

**First observation of the
semileptonic decay**

$B^- \longrightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ with the
BABAR detector

D I S S E R T A T I O N
zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt von

Heiko Jasper

geboren am 06.02.1980 in Dresden

Lehrstuhl für Experimentelle Physik 5
Fakultät Physik
der Technischen Universität Dortmund

2009

1. Gutachter : Prof. Dr. Bernhard Spaan

2. Gutachter : Dr. Jochen Dingfelder

Datum des Einreichens der Arbeit : 14.08.2009

Abstract

This thesis reports on the search for the semileptonic decay $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ including final states with electrons and muons. The analysis uses a data sample of about 377 million $B\bar{B}$ pairs, recorded with the *BABAR* detector at the PEP-II asymmetric energy e^+e^- collider at the *SLAC National Laboratory* in the years 1999-2006. The D_s meson is reconstructed exclusively in three channels: $D_s \rightarrow \phi\pi$, with the ϕ decaying into K^+K^- , $D_s \rightarrow \bar{K}^{*0}K$ with the \bar{K}^{*0} being reconstructed in the charged $K\pi$ final state and $D_s \rightarrow \bar{K}_S^0 K$, where only the \bar{K}_S^0 decay into $\pi^+\pi^-$ is reconstructed. Several Multi Layer Perceptron neural networks are used for the event selection. The missing mass with respect to the nominal B meson mass is used for the extraction of the signal yield. An extended simultaneous maximum likelihood fit of the D_s signal-regions and sidebands of the three D_s reconstruction channels leads to a final result with a significance larger than 5σ , thus claiming the first observation of the signal decay. The branching ratios for the decay $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ are calculated to be

$$\begin{aligned}\mathcal{BR}(B^- \rightarrow D_s^+ K^- e^- \bar{\nu}_e) &= (5.81_{-1.30}^{+1.30}(\text{stat.}) \pm 0.54(\text{syst.}) \pm 0.49(\text{BR}(D_s))) \times 10^{-4}, \\ \mathcal{BR}(B^- \rightarrow D_s^+ K^- \mu^- \bar{\nu}_\mu) &= (6.68_{-1.69}^{+1.72}(\text{stat.}) \pm 0.69(\text{syst.}) \pm 0.56(\text{BR}(D_s))) \times 10^{-4}, \\ \mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) &= (6.13_{-1.24}^{+1.26} \pm 0.51(\text{BR}(D_s))) \times 10^{-4},\end{aligned}$$

where the errors are statistical, systematic and due to the limited knowledge of the branching ratios of the D_s and its daughters, respectively. The first error of last branching ratio represents the total statistical and systematic error of the combined measurement.

Contents

Contents	i
Pictures	iv
Tables	viii
1 Introduction	1
2 Theoretical Background	3
2.1 The Standard Model of Particle Physics	3
2.1.1 Electroweak Interaction	4
2.1.2 CKM-Matrix	6
2.1.3 Strong Interaction	8
2.2 B meson physics	9
2.2.1 Semileptonic B decays	9
2.2.2 The decay $B \rightarrow D_s K \ell \nu_\ell$	11
3 The $BABAR$ experiment	15
3.1 PEP-II asymmetric energy e^+e^- collider	16
3.2 The $BABAR$ detector	16
3.2.1 Silicon Vertex Tracker (SVT)	17
3.2.2 Drift Chamber (DCH)	19
3.2.3 Cherenkov Detector (DIRC)	19
3.2.4 Electromagnetic Calorimeter (EMC)	21
3.2.5 Instrumented Flux Return (IFR)	22
3.3 Track reconstruction	23
3.4 Monte Carlo simulation	24
3.5 Particle identification	25
4 Data sample	29
4.1 Data	29
4.2 Monte Carlo samples	29
4.2.1 Generic $B\bar{B}$ Monte Carlo	30
4.2.2 Generic continuum Monte Carlo	30
4.2.3 Signal Monte Carlo simulation	30

5	Reconstruction and event selection	33
5.1	Optimization of the selection	34
5.2	Reconstruction of composite particles	34
5.3	Separation of continuum events	36
5.4	Separation of cascade leptons and fake leptons	39
5.5	Combinatorial background suppression	40
5.5.1	Selection of $D_s \rightarrow K_S^0 K$ decays	40
5.5.2	Selection of $D_s \rightarrow \phi\pi$ and $D_s \rightarrow K^{*0}K$ decays	42
5.5.3	Reconstruction of the B meson	49
5.5.4	True- D_s background rejection	52
5.5.5	Multiple candidates	53
6	Data - Monte Carlo comparison of cut variables	59
6.1	Event shape Variables	60
6.2	$D_s \rightarrow K_S^0 K$ reconstruction	62
6.3	$D_s \rightarrow \phi\pi$ reconstruction	65
6.4	$D_s \rightarrow K^{*0} K$ reconstruction	68
6.5	B candidate reconstruction	71
6.6	Suppression of Upper vertex decays	73
6.7	Lepton momentum distributions	74
7	Analysis method and fit	75
7.1	Neutrino reconstruction	75
7.2	Fit technique	79
7.2.1	Development of the simultaneous fit	79
7.2.2	Fit validation	90
7.2.3	Significance expectation from Toy Monte Carlo results	97
7.3	Fitting the Data	98
8	Systematic uncertainties	103
8.1	Systematic uncertainties due to the fit method	103
8.1.1	Parametrization of the True- D_s fraction	103
8.1.2	Fit error of the $m(D_s)$ fits	104
8.1.3	Width of the signal gaussian	104
8.1.4	Constraints on the single channels signal yields	105
8.1.5	Bias correction	105
8.1.6	Total systematic uncertainties arising from the fitting method	106
8.2	Systematic uncertainties due to MC simulation and reconstruction	106
8.2.1	Signal Monte Carlo model	106
8.2.2	Signal Monte Carlo statistics	108
8.2.3	Particle identification	109
8.2.4	Track and Photon reconstruction	110
8.2.5	K_S^0 corrections	110
8.2.6	Radiative corrections - PHOTOS	111
8.2.7	B counting	111
8.2.8	D_s branching ratios	111
8.3	Total systematic uncertainties relative to efficiencies	112

9 Results	113
9.1 Total significance of the result	113
9.2 Calculation of the branching ratio	114
10 Conclusions	117
Appendix	119
A	119
A.1 Study of D_s^* contributions to the signal	119
A.2 Fit of sideband subtracted data	121
A.3 Toy MC validation plots	124
References	ix
Acknowledgments	xv

List of Figures

2.1	Summary of the Standard Model of Particle Physics.	4
2.2	The CKM unitary triangle	8
2.3	Feynman diagram of $B^{-/0} \rightarrow D^{(*)0/+} \ell \bar{\nu}_\ell$	9
2.4	Spectator diagram of the $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ decay.	11
2.5	Feynman diagram for B^0 oscillation.	12
3.1	PEP-II area at SLAC.	16
3.2	Integrated luminosity of the <i>BABAR</i> experiment.	17
3.3	Overview of the <i>BABAR</i> detector.	18
3.4	Schematic view of the Silicon Vertex Tracker (SVT).	18
3.5	Schematic view of the drift chamber (DCH).	19
3.6	Schematic view of the Cherenkov detector (DIRC).	20
3.7	Layout and schematic view of the crystals of the electromagnetic calorimeter (EMC).	21
3.8	Schematic view of the Instrumented Flux Return.	23
4.1	Comparison of the lepton momentum spectrum of different Signal MC simulation models.	31
5.1	Distributions of the event shape MLP input parameters $R2$, $\cos \Theta_{Thrust}$ and $L2$	37
5.2	Comparison of background rejection vs. signal efficiency between different algorithms.	38
5.3	Response and cut efficiencies for the MLP of the event shape parameters.	39
5.4	Lepton momentum spectrum of $D_s K \ell$ combinations in the CMS.	39
5.5	Cut variables for the $D_s \rightarrow K_S^0 K$ reconstruction chain.	41
5.6	Vertex probability of the D_s candidate.	41
5.7	Sources of misreconstructed kaons in the $D_s \rightarrow K_S^0 K$ reconstruction.	42
5.8	Mass distributions of the ϕ -veto for the $D_s \rightarrow K_S^0 K$ reconstruction.	43
5.9	Definition of the helicity angle.	43
5.10	Distributions of the $D_s \rightarrow \phi \pi$ reconstruction chain MLP input parameters.	44
5.11	Helicity angle distribution of combinatorial background events of $D_s \rightarrow K^{*0} K$ reconstruction chain and applied veto.	45
5.12	Distributions of the $D_s \rightarrow K^{*0} K$ reconstruction chain MLP input parameters.	46
5.13	Comparison of background rejection vs. signal efficiency between different algorithms for the D_s reconstruction.	46
5.14	Response and cut efficiencies for the MLP of the D_s reconstruction.	47

5.15	Significance of cut on the MLP output vs. significance of the lepton momentum cut.	48
5.16	Profile plots of significance of cut on the MLP output vs. significance of the lepton momentum cut at $p_\ell = 1 \text{ GeV}/c$	48
5.17	Mass distributions of the ϕ -vetoes for the $D_s \rightarrow K^{*0}K$ reconstruction.	49
5.18	Distributions of the TMVA input parameters $m(B)$, $Sphericity(B)$ and $Thrust(\text{Rest of the event})$	50
5.19	Response and cut efficiencies for the MLP of the B reconstruction.	51
5.20	Significance vs. Cut on the MLP(B shape) response.	51
5.21	Significance vs. cut on the B candidates vertex probability.	52
5.22	Spectator diagram of upper vertex decays.	52
5.23	Angular distributions of $\cos \Theta(D_s K)$	53
5.24	Multiplicity and quality of best candidate choice.	54
6.1	Data-MC comparison of event shape variables.	61
6.2	Data-MC comparison of the MLP output from the event shape variables.	62
6.3	Data-MC comparison of variables, used for the $D_s \rightarrow K_S^0 K$ reconstruction, part 1.	63
6.4	Data-MC comparison of variables, used for the $D_s \rightarrow K_S^0 K$ reconstruction, part 2.	64
6.5	Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K_S^0 K$ reconstruction chain.	64
6.6	Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K_S^0 K$ reconstruction chain, corrected for D_s branching ratio.	65
6.7	Data-MC comparison of variables, used for the $D_s \rightarrow \phi\pi$ reconstruction, part 1.	66
6.8	Data-MC comparison of variables, used for the $D_s \rightarrow \phi\pi$ reconstruction, part 2.	67
6.9	Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow \phi\pi$ reconstruction chain.	67
6.10	Data-MC comparison of variables, used for the $D_s \rightarrow K^{*0}K$ reconstruction, part 1.	68
6.11	Data-MC comparison of variables, used for the $D_s \rightarrow K^{*0}K$ reconstruction, part 2.	69
6.12	Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K^{*0}K$ reconstruction chain.	70
6.13	Data-MC comparison of variables, used for the $D_s \rightarrow K^{*0}K$ reconstruction, part 3.	70
6.14	Data-MC comparison for input variables of the MLP, used for the B reconstruction.	71
6.15	Data-MC comparison of the MLP-output of the B-reconstruction and vertex probability of the $D_s K \ell$ combinations.	72
6.16	Data-MC comparison of the angles between the reconstructed D_s and additional kaons.	73
6.17	Data-MC comparison of the lepton momenta.	74
7.1	$\cos \Theta_{BY}$ distribution for signal events.	76
7.2	MM^2 distribution for signal events.	76
7.3	$\cos \Theta_{BY}$ distribution for background events.	77

7.4	MM^2 distribution for background events.	78
7.5	MM^2 and $\cos \Theta_{BY}$ distribution for signal and background events.	78
7.6	D_s mass distributions for data and MC, electron channel.	80
7.7	D_s mass distributions for data and MC, muon channel.	81
7.8	MM^2 data-MC comparison for the $m(D_s)$ sidebands.	83
7.9	MM^2 data-MC comparison for the $m(D_s)$ signal regions.	84
7.10	Fits of the True- D_s background MM^2 distributions for electron and muon channel.	86
7.11	Fits of the Signal MC MM^2 distributions for electron and muon channel.	88
7.12	Generic MC MM^2 validation fits for the electron channel.	91
7.13	Generic MC MM^2 validation fits for the muon channel.	92
7.14	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = (1/3/6) \times 10^{-4}$ for the electron channel.	94
7.15	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = (1/3/6) \times 10^{-4}$ for the muon channel.	95
7.16	Bias vs. Signal BR for electron and muon channel.	96
7.17	Expected significance of the measurement vs. Signal BR for electron and muon channel.	97
7.18	Final fits on the electron channel data sample.	98
7.19	Final fits on the muon channel data sample.	99
7.20	Toy MC results for signal branching ratios of $\mathcal{BR}_{\text{signal}} = 6.5/6.8 \times 10^{-4}$ for the muon channel.	101
8.1	2dimensional distributions of Toy Monte Carlo for True- D_s background PDF parameters.	104
8.2	Results of Toy MC fits, determining the systematics arising from the fixed True- D_s background PDF parameters.	104
8.3	Results of Toy MC fits, determining the systematics arising from the constraints on the single channels signal yields.	105
8.4	Efficiency vs. momentum for used lepton selectors.	107
8.5	Comparison of the lepton momentum distributions between different Signal MC models.	108
9.1	Sideband subtracted MM^2 distributions with fitted PDFs for electron and muon channel.	115
A.1	Comparison of MM^2 distributions for signal decays with D_s and D_s^* in the final state.	120
A.2	Simultaneous χ^2 fits to sideband subtracted generic MC for electron and muon channel.	122
A.3	Toy MC results for of the fit to the sideband subtracted MC sample electron and muon channel.	123
A.4	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = (1/2/3) \times 10^{-4}$ for electron channel.	124
A.5	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = (4/5/6) \times 10^{-4}$ for electron channel.	125
A.6	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = (7/8/9) \times 10^{-4}$ for electron channel.	126
A.7	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = (1/2/3) \times 10^{-4}$ for muon channel.	127

A.8	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = (4/5/6) \times 10^{-4}$ for muon channel.	128
A.9	Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = (7/8/9) \times 10^{-4}$ for muon channel.	129

List of Tables

2.1	Properties of orbitally excited D^{**}	13
3.1	Cross-sections at $\sqrt{s} = m(\Upsilon(4S))$	15
3.2	Definition of track quality requirements.	24
4.1	Monte Carlo event samples	32
5.1	Relevant particle parameters for D_s reconstruction.	35
5.2	Reconstruction cuts on the $D_s \rightarrow \phi\pi$ decay chain.	55
5.3	Reconstruction cuts on the $D_s \rightarrow K^{*0}K$ decay chain.	56
5.4	Reconstruction cuts on the $D_s \rightarrow K_S^0K$ decay chain.	57
7.1	Results of the simultaneous fits with the Fermi-function PDF for electron and muon channel.	85
7.2	Results of the simultaneous fits to Signal MC for electron and muon channel.	87
7.3	Results of the simultaneous validation fits to full generic MC with a defined signal fraction.	93
7.4	Results of the Toy MC fits.	96
7.5	Results of the fits to data.	100
8.1	Overview of the additive systematic uncertainties.	106
8.2	Relative change of reconstructed events for the different Signal MC models.	107
8.3	Systematic uncertainties arising from the limited amount of Signal MC.	108
8.4	Systematic uncertainties arising from the particle identification.	109
8.5	Killing probability for charged tracks in different <i>BABAR</i> runs.	110
8.6	Systematic uncertainties arising from the track and cluster reconstruction.	110
8.7	Systematic uncertainties arising from the simulation of final state radiation by PHOTOS.	111
8.8	Overview of the multiplicative systematic uncertainties.	112
9.1	Significance of the result.	114
A.1	Results of the simultaneous validation fits to the sideband subtracted generic MC with a defined signal fraction.	121

Chapter 1

Introduction

In the past century, the discovery of ever smaller constituents of matter led to the foundation of a new field of physics, the physics of elementary particles. Its subject is the composition and interactions of all known particles in the universe. By colliding particles at high energies, other particles, which are not present in the environment now, can be produced. These particles once existed in the universe like the ones, that are observable now, but vanished due to the decrease of the energy density in the universe. The decrease led to a thermal disequilibrium between the processes creating these particles and their decay, so that more particles decayed than could be reproduced. Thus, the acceleration and collision of particles create conditions similar to a distinct time in the universe's history. The knowledge, gained through these reactions, together with information derived from observations of radiation of the space over a large bandwidth can basically explain the whole evolution of the world from a short time after the big bang to the state we observe now. Nevertheless, it cannot explain why the universe is filled with matter at all. It is expected, that the big bang creates not only time, space and matter, but also anti-matter, consisting of anti-particles, in the same amount like matter. Since matter and anti-matter annihilate into radiation whenever they meet, there should be no matter left to build the solid structures of the universe, like galaxies or planets. This is obviously not the case.

There are two scenarios, which can explain the existence of solid structures. The first would be a spatial separation, which means that somewhere in the universe there are regions, filled with anti-matter without having any contact to regions filled with matter. Since such a separation is never perfect, from time to time some pieces of matter and anti-matter would meet and annihilate producing radiation of a specific energy, which could be detected by the telescopes on earth. No evidence of such radiation has been found up to now, making the theory of a spatial separation highly unlikely. This also leads to the conclusion, that today there is no anti-matter left in the universe. In 1967, Andrei Sakharov was the first to recognize that three conditions are needed to dynamically create the observed matter anti-matter asymmetry [1], providing a second possible explanation for the existence of matter in the universe. The first condition is the violation of the Baryon number B , basically describing the amount of strong interacting matter in the universe. This is obviously the case, since at the beginning of the universe $B = 0$ whereas now $B \neq 0$. Another condition is the presence of a thermal disequilibrium. This is necessary because an equilibrium would eliminate a Baryon number violation, by inverting the reaction, which created the violation.

In addition, the charge conjugation symmetry \mathcal{C} and combination of \mathcal{C} with the parity inversion \mathcal{P} , being equal to the flip of the sign of the spatial coordinates, \mathcal{CP} has to be violated. This kind of violation is already contained in the *Standard Model of Particle Physics*, which is the theory describing the known particles and the interactions between each other. As a matter of fact, this \mathcal{CP} violation can be measured by analyzing the decays of various particles. Its first observation can be dated back to the year 1964 [2]. It turned out, that the measured \mathcal{CP} violation, which can be described perfectly by the standard model, is not large enough to describe the observed asymmetry between matter and anti-matter. Thus, other theories and mechanism are needed for an accurate description. Since then, several experiments extended the knowledge of the \mathcal{CP} violation. The *BABAR* experiment, whose data this thesis used, is one of them. Besides the opportunity of measuring the \mathcal{CP} violation in a specific system, delivering the most precise measurement so far, the data can be used to probe the standard model by measuring fundamental parameters, which cannot be predicted by the theory. The experiment also provides enough data for extending the present knowledge of particle decays, which have not been observed yet, but may affect the accuracy of the measurements of the standard model parameters. The analysis, described by this thesis performed such a search for an unmeasured particle decay, the transition $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$. The number of B mesons, decaying through this particular transition, is filtered out of a high number of B meson decays, recorded with the *BABAR* detector, in order to calculate the probability of this decay to happen. This probability is expressed as a branching ratio \mathcal{BR} of the signal decay.

The analysis presented in this thesis is embedded into a theoretical framework, being described in the following chapter. It contains a short introduction into the field of particle physics, as well as the motivation for this analysis. Chapter 3 gives an overview of the *BABAR* experiment, followed by a description of the data used in this analysis. The candidate selection process and a detailed comparison between real and simulated events is presented in the chapters 5 and 6. Then, the analysis technique featuring the neutrino reconstruction and the fit determining the signal yield is described, followed by a discussion of the evaluated systematic uncertainties in the 8th chapter. The presentation of results and a conclusion finalizes this thesis.

Chapter 2

Theoretical Background

This work is embedded into a theoretical framework, which describes most of the present knowledge of particle physics, the *Standard Model of Particle Physics*. It contains all known particles and (almost) all important interactions between them as well as the CKM-Matrix, which plays a key role in the description of \mathcal{CP} violation in the standard model. But what is this matrix, why is it so important to measure its matrix elements and how can semileptonic decays (and especially this work) contribute to the determination of them? Another important question, being answered in the following chapter is: How can the behavior and decays of B mesons be described?

2.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a quantum field theory and a gauge field theory, describing today's knowledge of the fundamental particles and three of the four fundamental interactions, namely the electromagnetic interaction, the weak interaction and the strong interaction. The fourth force, gravity, is not included, due to its geometrical nature¹.

All fundamental particles, forming *matter* in the SM are fermions with spin $1/2$, while the interaction fields are mediated by bosons with an integer spin quantum number. The fermions can be divided into two classes, quarks and leptons. Quarks have a strong charge (the so called *color*) and thus can interact through the strong force, while the leptons are color neutral and not affected by the strong interaction. Both classes consist of 6 particles (*flavors*)², divided into three hierarchic ordered generations, which differ only in the mass of the constituents but not in any other quantum number. The first family of the leptons is formed by the electron (e) and the electron-neutrino (ν_e), the second family consists of the muon (μ) and the μ -neutrino (ν_μ), whereas the members of the heaviest generation are the tau lepton τ and the τ -neutrino ν_τ . At the quark sector, the up (u) and down (d) quarks form the first generation, the charm (c) and strange (s) quarks the second one and the top (t) and bottom (b) the third one³.

¹Gravity has a tensor-structure. The carriers of this force therefore must have a spin of 2, which leads to non-renormalizable terms in the theoretical description of the interaction [3].

²All through this document, the notation "particle" includes anti-particles too, unless quoted otherwise explicitly.

³Up, charm and top quarks are often called *up-type* quarks, whereas down, strange and bottom quarks are *down-type* quarks.

The three interactions are introduced into the SM by forming a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge group. Here, C denotes the color of the strong interaction and L represents the fact, that the electroweak force couples only to the left-handed states of the particles. Y denotes the weak hyper-charge, which results of the unification of the weak and electromagnetic interaction in the electroweak unified theory by Glashow, Weinberg and Salam. The interactions are mediated by interaction bosons. The γ acts as the carrier of the electromagnetic force, whereas W^\pm and Z^0 mediate the weak force. The strong interaction features eight interaction bosons, the gluons. Figure 2.1 summarizes the content of the Standard Model of Particle Physics.

Three Generations of Matter (Fermions)				
	I	II	III	
mass→	2.4 MeV/c ²	1.27 GeV/c ²	171.2 GeV/c ²	0
charge→	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
name→	u up	c charm	t top	γ photon
Quarks	4.8 MeV/c ²	104 MeV/c ²	4.2 GeV/c ²	0
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	d down	s strange	b bottom	g gluon
Leptons	<2.2 eV/c ²	<0.17 MeV/c ²	<15.5 MeV/c ²	91.2 GeV/c ²
	0	0	0	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	Z⁰ weak force
Bosons (Forces)	0.511 MeV/c ²	105.7 MeV/c ²	1.777 GeV/c ²	80.4 GeV/c ²
	-1	-1	-1	±1
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	e electron	μ muon	τ tau	W[±] weak force

Figure 2.1: Summary of the Standard Model of Particle Physics.

2.1.1 Electroweak Interaction

As already mentioned, the electroweak interaction represents the unification of the electromagnetic interaction and the weak force and affects only the left-handed particle states. The carriers of the forces are the W and Z bosons for the weak interaction and the γ for the electromagnetism. While the W^\pm and the Z^0 have an effective mass, generated via the Higgs mechanism, the photon is massless. The structure of the coupling of the weak interaction causes left-handed lepton and quark fields being weak isospin-doublets $\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$ ⁴, while all right handed fields are singlets with respect to the weak interaction (ν_{eR}, e_R) .

⁴The fields, denoted with bold characters are all kinds of fermion-fields, leptons and quarks. They should not be confused with electron and electron-neutrino field, with are denoted by the same non-bold characters.

Since the Lagrangian of the weak interaction has to be invariant under rotations in the weak isospin space, the interactions between the fermion fields can be written as:

$$\begin{aligned}\mathcal{L}_{EW} &= -q_e \{ \mathcal{J}_{em}^\lambda A_\lambda \\ &+ \frac{1}{\sin \vartheta_W \cos \vartheta_W} Z_\lambda \mathcal{J}_{NC}^\lambda \\ &+ \frac{1}{\sqrt{2} \sin \vartheta_W} (W_\lambda^+ \bar{\nu}_{eL} \gamma^\lambda \mathbf{e}_L + W_\lambda^- \bar{\mathbf{e}}_L \gamma^\lambda \nu_{eL}) \}\end{aligned}$$

using

$$\begin{aligned}q_e &= g \sin \vartheta_W; \\ \mathcal{J}_{em}^\lambda &= -\bar{\mathbf{e}} \gamma^\lambda \mathbf{e}; \\ \mathcal{J}_{NC}^\lambda &= 1/2 \bar{\nu}_{eL} \gamma^\lambda \nu_{eL} - 1/2 \bar{\mathbf{e}}_L \gamma^\lambda \mathbf{e}_L - \sin^2 \vartheta_W \mathcal{J}_{em}^\lambda\end{aligned}$$

A_λ , Z_λ , and W_λ^\pm are the electroweak gauge fields, while q_e is the elementary electric charge, g the weak coupling constant and ϑ_W is the *weak mixing angle* or *Weinberg angle* [4]. In this Lagrangian the W and Z fields are still massless. Naively, one could include explicit mass terms for the fields (e.g. $W_\lambda^\pm W^{\lambda\pm} m_W$), but that leads to a non-gauge-invariant, non-renormizable theory, containing an infinite number of parameters. The most popular proposal to explain the mass of the bosons is the so called *Higgs mechanism*, introduced by Peter Higgs [5, 6]. It inserts two complex, scalar fields $\Phi(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}$, forming a doublet in the weak isospin space, into the Lagrangian. This leads to an additional term in the SM Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (\delta^\mu \Phi)^\dagger (\delta_\mu \Phi) - V(\Phi); \quad \text{with } V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2,$$

which is constructed in such a way, that the ground state has a non-zero expectation value at low energies, but the potential $V(\Phi)$ is completely symmetric. Below a certain temperature (or energy), this leads to a spontaneous symmetry breaking and to a non-zero vacuum expectation value, which can be chosen as:

$$\Phi(x) = 1/\sqrt{2} \begin{pmatrix} 0 \\ \rho_0 + h(x) \end{pmatrix}; \quad \text{with } \rho_0 = \sqrt{\frac{\mu^2}{\lambda}}.$$

Using the newly introduced Higgs-field and its ground state expectation value, the interaction bosons become massive particles with masses of:

$$\begin{aligned}m_W^2 &= \frac{q_e^2 \rho_0^2}{4 \sin^2 \vartheta_w} \text{ and} \\ m_Z^2 &= \frac{q_e^2 \rho_0^2}{4 \sin^2 \vartheta_w \cos^2 \vartheta_w}.\end{aligned}$$

The coupling of fermions to the Higgs-field is named *Yukawa-coupling* after Hideki Yukawa. It also creates the masses of the fermions:

$$\begin{aligned}\mathcal{L}_{\text{Yuk}} &= -c_e \bar{\mathbf{e}}_R \Phi^\dagger \begin{pmatrix} \nu_{eL} \\ \mathbf{e}_L \end{pmatrix} + h.c. \\ &= -c_e \frac{\rho_0}{\sqrt{2}} (\bar{\mathbf{e}}_R \mathbf{e}_L + \bar{\mathbf{e}}_L \mathbf{e}_R) \\ &= m \bar{\mathbf{e}} \mathbf{e}\end{aligned}$$

with c_e being an arbitrary coupling constant.

Extending the Yukawa-coupling to all three generations of leptons results in

$$\mathcal{L}_{\text{Yuk, leptons}} = (\bar{e}_R, \bar{\mu}_R, \bar{\tau}_R) \mathbf{C}_\ell \begin{pmatrix} \Phi^\dagger \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \\ \Phi^\dagger \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix} \\ \Phi^\dagger \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \end{pmatrix} + h.c.$$

with $\begin{pmatrix} \nu_{(e/\mu/\tau)L} \\ e/\mu/\tau_L \end{pmatrix}$ being the eigenstates of the left-handed leptons, $\bar{e}_R, \bar{\mu}_R, \bar{\tau}_R$ being the eigenstates of the right-handed leptons and \mathbf{C}_ℓ an arbitrary complex 3×3 matrix. Through matrix transformations like $\mathbf{C}_\ell \rightarrow U_1^\dagger \mathbf{C}_\ell V_1$, with U_1 and V_1 being unitary matrices, \mathbf{C}_ℓ can be diagonalized and the lepton masses are given by the elements of the matrix:

$$\begin{aligned} \mathcal{L}_M^{\text{leptons}} &= -(\bar{e} \ \bar{\mu} \ \bar{\tau}) \mathbf{C}_\ell \begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix} \\ &= -(\bar{e} \ \bar{\mu} \ \bar{\tau}) \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} \begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix}. \end{aligned}$$

For the quarks, the equations are very similar, but there exist two matrices \mathbf{C}_q and \mathbf{C}'_q due to both isospin partners having mass.

2.1.2 CKM-Matrix

Both matrices \mathbf{C}_q and \mathbf{C}'_q cannot be diagonalized at the same time, because they're affecting the same doublets. Only transformations like

$$\begin{aligned} \mathbf{C}'_q &\rightarrow U_2^\dagger \mathbf{C}'_q V_2 \text{ and} \\ \mathbf{C}_q &\rightarrow U_3^\dagger \mathbf{C}_q V_2, \end{aligned}$$

where U_2, U_3 and V_2 are unitary matrices, can be made. Thus, the diagonalization of \mathbf{C}_q requires an additional matrix V in the mass term of the Lagrangian for quarks:

$$\begin{aligned} \mathcal{L}_M^{\text{quarks}} &= -(\bar{u} \ \bar{c} \ \bar{t}) \mathbf{C}'_q \begin{pmatrix} u \\ c \\ t \end{pmatrix} - (\bar{d}' \ \bar{s}' \ \bar{b}') V \mathbf{C}_q V^\dagger \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \\ &= -(\bar{u} \ \bar{c} \ \bar{t}) \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \begin{pmatrix} u \\ c \\ t \end{pmatrix} \\ &\quad - (\bar{d}' \ \bar{s}' \ \bar{b}') V \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} V^\dagger \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \end{aligned}$$

with u, d', c, s', t, b' being the quark fields of the electroweak theory and V is the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix.

