CDF/ANAL/BOTTOM/CDFR/10619 Version 1.3
February 7, 2012

Measurement of the Branching Fractions of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ Decays and the Relative Decay Width Difference $\Delta\Gamma_s/\Gamma_s$ Using 6.8 fb^{-1} of CDF Run II Data

Michael Feindt, Dominik Horn[1], Thomas Kuhr[2] *Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie (KIT)*

Michal Kreps[3] *Department of Physics, University of Warwick*

Abstract

Based on an exclusive 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 $B_s^0 \to D_s^{*+} D_s^{*-}$ we perform a 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 $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ using 6.8 fb^{-1} of data collected by the CDF II detector at the $p\bar{p}$ collider Tevatron. The absolute branching fractions are determined by measuring the ratios of branching fractions with respect to $\mathcal{B}(B_d^0 \to D^+ D_s^-)$, multiplied by the ratio of the quark fragmentation fractions f_s/f_d. The D^- meson is reconstructed via the hadronic decay $D^- \to K^+\pi^-\pi^-$ (charge conjugate states are implied throughout), the D_s^+ meson by selecting two narrow mass bands of phase space of the $K^+K^-\pi^+$ final state and accounting for the full Dalitz structure of $D_s^+ \to K^+K^-\pi^+$ in the estimation of reconstruction efficiencies. Using an simultaneous unbinned maximum likelihood fit to all studied decay channels we report first observation of the exclusive mode $B_s^0 \to D_s^{*+} D_s^{*-}$, and measure the ratio of each of the exclusive and the semi-inclusive branching fraction as:

$$
\begin{aligned}
f_{D_s D_s} &= 0.183^{+0.021}_{-0.020}(stat) \pm 0.017(sys) \\
f_{D_s^* D_s} &= 0.424^{+0.047}_{-0.045}(stat) \pm 0.035(sys) \\
f_{D_s^* D_s^*} &= 0.654^{+0.073}_{-0.071}(stat) \pm 0.065(sys) \\
f_{D_s^{(*)} D_s^{(*)}} &= 1.261^{+0.096}_{-0.093}(stat)^{+0.111}_{-0.112}(sys)
\end{aligned}
$$

[1]dhorn@ekp.uni-karlsruhe.de
[2]tkuhr@ekp.uni-karlsruhe.de
[3]M.Kreps@warwick.ac.uk

Using the most recent world average values for $\mathcal{B}(B_d^0 \to D^+ D_s^-)$ and f_s/f_d published by the Particle Data Group [1] we obtain

$$\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-}) = (3.38 \pm 0.25(stat) \pm 0.30(sys) \pm 0.56(norm))\%$$

for the absolute semi-inclusive branching fraction. Given certain theoretical assumptions, this measurement is directly sensitive to the relative decay width difference $\Delta\Gamma_s/\Gamma$ in the B_s-$\bar{B}_s$ system:

$$\frac{\Delta\Gamma_s}{\Gamma_s} = (6.99 \pm 0.54(stat) \pm 0.64(sys) \pm 1.20(norm) \pm 0.34(theo))\%$$

Documentation of the changes with respect to the previous draft:

- Fit with mass resolution scale factor
- D^+ Dalitz model systematic uncertainty

Contents

1 Introduction

1.1 Theoretical Motivation

The $B_s^0 - \bar{B}_s^0$ system exhibits rapid oscillations, with the two mass eigenstates B_s^H (heavy) and B_s^L (light) having a sizable mass difference $\Delta m = M_H - M_L = 2\,|M_{12}|$ [2]. These mass eigenstates have distinct decay widths Γ_s^H and Γ_s^L, where $\Gamma_s = (\Delta\Gamma_s^H + \Delta\Gamma_s^L)/2$. In the standard model (SM) of particle physics, the decay width difference $\Delta\Gamma_s = \Gamma_s^L - \Gamma_s^H$ is predicted to be non-zero, of the order of $\sim 15\%$ in terms of the relative decay width difference $\Delta\Gamma_s/\Gamma_s$. However, this has not yet been experimentally established beyond doubt.

The time evolution of the $B_s^0 - \bar{B}_s^0$ mixing and decay problem is governed by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix}|B_s\,(t)\rangle\\ |\bar{B}_s\,(t)\rangle\end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right)\begin{pmatrix}|B_s\,(t)\rangle\\ |\bar{B}_s\,(t)\rangle\end{pmatrix} \tag{1}$$

Here the Hamiltonian

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

is composed of the mass matrix $\mathbf{M} = \mathbf{M}^\dagger$ and the decay matrix $\mathbf{\Gamma} = \mathbf{\Gamma}^\dagger$, which are complex 2×2 matrices having a phase $\phi_s = \arg\left(-M_{12}/\Gamma_{12}\right)$. In the Standard Model, mixing is realized/induced by second order transitions of weak interaction with the b ($\bar{b}$) quark going into an s ($\bar{s}$) and vice versa by W^+ and W^- exchange. These loop processes that are commonly illustrated in Feynman box diagrams are strongly dominated by the *top* quark with small contributions stemming from u and c quarks.

The flavor eigenstates B_s and $\bar{B}_s$ are not invariant under the consecutive application of the charge (C) and parity (P) transformation:

$$CP\,|B_s\rangle \;=\; -\,|\bar{B}_s\rangle \tag{3}$$

Thus, B_s and $\bar{B}_s$ are not eigenstates of CP, yet their linear combinations

$$|B_s^{CP+}\rangle \;=\; \frac{1}{\sqrt{2}}\left(|B_s\rangle \;-\; |\bar{B}_s\rangle\right)\,, \quad \text{and} \quad |B_s^{CP-}\rangle \;=\; \frac{1}{\sqrt{2}}\left(|B_s\rangle \;+\; |\bar{B}_s\rangle\right) \tag{4}$$

are. In this notation, B_s^{CP+} and B_s^{CP-} represent the CP even and odd eigenstates, respectively. With the distinct decay widths Γ_s^{CP+} and Γ_s^{CP-}, the CP decay width difference $\Delta\Gamma_s^{CP} = \Gamma_s^{CP+} - \Gamma_s^{CP-}$ is defined. Given the standard model scenario which predicts a negligibly small phase $\phi_s \approx 0.004$, corresponding to vanishing CP violation in the $B_s^0 - \bar{B}_s^0$ system, the mass eigenstates coincide with the CP eigenstates and the CP width difference equals the decay width difference, as will be shown below. However, additional non-standard model contributions in the mixing box diagrams could add a sizable and measurable phase ϕ_s^{NP} both to ϕ_s and $\beta_s = \arg[-V_{ts}V_{tb}^*/V_{cs}V_{cb}] = 0.02$.

Assuming Γ_{12} to receive its dominant contribution from Cabibbo-favored $b \to c\bar{c}s$ tree-level transitions into CP specific final states, $\Delta\Gamma_s^{CP}$ can directly be related to Γ_{12} by $\Delta\Gamma_s^{CP} = 2\,|\Gamma_{12}|$. With this, it can be shown [3]

$$\Delta\Gamma_s = 2\,|\Gamma_{12}|\cos\phi_s = \Delta\Gamma_s^{CP}\cos\phi_s. \tag{5}$$

Given a sample of tagged (i.e. the flavor of the B_s meson at production time was determined) $b \to c\bar{c}s$ decays, $\Delta\Gamma_s^{CP}$ and $\cos\phi_s$ can be extracted in angular and time dependent studies. The $\sin(2\beta_s)$ analysis using the decay $B_s \to J/\psi\phi$ is currently one the most prominent examples of current efforts in the B_s meson sector.

Hence, $b \to c\bar{c}s$ transitions into CP-specific final states provide an excellent laboratory for probing the decay width difference $\Delta\Gamma_s$ of the B_s meson system. In principle, the same experimental angular- and time-dependent methods could be applied to $B_s \to D_s^{(*)+}D_s^{(*)-}$ decays which hold the same quark content ($c\bar{c}s\bar{s}$) as $J/\psi\phi$ in the final state. Existing measurements (see section 1.2) however suggest that experiments will not yield sufficient signal statistics needed for adopting experimental angular and time-dependent techniques in the near future.

However, there is an alternative approach to access $\Delta\Gamma_s$ which is related to the fact that $B_s \to D_s^{(*)+}D_s^{(*)-}$ decays are believed to account for the main contribution to the sizable decay width difference $\Delta\Gamma_s$. Under certain theoretical assumptions $D_s^{(*)+}D_s^{(*)-}$ can be approximated to have a defined, predominantly even CP content, while Γ_s^{CP-} is neglected [4, 5]. With this, the CP width difference $\Delta\Gamma_s^{CP}$ is approximately equal to the CP even decay width:

$$\Delta\Gamma_s^{CP} = \Gamma_s^{CP+} - \Gamma_s^{CP-} \approx \Gamma_s^{CP+} \tag{6}$$

Due to color suppression the $c\bar{c}s\bar{s}$ final states $D_s^{(*)+}D_s^{(*)-}$ are preferred over $(c\bar{c})(s\bar{s})$ final states like $B_s \to J/\psi\phi$ or $B_s \to J/\psi\eta$. Therefore, the contribution of these color suppressed modes to $\Delta\Gamma_s/\Gamma_s$ can safely be neglected. For further details on these considerations please refer to the literature [4, 5]. Under these various assumptions the semi-inclusive branching fraction of $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ can be related to the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ by [3]

$$\begin{aligned} 2\mathcal{B}(B_s \to D_s^{+(*)}D_s^{-(*)}) &\cong \frac{\Delta\Gamma_s^{CP}}{2}\left[\frac{\frac{1}{1-2x_f}+\cos\phi_s}{\Gamma_s^L} + \frac{\frac{1}{1-2x_f}-\cos\phi_s}{\Gamma_s^H}\right] \\ &\cong \frac{\Delta\Gamma_s^{CP}}{2}\left[\frac{1+\cos\phi_s}{\Gamma_s^L} + \frac{1-\cos\phi_s}{\Gamma_s^H}\right], \end{aligned} \tag{7}$$

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 (7) simplifies to

$$2\mathcal{B}(B_s \to D_s^{+(*)}D_s^{-(*)}) \cong \frac{\Delta\Gamma_s}{\Gamma_s + \Delta\Gamma_s/2}, \tag{8}$$

which we solve for the relative decay width difference:

$$\frac{\Delta\Gamma_s}{\Gamma_s} \cong \frac{2\mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})}{1 - \mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})} \tag{9}$$

Thus, in first order a measurement of $\mathcal{B}(B_s \to D_s^{+(*)} D_s^{-(*)})$ is directly sensitive to the strongly SM dominated observable $\Delta\Gamma_s^{CP} = 2\,|\Gamma_{12}|$, independent of CP violation. 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 this analysis can make an important and complementary contribution to this field by providing a measurement of $\Delta\Gamma_s^{CP}$ using an approach that does not rely on angular and time-dependent 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 the $B_s^0 - \bar{B}_s^0$ meson system in untagged analyses.

1.2 Experimental Status

$B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ decays already have been subject to several measurements. The first experiment to report evidence for this decay was ALEPH [6]. In this measurement $D_s^{(*\pm)}$ mesons were reconstructed by evaluating correlations among two ϕ_s 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 $\mathcal{B}(B_s \to D_s^{(*)+} D_s^{(*)-})$ and a non-vanishing decay width difference using data corresponding to 2.8 fb^{-1} [7]. Based on a reconstruction of the semi-leptonic decay $D_s^{*+}(\to D_s^+\gamma/\pi^0)D_s^{*-}(\to D_s^-\gamma/\pi^0)$, $D_s^+(\to \phi\pi^+)D_s^-(\to \phi\mu^-\bar{\nu}_\mu)$, 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 being made to distinguish between decays of a D_s^+ or D_s^{*+} meson.

First observation of the decay $B_s \to D_s^+ D_s^-$ on the basis of an exclusive measurement was reported by the CDF collaboration using 355 pb^{-1} of data [8]. 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 a 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 stems from the Belle collaboration [9]. With 23.6 fb^{-1} of data recorded at the Belle experiment running on the $\Upsilon(5S)$ resonance, the decays $B_s \to D_s^{(*)+} D_s^{(*)-}$ were reconstructed in several hadronic decay channels. Overall, 23 signal events were found. Belle was the first experiment to disentangle the decay modes $B_s \to D_s^{(*)+} D_s^{(*)-}$, thus confirming observation of $B_s \to D_s^+ D_s^-$ ($\mathcal{B}(B_s \to D_s^+ D_s^-) = 0.0103^{+0.0039+0.0015}_{-0.0032-0.0013} \pm 0.0021$), plus claiming observation of $B_s \to D_s^{*+} D_s^-$ and

	ALEPH	DØ	CDF	Belle
N_s	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: 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 the estimate on the relative decay width difference is given using the theoretical considerations outlined in section 1.1.

	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 2: Overview of the results of $\Delta\Gamma_s/\Gamma_s$ depending on the way of combining individual results [10].

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 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 well-defined, the *Heavy Flavor Averaging Group* (HFAG) presents different results depending on the way individual results were measured and combined [10]. According to Table 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 ranges from about 9% to 15%.

1.3 Aim and Scope of This Analysis

It is the intention of this work to provide a new measurement of the both the exclusive and the semi-inclusive relative branching fractions $f_{D_s^{(*)} D_s^{(*)}} = f_s/f_d \mathcal{B}(B_s \to B_s^0 \to D_s^{(*)+} D_s^{(*)-}/\mathcal{B}(B^0 \to D^+ D_s^-)$ using 6.8 fb^{-1} of CDF Run II Data. From these figures 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 – and $\mathcal{B}(B_d^0 \to$

$D^+D_s^-$). Whenever more accurate measurements of f_s/f_d and $\mathcal{B}(B_d^0 \to D^+D_s^-)$ are available, 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.

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

2 Preparation of Data Samples

2.1 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 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 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.

2.2 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. The charged D^+ meson is reconstructed from $D^+ \to K^-\pi^+\pi^+$. 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[4], while the D^+ meson is reconstructed from its decay into one charged kaon and two equally charged pions ($K^-\pi^+\pi^+$).

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 tracks assumed to be kaons (kaon and

[4] 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.

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 GeV/c^2 $< M_{K^+K^-} <$ 1.035 GeV/c^2 (0.837 GeV/c^2 $< M_{K^-\pi^+} <$ 0.947 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 $\pm$100 MeV/c^2 around the D_s mass of 1.968 GeV/c^2 published by the *Particle Data Group* (PDG) [1]. When performing the B_s^0 vertex fit the D_s^+ mass is constrained to its PDG value. Tables 3 and 4 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, data files holding information about B meson decays are called *BStntuples* at CDF.

From the BStntuples, which to 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. The *skimmed* or *flat* Ntuples are a streamlined version of the BStntuples and contain the hadronic decay modes of interest only.

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.

2.3 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 $c\tau_{B_s^0} = 441$ μm ($\tau = 1.471$ ps) and $c\tau_{B_d^0} = 458.7$ μm ($\tau = 1.530$ ps), respectively. These values correspond to the state of knowledge of the year 2008 [11]. For the B_d^0 meson, the current edition of the Review of Particles Physics [1] quotes a

$B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$			$B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$		
$3.0 <$	$m(B_s^0)$	< 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		$\Delta Z_0(\phi)$	< 1.5

Table 3: 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 in the Appendix for definition of variables.

$B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$			$B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0} K^-$		
$3.0 <$	$m(B_s^0)$	< 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			
	$\Delta Z_0(K^{*0})$	< 1.5		$\Delta Z_0(K^{*0})$	< 1.5

Table 4: Offline reconstruction requirements placed on the studied decay channels with one D_s^- decaying into $K^{*0}K^+$.

slightly smaller world average value of

$$\begin{aligned} \tau_{B_d^0} &= 1.519 \pm 0.007 \text{ ps} \\ c\tau_{B_d^0} &= 455.4 \pm 2.1 \ \mu\text{m} \end{aligned} \tag{10}$$

Therefore, all B_d^0 Monte Carlo samples are re-weighted using the central value of the more recent result (10). Re-weighting of Monte Carlo is done according to the relation

$$w_{c\tau} = \exp\left(\frac{ct}{c\tau_{MC}} - \frac{ct}{c\tau}\right) \tag{11}$$

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. [1] quotes

$$\begin{aligned} \tau_{B_s^0} &= 1.472^{+0.024}_{-0.026} \text{ ps} \\ c\tau_{B_s^0} &= 441^{+7}_{-8} \ \mu\text{m} \end{aligned} \tag{12}$$

as mean lifetime. This is still 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 1.1. 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) [10] the mean lifetime of the short-living light eigenstate amounts

$$\begin{aligned} \tau_{B_s^L} &= 1.408^{+0.033}_{-0.030} \text{ ps} \\ c\tau_{B_s^L} &= 422^{+10}_{-9} \ \mu\text{m} \end{aligned} \tag{13}$$

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, $c\tau_{B_s^L}$ (13). Changes in the both the B_d^0 and B_s^0 lifetime assumptions will be subject to systematic studies (Section 5.3.1).

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

As described in Section 2.2, 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 combinatoric

background events. However, methods will be put in place to effectively separate signal meson events from combinatorics (Section 3).

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 [1] 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, i.e. the CLEO collaboration. Based on the CLEO measurement, in the following sections the model implementation is described step-by-step.

2.4.1 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 [12]. 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^2_{ab}, m^2_{ac}, or m^2_{cb}.

If decay dynamics were governed solely by phase space, events in the plane of m^2_{ab} and m^2_{ac}, 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^2_{ab}, m^2_{ac}) 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

$$d\Gamma = \frac{|\mathcal{M}|^2}{(2\pi)^3 32 (M_{D_s})^3} dm^2_{ab} dm^2_{bc}, \tag{14}$$

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,

$$\mathcal{M} = \sum_R c_R \times \Omega_R \times \mathcal{F}^L_D \times \mathcal{F}^L_R \times \mathcal{W}_R, \tag{15}$$

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:

$$\Omega^{L=0}_R = 1 \tag{16}$$

$$\Omega^{L=1}_R = m^2_{bc} - m^2_{ac} + \frac{(m^2_D - m^2_c)(m^2_a - m^2_b)}{m^2_{ab}} \tag{17}$$

$$\begin{aligned}\Omega^{L=2}_R =& [\Omega^{L=1}_R]^2 - \frac{1}{3}\left(m^2_{ab} - 2m^2_D - 2m^2_c + \frac{(m^2_D - m^2_c)^2}{m^2_{ab}}\right) \\ &\times \left(m^2_{ab} - 2m^2_a - 2m^2_b + \frac{(m^2_a - m^2_b)^2}{m^2_{ab}}\right)\end{aligned} \tag{18}$$

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

$\mathcal{F}^L_D$ and $\mathcal{F}^L_R$ 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

[13]. Here, the cases of a zero meson radius, $r = 0$, giving rise to $\mathcal{F}_V^L = 1$, and non-zero meson radii

$$\mathcal{F}_R^0 = 1 \tag{19}$$
$$\mathcal{F}_D^0 = e^{-(z-z_R)/12} \tag{20}$$
$$\mathcal{F}_{R,D}^1 = \sqrt{\frac{1+z_R}{1+z}} \tag{21}$$
$$\mathcal{F}_{R,D}^2 = \sqrt{\frac{9+3z_R+z_R^2}{9+3z+z^2}} \tag{22}$$

where

$$z = r_{R,D}^2 p^2 \tag{23}$$
$$z_R = r_{R,D}^2 p_R^2 \tag{24}$$

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

$$\mathcal{W}_R(m) = \frac{1}{m_R^2 - m^2 - i m_R \Gamma(m)}. \tag{25}$$

The mass-dependent width can be expressed as [14]

$$\Gamma(m) = \Gamma_R \frac{m_R}{m} \left(\frac{p}{p_R}\right)^{2L+1} [\mathcal{F}_R^L(z)]^2. \tag{26}$$

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 [15] formula is commonly used:

$$\mathcal{W}_R(m) = \frac{1}{m_R^2 - m^2 - i \sum_{ab} g_{Rab}^2 \rho_{ab}(m)} \tag{27}$$

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

$$\mathcal{W}_\kappa(m) = \frac{1}{m_\kappa^2 - m^2} \tag{28}$$

is a common choice.

2.4.2 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 [16]. Using the CLEO-c data set corresponding to 548pb^{-1} of e^+e^- collisions at $\sqrt{s} = 4.17\text{GeV/c}^2$, 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:

$$FF_R = \frac{\oint |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}}{\oint \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}} \tag{29}$$

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 5. 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 it should be referred to [1, 17, 18].

2.4.3 Reproducing CLEO's Dalitz Plot Results

Based on the Dalitz formalism outlined above, additional valuable information provided by the authors [19], and by inserting parameter values (masses, widths, magnitudes, and phases of resonances) quoted in the CLEO paper (in particular results quoted in Table 5), two routines are coded: The first one provides a calculation of the squared

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

Table 5: 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 [16]. 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.

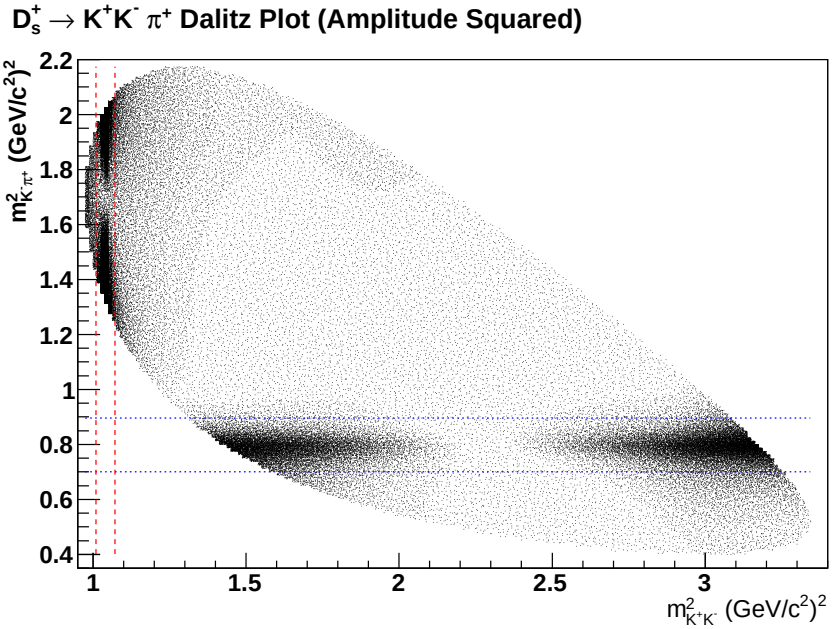


Figure 1: Graphical representation of the squared decay amplitude $\mathcal{M}(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ for the decay $D_s^+ \to K^+K^-\pi^+$ calculated from the results of CLEO's Dalitz plot analysis [16].

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 2.6 and 3.5.

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 of the pdf (i.e. the squared amplitude $|\mathcal{M}|^2$) is set up statically[5] only, with the values of magnitudes and phases determined by the CLEO Collaboration being inserted. Figure 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 $K^{*0}(892)$ 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

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

	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 6: 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.

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 6. 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.

2.4.4 Remarks and Nomenclature

The considerations made in Section 2.4 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 10% 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 75% 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 hadronic 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.

2.4.5 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:

$$f_{\phi^\dagger\pi^+} = \frac{\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = \frac{\int_{\phi(1020)} \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}}{\oint \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}} \tag{30}$$

$$f_{K^{*0\dagger}K^+} = \frac{\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = \frac{\int_{K^{*0}(892)} \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}}{\oint \sum_R |\mathcal{M}_R|^2 \, dm^2_{K^+K^-} dm^2_{K^-\pi^+}} \tag{31}$$

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 (30) and (31) various times. However, in each calculation the 12 CLEO Dalitz plot parameters (those quoted in Table 5) 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 B.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 [20]). The result is a vector of parameter uncertainties which are now correlated among each other according to

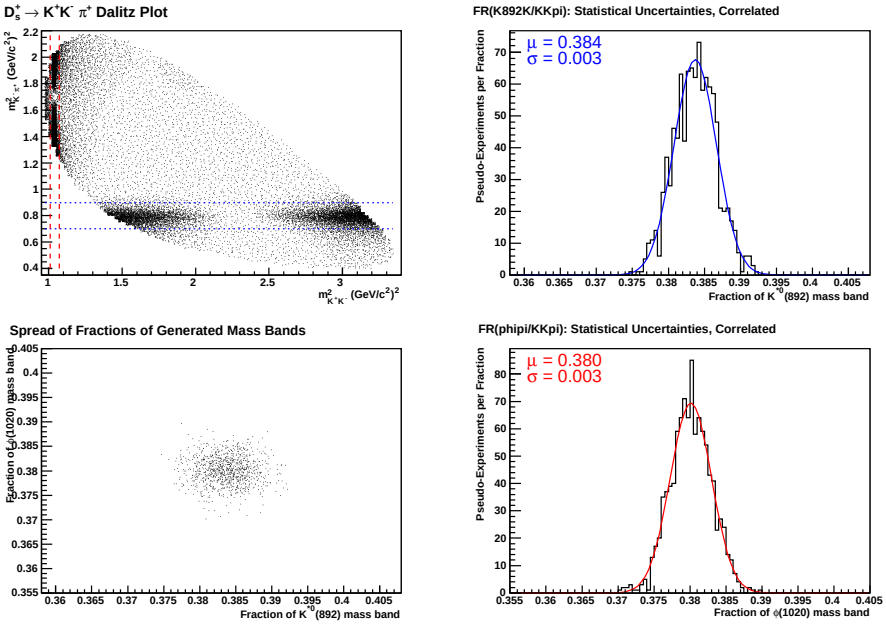


Figure 2: Computation of the fraction of $D_s^+ \to K^+K^-\pi^+$ decays as defined by equations (30) and (31). In the 1000 trials the statistical uncertainties of the CLEO Dalitz model parameters and their correlations were accounted for.

