.2 Signal

The following section focusses on $P_{Sig}(ct, \sigma_{ct}, \xi_D, \xi_P, D)$ (first line in equation 5.14), the probability density function for the signal in each variable except for mass. Based on the fact that the decay flavour of each candidate can flawlessly be determined, this expression can be separated into two terms:

$$P_{Sig}(ct, \sigma_{ct}, \xi_D, \xi_P, D) = \begin{cases} P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to B_s^0) & \text{if } \xi_D = +1 \\ P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to \bar{B}_s^0) & \text{if } \xi_D = -1 \end{cases} . \tag{5.15}$$

In this equation $P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to B_s^0)$ (the first case) denotes the probability density function for a particle which decayed as B_s^0. Thus, it has either been B_s^0 or $\bar{B}_s^0$ at production time. An analog argumentation can be made for the second case. Therefore the following ansatz can be made for both:

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to B_s^0) = {} & f_{B_s^0 \to B_s^0} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) \\ & + f_{\bar{B}_s^0 \to B_s^0} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to B_s^0), \end{aligned} \tag{5.16}$$

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to \bar{B}_s^0) = {} & f_{B_s^0 \to \bar{B}_s^0} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to \bar{B}_s^0) \\ & + f_{\bar{B}_s^0 \to \bar{B}_s^0} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0). \end{aligned} \tag{5.17}$$

The four fractions introduced inside both equations are further exemplified in the following: $f_{B_s^0 \to B_s^0}$, as an example, is the fraction of detected mesons which were produced as B_s^0 and also decayed as B_s^0. Conservation of probability requires both fractions within each equation to have a sum of 1:

$$f_{B_s^0 \to B_s^0} + f_{\bar{B}_s^0 \to B_s^0} = 1, \tag{5.18}$$

$$f_{B_s^0 \to \bar{B}_s^0} + f_{\bar{B}_s^0 \to \bar{B}_s^0} = 1. \tag{5.19}$$

Alternatively, those fractions can be interpreted as Bayesian probabilities. In this picture the previous two equations become

$$P(B_s^0 \text{ produced}|B_s^0 \text{ detected}) + P(\bar{B}_s^0 \text{ produced}|B_s^0 \text{ detected}) = 1, \tag{5.20}$$
$$P(B_s^0 \text{ produced}|\bar{B}_s^0 \text{ detected}) + P(\bar{B}_s^0 \text{ produced}|\bar{B}_s^0 \text{ detected}) = 1. \tag{5.21}$$

In this way the number of parameters are reduced from four to two. Further it is assumed that the fraction of mesons having the same (a different) flavour at production and decay time are the same for detected B_s^0 and $\bar{B}_s^0$.

$$f_{B_s^0 \to B_s^0} = f_{\bar{B}_s^0 \to \bar{B}_s^0} \tag{5.22}$$
$$f_{\bar{B}_s^0 \to B_s^0} = f_{B_s^0 \to \bar{B}_s^0} =: f_{mix} \tag{5.23}$$

Again, those equations can be interpreted as Bayesian probabilities:

$$P(B_s^0 \text{ produced}|B_s^0 \text{ detected}) = P(\bar{B}_s^0 \text{ produced}|\bar{B}_s^0 \text{ detected}), \tag{5.24}$$
$$\begin{aligned} P(\bar{B}_s^0 \text{ produced}|B_s^0 \text{ detected}) &= P(B_s^0 \text{ produced}|\bar{B}_s^0 \text{ detected}) \\ &=: P(\text{anti-particle produced}|\text{particle detected}) \end{aligned} \tag{5.25}$$

This implicitly assumes the same production rate and the same detection efficiency for particle and anti-particle. Further any kind of $\mathcal{CP}$ violation is not considered.

Now only one parameter remains and equations 5.16 and 5.17 become

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to B_s^0) = (1 - f_{mix}) \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) \\ + f_{mix} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to B_s^0), \end{aligned} \tag{5.26}$$
$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \to \bar{B}_s^0) = f_{mix} \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to \bar{B}_s^0) \\ + (1 - f_{mix}) \cdot P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0). \end{aligned} \tag{5.27}$$

Here f_{mix} is the fraction of detected mesons which were produced with a production flavour different from their decay flavour. From this definition it is clear that this quantity cannot be a fit parameter. Instead it is an inherent property of the analysed particle itself: for example, the periodic time in comparison to its mean lifetime is far smaller for the B_s^0 than for the B^0(see table 2.2 on page 11). Therefore, it is more likely for the B_s^0 to alter its flavour than for the B^0.

In order to define f_{mix} the probability density function for a given initial and a given final state (here $B_s^0 \to B_s^0$) must be further analysed. It can be factorised in the following way:

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) = P_{Sig}(ct|\sigma_{ct}, \xi_P, D, B_s^0 \to B_s^0) \\ \cdot P_{Sig}(\sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0), \end{aligned} \tag{5.28}$$

where $P_{Sig}(ct|\sigma_{ct}, \xi_P, D, B_s^0 \to B_s^0)$ is the probability density function in proper decay time. A basic expression for it was already derived in equations 2.61 and 2.62 (page 15): in true proper decay time $\tilde{ct}$, it consists of an exponential function (decay) with cosine modulation (mixing). The detector resolution is represented by the convolution of this function with a Gaussian $\mathcal{G}$. The result has the measured decay time ct as variable. In order to account for the limited acceptance of the detector

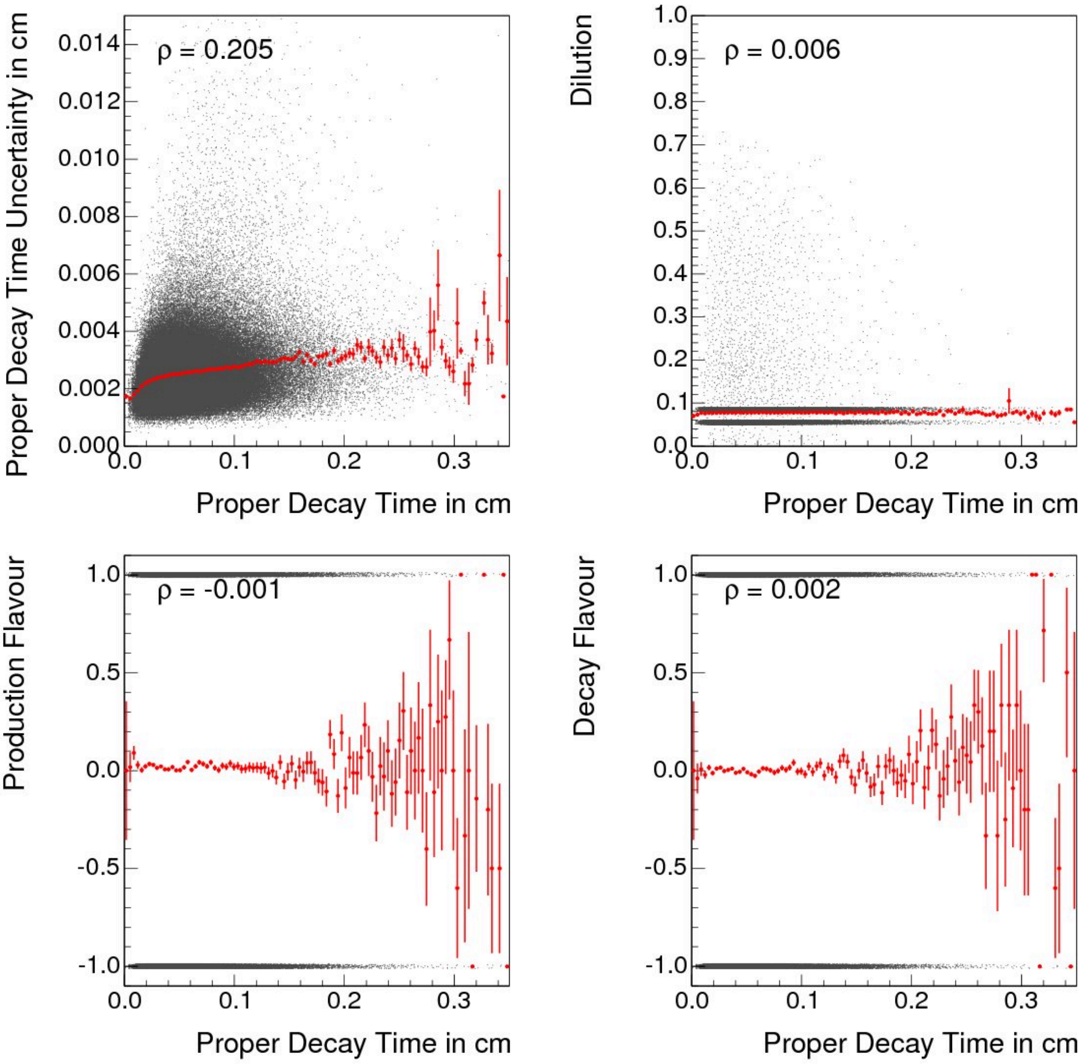


Figure 5.3: The correlation between proper decay time and other variables is illustrated by combined scatter and profile plots. A data sample consisting of simulated B_s^0 mesons was used for their creation.

or, more precise, of the Two Track Trigger, it must be multiplied by an efficiency function ϵ:

$$P_{Sig}(ct|\sigma_{ct},\xi_P,D,B_s^0\to B_s^0) = \frac{1}{N_{B_s^0\to B_s^0}}\cdot\left\{\left[\frac{1}{\tau}e^{-\tilde{t}/\tau}\cdot\frac{1}{2}(1+\cos(\Delta m\tilde{t}))\right]\otimes\mathcal{G}(c\tilde{t}-ct|\sigma_{ct})\right\} \cdot\epsilon(ct|\sigma_{ct},\xi_P,D,B_s^0\to B_s^0). \tag{5.29}$$

The efficiency function is derived from simulated data using the measured decay time and its resolution.

Figure 5.3 illustrates the correlation between the proper decay time ct and the remaining variables. Similar plots can be found in figure 5.4 for the proper decay time resolution σ_{ct}.

A strong correlation between proper decay time and proper decay time uncertainty is clearly recognisable. The remaining correlations are negligible. The efficiency function contains therefore only σ_{ct} as parameter. In order to handle this correlation the

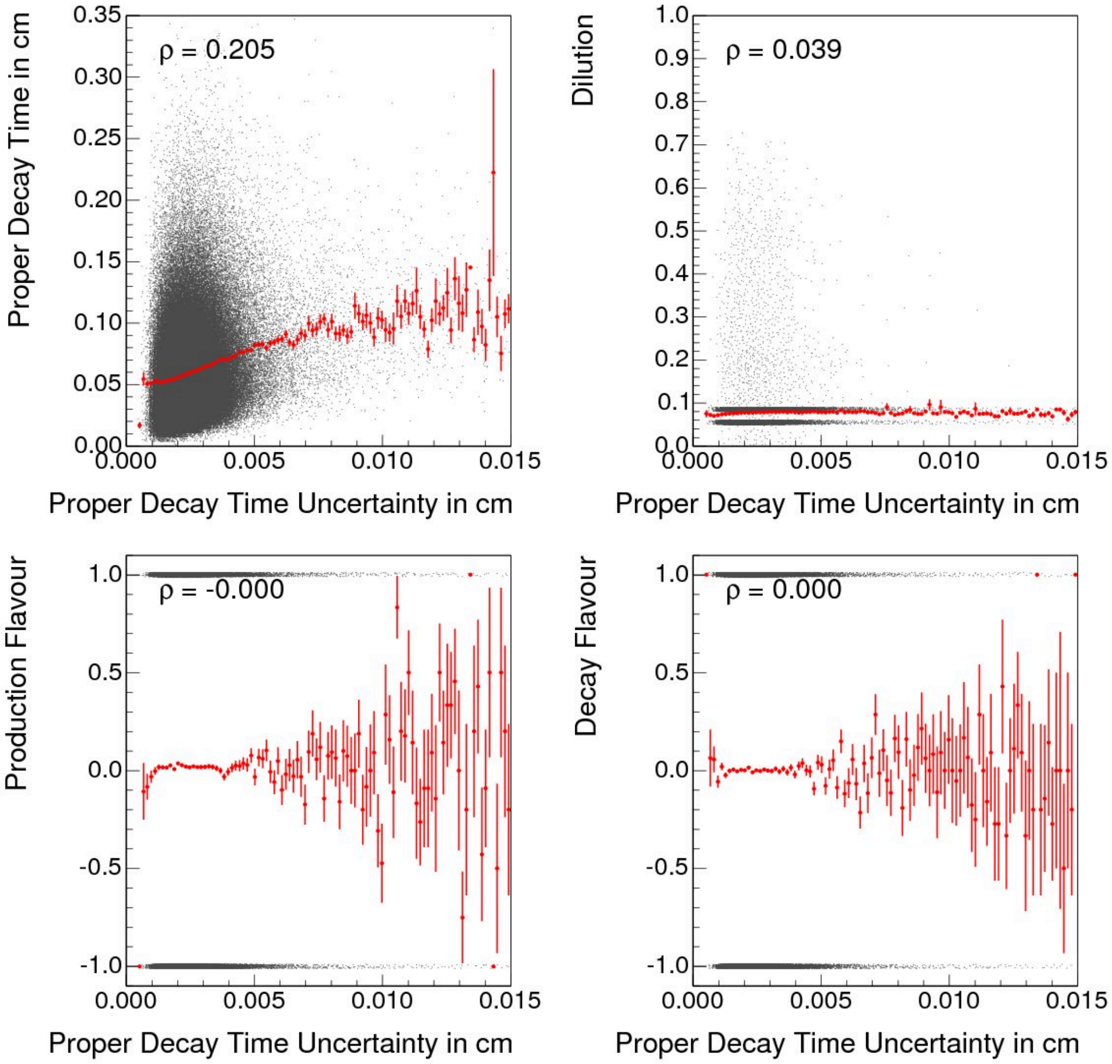


Figure 5.4: The correlation between proper decay time resolution and other variables is illustrated by combined scatter and profile plots. A data sample consisting of simulated B_s^0 mesons was used for their creation.

range of values of the proper decay time resolution is split into ten intervals. Within each interval, a separate efficiency function is determined. They are obtained by dividing the distribution of the measured proper decay time after trigger requirements and cuts by the sum of the proper decay time probability density functions assuming perfect acceptance:

$$\epsilon(ct) = \frac{\text{measured decay time after trigger and cuts}}{\sum_i \frac{1}{\tau} e^{-c\tilde{t}/c\tau} \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct,i})}. \tag{5.30}$$

The sum in the denominator runs hereby over all simulated events available. In each case the decay time uncertainty is used as width of the Gaussian resolution function. The result is a function of ct.

Summarising this, two equations are obtained, one for mesons changing their flavour (subscript *mix*) and one for mesons remaining unchanged (subscript *unmix*):

$$P_{Sig}(ct|\sigma_{ct}, unmix) = \frac{1}{N_{unmix}} \cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot \frac{1}{2}(1 + \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}), \tag{5.31}$$

$$P_{Sig}(ct|\sigma_{ct}, mix) = \frac{1}{N_{mix}} \cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot \frac{1}{2}(1 - \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}). \tag{5.32}$$

N_{mix} and N_{unmix} are normalisation constants. They are defined as

$$N_{\substack{unmix \\ mix}} = \int_{ct=-\infty}^{\infty} \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot \frac{1}{2}(1 \pm \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}) \; d(ct). \tag{5.33}$$

They can also be regarded as the area under the (unnormalised) probability density functions in proper decay time. N_{mix} and N_{unmix} can therefore be used to express f_{mix} (defined in equation 5.23):

$$f_{mix} = \frac{N_{mix}}{N_{mix} + N_{unmix}}. \tag{5.34}$$

The four equations for the different production and decay flavours can now be summarised using a simpler notation:

$$P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) = P_{Sig}(ct|\sigma_{ct}, unmix) \cdot P_{Sig}(\sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0), \tag{5.35}$$

$$P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to \bar{B}_s^0) = P_{Sig}(ct|\sigma_{ct}, mix) \cdot P_{Sig}(\sigma_{ct}, \xi_P, D|B_s^0 \to \bar{B}_s^0), \tag{5.36}$$

$$P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to B_s^0) = P_{Sig}(ct|\sigma_{ct}, mix) \cdot P_{Sig}(\sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to B_s^0), \tag{5.37}$$

$$P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0) = P_{Sig}(ct|\sigma_{ct}, unmix) \cdot P_{Sig}(\sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0). \tag{5.38}$$

The ct-independent terms on the right hand side of each equation can be factorised into two parts. The $B_s^0 \to B_s^0$-term is taken as example here:

$$P_{Sig}(\sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) = P_{Sig}(\sigma_{ct}|\xi_P, D, B_s^0 \to B_s^0) \cdot P_{Sig}(\xi_P, D|B_s^0 \to B_s^0). \tag{5.39}$$

Figure 5.4 shows that the proper decay time resolution σ_{ct} is independent from production flavour, decay flavour and dilution:

$$P_{Sig}(\sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) = P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|B_s^0 \to B_s^0). \tag{5.40}$$

The distributions of $P_{Sig}(\xi_P, D)$ are different for different initial states B_s^0 and $\bar{B}_s^0$. Otherwise tagging would have no effect.