As can be seen in eqn. (2.1), the CKM matrix relates the interaction eigenstates to the mass eigenstates of the quarks.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.1)$$

By requiring the CKM matrix to be unitary⁵, there are nine constraints on the eighteen free parameters of the complex 3×3 matrix:

$$\sum_{k=1}^3 V_{ki} V_{kj}^* = \delta_{ij} \quad ; \quad i, j = 1, 2, 3. \quad (2.2)$$

Five more parameters are absorbed in the six phases of the wave functions of the quark fields, which are not observable. Thus, there remain four free parameters, three angles ϑ_{ij} (were $i \neq j$) and one phase δ , which have to be determined by measurements. The Wolfenstein expansion represents a very common parametrization of the CKM-Matrix [7], taking into account the results of measurements of matrix elements:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (2.3)$$

with $\lambda = 0.2257_{-0.0010}^{+0.0009}$, $A = 0.814_{-0.022}^{+0.021}$, $\rho = 0.135_{-0.016}^{+0.031}$ and $\eta = 0.349_{-0.017}^{+0.015}$ [8]. Obviously, the diagonal elements of the CKM matrix are close to 1, whereas transitions between the first and the third quark generation are very much suppressed.

Physics with B mesons can contribute directly to the determination of the CKM matrix elements V_{cb} , which essentially can measure A , when divided by the $|V_{us}|^2$, and V_{ub} , which introduces \mathcal{CP} violation in the quark sector if $V_{CKM}^* \neq V_{CKM} \rightarrow \eta \neq 0$. Besides that, a measurement of $|\frac{V_{ub}^*}{V_{cd}V_{cb}}| = \sqrt{\rho^2 + \eta^2}$ defines a circle in the ρ, η plane, constraining the most important unitary triangle of particle physics. The unitary triangle arises from the interpretation of eqn. (2.2) for $i \neq j$ as a definition of triangles. For $i = 1$ and $j = 3$, one gets $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$. Dividing each side by $V_{cd}V_{cb}^*$, results in the unit-length for the base. Figure 2.2 shows a graphic impression of the unitary triangle in the complex $\bar{\rho}, \bar{\eta}$ plane. The definitions of $\bar{\rho}$ and $\bar{\eta}$ reproduce all approximate results in the literature, e.g. $\bar{\rho} = \rho(1 - \lambda^2/2 + \dots)$ [8]. A primary goal of high energy physics is to over-constrain this triangle, because any contradiction to the unitary constraints would hint to physics beyond the Standard Model of Particle Physics, commonly denoted as *New Physics*.

⁵Theory defines the CKM matrix to be an unitary matrix, describable as the product of two unitary matrices. From the experimental point of view, it is not unitary by definition. This has to be verified by measurements. Proving unitary of the CKM matrix experimentally, means proving, that the theory of the electroweak interaction sufficiently describes **all** observed flavor changing transitions. This is a major non-trivial aspect of the Standard Model of Particle Physics.

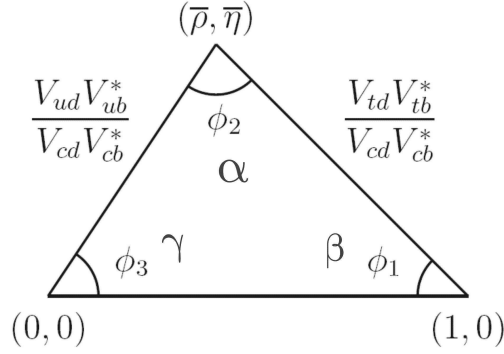


Figure 2.2: The rescaled version of the unitary triangle in which each side has been divided by $V_{cd}V_{cb}^*$, so that the base has unit length.

2.1.3 Strong Interaction

In addition to the electric charge, Quarks also carry a strong charge, which has six possible values. These values are red, green and blue for the quarks and anti-red, anti-green and anti-blue for the anti-quarks. An important property of the strong interaction, which is theoretically described by the *Quantum-Chromo-Dynamic* (QCD), is the *Confinement*. It describes the fact, that it is not possible to observe colored objects in nature, only color-neutral particles are visible. Such particles can consist of three different colored constituents with red + green + blue = white in a qqq object (*Baryon*) or of a color + anti-color = white combination in a $q\bar{q}$ particle, a *Meson* - in analogy to the optical color theory in biology. Related to the confinement is the property of the effective coupling constant of the QCD, α_S being energy dependent. At low energies, which es equivalent to large length scales (or distances between the quarks in a bound system), α_S becomes very strong. At small length scales the strength of the strong interactions becomes small. Thus, within a shortdistance bound state, the constituents can act as quasi-free particles, which is called *asymptotic freedom*. Since the QCD is a non-Abelian $SU(3)$ gauge theory, there are eight interaction bosons, the *Gluons*, which are flavor-blind, but carry color charge (actually one color and one anti-color) themselves. Thus, they couple not only to the quarks, but also to each other. This self coupling and the QCD confinement makes it very difficult to calculate QCD processes. The calculations become easier, when a heavy quark forms a meson together with a light quark. Heavy quarks are defined by $m_Q \gg \Lambda_{\text{QCD}}$, with Λ_{QCD} being the characteristic energy scale, that separates regions of large couplings from those, where the coupling constant is small. Its value is about 200 MeV, resulting in the treatment of the c , b and t quarks as heavy quarks and the u , d and s as light quarks. Calculations of QCD processes then become possible for the heavy quarks, due to the length scales being comparable or smaller than the Compton wavelength of the quark $\Lambda_Q \sim 1/m_Q$. Within this limit, non-pertubative strong interactions have the size of Λ_{QCD} and the heavy quark doesn't recoil. Thus, the QCD becomes pertubative and the heavy quark acts as a static source of electric and chromoelectric field. The *Heavy Quark Effective Theory* (HQET) [9, 10] uses such an approach to calculate strong interactions, resulting in very successful calculations in the field of semileptonic B decays.

2.2 *B meson physics*

As already mentioned in the previous sections, studies of *B* meson decays offer a wide field of important measurements of parameters of the SM. In addition to that, processes where New Physics could show up are accessible. Analyses of semileptonic *B* meson decays provide an excellent opportunity to determine the CKM-Matrix elements V_{ub} and V_{cb} , due to the strong interaction being isolated to the hadronic current. Since this thesis describes the extension of the current knowledge of these kind of decays, a short overview of the opportunities in extracting V_{cb} using $b \rightarrow c\ell\bar{\nu}$ transitions is given, followed by a description of the decay, that has been analyzed.

2.2.1 Semileptonic *B* decays

The determination of CKM-Matrix elements through semileptonic decays is the easiest approach, since the branching ratios of the decays are directly proportional to their absolute square. Furthermore, calculations for $b \rightarrow c\ell\bar{\nu}_\ell$ transitions, with ℓ being an electron or a muon and $\bar{\nu}_\ell$ the according anti-neutrino, are very accurate in the HQET. There are two major approaches to extract the matrix elements, the analysis of exclusive decays (such as $B \rightarrow D^*\ell\nu$ and $B \rightarrow \pi\ell\nu$) and inclusive measurements of $B \rightarrow X_{c/u}\ell\nu$ decays, with $X_{c/u}$ being any possible final state with a charmed/up quark in the final state. Both approaches suffer from requiring several approximations and from a lack of information about background sources and/or unknown contributions, resulting in an error of the measurements. The following sections feature a more detailed description of both approaches, restricted to the V_{cb} determination, as only it is relevant for this analysis.

Exclusive semileptonic *B* decays

The Feynman diagram of a semileptonic $b \rightarrow c\ell\bar{\nu}$ decay is shown in figure 2.3.

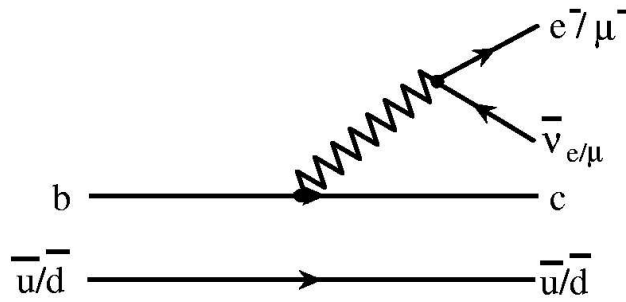


Figure 2.3: Feynman diagram of $B^{-/0} \rightarrow D^{(*)0/+}\ell\bar{\nu}_\ell$ decays as example of a semileptonic $b \rightarrow c\ell\nu$ transition.

As pointed out in subsection 2.1.3, the b , as well as the c quark are heavy quarks, which act only as static source of a chromoelectric field. The B and $D^{(*)}$ mesons also contain a light quark, which just sees the gluon field of the heavy quark, but is blind to its the flavor. On average, the velocity of the B meson and the b quark is the same. In the limit of $m_Q = m_b = m_c \rightarrow \infty$, the decay of the heavy quark just results in a change of velocity v of the source of the gluon field for the light quark.

Soft gluons then need to be exchanged to form a meson moving at a velocity of v' , which results in a suppression of the form factor for large velocity differences. Hence, the probability for an elastic transition decreases. The form factor for the transition then only depends on the Lorentz boost $w = v \cdot v'$ between the initial and final state mesons. The *Isgur-Wise-Function* $\mathcal{F}(w) = \mathcal{F}(v \cdot v')$, a dimensionless probability amplitude, describes the transition [10]. Only $\mathcal{F}(1)$ can be determined in the HQET, which is practically inaccessible for measurements, because the phase-space for $v \cdot v' = 1$, where the B and the daughter meson have a common rest frame, vanishes. The determination of V_{cb} is done by measuring the differential decay rate $d\Gamma/dw \sim |V_{cb}|^2 \mathcal{F}(w)^2$, and extracting $|V_{cb}|$ from an extrapolation of the data to the zero recoil point $w = 1$.

Over the last decade, several analyses of $B \rightarrow D\ell\nu$ [11] and $B \rightarrow D^*\ell\nu$ [12–14] decays have been performed leading to a weighted average of $V_{cb, excl.} = (38.6 \pm 0.9_{exp} \pm 1.0_{theo}) \times 10^{-3}$ [8], where the first error is due to statistical limitations and systematics arising from the experimental side and the second error is reflecting the accuracy of the calculations, including corrections to the approximations used within the calculation described.

Inclusive semileptonic B decays

The inclusive semileptonic branching fraction into charmed mesons depends on the CKM-Matrix element V_{cb} and the masses of the participating quarks, m_b and m_c , as well as several terms characterizing the strong interaction between the heavy and the light quark. An *Operator Product Expansion* (OPE) of the decay rate $\Gamma(B \rightarrow X_c \ell \nu)$ can be performed in terms of the inverse heavy quark masses and the strong interaction coupling constant α_S . In this expansion, there are no terms of the order of $1/m_b$, because they correspond to a decay of a free quark. Terms of the order of $1/m_b^2$ feature one-gluon corrections, a contribution from non-perturbative QCD effects, which can be interpreted as the energy of the b quark (proportionality factor μ_π^2), and a contribution related to the mass difference between the B and the B^* (proportionality factor μ_G^2) [10, 15]. The variables m_b , m_c , μ_π and μ_G need to be extracted from the experiment to measure V_{cb} . This happens by using the shape of the lepton energy spectrum (E_ℓ) and the hadronic mass spectrum M_{X_c} , which are connected to these parameters:

$$\frac{\int E_\ell^n (d\Gamma/dE_\ell) dE_\ell}{\int (d\Gamma/dE_\ell) dE_\ell} = f_n(m_b, m_c, \mu_\pi^2, \mu_G^2, \dots)$$

(with Γ representing the inclusive decay rate and E_ℓ^n the n -th moment of the lepton energy spectrum). Several momenta are fitted simultaneously, deriving a complete set of variables at the same time. It has to be noted that the differential distribution $d\Gamma/dE_\ell$ can be calculated with high precision only for values of E_ℓ not too close to the endpoint. In this region, a few resonances dominate the spectrum and account for a large fraction of the full spectrum. Detailed knowledge of the size of even small semileptonic decay rates is essential for comparisons between theory and experiment and an estimation of the error of $V_{cb, incl.}$. It is expected that the sum of the branching ratios of all exclusive decays should be consistent with the measured inclusive branching ratio. However, this is not the case at the current state of the art of B meson physics. The known exclusive semileptonic branching ratios only account for about $\sim 98\%$ of the inclusive rate. Probably, the missing 2% are covered by processes, which are not measured up to now. Candidates for such “missing” decays are, for example, baryonic semileptonic decays and the decay, analyzed in this thesis, $B \rightarrow D_s K \ell \nu_\ell$ ⁶.

⁶This decay is referred to as “signal decay” in this thesis.

The world average for the determination of V_{cb} through inclusive semileptonic measurements is $V_{cb, incl.} = (41.68 \pm 0.39_{exp} \pm 0.58_{theo}) \times 10^{-3}$ [8, 16]. While the errors on both, the experimental and the theoretical side, are smaller than for the exclusive measurements, there remains a discrepancy of about 2σ between both measurements. The main part of the theoretical error arises from the estimated accuracy of the OPE for the total semileptonic rate.

2.2.2 The decay $B \rightarrow D_s K \ell \nu_\ell$

This decay has not been observed yet, but represents an analog transition to $B \rightarrow D\pi\ell\nu_\ell$. The latter decay forms the $D\pi$ system by a fragmentation with $u\bar{u}$ or $d\bar{d}$ quarks, whereas the $B \rightarrow D_s K \ell \nu_\ell$ fragments with $s\bar{s}$ quarks in the final state, forming a D_s meson and a kaon. The spectator diagram of this decay is shown in figure 2.4. In addition to the shown

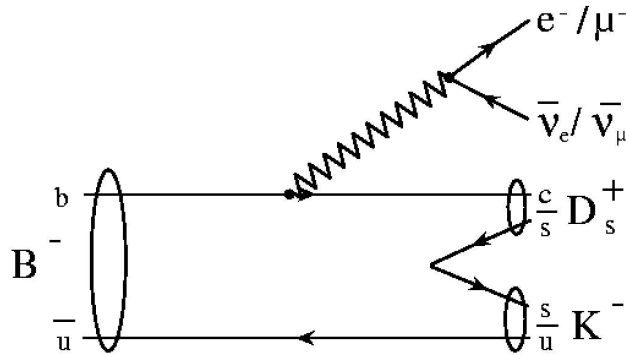


Figure 2.4: Spectator diagram of the $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ decay.

decay, also resonant contributions are possible, where the B first decays into a $D^{**}\ell\nu$ system with the D^{**} then decaying into a $D_s K$ final state.

Motivation for the analysis

In general, the ratio of fragmentations with an $s\bar{s}$ quark pair in the final state to the ones with quarks of the first family are interesting, because it increases the knowledge of the background for measurements of $b \rightarrow s\gamma$, which can be used to determine the CKM matrix element $|V_{ts}|$ and the Wolfenstein expansion value ρ .

Another interesting aspect of measuring the decay $B \rightarrow D_s K \ell \nu_\ell$ are possible dilutions for measurements of B_s oscillations, arising from that decay. These oscillations are similar to $B^0 - \bar{B}^0$ oscillations, where one meson can change its flavor into another through W boson exchange, illustrated by a so-called *Box-Diagram* as shown in figure 2.5. The oscillation frequency of B_s oscillations is proportional to the mass difference between the two B_s mesons. From this mass difference, it is possible to derive a measurement of the ratio of CKM matrix elements V_{td}/V_{ts} . Since the identification of the flavor of the B_s mesons often happens using the decay $B_s \rightarrow D_s \ell \nu$, it is possible to confuse a B_s with a B , if the latter decays into the signal decay.

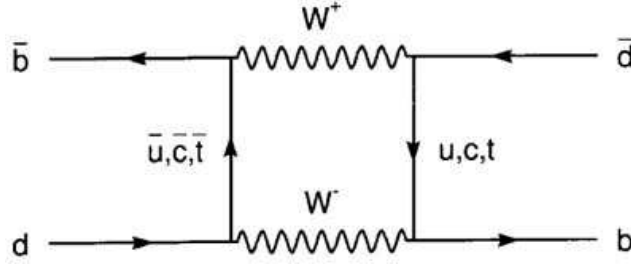


Figure 2.5: Example of a Feynman diagram (Box-Diagram) for B^0 oscillation. For B_s oscillation, the d quark needs to be replaced by an s quark.

In addition, the signal decay is supposed to represent a background contribution to exclusive measurements of $|V_{cb}|$ and to fill the gap between the inclusive and exclusive semileptonic $B \rightarrow X_c \ell \nu$ decays. As mentioned before, it is desirable to understand all contributions to the inclusive lepton momentum spectrum, in order to reduce the uncertainty of V_{cb} measurements and maybe to reduce the discrepancies between the values determined from inclusive and exclusive studies, too.

There is only one previously published measurement of the signal decay, an early ARGUS paper [17], reporting an upper limit of $\mathcal{BR}(B \rightarrow D_s K \ell \nu X) < 8 \times 10^{-3}$, which has been rescaled to $\mathcal{BR} < 5 \times 10^{-3}$ by the PDG, due to updates on the D_s branching ratios [8]. An unpublished, preliminary *BABAR* analysis from the author, using less data and only two D_s final states measured $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = (5.51 \pm 1.94) \times 10^{-4}$ at a statistical significance of 3.2σ .

Theoretical predictions of the branching fraction of the signal decay are difficult, because there are no decay models, describing either the resonant or the non-resonant component. Naively, one can use the results for $B \rightarrow D \pi \ell \nu_\ell$ decays ($\mathcal{BR} = (1.5 \pm 0.6)\%$ [8, 18, 19]), to compute a rough estimate, taking into account the different phase-space due to the fragmentation with the $s\bar{s}$ pair. This ends up in an expected signal branching fraction of $1 \times 10^{-4} \leq \mathcal{BR}_{Signal} \leq 8 \times 10^{-4}$. But since the fraction of the resonant component is completely unknown (and cannot be estimated easily), this is not a precise prediction.

A proper theoretical description of the signal decay is not provided by theorists yet. Nevertheless, there are two decay models, which can be adopted for describing the nature of the decay. The decay model of Goity and Roberts [20] can be used to describe the non-resonant fraction of the signal decay, whereas possible resonant contributions must use the ISGW2 model [21].

Goity-Roberts decay model

The decay model of Goity and Roberts originally describes only the soft pion emission in semileptonic decays. Basically, the four body decay is divided into the weak interaction of the heavy quark and the pion emission of the light quark. HQET is used to construct an effective Lagrangian by implementing inverse heavy mass expansions. Low energy constants, describing the emission of soft pions are estimated in a model, where pions only couple to the light constituent quark of the meson. Resonant contributions of D^* mesons and narrow D^{**} states are taken into account, too.

Table 2.1: *Properties of orbitally excited D^{**} . All mesons, except the D_1 have masses larger than the sum of the masses of the D_s and Kaon, $m(D_s) + m(K) = (2462.2 \pm 0.3) \text{ MeV}/c^2$*

meson	J^P	mass [MeV/c^2]	width [MeV/c^2]
$D_0^*(2400)$	0^+	2352 ± 50	261 ± 50
$D_1^*(2430)$	1^+	2427 ± 36	384_{-105}^{+130}
$D_2^*(2460)$	2^+	2461.1 ± 1.6	43 ± 4
$D_1(2420)$	1^+	2422.3 ± 1.3	20.4 ± 1.7

The calculations end up with a prediction of the branching ratios for $B \rightarrow D\pi\ell\nu_\ell$ of 2.1% and for $B \rightarrow D^*\pi\ell\nu_\ell$ of 0.3%, depending on the HQET input variables used (V_{cb} and the shape of the Isgur-Wise function). These values are in good agreement with experimental measurements, as reported in references [18] and [19].

Unfortunately, this model doesn't describe the $D_s K$ final state, where the narrow resonances play only a small role because their mass is too small. In addition, an estimate of the fraction of decays with a D_s^* in the final state is unknown. Naively, one can expect a ratio of $N(D_s^*) : N(D_s) = 3 : 1$, due to the different possible spin constellations, but this is only valid for non-resonant decays. Thus, there are no reliable predictions for the decay rates, which can be derived from this decay model.

This model could also be used for describing decays with a K^{*0} in the final state instead of a kaon. Due to the larger mass of the K^{*0} , such decays should play only a secondary role, compared to the signal decay. No measurements or calculations for the production of heavier mesons, even for the $D\pi$ final state, have been done yet.

Nevertheless, the modified Goity-Roberts model describes the kinematics of the non-resonant fraction of the decay in an appropriate way and therefore can be used for the simulation of signal events.

Resonant decays and the ISGW2 model

A resonant contribution for the signal decay can arise from semileptonic B decays into the two broad D^{**} states, which then decay into a $D_s K$ system. Both orbitally excited mesons are lighter, than the sum of the masses of the D_s and the kaon, but by having a width of several hundreds of MeV/c^2 , it is possible to contribute to the signal. There might also be a small contribution from the $D_2^*(2460)$ state, which has a mass just above the kinematic limit for $D_s K$ production. An overview of the relevant D^{**} states and their properties is given in table 2.1. In conclusion, all states of the spin triplett ($D_{(0,1,2)}^*$) can contribute to the signal. The different total angular momentum J results in different polarizations of their decay products, which affects angular distributions of the decay products. It may as well change the probability of a creation of final states with a D_s^* instead of a D_s , which affects the signal extraction technique, as will be described in chapter 7. Since there are no accurate estimates on the contributions of the resonant decays, it is not possible to concentrate on one polarization, being equivalent to one spin dominating the others and adjust the analysis to that. Furthermore, it is hard to derive any information about angular distributions of the decay products, if the $J^P = 2^+$ state is the dominant one, because the transfer of an angular momentum of 2 is not easy to predict.

The decay model created by Isgur, Scora, Grinstein and Wise (ISGW2) is used for all semileptonic 3-body decays, hence also for $B \rightarrow D^{**}\ell\nu$ transitions [22]. It is based on a non-relativistic description of the B meson decay, which is modeled by a potential consisting of a Coulomb and a linear component. The solutions of the Schrödinger equation for this potential are used to calculate the matrix elements for the transition. They depend on form factors, being dependent on the 4-momentum transfer q between the initial and the final state meson. This first decay model, called ISGW, was modified in 1995 [21], including relativistic corrections, hyperfine interactions of the meson wave functions and modifications to form factors for meeting the requirements of the HQET for $q_{max}^2 - q^2 > 0$.

In this work, the ISGW2 model is used as an alternative to the Goity-Roberts decay model to estimate systematic uncertainties, arising from the choice of the signal model and thus from the description of the dominant signal fraction.

Chapter 3

The *BABAR* experiment

The *BABAR* experiment at the *SLAC National Laboratory* (previously known as *Stanford Linear Accelerator Center*) in Stanford, USA, as well as the *BELLE* experiment at the *KEK* in Japan, are designed to measure the \mathcal{CP} violation in the system of neutral B mesons. Both experiments produce these mesons through the process

$$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B} \quad (3.1)$$

at a center of mass energy of $\sqrt{s} = 10.58 \text{ GeV}$, which is equivalent to the mass of the $\Upsilon(4S)$ resonance [8]. This resonance decays almost exclusively into a pair of B mesons, because its mass is right above the threshold for the two B production. The large cross-section for the B meson production at the energy of the $\Upsilon(4S)$, shown in table 3.1 together with the cross-sections for other reactions at $\sqrt{s} = m(\Upsilon(4S))$, is the reason for calling the two accelerators B factories.

Table 3.1: e^+e^- production cross-sections at $\sqrt{s} = m(\Upsilon(4S))$ within the experimental acceptance of the *BABAR* detector.

$e^+e^- \rightarrow$	$b\bar{b}$	$c\bar{c}$	$s\bar{s}$	$u\bar{u}$	$d\bar{d}$	$\tau^+\tau^-$	$\mu^+\mu^-$	e^+e^-
Cross-section [nb]	1.05	1.30	0.35	1.39	0.35	0.94	1.16	~ 40

In the center of mass system (CMS), both B mesons are created with only small momenta of $|\vec{p}_B| = 325 \text{ MeV}/c$. The high luminosity of the experiments allows the reconstruction of \mathcal{CP} eigenstates of the B decays which typically have only small branching fractions and the search for New Physics in rare B decays. To measure time dependent \mathcal{CP} asymmetry, two conditions are essential. First, the flavor of the analyzed B meson must be known. Fortunately, the $B\bar{B}$ system produced from the decay of $\Upsilon(4S)$ forms an entangled quantum system, evolving in phase, so that at any time there is exactly one B^0 and one \bar{B}^0 meson present until one of them decays. Thus, by identifying the flavor of the first decaying B , one gets the information about the flavor of the second B , a technique known as *tagging*. The second condition to be fulfilled is a Lorentz boost of the rest frame versus the laboratory frame, so that the second B meson has a measurable flight length, being translated in a lifetime. Such a boost can be brought in by using asymmetric beam energies for the production of the $\Upsilon(4S)$.

3.1 PEP-II asymmetric energy e^+e^- collider

The linear accelerator of the PEP-II asymmetric energy e^+e^- collider accelerates electrons and positrons to energies of $E_{e^-} = 9 \text{ GeV}$ and $E_{e^+} = 3.1 \text{ GeV}$. The particles are inserted into the storage ring and collided in interaction region two. Each collision, which is accepted by the triggers¹ is called *event*. Figure 3.1 shows the schematic setup of the facility.

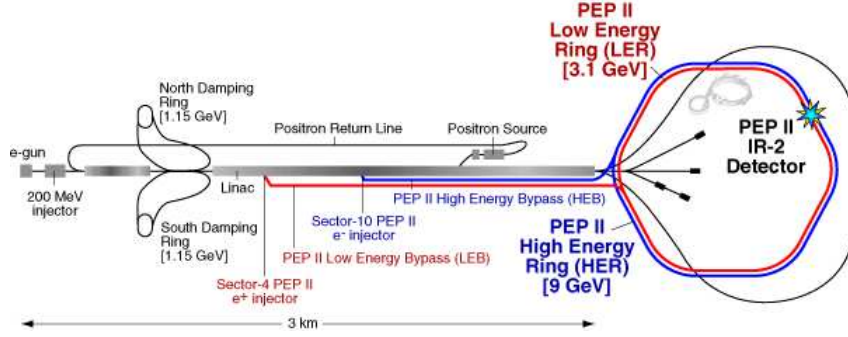


Figure 3.1: The PEP-II storage ring at SLAC National Laboratory.

The different beam energies result in a boost of the CMS of $\beta\gamma = 0.56$ with respect to the laboratory frame along the positive z -axis. The expected event rate in a high energy experiment with colliding beams is proportional to the cross-section of the process. The proportionality factor is called *luminosity* \mathcal{L} : $\dot{N} = \mathcal{L} \cdot \sigma$. The total data collected is expressed by the luminosity integrated over time: $L = \int \mathcal{L} dt$. The design luminosity of the collider $\mathcal{L} = 3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ was reached already in 2001 and has been increased up to a peak luminosity of $12 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ since then. In addition to accumulating $B\bar{B}$ pairs while running at the $\Upsilon(4S)$ resonance energy “on-peak” for 9 years, also “off-peak” data was taken at an energy of about 40 MeV below the $\Upsilon(4S)$ mass during that time. In 2007, an extension of the physics program led to additional data, taken at the $\Upsilon(3S)$ and $\Upsilon(5S)$ resonances. In total, the *BABAR* experiment has recorded an integrated luminosity of about $L = 531 \text{ fb}^{-1}$ of data at different energies, as shown in figure 3.2. As can be seen, the data delivered from PEP-II has been recorded by the *BABAR* detector with high efficiency $\epsilon = \frac{\mathcal{L}_{\text{recorded}}}{\mathcal{L}_{\text{delivered}}}$.

3.2 The *BABAR* detector

The main goals of the *BABAR* experiment, the study of \mathcal{CP} -violating decays of B mesons and the large physics program beyond \mathcal{CP} -violation, e.g. the search for rare decays and New Physics, have had a large influence on the detector design. For \mathcal{CP} physics, a very good vertex resolution, as well as a good separation between kaons and pions is necessary. In addition, a very good lepton identification, especially for electrons is desirable. Furthermore, an efficient detection and high angular acceptance of photons in the whole energy range from 30 MeV up to 4.5 GeV is necessary. An overview of the *BABAR* detector is shown in figure 3.3.

¹A trigger selects only physically interesting collisions before they are analyzed and recorded in order to reduce the computing requirements.

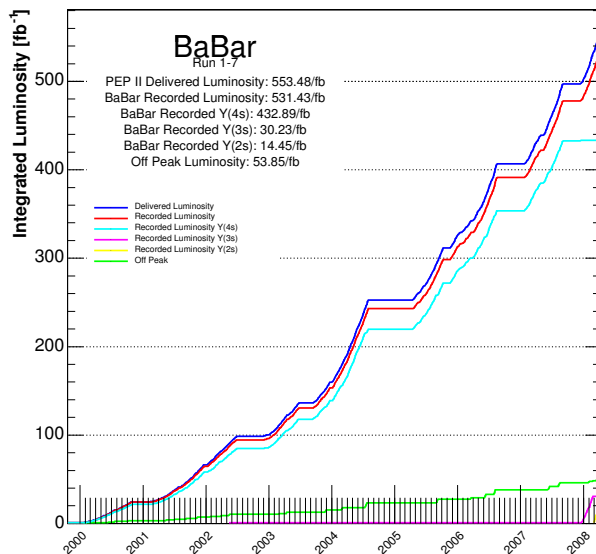


Figure 3.2: Integrated luminosity of the BABAR experiment.

The flight direction of the high energy electrons defines the direction of the boost and is referred to as the forward direction. Due to the asymmetric setup, the interaction point is not exactly in the geometrical center of the detector, but is displaced in the direction of the higher energetic electron beam to guarantee a higher acceptance. The Silicon Vertex Tracker (SVT) is the first detector the collision products pass. Next is the Drift Chamber (DCH), then the Cherenkov Detector (DIRC), and the Electromagnetic Calorimeter (EMC). The Instrumented Flux Return (IFR) is the outermost sub-detector and outside the superconducting magnet coil, which provides a solenoidal magnetic field of 1.5 T.

The following sections are a brief description of the major sub-systems of the BABAR detector. All given information is taken from [23] and [10], where a more detailed description of the various components can be found.

3.2.1 Silicon Vertex Tracker (SVT)

The main tasks of the SVT are the reconstruction of the decay vertices of the two B mesons, as well as providing tracking information for particles with low transverse momenta ($p_t < 100 \text{ MeV}/c$), which do not reach the drift chamber. The former is necessary to gain information about the time between the two B decays for performing time dependent measurements. The latter is important in analyses of decays with slow pions, like $B^\pm \rightarrow D^{*0}\pi^\pm$ or $B \rightarrow D^*\ell\nu$ with $D^* \rightarrow D\pi$. Figure 3.4 shows the schematic view of the SVT.