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 44 and 45 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, 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

$$f_{\phi^\dagger\pi^+} = \frac{\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = 0.380 \pm 0.003(stat) \pm 0.010(sys) = 0.380 \pm 0.010 \quad (32)$$

$$f_{K^{*0\dagger}K^+} = \frac{\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} = 0.384 \pm 0.003(stat) \pm 0.011(sys) = 0.384 \pm 0.011 \quad (33)$$

The top right and bottom right plots in Figures 2 and 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.

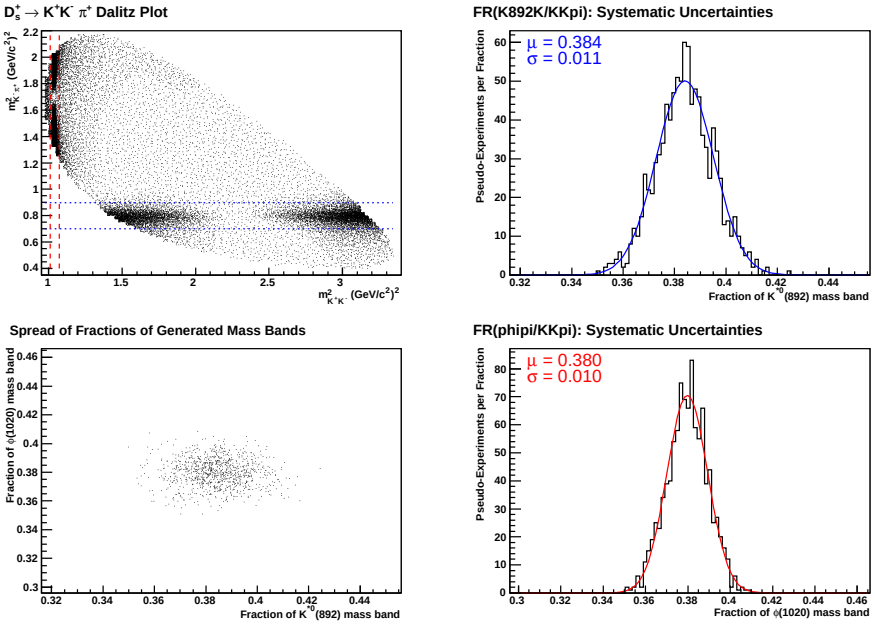


Figure 3: Computation of the fraction of $D_s^+ \to K^+K^-\pi^+$ decays as defined by equations (30) and (31) when accounting for systematic parameter uncertainties CLEO quotes.

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

$$\begin{matrix} 1 & -0.035 \\ -0.035 & 1 \end{matrix} \tag{34}$$

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

$$\begin{matrix} 1 & -0.013 \\ -0.013 & 1 \end{matrix} \tag{35}$$

We consider the tiny anti-correlations between the two relative branching fractions as being negligible. From the relative fractions (32) and (33), 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 $\mathcal{B}(D_s^+ \to K^+K^-\pi^+) = 0.0549 \pm 0.0027$:

$$\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+) = 0.0209 \pm 0.0012 \tag{36}$$

$$\mathcal{B}(D_s^+ \to K^{*0\dagger}K^+) = 0.0211 \pm 0.0012 \tag{37}$$

For the reasons outlined in Section 4.4, 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 (36) and (37) will be again factorized into (32) and (33) times the branching fraction of $D_s^+ \to K^+K^-\pi^+$.

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

The elaborations on the off-line reconstruction given in Section 2.2 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. 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) [1]:

$$\mathcal{B}(D_s^{*+} \to D_s^+ \gamma) = (94.2 \pm 0.7) \times 10^{-2} \tag{38}$$

$$\mathcal{B}(D_s^{*+} \to D_s^+ \pi^0) = (5.8 \pm 0.7) \times 10^{-2} \tag{39}$$

Furthermore, simulation of the detector response is implemented in the Monte Carlo software package (see Section 2.6). 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 [21] and the derived pedagogical reviews [22, 23, 24], the subsequent sections provide a brief introduction to this topic.

2.5.1 Decay Amplitudes in the Helicity Basis

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

$$\begin{aligned} 1 &\to 2\ 3 \\ 2 &\to 4\ 5 \\ 3 &\to 6\ 7 \end{aligned} \tag{40}$$

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 [21]. 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 $|jm'\rangle$:

$$R(\alpha, \beta, \gamma)\,|jm\rangle = \sum_{m'=-j}^{j} D^j_{m',m}(\alpha, \beta, \gamma)\,|jm'\rangle \tag{41}$$

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

$$D^j_{m',m}(\alpha, \beta, \gamma) = e^{-im'\alpha} d^j_{m',m}(\beta) e^{-im\gamma} \tag{42}$$

$d^j_{m',m}$ are commonly referred to as the reduced Wigner functions, some of which are tabulated in the literature, in [1] for instance. From the identity

$$d^j_{m',m} = (-1)^{m-m'} d^j_{m,m'} = d^j_{-m,-m'} \tag{43}$$

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 $1 \rightarrow 2\,3$, in the helicity frame the decay amplitude is calculated by

$$A(\lambda_1, \lambda_2, \lambda_3, \Omega) = \sqrt{\frac{2s_1+1}{4\pi}} D^{s_1 *}_{\lambda_1, \lambda_2 - \lambda_3}(\Omega) A_{\lambda_2 \lambda_3}, \tag{44}$$

where the set of angles $\phi, \theta, -\phi$ are encapsulated in Ω. Equation (44) 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

$$|\lambda_2 - \lambda_3| \leq s_1 \tag{45}$$

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

$$A_{-\lambda_2 - \lambda_3} = \eta_2 \eta_3 \eta_1 (-1)^{s_2 + s_3 - s_1} A_{\lambda_2 \lambda_3}, \tag{46}$$

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 $3 \rightarrow 6\,7$. Constraining particle 1 to be a spin-0 particle with fixed helicity $\lambda_1 = 0$, equation (45) implies that $\lambda_2 = \lambda_3 = \lambda$. Therefore, the general expression for the decay amplitude of a sequence of two-body decays

$$\begin{aligned} A(\lambda_1, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \Omega, \Omega', \Omega'') = & \sqrt{\frac{(2s_1+1)}{4\pi}\frac{(2s_2+1)}{4\pi}\frac{(2s_3+1)}{4\pi}} \\ & \times \sum_{|\lambda_2 - \lambda_3| = 0}^{s_1} D^{s_1 *}_{\lambda_1, \lambda_2 - \lambda_3}(\Omega) A_{\lambda_2 \lambda_3} D^{s_2 *}_{\lambda_2, \lambda_4 - \lambda_5}(\Omega') B_{\lambda_4 \lambda_5} D^{s_3 *}_{\lambda_3, \lambda_6 - \lambda_7}(\Omega'') C_{\lambda_6 \lambda_7} \end{aligned} \tag{47}$$

simplifies to

$$\begin{aligned} A(\lambda_4, \lambda_5, \lambda_6, \lambda_7, \Omega', \Omega'') = & \sqrt{\frac{1}{4\pi}\frac{(2s_2+1)}{4\pi}\frac{(2s_3+1)}{4\pi}} \\ & \times \sum_{|\lambda| = 0}^{1} A_{\lambda\lambda} D^{s_2 *}_{\lambda, \lambda_4 - \lambda_5}(\Omega') B_{\lambda_4 \lambda_5} D^{s_3 *}_{\lambda, \lambda_6 - \lambda_7}(\Omega'') C_{\lambda_6 \lambda_7}, \end{aligned} \tag{48}$$

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

$$|A(\theta,\psi,\chi)|^2 = \sum_{|\lambda_4-\lambda_5|}^{s_1} \sum_{|\lambda_6-\lambda_7|}^{s_1} \left| \sum_{|\lambda|=0}^{1} H_\lambda e^{i\lambda\chi} d^1_{\lambda,\lambda_4-\lambda_5}(\theta) B_{\lambda_4\lambda_5} d^1_{\lambda,\lambda_6-\lambda_7}(\psi) C_{\lambda_6\lambda_7} \right|^2 \tag{49}$$

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 [25, 26] by

$$T_\parallel = \frac{H_+ + H_-}{\sqrt{2}}, \quad T_0 = H_0, \quad T_\perp = \frac{H_+ - H_-}{\sqrt{2}} \tag{50}$$

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

$$f_L = \frac{|H_0|^2}{|H_+|^2 + |H_0|^2 + |H_-|^2} = \frac{|T_0|^2}{|T_\parallel|^2 + |T_0|^2 + |T_\perp|^2} \tag{51}$$

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:

$$f_{CP+} = f_\parallel + f_0 = \frac{|T_\parallel|^2 + |T_0|^2}{|T_\parallel|^2 + |T_0|^2 + |T_\perp|^2} \tag{52}$$

$$f_{CP-} = f_\perp = \frac{|T_\perp|^2}{|T_\parallel|^2 + |T_0|^2 + |T_\perp|^2} \tag{53}$$

2.5.2 $B_s^0 \to D_s^{*+} D_s^-$

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

$$0 \quad \to \quad \begin{matrix} 1 \\ 0 \\ -1 \end{matrix} \quad 0 \tag{54}$$

Being a spin-1 particle, the D_s^{*+} meson has three possible helicities -1, 0, -1. Due to conservation of angular momentum (45) only one helicity projection onto the decay

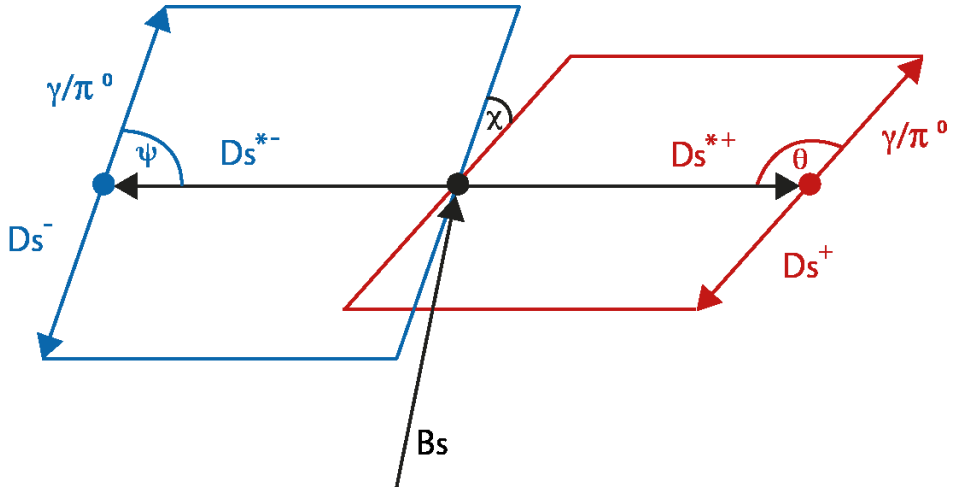


Figure 4: The decay $B_s^0 \to D_s^{*+} D_s^{*-}$ in the helicity frame. D_s^{*+} de-excitation proceeds via $D_s^{*+} \to D_s^+ \gamma$ or $D_s^{*+} \to D_s^+ \pi^0$

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 4.2.3. 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 2.6) are relied upon, including event decay simulation using the *EvtGen* package [27, 28, 29].

2.5.3 $B_s^0 \to D_s^{*+} D_s^{*-}$

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

$$0 \quad \to \quad \begin{matrix} 1 & 1 \\ 0 & 0 \\ -1 & -1 \end{matrix} \tag{55}$$

in terms of the helicities $\lambda_1 \to \lambda_2 \lambda_3$. Figure 4 illustrates the decay in the helicity frame. Because of (45), 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 (49) is given by

$$|A(\theta,\psi,\chi)|^2 = \sum_{|\lambda_4-\lambda_5|}^{1} \sum_{|\lambda_6-\lambda_7|}^{1} \left| \sum_{|\lambda|=0}^{1} H_\lambda e^{i\lambda\chi} d^1_{\lambda,\lambda_4-\lambda_5}(\theta) B_{\lambda_4\lambda_5} d^1_{\lambda,\lambda_6-\lambda_7}(\psi) C_{\lambda_6\lambda_7} \right|^2 , \qquad (56)$$

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 C. 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 (46):

$$B_{01} = -B_{0-1} \qquad (57)$$

$$C_{01} = -C_{0-1} \qquad (58)$$

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 effect the resulting angular shape.

Equation (56) 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 [30] it is however reasonable to use the helicity amplitudes measured for $B_d^0 \to D^{*+} D_s^{*-}$, $\{H_+, H_0, H_-\} = \{0.4904, 0.7204, 0.4904\}$ [31]: 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 5.3.4). 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 (52-53), the fraction of the CP-odd component, if significantly different from zero. This will be investigated in Section 4.5.

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

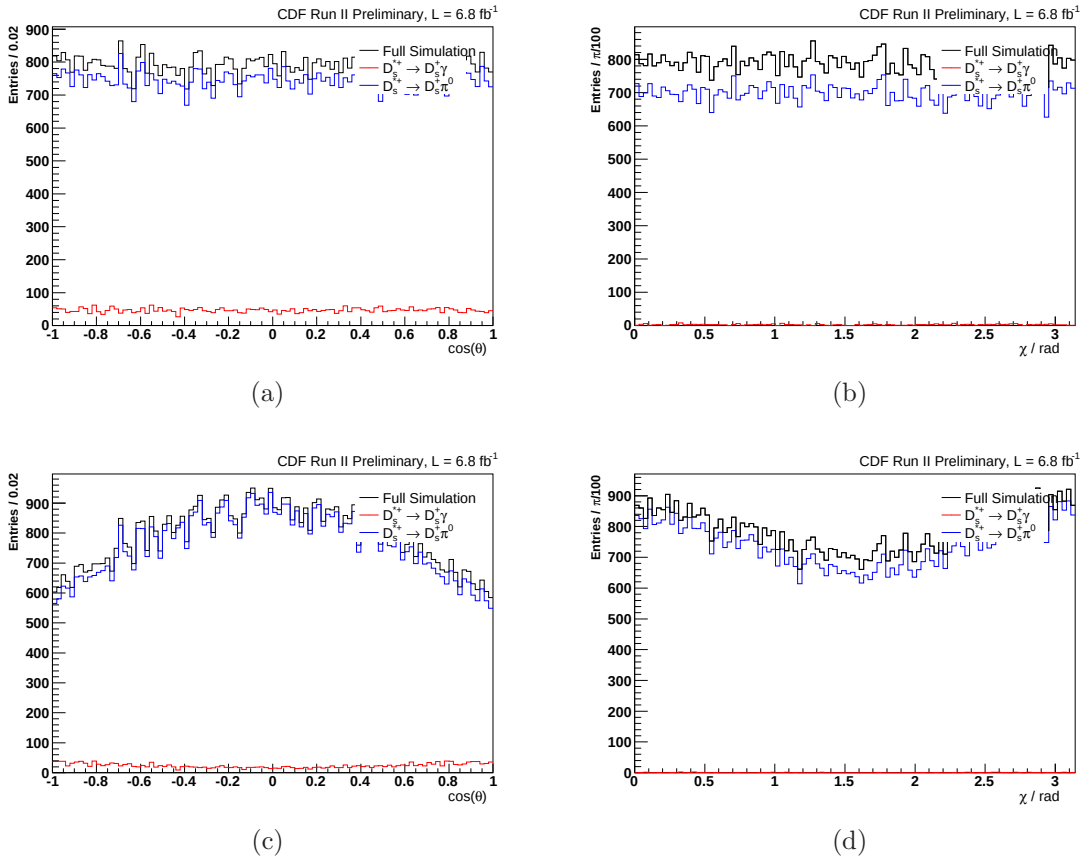


Figure 5: Monte Carlo $\cos(\theta)$ (a,c) and χ (b,d) distributions of $B_s^0 \to D_s^{*+} D_s^{*-}$ phase space (upper row) and re-weighted Monte Carlo (bottom row), where in the latter case weights were calculated from the standard helicity amplitude configuration $\{H_+, H_0, H_-\} = \{0.4904, 07204, 0.4904\}$.

data, where the weights are given by the squared decay amplitudes calculated from equation (56). 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 5 the angular distributions of the helicity angles $\cos(\theta)$ and χ are shown for the phase space case and the standard helicity scenario $\{H_+, H_0, H_-\} = \{0.4904, 0.7204, 0.4904\}$.

2.6 Monte Carlo Simulation

Monte Carlo Ntuple production comprises several successive stages, starting with the creation of B mesons using the event generator *BGenerator* (BGen)[32]. In contrast to the widely-used PYTHIA [33] 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 [34].

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* [27, 28, 29] 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 run the complete sequence of event generation, event decay, and a GEANT3 based [35] detector simulation the *cdfSim* [36] program that invokes the corresponding sub-modules one after another. To apply the CDF Two Track Trigger simulation the *TRGSim++* [37] 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 2.4, 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 daughter four-momenta that are randomly generated. 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 $f_{\phi\dagger\pi^+}$ 30 and $f_{K^{*0}\dagger K^+}$ 31. Rather than integrating over Dalitz plot squared 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 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 Monte Carlo simulations will adequately reflect 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 D

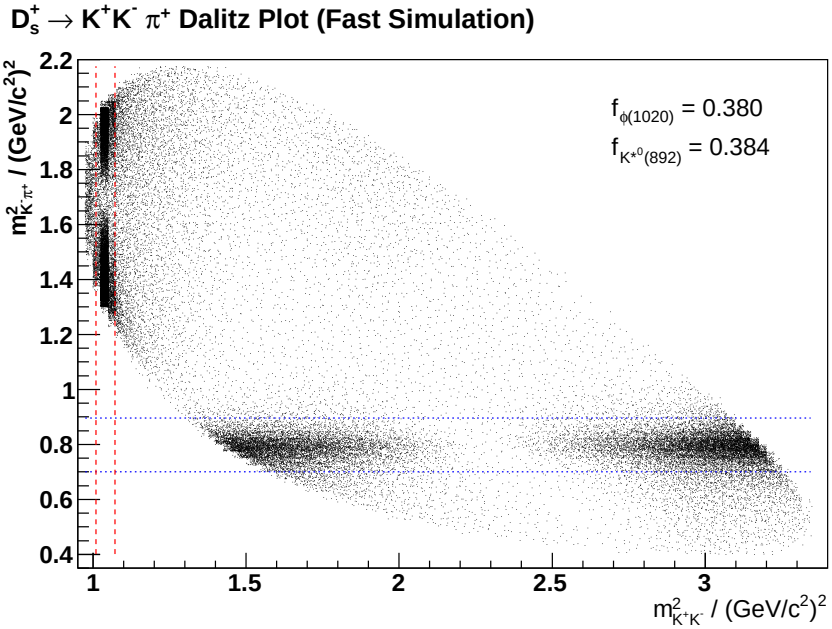


Figure 6: Scatter plot of simulated events populating the $D_s^+ \to K^+K^-\pi^+$ Dalitz $(m^2_{K^+K^-}, m^2_{K^-\pi^+})$ plane.

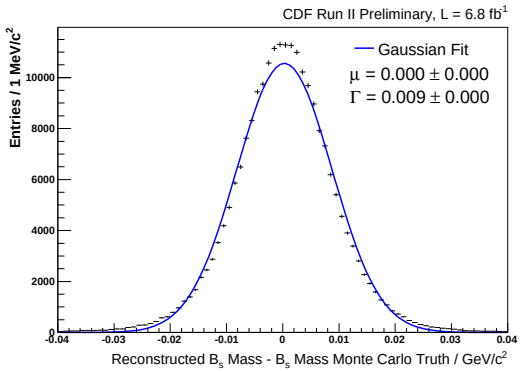
of the Appendix. In total, around 4.5 billion B_s^0 mesons and 1.8 billion B_d^0 mesons are generated and decayed. The numbers of events to be generated per run are determined by means of weights derived from integrated luminosity acquired for each run. Overall, simulations are run for a wide range of 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 2.2 are applied to the samples of simulated events.

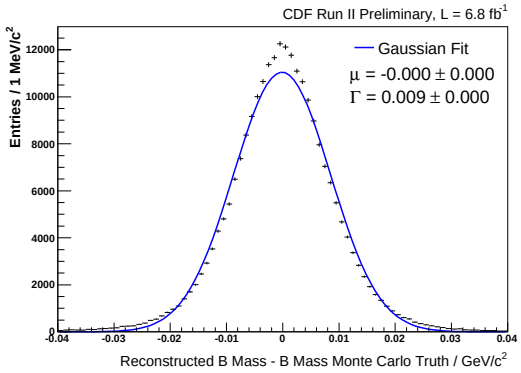
2.7 B Meson Transverse Momentum and Rapidity Spectrum

We conclude this section 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 driven by the uncertainties of reconstructed particle four-momenta, since the natural B meson width is tiny [1]. 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



(a)



(b)

Figure 7: Detector resolution of reconstructed B_s^0 meson mass (a) and B_d^0 meson mass (b) estimated from simulated data.

function to the distribution. The detector resolution is then given by the Gaussian width. Figure 7 displays the mass resolution fits both for the B_s^0 and the B_d^0 meson using the $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ 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 4.3), 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

$$w(m) = \frac{P_{bg}^i(m - \Delta m)}{P_{bg}^i(m)}, \tag{59}$$

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.

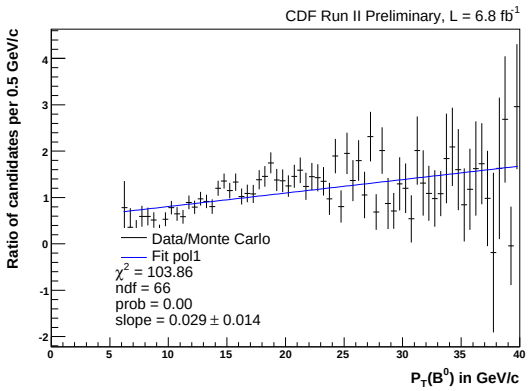


Figure 8: Normalized ratio of $p_T(B^0)$ distributions observed in side-band subtracted data and Monte Carlo. Simulated data has not been corrected for true Two-Track Trigger fractions.

	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 7: 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.

The parameters (slope, constant offset) of the background parameterization function P^i_{bg} are obtained from a fit to the channel $B^0_d \to D^+ D^-_s \to K^-\pi^+\pi^+\phi\pi^-$ (for fit details see Section 4.3). 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 selection please refer to sections 3.2 and 3.3). 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 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 section. 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 9). According to Table 7, 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 7. If not stated otherwise, this kind of Monte Carlo correction is applied whenever utilizing simulated data.

