Karlsruhe Institute of Technology

Measurement of the Exclusive and Inclusive Branching Fractions of $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ Decays at CDF and its Implications on the Decay Width Difference in the $B_s^0 - \bar{B}_s^0$ Meson System

---

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. Dominik E. Horn
aus Offenburg

Tag der mündlichen Prüfung:	18.11.2011
Referent:	Prof. Dr. M. Feindt
Korreferent:	Prof. Dr. U. Husemann

“Insight is not the same as scientific
deduction, but even at that it may be more
reliable than statistics.”

Anthony Standen, “*Science is a sacred cow*”.

Synopsis

In recent years, B_s^0 meson decays into non-leptonic CP specific final states have received considerable interest in the heavy flavor physics community. This is primarily due to the rich phenomenology inherent in the rapidly oscillating $B_s^0 - \bar{B}_s^0$ system, and the possibility to probe observables that are directly linked to fundamental parameters of the underlying Kobayashi-Maskawa (KM) model. Being embedded in the Standard Model (SM) of elementary particle physics, the Kobayashi-Maskawa mechanism describes quark flavor transitions through charged currents of weak interaction. While certain observables even offer the chance to reveal manifestations of physics beyond the Standard Model, others point to strongly KM-dominated processes that barely leave any room for competing new physics processes. Nevertheless, the latter provide important complementary information in order to constrain the associated parameter space.

A prominent example for the latter category of observables is given by the CP decay width difference $\Delta\Gamma_s^{CP}$, which is a measure of the lifetime difference between the CP-even and CP-odd B_s^0 eigenstate. Standard Model predictions suggest that the lifetime difference is unusually high compared to the neutral $B_d^0 - \bar{B}_d^0$ meson system, which has very similar mass and mean lifetime though. However, the size of $\Delta\Gamma_s^{CP}$ has not yet been experimentally established beyond doubt. In the Standard Model scenario $\Delta\Gamma_s^{CP}$ equals the decay width difference $\Delta\Gamma_s$, which is a measure for the lifetime difference between the light and heavy B_s^0 mass eigenstate. The mass eigenstates immediately follow from the $B_s^0 - \bar{B}_s^0$ mixing and decay eigenvalue problem in the framework of the KM model. In scenarios excluded by the Standard Model, the light and heavy B_s^0 mass states are no longer eigenstates of CP, and $\Delta\Gamma_s^{CP}$ can be regarded as an important input parameter to constrain the allowed parameter range of new physics models. Among other options, an interesting possibility to estimate $\Delta\Gamma_s^{CP}$ – and hence $\Delta\Gamma_s$ in the Standard Model scenario – is by measuring the partial decay widths of those decays mainly responsible for the decay width difference being non-zero.

It has been suggested that a measurement of the branching fraction of the B_s^0 decay into the inclusive combination of $D_s^{+(*)}D_s^{-(*)}$ final states (throughout this thesis these will be referred to as "semi-inclusive" decays to distinguish this sub-class from the full ensemble of final state combinations) may provide a reasonable estimate of the relative decay width difference $\Delta\Gamma_s/\Gamma$ in the $B_s^0 - \bar{B}_s^0$ system: assuming the mainly *CP*-even decay $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ to dominate over other *CP*-specific $B_s^0 - \bar{B}_s^0$ final states and taking *CP* violation in the $B_s^0 - \bar{B}_s^0$ system to be negligibly small, its branching fraction is directly sensitive to $\Delta\Gamma_s/\Gamma$ through

$$\frac{\Delta\Gamma_s}{\Gamma_s} \cong \frac{2\mathcal{B}(B_s \to D_s^{+(*)}D_s^{-(*)})}{1 - \mathcal{B}(B_s \to D_s^{+(*)}D_s^{-(*)})}$$

Among other theoretical pre-conditions, the size of a potentially non-zero *CP*-odd admixture in $D_s^{+(*)}D_s^{-(*)}$ is critical to the applicability of the latter relation.

The purpose of this thesis is threefold: Firstly, new measurements of both the exclusive and semi-inclusive partial decay widths of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ meson decays are presented. Secondly, the feasibility of extracting the unknown polarization components in $B_s^0 \to D_s^{*+}D_s^{*-}$ by partial reconstruction of this pseudo-scalar to vector-vector decay in a Monte Carlo driven analysis scheme is studied. Finally, based on the suggestions contributed by the theory community this study discusses how a measurement of the branching fraction of semi-inclusive decays $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ can contribute to gain insight about the relative decay width difference in the $B_s^0 - \bar{B}_s^0$ meson system.

The measurement of the exclusive branching fractions $\mathcal{B}(B_s^0 \to D_s^+D_s^-)$, $\mathcal{B}(B_s^0 \to D_s^{*+}D_s^-)$, $\mathcal{B}(B_s^0 \to D_s^{*+}D_s^{*-})$, and the semi-inclusive branching fraction $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$ is performed based on full reconstruction of the decay $B_s^0 \to D_s^+D_s^-$ and partial reconstruction of $B_s^0 \to D_s^{*+}D_s^-$ and $B_s^0 \to D_s^{*+}D_s^{*-}$ from a hadronic data sample 6.8 fb^{-1} in size. The data used in this thesis were collected by the CDF II detector at the $p\bar{p}$ collider Tevatron located at the Fermi National Accelerator Laboratory near Chicago, Illinois, USA. Absolute branching fractions are determined by measuring ratios of branching fractions using $B_d^0 \to D^+D_s^-$ as normalization channel:

$$f_{D_s^{(*)}D_s^{(*)}} = \frac{f_s}{f_d}\frac{\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})}{\mathcal{B}(B_d^0 \to D^+D_s^-)}$$

The branching fraction ratio is multiplied by the ratio of the quark fragmentation fractions f_s/f_d to account for the different probabilities of B_s^0 and B_s^0 meson production at the Tevatron $p\bar{p}$ collision energy.

The analysis described in this thesis contains several improvements over previous measurements: in the estimation of efficiencies for the reconstruction of the intermediate D_s^+ meson from two narrow mass bands of $K^+K^-\pi^+$ phase space for the first time the full underlying Dalitz structure of $D_s^+ \to K^+K^-\pi^+$ is accounted for. Furthermore, for efficiently separating signal from the vast amount of combinatorial background a neural network based multivariate approach is used. Observables are directly extracted from all studied decay channels simultaneously by using a maximum likelihood based parameter estimation model that is optimized towards describing the complicated multi-component structure of the selected invariant mass spectra in a consistent manner. This is mainly achieved by consistently sharing common parameters across the full fit function. In doing so, correlations among free fit parameters are already accounted for at the fitting stage and the effort of systematic studies is reduced considerably.

The thesis outline is as follows: In the first chapter, we motivate this analysis in the light of the conceptual framework of the Kobayashi-Maskawa model, and provide an overview of relevant experimental studies. Chapters 2 and 3 give a brief technical description of the experimental apparatus the used data were collected with. After detailing pre-processing of real data samples and production of simulated data in Chapter 4, Chapter 5 specifies the selection of signal candidates based on a multivariate analysis technique. In Chapter 6 a parametric model is set up to extract the figures of interest from data. To estimate systematic uncertainties, in Chapter 7 model assumptions made are scrutinized, followed by a presentation of branching fractions results and an estimate of the observable $\Delta\Gamma_s/\Gamma_s$ in Chapter 8. Chapter 9 concludes by summarizing the achievements of this study and pointing to the limitations of the $\Delta\Gamma_s/\Gamma_s$ estimation. In this context an outlook on complementary experimental strategies that might be realized in the near future is given.

Contents

Chapter 7 Systematic Studies 121

Chapter 8 Results and Discussion 145

Chapter 9 Concluding Remarks 151

Chapter A Appendix 155

List of Tables

List of Figures

1

Conceptual Framework and Motivation

1.1 Preface

In the Standard Model of elementary particle physics (see e.g. [1] for a pedagogical review), quarks, the fundamental constituents of hadronic matter, never appear isolated, but in bound states only. This phenomenon, that is referred to as *quark confinement*, is a consequence of gluons, the massless gauge bosons mediating the strong force, carrying color charge themselves. When two quarks are separated, the energy density in the "gluon tubes" connecting the two quarks increases continuously, and at some point it becomes energetically more favorable to form a new quark-antiquark pair out of the vacuum. This process is called *hadronization*.

Compound hadronic objects made up of three quarks are called *baryons*, while *mesons* consist of a quark-antiquark pair. While for a long time baryons and mesons have been regarded as the only manifestations of bound quark states realized in nature, only very recently the first tetra-quark state was experimentally established through the observation of the Z^+ particle [2]. As a consequence of confinement, the properties of quarks and the phenomenology of their participation in the three elementary forces embraced by the Standard Model – the strong, the electromagnetic, and the weak force – can not be studied on the basis of isolated objects, but indirectly at hadron level only. However, whenever self-interaction among the hadron constituents can be neglected the observations made at hadron level can be attributed to one individual quark, while the other quark(s) – depending on whether one is dealing with a baryon or a meson – can be considered as not participating in the interaction. The latter quarks are then called *spectator* quarks.

Because of the short lifetime of the *top* quark, the heaviest quark capable of forming hadronic objects existing for a measurable time duration is the *bottom* (or *beauty*) quark b. According to the commonly used naming convention the combination $(\bar{b}q)$ is called B_q meson, while the antiparticle $\bar{B}_q$ consists of $(b\bar{q})$, where q is any of the other quarks lighter than the b quark. With a typical lifetime at the scale of 10^{-12} s, B mesons are relatively long-living and thus provide an excellent laboratory to study the phenomenology of the weak interaction. As all neutral mesons, B_q^0 mesons do not only decay, but in addition periodically change from one flavor state into another. This phenomenon, that is commonly called $B_q^0 - \bar{B}_q^0$ *mixing* or *oscillation*, introduces further physical observables that allow to overconstrain and confirm the parameters of the Standard Model, or to search for hints of processes beyond this exceptionally successful theoretical framework. The character of related experimental studies is somewhat subtle, since they are less concerned with the direct detection of new particles, but rather how decay processes and related observables are influenced by them.

In recent years, a great deal of attention has particularly been paid to the properties of the neutral B_s^0 meson. The reasons for these scientific efforts are twofold: First of all, compared to B_d^0 meson decays the list of verified B_s^0 final states is still relatively fragmentary [3]. Secondly, the $B_s^0 - \bar{B}_s^0$ meson system has predicted – and partially confirmed – properties considerably different from those of the $B_d^0 - \bar{B}_d^0$ meson system. While the rapid $B_s^0 - \bar{B}_s^0$ oscillation frequency has already been verified at the 5σ level [4], the expected sizable difference between the lifetimes of the light and heavy B_s^0 mass eigenstates has not yet been established beyond doubt. Apart from that, considerable focus is currently placed on the possible violation of the symmetry against simultaneous charge and parity transformation in the $B_s^0 - \bar{B}_s^0$ system, as this would clearly point to new physics.

This introductory chapter is organized as follows: In Section 1.2 the *Charge*, *Parity*, and *Time* transformation are formally introduced. Section 1.3 outlines the mechanism of quark flavor exchange in the conceptual framework of the Standard Model, while Section 1.4 motivates the measurement carried out by introducing the key observables of the $B_s^0 - \bar{B}_s^0$ meson mixing and decay eigenvalue problem. We conclude this chapter by briefly reviewing existing measurements (Section 1.6) and defining the scope of the present study (Section 1.7).

1.2 Charge, Parity, and Time Symmetry

In natural science, symmetries play a special role. The reason for this is that according to Noether's theorem [5, 6] every invariance under a symmetry transformation is directly linked to an associated conserved quantity. In the field of particle physics three symmetry transformations are of particular interest:

- The *Charge* transformation induced by the C operator. Under C, all internal quantum numbers of quantum mechanic state are conjugated, $Q \rightarrow -Q$, converting a particle into its antiparticle,

  $$C\,|\Psi(p)\rangle = |\bar{\Psi}(p)\rangle\,, \tag{1.1}$$

  where Ψ is a particle's wave function and p its momentum.

- By application of the *Parity* operator P the handedness of space is inverted by mirroring all spatial coordinates, $\vec{x} \rightarrow -\vec{x}$:

  $$P\,|\Psi(p)\rangle = |\Psi(-p)\rangle \tag{1.2}$$

- The *Time* transformation T induces a reversion in time, $t \rightarrow -t$. Due to $\hat{p} = (1/i)\nabla \rightarrow -(1/i)\nabla$, momentum is inverted and the wave function complex-conjugated:

  $$P\,|\Psi(p)\rangle = |\Psi(-p)\rangle^{*} \tag{1.3}$$

The invariance of a physical state under the simultaneous application of all three operators, CPT, is postulated as a fundamental law in current physical models. The CPT symmetry theorem was first proved by G. Lüders and W. Pauli [7], and up to the present day all observations indicate that CPT indeed is a preserved symmetry.

The situation is different for the separate action of C, P, and T: The weak interaction separately violates C and P in a maximal way. While the concatenation CP is preserved in most weak processes, it is violated in certain neutral meson systems. CP violation was for the first time discovered in 1964 through the observation of the 2π decay of the neutral K_2 meson [8]. No other fundamental force considered in the Standard Model is known to violate C, P, or T. Violation of the C and CP symmetries is one of the three necessary conditions proposed by A. Sakharov [9] to explain baryogenesis within existing cosmological models. The size of all CP violating effects in the weak sector is, however, by far not enough to explain the matter-antimatter imbalance that we observe in today's universe. Finding evidence of sizable sources of C and CP violation is therefore one of the most important fields of investigation in particle physics. For a compact summary of the phenomenology and history of *CP* violation in meson decays in general the interested reader may refer to the review in Ref. [10].

Throughout the description of the $B_s^0 - \bar{B}_s^0$ mixing and decay eigenvalue problem in Section 1.4 we will introduce states that are invariant under the *CP* transformation by construction, i.e. that are eigenstates of *CP*.

1.3 Cabibbo-Kobayashi-Maskawa Matrix

The Kobayashi-Maskawa mechanism [11] has been proven to be a consistent and very successful phenomenological model to describe both quark-mixing and *CP* violation in terms of Yukawa couplings. Quark mixing describes the phenomenon of the transition of a given quark flavor into a different one. Within the Standard Model description of the weak interaction [12, 13, 14], the only flavor changing process is the charged-current $W^{\pm}$ coupling to the physical left-handed up-type antiquarks $\overline{u_{Li}}$ and down-type quarks d_{Lj}, where i and j label the quark generation. The charged current of weak interaction is described by the Lagrangian

$$\begin{aligned} \mathcal{L}_{W^{\pm}} &= -\frac{g}{\sqrt{2}}\,\overline{u_{Li}}\,\gamma^{\mu}\,(V_{CKM})_{ij}\,d_{Lj}\,W_{\mu}^{\dagger} + \text{h.c.} \qquad (1.4) \\ &= -\frac{g}{\sqrt{2}}\,(\overline{u_L}, \overline{c_L}, \overline{t_L})\,\gamma^{\mu}\,(V_{CKM})_{ij} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_{\mu}^{\dagger} + \text{h.c.} \end{aligned}$$

while the coupling strengths are given by the elements of the *Cabibbo-Kobayashi-Maskawa (CKM)* quark-mixing matrix V_{CKM} [11, 15]. Here, g denotes the weak coupling constant and γ^{μ} the Dirac matrices. In the Standard Model, just as the masses, mixing of quarks originates from Yukawa interactions with the scalar Higgs field ϕ. The interaction term of the Yukawa Lagrangian $\mathcal{L}_Y$ reads

$$\mathcal{L}_Y = -Y_{ij}^d\,\overline{Q_{Li}^I}\,\phi\,d_{Rj}^I - Y_{ij}^u\,\overline{Q_{Li}^I}\,\epsilon\,\phi^*\,u_{Rj}^I + \text{h.c.} \qquad (1.5)$$

where ϵ is the 2×2 antisymmetric tensor and $Y^{u,d}$ are general complex-valued 3×3 matrices. The quark fields Q_{Li}^I are left-handed doublets, while d_R^I and u_R^I are right-handed singlets in the weak-eigenstate basis. The CKM matrix arises from the diagonalization of $Y^{u,d}$ by virtue of four unitary transformation matrices, $V_{L,R}^{u,d}$, yielding the quark mass terms $M_{diag}^f = V_L^f\,Y^f\,V_R^{f\dagger}(v/\sqrt{2})$, $f = u, d$. The factor $(v/\sqrt{2})$ stems from the ϕ vacuum expectation value $\langle\phi\rangle = (0, v/\sqrt{2})$. The CKM matrix has the explicit form

$$V_{CKM} \equiv V_L^u V_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.6}$$

and the weak eigenstates of the down-type quarks, d'_j, can be represented as linear superpositions of the mass eigenstates of down-type quarks d_j,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.7}$$

Being a unitary 3×3 matrix, V_{CKM} depends on three real parameters and six phases. The parameter freedom is substantially reduced by the possibility of arbitrarily redefining the phases of the quark mass eigenstates, leaving one single phase only. The four remaining parameters are interpreted as three real rotation angles and the Kobayashi-Maskawa phase. In the Standard Model, the latter is the only source of all CP-violating phenomena in flavor changing processes.

For further discussions, it is worthwhile introducing an explicit parameterization of V_{CKM}. Having three angles θ_{ij} and a complex phase δ, a manifest parameterization is [16]

$$\mathbf{V}_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.8}$$

where s_{ij} and c_{ij} stand for $\sin\theta_{ij}$ and $\cos\theta_{ij}$, respectively. This parameterization of V_{CKM} can be substantially simplified by using the Wolfenstein parameterization [17], which takes advantage of the experimentally known hierarchy $s_{13} \ll s_{23} \ll s_{12} \ll 1$. With

$$\begin{aligned} s_{12} =& \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \cong\ 0.22,\ s_{23} = A\lambda^2 = \lambda \left|\frac{V_{cb}}{V_{us}}\right|, \\ s_{13}e^{e\delta} =& V_{ub}^* = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar\rho + i\bar\eta)\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\bar\rho + i\bar\eta)]} \end{aligned} \tag{1.9}$$

one can write the CKM matrix up to $\mathcal{O}(\lambda^3)$,

$$\mathbf{V}_{CKM} \cong \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right) \tag{1.10}$$

where in this notation the traditional representation in terms of ρ and η was used. For a geometrical representation of V_{CKM} in the complex plane, that is briefly discussed in the next section, $\bar{\rho} + i\bar{\eta} = -(V_{us}V_{ub}^*)/(V_{cs}V_{cb}^*)$ is a more common choice.

From (1.10) the hierarchy of coupling strengths among the quark generations becomes easily visible: The highest quark transition probability ($\mathcal{O}(1)$) is found for quarks belonging to the same generation. The coupling strengths among the first and the second quark generation is still of the order λ, while it is $\mathcal{O}(\lambda^2)$ among quark generations two and three. In contrast, transitions from the first to the third quark generation are suppressed by a factor of $\sim \lambda^3$.

1.3.1 Geometrical Interpretation

The unitarity condition of the CKM matrix,

$$\sum_{i=1}^{3} V_{ij}V_{ik}^* = \delta_{jk}, \qquad \sum_{j=1}^{3} V_{ij}V_{kj}^* = \delta_{ik}, \quad k = 1,2,3, \tag{1.11}$$

manifests in six vanishing relations, where three arise from the scalar products of two out of the three columns, and three relations from row products. The column products have the explicit form

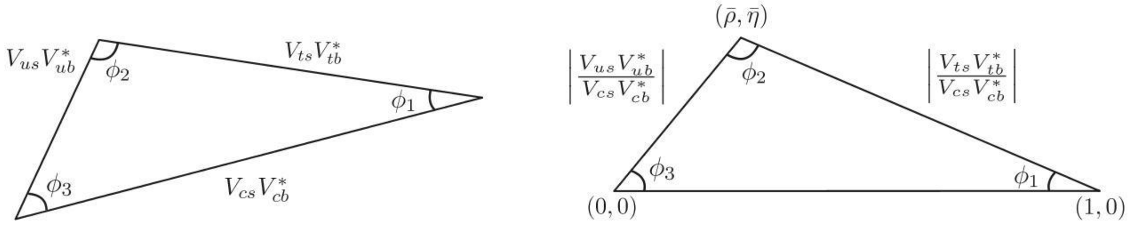
$$(23): \quad V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{cb}^* = 0 \tag{1.12}$$

$$(13): \quad V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.13}$$

$$(12): \quad V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{1.14}$$

Each of these equations can be geometrically interpreted as a triangle in the complex plane, where the length of each side represents the strength of the quark couplings, and the angles correspond to the relative phases among them. The area of each of the triangles equals $J/2$, where J is the Jarlskog invariant [18], and is identical for all triangles. Using the Wolfenstein parameterization, the triangle basis has an exact length of 1 and phase 0, and the other two sides meet in the apex ($\bar{\rho}$, $\bar{\eta}$). A general goal of flavor physics is to find observables that help to overconstrain the CKM elements. The graphical representation of CKM unitarity provides a convenient way to illustrate and compare results and to do consistency checks.

For the following discussions of B_s^0 meson mixing and decay, we focus on the first unitarity condition, equation (1.12), which is the scalar product of the second and the third column. The left plot of Figure 1.1 displays the associated unitary triangle, while in the right hand plot the same unitarity triangle is shown in the



(a) (b)

Figure 1.1: (a): Graphical representation of the unitarity condition (1.12), called "unitary triangle". (b) The same unitary triangle rotated into the $(\bar{\rho},\ \bar{\eta})$ plane.

complex $(\bar{\rho},\ \bar{\eta})$ plane. From this representation the three angles can be read off:

$$\phi_1 \equiv \beta = \arg\left(-\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*}\right) \tag{1.15}$$

$$\phi_2 \equiv \alpha = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{us}V_{ub}^*}\right) \tag{1.16}$$

$$\phi_3 \equiv \gamma = \arg\left(-\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*}\right), \tag{1.17}$$

where "arg" denotes the argument or phase of the complex number, so e.g. for $(V_{cs}V_{cb}^*)/(V_{ts}V_{tb}^*)$ the phase β is defined by

$$\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*} = -\left|\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*}\right| e^{i\beta} \tag{1.18}$$

In the Standard Model, the phase β is close to zero. Therefore, in contrast to the symbolic depictions of Figure 1.1 the unitarity triangle associated with (1.12) is almost degenerate.

1.4 B_s Meson Mixing and Decay Eigenvalue Problem

Throughout this section, different bases of eigenstates used to describe the $B_s^0 - \bar{B}_s^0$ meson mixing and decay problem are often referred to. The flavor eigenstates reflect the flavor of the quark content of the B_s^0 meson, where we follow the common convention

$$|B_s\rangle = |\bar{b}s\rangle\,, \qquad |\bar{B}_s\rangle = |b\bar{s}\rangle \tag{1.19}$$

The flavor eigenstates are not invariant under the concatenation of the Charge C (1.1) and Parity P (1.2) operation,

$$CP\,|B_s\rangle = -\,|\bar{B}_s\rangle\,. \tag{1.20}$$

B_s and $\bar{B}_s$ are hence no eigenstates of CP, yet their linear combinations

$$|B_s^{CP+}\rangle = \frac{1}{\sqrt{2}}\left(|B_s\rangle - |\bar{B}_s\rangle\right), \quad \text{and} \quad |B_s^{CP-}\rangle = \frac{1}{\sqrt{2}}\left(|B_s\rangle + |\bar{B}_s\rangle\right), \tag{1.21}$$

by construction are eigenstates of CP. $CP+$ and $CP-$ label the even and odd CP eigenstate, respectively. The factor of $1/\sqrt{2}$ ensures normalization.

The mass eigenstates B_s^L (light) and B_s^H (heavy) do not coincide with the flavor eigenstates either, but are as well linear superpositions of them:

$$|B_s^L\rangle = p\,|B_s\rangle + q\,|\bar{B}_s\rangle, \quad \text{and} \quad |B_s^H\rangle = p\,|B_s\rangle - q\,|\bar{B}_s\rangle, \tag{1.22}$$

with the normalization condition $\left|p\right|^2 + \left|q\right|^2 = 1$. While the CP eigenstates are constructed by symmetry considerations, the mass eigenstates (1.22) immediately follow from the mixing and decay eigenvalue problem formulated for the $B_s^0 - \bar{B}_s^0$ system.

As a compound object consisting of the elementary quark material, the flavor of a meson that is a B_s^0 at time $t = t_0$ can change over time by virtue of weak charged currents. This manifests in three ways : Either by decay into a different particle, or, being a neutral meson, by transition into its own anti-particle $\bar{B}_s^0$, where the latter phenomenon is referred to as $B_s^0 - \bar{B}_s^0$ mixing, and thirdly by first mixing into $\bar{B}_s^0$ and decaying afterwards. There is a wealth of pedagogical literature [19, 20, 21] concerning the mixing and decay of neutral $B_q^0 - \bar{B}_q^0$ meson systems the following discussion is extracted from.

The evolution in time of a $B_s^0 - \bar{B}_s^0$ system driven by both mixing and decay can be described in the Schrödinger formalism,

$$i\frac{d}{dt}\begin{pmatrix}|B_s(t)\rangle \\ |\bar{B}_s(t)\rangle\end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix}|B_s(t)\rangle \\ |\bar{B}_s(t)\rangle\end{pmatrix}, \tag{1.23}$$

using a 2×2 non-Hermitian effective Hamiltonian $\mathbf{H}$ which is constructed as the sum of the complex Hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$

$$\mathbf{H} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right) = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22}\end{pmatrix}. \tag{1.24}$$

$\mathbf{M}$ and $\mathbf{\Gamma}$ are identified as the mass (or mixing) matrix and the decay matrix, respectively. While the diagonal elements $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$ correspond to flavor-conserving transitions, only the off-diagonal elements $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$ cause a change in flavor. To lowest order, $(\bar{b}s) - (b\bar{s})$ transitions are mediated by flavor changing charged current loops with at least two $W^\pm$ bosons involved. These lowest order Feynman diagrams are commonly illustrated by so called *box diagrams* (Figure 1.2). As far as Γ_{12} is concerned, the dominant

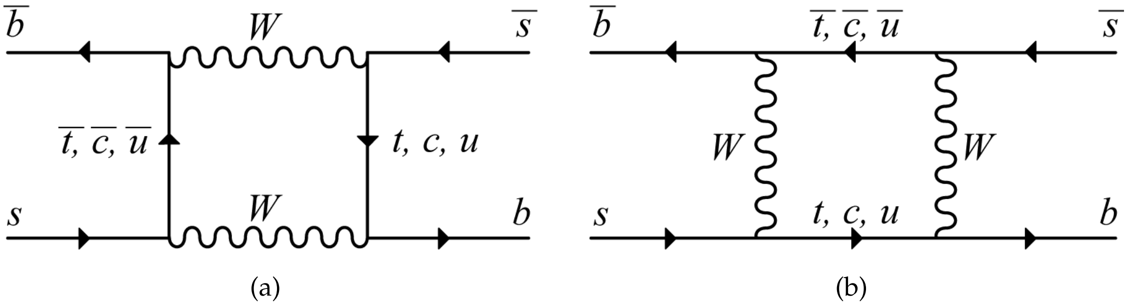


Figure 1.2: Lowest order Feynman diagrams, so called box diagrams, illustrating $B_s^0 - \bar{B}_s^0$ mixing. The main loop contribution is given by the *top* quark.

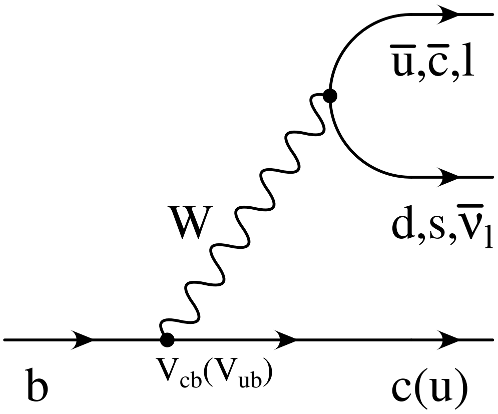


Figure 1.3: Feynman diagram displaying tree level b decays in the Standard Model. The main contribution stems from $b \to c\bar{c}s$.

contribution stems from CKM-favored $b \to c\bar{c}s$ tree-level decays (Figure 1.3) into *CP*-specific final states [22]. Owing to the small couplings $V_{ub}V_{us}^*$, the doubly Cabibbo-suppressed $b \to u\bar{u}s$ contributions can safely be neglected.

Using

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}}, \tag{1.25}$$

diagonalizing **H** yields the complex eigenvalues

$$\omega_{L,H} = (M - \frac{i}{2}\Gamma) \pm \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}), \tag{1.26}$$

and the associated eigenvectors, i.e. the mass eigenstates $|B_s^H\rangle$ and $|B_s^L\rangle$, that were introduced by (1.22). The real parts $m_{L,H}$ and the negative half of the imaginary parts $(i/2)\Gamma_{L,H}$ of $\omega_{L,H}$,

$$\omega_{L,H} = m_{L,H} - \frac{i}{2}\Gamma_{L,H} \tag{1.27}$$

are interpreted as the masses and the widths of the light and the heavy mass eigenstate. The relation between the partial widths $\Gamma_{L,H}$ and the total width Γ is given by

$$\Gamma = \frac{1}{\tau_{B_s^0}} = \frac{\Gamma_H + \Gamma_L}{2}, \tag{1.28}$$

where $\tau_{B_s^0}$ is the mean B_s^0 lifetime. The splittings in mass and width are defined as

$$\Delta m_s \equiv m_H - m_L = \Re(\omega_H - \omega_L) \tag{1.29}$$

$$\Delta\Gamma_s \equiv \Gamma_L - \Gamma_H = -2\Im(\omega_L - \omega_H) \tag{1.30}$$

Δm is positive by definition, while $\Delta\Gamma_s$ is made positive by convention. The physical meanings of Δm and $\Delta\Gamma_s$ are the following: While the mass difference determines the frequency the $B_s^0 - \bar{B}_s^0$ system oscillates with, considering (1.28) a decay width difference reflects the fact that the distinct mass eigenstates have different lifetimes. Or, to phrase it differently, the decay width difference is non-zero if there are more final states accessible to one of the two mass eigenstates than to the other. The size of mass splitting has been measured with high precision by the CDF experiment as $\Delta m = 17.77 \pm 0.10 \pm 0.07\,\mathrm{ps}^{-1}$ [4], a value which corresponds to rapid oscillations when compared to the B_s^0 lifetime. With the consistent result $\Delta m = 17.63 \pm 0.11 \pm 0.04\,\mathrm{ps}^{-1}$ the LHCb experiment [23] has already achieved a similar level of precision. Both results are in good agreement with the theoretical prediction $\Delta m = 17.3 \pm 2.6\,\mathrm{ps}^{-1}$ [24]. Despite the naming scheme using the attributes "light" and "heavy", it is worth noting that the mass difference expressed in the energy dimension is tiny, as low as $\Delta m = (1.170 \pm 0.008) \times 10^{-2}\,eV$ [3]. In contrast to the precise state of knowledge with respect to Δm, the size of the decay width difference is experimentally still not well established at a reasonable precision level, as will be highlighted later by Table 1.2. The theory prediction for the relative decay width difference [25],

$$\frac{\Delta\Gamma_s}{\Gamma_s} = 0.147 \pm 0.060, \tag{1.31}$$

has large uncertainties though as well.

By a closer examination of the $\mathbf{M} - (i/2)\,\mathbf{\Gamma}$ eigenvalue problem and using $|\Gamma_{12}| \ll |M_{12}|$, one finds that these observables are directly connected to the off-diagonal elements of the mass and decay matrix:

$$\Delta m_s = 2\,|M_{12}| \tag{1.32}$$

$$\Delta\Gamma_s = 2\,|\Gamma_{12}| \cos\phi_s = \Delta\Gamma_{CP}^s \cos\phi_s, \tag{1.33}$$

where in (1.33) one takes advantage of the fact that $b \to c\bar{c}s$ decays into CP-specific final states constitute the dominant contribution to Γ_{12}, implying $2\,|\Gamma_{12}| = \Delta\Gamma_{CP}^s$. As pointed out in Refs. [22, 24] it must be stressed that the possibility to relate $|\Gamma_{12}|$ to the measurable quantity $\Delta\Gamma^s_{CP}$ crucially depends on the fact that Γ_{12} is dominated by a single weak phase. Omitting $b \to u\bar{u}s$ transitions involving a weak phase different from $b \to c\bar{c}s$ introduces a theoretical uncertainty of about 5%. In (1.33), ϕ_s is the relative phase of the complex off-diagonal elements M_{12} and Γ_{12}, the phases of which are defined by

$$M_{12} = |M_{12}|\, e^{i\phi_M} \tag{1.34}$$

$$\Gamma_{12} = -\,|\Gamma_{12}|\, e^{i\phi_\Gamma} \tag{1.35}$$

Dividing (1.34) by (1.35) yields the relative phase ϕ_s

$$\frac{M_{12}}{\Gamma_{12}} = -\left|\frac{M_{12}}{\Gamma_{12}}\right| e^{i(\phi_M-\phi_\Gamma)} = -\left|\frac{M_{12}}{\Gamma_{12}}\right| e^{i\phi_s}. \tag{1.36}$$

While the phases in (1.34) and (1.35) are convention-dependent, the combination $\phi_s = \phi_M - \phi_\Gamma$ has physical meaning. ϕ_s is the measure of the size of CP violation in the interference between mixing and decay. This classification of CP violation describes the phenomenon that the absolute amplitude of the interference between a mixed and an unmixed B^0_s decay into a CP final state f does not equal the one of the $\bar{B}^0_s$ into a opposite CP final state $\bar{f}$:

$$\left|\langle f \mid \bar{B}^0_s \to B^0_s\rangle + \langle f \mid B^0_s\rangle\right| \neq \left|\langle \bar{f} \mid B^0_s \to \bar{B}^0_s\rangle + \langle \bar{f} \mid \bar{B}^0_s\rangle\right| \tag{1.37}$$

In the Standard Model, the phase ϕ_s is predicted to be close to zero, $\phi_s = 0.22^\circ \pm 0.06^\circ$ [24], corresponding to a vanishing CP violation in the $B^0_s - \bar{B}^0_s$ system. The measurement of a sizable non-zero phase ϕ_s would therefore be an unambiguous hint to new physics beyond the CKM mechanism. Since Γ_{12} is strongly dominated by CKM-favored tree-level processes that barely leave any room for competing new physics processes, one expects that any observed deviation in ϕ_s is due to new contributions in mixing. This, plus the poor knowledge of the decay width difference, led the B physics community focus their research efforts on measurements aiming at constraining relation (1.33), or even find a "smoking gun" for sizable CP violation in $B^0_s - \bar{B}^0_s$. One of the most promising decay channels to pursue this goal is the decay of $B^0_s \to J/\psi\phi$, which in recent years has been subject to intensive studies carried out by both the Fermilab experiments CDF [26] and DØ [27], and the LHCb experiment [28] located at CERN. This "golden" decay channel allows to probe the decay width difference and a potential sizable new-physics phase ϕ_s^{NP} simultaneously. However, this measurement is experimentally elaborate, since it requires angular and time-dependent analysis techniques plus tagging the flavor of the B^0_s meson at production time. New physics scenarios are not favored by the most recent results, can however not yet be ruled out given the current statistical precision.

Out of the $B_s^0-\bar{B}_s^0$ mixing and decay observables m_s, Δm_s, Γ_s, $\Delta\Gamma_s$ and ϕ_s, in the following we will focus on the decay width difference $\Delta\Gamma_s$ only and discuss how this observable can be accessed experimentally. In this regard, the measurement of the branching fraction of a yet another type of $b \to c\bar{c}s$ decays, $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ is motivated.

1.5 Accessing $\Delta\Gamma_s$

To access a decay width difference basically two distinct experimental strategies requiring different analysis techniques may be pursued: Through lifetime measurements of final states common to B_s^0 and $\bar{B}_s^0$ by performing fits to decay time distributions, or by measuring relative partial decay widths – also referred to as branching fractions – by quantifying signal yields. Both approaches require a final state to be unambiguously identified as *CP*-even or odd. Assuming the Standard Model scenario in which the phase ϕ_s is tiny, the *CP* eigenstates coincide with the mass eigenstates, and equation (1.33) implies that the *CP* width difference $\Delta\Gamma_s^{CP}$ equals the decay width difference $\Delta\Gamma_s$. In this sense, the terms "*CP* eigenstate" and "mass eigenstate" can be used interchangeably. Lifetime-related analysis methods comprise:

i Measurement of the lifetimes τ_{CP+} and τ_{CP-} by fits to the decay time distributions of pure *CP*-even and odd final states. The decay width difference is then given by $\Delta\Gamma_s = 1/\tau_{CP+} - 1/\tau_{CP-} = 1/\tau_L - 1/\tau_H$. Promising candidates are the pure *CP*-even final state $B_s^0 \to D_s^+D_s^-$ and the pure *CP*-odd decay $B_s^0 \to J/\psi f_0$, where for the latter decay both CDF [29] and LHCb [30] recently provided first results.

ii Extracting $\tau_{CP+,CP-}$ out of one single B_s^0 decay with the final state not being pure with regard to its *CP* content, but a mixture of *CP*-odd and even components, that can be disentangled by angular analyses. Typical decay modes are the $P \to VV$ (pseudo-scalar to vector-vector) decays $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s^{*+}D_s^{*-}$.

Further constraints can be introduced by fitting decay time distributions of flavor-specific decays. This technique directly determines the mean B_s^0 decay width Γ_s. Depending on the final state and data recording methods, decay time measurements and angular analyses can get very challenging though, if not impossible. At CDF, this particularly applies to $B_s^0 \to D_s^{*+}D_s^{*-}$, for two reasons: Firstly, the event recording trigger (see Section 3.6) places requirements on kinematic quantities highly correlated to the decay time, that introduce a significant lifetime bias. Secondly, owing to their low energies, the neutral particles emanating from de-excitation of the D_s^{*+} meson have a very low detection efficiency. By that, a data-driven angular analysis of $B_s^0 \to D_s^{*+}D_s^{*-}$ cannot be carried out using CDF data.

Branching fraction measurements represent an interesting alternative to direct lifetime measurements to infer $\Delta\Gamma_s$. Specifically, the methods we consider are:

i Determining the partial decay widths of all possible *CP*-specific final states common to B_s^0 and $\bar{B}_s^0$. While this method relies on the least assumptions and gives a very precise estimate of the relative decay width difference $\Delta\Gamma_s/\Gamma_s$, it is experimentally most tedious. Furthermore, for modes involving an admixture of *CP*-odd and *CP*-even components, again angular analyses are required to separate these eigenstates.

ii This method is a special case of (i): If there is one *CP*-specific class of decay modes that place the main contribution on $\Delta\Gamma_s$, measuring its semi-inclusive partial decay width should give a reasonable estimate of $\Delta\Gamma_s$. This method, however, requires more stringent assumptions: All other contributions need to be regarded as negligible, and the given final state must have one well-defined *CP* content.

As already indicated, each category comes with its own theoretical and experimental challenges. The combination of both however offers the chance to draw a consistent picture of the decay width observable.

This thesis focusses on branching fraction-based measurements, in particular on the special case (ii). It has been argued [31] that $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ is a promising candidate to adopt the latter experimental method. In the following sections the key assumptions that allow to infer $\Delta\Gamma_s/\Gamma_s$ from a branching fraction measurement of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ are summarized, while pointing to the pitfalls of this approach at the same time.

1.5.1 Accessing $\Delta\Gamma_s/\Gamma_s$ Using $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$

Decays of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ proceed via strongly CKM-favored $b \to c\bar{c}s$ tree transitions, where the strange quark s is regarded as *spectator* not participating in the weak current. Figure 1.4 shows the associated Feynman diagram. Although other final states like $B_s^0 \to J/\psi\phi$ or $B_s^0 \to J/\psi\eta$ have the same quark content as $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$, they are suppressed by kinematics and a color factor of 3. Color

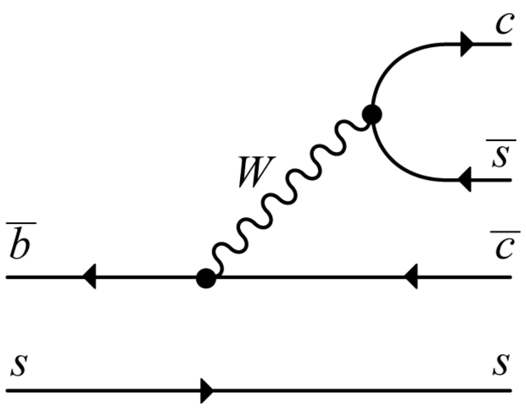


Figure 1.4: Feynman diagram displaying the decay of a neutral B_s ($\bar{b}s$) into the charged mesons $D_s^{+(*)}$ ($c\bar{s}$) and $D_s^{-(*)}$ ($\bar{c}s$). The strange quark is regarded as spectator quark not participating in the weak interaction process.

suppression in weak currents is a consequence of mesons being color-neutral objects and weak interaction being color-conserving. As a consequence, neglecting multi-body decays with final states involving more than two particles, $D_s^{(*)+}D_s^{(*)-}$ is the dominant final state triggered by $b \to c\bar{c}s$ tree transitions. If this final state can furthermore be shown to be a pure CP eigenstate, it would provide an excellent laboratory for probing the decay width difference $\Delta\Gamma_s$ in the $B_s^0-\bar{B}_s^0$ meson system.

While the exclusive decay $B_s^0 \to D_s^+D_s^-$ is intrinsically a pure CP-even final state, the CP content of $B_s^0 \to D_s^{*+}D_s^-$ and $B_s^0 \to D_s^{*+}D_s^{*-}$ can only be approximated to be even under certain assumptions [31]: Under the Shifman-Voloshin (SV) limit [32] $m_{c,b} \to \infty$ with $m_b - 2m_c \to 0$, and under the limit of infinite colors $N_{color} \to \infty$, it can be shown that both $B_s^{CP-} \to D_s^{*+}D_s^-$ and $B_s^{CP-} \to D_s^{*+}D_s^{*-}$ vanish and $\Delta\Gamma_s^{CP}$ is saturated by $\Gamma(B_s^{CP+} \to D_s^{(*)+}D_s^{(*)-})$. Thus, with $\Gamma(B_s^{CP-} \to D_s^{(*)+}D_s^{(*)-}) = 0$, the CP width difference equals the CP-even decay width:

$$\Delta\Gamma_s^{CP} = \Gamma(B_s^{CP+} \to D_s^{(*)+}D_s^{(*)-}). \tag{1.38}$$

The partial decay width of the semi-inclusive decay $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ therefore gives an estimate on the full decay width difference, or to phrase it in terms of the branching fraction $\mathcal{B}$ [22],

$$\begin{aligned} 2\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) &\cong \frac{\Delta\Gamma_s^{CP}}{2}\left[\frac{\frac{1}{1-2x_f}+\cos\phi_s}{\Gamma_s^L}+\frac{\frac{1}{1-2x_f}-\cos\phi_s}{\Gamma_s^H}\right] \\ &\cong \frac{\Delta\Gamma_s^{CP}}{2}\left[\frac{1+\cos\phi_s}{\Gamma_s^L}+\frac{1-\cos\phi_s}{\Gamma_s^H}\right], \end{aligned} \tag{1.39}$$

where in the second step the CP-odd fraction x_f, defined by $\Gamma_s^{CP-}/\Gamma_s^{CP+} = x_f/(1-x_f)$, has been set to zero. Substituting $\Gamma_s^{L,H} = \Gamma_s \pm \Delta\Gamma_s/2$ and $\Delta\Gamma_s^{CP} = \Delta\Gamma_s/\cos\phi_s$, and further assuming ϕ_s to be negligibly small according to the Standard Model scenario, equation (1.39) simplifies to

$$2\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-}) \cong \frac{\Delta\Gamma_s}{\Gamma_s + \Delta\Gamma_s/2}, \tag{1.40}$$

which one can solve for the relative decay width difference:

$$\frac{\Delta\Gamma_s}{\Gamma_s} \cong \frac{2\mathcal{B}(B_s \to D_s^{(*)+} D_s^{(*)-})}{1 - \mathcal{B}(B_s \to D_s^{(*)+} D_s^{(*)-})} \tag{1.41}$$

Thus, under the theoretical assumptions made, a measurement of $\mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})$ gives a direct estimate of the strongly Standard Model dominated observable $\Delta\Gamma_s^{CP} = 2\,|\Gamma_{12}|$, independently of CP violation. In case of the Standard Model scenario predicting a negligible small phase $\phi_s \approx 0$, it is furthermore sensitive to $\Delta\Gamma_s/\Gamma_s$. As a matter of fact, an analysis of this kind is not a "smoking gun" measurement in terms of new physics effects in $B_s^0 - \bar{B}_s^0$; nevertheless it can make an important complementary contribution to this field by providing an estimate of $\Delta\Gamma_s$ using an approach that does not rely on angular and decay time studies. Furthermore, if $\Delta\Gamma_s^{CP}$ was realized vanishing in nature, there would be no way to measure a non-zero CP-violating phase ϕ_s in untagged analyses, as can be seen from equation (1.33).

1.5.2 Theoretical Caveat

The previous discussions have shown that the reduction in experimental complexity comes at the cost of predictive robustness owing to the theoretical assumptions that need to be made. As the theoretical caveats and possible experimental remedies already have been exhaustively discussed in Ref. [22] the key issues are pointed out only.

Among the theoretical assumptions made, two are of particular concern, since they can significantly increase the theoretical uncertainty in the estimation of $\Delta\Gamma_s/\Gamma_s$, or even spoil the whole ansatz: Neglecting three-body modes, and the cleanliness of the *CP* content of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$. Both issues are controversial, as only theoretical predictions exist to date.

While the preference of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ over other two-body decay modes can be regarded as a safe assumption, the possibility to neglect other final states crucially depends on the branching fractions of multi-body modes, such as $B_s^0 \to D_s^{(*)} D^{(*)} K^{(*)}$. Recently published theoretical calculations [33] suggest that the partial widths of three-body modes can be sizable. According to the new estimations, their contribution to $\Delta\Gamma_s$ is not only greater than hitherto assumed, but almost equal to the contribution of two-body modes. As a consequence, approximating $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decays to saturate $\Delta\Gamma_s$ needs a considerable correction. However, it should be stressed that the theoretical uncertainties involved are very large and up to the present day none of the B_s^0 three-body decay modes has been experimentally established. Hence their true branching fractions, not to mention their *CP* content, remain unknown.

Another concern is related to the *CP* content of the $D_s^{(*)+} D_s^{(*)-}$ final state. The presence of a sizeable *CP*-odd component would be serious, since it would add to $\Delta\Gamma_s^{CP}$ in (1.39) with the wrong sign. For the exclusive $P \to VV$ decay $B_s^0 \to D_s^{*+} D_s^{*-}$ theory predicts a small *CP*-odd fraction of $\sim 5\%$ [34]. Depending on the size of $\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)$ and $\mathcal{B}(B_s^0 \to D_s^+ D_s^-)$, the share of the *CP*-odd component on the semi-inclusive final state $D_s^{(*)+} D_s^{(*)-}$ is correspondingly smaller. The true size of the *CP*-odd component still remains unknown, because limited statistics of data samples or other experimental limitations yet prohibit the extraction of f_{CP-} in $B_s^0 \to D_s^{*+} D_s^{*-}$.

In conclusion, the possibility to relate the branching fraction of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ to the decay width difference in the $B_s^0 - \bar{B}_s^0$ system strongly depends on the theoretical assumptions made. In this regard, any results quoted should therefore include a clear specification of the theoretical regime chosen.

1.6 Experimental Status

The first experiment to report evidence for $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ was the ALEPH collaboration [35]. In this measurement $D_s^{(*\pm)}$ mesons were reconstructed by evaluating correlations among two ϕ mesons detected in the same hemisphere. The semi-inclusive branching fraction was quoted as $\mathcal{B}(B_s \to D_s^{(*)+} D_s^{(*)-}) = 0.14 \pm 0.06 \pm 0.03$, giving rise to $\Delta\Gamma_s^{CP}/\Gamma_s = 0.25^{+0.21}_{-0.14}$.

DØ reported evidence for the decay $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ and a non-vanishing decay width difference using data corresponding to 2.8 fb^{-1} [36]. Based on a reconstruction of the semi-leptonic final state $D_s^{*+}(\to D_s^+ \gamma/\pi^0) D_s^{*-}(\to D_s^- \gamma/\pi^0)$, $D_s^+(\to \phi\pi^+) D_s^-(\to \phi\mu^-\bar{\nu}_\mu)$, 27 signal events were found. With $\mathcal{B}(B_s \to D_s^{(*)+} D_s^{(*)-}) = 0.035 \pm 0.010(stat) \pm 0.011(syst)$ and by assuming the semi-inclusive final state to be predominantly CP even, $\Delta\Gamma_s/\Gamma_s = 0.072 \pm 0.021(stat) \pm 0.022(syst)$ in the standard model scenario was derived. Both the ALEPH and the DØ measurement have in common that no attempt was made to distinguish between decays of a D_s^+ or a D_s^{*+} meson.

First observation of the decay $B_s \to D_s^+ D_s^-$ on the basis of an exclusive mea-

	ALEPH	DØ	CDF	Belle
Signal Yield	14	27	24	23
$\mathcal{B}(B_s^0 \to D_s^+ D_s^-)$ (%)	-	-	$1.04^{+0.35+1.1}_{-0.32-1.1}$	$1.03^{+0.39+0.26}_{-0.32-0.25}$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)$ (%)	-	-	-	$2.75^{+0.83+0.69}_{-0.71-0.69}$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})$ (%)	-	-	-	$3.08^{+1.22+0.84}_{-1.04-0.84}$
$\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ (%)	$14 \pm 6 \pm 3$	$3.5 \pm 1.0 \pm 1.1$	-	$6.85^{+1.53+1.89}_{-1.30-1.89}$
$\Delta\Gamma_s/\Gamma_s$ (%)	25^{+21}_{-14}	$7.2 \pm 2.1 \pm 2.2$	> 1.2	$14.7^{+3.6+4.4}_{-3.0-4.2}$

Table 1.1: Results of $B_s^0 \to D_s^{*+} D_s^-$, $B_s^0 \to D_s^{*+} D_s^-$, and $B_s^0 \to D_s^{*+} D_s^{*-}$ analyses. The next to last row gives the branching fraction of the semi-inclusive decays $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ with no attempt being made to distinguish between D_s^+ and D_s^{*+}. In the last row an estimate on the relative decay width difference is given using the theoretical considerations outlined in section 1.5.1.

surement was reported by the CDF collaboration using 355 pb^{-1} of data [37, 38]. By analyzing the hadronic decay modes $D_s^+(\to \phi\pi^+)D_s^-(\to \phi\pi^-/K^{0*}K^-/\pi^+\pi^-\pi^-)$, 24 signal events were reconstructed. Accounting for the fact that $B_s \to D_s^+ D_s^-$ is fully CP-even and the measurement lacks reconstruction of $B_s \to D_s^{(*)+} D_s^{*-}$ decays, with $\mathcal{B}(B_s \to D_s^+ D_s^-) = 0.0103^{+0.0037}_{-0.0034}$ a lower bound of $\Delta\Gamma_s^{CP}/\Gamma > 0.012$ at 95% C.L. was obtained.

To date, the most recent result published stems from the Belle collaboration [39]. With 23.6 fb^{-1} of data recorded at the Belle experiment running at the $\Upsilon(5S)$ resonance, several hadronic decay channels were used to reconstruct decays of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$. In total, 23 signal events were found. Belle was the first experiment to disentangle the individual decay modes $B_s \to D_s^{(*)+} D_s^{(*)-}$, thus confirming observation of $B_s \to D_s^+ D_s^-$, plus claiming observation of $B_s \to D_s^{*+} D_s^-$ and evidence for $B_s \to D_s^{*+} D_s^{*-}$. Again, assuming $B_s \to D_s^{(*)+} D_s^{(*)-}$ to saturate CP-even final states and taking *CP* violation to be negligibly small, Belle obtained a relative decay width difference of $\Delta\Gamma_s/\Gamma = 0.147^{+0.036}_{-0.030}(stat)^{+0.044}_{-0.042}(syst)$. Table 1.1 gives an overview of the existing measurements described in this section.

Since the way of calculating the world average value of $\Delta\Gamma_s/\Gamma_s$ is not uniquely defined, the *Heavy Flavor Averaging Group* (HFAG) presents different estimations of $\Delta\Gamma_s/\Gamma_s$ depending on the way individual results were measured and combined [40]. According to Table 1.2, which is an excerpt from a table that can be found in the referenced source, depending on the result combination technique, the current world average value of $\Delta\Gamma_s/\Gamma_s$ ranges from about 9% to 15%.

	W/o constraint from $\tau(B_s \to$ flavor specific) nor $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$	W/ constraint from $\tau(B_s \to$ flavor specific) only	W/ constraint from $\tau(B_s \to$ flavor specific) and $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$
$\Delta\Gamma_s/\Gamma_s$ (%)	$15.4^{+6.7}_{-6.5}$	$9.2^{+5.1}_{-5.4}$	$9.3^{+3.2}_{-3.3}$

Table 1.2: Overview of the estimations of $\Delta\Gamma_s/\Gamma_s$ depending on the way of combining individual results [40].

1.7 Aim and Scope of This Analysis

It is the intention of this work to provide new measurements of both the exclusive and the semi-inclusive branching fraction ratios $f_{D_s^{(*)}D_s^{(*)}} = (f_s/f_d)\mathcal{B}(B_s \to B_s^0 \to D_s^{(*)+}D_s^{(*)-}/\mathcal{B}(B^0 \to D^+D_s^-)$. From the ratios the absolute branching fractions of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ can be derived using world average values of f_s/f_d – i.e. the ratio of the s and d quark production fractions at the Tevatron collision energy – and the branching fraction of the normalization decay $B_d^0 \to D^+D_s^-$. Whenever more accurate measurements of f_s/f_d and $\mathcal{B}(B_d^0 \to D^+D_s^-)$ are available in the future, providing $f_{D_s^{(*)}D_s^{(*)}}$ enables a smooth re-calculation of the absolute $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ branching fractions independently of the presented analysis. Depending on the theoretical regime relied upon, the implications of this measurement on the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ will be discussed.

Higher statistics, improved selection techniques and more in-depth considerations concerning the properties of intermediate decays involved will help to improve the accuracy of this measurement. This thesis reports on the steps taken to achieve this ambition. Our studies concentrate on the decay channels $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ and $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$, as well as $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ and $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+K^{*0\dagger}K^-$ as normalization channels[1]. Additional hadronic D_s^+ decay channels might be added in the future to increase statistics.

Apart from measuring branching fractions, in the light of the discussions in Section 1.5.2 the separate reconstruction of $B_s^0 \to D_s^+D_s^-$, $B_s^0 \to D_s^{*+}D_s^-$, and $B_s^0 \to D_s^{*+}D_s^{*-}$ represents a useful preparatory step for later fits to the decay time distributions of these decays. As a decay time fit to $B_s^0 \to D_s^+D_s^-$ directly determines $1/\tau_{CP+} = \Gamma_{CP+}$, and given the current experimental efforts [29, 30] in measuring Γ_{CP-} from $B_s^0 \to J/\psi f_0$, a $B_s^0 \to D_s^+D_s^-$ lifetime analysis is a very interesting option to infer $\Delta\Gamma_s^{CP}$. When it comes to $B_s^0 \to D_s^{*+}D_s^{*-}$, a decay time fit might help to constrain

---

[1] The meaning of the "dagger" the ϕ and K^{*0} resonance are labeled with will be elucidated in Section 4.4

the *CP*-odd fraction in this decay, as described in the guidance given in Ref. [22]. Due to experimental limitations in the reconstruction of $B_s^0 \to D_s^{*+} D_s^{*-}$, in the course of Monte Carlo studies an attempt will be made to infer a potential non-zero *CP*-odd fraction in $B_s^0 \to D_s^{*+} D_s^{*-}$ without relying on lifetime or angular fits to data. Though beyond the scope of this analysis and despite the experimental challenges associated to decay time fits, a lifetime analysis of the used double-charm data sample should however not be impossible to achieve in the near future.

This analysis is carried out using a large set of collision data that was recorded by the CDF II detector at the proton-antiproton collider Tevatron located at the North American Fermi National Accelerator Laboratory. Before detailing the analysis steps, a technical description of both the accelerator and detection apparatus is given first.

2

The Tevatron Collider at Fermilab

2.1 Introduction

To gain knowledge about the nature of the fundamental constituents of matter and the forces they participate in, particle physicists probe the structure of matter at very small scales. Due to the inverse relationship between wave length and energy, resolving smaller scales requires experimental setups to operate at high energies. For a particle physics experiment in which high-energetic particles like protons are brought to collision, this means that physics interactions do not take place at hadron level, but rather among the very elementary particles these compound objects consist of, like quarks and gluons in the case of proton collisions. Through quark or gluon interactions new unstable compound objects are formed, the decay products of which can be used to track their properties. According to Albert Einstein's famous mass-energy relation $m = E/c^2$, the mass of new compound objects that can be produced depends on the kinetic energy of the colliding particles – the more energy available in an inelastic scattering process, the heavier the masses newly formed compound objects can receive. Thus, the role of energy in a particle physics experiment is twofold: On the one hand, it sets the resolution scale on which the sub-structure of matter can be probed, on the other hand it sets the mass scale on which new particles can be produced.