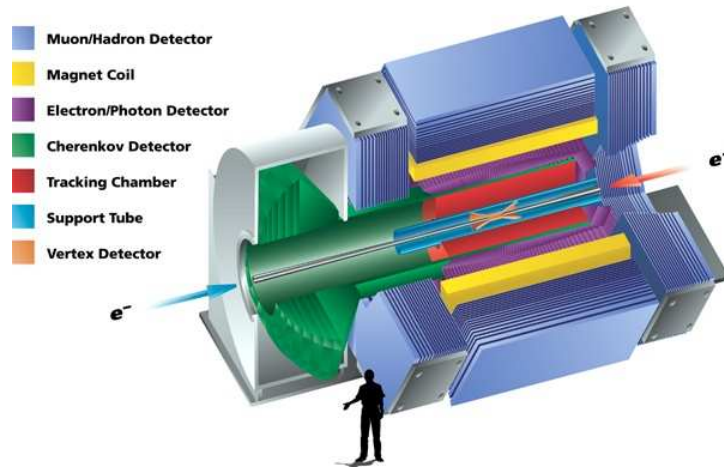


Figure 3.3: Overview of the BABAR detector.

It is built of five concentric cylindrical layers of double sided silicon strip detectors, arranged around the beam pipe. While the inner side of the strips is oriented perpendicular to the beam axis to measure the z -coordinate, the outer side is aligned along the z -axis, providing a precise ϕ measurement. The SVT covers polar angles between 20.1° and 150.2° and provides a vertex resolution of $\sim 70 \mu\text{m}$ for a fully reconstructed B meson, which matches well the average separation of the two B meson decays of approximately $240 \mu\text{m}$. The outermost layer, having a distance of about 11.4 cm to the beam pipe, consists of 18 strip detectors, followed by a layer with 16 modules. The ends of these two layers are tilted towards the beam pipe in order to save material, while the inner three layers, carrying 6 detectors each are aligned completely parallel to the z -axis. The innermost layer is attached directly to the beam pipe at a radius of about 32 mm.

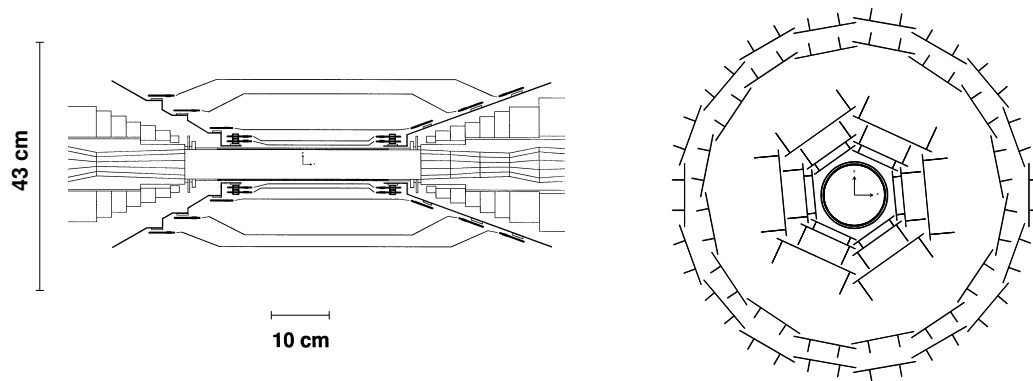


Figure 3.4: Schematic view of the Silicon Vertex Tracker. Left figure shows the side view, while the right figure shows the front view.

3.2.2 Drift Chamber (DCH)

BABAR's multi-wire drift chamber has an outer radius of 80.9 cm, an inner radius of 23.6 cm and a length of 280 cm. It consists of 7104 hexagonal drift cells, each of them containing a 20 μm gold-plated Tungsten-Rhenium signal wire and six gold-plated Aluminum field wires. The cells are filled with a gas mixture of 80% Helium and 20% Isobutane. The high voltage of the signal wires has initially been set to 1900V to save the chamber from aging. The voltage has been raised to 1960V for the second half of the year 2000 in order to improve the single hit efficiency and was lowered again to 1930V as a compromise between aging and hit efficiency. The drift cells are arranged in 10 super-layers, consisting of four layers each. Four super-layers are arranged as axial layers, the other six are stereo layers, three with positive and three with negative stereo angle ((45 – 76) mrad) to achieve a stereoscopic resolution in the z -axis. Figure 3.5 shows a side view of the DCH, as well as the setup of the wires. The

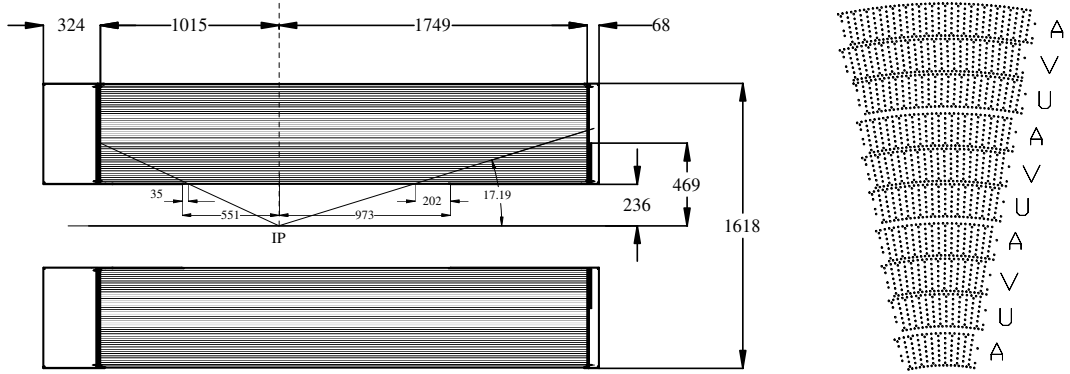


Figure 3.5: The left figure shows a schematic view of the drift chamber (DCH). On the right hand side, the pattern of axial (A) and stereo (U,V) layers is shown.

task of the drift chamber is the track reconstruction of particles with transverse momenta larger than 100 MeV/c. In combination with the track information of the SVT, the *BABAR* detector achieves a resolution of $\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% \cdot p_T / \text{GeV}/c + (0.45 \pm 0.03)\%$. Furthermore, the specific ionization energy loss dE/dx is measured for particles with momenta smaller than 700 MeV/c in order to provide additional particle identification information for the K/π separation in the Cherenkov Detector.

3.2.3 Cherenkov Detector (DIRC)

The DIRC (*Detection of Internally Reflected Cherenkov light*) is built using a novel design for ring imaging Cherenkov detectors, providing a very good K/π separation to momenta up to 4 GeV/c. This is of central importance not only to tag the flavor of the non- \mathcal{CP} B meson, but also for the most other analyses. An excellent kaon identification is a precondition for even performing the analysis, this thesis is about because up to three kaons need to be identified with a misidentification rate as low as possible (as described in chapter 5.5).

The detector uses the Cherenkov effect: When charged particles with a velocity v traverse through a medium with a refractive index $n > c/v$ (c is the speed of light in vacuum), Cherenkov light is emitted at the angle $\cos\vartheta_C = \frac{\sqrt{1+(m/p)^2}}{n} = \frac{1}{\beta n}$, which can be used to differentiate between particle types if the momentum is known (e.g. through the curvature in a magnetic field). The refractive index of the synthetic quartz medium of the DIRC is $n = 1.473$. Thus, kaons start radiating Cherenkov light at momenta above $p \sim 460 \text{ MeV}/c$. At such low momenta, it is not possible to confuse kaons and pions through the specific ionization energy loss dE/dx , measured by the drift chamber. Above momenta of $\sim 700 \text{ MeV}/c$, only the Cherenkov detector can separate kaons and pions from each other. The active

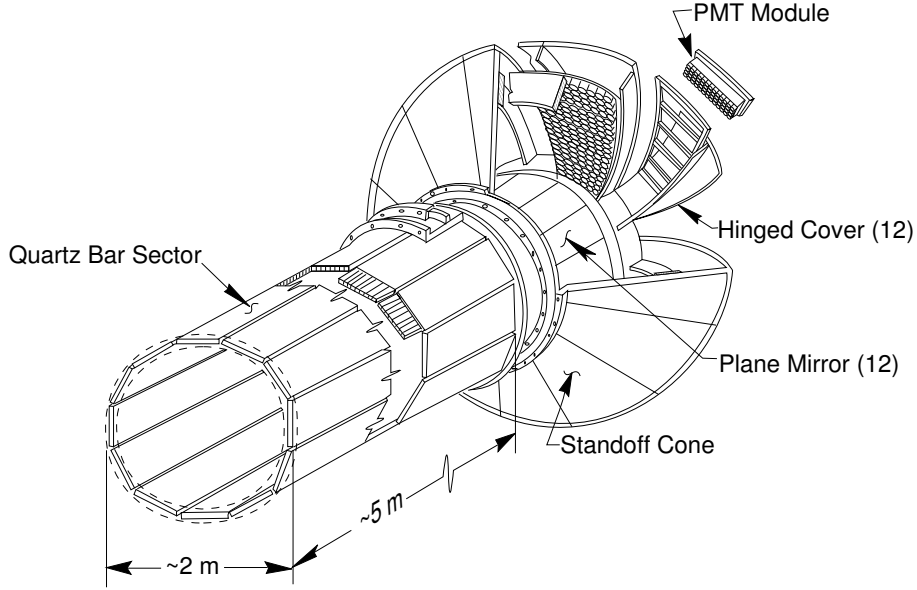


Figure 3.6: Schematic view of the Cherenkov detector (DIRC).

detector material, 144 bars of fused silica with a length of 4.9 m and a cross section of $1.7 \text{ cm} \times 3.5 \text{ cm}$ are arranged as shown in figure 3.6. There are twelve boxes, each containing 12 bars in a dodecagonal shape. Because of gaps between the boxes, the azimuthal coverage of the detector is reduced to 93%, while the polar coverage is about $\sim 87\%$ in the center of mass system. The emitted Cherenkov light is reflected many times inside the bar, keeping the emission angle through total internal reflection, until it reaches the “standoff box” on the cone, a readout reservoir, filled with about 6000 liters of purified water. The inner surface of the reservoir is covered with 10572 photo-multiplier tubes (PMT), which finally detect fractions of the Cherenkov angles. One main advantage of this setup is the low amount of material inside the detector, which leads to a decrease of costs and size of the outer components of the *BABAR* detector.

The information of the drift chamber is used to determine the impact point and angle for each reconstructed track in the fused silica bars. For each of the five particle types (pion, kaon, electron, muon, proton), the measured DIRC signal is compared to the expected signal to calculate a final particle hypothesis.

3.2.4 Electromagnetic Calorimeter (EMC)

In average, a B meson decay produces about 5.5 photons, half of them having an energy below 250 MeV. Thus, a major requirement to the electromagnetic calorimeter was to have an energy threshold which is much below that value. In addition to the detection of gammas, resulting in the reconstruction of neutral pions, the EMC contributes to the electron selection by measuring shower energy and shape, produced by the electrons in the calorimeter. Using this information together with the measurement of dE/dx and the Cherenkov angle by the other sub-detectors, an electron selection with more than 90% efficiency and less than 0.001% pion misidentification rate is achieved. A very good electron identification is a major ingredient for all semileptonic analysis, and thus also for the analysis, this thesis describes.

As a result of the asymmetric beam energies, the electromagnetic calorimeter consists of two parts, a cylindrical barrel around the central detector region, covering $26.9^\circ < \vartheta < 140.8^\circ$ and a conic forward region endcap, covering $15.8^\circ < \vartheta < 26.9^\circ$. The central barrel section is about 3 m long and has a minimum distance to the interaction point of 91 cm. It holds 5760 CsI(Tl) crystals, distributed in 48 rings with 120 crystals each. The endcap section consists of 8 rings, three adjacent to the barrel also containing 120 crystals each, three rings with 100 crystals each and the two innermost rings containing 80 crystals each, summing up to 6580 crystals in 56 rings in total. The crystals have a $4.7\text{ cm} \times 4.7\text{ cm}$ front size, equivalent to a Molière radius of about $R_m = 3.6\text{ cm}$. The short radiation length of $X_0 = 1.85\text{ cm}$ gives the opportunity of precise angular measurements and results in fully contained showers in the EMC. The length of the crystals increases from $16.1 X_0$ (29.76 cm) for all crystals in the backward half of the barrel in steps of $0.5 X_0$ every 7 crystals up to $17.6 X_0$ (32.55 cm) in the most forward barrel module and the endcap (except the two innermost rings, which are shortened by $1 X_0$ because of space limitations).

The crystal itself is transparent for the scintillation light, produced by the primary particle and thus, is working as fiber optic. Every crystal is enveloped with a reflecting foil, directing the photons to the back-side, where two photo diodes are used for the read-out. Figure 3.7 shows the EMC layout and a schematic overview of a single crystal.

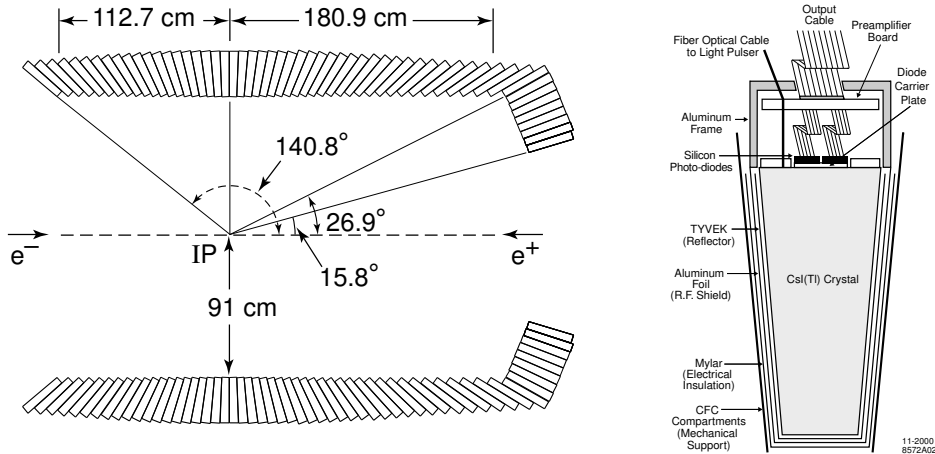


Figure 3.7: EMC layout (left) and schematic view of the EMC crystals (right).

The energy resolution and the angular resolution of the EMC is

$$\begin{aligned}\frac{\sigma_E}{E} &= \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E/\text{GeV}}} \oplus (1.85 \pm 0.12)\% \\ \sigma_\varphi = \sigma_\theta &= \left(\frac{3.87 \pm 0.07}{\sqrt{E/\text{GeV}}} \oplus 0.04 \right) \text{ mrad.}\end{aligned}$$

Due to response of the crystals being dependent on the photon energy and aging effects, a continuous calibration of the EMC is necessary. Three major calibrations at different energy scales are performed:

- at an energy of $E = 6.13 \text{ MeV}$, a radioactive source calibration is done by using γ s from the ^{16}N $\beta - \gamma$ decay cascade,
- $E \sim 5 \text{ GeV}$ Bhabha² events are used to deliver a high energy calibration point,
- $\mu^+ \mu^- \gamma$ events are used for calibrating the region between the low and high energy point (since 2006).

3.2.5 Instrumented Flux Return (IFR)

A toroidal super-conducting magnet coil, creating a 1.5 T magnetic field, is located next to the EMC to allow a momentum measurement from the track curvature. The flux return is located outside the toroid and is segmented into 18 plates of steel, carrying a hadron and muon detection system in the spaces between the plates. The nine innermost plates are 2 cm thick, the outermost 10 cm. Between them, there are four plates with 3 cm and three with 5 cm thickness. The endcap is built by only two plates, one 5 cm thick and one 10 cm. The space between the plates is 3.2–3.5 cm and carries the active component of this detector, the steel works as an absorbing material for hadrons and is helpful in achieving a good hadron muon separation. The geometry of the IFR is shown in figure 3.8

In the years 1998-2004, *Resistive Plate Chambers* (RPC) have been used for the particle detection. RPCs detect streamers from ionizing particles via capacitive readout strips. These chambers are built of two parallel plates of Bakelite with a 2 cm space in-between them. The space is filled with a gas mixture of about 55% argon, 40% freon and 5% isobutane. The two plates are connected to a high voltage of about 8 kV and ground. Thus, the RPCs are working in streamer mode. To provide a spatial resolution, the readout strips are oriented orthogonally. The endcap carries 18 layers of RPCs, while the central region has 19. Two additional, cylindrical layers of RPCs are located between the EMC and the magnet coil to detect hadrons and muons before passing the magnet material.

Starting with the upper sextant in 2004, the Resistive Plate Chambers have been replaced by *Limited Streamer Tubes* (LST) due to problems with the efficiency. The last LSTs were inserted during the long shutdown in 2006. Since almost all data used for this analysis was taken before this upgrade, the LSTs are not explained in detail. More detailed information about LSTs can be found in [24].

²The electromagnetic scattering process $e^+e^- \rightarrow e^+e^-$ is called Bhabha scattering.

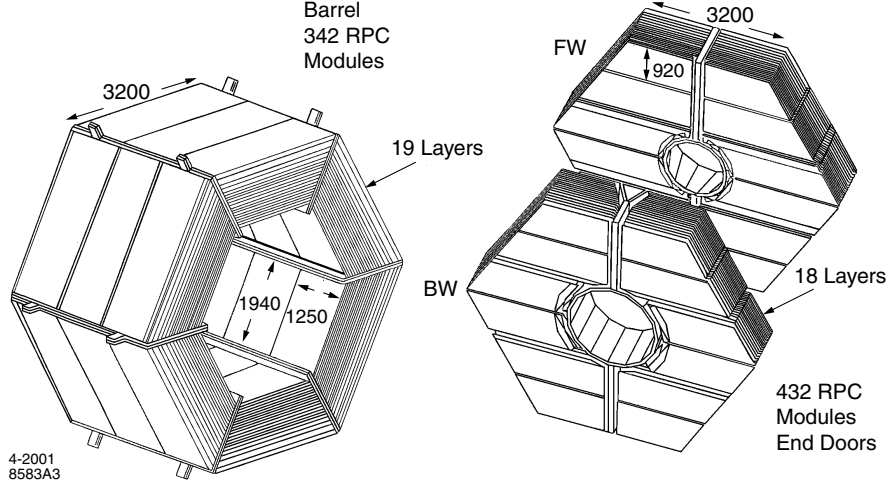


Figure 3.8: Geometry of the Instrumented Flux Return.

The muon identification is almost completely provided by the IFR information (as will be described in section 3.5), since muons penetrate through the whole material, whereas other particles are stopped by it, depending on the transverse momentum of the particles. The complete amount of material, traversed by the particle in radiation lengths is a major criterion for the particle identification. Misidentification of muons as pions is a major source of uncertainty in case of soft particles (= small momenta). The SVT and DCH hits are used to reconstruct the track and extrapolate it into the IFR. Then, all clusters within a defined distance to the track are associated with it. For the reconstruction of K_L^0 and other neutral Hadrons, all IFR clusters, which cannot be associated to a charged track are combined with the information of the EMC.

3.3 Track reconstruction

Tracks are reconstructed by taking into account all available information of interaction of particles with the sub-detectors, e.g. hits in the SVT (and DCH), energy deposit in the EMC, etc. The reconstruction of charged particles starts with a pattern recognition algorithm in the DCH, where parts of the tracks are determined. Then, these tracks are extrapolated into the SVT and hits, consistent with the errors of the extrapolation are added. Remaining hits in these two sub-detectors are combined [25]. Information like the distance to the interaction point and the transverse momentum are added. All reconstructed tracks are stored in the *ChargedTracks* particle list. Since most of the analyses use track quality criteria for background subtraction, additional lists with different track qualities are constructed by using track specific kinematic variables, such as the transverse momentum p_T or the distance of closest approach to the interaction point in the z direction or xy plane. The rigidity of the requirements is represented by suffixes: *VeryLoose*, *Loose* and *Tight*³. Table 3.2 shows the definition of the charged particle lists.

³This kind of classification of selection quality is also used for the particle identification and a naming convention all through the *BABAR* computing environment. In addition to the given names, there's also a *VeryTight* list, which is not defined in all cases - e.g. the track quality.

Table 3.2: Definition of track quality requirements: p_{max} = maximum momentum, d_{IP} = distance of closest approach to the interaction point in xy plane and z direction, p_T = transverse momentum [26, 27].

requirements	ChargedTracks	GoodTracks		
		VeryLoose	Loose	Tight
$p_{max} < 10 \text{ GeV}/c$	-	x	x	x
$d_{IP,z} < 2.5 \text{ cm}$	-	x	x	x
$d_{IP,xy} < 1.5 \text{ cm}$	-	x	x	x
$p_T > 0.05 \text{ GeV}/c$	-	-	x	x
min # of DCH hits	-	-	12	20

The reconstruction of neutral particles is done using the information from the EMC. Clusters are formed around crystals having at least 10 MeV of deposited energy. Adjacent crystals are added to the cluster, if their energy is larger than 1 MeV. If the energy of the newly added crystal exceeds 3 MeV, its neighbors are added to the cluster, too. If several particles enter the calorimeter very close to each other (e.g. in case of high energetic π^0 s), one cluster can contain more than one local maximum and needs to be divided into several pieces. To determine whether a cluster comes from a charged or neutral particle, all reconstructed charged tracks from the inner detector components are projected to the EMC crystals front. If no charged track can be associated with a bump, it is taken as neutral particle (mostly as a γ , depending on the deposited energy and the shower shape) [25].

3.4 Monte Carlo simulation

A Monte Carlo (MC) simulation of the BABAR detector is useful for the determination of efficiencies of the reconstruction, as well as for studies of background processes and the development of selection criteria. The MC production uses the program *EvtGen* [28] to simulate exclusive particle decays like $B\bar{B}$ production and decay. The probability of the decays is given by the known branching ratios of all (composite) particles. Inclusive reactions are simulated using the *JetSet* program [29]. It is used for the production of continuum MC ($e^+e^- \rightarrow q\bar{q}$, $q = u, d, s, c$) and for hadronization processes. The evolution of the produced particles, like their interaction with the detector material or their decays are simulated with a program based on GEANT4 [30]. The signals are read from the electronics and mixed with randomly taken background events. Such events are recorded during the running of the detector when no physically interesting event is noticed. They describe machine background like noise from the sub-detectors, or the PEP-II beams.

After this production cycle, the simulated event basically looks like a real data, except for small differences (due to the incomplete knowledge about the underlying physics) and the existence of the information what kind of process has been produced. This information can be accessed in analysis when running on MC and is called *Truth-Matching*. If the analyst studies a particular decay, e.g. $D_s \rightarrow \phi\pi$, $\phi \rightarrow K^+K^-$ it is possible to check, whether the reconstructed D_s has been produced as D_s using exactly these three tracks by iterating through the Monte Carlo decay tree in the following way:

1. Check if the K^+ , K^- and π are real tracks (they could be fakes, produced by the background noise),
2. Check if the K^+ , K^- and π are produced as K^+ , K^- and π ,
3. Check if the K^+ and the K^- originate from a common mother and if the mother is produced as ϕ ,
4. Check if the ϕ and π originate from a common mother and if the mother is a D_s .

Only if all of these checks return a “true”, the D_s has a MC-Truth partner, it is *truth-matched*.

3.5 Particle identification

By default, all detected particles are treated as pions due to the fact, that about 2/3 of all detected particles are indeed pions. For the final determination of the particle type, several algorithms are used and are free to choose for the analyst. Neural network based algorithms (NN) are available for the selection of kaons and muons only. A selection based on likelihood (LH) ratios of several particle hypotheses can be used for pions, kaons, protons and electrons⁴. In the following, relevant selection algorithms are described briefly.

- The likelihood selection algorithm computes likelihoods for the difference between expected and measured mean of various variables for several particle hypotheses, based on Gaussian distributions with characteristic resolutions. For the kaon, pion and proton selection, dE/dx of the DCH and SVT are used in the low momentum region ($p < 0.6$ GeV/c), while the number of Cherenkov photons and the Cherenkov angle of the DIRC are used in the high momentum region. The selection quality thus represents a cut on the likelihood ratios, which are associated with a defined misidentification rate [25].
- The likelihood selector for electrons *PidLHElectron* uses information from the EMC, the DCH and the DIRC as input. The variables, measured by the EMC are deposited energy divided by the momentum E_{dep}/p , the lateral shower shape and the longitudinal shower shape. The DCH provides information about the specific ionizing energy loss dE/dx . The DIRC information about the Cherenkov angles for pion and kaon hypothesis as well as the number of detected photons is used as well. For all variables, probability density functions are constructed and combined to likelihoods for each particle hypothesis (electron, pion, kaon, proton). As final step, a likelihood fraction (f_L) is computed by weighting the individual likelihood with *a priori* probabilities, taking into account the production rates of the particles. The likelihood fraction can vary between 0 and 1, cuts at specific values represent the selection quality. By default, the quality *Tight* is chosen, being equivalent to a cut at 0.98. Other choices can be *Loose* ($f_l > 0.5$) and *VeryLoose* ($f_l > 0.05$) [15]. It has to be noted, that soft electrons, which do not reach the detector components used for the particle identification due to the low transverse momentum cannot be identified. Thus, the lower limit for identifying electrons is $p \sim 0.5$ GeV/c. The larger the momentum is the larger is the probability of an electron passing the selector’s criteria.

⁴In the document, the terminology ‘electrons’ includes also positrons, unless defined otherwise explicitly.

- The Neural Network selection algorithm for hadrons uses the same input likelihoods as the LH hadron selection, except for the DIRC information. Instead of deriving the Cherenkov angle by fitting a circle to the photo-multiplier hits for each single particle (which is problematic, when only fractions of the Cherenkov ring are recorded), an iterative process is used. An “event likelihood” is computed, which is the overall probability to detect N photo electrons arranged in the same like they are detected. So, it is necessary to evaluate the hypotheses for each track, which maximize the event likelihood (because each track is background for all other tracks). Obviously the algorithm is very complex, because it needs the information from all detected tracks for determining the likelihood for only one track. The subsystem likelihoods are merged with the track momentum in the laboratory frame using a neural network. (A general description of neural networks is given in section 5.3.) Although the output of neural networks in general is a continuous variable between 0 and 1, a single cut on the output defines the selection quality. The five kaon lists⁵, produced are *KNNVeryTight* (NN-output > 0.68), *KNNTight* (NN-output > 0.62), *KNNLoose* (NN-output > 0.50), *KNNVeryLoose* (NN-output > 0.45) and *KNNNotPion* (NN-output > 0.06, this list only suppresses the dilution from pions for events passing the selector) [31].
- The muon neural network selector works very similar as the one for kaons, but uses different input variables: Except for the energy, deposited in the EMC E_{cal} , all other information come from the IFR:
 1. measured number of interaction lengths traversed by the track in the *BABAR* detector λ_{meas} (estimated from the last layer hit by the track in the IFR),
 2. the difference between λ_{meas} and the expectation, $\Delta\lambda$,
 3. the continuity of the track in the IFR T_C ,
 4. the average multiplicity of hit strips per layer \bar{m} ,
 5. the standard deviation of the average multiplicity of hit strips per layer σ_m ,
 6. the χ^2/NDF values of the IFR strips w.r.t. a 3rd order polynomial fit of the cluster χ_{fit}^2 and
 7. the χ^2/NDF values of the IFR strips in the cluster with respect to the track extrapolation χ_{mat}^2 .

The dependency of the selection efficiency on the polar angle Θ and the muon momentum has to be taken into account. In order to get a continuous sample, the requirement for the selection quality is the selection efficiency, which remained constant for all Θ and p_μ intervals. Muons pass the *muNNVeryTight* selection with $\sim 60\%$ efficiency, the *muNNTight* selection with $\sim 70\%$, the *muNNLoose* selection with $\sim 80\%$ and the *muNNVeryLoose* selection with $\sim 85\%$. Basically the same variables are used in the cut based selection, but by using the neural network, the efficiency of the selection increases by $\sim 8\%$ [32]. Since the most important detector component for the muon identification is the IFR, the minimum momentum, a muon needs for being identified ($p_\mu > 0.8 \text{ GeV}/c$) is larger compared to the electron selection. The selection efficiency rises with the momentum and reaches its maximum at $p_\mu > 1.3 \text{ GeV}/c$.

⁵The Neural Network based selectors are created only for the kaon selection.

The most important variables, describing the quality of a selection are the selection efficiency and the misidentification rate. Both variables depend on the momenta of the particles. The determination of these quantities is done by the *BABAR Particle Identification Group* (PID-Group) using samples with clean signatures for the single particle types, e.g. (radiative) bhabha events for electrons, initial state radiation $e^+e^- \rightarrow \mu^+\mu^-\gamma$ for muons, K_S^0 decays for pions, D^* decays for slow pions and kaons and $\Lambda \rightarrow p\pi$ decays for protons [33]. In general, a high efficiency at a low misidentification rate is desirable, but more often one has to come to a trade-off between both values, because an increasing efficiency usually leads to an increasing misidentification rate, too. The different selection qualities basically represent such trade-offs and give the analyst the opportunity to optimize the selection for his particular decay.

Differences in the performance of the particle selectors between data and Monte Carlo can lead to wrong results and/or large systematic uncertainties, arising from deviations in the reconstruction efficiencies for different decays. Using the clean control samples again, the selection performance in MC in dependency of the particle momentum and polar angle ϑ is compared to the one in data. Two methods for correcting the differences⁶ were developed by the PID-Group [34]:

- The *PIDWeighting* performs a re-weighting of tracks passing the selectors with the weight $\epsilon_{data}/\epsilon_{MC}$, depending on the momentum, the angle ϑ and the information about the type of the created particle (coming from the Monte Carlo information).
- The *PIDTweaking* also performs a re-weighting of tracks passing the selection with the weight $\epsilon_{data}/\epsilon_{MC}$, but additionally flags tracks as “selected” and adds them to the according list, if $\epsilon_{data} > \epsilon_{MC}$. In this case, the probability $1 - \epsilon_{data}/1 - \epsilon_{MC}$ is used. If $\epsilon_{data} = \epsilon_{MC}$, the original state of the track is kept. It is noticeable that fake rates with $\epsilon_{data} > \epsilon_{MC}$ happen very often.

This analysis uses the *PIDTweaking* for the correction of the selection efficiency. The evaluation of the systematic uncertainties arising from the correction itself is described in detail in section 8.2.3.

⁶In most of the cases MC overestimates data, due to better track quality through insufficient simulation of material interaction and high uncertainties in the simulation of bremsstrahlung.

Chapter 4

Data sample

This section gives an overview of the data and simulated Monte Carlo (MC) samples, used in the analysis. It presents the different types of simulated events and describes the challenges of producing signal Monte Carlo events without a decay model, designed for the signal decay.

4.1 Data

This analysis is based on an integrated on-peak luminosity of 341.8 fb^{-1} , recorded in the runs 1 to 5 in the years 1999-2006 (see figure 3.2). This corresponds to about 377 million $B\bar{B}$ pairs. In addition, the off-peak data, recorded within the same time period ($\mathcal{L} \simeq 36.6 \text{ fb}^{-1}$), has been used for studies to choose the appropriate analysis technique. The used sample represents a fraction of about 80% of the total $BABAR \Upsilon(4S)$ data sample¹.