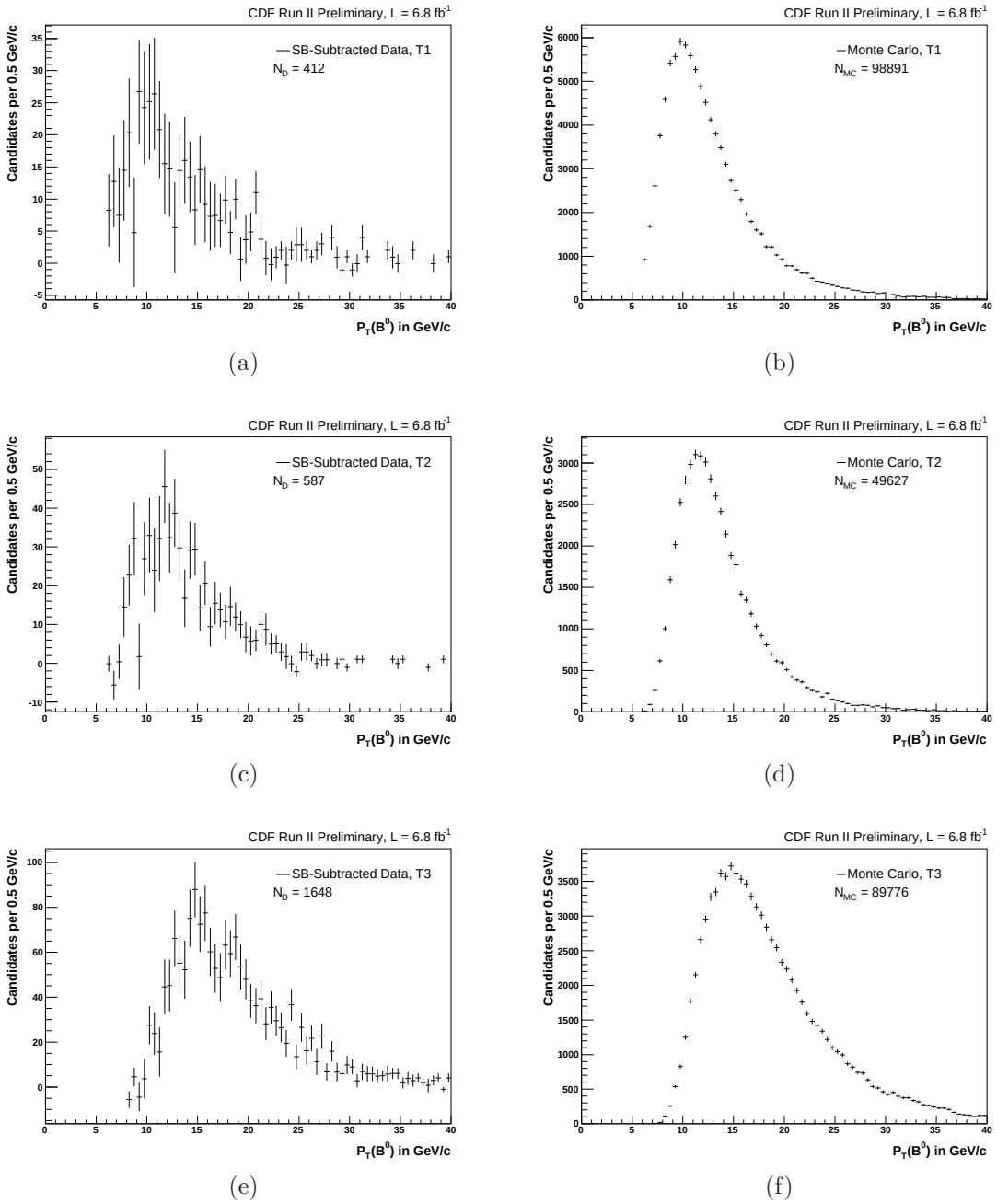
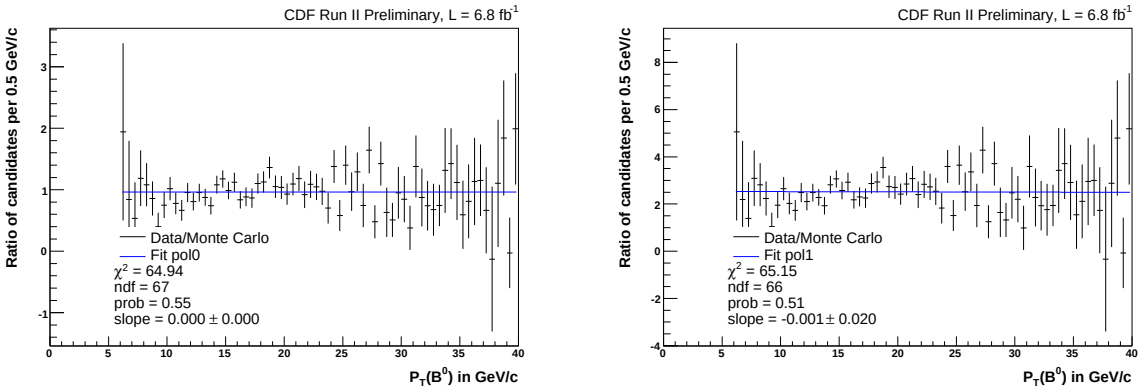


Figure 9: From first to third row, $p_T(B^0)$ distributions of the exclusive Two-Track Trigger sub-samples (labeled by TriggerFlags $T1$, $T2$, $T3$) in side-band subtracted data (left column) and Monte Carlo (right column).



(a) (b)

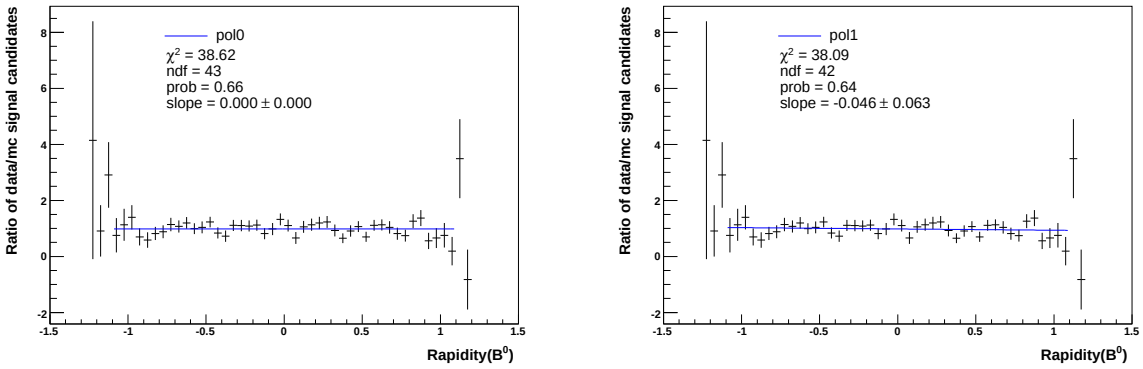
Figure 10: Normalized ratio of $p_T(B^0)$ distributions observed in side-band subtracted data and Two-Track Trigger corrected Monte Carlo using the decay channel $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$. We test the hypothesis of data – Monte Carlo agreement by fitting a constant to the distribution (a). Fitting a 1st order polynomial to the same distribution (b) yields a slope being compatible with 0 and almost identical fit probability.

Figures 10(a) and 10(b) show the normalized ratio of the number of data over corrected Monte Carlo events in p_T bins of 500 MeV/c. We test the hypothesis of p_T compliance in Monte Carlo and data by fitting a straight line to the ratio of p_T ratio distribution (Figure 10(a)). Fitting a 1st order polynomial to the same distribution (Figure 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. In the same way the agreement of the $Y(B^0)$ distributions observed in data and Monte Carlo (Figures 11(a) and 11(b)) is validated.

3 Candidate Selection

3.1 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 are large fraction of combinatorial noise, 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



(a) (b)

Figure 11: Ratio of $Y(B^0)$ distributions observed in side-band subtracted data and Monte Carlo using the decay channel $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$. We test the hypothesis of data – Monte Carlo agreement by fitting a constant to the distribution (a). Fitting a 1st order polynomial to the same distribution (b) gives a slope being compatible with 0 and almost the same fit probability.

that accounts for variable correlations.

3.2 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 kinematical variables are applied to sort out obvious background events. This enables a much smoother operation of artificial 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 8 and 9 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 nominal 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

$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$		$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$	
$q\left(\pi_{D_s^{(1)}}\right) \times q\left(\pi_{D_s^{(2)}}\right)$	$= -1$	$q\left(\pi_{D_s}\right) + q\left(K_D\right) + q\left(\pi_D^{(1)}\right) + q\left(\pi_D^{(2)}\right)$	$= 0$
		$q\left(\pi_D^{(1)}\right) \times q\left(\pi_D^{(2)}\right)$	$= 1$
$p_T\left(B_s^0\right)$	> 6.0 GeV/c	$p_T\left(B_d^0\right)$	> 6.0 GeV/c
$\lvert d_0\left(B_s^0\right)\rvert$	< 0.012 cm	$\lvert d_0\left(B_d^0\right)\rvert$	< 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 8: Applied pre-cuts for the studied decay channels having $\phi\pi^+$ in the final state. See Section A in the Appendix for definition of variables.

$B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$		$B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$	
$q\left(\pi_{D_s^{(1)}}\right) \times q\left(K_{D_s^{(2)}}\right)$	$= -1$	$q\left(\pi_{D_s}\right) + q\left(K_D\right) + q\left(\pi_D^{(1)}\right) + q\left(\pi_D^{(2)}\right)$	$= 0$
$q\left(K_{K^{*0}}\right) \times q\left(K_{D_s}\right)$	$= -1$	$q\left(K_{K^{*0}}\right) \times q\left(K_{D_s}\right)$	$= -1$
		$q\left(\pi_D^1\right) \times q\left(\pi_D^2\right)$	$= 1$
$p_T\left(B_s^0\right)$	> 6.0 GeV/c	$p_T\left(B_d^0\right)$	> 6.0 GeV/c
$\lvert d_0\left(B_s^0\right)\rvert$	< 0.012 cm	$\lvert d_0\left(B_d^0\right)\rvert$	< 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 9: Applied pre-cuts for the studied decay channels having $K^{*0}K^+$ in the final state. See Section A in the Appendix for definition of variables.

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 $\chi^2_{r\phi}(D_s^{(1,2)}) < 20$.

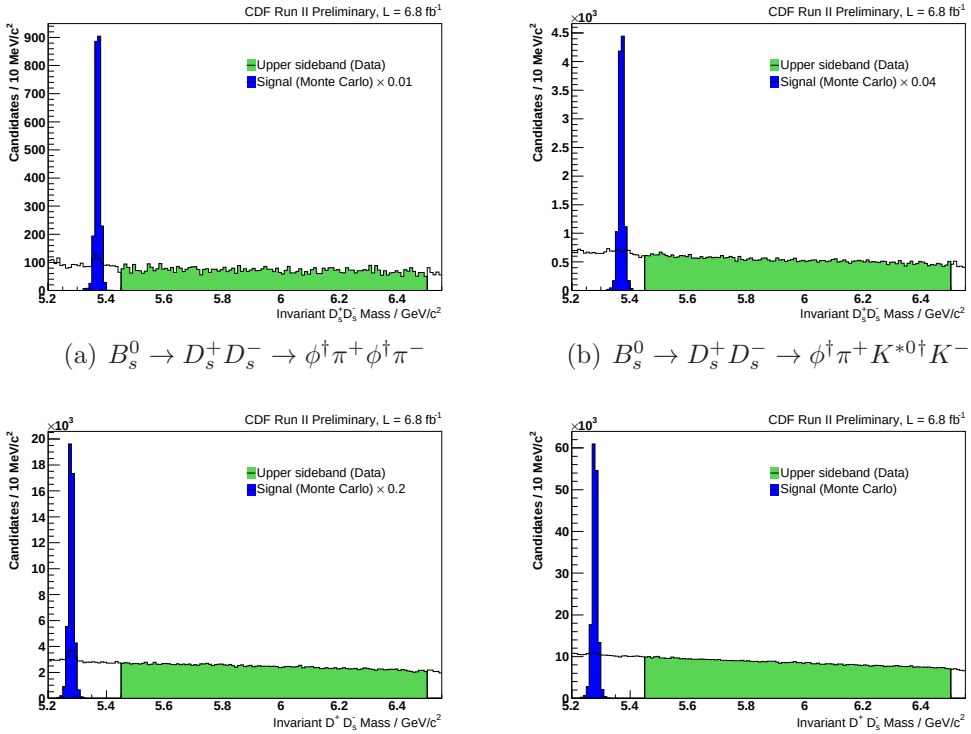
3.3 Neural Network Selection

For final event selection this analysis makes use of the NeuroBayes® [38, 38] program. This multivariate analysis software package combines a sophisticated variable pre-processing algorithm with a feed-forward three-layer artificial neural network.

3.3.1 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 4.2 and 4.3), 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 studied decay channels upper sideband data ranging from 5.45 to 6.5 GeV/c^2 are used as background training patterns. Figure 12 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 $B^0_{(s)}$ 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 (please refer to Chapter 4 for



(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 12: 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.

details),

$$f_{D_s^{(*)}D_s^{(*)}} = \frac{f_s}{f_d}\frac{\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})}{\mathcal{B}(B_d^0 \to D^+D_s^-)}, \tag{60}$$

where $B_d^0 \to D^+D_s^-$ is used as normalization channel. Having common selections for the signal and normalization channels will make systematic studies easier, because selection-related systematic effects cancel in the ratio of topologically similar decays.

3.3.2 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 [8] 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 a common set 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 ($D_s^+ \to \phi\pi^+ \to K^+K^-\pi^+$ on the one hand, $D_s^+ \to K^{*0}K^+ \to K^+\pi^-K^+$ 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 10 and 10 give a schematic overview of the input variables entering the network trainings.

3.3.3 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: First, training variables are presented ordered by their significance (Tables 12 and 13). 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.

Index	B_s^0 B_d^0	$D_s^{(1)}$ $D_s^{(1)}$	$D_s^{(2)}$ D	$\phi^\dagger\pi$ $\phi^\dagger\pi$	$\phi^\dagger\pi$ $K\pi\pi$
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				$\min(d_0/\sigma_{d_0})$	
22				$\mathrm{m}(K^{(1)}K^{(2)})$	
23				$\mathrm{m}(K^{(1)}\pi)$	
24				$\mathrm{m}(K^{(2)}\pi)$	
25			$m_{rec} - m_{pdg}$		

Table 10: 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 for definition of variables. Index number 1 is reserved for the target variable.

Index	B_s^0 B_d^0	$D_s^{(1)}$ $D_s^{(1)}$	$D_s^{(2)}$ D	$K^{*0\dagger}K$ $K^{*0\dagger}K$	$\phi^\dagger\pi$ $K\pi\pi$
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				$L_{xy}(D_s \leftarrow K^{*0\dagger})$	
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				$\min(d_0/\sigma_{d_0})$	
22				m$(K^{(1)}K^{(2)})$	
23				m$(K^{(1)}\pi)$	
24				m$(K^{(2)}\pi)$	
25			$m_{rec} - m_{pdg}$		

Table 11: 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	$\text{prob}(B^0_{(s)})$	411.13	182.55	40.61	85.87	6
4	$\text{PID.ratioKaon}(K^{(2)}_{D^{(2)}_s})$	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	$\text{PID.ratioKaon}(K^{(2)}_{D^{(2)}_s})$	431.52	68.18	50.40	72.93	19
7	$\left\lvert d_0(B^0_{(s)})\right\rvert$	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	$\text{m}(K^{(2)}K^{(2)})$	266.26	33.89	29.07	43.30	22
11	$\Delta\text{m}(D^{(2)}_{(s)})$	263.56	31.94	27.61	54.64	25
12	$\text{m}(K^{(2)}\pi)$	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	$\text{m}(K^{(2)}\pi)$	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	$\left\lvert d_0(D^{(2)}_s)\right\rvert$	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	$\text{prob}(D^{(2)}_s)$	153.63	2.42	2.41	79.72	12
-	2σ threshold	-	-	-	-	-
24	$\text{PID.ratioPion}(\pi_{D^{(2)}_s})$	288.52	0.68	0.68	62.85	15

Table 12: 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 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.

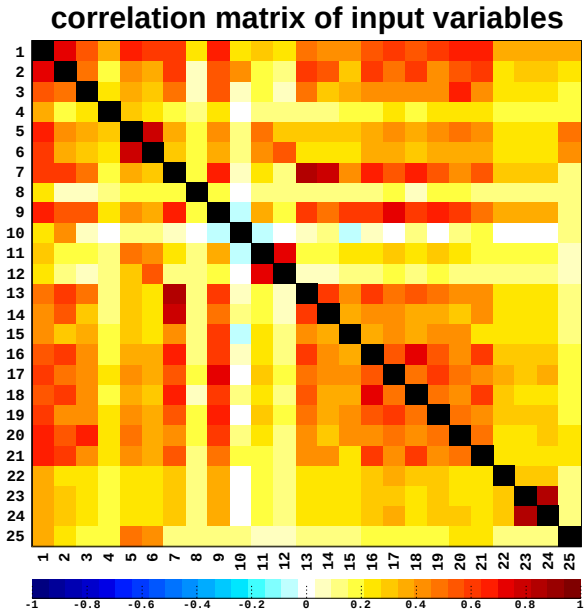
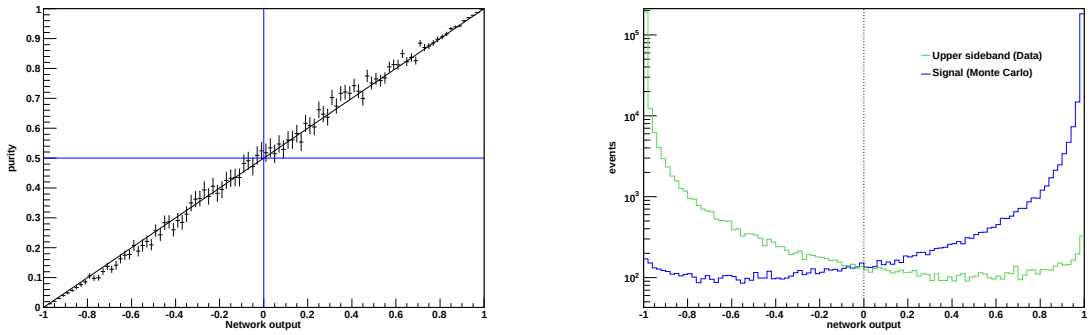


Figure 13: Correlation matrix of the input variables used for the combined network training of the decays $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$



(a) Purity over network output.

(b) Network classification of training data.

Figure 14: Two of the training graphs indicating the quality and the performance of the network trained for $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi\pi^-$

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 decorrelation 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 [38] and the supporting technical documents coming with the NeuroBayes® software package [39].

The correlations between the variables are illustrated by color-coded correlation matrices (Figures 13 and 15). Finally, two of the characteristic graphs indicating a neural network's quality and performance are given (Figures 14 and 16): In the graph to the left hand side 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, in the second one a network's capability to separate between signal and background events is verified on the training samples.

3.4 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 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 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 4. 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

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	$\mathrm{m}(K^{(2)}\pi)$	469.71	83.06	12.34	95.10	24
9	$\min(d_0/\sigma_{d_0})$	555.11	57.50	69.10	71.75	21
10	$\left\lvert d_0(B^0_{(s)})\right\rvert$	316.76	63.06	50.71	38.88	4
11	$\mathrm{prob}(B^0_{(s)})$	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	$\mathrm{m}(K^{(1)}\pi)$	249.91	34.05	33.38	27.27	23
16	$\Delta\mathrm{m}(D^{(2)}_{(s)})$	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	$\mathrm{m}(K^{(1)}K^{(2)})$	465.15	12.34	12.49	95.06	22
21	$\left\lvert d_0(D_s^{(1)})\right\rvert$	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	$\mathrm{prob}(D_s^{(1)})$	235.88	3.59	3.59	79.16	12

Table 13: Significance-ranked input variables of 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. See Appendix A for variable definitions.

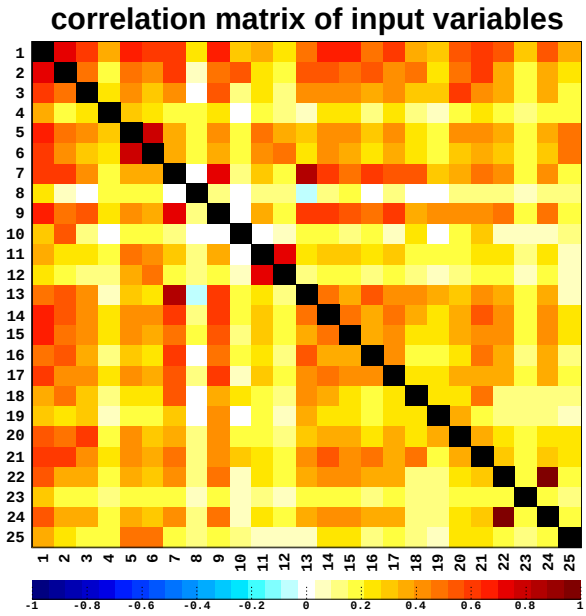
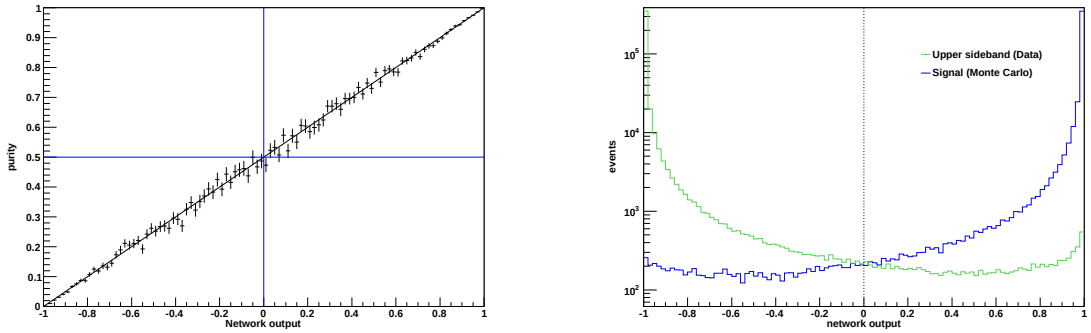


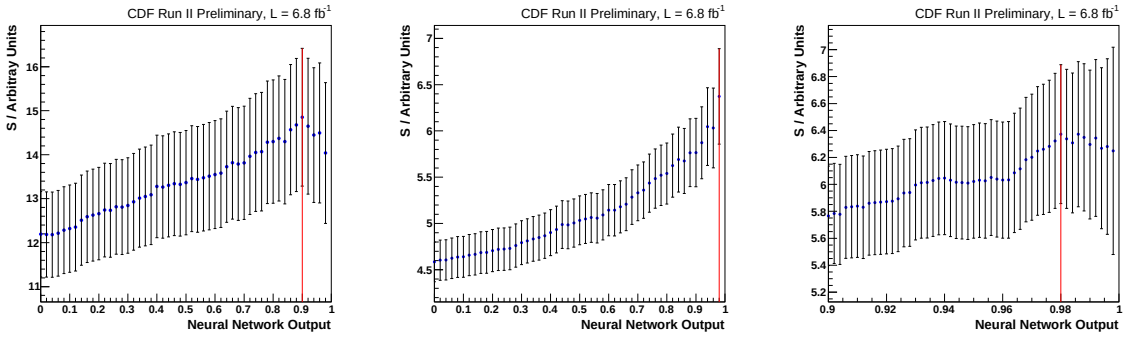
Figure 15: Correlation matrix of the input variables used for the combined network training of the decays $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0} K^-$.



(a) Purity over network output.

(b) Network classification of training data.

Figure 16: Graphs indicating the quality and the performance of the network trained for $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$ and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0} K^-$.



(a) (b) (c)

Figure 17: Signal significance $S = N_S^{exp}/\sqrt{N_S + N_B}^{Data}$ as a function of neural network output for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ (a) and $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ (b,c). The vertical red lines indicate the neural network requirements finally chosen. The error bars give the combined uncertainty due to the uncertainty of number of data events lying in the signal region and due to limited Monte Carlo statistics.

N_S^{exp} is chosen to be the expected number of signal events for a given network cut, deduced from the general relation (74):

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

$\epsilon_{B_s^0 \to D_s^+ D_s^-}$ 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 3.5). 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^{Data}_{S+B}. 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 17 shows the figure of merit S as a function of the Neural Network threshold for the decay $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ (left) and $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ (middle and right-hand plot). Because no maximum significance is found for $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ 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.

3.5 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 established in the previous section, and the number of generated simulation events:

$$\epsilon = \frac{N_{rec}^{MC}}{N_{gen}^{MC}} = \frac{N_{rec}^{MC}}{f^{MC} N_{tot}^{MC}} = \sum_{t}^{3} \frac{w_t N_{rec,t}^{MC}}{f^{MC} N_{tot}^{MC}} \tag{61}$$

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 Monte-Carlo-Data comparison (Section 2.7). 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 D), 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 2.4.5: The fraction of generated events in a particular mass window with respect to the number of events of the full Dalitz plot is equal to the ratio of the corresponding mass window integrals of squared amplitudes (calculated from the Dalitz model described in Section 2.4) and the full Dalitz plot integral, equations (30), (31). The numerical values are given by equations (32), (33). 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 14 the combined reconstruction and selection efficiencies estimated from simulated data are tabulated channel by channel.

4 Parameter Estimation

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* [1, 40, 41] is used. The negative log likelihood function to be minimized reads

$$F(\vec{a}) = -2 \sum_{k=1}^{n} \ln \left[\sum_{l}^{q} N_l h_l (m_k \mid \vec{a}) \right] + 2 \sum_{l=1}^{q} N_l (\vec{a}) \tag{62}$$

The factor of 2 is convention. The first task is hence to set up the negative log likelihood function (62) and its constituting pdfs. This is now being worked out

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 14: 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.

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.

4.1 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.