Particle synthesis through inelastic scattering also takes place in the earth atmosphere, when a high energy cosmic ray particle strikes a nitrogen or oxygen nucleus, producing a shower of decay products. As a matter of fact, in the early days of experimental particle physics atmospheric cosmic rays were the only available source of high energy particles, leading to the discovery of muons in 1936 [41] and charged pions in 1947 [42, 43], to name a few, and for a series of important physical disciplines the study of cosmic rays yet provides important insights, particularly when it comes to cosmology, astroparticle and neutrino physics.

Due to the steeply falling differential energy spectrum of cosmic rays [44], events at the highest energies occur at extremely small rates only. Hence, for searches of rare decay processes or heavy and short-living particles of interest today, neither do high energy cosmic rays provide reasonable particle fluxes, nor does the earth atmosphere represent a good and controlled laboratory environment. Therefore, eminent experimental endeavors have been made to design artificial earth-bound particle accelerators capable of colliding particles at high rates and well-defined kinetic energies, and particle detector devices hermetically enclosing the interaction regions.

In the particle physics community, scientific efforts of this kind are commonly embraced under the collective term *Energy Frontier*. The energy frontier is also an integral part of the scientific program [45] of the *Fermi National Accelerator Laboratory* (common abbreviations are FERMILAB or FNAL), a high energy physics laboratory located about 35 miles west of Chicago, Illinois, USA. Fermilab, founded in 1967 as the *National Accelerator Laboratory* and renamed in honor of Enrico Fermi in 1974, is host to a variety of both theoretical and experimental research groups as well as on-site experiments covering a wide range of scientific programs.

Besides various fixed target, muon and neutrino experiments designed to cover the scientific program summarized under the term *intensity frontier*, since 1985 Fermilab's main scientific instrument at the energy frontier is the *Tevatron* collider with its two experiments DØ and CDF. This landmark particle accelerator, the contour of which is well visible from a bird's eye view (see Figure 2.1), lay the ground for groundbreaking scientific findings in high energy physics, like the discovery of the *top* quark in 1995 [46, 47], and still provides important insights into the fields of high p_T and heavy flavor physics. Together with the KEKb high energy accelerator complex in Tsukuba, Japan, and the *Large Hadron Collider* (LHC) at CERN[1] near Geneva, Switzerland, the Tevatron is one out of three major particle accelerators to provide the technical requirements for the production of B_s^0 mesons and the investigation of their properties. In recent years, this enabled Fermilab to make important contributions in the field of $B_s^0 - \bar{B}_s^0$ meson mixing and the search for CP violation in this system.

While circulating in opposite directions around an underground ring about 6.3 km in circumference, inside the Tevatron [48] beams of protons (p) and an-

---

[1] This acronym is derived from the founding committee's name *Conseil Organisation Europeéenne pour la Recherche Nucléare*

tiprotons ($\bar{p}$) are accelerated to an energy of 980 GeV. At two positions, where the general purpose particle detectors DØ and CDF are located, protons and antiprotons collide at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. Until the proton-proton collider LHC set the new world record in particle energy (1.18 TeV per proton beam[2]) in late November 2009, this represented the highest collision energy ever reached in an artificial earth-bound particle accelerator.

The first two periods of Tevatron operation, usually referred to as *Run 0* and *Run I*, began in 1985 with $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV and ended in 1996. During this first 11 year operation period an integrated luminosity[3] of about 140 pb^{-1} was accumulated. During the following years the Tevatron and its experiments DØ and CDF underwent major upgrade work. In 2001, the second operation period *Run II* [49] was initialized, with the $p\bar{p}$ beams now colliding at a center-mass-energy of $\sqrt{s} = 1.96$ TeV at higher and ever increasing instantaneous luminosities. The second operation phase *Run IIb* was just terminated very recently in late September 2011.

2.2 The Tevatron Accelerator Complex

The still remarkable center-of-mass collision energy of 1.96 TeV is the result of gradual energy gains achieved by a sequence of sophisticated accelerator devices [50]. Figure 2.2 gives an schematic overview of the Tevatron accelerator chain.

2.2.1 Proton Pre-Acceleration

The first energy gain within the acceleration process of protons is the result of the production and pre-acceleration of negatively charged hydrogen ions in a static 750 kV high voltage field. This is performed inside the Cockroft-Walton [51] pre-accelerator: First, high-purity hydrogen gas is injected into the volume between a central cylindrical cathode surrounded by an anode. Electrical and magnetic fields inside this cavity set the conditions for the formation of a dense plasma of positively ionized H^+ atoms, that travel towards the cathode and occasionally convert into H^- ions by surface ionization. The cathode surface is coated with cesium to ease electron release. The H^- ions are extracted, accelerated to an energy

---

[2] This corresponds to a center-of-mass energy of $\sqrt{s} = 2.36$ TeV in pp collisions. With a design energy of $\sqrt{s} = 14$ TeV which is intended to be attained by 2013, the LHC is currently running at a collision energy of $\sqrt{s} = 7$ TeV.

[3] Luminosity is a common measure of the particle collision rate and thus of a collider's performance. Please refer to Section 3.1 for details.

Figure 2.1: Aerial photograph of the Fermilab accelerator complex taken from the west. In front of the photograph the contours of the Main Injector tunnel, which as well houses the Recycler, can be seen. In the background the maintenance road on top of the underground Tevatron tunnel is visible. Please note that the real dimensions of the pictured synchrotron rings are distorted by the perspective of the photograph: In reality the Tevatron's circumference is about twice as large as the one of the Main Injector.

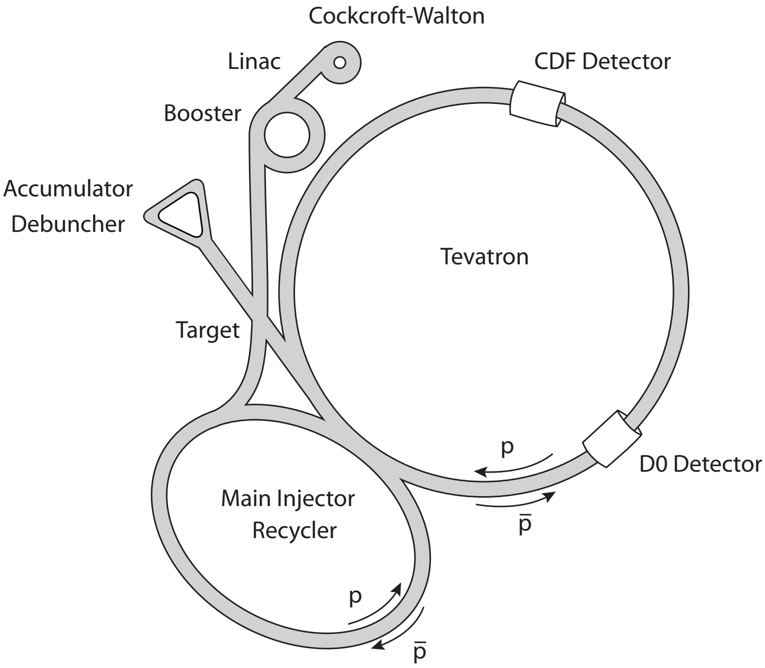


Figure 2.2: Schematic representation of the accelerator complex at Fermilab.

of 750 keV, magnetically focused and then transferred to the *Linac* [52, 53], a linear accelerator approximatively 150 m in length.

The first 75 m long section of the Linac provides a H^- acceleration up to an energy of 116 MeV by means of oscillating electromagnetic fields generated by radio frequency (RF) resonators. While the first, low energy section of the Linac consists of drift tube cavities made of copper, the second, high energy section houses a series of side-coupled cavity modules which again are fed by RF resonator stations. The design of both the drift tubes and the side-coupled cavities, as well as the gaps between them is such that the oscillating fields only have an accelerating effect on the hydrogen ions. As another significant result of the Linac configuration the beam of hydrogen ions is sub-divided into bunches each consisting of about 1.5×10^9 particles. Having reached an energy of 400 MeV, the H^- bunches are sent to the *Booster* [54, 55], a circular accelerator about 475 m in circumference, representing the first synchrotron in the accelerator chain.

The bunches of H^- ions entering – and the remaining ones circulating – the Booster pass through a carbon foil stripping off the electrons from the ions, leaving bare protons only. While circulating the Booster, protons are accelerated using radio frequency cavities arranged along the ring. After about 20,000 revolutions the energy of the protons has increased to 8 GeV.

The final proton acceleration phase outside the Tevatron collider takes place inside the *Main Injector* [56, 57], a 3 km circumference oval synchrotron, that commenced operations in 1999 as a result of the Run II upgrade work. Here, the proton bunches coming from the Booster are coalesced (seven Booster bunches into one bunch) and accelerated to a an energy of 150 GeV using several accelerating and focusing conventional water-cooled dipole and quadrupole electromagnets. These protons are now ready for injection into the Tevatron. Another important operation mode of the Main Injector is to prepare bunches of 120 GeV protons to be sent to the Antiproton Source and to provide protons for fixed target experiments located outside the Tevatron.

2.2.2 Antiproton Production and Buffering

The 120 GeV protons coming from the Main Injector are steered towards the *Antiproton Source* [58], the core element of which is a stack of nickel targets [59]. The particle bombardment on the nickel targets results in the production of a variety of secondary particles, including protons, pions, neutrons and antiprotons, which are focused into a beam by means of lithium lens positioned behind the target. Using a pulsed magnet antiprotons having a mean energy of 8 GeV are isolated from the spray of produced particles.

Antiproton production and storing efficiency is critical to Tevatron performance. However, for every 1 million protons hitting the nickel target, on average 20 antiprotons with an energy of 8 GeV are produced only. Furthermore, because of the bunch configuration of the protons impacting the Antiproton Source, the extracted antiprotons exhibit a large spread in momentum space. Unifying the antiprotons' momentum spectrum is accomplished by an expansion in space, i.e. dispersing the antiproton bunches. Debunching of the antiproton beam is performed inside the *Debuncher* [60], a triangular-shaped synchrotron with a mean diameter of 180 m, using a RF manipulation technique called bunch rotation. By stochastic cooling techniques [61, 62] the 8 GeV antiproton beam is stabilized and remains inside the Debuncher until the next antiproton production cycle starts.

The Debuncher tunnel is host to another important device in the antiproton production and accumulation chain, the *Accumulator* [63], a storage ring which is used to stack the 8 GeV antiprotons coming from the Debuncher. In order to minimize the antiproton loss rate, stochastic cooling is applied to wipe out momentum fluctuations, thus confining the antiproton beam to a smaller volume in phase space. When the cycle of antiproton stacking is complete, i.e. once enough antiprotons have accumulated, the continuous antiproton beam is again segmented into bunches using RF fields and sent to the *Recycler* [56, 64], a ring of permanent magnets that was mounted to the ceiling of the Main Injector tunnel as part of the Run II upgrades.

The role of the Recycler changed over time: As its name implies, its original purpose was to collect and recycle antiprotons left over after a Tevatron *store*. A store is a full Tevatron operation cycle comprising $p\bar{p}$ charging and collision operation until $p\bar{p}$ re-filling. This operation mode turned out to be ineffective, so in the early Run II phase it was decided to dismiss these original plans and to use the Recycler as a means of intermediate storage for antiprotons coming from the Accumulator. It turned out that stashing the antiprotons in the Recycler helped to increase antiproton stacking in the Accumulator and thus the overall antiproton production rate, representing one of the bottlenecks to Tevatron performance. While circulating the Recycler, stochastic and electron cooling techniques [65, 66] are applied, further reducing the antiprotons' spread in momentum space. Electron cooling is the preferred particle cooling technique for higher beam intensities. Once the antiproton beam has reached a certain density, it is overlaid with a 4.3 MeV electron beam along a stretch of about 20 m. The electron beam acts as a momentum damper on the antiprotons: By Coulomb scattering the antiprotons' momentum is transferred to the much lighter electrons, providing an substantial reduction in longitudinal momentum and thus resulting in more compact antiproton bunches. During intermediate storing in the Recycler the antiprotons are kept at an energy of 8 GeV by electrically powered energy correction magnets. As soon as the Tevatron is ready for injection, the antiproton bunches are directed to the Main Injector where they circulate contrariwise to the protons' direction of rotation. Within two seconds they are accelerated to an energy of 150 GeV and are now ready for Tevatron injection.

2.2.3 Tevatron

While *Tevatron* [67] is often used as a pars pro toto term for the full proton-antiproton accelerator complex at Fermilab, the Tevatron in fact is only the final, albeit the largest and most important element of the accelerator chain. Here, protons and antiprotons are accelerated to the final operational beam energy that is almost on the TeV scale – this is where the naming of this circular synchrotron originates from.

The Tevatron does not only provide for the final acceleration, it is primarily designed to serve as a storage ring in order to maintain stable collision conditions over a longer period of time. Both the protons and antiprotons circulate inside the same vacuum beam pipe in opposite directions, while spiralling around each other on helical paths. 774 dipole magnets and multiple correction magnets are used to keep the particles on track inside the slightly curved beam pipe. 240 quadrupole magnets narrow the beam into a thin line, while horizontally and vertically oriented electrostatic separators constrain the protons and antiprotons to helices. To sustain the strong magnetic field of 4.2 T the dipole magnets operate at, these magnets are superconducting with the coils made of niobium-titanium alloy wire. Superconductivity is sustained by cryogenic cooling using liquid helium, keeping the operation temperature at 4.3 K. As a matter of fact, the Tevatron was the world's first synchrotron involving superconducting magnets. At Fermilab, the Tevatron is the only cryogenically cooled accelerator device.

A Tevatron store starts with the *shot setup*, where proton and antiproton bunches are subsequently injected into the Tevatron: First, 36 bunches of 150 GeV protons are extracted from the Main Injector, followed by a ninefold injection of four antiproton bunches. Once the Tevatron is populated by an equal number of 36 proton and antiproton bunches, the *ramping* process starts: This is the final acceleration phase to the operation energy of 980 GeV, which is achieved by eight accelerating RF cavities. Tevatron loading can take up to 45 minutes, while ramping is usually done in a few minutes.

Before the injection helix is switched to collision mode, the beam halo is removed by collimators, since these outer bunch section particles do not significantly contribute to the inelastic collisions, but might potentially cause damage to the beam confinement material or, even worse, to sensitive detector parts. Following the cleaning of the beam environment, a process called *scraping*, the beams are focused to a minimal transverse size. By switching the helices to collision mode the proton and antiprotons are brought to collision at two designated positions of the Tevatron ring: Position B0 where the COLLIDER DETECTOR AT FERMILAB (CDF) is located, and D0 marking the position of the identically named DØ experiment.

At the end of a Tevatron store, which typically allows a continuous data taking period of 10 to 20 hours, or in case of malfunctions, the beams are dumped into absorber material located at Tevatron position A0 by so called abort kicker magnets, making room for a new store. To provide for a safe beam removal, already during shot setup the 36 bunches are grouped into three *bunch trains* with each train separated by 2.6 μs *abort gaps* to give the kicker magnets enough time to ramp up.

2.3 Performance and Acquired Data

In collider physics, the key indicator characterizing a particle accelerator performance is given by the *luminosity* $\mathcal{L}$ [49, 68], which is a measure of the rate of – in case of the Tevatron – proton-antiproton collisions. The luminosity depends on a number of collider-specific operating figures,

$$\mathcal{L} = f \cdot N_b \cdot \frac{N_p N_{\bar{p}}}{2\pi \left(\sigma_p^2 + \sigma_{\bar{p}}^2\right)} \cdot F\left(\frac{\sigma_l}{\beta^*}\right), \tag{2.1}$$

where f is the revolution frequency, N_b is the number of bunches, and N_p and $N_{\bar{p}}$ are the average number of protons and antiprotons, respectively, per bunch. σ_p^2 and $\sigma_{\bar{p}}^2$ denote the spatial bunch widths, while F is a form factor efficiency function accounting for the non-ideal bunch structure at the interaction point. The dimension of $\mathcal{L}$ is that of a particle flux, measured in cm^{-2}s^{-1}. From the instantaneous luminosity the rate of a certain process, e.g. the number of inelastic proton-antiproton interactions per time unit, can be inferred:

$$\dot{N} = \sigma_{in} \mathcal{L}, \tag{2.2}$$

where σ_{in} is the cross-section of inelastic proton-antiproton scattering. The cross-section is a measure of the probability of a given process to occur. Its dimension is that of an geometrical area, commonly quoted in units of 1 barn (b) (1 b = 10^{-28} cm^2).

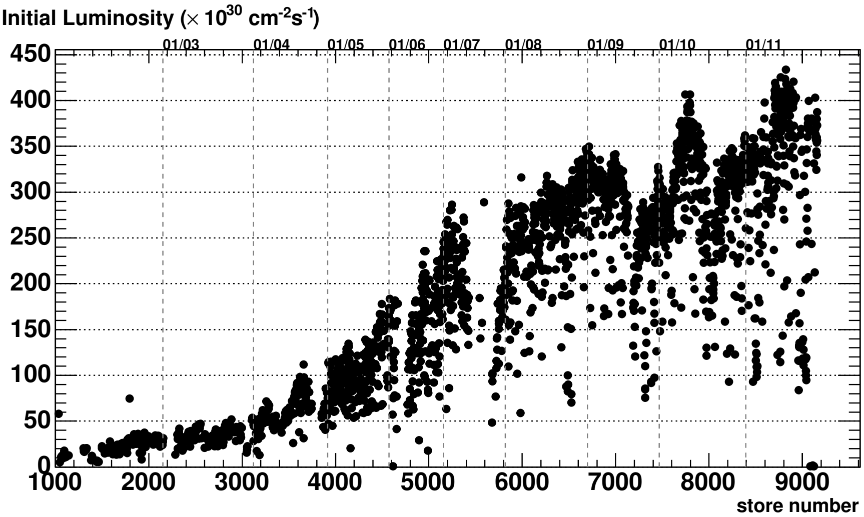


Figure 2.3: Development of the Tevatron Run II initial instantaneous luminosities [69].

Figure 2.3 shows the instantaneous luminosities measured at the beginning of each store since the beginning of Run II up to time of writing. According to the shown statistics the delivered instantaneous luminosity is steadily increasing. However, it took years to reach and exceed the ambitious design goals for Run II [70, 71]. The ever increasing instantaneous luminosities are the result of several technical improvements realized in the course of Run II upgrade work, with some of the measures being the increase of the number of bunches N_b and thus the effective revolution frequency f, and the number of particles $N_{p,\bar{p}}$ per bunch, just to name a few. The key improvement, the increase in antiproton yields, was made possible particularly by increasing the antiproton stacking rate through antiproton stashing in the Recycler.

The amount of data collected over a given period of time is commonly given in terms of the integrated luminosity, $\mathcal{L}_{int} = \int \mathcal{L}dt$, which is preferably quoted in the dimension of an inverse cross-section, b^{-1}. The number of events of a given process contained in a data sample can thus be calculated by $N = \sigma \cdot \mathcal{L}_{int}$. Figure 2.4 shows how the integrated luminosity collected per store developed over time. During the Run II operation period ending in late September 2011, a total of $\mathcal{L}_{int}$ = 11.9 fb^{-1} has been delivered by the Tevatron, whereof CDF managed to record 9.9 fb^{-1}, corresponding to an average data taking efficiency of 83.1%. The highest instantaneous luminosity reached during the Tevatron Run II operation period amounted 4.4×10^{32} cm^{-2}s^{-1}. Table 2.1 summarizes some of the CDF data acquisition facts.

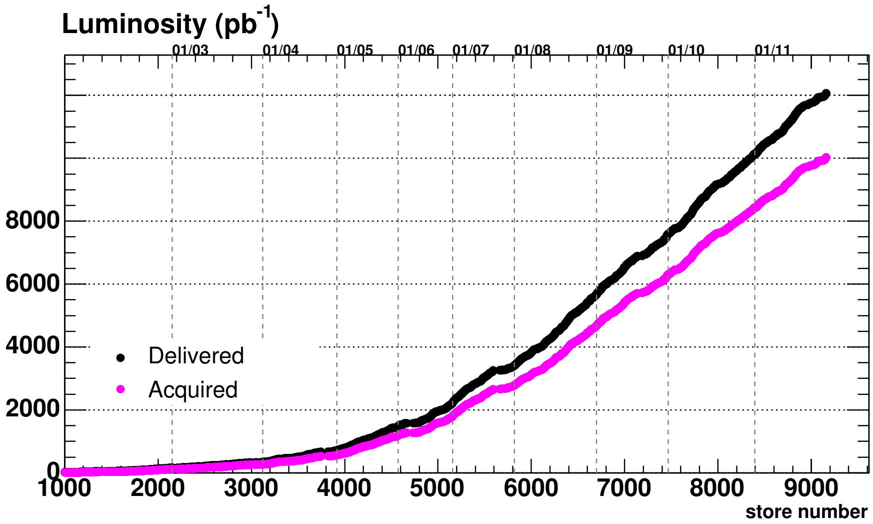


Figure 2.4: Development of the integrated luminosity per store since the beginning of Tevatron Run II [69].

	Time (hr)	$\mathcal{L}$ (10^{30} cm^{-2}s^{-1})	$\mathcal{L}_{int}^{del}$ (pb^{-1})	$\mathcal{L}_{int}^{acq}$ (pb^{-1})	Efficiency (%)
Average	16.5	186.3	4.5	3.7	80.0
Record	62.4	441.9	12.2	10.7	98.5
Sum	44016		11892.3	9885.2	83.1

Table 2.1: Average, record and sum of duration, delivered initial luminosity, delivered integrated luminosity, recorded integrated luminosity and CDF data acquisition efficiency per Tevatron Run II store [69].

3

The Collider Detector at Fermilab

The *Collider Detector at Fermilab* [72, 73], which first commenced operations in 1985, is a multipurpose collider detector designed for covering a broad range of high energy physics regimes. In the course of the Tevatron Run II upgrades it underwent major upgrade work, primarily to prepare for higher luminosity operations as well as to enhance particle tracking and identification. Since then, the Collider Detector at Fermilab is commonly abbreviated as CDF II.

Typical to classical collider detectors, the CDF II detector exhibits a azimuthal and forward-backward symmetry with most of the sub-detector systems arranged in layers coaxially around the beam pipe. Due to symmetric proton-antiproton head-on collisions (both the protons and the antiprotons have an energy of 980 GeV), on average the spatial distribution of the collision products is expected to be forward-backward symmetric – this is why the detector also features a forward-backward symmetry. With a length of 12 m and a radius of 6 m, CDF II weighs around $4,500$ tons. Figure 3.1 shows a photograph of CDF II taken during Run II upgrade installations carried in the CDF assembly hall, located next to the detector's nominal operational position at Tevatron ring coordinate B0, eight meters below the surface.

In global terms, from the inside out the detector is composed of six main detector components arranged around the beam pipe: A tracking system (1) comprising the Silicon Detector and the Central Outer Tracker, a Time-of-Flight System (2) enclosed by a superconducting solenoid magnet (3) generating a 1.4 T field orientated parallel to the beam pipe, electromagnetic calorimeters (4), hadronic calorimeters (5), and finally a multi-component muon detector system (6). A schematic depiction of CDF II is given in Figure 3.2.

According to the detector's symmetries, in the following brief discussion of the detector components an admixture of a polar and cylindrical coordinate

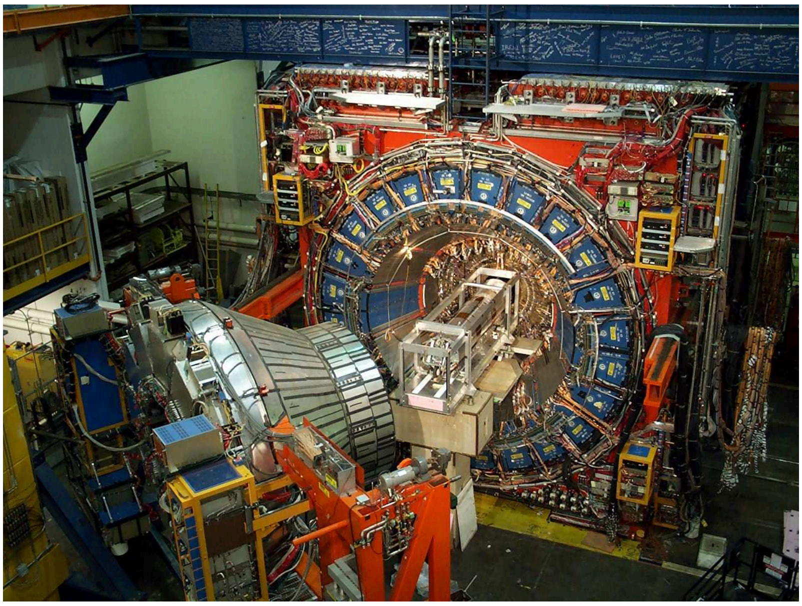


Figure 3.1: The CDF II detector. The photograph shows the installation of the new tracking system as part of the Run II upgrade work.

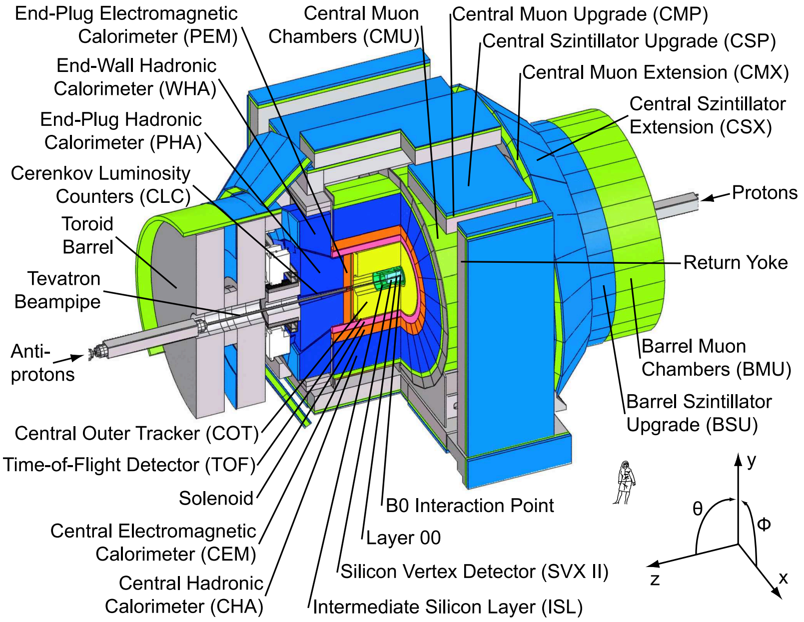


Figure 3.2: Cutaway view of the detector showing its sub-detector devices [74, 75].

system, indicated in the bottom right of Figure 3.2, is used:

- The nominal $p\bar{p}$ interaction position marks the point of origin of the coordinate system
- The proton beam direction determines the positive direction in z.
- r measures the distance from the beamline.
- The azimuthal angle ϕ is measured upwards within the x-y-plane where the x axis lies in the accelerator plane, radially directing outwards off the center of the ring; the Cartesian coordinate y points upright with respect to the accelerator plane.
- θ represents the polar angle lying in the y-z-plane.

In the field of experimental particle physics the *pseudorapidity* η, which is an approximation of the Lorentz invariant rapidity $\Theta = \tanh^{-1}(v/c)$ for rest masses $m_0 \ll p$, is commonly used: $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$. Moreover, in later discussions the physical quantities *transverse energy* E_T and *transverse momentum* p_T will be used frequently. These quantities describe the transverse component of a particle's energy or momentum, defined as $E_T = E \cdot sin\theta$ and $p_T = p \cdot sin\theta$, respectively.

3.1 Luminosity Counter

Before outlining the individual detector apparatus, the *Cerenkov Luminosity Counter* (CLC) [76] is briefly introduced. This device is needed to estimate the rate at which protons and antiprotons inelastically collide while crossing interaction point B0. As introduced in Section this rate is expressed in terms of the instantaneous luminosity $\mathcal{L}$.

CDF has two CLCs, each mounted around the beam pipe in the end-plug regions, covering the forward-backward region $3.7 \leq |\eta| \leq 4.7$. Each CLC is made of 48 isobutane Cerenkov light detectors that collect the burst of Cerenkov light generated by a particle travelling through the gaseous medium faster than the speed of light. From the rate the CLCs fire and the bunch crossing rate f the instantaneous luminosity $\mathcal{L}$ can then be calculated with an accuracy of about 6%.

3.2 Tracking System

The precise construction of charged particle tracks is an outstanding feature of the CDF II detector: Various kinds of analyses at CDF benefit from the high primary and secondary vertex reconstruction resolution, as well as the good momentum resolution – so does the present work. The momentum of a charged particle can be inferred by measuring the bending of its track inside the 1.4 T magnetic field. Offline reconstruction of particle tracks is performed by employing tracking algorithms provided by the CDF II offline software. In principle, track reconstruction works by extrapolating hits – i.e. spots in the detector material which have an energy deposit caused by an ionizing particle passing through – found in a certain detector component and finding matches in detector material of a different tracking unit.

The CDF II tracking system consists of two main detector apparatus: A *Silicon Microstrip Detector* providing high resolution tracking in the immediate vicinity of the interaction point, where the highest track densities are to be expected, and the *Central Outer Tracker* (COT), a drift chamber surrounding the silicon device.

3.2.1 Silicon Detector

From the inside out, the Silicon Detector comprises three sub-systems: The *Layer 00* (L00) [77] which is directly mounted onto the beam pipe, the *Silicon Vertex Detector* (SVX II) [74], and the *Intermediate Silicon Layers* (ISL) [78]. The CDF Tracker unit as well as the dimensions and η coverages of its sub-systems are illustrated in Figure 3.3.

The innermost detector, *L00*, is made of two overlapping layers of radiation-hard silicon microstrips positioned at radii 1.35 and 1.62 cm, each consisting of twelve sensor modules. With a total length of 94 cm in z direction, L00 provides a pseudorapidity coverage of $|\eta| \leq 4$. L00 was installed as part of the Run II upgrade and due to its vicinity to the primary interaction point it significantly contributed in improving the precision of track measurements and impact parameter resolution.

SVX II is a double-sided micro strip detector ranging from $r = 2.1$ cm to $r = 17.3$ cm. It basically consists of three barrel like devices, each of them 29 cm in length. The overall geometry of SVX II allows for a track reconstruction with a maximum pseudorapidity of $|\eta| = 2$. Due to the special layout of the five layers inside each device, a high impact parameter resolution can be achieved: On one side of each of the layers the silicon microstrips are axially aligned for a precise measurement of the r and ϕ coordinate of a particle track. In addition,

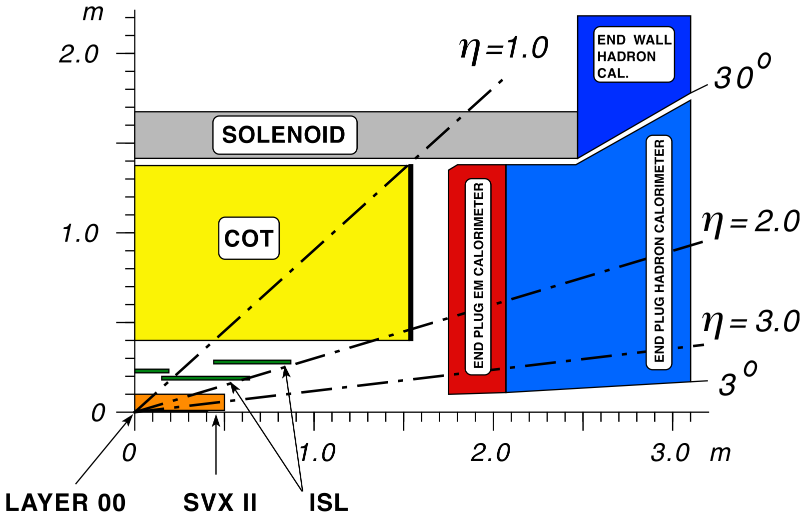


Figure 3.3: The components of the CDF II tracking apparatus. On the right hand side of the depiction three of the five end-plug calorimeters can be seen [79].

layers 0, 1 and 3 feature orthogonally attached stereo strips on the back side, supplementing the $r - \phi$ measurement with a z-measurement, whereas the strips on the back side of layers 2 and 4 are tilted by 1.2 degrees (SAS, Small Angle Stereo), thereby avoiding combinatorial ambiguities introduced by multiple hit events in the Silicon Detector and thus providing unique three-dimensional information.

The outermost component of the silicon detector system is the 1.9 m long *Intermediate Silicon Layer*. The main function of the ISL is to provide precise three-dimensional track information and to act as a link between SVX II and COT in later offline track reconstruction, where hits detected in the Silicon Detector and the COT are matched to each other. It consists of three double-sided silicon microstrip layers having the same SAS design as layers 2 and 4 of SVX II. One layer positioned in the central region at r = 22 cm covers $\left|\eta\right| < 1$, whereas two of the three layers are located at r = 20 cm and r = 28 cm covering the forward/backward region $(1 < \left|\eta\right| < 2)$. Figure 3.4 shows an schematic illustration of the Silicon Detector in the $r - \phi$ plane.

3.2.2 Drift Chamber

The precise spatial information provided by the Silicon Detector system is complemented by the excellent momentum resolution achieved with the *Central Outer Tracker* (COT) [80], an open-cell wire drift chamber surrounding the silicon devices. The active gas volume of the 3.1 m long cylindrical device radially ranges from 43.4 cm to 132.3 cm, resulting in a pseudorapidity coverage of $\eta \leq 1$. The drift chambers are filled by an admixture of argon and ethane in equal shares. A

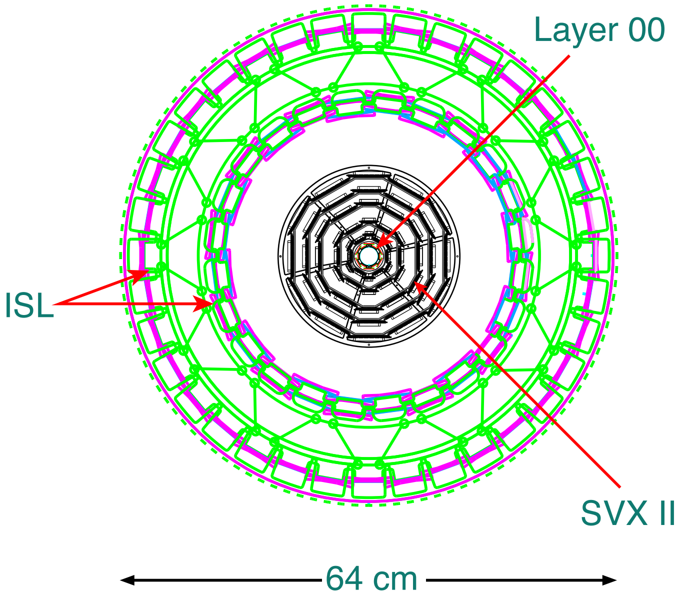


Figure 3.4: $r-\phi$ view of the CDF II Silicon Detector [74].

charged particle passing through the chamber ionizes gas molecules it encounters on its flight path. An electronic drift field causes the electron signal to amplify by secondary interactions and to drift towards sense wires where the charge deposition is detected. The COT is designed to have drift times not exceeding 100 ns, in order to handle the Run II luminosities.

The COT wire system is segmented into eight coaxially arranged *super-layers* with equal numbers of alternating axial and stereo angle super-layers. The axial super-layers responsible for measurements in the $r-\phi$ plane have wires parallel to the beam line, whereas the wires of the other four super-layers, providing z coordinate information, are tilted by an stereo angle of 2 degrees with respect to the z axis. Each super-layer in turn consists of 13 potential wires, shaping the electrical drift field, and 12 sense wires. By this configuration, the COT is capable of providing a maximum of 96 position measurements per ionizing particle.

The spatial resolution of the COT is poor compared to one delivered by the Silicon Detector. However, thanks to its large volume and the intrinsic low track densities it yields high precision in measurements of transverse momenta. The momentum resolution is approximately $\sigma(p_T)/p_T^2 = 0.0015$ (GeV/c)$^{-1}$.

In addition, the COT provides information about a particle's specific energy loss dE/dx which is proportional to the logarithm of the charge Q deposited in the detector's active volume, which in turn is proportional to the readout pulse width Δt. By using Bethe's stopping power formula [81], for a given $\beta\gamma = p/m$ the dE/dx information can be used to infer a particle's mass. The Kaon-Pion discrimination is shown in Figure 3.5 as the dashed line. The energy loss discrimination power

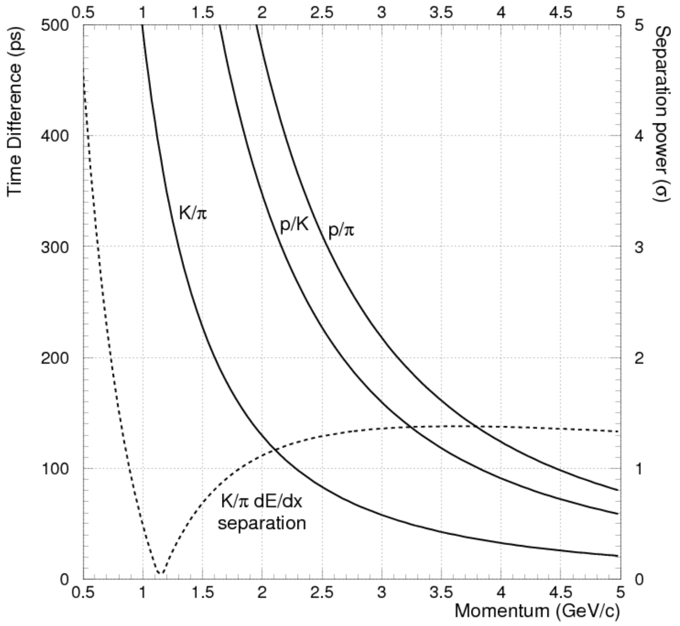


Figure 3.5: Discrimination power for kaons, pions, and protons using TOF (solid lines) and COT dE/dx information (dashed line) as a function of particle momenta [82].

is above 1σ for momenta larger than 1.7 GeV/c.

3.3 Time of Flight Detector

Another important input to particle identification is given by the measurement of the time it takes a particle to travel from the production point to the *Time of Flight* (TOF) [83] detector. The flight duration is calculated as the difference between the time of arrival at the TOF detector and its production time, $t = t_{TOF} - t_0$. Using the COT momentum measurement derived from the reconstructed track curvature, a particle's mass can be inferred by

$$m = \frac{p}{c}\sqrt{\frac{(ct)^2}{L^2} - 1}, \tag{3.1}$$

where L is the length of the flight path. The TOF particle separation power is high ($> 2\sigma$) for low momenta ($p \lesssim 1.6$ GeV/c) and declines for increasing momenta (see Figure 3.5).

The Time of Flight Detector is located at $r = 140$ cm just between the COT and the cryostat of the super-conducting solenoid magnet. It consists of 216 plastic scintillator bars 279 cm in length. Each scintillator has a photomultiplier attached for signal readout. The TOF system provides a time resolution of ~ 100 ps.

3.4 Calorimetry

The CDF II calorimeter system [84] consists of several segmented detector units comprising alternating layers of scintillators and absorbing material. It is located outside the solenoid and in the forward-backward end wall regions. It is dedicated for measuring energy deposits of electromagnetic and hadronic particles or particle showers, respectively, by fully absorbing all particles except neutrinos and muons. Since it is primarily designed for particles or jets with high transverse momenta ($p_T > 20\,\mathrm{GeV/c}$) the calorimeter system provides important information in the regime of high p_T physics and is therefore of major importance for analyses in the *top* quark physics sector, for instance. At CDF, b physics analyses do usually not rely on calorimetry – nor does the present thesis – except for studies of leptonic decay channels where electron or muon identification is needed. In the latter case one takes advantage of the screening effect the calorimeter system has on all particles except for muons. A description of the electromagnetic and hadronic calorimetry systems can be found in the Refs. [85, 86].

3.5 Muon Detector System

The muon system represents the outermost CDF detector device. It comprises four systems consisting of several multi- or single-wire proportional chambers filled with argon and ethane at equal shares with scintillators attached for signal readout. Because of the decay channels subject to this analysis, the muon detectors do not add any information needed throughout the given analysis. They do however provide important input for many other b physics studies of decay channels having muons in the final state. The interested reader may refer to Refs. [87, 88, 89, 90] for a detailed description of the muon detector and counter components.

3.6 Event Recording

With an effective Tevatron bunch crossing rate of 1.7 MHz and a required per-event detector readout storage space of about 250 kB it is neither possible nor desirable to record every collision event. A hadron collider yields a higher $b\bar{b}$ production cross-section compared to an electron machine, there is however an overwhelming amount of collision events not worth of being recorded: At the

Tevatron center-of-mass energy, the $b\bar{b}$ production cross-section is three orders of magnitude lower than the total inelastic $p\bar{p}$ cross section. $b\bar{b}$ events are therefore overlayed by very high noise level. For this reason, CDF II has an online system implemented that automatically triggers event recording if an observed detector signature matches one or multiple pre-set trigger schemes according to different physics regimes. The CDF *Data Acquisition System* (DAQ) [91], illustrated in Figure 3.6, is a successive three-level trigger system [92, 93] with Level 1 being entirely hardware-based, while Level 2 is a hardware-software implementation. The third trigger stage is implemented in software running on a Linux computer farm. By the three-level implementation of the DAQ the event recording rate is reduced to $\sim$ 100 Hz. Depending on the trigger scheme an event has fired, recorded data is split up into eight different streams and written to tape. A fraction of data is copied out for online monitoring. Based on certain quality requirements the CDF Control Room crew assesses data taking on a regular basis and marks a *run* – the smallest unit of data taking – bad in case of detector or data acquisition malfunctions.

All data used in this analysis were recorded by a trigger scheme called the *Two-Track Trigger* (TTT). The TTT basically selects events containing a track pair having a large displacement from the primary interaction vertex. In this sense, the TTT is purpose-built for finding detector signatures typical to decay chains of unstable, but relatively long-living particles like bottom or charm mesons that decay hadronically into kaons and pions in particular. Sometimes, the Two-Track Trigger is also referred to as “Two Displaced Tracks” Trigger or “Secondary Vertex” Trigger. Accounting for the requirements of the TTT, in the following section the signal processing of the CDF DAQ is outlined in a nutshell.

3.6.1 Two Track Trigger

The first trigger level, **Level 1 Trigger (L1)**, runs synchronous to the clock cycle and seeks for patterns of raw physics objects based on information provided by a set of sub-detectors, whereas in case of the Two-Track Trigger COT information is used only. The fixed latency time amounts 5.5 μs. In addition to the synchronous one clock L1 pipeline, digitized detector information is fed into a 42 clock cycle pipeline for further readout and for later use on the next trigger stage. Already at the early L1 stage, approximate COT tracking information with p_T and ϕ values provided by the *eXtremely Fast Tracker* (XFT) [95] is available for decision making. The criteria for TTT L1 acceptance are at least two oppositely charged COT tracks with an opening angle of $\Delta\phi^{(1,2)} < 135\ ^{\circ}$ and each track having a transverse momentum larger than 2 GeV/c.

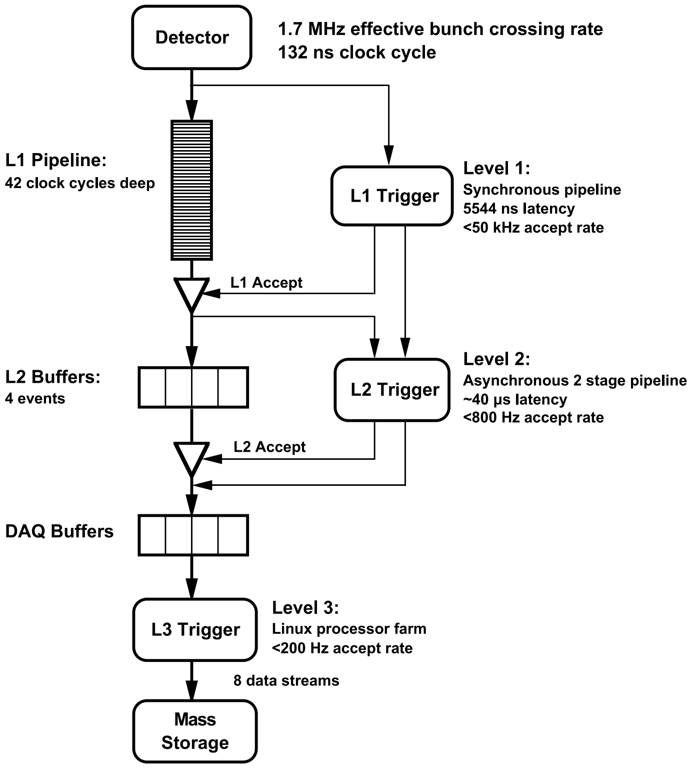


Figure 3.6: Flow chart illustrating the signal processing of the three-level data acquisition system at CDF II [94].

Level 2 Trigger (L2) is an asynchronous two-stage pipeline with a latency time of 40 μs. The longer time available for L2 decision making is obtained by keeping up to four events in an onboard buffer. The key function of the second trigger level is to link silicon tracking information provided by the *Silicon Vertex Trigger* (SVT) [96, 97] with XFT tracks. This functionality is of major importance in the heavy flavor physics regime where decays involving long-living b, c and s mesons are analyzed, because already at trigger stage events containing secondary vertices can be identified. The larger latency allows for executing vertex fits and calculating more complex track quantities like the impact parameter d_0 and decay length L_{xy}. The Level 2 Trigger requires the oppositely charged SVT tracks to have a transverse momentum of at least 2 GeV/c, to match the XFT tracks and to meet the vertex fit quality criterion of $\chi^2 < 25$. At this trigger stage, further pass requirements depend on particular sub-schemes that are implemented by the Level 2 sub-paths L2_B_LOWPT, L2_B_CHARM, and L2_B_HIGHPT:

- **L2_B_LOWPT**: The matched SVT tracks must have a minimum scalar transverse momentum sum of $p_T^{(1)} + p_T^{(2)} > 4.0$ GeV/c, an opening angle in the range $2° < \Delta\phi^{(1,2)} < 90°$, an impact parameter of $0.1\ \text{mm} < |d_0| < 1$ mm, and a decay length in the xy plane of $L_{xy} > 0.1$ mm.

- **L2_B_CHARM**: The matched SVT tracks are required to have a minimum scalar transverse momentum sum of $p_T^{(1)} + p_T^{(2)} > 5.5$ GeV/c, an opening angle in the range $2° < \Delta\phi^{(1,2)} < 135°$, an impact parameter of 0.1 mm$< |d_0| < 1$ mm, and transverse travel distance of $L_{xy} > 0.1$ mm.
- **L2_B_HIGHPT**: The requirements in this scenario are the same as for *L2_B_ CHARM*, except for a higher transverse momentum cut of $p_T^{(1)} + p_T^{(2)} > 6.5$ GeV/c.

All events accepted by **L1** and **L2** are kept in the DAQ buffers before entering **Level 3 Trigger (L3)**. Using software-implemented algorithms similar to those of the offline reconstruction software a full event reconstruction is performed. In doing so, decisions of the previous levels can be re-confirmed at a higher precision using an event's full decay topology.

At trigger level, the amount of recorded data can be artificially scaled down by a dynamically adjusted quantity p, called *prescale factor*. This is especially useful during periods of high luminosity, since otherwise certain trigger paths would fire perpetually, consuming a lot of bandwidth. Prescaling of the trigger paths means that only one out of $N = 1/p$ events having met the trigger requirements is accepted. This procedure frees bandwidth now available to other trigger paths by decreasing the deadtime of the trigger system. The prescale factors are stored in a database to be accessible in offline software.

The Two-Track Trigger is a prominent example of a trigger that is scaled down at high luminosities to allocate more bandwidth to triggers aimed for high p_T physics, like top quark physics or Higgs searches. Over the years, this began to negatively impact the gain in net Two-Track Trigger data acquired, since the Tevatron operates at by far higher instantaneous luminosities than in the early days of Run II. Tevatron stores are often dropped before reaching the low luminosity regime in which the Two-Track Trigger is not prescaled. The amount of recorded Two-Track Trigger data does therefore only scale well with integrated luminosity for the initial Run II operation periods, while for later periods an increase in integrated luminosity does not add that much of Two-Track Trigger data.

4

Preparation of Data Samples

4.1 Data Base

This study uses hadronic Two-Track Trigger data gathered between February 2002 and August 2010, corresponding to CDF II operation periods 0 to 31 (ending at run number 294777). Used data comprise generic hadronic streams extracted from the CDF SAM [98] database sets *xbhdid* (p0), *xbhdih* (p1-4), *xbhdii* (p5-10), *xbhdij* (p11-13), *xbhdik* (p14-17) *xbhdfm* (p18-28), and *xbhdfp* (p29-31). The Two-Track Trigger (see previous section) is an online run selection algorithm that triggers recording of detector events with at least two oppositely charged particle tracks, whose common decay vertex has a minimum transverse displacement from the primary vertex. Additional requirements are placed on the tracks' minimum transverse momenta and opening angles. The hadronic modes studied throughout this analysis were collected by selecting events having passed the trigger sub-paths *B_CHARM_LOWPT*, *B_CHARM*, and *B_CHARM_HIGHPT*. Total available statistics are equivalent to an integrated luminosity of 6.8 fb^{-1}.

Throughout this thesis, three different data sub-samples will sometimes be referred to. These correspond to different exclusive trigger configurations of data taking and are defined in the following way:

- **TriggerFlag 1 (T1)**: Candidates selected by *B_CHARM_LOWPT* but not by *B_CHARM* or *B_CHARM_HIGHPT*.

- **TriggerFlag 2 (T2)**: Candidates selected by *B_CHARM* but not by *B_CHARM_HIGHPT*.

- **TriggerFlag 3 (T3)**: Candidates selected by *B_CHARM_HIGHPT* only.

4.2 Offline Reconstruction

Since the total number of B_s^0 mesons produced at the Tevatron is a priori not known, the branching fractions $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ cannot be measured directly, but in ratio to a normalization channel only. In doing so, the unknown number of produced B_s^0 mesons cancels out. Due to similar decay topologies the decay $B_d^0 \to D^+ D_s^-$ is chosen for this purpose. Technically – in terms of offline reconstruction code – D_s^- mesons are constructed from $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$, respectively[1], while the D^+ meson is reconstructed from its decay into one charged kaon and two equally charged pions ($D^+ \to K^-\pi^+\pi^+$).

Offline reconstruction of the studied decay modes is performed from the bottom up. Reconstruction of a B_s^0 candidate in the decay mode $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ ($B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$) starts with the selection of a ϕ (K^{*0}) candidate. For this, in a first step two oppositely charged particle tracks assumed to be kaons (kaon and pion) are combined. To reject a large amount of combinatorial background events not stemming from a real ϕ (K^{*0}), a full vertex fit is performed and from that a candidate's invariant mass based on the sum of the four momenta of the kaon (kaon and pion) track candidates is calculated. Since the computation needed for a full vertex fit is rather time-consuming, prior to performing the vertex fit a soft pre-selection on the estimated raw invariant mass of the track pair is applied. A selected track pair is accepted as a ϕ (K^{*0}) candidate if the vertex fit has succeeded and the invariant mass lies in the mass window $1.005\ \text{GeV/c}^2 < M_{K^+K^-} < 1.035\ \text{GeV/c}^2$ ($0.837\ \text{GeV/c}^2 < M_{K^-\pi^+} < 0.947\ \text{GeV/c}^2$). By adding a further charged track which is assumed to be a pion (kaon), a D_s^+ candidate is formed. In the following vertex fit the tracks belonging to the selected candidates are required to come from one common vertex. $B_s^0 \to D_s^+ D_s^-$ candidates are accepted if the vertex fit has succeeded and the D_s^+ invariant lies within 1.87 and 2.07 GeV/c^2, corresponding to a mass window of ±100 MeV/c^2 around the D_s mass of 1.968 GeV/c^2 published by the *Particle Data Group* (PDG) [3]. When performing the B_s^0 vertex fit the D_s^+ mass is constrained to its PDG value. Tables 4.1 and 4.2 summarize the requirements placed on several kinematic and fit quality variables. At this stage, in addition B_s^0 candidates are required to fulfill the specifications of the Two-Track Trigger and the requirements of the *B_CHARM_LOWPT*, *B_CHARM*, and *B_CHARM_HIGHPT* trigger sub-paths. Reconstructed decay chains are stored in a hierarchical tree structure, which contains all the kinematical and other detector or reconstruction related quantities. Due their *tuple*-like structure, at CDF data files holding information about B meson decays are called *BStntuples*.

---

[1] As later discussions will show, this is only correct in technical respects. To correctly account for the full underlying decay dynamics this statement needs to be put under scrutiny.

$B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$			$B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$		
$3.0 <$	$m(B_s^0)$	< 6.6 GeV/c^2	$3.0 <$	$m(B_d^0)$	< 6.6 GeV/c^2
$1.870 <$	$m(D_s^{(1)})$	< 2.070 GeV/c^2	$1.870 <$	$m(D_s^{(1)})$	< 2.070 GeV/c^2
$1.870 <$	$m(D_s^{(2)})$	< 2.070 GeV/c^2	$1.770 <$	$m(D)$	< 1.970 GeV/c^2
$1.005 <$	$m(\phi^{(1,2)})$	< 1.035 GeV/c^2	$1.005 <$	$m(\phi)$	< 1.035 GeV/c^2
	$\chi^2_{r\phi}(B_s^0)$	< 40		$\chi^2_{r\phi}(B_d^0)$	< 40
	$\chi^2_{r\phi}(D_s^{(1)})$	< 40		$\chi^2_{r\phi}(D_s)$	< 40
	$\chi^2_{r\phi}(D_s^{(2)})$	< 40		$\chi^2_{r\phi}(D)$	< 20
	$\chi^2_{r\phi}(\phi^{(1,2)})$	< 15		$\chi^2_{r\phi}(\phi)$	< 15
				$L_{xy}/\sigma_{L_{xy}}(D_s)$	> 3.0
				$L_{xy}/\sigma_{L_{xy}}(D)$	> 3.0
	$\Delta Z_0(\phi^{(1,2)})$	< 1.5 cm		$\Delta Z_0(\phi)$	< 1.5 cm

Table 4.1: Requirements the CDF offline reconstruction software places when selecting the decays $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ (left column) and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ (right column). See Section A.1 in the Appendix for definitions of variables.

From the BStntuples, which do not only contain the described decays, but a variety of reconstructed hadronic channels, *flat* Ntuples are generated using the *BottomMods* library (Version 6.1.4) and the most recent hadronic goodrun lists, energy loss corrections and COT calibrations. This process is called *skimming* in technical terminology. The *skimmed* or *flat* Ntuples are a streamlined version of the BStntuples and contain the hadronic decay modes of interest only. The file format of flat Ntuples is that of the *ROOT* software [99], an object-orientated data analysis framework that is widely used in the particle physics community.

In addition to experimental data, realistic simulations reflecting the physics behavior of true B_s^0 and B_d^0 events in the studied decay modes are an essential input. Simulated data, that are generated by means of Monte Carlo (MC) techniques, are needed for various steps in the analysis. Before describing Monte Carlo production in technical respects, some physics-related issues that need to be addressed prior to simulation are pointed out.