Although there are no striking differences expected, positive and negative decay flavour are still distinguished in this dimension. Taking $P_{Sig}(\xi_P, D|B_s^0 \to B_s^0)$ as example, it can be seen as the conditional probability to observe a certain (ξ_P, D) value under the assumption of B_s^0 as true initial state flavour and B_s^0 as final state. Using Bayes' Law it can be expressed by the tagger output $P_{Sig}(B_s^0 \to |\xi_P, D)$:

$$P_{Sig}(\xi_P, D|B_s^0 \to B_s^0) = \frac{P_{Sig}(B_s^0 \to |\xi_P, D, \to B_s^0) \cdot P_{Sig}(\xi_P, D| \to B_s^0)}{P_{Sig}(B_s^0 \to B_s^0)}. \tag{5.41}$$

$P_{Sig}(B_s^0 \to B_s^0)$ represents the tagger's training sample composition of B_s^0 and $\bar{B}_s^0$. It can be written as $P_{Sig}(B_s^0 \to)$ because no information from the final state is used. In addition, a self-evident consensus is existing to set it to 50 %. Using the fact that the tagger output is a probability scaled between -1 and +1,

$$P_{Sig}(B_s^0 \to |\xi_P, D, \to B_s^0) = P_{Sig}(B_s^0 \to |\xi_P, D, \to \bar{B}_s^0) = \frac{1 + \xi_P \cdot D}{2}, \tag{5.42}$$

$$P_{Sig}(\bar{B}_s^0 \to |\xi_P, D, \to B_s^0) = P_{Sig}(\bar{B}_s^0 \to |\xi_P, D, \to \bar{B}_s^0) = \frac{1 - \xi_P \cdot D}{2}, \tag{5.43}$$

the following expressions are acquired:

$$P_{Sig}(\xi_P, D|B_s^0 \to B_s^0) = 2 \cdot P_{Sig}(\xi_P, D| \to B_s^0)\ \frac{1 + \xi_P \cdot D}{2}, \tag{5.44}$$

$$P_{Sig}(\xi_P, D|B_s^0 \to \bar{B}_s^0) = 2 \cdot P_{Sig}(\xi_P, D| \to \bar{B}_s^0)\ \frac{1 + \xi_P \cdot D}{2}, \tag{5.45}$$

$$P_{Sig}(\xi_P, D|\bar{B}_s^0 \to B_s^0) = 2 \cdot P_{Sig}(\xi_P, D| \to B_s^0)\ \frac{1 - \xi_P \cdot D}{2}, \tag{5.46}$$

$$P_{Sig}(\xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0) = 2 \cdot P_{Sig}(\xi_P, D| \to \bar{B}_s^0)\ \frac{1 - \xi_P \cdot D}{2}. \tag{5.47}$$

Using these results, the equations for the four possible transitions can again be summarised:

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to B_s^0) = {} & P_{Sig}(ct|\sigma_{ct}, unmix) \cdot P_{Sig}(\sigma_{ct}) \\ & \cdot P_{Sig}(\xi_P, D| \to B_s^0) \cdot (1 + \xi_P D), \end{aligned} \tag{5.48}$$

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D|B_s^0 \to \bar{B}_s^0) = {} & P_{Sig}(ct|\sigma_{ct}, mix) \cdot P_{Sig}(\sigma_{ct}) \\ & \cdot P_{Sig}(\xi_P, D| \to \bar{B}_s^0) \cdot (1 + \xi_P D), \end{aligned} \tag{5.49}$$

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to B_s^0) = {} & P_{Sig}(ct|\sigma_{ct}, mix) \cdot P_{Sig}(\sigma_{ct}) \\ & \cdot P_{Sig}(\xi_P, D| \to B_s^0) \cdot (1 - \xi_P D), \end{aligned} \tag{5.50}$$

$$\begin{aligned} P_{Sig}(ct, \sigma_{ct}, \xi_P, D|\bar{B}_s^0 \to \bar{B}_s^0) = {} & P_{Sig}(ct|\sigma_{ct}, unmix) \cdot P_{Sig}(\sigma_{ct}) \\ & \cdot P_{Sig}(\xi_P, D| \to \bar{B}_s^0) \cdot (1 - \xi_P D). \end{aligned} \tag{5.51}$$

By using the relations 5.34, 5.48 and 5.50, equation 5.26 can be expressed as:

$$\begin{aligned}
&P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \rightarrow B_s^0) \\
&\quad = \frac{1}{N_{mix} + N_{unmix}} \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D| \rightarrow B_s^0) \\
&\quad \cdot [N_{unmix} \cdot P_{Sig}(ct|\sigma_{ct}, unmix) \cdot (1 + \xi_P D) + N_{mix} \cdot P_{Sig}(ct|\sigma_{ct}, mix) \cdot (1 - \xi_P D)].
\end{aligned} \tag{5.52}$$

The last line of this equation can be explicitly written as

$$\begin{aligned}
&N_{unmix} \cdot P_{Sig}(ct|\sigma_{ct}, unmix) \cdot (1 + \xi_P D) \\
&+ N_{mix} \cdot P_{Sig}(ct|\sigma_{ct}, mix) \cdot (1 - \xi_P D) \\
&= \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot \frac{1}{2}(1 + \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}) \cdot (1 + \xi_P D) \qquad (5.53) \\
&+ \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot \frac{1}{2}(1 - \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}) \cdot (1 - \xi_P D)
\end{aligned}$$

By using the distributivity and associativity with scalar multiplication of the convolution, the following expression is acquired:

$$= \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 + \xi_P D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}). \tag{5.54}$$

In this way, a final result can be given for equation 5.26:

$$\begin{aligned}
P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \rightarrow B_s^0) &= \frac{1}{N_{mix} + N_{unmix}} \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D| \rightarrow B_s^0) \\
&\cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 + \xi_P D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}).
\end{aligned} \tag{5.55}$$

A similar derivation can be made for final state $\bar{B}_s^0$ mesons (equation 5.27) leading to

$$\begin{aligned}
P_{Sig}(ct, \sigma_{ct}, \xi_P, D| \rightarrow \bar{B}_s^0) &= \frac{1}{N_{mix} + N_{unmix}} \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D| \rightarrow \bar{B}_s^0) \\
&\cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 - \xi_P D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}).
\end{aligned} \tag{5.56}$$

Both expressions only differ in the sign in front of the $\xi_P D \cdot \cos(\Delta m t)$ term. They can therefore be merged into a common equation

$$\begin{aligned}
P_{Sig}(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= \frac{1}{N_{mix} + N_{unmix}} \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D) \\
&\cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 + \xi_D \xi_P D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct})
\end{aligned} \tag{5.57}$$

$$=: P_{Sig}(ct|\sigma_{ct}, \xi_D, \xi_P, D) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D). \tag{5.58}$$

This allows for the definition of the probability density function in proper decay time for a given B_s^0 meson:

$$P_{Sig}(ct|\sigma_{ct}, \xi_D, \xi_P, D) = \frac{1}{N_{mix} + N_{unmix}} \cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 + \xi_D \xi_P D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}). \tag{5.59}$$

The product of production and decay flavour $\xi_D \cdot \xi_P$ carries the information whether a given particle has mixed or not. It can therefore be merged into a new variable:

$$\xi := \xi_D \cdot \xi_P. \tag{5.60}$$

Both normalisations N_{mix} and N_{unmix} can also be merged into a common quantity $N = N_{mix} + N_{unmix}$. Apart from that, D is replaced by $\mathcal{A} \cdot D$. In this way the mixing amplitude is introduced. It is a fit parameter compensating for wrong dilution assessments. From its definition the following interpretation can be derived:

$$\mathcal{A} \begin{cases} < 1 & \text{tagger overestimates itself} \\ = 1 & \text{optimal value} \\ > 1 & \text{tagger underestimates itself} \end{cases} \tag{5.61}$$

The final expression which is used in this analysis is then

$$P_{Sig}(ct|\sigma_{ct}, \xi_D, \xi_P, D) = \frac{1}{N} \cdot \left\{ \left[\frac{1}{\tau} e^{-\tilde{t}/\tau} \cdot (1 + \xi \mathcal{A} D \cdot \cos(\Delta m \tilde{t})) \right] \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}). \tag{5.62}$$

Nearly all functions introduced are acquired by the parametrisation of a corresponding distribution. This is described in section 7.1. However, for $P_{Sig}(\xi_P, D|\xi_D)$ two histograms are used. Each one is responsible for a given decay flavour and carries $\xi_P \cdot D$ on the x-axis. In addition, variable bin widths are used to maintain a sufficient amount of statistics in each interval.

5.3.3 Combinatorial Background

While the probability density function for the signal contains physics motivations, the parametrisation of the combinatorial background is purely empirical. Nevertheless, correlations must also be considered here. Figure 5.5 illustrates the correlations between proper decay time and other variables for events taken from the combinatorial background. Apart from the strong correlation between proper decay time and proper decay time resolution, no significant value can be observed. This leads to the following equation:

$$\begin{aligned} P_{Comb}(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_{Comb}(ct|\sigma_{ct}, \xi_D, \xi_P, D) \cdot P_{Comb}(\sigma_{ct}, \xi_D, \xi_P, D) \qquad (5.63) \\ &= P_{Comb}(ct|\sigma_{ct}) \cdot P_{Comb}(\sigma_{ct}, \xi_D, \xi_P, D). \qquad (5.64) \end{aligned}$$

Figure 5.6 shows the correlation between the proper decay time resolution and production flavour, decay flavour and dilution, respectively. The biggest correlation

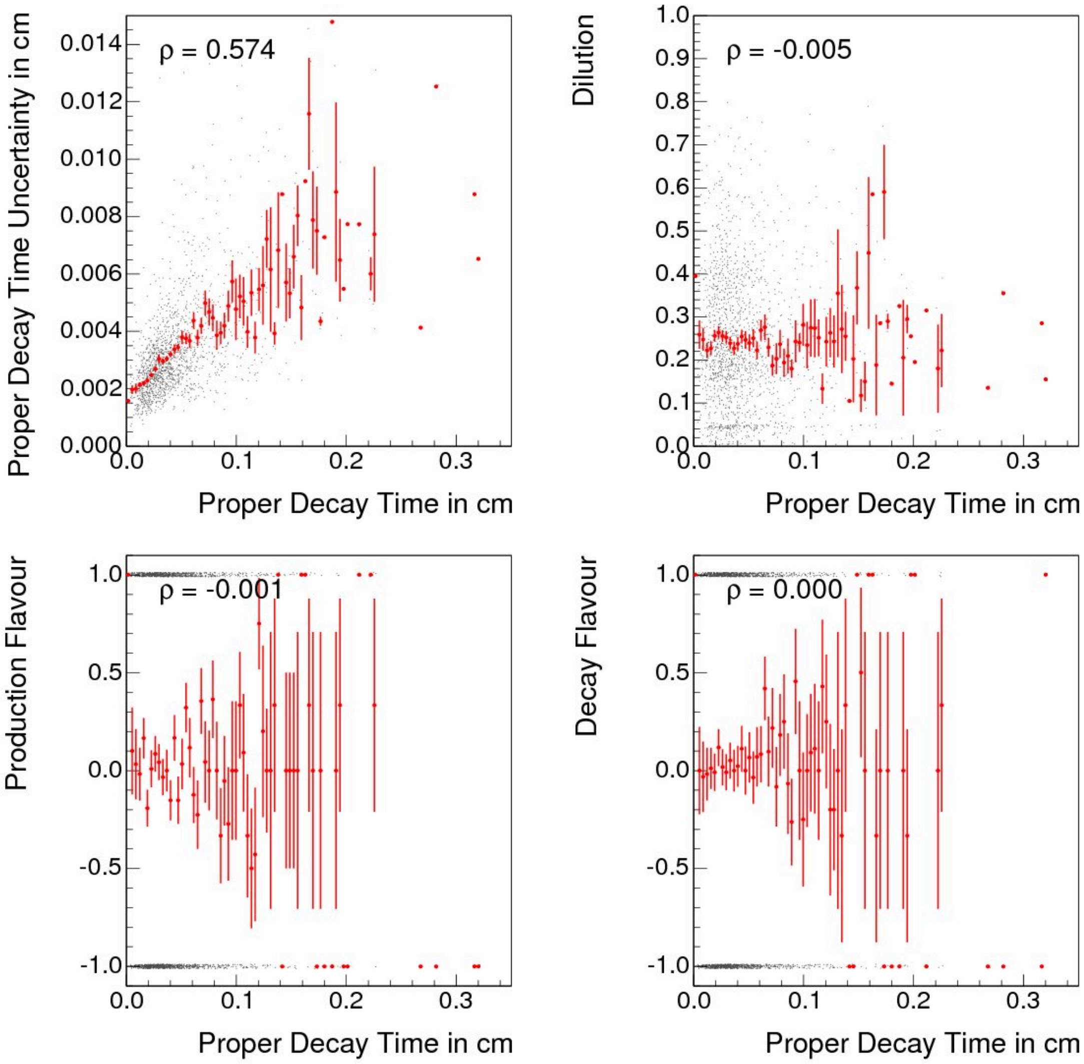


Figure 5.5: The correlation between proper decay time and other variables is illustrated by combined scatter and profile plots. Events taken from the combinatorial background of a B_s^0 dataset was used for their creation.

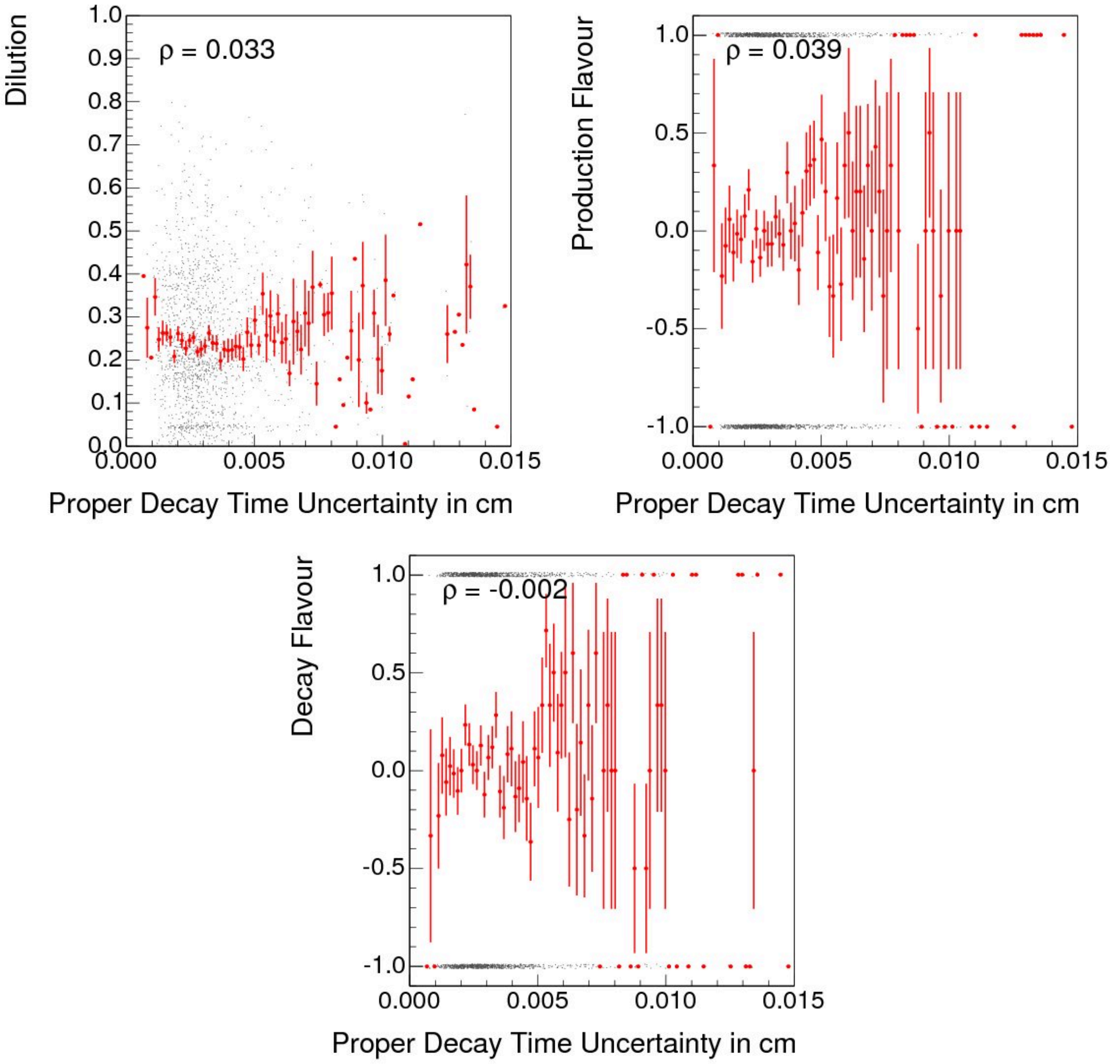


Figure 5.6: The correlation between proper decay time resolution and other variables is illustrated by combined scatter and profile plots. Events taken from the combinatorial background of a B_s^0 dataset was used for their creation.

amounts 3.9 %. They can therefore be neglected and in this way the following expression can be acquired:

$$\begin{aligned} P_{Comb}(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_{Comb}(ct|\sigma_{ct}) \cdot P_{Comb}(\sigma_{ct}, \xi_D, \xi_P, D) & (5.65)\\ &= P_{Comb}(ct|\sigma_{ct}) \cdot P_{Comb}(\sigma_{ct}) \cdot P_{Comb}(\xi_D, \xi_P, D). & (5.66)\end{aligned}$$

The combinatorial background originates mainly from prompt tracks. Its proper decay time distribution is therefore basically a Gaussian around zero with a width corresponding to the proper decay time resolution σ_{ct}. This means, for example, that in events with large σ_{ct} there can be large proper decay time values ct. This correlation between ct and σ_{ct} is to first order removed if one looks at the proper decay time significance

$$ct_{Signi} := \frac{ct}{\sigma_{ct}}. \tag{5.67}$$

This is shown in figure 5.7. As expected, the correlation between proper decay time significance and proper decay time resolution is significantly reduced. The miscellaneous correlations are negligible. The approach is therefore to parametrise the significance distribution and transform it event-by-event into a proper decay time distribution according to the following law:

$$P_{Comb}(ct|\sigma_{ct}) = P_{Comb}(ct_{Signi}) \cdot \left|\frac{d(ct_{Signi})}{d(ct)}\right| = \frac{P_{Comb}(ct_{Signi})}{\sigma_{ct}}. \tag{5.68}$$

The tagging term $P_{Comb}(\xi_P, D|\xi_D)$ is treated in a similar way as for the signal: two histograms, one for each decay flavour, are used.