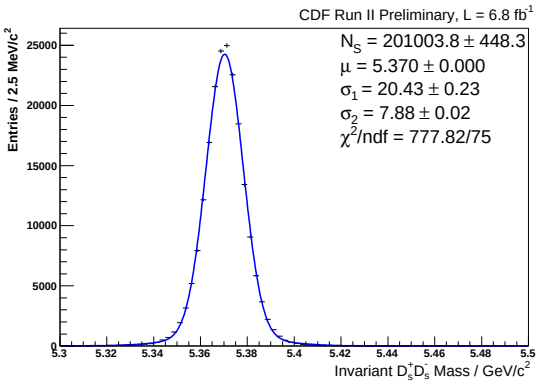
Partially reconstructed signal Events originating from true B mesons where the full decay chain cannot be fully reconstructed because because neutral particles are note detected or otherwise lost in reconstruction.

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.

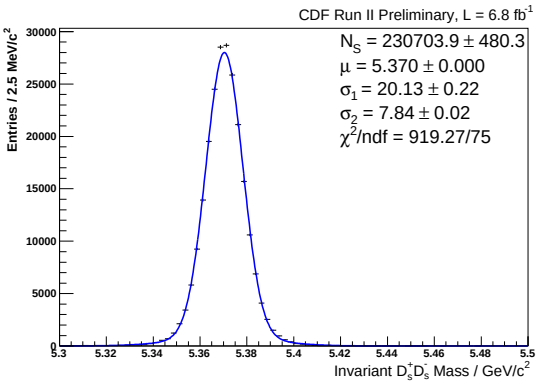
The individual fit model contributions belonging to each of these categories are discussed in the upcoming sections in more detail.

4.2 $B_s^0 \to D_s^+ D_s^-$ 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 parameters that are impossible to be all simultaneously determined during the final EML fit on data. 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 14. The shape parameters are determined by binned ML fits to the individual components.



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

Depending on the signal-like component, the parameters can then be chosen to be fixed, completely free or allowed to float within Gaussian constraints in the EML fit to real data.

4.2.1 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,

$$P_{bg}^i = \exp\left(m \mid s^i\right) + c^i, \tag{63}$$

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.

4.2.2 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

$$P_{sig}^i = f_{\mathcal{G}1}^i \mathcal{G}\left(m \mid \mu\sigma_1^i\right) + \left(1 - f_{\mathcal{G}1}^i\right) \mathcal{G}\left(m \mid \mu\sigma_2^i\right), \tag{64}$$

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 18(a), 18(b)). In the final EML fit to real data the widths are kept fixed, but scaled with a common factor that is floating and accounts for different mass resolutions in

data and simulation. The mean is shared among to the full fit function (for details please refer to Section 4.4) and allowed to float within a Gaussian constraint.

4.2.3 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 $D_s^{*+} \to D_s^+ \pi^0$ $(5.8 \pm 0.7\%)$ [1].

- $B_s^0 \to D_s^{*+} D_s^-$
  The physics observables related to the decay dynamics of this decay are well established today (see Section 2.5.2), 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,

$$\begin{aligned} P_{p1}^i &= f_{\mathcal{G}1}^i \mathcal{G}\left(m \mid \mu_1^i \sigma_1^i\right) + f_{\mathcal{G}2}^i \mathcal{G}\left(m \mid \mu_2^i \sigma_2^i\right) + \left(f_{\mathcal{G}3}^i\right) \mathcal{G}\left(m \mid \mu_3^i \sigma_3^i\right) \\ &\quad + \left(1 - f_{\mathcal{G}1}^i - f_{\mathcal{G}2}^i - f_{\mathcal{G}3}^i\right) \mathcal{G}\left(m \mid \mu_4^i \sigma_4^i\right), \end{aligned} \tag{65}$$

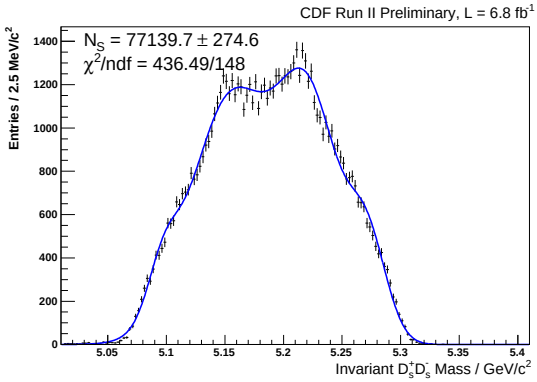
  is found to fit reasonably well (Figures 19(a), 19(b)). The small admixture of $D_s^{*+} \to D_s^+ \pi^0$ produces a smooth double peak structure: From helicity amplitude considerations and conservation of angular momentum (see Section 2.5.2) it follows that the D_s^{*+} is 100% longitudinally polarized. 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 $\{H_+, H_0, H_-\} = \{0.4904, 07204, 0.4904\}$ [31]. In-detail discussions can be found in Section 2.5.3.

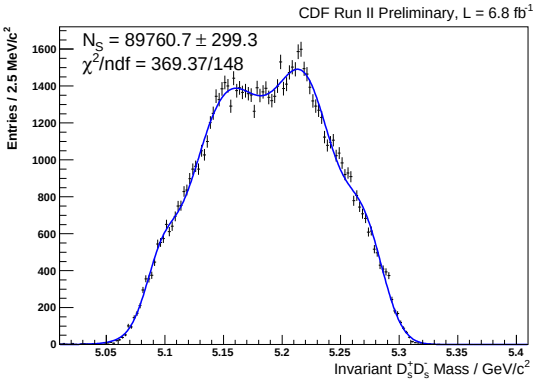
  Simulation generates a broad, featureless shape, that is fitted by the sum of three Gaussians (Figures 20(a), 20(b)):

$$P_{p2}^i = f_{\mathcal{G}1}^i \mathcal{G}\left(m \mid \mu_1^i \sigma_1^i\right) + f_{\mathcal{G}2}^i \mathcal{G}\left(m \mid \mu_2^i \sigma_2^i\right) + \left(1 - f_{\mathcal{G}1}^i - f_{\mathcal{G}2}^i\right) \mathcal{G}\left(m \mid \mu_3^i \sigma_3^i\right) \tag{66}$$

The choice of which functions to use to describe the partially reconstructed events indeed seems quite arbitrary. However, since this study does not intend to extract any

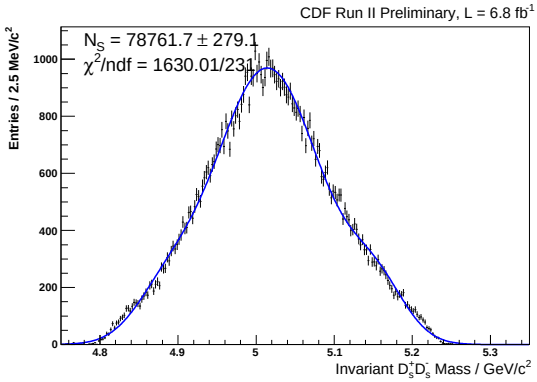


(a) $B_s^0 \to D_s^{*+} D_s^- \to \phi\pi^+\phi\pi^-$ template.

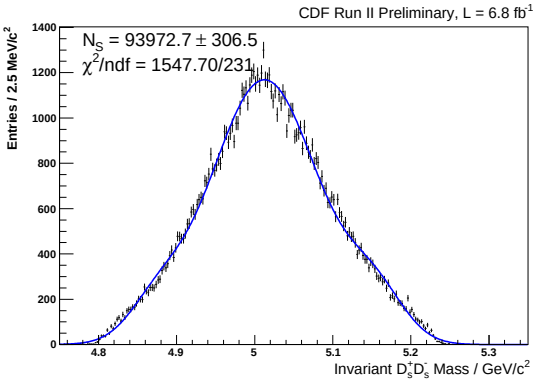


(b) $B_s^0 \to D_s^{*+} D_s^- \to \phi\pi^+ K^{*0} K^-$ template.

Figure 19: 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 20: 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).

physics parameters from these shapes those functions are chosen that do best describe our simulated data. For both of templates of partially reconstructed events described in this section the shape parameters and fractions of the individual single Gaussians are fixed in the final fit.

4.2.4 Reflections

The decay $B_s^0 \to D_s^+ (\phi\pi^+) D_s^- (K^{*0}K^-)$ has a strong reflection from $B_d^0 \to D_s^+ (\phi\pi^+)$ $D^-(K^+\pi^-\pi^-)$, 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 $B_d^0 \to D_s^+ (\phi\pi^+) D^- (K^+\pi^-\pi^-)$ Monte Carlo samples were reconstructed as $B_s^0 \to D_s^+ (\phi\pi^+) D_s^- (K^{*0}K^-)$ 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 21(a)):

$$P^i_{reflsig} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1\sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2\sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3\sigma^i_3\right) \quad (67)$$

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 21(b) to 21(d)):

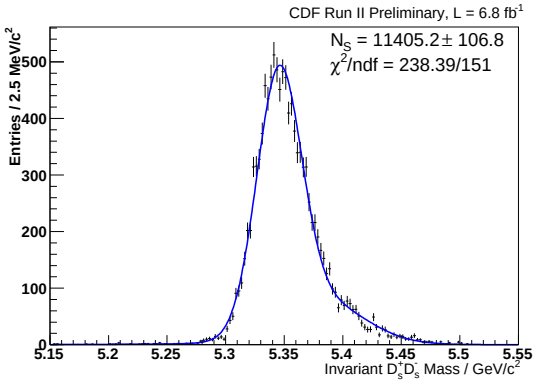
$$P^i_{rp1,2,3} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1\sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2\sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3\sigma^i_3\right) \quad (68)$$

4.2.5 Full Fit Function

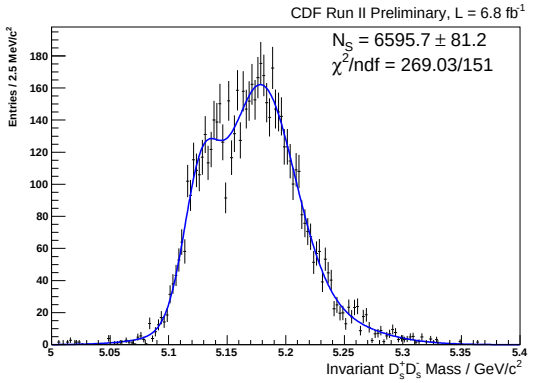
Following equation (62), the individual pdfs are combined to give the following negative log likelihood function used for EML fit to the $B_s^0 \to D_s^+ D_s^-$ invariant mass spectra:

$$\begin{aligned} F^i(\vec{a}) \;= -2\textstyle\sum_{k=1}^{n} \ln \;& [N^i_1 P^i_{sig}(m_k \mid \vec{a_{sig}}) + N^i_2 P^i_{p1}(m_k \mid \vec{a}_{p1}) + N^i_3 P^i_{p2}(m_k \mid \vec{a}_{p1}) \\ & + N^i_4 P^i_{rsig}(m_k \mid \vec{a}_{rsig1}) + N^i_5 P^i_{rp1}(m_k \mid \vec{a}_{rp1}) + N^i_6 P^i_{rp2}(m_k \mid \vec{a}_{rp2}) \\ & + N^i_8 P^i_{rp3}(m_k \mid \vec{a}_{rp3}) + N^i_8 P^i_{bg}(m_k \mid \vec{a}_{bg})] \\ & + 2\sum_{l=1}^{8} N^i_l, \end{aligned} \quad (69)$$

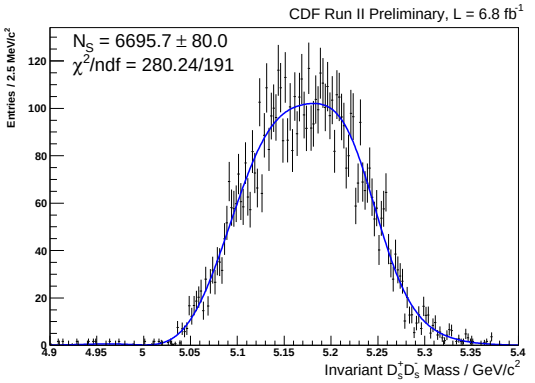
n is the total number of events of the fitted data sample and i labels the given B_s^0 decay mode. The factors N_l^i in front of each probability density function are the numbers of



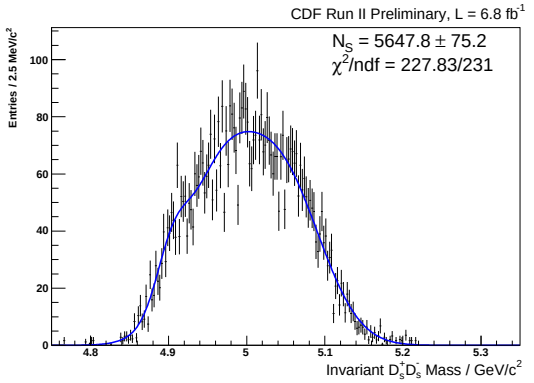
(a) Template for $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi \pi^-$ signal mis-reconstructed as $B_s^0 \to D_s^+ D_s^- \to \phi \pi^+ K^{*0} K^-$.



(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 21: 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 $B_d^0 \to D^+ D_s^{*-}$.

events of the given component:

$$\begin{aligned} N_1^i &= N^i_{B_s^0 \to D_s^+ D_s^-} \\ N_2^i &= N^i_{B_s^0 \to D_s^{*+} D_s^-} \\ N_3^i &= N^i_{B_s^0 \to D_s^{*+} D_s^{*-}} \\ N_4^i &= N^i_{B_d^0 \to D^+ D_s^-} \\ N_5^i &= N^i_{B_d^0 \to D^{*+} D_s^-} \\ N_6^i &= N^i_{B_d^0 \to D^+ D_s^{*-}} \\ N_7^i &= N^i_{B_d^0 \to D^{*+} D_s^{*-}} \\ N_8^i &= N^i_{bkg} \end{aligned} \tag{70}$$

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

4.2.6 Exclusive Ratios: Function Re-Parameterization

Based on the per-channel fit function (69), 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

$$N^i_{B_s^0 \to D_s^+ D_s^-} = N^{tot}_{B_s^0} \mathcal{B}(B_s^0 \to D_s^+ D_s^-) \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B} \begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-}, \tag{71}$$

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 3.5). 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, $N^j_{B_d^0 \to D^+ D_s^-}$, is given by

$$N^j_{B_d^0 \to D^+ D_s^-} = N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^+ D_s^-) \mathcal{B} \begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^j_{B_d^0 \to D^+ D_s^-} \tag{72}$$

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

In the next step we intend to introduce a common factor equations (71) and (72) share. For this, we eliminate $N_{B_s^0}^{tot}$ in equation (71) by introducing the ratio of the fragmentation fractions, f_s/f_d, of *strange* and *down* quarks produced at the Tevatron:

$$\frac{N_{B_s^0}^{tot}}{N_{B_d^0}^{tot}} = \frac{f_s}{f_d} \tag{73}$$

With this, equation (71) now reads

$$N^i_{B_s^0 \to D_s^+ D_s^-} = N^{tot}_{B_d^0} f_{D_s D_s} \mathcal{B}(B_d^0 \to D^+ D_s^-) \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B} \begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-}, \tag{74}$$

where in the same step the quantity

$$f_{D_s D_s} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^+ D_s^-)}{\mathcal{B}(B_d^0 \to D^+ D_s^-)} \tag{75}$$

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 $N_{B_d^0}^{tot}$ is globally shared among the combined log likelihood function.

Following the formulation outlined above, the parameterization of the expected yields of $B_s^0 \to D_s^{*+} D_s^-$ and $B_s^0 \to D_s^{*+} D_s^{*-}$ decay events is straightforward:

$$\begin{aligned} N^i_{B_s^0 \to D_s^{*+} D_s^-} =& N^{tot}_{B_d^0} f_{D_s^* D_s} \mathcal{B}(B_d^0 \to D^+ D_s^-) \mathcal{B}(D_s^{*+} \to D_s^+ X) \\ & \times \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B} \begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^{*+} D_s^-} \end{aligned} \tag{76}$$

$$\begin{aligned} N^i_{B_s^0 \to D_s^{*+} D_s^{*-}} =& N^{tot}_{B_d^0} f_{D_s^* D_s^*} \mathcal{B}(B_d^0 \to D^+ D_s^-) \mathcal{B}^2(D_s^{*+} \to D_s^+ X) \\ & \times \mathcal{B}(D_s^+ \to \phi^\dagger \pi^+) \mathcal{B} \begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^{*+} D_s^{*-}} \end{aligned} \tag{77}$$

$\mathcal{B}(D_s^{*+} \to D_s^+ X) \equiv 1$ 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 (75):

$$f_{D_s^* D_s} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)}{\mathcal{B}(B_d^0 \to D^+ D_s^-)} \tag{78}$$

$$f_{D_s^* D_s^*} = \frac{f_s}{f_d} \frac{\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})}{\mathcal{B}(B_d^0 \to D^+ D_s^-)} \tag{79}$$

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 4.2.4, 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 casual misreconstruction 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

$$N^{mis}_{B_d^0\to D^+D_s^-} = N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^+D_s^-)\mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^{mis}_{B_d^0\to D^+D_s^-}, \tag{80}$$

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:

$$\begin{aligned} N^{mis}_{B_d^0\to D^{*+}D_s^-} =& N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^{*+}D_s^-)\mathcal{B}(D^{*+} \to D^+X) \\ &\times \mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^{mis}_{B_d^0\to D^{*+}D_s^-} \end{aligned} \tag{81}$$

$$\begin{aligned} N^{mis}_{B_d^0\to D^{+}D_s^{*-}} =& N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^{+}D_s^{*-}) \\ &\times \mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^{mis}_{B_d^0\to D^{+}D_s^{*-}} \end{aligned} \tag{82}$$

$$\begin{aligned} N^{mis}_{B_d^0\to D^{*+}D_s^{*-}} =& N^{tot}_{B_d^0}\mathcal{B}(B_d^0 \to D^{*+}D_s^{*-})\mathcal{B}(D^{*+} \to D^+X) \\ &\times \mathcal{B}(D_s^+ \to \phi^\dagger\pi^+)\mathcal{B}(D^+ \to K^-\pi^+\pi^+)\epsilon^{mis}_{B_d^0\to D^{*+}D_s^{*-}} \end{aligned} \tag{83}$$

Here $\mathcal{B}(D^{*+} \to D^+X) = 0.323 \pm 0.006$ [1] 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.

4.2.7 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:

$$f_{D_s^{(*)}D_s^{(*)}} = f_{D_sD_s} + f_{D_s^*D_s} + f_{D_s^*D_s^*} \tag{84}$$

In principle the semi-inclusive branching fraction could be easily calculated from equation (84), 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 (84) 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:

$$f_{D_s D_s} = f_{D_s^{(*)} D_s^{(*)}}(1 - p_{D_s^* D_s} - p_{D_s^* D_s^*}) \tag{85}$$

$$f_{D_s^* D_s} = f_{D_s^{(*)} D_s^{(*)}} p_{D_s^* D_s} \tag{86}$$

$$f_{D_s^* D_s^*} = f_{D_s^{(*)} D_s^{(*)}} p_{D_s^* D_s^*} \tag{87}$$

Using these definitions, we may re-write the expected number of reconstructed events defined in equations 74 through 77 as

$$N^i_{B_s^0 \to D_s^+ D_s^-} = N^{tot}_{B_d^0} f_{D_s^{(*)} D_s^{(*)}}(1 - p_{D_s^* D_s} - p_{D_s^* D_s^*}) \times \mathcal{B}(B_d^0 \to D^+ D_s^-)\mathcal{B}(D_s^+ \to \phi^\dagger \pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \tag{88}$$

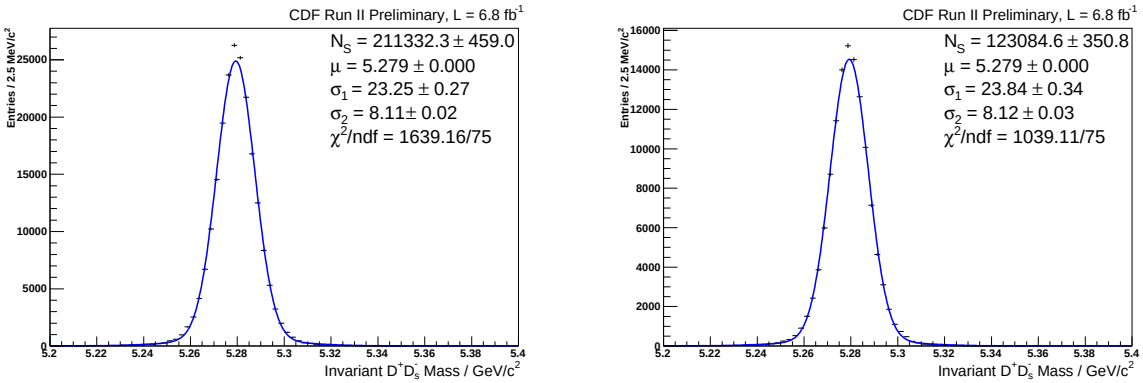
$$N^i_{B_s^0 \to D_s^{*+} D_s^-} = N^{tot}_{B_d^0} f_{D_s^{(*)} D_s^{(*)}} p_{D_s^* D_s} \times \mathcal{B}(B_d^0 \to D^+ D_s^-)\mathcal{B}(D_s^+ \to \phi^\dagger \pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \tag{89}$$

$$N^i_{B_s^0 \to D_s^{*+} D_s^{*-}} = N^{tot}_{B_d^0} f_{D_s^{(*)} D_s^{(*)}} p_{D_s^* D_s^*} \times \mathcal{B}(B_d^0 \to D^+ D_s^-)\mathcal{B}(D_s^+ \to \phi^\dagger \pi^+)\mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \epsilon^i_{B_s^0 \to D_s^+ D_s^-} \tag{90}$$

Instead of the exclusive ratios $f_{D_s D_s}$, $f_{D_s^* D_s}$, 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.

4.3 $B_d^0 \to D^+ D_s^-$ Fit Function

The log likelihood function and the individual templates involved are set up in a very similar way to those discussed in Section 4.2. The decay $B_d^0 \to D_s D$ 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 22(a), 22(b)) having a common mean value and a floating width scale factor in the final fit, 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.



(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.

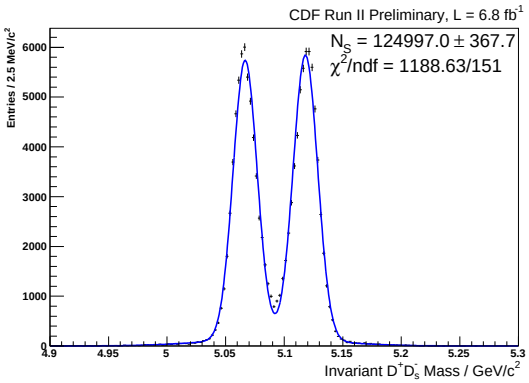
Figure 22: Signal template for the decay $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+\phi^\dagger\pi^-$ (a) and $B_d^0 \to D^+ D_s^- \to K^-\pi^+\pi^+ K^{*0\dagger} K^-$ (b).

4.3.1 Partially Reconstructed Signal

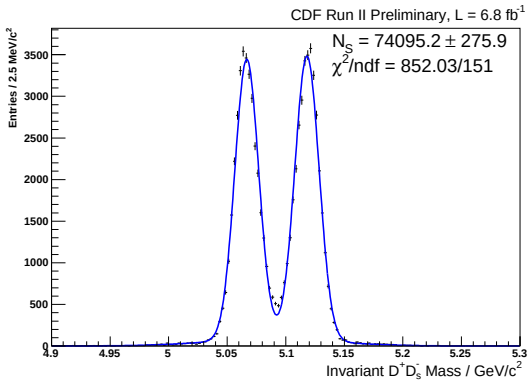
The decays $B_d^0 \to D_s^{*+} D^{*-}$ produce a series of satellite peaks shifted to smaller masses approximately by one to two pion masses with respect to the position of the signal peak of fully reconstructed $B_d^0 \to D^+ D_s^-$ signal. There are three possible combinations:

- $B_d^0 \to D^{*+} D_s^-$
  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 4.2.3: 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 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 23(a), 23(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 24(a), 24(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 23: 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).

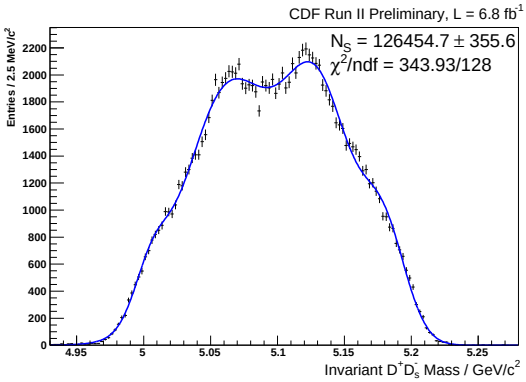
- $B_d^0 \to D^{*+} D_s^{*-}$
  This mode generates a wide, featureless structure located about two pion masses below the $B_d^0 \to D_s^+ D^-$ signal. A triple Gaussian is used to fit the resulting shape (Figures 25(a), 25(b)).