4.3 Additional Considerations on B Meson Lifetimes

The B_s^0 and B_d^0 Monte Carlo samples used in this analysis were generated with mean decay lengths of $c\tau_{B_s^0} = 441\ \mu\text{m}$ ($\tau = 1.471$ ps) and $c\tau_{B_d^0} = 458.7\ \mu\text{m}$ ($\tau = 1.530$

$B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$			$B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0} K^-$		
$3.0 <$	$m(B_s^0)$	< 6.6 GeV/c^2	$3.0 <$	$m(B_d^0)$	< 6.6 GeV/c^2
$1.870 <$	$m(D_s^{(1)})$	< 2.070 GeV/c^2	$1.870 <$	$m(D_s^{(1)})$	< 2.070 GeV/c^2
$1.870 <$	$m(D_s^{(2)})$	< 2.070 GeV/c^2	$1.770 <$	$m(D)$	< 1.970 GeV/c^2
$1.005 <$	$m(\phi)$	< 1.035 GeV/c^2			
$0.837 <$	$m(K^{*0})$	< 0.947 GeV/c^2	$0.837 <$	$m(K^{*0})$	< 0.947 GeV/c^2
	$\chi^2_{r\phi}(B_s^0)$	< 40		$\chi^2_{r\phi}(B_d^0)$	< 40
	$\chi^2_{r\phi}(D_s^{(1)})$	< 40		$\chi^2_{r\phi}(D_s)$	< 20
	$\chi^2_{r\phi}(D_s^{(2)})$	< 20		$\chi^2_{r\phi}(D)$	< 20
	$\chi^2_{r\phi}(\phi)$	< 15			
	$\chi^2_{r\phi}(K^{*0})$	< 15		$\chi^2_{r\phi}(K^{*0})$	< 15
	$L_{xy}/\sigma_{L_{xy}}(D_s^{(1)})$	> 3.0		$L_{xy}/\sigma_{L_{xy}}(D_s)$	> 3.0
	$L_{xy}/\sigma_{L_{xy}}(D_s^{(2)})$	> 3.0		$L_{xy}/\sigma_{L_{xy}}(D)$	> 3.0
	$L_{xy}/\sigma_{L_{xy}}(K^{*0})$	> 2.0		$L_{xy}/\sigma_{L_{xy}}(K^{*0})$	> 2.0
	$\Delta Z_0(\phi)$	< 1.5 cm			
	$\Delta Z_0(K^{*0})$	< 1.5 cm		$\Delta Z_0(K^{*0})$	< 1.5 cm

Table 4.2: Offline reconstruction requirements placed on the studied decay channels with one D_s^- decaying into $K^{*0}K^+$.

ps), respectively. These values correspond to the state of knowledge of the year 2008 [100]. For the B_d^0 meson, the current edition of the Review of Particles Physics [3] quotes a slightly smaller world average value of

$$\begin{aligned} \tau_{B_d^0} &= 1.519 \pm 0.007 \text{ ps} \\ c\tau_{B_d^0} &= 455.4 \pm 2.1\ \mu\text{m} \end{aligned} \tag{4.1}$$

Therefore, all B_d^0 Monte Carlo samples are re-weighted using the central value of the more recent result (4.1). Re-weighting of Monte Carlo is done according to the relation

$$w_{c\tau} = \exp\left(\frac{ct}{c\tau_{MC}} - \frac{ct}{c\tau}\right) \tag{4.2}$$

where $c\tau_{MC}$ denotes the mean decay length Monte Carlo was originally generated with, while $c\tau$ is the new mean decay length simulated events shall be re-weighted with. ct is the per-event proper decay length at generator level. In the context of Monte Carlo simulation, generator level quantities are usually referred to as *Monte Carlo Truth* quantities.

For the B_s^0 meson the current edition of Ref. [3] quotes

$$\begin{aligned} \tau_{B_s^0} &= 1.472^{+0.024}_{-0.026} \text{ ps} \\ c\tau_{B_s^0} &= 441^{+7}_{-8}\ \mu\text{m} \end{aligned} \tag{4.3}$$

as mean lifetime. This is identical to the input value used in simulation. However, the attentive reader may recall that $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ is believed to be predominantly CP-even. This has been the pre-condition for relating the branching fraction of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ to the relative decay width difference $\Delta\Gamma_s/\Gamma_s$, as described in the theoretical introduction, Section 1.5.1. Given the Standard Model expectation of a vanishing CP-violating phase ϕ_s, the CP-even state B_s^{CP+} coincides with the light mass eigenstate B_s^L. According to the *Heavy Flavor Averaging Group* (HFAG) [40] the mean lifetime of the short-living light eigenstate amounts

$$\begin{aligned} \tau_{B_s^L} &= 1.408^{+0.033}_{-0.030} \text{ ps} \\ c\tau_{B_s^L} &= 422^{+10}_{-9}\ \mu\text{m} \end{aligned} \tag{4.4}$$

In order to bring the $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ Monte Carlo sample in line with the CP assumption made, simulated data is re-weighted using the decay length of the light eigenstate, $c\tau_{B_s^L}$ (4.4). Changes in the both the B_d^0 and B_s^0 lifetime assumptions will be subject to systematic studies (Section 7.3.1).

4.4 Additional Considerations on $D_s^+ \to \phi\pi^+$ and $D_s^+ \to K^{*0}K^+$

As described in Section 4.2, reconstruction of a D_s^+ meson candidate starts with a ϕ (K^{*0}) candidate which is formed from two oppositely charged kaons (one kaon and one pion of opposite charge). In doing so a large amount of random kaon and pion tracks are being picked up in reconstruction, leading to a huge amount of combinatorial background events. However, methods will be put in place to effectively separate signal meson events from combinatorics (Section 5).

There is yet an additional issue that needs to be addressed: The decays $D_s^+ \to \phi\pi^+$ and $D_s^+ \to K^{*0}K^+$ are not the only possibilities for a D_s^+ meson to proceed into the final state $K^+K^-\pi^+$. In fact, D meson three-body decays are expected to proceed through a variety of resonant two-body decays, where, in our example, ϕ and K^{*0} only represent two possible resonances among other intermediate states that are kinematically allowed for $D_s^+ \to K^+K^-\pi^+$. Conversely, when forming a ϕ candidate according to the offline reconstruction algorithm, there is a non-vanishing chance for any other intermediate state fulfilling the invariant mass requirement $1.005\ \mathrm{GeV/c^2} < M_{K^+K^-} < 1.035\ \mathrm{GeV/c^2}$ to be falsely reconstructed as a ϕ. The same arguments hold for the reconstruction of K^{*0} mesons.

In the end, no attempt is made to identify the $\phi\pi^+$ or the $K^{*0}K^+$ component as such; any D_s^+ meson reconstructed from the final state $K^+K^-\pi^+$, where a K^+K^- ($K^-\pi^+$) pair and the D_s^+ meson itself fulfill the invariant mass and vertex fit quality requirements, is accepted as a D_s^+ candidate. However, when it comes to the extraction of observables (particularly relative branching fractions) from data, one has to take into account that through the reconstruction technique one reconstructs $\phi\pi^+$ ($K^{*0}K^+$) *plus* additional contributions. It is therefore not correct to use the published [3] final state branching fractions of pure $D_s^+ \to \phi\pi^+$ or $D_s^+ \to K^{*0}K^+$ decays. Besides, more in-depth considerations on $D_s^+ \to K^+K^-\pi^+$ might allow to reduce systematic uncertainties introduced by final state branching fractions.

Based on a well established kinematical model describing three-body decays a formalism will be implemented that allows to simulate the full underlying decay dynamics of $D_s^+ \to K^+K^-\pi^+$. This formalism requires a set of parameters to be determined to correctly model all components contributing to this decay. The determination of these model parameters is, however, beyond the scope of the present analysis. Instead, this study makes use of model parameters measured by another experimental flavor physics group, the CLEO collaboration. Based on the CLEO measurement, in the following sections the model implementation is described step-by-step.

4.4.1 Dalitz Plot Formalism

An approach that allows to systematically identify all the significantly contributing intermediate states of a three-body decay of the form $D \to Rc$ with $R \to ab$ is given by the *Dalitz Plot* technique [101]. In the Dalitz technique a minimum set of independent observable quantities is used only. For a spin-zero particle like the D_s^+ meson decaying into three daughters, only two degrees of freedom are required to completely describe the kinematics of the decay chain $D \to Rc$, $R \to ab$. This can be derived from simple kinematical considerations: Due to mass constraints, conservation of energy and momentum, and rotational invariance of the $D \to Rc$ system, the 12 unknown quantities corresponding to the four-momenta of the three daughters in the D_s^+ rest frame can be reduced to two. A good choice for these two independent Dalitz variables are two out of the three possible squared invariant masses of each pair of the final state particles abc, m_{ab}^2, m_{ac}^2, or m_{cb}^2.

If decay dynamics were governed solely by phase space, events in the plane of m_{ab}^2 and m_{ac}^2, for instance, would be uniformly distributed within the kinematically allowed borders. However, since weak non-leptonic decays of D (and as well B) mesons proceed dominantly through resonant two-body decays one expects the scatter plot of events in the (m_{ab}^2, m_{ac}^2) plane to have characteristic structures and shapes that reflect the dynamics of the decay $D \to abc$: For instance, strongly localized enhancements in the scatter plot indicate sharp intermediate resonances $R \to ab$.

For three-body decays of a spin-zero D meson to pseudo-scalar final states, $D \to abc$, the decay fraction reads

$$d\Gamma = \frac{|\mathcal{M}|^2}{(2\pi)^3 32 (M_{D_s})^3} dm_{ab}^2 dm_{bc}^2, \tag{4.5}$$

where $\mathcal{M}$ is the decay matrix element incorporating all the decay dynamics. A common approach to compute $\mathcal{M}$ is given by the *isobar model*: Each contribution to the Dalitz plot is modeled as a separate amplitude with a complex coefficient parameterized by magnitude and phase. The complex coefficients thus contain information about relative magnitudes and phases among the contributions. The total amplitude is given by the sum of all contributions. Hence, by construction the squared amplitude $\mathcal{M}^2$ contains diagonal and interference terms.

In the following the individual components contributing to $\mathcal{M}$ are being detailed. The matrix element can be parameterized as a sum of partial amplitudes,

$$\mathcal{M} = \sum_R c_R \times \Omega_R \times \mathcal{F}_D^L \times \mathcal{F}_R^L \times \mathcal{W}_R, \tag{4.6}$$

where $c_R = a_R e^{i\phi_R}$ is the complex coefficient.

Ω_R is the angular distribution depending on the spin of a given resonance R:

$$\Omega_R^{L=0} = 1 \tag{4.7}$$

$$\Omega_R^{L=1} = m_{bc}^2 - m_{ac}^2 + \frac{(m_D^2 - m_c^2)(m_a^2 - m_b^2)}{m_{ab}^2} \tag{4.8}$$

$$\Omega_R^{L=2} = [\Omega_R^{L=1}]^2 - \frac{1}{3}\left(m_{ab}^2 - 2m_D^2 - 2m_c^2 + \frac{(m_D^2 - m_c^2)^2}{m_{ab}^2}\right) \times \left(m_{ab}^2 - 2m_a^2 - 2m_b^2 + \frac{(m_a^2 - m_b^2)^2}{m_{ab}^2}\right) \tag{4.9}$$

Particular attention must be given permutation of the axes ab, bc, ac.

$\mathcal{F}_D^L$ and $\mathcal{F}_R^L$ are the angular momentum barrier-penetration factors for the D meson or the resonance decay vertex R. They depend both on spin and the effective meson radius r and are commonly parameterized in the (normalized) Blatt-Weisskopf form [102]. Here, the cases of a zero meson radius, $r = 0$, giving rise to $\mathcal{F}_V^L = 1$, and non-zero meson radii

$$\mathcal{F}_R^0 = 1 \tag{4.10}$$

$$\mathcal{F}_D^0 = e^{-(z - z_R)/12} \tag{4.11}$$

$$\mathcal{F}_{R,D}^1 = \sqrt{\frac{1 + z_R}{1 + z}} \tag{4.12}$$

$$\mathcal{F}_{R,D}^2 = \sqrt{\frac{9 + 3z_R + z_R^2}{9 + 3z + z^2}} \tag{4.13}$$

where

$$z = r_{R,D}^2 p^2 \tag{4.14}$$

$$z_R = r_{R,D}^2 p_R^2 \tag{4.15}$$

are distinguished. p is the decay products' momentum in the decaying particle's rest frame, and p_R the products' momentum at $m = m_R$.

The resonance term is introduced by $\mathcal{W}_R$. Its parameterization depends on the particular kind of the resonance: For regular resonances like $K^{*0}(892)$, $K_0^*(1430)$ or $\phi(1020)$ a common formulation is given by the Breit-Wigner function

$$\mathcal{W}_R(m) = \frac{1}{m_R^2 - m^2 - i m_R \Gamma(m)}. \tag{4.16}$$

The mass-dependent width can be expressed as [103]

$$\Gamma(m) = \Gamma_R \frac{m_R}{m} \left(\frac{p}{p_R}\right)^{2L+1} [\mathcal{F}_R^L(z)]^2. \tag{4.17}$$

If a resonance's mass is close to a meson pair production threshold, which is true for the $f_0(980)$ resonance ($m_{f_0(980)}$ is close to $K\bar{K}$ production threshold), the Flatté [104] formula is commonly used:

$$\mathcal{W}_R(m) = \frac{1}{m_R^2 - m^2 - i\sum_{ab} g_{Rab}^2 \rho_{ab}(m)} \tag{4.18}$$

where $a, b \in [\pi^0, \pi^\pm, K^\pm, K^0, \bar{K}^0]$. $\rho_{ab}(m) = 2p_a/m$ is a phase space factor and g_{Rab}^2 is a coupling constant of resonance R to the final state ab.

For a low mass $K^+\pi^-$ S wave, known as κ or $K(800)$, a complex pole amplitude

$$\mathcal{W}_\kappa(m) = \frac{1}{m_\kappa^2 - m^2} \tag{4.19}$$

is a common choice.

4.4.2 CLEO $D_s^+ \to K^+K^-\pi^+$ Dalitz Plot Analysis

At the time of model implementation, the most recent and most precise Dalitz plot analysis of the decay $D_s^+ \to K^+K^-\pi^+$ is provided by the CLEO Collaboration [105]. Using the CLEO-c data set corresponding to 548pb^{-1} of e^+e^- collisions at $\sqrt{s} = 4.17\text{GeV}/\text{c}^2$, the CLEO Collaboration reconstructs $14,400$ $D_s^+ \to K^+K^-\pi^+$ candidates (the charge conjugate decay is implied throughout) at a background level of 15.1%, corresponding to a yield of $12,200$ signal events.

Following the Dalitz formalism outlined above, magnitudes and phases of all kinematically allowed $K^-\pi^+$ and K^+K^- resonances recognized by the PDG are determined in a unbinned maximum likelihood fit. The signal component of the fit function essentially contains the squared amplitude $|\mathcal{M}|^2$. In addition to the magnitudes and phases, whose values CLEO sets relative to the magnitude $a_{K^{*0}(892)} = 1$ and the phase $\phi_{K^{*0}(892)} = 0$, the mass and width of the $K^{*0}(892)$ resonance are kept free in the fit. This gives a total of 12 floating fit parameters.

Given a successful Dalitz plot fit, CLEO calculates fit fractions of contributing resonances. The fit fraction of a given resonance R is defined as the full Dalitz plot integral of squared amplitudes for a given resonance divided by the integral of squared amplitudes for the sum of all considered resonances:

$$FF_R = \frac{\oint |\mathcal{M}_R|^2 \, dm_{K^+K^-}^2 dm_{K^-\pi^+}^2}{\oint \sum_R |\mathcal{M}_R|^2 \, dm_{K^+K^-}^2 dm_{K^-\pi^+}^2} \tag{4.20}$$

CLEO concludes that a six-resonance model, referred to as Model A, consisting of $K^{*0}(892)K^+$, $K_0^*(1430)K^+$, $f_0(980)\pi^+$, $\phi(1020)\pi^+$, $f_0(1370)\pi^+$, and $f_0(1710)\pi^+$ describes their data best. In addition, CLEO reports that the consistency with data is larger than in previous analyses that did not account for the $f_0(1370)\pi^+$ contribution.

All remaining allowed resonances are included one by one in the signal component of the fit function, and the fit to data is repeated. It is found that all other resonances do not significantly improve fit quality. The final CLEO results are summarized in Table 4.3. From that it can be seen that the sharp resonances $K^{*0}(892)$ and $\phi(1020)$ and the threshold resonance $f_0(980)$ are the most relevant resonances observed in $D_s^+ \to K^+K^-\pi^+$ phase space. It is worth noting that the fit fractions do not add up to 100%, an intrinsic property of the isobar ansatz: Contributions can interfere constructively or destructively. In the given case constructive interference seems to be prevalent.

For further details on the Dalitz formalism, experimental techniques, and parameter values used in CLEO Dalitz analyses please refer to [3, 106, 107].

4.4.3 Reproducing CLEO's Dalitz Plot Results

Based on the Dalitz formalism outlined above, additional valuable information provided by the authors [108], and by inserting parameter values (masses, widths, magnitudes, and phases of resonances) quoted in the CLEO paper (in particular results quoted in Table 4.3), two routines are coded: The first one provides a calculation of the squared amplitude $|\mathcal{M}|^2$ at a given coordinate in the $(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ plane for a given or the sum of all resonances of Model A. With the second routine integrations of $|\mathcal{M}|^2$ over the full or particular regions of the Dalitz plot can be carried out. This enables us to reproduce CLEO's fit fraction results, for instance. The whole purpose of these two routines will become evident in Sections 4.6 and 5.5.

Please note that in contrast to the CLEO analysis no Dalitz fits to data are performed in the course of the present study. Instead, the signal component (i.e. the squared amplitude $|\mathcal{M}|^2$) of the probability density function used in the CLEO fit is set up statically[2] only, with the values of magnitudes and phases determined by the CLEO Collaboration being inserted. Figure 4.1 shows the Dalitz plot of $D_s^+ \to K^+K^-\pi^+$. The plot does not show a scatter plot of events, but is rather a graphical representation of the square of the absolute value of the complex function $\mathcal{M}(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ that was computed using CLEO's specifications. A higher density of plotted points indicates a higher $|\mathcal{M}|^2$. As a matter of fact, an event scatter plot exhibits the same topology, since the probability for events to populate a particular region of phase space is given by $|\mathcal{M}|^2$.

The two dashed vertical red lines and the dotted horizontal blue lines indicate the invariant mass requirements applied in the reconstruction of ϕ and the $K^{*0}(892)$

---

[2] In the course of later discussions this statement will be revised.

Parameter	**Parameter Index**	**CLEO-c Model A**
$m_{K^{*0}(892)}$ (MeV/c^2)	0	894.9 ± 0.5 ± 0.7
$\Gamma_{K^{*0}(892)}$ (MeV/c^2)	1	45.7 ± 1.1 ± 0.5
$a_{K^{*0}(892)}$	–	1 (fixed)
$\phi_{K^{*0}(892)}$ (°)	–	0 (fixed)
$a_{K_0^*(1430)}$	2	1.51 ± 0.11 ± 0.09
$\phi_{K_0^*(1430)}$ (°)	3	146 ± 8 ± 8
$a_{f_0(980)}$	4	4.72 ± 0.18 ± 0.17
$\phi_{f_0(980)}$ (°)	5	157 ± 3 ± 4
$a_{\phi(1020)}$	6	1.13 ± 0.02 ± 0.02
$\phi_{\phi(1020)}$ (°)	7	−8 ± 4 ± 4
$a_{f_0(1370)}$	8	1.15 ± 0.09 ± 0.06
$\phi_{f_0(1370)}$ (°)	9	53 ± 5 ± 6
$a_{f_0(1710)}$	10	1.11 ± 0.07 ± 0.10
$\phi_{f_0(1710)}$ (°)	11	89 ± 5 ± 5
FF[$K^{*0}(892)$] (%)		47.4 ± 1.5 ± 0.4
FF[$K_0^*(1430)$] (%)		3.9 ± 0.5 ± 0.5
FF[$f_0(980)$] (%)		28.2 ± 1.9 ± 1.8
FF[$\phi(1020)$] (%)		42.2 ± 1.6 ± 0.3
FF[$f_0(1370)$] (%)		4.3 ± 0.6 ± 0.5
FF[$f_0(1710)$] (%)		3.4 ± 0.5 ± 0.3
$\sum_R \mathrm{FF}_R$ (%)		129.5 ± 4.4 ± 2.0

Table 4.3: Magnitudes, phases, and fit fractions of the six most significantly contributing resonances (Model A) observed in $D_s^+ \to K^+K^-\pi^+$ decays as measured by CLEO [105]. The magnitudes a_R are given in units of the $K^{*0}(892)$ magnitude which is fixed to 1. The uncertainties quoted are statistical and systematic.

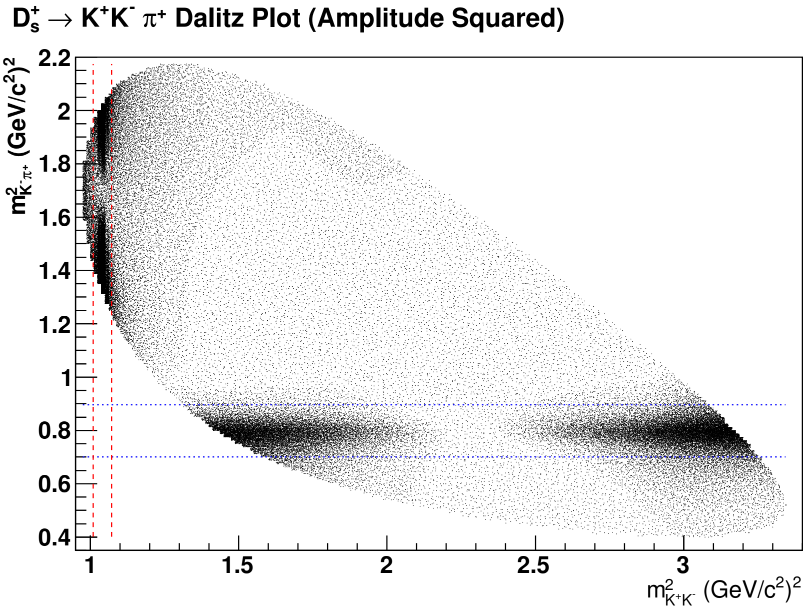


Figure 4.1: Graphical representation of the squared decay amplitude $\mathcal{M}(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ for the decay $D_s^+ \to K^+K^-\pi^+$ calculated from the results of CLEO's Dalitz plot analysis [105].

mesons, respectively. Inside the area marked by these lines the sharp ϕ resonance (dashed red lines) and the broader $K^{*0}(892)$ resonance (dotted blue lines) are clearly visible. Their twin-peak structure is a consequence of both particles being spin-1 states. The thin vertical enhancement to the very left of the Dalitz plot can be explained by the presence of the $f_0(980)$ meson. All the other resonances considered in the CLEO-c Model A are fairly uniformly distributed across $D_s^+ \to K^+K^-\pi^+$ phase space.

After qualitatively reproducing CLEO's Dalitz analysis results, the next aim is to quantitatively reproduce the fit fractions presented in the CLEO paper. For this, Dalitz plot integrations are performed for each of the resonances and divided by the integral of the full Dalitz plot. Apart from that the fit fractions within the mass windows introduced in reconstruction are of interest. The results are quoted in Table 4.4. Inside the K^{*0} mass window the $K^{*0}(892)$ accounts for a overwhelming fraction of 99.0% with very small contributions from the other resonances. The second largest contribution stems from the $f_0(980)$, accounting for 4.5%. In the ϕ mass band the major contribution comes from the $\phi(1020)$ itself (93.8%), but there is still a significant contribution from the $f_0(980)$ (8.8%). All the other resonances in this mass band are found to be almost negligible.

	Fit Fraction FF (%) In		
Resonance	**Full Dalitz Region**	**ϕ Mass Band**	**K^{*0} Mass Band**
$K^{*0}(892)$	47.6	0.3	99.0
$K_0^*(1430)$	4.0	0.4	0.5
$\phi(1020)$	41.8	8.8	4.5
$f_0(980)$	28.4	93.8	0.9
$f_0(1370)$	4.3	0.1	1.5
$f_0(1710)$	3.4	0.0	2.2
$\sum_R \mathrm{FF}_R$	129.5	103.3	108.6

Table 4.4: Fit fractions calculated from squared Dalitz plot amplitudes using CLEO's $D_s^+ \to K^+K^-\pi^+$ Dalitz model specifications. From left to right the fit fractions in full $D_s^+ \to K^+K^-\pi^+$ phase space, inside the ϕ mass band, and inside the K^{*0} mass band are shown.

4.4.4 Remarks and Nomenclature

The considerations made in Section 4.4 and the subsequent discussions might suggest that the statements with regards to reconstruction of $B_s^0 \to D_s^+D_s^-$ via $D_s^+ \to \phi\pi^+$ and $D_s^+ \to K^{*0}K^+$ need to be overhauled; in actual fact, $B_s^0 \to D_s^+D_s^-$ is reconstructed via $D_s^+ \to K^+K^-\pi^+$ by choosing two narrow mass bands of the $D_s^+ \to K^+K^-\pi^+$ Dalitz plot for which one expects to reconstruct $\phi(1020)$ and $K^{*0}(892)$ mainly. Neither in offline reconstruction, nor in the final selection any attempt is being made to distinguish between $\phi(1020)$ ($K^{*0}(892)$) and any other K^+K^- ($K^-\pi^+$) resonance. To avoid confusion and to have a formal means of distinguishing between the exclusive decays $D_s^+ \to \phi\pi^+$ and $D_s^+ \to K^{*0}K^+$ on the one hand and the $D_s^+ \to K^+K^-\pi^+$ decays lying in the invariant mass regions introduced by reconstruction on the other hand, from now on the phrases $D_s^+ \to \phi^\dagger\pi^+$ and $D_s^+ \to K^{*0\dagger}K^+$ are used as abbreviations for $[D_s^+ \to K^+K^-\pi^+]_{\phi(1020)\text{MASS BAND}}$ and $[D_s^+ \to K^+K^-\pi^+]_{K^{*0}(892)\text{MASS BAND}}$, respectively, to increase the readability of the text.

Finally, the question may arise why reconstruction does not use the full $D_s^+ \to K^+K^-\pi^+$ Dalitz plot region in the first place; the area of the chosen mass regions does indeed only account for about 14% of the full kinematically allowed region. Previous discussions have however shown that, due to the observed Dalitz structure of $D_s^+ \to K^+K^-\pi^+$, around 76% of all $D_s^+ \to K^+K^-\pi^+$ events are concentrated inside these two mass bands. The presence of large combinatorial background all across $D_s^+ \to K^+K^-\pi^+$ phase space one expects for a hadron collider, combined with the fact of strong localization of resonances in certain Dalitz plot regions, strongly suggests to use those two mass bands only where a high signal purity is expected.

4.4.5 Modified Final State Branching Fractions

For the decays $D_s^+ \to \phi^\dagger\pi^+$ and $D_s^+ \to K^{*0\dagger}K^+$ the relative branching fraction with respect to $D_s^+ \to K^+K^-\pi^+$ are re-defined as either the integral of the squared amplitudes over the $\phi(1020)$ or the $K^{*0}(892)$ mass band (specified by offline reconstruction) of the $D_s^+ \to K^+K^-\pi^+$ Dalitz plot divided by the integral over the full Dalitz plot region:

$$f_{\phi^\dagger\pi^+} = \frac{\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = \frac{\int_{\phi(1020)} \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}}{\oint \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}} \tag{4.21}$$

$$f_{K^{*0\dagger}K^+} = \frac{\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = \frac{\int_{K^{*0}(892)} \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}}{\oint \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}} \tag{4.22}$$

To estimate the impact of the Dalitz model parameter uncertainties on these mass band fractions and to evaluate possible correlations among them parameter toy studies are carried out. This is done by repeating the calculations (4.21) and (4.22) various times. However, in each calculation the 12 CLEO Dalitz plot parameters (those quoted in Table 4.3) are now allowed to take any values according to Gaussian errors, where the cases of statistical and systematic uncertainties are treated separately.

For the category of systematic parameter uncertainties, the toy study procedure is straightforward: We generate uncorrelated Gaussian random numbers according to the systematic uncertainties CLEO quotes and add them to the central parameter values. Statistical uncertainties are a little bit more difficult to handle, since one has to take into account correlations among the 12 free fit parameters. The CLEO authors kindly provided us the full covariance matrix of fit parameters (see Section A.2.1), which enables us to compute correlated Gaussian random numbers. This is done as follows: First, a vector of uncorrelated Gaussian random numbers having mean 0 and width 1 is generated and then multiplied with the decomposed lower-left covariance matrix (obtained by Cholesky decomposition [109]). The result is a vector of parameter uncertainties which are now correlated among each other according to the covariance matrix. Afterwards the individual uncertainty projections are verified to keep their Gaussian shape, with their respective Gaussian widths being equal to the published parameter uncertainties (see Figures A.1 and A.2 in the Appendix). To obtain the absolute random parameter values, the randomly generated (and now correlated) uncertainties are added to the central parameter values. According to the variation of Dalitz parameters,

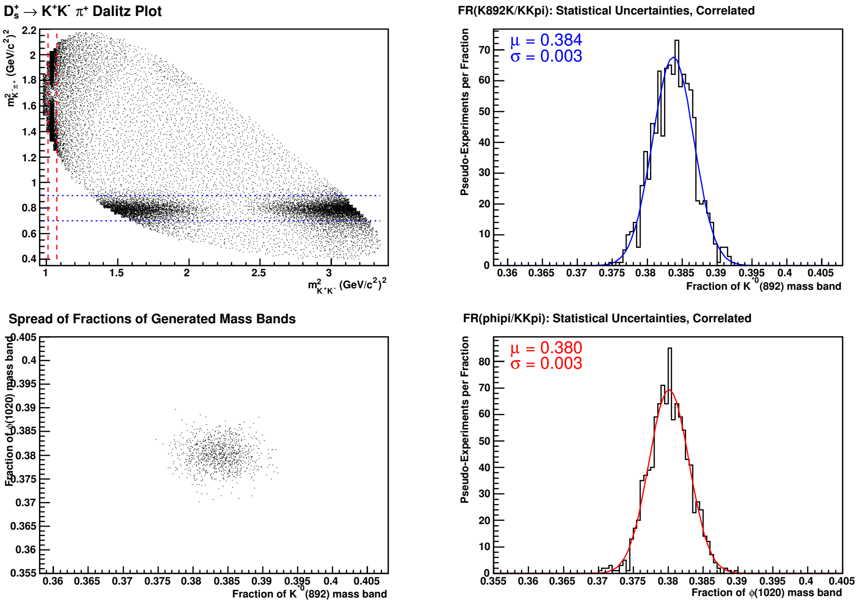


Figure 4.2: Computation of the fraction of $D_s^+ \to K^+K^-\pi^+$ decays as defined by equations (4.21) and (4.22). In the 1000 trials the statistical uncertainties of the CLEO Dalitz model parameters and their correlations were accounted for.

each parameter configuration represents a slightly differing $D_s^+ \to K^+K^-\pi^+$ Dalitz model scenario.

Considering 1,000 Dalitz model variations separately for both the categories of statistical and systematic Dalitz model parameter uncertainties, the computed $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ mass band fractions are plotted into histograms and the widths of the distributions are assigned as uncertainties. The parameter toy studies give rise to

$$f_{\phi^\dagger\pi^+} = \frac{\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = 0.380 \pm 0.003(stat) \pm 0.010(sys) = 0.380 \pm 0.010 \quad (4.23)$$

$$f_{K^{*0\dagger}K^+} = \frac{\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = 0.384 \pm 0.003(stat) \pm 0.011(sys) = 0.384 \pm 0.011 \quad (4.24)$$

The top right and bottom right plots in Figures 4.2 and 4.3 illustrate the outcomes of the $D_s^+ \to \phi^\dagger\pi^+$ and $D_s^+ \to K^{*0\dagger}K^+$ relative mass band fraction estimations. The top left plot shows the two mass bands of the $D_s^+ \to K^+K^-\pi^+$ Dalitz plot chosen in reconstruction.

In addition to the central values and uncertainties obtained in the toy experiments, correlations (illustrated by the bottom left scatter plots) among them are investigated. When correlated statistical Dalitz parameter uncertainties are

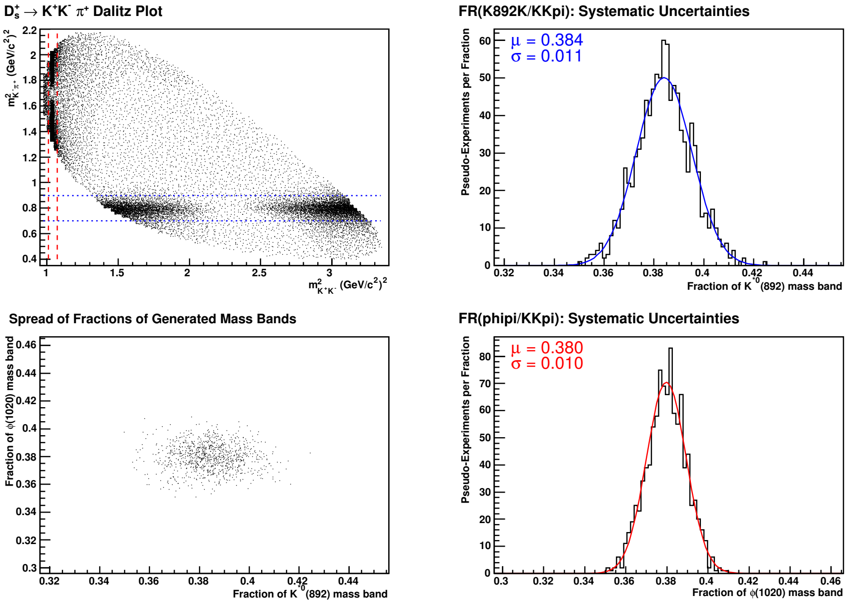


Figure 4.3: Computation of the fraction of $D_s^+ \to K^+K^-\pi^+$ decays as defined by equations (4.21) and (4.22) when accounting for systematic parameter uncertainties CLEO quotes.

accounted for, the correlation matrix for $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ reads:

$$\begin{array}{cc} 1 & -0.035 \\ -0.035 & 1 \end{array} \tag{4.25}$$

Considering systematic Dalitz parameter uncertainties, the following correlation matrix is obtained:

$$\begin{array}{cc} 1 & -0.013 \\ -0.013 & 1 \end{array} \tag{4.26}$$

We consider the tiny anti-correlations between the two relative branching fractions negligible. From the relative fractions (4.23) and (4.24), the absolute mass band fractions of $D_s^+ \to \phi^\dagger\pi^+$ and $D_s^+ \to K^{*0\dagger}K^+$ decays can be calculated by inserting the PDG value $\mathcal{B}(D_s^+ \to K^+K^-\pi^+) = 0.0549 \pm 0.0027$:

$$\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+) = 0.0209 \pm 0.0012 \tag{4.27}$$

$$\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+) = 0.0211 \pm 0.0012 \tag{4.28}$$

For the reasons outlined in Section 6.5, when extracting the quantities of interest $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ from data, the figures given in equations (4.27) and (4.28) will be again factorized into (4.23) and (4.24) times the branching fraction of $D_s^+ \to K^+K^-\pi^+$.

4.5 Additional Considerations on $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$

The elaborations on the off-line reconstruction given in Section 4.2 did not cover the cases where in the decay of the B_s^0 meson one or two excited $D_s^{*\pm}$ mesons are involved. Theses cases introduce a slight complication to the reconstruction process since the $D_s^{*\pm}$ de-excitation into the ground state $D_s^{\pm}$ is accompanied by the emittance of a slow neutral pion or a photon. Both the neutral pion and the photon are however not detected, because their energies are below the energy threshold of the CDF calorimetry system. Thus, the four-momentum of either the pion or the photon are missing and the bottom-up reconstruction chain is insofar incomplete as the chain element next to the B_s^0, i.e. the $D_s^{*\pm}$, is skipped over. In this case one speaks of *partially* reconstructed B_s^0 events, which manifest as differently shaped structures shifted towards lower values in the invariant $B_s^0 \to D_s^+ D_s^-$ mass spectrum.

The mass line shapes of these partially reconstructed sequential two-body decays are non-trivial since they can depend on three different aspects: First of all, as stated above, $D_s^{*\pm}$ de-excitation can proceed via both a neutral pion or a photon. Being non-spin-0 particles, on the other hand the angular decay distributions of the $D_s^{*\pm}$ decay products are not flat. In this context, "flat" means that there is no preferred direction of particle release. Different angular distributions might result in variously formed mass line shapes. Finally, the aforementioned aspects might in addition be sculptured by detector acceptance effects.

The mass line shapes can be estimated from Monte Carlo simulations, as long as the required input information are available. This is true for the former aspect: The branching ratios of the $D_s^{*\pm}$ de-excitation modes are precisely measured and published by the *Particle Data Group* (PDG) [3]:

$$\mathcal{B}(D_s^{*+} \to D_s^+ \gamma) = (94.2 \pm 0.7) \times 10^{-2} \tag{4.29}$$

$$\mathcal{B}(D_s^{*+} \to D_s^+ \pi^0) = (5.8 \pm 0.7) \times 10^{-2} \tag{4.30}$$

Furthermore, simulation of the detector response is implemented in the Monte Carlo software package (see Section 4.6). For the discussion of angular distributions, the computation of decay amplitudes in the helicity basis is outlined first. The helicity basis represents a commonly used frame well suited for describing sequential two-body decays. While concentrating on the critical aspects and expressions only, which were extracted from [110] and the derived pedagogical reviews [111, 112, 113], the subsequent sections provide a brief introduction to this topic.

4.5.1 Decay Amplitudes in the Helicity Basis

We consider a sequential two-body decay of particles having spins s_i

$$\begin{array}{ccll} 1 & \rightarrow & 2 & 3 \\ 2 & \rightarrow & 4 & 5 \\ 3 & \rightarrow & 6 & 7 \end{array} \tag{4.31}$$

in the helicity basis. In this frame the dynamics of this sequential decay are fully described by the given helicities $\lambda_i, i = 1...7$, where λ_i is defined as the particle's i spin-projection with the spin quantization axis $\tilde{z}$ given by its own momentum $\vec{p}_i^{(j)}$ measured in the parent's j rest frame, and the particle's i flight direction in its parent's rest frame, determined by a set of angles $\tilde{\Omega} = (\tilde{\theta}, \tilde{\phi})$.

For a sequence of two two-body decays one therefore has to construct three coordinate systems to fully describe the problem:

- The rest frame of particle 1 which is identified by the unprimed coordinate system $X = (x, y, z)$. In this rest frame the set of angles $\Omega = (\theta, \phi)$ is defined.
- The rest frame of particle 2 which is identified by the primed coordinate system $X' = (x', y', z')$. In this rest frame the set of angles $\Omega' = (\theta', \phi')$ is defined.
- The rest frame of particle 3 which is identified by the double primed coordinate system $X'' = (x'', y'', z'')$. In this rest frame the set of angles $\Omega'' = (\theta'', \phi'')$ is defined.

Each of the coordinate systems forms a right-handed orthonormal basis.

Given particle momenta $\vec{p}_i^{lab}$ specified in the lab frame, a series of rotations and boosts are needed to proceed from the lab frame to X, X', and X''. First, to obtain the unprimed coordinate system the Euler operator $R(\phi^{lab}, \theta^{lab}, -\phi^{lab})$ is applied to the lab system, effecting a rotation of the z^{lab} axis onto the flight direction of particle 1. With the choice of the third Euler angle $\gamma = -\alpha$ we follow the Jacob-Wick convention [110]. In a second step, the z axis is rotated by $R(\phi, \theta, -\phi)$ onto the flight direction of particle 2, followed by a boost into the rest frame of particle 2, yielding the primed coordinate system X'. Again starting from X, in a similar way the double primed system X'' is obtained by rotating and boosting into the rest frame of particle 3.

The Euler operator $R(\alpha, \beta, \gamma)$ performs an irreducible rotation by a set of rotation angles (α, β, γ) of a system with total angular momentum j and a definite third component m along a quantization axis z onto a new system characterized by the axis z'. In applying the rotation operator, the basis state $|jm\rangle$ is transformed into a linear combination of new basis states $|jm'\rangle$:

$$R(\alpha,\beta,\gamma)\,|jm\rangle = \sum_{m'=-j}^{j} D^{j}_{m',m}(\alpha,\beta,\gamma)\,|jm'\rangle \tag{4.32}$$

Here, $D^{j}_{m,m'}(\alpha,\beta,\gamma)$ are the Wigner D-functions, which can be explicitly represented as

$$D^{j}_{m',m}(\alpha,\beta,\gamma) = e^{-im'\alpha} d^{j}_{m',m}(\beta) e^{-im\gamma} \tag{4.33}$$

$d^{j}_{m',m}$ are commonly referred to as the reduced Wigner functions, some of which are tabulated in the literature, in [3] for instance. From the identity

$$d^{j}_{m',m} = (-1)^{m-m'} d^{j}_{m,m'} = d^{j}_{-m,-m'} \tag{4.34}$$

the reduced Wigner functions obey, all the other elements of $d^{j}_{m',m}$ that are not listed in the relevant literature can be obtained.

Given a two-body decay $1 \to 2\,3$, in the helicity frame the decay amplitude is calculated by

$$A(\lambda_1,\lambda_2,\lambda_3,\Omega) = \sqrt{\frac{2s_1+1}{4\pi}} D^{s_1 *}_{\lambda_1,\lambda_2-\lambda_3}(\Omega) A_{\lambda_2\lambda_3}, \tag{4.35}$$

where the set of angles $\phi, \theta, -\phi$ are encapsulated in Ω. Equation (4.35) represents the amplitude for the spin s_1 of a decaying particle 1 to have projection $\lambda_2 - \lambda_3$ along the decay axis, multiplied by the coupling to the final state helicities, given by the complex amplitude $A_{\lambda_2\lambda_3}$. Because of conservation of angular momentum, only helicity amplitudes satisfying the selection rule

$$|\lambda_2 - \lambda_3| \leq s_1 \tag{4.36}$$

are valid. If a decay is parity conserving, the number of independent helicity amplitudes is further reduced by the condition

$$A_{-\lambda_2 - \lambda_3} = \eta_2\eta_3\eta_1(-1)^{s_2+s_3-s_1} A_{\lambda_2\lambda_3}, \tag{4.37}$$

where η_i denotes a particle's intrinsic parity.

We now add another sequence of two-body decays, with particle 2 decaying to 4 and 5, and $3 \to 6\,7$. Constraining particle 1 to be a spin-0 particle with fixed helicity $\lambda_1 = 0$, equation (4.36) implies that $\lambda_2 = \lambda_3 = \lambda$. Therefore, the general expression for the decay amplitude of a sequence of two-body decays

$$\begin{aligned} A(\lambda_1,\lambda_4,\lambda_5,\lambda_6,\lambda_7,\Omega,\Omega',\Omega'') &= \sqrt{\frac{(2s_1+1)}{4\pi}\frac{(2s_2+1)}{4\pi}\frac{(2s_3+1)}{4\pi}} \\ &\times \sum_{|\lambda_2-\lambda_3|=0}^{s_1} D^{s_1 *}_{\lambda_1,\lambda_2-\lambda_3}(\Omega) A_{\lambda_2\lambda_3} D^{s_2 *}_{\lambda_2,\lambda_4-\lambda_5}(\Omega') B_{\lambda_4\lambda_5} D^{s_3 *}_{\lambda_3,\lambda_6-\lambda_7}(\Omega'') C_{\lambda_6\lambda_7} \end{aligned} \tag{4.38}$$

simplifies to

$$A(\lambda_4, \lambda_5, \lambda_6, \lambda_7, \Omega', \Omega'') = \sqrt{\frac{1}{4\pi}\frac{(2s_2+1)}{4\pi}\frac{(2s_3+1)}{4\pi}} \times \sum_{|\lambda|=0}^{1} A_{\lambda\lambda} D^{s_2 *}_{\lambda,\lambda_4-\lambda_5}(\Omega') B_{\lambda_4\lambda_5} D^{s_3 *}_{\lambda,\lambda_6-\lambda_7}(\Omega'') C_{\lambda_6\lambda_7}, \tag{4.39}$$

where one coherently sums over the allowed helicities of the intermediate particles 1 and 2, because they cannot be measured. The probability of particles stemming from a sequence of two-body decays to have flight direction in the angle elements $d\Omega'$,$d\Omega''$ is then simply found by squaring the decay amplitude. If an experiment does not measure the final state helicities one in addition has to incoherently sum over them. Neglecting constant factors, the angular distribution is then given by

$$\left|A(\theta, \psi, \chi)\right|^2 = \sum_{|\lambda_4-\lambda_5|}^{s_1} \sum_{|\lambda_6-\lambda_7|}^{s_1} \left| \sum_{|\lambda|=0}^{1} H_\lambda e^{i\lambda\chi} d^1_{\lambda,\lambda_4-\lambda_5}(\theta) B_{\lambda_4\lambda_5} d^1_{\lambda,\lambda_6-\lambda_7}(\psi) C_{\lambda_6\lambda_7} \right|^2 \tag{4.40}$$

For the upcoming discussions we demand that the maximum spin of particles 2 and 3 is $s_{2,3} \leq 1$, and simplify the notation of the three allowed helicity amplitudes A_{11}, A_{00}, A_{-1-1} to H_+, H_0, H_-. The helicity amplitudes are related to the amplitudes of the transversity basis [114, 115] by

$$T_{\|} = \frac{H_+ + H_-}{\sqrt{2}}, \quad T_0 = H_0, \quad T_\perp = \frac{H_+ - H_-}{\sqrt{2}} \tag{4.41}$$

The fraction of the square of the amplitude T_0 (or H_0) is called the longitudinal polarization fraction f_L:

$$f_L = \frac{|H_0|^2}{|H_+|^2 + |H_0|^2 + |H_-|^2} = \frac{|T_0|^2}{\left|T_{\|}\right|^2 + |T_0|^2 + |T_\perp|^2} \tag{4.42}$$

From the transversity amplitudes the fractions of the *CP* eigenstates can be derived easily. The fractions of the parallel and the longitudinal amplitude determine the *CP* even component, while the fraction of the perpendicular amplitude corresponds to the *CP* odd component:

$$f_{CP+} = f_{\|} + f_0 = \frac{\left|T_{\|}\right|^2 + |T_0|^2}{\left|T_{\|}\right|^2 + |T_0|^2 + |T_\perp|^2} \tag{4.43}$$

$$f_{CP-} = f_\perp = \frac{|T_\perp|^2}{\left|T_{\|}\right|^2 + |T_0|^2 + |T_\perp|^2} \tag{4.44}$$

4.5.2 $B_s^0 \to D_s^{*+} D_s^-$

We now consider the decay $B_s^0 \to D_s^{*+} D_s^-$, with $D_s^{*+} \to D_s^+ \gamma$ or $D_s^{*+} \to D_s^+ \pi^0$. The first decay of the two-body decay chain corresponds to the decay of a pseudo-scalar (P) to a vector (V) and a pseudo-scalar (P) particle. In terms of helicities $\lambda_1 \to \lambda_2\, \lambda_3$, it can be expressed as

$$\begin{array}{cccc} & & 1 & \\ 0 & \to & 0 & 0 \\ & & -1 & \end{array} \tag{4.45}$$

Being a spin-1 particle, the D_s^{*+} meson has three possible helicities $-1, 0, -1$. Due to conservation of angular momentum (4.36) only one helicity projection onto the decay axis, $\lambda = \lambda_2 - \lambda_3 = 0$, is allowed though. This implies that this decay necessarily has only one helicity amplitude H_0, or in other words, the D_s^{*+} meson is 100% longitudinally polarized. As a consequence, $D_s^{*+} \to D_s^+ \pi^0$ only has one amplitude B_{00}. The impact of this on invariant mass line shapes will be discussed in Section 6.3.3. Since the photon is a massless spin-1 particle which has no longitudinal polarization component ($\lambda_5 = \pm 1$) there are two amplitudes C_{01}, C_{0-1} for $D_s^{*+} \to D_s^+ \gamma$.

This brief discussion demonstrates that all requisites (helicity amplitudes, final state branching fractions) are in place to simulate this decay based on decay amplitude calculations. Thus, for the simulation of $B_s^0 \to D_s^{*+} D_s^-$ common Monte Carlo generation procedures (please refer to Section 4.6) are relied upon, including event decay simulation using the *EvtGen* package [116, 117, 118].

4.5.3 $B_s^0 \to D_s^{*+} D_s^{*-}$

The situation is different in the case of the $P \to VV$ decay $B_s^0 \to D_s^{*+} D_s^{*-}$, which can be expressed as

$$\begin{array}{cccc} & & 1 & 1 \\ 0 & \to & 0 & 0 \\ & & -1 & -1 \end{array} \tag{4.46}$$

in terms of the helicities $\lambda_1 \to \lambda_2\, \lambda_3$. Figure 4.4 illustrates the decay in the helicity frame. Because of (4.36), there are three allowed helicity amplitudes $\{H_{11}, H_{00}, H_{-1-1}\} = \{H_+, H_0, H_-\}$. For the full sequence of decays, with $D_s^{*+} \to D_s^+ \gamma$ and $D_s^{*+} \to D_s^+ \pi^0$, the squared decay amplitude (4.40) is given by

$$\left|A(\theta, \psi, \chi)\right|^2 = \sum_{|\lambda_4 - \lambda_5|}^{1} \sum_{|\lambda_6 - \lambda_7|}^{1} \left| \sum_{|\lambda|=0}^{1} H_\lambda e^{i\lambda\chi} d^1_{\lambda, \lambda_4 - \lambda_5}(\theta) B_{\lambda_4 \lambda_5} d^1_{\lambda, \lambda_6 - \lambda_7}(\psi) C_{\lambda_6 \lambda_7} \right|^2 , \tag{4.47}$$

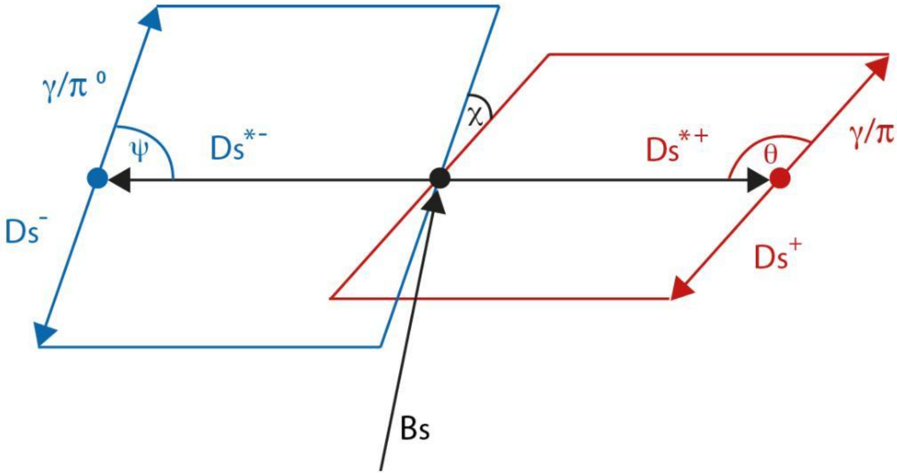


Figure 4.4: The decay $B_s^0 \to D_s^{*+} D_s^{*-}$ in the helicity frame. D_s^{*+} de-excitation proceeds via $D_s^{*+} \to D_s^+ \gamma$ or $D_s^{*+} \to D_s^+ \pi^0$

where constant factors have been omitted. In addition, the substitutions $\theta' \equiv \theta$, $\theta'' \equiv \psi$, and $\chi \equiv \phi' + \phi''$ have been made. The reduced Wigner d-functions for a particle with total angular momentum of 1 are tabulated in Section A.3. The external incoherent summation over helicities depends on the final state particles involved in the decay of the spin-1 particle D_s^{*+}. The pseudo-scalars D_s^+ and π^0 have fixed helicities of 0, while the photons' helicities are $\lambda_{5,7} = \pm 1$. The coupling amplitudes are then $B_{00} = C_{00}$ for pion decays, and B_{01}, B_{0-1}, C_{01}, and C_{0-1} for radiative D_s^{*+} decays. Because electromagnetic decays are parity conserving, it follows from (4.37):

$$B_{01} = -B_{0-1} \tag{4.48}$$

$$C_{01} = -C_{0-1} \tag{4.49}$$

Hence, the coupling amplitudes $B_{\lambda_4\lambda_5}$ and $C_{\lambda_6\lambda_7}$ can be factorized out as constant factors for every summand of the external helicity summation, and do not affect the resulting angular shape.

Equation (4.47) lacks one essential piece of input information: The helicity amplitudes $\{H_+, H_0, H_-\}$ are unknown, since they have not yet been measured before for this particular decay. From a theoretical point of view [34] it is however reasonable to use the helicity amplitudes measured for $B_d^0 \to D^{*+} D_s^{*-}$, $\{H_+, H_0, H_-\} = \{0.4904, 0.7204, 0.4904\}$ [119]: Both $B_s^0 \to D_s^{*+} D_s^{*-}$ and $B_d^0 \to D^{*+} D_s^{*-}$ proceed via a Cabibbo- and color-favored $b \to c\bar{c}s$ tree diagram, with the D_s^{*+} meson produced by the same weak current and the initial and final state meson masses being very close. The only difference arises by the spectator quark, s or d. When simulating $B_s^0 \to D_s^{*+} D_s^{*-}$ decays the helicity amplitudes of $B_d^0 \to D^{*+} D_s^{*-}$ are thus relied upon. Because different angular distributions can affect reconstructed

B_s^0 mass line shapes and efficiencies, not only the default $B_d^0 \to D^{*+}D_s^{*-}$ helicity configuration will be used, but in addition several different configurations in order to assess the effects on the final results as a systematic check (Section 7.3.3). Despite the fact that $B_s^0 \to D_s^{*+}D_s^{*-}$ is only partially reconstructed, changing parameter expectations in Monte Carlo might even allow to infer the true helicity amplitudes of $B_s^0 \to D_s^{*+}D_s^{*-}$ and, using (4.43-4.44), the fraction of the CP-odd component, if significantly different from zero. This will be investigated in Section 6.6.

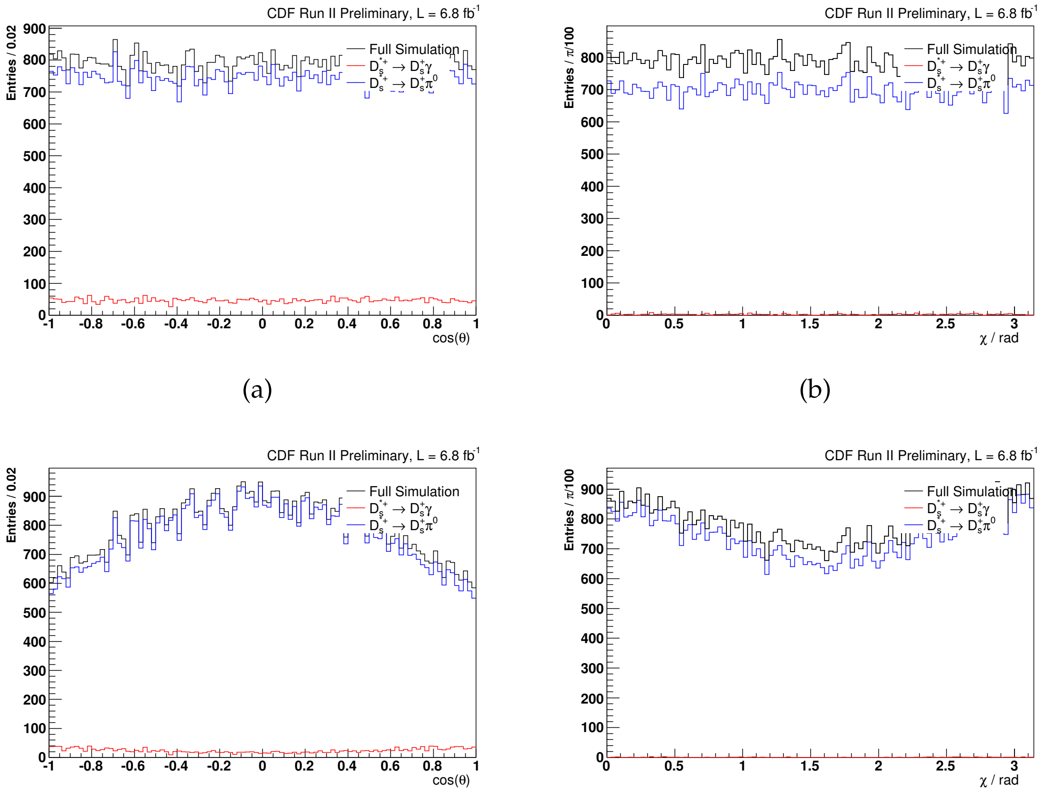
From a technical point of view there is one difficulty that has to be overcome: The full procedure of generating a Monte Carlo sample – including event generation, event decay, detector simulation, and further data processing – with reasonable statistics takes a considerable amount of time. In order to carry out studies with different sets of helicity amplitudes within a reasonable time scale only one Monte Carlo sample is therefore produced, for which phase space configuration is used in simulation. Other scenarios are obtained by doing an event-by-event weighting of simulated data, where the weights are given by the squared decay amplitudes calculated from equation (4.47). For the final state $\phi^{\dagger}\pi^{+}\phi^{\dagger}\pi^{-}$ (qualitatively the same results are obtained for $\phi^{\dagger}\pi^{+}K^{*0\dagger}K^{-}$), in Figure 4.5 the angular distributions of the helicity angles $\cos(\theta)$ and χ are shown for the phase space case and the standard helicity scenario $\{H_+, H_0, H_-\} = \{0.4904, 0.7204, 0.4904\}$.

4.6 Monte Carlo Simulation

Monte Carlo Ntuple production comprises several successive stages, starting with the creation of B mesons using the event generator *BGenerator* (BGen)[120]. In contrast to the widely-used PYTHIA [121] event generator which creates $b\bar{b}$ pairs, BGen only simulates single B mesons without the anti–bottom quark and fragmentation products. Fragmentation processes are implemented via the Peterson fragmentation function [122].

Both for B_s^0 and B_d^0 mesons the generated p_T spectrum follows a reference spectrum derived from 2 fb^{-1} of exclusive B_d^0 decays. It is nominally valid for B^+, B_s^0 and B_d^0 mesons over the kinematic range $p_T > 6$ GeV/c and the rapidity range $|Y| < 1.5$.