5.3.4 Physics Reflections

The ansatz for the probability density functions of the physics reflections (B and Λ_b) is the following:

$$\begin{aligned} P_{\Lambda_b}(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_{\Lambda_b}(ct) \cdot P_{\Lambda_b}(\sigma_{ct}) \cdot P_{\Lambda_b}(\xi_P, D|\xi_D), & (5.69)\\ P_B(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_B(ct) \cdot P_B(\sigma_{ct}) \cdot P_B(\xi_P, D|\xi_D). & (5.70)\end{aligned}$$

As described in subsection 4.4, simulated datasets are created containing simulated B^0 and Λ_b events which were reconstructed as B_s^0. Parametrisations in mass and proper decay time are acquired by those datasets. For the proper decay time resolution and the production flavour/dilution terms the same distributions as for the signal are used:

$$\begin{aligned} P_{\Lambda_b}(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_{\Lambda_b}(ct) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D), & (5.71)\\ P_B(ct, \sigma_{ct}, \xi_D, \xi_P, D) &= P_B(ct) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D). & (5.72)\end{aligned}$$

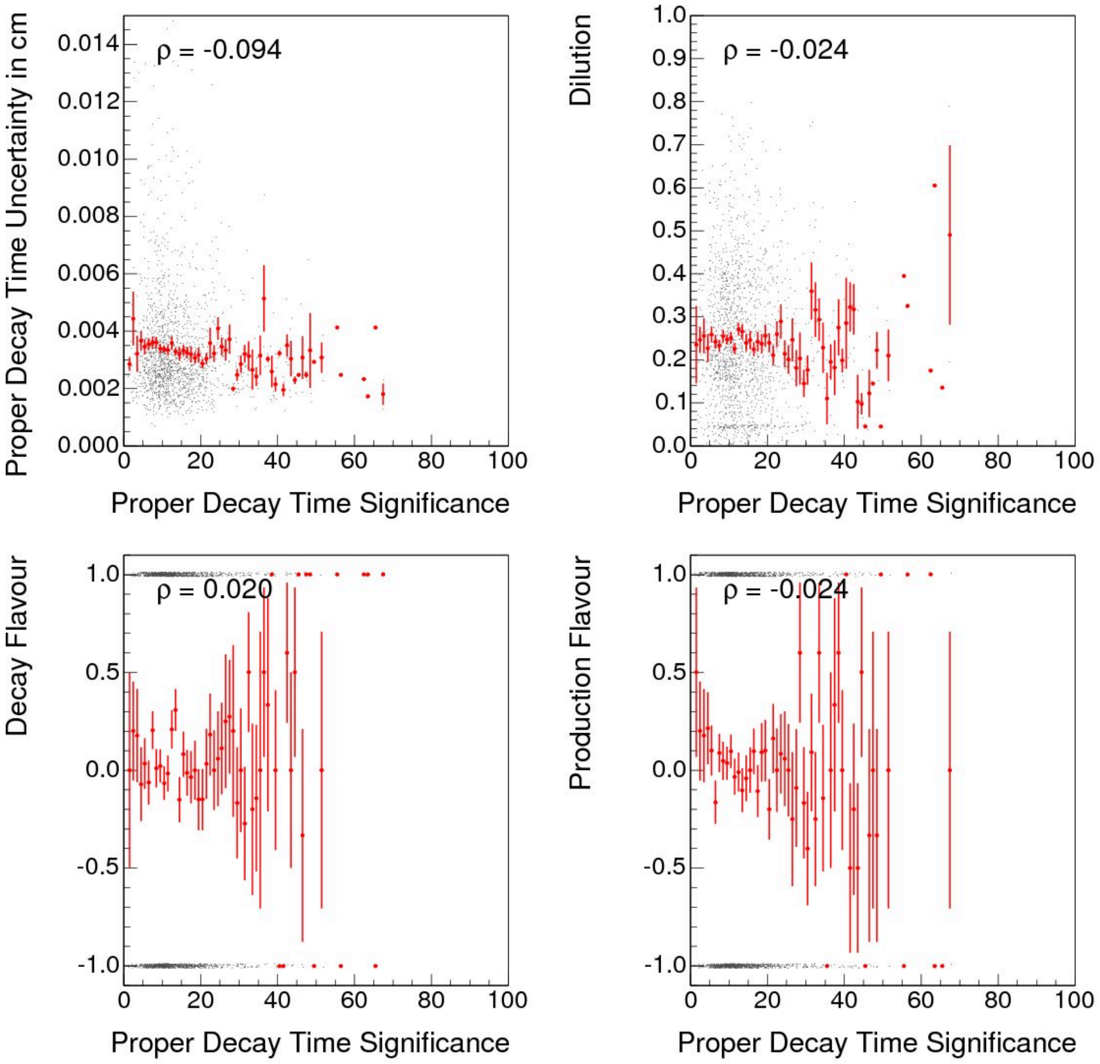


Figure 5.7: The correlation between proper decay time significance and other variables is illustrated by combined scatter and profile plots. Events taken from the combinatorial background of a B_s^0 dataset was used for their creation.

5.3.5 Summary

Summarising the last subsection, the following probability density function is obtained:

$$\begin{aligned} P(m, ct, \sigma_{ct}, \xi_D, \xi_P, D) = {} & f_{Sig} \cdot P_{Sig}(m) \cdot P_{Sig}(ct|\sigma_{ct}, \xi_D, \xi_P, D) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D) \\ & + f_{Comb} \cdot P_{Comb}(m) \cdot P_{Comb}(ct|\sigma_{ct}) \cdot P_{Comb}(\sigma_{ct}) \cdot P_{Comb}(\xi_P, D|\xi_D) \\ & + f_{\Lambda_b} \cdot P_{\Lambda_b}(m) \cdot P_{\Lambda_b}(ct) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D) \\ & + f_B \cdot P_B(m) \cdot P_B(ct) \cdot P_{Sig}(\sigma_{ct}) \cdot P_{Sig}(\xi_P, D|\xi_D). \end{aligned} \tag{5.73}$$

6 Analysis

In this analysis, various fits are performed successively. In some cases the result of a given fit depends on other fits done before. At the very top of this hierarchy stands the final fit where mixing frequency and scale factor are determined. In this sense the final results depend on every step taken before.

The best estimations of the mixing frequency and the mixing amplitude are done using different tagger configurations. However except for this last step, both measurements are identical. The first, tagger independent part is described in this chapter.

It was a trait of the analysis at hand to treat all decay channels in a similar way. Because of this and for reasons of clarity, the following section focusses only on the most yielding B_s^0 channel

$$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-.$$

Results for the other decay modes will be discussed later. Further technical details can be found in Ref. [65].

6.1 Mass Template Fits

The very first step in the whole fit procedure are mass template fits. Here various simulated datasets, one for each contribution expected, are used. The distribution of the invariant mass in each sample is fitted by an adequate phenomenological function. This means that no physics parameters are contained inside those functions. It is also noteworthy that this step only determines the shape of the corresponding function - not the fraction in the data sample. If it is not explicitly mentioned, all parameters will be kept fixed after the fit throughout the rest of the whole procedure.

6.1.1 Proper Signal

The simulated data sample for the signal also contains final state radiation events. This leads to a slightly asymmetric distribution. The sum of three Gaussians with

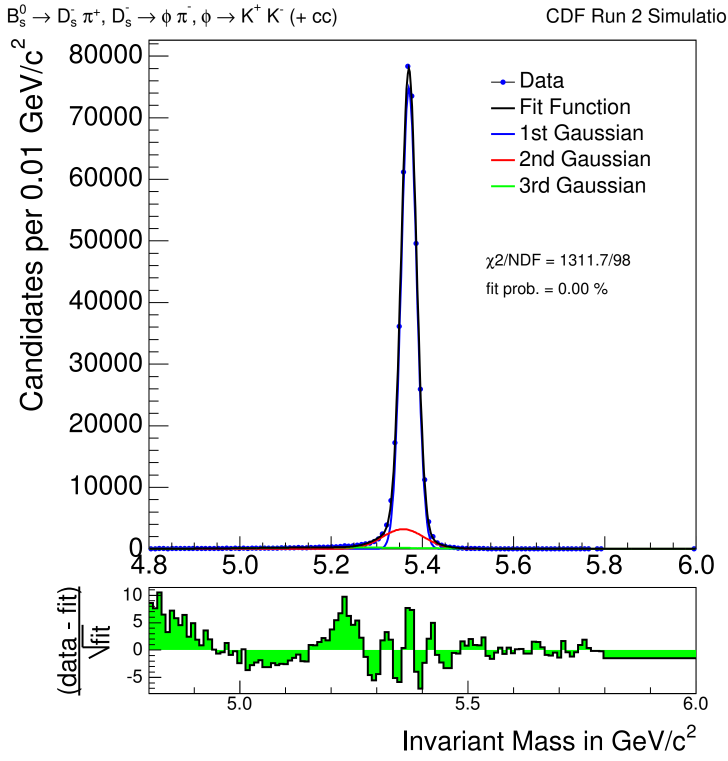


Figure 6.1: Fit to simulated signal events.

different standard deviation and mean value is therefore used as fit function here:

$$\begin{aligned}&P_{Sg}(m|N_{Sg}, f_1, f_2, \mu_1, \sigma_1, \mu_2, \sigma_2, \mu_3, \sigma_3)\\ &\qquad = N_{Sg} \cdot \Big[(1-f_1)\cdot \mathcal{G}(m|\mu_1,\sigma_1) + f_1 \cdot [f_2 \cdot \mathcal{G}(m|\mu_2,\sigma_2) + (1-f_2)\cdot \mathcal{G}(m|\mu_3,\sigma_3)]\Big].\end{aligned} \tag{6.1}$$

The Gaussians are normalised to 1.0 within the wide mass boundaries 4.8 GeV/c^2 and 6.0 GeV/c^2. The result of this binned fit can be seen in figure 6.1. The ratio of χ^2 and the number of degrees of freedom is not satisfactory for this fit. However it should noted that the number of events used for these template fits is higher than the number of events expected in the data sample by at least one order of magnitude. The quality of the fits are therefore considered to be sufficient for the expected signal yield.

6.1.2 Cabibbo Reflection

A phenomenological approach is also used for the Cabibbo suppressed signal contribution:

$$\begin{aligned}&P_{Cb}(m|N_{Cb}, f, \mu_1, \sigma_1, \mu_2, \sigma_2, \tau_2)\\ &\qquad = N_{Cb} \cdot [f \cdot \mathcal{G}(m|\mu_1,\sigma_1) + (1-f)\cdot \mathcal{H}(m|\mu_2,\sigma_2,\tau_2)].\end{aligned} \tag{6.2}$$

$\mathcal{G}$ denotes a Gaussian, $\mathcal{H}$ denotes a Gaussian convoluted with an exponential function. Both are normalised to 1.0 within the wide mass boundaries 4.8 GeV/c^2 and

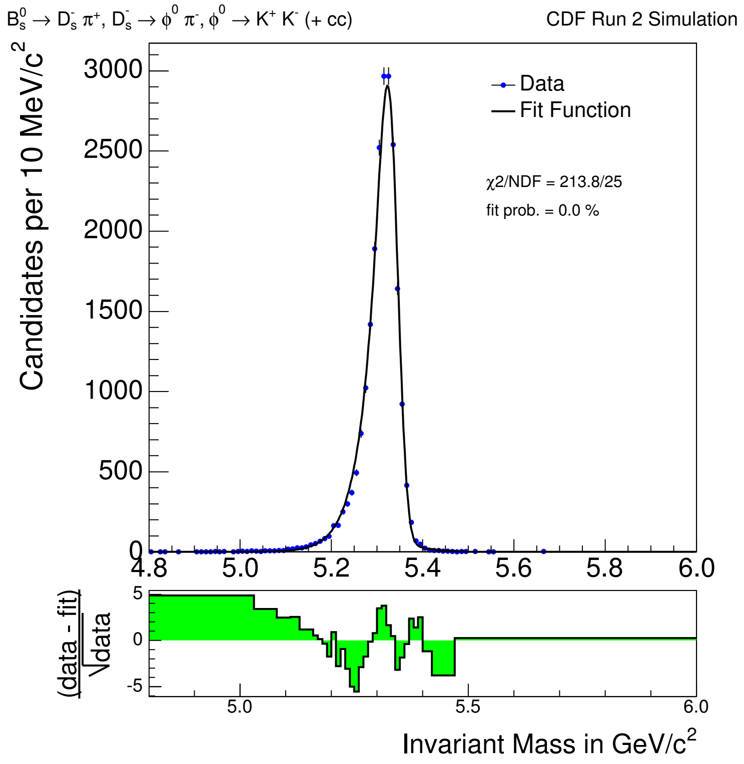


Figure 6.2: Fit to simulated Cabibbo signal events.

6.0 GeV/c^2. The result of the fit can be seen in figure 6.2. Regarding the low fit quality it must be noted here as well that the number of simulated events used for this fit exceeds the number of events expected in measured data by at least one order of magnitude.

6.1.3 Physics Reflections

For the parametrisation of the physics reflections, $B^0 \to D^- \pi^+$ and $\Lambda_b^0 \to \Lambda_c^+ \pi^-$, which are expected in some of the B_s^0 decay channels, similar functions as for the Cabibbo suppressed signal are used (equation 6.2).

6.1.4 Partially Reconstructed Decays

As mentioned earlier two partially reconstructed contributions possess a distinct shape: the $B_s^0 \to D_s^{*-} \pi^+$ contribution (P_{DS}) is assumed as Gaussian. A function similar to equation 6.2 is used for the $B_s^0 \to D_s^- \rho^+$ contribution (P_{Rh}). The $B \to D_s^- X^+$ contribution is described by the sum of two linear functions:

$$P_{Ln}(m|N_{Ln}, f, m_{01}, m_{02}) = N_{Ln} \cdot (f \cdot \mathcal{P}(m|m_{01}) + (1-f) \cdot \mathcal{P}(m|m_{02})). \qquad (6.3)$$

The parameters m_{01} and m_{02} are positions where the linear functions first cross $m = 0$.

The result of the fit to the combined sample of partially reconstructed decays can be seen in figure 6.3.

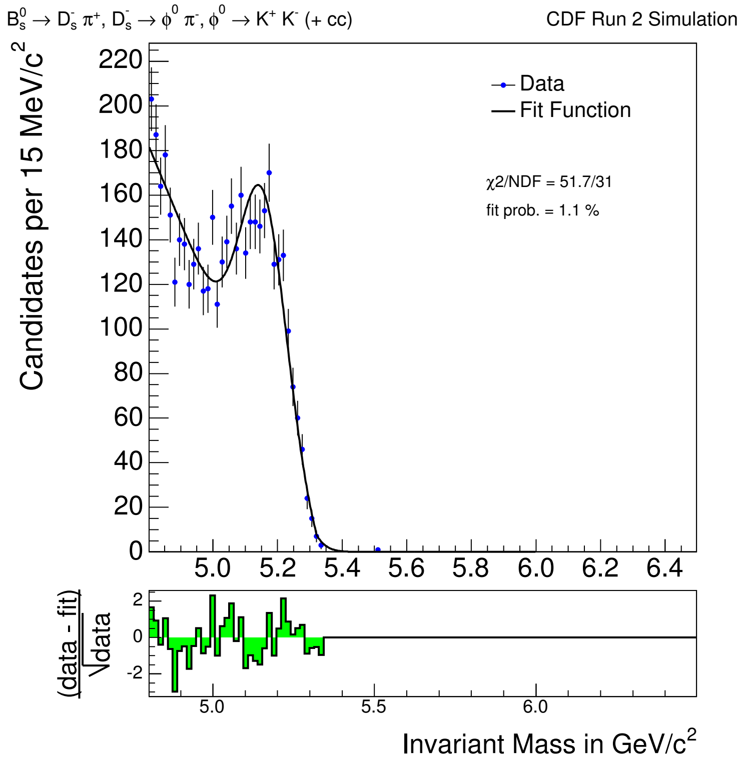


Figure 6.3: Plot showing a fit to partially reconstructed B_s decays which were acquired by a simulation.

6.2 Wide Mass Fit

The shapes which were acquired by separate mass fits to simulated events are now combined and applied on measured data. This is done in order to get a first understanding of the different contributions the sample is composed of. Another reason is to acquire good starting values for parameters, which are left free inside the subsequent unbinned fit. The mass window used for this goes from 4.8 GeV/c^2 to 6.0 GeV/c^2.

For the generation of the simulated dataset, which were used in the last section, a certain magnetic field was assumed. In past analyses differences between the assumed and the actual magnetic field were observed. In the analysis at hand a parameter k is introduced to compensate for such discrepancies. Every localisation parameter of every mass template function is multiplied by this quantity.

So far the shape of every contribution except for the combinatorial background has been modelled. For this the following function is now assumed:

$$P_{Co}(m|N_{Co}, f, \lambda) = N_{Co} \cdot [f \cdot \mathcal{E}(m|\lambda) + (1-f) \cdot \mathcal{K}(m)]. \tag{6.4}$$

$\mathcal{E}$ denotes hereby an exponential function and $\mathcal{K}$ a constant function.

In order to handle the different normalisations, a binary tree is used. It can be seen in figure 6.4. Here the fractions f^*_{Cb}, f^*_{Ph}, f^*_{Lm}, f_{Bg}, f_{Co}, f_{Ln} and f_{DS} are introduced. They can be divided into two classes: f_{Bg}, f_{Co}, f_{Ln} and f_{DS}, the fractions without a star, are direct fit parameters. The fractions which are designated by a star (f^*_{Cb},

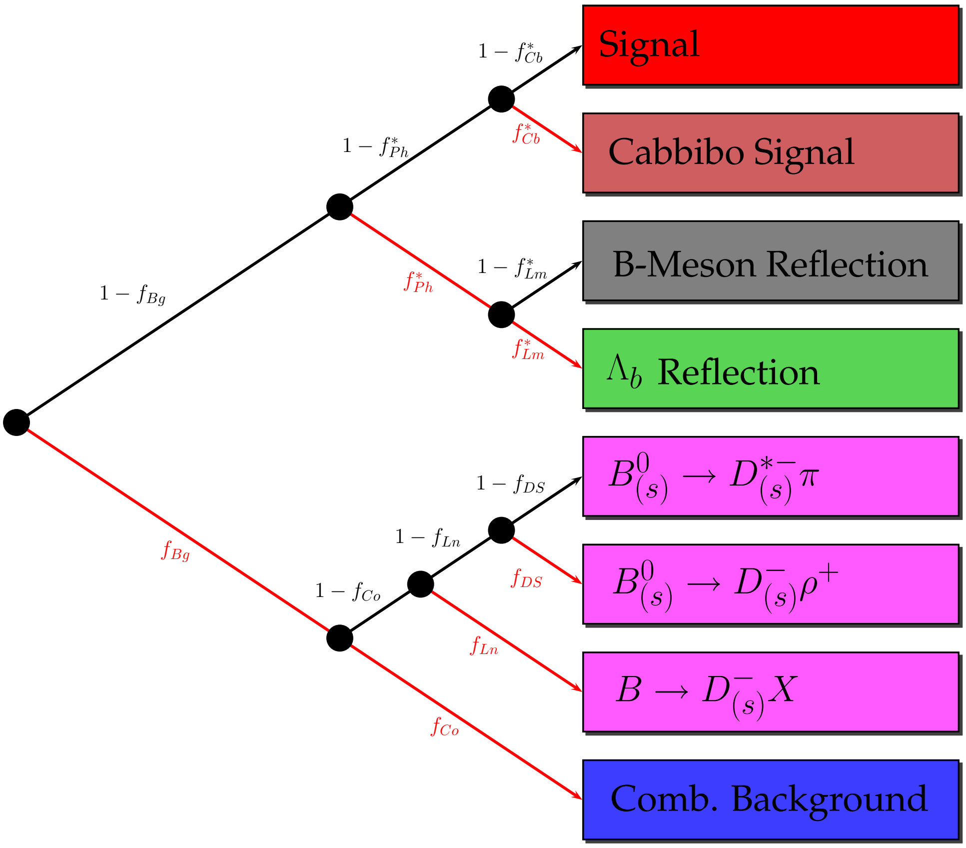


Figure 6.4: Binary tree used for the determination of fractions and normalisations. Fractions designated by a star are indirect fit parameters, which means that they are calculated by other (direct) fit parameters. The colours of the different contributions are the same as in the projections show in section 6.6.

f^*_{Ph} and f^*_{Lm}) are indirect fit parameters. This means that they are calculated using the direct fit parameters A_{Cb}, A_{Lm} and A_{B0}, which were defined in section 4.3:

$$f^*_{Cb} = \frac{A_{Cb}}{1.0 + A_{Cb}}, \tag{6.5}$$

$$f^*_{Lm} = \begin{cases} 0 & \text{if } A_{Lm} = 0 \\ \frac{A_{Lm}}{A_{Lm}+A_{B0}} & \text{if } A_{Lm} > 0 \end{cases}, \tag{6.6}$$

$$f^*_{Ph} = \frac{A_{B0} + A_{Lm}}{1.0 + A_{Cb} + A_{Lm} + A_{B0}}. \tag{6.7}$$

By utilising such a binary tree only one global normalisation parameter N is further necessary. It corresponds to the area enclosed under the total mass function. The other normalisation values can easily be acquired by following the corresponding path from trunk to leaf and multiplying the dedicated fractions.