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.

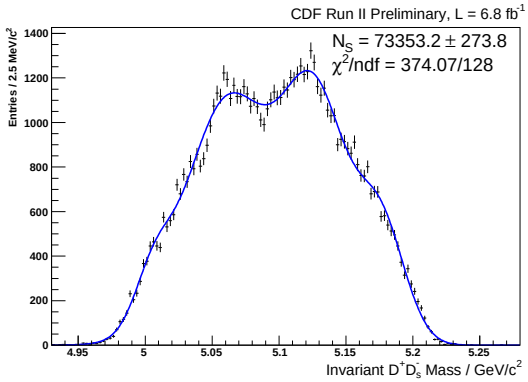
4.3.2 Full Fit Function

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

$$\begin{aligned} F^j(\vec{b}^j) \;= -2 \sum_{k=1}^{n} \ln \; & [N_1^j P_{sig}^j(m_k \mid \vec{b}_{sig}^j) + N_2^j P_{p1}^j(m_k \mid \vec{b}_{p1}^j) + N_3^j P_{p2}^j(m_k \mid \vec{b}_{p2}^j) \\ & + N_4^j P_{p3}^j(m_k \mid \vec{b}_{p3}^j) + N_5^i P_{bg}^j(m_k \mid \vec{b}_{bg}^j)] \\ & + 2 \sum_{l=1}^{5} N_l^j, \end{aligned} \tag{91}$$

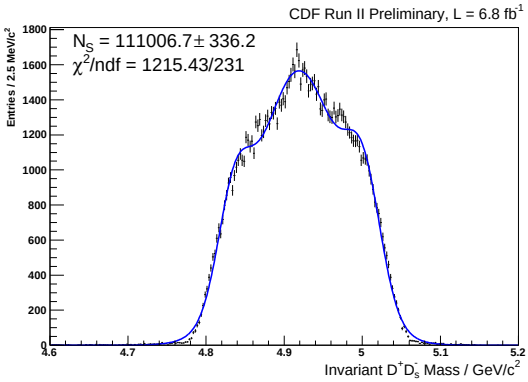


(a) $B_d^0 \to D^+ D_s^{*-} \to K^-\pi^+\pi^+\phi\pi^-$ template.

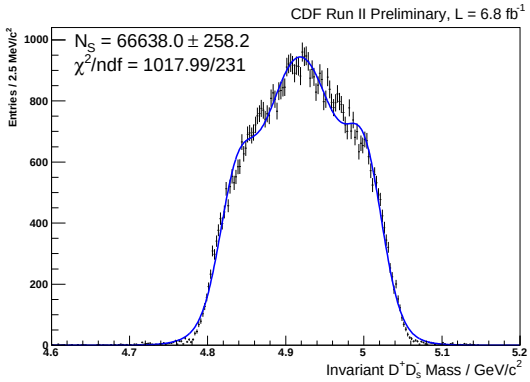


(b) $B_d^0 \to D^+ D_s^{*-} \to K^-\pi^+\pi^+K^{*0}K^-$ template.

Figure 24: 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 25: 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).

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

$$\begin{aligned} N_1^j &= N^j_{B_d^0 \to D^+ D_s^-} \\ N_2^j &= N^j_{B_d^0 \to D^{*+} D_s^-} \\ N_3^j &= N^j_{B_d^0 \to D^+ D_s^{*-}} \\ N_4^j &= N^j_{B_d^0 \to D^{*+} D_s^{*-}} \\ N_5^j &= N^j_{bkg} \end{aligned} \tag{92}$$

(93)

4.3.3 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 (72):

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

$N^{tot}_{B_d^0}$ 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:

$$\begin{aligned} N^j_{B_d^0 \to D^{*+} D_s^-} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^{*+} D_s^-) \mathcal{B}(D^{*+} \to D^+ X) \\ & \times \mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^j_{B_d^0 \to D^{*+} D_s^-} \end{aligned} \tag{94}$$

$$\begin{aligned} N^j_{B_d^0 \to D^+ D_s^{*-}} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^+ D_s^{*-}) \\ & \times \mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^j_{B_d^0 \to D^+ D_s^{*-}} \end{aligned} \tag{95}$$

$$\begin{aligned} N^j_{B_d^0 \to D^{*+} D_s^{*-}} =& N^{tot}_{B_d^0} \mathcal{B}(B_d^0 \to D^{*+} D_s^{*-}) \mathcal{B}(D^{*+} \to D^+ X) \\ & \times \mathcal{B}\begin{pmatrix} D_s^+ \to \phi^\dagger \pi^+ \\ D_s^+ \to K^{*0\dagger} K^+ \end{pmatrix} \mathcal{B}(D^+ \to K^- \pi^+ \pi^+) \epsilon^j_{B_d^0 \to D^{*+} D_s^{*-}} \end{aligned} \tag{96}$$

The parameterizations (72) and (94) through (96) 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 4.2.6) and the semi-inclusive branching fraction ratio $f_{D_s^{(*)} D_s^{(*)}}$ (Section 4.2.7).

4.4 Extraction of Branching Fractions

In order to extract the observables of interest simultaneously from all decay channels, the negative log likelihood functions (69) and (91) are combined to give one joint likelihood that is minimized. The parameterizations of the signal yields in (69) 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 $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$ enters twice, once as correctly reconstructed signal in the invariant spectrum of the normalization channel, and once as background (mis-reconstructed signal) in $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$. The simultaneous fitting approach and the consistent parameterizations of both correctly and falsely reconstructed signal ensure that existing 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 analogous argumentation applies when it comes to external parameters like intermediate or final state branching fractions that are required to parameterize signal yields.

Table 15 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, $N_{B_0^{tot}}$, is globally shared among all decay modes. By equations (72), (74), (76), (77), (80 - 83), and (94 - 96) $N_{B_0^{tot}}$ 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 $f_{D_s D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, 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 $\mathcal{B}(B_d^0 \to D^{*+} D_s^{*-})$ [1] in later systematic studies.

$B_s^0 \to D_s^+ D_s^-$ $\to \phi^\dagger \pi^+ \phi^\dagger \pi^-$	$B_s^0 \to D_s^+ D_s^-$ $\to \phi^\dagger \pi^+ K^{*0\dagger} K^-$	$B_d^0 \to D^+ D_s^-$ $\to K^- \pi^+ \pi^+ \phi^\dagger \pi^-$	$B_d^0 \to D^+ D_s^-$ $\to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$	Parameter Freedom
$N_{B_0^{tot}}$				F
$f_{D_s D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, $f_{D_s^{(*)} D_s^{(*)}}$		-		F
-	$\mathcal{B}(B_d^0 \to D^{*+} D_s^-)$, $\mathcal{B}(B_d^0 \to D^+ D_s^{*-})$, $\mathcal{B}(B_d^0 \to D^{*+} D_s^{*-})$			F
-	$\mathcal{B}(D^{*+} \to D^+ X)$			N, GC
mass resolution scale		-		F
μ_{B_s}		-		GC
-		mass resolution scale		F
-		μ_{B^0}		GC
$\mathcal{B}(B_d^0 \to D^+ D_s^-)$				N, GC
-	$\mathcal{B}(D^+ \to K^- \pi^+ \pi^+)$			N, GC
$\frac{\mathcal{B}(D_s^+ \to \phi \pi^{+\dagger})}{\mathcal{B}(D_s^+ \to K^+ K^- \pi^+)}$, $\frac{\mathcal{B}(D_s^+ \to K^{*0} K^{+\dagger})}{\mathcal{B}(D_s^+ \to K^+ K^- \pi^+)}$				N
$\mathcal{B}(D_s^+ \to K^+ K^- \pi^+)$				N, GC

Table 15: 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).

Figures 26(a) - 26(f) show the significance-optimized (see Section 3.4) invariant mass spectra that have been fitted using the simultaneous parameter estimation procedure. From the simultaneous fits the following branching fractions ratios are extracted:

$$f_{D_s D_s} = 0.183^{+0.021}_{-0.020} \tag{97}$$

$$f_{D_s^* D_s} = 0.424^{+0.047}_{-0.045} \tag{98}$$

$$f_{D_s^* D_s^*} = 0.654^{+0.073}_{-0.071} \tag{99}$$

$$f_{D_s^{(*)} D_s^{(*)}} = 1.261^{+0.096}_{-0.093} \tag{100}$$

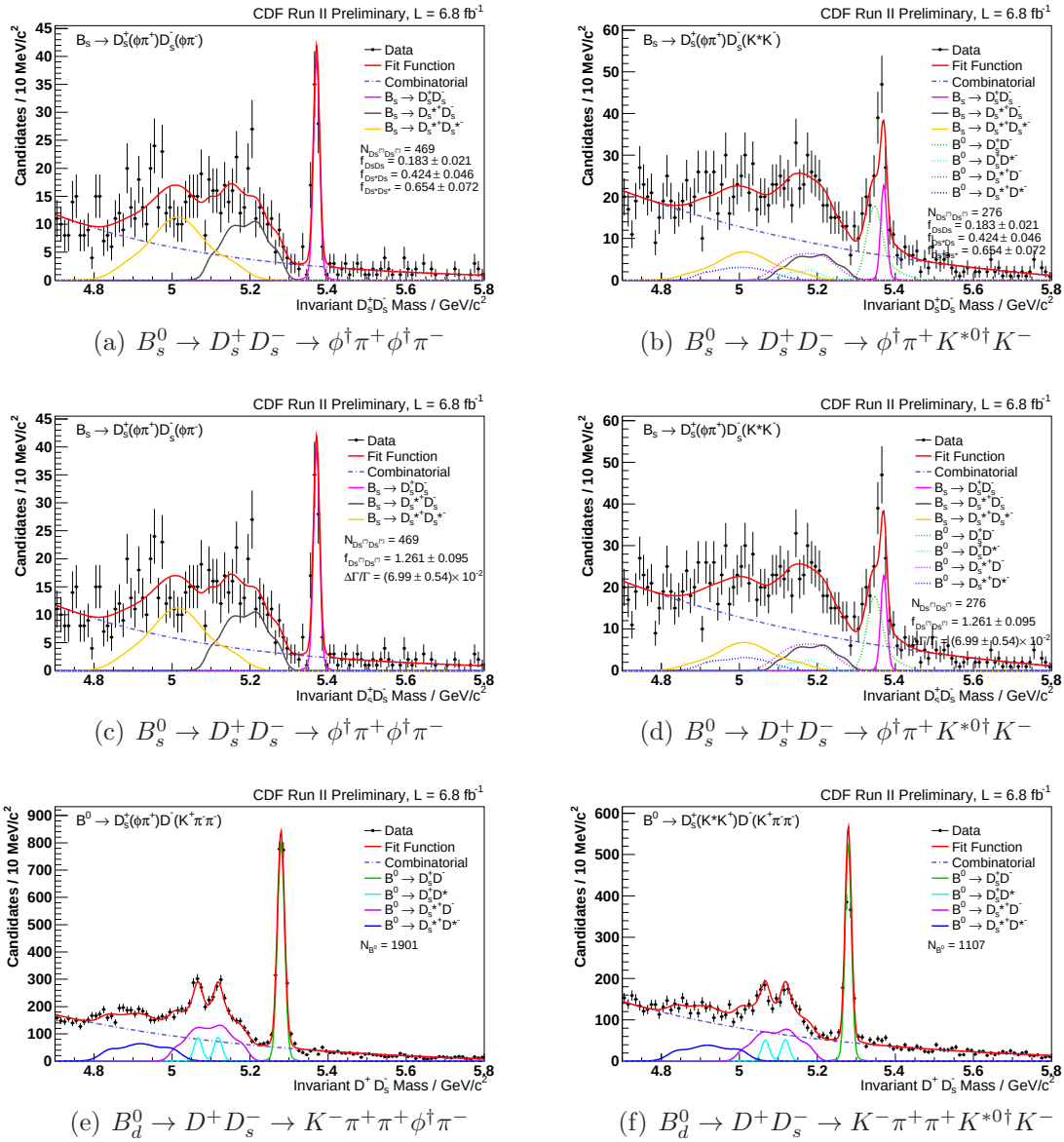
The uncertainties given are statistical only. As expected, the lowermost value 99 extracted from a separate simultaneous fit using an alternate signal parameterization (Section 4.2.7) is identical to the sum of the exclusive branching fraction ratios, (97) through (99).

The matrix of correlation coefficients reads:

$$\begin{array}{c|ccc} & f_{D_s D_s} & f_{D_s^* D_s} & f_{D_s^* D_s^*} \\ f_{D_s D_s} & 1 & 0.095 & 0.104 \\ f_{D_s^* D_s} & 0.095 & 1 & 0.04 \\ f_{D_s^* D_s^*} & 0.104 & 0.04 & 1 \end{array} \tag{101}$$

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

As indicated by Table 15, 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



(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 26: 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 $f_{D_s D_s}$, $f_{D_s^* D_s}$, 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
$f_{D_s D_s}$	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 16: 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.

16 the resulting values are tabulated and compared to the current world average values extracted from Ref. [1]. 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.

4.5 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 (7) 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 (53). 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 oder 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 (56). 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 27) 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 preparative 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 [31], 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

$$|H_+|^2 + |H_0|^2 + |H_-|^2 = 1 \tag{102}$$

and equation (53):

$$H_+^{(1)} = H_-^{(2)} = \frac{2\sqrt{2f_{CP-}} + \sqrt{8f_{CP-} - 8(-1 + |H_0|^2 + 2f_{CP-})}}{4} \tag{103}$$

$$H_-^{(1)} = H_+^{(2)} = \frac{2\sqrt{2f_{CP-}} - \sqrt{8f_{CP-} - 8(-1 + |H_0|^2 + 2f_{CP-})}}{4} \tag{104}$$

The Monte Carlo distributions shown for the cases $f_{CP-} = 0.00, 0.08, 0.16, 0.24$ (Figure 28) clearly indicate that for a fixed longitudinal polarization both the angular distributions and the invariant mass line shapes of partially reconstructed $B_s^0 \to D_s^{*+} D_s^{*-}$

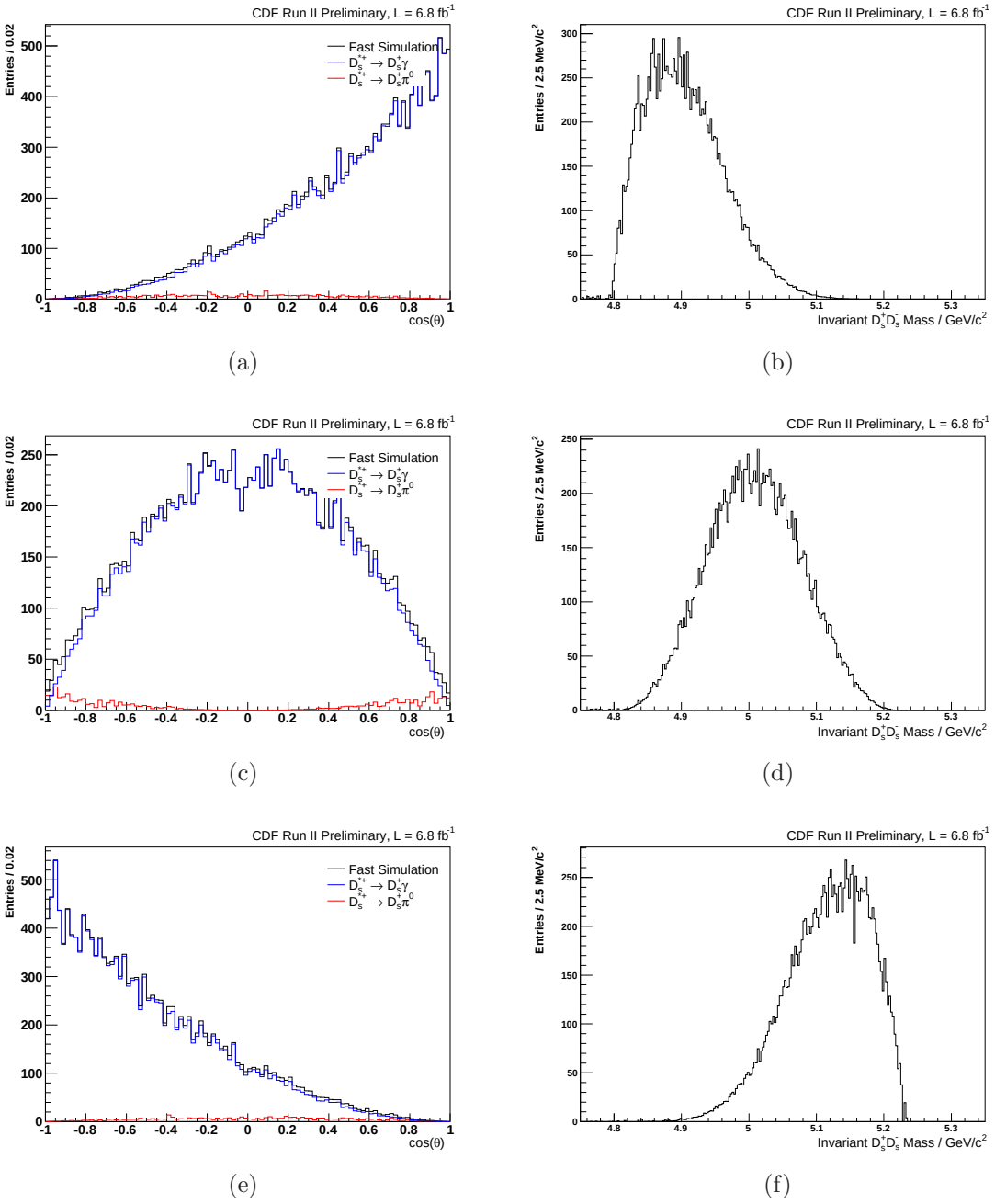


Figure 27: Monte Carlo truth $\cos(\theta)$ (left column) and resulting invariant mass (right column) distributions of $B_s^0 \to D_s^{*+} D_s^{*-}$ decays for the extreme helicity amplitude configurations $\{H_+,\ H_0,\ H_-\} = \{1,\ 0,\ 0\}$ (a,b), $\{0,\ 1,\ 0\}$ (b,c), and $\{0,\ 0,\ 1\}$ (e,f) obtained from fast EvtGen simulations.

events are completely insensitive to changes in H_+ and H_-, and thus to f_{CP-}. Since the Monte Carlo templates are all 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.

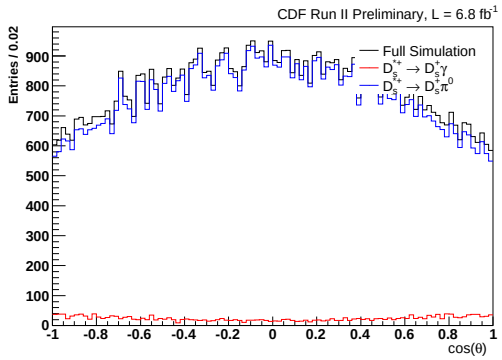
We may speculate that these findings can be explained by the general form of the decay amplitude (56), the structure of the Wigner rotation functions, and the parity-conserving nature of the electro-magnetic interaction. In decays of $D_s^{*+} \to D_s^+\gamma$ this causes the photon to have two preferred decay directions 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 29 shows the resulting $\cos(\theta)$ 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 (56), under scrutiny. In fact, by valuable consultations with one of the EvtGen authors [42] 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 30 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.

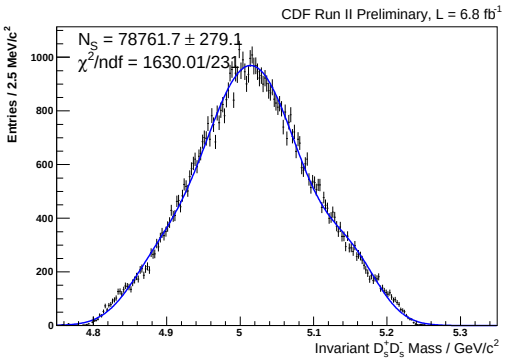
5 Systematic Studies

5.1 Intermediate and Final State Branching Fractions

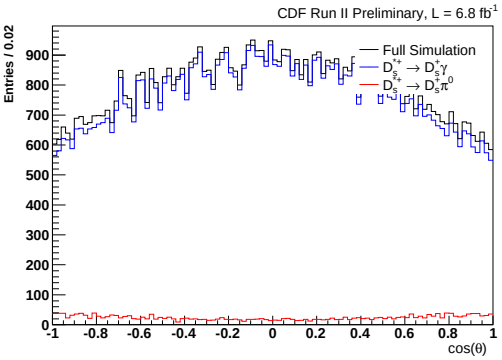
The results presented in Section 4.4 were obtained with some of the external parameters (intermediate and final state branching fractions $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$, $\mathcal{B}(B_d^0 \to D^+D_s^-)$, $\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$, and $\mathcal{B}(D^{*+} \to D^+X)$) kept fixed to their world average values [1]. For this reason, the uncertainties quoted are of statistical nature only. In



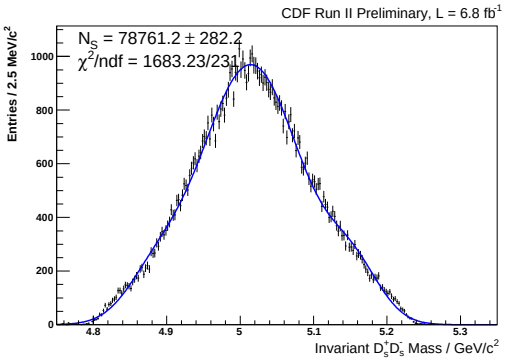
(a) $f_{CP-} = 0$: $H_+ = H_- = 0.4904$



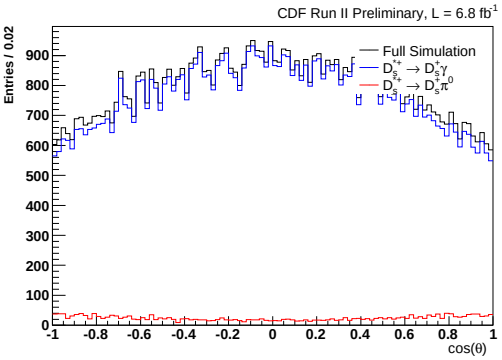
(b) $f_{CP-} = 0$: $H_+ = H_- = 0.4904$



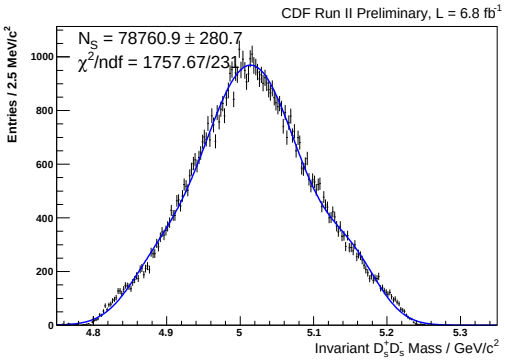
(c) $f_{CP-} = 0.08$: $H_\pm = 0.6478$, $H_\mp = 0.2478$



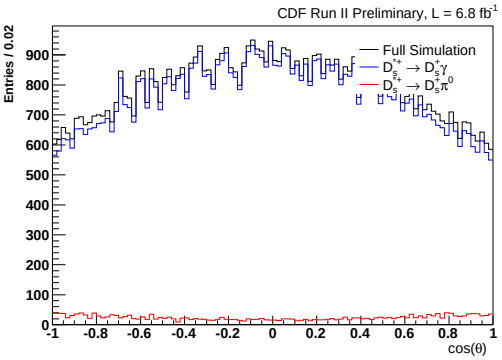
(d) $f_{CP-} = 0.08$: $H_\pm = 0.6478$, $H_\mp = 0.2478$



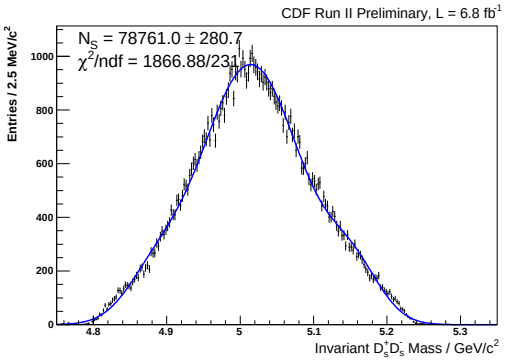
(e) $f_{CP-} = 0.16$: $H_\pm = 0.6835$, $H_\mp = 0.1172$



(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 28: Monte Carlo truth $\cos(\theta)$ (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.