For the simulation of the B decay chains the *EvtGen* [116, 117, 118] software package is employed. EvtGen is an versatile event decayer containing many detailed physics models that allow to simulate a wide range of decays. To perform the complete sequence of event generation, event decay, and a GEANT3 based [123] detector simulation the *cdfSim* [124] program is run, that invokes the corre-



(a) (b)

(c) (d)

Figure 4.5: Monte Carlo cos(θ) (a,c) and χ (b,d) distributions of $B_s^0 \to D_s^{*+}D_s^{*-}$ phase space (upper row) and re-weighted Monte Carlo (bottom row), where in the latter case weights were calculated from the standard helicity amplitude configuration $\{H_+, H_0, H_-\} = \{0.4904, 07204, 0.4904\}$.

sponding sub-modules one after another. To apply the CDF Two Track Trigger simulation the *TRGSim++* [125] program runs over the output of *cdfSim*. Finally, the *TRGSim++* output is fed through *ProductionExe* which performs event reconstruction. All programs used for simulation are built from CDF software version 6.1.4mc.m, patch *r*.

For the reasons outlined in Section 4.4, it is not correct to estimate reconstruction and selection efficiencies from Monte Carlo simulations with the D_s^+ meson exclusively decaying to $\phi\pi^+$ or $K^{*0}K^+$. Thus, existing Monte Carlo samples which already had been produced for the studied decay channels cannot be used. To have a simulation correctly reflecting the sub-resonance structure of $D_s^+ \to K^+K^-\pi^+$ decays of $B_s^0 \to D_s^+(K^+K^-\pi^+)D_s^-(K^-K^+\pi^-)$ and $B_d^0 \to D_s^+(K^+K^-\pi^+)D^-(K^+\pi^-\pi^-)$ are simulated considering the full Dalitz structure of $D_s^+ \to K^+K^-\pi^+$. Since this particular decay mode has not yet been included in the official *EvtGen* software it has to be implemented first.

Adding a new decay model into *EvtGen* is straightforward: All the program needs to calculate the decay probability at the D_s decay vertex is the squared decay amplitude as a function of randomly generated daughter four-momenta. Therefore, the module calculating the squared amplitude $|\mathcal{M}|^2$ at a given coordinate in the $(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ Dalitz plane is integrated into the *EvtGen* framework. By running fast simulations – i.e. simulations without simulating trigger and detector response – it is verified that Monte Carlo generation does not introduce any bias to decay dynamics, where we are particularly interested in the size of the mass band fractions $f_{\phi^\dagger\pi^+}$ (4.21) and $f_{K^{*0\dagger}K^+}$ (4.22). Rather than integrating over squared Dalitz plot amplitudes, these fractions are now determined by counting events lying within the corresponding mass bands and dividing them by the total number of simulated events populating the full Dalitz plane. If the total number of events generated in fast simulation is very large, the fraction of events lying inside the two mass bands should be as large as calculated directly from the model. Therefore a total number of 10^6 events populating the kinematically allowed Dalitz plot region are requested in simulation. Figure 4.6 shows the scatter plot of simulated events. The mass band fractions determined from counting simulated events agree well with those calculated from integrating the two-dimensional real function of squared amplitudes. We conclude that EvtGen simulation does not introduce any bias and are now confident that the Monte Carlo simulation will adequately reflect the decay dynamics according to the used Dalitz model.

We run full Monte Carlo simulations for $B_s^0 \to D_s^+(K^+K^-\pi^+)D_s^-(K^-K^+\pi^-)$ and $B_d^0 \to D_s^+(K^+K^-\pi^+)D^-(K^+\pi^-\pi^-)$ according to the decay tables shown in Section A.4 of the Appendix. In total, around 4.5 billion B_s^0 mesons and 1.8 billion B_d^0 mesons

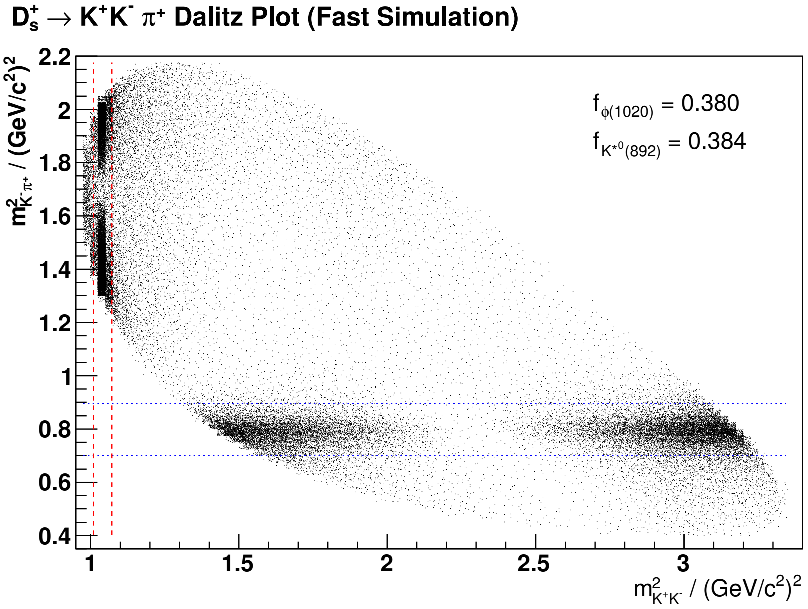


Figure 4.6: Scatter plot of simulated events populating the $D_s^+ \to K^+K^-\pi^+$ Dalitz $(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ plane.

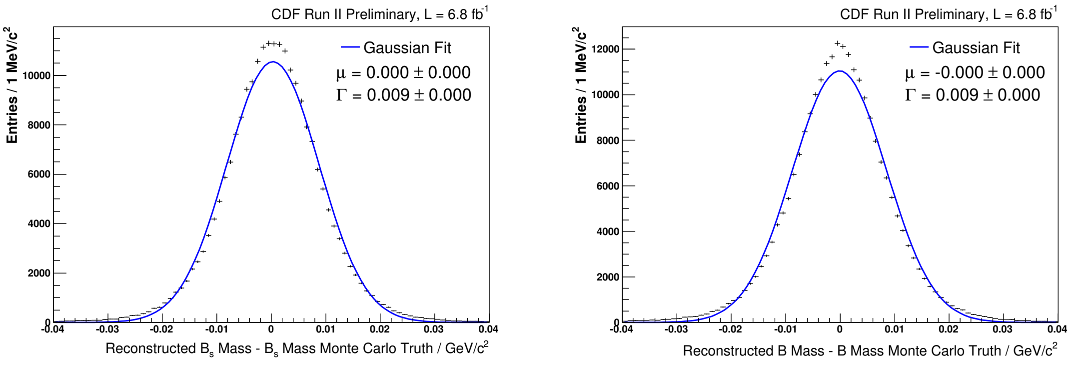
are generated and decayed. The numbers of events to be generated per run are determined by means of weights derived from integrated luminosity acquired for each run. Simulations are run for a wide range of runs (numbers 138809 to 267718, corresponding to CDF operation periods 0 through 20).

After full simulation of detector and trigger response, the same offline reconstruction and skimming procedures as discussed in Section 4.2 are applied to the samples of simulated events.

4.7 *B* Meson Transverse Momentum and Rapidity Spectrum

This chapter concludes with a comparison of the signal B^0 transverse momentum ($p_T(B^0)$) and rapidity ($Y(B^0)$) distributions observed in real data and Monte Carlo. For this purpose the decay mode $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ is used. Among the decay modes being subject to this analysis this channel is expected to hold the largest amount of signal events.

Signal candidates in data are taken from the expected $B_d^0 \to D^+D_s^-$ signal region in invariant mass space, which is defined as $\pm 3\sigma$ around the nominal mass $m_{B_d^0} = 5.279$ GeV/c^2. $1\sigma = 9$ MeV/c^2 is the effective $B_d^0 \to D^+D_s^-$ width mainly



(a) (b)

Figure 4.7: Detector resolution of reconstructed B_s^0 meson mass (a) and B_d^0 meson mass (b) estimated from simulated data.

driven by the uncertainties of reconstructed particle four-momenta, since the natural B meson width is tiny [3]. The detector resolution can be estimated from simulated data by subtracting the Monte Carlo truth B mass from the reconstructed mass and fitting a single Gaussian function to the distribution. The detector resolution is then given by the Gaussian width. Figure 4.7 displays the mass resolution fits both for the B_s^0 and the B_d^0 meson using the $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+\phi^+\pi^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^+\pi^-$ Monte Carlo samples.

To obtain a clean $p_T(B^0)$ signal distribution, the background contribution needs to be subtracted first. Background events lying within an equally-sized mass band ($m_{B_d^0 \to D^+ D_s^-} \in [5.356, 5.410]$) slightly shifted to higher invariant masses should give a reasonable representation of the background contribution in the signal region. As upcoming discussions will show (Section 6.4), the background contribution in invariant mass is not flat but rather has a decreasing exponential shape. Therefore, prior to background subtraction upper sideband events are re-weighted using

$$w(m) = \frac{P^i_{bg}(m - \Delta m)}{P^i_{bg}(m)}, \tag{4.50}$$

where m is the floating invariant B_d^0 mass in upper sideband and Δm is the offset between the upper edge of the signal region and the lower edge of the sideband region. The parameters (slope, constant offset) of the background parameterization function P^i_{bg} are obtained from a fit to the channel $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ (for fit details see Section 6.4). To retain a sufficient amount of background, both for the purpose of fitting and background subtraction a very soft neural network cut of $NN > -0.8$ is applied (for details on the pre- and the Neural Network

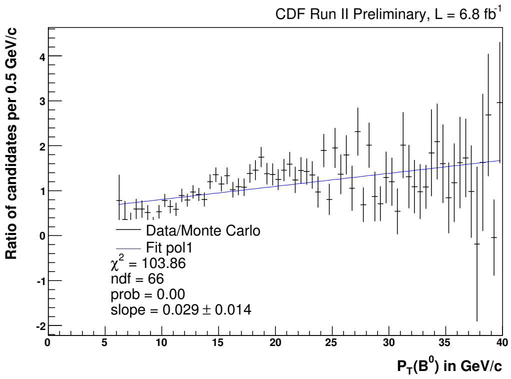


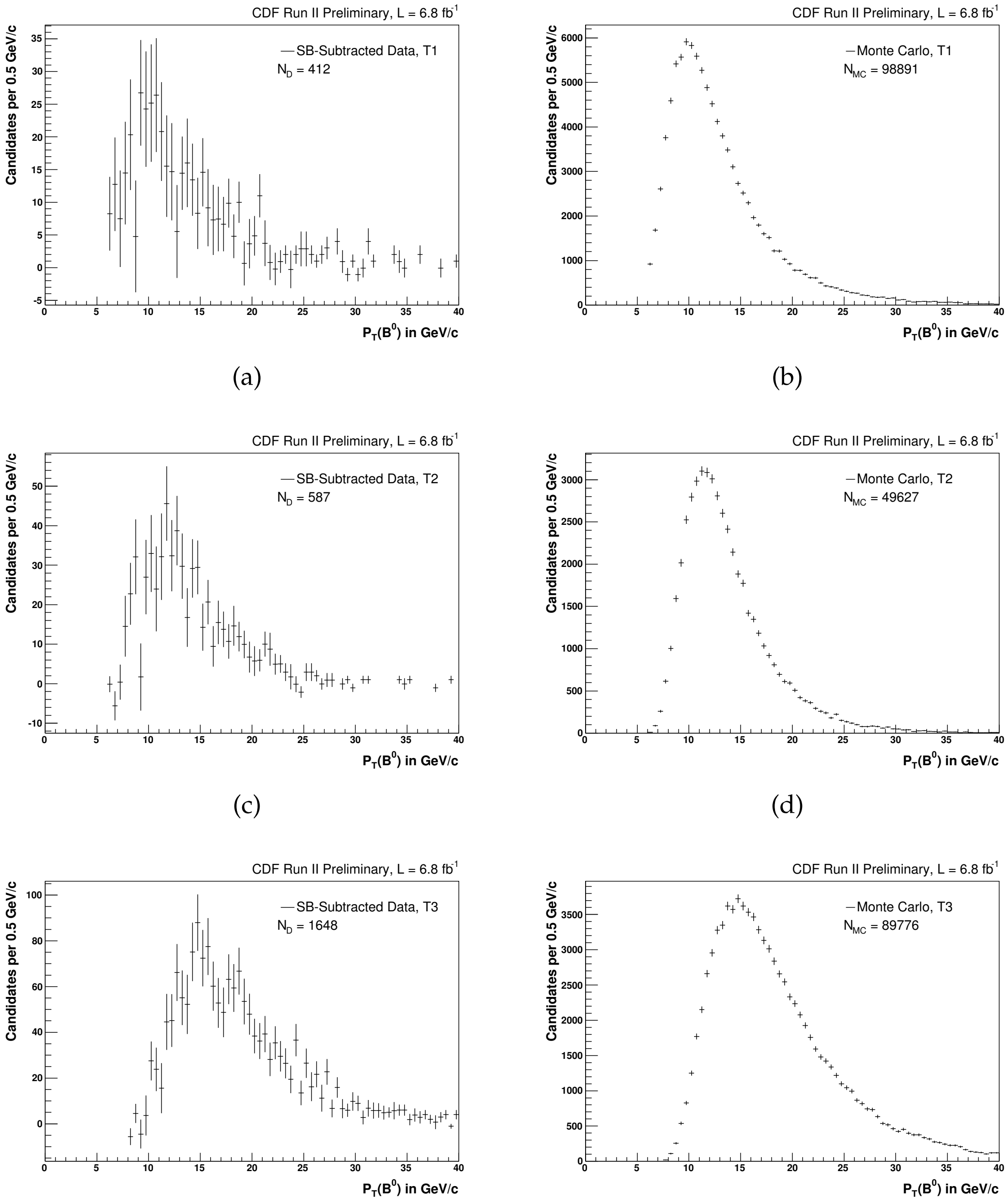
Figure 4.8: Normalized ratio of $p_T(B^0)$ distributions observed in side-band subtracted data and Monte Carlo. Simulated data has not yet been corrected for true Two-Track Trigger fractions.

	Data		**Monte Carlo**		
Trigger Sub-Path	**Events**	**Fraction**	**Events**	**Fraction**	**Weight w_t**
1	412	0.156	98891	0.415	0.375
2	587	0.222	49627	0.208	1.065
3	1648	0.622	89776	0.377	1.652

Table 4.5: Comparison of the two-track trigger sub-path fractions in Data and Monte Carlo. The right-most column gives the weight w_t simulated data is corrected with.

selection please refer to sections 5.2 and 5.3). After background subtraction, the data p_T histogram distribution is divided by the one observed in Monte Carlo. If simulated data provides a reasonable representation of the true p_T signal distribution, one expects the p_T ratio to be flat. Apparently (Figure 4.8), transverse momentum is not well described by simulated data. To follow up this matter both data and Monte Carlo are sub-divided into the three exclusive sub-samples according to the different trigger scenarios of data taking. These were defined at the very beginning of this chapter. For each of the sub-samples the $p_T(B^0)$ distributions of side-band subtracted data and Monte Carlo are plotted and the number of events extracted (Figure 4.9). According to Table 4.5, particularly the shares of the trigger sub-samples 1 and 3 do not agree well in data and Monte Carlo. We thus re-weight simulated data using the weights given in the last column of Table 4.5. If not stated otherwise, this kind of Monte Carlo correction is applied whenever utilizing simulated data.

Figures 4.10(a) and 4.10(b) show the normalized ratio of the number of data over corrected Monte Carlo events in p_T bins of 500 MeV/c. The hypothesis

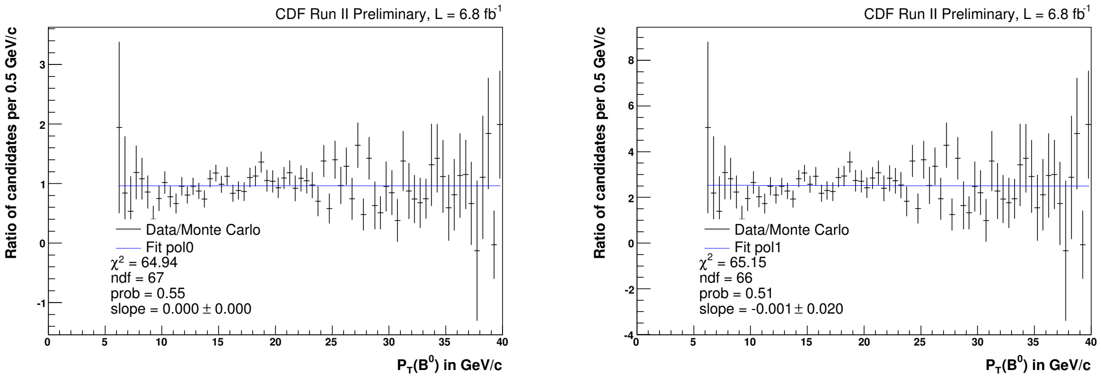


(a) (b)

(c) (d)

(e) (f)

Figure 4.9: From first to third row, $p_T(B^0)$ distributions of the exclusive Two-Track Trigger sub-samples (labeled by TriggerFlags $T1$, $T2$, $T3$) in side-band subtracted data (left column) and Monte Carlo (right column).

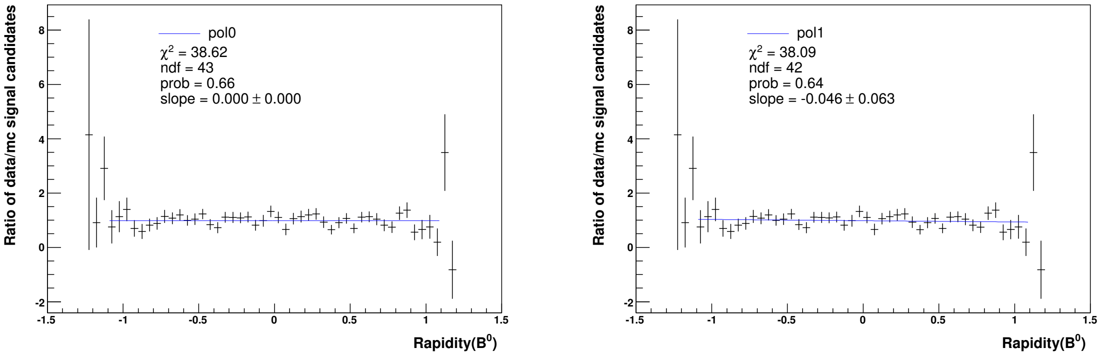


(a) (b)

Figure 4.10: Normalized ratio of $p_T(B^0)$ distributions observed in side-band subtracted data and Two-Track Trigger corrected Monte Carlo using the decay channel $B^0_d \rightarrow D^+D^-_s \rightarrow K^-\pi^+\pi^+\phi^+\pi^-$ with a constant (a) and a 1st order polynomial fitted to it.

of p_T compliance in Monte Carlo and data is tested by fitting a straight line to the ratio of p_T ratio distribution (Figure 4.10(a)). Fitting a 1st order polynomial to the same distribution (Figure 4.10(b)) yields a slope being compatible with 0 and a comparable fit quality. We therefore conclude that after Two-Track Trigger correction Monte Carlo does describe the $p_T(B^0)$ distribution observed in data reasonably well. The change in p_T slope can be attributed to the p_T dependance of the Two-Track Trigger sub-paths.

In the same way the agreement of the $Y(B^0)$ distributions observed in data and Monte Carlo (Figures 4.11(a) and 4.11(b)) is validated.



(a) (b)

Figure 4.11: Ratio of $Y(B^0)$ distributions observed in side-band subtracted data and Monte Carlo using the decay channel $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ with a constant fit (a) and 1st order polynomial fit (b).

5

Signal Selection

5.1 Overview

After building skimmed Ntuples from reconstructed hadronic streams, data and Monte Carlo are ready for the next analysis stage: classification of data and selection of a clean sample of signal candidates to be used for the extraction of branching fractions. Data selection is a two-step process: Firstly, data is required to pass loose pre-cuts applied to several kinematical quantities. This already removes a large fraction of combinatorial background, substantially reducing the size of the data samples to be further processed. In contrast to pre-selection, final selection of signal candidates is not done on the basis of a rectangular cut optimization but using a neural network based multivariate approach that accounts for variable correlations.

5.2 Track Quality and Pre-selection Requirements

In order to improve track quality, both data and Monte Carlo flat Ntuples share the following pre-cuts per track:

- Number of COT stereo hits ≥ 10
- Number of COT axial hits ≥ 10
- Number of Silicon Stereo + Silicon axial hits ≥ 3
- Minimum track transverse momentum $p_T \geq 0.35$ GeV/c

Additionally, loose pre-cuts on certain kinematic variables are applied to sort out obvious background events. This enables a much smoother operation of artificial

$B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+\phi^{\dagger}\pi^-$		$B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^{\dagger}\pi^-$	
$q\left(\pi_{D_s^{(1)}}\right) \times q\left(\pi_{D_s^{(2)}}\right)$	$= -1$	$q\left(\pi_{D_s}\right) + q\left(K_D\right) + q\left(\pi_D^{(1)}\right) + q\left(\pi_D^{(2)}\right)$	$= 0$
		$q\left(\pi_D^{(1)}\right) \times q\left(\pi_D^{(2)}\right)$	$= 1$
$p_T\left(B_s^0\right)$	> 6.0 GeV/c	$p_T\left(B_d^0\right)$	> 6.0 GeV/c
$\left\|d_0\left(B_s^0\right)\right\|$	< 0.012 cm	$\left\|d_0\left(B_d^0\right)\right\|$	< 0.012 cm
$L_{xy}/\sigma_{L_{xy}}\left(B_s^0\right)$	> -2.0	$L_{xy}/\sigma_{L_{xy}}\left(B_d^0\right)$	> -2.0
$L_{xy}/\sigma_{L_{xy}}\left(D_s^{(1)}\right)$	> 3.0	$L_{xy}/\sigma_{L_{xy}}\left(D_s\right)$	> 3.0
$L_{xy}/\sigma_{L_{xy}}\left(D_s^{(2)}\right)$	> 3.0	$L_{xy}/\sigma_{L_{xy}}\left(D\right)$	> 3.0
$\chi^2_{r\phi}(D_s^{(1)})$	< 20	$\chi^2_{r\phi}(D_s)$	< 20
$\chi^2_{r\phi}(D_s^{(2)})$	< 20	$\chi^2_{r\phi}(D)$	< 20

Table 5.1: Applied pre-cuts for the studied decay channels having $\phi^{\dagger}\pi^+$ in the final state. See Section A.1 in the Appendix for definition of variables.

Neural Networks that will be used for final classification and selection of $B_s^0 \to D_s^+ D_s^-$ and $B_d^0 \to D^+ D_s^-$ events. Moreover, candidates reconstructed from tracks with wrong charge combinations are rejected. The charge constraints depend on the final pion and kaon states of the particular decay mode and are listed in Tables 5.1 and 5.2 along with all the other applied pre-selection requirements.

The selection of variables and cut values were chosen on the basis of the following criteria: The cut on the B meson transverse momentum was selected because of the nominally valid kinematic region of the input p_T spectrum used for Monte Carlo generation. The $L_{xy}/\sigma_{L_{xy}}$ requirement placed on the $D_{(s)}$ mesons is a confirmation of the offline reconstruction cuts, except for the decay channel $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ where this cut had not been made on reconstruction level. The standard off-line reconstruction criteria exhibited another issue related to the χ^2 cut of the D_s vertex fit in the r-ϕ plane. This inconsistency was eliminated by choosing a common requirement of $\chi^2_{r\phi}(D_s^{(1,2)}) < 20$.

5.3 Neural Network Selection

For final event selection this analysis makes use of the NeuroBayes [126] program. This multivariate analysis software package combines a sophisticated variable pre-processing algorithm with a feed-forward three-layer artificial neural network. Each network layer comprises a set of artificial neurons, or nodes, that are interconnected with every node of the neighboring layer. The number of input

$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$		$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$	
$q\left(\pi_{D_s^{(1)}}\right) \times q\left(K_{D_s^{(2)}}\right)$	$= -1$	$q\left(\pi_{D_s}\right) + q\left(K_D\right) + q\left(\pi_D^{(1)}\right) + q\left(\pi_D^{(2)}\right)$	$= 0$
$q\left(K_{K^{*0\dagger}K}\right) \times q\left(K_{D_s}\right)$	$= -1$	$q\left(K_{K^{*0\dagger}}\right) \times q\left(K_{D_s}\right)$	$= -1$
		$q\left(\pi_D^1\right) \times q\left(\pi_D^2\right)$	$= 1$
$p_T\left(B_s^0\right)$	> 6.0 GeV/c	$p_T\left(B_d^0\right)$	> 6.0 GeV/c
$\left\vert d_0\left(B_s^0\right)\right\vert$	< 0.012 cm	$\left\vert d_0\left(B_d^0\right)\right\vert$	< 0.012 cm
$L_{xy}/\sigma_{L_{xy}}\left(B_s^0\right)$	> -2.0	$L_{xy}/\sigma_{L_{xy}}\left(B_d^0\right)$	> -2.0
$L_{xy}/\sigma_{L_{xy}}\left(D_s^{(1)}\right)$	> 3.0	$L_{xy}/\sigma_{L_{xy}}\left(D_s\right)$	> 3.0
$L_{xy}/\sigma_{L_{xy}}\left(D_s^{(2)}\right)$	> 3.0	$L_{xy}/\sigma_{L_{xy}}\left(D\right)$	> 3.0
$\chi^2_{r\phi}(D_s^{(1)})$	< 20	$\chi^2_{r\phi}(D_s)$	< 20
$\chi^2_{r\phi}(D_s^{(2)})$	< 20	$\chi^2_{r\phi}(D)$	< 20

Table 5.2: Applied pre-cuts for the studied decay channels having $K^{*0\dagger}K$ in the final state. See Section A.1 in the Appendix for definition of variables.

nodes corresponds to the amount of discriminating variables chosen. Since the discrimination between signal and background in data poses a binary classification problem the output layer consists of a single node only[1]. Figure 5.1 gives a schematic view of a simple three-layer feed-forward network. In applying a neural network the N-dimensional space spanned by the input variables is mapped onto the single output node, providing a powerful one-dimensional discriminator. Instead of cutting on several discriminating variables, this single discriminator is then used to select events having a particular signature from an unclassified data sample. A striking advantage over rectangular cut-based analysis techniques comes by the fact that correlations among the individual input variables are taken into account.

Before a neural network is ready to operate, an optimal net configuration needs to be established by machine learning, which in this context is also called network training. For a binary separation problem, this procedure requires training patterns with known classification to be in place. During the learning process the connection strengths among the individual nodes are iteratively adjusted until the network output matches the known classification encoded in the *target* variable. To further optimize the prediction and generalization capabilities, at the same time preventing the network from learning certain patterns by heart, advanced network learning algorithms like the one embedded in the NeuroBayes

---

[1] As an additional program function, NeuroBayes also provides the estimation of continuous probability densities. For this work however binary classification is relevant only.

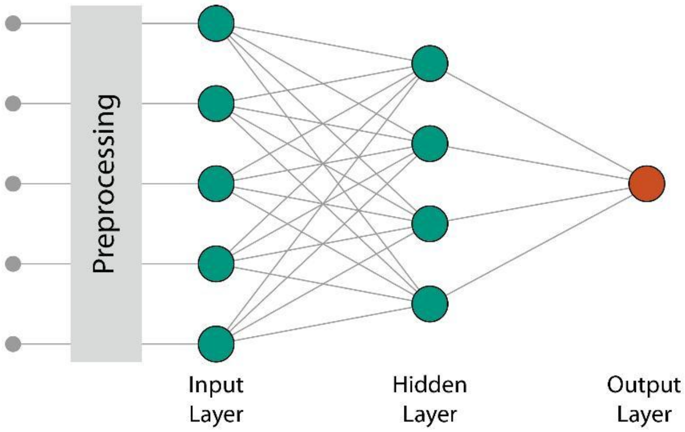


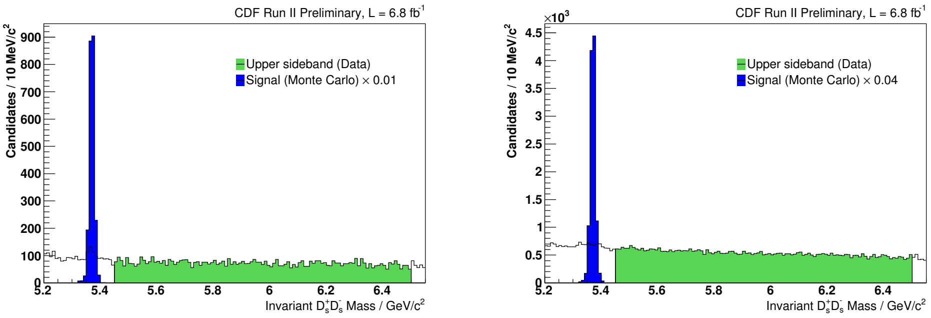
Figure 5.1: Schematic representation of a three-layer feed-forward neural network. Prior to network training the input variables can be preprocessed to ease the network training process.

package incorporate further sophisticated measures. These are described in detail in Ref. [126].

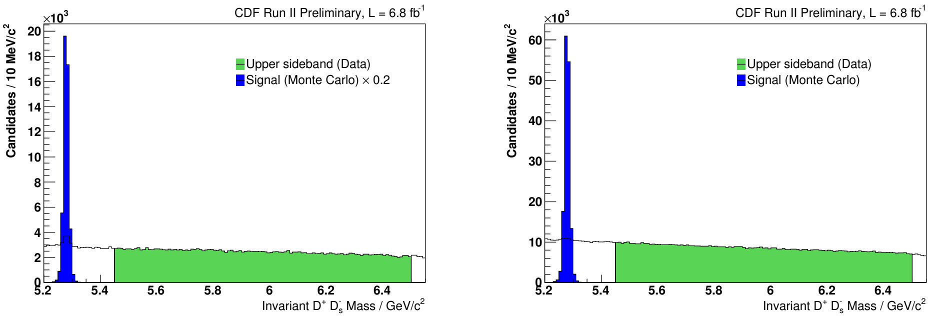
Once the training has succeeded, the network topology and the weights among the nodes are saved to what is commonly referred to as the *expertise* of the neural network. When applied to unclassified data, the vector of input variables is fed into the network and transformed into a single floating number based on the expertise.

5.3.1 Composition of Training Data

Training patterns comprise two distinctive samples which are a priori known to be composed of signal or background events only. Signal events are provided by means of Monte Carlo simulations, where only fully and correctly reconstructed signal is used. The other sample has to reflect the random combinatorial behavior of non-signal events. Since current models used in Monte Carlo simulations are not able to adequately model the complex quark production and hadronization processes, background patterns are taken from regions in real data which are known to be completely free of any signal like component. A common approach is to use data from the lower and upper sidebands (i.e. events to the left and to the right of fully reconstructed $B_{(s)}$ meson signal) of the invariant mass spectra in order to emulate the behavior of combinatorial background lying inside the signal region. However, as later discussions will show (Sections 6.3 and 6.4), the lower mass region is populated by so called partially reconstructed events, which exhibit signal-like signatures. The upper sideband region does not contain any signal-like events and thus can be used for network training. For all the



(a) $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ (b) $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$



(c) $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ (d) $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$

Figure 5.2: Invariant mass distributions of pre-selected data. The upper mass regions used as background training patterns are highlighted by the green color. Fully reconstructed simulated signal used as signal training patterns are overlayed in blue color.

studied decay channels upper sideband data ranging from 5.45 to 6.5 GeV/c^2 are used as background training patterns. Figure 5.2 shows the reconstructed mass distributions containing all events having passed the pre-selection requirements. The background training regions are highlighted by the green mass band, with the samples of fully reconstructed Monte Carlo samples overlayed in blue.

Training and selection are not done on the basis of four networks (one network per studied decay channel), but using two networks only: Exploiting the similar decay topologies of particular modes, one joint network is trained for those $B_{(s)}^0$ decay modes reconstructed from $D_s^+ \to \phi^\dagger \pi^+$ ($B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$), and one training for those decay modes with the (second) D_s decaying into $K^{*0\dagger} K$ ($B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$). This reduces the amount of networks required from four to two. The main motivation for using joint B_s^0/B_d^0 networks is to be in line with the way the figures of interest will be determined: The branching fractions of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$

decays will be extracted from data by measuring the ratio of branching fractions

$$f_{D_s^{(*)}D_s^{(*)}} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})}{\mathcal{B}(B_d^0 \to D^+ D_s^-)}, \tag{5.1}$$

where $B_d^0 \to D^+ D_s^-$ is used as normalization channel (please refer to Chapter 6 for details). Having common selections for the signal and normalization channels will make systematic studies easier, and selection efficiency related systematic effects partly cancel out in the ratio of topologically similar decays.

5.3.2 Input Variables

To achieve an optimal neural network training result, a set of variables that allows to efficiently exploit the information contained in the training patterns needs to be selected. The set of variables that had been used for the rectangular cut-based optimization procedure in the former CDF analysis [37] certainly forms a sound basis for a neural network training. In addition to the $B_{(s)}$ and $D_{(s)}$ meson variables originally used, kinematical quantities of the final state daughter particles and *Particle Identification* (PID) variables, that reflect the probability of a given track to be a pion, kaon or proton, are included. A full list of the training variables will be given below.

The B_s^0 and the B_d^0 meson both have very similar masses, and exhibit long and virtually equal lifetimes, thus the kinematical $B_{(s)}^0$ quantities have very similar features. Therefore, the joint B_s^0/B_d^0 neural network training setups use common sets of kinematical and fit quality $B_{(s)}^0$ variables. The second type of input variables are kinematical, vertex fit quality and PID variables of one D_s meson and its daughter particles ($D_s^+ \to \phi^\dagger \pi^+ \to K^+ K^- \pi^+$ on the one hand, $D_s^+ \to K^{*0\dagger} K K^+ \to K^+ \pi^- K^+$ on the other hand). The drawback of performing one joint B_s^0/B_d^0 network training for two channels is given by the fact that one can only consider one branch of the meson decay chain, since the second charm meson is a D_s decaying into $\phi^\dagger \pi$ ($K^{*0\dagger} K$) in the case of B_s^0 decays and a D decaying into $K\pi\pi$ for B_d^0 decays. This is partially compensated by adding the deviation in reconstructed mass of the second $D_{(s)}$ meson with respect to the corresponding PDG value. Tables 5.4 and 5.4 give a schematic overview of the input variables entering the network trainings.

5.3.3 Neural Network Training Results

In this section the results of the neural network trainings are presented. The compilation of results is limited to the essential input and output information.

	B_s^0	$D_s^{(1)}$	$D_s^{(2)}$	$\phi^\dagger\pi$	$\phi^\dagger\pi$
Index	B_d^0	$D_s^{(1)}$	D	$\phi^\dagger\pi$	$K\pi\pi$
2	$L_{xy}/\sigma_{L_{xy}}$				
3	p_T				
4	$\lvert d_0\rvert$				
5	$\chi^2_{r\phi}$				
6	prob				
7		$L_{xy}/\sigma_{L_{xy}}$			
8		$L_{xy}(B_{(s)} \leftarrow D_s)$			
9		p_T			
10		$\lvert d_0\rvert$			
11		$\chi^2_{r\phi}$			
12		prob			
13				$L_{xy}(D_s \leftarrow \phi^\dagger)$	
14				$d_0^{lts}/\sigma_{d_0}\,(\pi)$	
15				PID.ratioPion(π)	
16				$d_0^{lts}/\sigma_{d_0}\left(K^{(1)}\right)$	
17				PID.ratioKaon($K^{(1)}$)	
18				$d_0^{lts}/\sigma_{d_0}\left(K^{(2)}\right)$	
19				PID.ratioKaon($K^{(2)}$)	
20				min(p_T)	
21				min(d_0/σ_{d_0})	
22				m($K^{(1)}K^{(2)}$)	
23				m($K^{(1)}\pi$)	
24				m($K^{(2)}\pi$)	
25			$m_{rec} - m_{pdg}$		

Table 5.3: Schematic overview of the variables entering the joint $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^- / B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ network. See Appendix A.1 for definition of variables. Index number 1 is reserved for the target variable.

	B_s^0	$D_s^{(1)}$	$D_s^{(2)}$	$K^{*0\dagger}K$	$\phi^\dagger\pi$
Index	B_d^0	$D_s^{(1)}$	D	$K^{*0\dagger}K$	$K\pi\pi$
2	$L_{xy}/\sigma_{L_{xy}}$				
3	p_T				
4	$\lvert d_0 \rvert$				
5	$\chi^2_{r\phi}$				
6	prob				
7		$L_{xy}/\sigma_{L_{xy}}$			
8		$L_{xy}(B_{(s)} \leftarrow D_s)$			
9		p_T			
10		$\lvert d_0 \rvert$			
11		$\chi^2_{r\phi}$			
12		prob			
13				$L_{xy}(D_s \leftarrow K^{*0\dagger})$	
14				$d_0^{lts}/\sigma_{d_0}\left(K^{(1)}\right)$	
15				PID.ratioKaon($K^{(1)}$)	
16				$d_0^{lts}/\sigma_{d_0}\left(K^{(1)}\right)$	
17				PID.ratioKaon($K^{(2)}$)	
18				$d_0^{lts}/\sigma_{d_0}(\pi)$	
19				PID.ratioPionπ)	
20				min(p_T)	
21				min(d_0/σ_{d_0})	
22				m($K^{(1)}K^{(2)}$)	
23				m($K^{(1)}\pi$)	
24				m($K^{(2)}\pi$)	
25			$m_{rec} - m_{pdg}$		

Table 5.4: Schematic overview of the variables entering the joint $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+ K^{*0\dagger}K^-/B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0\dagger}K^-$ network. The list of variables used is formally identical to one of the $\phi^\dagger\pi$ network. The difference arises by the final states $\phi^\dagger\pi$ and $K^{*0\dagger}K$, which are reconstructed in two non-overlapping mass bands of $D_s^+ \to K^+K^-\pi^+$ phase space.

Rank	Name	This only	Add Signi	Loss	Global Corr	Index
-	Target	-	-	-	-	1
1	$L_{xy}/\sigma_{L_{xy}}(B^0_{(s)})$	507.98	507.98	117.06	83.51	2
2	$\min(p_T)$	478.58	243.26	78.49	74.71	20
3	$\mathrm{prob}(B^0_{(s)})$	411.13	182.55	40.61	85.87	6
4	$\mathrm{PID.ratioKaon}(K^{(2)}_{D^{(2)}_s})$	443.15	130.66	51.40	75.61	17
5	$\min(d_0/\sigma_{d_0})$	454.80	93.23	73.75	73.92	21
6	$\mathrm{PID.ratioKaon}(K^{(2)}_{D^{(2)}_s})$	431.52	68.18	50.40	72.93	19
7	$\left\|d_0(B^0_{(s)})\right\|$	245.80	60.59	47.17	37.65	4
8	$\chi^2_{r\phi}(B^0_{(s)})$	446.08	49.39	35.36	85.41	5
9	$L_{xy}(B^0_{(s)} \leftarrow D^{(2)}_s)$	159.00	41.53	39.99	35.71	8
10	$\mathrm{m}(K^{(2)}K^{(2)})$	266.26	33.89	29.07	43.30	22
11	$\Delta\mathrm{m}(D^{(2)}_{(s)})$	263.56	31.94	27.61	54.64	25
12	$\mathrm{m}(K^{(2)}\pi)$	266.32	29.42	10.00	83.26	24
13	$L_{xy}(D^{(2)}_s \leftarrow \phi)$	347.85	27.89	28.94	84.49	13
14	$p_T(B^0_{(s)})$	395.91	24.94	20.62	71.76	3
15	$p_T(D^{(2)}_s)$	454.46	16.95	20.13	86.89	9
16	$\chi^2_{r\phi}(D^{(1)}_2)$	205.11	14.77	9.13	78.83	11
17	$\mathrm{m}(K^{(2)}\pi)$	265.17	9.62	9.63	83.26	23
18	$d^{lts}_0/\sigma_{d_0}\left(K^{(2)}_\phi\right)$	391.26	8.99	8.37	79.40	18
19	$\left\|d_0(D^{(2)}_s)\right\|$	148.41	7.21	8.03	58.61	10
20	$d^{lts}_0/\sigma_{d_0}\left(K^{(2)}_\phi\right)$	398.51	5.25	7.51	81.19	16
21	$L_{xy}/\sigma_{L_{xy}}(D^{(2)}_s)$	407.23	5.69	8.41	92.16	7
22	$d^{lts}_0/\sigma_{d_0}\left(\pi_{D^{(2)}_s}\right)$	312.91	6.70	6.76	81.21	14
23	$\mathrm{prob}(D^{(2)}_s)$	153.63	2.42	2.41	79.72	12
-	2σ threshold	-	-	-	-	-
24	$\mathrm{PID.ratioPion}(\pi_{D^{(2)}_s})$	288.52	0.68	0.68	62.85	15

Table 5.5: Input variables of the combined network trained for the decays $B^0_s \to D^+_s D^-_s \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ and $B^0_d \to D^+D^-_s \to K^-\pi^+\pi^+\phi^\dagger\pi^-$, ranked by their significance. $K^{1,2}$ denotes the first or second kaon from a ϕ meson decay. See Appendix A.1 for variable definitions. The quantity *This only* is the correlation of a variable to the target multiplied by $\sqrt{n}$ (with n being the training samples size), without taking into account other variables. The ranking of variables is made based on the quantity *Add Signi* which gives the amount of information this variable adds to the overall performance. Variables ranked below the 2σ threshold are not used in the network training. *Loss* indicates the correlation of a variable to the target multiplied by $\sqrt{n}$ when the given variable is removed from the variable set. *Global Corr* denotes the global correlation of a given variable to all the others given in percent. *Index* marks a variable's column / row position in the correlation matrix.

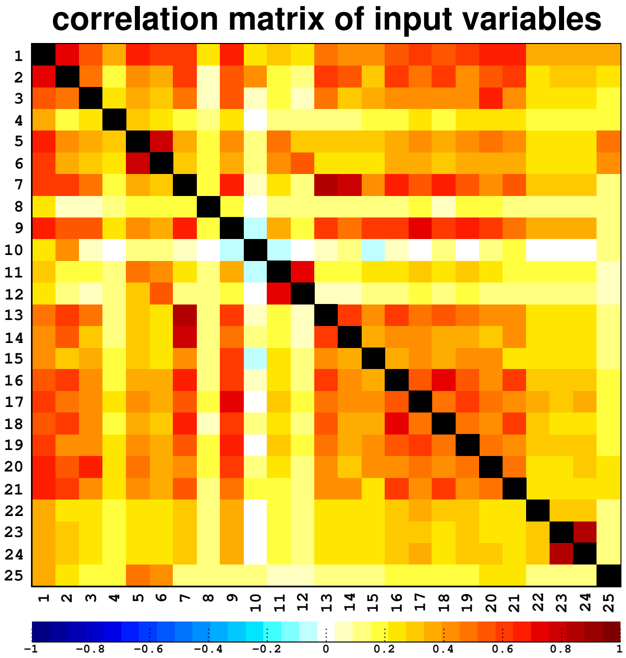
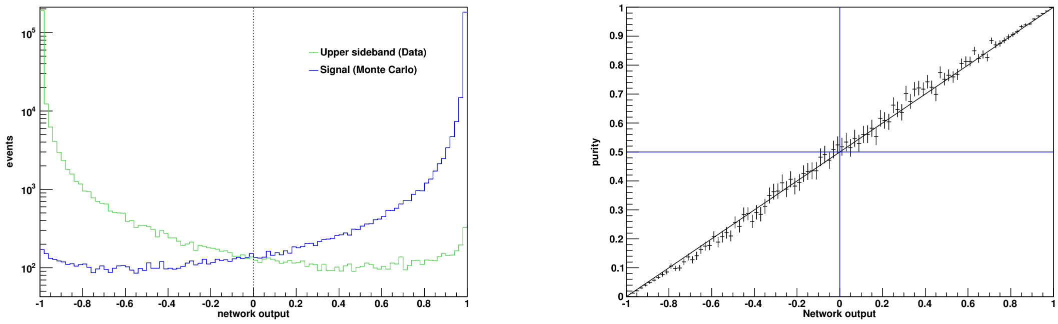


Figure 5.3: Correlation matrix of the input variables used for the combined network training of the decays $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+\phi^{\dagger}\pi^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^{\dagger}\pi^-$



(a) Network classification of training data.

(b) Purity over network output.

Figure 5.4: Two of the training graphs indicating performance and quality of the network trained for $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+\phi^{\dagger}\pi^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^{\dagger}\pi^-$

First, training variables are presented ordered by their significance (Tables 5.5 and 5.6). The ranking of the input variables according to their significance is part of the pre-processing of the NeuroBayes learning program (Teacher), and is done in the following way: After the correlation matrix is computed for the full set of N input variables, one variable at a time is removed and the correlation to the target is re-calculated. The least significant variable, i.e. the variable causing the least loss of information, is detached from the input set. The correlation matrix is re-computed for the set of $N-1$ input variables, and the procedure of removing one variable one after another is repeated. Again, the least significant variable is discarded. This procedure is repeated until one single variable, i.e. the variable holding the strongest discriminating power, remains. For what concerns the ranking shown here, all linear dependence with the target has been rotated into the first variable (the first training variable, not the target) after de-correlation of input variables. For the network training only variables were used which have a statistical significance higher than 2σ. The variables used in the training are computed as a combination of the original input variables after the correlation matrix has been diagonalized by means of iterative Jacobi rotations. For more information it should be referred to [126] and the supporting technical documents coming with the NeuroBayes software package [127].

The correlations between the variables are illustrated by color-coded correlation matrices (Figures 5.3 and 5.5). Finally, two of the characteristic graphs indicating a neural network's quality and performance are given (Figures 5.4 and 5.6): In the first graph (a) a network's capability to separate between signal and background events is verified on the training samples. In the second plot (b) the linearity of the purity $P(NNout) = N_s(NNout)/(N_S + N_B)(NNout)$ as a function of the neural network threshold *NNout* is tested. For a well trained network one expects the training data points to lie on the diagonal.

5.4 Finding Optimal Neural Network Working Points

By running a routine called *Expert* the network expertise is applied to unclassified data. In doing so, an additional variable *NNout* holding an event-by-event classification encoded in a real number in the range $[-1, 1]$ is written onto the flat ntuples. In this notation, a classification of $NNout < 0$ means that a given event is more likely to be a background event, while signal-like events are classified by $NNout > 0$. The selection of signal candidates is done by only retaining candidates having a neural network classification larger than a certain threshold, providing

Rank	Name	This only	Add Signi	Loss	Global Corr	Index
-	Target	-	-	-	-	1
1	$L_{xy}/\sigma_{L_{xy}}\left(B^0_{(s)}\right)$	650.99	650.99	126.70	84.78	2
2	$\chi^2_{r\phi}\left(B^0_{(s)}\right)$	598.25	321.76	53.43	85.72	5
3	PID.ratioKaon($K_{D_s^{(1)}}$)	568.95	223.18	85.28	64.62	15
4	PID.ratioKaon($K_{K^{*0}}$)	532.82	154.14	69.13	64.62	17
5	$d_0^{lts}/\sigma_{d_0}\left(K_{D_s^{(1)}}\right)$	579.21	116.49	50.87	72.84	14
6	min(p_T)	496.02	98.86	52.80	68.23	20
7	$L_{xy}(B^0_{(s)} \leftarrow D_s^{(1)})$	230.78	41.05	67.87	35.10	8
8	m($K^{(2)}\pi$)	469.71	83.06	12.34	95.10	24
9	min(d_0/σ_{d_0})	555.11	57.50	69.10	71.75	21
10	$\left\|d_0(B^0_{(s)})\right\|$	316.76	63.06	50.71	38.88	4
11	prob($B^0_{(s)}$)	542.94	50.76	41.77	85.53	6
12	$p_T(D_s^{(1)})$	582.27	43.70	44.62	83.54	9
13	$p_T(B^0_{(s)})$	520.21	35.68	40.38	72.22	3
14	$L_{xy}(D_s^{(1)} \leftarrow K^{*0})$	433.98	38.19	43.65	84.67	13
15	m($K^{(1)}\pi$)	249.91	34.05	33.38	27.27	23
16	Δm($D^{(2)}_{(s)}$)	334.59	26.79	24.27	54.01	25
17	$L_{xy}/\sigma_{L_{xy}}(D_s^{(1)})$	541.99	21.83	32.01	90.23	7
18	$L_{xy}(D_s^{(1)} \leftarrow K_{K^{*0})}$	417.11	19.29	24.13	71.97	16
19	$d_0^{lts}/\sigma_{d_0}\,(\pi_{K^{*0}})$	344.83	17.50	17.50	65.42	18
20	m($K^{(1)}K^{(2)}$)	465.15	12.34	12.49	95.06	22
21	$\left\|d_0(D_s^{(1)})\right\|$	253.10	12.34	8.68	66.49	10
22	$\chi^2_{r\phi}\left(D_s^{(1)}\right)$	321.29	8.82	8.05	79.16	11
23	PID.ratioPion($\pi_{K^{*0}}$)	258.74	7.59	3.86	48.82	19
24	prob($D_s^{(1)}$)	235.88	3.59	3.59	79.16	12

Table 5.6: Significance-ranked input variables of the joint $B^0_s \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ / $B^0_d \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$ network. See Appendix A.1 for variable definitions.

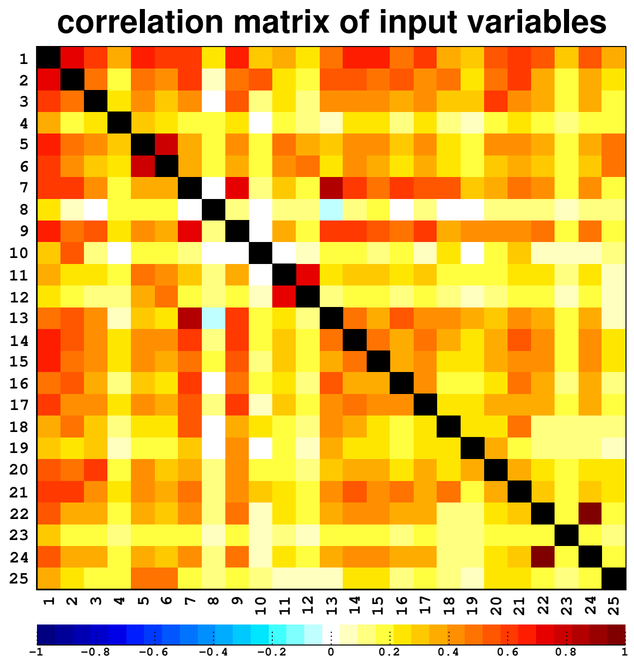
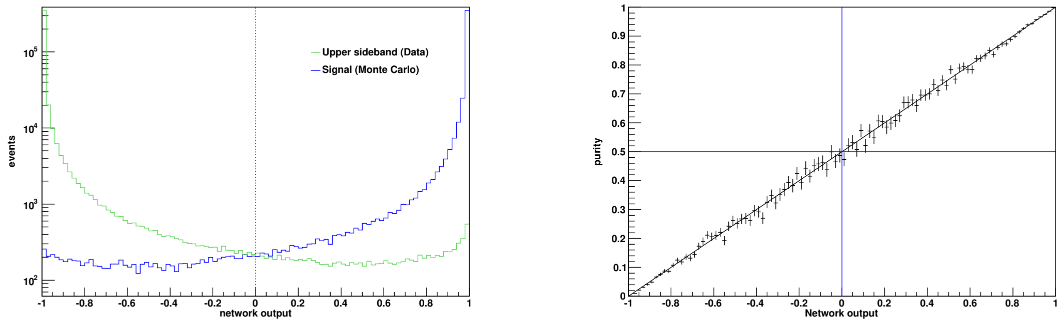


Figure 5.5: Correlation matrix of the input variables used for the combined network training of the decays $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0\dagger} K^-$.



(a) Network classification of training data.

(b) Purity over network output.

Figure 5.6: Graphs indicating performance and quality of the network trained for $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0\dagger} K^-$.

for an effective elimination of the vast amount of background events.

Since one is interested in minimizing the statistical uncertainty of the branching fraction measurement a common approach is to choose a cut on the neural network output that maximizes the figure of merit $S = N_S^{exp}/\sqrt{N_{S+B}^{Data}}$, which is usually referred to as *Significance*. Here, N_{S+B}^{Data} denotes the total yield comprising signal and background in a pre-defined B_s^0 signal region, $[5.343 < m(B_s^0) < 5.397]$. The mass window corresponds to the central B_s mass $\pm 3\sigma$, where $1\sigma = 9$ MeV is the effective width mainly driven by detector resolution derived from Monte Carlo. Since the signal yields for the studied $B_s^0 \to D_s^+ D_s^-$ channels are expected to be by far smaller than those for $B_d^0 \to D^+ D_s^-$ the statistical uncertainty of the branching fraction result is expected to be dominated by the former decay modes. Therefore, the signal optimization procedure is done for $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ and $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ only.

When optimizing the figure of merit no use is made of the fitting method that will be described in Section 6. This is to avoid a potential bias which could be introduced by fluctuations of the fit result, particularly in the case of low signal statistics. Therefore N_S^{exp} is chosen to be the expected number of signal events for a given network cut, deduced from the general relation (6.18):

$$N^{i,exp}_{B_s^0 \to D_s^+ D_s^-} = N^{tot}_{B_s^0} \mathcal{B}(B_s^0 \to D_s^+ D_s^-) \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-},$$

$\epsilon_{B_s^0 \to D_s^+ D_s^-}$ is the combined reconstruction and selection efficiency which is calculated from simulated data having a larger Neural Network classification than a given value (for a more accurate definition of the efficiency term please refer to the immediately following Section 5.5). In this sense, the figure of merit S is the simulated signal yield scaled down by the current world average intermediate and final state branching fractions, divided by the square root of the total number of candidates inside the specified signal region in data, N_{S+B}^{Data}. To obtain a reasonable scaling for S, $N^{tot}_{B_d^0}$, the total number of B_d^0 mesons produced at the Tevatron, is arbitrarily set to 10^{11}.

Figure 5.7 shows the figure of merit S as a function of the Neural Network threshold for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$. Because no maximum significance is found for $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ (Figure 5.8) in the examined Neural Network range the network scan is focussed on the upper region, while the granularity of the scan is increased. From the significance curves the values $NNout > 0.9$ and $NNout > 0.98$, respectively, are extracted as Neural Network working points.

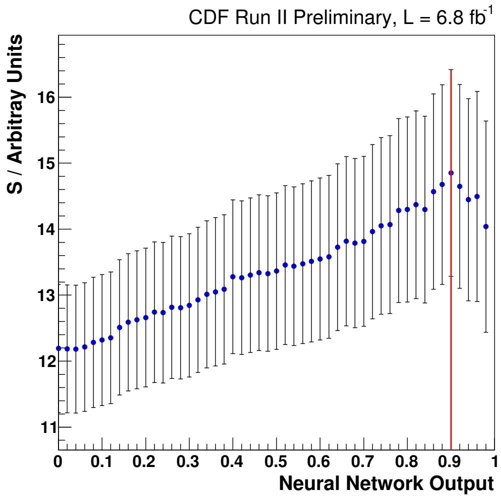
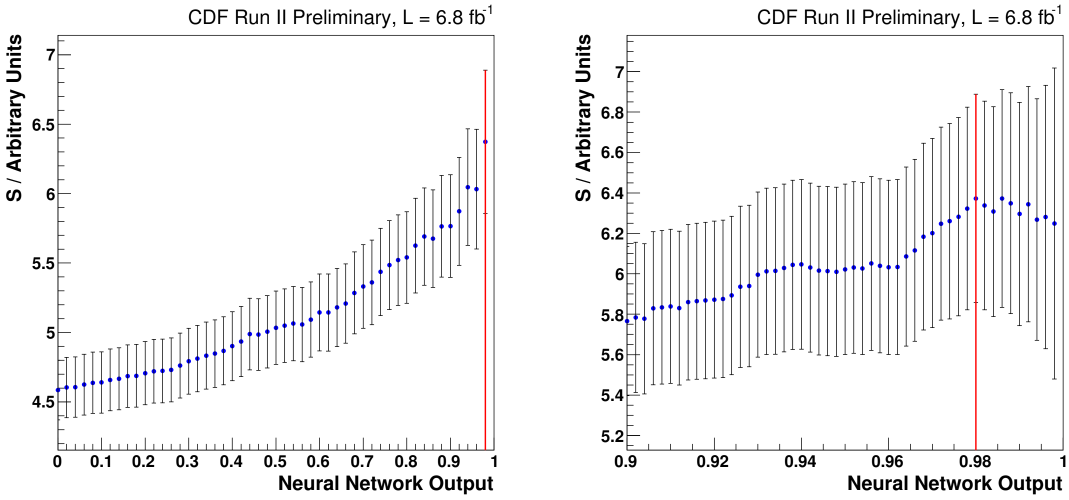


Figure 5.7: Signal significance $S = N_S^{exp} / \sqrt{N_S + N_B}^{Data}$ as a function of neural network output for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+\phi^+\pi^-$. The vertical red line indicates the neural network requirements finally chosen. The error bars give the combined uncertainty due to the uncertainty of the number of data events lying in the signal region and due to limited Monte Carlo statistics.