Taking the Λ_b normalisation as example, it is acquired by the following calculation:

$$N_{\Lambda_b} = N \cdot (1 - f_{Bg}) \cdot (1 - f^*_{Ph}(A_{Cb}, A_{Lm}, A_{B0})) \cdot f^*_{Lm}(A_{Lm}, A_{B0}). \tag{6.8}$$

There are two advantages in using such a tree: the first one is that the range of values of the tree fractions are trivial to provide: every value between 0 and 1 is allowed. This is for example not the case for a non-binary tree. The second advantage is that every tree configuration meets normalisation constraints like equation 5.11 on page 50.

The fractions of both the physics backgrounds and the Cabibbo suppressed signal are fixed to the values presented in table 4.2 on page 45. f_{Bg} and f_{Co} are left free. Further both remaining parameters λ and f of the combinatorial background are released. In addition the compensation parameter for discrepancies between simulated and measured data is also left free. Altogether this makes six parameters, which are left free: N, f_{Bg}, f_{Co}, λ, f and k.

Using this setup a binned fit is performed. Amongst other things the following value for the shift parameter is obtained:

$$k = 0.999373 \pm 0.000078. \tag{6.9}$$

Throughout the rest of the whole fit procedure this value is kept fixed.

In the second step, the position and standard deviation of the dominant Gaussian, the signal is composed of, is also released and another binned fit is performed.

The result can be seen in figure 6.5. The dashed vertical lines contain the signal range. It was chosen from 5.32 GeV/c^2 to 5.42 GeV/c^2. The quality of those fits is not satisfying, however sufficient for getting good starting values.

It is clear to see that events with an invariant mass higher than 5.55 GeV/c^2, belong exclusively to combinatorial background. The interval from 5.55 GeV/c^2 to 6.0 GeV/c^2, which is in the following referred to as upper side band, will therefore be used to derive models for it, e.g. for the resolution of the proper decay time.

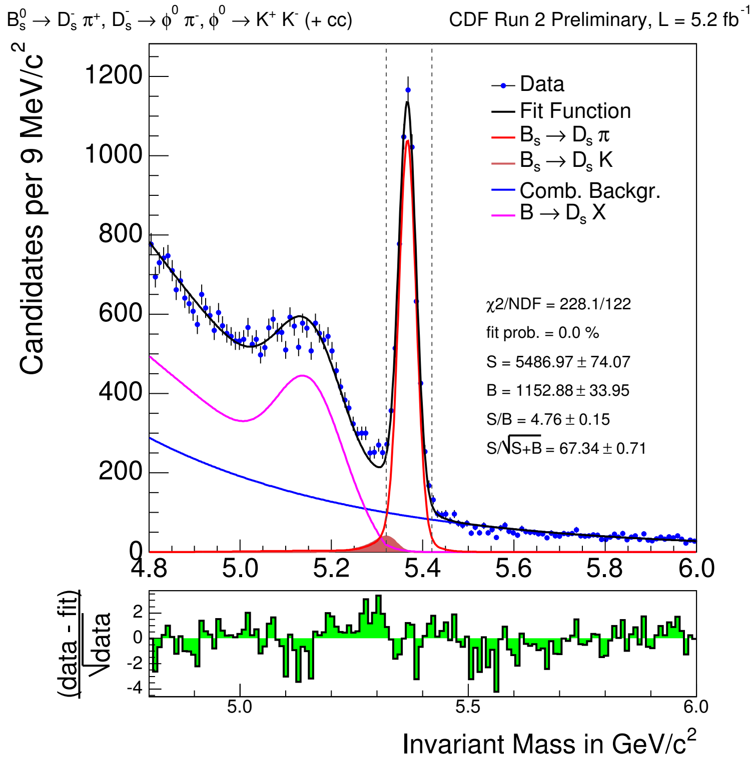


Figure 6.5: A binned fit is performed on data in a wide mass window. S and B denote the number of events between the dashed lines: S is the number of proper signal events, B includes the Cabibbo suppressed signal, partially reconstructed decays, physics reflections and combinatorial background. The root of the respective number is taken as standard deviation and does therefore not include fit uncertainties.

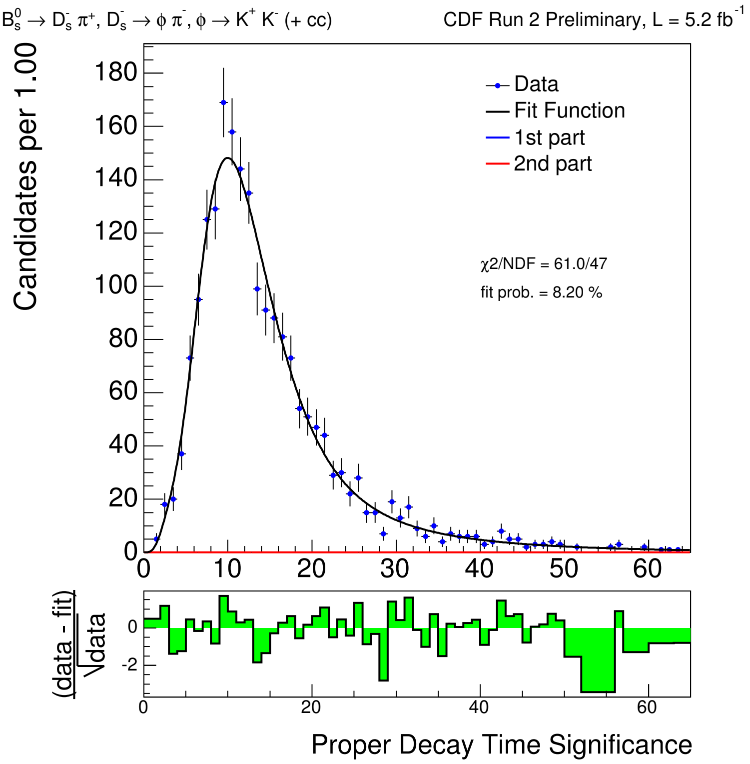


Figure 6.6: Fit to the significance of the proper decay time. The events were taken from the upper sideband in data.

6.3 Description of the Combinatorial Background in Proper Decay Time

As described in section 5.3.3, a strong correlation between the proper decay time and its uncertainty is observed. Therefore a parametrisation of the significance of the proper decay time is determined. The distribution is taken from the upper sideband in data. The sum of two log-normal functions is used for parametrisation:

$$\begin{aligned} &P_{Comb}(ct_{Signi}|N, f, \mu_1, \theta_1, \sigma_1, \mu_2, \theta_2, \sigma_2) \\ &\quad = N \cdot [f \cdot \mathcal{N}(m|\mu_1, \theta_1, \sigma_1) + (1-f) \cdot \mathcal{N}(m|\mu_2, \theta_2, \sigma_2)]. \end{aligned} \tag{6.10}$$

The result of the corresponding binned maximum likelihood fit can be seen in figure 6.6.

For a given event with a given proper decay time resolution σ, the corresponding probability density function is obtained by the following law:

$$P_{Comb}(ct|\sigma_{ct}) = P_{Comb}(ct_{Signi}(ct)) \left| \frac{dct_{Signi}(ct)}{dct} \right| = \frac{P_{Comb}(ct_{Signi}(ct))}{\sigma_{ct}}. \tag{6.11}$$

In order to check whether this event-by-event probability density function fits in proper decay time for each resolution value, a test quantity is designed. It is defined for each event as the integral over the function $P_{Comb}(ct|\sigma_{ct})$ from $-\infty$ to ct:

$$\int_{-\infty}^{ct} P_{Comb}(ct'|\sigma_{ct}) d(ct'). \tag{6.12}$$

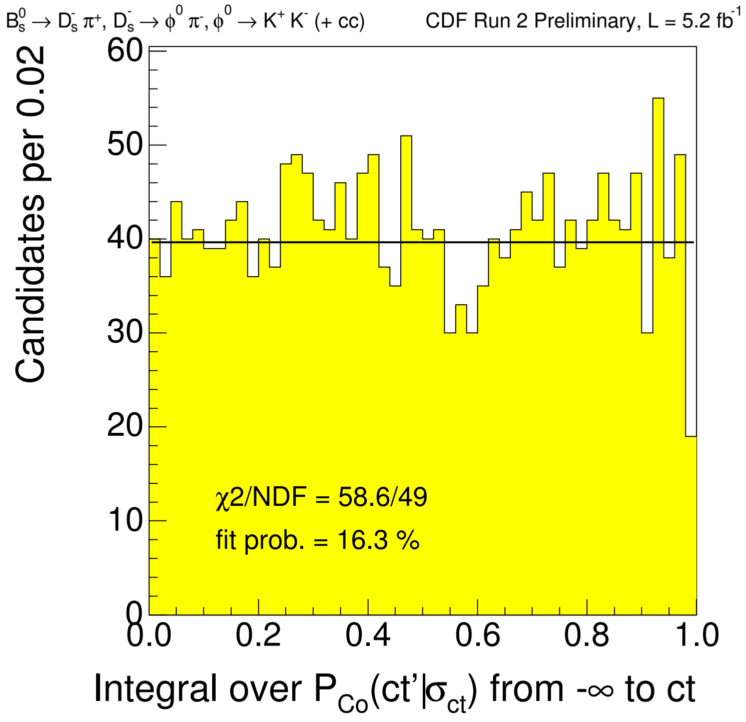


Figure 6.7: Plot showing the integral over the event-by-event probability density function in proper decay time for the combinatorial background. The boundaries are $-\infty$ and ct. A flat distribution is expected.

$\sigma_{ct,1}$	$\sigma_{ct,2}$	$\sigma_{ct,3}$	$\sigma_{ct,4}$	$\sigma_{ct,5}$
0.0016	0.0019	0.0021	0.0023	0.0025

$\sigma_{ct,6}$	$\sigma_{ct,7}$	$\sigma_{ct,8}$	$\sigma_{ct,9}$	$\sigma_{ct,10}$
0.0027	0.0029	0.0033	0.0038	0.0150

Table 6.1: Upper boundaries for the different efficiency functions in proper decay time resolution for the B_s^0 data sample. All values are given in cm.

Only events from the upper side band, which meet all selection criteria, are used for this test. In an ideal case, a uniform distribution is expected. Because of the normalisation of P_{Comb}, its boundaries are 0.0 and 1.0. The result can be seen in figure 6.7. It looks looks flat as expected.

6.4 Efficiency Fits

The efficiency functions are determined using simulated data. Because of the correlation between proper decay time and proper decay time uncertainty, ten efficiency functions are used. Each one is responsible for a certain resolution range. The corresponding boundaries can be found in table 6.1 and in figure 6.8. They were chosen in a way that the same number of events are contained within each interval.

Within each of those intervals, an efficiency function is determined. They are obtained by dividing the distribution of the measured proper decay time after trigger requirements and cuts by the sum of the proper decay time probability density functions assuming perfect acceptance:

$$\epsilon(ct) = \frac{\text{measured decay time after trigger and cuts}}{\sum_i \frac{1}{\tau} e^{-\tilde{ct}/c\tau} \otimes G(\tilde{ct}|\sigma_{ct,i})}. \tag{6.13}$$

Each efficiency function is fitted by the phenomenological function

$$\epsilon(ct|N, f_1, f_2, \beta_{1\ldots3}, \tau_{1\ldots3}) = N \cdot \sum_{i=1}^{3} f_i \cdot e^{-\frac{ct}{\tau_i}} (\beta_i + ct)^2 \cdot \Theta(ct - \beta_i) \tag{6.14}$$

with $f_3 = 1 - f_1 - f_2$.

The distributions in the different intervals and the corresponding fit functions can be found in figure 6.8.

6.5 Proper Decay Time Resolution

In order to acquire a description of the uncertainty of the proper decay time $P(\sigma_{ct})$, a sideband subtraction on measured data is performed: this means that the distribution is drawn for all events contained inside the signal range. Afterwards a second distribution using events taken from the upper side band is drawn and normalised to the number of combinatorial background events which are expected inside the signal range. This number is acquired by the wide mass fit, which was described in section 6.2. The distribution obtained from the sideband is subtracted from the distribution of the signal range. Besides of signal, the distribution may also contain some partially reconstructed events and also physics reflections. However this is not considered as contamination for the distribution acquired is used for all classes except for combinatorial background.

The distribution of the proper decay time uncertainty of the combinatorial background is obtained by plotting this variable only for events, which are located inside the mass upper side band.

Both distributions are fitted by the sum of two log-normal distributions, which are normalised between 0.0 cm and 0.015 cm:

$$P(\sigma_{ct}|N, f, \mu_1, \theta_1, \sigma_1, \mu_2, \theta_2, \sigma_2) = N \cdot [f \cdot \mathcal{N}(m|\mu_1, \theta_1, \sigma_1) + (1 - f) \cdot \mathcal{N}(m|\mu_2, \theta_2, \sigma_2)]. \tag{6.15}$$

The result is shown in figure 6.9 for the combinatorial background and in figure 6.10 for the signal.

6.6 Unbinned Maximum Likelihood Fit

The unbinned maximum likelihood fit is performed using only events which are located inside a narrow window in mass. It is chosen from 5.31 GeV/c^2 to 5.60 GeV/c^2 in invariant mass.

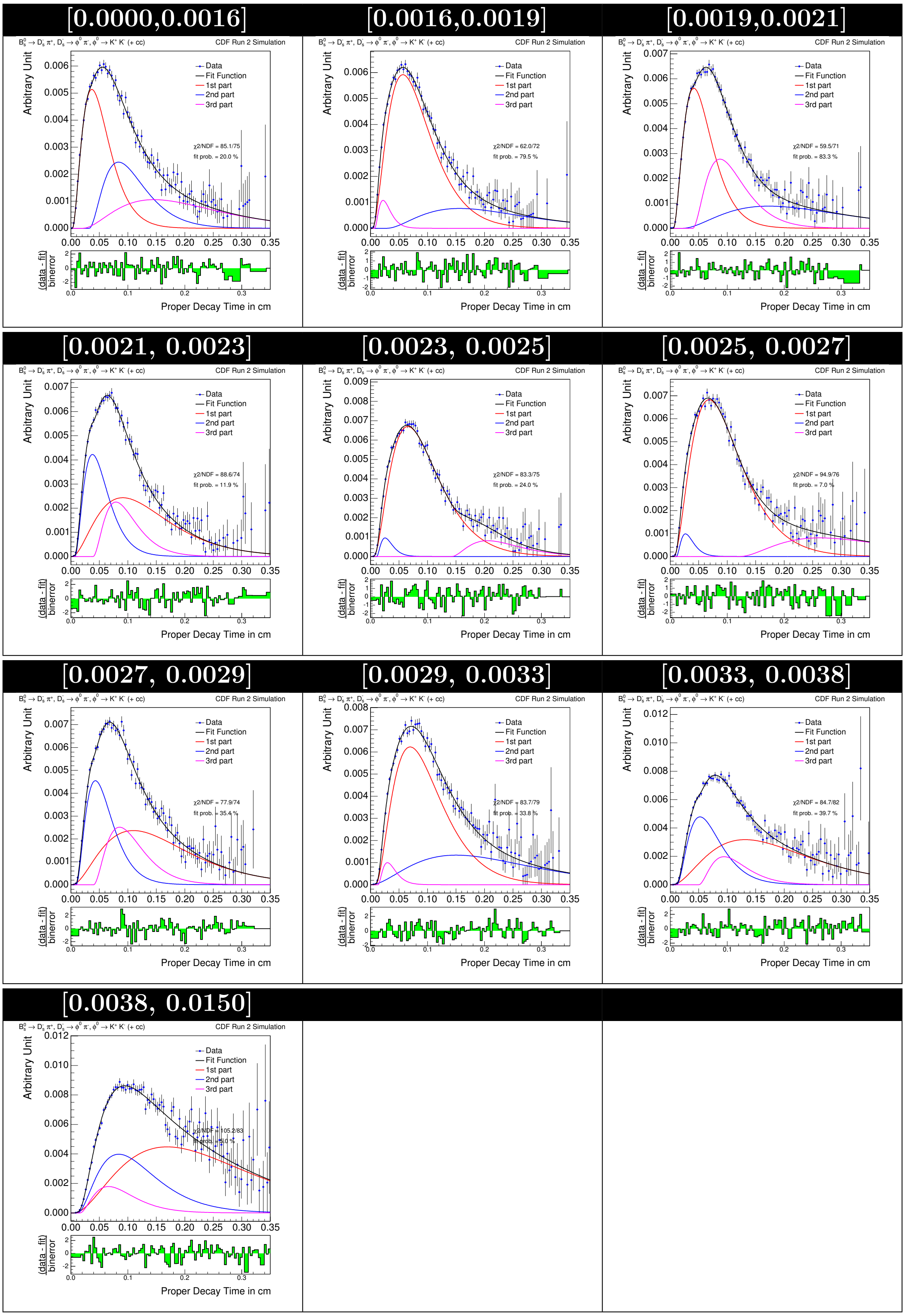


Figure 6.8: Efficiency functions of the B_s^0 analysis. Each function is responsible for a certain interval in proper decay time uncertainty. It is given in cm above each function.

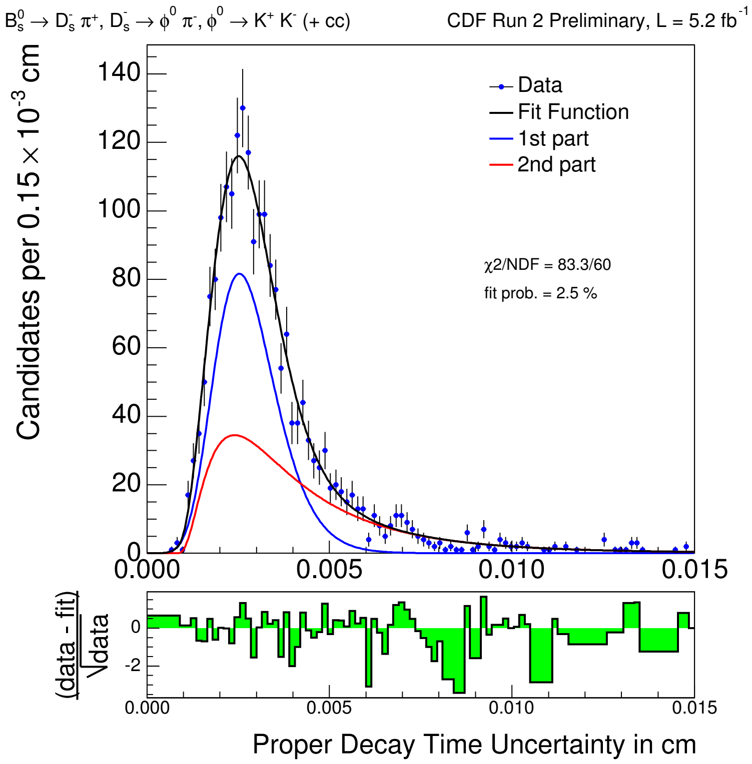


Figure 6.9: Distribution of the proper decay time resolution of the combinatorial background.