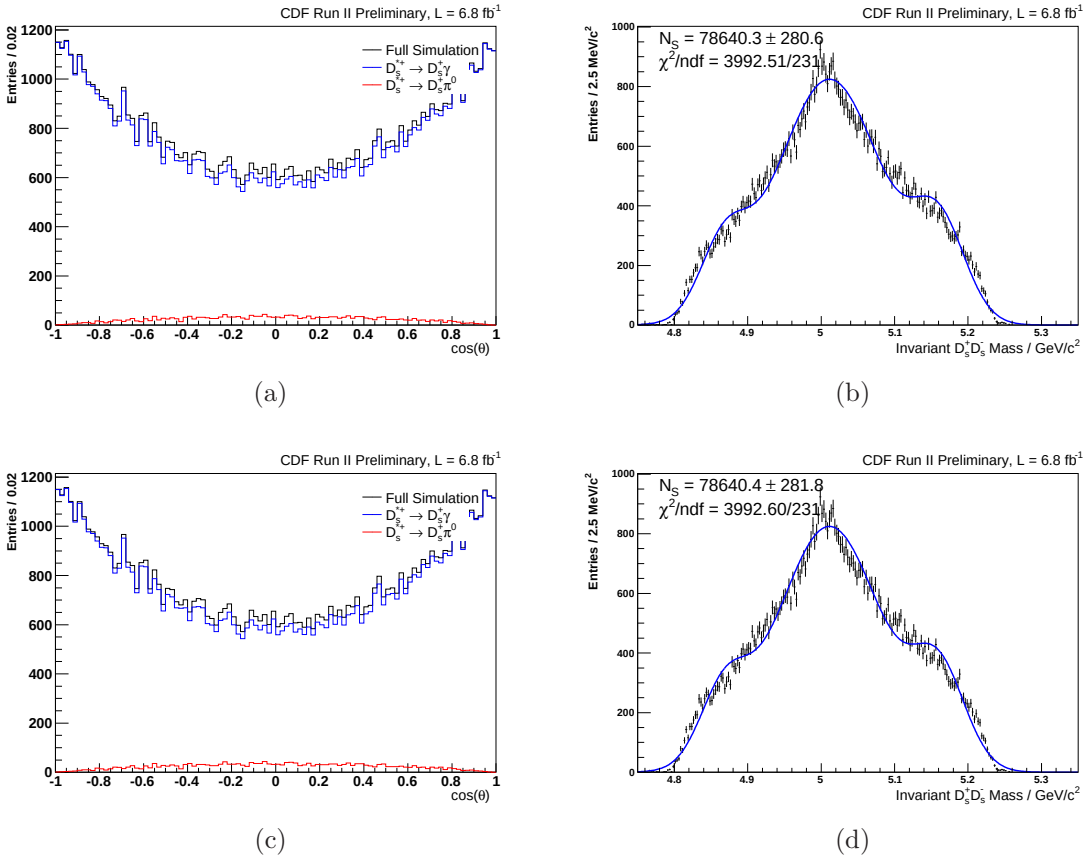


Figure 29: Monte Carlo truth $\cos(\theta)$ (left column) and resulting invariant mass (right column) $B_s^0 \to D_s^{*+} D_s^{*-}$ distributions for the extreme helicity amplitude configurations $\{H_+,\, H_0,\, H_-\} = \{1,\, 0,\, 0\}$ (a,b) and $\{0,\, 0,\, 1\}$ (c,d) obtained by weighting phase space Monte Carlo (full simulation). In contrast to direct EvtGen simulations the distributions are all symmetric.

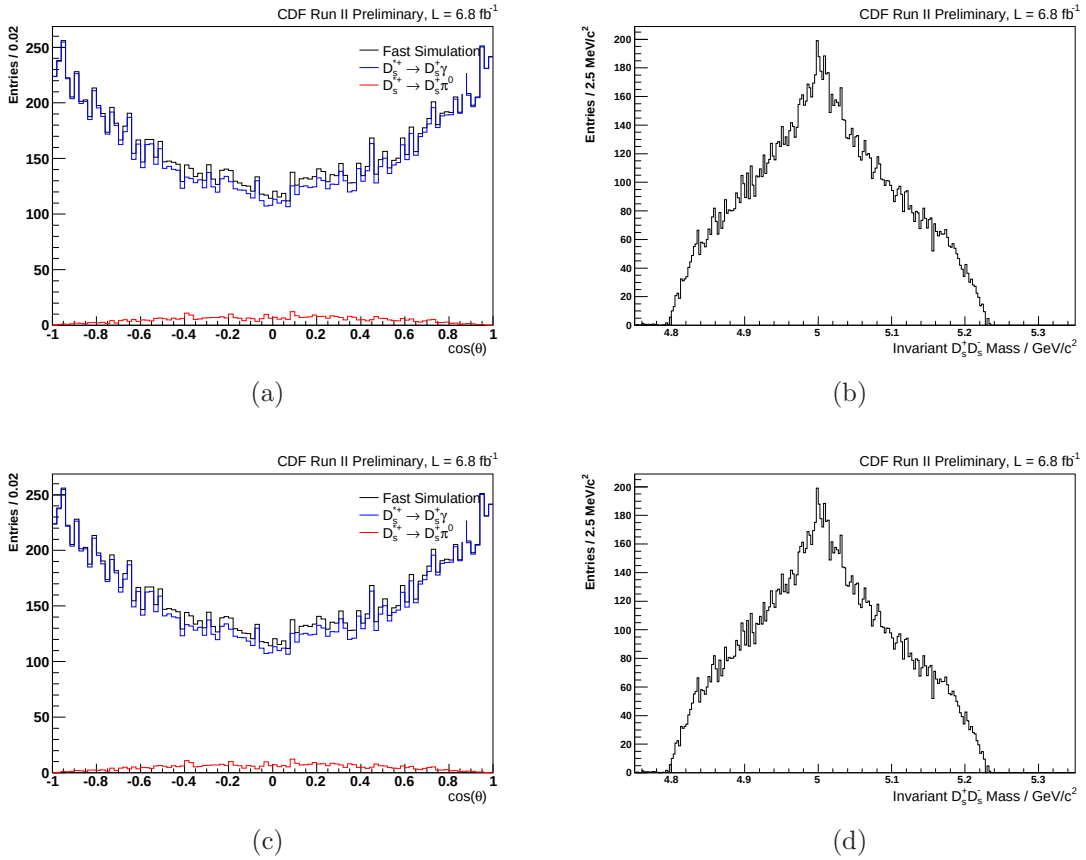


Figure 30: Monte Carlo truth $\cos(\theta)$ (left column) and resulting invariant mass (right column) $B_s^0 \to D_s^{*+} D_s^{*-}$ distributions for the extreme helicity amplitude configurations $\{H_+,\ H_0,\ H_-\} = \{1,\ 0,\ 0\}$ (a,b) and $\{0,\ 0,\ 1\}$ (c,d) obtained from fast simulations after fixing the programming error in the responsible EvtGen module.

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

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. The systematic uncertainties assigned to $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ due to uncertainties of intermediate and final state branching fraction are summarized in Table 17. 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 4.2.6 and 4.3.3): $\mathcal{B}(B_d^0 \to D^+D_s^-)$ is a common factor to all relevant signal yield parameterizations, any variation thus cancels out.

5.2 Reconstruction and Selection

5.2.1 Two-Track Trigger Correction

Recording of B meson decay candidates of the fully hadronic meson decays studied in this analysis is triggered by an on-line run 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 2.7) that the proportions of these sub-samples are not well reflected in simulated data. To provide a more realistic description of real data, all the Monte Carlo samples have been re-weighted with the weights calculated from a comparison of real and simulated data. Since the branching fractions if 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 $f_{D_s^{(*)}D_s^{(*)}}$ is repeated, however this time keeping the (poorly matching) proportions of

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 18: 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).

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 18).

5.2.2 $K \rightleftarrows \pi$ Swap in $D_{(s)}$ Reconstruction

We consider several possibilities of double-mis-identification of kaons and pions, that is, reconstructing a true kaon as a pion and vice versa, faking a true event whenever a given candidate fulfills the applied mass window requirements.

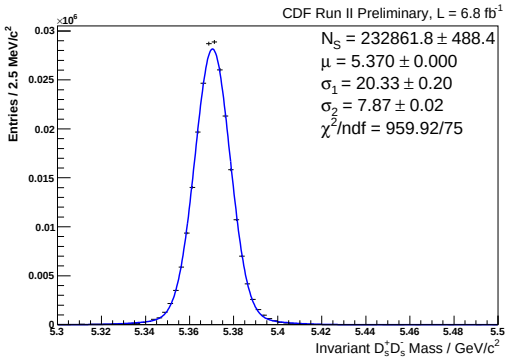
For $D_s^+ \to K^{*0\dagger}K^+$ the $K^{*0\dagger}$ resonance is reconstructed from a charged kaon and a pion of opposite charge. In principle, a reconstructed $K^-\pi^+$ pair could in reality have been a π^-K^+ combination in case of $K \rightleftarrows \pi$ mis-identification. However, in the considered scenario this is not possible since the K^{*0} flavor[6] and the daughter charges are fixed by the charge of the D_s mother particle.[7] This was checked by swapping the kaon and pions mass assignments when generating flat ntuples from reconstructed Monte Carlo Bstntuples. When placing the same charge requirements also applied in the pre-selection of data, exactly zero events remain.

Similar considerations hold for $D^+ \to K^-\pi^+\pi^+$: The kaon cannot be interchanged with an oppositely charged pion, which themselves are always of equal charge. This is because for a given D meson charge no other daughter charge combination is possible. This was again checked by swapping the kaon with any of two the pions when skimming Monte Carlo, leading to zero reconstructed events.

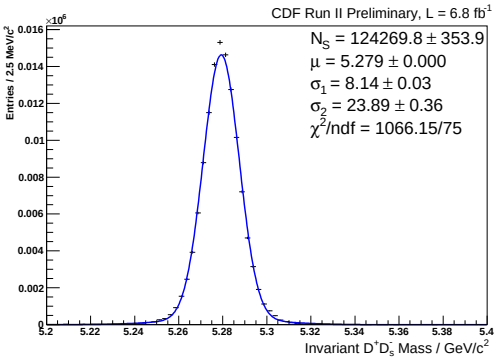
In reconstruction of $D_s^+ \to \overline{K^{*0}}^\dagger K^+$ with $\overline{K^{*0}}^\dagger \to K^-\pi^+$, the π^+ originating from the $\overline{K^{*0}}^\dagger$ might be interchanged with the K^+ directly coming from the D_s^+ mother particle. The same considerations apply to $D_s^+ \to \phi^\dagger\pi^+$, where π^+ coming from the D_s^+ could be mis-identified as the K^+ coming from the $\phi^\dagger$, and vice versa. We expect

[6] Well, correctly speaking the decay actually is $D_s^+ \to \overline{K^{*0}}(\to K^-\pi^+)K^+$

[7] The situation would be different in the decay of $B_d^0 \to J/\psi K^{*0}$ for instance. Here, both $K^{*0} \to K^+\pi^-$ and $\overline{K^{*0}} \to K^-\pi^+$ are allowed, potentially leading to a self-cross-feed in the case of $K \rightleftarrows \pi$ mis-reconstruction.



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



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

Figure 31: Signal templates for the decays $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ (a) and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$ (b) obtained by fitting Monte Carlo that include the cases of occasional $\pi^+ \rightleftarrows K^+$ mis-identifications in $D_s^+ \to K^{*0\dagger} K^+$. To be compared with 18(b) and 22(b), respectively.

Quantity	**Central**	$K \rightleftarrows \pi$	**Assigned Uncertainty**
$f_{D_s D_s}$	0.184	0.184	0.000
$f_{D_s^* D_s}$	0.421	0.422	0.001
$f_{D_s^* D_s^*}$	0.654	0.656	0.002
$f_{D_s^{(*)} D_s^{(*)}}$	1.259	1.263	0.004

Table 19: Assigned systematic uncertainties (fourth column) due to possible $K^+ \rightleftarrows \pi^+$ mis-identification in $D_s^+ \to K^{*0\dagger} K^+$ with $K^{*0\dagger} \to K^- \pi^+$, calculated from results determined in the swap scenario (third column) and the default scenario (second column).

the effect of the latter case to be rather small due to the very narrow mass band applied in the reconstruction of $\phi^\dagger \to K^+ K^-$.

First of all, checks were performed for the $D_s^+ \to \overline{K^{*0\dagger}} K^+$ case by swapping $K^+ \rightleftarrows \pi^+$ in Monte Carlo. For the given selection, both for fully reconstructed $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ K^{*0\dagger} K^-$ and $B_d^0 \to D^+ D_s^- \to K^- \pi^+ \pi^+ K^{*0\dagger} K^-$ the yields of fitted Monte Carlo are larger by about 1% when allowing for $K^+ \rightleftarrows \pi^+$ swaps (compare Figure 31(a) with Figure 18(b), and Figure 31(b) with Figure 22(b), respectively). Using the Monte Carlo samples containing the swap cases (for all signal-like components), we repeat the simultaneous fits for $f_{D_s D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, and $f_{D_s^{(*)} D_s^{(*)}}$, and assign the deviations with respect to central values as systematic uncertainties (Table 19). Since these deviations are compatible with zero (largest deviation is $0.04/\sigma_{stat}$), we assume the swap in the case of $D_s^+ \to \phi^\dagger \pi^+$ as negligible as well.

We conclude that double-mis-identification of kaons and pions in the studied decays has marginal impact on final results.

Quantity	$p_T > 350$ MeV/c	$p_T > 400$ MeV/c	Deviation	Deviation/$\sigma^{corr}_{A,B}$
$f_{D_sD_s}$	0.184	0.181	0.003	0.7
$f_{D_s^*D_s}$	0.421	0.419	0.002	0.2
$f_{D_s^*D_s^*}$	0.654	0.635	0.019	1.3
$f_{D_s^{(*)}D_s^{(*)}}$	1.259	1.235	0.024	1.3

Table 20: Systematic checks on the track transverse momentum requirement. Branching fraction ratios obtained in the $p_T > 400$ MeV/c pre-cut scenario (third column) are compared with the default cut scenario (second column). Absolute deviations are shown in the fourth column. The last column shows that the deviations are compatible with being of statistical nature, though the two data samples the branching fraction ratios were extracted from are highly correlated.

5.2.3 Track p_T Cuts

This analysis used a lower transverse momentum threshold of 350 MeV/c per track for the pre-selection of candidates. It was suggested that efficiency estimation using Monte Carlo may not be reliable below 400 MeV/c. To study this issue we increase the track p_T requirement to the latter value (both for data and Monte Carlo) and repeat the standard simultaneous fit for the branching fractions. Table 20 shows the absolute deviations of the fitted values with respect to the central values. We now check if there is a systematic shift among the two scenarios, or if the difference in results is due to statistics only.

Generally speaking, 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

$$\Delta R_{A,B} = n\sigma_{A,B}\sqrt{\frac{s_A}{s_B} - 1} = n\sigma^{corr}_{A,B} \tag{105}$$

Given the total candidate yields in the studied data samples of $s_A = 24,760$ ($p_T > 350$ MeV/c) and $s_B = 23,707$ ($p_T > 400$ MeV/c), according to Table 20 the deviations are compatible within 1.3 $\sigma^{corr}_{A,B}$ in the worst case. Therefore, we see no hints to a systematic shift in results due to the chosen lower transverse momentum threshold of 350 MeV/c. No systematic uncertainties are applied due to the track p_T cuts.

5.2.4 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 by the cancellation of the major fraction of acceptance and selection related effects. On the other hand, a given candidate might be assigned to more than one exclusive decay channel during reconstruction. As a consequence, this candidate might populate the invariant

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

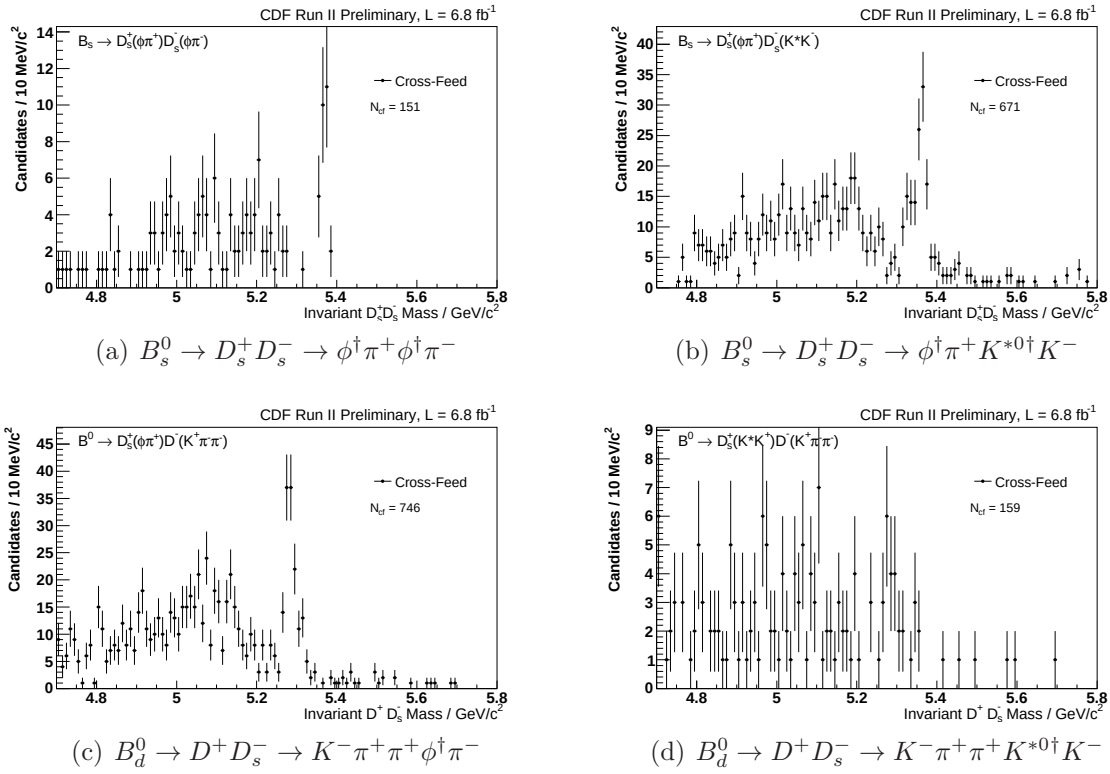
mass spectra of more than one channel, a fact that might spoil the simultaneous fitting approach. If the fraction of multiple candidates among the different decay channels is significant the distributions entering the simultaneous fit cannot be treated as statistically independent, since the invariant 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,

$$pdf(m_1, m_2) \neq pdf(m_1) \times pdf(m_2), \tag{106}$$

where equation (106) 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 21 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 32. Table 21 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 32, 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: The cross-feed due to false reconstruction has 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 placing a veto on the 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 (Table 22) indicate that the deviations with respect to the standard scenario (no multiple candidate veto) are below 0.2 statistical standard deviations. Given 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 32: 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.

Quantity	No Cross-Feed Veto		
	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 22: 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.

overall precision of this measurement the deviation is thus negligible and could in principle be ignored. Nevertheless, it is more correct to perfrom the simultaneous fit on data samples where cross-correlations are eliminated. Therefore, the branching fraction ratios obtained from the veto-fit will be used as central values. In the following we will show that it is, however, still reasonable to perform all the systematic studies in the non-veto scenario.

While the central values can be unaffected, mass correlations inside the total pdf might still introduce biases to the estimation of systematic uncertainties. To address this concern, one particular variation of the $D_s^+ \to K^+K^-\pi^+$ Dalitz model[8] is considered that introduces 1σ 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 22, 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 significant bias to the analysis results obtained from a joint likelihood fit, neither in terms of central values, nor in terms of systematics. 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 throughout this systematics chapter.

[8]The Dalitz model represents one of the leading sources of the overall systematic uncertainty. See Section 5.3.2 for details.

5.3 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 described in Section 5.1. 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.

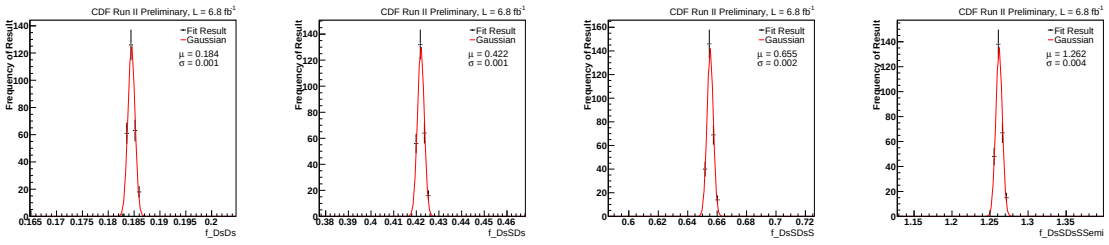
5.3.1 B and D Meson Lifetimes

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 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 a differing mean decay length $c\tau$. The per-event weights are calculated from the Monte Carlo truth information for the proper decay length ct, using relation (11)

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

The B_s^0 and B_d^0 Monte Carlo samples used in this analysis were generated with mean decay lengths of $c\tau_{B_s^0} = 441$ μm ($\tau_{B_s^0} = 1.471$ ps) and $c\tau_{B_d^0} = 458.7$ μm ($\tau = 1.530$ 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, the B_s^0 meson gets assigned the mean B_s^0 decay length (plus 1σ uncertainty), while the B_d^0 decay length is varied downwards by 1σ.



(a) (b) (c) (d)

Figure 33: 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.

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**
$f_{D_s D_s}$	$+0.001$	-0.002	0.001	$^{+0.001}_{-0.002}$
$f_{D_s^* D_s}$	$+0.002$	-0.004	0.001	$^{+0.002}_{-0.004}$
$f_{D_s^* D_s^*}$	$+0.002$	-0.006	0.002	$^{+0.003}_{-0.006}$
$f_{D_s^{(*)} D_s^{(*)}}$	$+0.005$	-0.011	0.004	$^{+0.006}_{-0.012}$

Table 23: Fluctuations of the ratios of branching fractions $f_{D_s D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, and $f_{D_s^{(*)} D_s^{(*)}}$ induced by 1 sigma variations of the B meson decay lengths (second and third column; in the latter case in addition the mean $c\tau_{B_s^0}$ lifetime was assigned to B_s^0), and by randomly varying the D meson lifetimes in 250 trial fits (fourth column). The fifth column quotes the combined systematic uncertainties with the individual uncertainties added in quadrature.

This procedure should cover any systematic effects related to B meson kinematics. By replacing the lifetime of the shorter-living light eigenstate B_s^L by the mean B_s^0 lifetime (which does virtually coincide with the B_d^0 lifetime), the decay width difference $\Delta\Gamma_s$ is allowed to vanish. 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 23 summarizes the variations in the branching fraction ratios observed in the two scenarios. The down- and upward fluctuations are assigned as asymmetric systematic uncertainties, that are fairly low compared to statistical and other systematic uncertainties.

The studied decay chains contain charged D and D_s mesons as long-living intermediate states. To study effects of variations the D and D_s mean decay lengths are allowed to float within published Gaussian uncertainties. After re-weighting Monte Carlo using these random values, the standard likelihood fits to data are repeated. Figure 33 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 $f_{D_s D_s}$, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, 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.

5.3.2 $D_s^+ \to K^+K^-\pi^+$ Dalitz Model

As described in Sections 2.4 and 2.6, decays of $D_s^+ \to K^+K^-\pi^+$ are simulated according to the Dalitz model using parameters (magnitudes, phases) measured by the CLEO collaboration [16]. The model parameters are afflicted with uncertainties though (Table 5). 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 $\phi(1020)$ and $K^{*0}(892)$ mass bands undergo changes (cf. equations (32) and (33)). This is a direct effect of any change in the Dalitz parameter configuration. Secondly, as the proportions of the resonances considered in the $D_s^+ \to K^+K^-\pi^+$ decay model can change, due to (second-order) acceptance effects a variation of the Dalitz model can influence the number of reconstructed events that enter reconstruction efficiencies. 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 2.4.5 the parameters are correlated to each other according to the parameters' covariance matrix (see Section B 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 (74), 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 (72), (76-77), (80-83, and (94-96)). For ease of discussion equation (74) 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 $f_{\phi^\dagger\pi^+}$ (30), $f_{K^{*0\dagger}K^+}$ (31) and writing out the efficiency term,

$$\begin{aligned} N^i_{B_s^0\to D_s^+D_s^-} =& N^{tot}_{B_d^0} f_{D_sD_s} \mathcal{B}(B_d^0 \to D^+D_s^-) \\ & \times f_{\phi^\dagger\pi^+}\mathcal{B}(D_s^+ \to K^+K^-\pi^+)\begin{pmatrix} f_{\phi^\dagger\pi^+} \\ f_{K^{*0\dagger}K^+} \end{pmatrix}\mathcal{B}(D_s^+ \to K^+K^-\pi^+) \\ & \times \frac{N^i_{rec}}{f_{\phi^\dagger\pi^+}\begin{pmatrix} f_{\phi^\dagger\pi^+} \\ f_{K^{*0\dagger}K^+} \end{pmatrix} N^{gen}_{KK\pi}}, \end{aligned} \tag{107}$$

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

$$W(m^2_{K^+K^-}, m^2_{K^-\pi^+}) = \frac{\left|\tilde{\mathcal{M}}(m^2_{K^+K^-}, m^2_{K^-\pi^+})\right|^2}{\left|\mathcal{M}(m^2_{K^+K^-}, m^2_{K^-\pi^+})\right|^2}. \tag{108}$$

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 2.4.5. 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 (107) 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 (107) 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 $f_{D_sD_s}$, $f_{D^*_sD_s}$, $f_{D^*_sD^*_s}$, and $f_{D^{(*)}_sD^{(*)}_s}$ is hence the combined result of direct and second-order effects due to changes of the Dalitz model parameters. It must be emphasized though that direct and second-order (acceptance) effects had to be studied separately if external values for $f_{\phi^\dagger\pi^+}$ and $f_{K^{*0\dagger}K^+}$ differing from the ones in the efficiency term 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 2.4.5.