4.2 Monte Carlo samples

All Monte Carlo samples were produced within the $BABAR$ SP9 (*Simulation Production 9*) production series [35], being the most recent at the time the analysis was performed. The production series differ only in terms of updates in the detector description and the underlying software (e.g. GEANT), newly observed decays or resonances and updated measurements of single branching fractions. One distinguishes three major MC samples:

1. Generic $B\bar{B}$ Monte Carlo, containing a simulation of the known B decays,
2. Generic continuum Monte Carlo, simulating the generation and decay of $e^+e^- \rightarrow q\bar{q}$, $q = u, d, s, c$ processes and
3. Signal Monte Carlo, where one B is forced to decay exclusively in the signal channel whereas the other B decay remains unchanged.

There is about three times the data luminosity in generic $B\bar{B}$ Monte Carlo, to allow for sufficient statistics for the analysis of backgrounds for rare decays searches. In this analysis, all available generic Monte Carlo events have been used.

In the following, a brief description of the three samples is given. Table 4.1 at the end of this section summarizes the size and type of all Monte Carlo samples used.

¹The last $BABAR$ run finished, when the analysis already was in an advanced state, so it was decided to complete the analysis, rather than adding the remaining data.

4.2.1 Generic $B\bar{B}$ Monte Carlo

The generic $B\bar{B}$ Monte Carlo simulates the formation of the $\Upsilon(4S)$ and its decay in pairs of charged or neutral B mesons. All known B meson decays are included and their measured branching ratios are used to describe the $BABAR$ data as good as possible. Semileptonic B meson decays into mesons containing a charm quark are simulated using different kinematic models, e.g. the ISGW2 model, introduced in section 2.2.2. For non-resonant decays with an additional pion in the final state, the Goity-Roberts model is used.

Since the signal decay has not been observed yet, it is not contained in the generic $B\bar{B}$ Monte Carlo and was produced separately.

Generic $B\bar{B}$ Monte Carlo is used for the development of selection criteria, for background studies and to validate the fit method as described in chapters 5 and 7.

4.2.2 Generic continuum Monte Carlo

Although data recorded at the $\Upsilon(4S)$ resonance has a high cross section for $B\bar{B}$ production, there is also a large fraction of continuum events in the on-peak data sample. Thus, simulated samples of continuum events are necessary to provide an accurate simulation of the $BABAR$ data. Alternatively, it would be possible to use off-peak data to study the contribution of continuum events. But since running at the lower energy reduces the available $\Upsilon(4S)$ data samples and therefore the luminosity for B -physics analyses, a trade-off between running at the different energies had to be found. Two samples of continuum events are provided, one containing only $c\bar{c}$ events and one containing a cocktail of $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$ quarks, mixed with the cross sections at the beam energy, as given in table 3.1.

In this analysis, continuum MC samples are used to study the contribution of continuum events to the expected signal yield.

4.2.3 Signal Monte Carlo simulation

As described in section 2.2.2, possible signal contributions arise from non-resonant decays and decays through an orbitally excited D^{**} -meson. Thus, several samples of signal MC have been produced independently from the $BABAR$ Monte Carlo production cycle.

As already pointed out, the decay model of Goity and Roberts is originally designed for the description of $B \rightarrow D^{(*)}\pi\ell\nu$ decays, but no calculations have been done for the $D_s K$ final state. Thus, a modification of the decay model was necessary to adopt the model for the signal decay simulation. The major changes with respect to the $D^{(*)}\pi$ final state appear due to the different masses and decay constants of the participating mesons. A validation of the modified Goity-Roberts model has been done by comparing the lepton momentum distributions for the $D\pi$ final state with the $D_s K$ final state. Both distributions should show the same shape, if the adoption was successful. The comparison is shown in figure 4.1.

As an alternative approach to simulate signal decays, a manipulation of a semileptonic B decay into a doubly excited D meson, a D_0^{*0} , has been used. This decay is described appropriately within the Isgur Wise decay model (ISGW2). The mass of the D_0^{*0} mesons, produced through the decay $B \rightarrow D_0^{*0}\ell\nu$ have been forced to be larger than the sum of the masses of D_s and kaon and the D_0^{*0} is forced to decay into the $D_s K$ system.

Again, the lepton momentum distribution is compared to the ones produced by the two Goity-Roberts decay models to validate this type of signal Monte Carlo. Figure 4.1 shows this comparison of the lepton momentum spectra of both decay simulations.

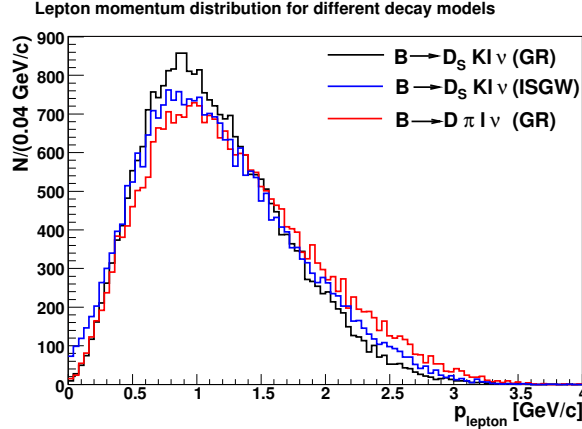


Figure 4.1: Comparison of the momenta of MC simulated leptons in the lab frame, black: signal MC, produced with the modified Goity-Roberts Model, blue: signal MC, produced by a manipulation of the ISGW2 Model, red: $B \rightarrow D\pi l\nu$ MC, produced with the Goity-Roberts Model. The distributions are normalized to the same number of events.

As can be seen, the spectrum of the signal events simulated by using the Goity-Roberts model shows a good agreement with the $B \rightarrow D\pi l\nu$ reference, especially in the low energy region. The differences between both distributions can be explained by the mass difference between the particles in the final states of the both decay channels. The ISGW2 simulated model shows an non-physically high amount of leptons with very low momenta, probably a relict of the decay manipulation. At high energies, there remains a large difference between this type of signal decay simulation and the more reasonable Goity-Roberts like simulation.

In conclusion, the modified Goity-Roberts model is the most accurate simulation of the signal decay and is used for optimizing the candidate selection and the determination of the reconstruction efficiency. This MC sample is referred to as default signal MC sample throughout this document. The manipulated ISGW2 simulation and an additional signal MC sample, produced by using only a simple phase-space model² have been used for the evaluation of systematic uncertainties, arising from the choice of the signal Monte Carlo decay model (see section 8.2.1). A signal MC simulation with the default decay model, but without Bremsstrahlung emission is used to study systematic uncertainties, arising from the simulation of Bremsstrahlung (see section 8.2.6).

²This kind of simulation neglects all physical knowledge about the simulated process, except 4-momentum conservation.

The default signal Monte Carlo sample produces an $\Upsilon(4S)$, which decays into a charged pair of B mesons. One of them is forced to decay into $D_s K \ell \nu_\ell$ using the modified Goity-Roberts decay model, while the other B meson decays randomly by using the same information as generic $B\bar{B}$ Monte Carlo. The D_s meson produced is forced to decay exclusively into one of the three reconstructed final states, $\phi\pi$, $\bar{K}^{*0}K$ or $\bar{K}_S^0 K$. The first two decay modes produce a $K^+ K^- \pi^\pm$ final state, while the K^0 decays as K_S^0 into a $\pi^+ \pi^-$ system, forming a $\pi^+ \pi^- K^\pm$ final state. The branching fractions of all described daughter decays are set to 100%. Each particle produced is allowed to emit Bremsstrahlung. The latter is simulated using a software package called *PHOTOS* [36].

For each D_s decay channel, each lepton channel ($\ell = e/\mu$) and each type of signal MC, 50000 events have been produced. At a signal branching ratio of $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = 5 \times 10^{-4}$, this is equivalent to 13-35 times the data luminosity, used in this analysis. The exact numbers depend on the D_s decay channel and can be seen in table 4.1. In addition, the same amount of simulated events has been produced for the decay $B^- \rightarrow D_s^{*+} (D_s^+ \gamma) K^- \ell^- \bar{\nu}_\ell$, in order to study the possibility to differentiate between the signal and this decay when the analysis doesn't consider the photon in the signal reconstruction.

Table 4.1: Monte Carlo event samples, used for cut development and validation of the fit method. The signal MC is produced in four different setups: a) Goity-Roberts decay model, Bremsstrahlung emission enabled, b) ISGW2 decay model with Bremsstrahlung, c) Phase-space decay model with Bremsstrahlung, d) Goity-Roberts decay model without emission of Bremsstrahlung. The equivalent luminosity of the signal MC samples is calculated using the branching ratios from [8] and $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = 5 \times 10^{-4}$.

Data Set	B/ \bar{B} decay modes	equiv. lumi.	$N_{B\bar{B}}$
Signal MC			
e channel	$B \rightarrow D_s(\phi\pi)K\ell\nu_\ell$ / <i>Generic</i>	8437.6 fb $^{-1}$	50000
e channel	$B \rightarrow D_s(K^{*0}K)K\ell\nu_\ell$ / <i>Generic</i>	6993.4 fb $^{-1}$	50000
e channel	$B \rightarrow D_s(K_S^0 K)K\ell\nu_\ell$ / <i>Generic</i>	17634.7 fb $^{-1}$	50000
μ channel	$B \rightarrow D_s(\phi\pi)K\mu\nu_\mu$ / <i>Generic</i>	8437.6 fb $^{-1}$	50000
μ channel	$B \rightarrow D_s(K^{*0}K)K\mu\nu_\mu$ / <i>Generic</i>	6993.4 fb $^{-1}$	50000
μ channel	$B \rightarrow D_s(K_S^0 K)K\mu\nu_\mu$ / <i>Generic</i>	17634.7 fb $^{-1}$	50000
Generic MC			
B^0 generic	<i>Generic</i> / <i>Generic</i>	1076.6 fb $^{-1}$	592.1×10^6
B^\pm generic	<i>Generic</i> / <i>Generic</i>	1094.9 fb $^{-1}$	602.2×10^6
$c\bar{c}$	-	712.3 fb $^{-1}$	926.0×10^6
$u\bar{u}, d\bar{d}, s\bar{s}$	-	352.6 fb $^{-1}$	736.9×10^6

Chapter 5

Reconstruction and event selection

Several types of background are important for this analysis and need to be reduced, before an extraction of the signal yield is possible. In this analysis, major sources of background are:

- continuum events from $e^+e^- \rightarrow q\bar{q}$, where $q = u, d, s, c$,
- cascade leptons, where the lepton is not coming from a B decay, but originates from one of its daughters or their decays, e.g. through the reaction $b \rightarrow c \rightarrow X\ell\nu$,
- combinatorial background from the reconstruction of composite particles, where tracks are combined, which are not originating from a common mother particle or from a D_s
- background with correctly reconstructed D_s
- fake kaons or leptons, constructed from hits in the DCH, where no real particle passed
- particles with a wrong particle identification

The most important background arise from cascade leptons, combined with a kaon and a correctly or even an incorrectly reconstructed D_s candidate. While the latter can be rejected through the development of various selection criteria, the background with correctly reconstructed D_s , in the following referred to as *True- D_s background*, is hard to suppress. In such a case, one B meson produces the D_s or a D_s^* , while the other one produces the lepton and kaon candidate. Thus, the reduction of the leptons, available for forming B candidates is essential. Fake kaons or leptons and particles with wrong identification represent only a small fraction of background events. This is also true for the background arising from continuum events.

5.1 Optimization of the selection

The reduction of background through the development of selection criteria for various variables is a common approach to extract a signal out of a large amount of background. The choice of the cut values is often crucial, since it determines the accuracy and the error of the measurement. Therefore, an optimization of the selection is needed to maximize the information extracted from the dataset available.

In general, the amount of background B and the signal yield S are the relevant quantities to evaluate the quality of a selection criterion¹. In some cases, the usage of the signal efficiency, calculated between the number of reconstructed signal events S_{rec} and the number of produced signal events S_{prod} , $\epsilon = S_{rec}/S_{prod}$ instead of S is more convenient. Depending on the analysis goal, different combinations of variables are used for the determination of the optimum cut value [37]:

1. The maximization of $S/\sqrt{S+B}$, often called *significance*, is widely used for the cut optimization. Its maximum represents the optimum choice of a cut, if one wants to reduce the uncertainty of a measurement with a known branching ratio for the signal decay. If the cross section of the signal being searched for is unknown, the expression shouldn't be used, since S is directly dependent on the expected signal branching ratio, while B is not.
2. ϵ/\sqrt{B} or S/B are both commonly used and are independent from the cross section of the new process, due to being linear in S . But these expressions have the problem of breaking down at small values of B . Imposing the maximization of ϵ/\sqrt{B} may push the reconstruction efficiency down to very small values.
3. $S/\sqrt{B+A}$ or $\epsilon/\sqrt{B+A}$, with A being an arbitrary value larger than zero, is the more sophisticated expression. It features an optimization for searches for new decays or particles through the linearity in S , while the insertion of A saves the quantity from breaking down for small background yields.

For this analysis, a quantity derived from expression 3. is the best choice for the cut optimization. Choosing $A = 5$, the expression $\epsilon/\sqrt{B+5}$ has been used and will be referred to as *significance* in the rest of this thesis.

5.2 Reconstruction of composite particles

The first step of the reconstruction of the signal decay from the tracks stored in the *BABAR* framework is to reduce the number of events, that have no chance at all to contribute to the signal. To save computing time, a couple of requirements must be satisfied, before a reconstruction is even attempted:

- There must be at least one lepton in the event.
- There must be at least one loosely identified kaon in the event.
- There must be at least five tracks in the event.
- There must be more than one hadron in the event.

¹The amount of both fractions is determined by using simulated events.

Since the D_s meson is reconstructed only in the three decay modes $D_s \rightarrow \phi\pi$, $D_s \rightarrow \bar{K}^{*0}K$ and $D_s \rightarrow \bar{K}^0K$, the reconstruction starts with a preselection. Candidates for the composite particles (ϕ , K^{*0} , K_S^0 and D_s) are reconstructed within broad mass windows, again reducing the number of events to be analyzed. The D_s candidates are then combined with an additional kaon satisfying the *KNNVeryTight* selection and a lepton with the appropriate charge constellation to form a B candidate. Only if this final combination into a B meson succeeds, the properties of all the candidates in the decay tree are stored for the further analysis.

Relevant parameters for the D_s reconstruction, such as the branching ratios for important decays, as well as the mass and width of the daughter particles and resonances in the three decay modes are shown in table 5.1.

Table 5.1: Relevant parameters for the D_s reconstruction. The branching fractions have been obtained from PDG 2008 [8].

particle	mass and width/lifetime	$J^{P(C)}$	relevant decay modes BF
ϕ	$(1019.46 \pm 0.02) \text{ MeV}/c^2$ $(4.26 \pm 0.04) \text{ MeV}/c^2$	1^{--}	$K^+K^-: (49.2 \pm 0.6)\%$ $K_L^0K_S^0: (34.0 \pm 0.5)\%$
K^{*0}	$(896.00 \pm 0.25) \text{ MeV}/c^2$ $(50.3 \pm 0.6) \text{ MeV}/c^2$	1^-	$K^+\pi^-: 66.67\%$ $K^0\pi^0: 33.33\%$
K_S^0	$(497.61 \pm 0.02) \text{ MeV}/c^2$ $(0.895 \pm 0.001) \times 10^{-10} \text{ s}$	0^-	$\pi^+\pi^-: 69.20 \pm 0.05\%$ $\pi^0\pi^0: 30.69 \pm 0.05\%$
D_s^+	$(1968.5 \pm 0.3) \text{ MeV}/c^2$ $(500 \pm 7) \times 10^{-15} \text{ s}$	0^-	$\phi\pi^+: (4.38 \pm 0.6)\%$ $K^{*0}K^+: (3.9 \pm 0.4)\%$ $\bar{K}_S^0K^+: (1.49 \pm 0.9)\%$

The ϕ reconstruction works as follows: Only the decay into a pair of charged kaons is reconstructed to form ϕ candidates. One track needs to satisfy the *KNNNotPion* selection criterion and the second one the *KNNLoose*. Both lists use the *GoodTracksVeryLoose* (GTVL) list as input [31]. In this analysis, this setup represents the combination giving the best significance, as detailed studies showed.

The K^{*0} is reconstructed by using the decay into a $K\pi$ pair. Only the decay into charged daughters is used, which has a fraction of $2/3$ of the whole $K^{*0} \rightarrow K\pi$ branching ratio. Because of the large width of the K^{*0} , the kaon list, used for its reconstruction, is tighter than the ones used in the ϕ reconstruction. A track of the *KNNTight* list is combined with a charge conjugated track of the *GoodTracksVeryLoose* selection. A third track of the *KNNLoose* list is then added to the K^{*0} candidate to build the D_s . This configuration provides the maximum significance, again.

The reconstruction of the K_S^0 is done by combining two oppositely charged tracks from the *ChargedTracks* list to form a K_S^0 candidate. Since the K_S^0 reconstruction is very clean, the major background source comes from the identification of the kaon which is used to form the D_s candidate. Thus, the *KNNTight* list has been chosen to get a reasonable amount of D_s candidates, while not creating a large fraction of combinatorial background.

5.3 Separation of continuum events

Continuum events can be differentiated from $B\bar{B}$ events by using the differences in the resulting momenta of the formed particles. In a $B\bar{B}$ event, the energy of the beams is just enough to produce the $\Upsilon(4S)$, decaying into two slow B mesons (in the $\Upsilon(4S)$ rest frame). Thus, the topology of each decaying B meson is nearly isotropic and no direction is favored. Events producing light quarks ($u\bar{u}/d\bar{d}/s\bar{s}$) will have a more jet-like shape, since their production consumes less energy than provided by the beams and the overshoot results in high momenta of the light particles. In addition, there will be a strongly preferred direction, characterizing the full event. Thus, such events will not have an isotropic event shape and there will be angular correlations between the two B candidates, created from the tracks. In $c\bar{c}$ events, this is valid too, but due to lesser energy available for the momenta of the mesons produced, the discrimination power of the variables will be worse.

Since it is not possible to determine the type of the quarks produced by the colliding beams before the reconstruction, the production of a $B\bar{B}$ pair is used as starting hypothesis. One B candidate is reconstructed and all remaining tracks (the *rest-event*) are interpreted as second B^2 . In this analysis, three variables are used to reject continuum events. Each of them uses either the topology of the B candidates or angular correlations between them and the rest-event. These variables are:

- the ratio between the second and zeroth Fox-Wolfram moment $R2 = \frac{H_2}{H_0}$, with $H_l = \sum_{i,j} \frac{|\vec{p}_i| \cdot |\vec{p}_j|}{E_{vis}^2} P_l(\cos \vartheta_{ij})$ [38], where P_l are the Legendre polynomials, $\vec{p}_{i/j}$ are the momenta of the particles, ϑ_{ij} is the opening angle between the particles i and j and E_{vis} is the total visible energy of the event.
- the cosine of the angle between the momentum of the B candidate and the direction which maximizes the sum of the longitudinal momenta, the *Thrust-Axis* of the rest event \hat{T} , $\cos \Theta_{Thrust}$ [10] and
- The second Legendre monomial $L2$, defined as $L2 = \sum_i |\vec{p}_i^*| \cos^2 \vartheta_i^*$, where the sum is over all tracks of the rest of the event and \vec{p}_i^* and ϑ_i^* are the momenta and angles measured with respect to the thrust axis of the B candidate.

While $R2$ uses the shape of the B mesons, the other two variables use angular correlations. The distributions for signal (i.e. $B\bar{B}$ events) and continuum background of all parameters is shown in figure 5.1. All variables are combined in one discriminating variable to cut on by using a novel neural network technique. The construction of the network is done using a software package called TMVA [39]. It provides various algorithms to construct one discriminating variable out of a set of input variables. Besides neural networks, which use weights to account for the different separation power of the input variables, there is also the possibility to use linear combinations of the variables like a fisher discriminant. A comparison of four algorithms has been done for the continuum background suppression.

²Each particle, which has not been identified as kaon, lepton or proton is treated as pion. Assigning a correct particle ID is essential for a correct calculation of the event shape variables.

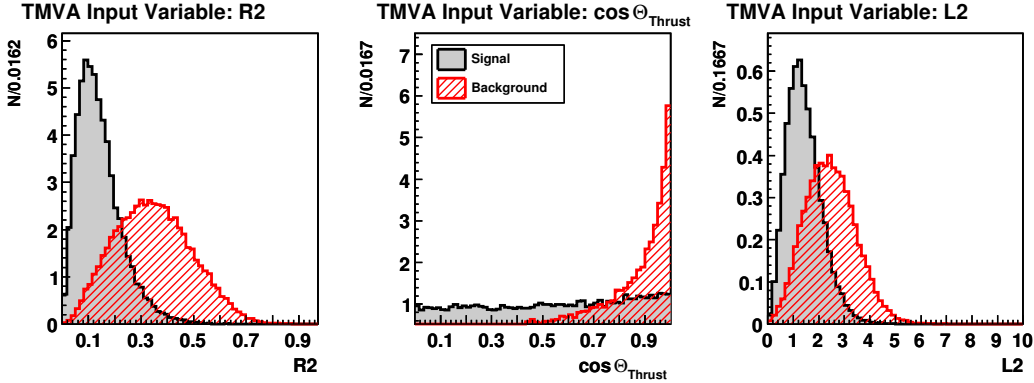


Figure 5.1: Scaled distributions of the MLP input variables $R2$, $\cos \Theta_{Thrust}$ and $L2$, black $B\bar{B}$ Monte Carlo, red: continuum Monte Carlo events

Following is a short description on four of the available methods, focusing on the neural network. A more detailed discussion can be found in reference [40].

- The *CutsD* method is just used for comparing a cut-based selection with the other methods. Unlike all other classifiers in TMVA, it returns only a binary response (signal or background).
- The (maximum) *Likelihood(D)* method basically consists of a model built out of probability density functions, that reproduce the input variables for signal and background.
- The *fisher* discriminant uses a transformed parameter space to distinguish between signal and background.
- The *Multi Layer Perceptron* (MLP) method is an artificial neural network algorithm, where the complexity is reduced by organizing the variables in layers. Only directional connections from one layer to the immediate next one are allowed. The input variables are the first layer and the output variable is the last layer. All layers in-between are hidden layers. A weight is assigned to each connection between the output of one neuron and the input of another neuron. The value of the output neuron is then multiplied by the weight and used as input value for the next neuron. The MLP for the continuum rejection uses four hidden layers. The default configuration requests $N + 1$ hidden layers, with N being the number of input parameters. The weights of the single layers are determined by training the neural network, for which a signal and a background sample needs to be provided. Each training iteration is equivalent to fitting a model represented by the network to the training data sample and determining the network error of this model. The error of a particular configuration is determined comparing the actual output for each input set with the desired or target output. The differences are combined together by an error function (e.g. squared errors) to give the network error. The training of the network is finished when a minimum of the neural network error has been found. The final set of weights represents the model with the smallest network error and therefore the best separation power between background and signal. The performance and quality of the training thus depends on the training samples, which need to be chosen wisely and should contain enough data sets to allow an effective training.

In general, neural networks already take care of possible correlations between the input variables (e.g. $L2$ and $\cos\Theta_{Thrust}$ are highly correlated in this case). Another advantage with respect to a cut based selection, where one needs to determine the optimum choice of the cut value for all three variables, is the saving of time and effort to optimize the order of the cuts. In iterative cut based analyses, the order of the cut often has an influence on the gradient of the significance. By changing the cut order and iterating cuts on the three variables over and over, searching for a new maximum of the significance in each step, an optimum sequence of cuts need to be found. This is only necessary, if the variables used are correlated. A third advantage of the usage of neural networks lies in the combination of variables, which all might have only low discrimination power and maybe even worsen the significance, when cutting on all of them. By combining them wisely using appropriate weights, it is possible to create a more powerful variable, which can be used for the background rejection. It has to be noted, that this only works if there's at least a small correlation between the input variables.

As can be seen in figure 5.2, the MLP and the LikelihoodD method give similar results regarding background rejection vs. signal efficiency, whereas the performance of the fisher discriminant is significantly worse. The CutsD method shows a typical behavior (in terms of fluctuations) for a combination returning a binary response. TMVA neural networks are

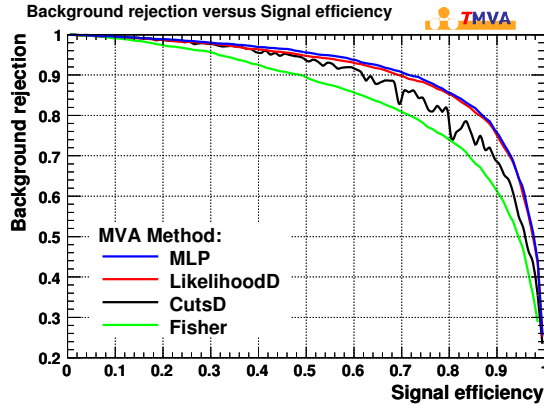


Figure 5.2: Comparison of the background rejection vs. signal efficiency between different TMVA algorithms.

used several times in this analysis for combining appropriate variables and providing an optimal background rejection. In order of being consistent concerning the optimal choice through the complete analysis, the MLP algorithm has been chosen. The MLP response and cut efficiencies for signal and background are shown in figure 5.3. Requiring the MLP response to be $MLP_{Eventshape} > 0.5$ results in a background rejection of $\sim 80\%$, while the signal efficiency remains at $\sim 85\%$. This optimal cut value has been derived by maximizing the significance. An overview of all cuts and their efficiencies for all three D_s reconstruction chains is given in the tables 5.2-5.4 at the end of this chapter.

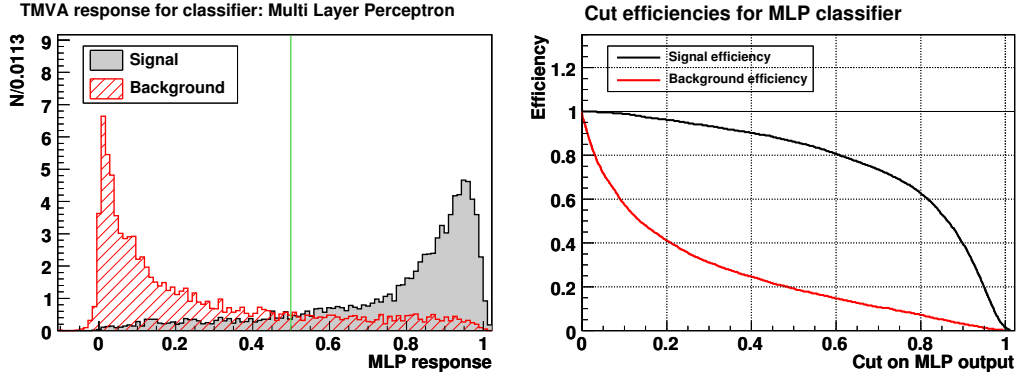


Figure 5.3: MLP response and cut efficiencies for signal (black) and background (red) for the MLP of event shape variables. The green line indicates the cut value.

5.4 Separation of cascade leptons and fake leptons

The suppression of cascade leptons, as well as of fake leptons is a major issue for this analysis. For the latter, a proper lepton identification is very important. Thus, selectors with a low misidentification rate are chosen, leading to a smaller, but cleaner sample in terms of lepton identification. The selectors used are the *PidLHElectron* selector for electrons and the *muNNVeryTight* selector for muons (see section 3.5 for the definition of the selectors). A further reduction of the fake rate for the lepton Id, as well as a reduction of the cascade leptons is done by introducing a cut on the lepton momentum. Cascade leptons show a softer momentum spectrum than leptons directly coming from B decays, due to less energy available, which can be seen in figure 5.4.

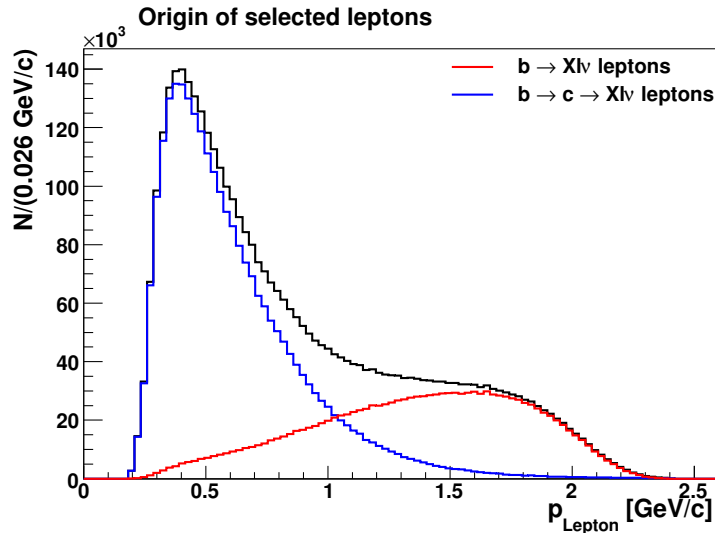


Figure 5.4: Content of the lepton momentum spectrum of $D_s K \ell$ combinations in generic B MC, red: leptons, coming from B decays, blue: leptons, coming from cascade decays, black: sum of both fractions

Since the cut on the lepton momentum, as well as the quality of the reconstruction of the D_s has a large impact on the number of signal and background B candidates in the final data sample (and therefore on the efficiency), it is a non-trivial question whether the cut has to be applied before or after the D_s reconstruction to achieve the best significance possible. Thus, the determination of the optimum cut values is done using a two-dimensional cut optimization of the lepton momentum and the MLP used for the reconstruction of the D_s daughter particles. A more detailed description of the optimization process, together with the final cut values is presented in section 5.5.2.

5.5 Combinatorial background suppression

While combinatorial background suppression for the $K_S^0 K$ decay chain is done by using one of the standard *BABAR* K_S^0 recipes [41]³, a Multi Layer Perceptron neural network is used for reconstructing the $D_s \rightarrow \phi\pi$ and $D_s \rightarrow K^{(*)0}K$ decay chains.