(a) (b)

Figure 5.8: Signal significance $S = N_S^{exp} / \sqrt{N_S + N_B}^{Data}$ as a function of neural network output for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+K^{*0+}K^-$ (a). Zoom into the upper network scan region (b).

5.5 Associated Efficiencies

The combined reconstruction and signal selection efficiency for a given decay mode is estimated from the ratio of the number of reconstructed simulated events passing the Neural Network requirement, and the number of generated simulation events:

$$\epsilon = \frac{N_{rec}^{MC}}{N_{gen}^{MC}} = \frac{N_{rec}^{MC}}{f^{MC} N_{tot}^{MC}} = \sum_{t}^{3} \frac{w_t N_{rec,t}^{MC}}{f^{MC} N_{tot}^{MC}} \tag{5.2}$$

The number of reconstructed events in simulated data is the sum of re-weighted Two-Track Trigger sub-samples numbered by $t \in [1, 2, 3]$. The weights w_t are calculated from a data/Monte Carlo comparison (Section 4.7). The number of generated signal events of a particular decay mode is calculated from the total number of generated events of a particular Monte Carlo sample, N_{tot}^{MC}, and the decay fraction f^{MC} of a simulated decay. Following a particular decay chain, f^{MC} results from multiplying the individual decay fractions at each decay stage defined in the Monte Carlo decay tables (Section A.4), with the following exception: The fractions of generated events in the $\phi(1020)$ and $K^{*0}(892)$ mass windows are not immediately apparent from the decay tables, since $D_s^+ \to K^+K^-\pi^+$ events are simulated all across the Dalitz plot. Nevertheless, having full knowledge of the model used in simulation it is straightforward to calculate $f_{\phi}^{MC}(1020)$ and $f_{K^{*0}(892)}^{MC}$. The procedure is identical to the one outlined in Section 4.4.5: The fraction of generated events lying in a particular mass window is equal to the ratio of the corresponding mass window integrals of squared amplitudes (calculated from the Dalitz model described in Section 4.4) and the full Dalitz plot integral, equations (4.21), (4.22). The numerical values are given by equations (4.23), (4.24). In the case of the decay $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$, for combinatorial reasons the resulting Monte Carlo decay probability $f_{\phi}^{MC}(1020) \times f_{K^{*0}(892)}^{MC}$ has to be multiplied by two, because – neglecting daughter particle charges – there are two possible combinations to extract this final state from $B_s^0 \to D_s^+ D_s^- \to K^+K^-\pi^+K^-K^+\pi^-$, namely $\phi^{\dagger}\pi^+K^{*0\dagger}K^-$ or $K^{*0\dagger}K^-\phi^{\dagger}\pi^+$.

In Table 5.7 the combined reconstruction and selection efficiencies estimated from simulated data are tabulated channel by channel.

Decay Channel	$N_{gen}^{tot}(10^9)$	$N_{gen}^{ch}(10^6)$	N_{rec}	$\epsilon(10^{-3})$
$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$	1.773	175.658	201004 ± 448	1.144 ± 0.003
$B_s^0 \to D_s^{*+} D_s^-$	1.795	79,450	77140 ± 276	0.971 ± 0.003
$B_s^0 \to D_s^{*+} D_s^{*-}$	0.884	87.626	78762 ± 281	0.899 ± 0.003
$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$	1.773	355.014	230704 ± 480	0.650 ± 0.001
$B_s^0 \to D_s^{*+} D_s^-$	1.795	160.572	89761 ± 299	0.560 ± 0.002
$B_s^0 \to D_s^{*+} D_s^{*-}$	0.884	177.096	93973 ± 307	0.531 ± 0.002
$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$	1.803	189.142	211334 ± 460	1.117 ± 0.002
$B_d^0 \to D^{*+} D_s^-$	1.803	126.094	124997 ± 354	0.991 ± 0.003
$B_d^0 \to D^+ D_s^{*-}$	1.803	126.094	126455 ± 356	1.003 ± 0.003
$B_d^0 \to D^{*+} D_s^{*-}$	1.803	126.094	111007 ± 333	0.880 ± 0.003
$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$	1.803	191.132	123085 ± 351	0.644 ± 0.002
$B_d^0 \to D^{*+} D_s^-$	1.803	127.422	74095 ± 272	0.581 ± 0.002
$B_d^0 \to D^+ D_s^{*-}$	1.803	127.422	73353 ± 272	0.576 ± 0.002
$B_d^0 \to D^{*+} D_s^{*-}$	1.803	127.422	66638 ± 258	0.523 ± 0.002
$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$				
mis-reconstructed as				
$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$	1.803	189.142	11405 ± 107	0.006 ± 0.000
$B_d^0 \to D^{*+} D_s^-$	1.803	126.094	6595 ± 81	0.005 ± 0.000
$B_d^0 \to D^+ D_s^{*-}$	1.803	126.094	6742 ± 87	0.005 ± 0.000
$B_d^0 \to D^{*+} D_s^{*-}$	1.803	126.094	5648 ± 75	0.004 ± 0.000

Table 5.7: Combined reconstruction and selection efficiencies determined from binned likelihood fits to simulated data (please refer to description of Monte Carlo templates in the following sections). The second column quotes the total number of delivered simulated events per Monte Carlo sample, while in the third column the number of simulated events of a particular decay channel is given.

6

Parameter Estimation

6.1 Maximum Likelihood

In order to extract the figures of interest, i.e. the ratios of branching fractions $f_{D_s^* D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, and $f_{D_s^{(*)} D_s^{(*)}}$, from the selected invariant mass spectra, a statistical parameter estimation method and a robust ansatz for the fit function need to be in place. For all fits to real data carried out throughout this analysis the *unbinned extended maximum likelihood method* is used.

The underlying principle of the maximum likelihood (ML) method [3, 128] is the following: Given a set of n independent measurements (or events) of the vector of variables $\vec{x}$, which follows a normalized probability density function (pdf) $f(\vec{x} \mid \vec{a})$, the best estimation of the vector of unknown parameters $\vec{a}$ is the one that maximizes the joint probability density function, called *likelihood function*

$$L(\vec{a}) = \prod_{k=1}^{n} f(\vec{x}_k \mid \vec{a}) \tag{6.1}$$

Instead of maximizing the likelihood function, for numerical reasons common computational maximum likelihood libraries, like the one implemented in the *ROOT* framework, utilize algorithms that aim for minimizing the negative logarithm of (6.1), the *negative log likelihood function* $F(\vec{a})$

$$F(\vec{a}) = -\ln L(\vec{a}) = -\sum_{k=1}^{n} \ln f(\vec{x}_k \mid \vec{a}) \tag{6.2}$$

Within this analysis, the branching fractions to be measured do not only depend on parameters determining the expected shape in $\vec{x}$, but mainly on the normalization of signal-like components. For this purpose, instead of a probability density $f(\vec{x} \mid \vec{a})$ with fixed normalization a function $g(\vec{x} \mid \vec{a})$ is used, where the integral over $g(\vec{x} \mid \vec{a})$ in a measuring range Ω corresponds to the quantity of expected events $N(\vec{a})$

$$N(\vec{a}) = \int_{\Omega} g(\vec{x} \mid \vec{a})\, \mathrm{d}x, \tag{6.3}$$

that are allowed to vary. Furthermore, the observables will be determined from the one-dimensional distribution of reconstructed invariant mass m. Thus the vector of variables $\vec{x}$ simplifies to the scalar quantity m. Within this *extended maximum likelihood* (EML) [129] regime, the negative log likelihood function to be minimized reads

$$F(\vec{a}) = -\sum_{k=1}^{n} \ln g(m_k \mid \vec{a}) + N(\vec{a}) \tag{6.4}$$

We now assume that the shape and normalization of the distribution in m can be described by the sum of individual components that are expected to populate the analyzed range in m. The probability density $g(m \mid \vec{a})$ can then be represented as the sum of q normalized probability densities $h_l(m \mid \vec{a})$ multiplied by their expected yields N_l:

$$g(m \mid \vec{a}) = \sum_{l=1}^{q} N_l h_l(m \mid \vec{a}) \tag{6.5}$$

With this substitution, equation (6.4) becomes

$$F(\vec{a}) = -2\sum_{k=1}^{n} \ln\left[\sum_{l}^{q} N_l h_l(m_k \mid \vec{a})\right] + 2\sum_{l=1}^{q} N_l(\vec{a}) \tag{6.6}$$

The factor of 2 is convention and does not change the position of the minimum. The first task is hence to set up the negative log likelihood function (6.6) and its constituting pdfs. In the following sections the likelihood function that is used to extract the ratios of branching fractions from data is worked out step-by-step. Firstly, the per-channel pdfs are established. Afterwards, these are modified by re-parameterizing and interrelating certain parameters. The function re-parameterizations are targeted towards simultaneously extracting the ratios of branching fractions in one fit that is run on all the studied decay modes in parallel by sharing certain parameters.

6.2 Basic Features of the Fit Model

In order to formulate a robust fit model, the key features of the invariant mass spectra in the analyzed range need to be understood and taken into account. The components contributing to the invariant mass spectra can be roughly classified into four categories:

Combinatorial background Random combinations of tracks not originating from decays of true B mesons that happened to pass the reconstruction and selection requirements.

Fully reconstructed signal Fully reconstructed B_s^0 or B_d^0 mesons characterized by a sharp resonant peak.

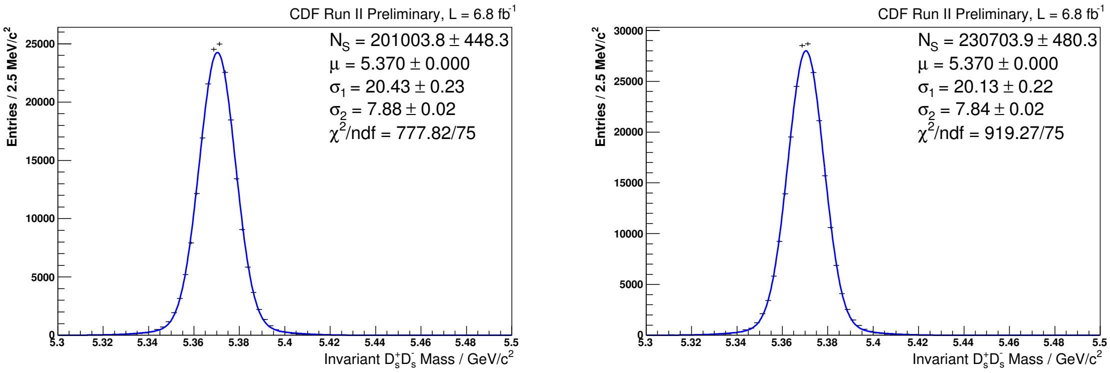
Partially reconstructed signal Events originating from true B mesons where the full decay chain cannot be fully reconstructed because because neutral particles are note detected or otherwise lost in reconstruction.

Physics background Events sharing several common features with true signal events. For the studied decay channels these are essentially given by reflections of other decays, i.e. fully or partially reconstructed decays of mesons that happen to occur in the invariant mass spectrum of a given decay, whenever a wrong particle hypothesis has been assigned to one of the particles in the final state, hence leading to mis-reconstruction of a topologically similar decay.

The individual fit model contributions belonging to each of these categories are discussed in the upcoming sections in more detail.

6.3 $B_s^0 \to D_s^+ D_s^-$ Fit Function

Throughout this section, the individual probability densities constituting to the full pdf of the decay $B_s^0 \to D_s^+ D_s^-$ and the studied sub-channels are set up. The description of the shapes of the various components involve a large amount of parameters that are impossible to be all simultaneously determined during the final EML fitting stage. The remedy for this issue is simulated data: All signal-like components that are expected to contribute to the studied invariant mass spectra have been simulated by means of Monte Carlo techniques. An overview of all signal channels was given in Table 5.7. The shape parameters are determined by binned ML fits to the individual components and can be chosen to be fixed, completely free or allowed to float within Gaussian constraints in the EML fit to real data.



(a) $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ signal template.

(b) $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+K^{*0}K^-$ signal template.

Figure 6.1: Signal template for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+\phi^+\pi^-$ (a) and $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+K^{*0+}K^-$ (b).

6.3.1 Combinatorial Background

Combinatorial background and further background contributions not being explicitly considered in separate templates are well described by a single exponential function plus a constant,

$$P_{bg}^i = \exp\left(m \mid s^i\right) + c^i, \tag{6.7}$$

where s denotes the slope of the exponential and c the constant offset. For the lack of a pure combinatorial background sample the parameters of this template can not be determined by any pre-fitting and are left free in the final unbinned fit to data.

6.3.2 Fully Reconstructed Signal

The shape of the signal peak is modeled by two Gaussian distributions with different widths σ_1^i but sharing a common mean value μ. With i being the index of a particular B_s decay channel, the probability density function for the signal reads

$$P_{sig}^i = f_{\mathcal{G}1}^i \mathcal{G}\left(m \mid \mu\sigma_1^i\right) + \left(1 - f_{\mathcal{G}1}^i\right) \mathcal{G}\left(m \mid \mu\sigma_2^i\right), \tag{6.8}$$

where m is the reconstructed invariant $D_s^+D_s^-$ mass, and $f_{\mathcal{G}1}^i$ is the fraction of one of the Gaussian distributions. For each channel the widths and the mean of the double Gaussians are determined by a binned fit to the signal Monte Carlo samples (Figures 6.1(a), 6.1(b)). In the final EML fit to real data the widths are kept fixed, while the mean is shared among to the full fit function (for details please refer to Section 6.5) and allowed to float within a Gaussian constraint.

6.3.3 Partially Reconstructed Signal

On top of combinatorial background, the lower sideband region is populated by $B_s^0 \to D_s^{(*)+} D_s^{*-}$ decays which arise in the $D_s^+ D_s^-$ invariant mass spectrum as broad satellite peaks. They stem from $D_s^{*+} \to D_s^+ \gamma\,(\pi^0)$ decays, where the low-energy π^0 or γ are not reconstructed since their energies are below the CDF calorimetry system's threshold. Therefore these events can only be partially reconstructed. The main decay mode for D_s^{*+} is given by $D_s^{*+} \to D_s^+ \gamma$ (94.2 ± 0.7%) with a minor contribution from $D_s^{*+} \to D_s^+ \pi^0$ (5.8 ± 0.7%) [3].

- $B_s^0 \to D_s^{*+} D_s^-$
  The physics observables related to the decay dynamics of this decay are well established today (see Section 4.5.2), so there is no intention to make any inference on them by means of parameterization. The fit pdf thus lacks any physics motivation but is rather of phenomenological nature. A function composed of four Gaussians each having its own mean and width,

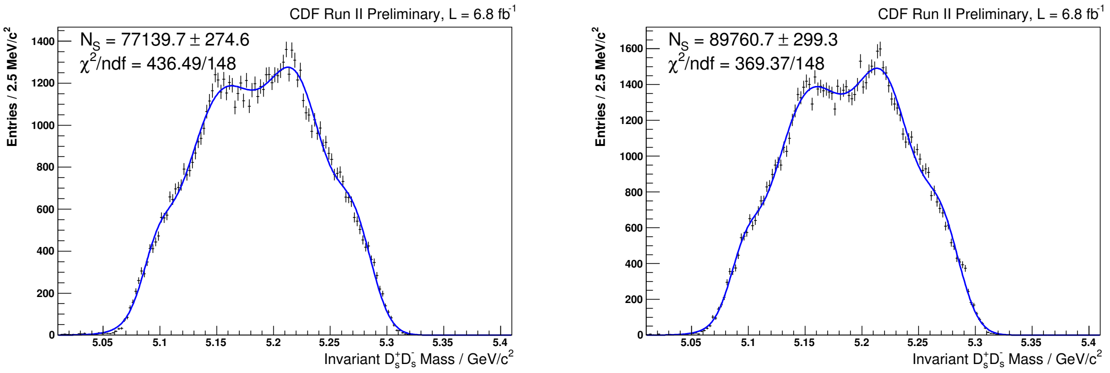
$$\begin{aligned} P^i_{p1} \;=\;& f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1 \sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2 \sigma^i_2\right) + \left(f^i_{\mathcal{G}3}\right)\mathcal{G}\left(m \mid \mu^i_3 \sigma^i_3\right) \\ & + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2} - f^i_{\mathcal{G}3}\right)\mathcal{G}\left(m \mid \mu^i_4 \sigma^i_4\right), \end{aligned} \tag{6.9}$$

  is found to fit reasonably well (Figures 6.2(a), 6.2(b)). The small admixture of $D_s^{*+} \to D_s^+ \pi^0$ produces a smooth double peak structure: From helicity amplitude considerations and conservation of angular momentum (see Section 4.5.2) it follows that the D_s^{*+} is 100% longitudinally polarized. Geometrically speaking, in the subsequent decay of $D_s^{*+} \to D_s^+ \pi^0$ the pion is either emitted in the flight direction of the D^{+*} or opposite to it. For the partial reconstruction of the D^{+*} this means that slightly more or less four-momentum is missing. Since $D_s^{*+} \to D_s^+ \pi^0$ contributes to a much lesser extent than $D_s^{*+} \to D_s^+ \gamma$ the double-peak structure is strongly smeared out.

- $B_s^0 \to D_s^{*+} D_s^{*-}$
  To simulate the dynamics of this decay $B_s^0 \to D_s^{*+} D_s^{*-}$ phase space Monte Carlo were re-weighted according to the helicity amplitudes $\{H_+, H_0, H_-\} = \{0.4904, 07204, 0.4904\}$ [119]. In-detail discussions can be found in Section 4.5.3.

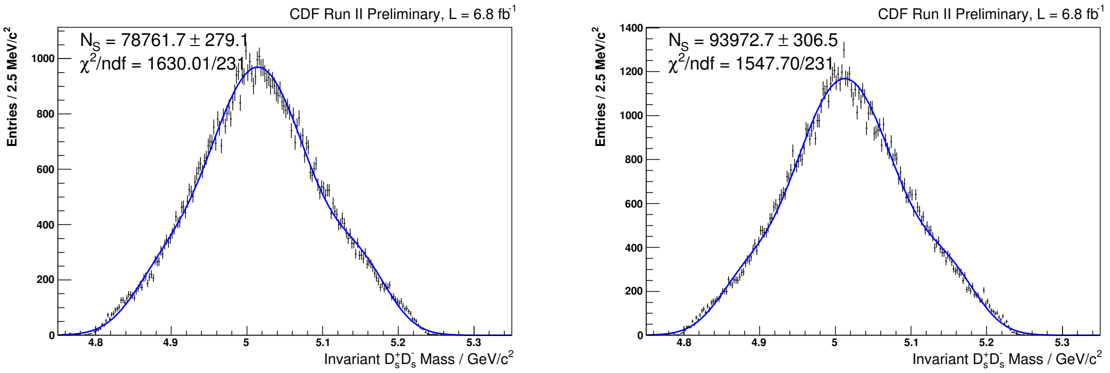
  Simulation generates a broad, featureless shape, that is fitted by the sum of three Gaussians (Figures 6.3(a), 6.3(b)):

$$P^i_{p2} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1 \sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2 \sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3 \sigma^i_3\right) \tag{6.10}$$



(a) $B_s^0 \to D_s^{*+}D_s^- \to \phi\pi^+\phi\pi^-$ template. (b) $B_s^0 \to D_s^{*+}D_s^- \to \phi\pi^+K^{*0}K^-$ template.

Figure 6.2: Templates for the partially reconstructed decay $B_s^0 \to D_s^{*+}D_s^-$ in the final states $\phi^\dagger\pi^+\phi^\dagger\pi^-$ (a) and $\phi^\dagger\pi^+K^{*0\dagger}K^-$ (b).



(a) $B_s^0 \to D_s^{*+}D_s^{*-} \to \phi\pi^+\phi\pi^-$ template. (b) $B_s^0 \to D_s^{*+}D_s^{*-} \to \phi\pi^+K^{*0}K^-$ template.

Figure 6.3: Templates for the partially reconstructed decay $B_s^0 \to D_s^{*+}D_s^{*-}$ in the final states $\phi^\dagger\pi^+\phi^\dagger\pi^-$ (a) and $\phi^\dagger\pi^+K^{*0\dagger}K^-$ (b).

For both templates of partially reconstructed events described in this section the shape parameters and fractions of the individual single Gaussians are kept fixed in the final fit and only the normalizations are allowed to float.

6.3.4 Reflections

The decay $B_s^0 \to D_s^+ \left(\phi\pi^+\right) D_s^- \left(K^{*0} K^-\right)$ has a strong reflection from $B_d^0 \to D_s^+ \left(\phi\pi^+\right)$ $D^-(K^+\pi^-\pi^-)$, which arises if one pion in the final state happens to be mis-reconstructed as a kaon, leading to mis-reconstruction of a B_d^0 as a B_s^0 meson. To estimate the shape of this reflection $B_d^0 \to D_s^+ \left(\phi\pi^+\right) D^- \left(K^+\pi^-\pi^-\right)$ Monte Carlo samples were reconstructed as $B_s^0 \to D_s^+ \left(\phi\pi^+\right) D_s^- \left(K^{*0} K^-\right)$ and then subjected to the same pre-selection and final selection procedure as the decay channel $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$. The reflection signal component has the shape of a resonant peak with a tail to the right. We find that a combination of three Gaussians, each having its own value for mean and width, best fits the mis-reconstructed signal Monte Carlo sample (see Figure 6.4(a)):

$$P^i_{reflsig} = f^i_{\mathcal{G}1} \mathcal{G}\left(m \mid \mu^i_1 \sigma^i_1\right) + f^i_{\mathcal{G}2} \mathcal{G}\left(m \mid \mu^i_2 \sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right) \mathcal{G}\left(m \mid \mu^i_3 \sigma^i_3\right) \qquad (6.11)$$

Again, the shape parameters determined by means of binned fits to simulated data are kept fixed in the final unbinned fit to real data.

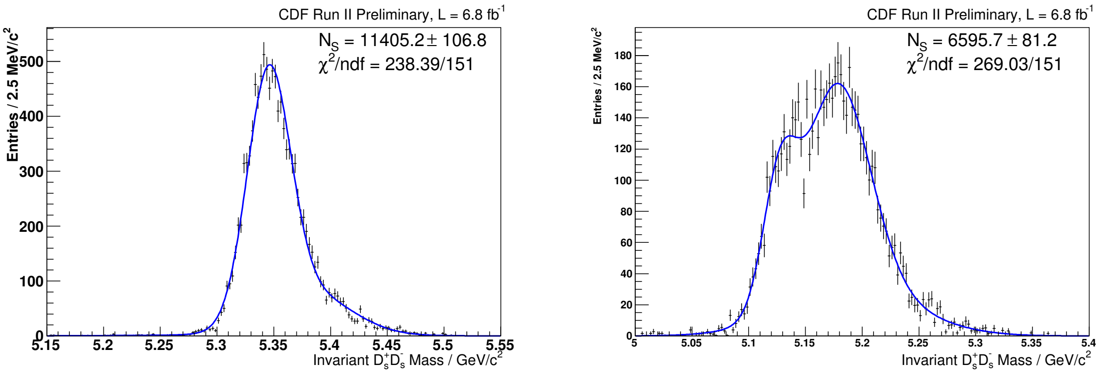
This reflection entails additional satellite contributions, which arise due to false reconstruction of the partially reconstructed $B_d^0 \to D_s^{+(*)} D^{-(*)}$ decays for the same reasons outlined above (lost photon or neutral pion). Each of these contributions is fitted by a triple Gaussian (Figures 6.4(b) to 6.4(d)):

$$P^i_{rp1,2,3} = f^i_{\mathcal{G}1} \mathcal{G}\left(m \mid \mu^i_1 \sigma^i_1\right) + f^i_{\mathcal{G}2} \mathcal{G}\left(m \mid \mu^i_2 \sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right) \mathcal{G}\left(m \mid \mu^i_3 \sigma^i_3\right) \qquad (6.12)$$

6.3.5 Full Fit Function

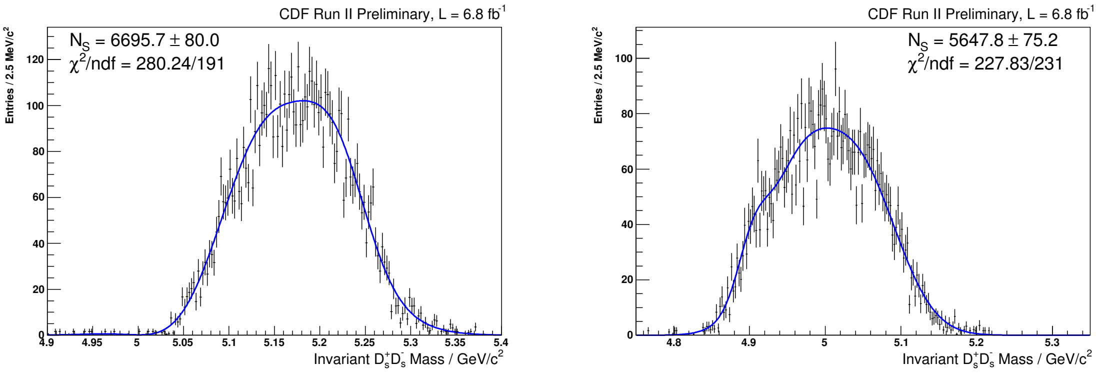
Following equation (6.6), the individual pdfs are combined to give the following negative log likelihood function used for EML fit to the $B_s^0 \to D_s^+ D_s^-$ invariant mass spectra:

$$\begin{aligned} F^i(\vec{a}) \quad = -2 \textstyle\sum_{k=1}^{n} \ln \quad & [N^i_1 P^i_{sig}(m_k \mid \vec{a_{sig}}) + N^i_2 P^i_{p1}(m_k \mid \vec{a}_{p1}) + N^i_3 P^i_{p2}(m_k \mid \vec{a}_{p1}) \qquad (6.13) \\ & + N^i_4 P^i_{rsig}(m_k \mid \vec{a}_{rsig1}) + N^i_5 P^i_{rp1}(m_k \mid \vec{a}_{rp1}) + N^i_6 P^i_{rp2}(m_k \mid \vec{a}_{rp2}) \\ & + N^i_8 P^i_{rp3}(m_k \mid \vec{a}_{rp3}) + N^i_8 P^i_{bg}(m_k \mid \vec{a}_{bg})] \\ & + 2 \sum_{l=1}^{8} N^i_l, \end{aligned}$$



(a) Template for $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ signal mis-reconstructed as $B_s^0 \to D_s^+D_s^- \to \phi\pi^+K^{*0}K^-$.

(b) Reflection template for $B_d^0 \to D^{*+}D_s^-$.



(c) Reflection template for $B_d^0 \to D^+D_s^{*-}$.

(d) Reflection template for $B_d^0 \to D^{*+}D_s^{*-}$.

Figure 6.4: Templates used for the decay $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ falsely reconstructed as $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$. Clockwise shown are the templates for mis-reconstructed $B_d^0 \to D^+D_s^-$ signal and the templates of the falsely and partially reconstructed decays $B_d^0 \to D^{*+}D_s^-$, $B_d^0 \to D^{*+}D_s^{*-}$, and $B_d^0 \to D^+D_s^{*-}$.

where *rsig* label the reflection of the fully reconstructed $B_d^0 \to D^+ D_s^-$ signal, while *rp1,2,3* denote the reflections of partially reconstructed $B_d^0 \to D^{(*)+} D_s^{(*)-}$ events. n is the total number of events of the fitted data sample and i labels the given B_s^0 decay mode. The factors N_l^i in front of each probability density function are the numbers of events of the given component:

$$\begin{aligned}
N_1^i &= N^i_{B_s^0 \to D_s^+ D_s^-} \\
N_2^i &= N^i_{B_s^0 \to D_s^{*+} D_s^-} \\
N_3^i &= N^i_{B_s^0 \to D_s^{*+} D_s^{*-}} \\
N_4^i &= N^i_{B_d^0 \to D^+ D_s^-} \\
N_5^i &= N^i_{B_d^0 \to D^{*+} D_s^-} \\
N_6^i &= N^i_{B_d^0 \to D^+ D_s^{*-}} \\
N_7^i &= N^i_{B_d^0 \to D^{*+} D_s^{*-}} \\
N_8^i &= N^i_{bkg}
\end{aligned} \tag{6.14}$$

For decay channels in which no reflections are expected, the expected yields are set to zero, $N_4^i = N_5^i = N_6^i = N_7^i = 0$.

6.3.6 Exclusive Ratios: Function Re-Parameterization

Based on the per-channel fit function (6.13), in a next step a modified negative log likelihood function is deduced that allows to fit all the studied decay channels simultaneously. This is achieved by an advantageous re-parameterization of the expected number of signal-like events.

The expected yield of reconstructed $B_s^0 \to D_s^+ D_s^-$ events in the ith studied decay channel, $N^i_{B_s^0 \to D_s^+ D_s^-}$, is given by

$$N^i_{B_s^0 \to D_s^+ D_s^-} = N^{tot}_{B_s^0} \mathcal{B}(B_s^0 \to D_s^+ D_s^-) \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-}, \tag{6.15}$$

where $N^{tot}_{B_s^0}$ is the (a priori unknown) total number of produced B_s^0 mesons at the Tevatron, and $\epsilon^i_{B_s^0 \to D_s^+ D_s^-}$ is the combined reconstruction and selection efficiency for this particular decay channel estimated from simulated events (see Section 5.5). The term $\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix}$ represents the non-overlapping $D_s^+ \to K^+ K^- \pi^+$ mass band fractions introduced by reconstruction.

Likewise, the expected number of reconstructed $B_d^0 \to D^+ D_s^-$ events in the jth studied decay channel, $N^j_{B^0_d \to D^+ D^-_s}$, is given by

$$N^j_{B^0_d \to D^+ D^-_s} = N^{tot}_{B^0_d} \mathcal{B}(B^0_d \to D^+ D^-_s) \mathcal{B}\begin{pmatrix} D^+_s \to \phi^{\dagger}\pi^+ \\ D^+_s \to K^{*0\dagger}K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^j_{B^0_d \to D^+ D^-_s} \tag{6.16}$$

$N^{tot}_{B^0_d}$ represents the (a priori unknown) total number of produced B^0_d mesons at the Tevatron, and $\epsilon^j_{B^0_d \to D^+ D^-_s}$ again is the combined reconstruction and selection efficiency estimated from Monte Carlo.

In the next step we intend to introduce a common factor equations (6.15) and (6.16) share. For this, we eliminate $N^{tot}_{B^0_s}$ in equation (6.15) by introducing the ratio of the fragmentation fractions, f_s/f_d, of *strange* and *down* quarks produced at the Tevatron:

$$\frac{N^{tot}_{B^0_s}}{N^{tot}_{B^0_d}} = \frac{f_s}{f_d} \tag{6.17}$$

With this, equation (6.15) now reads

$$N^i_{B^0_s \to D^+_s D^-_s} = N^{tot}_{B^0_d} f_{D_s D_s} \mathcal{B}(B^0_d \to D^+ D^-_s) \mathcal{B}(D^+_s \to \phi^{\dagger}\pi^+) \mathcal{B}\begin{pmatrix} D^+_s \to \phi^{\dagger}\pi^+ \\ D^+_s \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B^0_s \to D^+_s D^-_s}, \tag{6.18}$$

where in the same step the quantity

$$f_{D_s D_s} = \frac{f_s}{f_d} \frac{\mathcal{B}(B^0_s \to D^+_s D^-_s)}{\mathcal{B}(B^0_d \to D^+ D^-_s)} \tag{6.19}$$

has been introduced. By this parameterization, $f_{D_s D_s}$ can now be directly estimated in one simultaneous fit to all four studied decay channels, where $N^{tot}_{B^0_d}$ is globally shared among the combined log likelihood function.

Following the principles outlined above, the parameterization of the expected yields of $B^0_s \to D^{*+}_s D^-_s$ and $B^0_s \to D^{*+}_s D^{*-}_s$ decay events is straightforward:

$$\begin{aligned} N^i_{B^0_s \to D^{*+}_s D^-_s} =& N^{tot}_{B^0_d} f_{D^*_s D_s} \mathcal{B}(B^0_d \to D^+ D^-_s) \mathcal{B}(D^{*+}_s \to D^+_s X) \\ & \times \mathcal{B}(D^+_s \to \phi^{\dagger}\pi^+) \mathcal{B}\begin{pmatrix} D^+_s \to \phi^{\dagger}\pi^+ \\ D^+_s \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B^0_s \to D^{*+}_s D^-_s} \end{aligned} \tag{6.20}$$

$$\begin{aligned} N^i_{B^0_s \to D^{*+}_s D^{*-}_s} =& N^{tot}_{B^0_d} f_{D^*_s D^*_s} \mathcal{B}(B^0_d \to D^+ D^-_s) \mathcal{B}^2(D^{*+}_s \to D^+_s X) \\ & \times \mathcal{B}(D^+_s \to \phi^{\dagger}\pi^+) \mathcal{B}\begin{pmatrix} D^+_s \to \phi^{\dagger}\pi^+ \\ D^+_s \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B^0_s \to D^{*+}_s D^{*-}_s} \end{aligned} \tag{6.21}$$

$\mathcal{B}(D^{*+}_s \to D^+_s X) \equiv 1$ is the branching fraction of the excited D^{*+}_s meson to D^+_s and a neutral particle X (either a photon or π^0). $f_{D^*_s D_s}$ and $f_{D^*_s D^*_s}$ are defined analogously to equation (6.19):

$$f_{D_s^* D_s} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)}{\mathcal{B}(B_d^0 \to D^+ D_s^-)} \tag{6.22}$$

$$f_{D_s^* D_s^*} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})}{\mathcal{B}(B_d^0 \to D^+ D_s^-)} \tag{6.23}$$

Just like $f_{D_s D_s}$, these parameters are also determined in the same simultaneous fit to all data available.

As discussed in Section 6.3.4, the decay $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ has a strong reflection from $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ which is due to occasional mis-reconstruction of one pion as a kaon in the final decay $D^+ \to K^- \pi^+ \pi^+$. Following the same principles outlined in this section, the expected yield of falsely reconstructed $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ events reads

$$N^{mis}_{B_d^0 \to D^+ D_s^-} = N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^+ D_s^-) \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^{mis}_{B_d^0 \to D^+ D_s^-}, \tag{6.24}$$

where $\epsilon^{mis}_{B_d^0 \to D^+ D_s^-}$ is the efficiency of mis-reconstructing $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi \pi^-$ decays as $B_s^0 \to D_s^+ D_s^- \to \phi \pi^+ K^{*0} K^-$ estimated from Monte Carlo. In the same way the number of the falsely and partially reconstructed $B_d^0 \to D_s^{+(*)} D^{-(*)}$ decays is parameterized:

$$\begin{aligned} N^{mis}_{B_d^0 \to D^{*+} D_s^-} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^{*+} D_s^-) \mathcal{B}(D^{*+} \to D^+ X) \\ & \times \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^{mis}_{B_d^0 \to D^{*+} D_s^-} \end{aligned} \tag{6.25}$$

$$\begin{aligned} N^{mis}_{B_d^0 \to D^+ D_s^{*-}} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^+ D_s^{*-}) \\ & \times \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^{mis}_{B_d^0 \to D^+ D_s^{*-}} \end{aligned} \tag{6.26}$$

$$\begin{aligned} N^{mis}_{B_d^0 \to D^{*+} D_s^{*-}} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^{*+} D_s^{*-}) \mathcal{B}(D^{*+} \to D^+ X) \\ & \times \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^{mis}_{B_d^0 \to D^{*+} D_s^{*-}} \end{aligned} \tag{6.27}$$

Here $\mathcal{B}(D^{*+} \to D^+ X) = 0.323 \pm 0.006$ [3] denotes the branching fraction of the excited D^{*+} meson into a charged D^+ and a neutral particle X either being a π^0 or a photon. As in the case of $D_s^{*+} \to D_s^+ \gamma$ and $D_s^{*+} \to D_s^+ \pi^0$ the neutral particles originating from D^{*+} de-excitation are not seen in the CDF detector since their energies are below the energy threshold of the calorimetry system.

6.3.7 Semi-Inclusive Ratio: Function Re-Parameterization

To take further advantage of the simultaneous parameter estimation technique the semi-inclusive branching fraction ratio $f_{D_s^{(*)}D_s^{(*)}}$ is also determined directly in one parallel fit to all decay channels. Trivially, $f_{D_s^{(*)}D_s^{(*)}}$ is defined as the sum of the ratios of the exclusive branching fractions:

$$f_{D_s^{(*)}D_s^{(*)}} = f_{D_sD_s} + f_{D_s^*D_s} + f_{D_s^*D_s^*} \tag{6.28}$$

In principle the semi-inclusive branching fraction could be easily calculated from equation (6.28), without the need of performing a separate EML fit for $f_{D_s^{(*)}D_s^{(*)}}$. As all exclusive branching fraction ratios are extracted from one and the same data sample there might be high correlations among them. When it comes to the estimation of systematic uncertainties for the semi-inclusive sum (6.28) it is thus advantageous to already account for correlations at the fitting stage without the need to treat them afterwards by hand.

The individual branching fraction ratios can be expressed as the semi-inclusive branching fraction ratio times a proportion factor p expressing its share on the total branching fraction:

$$f_{D_sD_s} = f_{D_s^{(*)}D_s^{(*)}}(1 - p_{D_s^*D_s} - p_{D_s^*D_s^*}) \tag{6.29}$$

$$f_{D_s^*D_s} = f_{D_s^{(*)}D_s^{(*)}} p_{D_s^*D_s} \tag{6.30}$$

$$f_{D_s^*D_s^*} = f_{D_s^{(*)}D_s^{(*)}} p_{D_s^*D_s^*} \tag{6.31}$$

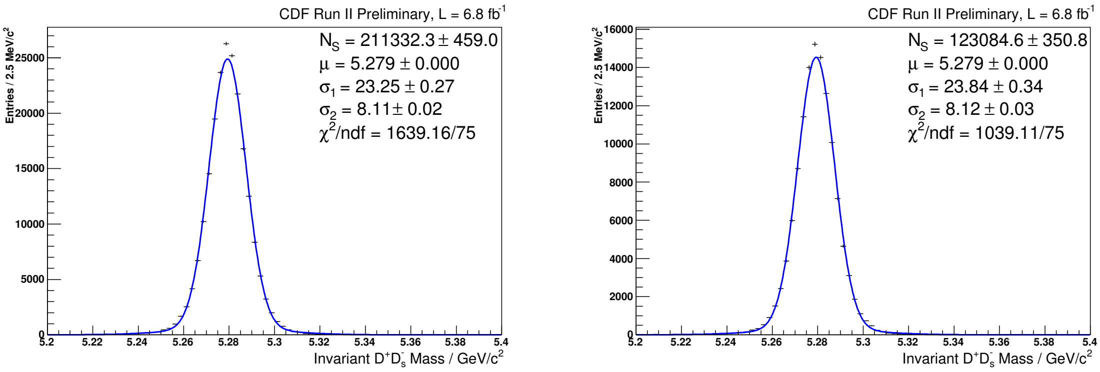
Using these definitions, we may re-write the expected number of reconstructed events defined in equations (6.18) through (6.21) as

$$\begin{aligned} N^i_{B_s^0 \to D_s^+ D_s^-} = {} & N^{tot}_{B_d^0} f_{D_s^{(*)}D_s^{(*)}}(1 - p_{D_s^*D_s} - p_{D_s^*D_s^*}) \\ & \times \mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}(D_s^+ \to \phi^{\dagger}\pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^{\dagger}\pi^+ \\ D_s^+ \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \end{aligned} \tag{6.32}$$

$$\begin{aligned} N^i_{B_s^0 \to D_s^{*+} D_s^-} = {} & N^{tot}_{B_d^0} f_{D_s^{(*)}D_s^{(*)}} p_{D_s^*D_s} \\ & \times \mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}(D_s^+ \to \phi^{\dagger}\pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^{\dagger}\pi^+ \\ D_s^+ \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \end{aligned} \tag{6.33}$$

$$\begin{aligned} N^i_{B_s^0 \to D_s^{*+} D_s^{*-}} = {} & N^{tot}_{B_d^0} f_{D_s^{(*)}D_s^{(*)}} p_{D_s^*D_s^*} \\ & \times \mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}(D_s^+ \to \phi^{\dagger}\pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^{\dagger}\pi^+ \\ D_s^+ \to K^{*0\dagger}K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \end{aligned} \tag{6.34}$$

Instead of the exclusive ratios $f_{D_sD_s}$, $f_{D_s^*D_s}$, and $f_{D_s^*D_s^*}$, now the semi-inclusive ratio $f_{D_s^*D_s^*}$ is the parameter that is determined in the simultaneous EML fit to all available channels.



(a) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ signal template. (b) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+K^{*0}K^-$ signal template.

Figure 6.5: Signal template for the decay $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ (a) and $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+K^{*0\dagger}K^-$ (b).

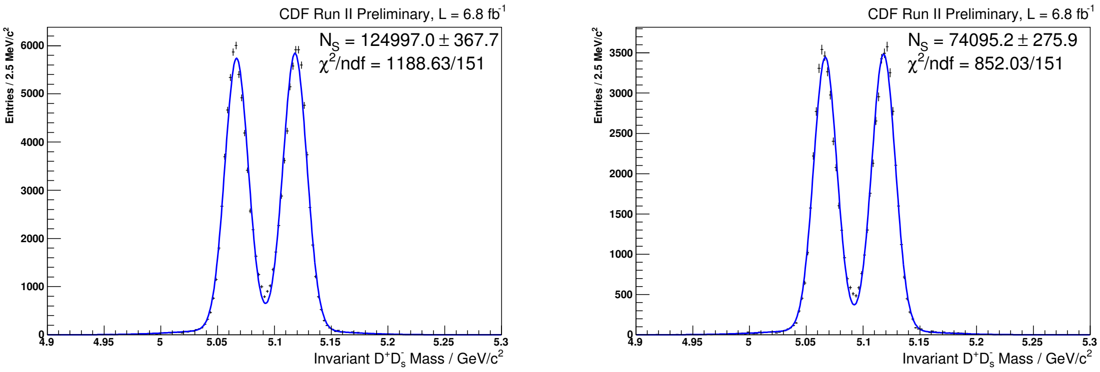
6.4 $B_d^0 \to D^+D_s^-$ Fit Function

The log likelihood function and the individual templates involved are set up in a very similar way to those discussed in Section 6.3. The decay $B_d^0 \to D_sD$ is studied because it serves as the normalization mode for the calculation of the relative branching fraction $\mathcal{B}(B_s^0 \to D_s^+D_s^-)/\mathcal{B}(B_d^0 \to D^+D^-)$. The functions used to describe fully reconstructed signal and combinatorial background are the same used for $B_s^0 \to D_s^+D_s^-$, namely the sum of two Gaussian distributions (Figures 6.5(a), 6.5(b)) having a common mean value, and a single exponential function plus constant, respectively. The topology of the lower sideband region of the $D^+D_s^-$ invariant mass spectrum is very similar to that of $D_s^+D_s^-$. It is as well populated by several satellite peaks originating from partial reconstruction of $B_d^0 \to D_s^{(*)+}D_s^{(*)-}$ decays.

6.4.1 Partially Reconstructed Signal

The decays $B_d^0 \to D_s^{*+}D^{*-}$ produce a series of satellite peaks shifted to smaller masses approximately by one to two pion masses with respect to the position of the signal peak of fully reconstructed $B_d^0 \to D^+D_s^-$ signal. There are three possible combinations:

- $B_d^0 \to D^{*+}D_s^-$
  The D^{*-} can decay either into D^- and a lost γ/π^0, or into $D^0\pi^-$ which is not subject to reconstruction. $B_d^0 \to D^{*-}D_s^+$ creates a distinct double peak structure. The explanation is similar to the one given in Section 6.3.3: Again, we are dealing with a pseudo-scalar decaying into a vector and a pseudo-scalar particle ($P \to VP$), and due to conservation of angular momentum the



(a) $B_d^0 \to D^{*+}D_s^- \to K^-\pi^+\pi^+\phi\pi^-$ template. (b) $B_d^0 \to D^{*+}D_s^- \to K^-\pi^+\pi^+K^{*0}K^-$ template.

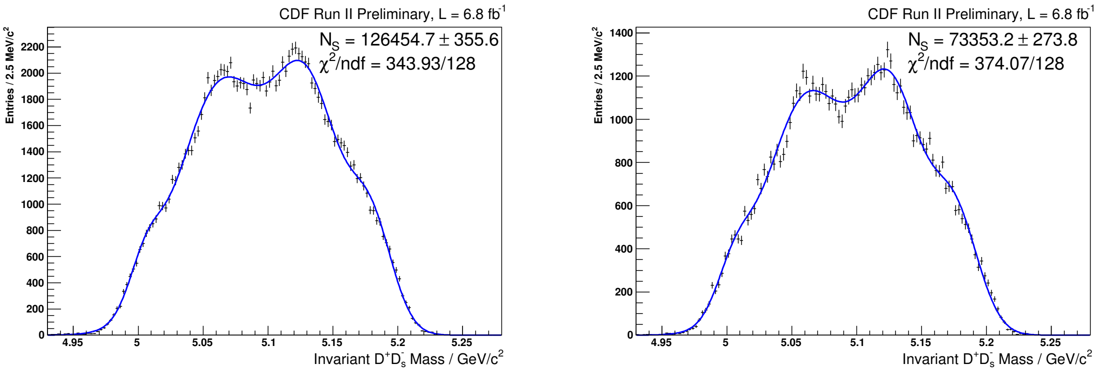
Figure 6.6: Template for the partially reconstructed decay $B_d^0 \to D^{*+}D_s^-$ going into the final states $K^-\pi^+\pi^+\phi^+\pi^-$ (a) and $K^-\pi^+\pi^+K^{*0+}K^-$ (b).

decaying D^{*-} is 100% longitudinally polarized. Therefore the angle between the π^0 and the momentum helicity of the D^* follows a $\cos^2\theta$ distribution, which corresponds to a preferred pion release in the direction of the D^{+*} or opposite to it. In contrast to $D_s^{*+} \to D_s^+\pi^0/\gamma$, the main contribution is given by $D^{*-} \to D^-\pi^0$ decays which results in the double peak shape being not diluted by other contributions. The resulting shape is fitted by a combination of three Gaussians (Figures 6.6(a), 6.6(b)).

- $B_d^0 \to D^+D_s^{*-}$
  The invariant mass distribution of this partially reconstructed $P \to VP$ decay exhibits a weakly pronounced double peak structure very similar to the one observed for $B_s^0 \to D_s^{*+}D_s^-$. As in the cases discussed before, the reason for the resulting weak double-horn structure is the full longitudinal polarization of the vector D_s^{*+} particle, which decays into $D_s^+\pi^0$ in 5.8% of the cases only. The shape is fitted by the sum of four Gaussian distributions, each having its own mean and width (Figures 6.7(a), 6.7(b)).

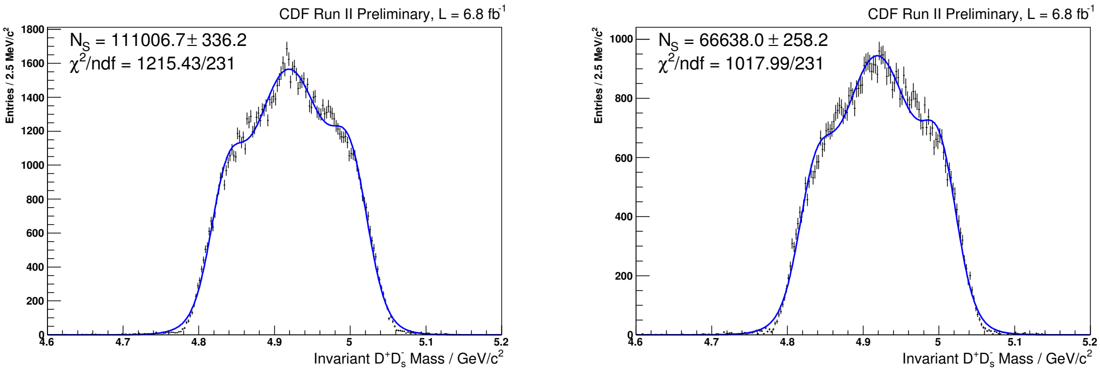
- $B_d^0 \to D^{*+}D_s^{*-}$
  This mode generates a wide, featureless structure located about two pion masses below the $B_d^0 \to D_s^+D^-$ signal. A triple Gaussian is used to fit the resulting shape (Figures 6.8(a), 6.8(b)).

The shapes of all signal templates determined by means of binned fits using simulated data are kept fixed in the final EML fit to real data.



(a) $B_d^0 \to D^+D_s^{*-} \to K^-\pi^+\pi^+\phi\pi^-$ template. (b) $B_d^0 \to D^+D_s^{*-} \to K^-\pi^+\pi^+K^{*0}K^-$ template.

Figure 6.7: Template for the partially reconstructed decay $B_d^0 \to D^+D_s^{*-}$ going into the final states $K^-\pi^+\pi^+\phi^\dagger\pi^-$ (a) and $K^-\pi^+\pi^+K^{*0\dagger}K^-$ (b).



(a) $B_d^0 \to D^{*+}D_s^{*-} \to K^-\pi^+\pi^+\phi\pi^-$ template. (b) $B_d^0 \to D^{*+}D_s^{*-} \to K^-\pi^+\pi^+K^{*0}K^-$ template.

Figure 6.8: Template for the partially reconstructed decay $B_d^0 \to D^{*+}D_s^{*-}$ going into the final states $K^-\pi^+\pi^+\phi^\dagger\pi^-$ (a) and $K^-\pi^+\pi^+K^{*0\dagger}K^-$ (b).

6.4.2 Full Fit Function

Combining the pdfs set up above gives rise to the total negative log likelihood function for the jth B_d^0 mode,

$$\begin{aligned} F^j(\vec{b}^j) &= -2\textstyle\sum_{k=1}^{n} \ln \ [N_1^j P_{sig}^j(m_k \mid \vec{b}_{sig}^j) + N_2^j P_{p1}^j(m_k \mid \vec{b}_{p1}^j) + N_3^j P_{p2}^j(m_k \mid \vec{b}_{p2}^j) \\ &\quad +N_4^j P_{p3}^j(m_k \mid \vec{b}_{p3}^j) + N_5^i P_{bg}^j(m_k \mid \vec{b}_{bg}^j)] \\ &\quad +2\sum_{l=1}^{5} N_l^j, \end{aligned} \tag{6.35}$$

where the parameters N_l^i in front of each probability density function are the expected yields of the given component:

$$\begin{aligned} N_1^j &= N^j_{B_d^0\to D^+D_s^-} \\ N_2^j &= N^j_{B_d^0\to D^{*+}D_s^-} \\ N_3^j &= N^j_{B_d^0\to D^+D_s^{*-}} \\ N_4^j &= N^j_{B_d^0\to D^{*+}D_s^{*-}} \\ N_5^j &= N^j_{bkg} \end{aligned} \tag{6.36}$$

6.4.3 Function Re-Parameterization

The parameterization of the expected yield of reconstructed $B_d^0 \to D^+D_s^-$ events $N^j_{B_d^0\to D^+D_s^-}$ in the jth studied decay channel has already been introduced (6.16):

$$N^j_{B_d^0\to D^+D_s^-} = N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^+\pi^+ \\ D_s^+ \to K^{*0+}K^+ \end{pmatrix}\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^j_{B_d^0\to D^+D_s^-}$$

$N^{tot}_{B_d^0}$ represents the (a priori unknown) total number of produced B_d^0 mesons at the Tevatron, and $\epsilon^j_{B_d^0\to D^+D_s^-}$ is the combined reconstruction and selection efficiency estimated from simulated events. Accordingly, the yields of partially reconstructed events are parameterized as:

$$\begin{aligned} N^j_{B_d^0\to D^{*+}D_s^-} =& N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^{*+}D_s^-)\mathcal{B}(D^{*+} \to D^+X) \\ &\times\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^+\pi^+ \\ D_s^+ \to K^{*0+}K^+ \end{pmatrix}\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^j_{B_d^0\to D^{*+}D_s^-} \end{aligned} \tag{6.37}$$

$$\begin{aligned} N^j_{B_d^0\to D^+D_s^{*-}} =& N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^+D_s^{*-}) \\ &\times\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^+\pi^+ \\ D_s^+ \to K^{*0+}K^+ \end{pmatrix}\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^j_{B_d^0\to D^+D_s^{*-}} \end{aligned} \tag{6.38}$$

$$
\begin{aligned}
N^{j}_{B^0_d \to D^{*+} D^{*-}_s} =& N^{tot}_{B^0_d} \mathcal{B}(B^0_d \to D^{*+} D^{*-}_s) \mathcal{B}(D^{*+} \to D^+ X) \\
& \times \mathcal{B}\begin{pmatrix} D^+_s \to \phi^\dagger \pi^+ \\ D^+_s \to K^{*0\dagger} K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^{j}_{B^0_d \to D^{*+} D^{*-}_s}
\end{aligned}
\tag{6.39}
$$

The parameterizations (6.16) and (6.37) through (6.39) are valid for both the simultaneous fit for the exclusive ratios $f_{D_s D_s}$, $f_{D^*_s D_s}$, $f_{D^*_s D^*_s}$ (Section 6.3.6) and the semi-inclusive branching fraction ratio $f_{D^{(*)}_s D^{(*)}_s}$ (Section 6.3.7).

6.5 Extraction of Branching Fractions

In order to extract the observables of interest simultaneously from all decay channels, the negative log likelihood functions (6.13) and (6.35) are combined to give one joint likelihood that is minimized. The parameterizations of the signal yields in (6.13) depend on whether one aims for measuring the exclusive branching fraction ratios or the semi-inclusive ratio.

There is a compelling case for the joint likelihood approach: The decay $B^0_d \to D^+ D^-_s \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ enters twice, once as correctly reconstructed signal in the invariant spectrum of the normalization channel, and once as background (mis-reconstructed signal) in $B^0_s \to D^+_s D^-_s \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$. The simultaneous fitting approach and the consistent parameterizations of both correctly and falsely reconstructed signal ensure that correlations are taken into account already at the fittings stage. Secondly, the total amount of parameters needed to describe the problem is not unnecessarily inflated beyond the minimum required parameter set. The same argument applies when it comes to external parameters like intermediate or final state branching fractions that are required to parameterize signal yields.

Table 6.1 gives an overview of the parameters that are shared among the decay modes. Parameters enter the joint likelihood fit with a certain attribute assigned: 'None' (parameter fixed), 'Gaussian Constraint' (parameter allowed to vary around, but strongly constrained to its default central value within Gaussian uncertainties), and 'Free' (parameter completely free within wide physical boundaries). Unique parameters (template shape parameters, efficiencies) that are fixed in the EML fit to data are not listed. In the standard fit, the default preference for parameters that have both the attributes 'None' and 'Gaussian Constraint' is 'None'. The 'Gaussian Constraint' attribute is used for systematic studies only to propagate uncertainties.