We perform a total of 500 simultaneous fits with Monte Carlo re-weighted according to 500 different Dalitz model scenarios, 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 $f_{D_sD_s}$, $f_{D^*_sD_s}$, $f_{D^*_sD^*_s}$, and $f_{D^{(*)}_sD^{(*)}_s}$. Table 24 lists the uncertainties derived from the distributions shown in Figure 34.

5.3.3 $D^+ \to K^-\pi^+\pi^+$ Dalitz Model

Just like $D^+_s \to K^+K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ is a hadronic three-body decay that can be well described by the Dalitz model. The model implemented in the event decay package EvtGen the used $B^0_d \to D^+D^-_s$ Monte Carlo were produced with is based on a measurement performed by the E691 experiment [43], published in 1993. Since then, the according EvtGen module has not been updated, although much newer

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 24: 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.

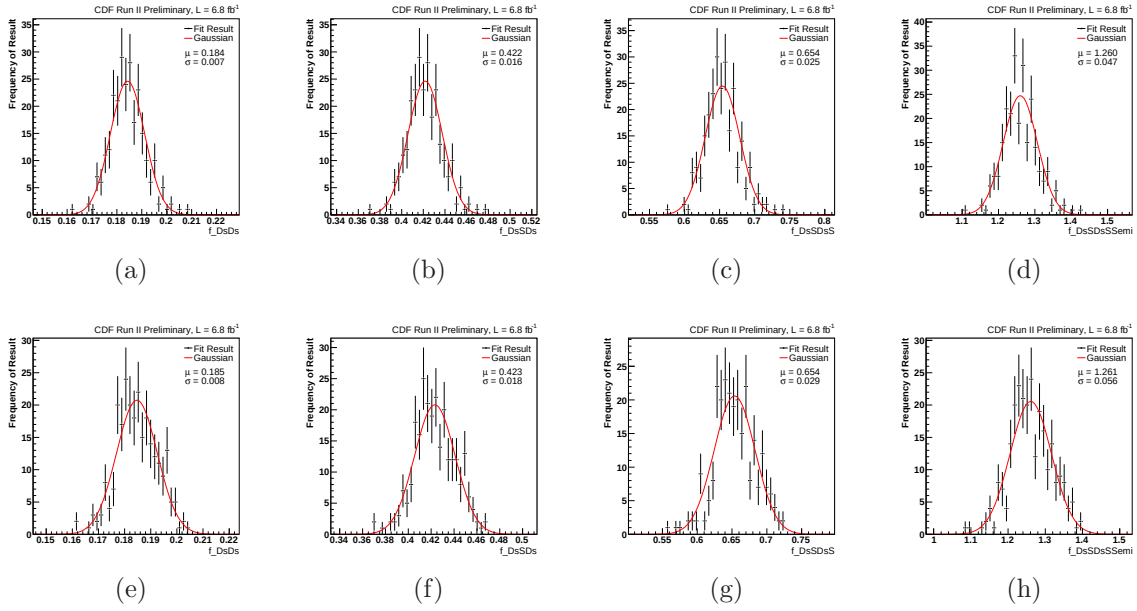
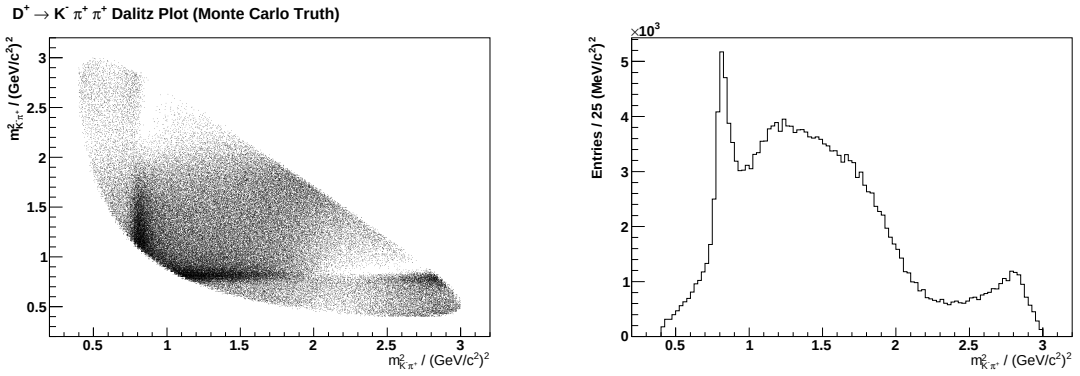


Figure 34: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ (from left to right) induced by variations of the CLEO Dalitz model parameters within correlated statistical (upper row) and uncorrelated systematic uncertainties (bottom row).



(a) (b)

Figure 35: Simulation of $D^+ \to K^-\pi^+\pi^+$ decays using the E691 model implemented in EvtGen. Full Dalitz plot (a) and projection onto $m^2_{K^-\pi^+}$ (b).

Resonance	Amplitude	Phase	Fit Fraction
Non-resonant	1	0	0.838
$K^{*0}(892)$	0.78 ± 0.02	-60 ± 3	0.170 ± 0.009
$K^{*0}(1430)$	0.53 ± 0.02	132 ± 2	0.248 ± 0.019
$K^{*0}(1680)$	0.47 ± 0.03	-51 ± 4	0.030 ± 0.004

Table 25: $D^+ \to K^-\pi^+\pi^+$ resonances and associated amplitudes, phases and fit fractions measured by E691 [43]

measurements being available, like a CLEO measurement from 2008 [44]. Figures 35(a) and 35(b) show the Dalitz plot and the projection onto $m^2_{K^-\pi^+}$ resulting from event decay simulation, detector simulation and candidate selection applied in this analysis.

As with $D_s^+ \to K^+K^-\pi^+$, when considering different Dalitz models, acceptance effects might result in changes in estimated efficiencies. Since the D^+ is reconstructed from full $D^+ \to K^-\pi^+\pi^+$ phase space we expect acceptance-induced variations in estimated efficiencies to be small. We first consider variations of the implemented model only, without taking into account newer models. If observed variations turn out to be small, there will be no need to consider newer models.

Table 25 gives an overview of the resonances and associated Dalitz parameters (amplitudes and phases) measured by E691. We use a slightly modified version of the original EvtGen code[9] to calculate squared Dalitz decay amplitudes. In Figure 36 a graphical representation of the squared amplitude distribution over the Dalitz plot is

[9]The modules EvtDDalitz and EvtResonance use particle four momenta as input parameters for the amplitude calculation. To perform Dalitz plot integrations it is more practicable to directly use the Dalitz variables $m^2_{K^-\pi^+}$, $m^2_{\pi^+K^-}$ as function input parameters. We adjusted parts of the code accordingly.

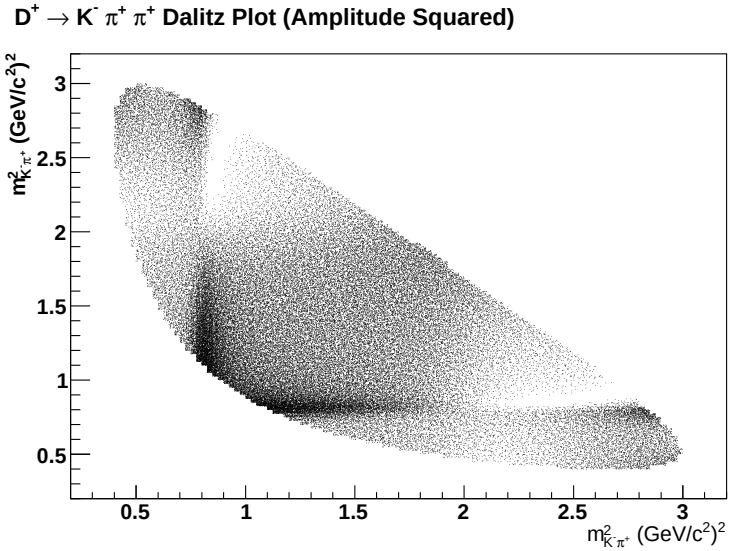


Figure 36: Graphical representation of the squared $D^+ \to K^-\pi^+\pi^+$ decay amplitudes calculated from the E691 Dalitz model.

shown. To estimate systematic uncertainties on $f_{D_s^{(*)}D_s^{(*)}}$ introduced by $D^+ \to K^-\pi^+\pi^+$ Dalitz model variations we perform similar toy parameter studies as described in the last section: $B_d^0 \to D^+D_s^-$ Monte Carlo are re-weighted using the per-event weight

$$w(m^2_{K^-\pi^+}, m^2_{\pi^+K^-}) = \frac{|\mathcal{M}'|^2}{|\mathcal{M}|^2}, \tag{109}$$

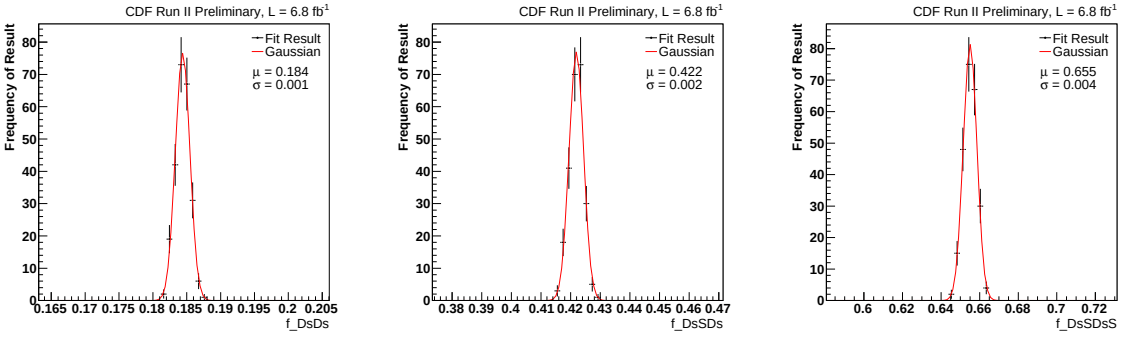
where the amplitude $\mathcal{M}$ is calculated by

$$\mathcal{M} = 1 + \sum_{R=1}^{n} c_R e^{i\theta_R} F_R^{BW}(m^2_{K^-\pi^+}, m^2_{\pi^+K^-}) D_R^{ang}(m^2_{K^-\pi^+}, m^2_{\pi^+K^-}). \tag{110}$$

Here, F_R^{BW} is a non-relativistic Breit-Wigner function and D_R^{ang} an angular factor accounting for non-isotropic emission of daughters in case of resonances which have spin not equal to zero. $\mathcal{M}'$ represents an arbitrary Dalitz model scenario (created by random Gaussians variations of the Dalitz model parameters according to the uncertainties given in Table 25), while $\mathcal{M}$ stands for the default case the Monte Carlo samples used where generated with. When calculating the efficiency using a different Dalitz model, the number of generated events in addition have to corrected by

$$W = \frac{\oint |\mathcal{M}'|^2 \, dm^2_{K^-\pi^+} dm^2_{\pi^+K^-}}{\oint |\mathcal{M}|^2 \, dm^2_{K^-\pi^+} dm^2_{\pi^+K^-}} \tag{111}$$

For ~ 250 $D^+ \to K^-\pi^+\pi^+$ Dalitz worlds efficiencies are re-estimated and the simultaneous fit for $f_{D_s^{(*)}D_s^{(*)}}$ is repeated. The systematic uncertainties, derived from the Gaussian widhts of the result distributions (Figure 37) are negligible.



(a) (b) (c)

Figure 37: Fluctuations of fit results induced by variations of the E691 $D^+ \to K^-\pi^+\pi^+$ Dalitz model.

Within this systematic study variations of the E691 Dalitz model have been considered only. Newer measurements like [44] are not covered. However, since the observed effects when varying the Dalitz model are so tiny we conclude that there's no need to consider other $D^+ \to K^-\pi^+\pi^+$ Dalitz model measurements.

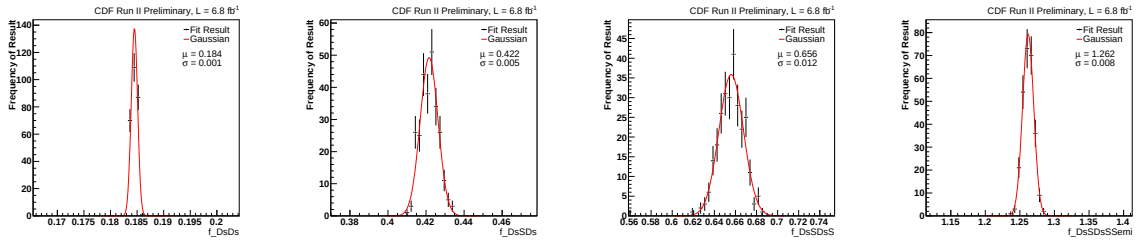
5.3.4 $B_s^0 \to D_s^{*+}D_s^{*-}$ Helicity Amplitudes

Section 4.5 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 (102). The longitudinal polarization fraction is varied by generating random numbers around the central value $f_L(B_d^0 \to D^{*+}D_s^{*-}) = 0.519$[10] [31] 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 for $B_s^0 \to D_s^{*+}D_s^{*-}$ 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 (56) 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 38) and the widths of Gaussian fits are taken as systematic uncertainties (Table 26).

[10] For the reasons outlined in Section 2.5.3 this is a reasonable estimate for $f_L(B_s^0 \to D_s^{*+}D_s^{*-})$.

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 26: 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 4.5). Compared to statistical uncertainties, systematic uncertainties originating from variations of f_L (third column) are insignificant.



(a) (b) (c) (d)

Figure 38: Systematic uncertainties of the ratios of branching fractions $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ (from left to right) 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	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 27: Systematic uncertainties due to the variation of the parameterization of signal-like components in data.

5.4 Fit

5.4.1 Signal Parameterization

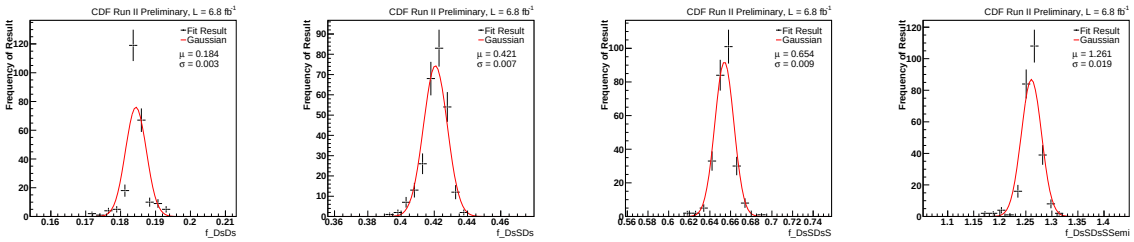
The parameterizations of the signal-like contributions (fully and partially reconstructed signal events) are determined by means of fits to the Monte Carlo samples generated for the respective components. Differences in mass resolution between data and simulation are already taken into account to a large extent by the floating width scale factor. Except for the width scale factor and the shared means of fully reconstructed signal all the shape parameters a 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 5.1) 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 27 and Figure 39 for details. Compared to statistical and the leading systematic uncertainties the assigned uncertainties are insignificant.

5.4.2 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 (97) through (99).

Since the parameterization chosen lacks any physical motivation – and cannot be verified on a clean statistically independent sample – a function having a similar slope



(a) (b) (c) (d)

Figure 39: 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 varying the parameterization of all signal-like components.

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 28: 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).

and amount of parameters may describe the background component in data equally well, but might lead to different results in terms of the ratios branching fractions. To investigate the influence of a slightly differently shaped background the standard simultaneous fit is repeated using a 2nd order polynomial function, that needs an equal amount of parameters to be described. As systematic uncertainties the full deviations from our 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 28).

5.4.3 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 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 values inside the original 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 studied decay channels. For each combination of random invariant mass distributions a simultaneous fit for $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ is carried out. In this way, 250 toy experiments are run in total, and the fit results are filled into histograms (upper row of Figure 40). 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

$$p_{trial} = \frac{x_{trial} - x_{central}}{\sigma_{x_{trial}}}. \tag{112}$$

$x_{central}$ 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 $f_{D_s^{(*)}D_s^{(*)}}$ are slightly asymmetric, the pull relation for asymmetric uncertainties $\sigma^+_{x_{trial}}$, $\sigma^-_{x_{trial}}$ has to be used:

$$\begin{aligned} p_{trial} &= \frac{x_{trial} - x_{central}}{\sigma^+_{x_{trial}}} \quad \text{for } x_{trial} \leq x_{central} \\ p_{trial} &= \frac{x_{trial} - x_{central}}{\sigma^-_{x_{trial}}} \quad \text{for } x_{trial} > x_{central} \end{aligned} \tag{113}$$

For a bias-free fit, i.e. a fit where 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. The lower row of Figure 40 shows the pull distributions for the fit 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 $2\sigma = 0.13$ the toy pull distributions are compatible with a bias-free fit. Therefore, no additional systematic uncertainty due to the fitting method is assigned.

5.5 Overview

The systematics considered throughout the foregone sections are finally summarized and combined. The individual absolute and relative uncertainties are presented in Table 29. The leading sources of uncertainty introduced by analysis techniques are the combined uncertainties of intermediate and final state branching fractions and the uncertainty due the parameterization of the $D_s^+ \to K^+K^-\pi^+$ Dalitz model.

6 Results and Discussion

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

By combining the individual and total systematic uncertainties estimated in Section 5 we may write the results of the exclusive branching fractions ratios and the semi-inclusive branching fraction ratio obtained from joint likelihood fits in the multiple

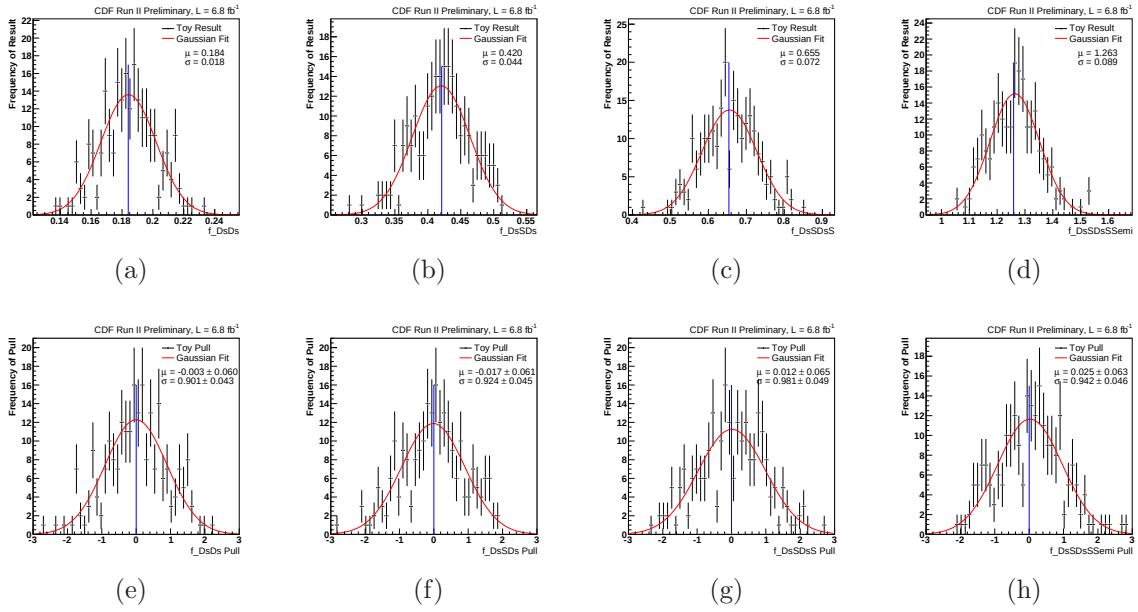


Figure 40: From left to right, distributions of the $f_{D_sD_s}$, $f_{D_s^*D_s}$, $f_{D_s^*D_s^*}$, and $f_{D_s^{(*)}D_s^{(*)}}$ results determined in 250 toy experiments (upper row), and corresponding pull distributions (bottom row).

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
$K \rightleftarrows \pi$	0.000	0.00	0.001	0.00	0.002	0.00	0.004	0.00
B, D Lifetimes	$^{+0.001}_{-0.002}$	$^{+0.01}_{-0.01}$	$^{+0.002}_{-0.004}$	$^{+0.00}_{-0.01}$	$^{+0.003}_{-0.006}$	$^{+0.00}_{-0.01}$	$^{+0.006}_{-0.012}$	$^{+0.00}_{-0.01}$
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	± 0.09	± 0.035	± 0.08	± 0.065	0.10	$^{+0.111}_{-0.112}$	0.09

Table 29: 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 $\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-})$.

candidate veto scenario (see Section 5.2.4):

$$f_{D_sD_s} = 0.183^{+0.021}_{-0.020}(stat) \pm 0.017(sys) \quad (114)$$

$$f_{D_s^*D_s} = 0.424^{+0.047}_{-0.045}(stat) \pm 0.035(sys) \quad (115)$$

$$f_{D_s^*D_s^*} = 0.654^{+0.073}_{-0.071}(stat) \pm 0.065(sys) \quad (116)$$

$$f_{D_s^{(*)}D_s^{(*)}} = 1.261^{+0.096}_{-0.093}(stat)^{+0.111}_{-0.112}(sys) \quad (117)$$

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 (114) through (117) absolute branching fractions are calculated by inserting the current world average values $\mathcal{B}(B_d^0 \to D^+D_s^-) = (7.2 \pm 0.8) \times 10^{-3}$ [1] and $f_s/f_d = 0.269 \pm 0.033$ into equations (75, 78, 79):

$$\mathcal{B}(B_s^0 \to D_s^+D_s^-) = (0.49 \pm 0.06(stat) \pm 0.05(sys) \pm 0.08(norm))\% \quad (118)$$

$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^-) = (1.13 \pm 0.12(stat) \pm 0.09(sys) \pm 0.19(norm))\% \quad (119)$$

$$\mathcal{B}(B_s^0 \to D_s^{*+}D_s^{*-}) = (1.75 \pm 0.19(stat) \pm 0.17(sys) \pm 0.29(norm))\% \quad (120)$$

$$\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (3.38 \pm 0.25(stat) \pm 0.30(sys) \pm 0.56(norm))\% \quad (121)$$

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 (118) - (121) 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 (114) - (117).

As discussed in Section 1.1, under certain theoretical assumptions $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays saturate decays to CP-even final states. Taking CP violation to be negligibly small, and using the relationship specified in equation (9),

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

the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ in the $B_s^0\bar{B}_s^0$ system can be estimated from the branching fraction of semi-inclusive $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays. Inserting the calculated branching fraction value (121) and propagating all uncertainties, one obtains:

$$\frac{\Delta\Gamma_s}{\Gamma_s} = (6.99 \pm 0.54(stat) \pm 0.64(sys) \pm 1.20(norm) \pm 0.34(theo))\% \quad (122)$$
$$= (7.08 \pm 1.42)\%$$

In the estimation of $\Delta\Gamma_s/\Gamma_s$ a theoretical uncertainty of $\pm 5\%$ has been added. This arises [4] 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 $\bar{B}_s^0$.

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

	DØ	CDF	Belle	CDF 2011
N_s	27	24	23	745
$\mathcal{B}(B_s^0 \to D_s^+ D_s^-)$ (%)	-	$1.04^{+0.35+1.1}_{-0.32-1.1}$	$1.03^{+0.39+0.26}_{-0.32-0.25}$	$0.49 \pm 0.06 \pm 0.09$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^-)$ (%)	-	-	$2.75^{+0.83+0.69}_{-0.71-0.69}$	$1.13 \pm 0.12 pm 0.21$
$\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-})$ (%)	-	-	$3.08^{+1.22+0.84}_{-1.04-0.84}$	$1.75 \pm 0.19 pm 0.34$
$\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ (%)	$3.5 \pm 1.0 \pm 1.1$	-	$6.85^{+1.53+1.89}_{-1.30-1.89}$	$3.38 \pm 0.25 \pm 0.63$
$\Delta\Gamma_s/\Gamma_s$ (%)	$7.2 \pm 2.1 \pm 2.2$	> 1.2	$14.7^{+3.6+4.4}_{-3.0-4.2}$	$6.99 \pm 0.54 \pm 1.36$

Table 31: Comparison of available results and the preliminary results presented in this document.