5.5.1 Selection of $D_s \rightarrow K_S^0 K$ decays

Selection of K_S^0 candidates

The *BABAR* Tracking Efficiency Task Force provides several proposals for the K_S^0 reconstruction, for which efficiency corrections are available. The corrections account for differences in the K_S^0 reconstruction efficiencies between data and MC. These proposals, also containing an approach for the determination of the correction are called K_S^0 recipes [41]. Since the proposals contain several setups of geometric cuts, but do not favor one specific setup, the significance and the signal-to-background ratio of all possible combinations have been studied. The cuts described below represent the optimal choice in terms of significance. First, a requirement on the mass of the $\pi^\pm\pi^\mp$ is applied, representing a 3σ region around the mean of the mass peak, resulting in the definition of the K_S^0 signal region of $0.4897 \text{ GeV}/c^2 < m(\pi^\pm\pi^\mp) < 0.5057 \text{ GeV}/c^2$ (see figure 5.5 a)). In addition, the K_S^0 vertex probability is cut at $P(\chi^2)(K_S^0) > 0.001$ (figure 5.5 b)). The vertex probability is evaluated as result of a kinematic fitting algorithm, which uses the information of the fitted tracks. In the vertexing algorithm, the sum of squares of the distance of closest approach of the tracks to a point is minimized. Because of the curvature of the tracks (and the quality of the determination of the track parameters through hits in the detector), this problem becomes non-linear. A local solution is obtained by linearization. This procedure is iterated from the local solution until convergence is achieved [10]. The χ^2 of the final fit result is then translated into a probability, called *Vertex Probability* $P(\chi^2)$ by including the information about the number of degrees of freedom. For good vertex fits, the distribution of the vertex probability is flat, whereas peaks at very small probabilities represent misfits.

A further reduction of the combinatorial background is achieved by requiring the 3-dimensional flight length of the K_S^0 to be $l > 1$ mm and the cosine of the angle between the K_S^0 momentum and the line connecting the K_S^0 decay vertex and the primary vertex of the event [41] to be larger than 0 (see figures 5.5 c) and 5.5 d)).

³For each of the recipes, a dedicated correction, as well as a systematic uncertainty for the reconstruction efficiency is available. Changes in the reconstruction make the efficiency correction more difficult and lead to larger systematic uncertainties.

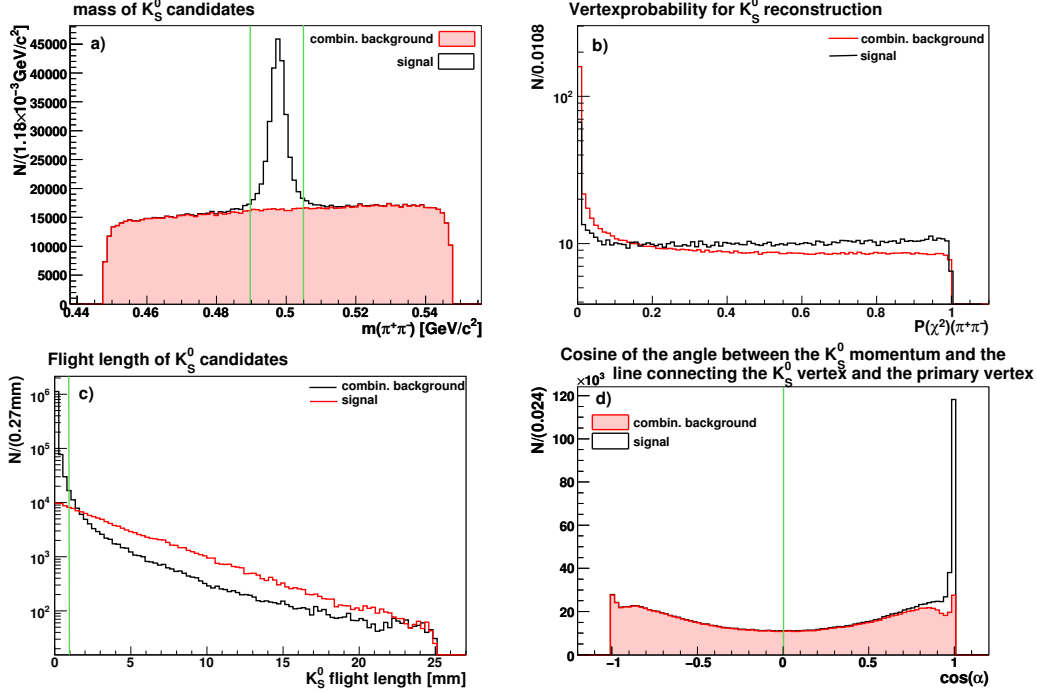


Figure 5.5: Cut variables for the $D_s \rightarrow K_S^0 K$ reconstruction chain: a) K_S^0 mass; b) K_S^0 vertex probability; c) K_S^0 flight length; d) cosine of the angle between the K_S^0 momentum and the line connecting the K_S^0 decay vertex and the events primary vertex. The green lines indicate the cuts applied.

Selection of $D_s \rightarrow K_S^0 K$ decays

After these cuts, the K_S^0 sample contains only very few combinatorial background events. Thus, the main source of combinatorial background of this reconstruction channel arises from the combination with a kaon to form the D_s . To suppress this kind of background, the vertex of the reconstructed D_s candidate is required to have a probability of $P(\chi^2)(D_s) > 0.001$. The distribution of the vertex probability of the D_s is shown in figure 5.6.

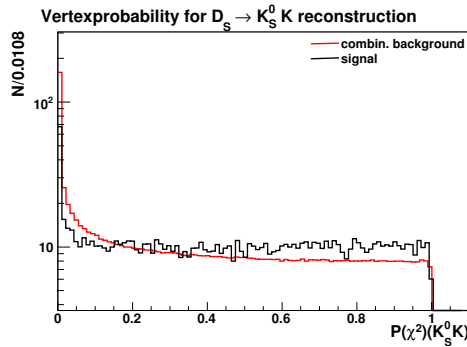


Figure 5.6: Vertex probability of the D_s candidate.

A further improvement of the purity can be achieved by checking whether the selected track can also originate from another composite particle. This can be done by combining the track with one or more other tracks and looking at the mass of the built composite candidate. In general, such *Veto*s can be constructed for all existing resonances, but it is only meaningful for narrow states, such as ϕ 's or D^0 mesons. The latter are the dominant misreconstruction sources in this analysis, as can be seen in figure 5.7.

A D^0 veto is constructed by combining the kaon, used for the D_s reconstruction, with one

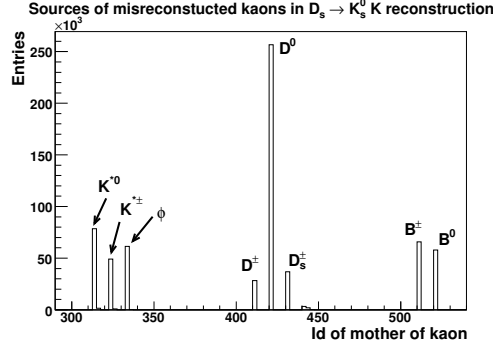


Figure 5.7: Source of misreconstructed kaons in the $D_s \rightarrow K_S^0 K$ reconstruction.

or more pions. As a preselection, D^0 veto-candidates are only considered if their mass lies within a mass window of $50 \text{ MeV}/c^2$ around the nominal D^0 mass [8]. If there are more than one D^0 veto-candidates, the one with the lowest difference to the D^0 mass is stored. The D^0 candidates are reconstructed using the following final states $K^\pm \pi^\mp$, $K^\pm \pi^\mp \pi^0$, $K^\pm \pi^\mp \pi^\pm \pi^\mp$, $K^\pm \pi^\mp \pi^\pm \pi^\mp \pi^0$, $K^\pm K^\mp$, $K^\pm K^\mp \pi^0$, $K^\pm K^\mp \pi^\pm \pi^\mp$ and $K^\pm K^\mp \pi^\pm \pi^\mp \pi^0$. The total inclusive branching ratio of all decays used for reconstructing the veto is $\sim 30\%$. Nevertheless, no improvement in terms of increasing significance of the D_s reconstruction has been found when applying this veto by rejecting events in a 3σ region around the D^0 mass peak. Thus, the veto has not been applied in this D_s reconstruction chain.

Candidates for the ϕ veto are reconstructed using the dominant decay, $\phi \rightarrow K^+ K^-$. It turns out that vetoing against kaons from ϕ decays helps to reduce the combinatorial background and increases the significance. A clear peak at the nominal ϕ mass is visible when looking into the mass distribution of ϕ veto-candidates of combinatorial background, as can be seen in figure 5.8. For signal decays, there are only very few events for which a veto candidate can be constructed. Table 5.4 at the end of this section gives an overview of all cuts used and their efficiencies on different Monte Carlo samples.

5.5.2 Selection of $D_s \rightarrow \phi \pi$ and $D_s \rightarrow K^{*0} K$ decays

To perform an efficient suppression of combinatorial background, the reconstruction of the decay chains $D_s \rightarrow \phi \pi$ and $D_s \rightarrow K^{*0} K$ is done by a Multi Layer Perceptron, created by the TMVA software package. Since both the ϕ and the K^{*0} are vector mesons and carry an angular momentum of $J = 1$, the D_s decays producing them are very similar. Thus, the MLP input variables are the same and can be described together.

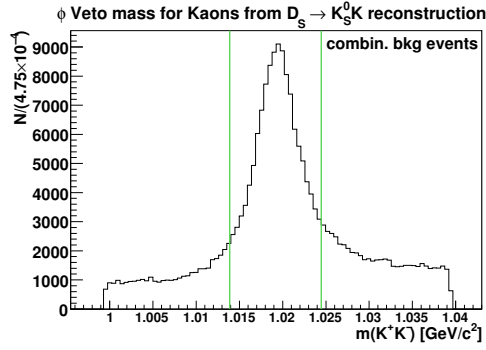


Figure 5.8: Mass distributions of the ϕ veto for the $D_s \rightarrow K_S^0 K$ reconstruction. The green lines indicate the cut region.

MLP input variables

The Truth-Matching (see section 3.4) is used to prepare the input samples for the neural network training. Events of the “signal” sample are required to contain a truth-matched D_s , while events of the background samples require the D_s and the D_s daughter to have none. Three input variables have been chosen as suitable for the neural network:

- the absolute value of the mass difference between the mass of the reconstructed D_s daughter and its PDG value: $|\Delta m| = |m(\phi/K^{*0}) - m_{PDG}(\phi/K^{*0})|$,
- the absolute value of cosine of the helicity angle (see figure 5.9 for definition) of the ϕ/K^{*0} candidates $|\cos \Theta_{\text{Hel}, \phi/K^{*0}}|$ and
- the vertex probability of the D_s candidate $P(\chi^2)(D_s)$.

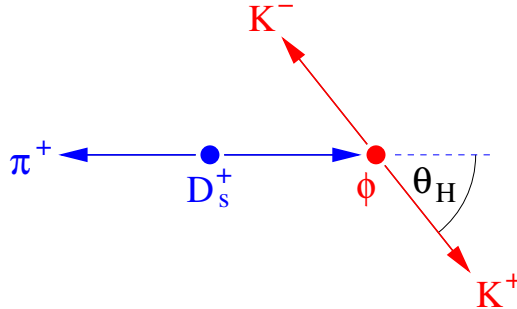


Figure 5.9: Definition of the helicity angle as the angle between the flight direction of the D_s and a ϕ (or K^{*0}) daughter in the ϕ (or K^{*0}) rest-frame.

The decay of the D_s into a vector meson (ϕ/K^{*0}) with $J = 1$ (see 3rd column of table 5.1) and a scalar meson with $J = 0$ results in a polarization of its daughters. The expected distribution for the helicity angle can be derived using spherical harmonics, leading to $\frac{dN}{d \cos \Theta_{\text{Hel}}} \sim \cos^2 \Theta_{\text{Hel}}$ for an angular momentum of $L = 1$ between the D_s daughters.

Further studies, concerning the setup and performance of the neural network have been conducted in order to find the optimal setup:

- The discrimination power of an alternative definition of the background sample, requiring the D_s daughter being successfully truth-matched but the D_s not, has been checked. It gave no improvement, since the weight of the mass of the D_s daughter and the helicity angle are much bigger than the weight of the D_s vertex probability.
- The insertion of the ϕ/K^{*0} vertex probability into the set of input variables has been studied. Since no improvement in the discrimination power of the MLP output was visible, it was decided to apply a *BABAR* standard cut on that variable ($P(\chi^2)(\phi/K^{*0}) > 0.001$) and use only three variables in the neural net to simplify the setup.
- Dropping the D_s vertex probability from the input variable dataset lead to a significantly worse discrimination between signal and background.
- Replacing Δm with the $m(\phi/K^{*0})$ results in a worse performance due to a more difficult shape of the mass input variable.
- Shifting $m_{\text{PDG}}(\phi/K^{*0})$ results in a negligible change of the MLP neural network performance and the efficiency of the cut on the output of the neural network.
- A preselection study of other TMVA methods came to the conclusion that the MLP, together with a couple of other algorithms, gives the best performance in terms of background rejection vs. signal efficiency. It could be shown, that using the Multi Layer Perceptron leads to major improvements in terms of the significance compared to a cut-based selection.

The input variables distributions for the $D_s \rightarrow \phi\pi$ reconstruction chain are shown in figure 5.10.

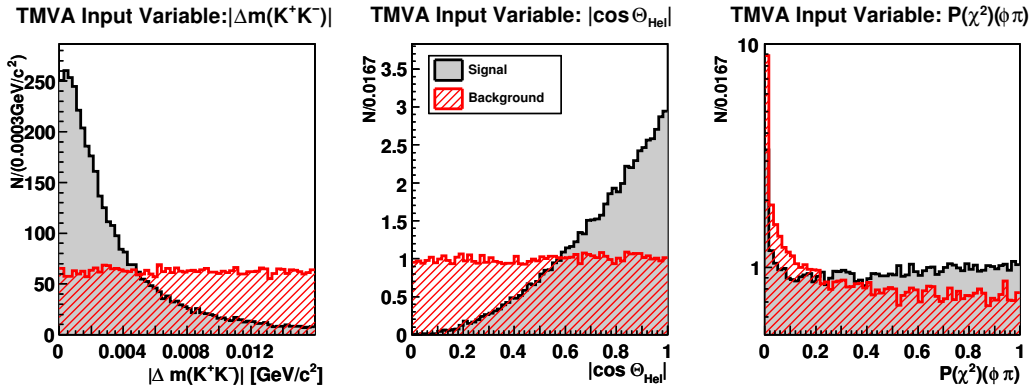


Figure 5.10: Scaled distributions of the MLP input variables Δm , $|\cos \Theta_{\text{Hel}}|$ and $P(\chi^2)(D_s)$ for the $\phi\pi$ decay chain: red combinatorial background, black: truth-matched D_s events.

As can be seen, all distributions show the expected behavior and especially the variables $|\Delta m|$ and $|\cos \Theta_{\text{Hel}}|$ already show a good separation between signal and background.

The helicity angle distribution of the $D_s \rightarrow K^{*0}K$ decay chain needs some special treatment. As can be seen in figure 5.11 a), a clear peak at $\cos \Theta_{\text{Hel}} = -1$ shows up in the combinatorial background as well. Thus, a training of the neural network with this sample would be less

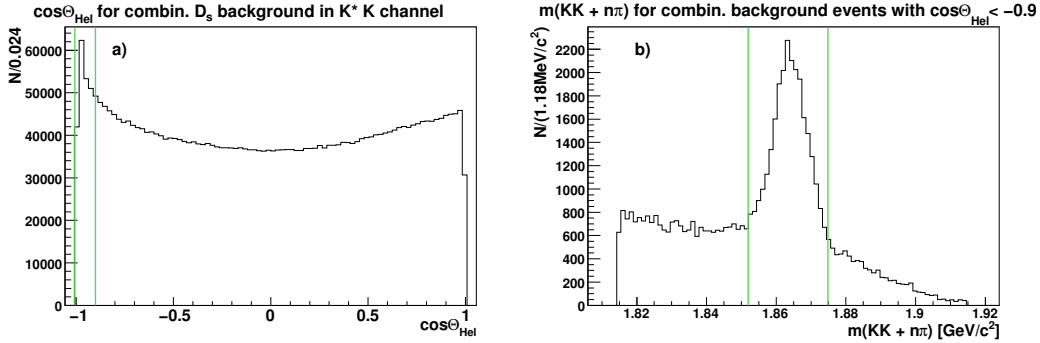


Figure 5.11: a): Helicity angle distribution of combinatorial background events of $K^{*0}K$ channel. The green lines indicate the region, for which the $m(KK n \cdot \pi)$ distribution has been examined. b): Scaled $KK(n \cdot \pi)$ mass distribution for the green region of figure a). The green lines indicate cut region of the applied D^0 veto.

effective than in the $\phi\pi$ reconstruction, where the background shows a flat distribution. It turns out that this peak appears due to the large width of the K^{*0} , which benefits wrongly reconstructed K^{*0} candidates. The two kaons, used for in the reconstruction of the $D_s \rightarrow K^{*0}K$ decay chain originate from other decays, producing two kaons with an angular correlation. The most prominent example for this are the decays $D \rightarrow KK(n \cdot \pi)$. The angular correlation between the two (almost) back-to-back emitted kaons propagates into the reconstruction of the D_s and leads to a peaking background component in the helicity angle distribution. Thus, a rejection of this kind of background is possible by introducing a veto on such decays. Combining the two kaons (and a couple of pions in case the resulting mass is much below the D mass), the invariant mass of these $KK(n \cdot \pi)$ candidates peaks at the D^0 mass, as shown in figure 5.11 b). Vetoing against these candidates by introducing a cut on the D^0 mass and tightening the preselection restrictions on the K^{*0} mass, the peak in the $\cos \Theta_H$ distribution vanishes. The input data sample for the training of the MLP neural net of the $D_s \rightarrow K^{*0}K$ channel is created using the cut, described above. The input variables distributions are shown in figure 5.12. As can be seen in the middle plot, there remains only a tiny enhancement at $|\cos \Theta_H| \sim 1$ in the distribution of combinatorial background candidates.

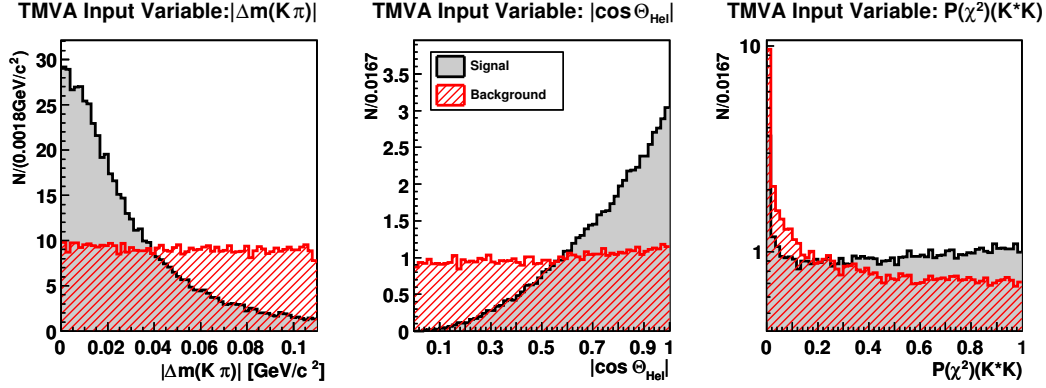


Figure 5.12: Scaled distributions of the MLP input variables Δm , $|\cos \Theta_H|$ and $P(\chi^2)(D_s)$ for the $K^{*0}K$ decay chain: red combinatorial background, black: truth-matched D_s events

MLP output

As previously mentioned, the neural network training is performed for different algorithms. Together with the (decorrelated) Likelihood and Fisher discriminant, the Multi Layer Perceptron provides the best signal efficiencies at a given background rejection level. In addition, the performance of an alternative realization of a Multi Layer Perceptron, the Clermont-Ferrand neural network has been checked. This network is an adaption of an older MLP realization, developed from at the Université Blaise Pascal in Clermont-Ferrand. In principle, the regular MLP method is much faster and more flexible than the Clermont-Ferrand MLP (CFMlpANN) [40]. Because of consistency reasons, the regular MLP has been chosen for further usage. The comparison plots for the different methods for both decay chains are shown in figure 5.13.

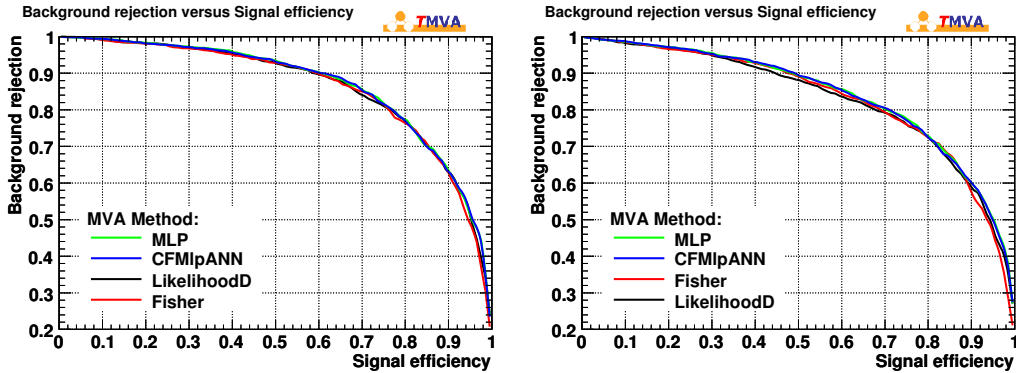


Figure 5.13: Comparison of the background rejection vs. signal efficiency between different TMVA algorithms for the $\phi\pi$ (left) and $K^{*0}K$ (right) reconstruction chain.

The MLP response and cut efficiencies for signal and background of both decay chains are shown in figure 5.14. As already mentioned in section 5.4, the cut values for the MLP of the D_s reconstruction and for the lepton momentum have been derived by a 2-dimensional cut optimization. The significance of a cut on the MLP output has been plotted against the significance of the lepton momentum cut.

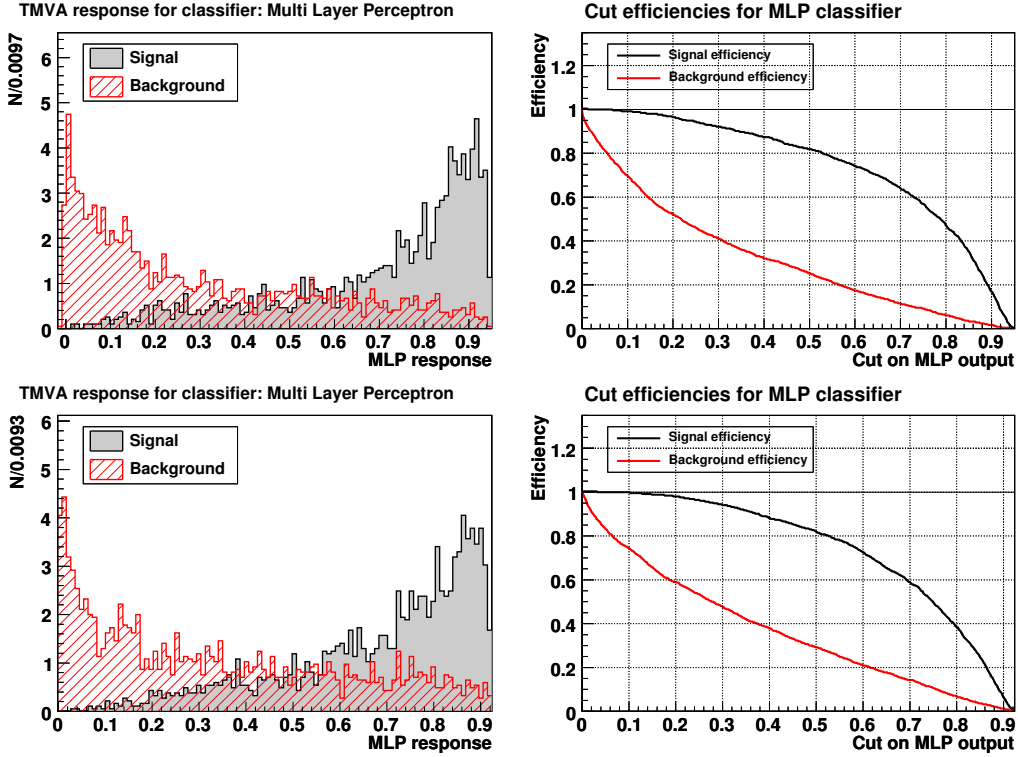


Figure 5.14: MLP response (left) and cut efficiencies (right) for signal (black) and background (red) for the MLP of $D_s \rightarrow \phi\pi$ (top) and $D_s \rightarrow K^{*0}K$ reconstruction chains (bottom).

As can be seen in figure 5.15, the significance increases with lower lepton momenta cuts and forms a plateau with the MLP output cut where it remains more or less constant. Thus, the lepton momentum cut should be as low as possible. A lower limit on this cut is given by the performance of the selectors used. They show a worse data-MC agreement of selection efficiency and a higher mis-identification rate at lower momenta, as mentioned in section 3.5. Thus, cutting too low at this quantity would introduce a high systematic uncertainty. In addition, the selection efficiency of the muon selector is very poor at momenta lower than $1.0 \text{ GeV}/c$.⁴ Taking all of this into account, the lepton momentum is required to be larger than $p_\ell > 1.0 \text{ GeV}/c$ for both lepton channels.

To determine optimal cut value for the MLP output, profiles of the 2D-histogram at the chosen lepton momentum cut are used. The profile plots for both D_s decay chains are shown in figure 5.16. As can be seen, the shape of the distribution is the same for the two lepton flavors. Therefore the same cuts are used for both final states. The final cut values are chosen to have the maximum significance. The MLP output for the $\phi\pi$ channel is cut at $\text{MLP}(\phi\pi) > 0.55$, whereas the cut for the $K^{*0}K$ channel is $\text{MLP}(K^{*0}K) > 0.65$.

An overview of all cuts applied together with their efficiencies is given in the tables 5.2 and 5.3 at the end of this chapter.

⁴More information on the efficiency of the selectors is given in section 8.2.3.

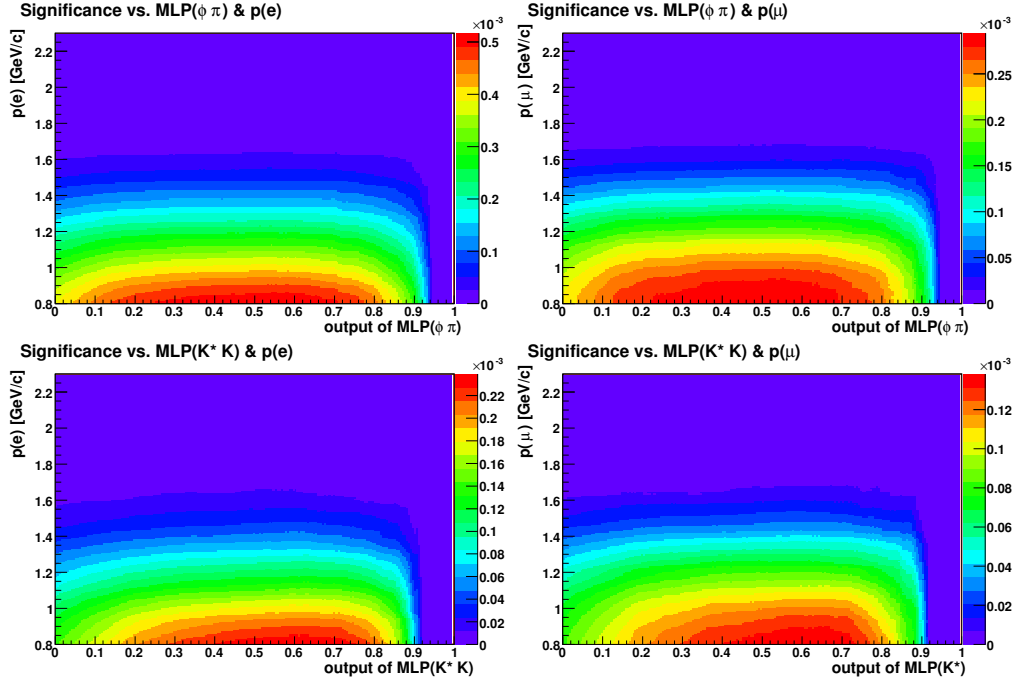


Figure 5.15: Significance of cut on the MLP output has been plotted against the significance of the lepton momentum cut for electrons (left) and muons (right) of the $D_s \rightarrow \phi\pi$ (top) and $D_s \rightarrow K^{*0}K$ reconstruction chains.

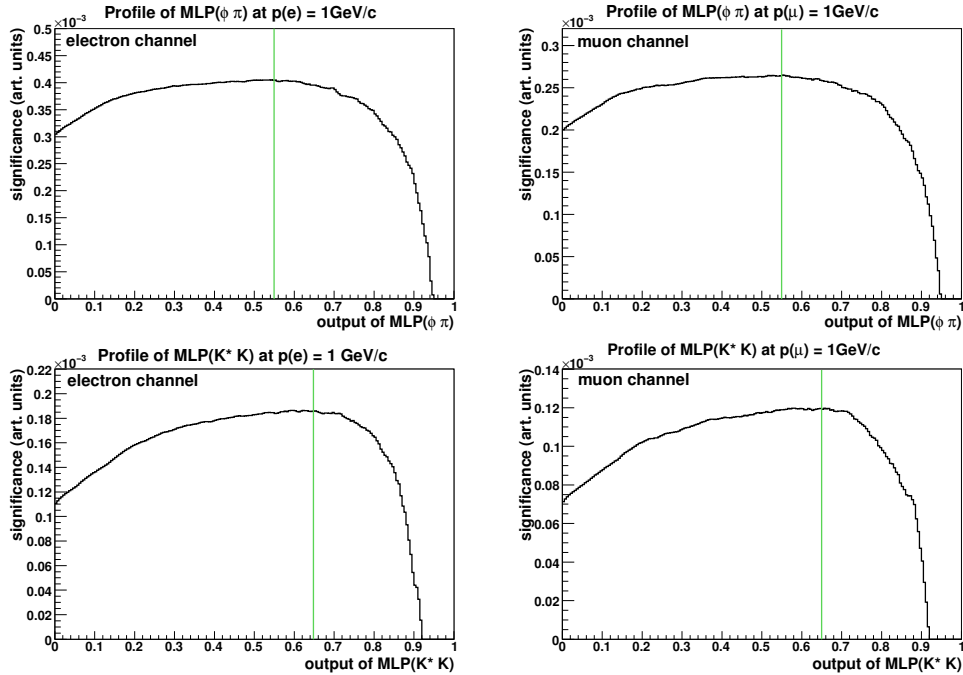


Figure 5.16: Profile plots of figure 5.15 at $p_\ell = 1 \text{ GeV/c}$ for electrons (left) and muons (right) of the $D_s \rightarrow \phi\pi$ (top) and $D_s \rightarrow K^{*0}K$ reconstruction chains.

Further cuts on the $D_s \rightarrow \phi\pi$ and $D_s \rightarrow K^{*0}K$ reconstruction

As already mentioned, the vertex probability of the ϕ/K^{*0} is not included in the set of MLP input variables. Nevertheless, the requirement of $P(\chi^2)(\phi/K^{*0}) > 0.001$ has been used to achieve a further background rejection. The possibility of increasing the purity of the sample by constructing vetoes against various particles (e.g. $D^{(*)0\pm}$ mesons) with the kaons used have been studied. It turned out that not applying any further cuts for the $D_s \rightarrow \phi\pi$ reconstruction is the optimum choice.