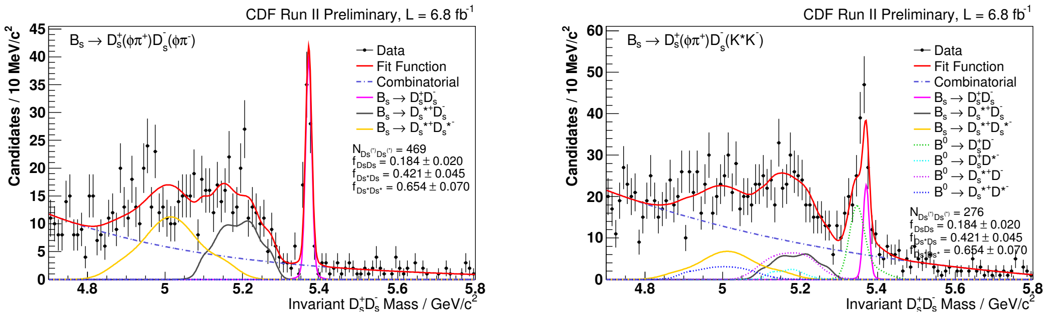
The total number of B_0 mesons produced at the Tevatron, $N_{B^{tot}_0}$, is globally

<table>
<tr><th>$B_s^0 \to D_s^+ D_s^-$
$\to \phi^{\dagger}\pi^+\phi^{\dagger}\pi^-$</th><th>$B_s^0 \to D_s^+ D_s^-$
$\to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$</th><th>$B_d^0 \to D^+ D_s^-$
$\to K^-\pi^+\pi^+\phi^{\dagger}\pi^-$</th><th>$B_d^0 \to D^+ D_s^-$
$\to K^-\pi^+\pi^+ K^{*0\dagger} K^-$</th><th>Parameter Freedom</th></tr>
<tr><td colspan="4">$N_{B_0^{tot}}$</td><td>F</td></tr>
<tr><td colspan="2">$f_{D_sD_s}, f_{D_s^*D_s}, f_{D_s^*D_s^*}, f_{D_s^{(*)}D_s^{(*)}}$</td><td colspan="2">-</td><td>F</td></tr>
<tr><td>-</td><td colspan="3">$\mathcal{B}(B_d^0 \to D^{*+} D_s^-), \mathcal{B}(B_d^0 \to D^+ D_s^{*-}), \mathcal{B}(B_d^0 \to D^{*+} D_s^{*-})$</td><td>F</td></tr>
<tr><td>-</td><td colspan="3">$\mathcal{B}(D^{*+} \to D^+ X)$</td><td>N, GC</td></tr>
<tr><td colspan="2">μ_{B_s}</td><td colspan="2">-</td><td>GC</td></tr>
<tr><td colspan="2">-</td><td colspan="2">μ_{B^0}</td><td>GC</td></tr>
<tr><td colspan="4">$\mathcal{B}(B_d^0 \to D^+ D_s^-)$</td><td>N, GC</td></tr>
<tr><td>-</td><td colspan="3">$\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$</td><td>N, GC</td></tr>
<tr><td colspan="4">$\frac{\mathcal{B}(D_s^+ \to \phi\pi^{+\dagger})}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)}, \frac{\mathcal{B}(D_s^+ \to K^{*0}K^{+\dagger})}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)}$</td><td>N</td></tr>
<tr><td colspan="4">$\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$</td><td>N, GC</td></tr>
</table>

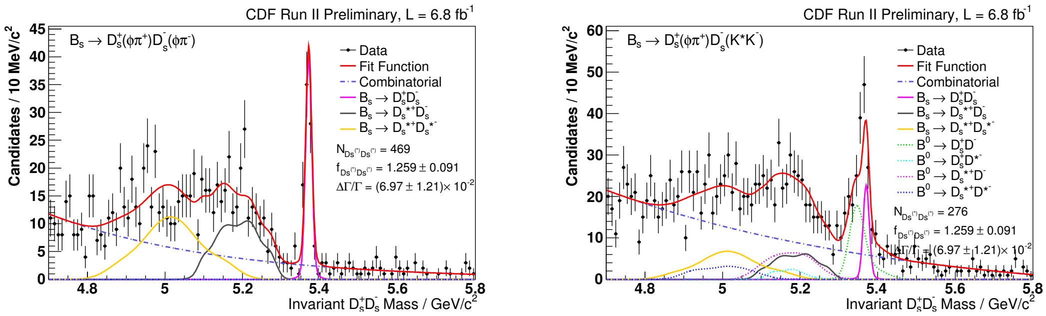
Table 6.1: Overview of the fit parameters which are shared in the simultaneous fit to all four studied decay channels. 'Parameter Freedom' indicates the degree of freedom a given parameter has in the fit: 'F' stands for 'Free', 'GC' for 'Gaussian Constraint', 'N' for 'None' (parameter fixed).

shared among all decay modes. By equations (6.16), (6.18), (6.20), (6.21), (6.24 - 6.27), and (6.37 - 6.39) $N_{B_0^{tot}}$ is related to the expected yields of all (both fully, partially, and falsely reconstructed) signal-like events in the decay modes. The yields in turn determine $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ through their respective intermediate and final state branching fractions and the combined reconstruction and selection efficiencies. The branching fractions $\mathcal{B}(B_d^0 \to D^{*+} D_s^-)$, $\mathcal{B}(B_d^0 \to D^+ D_s^{*-})$, $\mathcal{B}(B_d^0 \to D^{*+} D_s^{*-})$, and $\mathcal{B}(D^{*+} \to D^+ X)$ also enter the likelihood function of the decay mode $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$ due to partial reconstruction of the reflections of the channels $B_d^0 \to D^{(*)+} D_s^{(*)-} \to K^-\pi^+\pi^+\phi^{\dagger}\pi^-$. Despite the complicated multi-component structure in the lower mass region of $B_s^0 \to D_s^+ D_s^- \to \phi^{\dagger}\pi^+ K^{*0\dagger} K^-$, it is worth noting that the associated components are not fixed by any external knowledge. The constraints to these fractions are introduced "in situ" by sharing the respective parameters with the high-statistics normalization channels in the joint likelihood fit. In doing so, one does not have to rely on the relatively large uncertainties of the world average values of $\mathcal{B}(B_d^0 \to D^{*+} D_s^-)$, $\mathcal{B}(B_d^0 \to D^+ D_s^{*-})$, and $\mathcal{B}(B_d^0 \to D^{*+} D_s^{*-})$ [3] in later systematic studies.

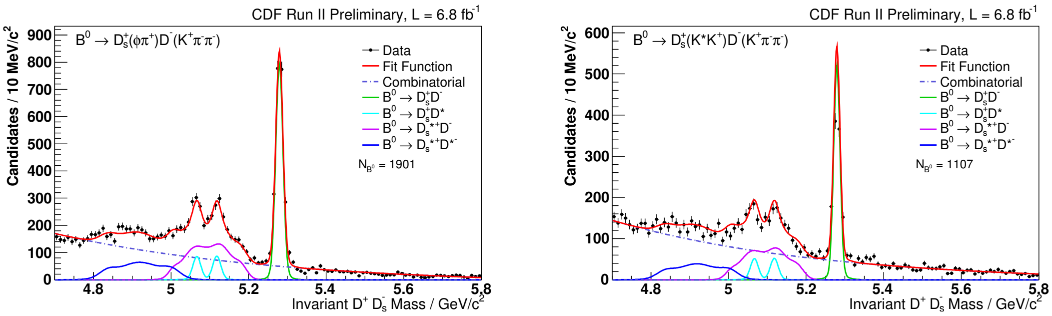
Figures 6.9(a) - 6.9(f) show the significance-optimized (see Section 5.4) invariant mass spectra that have been fitted using the simultaneous parameter estimation procedure. From the simultaneous fits the following branching fractions



(a) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ (b) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$



(c) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ (d) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$



(e) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ (f) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+K^{*0\dagger}K^-$

Figure 6.9: Invariant $D_s^+D_s^-$ (a-d) and $D^+D_s^-$ (e,f) mass spectra fitted under the joint likelihood scheme. The first row shows the results for the scenario where the fit was geared towards measuring $f_{D_sD_s}$, $f_{D_s^*D_s}$, and $f_{D_s^*D_s^*}$, the second row where $f_{D_s^{(*)}D_s^{(*)}}$ was the quantity to be directly determined. The bottom row shows the fit results of the decay modes that were used as normalization channels in either of the two fit scenarios. The complex multi-component structure in the lower mass region of $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$ (b,d) is due to reflections of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays into the final state $K^-\pi^+\pi^+\phi^\dagger\pi^-$.

	BF / 10^{-3}		**Deviation**
Quantity	**This Analysis**	**World Average**	**/ σ**
$f_{D_sD_s}$	6.2 ± 0.7	8.0 ± 1.1	1.4
$f_{D_s^*D_s}$	8.4 ± 0.4	7.4 ± 1.6	0.6
$f_{D_s^*D_s^*}$	17.0 ± 1.0	17.7 ± 1.4	0.4

Table 6.2: Comparison of branching fractions of $B_d^0 \to D^{(*)+}D_s^{(*)-}$ decays measured in the course of the joint likelihood fit with respect to current world average values. For the values determined by this analysis statistical uncertainties are given only.

ratios are extracted:

$$f_{D_sD_s} = 0.184^{+0.021}_{-0.020} \tag{6.40}$$

$$f_{D_s^*D_s} = 0.421^{+0.046}_{-0.045} \tag{6.41}$$

$$f_{D_s^*D_s^*} = 0.654^{+0.073}_{-0.071} \tag{6.42}$$

$$f_{D_s^{(*)}D_s^{(*)}} = 1.259^{+0.094}_{-0.092} \tag{6.43}$$

The uncertainties given are statistical only. As expected, the lowermost value (6.43) extracted from a separate simultaneous fit using an alternate signal parameterization (Section 6.3.7) is identical to the sum of the exclusive branching fraction ratios, (6.40) through (6.43).

The matrix of correlation coefficients reads:

$$\begin{array}{cccc} & f_{D_sD_s} & f_{D_s^*D_s} & f_{D_s^*D_s^*} \\ f_{D_sD_s} & 1 & 0.095 & 0.104 \\ f_{D_s^*D_s} & 0.095 & 1 & 0.04 \\ f_{D_s^*D_s^*} & 0.104 & 0.04 & 1 \end{array} \tag{6.44}$$

The correlations among the exclusive branching fractions are found to be fairly low, not larger than 10%.

As indicated by Table 6.1, the branching fractions of $B_d^0 \to D^{*+}D_s^-$, $B_d^0 \to D^+D_s^{*-}$, and $B_d^0 \to D^{*+}D_s^{*-}$ decays were allowed to float in the joint EML fit to data. In Table 6.2 the resulting values are tabulated and compared to the current world average values extracted from Ref. [3]. All the values are in good agreement within 1.5σ. This shows that the fitting method does not significantly bias values of experimentally established observables. Moreover, as a side product of this analysis we might even be able to provide new measurements of these branching fractions in the near future. For the time being, this is however beyond the scope of this analysis and will require further systematic studies.

6.6 Feasibility of Measuring f_{CP-} from Partial $B_s^0 \to D_s^{*+}D_s^{*-}$ Reconstruction

The assumption of a non-sizable CP-odd component in $B_s^0 \to D_s^{*+}D_s^{*-}$ is critical to the applicability of equation (1.39) which allows to infer the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ from a branching fraction measurement. The observation of a sizeable CP-odd fraction would put the described measuring concept into question.

The most promising way to disentangle CP-even and CP-odd components of the $P \to VV$ decay $B_s^0 \to D_s^{*+}D_s^{*-}$ would be through a full angular analysis of the decay products, be it in the helicity or the transversity frame. By fitting the angular distributions the helicity amplitudes H_+, H_0, and H_- can be quantified. The CP-odd fraction f_{CP-} is then determined by the difference of the helicity amplitudes H_+ and H_- through equation (4.44). This approach has already been successfully employed in other $B \to VV$ analyses. One of the most prominent examples in the B_s^0 sector is the decay $B_s^0 \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^+$ and $\phi \to K^+K^-$. In the case of $B_s^0 \to D_s^{*+}D_s^{*-}$ however, the fact that the low-energetic neutral decay products γ and π^0 are lost in reconstruction prohibits this approach. Depending on the direction the missing decay particles are emitted into, their angular distributions might influence the distribution in (partially) reconstructed invariant mass. Therefore, as an alternative to a full data-driven angular analysis, a Monte Carlo based approach is chosen: If the angular and reconstructed mass distributions do show a strong dependence on the helicity amplitudes used in simulation the true helicity amplitudes can be deduced by varying helicity amplitude expectations in simulated data. In practice, this can be realized by using a series of Monte Carlo templates reflecting different helicity amplitude scenarios in the EML fit to data. The resulting fit qualities can then be used to confine or rule out the most probable helicity amplitudes in $B_s^0 \to D_s^{*+}D_s^{*-}$.

As pointed out above, the high computing time prohibits running full simulation for *every* helicity amplitude scenario that one may want to test. Therefore, the study is performed by weighting the dedicated $B_s^0 \to D_s^{*+}D_s^{*-}$ phase space Monte Carlo sample according to set of desired helicity amplitudes H_+, H_0, and H_-, with the weights equal to the squared amplitude calculated in the helicity frame, equation (4.47). To have a quick means of evaluating the feasibility of the suggested approach, the decay $B_s^0 \to D_s^{*+}D_s^{*-}$ is generated in a fast simulation scheme using the extreme (and probably unphysical) helicity amplitudes $\{H_+, H_0, H_-\} = \{1, 0, 0\}$ and $\{0, 0, 1\}$ in order to emulate maximal differences between H_+ and H_-. For comparison the case of 100% longitudinal polarization,

$\{H_+, H_0, H_-\} = \{0, 1, 0\}$, is also shown. In fast simulations, the decay chain is simulated only, without detector and trigger simulation. This reduces the computing time required for simulation considerably. For the simulation of the $B_s^0 \to D_s^{*+} D_s^{*-}$ decay chain internal decay amplitude calculations inside the EvtGen program are relied upon. The simulated angular distributions and invariant mass spectra (Figure 6.10) show a clear dependence on helicity amplitudes. Compared to the case of full longitudinal polarization $\{H_+, H_0, H_-\} = \{0, 1, 0\}$ the mass distributions obtained for the two other extreme cases considered exhibit a strong asymmetric shape. Thus, simulation indicate that there is a clear correlation between helicity amplitudes and mass line shape of partially reconstructed $B_s^0 \to D_s^{*+} D_s^{*-}$ events. Thus, the suggested approach of inferring a potentially non-zero *CP*-odd fraction f_{CP-} from partial reconstruction of $B_s^0 \to D_s^{*+} D_s^{*-}$ looks very promising.

Following this preparatory test using fast simulations, the studies are continued in a more systematic way by re-weighting full phase space simulations of $B_s^0 \to D_s^{*+} D_s^{*-}$ decays for particular sets of helicity amplitudes. With the helicity amplitude H_0 kept fixed to 0.7204 [119], H_+ and H_- are adjusted in a way that the fraction of the *CP*-odd component ranges from 0.0 through 0.24 in steps of 0.08 (the upper boundary corresponds to the maximum *CP*-odd fraction for $H_0 = 0.7204$). For given values of H_0 and f_{CP-}, H_+ and H_- are determined by the normalization condition

$$|H_+|^2 + |H_0|^2 + |H_-|^2 = 1 \tag{6.45}$$

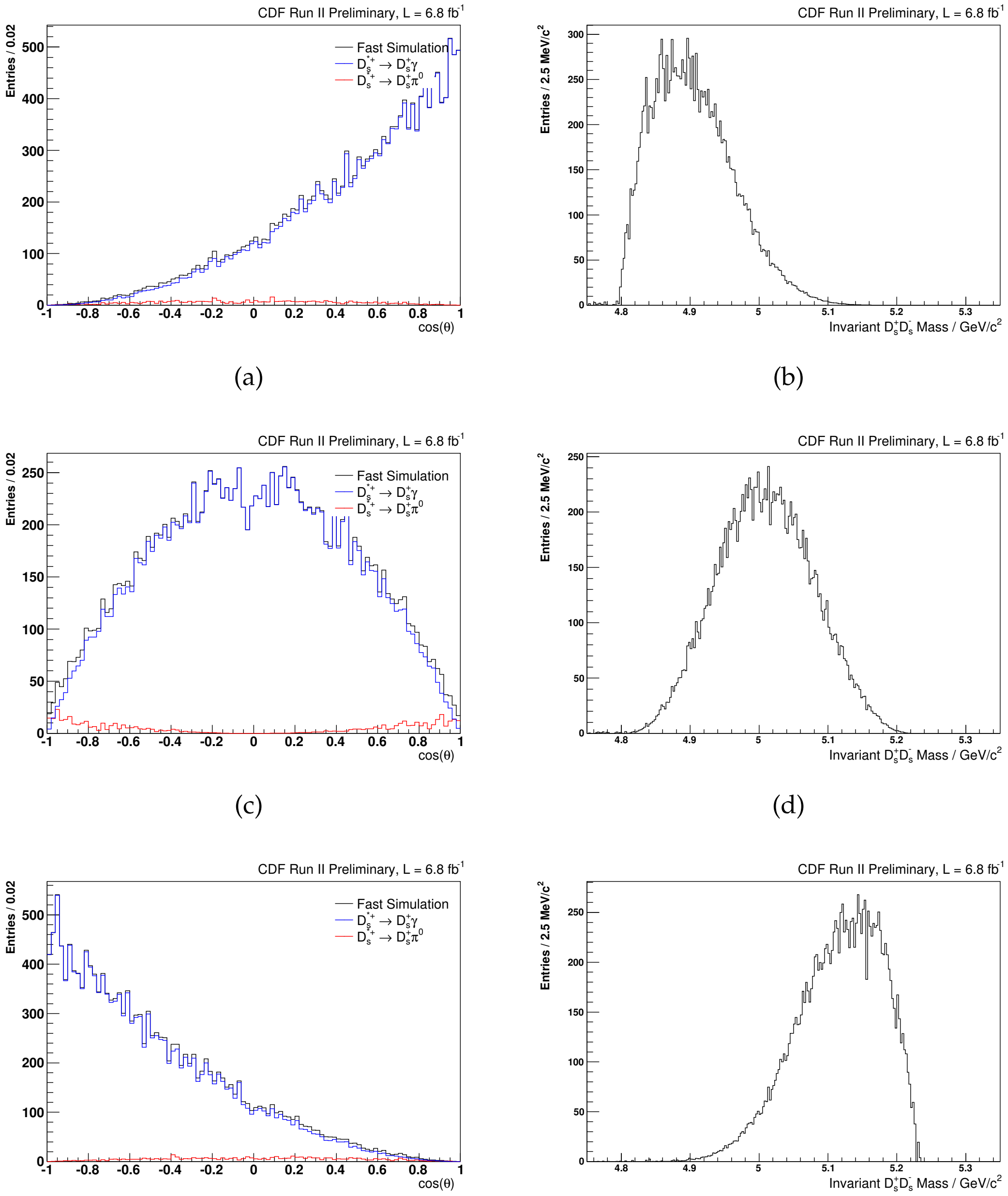
and equation (4.44):

$$H_+^{(1)} = H_-^{(2)} = \frac{2\sqrt{2f_{CP-}} + \sqrt{8f_{CP-} - 8(-1 + |H_0|^2 + 2f_{CP-})}}{4} \tag{6.46}$$

$$H_-^{(1)} = H_+^{(2)} = \frac{2\sqrt{2f_{CP-}} - \sqrt{8f_{CP-} - 8(-1 + |H_0|^2 + 2f_{CP-})}}{4} \tag{6.47}$$

The Monte Carlo distributions shown for the cases $f_{CP-} = 0.00, 0.08, 0.16, 0.24$ (Figure 6.11) clearly indicate that for a fixed longitudinal polarization both the angular distributions and the invariant mass line shapes of partially reconstructed $B_s^0 \to D_s^{*+} D_s^{*-}$ events are completely insensitive to changes in H_+ and H_-, and thus to f_{CP-}. Since all the Monte Carlo templates are virtually identical – despite changes in H_+ and H_- – the proposed procedure of performing fits to real data and evaluating fit qualities for different scenarios can be regarded as obsolete.

We may speculate that these findings can be explained by the general form of the decay amplitude (4.47), the structure of the Wigner rotation functions, and the parity-conserving nature of the electro-magnetic interaction. In decays of $D_s^{*+} \to D_s^+ \gamma$ this causes the photon to have two preferred decay directions

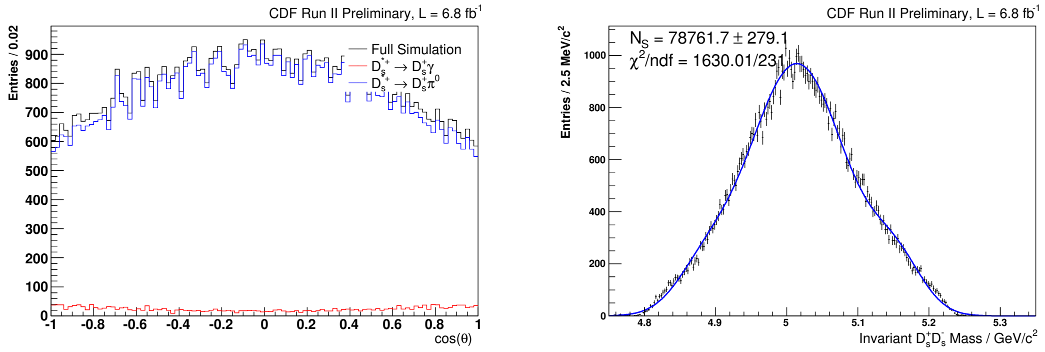


(a) (b)

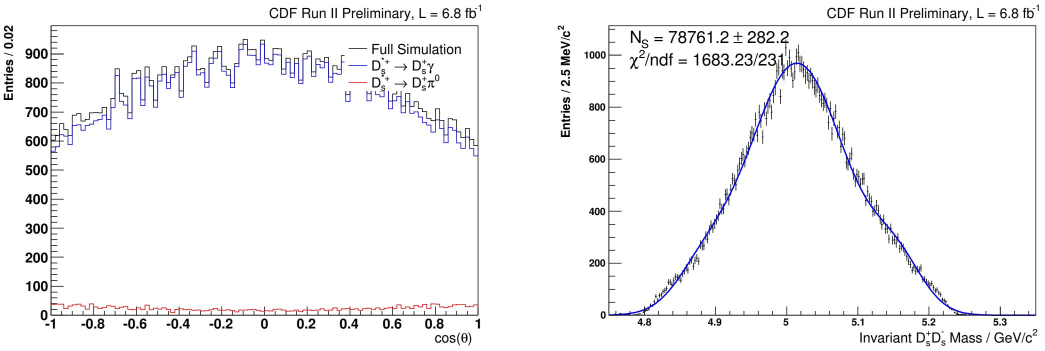
(c) (d)

(e) (f)

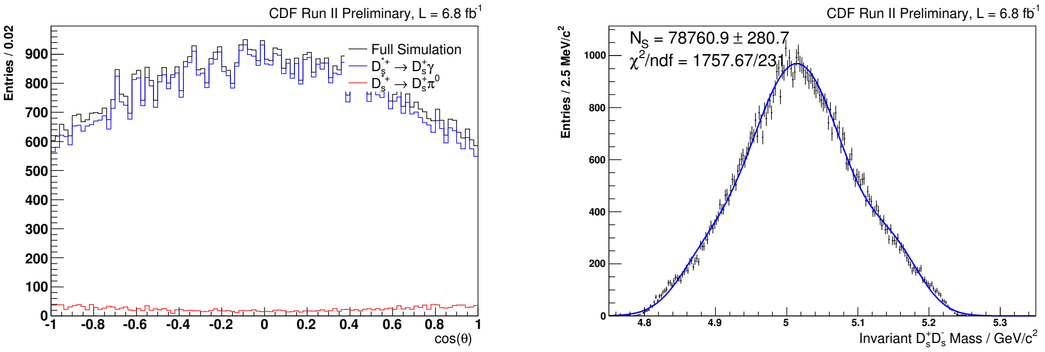
Figure 6.10: Monte Carlo truth $\cos(\theta)$ (left column) and resulting invariant mass (right column) distributions of $B_s^0 \to D_s^{*+} D_s^{*-}$ decays for the extreme helicity amplitude configurations $\{H_+, H_0, H_-\} = \{1, 0, 0\}$ (a,b), $\{0, 1, 0\}$ (b,c), and $\{0, 0, 1\}$ (e,f) obtained from fast EvtGen simulations.



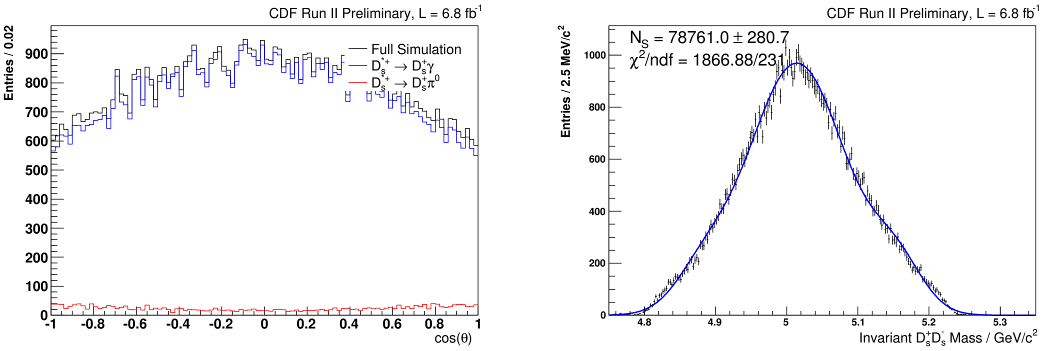
(a) $f_{CP-} = 0$: $H_+ = H_- = 0.4904$ (b) $f_{CP-} = 0$: $H_+ = H_- = 0.4904$



(c) $f_{CP-} = 0.08$: $H_\pm = 0.6478$, $H_\mp = 0.2478$ (d) $f_{CP-} = 0.08$: $H_\pm = 0.6478$, $H_\mp = 0.2478$

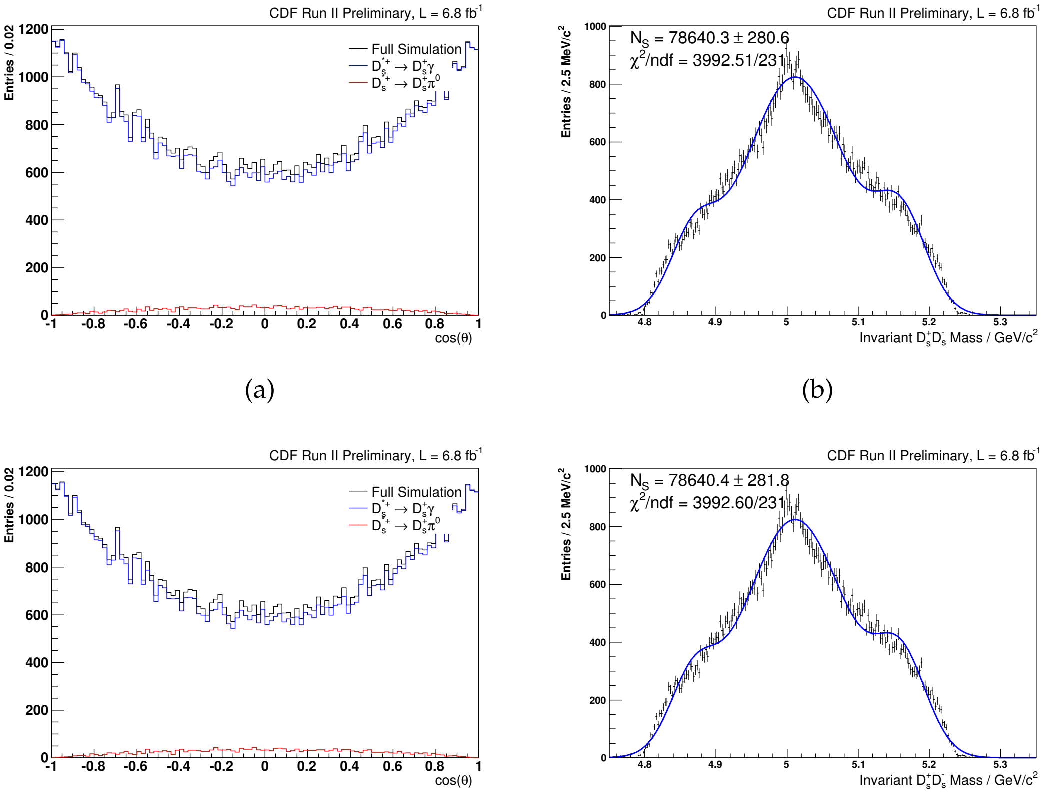


(e) $f_{CP-} = 0.16$: $H_\pm = 0.6835$, $H_\mp = 0.1172$ (f) $f_{CP-} = 0.16$: $H_\pm = 0.6835$, $H_\mp = 0.1172$



(g) $f_{CP-} = 0.24$: $H_\pm = 0.6936$, $H_\mp = 0.0000$ (h) $f_{CP-} = 0.24$: $H_\pm = 0.6936$, $H_\mp = 0.0000$

Figure 6.11: Monte Carlo truth cos(θ) (left column) and invariant mass (right column) distributions obtained by weighting $B_s^0 \to D_s^{*+}D_s^{*-}$ phase space Monte Carlo (full simulation) according to different CP-odd fractions.



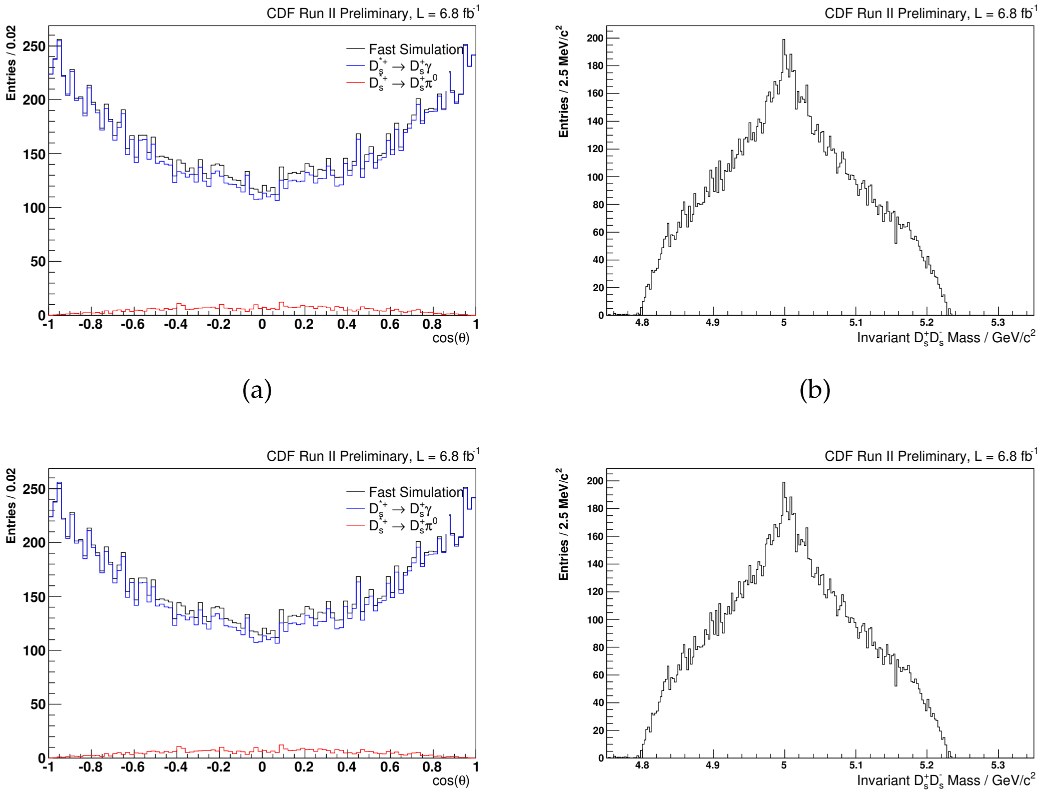
(a) (b)

(c) (d)

Figure 6.12: Monte Carlo truth cos(θ) (left column) and resulting invariant mass (right column) $B_s^0 \to D_s^{*+}D_s^{*-}$ distributions for the extreme helicity amplitude configurations $\{H_+, H_0, H_-\} = \{1, 0, 0\}$ (a,b) and $\{0, 0, 1\}$ (c,d) obtained by weighting phase space Monte Carlo (full simulation). In contrast to direct EvtGen simulations the distributions are all symmetric.

for each of the D_s^{*+} helicity projections H_+ and H_-: in the flight direction of the D_s^{*+} or opposite to it. This is because both left- and right-handiness of photons must be accounted for when incoherently summing over final state helicities. As a consequence, even if H_+ is unequal to the H_- amplitude the angular distribution in the helicity angle θ is always symmetric, and so is the mass line shape of partially reconstructed $B_s^0 \to D_s^{*+}D_s^{*-}$ decays. To illustrate this, phase space $B_s^0 \to D_s^{*+}D_s^{*-}$ Monte Carlo were weighted according to the extreme cases $\{H_+, H_0, H_-\} = \{1, 0, 0\}$ and $\{0, 0, 1\}$. Figure 6.12 shows the resulting cos(θ) and invariant $D_s^+D_s^-$ mass distributions. We therefore conclude that we are not able to make any inference on the *CP* odd fraction in $B_s^0 \to D_s^{*+}D_s^{*-}$ decays by partial reconstruction. The implication of this finding is two-edged: Neither can this study find any evidence of a non-zero *CP*-odd component, nor can $B_s^0 \to D_s^{*+}D_s^{*-}$ be confirmed to be *CP*-even by 100%.

However, these findings stand in clear contradiction to the preliminary evaluations made at the beginning of this section, where EvtGen $B_s^0 \to D_s^{*+} D_s^{*-}$ decay amplitude calculations were relied upon. The disagreement puts both the EvtGen modules responsible for the calculation of this process and the decay amplitude formula presented in this work, equation (4.47), under scrutiny. In fact, by valuable consultations with one of the EvtGen authors [130] a programming error in the EvtGen package could be revealed: Wrong summation over partial decay amplitudes led to incorrect computation of the $P \to VV$, $V \to P\gamma$ decay amplitude for helicity amplitude scenarios with $H_+ \neq H_-$. Figure 6.13 shows that by fixing this error angular and invariant mass distributions are now symmetric also for the considered extreme cases $\{H_+, H_0, H_-\} = \{1, 0, 0\}, \{0, 0, 1\}$. While the findings presented in this section are not satisfactory with regards to our efforts in determining f_{CP-} from partial reconstruction of $B_s^0 \to D_s^{*+} D_s^{*-}$, the in-depth studies made in the course of this survey at least helped to track down a programming error in a widely used decay simulator package.



(a) (b)

(c) (d)

Figure 6.13: Monte Carlo truth cos(θ) (left column) and resulting invariant mass (right column) $B_s^0 \to D_s^{*+} D_s^{*-}$ distributions for the extreme helicity amplitude configurations $\{H_+, H_0, H_-\} = \{1, 0, 0\}$ (a,b) and $\{0, 0, 1\}$ (c,d) obtained from fast simulations after fixing the programming error in the responsible EvtGen module.

7

Systematic Studies

7.1 Intermediate and Final State Branching Fractions

The results presented in Section 6.5 were obtained with some of the external parameters (intermediate and final state branching fractions $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$, $\mathcal{B}(B_d^0 \to D^+D_s^-)$, $\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$, and $\mathcal{B}(D^{*+} \to D^+X)$) kept fixed to their world average values [3]. For this reason, the uncertainties quoted are of statistical nature only. In a second step, uncertainties introduced by external parameters are accounted for by re-doing the simultaneous fit, but adding Gaussian constraints to the log likelihood fit function according to the uncertainties these external parameters are afflicted with. This allows these uncertainties to be propagated in the fit, while the parameters themselves are being strongly constrained to their central values. In this respect, one of the major benefits of the simultaneous fitting approach is the opportunity to directly evaluate impacts of variations of external parameters on the branching fraction result without the need for re-evaluating parameter correlations afterwards. As systematic uncertainties of $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ (Table 7.1) the root of the differences between the squared total uncertainties – those obtained with the uncertainties of intermediate and final state branching fractions propagated – and statistical fit uncertainties – those obtained in the standard fit – are assigned. It is not surprising that the branching fraction of the normalization channel $B_d^0 \to D^+D_s^-$ does not add any uncertainty. This can be traced back to the way the expected number of reconstructed events that are involved in the relative branching fraction measurement are parameterized (sections 6.3.6 and 6.4.3): $\mathcal{B}(B_d^0 \to D^+D_s^-)$ is a common factor to all relevant signal yield parameterizations, any variation thus cancels out.

External Input	$f_{D_sD_s}$	$f_{D_s^*D_s}$	$f_{D_s^*D_s^*}$	$f_{D_s^{(*)}D_s^{(*)}}$
$\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$	0.011	0.022	0.037	0.067
$\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$	0.006	0.010	0.012	0.031
$\mathcal{B}(B_d^0 \to D^+D_s^-)$	0.000	0.000	0.000	0.000
$\mathcal{B}(D^{*+} \to D^+X)$	0.000	0.000	0.000	0.000
Total $\mathcal{B}$	0.013	0.024	0.039	0.074

Table 7.1: Systematic uncertainties caused by uncertainties of intermediate and final state branching fractions.

7.2 Reconstruction and Selection

7.2.1 Two-Track Trigger Correction

Recording of B meson decay candidates of the fully hadronic meson decays studied in this analysis is triggered by an online algorithm, called Two-Track Trigger, that uses three different sub-scenarios basically depending on the transverse momentum of the decaying B meson. According to these scenarios recorded data can be sub-categorized into three exclusive sub-samples. Previous studies have shown (Section 4.7) that the proportions of these sub-samples are not well reflected in simulated data. To provide a more realistic description of real data, all Monte Carlo samples have been re-weighted with the weights calculated from a comparison of real and simulated data. Since the branching fractions of interest are measured relative to the branching fractions of topologically similar decays, to first order the effects of poorly simulated Two-Track Trigger fractions should cancel out in the ratios of reconstruction efficiencies. As a systematic check the standard simultaneous fits for $f_{D_sD_s}$, $f_{D_s^*D^*}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are repeated, however this time keeping the (poorly matching) proportions of the trigger sub-samples as simulated in the first place. The full deviations with respect to the central $f_{D_s^{(*)}D_s^{(*)}}$ results, that were obtained with reconstruction efficiencies computed from Two-Track Trigger corrected Monte Carlo, are conservatively assigned as systematic uncertainties (Table 7.2).

7.2.2 Multiple Candidates

The four studied decay channels have a very similar decay topology, each having a combination of six kaon or pion tracks in the final state. The similar topological decay structure is both a blessing and a curse. On the one hand, one benefits from the cancellation of the major fraction of acceptance and selection related effects.

Quantity	TTT Correction	No TTT Correction	Assigned Uncertainty
$f_{D_sD_s}$	0.184	0.185	0.001
$f_{D_s^*D_s}$	0.421	0.424	0.003
$f_{D_s^*D_s^*}$	0.654	0.644	0.010
$f_{D_s^{(*)}D_s^{(*)}}$	1.259	1.254	0.005

Table 7.2: Systematic uncertainties due to correction of Monte Carlo samples according to the fractions of two-track trigger sub-samples observed in side-band subtracted data. As systematic uncertainties the absolute differences between the default central values (second column) and the values determined from non-corrected Monte Carlo (third column) are conservatively assigned as uncertainties (last column).

On the other hand, during reconstruction a given candidate might be assigned to more than one exclusive decay channel. As a consequence, this candidate then populates the invariant mass spectra of other decay channels as well, a fact that might spoil the simultaneous fitting approach. If the fraction of multiple candidates among the different decay channels is significant the distributions entering the simultaneous fit cannot be treated as statistically independent, since the invariant mass in one sample will be correlated to the mass of another one. To phrase it differently, the total probability density function (pdf) does not factorize into the individual pdfs,

$$pdf(m_1, m_2) \neq pdf(m_1) \times pdf(m_2), \tag{7.1}$$

where equation (7.1) exemplarily covers the case of two non-independent data samples only.

To address this issue, in a first step the level of candidate cross-feed among the four invariant distributions is quantified. Once recorded and reconstructed, an event candidate is unambiguously identified by the number of the CDF run it was found in and a running event number. Table 7.3 shows the exclusive fraction of multiple events a particular data sample shares with any of the other three studied decay channels. The level of inclusive candidate cross-feed is illustrated by the plots in Figure 7.1. Table 7.3 indicates that already the individual channels contain a non-vanishing, but low amount of multiple candidates. To phrase it according to the geometrical arrangement of the plots in Figure 7.1, the level of "horizontal" candidate cross-feed is negligible. The "vertical" cross-feed is neither an issue of concern. Owing to the very similar decay topology there is however a remarkable level of candidate cross-feed in a "diagonal" sense among the channels $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$. The latter finding is not surprising, and the cross-feed due to false reconstruction has

Final State	$\phi^{\dagger}\pi^{+}\phi^{\dagger}\pi^{-}$	$\phi^{\dagger}\pi^{+}K^{*0\dagger}K^{-}$	$K^{-}\pi^{+}\pi^{+}\phi^{\dagger}\pi^{-}$	$K^{-}\pi^{+}\pi^{+}K^{*0\dagger}K^{-}$
$\phi^{\dagger}\pi^{+}\phi^{\dagger}\pi^{-}$	1.7%	0.1%	16.2%	0.0%
$\phi^{\dagger}\pi^{+}K^{*0\dagger}K^{-}$	0.1%	1.2%	37.3%	9.6%
$K^{-}\pi^{+}\pi^{+}\phi^{\dagger}\pi^{-}$	1.2%	4.6%	1.9%	0.6%
$K^{-}\pi^{+}\pi^{+}K^{*0\dagger}K^{-}$	0.0%	1.6%	0.8%	3.3%

Table 7.3: Fraction of multiple candidates a given decay channel (row) shares with any of the other studied decay channels (column).

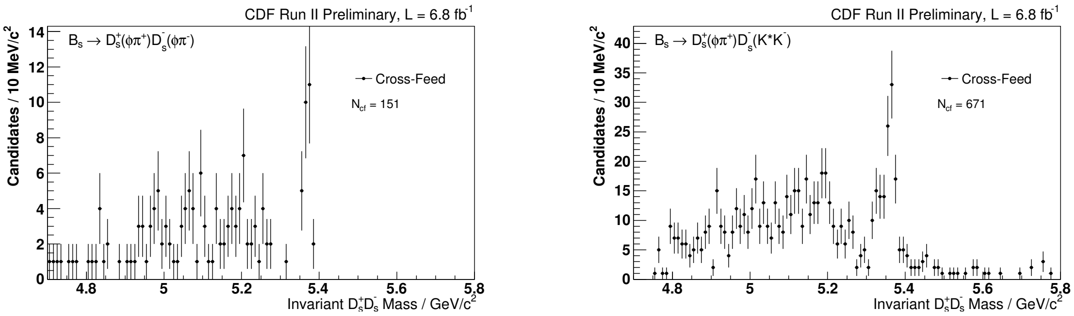
	No Cross-Feed Veto		
Quantity	**Central**	**Dalitz Fluct**	**Deviation**
$f_{D_sD_s}$	0.184 ± 0.021	0.166 ± 0.018	-0.018
$f_{D_s^*D_s}$	0.421 ± 0.046	0.379 ± 0.040	-0.042
$f_{D_s^*D_s^*}$	0.654 ± 0.073	0.586 ± 0.063	-0.068
	Cross-Feed Veto		
$f_{D_sD_s}$	0.186 ± 0.021	0.168 ± 0.018	-0.018
$f_{D_s^*D_s}$	0.428 ± 0.047	0.385 ± 0.041	-0.043
$f_{D_s^*D_s^*}$	0.664 ± 0.073	0.596 ± 0.064	-0.068

Table 7.4: Comparison of fit results obtained without a cross-feed veto (upper half) and with a cross-feed veto placed on the normalization channels (lower half). In addition to the central fit results in both scenarios one particular Dalitz model fluctuation is considered as systematic cross-check. The absolute deviations with respect to central values are equal for both scenarios.

already been accounted for by adequately parameterizing the respective yields.

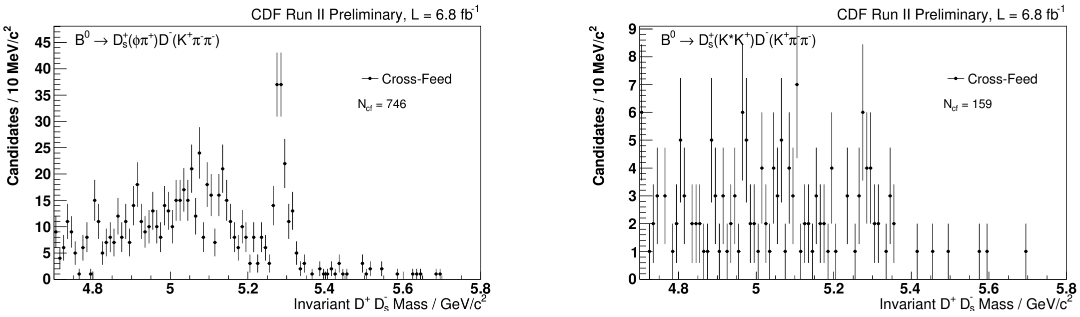
As a second step, the impact of the correlations inside the total fit pdf on the final results is evaluated. This is done by vetoing multiple candidates found in the normalization channels. Due to the structure of candidate cross-feeds described above, in doing so the level of mass correlation should be reduced down to a negligible level. To avoid a bias multiple candidates are also removed from simulated data. With the multiple candidate veto placed on both data and Monte Carlo of the normalization channels the standard simultaneous fitting procedure is repeated. The results (column labeled by "Central" in Table 7.4) indicate that the deviations with respect to the standard scenario (no multiple candidate veto) are below 0.2 statistical standard deviations. Given the overall precision of this measurement the deviation is thus negligible.

While the central values can be unaffected, mass correlations inside the total pdf might still introduce biases to the estimation of systematic uncertainties. To



(a) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$

(b) $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$



(c) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$

(d) $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+K^{*0\dagger}K^-$

Figure 7.1: Inclusive candidate cross-feeds. N_{cf} gives the number of cross-feed candidates a given data sample shares with any of the other decay channels. The level of cross-feed among the decays $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$ and $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ is remarkable, but not surprising.

address this concern, one particular variation of the $D_s^+ \to K^+K^-\pi^+$ Dalitz model[1] is considered that introduces $\sim 1\sigma$ downward fluctuations for $f_{D_sD_s}$, $f_{D_s^*D_s}$, and $f_{D_s^*D_s^*}$ in the standard non-veto scenario. Given the same Dalitz model variation, the simultaneous fit is repeated with the multiple candidate veto placed on the normalization channels. As can be seen from Table 7.4, the same absolute downward fluctuations with respect to the central values are observed.

We conclude that the presence of multiple candidates does not introduce a bias to the analysis results obtained from a joint likelihood fit, neither in terms of central values, nor in terms of systematic estimates. Therefore, no systematic uncertainty is assigned due to multiple candidate cross-feeds. Since removing multiple candidates from real and simulated data takes a considerable amount of time and would be mandatory for all systematic studies, for practical reasons the non-veto case is retained as the standard scenario.

7.2.3 $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$ Network Selection

The neural network working point for the decay $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$ determined by a significance scan over a range of neural network outputs was not very pronounced, particularly given the size of statistical uncertainties (see Figure 5.8). The question may arise if the neural network threshold chosen indeed represents a robust estimation of an optimal working point, and if the selection introduces biases to the final results. This can be verified by determining the ratios of branching fractions for a series of different network thresholds. This is done by scanning over the same neural network output region the neural network working point was determined from in steps of $\Delta NNout = 0.04$ and repeating the simultaneous likelihood fits for the ratios of branching fractions.

Figure 7.7 shows the result of the network scan. The left column of diagrams displays the branching fraction results, the right column of plots their relative uncertainties. From the latter group of plots it is clearly visible that the relative uncertainty of the measurement is virtually independent of the $K^{*0}K$ network cut chosen. In this sense, no neural network point is really preferred over another, and the choice of the network threshold is indeed somewhat arbitrary. Due to this freedom of choice we retain the original network working point.

When it comes to absolute values (left group of plots in Figure 7.7), care must be taken when interpreting the outcome of the network scan. At first glance, the results seem to be perfectly compatible within the statistical uncertainties indicated by the errors bars. However, the fitted samples are not statistically

---

[1] The Dalitz model represents one of the leading sources of the overall systematic uncertainty. See Section 7.3.2 for details.

Quantity	$NN > 0.00$	$NN > 0.98$	$\Delta R_{A,B}$	n
$f_{D_sD_s}$	0.168 ± 0.019	0.184 ± 0.020	0.016	0.41
$f_{D_s^*D_s}$	0.401 ± 0.044	0.421 ± 0.045	0.020	0.23
$f_{D_s^*D_s^*}$	0.589 ± 0.068	0.654 ± 0.070	0.065	0.48
$f_{D_s^{(*)}D_s^{(*)}}$	1.159 ± 0.092	1.259 ± 0.091	0.100	0.56

Table 7.5: Comparison of results determined from data samples selected from a low ($NN > 0.00$) and a high ($NN > 0.98$) neural network threshold. See text for further explanations.

independent, since every sample is a subset of any sample that was selected using a looser network cut. Given the results R_A, R_B with statistical uncertainties σ_A, σ_B determined from data samples having total candidate yields of s_A and s_B with $s_B \subset s_A$, the difference $\Delta R_{A,B}$ can be verified to be statistically compatible within $n\sigma$ using the relation

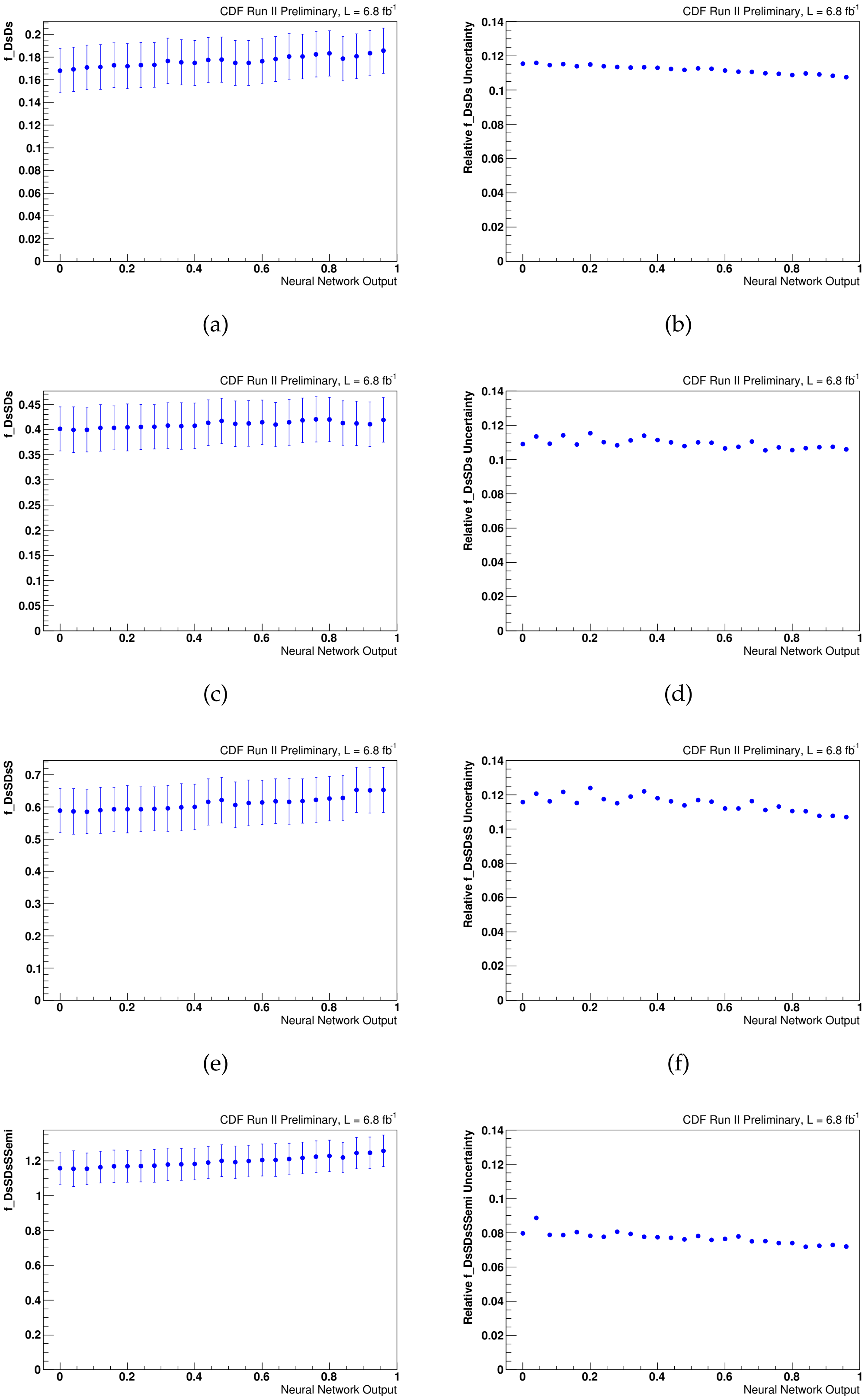
$$\Delta R_{A,B} = n\sigma_{A,B}\sqrt{\frac{s_A}{s_B} - 1} \tag{7.2}$$

Since the fluctuations of the results are not very high, only the values determined from the lowest and highest network working point are compared. The size of the associated data samples is $7,511$ and $1,569$ candidates, respectively. As shown in Table 7.5, the variations of the results are compatible with statistical fluctuations due to differences in the data samples.

7.3 Monte Carlo Simulation

As mass line shapes and selection efficiencies are estimated from simulated data, variations in Monte Carlo model assumptions might affect the measured branching fractions. In the upcoming sections the effects of variations of model assumptions entering Monte Carlo generation are scrutinized.

Beside Monte Carlo models, another source of uncertainty arises by the limited statistics of the Monte Carlo samples. To account for this aspect, in the simultaneous fits for $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ the statistical Monte Carlo uncertainties are propagated as Gaussian constraints to the efficiency parameters. As high-statistics simulations are at hand the relative uncertainties on the number of reconstructed Monte Carlo events are small, and the re-fitted values of $f_{D_s^{(*)}D_s^{(*)}}$ do not show any deviations from the central values. Therefore no systematic uncertainties are assigned due to Monte Carlo statistics.



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 7.2: Results (left column) and relative uncertainties (right column) of $f_{D_sD_s}$ (a,b), $f_{D_s^*D_s}$ (c,d), $f_{D_s^*D_s^*}$ (e,f), and $f_{D_s^{(*)}D_s^{(*)}}$ (g,h) obtained from changing the working points of the $K^{*0}K$ network in steps of 0.04.

7.3.1 *B* and *D* Meson Lifetimes

In Monte Carlo simulation *B* and *D* mesons are decayed following a decay law with a particular mean decay length $c\tau_{MC}$. The Two-Track Trigger as well as the pre- and final selection procedure place requirements on kinematical quantities that are highly correlated to the travel distance of the decaying particle. Therefore, different mean *B* meson lifetimes used in Monte Carlo generation may result in deviating trigger and reconstruction efficiencies. The effect of varying the mean meson lifetimes on the measured ratios $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ is evaluated by re-weighting simulated data according to different mean decay lengths $c\tau$. The per-event weights are calculated from the Monte Carlo truth information for the proper decay length ct by using relation (4.2)

$$w_{c\tau} = \exp\left(\frac{ct}{c\tau_{MC}} - \frac{ct}{c\tau}\right)$$

The B_s^0 and B_d^0 Monte Carlo samples used in this analysis were generated with mean decay lengths of $c\tau_{B_s^0} = 441\ \mu\text{m}$ ($\tau_{B_s^0} = 1.471$ ps) and $c\tau_{B_d^0} = 458.7\ \mu\text{m}$ ($\tau = 1.530$ ps). By default, prior to the determination of $f_{D_s^{(*)}D_s^{(*)}}$ all B_d^0 Monte Carlo have already been re-weighted using the more recent value $\tau_{B_d^0} = 1.519$. $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ Monte Carlo samples have been re-weighted using the lifetime of the short-living light mass eigenstate, $\tau_{B_s^L} = 1.408$ ps.

In order to quantify systematic effects on the measured branching fractions introduced by deviations of the central *B* meson lifetimes we proceed as follows: Each of the B_d^0 (including the channel $B_d^0 \to D^+D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ falsely reconstructed as $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$) and the B_s^0 Monte Carlo samples are re-weighted using deviating decay lengths $c\tau_{B_d^0}$ and $c\tau_{B_s^0}$, respectively, and the standard likelihood fits to data are repeated two times: In the first scenario the mean B_s^L decay length is varied downwards by 1σ, while the mean B_d^0 decay length is varied upwards by 1σ at the same time. In the second scenario, B_s^L/B_d^0 decay length assignments are swapped, with the B_d^0 decay length (now assigned to the B_s^0 meson) varied upwards by 1σ, and the B_s^L decay length (now assigned to the B_d^0 meson) varied downwards by 1σ. This procedure should cover any systematic effects related to *B* meson kinematics, and is motivated by these considerations:

- The final selection used joint networks trained for each combination of $B_s^0 \to D_s^+D_s^-$ and the topologically similar $B_d^0 \to D^+D_s^-$ normalization sub-channel. This procedure was motivated by the very similar lifetimes of the B_s^0 and B_d^0 meson. The swap in lifetimes (including 1 standard deviation) accounts for the non-perfect agreement by using the lifetime of the shorter-living light eigenstate B_s^L instead of the mean B_s^0 lifetime.

Quantity	$\sigma(c\tau_{B_s^L})/\sigma(c\tau_{B_d^0})$	$c\tau_{B_s^L} \leftrightarrow c\tau_{B_s^0}$	$\sigma(c\tau_{D_s^+})/\sigma(c\tau_{D^+})$	**Combined**
$f_{D_sD_s}$	+0.001	−0.003	0.001	$^{+0.001}_{-0.003}$
$f_{D_s^*D_s}$	+0.002	−0.008	0.001	$^{+0.002}_{-0.008}$
$f_{D_s^*D_s^*}$	+0.002	−0.012	0.002	$^{+0.003}_{-0.012}$
$f_{D_s^{(*)}D_s^{(*)}}$	+0.005	−0.023	0.004	$^{+0.006}_{-0.023}$

Table 7.6: Fluctuations of the ratios of branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ induced by 1 sigma variations of the B meson decay lengths (second column), by swapping $c\tau_{B_s^L}$ and $c\tau_{B_d^0}$ lifetime assignments (third column), and by randomly varying the D meson lifetimes in 250 trial fits (fourth column). The fifth column quotes the combined systematic uncertainties by adding the individual uncertainties in quadrature.