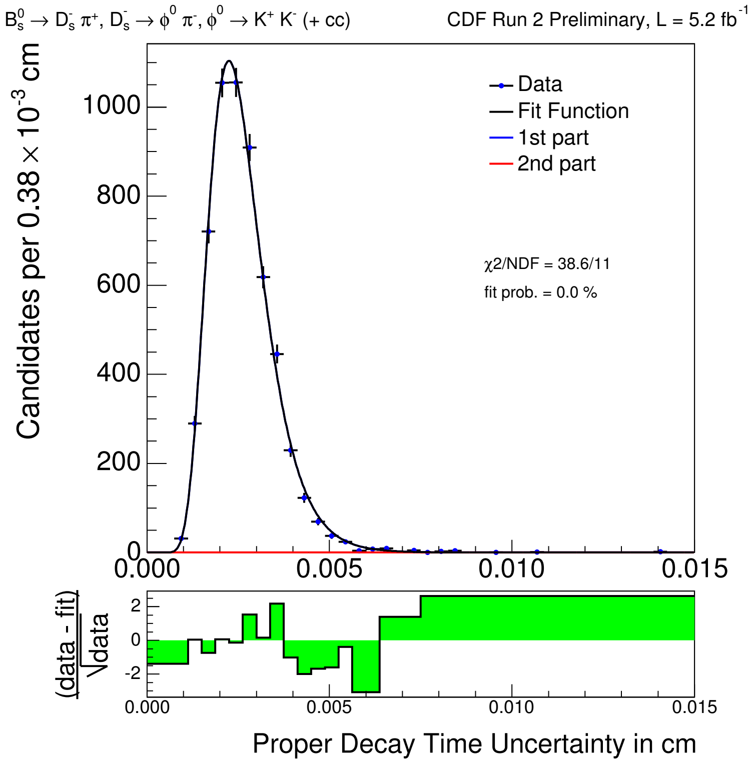


Figure 6.10: Distribution of the proper decay time resolution for the signal.

At first, a fit in mass, proper decay time and proper decay time resolution is performed. In other words the tagging decisions are not considered here. The subfunctions which contain ξ_P, ξ_D or D as variables (e.g. $P_{Sig}(\xi_P, D|\xi_D)$) are also neglected. In this case equation 5.29 simplifies to

$$P_{Sig}(ct|\sigma_{ct}) = N \cdot \left\{ \frac{1}{\tau} e^{-\tilde{t}/\tau} \otimes \mathcal{G}(c\tilde{t} - ct|\sigma_{ct}) \right\} \cdot \epsilon(ct|\sigma_{ct}). \tag{6.16}$$

Parameters which are left free are the background fraction f_{Bg}, width and position of the dominant signal Gaussian and the mean lifetime. The physics fraction $A_{Cb} = 0.055$, which was introduced in section 4.3, enters with Gaussian constraints. Its uncertainty is set to one third of the respective value to 0.018.

Projections into invariant mass, proper decay time and proper decay time uncertainty can be seen in figures 6.11, 6.12 and 6.13. The Cabibbo fraction is acquired as

$$A_{Cb} = 0.069 \pm 0.013. \tag{6.17}$$

The estimation on the mean lifetime of the B_s^0 amounts

$$c\tau = (455.9 \pm 8.4)\ \mu\text{m}. \tag{6.18}$$

A previous CDF measurement using 1.3 fb^{-1} of data taken in the same decay channel [66] obtained for fully reconstructed events

$$c\tau(B_s^0 \to D_s^-(\phi^0(K^+K^-)\pi^-)\pi^+) = (436.5 \pm 20.0\ (\text{stat.}) \pm 8.2\ (\text{syst.}))\ \mu\text{m}. \tag{6.19}$$

Both measured lifetimes are consistent with each other within one standard deviation.

6.7 Combined Results

Similar fits are done for the following B_s^0 decay channels:

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

Projections into mass, proper decay time and proper decay time resolution can be found in appendix A. The results for the mean lifetime can be found (together with the B_s^0 decay channel described in the previous sections) in table 6.2.

The individual results on the mean lifetime are inconsistent with each other. The reason for this is that a reliable measurement of this physics quantity relies on a precisely known efficiency function. As described above, this function is derived using simulated data. Thus it requires the knowledge of the interaction, the detector acceptance and selection. Especially the utilised two-track trigger introduces a considerable bias. Small variations of the input have an effect on the efficiency function and in this way on the measured lifetime.

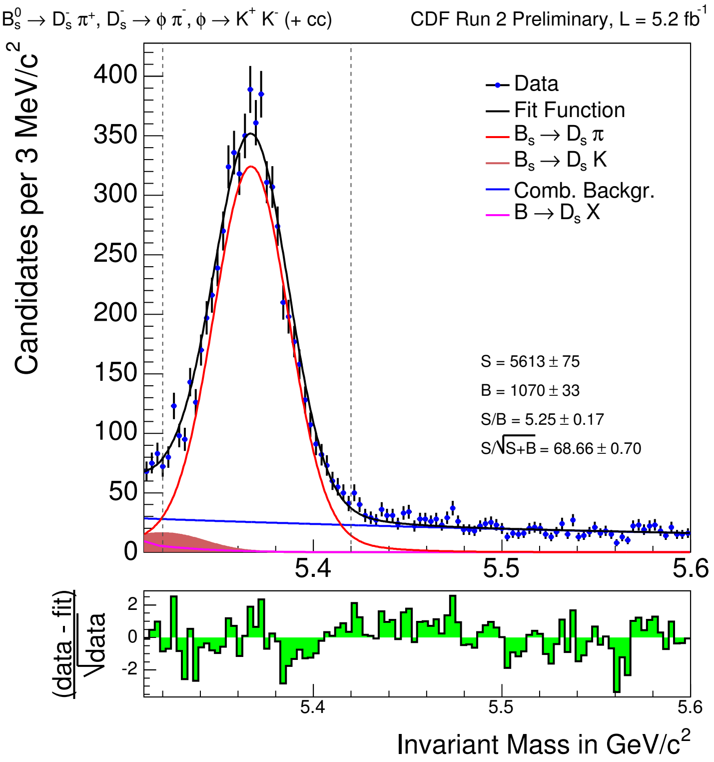


Figure 6.11: Projection of the result parameters of the unbinned maximum likelihood fit into invariant mass. The yields S and B are given for the signal range [5.32, 5.42] GeV/c^2 which is limited by dashed vertical lines inside the plot.

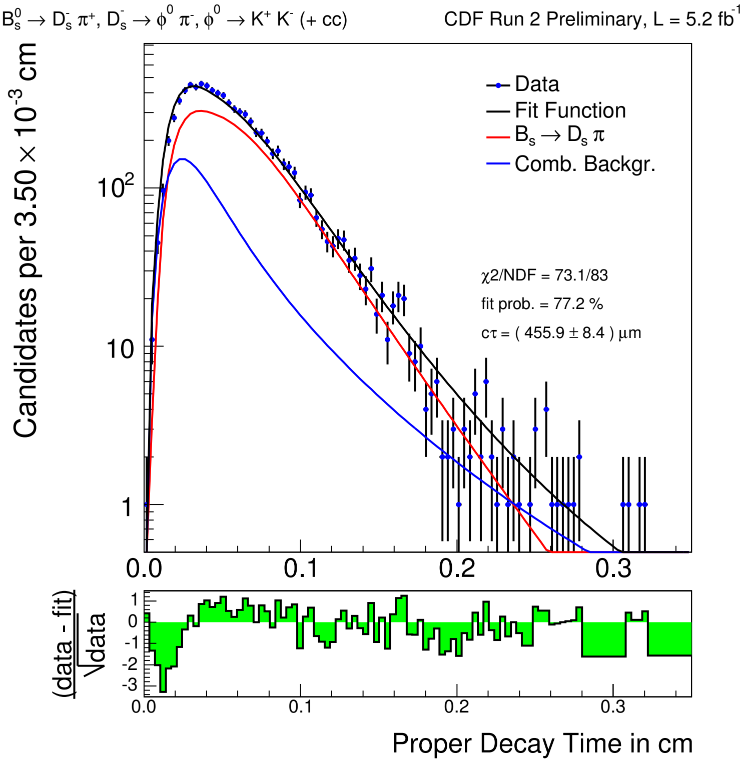


Figure 6.12: Projection of the result parameters of the unbinned maximum likelihood fit into proper decay time.

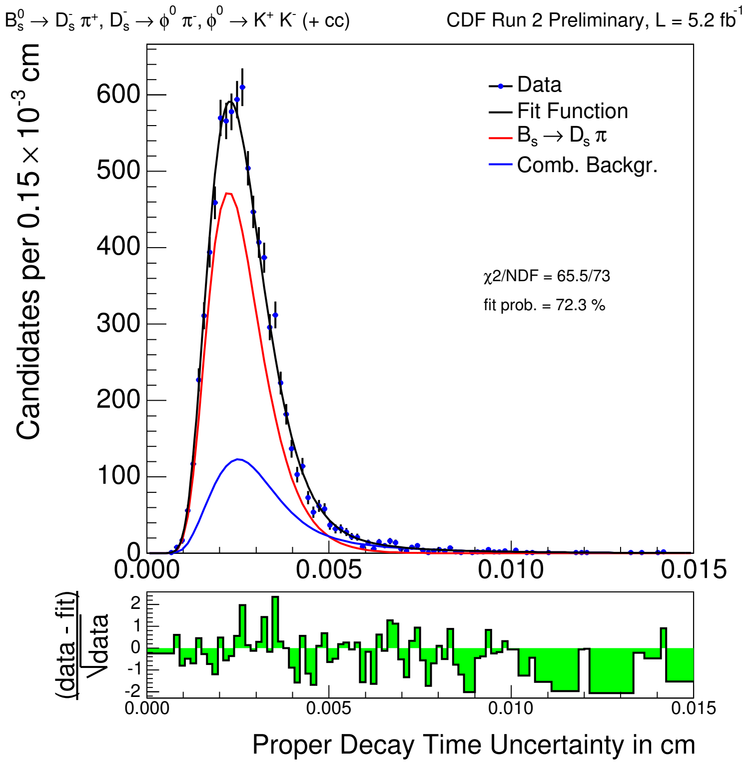


Figure 6.13: Projection of the result parameters of the unbinned maximum likelihood fit into proper decay time uncertainty.

Decay Channel	$c\tau$ [μm]
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-$	455.9 ± 8.4
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-$	437.3 ± 11.0
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+ \pi^- \pi^-$	421.7 ± 13.4
$B_s^0 \to D_s^- \pi^- \pi^+ \pi^+,\ D_s^- \to \phi^0 \pi^-$	489.7 ± 15.5

Table 6.2: Results for the mean lifetime for all four B_s^0 decay channels. The data amount corresponding to an integrated luminosity of 5.2 fb^{-1}.

Decay Channel	$\boldsymbol{S}$	$\boldsymbol{B}$	$\boldsymbol{S/B}$	$\boldsymbol{S/\sqrt{S+B}}$
$B_s^0 \to D_s^-\pi^+,\ D_s^- \to \phi\pi^-$	5613 ± 75	1070 ± 33	5.25 ± 0.17	68.66 ± 0.70
$B_s^0 \to D_s^-\pi^+,\ D_s^- \to K^*K^-$	2761 ± 53	1619 ± 40	1.71 ± 0.05	41.72 ± 0.74
$B_s^0 \to D_s^-\pi^+,\ D_s^- \to \pi^+\pi^-\pi^-$	2652 ± 52	3533 ± 59	0.75 ± 0.02	33.72 ± 0.68
$B_s^0 \to D_s^-\pi^-\pi^+\pi^+,\ D_s^- \to \phi\pi^-$	1852 ± 43	695 ± 26	2.66 ± 0.12	36.69 ± 0.73
Sum	$\mathbf{12877 \pm 113}$			

Table 6.3: Estimated number of signal events (S), background events (B), ratio of signal to background (S/B) and significance ($S/\sqrt{S+B}$) for all four B_s^0 decay channels. A data amount corresponding to an integrated luminosity of 5.2 fb^{-1} was hereby used. The evaluation was done inside the signal range, chosen from 5.32 GeV/c^2 to 5.42 GeV/c^2. For signal and background, the square root of the value is used as uncertainty. All other uncertainties are derived by Gaussian error propagation neglecting correlations.

Therefore it is no aim of this thesis to provide a full measurement of the mean lifetime. Instead it is shown in chapter 7 that a wrong lifetime measurement affects neither mixing frequency nor amplitude.

The different signal yields are given, together with other quantities of interest, in table 6.3. The four measurements in the B_s^0 decay channels are combined as described in section 5.2. The mean lifetime is determined in a first step by a simultaneous fit in mass, proper decay time and proper decay time resolution. The result amounts

$$c\tau = (451.2 \pm 5.5)\ \mu\text{m}. \tag{6.20}$$

This value is assigned to all decay channels before mixing frequency and amplitude are determined.

7 Mixing Amplitude

Up to now the same side kaon tagger could only be used on a data amount corresponding to the first 1.35 fb^{-1} of integrated luminosity. In the following chapter the range of validity of this powerful tagger is extended to 5.2 fb^{-1} [67]. This is done by measuring the mixing amplitude. The analysis flow continues hereby where it ended in the last chapter.

7.1 Tagging Decision and Dilution

The full probability density function (equation 5.73) including $P_{Sig}(\xi_P, D|\xi_D)$ and $P_{Comb}(\xi_P, D|\xi_D)$, the terms for the tagging decision and dilution, is now used. They are obtained in a similar way as the proper decay time resolution: the distribution for the signal is acquired by a sideband subtraction while the distribution for the combinatorial background is directly taken from the upper side band. The only difference is that instead of using an analytic parametrisation, the histogram is used directly in the unbinned maximum likelihood fit. In order to maintain a decent amount of statistics within each bin, variable bin widths are used.

7.2 Measurement

At first, mixing frequency and mixing amplitude are determined for each decay channel separately. The individual results can be found in table 7.1. An example projection in tagging decision times dilution for the decay channel $B_s^0 \rightarrow D_s^- \pi^+$, $D_s^- \rightarrow \phi^0 \pi^-$ can be found figure 7.1. The purity, defined as the number of signal events divided by the number of total events in each bin, is also shown.

Afterwards, a simultaneous fit for mixing frequency and amplitude is performed. The following results are acquired:

$$\mathcal{A} = 0.94 \pm 0.15. \tag{7.1}$$

$$\Delta m_s = (17.79 \pm 0.07)\ \text{ps}^{-1}. \tag{7.2}$$

Decay Channel	$\boldsymbol{\Delta m_s}$ **[$\boldsymbol{ps^{-1}}$]**	$\boldsymbol{\mathcal{A}}$ **[1]**
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-$	17.85 ± 0.09	1.13 ± 0.24
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-$	17.63 ± 0.19	0.72 ± 0.33
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+ \pi^- \pi^-$	18.07 ± 0.23	0.92 ± 0.40
$B_s^0 \to D_s^- \pi^- \pi^+ \pi^+,\ D_s^- \to \phi^0 \pi^-$	17.64 ± 0.14	1.02 ± 0.29

Table 7.1: Results for mixing frequency and amplitude for all four B_s^0 decay channels. The same side kaon tagger was used on a data amount corresponding to an integrated luminosity of 5.2 fb^{-1}.

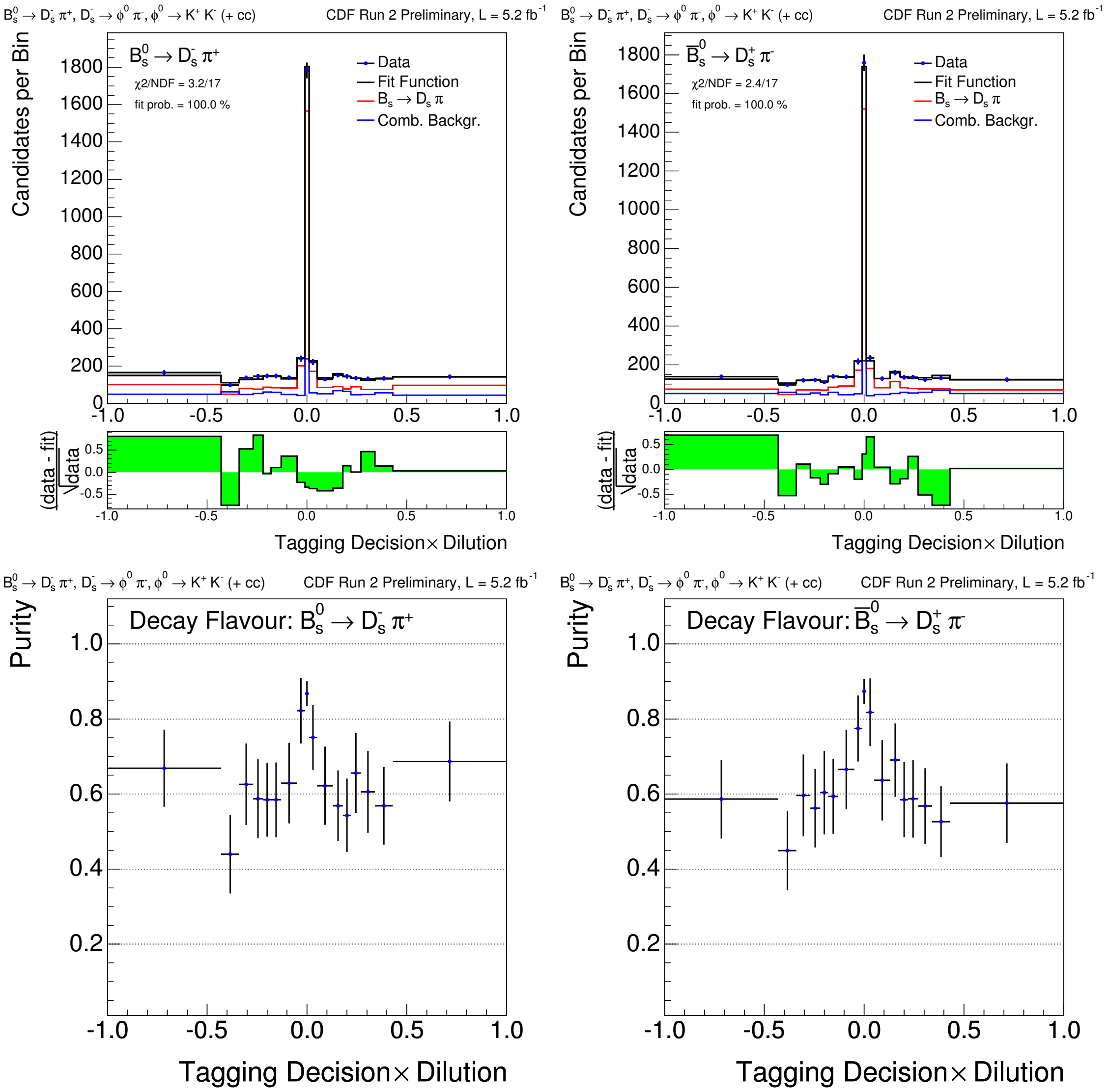


Figure 7.1: The output of the same side kaon tagger is illustrated using four plots. The histograms in the left column show B_s^0 candidates, $\bar{B}_s^0$ candidates are shown in the right column. Each histogram carries the tagging decision multiplied by the dilution on the x-axis. The plots in the upper row show the absolute frequency of the values using variable bin sizes. The histograms in the lower row show the signal purity in each such bin.