For the $D_s \rightarrow K^{*0}K$ reconstruction chain, it turned out that vetoing against $D^{(*)0\pm}$ is not useful, but constructing a ϕ candidate from the kaons and implementing a veto against it increases the significance of the D_s reconstruction. Figure 5.17 shows the invariant mass distributions of the veto ϕ candidates. A clear ϕ signal is visible for the combinatorial background sample whereas there's no peak when the D_s is reconstructed correctly.

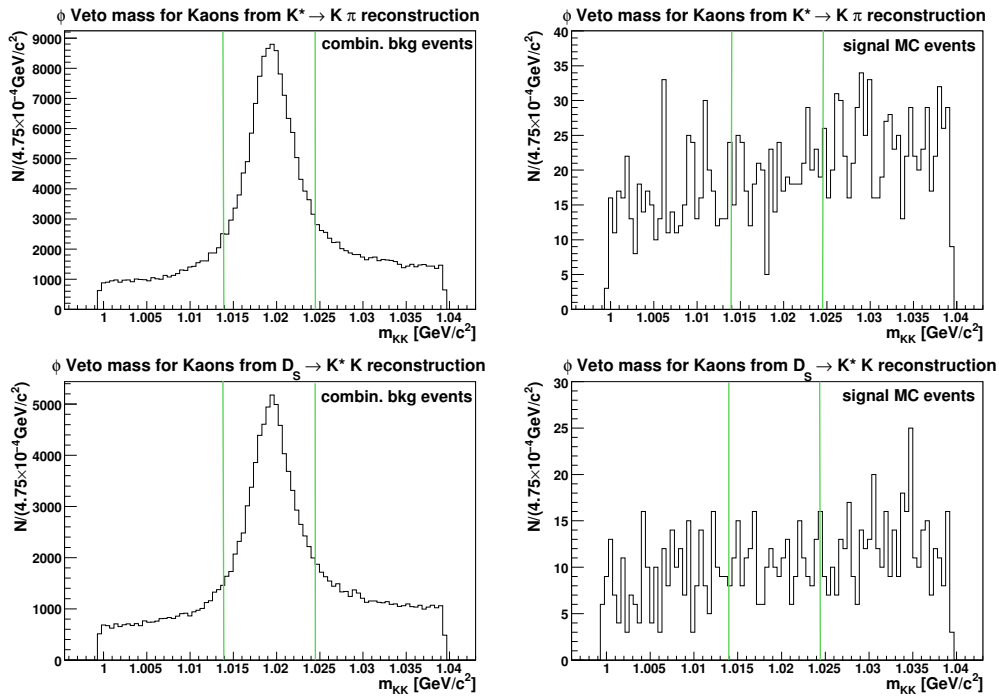


Figure 5.17: Mass distributions of the ϕ vetoes for the K^{*0} reconstruction (top) and for the $D_s \rightarrow K^{*0}K$ reconstruction (bottom) for combinatorial background events (left) and signal events (right).

5.5.3 Reconstruction of the B meson

To reduce the combinatorial background resulting from the reconstruction of the B meson, another Multi Layer Perceptron is used. The input variables are either sensitive to kinematic variables or to the shape of the B candidate or of the remaining tracks in the event. In addition, the vertex probability of the B candidates is studied. Finally, some physical cuts, like requiring four momentum conservation and a cut on the maximum B mass are applied.

The MLP for B reconstruction

The three input variables of the MLP of the B -shape are:

- The mass of the B candidate.
- The *Sphericity* of the B candidate, defined as the sum of the squares of transverse momenta for each track with respect to the event axis.
- The *Thrust* of the rest of the event, being related to the Thrust-axis \hat{T} (as defined in section 5.3) by $T = \frac{\sum_i |\hat{T} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$ with \vec{p}_i being the momenta of the tracks.

The input variables of the neural network are shown in figure 5.18.

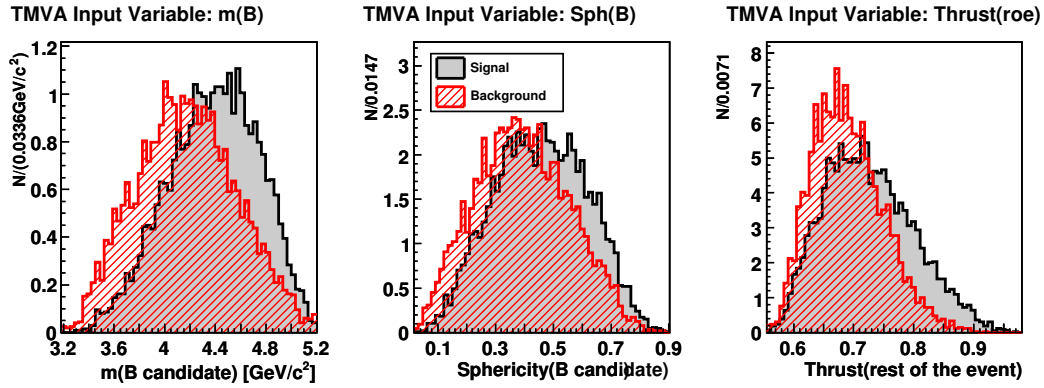


Figure 5.18: Scaled distributions of the MLP input variables $m(B)$, $Sphericity(B)$ and $Thrust(\text{Rest of the event})$ for the B reconstruction: red: combinatorial background, black: truth-matched signal B events

As can be seen, all variables show only small shape differences between signal and background. Nevertheless, a cut on the combination of them, constructed by the MLP algorithm results in an increase of the significance. It has to be noted, that the signal sample, used for the training of the neural network consists of reconstructed B candidates with a MC Truth-Match. Thus, the data sample available for the training is much smaller than for the construction of the other neural networks, leading to the MLP responses of signal and background showing large statistical fluctuations. As such, the reliability of a requirement arising from this MLP has to be checked by comparing the response distribution for the data with the one corresponding to Monte Carlo, scaled to data luminosity. The comparison is shown in the next chapter.

The use of additional variables has been tested but resulted in no improvement. The MLP response, as well as the background rejection vs. signal efficiency plots are shown in figure 5.19. The cut value of the MLP response has been determined by maximizing the significance. This procedure led to the requirement $MLP_{B\text{shape}} > 0.45$. The significance plots of the MLP cut for the electron and muon channel are shown in figure 5.20.

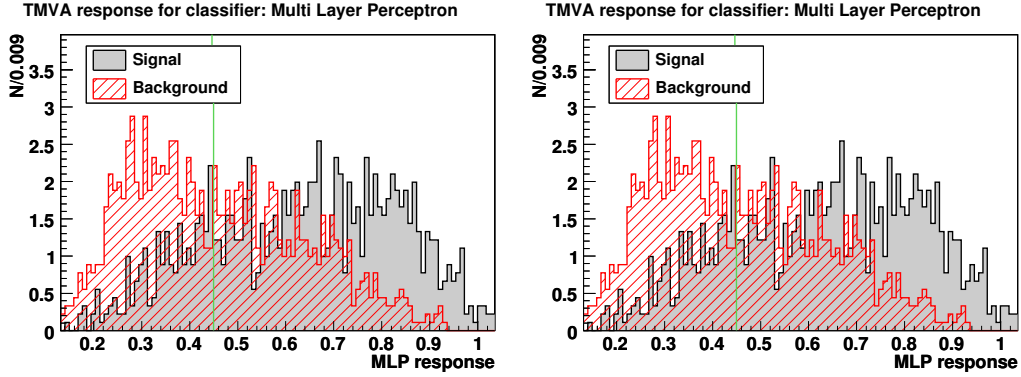


Figure 5.19: MLP response (left) and cut efficiencies (right) for signal (black) and background (red) for the MLP of B reconstruction.

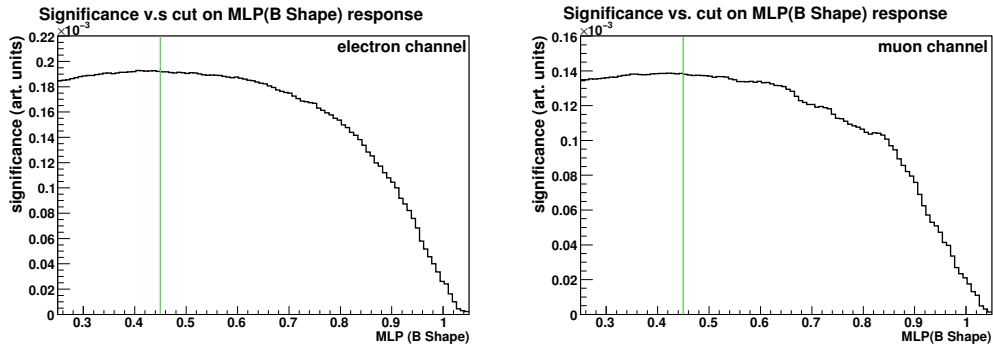


Figure 5.20: Significance vs. Cut on MLP(B shape) response for the electron (left) and muon channel (right). The red lines indicate the chosen cut value.

Further cuts on the B candidate

The vertex probability of the $D_s K \ell$ combinations is used to reject combinatorial background events, for which it is hard to fit the three tracks into one vertex point. A significance study showed, that the optimum choice of this cut is $P(\chi^2)(D_s K \ell) > 0.01$, which is the $BABAR$ standard cut for composite particles reconstructed from an intermediate composite particle (the D_s). The dependence of the significance from the cut on the vertex probability is shown in figure 5.21.

The CMS-momenta of the B daughter particles are required to be smaller than the maximum values, postulated by four momentum conservation. Thus, the momenta are cut at $p_e < 2.1 \text{ GeV}/c$, $p_K < 2.1 \text{ GeV}/c$ and $p_{D_s} < 2.25 \text{ GeV}/c$. Furthermore the mass of the $D_s K \ell$ combination is cut at $m(B) < 5.3 \text{ GeV}/c^2$. These cuts do not suppress much combinatorial background (almost only continuum events are rejected), but since they represent the physical boundaries on the signal case, no signal events are rejected.

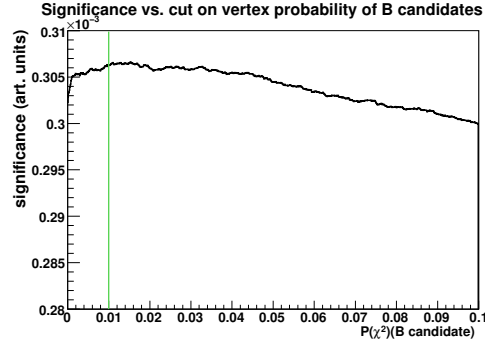


Figure 5.21: Significance vs. cut on the B candidates vertex probability, the red line indicates the chosen cut value.

5.5.4 True- D_s background rejection

As mentioned in the beginning of this chapter, True- D_s background consists of a correctly reconstructed D_s , combined with a kaon or lepton, coming from the second B . This type of background cannot be suppressed by the cuts, described above, due to its similarity to signal events. Thus, it represents a large fraction of the remaining candidates. Looking at the decays underlying these background events, it turns out that they are dominated by decays such as $B \rightarrow D_s^{(*)} D^{(**)} (+n\pi)$, where the D_s is produced by the upper vertex, in contradiction to the signal channel. Thus, this background is referred to as *Upper Vertex background* in the following. Figure 5.22 shows the spectator diagram of this kind of upper vertex decays.

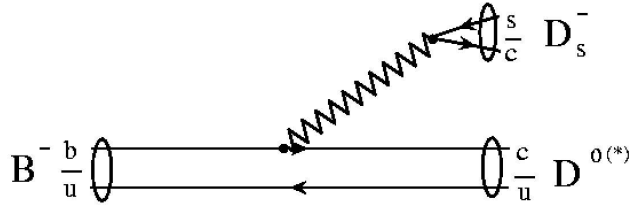


Figure 5.22: Spectator diagram of upper vertex decays into $D_s^{(*)} D^{(**)}$ systems.

The decays of the $D^{(**)}$ meson into charged or neutral kaons are used to suppress the amount of the upper vertex background, using the remaining angular correlation between the kaons and the reconstructed D_s . It has to be noted, that the rejection only uses tracks, which have not already been used to reconstruct the $B \rightarrow D_s K \ell \nu_\ell$ decay.

The inclusive branching fraction for $D^{(**)}$ decays into a final state with a charged kaon, where the charge is the same as the one of the D_s , is about 25% for a decay of a charged $D^{(**)}$ and about 53% for a decay of a neutral $D^{(**)}$. Neutral kaons are produced with a \mathcal{BR} of about 60% in the decay of a $D^{(*)\pm}$ and 42% in decays of $D^{(*)0}$ s. If the event contains a signal decay, the production rates for additional charged and neutral kaons are similar. Only the different production mechanisms can be used to introduce a selection criterion.

Since both the D_s and the additional kaon originate directly or via cascade decay from the same B , there remains an angular correlation between them. The B daughter particles $D^{(*)}$ and D_s are emitted more or less back-to-back. In the decay of the c quark of the $D^{(*)0}$, most of the momentum is transferred to the s quark, forming the additional kaon. This constrains its flight direction to be very similar to the one of the $D^{(*)}$. Thus, the cosine of the angle between the flight directions of the D_s and the additional kaon peaks at -1 . In the signal decay, additional kaons originate from the other B and therefore show no angular correlation with the D_s . Hence, the $\cos\Theta(D_s K)$ distribution is flat for such events. A comparison of the distributions of both cases is shown in figure 5.23.

Two oppositely charged tracks are used to form K^0 candidates, requiring the resulting composite to have a mass within $5 \text{ MeV}/c^2$ of the nominal K^0 mass [8]. Charged additional kaons are taken from the *KNNNotPion* list and are required to have the correct charge configuration $q(D_s) \cdot q(K_{\text{additional}}) > 0$.

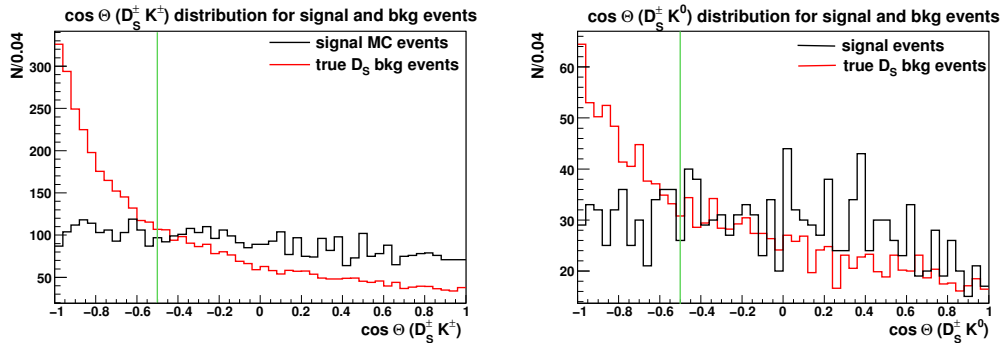


Figure 5.23: Angular distribution of $\cos\Theta(D_s^\pm K^\pm)$ (left) and $\cos\Theta(D_s^\pm K^0)$ (right) after all described cuts. Red curves show the distributions of true- D_s background while the black curves show the behavior for signal MC events. The green lines indicate the chosen cut value. Both distributions are scaled to equal areas.

The cut values of the $\cos\Theta(D_s K)$ selection criterion are chosen not to maximize the significance, but to reject as much D_s background as possible, since the size of this background fraction has a large impact on the systematic uncertainty. The more background can be rejected, the smaller is the final systematic (and also the statistical) error on the measurement.

5.5.5 Multiple candidates

After applying the selection criteria described, approximately 8% of the remaining events contain more than one successfully reconstructed B candidate. Since the expected branching fraction of the signal is at the order of 10^{-4} and the reconstruction efficiency is very small, the probability of both B mesons decaying into the signal channel is very small. Thus, the loss of signal events, by requiring only one B candidate per event is negligible. Most of the multiple candidates are double candidates, which can be seen in the figure 5.24a. A best candidate selection is done using the vertex probability of the B . The candidate with the highest vertex probability is chosen as best candidate. This method selects the correct candidate in about 60% of all cases. The quality of the best candidate selection in terms of right choice vs. wrong choice is shown in figure 5.24b.

Although the quality of this selection is not very good, it represents the only possibility for the rejection of multiple candidates. Since in most cases only one D_s candidate survives the selection and is combined with two or more kaons and/or leptons, only variables describing the combination of the three particles can be used. The vertex probability is the only reliable information reflecting the quality of the combination of the tracks.

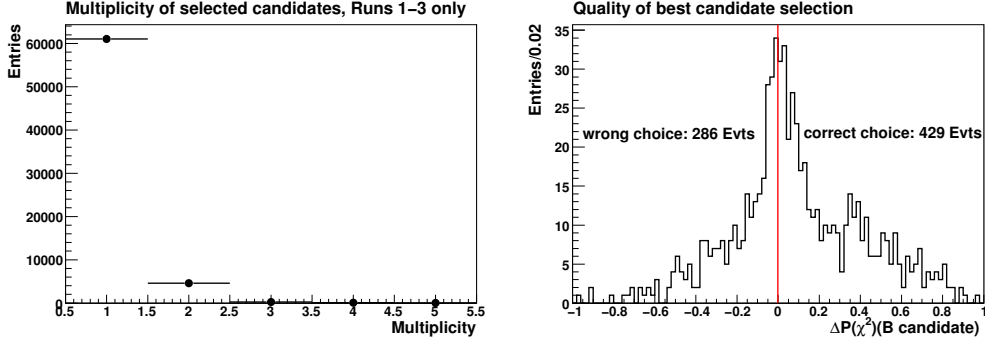


Figure 5.24: a) Multiplicity after all cuts applied b) $\Delta P(\chi^2)(D_s K \ell) = P(\chi^2)(D_s K \ell, \text{signal}) - P(\chi^2)(D_s K \ell, \text{combin. bkg})$. Candidates with $\Delta P(\chi^2)(D_s K \ell) > 0$ represent the correct choice.

Tables 5.2-5.4 show the efficiency of the event selection for different data samples.

It has to be noted, that no cut on the D_s mass distribution is made as part of the event selection, which would reduce the amount of combinatorial background. There are two particular reasons for that:

The first one is that the analysis of the remaining data sample needs both, signal region and sidebands of the D_s mass distribution for the extraction of the number of signal events, as will be described in chapter 7.

The second reason are differences in the position and width of the peak of the $m(D_s)$ distribution between data and the simulation, which is shown in the next chapter. A restriction to $m(D_s)$, developed using only simulated events would lead to additional systematic uncertainties due to the shift of the mean of the D_s mass distribution.

Table 5.2: Reconstruction cuts and their efficiencies on the $D_s \rightarrow \phi\pi$ decay chain, top: electron channel, bottom: muon channel.

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
					Events	ϵ	Events	ϵ
reconstructed	14682	100%	572374	100%	1.73×10^6	100%	213691	100%
MLP(Eventshape) < 0.5	12518	85.3%	125539	21.9%	1.42×10^6	82.1%	175153	82.0%
$p_{\text{electron}} > 1.0$ GeV/c	4882	33.3%	17410	3.0%	337468	19.5%	59275	27.7%
MLP(ϕ) > 0.55, $P(\chi^2)(\phi)$	3415	23.3%	5865	1.0%	84932	4.9%	42771	20.0%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	3350	22.8%	3638	0.6%	74583	4.3%	38045	17.8%
$P(\chi^2)(D_s K\ell) < 0.01$	2970	20.2%	2670	0.5%	56597	3.3%	28531	13.4%
MLP(Bshape) > 0.45	2212	15.1%	1471	0.3%	27899	1.6%	17833	8.3%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	1826	12.4%	1170	0.2%	23539	1.4%	12534	5.9%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	1724	11.7%	1148	0.2%	22809	1.3%	11982	5.6%

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
					Events	ϵ	Events	ϵ
reconstructed	4528	100%	132425	100%	437235	100%	67662	100%
MLP(Eventshape) < 0.5	3889	85.9%	27002	20.4%	353715	80.9%	54622	80.7%
$p_{\text{muon}} > 1.0$ GeV/c	2894	63.9%	14735	11.1%	235899	54.0%	40699	60.2%
MLP(ϕ) > 0.55, $P(\chi^2)(\phi)$	2002	44.2%	4939	3.7%	60098	13.7%	29455	43.5%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	1950	43.1%	3173	2.4%	52045	11.9%	25675	37.9%
$P(\chi^2)(D_s K\ell) < 0.01$	1697	37.5%	2554	1.9%	40101	9.2%	19346	28.6%
MLP(Bshape) > 0.45	1276	28.2%	1448	1.1%	20177	4.6%	12355	18.3%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	1073	23.7%	1157	0.9%	17190	3.9%	8681	12.8%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	1046	23.1%	1137	0.9%	16642	3.8%	8338	12.3%

Table 5.3: Reconstruction cuts and their efficiencies on the $D_s \rightarrow K^{*0}K$ decay chain, top: electron channel, bottom: muon channel.

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
					Events	ϵ	Events	ϵ
reconstructed	11752	100%	1.70×10^6	100%	8.81×10^6	100%	204694	100%
MLP(Eventshape) < 0.5	10230	87.0%	503475	29.6%	7.51×10^6	85.2%	170060	83.1%
$p_{\text{electron}} > 1.0$ GeV/c	3923	33.4%	71336	4.2%	1.72×10^6	19.5%	57387	28.0%
MLP(K^{*0}) > 0.65, $P(\chi^2)(K^{*0})$	2158	18.4%	8271	0.5%	159087	1.8%	33314	16.3%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	2115	18.0%	5933	0.3%	74583	0.8%	30025	14.7%
$P(\chi^2)(D_s K\ell) < 0.01$	1861	15.8%	4483	0.3%	56597	0.6%	22233	10.9%
MLP(Bshape) > 0.45	1399	11.9%	2575	0.2%	53758	0.6%	14193	6.9%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	1171	9.9%	2060	0.1%	44790	0.5%	9934	4.9%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	1108	9.4%	2003	0.1%	43374	0.5%	9568	4.7%

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
					Events	ϵ	Events	ϵ
reconstructed	3556	100%	409851	100%	2.26×10^6	100%	65013	100%
MLP(Eventshape) < 0.5	3085	86.8%	109327	26.7%	1.81×10^6	80.1%	53469	82.2%
$p_{\text{muon}} > 1.0$ GeV/c	2288	64.3%	60966	14.9%	1.21×10^6	53.5%	39794	61.2%
MLP(K^{*0}) > 0.65, $P(\chi^2)(K^{*0})$	1264	35.5%	7130	1.7%	111631	4.9%	23234	35.7%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	1237	34.8%	5044	1.2%	100239	4.4%	20480	31.5%
$P(\chi^2)(D_s K\ell) < 0.01$	1090	30.7%	4116	1.0%	77881	3.4%	16254	25.0%
MLP(Bshape) > 0.45	839	23.6%	2356	0.6%	38137	1.7%	10027	15.4%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	711	20.0%	1892	0.5%	31920	1.4%	7027	10.8%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	693	19.5%	1846	0.5%	30820	1.4%	6764	10.4%

Table 5.4: Reconstruction cuts and their efficiencies on the $D_s \rightarrow K_S^0 K$ decay chain, top: electron channel, bottom: muon channel.

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
	Events	ϵ	Events	ϵ	Events	ϵ	Events	ϵ
reconstructed	12888	100%	2.60×10^6	100%	6.58×10^6	100%	120237	100%
MLP(Eventshape) < 0.5	11912	85.3%	612755	21.9%	5.34×10^6	82.1%	100756	82.0%
$p_{\text{electron}} > 1.0$ GeV/c	4329	33.3%	78878	3.0%	1.36×10^6	19.5%	34348	27.7%
K_S^0 reconstr., $P(\chi^2)(D_s) > 0.001$	3329	23.3%	8652	1.0%	88566	4.9%	26531	20.0%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	3286	22.8%	6705	0.6%	81204	4.3%	23877	17.8%
$P(\chi^2)(D_s K\ell) < 0.01$	2907	20.2%	4884	0.5%	60909	3.3%	18065	13.4%
MLP(Bshape) > 0.45	1878	15.1%	2459	0.3%	31300	1.6%	11930	8.3%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	1878	12.4%	2051	0.2%	27185	1.4%	8430	5.9%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	1776	11.7%	2000	0.2%	26327	1.3%	8063	5.6%

Cut	Signal MC		continuum MC		$B\bar{B}$ MC			
	Events	ϵ	Events	ϵ	combin. D_s Bkg		D_s Bkg	
	Events	ϵ	Events	ϵ	Events	ϵ	Events	ϵ
reconstructed	4045	100%	583791	100%	1.69×10^6	100%	37135	100%
MLP(Eventshape) < 0.5	3486	86.2%	125204	21.4%	1.36×10^6	80.5%	31707	85.4%
$p_{\text{muon}} > 1.0$ GeV/c	2561	63.3%	68192	11.7%	943415	55.8%	23712	63.9%
K_S^0 reconstr., $P(\chi^2)(D_s) > 0.001$	1940	48.0%	7240	1.2%	61293	3.69%	18293	49.3%
$p_{\text{max},K}, p_{\text{max},D_s}, m_{D_s K\ell} < m_{B^\pm}$	1893	47.0%	5565	1.0%	55467	3.2%	16106	43.4%
$P(\chi^2)(D_s K\ell) < 0.01$	1674	41.4%	4347	0.7%	41857	2.5%	12245	33.0%
MLP(Bshape) > 0.45	1310	32.4%	2195	0.4%	22054	1.3%	8086	21.8%
$\cos\Theta_{D_s^\pm K^\pm} > -0.5$	1093	27.0%	1833	0.3%	19127	1.1%	5772	15.5%
$\cos\Theta_{D_s^\pm K^0} > -0.5$	1066	26.4%	1771	0.3%	18541	1.1%	5543	14.9%

Chapter 6

Data - Monte Carlo comparison of cut variables

Before the remaining candidates are fitted to extract the signal yield, it has to be proven, that the shapes of the cut variables are the same in data and the Monte Carlo simulation. If this is not the case, large systematic uncertainties can arise or selection criteria even can become unusable. Furthermore, it has to be checked if cuts on certain variables change the distributions of other variables. For the comparison, the MC samples have been rescaled to be equivalent to the data luminosity, used in this analysis. Differences between data and simulated events can appear due to insufficient simulation of interaction with detector material, varieties in the acceptances of simulated and real detector components or uncertainties in the branching ratios used for the simulation. In general, there are two kinds of differences, which need to be distinguished:

- Differences in the overall scaling of the distributions are uncritical for the analysis, since the total number of events in the sample is only relevant for the statistical error of the measurement, but doesn't affect the efficiency of the reconstruction of signal events.
- Differences in the shape of the distributions are critical. They alter the reconstruction efficiency for signal events in an unpredictable way and can introduce a bias into the measurements or even artificially create a signal by changing distributions, later used for the extraction of the signal yield.

The distributions of all variables used in the analysis have been compared. For each variable two comparisons have been made, one with only the lepton momentum cut applied and one after the whole event selection. The different contributions to the simulation are shown separately, using different colors. All comparison plots show the same color code, which is as following:

- red: combinatorial $B\bar{B}$ background from the D_s reconstruction
- yellow: True- D_s background
- blue: background from $c\bar{c}$ events
- green: background from $u\bar{u}/d\bar{d}/s\bar{s}$ events

Although the $c\bar{c}$ background contains also D_s , which might have been reconstructed correctly, these candidates are associated to the $c\bar{c}$ background contribution. Contributions of the signal decay are not contained in the Monte Carlo sample. But due to the smallness of the signal branching fraction they would not be visible anyway. Also, a possible signal contribution in the data cannot be seen at this stage of the analysis. The data is represented by black points.

Below the actual comparison plots, there are plots showing the relative difference between the entries of data and the simulation in every bin. The difference plots should show shape differences, which might not be visible in the comparison plots itself if the entries of data and Monte Carlo being very close to each other in several bins. Plots on the left hand side always show the variables with only the lepton momentum cut applied, while the plots on the right hand side show the distributions after all cuts have been applied.

6.1 Event shape Variables

The variables, used for the suppression of continuum background are $R2$, $\cos\Theta_{Thrust}$ and $L2$, as already introduced in chapter 5. Figure 6.1 shows the data versus Monte Carlo comparison for these three variables.

As can be seen, there's a scaling problem between data and MC. The simulation overestimates data by about $\sim 8\%$ before all cuts have been applied. This kind of scaling problem before applying the cuts is seen very often in *BABAR* analyses. It is not completely understood, but often related to the simulation of continuum events. Thus, the discrepancies should vanish with the reduction of continuum events by applying the cuts. Nevertheless, the plots on the right hand side of figure 6.1 show an even larger relative difference between data and Monte Carlo. It will be shown in one of the next figures, that the True- D_s background is the major source of this difference. Besides the scaling problem, there are no obvious differences in the shapes of the variables.

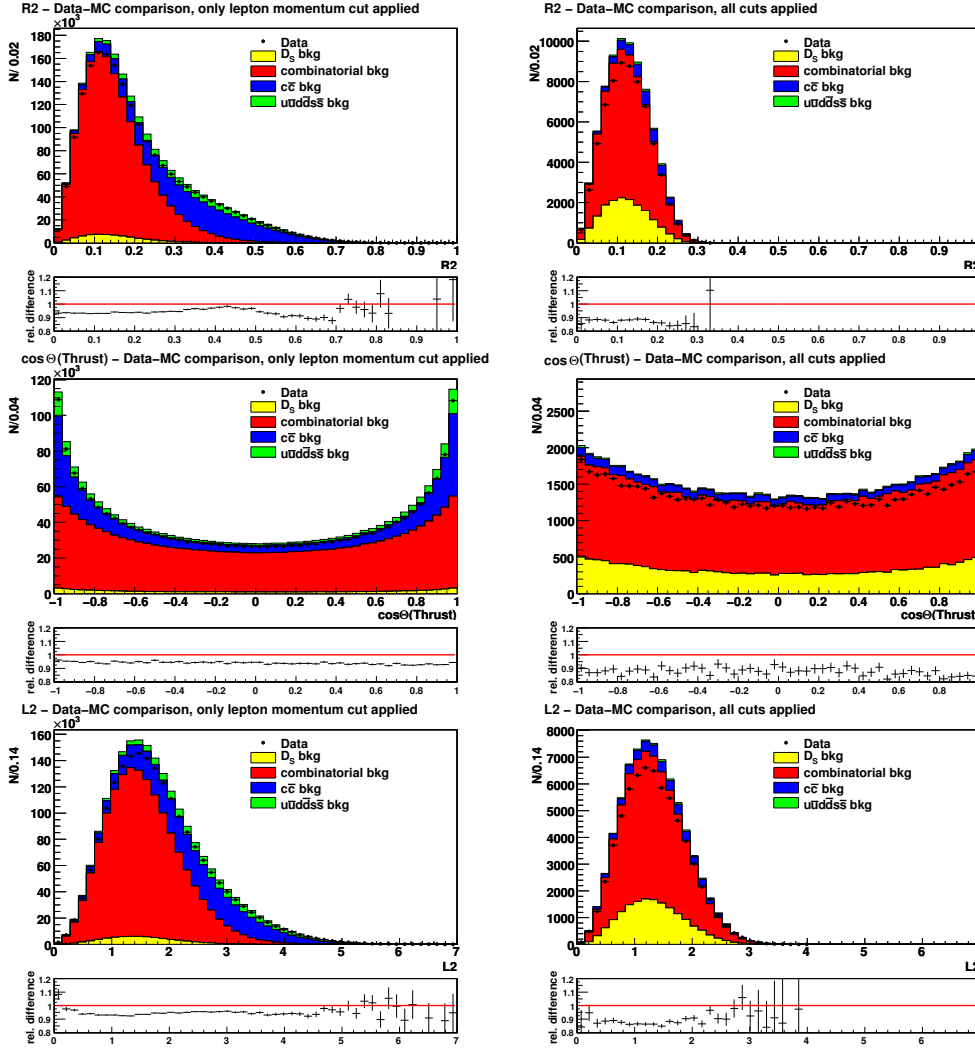


Figure 6.1: Data-MC comparison of the event shape variables $R2$ (top), $\cos \Theta_{\text{Thrust}}$ (middle) and $L2$ (bottom). Left: with lepton momentum cut only. Right: after the full selection.