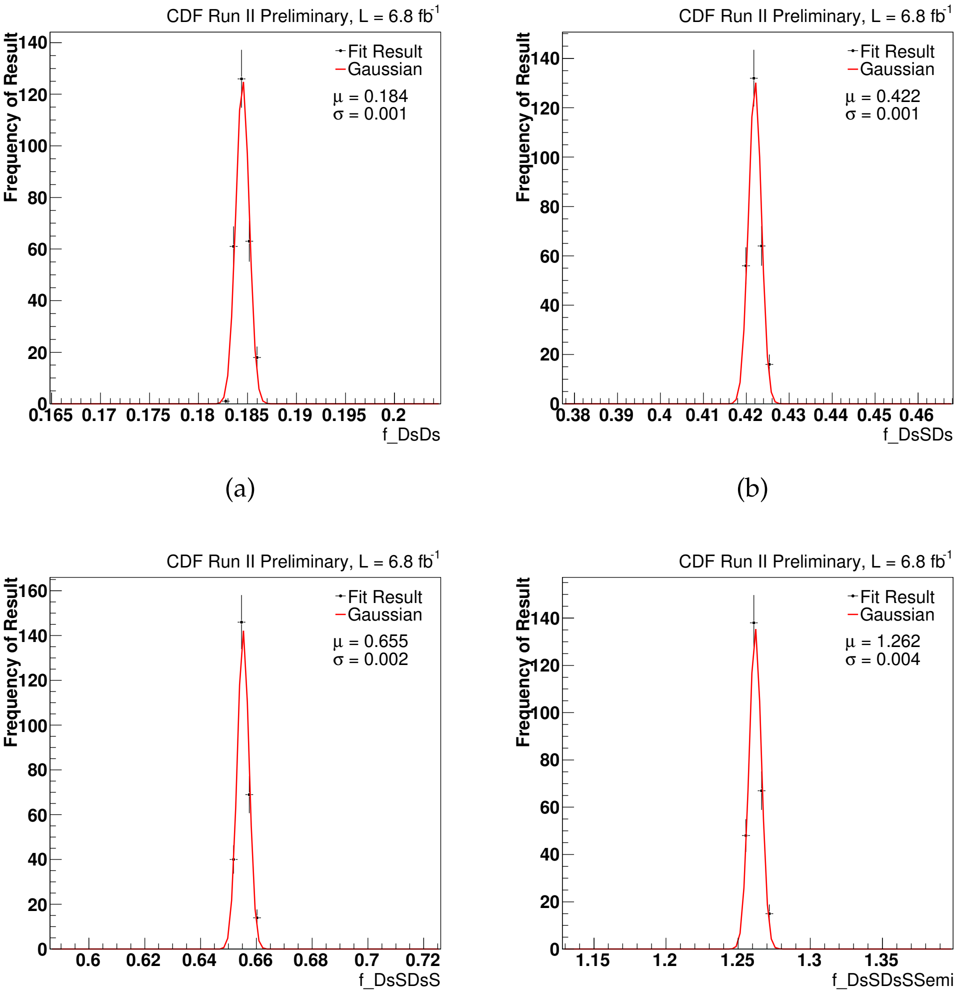
- By swapping lifetime assignments according to $c\tau_{B_s^L} \leftrightarrow c\tau_{B_d^0} \approx c\tau_{B_s^0}$ at the same time the B_s^0 meson is assigned the mean B_s^0 lifetime, which virtually coincides with the mean B_d^0 lifetime. In this way the scenario of a vanishing decay width difference $\Delta\Gamma_s$ is considered. Looking at $\Delta\Gamma_s = \Delta\Gamma_s^{CP} \cos\phi_s$, the case of a zero decay width difference could occur if the phase ϕ_s takes multiples of $\pi/2$, or if the CP width difference $\Delta\Gamma_s^{CP}$ vanishes.

Table 7.6 quotes the branching fraction ratios measured in the two scenarios. The down- and upward fluctuations are assigned as asymmetric systematic uncertainties.

The studied decay chains contain charged D and D_s mesons as long-living intermediate states. To study effects of variations in the D and D_s mean decay lengths these are allowed to randomly float within Gaussian uncertainties [3]. After re-weighting Monte Carlo using these random values, the standard likelihood fits to data are repeated. Figure 7.3 displays the results of 250 fit trials. The systematic uncertainties that are deduced from the Gaussian widths of the resulting distributions of the fitted ratios $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are negligible. We conclude that variations of the mean B and D meson decay lengths used in simulation have marginal impact on the measured branching fractions.

7.3.2 $D_s^+ \to K^+K^-\pi^+$ Dalitz Model

As described in Sections 4.4 and 4.6, decays of $D_s^+ \to K^+K^-\pi^+$ are simulated according to the Dalitz model using parameters (magnitudes, phases) measured by the CLEO collaboration [105]. The model parameters are afflicted with uncertainties though (Table 4.3). The effects of any variation of the Dalitz model are two-fold: Firstly, the fractions of $D_s^+ \to K^+K^-\pi^+$ events that populate the $\phi(1020)$



(a) (b)

(c) (d)

Figure 7.3: Fluctuations of the ratios of branching fractions $f_{D_s D_s}$ (a), $f_{D_s^* D_s}$ (b), $f_{D_s^* D_s^*}$ (c), and $f_{D_s^{(*)} D_s^{(*)}}$ (d) due to variations of the mean D_s^+ and D_+ decay lengths. See text for details.

and $K^{*0}(892)$ mass bands introduced by reconstruction undergo changes (cf. equations (4.23) and (4.24)). This is a direct effect of any change in the Dalitz parameter configuration. Secondly, as the proportions of the resonances in the considered $D_s^+ \to K^+K^-\pi^+$ decay model change, due to acceptance effects a variation of the Dalitz model can influence the reconstructed number of events lying inside the two mass bands. In this context, resolution effects at the boundaries of the mass bands may also play a role. In this section the effects of variations of the Dalitz model on the quantities $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ is quantitatively investigated.

For this purpose, the ratios of branching fractions are extracted using the standard simultaneous fitting method for a number of arbitrary but fixed configurations of the Dalitz model. Each model scenario is expressed in terms of a different set of the 12 Dalitz model parameters that are randomly generated. Following the toy parameter procedure described in Section 4.4.5 the parameters are correlated to each other according to the parameters' covariance matrix (see Section A.2 in the Appendix).

Before presenting the outcomes of the fit trials, the effects on $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are discussed in some more detail. For this it is worthwhile taking a closer look at equation (6.18), which specifies the number of expected fully reconstructed events in the decay channels $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ and $B_s^0 \to D_s^+D_s^- \to \phi^\dagger\pi^+K^{*0\dagger}K^-$ (though the discussion is analogous for equations (6.16), (6.20-6.21), (6.24-6.27) and (6.37-6.39). For ease of discussion equation (6.18) is repeated:

$$N^i_{B_s^0\to D_s^+D_s^-} = N^{tot}_{B_d^0} f_{D_sD_s} \mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger\pi^+ \\ D_s^+ \to K^{*0\dagger}K^+ \end{pmatrix}\epsilon^i_{B_s^0\to D_s^+D_s^-}$$

The equation is now re-formulated by introducing the relative fractions $f_{\phi^\dagger\pi^+}$ (4.21), $f_{K^{*0\dagger}K^+}$ (4.22) and writing out the efficiency term,

$$\begin{aligned} N^i_{B_s^0\to D_s^+D_s^-} =& N^{tot}_{B_d^0} f_{D_sD_s} \mathcal{B}(B_d^0 \to D^+D_s^-) \\ &\times f_{\phi^\dagger\pi^+}\mathcal{B}(D_s^+ \to K^+K^-\pi^+)\begin{pmatrix} f_{\phi^\dagger\pi^+} \\ f_{K^{*0\dagger}K^+} \end{pmatrix}\mathcal{B}(D_s^+ \to K^+K^-\pi^+) \\ &\times \frac{N^i_{rec}}{f_{\phi^\dagger\pi^+}\begin{pmatrix} f_{\phi^\dagger\pi^+} \\ f_{K^{*0\dagger}K^+} \end{pmatrix} N^{gen}_{KK\pi}}, \end{aligned} \tag{7.3}$$

where by $N^{gen}_{KK\pi}$ the number of simulated $B_s^0 \to D_s^+D_s^- \to K^+K^-\pi^+K^-K^+\pi^-$ decays is abbreviated. If the configuration of the considered Dalitz model changes, the quantities $f_{\phi^\dagger\pi^+}$, $f_{K^{*0\dagger}K^+}$, and N^i_{rec} are expected to change, too. Technically, the change in N^i_{rec} is induced by an event-by-event re-weighting of Monte Carlo.

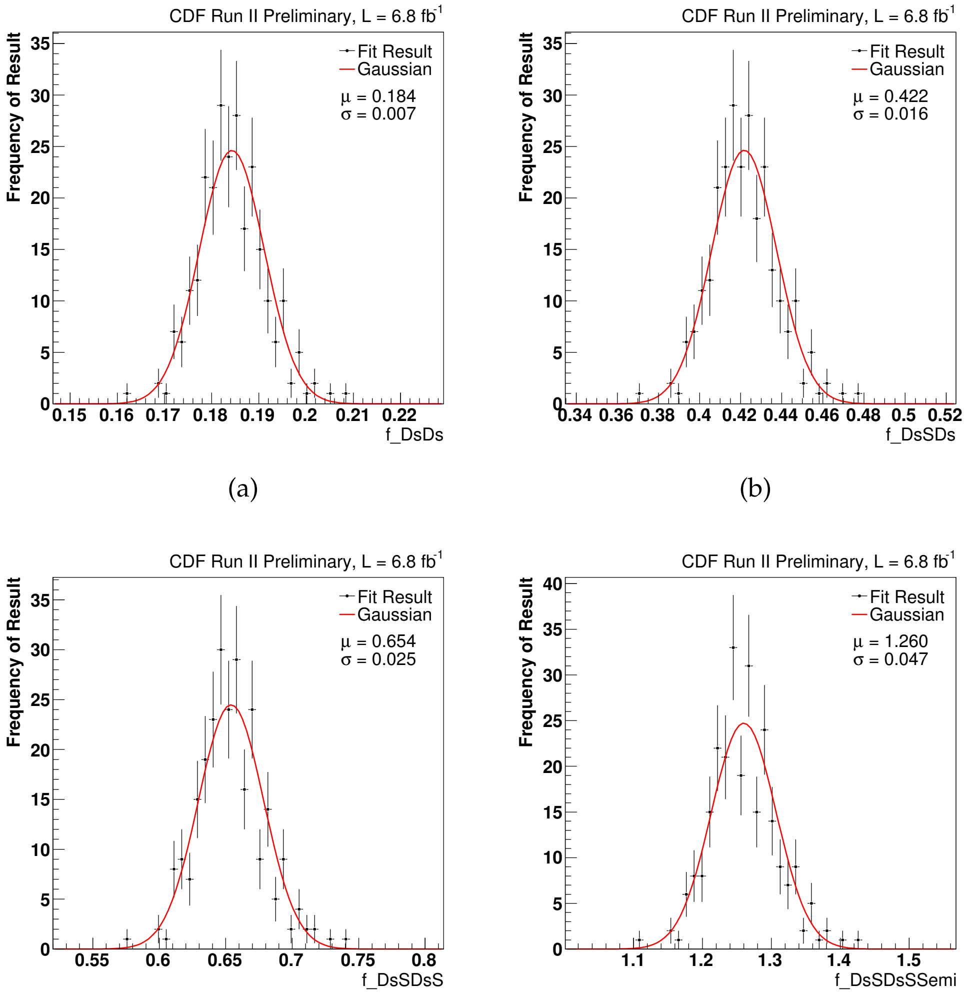
The weights per $D_s^+ \to K^+K^-\pi^+$ decay branch, $W(m^2_{K^+K^-}, m^2_{K^-\pi^+})$, are given by the squared decay amplitude using an arbitrary but fixed Dalitz parameter configuration, divided by the squared decay amplitude according to the Dalitz configuration simulated data have originally been generated with

$$W(m^2_{K^+K^-}, m^2_{K^-\pi^+}) = \frac{\left|\tilde{\mathcal{M}}(m^2_{K^+K^-}, m^2_{K^-\pi^+})\right|^2}{\left|\mathcal{M}(m^2_{K^+K^-}, m^2_{K^-\pi^+})\right|^2}. \tag{7.4}$$

The squared invariant masses of pairs of daughter particles, $m^2_{K^+K^-}$ and $m^2_{K^-\pi^+}$ are obtained from Monte Carlo Truth information. The variation in $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ can be evaluated using the procedure outlined in Section 4.4.5. However, care must be taken to ensure that the underlying Dalitz scenario has the same random but fixed parameter configuration the re-weighting procedure makes use of. Thereby the efficiency term in equation (7.3) solely reflects second-order effects in the form of variations in reconstruction efficiency, whereas changes in $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ reflect direct effects of Dalitz model variations.

From equation (7.3) it becomes apparent that a substantial simplification can be made: Since the quantities $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ cancel out for every arbitrary but fixed model variation, it is sufficient to vary N^i_{rec} only by re-weighting Monte Carlo. Any observed variation in the quantities $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ is hence the combined result of direct (variation in mass band fractions) and second-order effects (changing resonance fit fractions influencing acceptance) due to changes of the Dalitz model parameters. It must be emphasized though that direct and acceptance effects had to be studied separately if external values of $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ – e.g. stemming from a newer Dalitz plot analysis – differing from the those Monte Carlo were generated with were used. Secondly, the Dalitz plot mass band fractions might be correlated, a fact that also would need to be accounted for. However, the anti-correlation among $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ was found to be insignificant, as shown in Section 4.4.5.

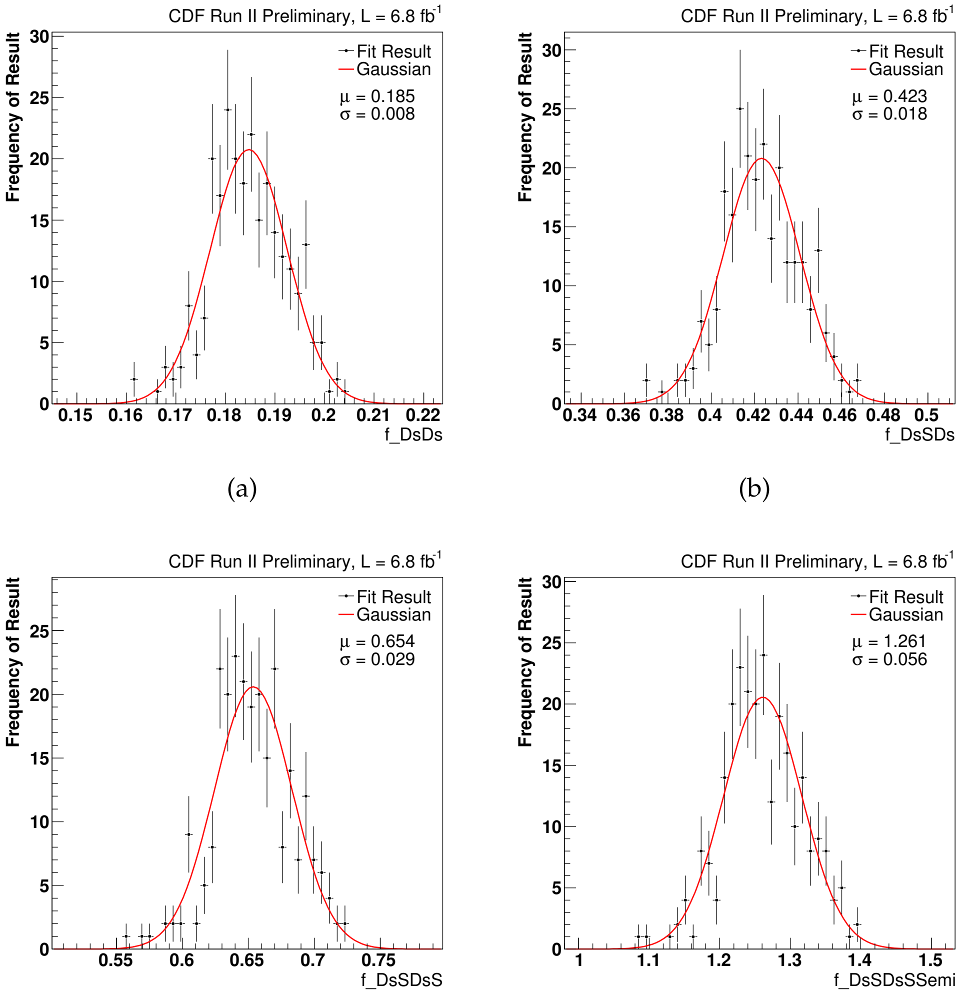
We perform a total of 500 simultaneous fits with Monte Carlo being re-weighted according to 500 different Dalitz universes, whereby variations within correlated Gaussian statistical and uncorrelated systematic uncertainties of the Dalitz model parameters are considered separately. The widths of the Gaussian distributions fitted to the corresponding result histograms are taken as systematic uncertainties of $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$. Table 7.7 lists the uncertainties derived from the distributions shown in Figures 7.4 and 7.5.



(a) (b)

(c) (d)

Figure 7.4: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$ (a), $f_{D_s^*D_s}$ (b), $f_{D_s^*D_s^*}$ (c), and $f_{D_s^{(*)}D_s^{(*)}}$ (d) induced by variations of the CLEO Dalitz model parameters within correlated statistical uncertainties.



(a)

(b)

(c)

(d)

Figure 7.5: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$ (a), $f_{D_s^*D_s}$ (b), $f_{D_s^*D_s^*}$ (c), and $f_{D_s^{(*)}D_s^{(*)}}$ (d) induced by variations of the CLEO Dalitz model parameters within uncorrelated systematic uncertainties.

Quantity	Dalitz Stat Errors	Dalitz Sys Errors	Combined
$f_{D_sD_s}$	0.007	0.008	0.011
$f_{D_s^*D_s}$	0.016	0.018	0.024
$f_{D_s^*D_s^*}$	0.025	0.029	0.038
$f_{D_s^{(*)}D_s^{(*)}}$	0.047	0.056	0.073

Table 7.7: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ induced by variations of the CLEO Dalitz model parameters within correlated statistical (second column) and uncorrelated systematic uncertainties (third column). The fourth column quotes the combined systematic uncertainties by adding the individual uncertainties in quadrature.

7.3.3 $B_s^0 \to D_s^{*+}D_s^{*-}$ Helicity Amplitudes

Section 6.6 demonstrated that both the distribution of helicity angles and the mass line shape of $B_s^0 \to D_s^{*+}D_s^{*-}$ do not depend on any changes of the helicity amplitudes H_+ and H_- as long the longitudinal polarization fraction is fixed. No systematic uncertainty is therefore assigned due to variations of f_{CP-}. For quantifying any systematic effect on the ratios of branching fractions it is thus sufficient to vary the longitudinal polarization fraction $f_L = |H_0|^2/(|H_+|^2 + |H_0|^2 + |H_-|^2)$ only, with $H_+ = H_-$ calculated from the normalization condition (6.45). The longitudinal polarization fraction is varied by generating random numbers around the central value $f_L(B_d^0 \to D^{*+}D_s^{*-}) = 0.519$[2] [119] by taking the assigned uncertainty ±0.057 as a Gaussian constraint. In doing so, the helicity amplitude H_0 fluctuates around the value $H_0 = 0.7204$ originally used in $B_s^0 \to D_s^{*+}D_s^{*-}$ Monte Carlo simulation. For each variation of H_0 the dedicated $B_s^0 \to D_s^{*+}D_s^{*-}$ phase space Monte Carlo sample is weighted using equation (4.47) and the standard simultaneous fit for $f_{D_s^{(*)}D_s^{(*)}}$ is repeated. The resulting values of $f_{D_s^{(*)}D_s^{(*)}}$ are plotted in histograms (Figure 7.6) and the widths of Gaussian fits are taken as systematic uncertainties (Table 7.8).

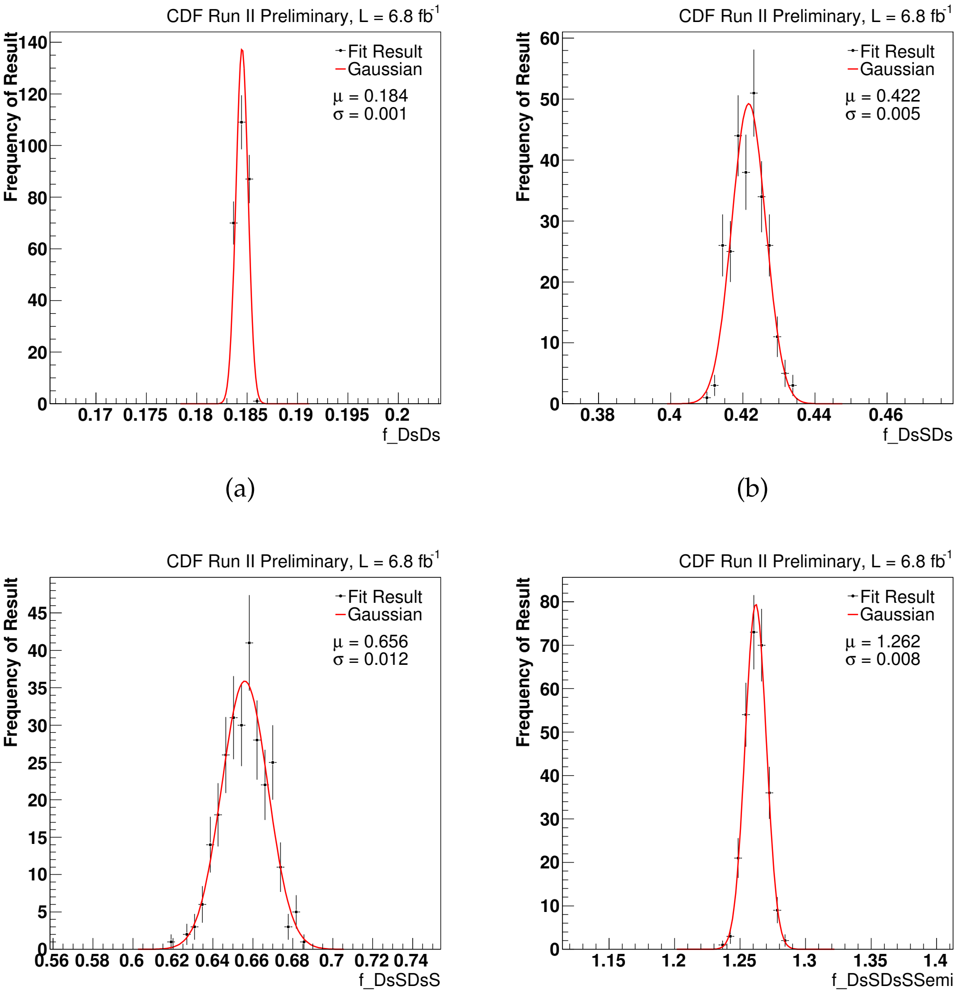
7.4 Fit

7.4.1 Signal Parameterization

The parameterizations of the signal-like contributions (fully and partially reconstructed signal events) are determined by means of fits to the Monte Carlo samples

---

[2] For the reasons outlined in Section 4.5.3 this is a reasonable estimate for $f_L(B_s^0 \to D_s^{*+}D_s^{*-})$.



(a) (b)

(c) (d)

Figure 7.6: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ induced by variations of the $B_s^0 \to D_s^{*+}D_s^{*-}$ helicity amplitudes H_+, H_0, and H_-. For the reasons set out above, in practice the amplitude $H_0 = \sqrt{f_L}$ is varied only.

Quantity	f_{CP-}	f_L	**Combined**
$f_{D_sD_s}$	0.000	0.001	0.001
$f_{D_s^*D_s}$	0.000	0.005	0.005
$f_{D_s^*D_s^*}$	0.000	0.012	0.012
$f_{D_s^{(*)}D_s^{(*)}}$	0.000	0.008	0.008

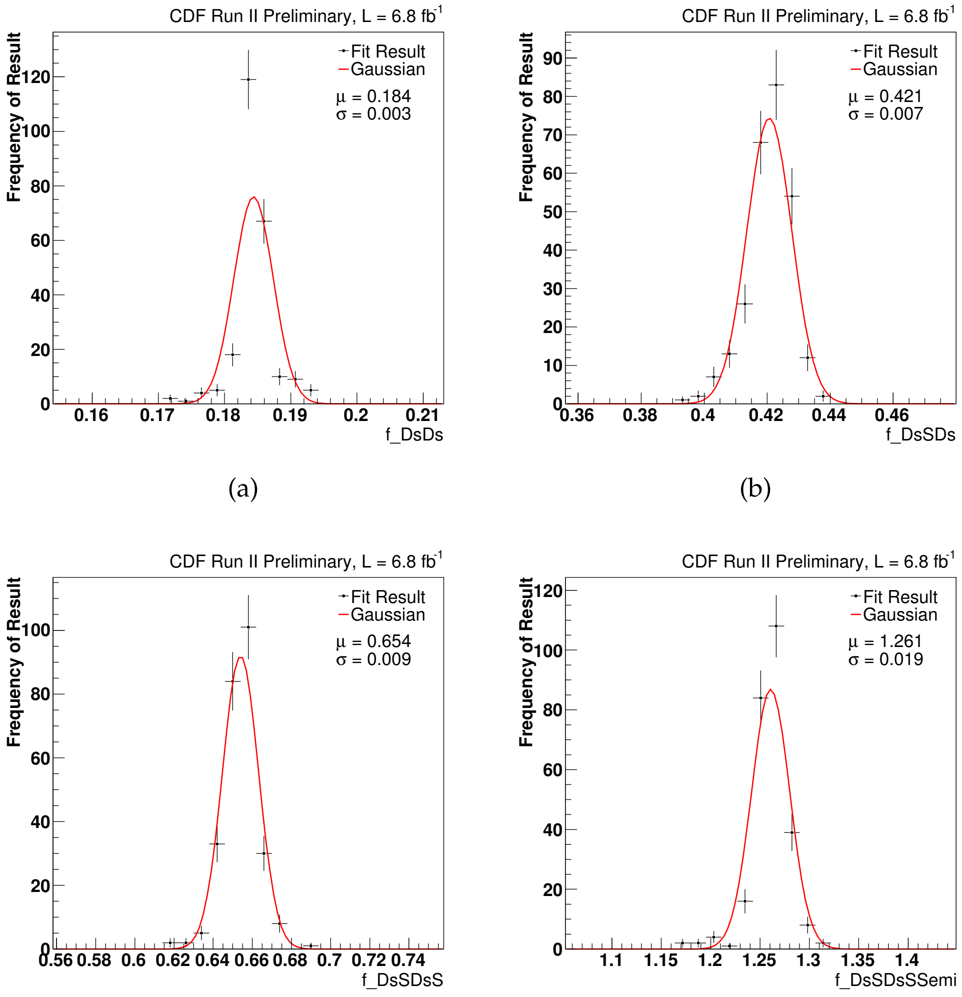
Table 7.8: Systematic uncertainties due to variations of $B_s^0 \to D_s^{*+}D_s^{*-}$ helicity amplitudes. Second column: No uncertainties are assigned due to changes of f_{CP-} (c.f. Section 6.6). Compared to statistical uncertainties, systematic uncertainties originating from variations of f_L (third column) are insignificant.

Quantity	**Assigned Uncertainty**
$f_{D_sD_s}$	0.003
$f_{D_s^*D_s}$	0.007
$f_{D_s^*D_s^*}$	0.009
$f_{D_s^{(*)}D_s^{(*)}}$	0.019

Table 7.9: Systematic uncertainties due to the variation of the parameterization of signal-like components in data.

generated for the respective components. Except for the shared means of fully reconstructed signal all the shape parameters are kept fixed in the final fit to data. This sections examines how variations in the shapes of the signal parameterizations affect the measured ratios of branching fractions.

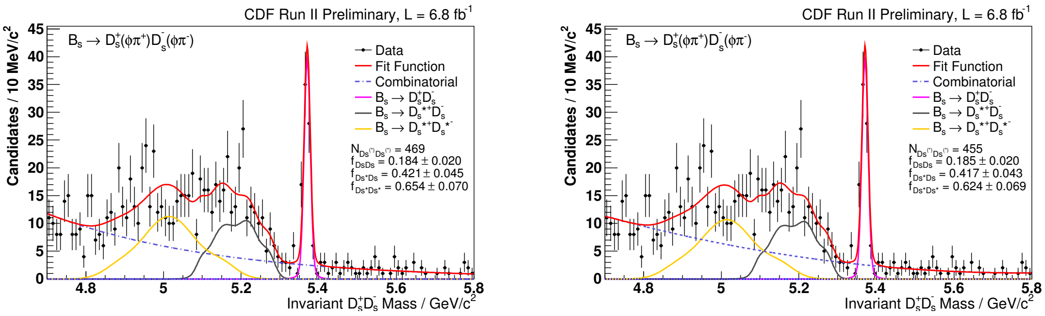
Propagating the uncertainties of the shape parameters in the full fit function the way it was done in the case of intermediate and final state branching fractions (Section 7.1) is not a practicable option: The amount of parameters needed to describe all the contributions of partially reconstructed events is huge, and so would be the number of free parameters in the full fit. To overcome this problem, the full simultaneous fit is run repeatedly, with each fit trial using fixed but slightly varied signal component parameters. This is achieved by generating correlated Gaussian random numbers taking into account the full covariance matrices resulting from the Monte Carlo template fits. In this way, 250 fits are performed, and the results of $f_{D_sD_s}$, $f_{D_s^*D_s^*}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are plotted into histograms. As systematic uncertainties the widths of Gaussian functions fitted to the result histograms are assigned. Please see Table 7.9 and Figure 7.7 for details. Compared to statistical and the leading systematic uncertainties the assigned uncertainties are insignificant.



(a) (b)

(c) (d)

Figure 7.7: Result distributions of $f_{D_sD_s}$ (a), $f_{D_s^*D_s}$ (b), $f_{D_s^*D_s^*}$ (c), and $f_{D_s^{(*)}D_s^{(*)}}$ (d) obtained from 250 trial fits, each time randomly varying the parameterizations of all signal-like components.



(a) Exponential plus constant (b) Second order polynomial

Figure 7.8: Comparison of the mass projections of the decay channel $B_s^0 \to D_s^+ D_s^- \to \phi^+\pi^+\phi^+\pi^-$ using the default parameterization for the background component (a) in the simultaneous fit and alternatively using a second order polynomial function to describe background (b).

7.4.2 Background Parameterization

As stated above, the combinatorial background component is parameterized by a sloping exponential function plus a constant. Both the shape parameters are left completely free in the fit. Consequently, the uncertainty of – this particular – background parameterization is already accounted for and included in the statistical uncertainties quoted in equations (6.40) through (6.43).

Since the parameterization chosen lacks any physical motivation – and cannot be verified on a clean statistically independent sample – a function having a similar slope and amount of parameters may describe the background component in data equally well, but might lead to different results in terms of branching fractions ratios. To investigate the influence of a slightly differently shaped background the simultaneous fits are repeated using a second order polynomial function, that needs an equal amount of parameters to be described. As systematic uncertainties the full deviations from the central results for $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are assigned (Table 7.10). The relatively large deviation in the case of $f_{D_s^*D_s^*}$ can be attributed to the different shapes of the exponential and the polynomial parameterization in the lower region of the fit range, as can be seen from Figure 7.8.

7.4.3 Fit Validity

The simultaneous fitting method might introduce a systematic bias, pushing the values of $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ in one direction or another. The potential

Quantity	Exponential	2nd Order Polynomial	Assigned Uncertainty
$f_{D_sD_s}$	0.184	0.185	0.001
$f_{D_s^*D_s}$	0.421	0.417	0.004
$f_{D_s^*D_s^*}$	0.654	0.624	0.030
$f_{D_s^{(*)}D_s^{(*)}}$	1.259	1.226	0.033

Table 7.10: Systematic uncertainties due to parameterization of the combinatorial background component in data. As systematic uncertainties the absolute differences between the default central values obtained in a fit using an exponential function plus constant (second column) and the values determined in a fit using a second order polynomial instead (third column) are conservatively assigned as uncertainties (last column).

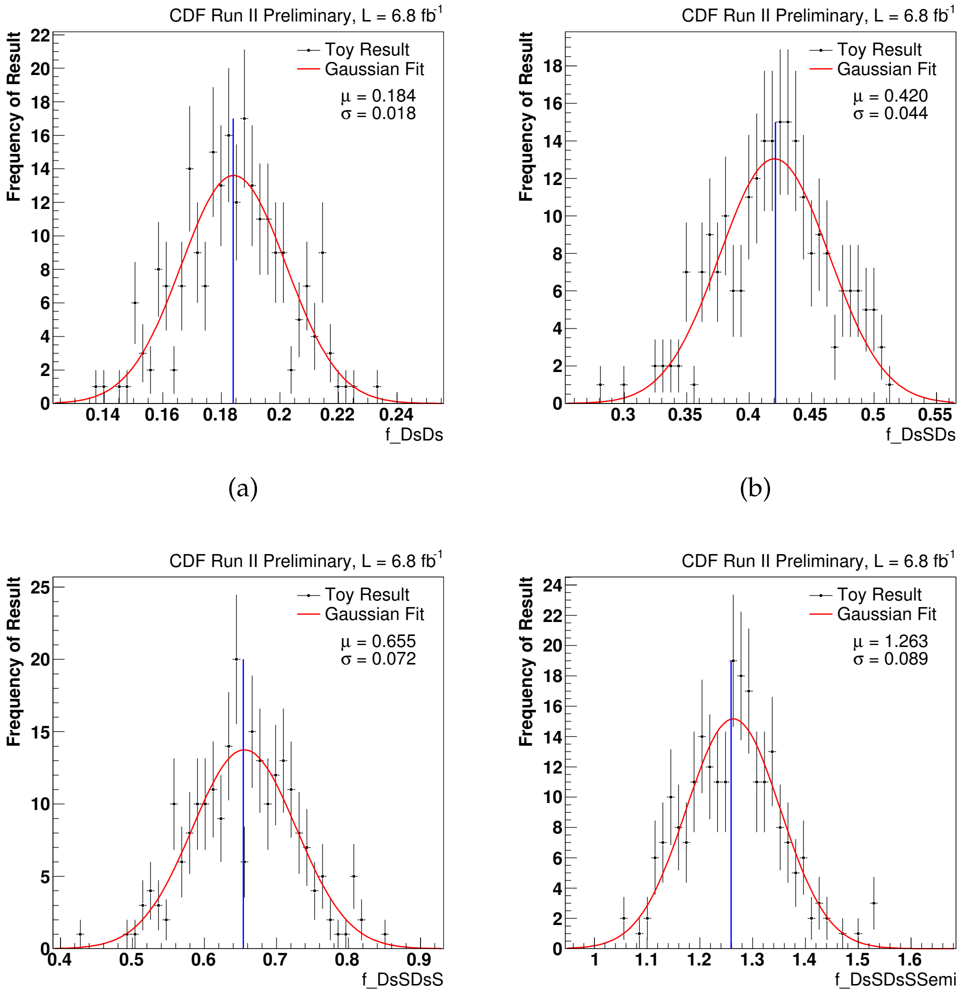
presence of a systematic shift can be investigated by means of toy experiments: According to the final parameter values determined in the simultaneous fit to real data, random mass distributions for all the studied channels are generated. This is done by calculating the per-event likelihood for a given random invariant mass lying inside the fitting range and employing the accept/reject method. The amount of events per decay mode randomly generated are equal to the number observed in the original real data sample. Generation of random mass spectra is performed for all four decay channels studied. For each combination of random invariant mass distributions simultaneous fits for $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ are carried out. In this way, 250 toy experiments are run in total, and the fit results are filled into histograms (Figure 7.9). The validity of the fit can be reviewed by evaluating *pull* distributions, where the per-trial pull p_{trial} of a given fit parameter x is defined as

$$p_{trial} = \frac{x_{trial} - x_{central}}{\sigma_{x_{trial}}}. \tag{7.5}$$

$x_{central}$ denotes the central parameter value having a symmetric uncertainty $\sigma_{x_{trial}}$ determined in a fit to real data. Since the statistical uncertainties of $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ determined in the original fits to data are slightly asymmetric, the pull relation for asymmetric uncertainties $\sigma^+_{x_{trial}}$, $\sigma^-_{x_{trial}}$ has to be used:

$$\begin{aligned} p_{trial} &= \frac{x_{trial} - x_{central}}{\sigma^+_{x_{trial}}} \quad \text{for } x_{trial} \leq x_{central} \\ p_{trial} &= \frac{x_{trial} - x_{central}}{\sigma^-_{x_{trial}}} \quad \text{for } x_{trial} > x_{central} \end{aligned} \tag{7.6}$$

For a bias-free fit, i.e. a fit in which result variations are driven by statistical fluctuations of the data sample only, one expects a pull distribution centered at 0 and having a width of 1. Figure 7.10 shows the pull distributions for the fit

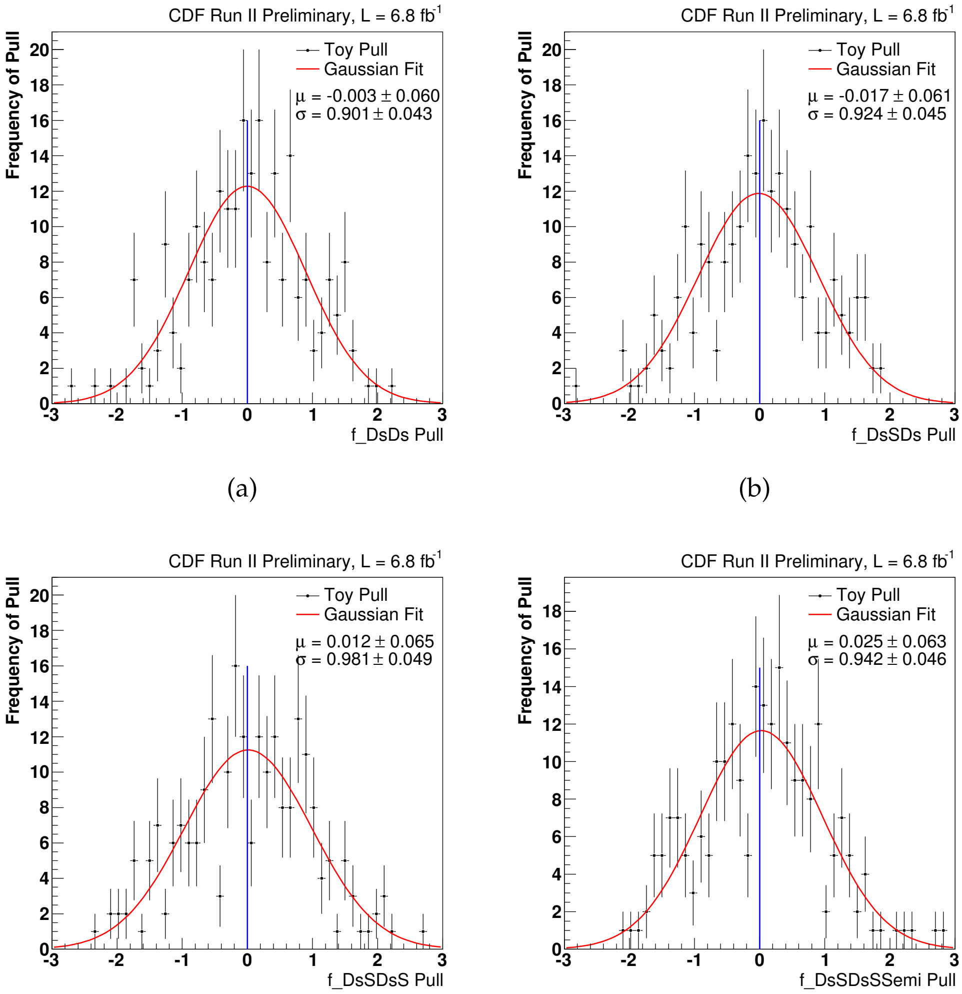


(a) (b)

(c) (d)

Figure 7.9: Distributions of the $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ results determined in 250 toy experiments.

parameters $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$. We conclude that, according to the amount of toy experiments run, within 0.5σ the pull means are compatible with a bias-free fit. The toy widths – particularly for $f_{D_sD_s}$, see Figures 7.9(a) and 7.10(a) – tend to be marginally smaller than the uncertainties determined in the standard fit to real data. This could be explained by the standard fit slightly overestimating the statistical error. No additional systematic uncertainties due to the fitting method are assigned.



(a) (b)

(c) (d)

Figure 7.10: Pull distributions of the $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ results determined in 250 toy experiments.

Source	$f_{D_sD_s}$	**Rel**	$f_{D_s^*D_s}$	**Rel**	$f_{D_s^*D_s^*}$	**Rel**	$f_{D_s^{(*)}D_s^{(*)}}$	**Rel**
Branching Fractions	0.013	0.07	0.024	0.06	0.039	0.06	0.074	0.06
MC Statistics	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
TTT Correction	0.001	0.01	0.003	0.01	0.010	0.02	0.005	0.00
B, D Lifetimes	$^{+0.001}_{-0.003}$	$^{+0.01}_{-0.02}$	$^{+0.002}_{-0.008}$	$^{+0.00}_{-0.02}$	$^{+0.003}_{-0.012}$	$^{+0.00}_{-0.02}$	$^{+0.006}_{-0.023}$	$^{+0.00}_{-0.02}$
Dalitz Model	0.011	0.06	0.024	0.06	0.038	0.06	0.073	0.06
Helicity Model	0.001	0.01	0.005	0.01	0.012	0.02	0.008	0.01
Signal Model	0.003	0.02	0.007	0.02	0.009	0.01	0.019	0.02
Background Model	0.001	0.01	0.004	0.01	0.030	0.05	0.033	0.03
Total	$^{+0.017}_{-0.018}$	$^{+0.09}_{-0.10}$	$^{+0.035}_{-0.036}$	$^{+0.08}_{-0.09}$	$^{+0.065}_{-0.066}$	0.10	$^{+0.111}_{-0.113}$	0.09

Table 7.11: Overview of absolute and relative systematic uncertainties. The total systematic uncertainties are calculated by adding the individual ones in quadrature. Due to the uncertainties of f_s/f_d and $\mathcal{B}(B_d^0 \to D^+D_s^-)$, further uncertainties will contribute to the absolute branching fractions $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$.

7.5 Overview

The systematics considered throughout the foregone sections are finally summarized and combined. The individual absolute and relative uncertainties are presented in Table 7.11. The leading sources of uncertainty introduced by analysis techniques are the combined uncertainties of intermediate and final state branching fractions and the uncertainty in the parameterization of the $D_s^+ \to K^+K^-\pi^+$ Dalitz model.

8

Results and Discussion

8.1 Absolute Branching Fractions and Implications on $\Delta\Gamma_s/\Gamma_s$

By combining the individual and total systematic uncertainties estimated in Section 7 we may write the results of the exclusive branching fractions ratios and the semi-inclusive branching fraction ratio:

$$f_{D_sD_s} = 0.184^{+0.021}_{-0.020}(stat)^{+0.017}_{-0.018}(sys) \tag{8.1}$$

$$f_{D_s^*D_s} = 0.421^{+0.046}_{-0.045}(stat)^{+0.035}_{-0.036}(sys) \tag{8.2}$$

$$f_{D_s^*D_s^*} = 0.654^{+0.073}_{-0.071}(stat)^{+0.065}_{-0.066}(sys) \tag{8.3}$$

$$f_{D_s^{(*)}D_s^{(*)}} = 1.259^{+0.094}_{-0.092}(stat)^{+0.111}_{-0.113}(sys) \tag{8.4}$$

For the exclusive ratios $f_{D_sD_s}$, $f_{D_s^*D_s}$, and $f_{D_s^*D_s^*}$ the total systematic uncertainties are slightly smaller than statistical uncertainties, while for the sum of them, $f_{D_s^{(*)}D_s^{(*)}}$, which is deduced from the inclusive yield, the opposite is observed.

From the ratios (8.1) through (8.4) absolute branching fractions are calculated by inserting the current world average value $\mathcal{B}(B_d^0 \to D^+D_s^-) = (7.2 \pm 0.8) \times 10^{-3}$ [3] and the relative B_s^0 to B_d^0 meson production rate at Tevatron energies, $f_s/f_d = 0.269 \pm 0.033$ [131], into equations (6.19, 6.22, 6.23):

$$\mathcal{B}(B_s^0 \to D_s^+D_s^-) = (0.49^{+0.06}_{-0.05}(stat) \pm 0.05(sys) \pm 0.08(norm))\% \tag{8.5}$$

$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^-) = (1.13 \pm 0.12(stat)^{+0.09}_{-0.10}(sys) \pm 0.19(norm))\% \tag{8.6}$$

$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^{*-}) = (1.75^{+0.20}_{-0.19}(stat)^{+0.17}_{-0.18}(sys) \pm 0.29(norm))\% \tag{8.7}$$

$$\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (3.37 \pm 0.25(stat) \pm 0.30(sys) \pm 0.56(norm))\% \tag{8.8}$$

The uncertainties of the branching fraction of the normalization channel $B_d^0 \to D^+D_s^-$ and the ratio of quark fragmentation fractions f_s/f_d contribute further uncertainties. These were added in quadrature and labeled by (*norm*). Equations (8.5) - (8.8) indicate that this is the leading source of uncertainty. Apparently, this analysis would considerably benefit from more precise measurements of $\mathcal{B}(B_d^0 \to D^+ D_s^-)$ and f_s/f_d. Whenever new values are available, it is straightforward to re-calculate the absolute branching fractions from the relative branching fractions given by equations (8.1) - (8.4).

As discussed in Section 1, under certain theoretical assumptions $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decays saturate decays to CP-even final states. Furthermore taking CP violation in the $B_s^0 - \bar{B}_s^0$ system to be negligibly small the relationship specified in equation (1.41),

$$\frac{\Delta\Gamma_s}{\Gamma_s} \cong \frac{2\mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})}{1 - \mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})},$$

may be used to estimate the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ system from the branching fraction of semi-inclusive $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decays. Inserting the calculated branching fraction value (8.8) and propagating all uncertainties, one obtains:

$$\begin{aligned} \frac{\Delta\Gamma_s}{\Gamma_s} &= (6.97^{+0.50}_{-0.49}(stat)^{+0.59}_{-0.60}(sys) \pm 1.11(norm) \pm 0.34(theo))\% \\ &= (6.97^{+1.39}_{-1.40})\% \end{aligned} \tag{8.9}$$

In the estimation of $\Delta\Gamma_s/\Gamma_s$ a theoretical uncertainty of ±5% has been added. This arises [31] from the potential presence of a small but non-zero CP-odd component in $B_s^0 \to D_s^{*+} D_s^-$ and $B_s^0 \to D_s^{*+} D_s^{*-}$ and contributions from two-body decay modes other than $b \to c\bar{c}s$ that are common to B_s^0 and $\bar{B}_s^0$.

8.2 Statistical Significance

The statistical significances of the relative branching fraction measurements are calculated from $\sqrt{-2\ln(\mathcal{L}_0^f/\mathcal{L}_{max})}$, where $\mathcal{L}_{max}$ is the value of the likelihood function for the central branching fraction values found in the standard fit, and $\mathcal{L}_0^f$ are the likelihood values when each of the branching fractions of $B_s^0 \to D_s^+ D_s^-$, $B_s^0 \to D_s^{*+} D_s^-$, $B_s^0 \to D_s^{*+} D_s^{*-}$, and $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ are fixed to zero one by one.

The statistical significances quoted in Table 8.1 clearly indicate that the present analysis confirms observation of the decay modes $B_s^0 \to D_s^+ D_s^-$, $B_s^0 \to D_s^{*+} D_s^-$, and $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$. In addition, this thesis claims first observation of $B_s^0 \to D_s^{*+} D_s^{*-}$ with a statistical significance above 10σ.

Quantity	$f_{D_sD_s}$	$f_{D_s^*D_s}$	$f_{D_s^*D_s^*}$	$f_{D_s^{(*)}D_s^{(*)}}$
Significance (σ)	16.8	12.2	10.9	20.6

Table 8.1: Statistical significances for the observations of the exclusive and the semi-inclusive decay modes.

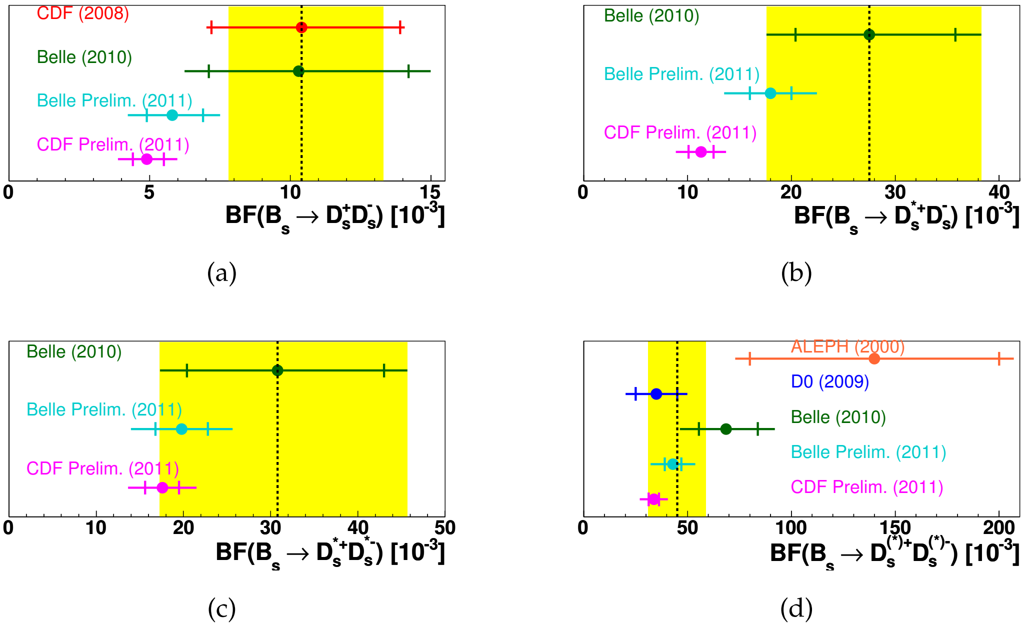
	DØ	CDF	Belle	CDF 2011
Signal Yield	27	24	23	745
$\mathcal{B}(B_s^0 \to D_s^+ D_s^-)$ (%)	-	$1.04^{+0.35+1.1}_{-0.32-1.1}$	$1.03^{+0.39+0.26}_{-0.32-0.25}$	$0.49^{+0.06+0.09}_{-0.05-0.09}$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)$ (%)	-	-	$2.75^{+0.83+0.69}_{-0.71-0.69}$	$1.13^{+0.12+0.21}_{-0.12-0.21}$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})$ (%)	-	-	$3.08^{+1.22+0.84}_{-1.04-0.84}$	$1.75^{+0.20+0.34}_{-0.19-0.34}$
$\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ (%)	$3.5 \pm 1.0 \pm 1.1$	-	$6.85^{+1.53+1.89}_{-1.30-1.89}$	$3.37^{+0.25+0.63}_{-0.23-0.63}$
$\Delta\Gamma_s/\Gamma_s$ (%)	$7.2 \pm 2.1 \pm 2.2$	> 1.2	$14.7^{+3.6+4.4}_{-3.0-4.2}$	$6.97^{+0.50+1.30}_{-0.49-1.31}$

Table 8.2: Comparison of available results and the preliminary results presented in this document. "Signal Yield" refers to the total number of any kind of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decay events a given experiment was able to reconstruct.

8.3 Comparison of Results

The presented results (8.5) - (8.8) are now being embedded in the context of the current experimental status. The estimated relative decay width difference (8.9) is directly comparable to the values obtained by the previous analyses published in the Refs. [35, 36, 39], where $\Delta\Gamma_s/\Gamma_s$ was determined within the same theoretical regime the present estimation makes use of. At the cost of the oldest result by the ALEPH collaboration, Table 8.2 gives an overview of previous results and the preliminary results contributed by the present analysis. To enhance comparability, Figure 8.1 illustrates the compilation of results in a graphical way.

The results contributed by this analysis tend to be lower than or at the lower edge of the world average values: Our value of $\mathcal{B}(B_s^0 \to D_s^+ D_s^-)$ is 1.9σ below the world average value, $\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})$ is off by 1.6σ. In the latter case the world average value yet consists of the Belle measurement only. Both $B_s^0 \to D_s^{*+} D_s^{*-}$ and $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ are compatible with the world average values within 1σ. Recently, the Belle collaboration presented new preliminary results on the basis of the full $\Upsilon(5S)$ dataset holding an integrated luminosity of 121.4 fb^{-1} [132], indicated by *Belle Prelim. (2011)* in Figure 8.1. While these recent values are not yet included in the calculated average (the yellow band), they show a clear tendency towards lower branching fraction values. Including these values in the average calculation



(a) (b)

(c) (d)

Figure 8.1: Comparison of $B_s^0 \to D_s^+ D_s^-$ (a), $B_s^0 \to D_s^{*+} D_s^-$ (b), $B_s^0 \to D_s^{*+} D_s^{*-}$ (c), and $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ (d) branching fraction measurements. The preliminary results *CDF Prelim. (2011)* (this analysis) and *Belle Prelim. (2011)* were not yet considered in the calculated averages (yellow bands).

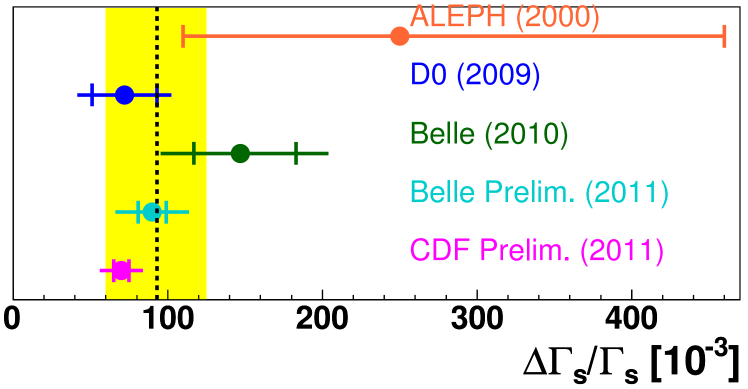


Figure 8.2: Comparison of results of the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ estimated from $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$. The average value (yellow band) is calculated from lifetime measurements and non-preliminary $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$ measurements [40].

would certainly reduce the tension the presented results partly have with the current world averages considerably.

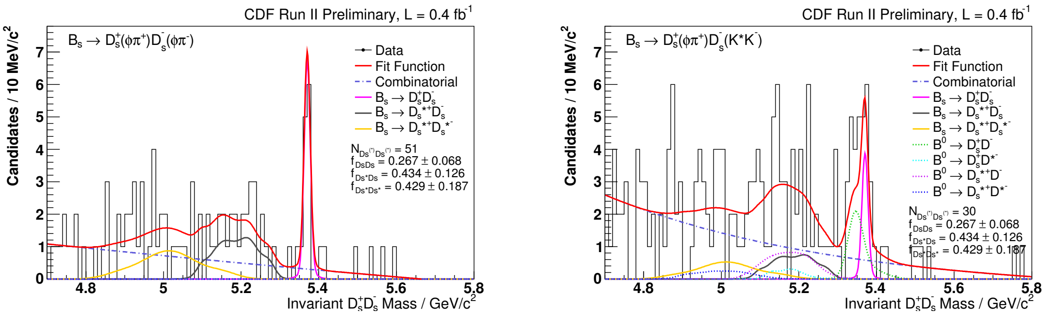
Even with the inclusion of the preliminary results presented by the Belle collaboration, the given measurements of both the individual exclusive branching fractions and the semi-inclusive branching fraction of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays represent the world's most precise measurements of these quantities. As the concluding discussions in the last Chapter to come will show, it is difficult to make any definite statement with regards to the accuracy and robustness of the estimated value of the relative decay width difference $\Delta\Gamma_s/\Gamma_s$. Nonetheless, under the theoretical regime described in Section 1.5.1, our estimate of $\Delta\Gamma_s/\Gamma_s$ is in good agreement with the current world average value $\Delta\Gamma_s/\Gamma_s = 9.3^{+3.2}_{3.3}$ [40] that is calculated from a variety of lifetime measurements and existing $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$ measurements. Figure 8.2 gives a graphical comparison of results.

8.3.1 Comparison with Previous CDF Measurement

Compared with the previous CDF analysis [37] this study introduced a series of differing approaches. We are curious if the tension with the old CDF result (please compare "CDF (2008)" with "CDF Prelim. (2011)" in Figure 8.1(a)) is model-driven or a statistical effect only. To gain confidence in the methods used throughout this analysis the simultaneous fit is repeated using CDF period 0 data only. This is the dataset that has been available to the former CDF analysis[1]. Figure 8.3 shows

---

[1] There might be slight differences in the period 0 datasets due to changes in reconstruction software.