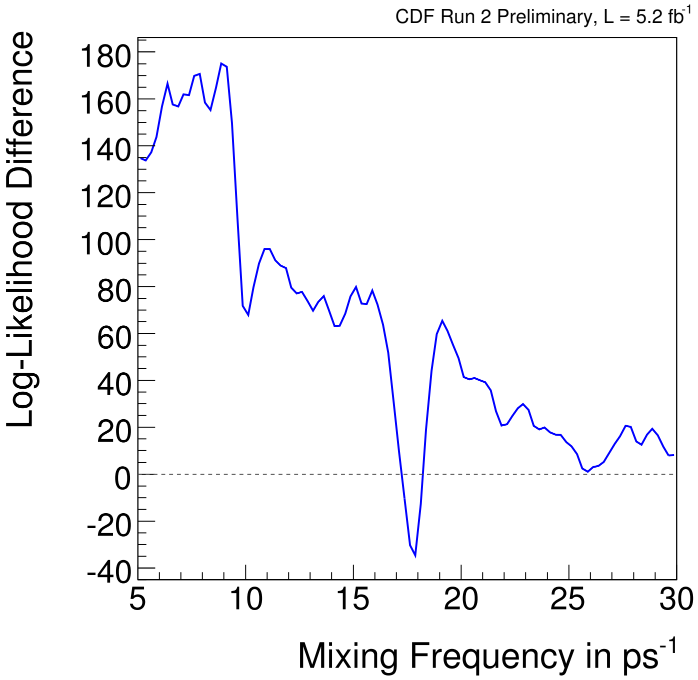


Figure 7.2: The difference of the logarithmic likelihood between the two assumptions $\mathcal{A} = 1$ and $\mathcal{A} = 0$ is drawn as a function of mixing frequency. All four decay channels were combined and the same side kaon tagger was applied.

The obtained mixing frequency is consistent within one standard deviation with the previous CDF measurement. In the context of the current chapter, this is only a cross-check. The result on the amplitude is consistent within one standard deviation with the optimal value of 1.0. All uncertainties are statistical only.

The significance of the measurement and the position of the mixing frequency is visualised by plotting the following quantity as a function of Δm_s:

$$-2 \cdot \left[\ln \Big(\mathcal{L}(\mathcal{A} = 1) \Big) - \ln \Big(\mathcal{L}(\mathcal{A} = 0) \Big) \right] \tag{7.3}$$

This term consists of the difference between two negative logarithmic likelihood expressions. The amplitude is set to 1 for the first one. For the second expression, the amplitude is set to zero. In this sense it can be seen as a quantitative comparison of the two hypotheses of mixing ($\mathcal{A} = 1$) and no mixing ($\mathcal{A} = 0$). The result plot created by combining all decay channels can be found in figure 7.2.

7.3 Systematic Uncertainties on the Amplitude

The following sources of systematic uncertainties on the mixing amplitude are considered:

- As described in section 4.6, it is known from several studies that the measured proper decay time resolution is underestimated at the CDF-II experiment. Therefore, a resolution scaling technique is employed for the study at hand. In order to evaluate the systematic effect of this scaling technique on the amplitude, it is replaced by a constant function with a mean value of 1.29. As a consequence, the measured amplitude is reduced by 0.11. This value is added to the list of systematic uncertainties.

- The proper decay time resolution is assumed as Gaussian. Section 4.6 showed that the sum of two Gaussians with different standard deviations gives a better description. Therefore, 1000 simulated experiments are used to examine the extend of this simplification on the measured amplitude. The simulated data hereby considers the refined model, the fit function does not. The obtained amplitude shows a deviation of 0.06 which is taken as systematic uncertainty.
- A given B_s^0 may not only go into $D_s^- K^+$, but also into $D_s^+ K^-$. While both processes are Cabibbo suppressed, the former is expected to occur more often [11]. Consequences of the presence of the latter, charge conjugated final state are evaluated by completely removing tagging information for the Cabibbo reflection. A change of 0.03 is observed for the mixing amplitude and used as systematic uncertainty.
- The actual fraction of Cabibbo suppressed decays enters with Gaussian constraints during the unbinned maximum likelihood fit. As a test, it is left unconstrained in the decay channel $B_s^0 \to D_s^- \pi^+$, $D_s^- \to \phi\pi^-$. In this way a value of 0.086 ± 0.020 is obtained instead of 0.069 ± 0.013 for the Cabibbo fraction. This deliberate increase had no effect on neither mixing frequency nor amplitude.
- The fit function, given in equation 5.62, does not take into account effects of the decay width difference $\Delta\Gamma$ in the B_s^0 system. This neglect is studied using 1000 simulated experiments which are generated with an assumed value of $\Delta\Gamma/\Gamma = 0.12$. The procedure follows here Ref. [68]: using the mean lifetime $\tau = 1/\Gamma$ and the assumed decay width $\Delta\Gamma$, the short and long lifetime can be calculated:

  $$\Gamma_L = \Gamma - \frac{1}{2}\Delta\Gamma, \qquad \tau_L = 1/\Gamma_L, \tag{7.4}$$

  $$\Gamma_S = \Gamma + \frac{1}{2}\Delta\Gamma, \qquad \tau_S = 1/\Gamma_S. \tag{7.5}$$

  The fraction short lived B_s^0 mesons with respect to the total number is

  $$f_S = \frac{\tau_S}{\tau_L + \tau_S}. \tag{7.6}$$

  The decision if a B_s^0 meson is mixed or not mixed at decay time can be derived from the probability

  $$p_{unmix} = \frac{\frac{1}{2}(\cosh\frac{\Delta\Gamma t}{2} + \cos(\Delta m_s t))}{\cosh\frac{\Delta\Gamma t}{2}}. \tag{7.7}$$

  The fit function itself is retained as it is. The absolute deviation of the pull is smaller than 0.01.
- As a test, the measurement of mixing frequency and amplitude is repeated with different values for the mean lifetime. The reason for doing this is to evaluate consequences of a wrong lifetime measurement. However, variations of the mean lifetime between 420 μm and 490 μm have no effect on neither mixing frequency nor amplitude.

Modification	**Systematic Uncertainty**
Proper decay time resolution scaling	0.11
Resolution model	0.06
Cabibbo reflection	0.03
Cabibbo fraction	negligible
Mass window	negligible
Selection of upper side band	negligible
Λ_b template	negligible
$\Delta\Gamma/\Gamma$	negligible
Mean Lifetime	negligible
Trigger Composition	negligible
Signal Mass Model	negligible
Total	**0.13**

Table 7.2: Systematic uncertainties on the mixing amplitude. The total number is the root of the quadrature sum.

- As mentioned above, a Λ_b reflection is present in some decay channels. In the tagging quantities, the same template is used for it as for the signal. The actual size and location of that reflection makes it hard to check if this modelling is appropriate. However, its effect on the actual result can be determined by replacing it by the distribution derived for the combinatorial background. In doing so the same result is obtained as above.
- Variations of the mass window used in the unbinned fit or the choice of the upper side band did not have any effects on the mixing amplitude as well.
- Different trigger paths are applied for the data taking. The information which one was responsible for the taking of a given event is available both on the measured and the simulated data. However, a comparison between both shows that the composition is different. Because of this a reweighting is performed in order to adjust the simulated events to the measured data. A measurement without that treatment revealed the same results for mixing frequency and amplitude.
- The model for the signal in invariant mass consists of three Gaussians. A simplification of it to one Gaussian had no effect on the mixing amplitude.

The different contributions are summarised in table 7.2. In total, a systematic uncertainty of 0.13 is achieved.

7.4 Results

Using the systematic uncertainty determined in the previous section, the result on the mixing amplitude of the same side kaon tagger on a data amount corresponding to an integrated luminosity of 5.2 fb^{-1} reads

$$\mathcal{A} = 0.94 \pm 0.15\ \text{(stat.)} \pm 0.13\ \text{(syst.)}. \tag{7.8}$$

With this value it is now possible to calculate a final value for the tagging power. In order to do so, both the systematic and the statistical uncertainty are merged. In

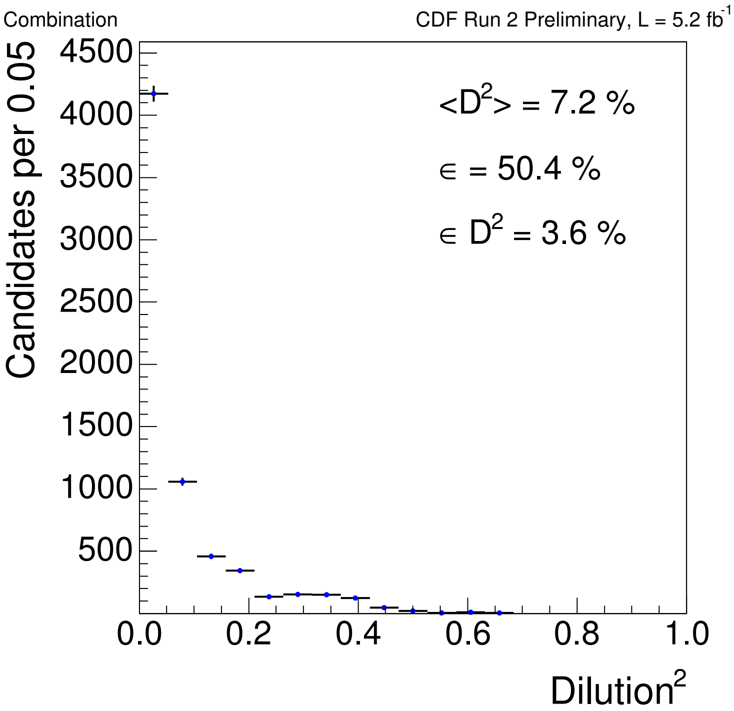


Figure 7.3: Distribution of the dilution squared for all four decay channels combined. Mean squared dilution, efficiency, tagging power and significance are plotted inside the histogram.

this way, a value of $\mathcal{A} = 0.94 \pm 0.20$ is obtained. The mixing amplitude is a scale factor for the dilution. Therefore, it enters the equation of the tagging power as square. The raw value of εD^2 is determined by performing a sideband subtraction channel-by-channel. Afterwards, all four distributions are added up and a mean value of 3.6 % is obtained. The corresponding histogram can be seen in figure 7.3. Using this result the final tagging power of the same side kaon tagger is acquired as

$$\mathcal{T} = \varepsilon \mathcal{A}^2 D^2 \approx (3.2 \pm 1.4)\ \%. \tag{7.9}$$

8 Mixing Phase

Every analysis which employs the same side kaon tagger benefits from the calibration described in the previous chapter. However, the primary intention behind this measurement was its application in the determination of β_s.

The measurement of this $\mathcal{CP}$ violating phase is an ongoing effort of both Tevatron experiments since years and many results were already published. Therefore, the current chapter starts with a chronology of previous measurements. One common feature of the results was that no point estimation on the $\mathcal{CP}$ violating phase could be provided. The reason for this is a correlation between β_s and the decay width difference $\Delta\Gamma$. Instead, confidence regions are given as result. They can be found in figure 8.1.

8.1 Analysis Outline

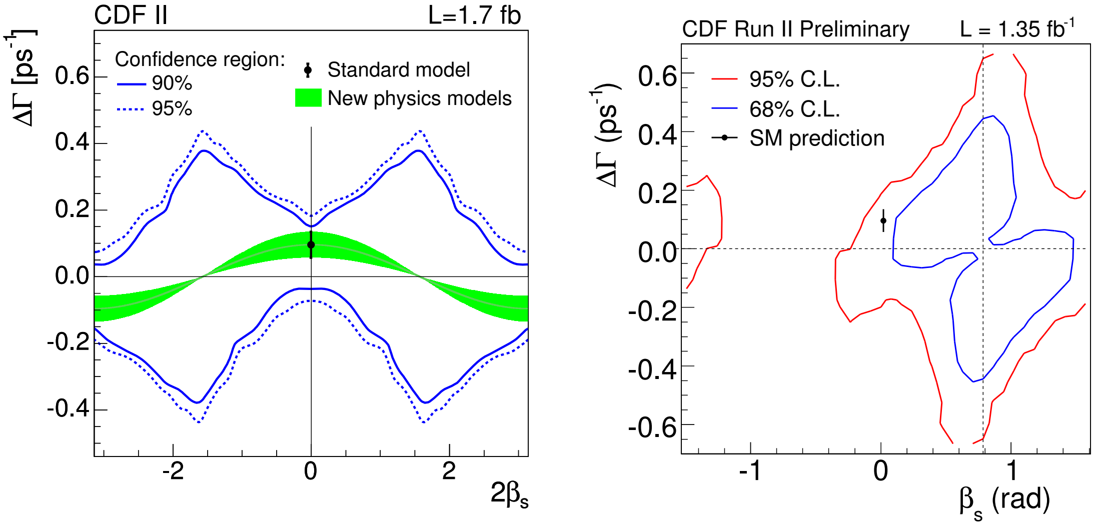
As described in chapter 2.7, the decay channel

$$B_s^0 \to J/\psi\phi,\ J/\psi \to \mu^+\mu^-,\ \phi \to K^+K^- \tag{8.1}$$

is primarily used for this measurement. Here the final state can be reached with and without mixing. The B_s^0 itself is a pseudo-scalar meson while its decay products J/ψ and ϕ are both vector mesons. The final state is a mixture of $\mathcal{CP}$ eigenstates which can be distinguished using the angular quantum number: the $L = 0, 2$ states are $\mathcal{CP}$-even and the $L = 1$ state is $\mathcal{CP}$-odd.

In order to extract the $\mathcal{CP}$ violating phase $\beta_s^{J/\psi\phi}$ and the decay width difference $\Delta\Gamma$, an unbinned maximum likelihood fit is employed. It uses mass, transversity angles, proper decay time and flavour tagging information of each event. Next to the parameters β_s and $\Delta\Gamma$, the probability density function contains the magnitudes of the polarisation amplitudes $|A_0|^2$, $|A_\perp|^2$ and $|A_\parallel|^2$, the mixing frequency Δm_s, the mean width $\Gamma = (\Gamma_L + \Gamma_H)/2$ and the strong phases $\delta_\parallel = \arg(A_\parallel^* A_0)$ and $\delta_\perp = \arg(A_\perp^* A_0)$. The transversity angles $\vec{\rho} = (\cos\theta_T, \phi_T, \cos\psi)$ [72] are used to separate the $\mathcal{CP}$-even and $\mathcal{CP}$-odd components of the $J/\psi\phi$ final state.

(a) September 2007, [21] **(b) December 2007, [69]**



(c) July 2008, [70] **(d) July 2009, [71]**

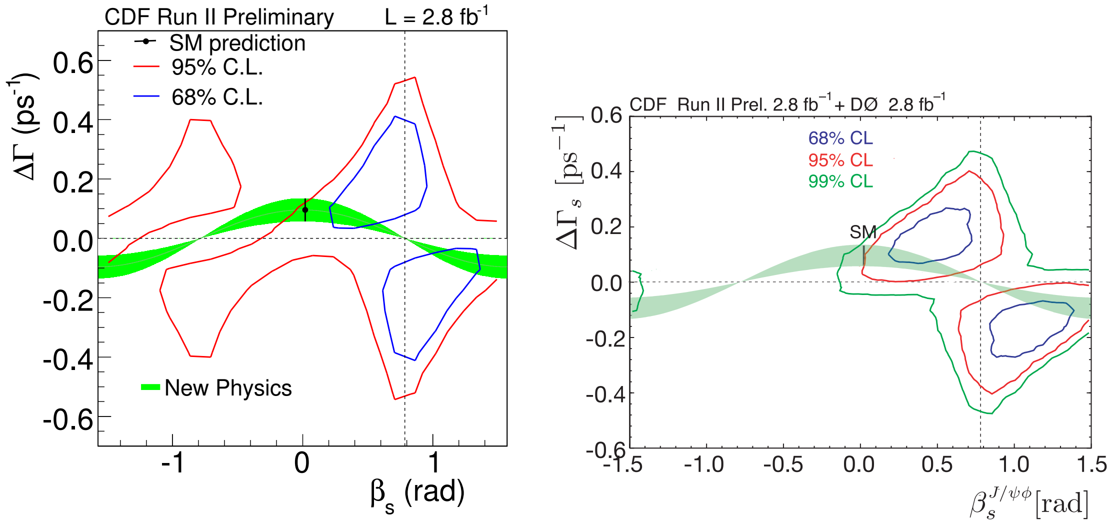


Figure 8.1: Chronology of the $\Delta\Gamma$-β_s-confidence regions published by the CDF collaboration. Remarks: **(a)** The plot was created without using tagging information, **(c+d)** The opposite side tagger was applied on the full dataset. The same side kaon tagger was applied on a dataset corresponding to the first 1.35 fb^{-1} of integrated luminosity, **(d)** the result was obtained by combining CDF and DØ results.

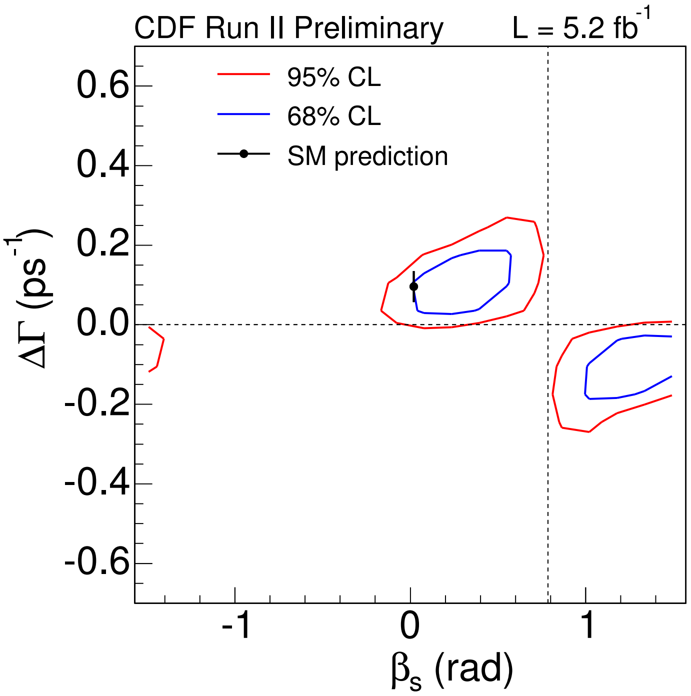


Figure 8.2: Up-to-date confidence region in the $\Delta\Gamma$-β_s-plane based on a dataset corresponding to an integrated luminosity of 5.2 fb^{-1}.