The combination of the three variables delivers the MLP(Event shape) variable, for which the data-MC comparison is shown in figure 6.2. As can be seen, there are no visible differences in the shape of the distributions. The scaling problem is present at the same order of magnitude as before and shows the same behavior when comparing the relative differences before and after the event selection.

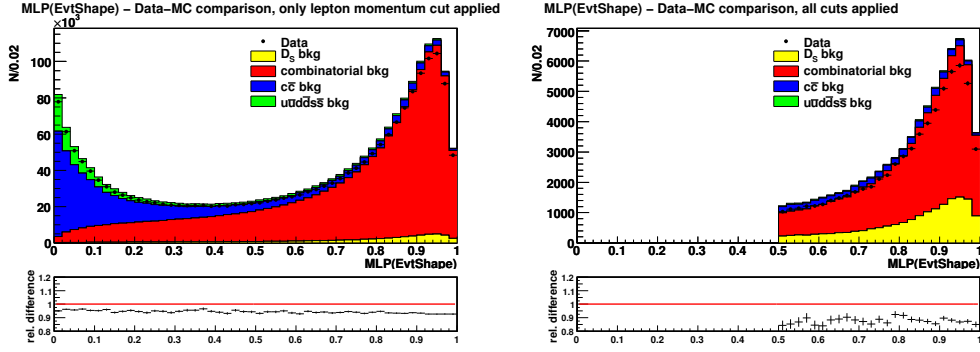


Figure 6.2: Data-MC comparison of the MLP output from the event shape variables. Left: with lepton momentum cut only. Right: after the full selection.

6.2 $D_s \rightarrow K_S^0 K$ reconstruction

The data-MC comparison for the variables used for the reconstruction of the $D_s \rightarrow K_S^0 K$ decay chain can be seen in figures 6.3 and 6.4.

The K_S^0 mass distribution shows a good agreement for the non-signal region, while there are more K_S^0 in Monte Carlo than in data. A comparison of the PDG08 [8] with the branching ratios used in the simulation shows, that the branching ratio of the decay $D_s \rightarrow K_S^0 K$ is about 35% too high in the simulation, while there's a good agreement between the values for the other two D_s decay chains, used in this analysis.¹ While the distributions of the angle between the K_S^0 momentum and the line connecting the K_S^0 production and decay vertex show a good agreement, the relative difference of the K_S^0 vertex probability shows a clear increment at lower values (which will be discussed later), as well as the flight length distribution. After applying all selection cuts, this increment vanishes and besides the usual scaling problem, both variables show a good agreement between data and MC.

The D_s vertex probability shows the same behavior like the K_S^0 vertex probability, while the veto mass shows a good agreement between data and MC.

¹During the time the analysis was performed, the Particle Data Group updated the branching fractions for the D_s decays several times. The MC simulated data was updated with a temporal delay due to computing capacities. Thus, the comparison can only be done for the MC production cycle used in the dedicated analysis.

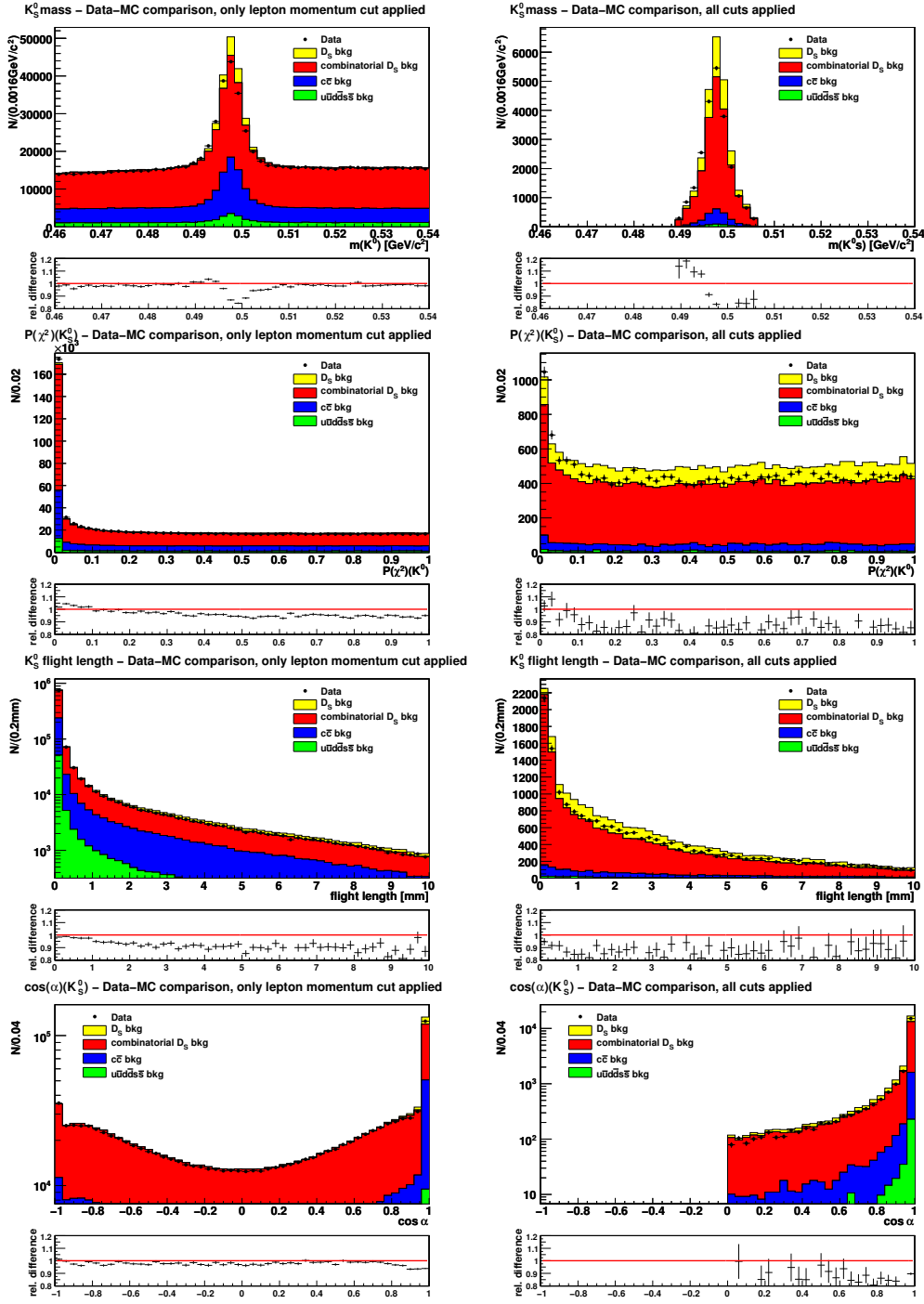


Figure 6.3: Data-MC comparison of the variables, used for the $D_s \rightarrow K_S^0 K$ reconstruction, part 1: K_S^0 mass (top), K_S^0 vertex probability (2nd from top), K_S^0 flight length (3rd from top) and angle between the K_S^0 momentum and the line connecting the K_S^0 production and decay vertex (bottom). Left: with lepton momentum cut only. Right: after the full selection.

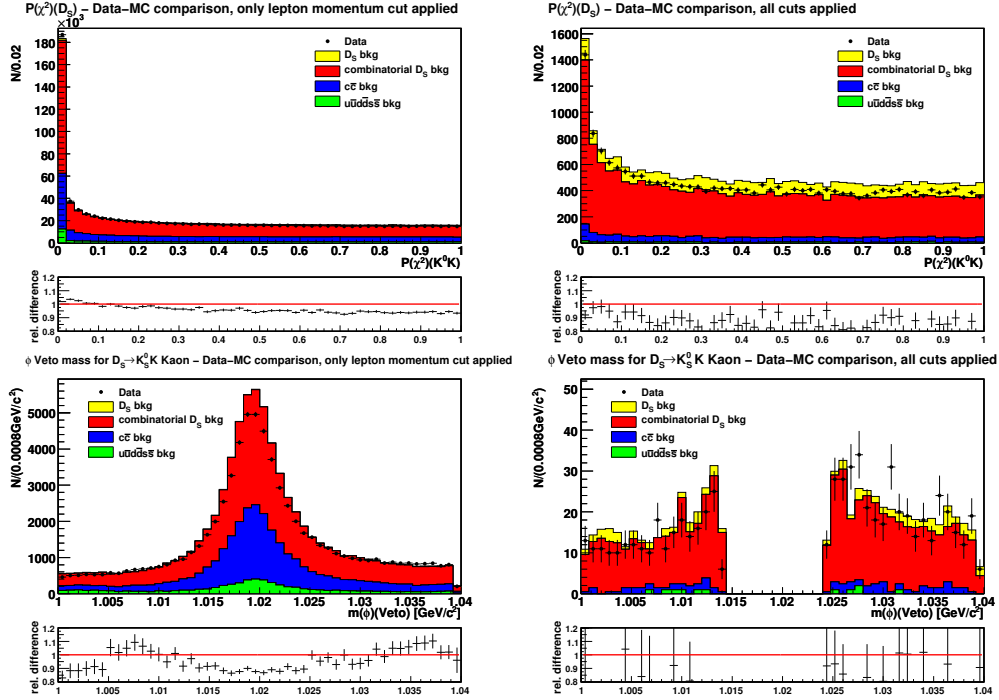


Figure 6.4: Data-MC comparison of the variables, used for the $D_s \rightarrow K_S^0 K$ reconstruction, part 2: vertex probability of reconstructed D_s from $D_s \rightarrow K_S^0 K$ decays (top) and veto mass for ϕ candidates, constructed from the charged kaon of the D_s decay (bottom). Left: with lepton momentum cut only. Right: after the full selection.

Figure 6.5 shows the data-MC comparison for the D_s mass distribution. Only events reconstructed in the $K_S^0 K$ decay channel are shown. Again, the combinatorial background

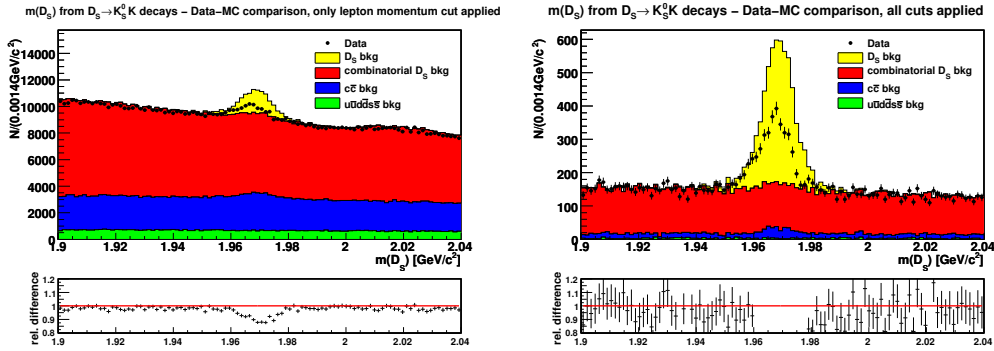


Figure 6.5: Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K_S^0 K$ reconstruction chain. Left: with lepton momentum cut only. Right: after the full selection.

events in data agree well with the scaled MC, while events with a correctly reconstructed D_s are overestimated. This is consistent with the observations in the distributions of the event shape variables concerning the change of the relative difference after applying the cuts and with comparison of the K_S^0 mass distribution.

Besides that, a slight shift in the D_s mass central value to lower values is visible. In reference [8], the reported branching ratio for this D_s decay is $\mathcal{BR}(D_s \rightarrow K_S^0 K)_{PDG} = (1.49 \pm 0.09)\%$, while the MC generator uses a value of $\mathcal{BR}(D_s \rightarrow K_S^0 K)_{MC} = 2.2\%$. A rescaling of the True- D_s background fraction has been done to achieve a better agreement between data and the Monte Carlo simulation by using the ratio of the two branching fractions. Figure 6.6 shows the comparison with the rescaled branching ratio resulting in a decreasing difference of the correct reconstructed D_s fraction. Nevertheless, there remains a clear difference after

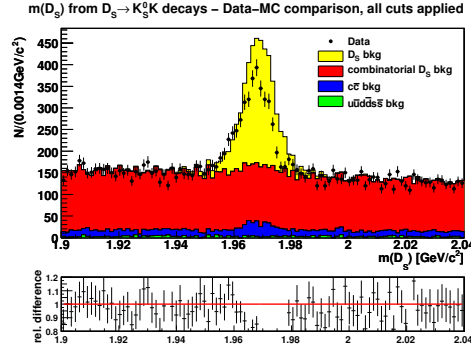


Figure 6.6: Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K_S^0 K$ reconstruction chain. The True- D_s background fraction has been corrected for D_s branching ratio used in the MC production after the full selection.

the correction has been applied. The source of this discrepancy probably lies in the simulated branching ratios for the upper vertex decays. The measured branching fractions for such decays contain large errors of about 25% [8]. These uncertainties can easily propagate into the simulation and create the differences observed. The other reconstructed D_s decay chains show a discrepancy at the same order. One can conclude, that the (corrected) D_s decays are simulated appropriately and the difference arises due to the limited knowledge of D_s production processes. Hence, the size of the True- D_s background is expected to be different in data and the simulation. This results in different normalizations of the description of this contribution in the fit, which is automatically covered by the fitting package used.

6.3 $D_s \rightarrow \phi\pi$ reconstruction

Figure 6.7 shows the data-MC comparison for the input variables of the MLP, used for the $D_s \rightarrow \phi\pi$ reconstruction. Again, the number of D_s events in the data sample is overestimated by Monte Carlo and the relative difference of the vertex probability distribution shows an increment at low values. Although the simulated $D_s \rightarrow \phi\pi$ branching ratio agrees with the actual PDG values [8]², there remains a high uncertainty in the number of the dominating True- D_s background events, as described in the previous section.

²The correction factor in this decay chain is 1.01.

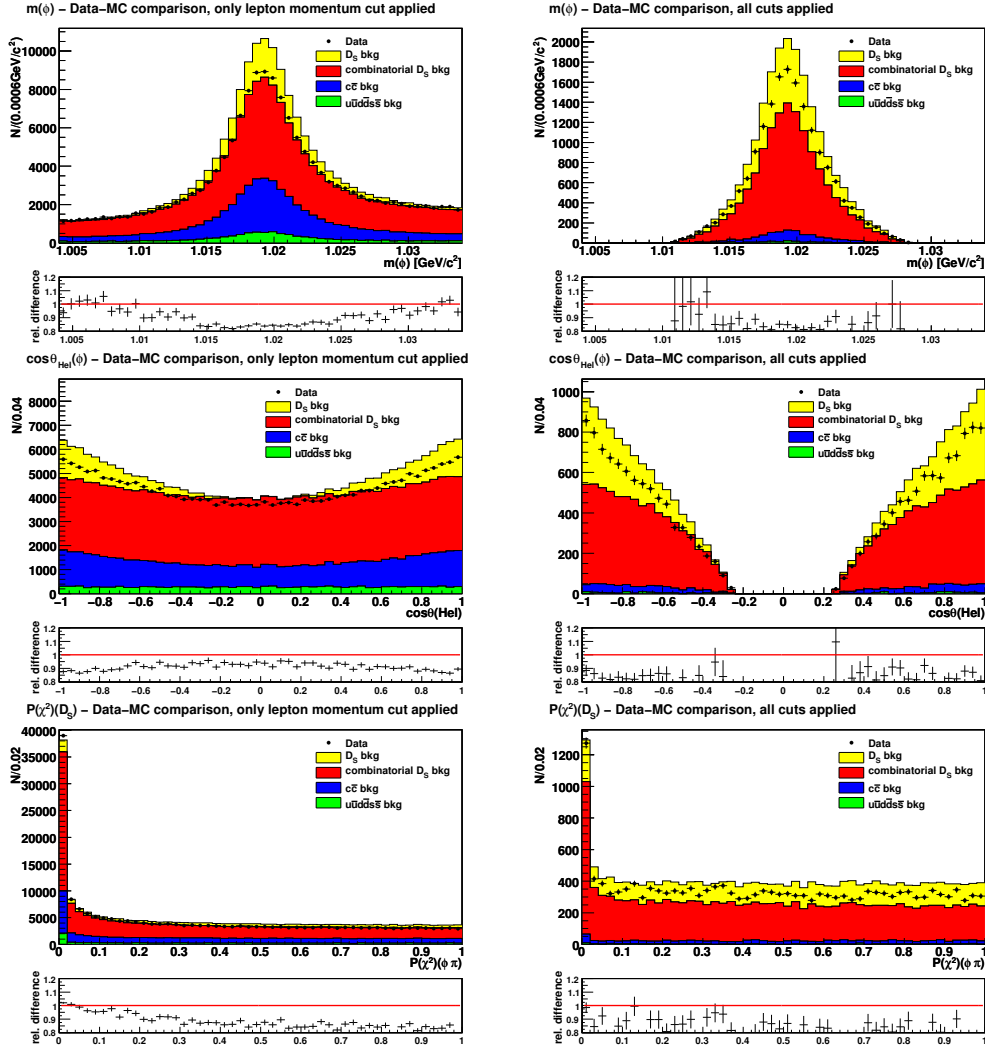


Figure 6.7: Data-MC comparison of the input variables of the MLP, used for the $D_s \rightarrow \phi\pi$ reconstruction: ϕ mass (top), helicity angle distribution (middle) and vertex probability of the D_s , reconstructed in the $\phi\pi$ final state (bottom). Left: with lepton momentum cut only. Right: after the full selection.

The comparison plots for the MLP output of the three combined variables, as well as the vertex probability of the ϕ candidates are shown in figure 6.8. The expected behavior, an increase of the relative difference for areas, containing more True- D_s events is visible in both histograms.

Figure 6.9 shows the data-MC comparison of reconstructed D_s events in the $\phi\pi$ final state. Again, the number of D_s in the peak is overestimated by MC while the background level agrees very well. Besides that, a small shift in the D_s central value is visible in this decay chain, too.

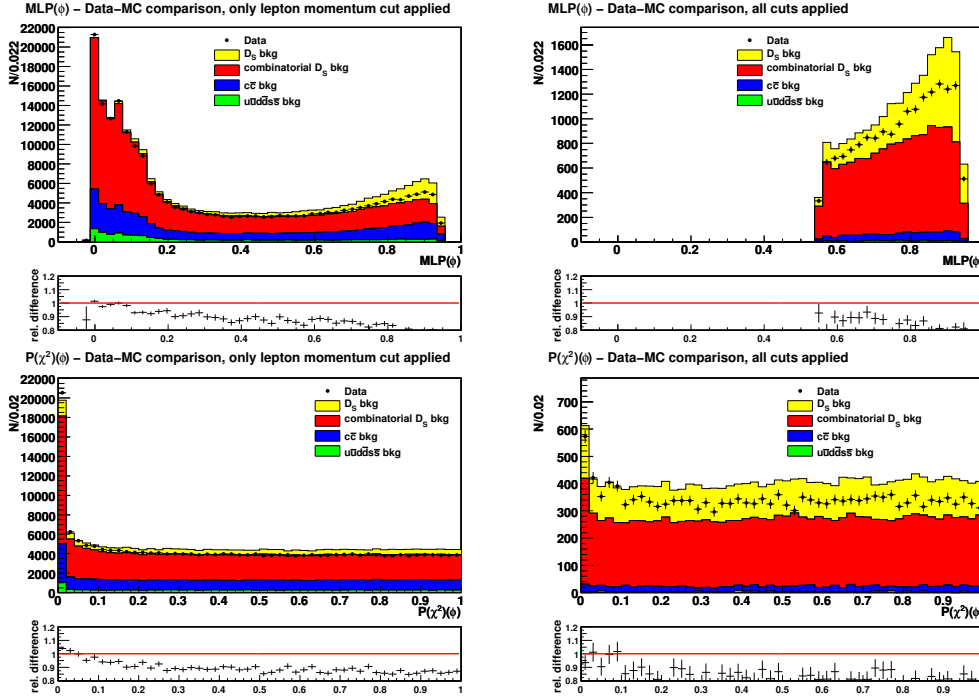


Figure 6.8: Data-MC comparison of the MLP output, used for the $D_s \rightarrow \phi\pi$ reconstruction (top), vertex probability of the ϕ candidates (bottom) Left: with lepton momentum cut only. Right: after the full selection.

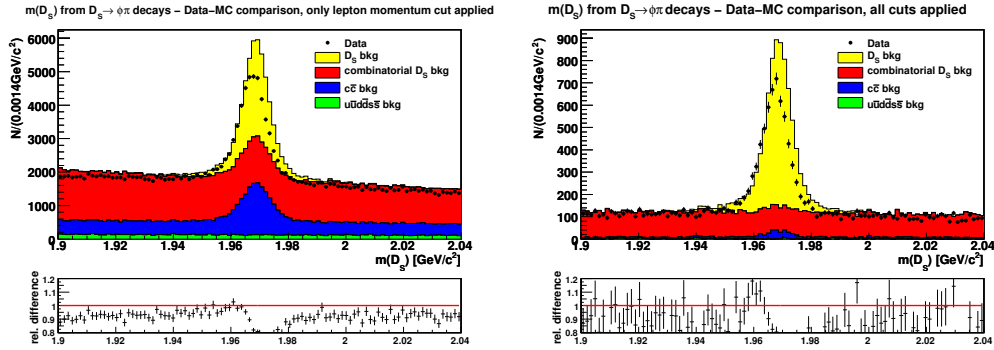


Figure 6.9: Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow \phi\pi$ reconstruction chain. Left: comparison with just the lepton momentum cut applied, right: comparison after the full selection.

6.4 $D_s \rightarrow K^{*0}K$ reconstruction

The data-Monte Carlo comparison plots for the MLP input variables of the $D_s \rightarrow K^{*0}K$ reconstruction chain can be seen in figure 6.10.

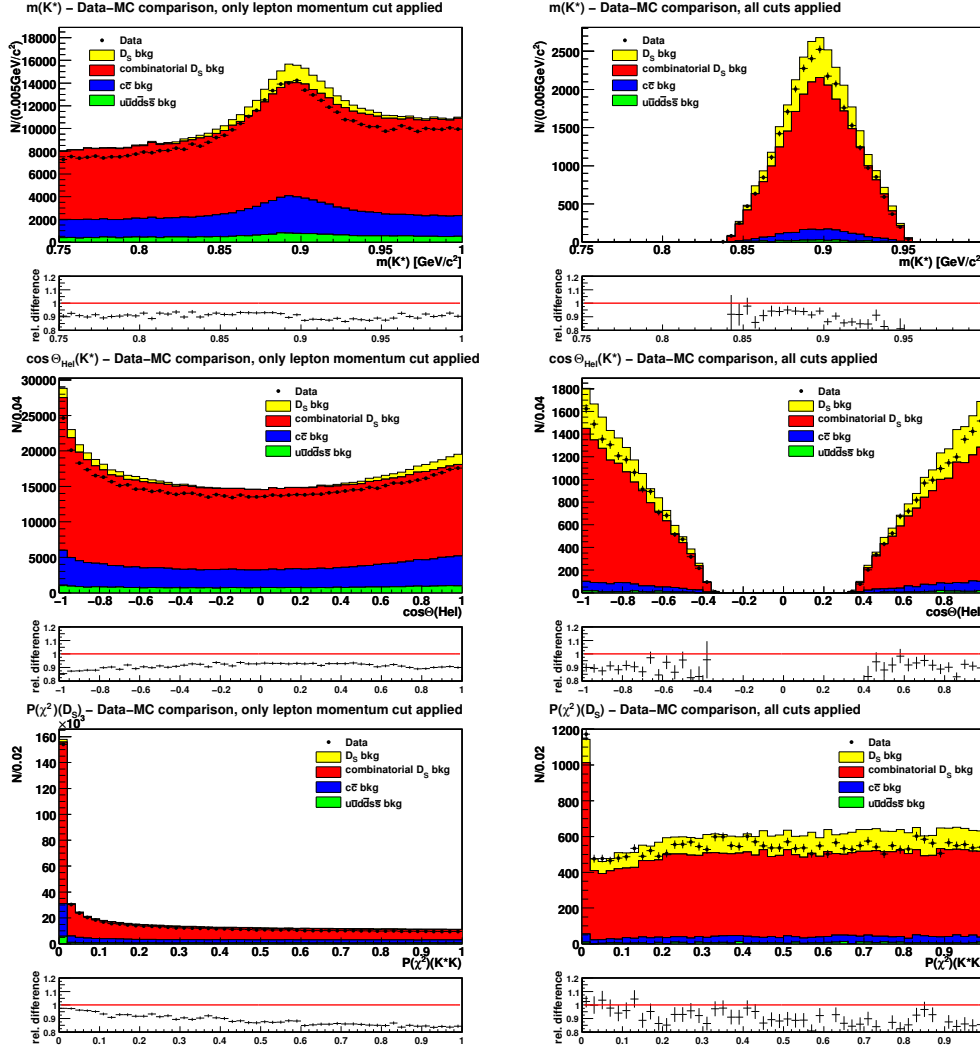


Figure 6.10: Data-MC comparison of the input variables of the MLP, used for the $D_s \rightarrow K^{*0}K$ reconstruction: K^{*0} mass (top), helicity angle distribution (middle) and vertex probability of the D_s , reconstructed in the $K^{*0}K$ final state (bottom). Left: with lepton momentum cut only. Right: after the full selection.

Besides the scaling problem and the behavior of the vertex probability already described in section 6.2, no significant differences are visible. Figure 6.11 shows the comparison for the MLP output variable and the K^{*0} vertex probability. Both show a reasonable agreement between data and MC. The mass of the reconstructed D_s events in this decay chain, depicted in figure 6.12, shows a disagreement in the combinatorial background region before all cuts have been applied.

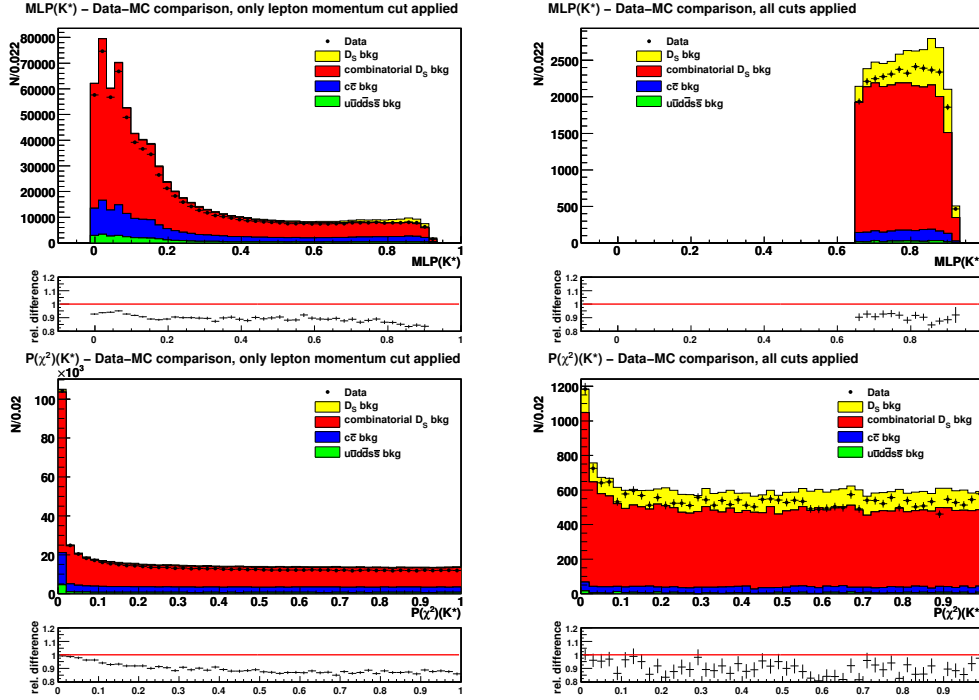


Figure 6.11: Data-MC comparison of the MLP output, used for the $D_s \rightarrow K^{*0}K$ reconstruction (top), vertex probability of the K^{*0} candidates (bottom) Left: with lepton momentum cut only. Right: after the full selection.

It vanishes after using the full reconstruction sequence, due to the reduction of continuum background. Again, a slight shift of the D_s mass central value and the overestimation of the correctly reconstructed D_s is visible³.

Figure 6.13 shows the data-MC comparison for the applied vetoes in this reconstruction chain. As can be seen, there's a good agreement for both vetoes. The plots after all cuts have been applied are shown for completeness reasons only.

³The scaling factor for the D_s decay in this reconstruction chain is 0.974.

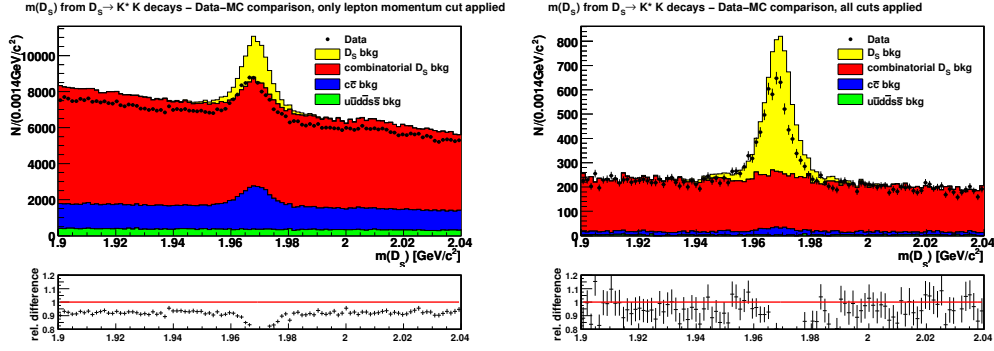


Figure 6.12: Data-MC comparison of the D_s mass distribution for the $D_s \rightarrow K^{*0} K$ reconstruction chain. Left: with lepton momentum cut only. Right: after the full selection.