6.2 Statistical Significance

The statistical significances of the relative branching fraction measurements are calculated from $\sqrt{-2 \ln \log(\mathcal{L}_0^f / \mathcal{L}_{max})}$, 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 30 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.

6.3 Comparison of Results

The presented results (118) - (121) are now being embedded in the context of the current experimental status. The estimated relative decay width difference (122) is directly comparable to the values obtained by the previous analyses published in the Refs. [6, 7, 9], 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 31 gives an overview of previous results and the preliminary results contributed by the present analysis. To enhance comparability, Figure 41 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

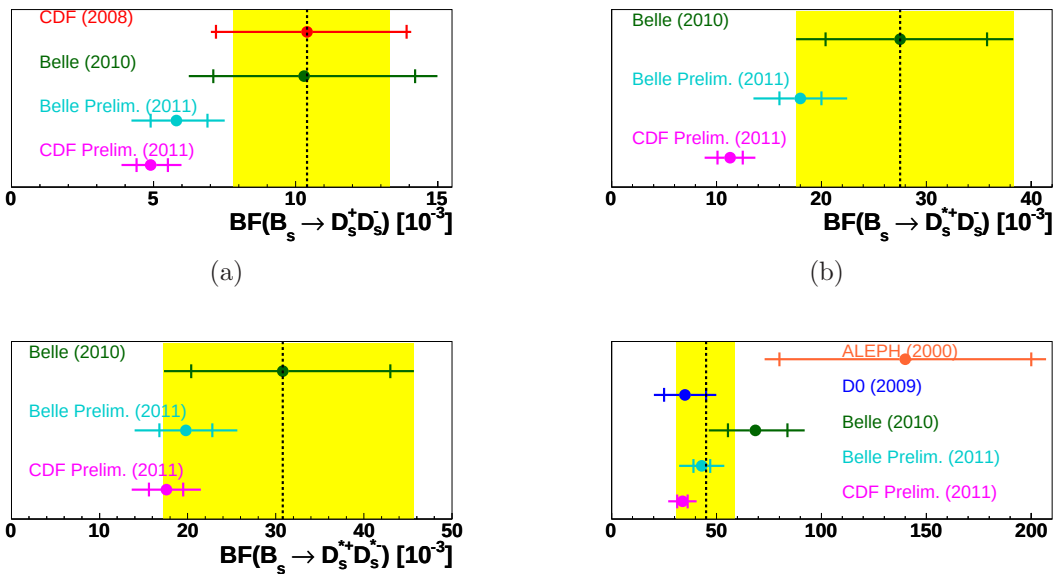


Figure 41: Comparison of $B_s^0 \to D_s^+ D_s^-$ (a), $B_s^0 \to D_s^{*+} D_s^-$ (b), $B_s^0 \to D_s^{*+} D_s^{*-}$ (c), and $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ (d) branching fraction measurements. The preliminary results *CDF Prelim. (2011)* (this analysis) and *Belle Prelim. (2011)* were not yet considered in the calculated averages (yellow bands).

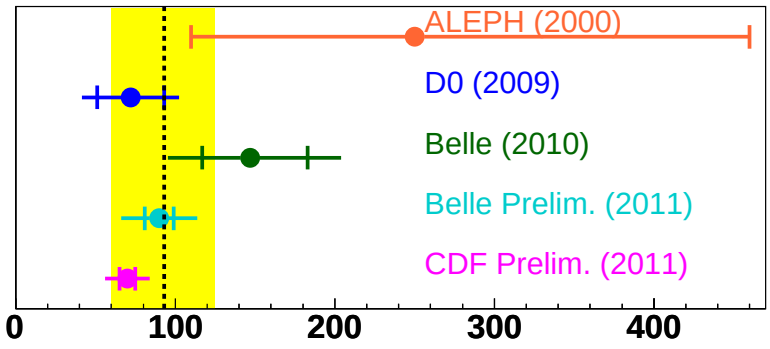


Figure 42: Comparison of results of the relative decay width difference $\Delta\Gamma_s/\Gamma_s$ estimated from $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$. The average value (yellow band) is calculated from lifetime measurements and non-preliminary $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ measurements [10].

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} [45], indicated by *Belle Prelim. (2011)* in Figure 41. 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 would certainly reduce the tension the presented results partly have with the current world averages considerably.

Even under 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 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 1.1, our estimate of $\Delta\Gamma_s/\Gamma_s$ is in good agreement with the current world average value $\Delta\Gamma_s/\Gamma_s = 9.3^{+3.2}_{3.3}$ [10] that is calculated from a variety of lifetime measurements and existing $\mathcal{B}(B_s^0 \to D_s^{(*)+} D_s^{(*)-})$ measurements. Figure 42 gives a graphical comparison of results.

We are curious if the tension with the old CDF result is method- or only data-driven. To gain confidence in the methods used throughout this analysis we repeat the simultaneous fit using CDF period 0 data only (this is the dataset available to the

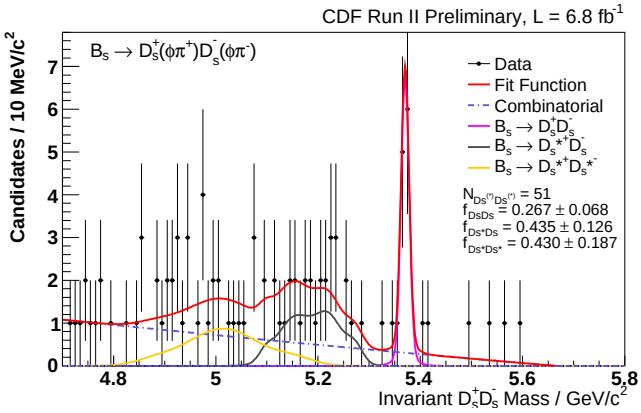


Figure 43: Fitted mass projection of $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$ using period 0 CDF data only.

former CDF analysis[11]). We obtain:

$$f_{D_s D_s} = 0.27 \pm 0.07(stat) \tag{123}$$

Figure 43 shows the mass projection of the decay channel $B_s^0 \to D_s^+ D_s^- \to \phi^\dagger \pi^+ \phi^\dagger \pi^-$. We compare (123) with the result presented in CDF Note 8085. Excluding $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\pi^+\pi^-\pi^-$, we use equation (31) in 8085 to calculate the weighted average of $f_{D_s D_s}$ out of $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^-$ and $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+ K^{*0} K^-$, where we insert yields and efficiencies as presented in 8085, but use today's external branching fractions. We calculate

$$f_{D_s D_s} = 0.26^{+0.08}_{-0.06}(stat) \tag{124}$$

This is in good agreement. 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 tension with the former CDF results seems to due to the two analyses being carried out on different data samples.

6.4 Concluding Remarks

As pointed out in [3], the theoretical assumptions made in [4] 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 any more give a good estimation of $\Delta\Gamma_s$. These concerns seem to be encouraged by recent theoretical calculations of $\Delta\Gamma_s$: According to Ref. [46] the effect of three-body modes on $\Delta\Gamma_s$ is comparable to that of two-body modes. Thus the assumption of two-body modes 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

[11] There might be slight differences due to changes in reconstruction software.

be interpreted as a lower limit to the decay width difference only. To provide a more accurate estimation of $\Delta\Gamma_s$ it would be necessary to account for 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 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 ($\sim 5\%$) [30] in $B_s^0 \to D_s^{*+} D_s^{*-}$, 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 comparable to the one predicted and measured in $B_d^0 \to D^{*+} D_s^{*-}$. As our discussions in Section 4.5 demonstrated, we are not able to confirm or rule out any CP-odd fractions lying within 0 and 24% by partial reconstruction of $B_s^0 \to D_s^{*+} D_s^{*-}$.

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 the detection of the tracks of the neutral pion and photon emitted in the decay of D_s^{*+} to determine their helicity angles in the D_s^{*+} rest frame. This however is beyond the technical possibilities of the CDF II detector. Given sufficient statistics of $\Upsilon(5S)$ data, in this respect the Belle collaboration might be able to make an important contribution, 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.

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. 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 and bias effects mostly cancel out. A promising candidate for this purpose could be the same decay mode used as normalization channel in this branching fraction analysis, $B_d^0 \to D^+ D_s^-$.

In the light of the measuring concepts introduced in Section 6 and the direct determination of the lifetime of the purely CP-odd decay $B_s^0 \to J/\psi f_0$, apart from shedding light on the CP puzzle in $B_s^0 \to D_s^{*+} D_s^{*-}$, a lifetime measurement is a very intersting option in other respects as well. Since the the decay $B_s^0 \to D_s^+ D_s^-$ represents the CP-even counterpart of $B_s^0 \to J/\psi f_0$, by measuring its lifetime both ingredients are at hand to determine the decay width difference $\Delta\Gamma_s$ from the lifetime measurements from two pure CP eigenstates.

7 Conclusions

In summary, we have 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 report first observation of the exclusive mode $B_s^0 \to D_s^{*+}D_s^{*-}$. The decay of $B_d^0 \to D^+D_s^-$ with $D^+ \to K^-\pi^+\pi^+$ 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 doing so, reconstruction efficiencies estimated from simulated data are more reliable and related uncertainties smaller 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.

Furthermore, in the course of in-depth systematic studies we have examined the feasibility of measuring of 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 however conclude that by partial reconstruction of this decay we are not able to make any inference on f_{CP-}.

Assuming the CP-even final state $D_s^{(*)+}D_s^{(*)-}$ 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

$$\mathcal{B}(B_s^0 \to D_s^{(*)+}D_s^{(*)-}) = (3.38 \pm 0.25(stat) \pm 0.30(sys) \pm 0.56(norm))\%$$

we obtain

$$\begin{aligned}\frac{\Delta\Gamma_s}{\Gamma_s} &= (6.99 \pm 0.54(stat) \pm 0.64(sys) \pm 1.20(norm) \pm 0.34(theo))\% \\ &= (6.99 \pm 1.46)\%\end{aligned}$$

This analysis used 6.8 fb^{-1} of CDF Run II data. 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. Nonetheless, a reduction of statistical uncertainty may be achieved by extending the ensemble of D_s^+ decay modes used.

Acknowledgments

We thank the technical staff of the Fermilab, the CDF collaboration and the participating institutions for their vital contributions. We owe a debt of gratitude to David Cinabro, Mikhail S. Dubrovin, Uli Nierste, Jonathan L. Rosner, Anders Ryd, and Anze Zupanc for valuable discussions.

A 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[12] 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 on $L_{xy}(P)$.
- $\chi^2_{R\phi}(P)$ is the χ^2 in the $R-\phi$ plane of the kinematical fit of the P candidate.
- 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.
- $d_0^{lts}(P)$ 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/σ_{d_0}) is the minimum of the significance of the final state particles' impact parameter .
- m(C^iC^j) is the invariant mass of two particles C^i, C^j.
- Δm(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.

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

B Dalitz Plot Parameter Toy Studies

The estimation of impacts of Dalitz model variations on the ratios of branching fractions (section 5.3.2) 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 [16]. The authors kindly provided us the covariance matrix, from which the decomposed lower left triangle matrix 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 5.

B.1 Correlated Statistical Uncertainties

B.1.1 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

B.1.2 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

B.1.3 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

B.1.4 Toy Parameter Distributions

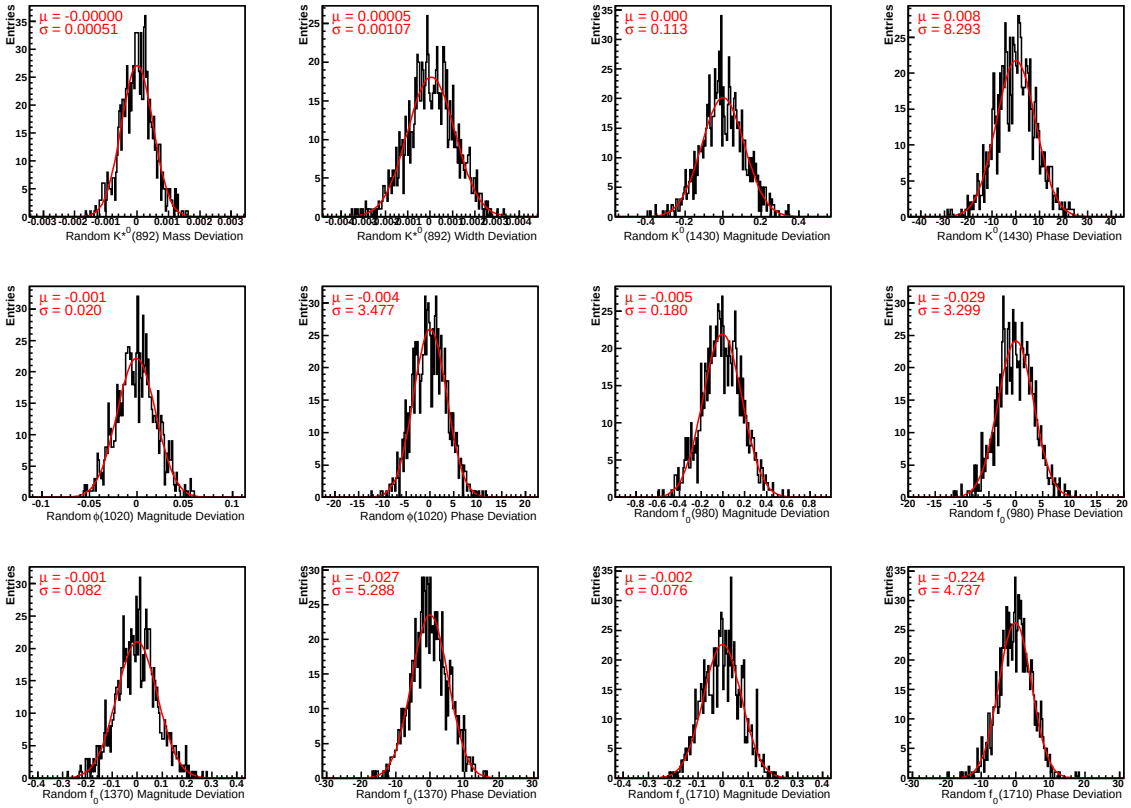


Figure 44: Randomly generated Dalitz parameter deviations for the scenario of correlated statistical Dalitz plot parameter uncertainties. The toy parameter distributions keep their Gaussian shape. The small deviations of the toy widths with respect to the uncertainties given by Table 5 are due to rounding errors.

B.2 Systematic Uncertainties

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

B.2.1 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

B.2.2 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

B.2.3 Toy Parameter Distributions

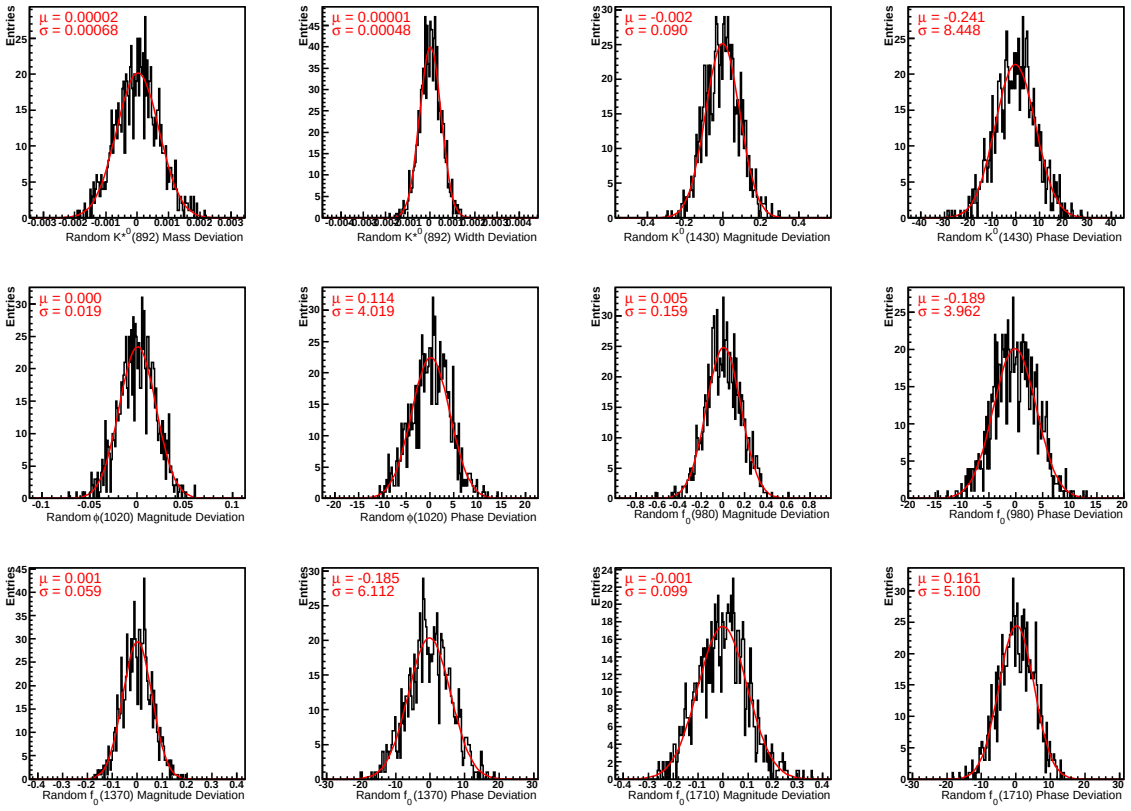


Figure 45: Randomly generated Dalitz parameter deviations for the scenario of systematic Dalitz plot parameter uncertainties.

C 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 [1] were derived using the identity $d^j_{m',m} = (-1)^{m-m'} d^j_{m,m'} = d^j_{-m,-m'}$.

$$
\begin{aligned}
d^1_{-1,-1} &= \frac{1+\cos\theta}{2} & d^1_{0,1} &= \frac{\sin\theta}{\sqrt{2}} & d^1_{1,-1} &= \frac{1-\cos\theta}{2} \\
d^1_{-1,0} &= \frac{\sin\theta}{\sqrt{2}} & d^1_{0,0} &= \cos\theta & d^1_{1,0} &= -\frac{\sin\theta}{\sqrt{2}} \\
d^1_{-1,1} &= \frac{1-\cos\theta}{2} & d^1_{0,-1} &= -\frac{\sin\theta}{\sqrt{2}} & d^1_{1,1} &= \frac{1+\cos\theta}{2}
\end{aligned}
$$

D Monte Carlo Decay Tables

D.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
```

D.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
```

D.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
```

D.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
```

References

[1] K. Nakamura et al. Review of Particle Physics. *J. Phys. G*, 37:075021, 2010 and 2011 partial update for the 2012 edition. 2, 13, 15, 16, 19, 26, 27, 34, 52, 56, 62, 68, 71, 74, 98, 111

[2] A. Abulencia et al. Observation of $B_s\bar{B}_s$ Oscillations. *Phys. Rev. Lett.*, 97:242003, 2006. 7

[3] I. Dunietz, R. Fleischer, U. Nierste. In Pursuit of New Physics with B_s Decays. *Phys. Rev. D*, 63:114015, 2001. 8, 102

[4] 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. 8, 98, 102

[5] M. A. Shifman, M. B. Voloshin. *Sov. J. Nucl. Phys.*, 47:511, 1988. 8

[6] R. Barate et al. *Phys. Lett. B*, 486:286, 2000. 9, 99

[7] V. M. Abazov et al. Evidence for the decay $B_s^0 \to {D_s}^{(*)}{D_s}^{(*)}$ and a measurement of ${\Delta\Gamma_s}^{CP}/\Gamma_s$. 2009. 9, 99

[8] 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. 9, 43

[9] 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 $\Delta\Gamma_s$. 2010. 9, 99

[10] D. Asner et al. b Hadron Lifetime averages – Results for the PDG 2011 Web Update. 10, 15, 101

[11] C. Amsler et al. Review of Particle Physics. *Physics Letters B*, 667:1, 2008. 13

[12] R. H. Dalitz. *Philos. Mag.*, 44:1068, 1953. 16

[13] J. M. Blatt, V. F. Weisskopf. *Theoretical Nuclear Physics*. John Wiley & Sons, New York, 1952. 18

[14] H. M. Pilkuhn. *The Interactions of Hadrons*. North-Holland Pub., Amsterdam, 1967. 18

[15] 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. 18

[16] R. E. Mitchell et al. Dalitz Plot Analysis of $D^+ \to K^+K^-\pi^+$. *arXiv:0903.1301v1 [hep-ex]*, 2009. 19, 20, 21, 87, 107

[17] G. Bonvicini. Dalitz Plot Analysis of the $D^+ \to K^-\pi^+\pi^+$ Decay. *Phys. Rev. D*, 78:052001, 2008. 19

[18] S. Kopp et al. Dalitz Plot Analysis of the Decay $D^0 \to K^-\pi^+\pi^0$. *Phys. Rev. D*, 63:092001, 2001. 19

[19] M. S. Dubrovin, D. Cinabro. Private Communication. 19

[20] V. Blobel, E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, chapter 6, page 175. B. G. Teubner, 1998. 23

[21] G. C. Wick M. Jacob. On the General Theory of Collisions for Particles with Spin. *Ann. Phys.*, 7:404–428, 1959. 26, 27

[22] S. U. Chung. Spin Formalisms. CERN Yellow Reports, CERN-71-08, 1971. 26

[23] J. D. Richman. An Experimenter's Guide to the Helicity Formalism. CALT-68-1148, DOE Research and Development Report, 1986. 26

[24] R. Kutschke. An Angular Distribution Cookbook. CLEO Internal Note, `http://home.fnal.gov/~kutschke/Angdist/angdist.ps`, 1986. 26

[25] S. M. Berman, M. Jacob. Systematics of Angular and Polarization Distributions in Three-Body Decays. *Phys. Rev. B*, B:1023, 1965. 29

[26] 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. 29

[27] L. Flores F. Wuerthwein R. J. Tesarek W. Bell, J. P. Fernandez. User Guide for EvtGen at CDF. CDF Note 5618. 30, 33

[28] D. J. Lange. The EvtGen Particle Decay Simulation Package. *Nuclear Instruments and Methods in Physics Research Section A*, 462:152–155, 2001. 30, 33

[29] A. Ryd et al. *EvtGen – A Monte Carlo Generator for B-Physics*, 2004. `http://robbep.home.cern.ch/robbep/EvtGen/GuideEvtGen.pdf`. 30, 33

[30] J. L. Rosner. Determination of Pseudoscalar-Charmed-Meson Decay Constants from B-Meson Decays. *Phys. Rev. D*, 42:3732, 1990. 31, 103

[31] 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. 31, 56, 72, 92

[32] K. Anikeev, C. Paus, P. Murat. Description of Bgenerator II. CDF Note 5092, 1999. 32

[33] T. Sjostrand, L. Lonnblad, S. Mra. *PYTHIA 6.2: Physics and Manual*, 2001. 32

[34] C. Peterson et al. *Phys. Rev. D*, 27:105, 1983. 32

[35] R. Brun, R. Hagelberg, M. Hansroul, J. Lassalle. CERN-DD-78-2-REV, 1978. 33

[36] The cdf detector simulation. `http://www-cdf.fnal.gov/cdfsim/cdfsim_main.html`. 33

[37] The trigger simulation project. `http://ncdf70.fnal.gov:8001/trigsim/trgsim.html`. 33

[38] M. Feindt. A Neural Bayesian Estimator for Conditional Probability Densities. *arXiv:physics/0402093*, 2004. 41, 48

[39] M. Feindt. *The NeuroBayes User's Guide*. `NEUROBAYES/doc/NeuroBayes-HowTo`. 48

[40] V. Blobel, E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, chapter 6, page 183. B. G. Teubner, 1998. 52

[41] R. Barlow. Extended Maximum Likelihood. *Nucl. Instrum. Meth. A*, 297:496–506, 1990. 52

[42] A. Ryd. Private Communication. 74

[43] J.C. Anjos et al. *Phys. Rev. D*, 48:56, 1993. 88, 90

[44] G. Bonvicini et al. *Phys. Rev. D*, 78:052001, 2008. 90, 92

[45] S. Esen for the Belle Collaboration. Observation of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ 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.* 101

[46] Chun-Khiang Chua, Wei-Shu Hou, Chia-Hsien Shen. Long Distance Contribution to $\Delta\Gamma$ of the $B_s\bar{B}_s$ system. 2011. 102

List of Figures

List of Tables