The up-to-date result is still provided as confidence region as described above. It can be seen in figure 8.2. It is determined using a frequentist method and features two distinct, symmetric minima. They are caused by the exact symmetry of the strong phases in the three decay amplitudes of the $L = 0, 1, 2$ final states. The standard model expectation value is marked as black point.

The tagging information of both same side kaon tagger and opposite side tagger is combined for this measurement. Both are now applied on the full data sample.

The first is calibrated by using the mixing amplitude as scale factor as described in the previous chapter. The latter is calibrated using measured $B^{\pm} \to J/\psi K^{\pm}$ data. Charge conservation prevents this particle from changing its flavour. The principle behind this calibration is therefore based on a comparison between the predicted production flavour and the measured one, which can flawlessly be determined from the final state. Two scale factors are employed there, one for the B^+ and one for the B^-, in order to allow for any asymmetry in the tagging algorithms. Both are used for the whole dataset. The measured tagging power for the opposite side tagger amounts 1.8 %. In both cases the event-by-event dilution values are multiplied by the respective scale factor.

The agreement with the standard model is quantified by the p-value, which can be seen as the probability for observing a result which is at least as “extreme” as the one obtained. In detail, it is defined as the integral over the distribution obtained in pseudo-experiments starting from $R_{data}(\Delta\Gamma, \beta_s^{J/\psi\phi})$ to infinity. $R_{data}(\Delta\Gamma, \beta_s^{J/\psi\phi})$ is the difference between the logarithmic likelihood where $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ is fixed in the fit to standard model expectation values and the logarithmic likelihood where both parameters float freely. Small p-values ~ 10 % were observed in the previous analyses, which roused a lot of interest in the community. The up-to-date result features a p-value of 44 % which speaks for the validity of the standard model.

9 Mixing Frequency

The statistical uncertainty of the mixing frequency which served as control variable during the calibration of the same side kaon tagger was already improved by 30 % in comparison with the published result from 2006.

This motivates a new measurement of this important physics quantity. However the reason for not just taking over the value derived during the calibration is that both significance and statistical uncertainty benefit from higher tagging power. The analysis flow therefore continues at the point, where chapter 6 finished.

9.1 Analysis Outline

Similar to the measurement described in the previous chapter, the tagging information of same side kaon tagger and opposite side tagger are combined. It is done assuming that tags are uncorrelated. If only one tagger provides information about the production flavour for a given event, the information is assigned as it is. If estimations (ξ_1, D_1) and (ξ_2, D_2) are given by both, the information is combined. In case the tagging decisions ξ_1 and ξ_2 agree, the dilution is obtained by

$$D = \frac{D_1 + D_2}{1 + D_1 \cdot D_2}. \tag{9.1}$$

If ξ_1 and ξ_2 disagree, the decision of the tagger with the larger dilution is assumed. Furthermore

$$D = \frac{D_1 - D_2}{1 - D_1 \cdot D_2} \tag{9.2}$$

is taken as dilution in this case.

Apart from different tagging information, the measurement of the mixing frequency works exactly like the calibration described in chapter 7. Mixing frequency and amplitude are determined at the same time, at first channel by channel. Afterwards they are combined. The individual results can be found in table 9.1.

Decay Channel	$\boldsymbol{\Delta m_s}$ $[\boldsymbol{ps^{-1}}]$	$\boldsymbol{\mathcal{A}}$ [1]
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-$	17.81 ± 0.08	1.10 ± 0.20
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-$	17.66 ± 0.21	0.63 ± 0.28
$B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+ \pi^- \pi^-$	17.94 ± 0.13	1.08 ± 0.40
$B_s^0 \to D_s^- \pi^- \pi^+ \pi^+,\ D_s^- \to \phi^0 \pi^-$	17.64 ± 0.11	1.22 ± 0.26

Table 9.1: Results for mixing frequency and amplitude for all four B_s^0 decay channels. The same side kaon tagger was combined with the opposite side tagger on a data amount corresponding to an integrated luminosity of 5.2 fb^{-1}.

A combined fit to all four decay channels yields

$$\Delta m_s = (17.78 \pm 0.06)\ \text{ps}^{-1}, \tag{9.3}$$

$$\mathcal{A} = 1.00 \pm 0.13. \tag{9.4}$$

The result on the frequency is consistent with the published CDF measurement. The statistical uncertainty is improved by 40 %. The mixing amplitude, which is a control quantity in the context of the current chapter, is consistent with the optimal value of one. Using the result on the amplitude, a tagging power of 5.0 % is obtained with an efficiency of 96 % for this tagger combination.

9.2 Amplitude Scan

One nice way to observe mixing is referred to as amplitude scan. It resembles a Fourier transform and is produced in the following way: frequencies are chosen in equidistant steps within a certain interval. For each such frequency an unbinned maximum likelihood fit is performed with mixing frequency fixed to the corresponding position and mixing amplitude as the only free parameter.

In this way a set of value pairs are acquired. They are plotted with frequency on the x-axis and amplitude on the y-axis. At the frequency measured in the previous section the amplitude assumes a value consistent with one.

Figure 9.1 shows the amplitude scan for all four B_s^0 decay channels combined.

The half width of the 90 % confidence interval of each measured value is shown as dashed line. The value at which this line meets the $\mathcal{A} = 1$ is referred to as sensitivity. Starting from this point it is not possible anymore to distinguish between zero and one with the given confidence. For the measurement presented here, this value is located at 34 ps^{-1} which lies beyond the upper boundary of the histogram.

9.3 Final Result

The systematic studies, shown in table 7.2, which were evaluated for the measurement of the mixing amplitude, are also checked for the measurement of the mixing frequency. However no significant deviation can be observed here. This agrees with previous measurements, where it was found that there is only one non-negligible

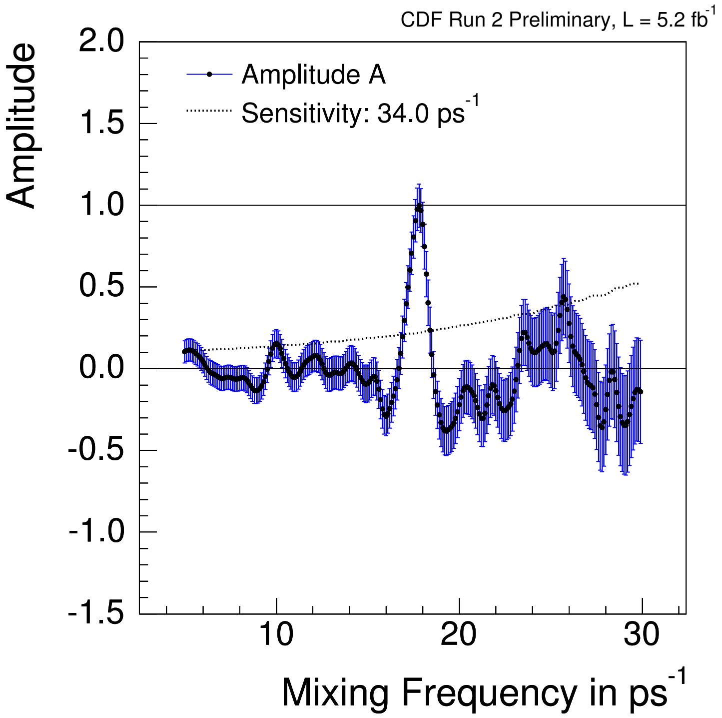


Figure 9.1: Amplitude scan in all decay channels using the same side kaon tagger and the opposite side tagger on a data amount corresponding to 5.2 fb^{-1}.

systematic uncertainty on Δm_s: it is the uncertainty on the absolute scale of the decay-time measurement ct which is caused by the level of uncertainty in the alignment of the silicon layers. A significant change is not expected for it. The previous mixing analyses assigned an uncertainty of 0.07 ps^{-1} to that systematic source. This value is taken over for this thesis. In this way, the final result is obtained as

$$\Delta m_s = (17.78 \pm 0.06 \text{ (stat.)} \pm 0.07 \text{ (syst.)}) \text{ ps}^{-1}. \tag{9.5}$$

The significance of this measurement is obtained by dividing the result on the amplitude (equation 9.4) by its statistical uncertainty. In this way a significance value of 7.7 standard deviations is acquired which establishes B_s mixing without doubt.

10 Conclusion

The thesis at hand makes two valuable contributions to the understanding of the B_s meson system.

The first one is indirect and consists of a calibration of the same side kaon tagger. It was performed using the decay channels

$$\begin{aligned}
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-,\\
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-,\ K^* \to K^+ \pi^-,\\
&B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+ \pi^- \pi^-,\\
&B_s^0 \to D_s^- \pi^+ \pi^+ \pi^-,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-,
\end{aligned}$$

on a data amount corresponding to the first 5.2 fb^{-1} of integrated luminosity. The observables of the B_s^0 were hereby described by a multi-dimensional model, which was derived using an extensive study of correlations. The calibration was performed by measuring the mixing amplitude, using an unbinned maximum likelihood fit, as

$$\mathcal{A} = 0.94 \pm 0.15\ (\text{stat.}) \pm 0.13\ (\text{syst.}). \tag{10.1}$$

The validity of this result is guaranteed by two main features. First of all, the final result is derived from measured data. This is important because the agreement between simulated events and real data is always limited. Second, the mixing frequency which was measured at the same time agrees with previous result. In addition, an extensive study of systematic uncertainties was performed.

While every analysis employing the same side kaon tagger benefits from the above result on $\mathcal{A}$ this thesis was developed in the context of the new measurement of the $\mathcal{CP}$-violating phase β_s. This physics quantity is expected to be small. Therefore a deviation from this expectation would be a clear sign of new physics. Because of a correlation between the decay width difference $\Delta\Gamma$ and β_s, a confidence region is provided as result here. Tagging information increases the sensitivity on these parameters and therefore helps constraining this confidence region. However it was a drawback of past measurements that the same side kaon tagger, the most powerful stand-alone tagger at the CDF experiment, could only be used on a data amount corresponding to the first 1.35 fb^{-1} of integrated luminosity. The determination of the mixing amplitude presented in this thesis allowed for the extension of that range up to 5.2 fb^{-1}. In this sense, it made an important contribution to the newest result.

The mixing frequency Δm_s itself was so far measured only once with a confidence level of five standard deviations. The availability of the formalism described above suggested an update on this frequency. This was the second aim of this thesis. The sensitivity of this physics parameter increases with tagging power. Therefore the same side kaon tagger was combined with the opposite side tagger. The new result is consistent with the previous one and was obtained as

$$\Delta m_s = (17.78 \pm 0.06 \text{ (stat.)} \pm 0.07 \text{ (syst.)}) \text{ ps}^{-1}. \tag{10.2}$$

With a statistical uncertainty, which was reduced by 40 %, this is the most precise measurement so far. At the same time, the result significance was increased from 5σ to 7.7σ.

Using this result it is possible to extract the ratio between the CKM matrix elements $|V_{td}|$ and $|V_{ts}|$. Presently, no improvement is gained by the above result on Δm_s because the result on this ratio is dominated by theoretical uncertainties stemming from lattice QCD. However, this theoretical field has made constant progress in the past years and expects further improvements in coming years. Thus, a lowering of the theoretical uncertainty is conceivable in the near future. The new measurement of Δm_s is therefore a safe investment and may result in a powerful constraint in the near future.

A Fit Projections

A.1 $B_s^0 \to D_s^- \pi^+,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-$

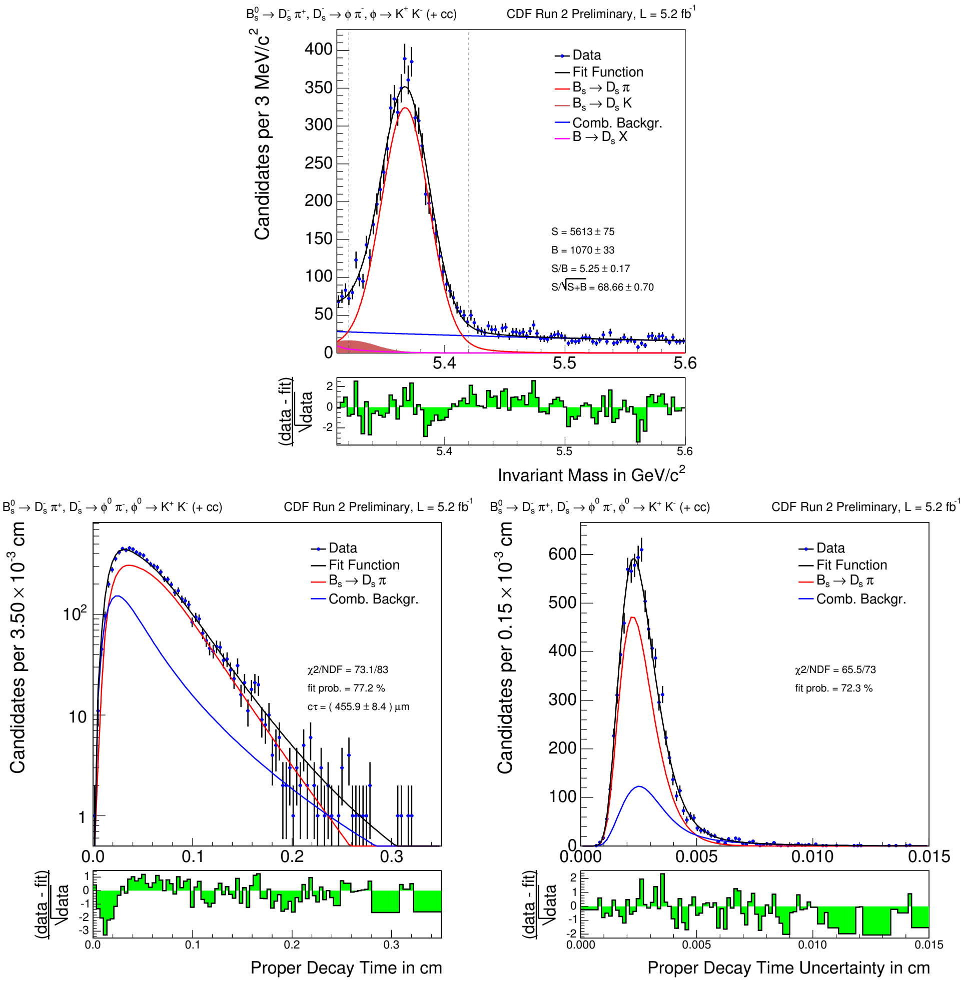


Figure A.1: Projections into invariant mass, proper decay timeand proper decay time resolution.

A.2 $B_s^0 \to D_s^- \pi^+,\ D_s^- \to K^* K^-,\ K^* \to K^+ \pi^-$

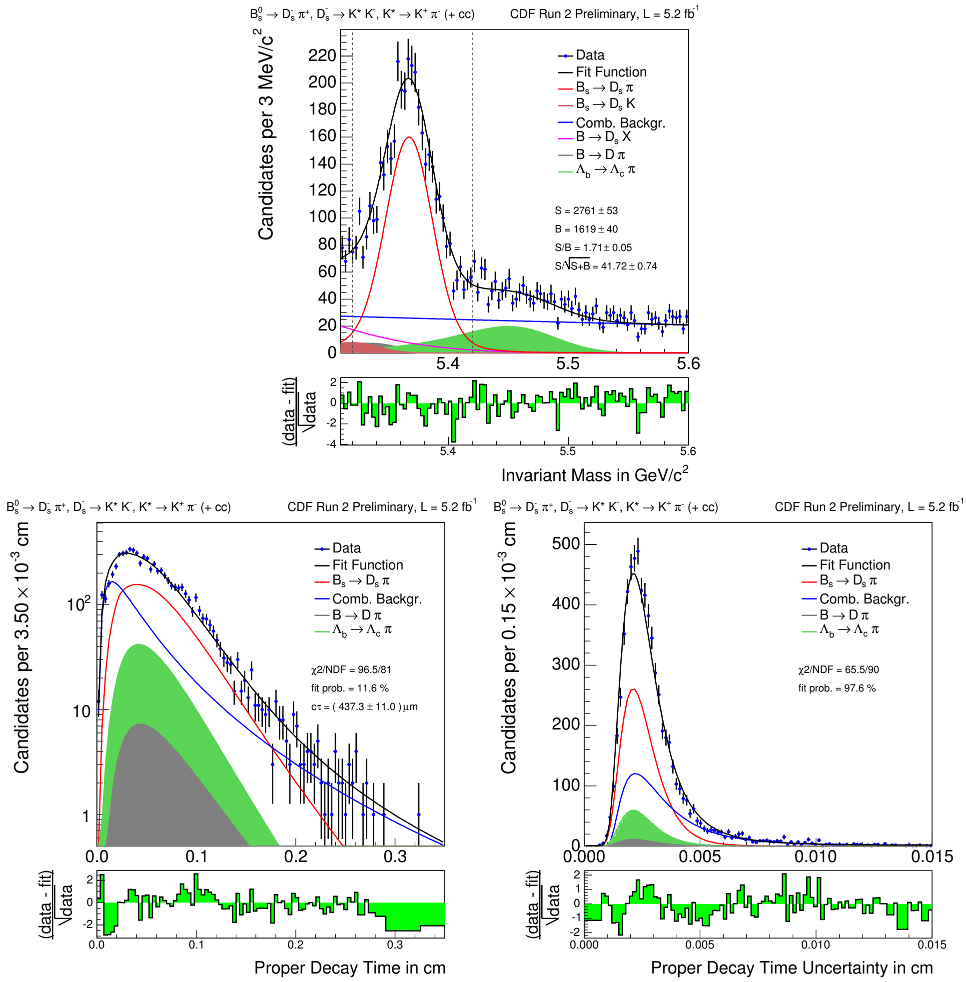


Figure A.2: Projections into invariant mass, proper decay timeand proper decay time resolution.