(a) (b)

Figure 8.3: Fitted mass projections of $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ (a) and $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+ K^{*0\dagger} K^-$ using period 0 CDF data only.

the fitted mass projections of the decay channels $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+\phi^\dagger\pi^-$ and $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger\pi^+ K^{*0\dagger} K^-$. From the simultaneous fit the ratio $f_{D_sD_s}$ is extracted to be

$$f_{D_sD_s} = 0.27 \pm 0.07(stat) \tag{8.10}$$

The fit result (8.10) is compared with the results extracted from Ref. [37] and the accompanying public documentation [38]. Excluding $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\pi^+\pi^-\pi^-$, equation (10) in [38] is used to calculate the weighted average of $f_{D_sD_s}$, where we insert yields and efficiencies as presented in [37], but use today's final state branching fractions. We calculate

$$f_{D_sD_s} = 0.26^{+0.08}_{-0.06}(stat) \tag{8.11}$$

This is in good agreement with (8.10). Neglecting the small correlation between the two datasets (the period 0 dataset is a subset of the full dataset used in this very analysis) the 2σ deviation with respect to the former CDF result can be explained by the two analyses being carried out on different data samples.

9

Concluding Remarks

In summary, this thesis presented a measurement of the ratios of the exclusive branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and the semi-inclusive relative branching fraction $f_{D_s^{(*)}D_s^{(*)}}$, and reports first observation of the exclusive mode $B_s^0 \to D_s^{*+}D_s^{*-}$. In the relative measurement the decay of $B_d^0 \to D^+D_s^-$ was chosen as normalization channel. The D_s^+ meson was reconstructed selecting two narrow mass bands in $K^+K^-\pi^+$ phase space, where for the first time the full Dalitz structure of $D_s^+ \to K^+K^-\pi^+$ decays was accounted for in an analysis of this kind. In doing so, reconstruction efficiencies estimated from simulated data are more reliable when compared to other analyses involving $D_s^+ \to K^+K^-\pi^+$ decays where K^+K^- and $K^-\pi^+$ resonances other than $\phi(1020)$ and $K^{*0}(892)$ are neglected. The relative branching fractions were determined in a joint likelihood fit to all signal and normalization channels by sharing certain parameters among the full likelihood function. Using the world average values of the quark production ratio f_s/f_d and $\mathcal{B}(B_d^0 \to D^+D_s^-)$ this analysis measures the absolute exclusive branching fractions as

$$\mathcal{B}(B_s^0 \to D_s^+D_s^-) = (0.49^{+0.06}_{-0.05}(stat) \pm 0.05(sys) \pm 0.08(norm))\%$$
$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^-) = (1.13 \pm 0.12(stat)^{+0.09}_{-0.10}(sys) \pm 0.19(norm))\%$$
$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^{*-}) = (1.75^{+0.20}_{-0.19}(stat)^{+0.17}_{-0.18}(sys) \pm 0.29(norm))\%$$

These values represent the world's most precise measurements of these observables.

Assuming the mostly CP-even final states $D_s^{(*)+}D_s^{(*)-}$ to saturate $\Delta\Gamma_s$, and taking CP violation to be negligible, the branching fraction of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ can be used to infer the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ in the B_s-$\bar{B}_s$ system. Using the absolute branching fraction value of semi-inclusive decays measured in this analysis,

$$\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-}) = (3.37 \pm 0.25(stat) \pm 0.30(sys) \pm 0.56(norm))\%$$

the relative decay width difference is estimated to be

$$\begin{aligned}\frac{\Delta\Gamma_s}{\Gamma_s} &= (6.97^{+0.50}_{-0.49}(stat)^{+0.59}_{-0.60}(sys) \pm 1.11(norm) \pm 0.34(theo))\% \\ &= (6.97^{+1.39}_{-1.40})\%\end{aligned}$$

The possibility of estimating the decay width difference in the $B_s^0 - \bar{B}_s^0$ system through a branching fraction measurement of the decays $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ critically depends on the assumptions contributed by the theory community. As with other scientific disciplines, theoretical heavy flavor physics is a continually evolving field. Over time, this gives rise to changing theoretical models and predictions. Deviating assumptions however may lead to a considerably different interpretations of experimental results. In the light of these thoughts and the considerations made in Section 1.5.2, this thesis closes with a brief discussion of concerns and possible solutions with regards to the interpretation of the presented estimate of $\Delta\Gamma_s/\Gamma_s$.

As pointed out in [22], the theoretical assumptions made in [31] do not account for contributions stemming from multi-body (more than two) final states and two-body final states other than $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$. If other contributions to $\Delta\Gamma_s$ are sizable, measuring $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ does not give any more a robust estimation of $\Delta\Gamma_s$. These concerns seem to be encouraged by recent theoretical calculations of $\Delta\Gamma_s$: in contrast to previous predictions, according to Ref. [33] the effect of three-body modes on $\Delta\Gamma_s$ is virtually comparable to that of two-body modes. Thus the assumption of the semi-inclusive two-body mode $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ saturating $\Delta\Gamma_s$ receives a considerable correction. With this in mind, $\Delta\Gamma_s$ estimated from $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ can be interpreted as a lower limit to the full decay width difference only. However, it must be stressed that the theoretical uncertainties of the newly calculated semi-inclusive branching fractions of three-body modes are still so sizable that their estimated contribution to $\Delta\Gamma_s$ can still be considered as compatible with zero. Therefore, further experimental efforts in this field are certainly required. To provide a more accurate estimation of $\Delta\Gamma_s$ it would be necessary to experimentally establish the full ensemble – both two-body and three-body – of *CP*-specific $B_s^0 - \bar{B}_s^0$ final states, or at least to identify the dominant modes other than $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$.

Of equivalent concern would be the existence of a sizeable *CP*-odd component in $B_s^0 \to D_s^{*+} D_s^{*-}$. Though theoretical calculations predict a small *CP*-odd fraction ($\sim$ 5%) [34], it could be easily as large as 24% given a scenario where the longitudinal polarization fraction in $B_s^0 \to D_s^{*+}D_s^{*-}$ is compatible with the one predicted and measured in $B_d^0 \to D^{*+}D_s^{*-}$ decays.

The only promising concept to overcome the latter concern is to exactly determine the CP-odd component in $B_s^0 \to D_s^{*+}D_s^{*-}$, either by angular studies or by a lifetime fit to $B_s^0 \to D_s^{*+}D_s^{*-}$. The first approach requires full reconstruction of $B_s^0 \to D_s^{*+}D_s^{*-}$, including detection of the tracks of the neutral pion and photon emitted in the decay of the excited D_s^{*+} meson in order to determine their helicity angles in the D_s^{*+} rest frame. Full reconstruction of $B_s^0 \to D_s^{*+}D_s^{*-}$ is however beyond the technical capabilities of the CDF II detector. In the course of in-depth systematic studies this thesis examined the feasibility of measuring a potentially non-zero CP-odd fraction f_{CP-} in $B_s^0 \to D_s^{*+}D_s^{*-}$ decays by varying helicity amplitude expectations in simulated data. We are however not able to confirm or rule out any CP-odd fractions lying within 0 and 24% by partial reconstruction of $B_s^0 \to D_s^{*+}D_s^{*-}$, though preparative studies relying upon decay amplitude calculations of the widely used EvtGen decay simulator package clearly pointed to this possibility. The EvtGen simulations could however be shown to give false results for the considered decay chain and certain input parameter configurations. Since the Belle detector allows for the detection of low-energetic neutral particles, and decays of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ already have been observed with 23.6 fb^{-1} of data, given sufficient statistics of $\Upsilon(5S)$ data the Belle collaboration might be able to make important contributions to angular analyses of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$.

Though beyond the scope of this very analysis, given the statistics available the second approach proposed might yet be in reach for the CDF collaboration: a lifetime fit to the decay time distribution of $B_s^0 \to D_s^{*+}D_s^{*-}$. It must however be pointed out that a study of this kind would be technically demanding, since biases to the proper decay time introduced by Two-Track Trigger requirements need to be corrected for. This argument holds true for both the CDF and the LHCb detector, that uses similar online trigger requirements to identify hadronic tracks with large replacements of secondary vertices. It should be possible to overcome this difficulty by measuring the $B_s^0 \to D_s^{*+}D_s^{*-}$ lifetime relative to a decay having similar decay kinematics and topology so that resolution effects and kinematical biases mostly cancel out. A promising candidate for this approach could be the same decay mode used as normalization channel in the presented branching fraction analysis, $B_d^0 \to D^+D_s^-$. In addition to the CDF experiment, both the Belle and the LHCb collaboration can certainly make valuable contributions in this field.

In the context of the measuring concepts introduced in Section 1.5, apart from shedding light on the CP puzzle in $B_s^0 \to D_s^{*+}D_s^{*-}$, a lifetime measurement is a very interesting option in other respects as well. Since the decay $B_s^0 \rightarrow D_s^+ D_s^-$ represents the *CP*-even counterpart of the pure *CP*-odd decay $B_s^0 \rightarrow J/\psi f_0$, by measuring the distinct lifetimes of these two *CP* eigenstates all required ingredients would be at hand to directly determine the decay width difference $\Delta\Gamma_s$. This approach is highly encouraged by the first successful determination of the $B_s^0 \rightarrow J/\psi f_0$ lifetime that was recently reported by the CDF collaboration.

The presented branching fraction analysis used 6.8 fb^{-1} of CDF Run II hadronic Two Track Trigger data. This represents the total amount of processed data available at the time of writing. Due to the poor scaling of the net amount of Two-Track Trigger data with the gross amount of integrated luminosity acquired, no significant reduction in statistical uncertainty is to be expected from future CDF measurements of $f_{D_s^{(*)} D_s^{(*)}}$ using hadronic D_s^+ decay channels, even when using the full available Tevatron Run II dataset of about 10 fb^{-1}. Nonetheless, as production of the last Tevatron operation periods is already in progress, the author is keen to use the full processed dataset as soon as available. A further reduction in statistical uncertainty may be achieved by extending the ensemble of D_s^+ decay modes used.

A

Appendix

A.1 Variable Definitions

With P being a non-stable parent particle (a meson decaying into stable final state particles or non-stable particles that decay further) and C_i[1] being a child particle (a decaying meson or a stable final state particle), throughout the analysis we refer to variables defined in the following way:

- $L_{xy}(P)$ is the transverse displacement (the displacement in the xy-plane) of the reconstructed P vertex with respect to the primary interaction point.
- $\sigma_{L_{xy}}(P)$ is the estimated uncertainty of $L_{xy}(P)$.
- $\chi^2_{R\phi}(P)$ is the χ^2 in the $R - \phi$ plane of the kinematic fit of the P candidate.
- $\Delta Z_0(P)$ is the difference in the Z coordinates of the starting points of the fitted daughter particle tracks prior to the vertex fit. This is to avoid accepting particle tracks that seem to be close in the transverse plane, but are far off in Z direction.
- prob(P) is the P candidate probability derived from $\chi^2_{R\phi}(P)$.
- $d_0(P)$ is the distance of closest approach (i.e. the impact parameter) of the P trajectory with respect to the beamline.
- $p_T(P)$ is the projection of the P momentum into the transverse plane.
- $L_{xy}(P \leftarrow C)$ is the transverse displacement (the displacement in the xy-plane) of the reconstructed C vertex with respect to the reconstructed P vertex.

---

[1] To avoid ambiguities, in some cases the parent particle P the child particle C is associated with is given as a subscript.

- $d_0^{lts}(P)$ is the lifetime-signed impact parameter of particle P.
- $\sigma_{d_0}(P)$ is the estimated uncertainty of $d_0^{lts}(P)$.
- $\min(p_T)$ is the minimum transverse momentum of the final state particles.
- $\min(d_0/\sigma_{d_0})$ is the minimum of the significance of the final state particles' impact parameter .
- $\mathrm{m}(C^iC^j)$ is the invariant mass of two particles C^i, C^j.
- $\Delta\mathrm{m}(P)$ is the difference of the reconstructed mass of particle P and the published world average P mass value
- PID.ratioC(C) is the probability of a child particle C to be a C candidate, divided by the probability for being a non-C candidate, derived from dE/dx and TOF information.

A.2 Dalitz Plot Parameter Toy Studies

The estimation of impacts of Dalitz model variations on the ratios of branching fractions (section 7.3.2) required knowledge of the full covariance matrix resulting from a fit to the $D_s^+ \to K^+K^-\pi^+$ Dalitz plot performed by the CLEO collaboration [105]. The authors kindly provided us the fit parameter covariance matrix, from which the decomposed lower left triangle matrix is calculated in order to generate correlated random parameter values. The parameters from left to right (or top to bottom) are: $m_{K^{*0}(892)}$, $\Gamma_{K^{*0}(892)}$, $a_{K_0^*(1430)}$, $\phi_{K_0^*(1430)}$, $a_{f_0(980)}$, $\phi_{f_0(980)}$, $a_{\phi(1020)}$, $\phi_{\phi(1020)}$, $a_{f_0(1370)}$, $\phi_{f_0(1370)}$,$a_{f_0(1710)}$, and $\phi_{f_0(1710)}$. The ordering of the parameters follows the indexing scheme given in Table 4.3.

A.2.1 Correlated Statistical Uncertainties

Covariance Matrix

```
     2.7e-07           0       1e-05     0.00084      -1e-05     0.00029          -0     0.00028      -1e-05    -0.00023       1e-05    -0.00039
           0   1.145e-06          -0    -0.00064      -6e-05      -9e-05      -2e-05    -0.00039      -1e-05    -0.00098      -2e-05       4e-05
       1e-05          -0     0.01222     0.20407    -0.00504     0.13979     0.00012     0.01612    -0.00238    -0.10867    -0.00144    -0.02375
     0.00084    -0.00064     0.20407      67.867    -0.48353      13.232     0.04142      18.927    -0.05537     -2.1973     0.04057     -7.5926
      -1e-05      -6e-05    -0.00504    -0.48353     0.03094     0.13005     0.00124     0.14102     0.00935     0.39917      0.0066     0.18934
     0.00029      -9e-05     0.13979      13.232     0.13005      10.225     0.01399      8.4925     0.01036       8.644     0.09497      3.1487
          -0      -2e-05     0.00012     0.04142     0.00124     0.01399     0.00043      0.0223     0.00026      0.0178     0.00053     0.01762
     0.00028    -0.00039     0.01612      18.927     0.14102      8.4925      0.0223      11.297     0.05329      6.2204     0.10004       1.318
      -1e-05      -1e-05    -0.00238    -0.05537     0.00935     0.01036     0.00026     0.05329      0.0073     0.02883     0.00038    -0.06801
     0.00023    -0.00098    -0.10867     -2.1973     0.39917       8.644      0.0178      6.2204     0.02883      26.473     0.22972      2.8634
       1e-05      -2e-05    -0.00144     0.04057      0.0066     0.09497     0.00053     0.10004     0.00038     0.22972     0.00544     0.04375
    -0.00039       4e-05    -0.02375     -7.5926     0.18934      3.1487     0.01762       1.318    -0.06801      2.8634     0.04375      22.887
```

Correlation Matrix

```
           1           0     0.17409     0.19623    -0.10941     0.17454          -0     0.16032    -0.22525   -0.086029     0.26093    -0.15689
           0           1          -0   -0.072602    -0.31878   -0.026303    -0.90135    -0.10844    -0.10938      -0.178    -0.25341   0.0078137
     0.17409          -0           1     0.22408     -0.2592     0.39547    0.052349    0.043385    -0.25199    -0.19106    -0.17661   -0.044909
     0.19623   -0.072602     0.22408           1    -0.33368     0.50232     0.24246     0.68354   -0.078665   -0.051838    0.066769    -0.19265
    -0.10941    -0.31878     -0.2592    -0.33368           1     0.23122     0.33996     0.23852     0.62214     0.44106     0.50873       0.225
     0.17454   -0.026303     0.39547     0.50232     0.23122           1     0.21099     0.79016     0.03792      0.5254     0.40268     0.20583
          -0    -0.90135    0.052349     0.24246     0.33996     0.21099           1     0.31995     0.14675     0.16683     0.34653     0.17761
     0.16032    -0.10844    0.043385     0.68354     0.23852     0.79016     0.31995           1     0.18556     0.35969     0.40354    0.081965
    -0.22525    -0.10938    -0.25199   -0.078665     0.62214     0.03792     0.14675     0.18556           1    0.065582    0.060301    -0.16638
    0.086029      -0.178    -0.19106   -0.051838     0.44106      0.5254     0.16683     0.35969    0.065582           1     0.60534     0.11633
     0.26093    -0.25341    -0.17661    0.066769     0.50873     0.40268     0.34653     0.40354    0.060301     0.60534           1     0.12399
    -0.15689   0.0078137   -0.044909    -0.19265       0.225     0.20583     0.17761    0.081965    -0.16638     0.11633     0.12399           1
```

Lower Left Triangle Covariance Matrix

```
  0.00051962           0           0           0           0           0           0           0           0           0           0           0
           0     0.00107           0           0           0           0           0           0           0           0           0           0
    0.019245          -0     0.10886           0           0           0           0           0           0           0           0           0
      1.6166     -0.5981      1.5889      7.8976           0           0           0           0           0           0           0           0
   -0.019245   -0.056072   -0.042897   -0.052902     0.15095           0           0           0           0           0           0           0
     0.55811   -0.084108      1.1855      1.3164      1.6997      1.9696           0           0           0           0           0           0
          -0   -0.018691   0.0011024   0.0036074   0.0028492  0.00077167   0.0075969           0           0           0           0           0
     0.53886    -0.36447    0.052819       2.248      1.6704      1.1679     0.21848      1.2713           0           0           0           0
   -0.019245  -0.0093454   -0.018461 -6.5282e-05    0.050746   -0.022322  -0.0028227   0.0025939    0.058424           0           0           0
    -0.44264    -0.91585    -0.92004   -0.071878      1.9611      3.3846    -0.82186    -0.56127    -0.51464      2.8337           0           0
    0.019245   -0.018691   -0.016631   0.0031281    0.035603    0.019163   0.0094087  -0.0056644   -0.018295    0.023469    0.042344           0
    -0.75056    0.037382   -0.085485    -0.78771     0.87215      1.6383      2.3043    -0.28481     -1.4409     -1.3526    -0.48223      2.9532
```

Toy Parameter Distributions

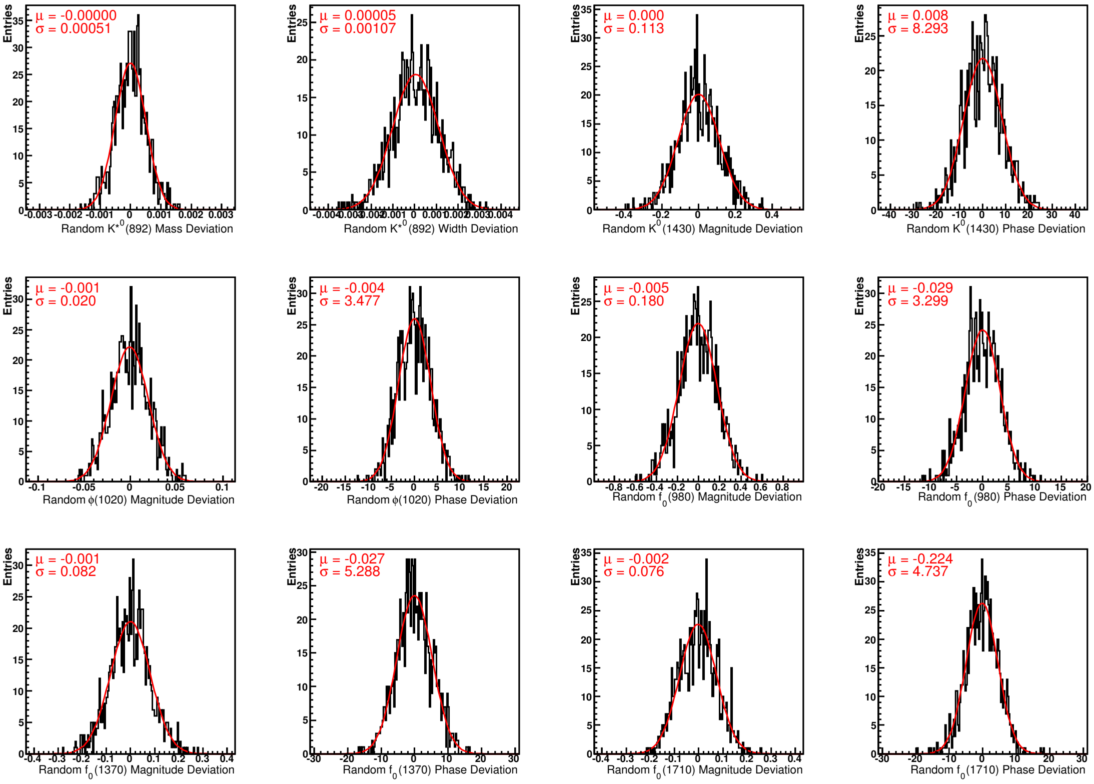


Figure A.1: Randomly generated Dalitz parameter deviations for the scenario of correlated statistical Dalitz plot parameter uncertainties. The toy parameter distributions keep their Gaussian shape. The small deviations of the toy widths with respect to the uncertainties given by Table 4.3 are due to rounding errors.

A.2.2 Systematic Uncertainties

Systematic uncertainties are not correlated to each other. Therefore the covariance matrix is a diagonal variance matrix.

Variance Matrix

```
   4.9e-07         0         0         0         0         0         0         0         0         0         0         0
         0   2.5e-07         0         0         0         0         0         0         0         0         0         0
         0         0    0.0081         0         0         0         0         0         0         0         0         0
         0         0         0        64         0         0         0         0         0         0         0         0
         0         0         0         0    0.0289         0         0         0         0         0         0         0
         0         0         0         0         0        16         0         0         0         0         0         0
         0         0         0         0         0         0    0.0004         0         0         0         0         0
         0         0         0         0         0         0         0        16         0         0         0         0
         0         0         0         0         0         0         0         0    0.0036         0         0         0
         0         0         0         0         0         0         0         0         0        36         0         0
         0         0         0         0         0         0         0         0         0         0      0.01         0
         0         0         0         0         0         0         0         0         0         0         0        25
```

Variance Triangle Matrix

```
    0.0007         0         0         0         0         0         0         0         0         0         0         0
         0    0.0005         0         0         0         0         0         0         0         0         0         0
         0         0      0.09         0         0         0         0         0         0         0         0         0
         0         0         0         8         0         0         0         0         0         0         0         0
         0         0         0         0      0.17         0         0         0         0         0         0         0
         0         0         0         0         0         4         0         0         0         0         0         0
         0         0         0         0         0         0      0.02         0         0         0         0         0
         0         0         0         0         0         0         0         4         0         0         0         0
         0         0         0         0         0         0         0         0      0.06         0         0         0
         0         0         0         0         0         0         0         0         0         6         0         0
         0         0         0         0         0         0         0         0         0         0       0.1         0
         0         0         0         0         0         0         0         0         0         0         0         5
```

Toy Parameter Distributions

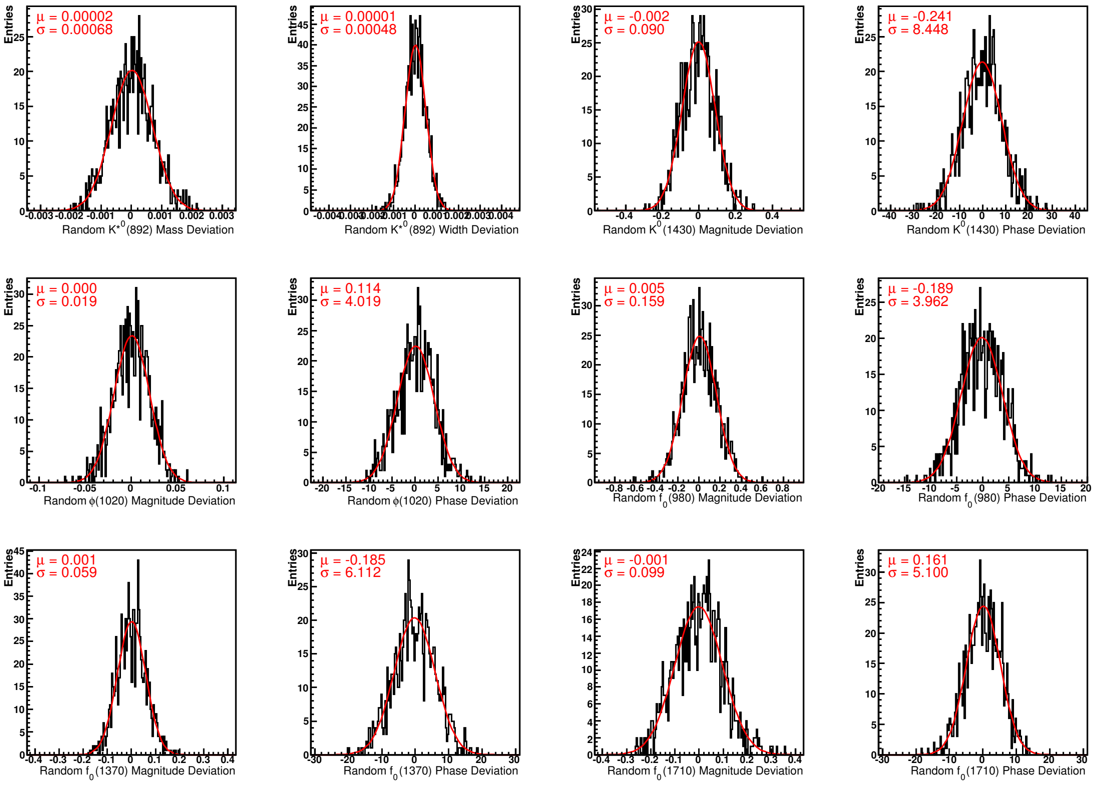


Figure A.2: Randomly generated Dalitz parameter deviations for the scenario of systematic Dalitz plot parameter uncertainties.

A.3 Spin-1 Wigner Rotation Functions

In this section the reduced Wigner d-functions $d^1_{m,m'}$ for the rotation of a particle with total angular momentum of $j = 1$ and a third projection component $m = 0, \pm$ into a final state with helicity $m' = 0, \pm$ are tabulated. The function elements not quoted in the literature [3] were derived using the identity $d^j_{m',m} = (-1)^{m-m'} d^j_{m,m'} = d^j_{-m,-m'}$.

$$\begin{array}{lll}
d^1_{-1,-1} = \dfrac{1+\cos\theta}{2} & d^1_{0,1} = \dfrac{\sin\theta}{\sqrt{2}} & d^1_{1,-1} = \dfrac{1-\cos\theta}{2} \\
d^1_{-1,0} = \dfrac{\sin\theta}{\sqrt{2}} & d^1_{0,0} = \cos\theta & d^1_{1,0} = -\dfrac{\sin\theta}{\sqrt{2}} \\
d^1_{-1,1} = \dfrac{1-\cos\theta}{2} & d^1_{0,-1} = -\dfrac{\sin\theta}{\sqrt{2}} & d^1_{1,1} = \dfrac{1+\cos\theta}{2}
\end{array}$$

Decay Model	Meaning	Parameters
PHSP	Phase Space	
SVS	Pseudo-Scalar to Vector and Pseudo-Scalar	
VSS	Vector to Pseudo-Scalar and Pseudo-Scalar	
VSP_PWAVE	P-Wave Decay of Vector to Pseudo-Scalar and Photon	
SVV_HELAMP	Pseudo-Scalar to Vector and Vector using Helicity Model	H_+, $\phi(H_+)$, H_0, $\phi(H_0)$, H_-, $\phi(H_-)$
D_DALITZ	Three-body Decay using Dalitz Model	

Table A.1: Decay models used in the simulation of decay chains.

A.4 Monte Carlo Decay Tables

The following tables were used as input to the EvtGen program to simulate different decay chains, starting from a B_s^0 or a B_d^0 meson. The decay tables also include decay chains with the three-pion final state. These were not subject to this very study but might be useful for later studies incorporating $D_s^+ \to \pi^-\pi^+\pi^+$ to extent the ensemble of hadronic $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays.

At each stage of the decay chain, the decay tables specify the mother particle (preceded by "Decay"), the daughter particles preceded by the decay probability requested for this final state, and the decay model used for this particular decay. Table A.1 provides an overview of the decay models relevant for Monte Carlo generation of the decay channels used in this analysis.

A.4.1 $B_d^0 \to D^{(*)+}D_s^{(*)-}$

```
#
# Delivered by simulation: 1802726904
#
#                 mass/Gev      ctau/mm
# B0              5.27953       0.4587
# anti-B-0        5.27953       0.4587
#
#-----------------------------------------------------------------------------
#
Decay B0
0.3   D-    D_s+              PHSP;
0.2   D*-   D_s+              SVS;
```

```
0.2   D_s*+ D-                SVS;
0.2   D_s*+ D*-               SVV_HELAMP 0.4904 0.0 0.7204 0.0 0.4904 0.0;
Enddecay
#
Decay anti-B0
0.3   D+    D_s-              PHSP;
0.2   D*+   D_s-              SVS;
0.2   D_s*- D+                SVS;
0.2   D_s*- D*+               SVV_HELAMP 0.4904 0.0 0.7204 0.0 0.4904 0.0;
Enddecay
#
#
#-----------------------------------------------------------------------------
# D*
#-----------------------------------------------------------------------------
Decay D*+
0.3060    D+  pi0                        VSS;
0.0110    D+  gamma                       VSP_PWAVE;
Enddecay
#
Decay D*-
0.3060    D-  pi0                        VSS;
0.0110    D-  gamma                       VSP_PWAVE;
Enddecay
#-----------------------------------------------------------------------------
# Ds*
#-----------------------------------------------------------------------------
Decay D_s*+
0.942   D_s+  gamma                       VSP_PWAVE;
0.058   D_s+  pi0                        VSS;
Enddecay
#
Decay D_s*-
0.942   D_s-  gamma                       VSP_PWAVE;
0.058   D_s-  pi0                        VSS;
Enddecay
#-----------------------------------------------------------------------------
# D
#-----------------------------------------------------------------------------
Decay D+
0.920   K-  pi+  pi+                    D_DALITZ;
Enddecay
#
Decay D-
0.920   K+  pi-  pi-                    D_DALITZ;
Enddecay
```

```
#------------------------------------------------------------------------------
# Ds
#------------------------------------------------------------------------------
Decay D_s-
0.0550  K-  K+  pi-                     D_DALITZ;
0.0004    rho0  pi-                     SVS;
0.0057    f_0   pi-                     PHSP;
0.0020    f_2   pi-                     PHSP;
0.0033    f'_0  pi-                     PHSP;
Enddecay
#
Decay D_s+
0.0550  K+  K-  pi+                     D_DALITZ;
0.0004    rho0  pi+                     SVS;
0.0057    f_0   pi+                     PHSP;
0.0020    f_2   pi+                     PHSP;
0.0033    f'_0  pi+                     PHSP;
Enddecay
#
#------------------------------------------------------------------------------
# Final decay products
#------------------------------------------------------------------------------
#
Decay f'_0
0.5200   pi+  pi-                       PHSP;
Enddecay
#
Decay f_0
0.5200   pi+  pi-                       PHSP;
Enddecay
#
Decay f_2
0.5650   pi+  pi-                       TSS;
Enddecay
#
Decay rho0
1.000    pi+ pi-                        VSS;
Enddecay
#
#------------------------------------------------------------------------------
#
End
```

A.4.2 $B_s^0 \to D_s^+ D_s^-$

```
#
# Delivered by simulation: 1773006604
#
#               mass/Gev      ctau/mm
# B_s0          5.3663        0.441
# anti-B_s0     5.3663        0.441
#
#----------------------------------------------------------------------------
#
Decay B_s0
0.0026      D_s-        D_s+          PHSP;

Enddecay
#
Decay anti-B_s0
0.0026      D_s+        D_s-          PHSP;
Enddecay
#
#----------------------------------------------------------------------------
Decay D_s-
0.0550 K-  K+ pi-D_DALITZ;
0.0004    rho0  pi-                     SVS;
0.0057    f_0   pi-                     PHSP;
0.0020    f_2   pi-                     PHSP;
0.0033    f'_0  pi-                     PHSP;
Enddecay
#
Decay D_s+
0.0550 K+  K-pi+ D_DALITZ;
0.0004    rho0  pi+                     SVS;
0.0057    f_0   pi+                     PHSP;
0.0020    f_2   pi+                     PHSP;
0.0033    f'_0  pi+                     PHSP;
Enddecay
#
#----------------------------------------------------------------------------
#
Decay f'_0
0.5200   pi+  pi-                       PHSP;
Enddecay
#
Decay f_0
0.5200   pi+  pi-                       PHSP;
Enddecay
```

```
#
Decay f_2
0.5650   pi+  pi-                       TSS;
Enddecay
#
Decay rho0
1.000    pi+ pi-                        VSS;
Enddecay
#
#------------------------------------------------------------------------------
#
End
```

A.4.3 $B_s^0 \to D_s^{*+} D_s^-$

```
#
# Delivered by simulation: 1795428397
#
#               mass/Gev      ctau/mm
#B_s0           5.3663        0.441
#anti-B_s0      5.3663        0.441
#
#------------------------------------------------------------------------------
#
Decay B_s0
0.0026      D_s-        D_s+          PHSP;
0.0090      D_s*+       D_s-          SVS;
0.0090      D_s*-       D_s+          SVS;
0.0197      D_s*-       D_s*+         SVV_HELAMP  0.4904 0.0 0.7204 0.0 0.4904 0.0;
Enddecay
#
Decay anti-B_s0
0.0026      D_s+        D_s-          PHSP;
0.0090      D_s*-       D_s+          SVS;
0.0090      D_s*+       D_s-          SVS;
0.0197      D_s*+       D_s*-         SVV_HELAMP  0.4904 0.0 0.7204 0.0 0.4904 0.0;
Enddecay
#
#------------------------------------------------------------------------------
# Ds*
#- ----------------------------------------------------------------------------
#Decay D_s*+
```

```
0.942   D_s+  gamma                         VSP_PWAVE;
0.058   D_s+  pi0                          VSS;
#Enddecay
#
#Decay D_s*-
0.942   D_s-  gamma                         VSP_PWAVE;
0.058   D_s-  pi0                          VSS;
#Enddecay
#-----------------------------------------------------------------------------
Decay D_s-
0.0550 K-  K+ pi-D_DALITZ;
0.0004    rho0  pi-                       SVS;
0.0057    f_0   pi-                       PHSP;
0.0020    f_2   pi-                       PHSP;
0.0033    f'_0  pi-                       PHSP;
Enddecay
#
Decay D_s+
0.0550 K+  K-pi+ D_DALITZ;
0.0004    rho0  pi+                       SVS;
0.0057    f_0   pi+                       PHSP;
0.0020    f_2   pi+                       PHSP;
0.0033    f'_0  pi+                       PHSP;
Enddecay
#
#-----------------------------------------------------------------------------
#
Decay f'_0
0.5200   pi+  pi-                         PHSP;
Enddecay
#
Decay f_0
0.5200   pi+  pi-                         PHSP;
Enddecay
#
Decay f_2
0.5650   pi+  pi-                         TSS;
Enddecay
#
Decay rho0
1.000    pi+ pi-                          VSS;
Enddecay
#
#-----------------------------------------------------------------------------
#
End
```

A.4.4 $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ Phase Space

```
#
# Delivered by simulation: 884452701
#
#                mass/Gev      ctau/mm
# B_s0           5.3663        0.441
# anti-B_s0      5.3663        0.441
#
#-----------------------------------------------------------------------------
#
Decay B_s0
0.0197      D_s*-       D_s*+         PHSP;
Enddecay
#
Decay anti-B_s0
0.0197      D_s*+       D_s*-         PHSP;
Enddecay
#
#-----------------------------------------------------------------------------
# Ds*
#- ---------------------------------------------------------------------------
Decay D_s*+
0.942   D_s+  gamma                       PHSP;
0.058   D_s+  pi0                        PHSP;
Enddecay
#
Decay D_s*-
0.942   D_s-  gamma                       PHSP;
0.058   D_s-  pi0                        PHSP;
Enddecay
#-----------------------------------------------------------------------------
Decay D_s-
0.0550 K-  K+ pi-D_DALITZ;
0.0004    rho0  pi-                     SVS;
0.0057    f_0   pi-                     PHSP;
0.0020    f_2   pi-                     PHSP;
0.0033    f'_0  pi-                     PHSP;
Enddecay
#
Decay D_s+
```

```
0.0550 K+  K-pi+ D_DALITZ;
0.0004    rho0  pi+                     SVS;
0.0057    f_0   pi+                     PHSP;
0.0020    f_2   pi+                     PHSP;
0.0033    f'_0  pi+                     PHSP;
Enddecay
#
#-----------------------------------------------------------------------------
#
Decay f'_0
0.5200   pi+  pi-                       PHSP;
Enddecay
#
Decay f_0
0.5200   pi+  pi-                       PHSP;
Enddecay
#
Decay f_2
0.5650   pi+  pi-                       TSS;
Enddecay
#
Decay rho0
1.000    pi+ pi-                        VSS;
Enddecay
#
#-----------------------------------------------------------------------------
#
End
```

Bibliography

[1] M. K. Gaillard, P. D. Grannis, F. J. Sciulli. The Standard Model of Particle Physics. *Rev. Mod. Phys.*, 71:96–11, 1999.

[2] S. K. Choi et al. Observation of a Resonance-Like Structure in the $\pi^{\pm}\psi'$ Mass Distribution in Exclusive $B \to K\pi^{\pm}\psi'$ Decays. *Phys. Rev. Lett.*, 100:142001, 2008.

[3] K. Nakamura et al. Review of Particle Physics. *J. Phys. G*, 37:075021, 2010 and 2011 partial update for the 2012 edition.

[4] A. Abulencia et al. Observation of $B_s\bar{B}_s$ Oscillations. *Phys. Rev. Lett.*, 97:242003, 2006.

[5] E. Noether. Invarianten beliebiger Differentialausdrücke. *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen*, pages 37–44, 1918.

[6] E. Noether. Invariante Variationsprobleme. *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen*, pages 235–257, 1918.

[7] G. Lüders. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Dan. Mat. Fys. Medd.*, 28:5, 1954.

[8] J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay. Evidence for the 2π Decay of the K_2 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[9] A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Journal of Experimental and Theoretical Physics*, 5:24–27, 1967.

[10] D. Kirkby et al. *CP* Violation in Meson Decays. *J. Phys. G*, 37:075021, 2010.

[11] M. Kobayashi, T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[12] S. L. Glashow. Partial-Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[13] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[14] A. Salam. *Elementary Particle Physics: Relativistic Groups and Analyticity*. 1968.

[15] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531, 1963.

[16] L. L. Chau, W. Y. Keung. Comments on the Parameterization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 53:1802–1805, 1984.

[17] L. Wolfenstein. Parameterization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, 1983.

[18] C. Jarlskog. Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation. *Phys. Rev. Lett.*, 55:1039–1024, 1985.

[19] K. Anikeev et al. B physics at the Tevatron: Run II and Beyond. 2002.

[20] U. Nierste. CP asymmetry in flavour-specific B decays. 2006.

[21] N. Nir. CP Violation in Meson Decays. 2005.

[22] I. Dunietz, R. Fleischer, U. Nierste. In Pursuit of New Physics with B_s Decays. *Phys. Rev. D*, 63:114015, 2001.

[23] R. Aaij et al. LHCb Conference Note LHCB-CONF-2011-00, 2011.

[24] A. Lenz, U. Nierste. Numerical Updates of Lifetimes and Mixing Parameters of B Mesons. *arXiv:1102.4274v1*, 2011.

[25] A. Lenz, U. Nierste. Theoretical Update of $B_s - \bar{B}_s$ mixing. *arXiv:hep-ph/0612167*, 2007.

[26] 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.

[27] V. M. Abazov et al. Measurement of B_s^0 Mixing Parameters from the Flavor-Tagged Decay $B_s^0 \to J/\psi\phi$. *Phys. Rev. Lett.*, 101:241801, 2008.

[28] R. Aaij et al. Flavor-untagged Analysis of $B_d^d \to J/\psi K^*$ and $B_s^0 \to J/\psi\phi$ Decays. LHCb-CONF-2011-002, 2011.

[29] T. Aaltonen et al. Measurement of Branching Ratio and B_s^0 Lifetime in the Decay $B_s^0 \to J/\psi f_0(980)$ at CDF. *arXiv:1106.3682*, 2011.

[30] R. Aaij et al. First observation of $B_s^0 \to J/\psi f_0(980)$ decays. *Phys. Lett. B*, 698:115, 2011.

[31] R. Aleksan, A. Le Yaouanc, L. Oliver. O. Pène, J.-C. Raynal. Estimation of $\Delta\Gamma$ for the $B_s\bar{B}_s$ system. Exclusive decays and the parton model. *Phys. Lett. B*, 316:567–577, 1993.

[32] M. A. Shifman, M. B. Voloshin. *Sov. J. Nucl. Phys.*, 47:511, 1988.

[33] Chun-Khiang Chua, Wei-Shu Hou, Chia-Hsien Shen. Long Distance Contribution to $\Delta\Gamma$ of the $B_s\bar{B}_s$ system. *arXiv:1107.4305v1*, 2011.

[34] J. L. Rosner. Determination of Pseudoscalar-Charmed-Meson Decay Constants from B-Meson Decays. *Phys. Rev. D*, 42:3732, 1990.

[35] R. Barate et al. *Phys. Lett. B*, 486:286, 2000.

[36] V. M. Abazov et al. Evidence for the decay $B_s^0 \to {D_s}^{(*)}{D_s}^{(*)}$ and a measurement of $\Delta{\Gamma_s}^{CP}/\Gamma_s$. *arXiv:0911.2173v2 [hep-ex]*, 2009.

[37] T. Aaltonen et al. First Observation of the Decay $B_s^0 \to {D_s}^-{D_s}^+$ and a Measurement of Its Branching Ratio. *Phys. Rev. Lett.*, 100:021803, 2008.

[38] T. Aaltonen et al. Observation of $B_s^0 \to {D_s}^-{D_s}^+$, $D_s \to \phi\pi(3\pi, K^*K)$, $D^+ \to K\pi\pi$ with $355pb^{-1}$ in Run II. CDF Note 7925, Preliminary Results for the Winter 2006 Conferences.

[39] S. Esen et al. Observation of $B_s^0 \to {D_s}^{(*)+}{D_s}^{(*)-}$ using e^+e^- collisions and a determination of the $B_s\bar{B}_s$ width difference $\Delta\Gamma_s$. *arXiv:1005.5177v1 [hep-ex]*, 2010.

[40] D. Asner et al. b Hadron Lifetime averages – Results for the PDG 2011 Web Update.

[41] S. H. Neddermeyer, C. D. Anderson. Note on the Nature of Cosmic-Ray Particles. *Phys. Rev.*, 51:884–886, 1937.

[42] D. H. Perkins. *Nature*, 159:126, 1947.

[43] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, C. F. Powell. Processes Involving Charged Mesons. *Nature*, 159:694–697, 1947.

[44] H. V. Klapdor-Kleingrothaus, K. Zuber. *Teilchenastrophysik*. Teubner, 1997.

[45] `http://www.fnal.gov/pub/science/frontiers/`.

[46] F. Abe et al. Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[47] S. Abache et al. Observation of the Top Quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

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

[49] *Run II Handbook*. `http://www-bdnew.fnal.gov/pbar/run2b/Documents/RunII_handbook.pdf`.

[50] Fermilab's Chain of Accelerators. `http://www-bd.fnal.gov/public/chain.html`.

[51] Cockroft Walton Voltage Multipliers. `http://home.earthlink.net/~jimlux/hv/cw1.htm`.

[52] `http://www-bd.fnal.gov/public/proton.html#Linac`.

[53] Fermilab Accelerator Devision. *Linac Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf`.

[54] `http://www-bd.fnal.gov/public/proton.html#Booster`.

[55] Fermilab Accelerator Devision. *Booster Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V4.1.pdf`.

[56] `http://www-bd.fnal.gov/public/maininj.html`.

[57] Fermilab Accelerator Devision. *Main Injector Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.1.pdf`.

[58] Fermilab Accelerator Devision. *Antiproton Source Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_v2.2.pdf`.

[59] `http://www-bd.fnal.gov/public/antiproton.html#target`.

[60] `http://www-bd.fnal.gov/public/antiproton.html#debuncher`.

[61] D. Mohl, G. Petrucci, L. Thorndahl, S. van der Meer. Physics and Technique of Stochastic Cooling. *Phys. Rept.*, 58:73, 1980.

[62] S. van der Meer. Stochastic Cooling and the Accumulation of Anti-Protons. *Rev. Mod. Phys.*, 57:689, 1985.

[63] `http://www-bd.fnal.gov/public/antiproton.html#accumulator`.

[64] Fermilab Accelerator Devision. *Recycler Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_RB_v1.42.pdf`.

[65] S. Nagaitsev et al. Experimental Demonstration of Relativistic Electron Cooling. *Phys. Rev. Lett.*, 96:044801, 2006.

[66] S. Nagaitsev et al. Status of the Fermilab Electron Cooling Project. *Nucl. Instrum. Meth. A*, 532:275, 2004.

[67] Fermilab Accelerator Devision. *Tevatron Rookie Book*. `http://www-bdnew.fnal.gov/operations/rookie_books/Tevatron_v2.3.pdf`.

[68] B. Povh, K. Rith, C. Scholz, F. Zetsche. *Teilchen und Kerne*. Springer, 2004.

[69] `http://www-cdfonline.fnal.gov/ops/opshelp/stores/`.

[70] D. F. Sutter et al. *Department of Energy Review Committee Report on the Run II Luminosity Performance of the Fermilab Tevatron*, 2002. `http://www-ad.fnal.gov/doereview02/Run_II_lum_review_Final.pdf`.

[71] P. H. Garbincius. Tevatron Collider Operations and Plans. *arXiv:hep-ex/0406013*.

[72] R. Blair et al. *The CDF II Detector – Technical Design Report*, 1996. FERMILAB-Pub-96/390-E.

[73] CDF Collaboration. A Brief Description the CDF Detector in Run II. `http://www-cdf.fnal.gov/internal/detectors/cdf_det.ps`, 2004.

[74] A. Sill et al. CDF Run II Silicon Tracking Projects. *Nucl. Instrum. Meth. A*, 447:1–8, 2000.

[75] J. Lueck. *Observation of Electroweak Single Top-Quark Production with the CDF II Experiment*. PhD thesis, Fakultät für Physik, Karlsruher Institut für Technologie (KIT), 2009.

[76] D. Acosta et al. The CDF Cherenkov Luminosity Monitor. *Nucl. Instrum. Meth. A*, 461:540–544, 2001.

[77] C. S. Hill et al. Initial Experience with the CDF Layer 00 Silicon Detector. *Nucl. Instrum. Meth. A*, 511:118–120, 2003.

[78] A. Affolder et al. Intermediate Silicon Layers Detector for the CDF Experiment. *Nucl. Instrum. Meth. A*, 453:84–88, 2000.

[79] http://www-cdf.fnal.gov/upgrades/tdr/doc/tracker_layout.eps.

[80] A. Affolder et al. COT Central Outer Tracker. *Nucl. Instrum. Meth. A*, 526:249, 2004.

[81] H. Bethe. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Ann. d. Phys.*, 397:325–400, 1930.

[82] S. Bailey, P. Maksimovic. Prospects for Measuring γ with $B_s \to D_s^{\mp} K^{\pm}$. CDF Note 4863, 2000.

[83] D. Acosta et al. A Time-of-Flight Detector in CDF-II. *Nucl. Instrum. Meth. A*, 518:605–608, 2004.

[84] S. Kuhlmann et al. The CDF Calorimetry Upgrade for Run IIb. *Nucl. Instrum. Meth. A*, 518:39–41, 2004.

[85] L. Balka et a. The CDF Central Electromagnetic Calorimeter. *Nucl. Instrum. Meth. A*, 267:272, 1988.

[86] S. Bertolucci et a. The CDF Central and Endwall Hadron Calorimeter. *Nucl. Instrum. Meth. A*, 267:301, 1988.

[87] C. M. Ginsburg. CDF Run 2 Muon System. *Eur. Phys. J. C*, 33:s1002–s1004, 2004.

[88] A. Artikov et al. Design and Construction of New Central and Forward Muon Counters for CDF-II. *Nucl. Instrum. Meth. A*, 538:358, 2005.

[89] G. Ascoli et al. CDF Central Muon Detector. *Nucl. Instrum. Meth. A*, 268:33, 1988.

[90] T. Dorigo et al. The Muon System Upgrade for the CDF II Experiment. *Nucl. Instrum. Meth. A*, 461:560–562, 2001.

[91] D. Amidei et al. A Two Level Fastbus Based Trigger System for CDF. *Nucl. Instrum. Meth. A*, 269:51–62, 1988.

[92] H. Frisch. *Conceptual Design of a Deadtimeless Trigger for the CDF Trigger Upgrade*, 1994. CDF Note 2038.

[93] H. Ray. *Level 2 and the L1/L2 Trigger System for Non-Experts*, 2002. CDF Note 5860.

[94] http://www-cdf.fnal.gov/upgrades/tdr/doc/electronics/cdf_dataflow.eps.

[95] E. J. Thomson. Online Track Processor for the CDF Upgrade. *Trans. Nucl. Sci.*, 49:1063–1070, 2002.

[96] A. Bardi et al. SVT: An Online Silicon Vertex Tracker for the CDF Upgrade. *Nucl. Instrum. Meth. A*, 409:658–661, 1991.

[97] B. Ashmanskas et al. The CDF Silicon Vertex Trigger. *Nucl. Instrum. Meth. A*, 518:532–536, 2004.

[98] *CDF SAM User Documentation*. http://www-cdf.fnal.gov/tiki/tiki-index.php?page=CdfSamUserDocumentation.

[99] R. Brun, F. Rademakers. ROOT - An Object Oriented Data Analysis Framework. *Nucl. Instrum. Meth. A*, 389:81–86, 1997.

[100] C. Amsler et al. Review of Particle Physics. *Physics Letters B*, 667:1, 2008.

[101] R. H. Dalitz. *Philos. Mag.*, 44:1068, 1953.

[102] J. M. Blatt, V. F. Weisskopf. *Theoretical Nuclear Physics*. John Wiley & Sons, New York, 1952.

[103] H. M. Pilkuhn. *The Interactions of Hadrons*. North-Holland Pub., Amsterdam, 1967.

[104] S. M. Flatté. Coupled-Channel Analysis of the pi eta and K anti-K Systems Near K anti-K Threshold. *Phys. Lett. B*, 63:224, 1976.

[105] R. E. Mitchell et al. Dalitz Plot Analysis of $D^+ \rightarrow K^+K^-\pi^+$. *arXiv:0903.1301v1 [hep-ex]*, 2009.

[106] G. Bonvicini. Dalitz Plot Analysis of the $D^+ \rightarrow K^-\pi^+\pi^+$ Decay. *Phys. Rev. D*, 78:052001, 2008.

[107] S. Kopp et al. Dalitz Plot Analysis of the Decay $D^0 \rightarrow K^-\pi^+\pi^0$. *Phys. Rev. D*, 63:092001, 2001.

[108] M. S. Dubrovin, D. Cinabro. Private Communication.

[109] V. Blobel, E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, chapter 6, page 175. B. G. Teubner, 1998.

[110] M. Jacob, G. C. Wick. On the General Theory of Collisions for Particles with Spin. *Ann. Phys.*, 7:404–428, 1959.

[111] S. U. Chung. Spin Formalisms. CERN Yellow Reports, CERN-71-08, 1971.

[112] J. D. Richman. An Experimenter's Guide to the Helicity Formalism. CALT-68-1148, DOE Research and Development Report, 1986.

[113] R. Kutschke. An Angular Distribution Cookbook. CLEO Internal Note, `http://home.fnal.gov/~kutschke/Angdist/angdist.ps`, 1986.

[114] S. M. Berman, M. Jacob. Systematics of Angular and Polarization Distributions in Three-Body Decays. *Phys. Rev. B*, B:1023, 1965.

[115] I. Dunietz, H. Quinn, A. Snyder, W. Toki, H. J. Lipkin. How to Extract *CP*-Violating Asymmetries from Angular Correlations. *Phys. Rev. D*, 43:2193, 1990.

[116] W. Bell, J. P. Fernandez, L. Flores, F. Wuerthwein, R. J. Tesarek. User Guide for EvtGen at CDF. CDF Note 5618.

[117] D. J. Lange. The EvtGen Particle Decay Simulation Package. *Nuclear Instruments and Methods in Physics Research Section A*, 462:152–155, 2001.

[118] A. Ryd et al. *EvtGen – A Monte Carlo Generator for B-Physics*, 2004. `http://robbep.home.cern.ch/robbep/EvtGen/GuideEvtGen.pdf`.

[119] B. Aubert et al. Measurement of the $B^0 \to D_s^{(*)+} D^{*-}$ Branching Fractions and $B^0 \to D_s^{*+} D^{*-}$ Polarization with a Partial Reconstruction Technique. *Phys. Rev. D*, 67:092003, 2003.

[120] K. Anikeev, C. Paus, P. Murat. Description of Bgenerator II. CDF Note 5092, 1999.

[121] T. Sjöstrand, L. Lönnblad, S. Mra. *PYTHIA 6.2: Physics and Manual*, 2001.

[122] C. Peterson et al. *Phys. Rev. D*, 27:105, 1983.

[123] R. Brun, R. Hagelberg, M. Hansroul, J. Lassalle. CERN-DD-78-2-REV, 1978.

[124] The CDF Detector Simulation. `http://www-cdf.fnal.gov/cdfsim/cdfsim_main.html`.

[125] The trigger simulation project. `http://ncdf70.fnal.gov:8001/trigsim/trgsim.html`.

[126] M. Feindt. A Neural Bayesian Estimator for Conditional Probability Densities. *arXiv:physics/0402093*, 2004.

[127] M. Feindt. *The NeuroBayes User's Guide*. `NEUROBAYES/doc/NeuroBayes-HowTo`.

[128] V. Blobel, E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, chapter 6, page 183. B. G. Teubner, 1998.

[129] R. Barlow. Extended Maximum Likelihood. *Nucl. Instrum. Meth. A*, 297:496–506, 1990.

[130] A. Ryd. Private Communication.

[131] T. Aaltonen et al. Observation of $B_s^0 \rightarrow J/\psi K^*(892)^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ Decays. *Phys. Rev. D*, 83:052012, 2011.

[132] S. Esen for the Belle Collaboration. Observation of $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ and Estimate of $\Delta\Gamma_s$ at Belle. *Presentation at the Meeting of the Division of Particles and Fields of the American Physical Society, August 9-13, 2011*.

Danksagung

An erster Stelle möchte ich mich bei Herrn Professor Dr. Michael Feindt dafür bedanken, dass ich als Doktorand am Institut für Experimentelle Kernphysik angenommen wurde und unter seiner Betreuung diese Dissertation anfertigen durfte. Ganz herzlich danken will auch Herrn Professor Dr. Ulrich Husemann für die Übernahme des Korreferats.

Nicht wegzudenken für das Gelingen dieser Arbeit sind Dr. Thomas Kuhr, Dr. Michal Kreps, Dr. Anze Zupanc und Dr. Martin Heck, die mich als hervorragende Betreuer und kritische Diskussionspartner durch die Promotion begleitet haben. Meinen herzlichen Dank möchte ich zudem Daniel Zander, Dr. Judith Horn und abermals Dr. Thomas Kuhr und Dr. Martin Heck für das Korrekturlesen meiner Arbeit aussprechen. Virginia Jani danke ich für die Anfertigung zahlreicher Graphiken.

Für die großartige gemeinsame Zeit in- und außerhalb des Instituts sei allen hiesigen Mitstreitern gedankt, allen voran den Mitgliedern der B-Physik-Gruppe und meinen Zimmerkollegen Dr. Andreas Gessler, Daniel Zander und Manuel Heider. Dank und Respekt gebührt zudem dem Administratorenteam für die Betreuung der Rechnerinfrastruktur sowie Bärbel Bräunling für die stets freundliche und geduldige Unterstützung im Kampf gegen die Mühlen der Bürokratie.

Diese Arbeit wurde mit Mitteln der Deutschen Forschungsgemeinschaft im Rahmen eines Stipendiums des Graduiertenkollegs für Hochenergie- und Teilchenastrophysik unterstützt. In diesem Zusammenhang gilt mein Dank der Auswahlkommission des Graduiertenkollegs für die Aufnahme als Stipendiat und Isabelle Junge für die Unterstützung in organisatorischen Fragen. Außerdem bedanken will ich mich auch bei Herrn Professor Dr. Thomas Müller, dass mir durch seine Funktion als Institutsleiter die Möglichkeit diverser Arbeitsaufenthalte und Konferenzbesuche geboten wurde.

Gar nicht genug Dank auszudrücken vermag ich meinen Eltern Peter und Elisabeth Horn sowie meinen Geschwistern Judith, Pirmin und Simon für ihre aufopferungsvolle, vorbehaltlose und immer humorvolle Unterstützung. Zutiefst dankbar bin ich Dir, Virginia, die Du mich als verlässliche Partnerin durch diese herausfordernde Zeit stets mit Geduld, kraftspendendem Zuspruch und unentwegtem tatkräftigem Beistand begleitet hast.