A.3 $B_s^0 \to D_s^- \pi^+,\ D_s^- \to \pi^+\pi^-\pi^-$

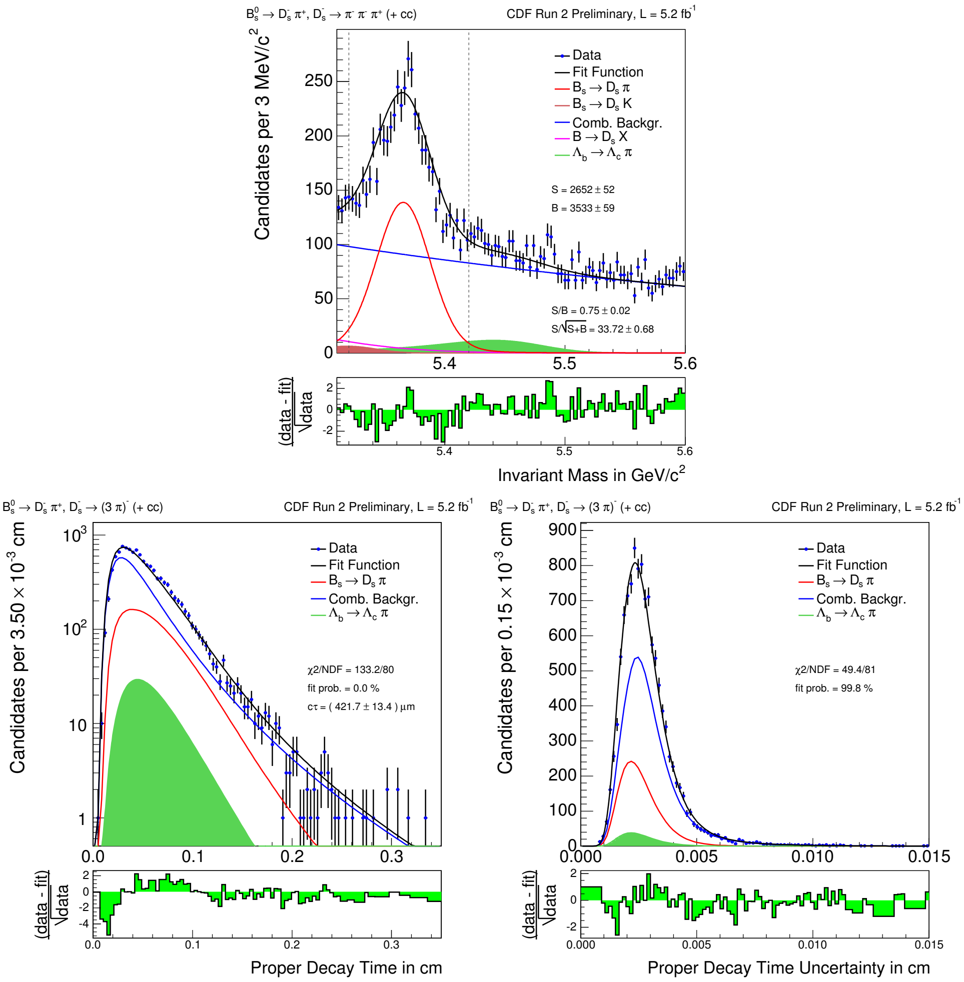


Figure A.3: Projections into invariant mass, proper decay timeand proper decay time resolution.

A.4 $B_s^0 \to D_s^- \pi^- \pi^+ \pi^+,\ D_s^- \to \phi^0 \pi^-,\ \phi^0 \to K^+ K^-$.

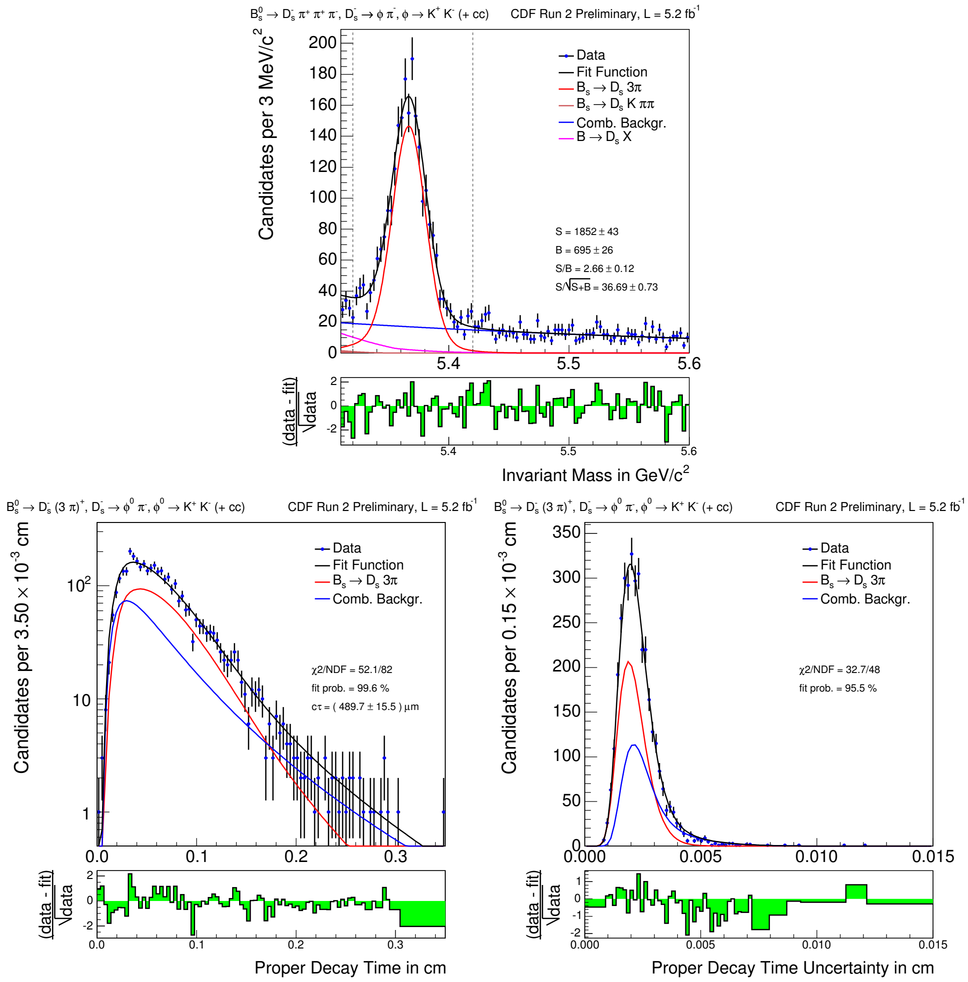


Figure A.4: Projections into invariant mass, proper decay timeand proper decay time resolution.

Bibliography

[1] C. Amsler et al. (Particle Data Group). Physics Letters B667. 1 (2008) and 2009 partial update for the 2010 edition, `http://pdg.lbl.gov/`.

[2] A. Lenz and U. Nierste. Theoretical update of B_s^0-$\bar{B}_s^0$ mixing. *JHEP*, Jun. 2007. arXiv:hep-ph/0612167v3.

[3] `http://www-cdf.fnal.gov/tiki/tiki-index.php?page=ProductionFarm.DataProcessing`.

[4] CKMfitter Group (Charles, J. et al.). CP violation and the CKM matrix: assessing the impact of the asymmetric B factories. *Eur. Phys. J.*, C41:1–131, 2005. hep-ph/0406184.

[5] A. Abulencia et al. Observation of B_s^0 anti-B_s^0 oscillations. *Phys. Rev. Lett.*, 97:242003, 2006.

[6] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli. The Standard Model of Particle Physics. *Rev. Mod. Phys.*, 71:96–111, 1999. arXiv:hep-ph/9812285v1.

[7] Donald H. Perkins. *Introduction to High Energy Physics*, chapter 1, page 11. Cambridge University Press, fourth edition, 2000.

[8] G. Lüders. *Dan. Mat. Fys. Medd.*, 28:5, 1954.

[9] Angelopoulos, A. et al. (CPLEAR Collaboration). Physics at CPLEAR. *Phys. Rep.*, 374:165–270, 2003.

[10] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. 1983. Phys. Rev. Lett. 51, 1945.

[11] K. Anikeev et al. B Physics at the Tevatron: Run II and Beyond. page 194, 2002. arXiv:hep-ph/0201071v2.

[12] C. Albajar et al. Search for B0 anti-B0 Oscillations at the CERN Proton - anti-Proton Collider. 2. *Phys. Lett.*, B186:247, 1987.

[13] H. Albrecht et al. Observation of B0 - anti-B0 Mixing. *Phys. Lett.*, B192:245, 1987.

[14] A Heister et al. Improved search for $B_s^0 - \overline{B}_s^0$ oscillations. *Eur. Phys. J. C*, 29:143–170. 40 p, Feb 2002.

[15] P. Abreu et al. Measurement of the B_s^0 lifetime and study of $B_s^0 - \bar{B}_s^0$ oscillations using $D_s \ell$ events. *Eur. Phys. J.*, C16:555, 2000.

[16] P. Abreu et al. Study of B0(S) - anti-B0(S) oscillations and B0(S) lifetimes using hadronic decays of B0(S) mesons. *Eur. Phys. J.*, C18:229–252, 2000.

[17] G. Abbiendi et al. A Study of B_s meson oscillation using D_s - lepton correlations. *Eur. Phys. J.*, C19:241–256, 2001.

[18] K. Abe et al. Time dependent B_s^0 - anti-B_s^0 mixing using inclusive and semileptonic B decays at SLD. 2000.

[19] T. Inami and C. S. Lim. Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes k(L) → mu anti-mu, K+ → pi+ Neutrino anti-neutrino and K0 <—> anti-K0. *Prog. Theor. Phys.*, 65:297, 1981.

[20] J. Morlock. Study of CP violation in $B_s \to J/\psi\phi$ decays at CDF. In *Proceedings of Science*, volume 171, 2009. PoS(EPS-HEP 2009)171.

[21] T. Aaltonen et al. Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays. *Phys. Rev. Lett.*, 100:121803, 2008.

[22] M. Milnik. *Measurement of the Lifetime Difference and CP-Violating Phase in $B_s \to J/\psi\phi$ Decays.* PhD thesis, Universität Karlsruhe, Nov. 2007. IEKP-KA/2007-16.

[23] A. Schmidt. *B Flavour Tagging with Artificial Neural Networks for the CDF II Experiment.* PhD thesis, Universität Karlsruhe, Jan. 2010. IEKP-KA/2010-1.

[24] N. Leonardo. *Analysis of B_s flavor oscillations at CDF.* PhD thesis, Massachusetts Institute of Technology (MIT), Sep. 2006.

[25] `http://www.fnal.gov/pub/about/whatis/history.html`.

[26] `http://www.fnal.gov/pub/science/accelerator/`.

[27] `http://www.fnal.gov/pub/inquiring/physics/accelerators/chainaccel.html`.

[28] `http://www-bd.fnal.gov/public/proton.html`.

[29] A. Boveia. Status and Performance of the CDF Run-II Silicon Detector. In *Proceedings of Science*, volume 377, 2005.

[30] M. Stanitzki. Operational experience with the CDF Run-II silicon detector. In *Nuclear Science Symposium Conference Record*, pages 1101–1105, 2005.

[31] A. Sill et al. *Nucl. Instrum. Methods*, A447:1, 2000.

[32] A. Affolder et al. *Nucl. Instrum. Methods*, A453:84, 2000.

[33] C. S. Hill. *Nucl. Instrum. Methods*, A530:1, 2004.

[34] K. Burkett. Design and Construction of the CDF Central Outer Tracker. *Nucl. Instrum. Methods*, A461:62–64, 2001.

[35] T. Affolder et al. *Nucl. Instrum. Methods*, A526:249, 2004.

[36] D. Acosta et al. Measurement of the J/ψ meson and b-hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ GeV. *Phys. Rev. D*, 71(3):032001, Feb. 2005. 10.1103/PhysRevD.71.032001.

[37] F. Azfar, L. Oakes, and D. Tonelli. Extended and improved dE/dx calibration for 3/fb analyses. 2008. Internal CDF note 9592.

[38] M. Feindt, M. Kreps, T. Kuhr, and J. Morlock. Particle Dependent Time-of-Flight Calibration. 2008. Internal CDF note 9618.

[39] D. Acosta et al. *Nucl. Instrum. Methods*, A518:605, 2004.

[40] G. Pauletta. Operation and performance of the CDF calorimeters. In *Journal of Physics: Conference Series*, volume 160, 2009. 10.1088/1742-6596/160/1/012007.

[41] L. Balka et al. *Nucl. Instrum. Methods*, A267:272–279, 1988.

[42] S. Bertolucci et al. *Nucl. Instrum. Methods*, A267:301–314, 1988.

[43] M. Albrow et al. *Nucl. Instrum. Methods*, A480:524–545, 2002.

[44] C. M. Ginsburg. CDF Run 2 muon system. *Eur. Phys. J.*, pages s1002–s1004, 2004.

[45] A. Artikov et al. Design and construction of new central and forward muon counters for CDF-II. 2004. arXiv:physics/0403079v1.

[46] G. Ascoli et al. *Nucl. Instrum. Methods*, A268:33, 1988.

[47] T. Dorigo et al. *Nucl. Instrum. Methods*, A461:560, 2001.

[48] A. Annovi. CDF: B Physics Performance and Trigger - Operational Experience . In *Nuclear Physics B - Proceedings Supplements*, volume 170, pages 283–287, 2007. Proceedings of the 11th International Conference on B-Physics at Hadron Machines - Beauty 2006.

[49] M. Kreps. Tevatron Experimental Issues at High Luminosities. In *Proceedings of Science*, 2009. arXiv:0912.0110v1.

[50] E. J. Thomson et al. *IEEE Trans. on Nucl. Science.*, 49:1063, 2002.

[51] W. Ashmanskas et al. *Nucl. Instrum. Methods*, A518:532, 2004.

[52] C. Dörr, M. Feindt, A. Gessler, M. Kreps, T. Kuhr, and P. Mack. Optimization of Signal Selection for $B_s \rightarrow D_s(3)\pi$ using a Neural Network. 2009. Internal CDF note 9751.

[53] C. Dörr. *Optimization of the signal selection of exclusively reconstructed decays of B^0 and B_s mesons at CDF-II.* PhD thesis, University of Karlsruhe (TH), 2006. IEKP-KA/2006-6.

[54] P. Mack. *Calibration of New Flavor Tagging Algorithms using B_s Oscillations.* PhD thesis, University of Karlsruhe (TH), 2007. IEKP-KA/2007-10.

[55] A. Gessler. Rekonstruktion der Zerfälle $B_s \rightarrow D_s\pi$ und $B_s \rightarrow D_s\pi\pi\pi$ mit dem CDF-II-Detektor, 2006. Diploma thesis. IEKP-KA/2006-12.

[56] E. Barberio, B. van Eijk, and Z. Wąs. PHOTOS : a universal Monte Carlo for QED radiative corrections; version 2.0. *Comput. Phys. Commun.*, 79:291–308, 1994. CERN-TH.7033/93.

[57] T. Sjöstrand, L. Lönnblad, and S. Mrenna. PYTHIA 6.2 Physics and Manual. 2001. hep-ph/0108264.

[58] K. Anikeev, P. Murat, and Ch. Paus. Description of Bgenerator II. 1999. Internal CDF note 5092.

[59] A. Annovi. CDF: B physics performance & trigger - operational experience. 2006. Public CDF note 8611.

[60] M. Paulini and B. Wicklund. Summary of Proposals for B Physics Triggers in Run II. 2000. Internal CDF note 5483.

[61] The CDF Collaboration. Scale Factors for Proper Time Uncertainties at CDF. 2005. Public CDF note 7944.

[62] Bs Mixing Group (Contact: A. Belloni, I. Furic, G. Gomez-Ceballos, I. Kravchenko, N. Leonardo, S. Menzemer, J. Miles, Ch. Paus, and J. Piedra). Scale Factors for Proper Time Uncertainties with Event-by-Event Primary Vertex Reconstruction. 2005. Internal CDF note 7818.

[63] V. Blobel and E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, chapter 6, pages 183–185. B. G. Teubner, 1998.

[64] F. James. *Minuit - Function Minimization and Error Analysis. Reference Manual.* Computing and Networks Division. CERN Geneva, Switzerland, version 94.1 edition, 1994. CERN Program Library Long Writeup D506.

[65] M. Feindt, M. Kreps, Th. Kuhr, J. Morlock, and A. Schmidt. Measurement of Mixing Frequency and Amplitude of Neutral B-Mesons. 2009. Internal CDF note 10025.

[66] The CDF Collaboration. Measurement of the B_s Lifetime in Hadronic Decays Using Partially and Fully Reconstructed Modes. 2008. Public CDF note 9203.

[67] The CDF Collaboration. Calibration of the Same Side Kaon Tagger using 5.2 fb^{-1} of $p\bar{p}$ Collisions. 2010. Public CDF note 10108.

[68] Bs Mixing Working Group. Study of Bs Oscillations in Bs → Ds pi. 2005. Internal CDF note 7481.

[69] T. Aaltonen et al. First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in $B_s^0 \to J/\psi\phi$ Decays. *Phys. Rev. Lett.*, 100:161802, 2008.

[70] The CDF Collaboration. An Updated Measurement of the CP Violating Phase $\beta_s^{J/\psi\phi}$. 2008. Public CDF note 9458.

[71] The CDF Collaboration. Combination of D0 and CDF results on $\Delta\Gamma_s$ and the CP-violating phase $\beta_s^{J/\psi\phi}$. 2009. Public CDF note 9787.

[72] A. S. Dighe, I. Dunietz, and R. Fleischer. *Eur. Phys. J.*, C6:647, 1999.

Danksagung

Ich möchte mich an dieser Stelle bei Prof. Dr. Michael Feindt für die Möglichkeit bedanken, dass ich diese Dissertation in seiner Gruppe anfertigen durfte. Desweiteren bedanke ich mich bei Prof. Dr. Günter Quast für die Übernahme des Korreferats.

Für die großartige Betreuung während der gesamten Zeit möchte ich mich bei Dr. Thomas Kuhr und Dr. Michal Kreps bedanken. Beim Korrekturlesen dieser Arbeit wurden sie ausserdem von Felix Wick unterstützt, dem hierfür ebenfalls vielen Dank gebührt.

Ermöglicht wurde die Dissertation durch ein Stipendium des Graduiertenkollegs Hochenergiephysik und Teilchenastrophysik der Deutschen Forschungsgemeinschaft.

Prof. Dr. Thomas Müller möchte ich in seiner Funktion als Institutsleiter dafür danken, dass er verschiedene Konferenzbesuche sowie diverser Forschungsaufenthalte ermöglichte.

Den derzeitigen und ehemaligen Doktoranden und Diplomanden aus Raum 9-2 danke ich für das angenehme Arbeitsklima. Hierbei sind besonders die langjährigen Mitstreiter Simon Honc, Felix Wick, Sebastian Neubauer und Dr. Andreas Schmidt zu erwähnen.

Vielen Dank gilt auch der Trampolingruppe und dem Fitness Basics Angebot des Karlsruher Hochschulsports für den passenden körperlichen Ausgleich.

Zu guter Letzt bedanke ich mich bei meinen Freunden und meiner Familie, insbesondere bei meinen Eltern Peter und Renate Morlock und meiner Schwester Ina Morlock nebst ihrer ganzen Familie.