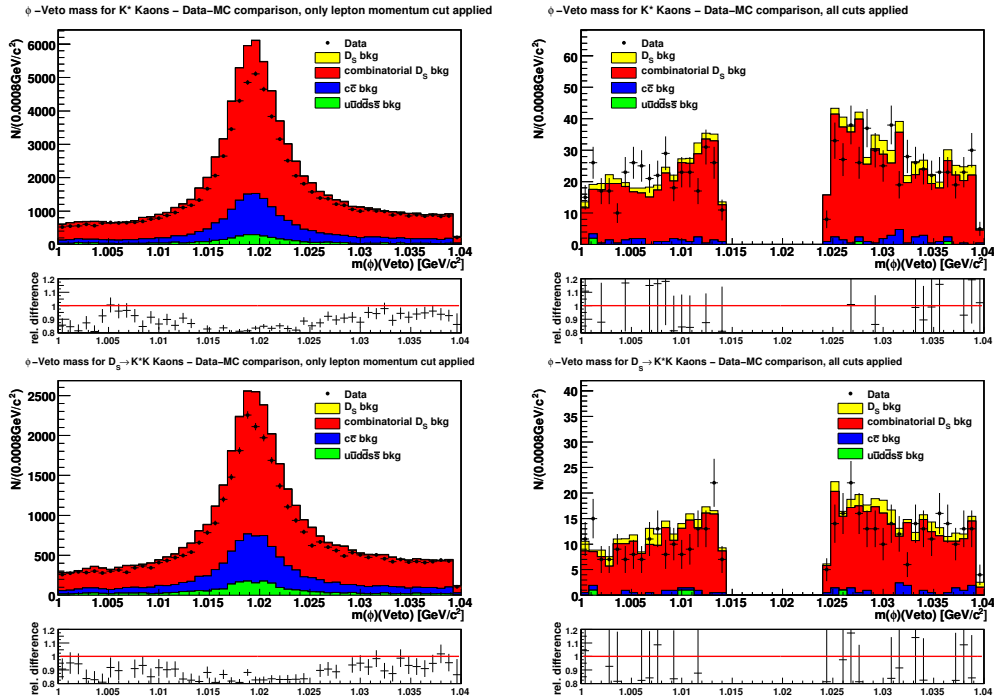


Figure 6.13: Data-MC comparison of the vetoes, used for the $D_s \rightarrow K^{*0} K$ reconstruction: top plot shows the ϕ veto mass of the kaon used for the K^{*0} reconstruction, the bottom plot the one of the kaon used to reconstruct the D_s . Left: with lepton momentum cut only. Right: after the full selection.

6.5 B candidate reconstruction

Data-Monte Carlo comparison plots of the input variables for the MLP of the B shape are shown in figure 6.14. As can be seen, the agreement between the simulation and the data is at a reasonable level.

Figure 6.15 shows the comparison plots for the output of the MLP, as well as for the vertex probability of the $D_s K \ell$ combination.

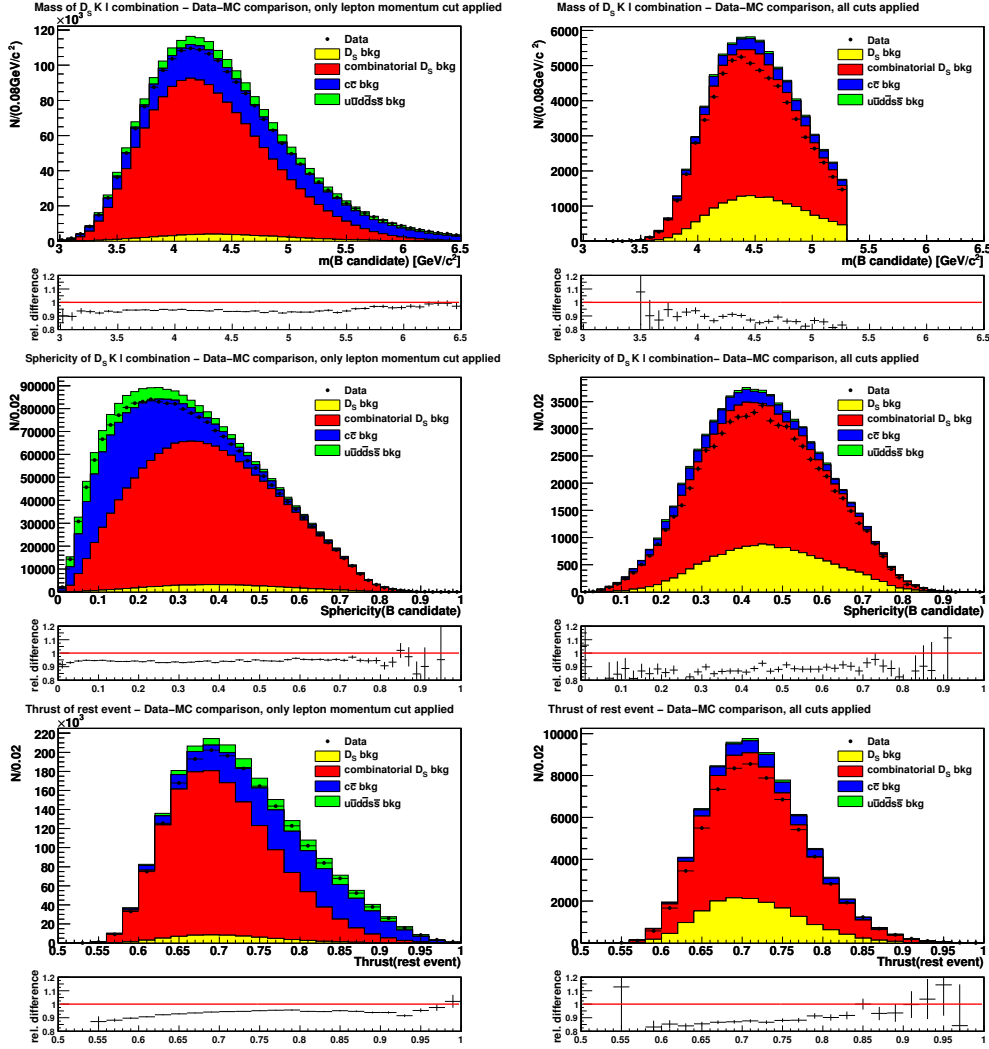


Figure 6.14: Data-MC comparison of the input variables of the MLP, used for the B -candidate reconstruction: mass of the $D_s K \ell$ combination (top), sphericity of the $D_s K \ell$ combination (middle) and Thrust of the rest of the event (bottom). Left: with lepton momentum cut only. Right: after the full selection.

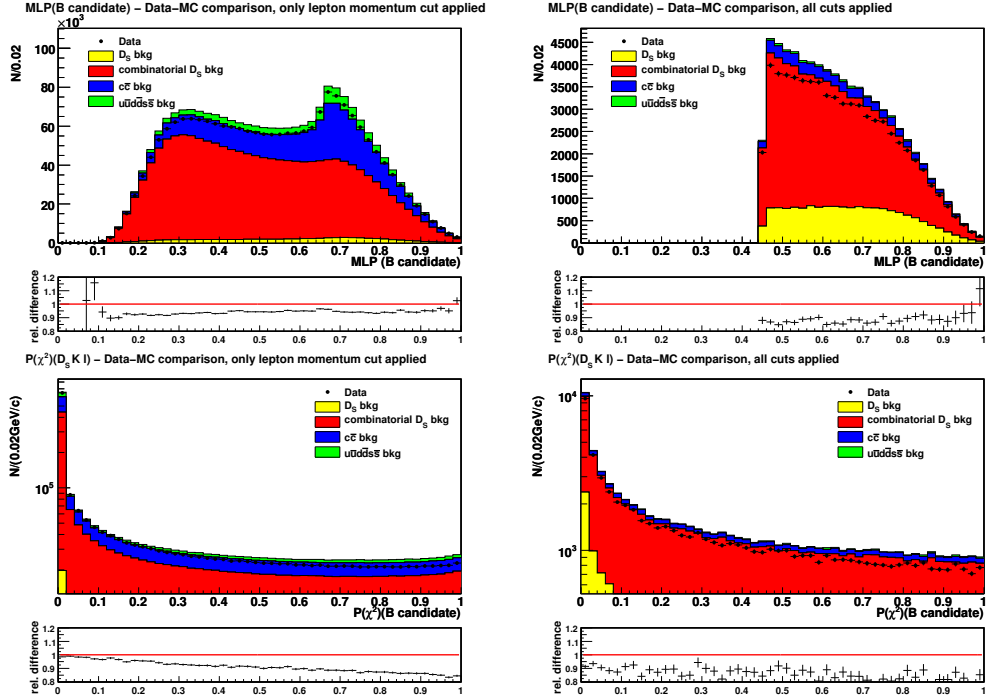


Figure 6.15: Data-MC comparison of the MLP output of the B -reconstruction (top) and of the vertex probability of the $D_s K l$ combinations. Left: with lepton momentum cut only. Right: after the full selection.

There's a good agreement for the output of the MLP, while the increment of the difference for the vertex probability is much stronger than for the other reconstructed composite particles, because this kinematic fit requires five tracks to come from the same vertex. In general, the more tracks are merged into one vertex, the more sensible the vertex fit will be to the errors assigned to the tracks from the track reconstruction. Since the resolution of the tracks is better in MC (because of undescribed material interaction, etc. which affect only data), the errors get smaller and the vertex probability increases. Therefore, this kind of slope of the difference plots is expected. A re-weighting of the MC distribution has been done as a systematics study in order to achieve a constant ratio between data and MC, leading to negligible uncertainties.

6.6 Suppression of Upper vertex decays

The data-MC comparison plots for the angles between the reconstructed D_s and the additional kaons are shown in figure 6.16.

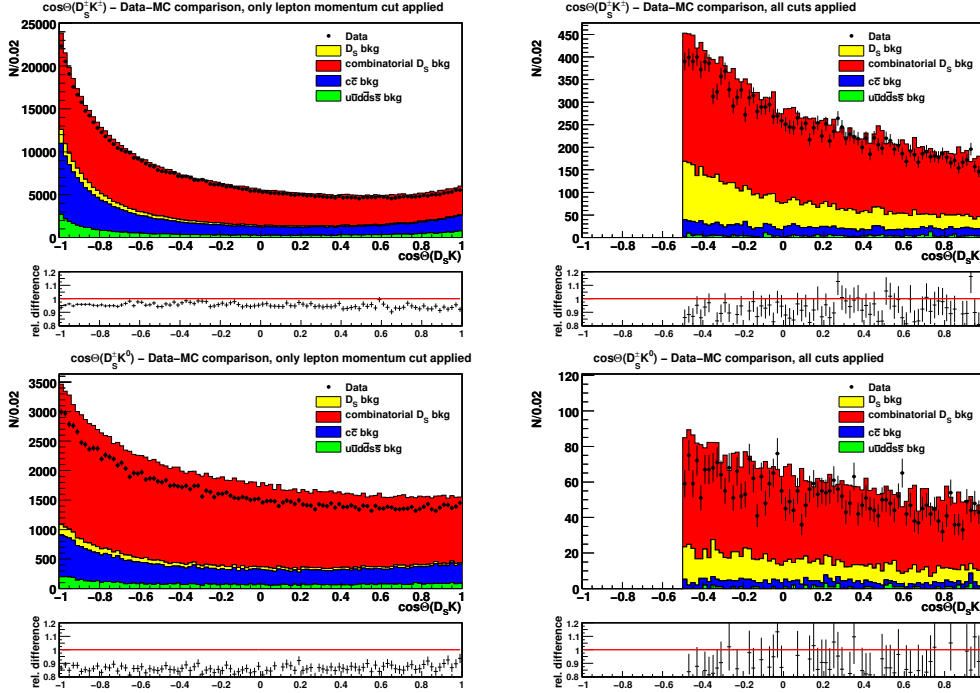


Figure 6.16: Data-MC comparison of the angle between the reconstructed D_s and an additional kaon of the event. The top plots show the distribution for charged kaons, the bottom plots for neutral kaons. Left: with lepton momentum cut only. Right: after the full selection.

Only differences due to scaling effects are visible, while the shape is in good agreement for both, charged and neutral kaons.

6.7 Lepton momentum distributions

The comparison of the lepton momentum distribution can be seen in figure 6.17. No differences besides the overall scaling problem are visible.

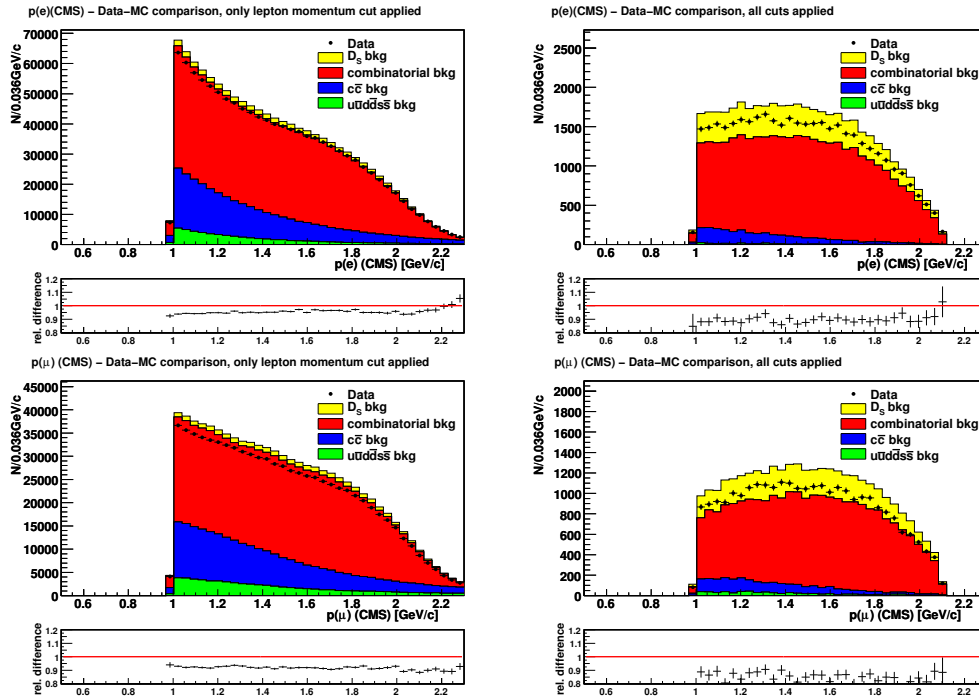


Figure 6.17: Data-MC comparison of the electron- (top) and muon-momenta (bottom). Left: with lepton momentum cut only. Right: after the full selection.

Chapter 7

Analysis method and fit

A major challenge of semileptonic analyses is the neutrino not being detectable. Thus, the B candidate can not be fully reconstructed and uses only the information of the D_s , the kaon and the electron/muon tracks. In the first part of this chapter, different possibilities of performing an indirect neutrino reconstruction are discussed. Secondly, the development of the fit technique by parametrizing the data with *Probability Density Functions* (PDFs), describing the shape of the distribution of the variables, is shown as well as the validation of the fit method on various Monte Carlo samples. The final fit of the data is presented at the end of this chapter.

7.1 Neutrino reconstruction

Indirect neutrino reconstruction can be done using kinematic consistency requirements, derived from the four momentum of the neutrino, produced by a decay $B \rightarrow Y\nu$, where the Y represents all detectable particles in the final state.

The variable $\cos \Theta_{BY}$

Based on four momentum conservation in the B decay, the neutrino's four momentum can be written as

$$\mathbf{p}_\nu^2 = 0 = (\mathbf{p}_B - \mathbf{p}_Y)^2 = M_B^2 + M_Y^2 - 2(E_B E_Y - |\vec{p}_B| |\vec{p}_Y| \cos \Theta_{BY}), \quad (7.1)$$

where $\mathbf{p}_{B/Y} = (E_{B/Y}, \vec{p}_{B/Y})$ and $M_{B/Y}$ denotes the four momentum and mass of the B meson and the Y -System, respectively. The neutrino mass is set to 0 due to its smallness. $\cos \Theta_{BY}$ is the angle between the B meson and the Y -System. This angle is now defined by measurable quantities only¹ as

$$\cos \Theta_{BY} = \frac{2}{2E_B E_Y - M_B^2 - M_Y^2} |\vec{p}_B| |\vec{p}_Y|$$

and needs to satisfy the mathematical definition of a cosine, $|\cos \Theta_{BY}| < 1.0$. Figure 7.1 shows the $\cos \Theta_{BY}$ distributions for the electron and muon channel Signal MC events².

¹Actually, only the quantities of the Y -System are taken from the reconstruction, while the B meson quantities are taken from the beam information. This leads to an ambiguity, since only the absolute value of the momentum of the B is known, but not its flight direction.

²Unless not quoted otherwise, the signal component in all shown plots is scaled to be equivalent to the luminosity of the data used in this analysis and by assuming a signal branching ratio of $\mathcal{BR} = 5 \times 10^{-4}$.

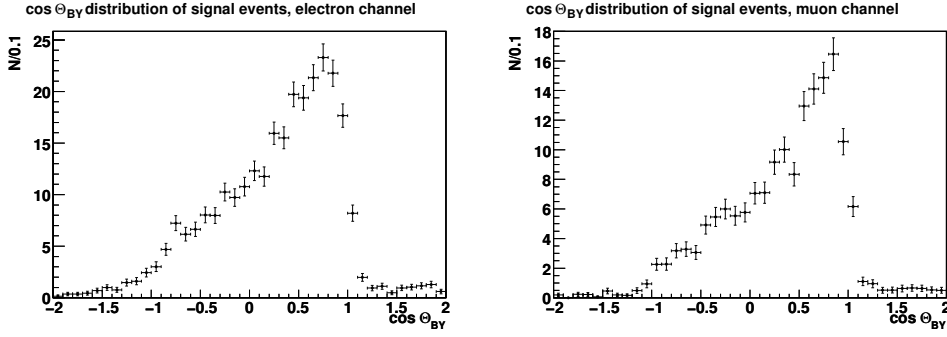


Figure 7.1: $\cos \Theta_{BY}$ distribution for signal events for electron (left) and muon channel (right). Signal MC is scaled to be equivalent to data luminosity, assuming $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = 5 \times 10^{-4}$.

As can be seen, most of the candidates are distributed between -1 and 1 , as expected. The edges at $\cos \Theta_{BY} = \pm 1$ are smeared due to detector resolution limits and fluctuations of the beam energy, which are not simulated accurately in the *BABAR* Monte Carlo. An additional effect, causing signal events to have $|\cos \Theta_{BY}| > 1$ is the emission of bremsstrahlung (resulting in a shift to lower values of $\cos \Theta_{BY}$).

The variable Missing Mass

A second approach for performing a neutrino reconstruction is to neglect the small momentum of the B meson for performing a neutrino reconstruction in equation 7.1 and interpret the resulting mass as a *Missing Mass* (MM), relative to the B meson mass:

$$MM^2 = (E_B - E_Y)^2 - |\vec{p}_Y|^2 = m_\nu^2. \quad (7.2)$$

If this equation would be exact, a discrete peak is expected at the neutrino mass, which is essentially 0. But due to the negligence of the B momentum ($|\vec{p}_B| \sim 300 \text{ MeV}/c$) and its unknown direction, the MM^2 distribution becomes a gaussian distribution at the same mean. The width of the gaussian depends on the size of the neglected B momentum and is, within the error limits, the same for both lepton channels. Figure 7.2 shows the Missing Mass squared distributions for both, electron and muon channel Signal MC events.

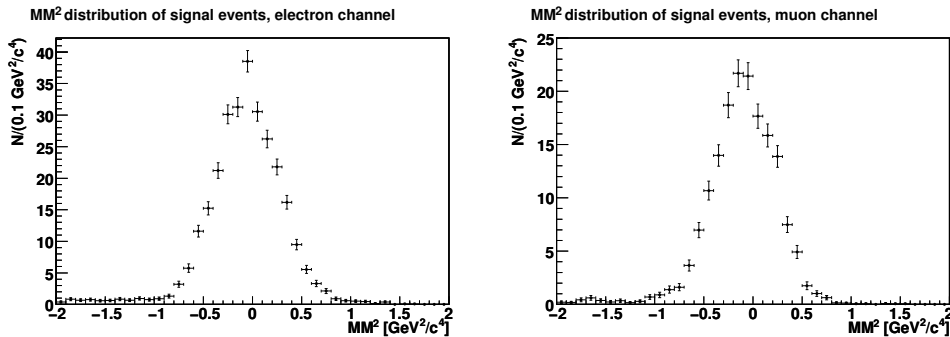


Figure 7.2: MM^2 distribution for signal events for electron (left) and muon channel (right). The Signal MC sample has been scaled to be equivalent to data luminosity, assuming $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = 5 \times 10^{-4}$.

The choice of the fit variable

The decision which of these variables to use for the extraction of the number of signal events, as well as the choice of the fit technique has been made considering the following three criteria:

- complexity of the PDF used for parametrizing the signal events
- complexity of the PDF used for parametrizing the background events (including both, combinatorial and True- D_s background)
- size of the signal with respect to underlying background events

As can be seen in figures 7.1 and 7.2, the $\cos \Theta_{BY}$ distribution shows a more complicated behavior than the MM^2 with an increase starting around -1 , just as expected for the mathematical meaningful region. Nevertheless, the steady increase between -1 and 1 and the sharp drop at $\cos \Theta_{BY} \sim 1$ is hard to parametrize. This can be done using a KEYS PDF, which provides an unbinned, non-parametric estimate of the PDF, underlying the data [42]. Nevertheless, developing a parametrization for the MM^2 is more straight forward than for $\cos \Theta_{BY}$.

The distribution of background events for both variables is shown in figures 7.3 and 7.4³.

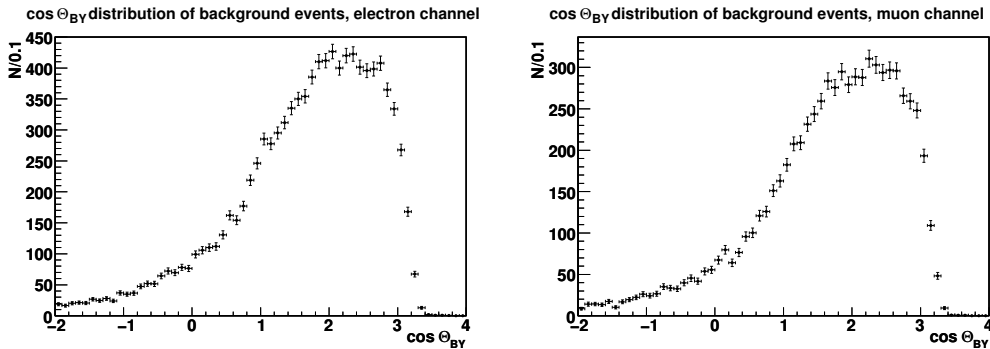


Figure 7.3: $\cos \Theta_{BY}$ distribution for background events for electron- (left) and muon channel (right). The distribution is scaled to be equivalent to data luminosity.

For both plots only events in the D_s mass signal region (within $\pm 2.5\sigma$ around the mean of a fit to the mass distribution) have been used. As can be seen, the distributions have a very similar shape, but are mirrored in terms of increasing and decreasing flanks. In both cases, the signal component is located on a flank, either on a decreasing (MM^2) or on an increasing flank ($\cos \Theta_{BY}$). Therefore none of the variables has an advantage here.

Adding both components, the expected background and the signal, gives an idea of the distribution which finally needs to be fitted to extract the signal yield. Figure 7.5 shows the distributions for electron and muon channel as a stack of the contributions of the different event types where all three D_s reconstruction channels have been added.

³Unless quoted otherwise, the background component in all plots is scaled to be equivalent to the luminosity of the data used in this analysis.

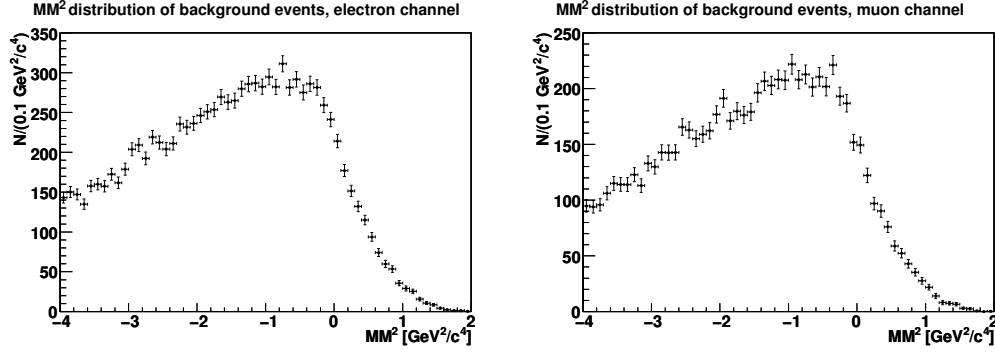


Figure 7.4: MM^2 distribution for background events for electron- (left) and muon channel (right). The distribution is scaled to be equivalent to data luminosity.

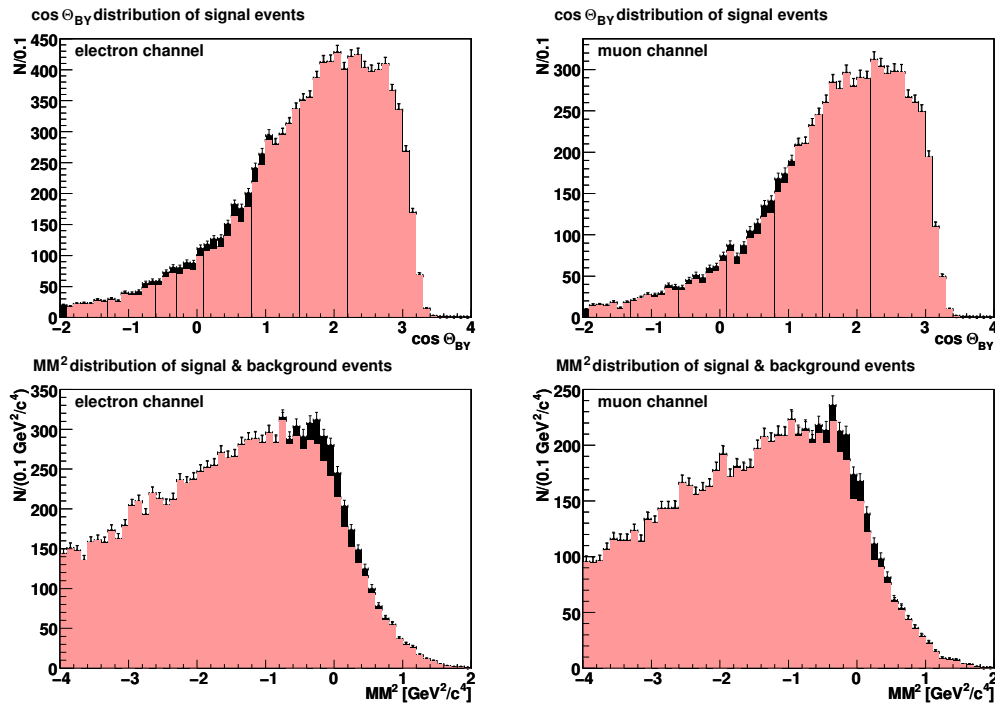


Figure 7.5: Distributions of possible fit variables for a neutrino reconstruction for signal (black) + background events (red) for electron (left) and muon channel (right). Top row: $\cos \Theta_{BY}$. Bottom row: MM^2 . Signal MC is scaled to be equivalent to data luminosity by assuming a signal branching ratio of 5×10^{-4} . The background component is scaled to be equivalent to data luminosity, as well.

The $\cos \Theta_{BY}$ distributions show only a tiny enhancement in the area between -1 and 1 on a slow increasing flank of the distribution. The signal amount increases with larger $\cos \Theta_{BY}$ values and is almost invisible for small values. In comparison to that, the signal is visible much clearer in the MM^2 distribution. Nevertheless, it's location almost on top of a decreasing flank makes the differentiation between signal and background very difficult.

Considering the three points mentioned, the MM^2 variable has been chosen for the signal extraction, due to the easier handling of the signal probability density function on a high background sample. It has to be noted, that both variables are highly correlated, since they rely on the same physics background. Thus, a cut on one of the variables translates directly into a cut on the opposite shoulder of the other variable and doesn't reject any background events in the particular signal region.

7.2 Fit technique

As mentioned in the previous section, fitting the MM^2 distribution is difficult since the signal is located close to the drop off point of the decreasing flank. The contributions of the single components (signal, combinatorial background, True- D_s background) cannot be distinguished from each other easily by the fit, resulting in unstable results, making the fit unusable. An approach to increase the fit stability is to parametrize the combinatorial background components separately. This can be achieved by using a simultaneous fit of regions containing signal (and background) events (*signal region*) and regions without any signal contribution (*sidebands*) for the signal extraction, where the parametrization of background events can be derived from the sidebands. The mass of the reconstructed D_s candidates is used for the classification. The samples are split up based on the underlying D_s decay chain. Each lepton flavor is treated separately and the results are combined later. In summary, there are six different samples to be analyzed simultaneously for each lepton channel: Signal region and sideband for each of the three D_s decay chains. The development of the simultaneous fit is described below. It starts with the definition of the signal region and sideband in the $m(D_s)$ distribution and the determination of the number of True- D_s events by a fit. Then, the candidates are sorted into the particular regions and the MM^2 distributions of both regions in the three D_s reconstruction chains are fitted simultaneously to retrieve the signal yield. After the development, the fit is validated on several kinds simulated event samples.

7.2.1 Development of the simultaneous fit

Definition of the $m(D_s)$ signal region

Each sample is divided into two parts, events in the $m(D_s)$ signal region and sideband region. The latter is defined as:

$$\begin{aligned} \text{upper sideband : } & 2.0 \text{ GeV}/c^2 \leq m(D_s) \leq 2.04 \text{ GeV}/c^2 \\ \text{lower sideband : } & 1.9 \text{ GeV}/c^2 \leq m(D_s) \leq 1.94 \text{ GeV}/c^2 \end{aligned}$$

While the sideband is the same for all six samples, the signal region is slightly different. It is defined by fitting the $m(D_s)$ distribution with a single gaussian for the signal fraction and a straight line for the combinatorial background and taking a 2.5σ region around the mean of the fit as signal region. As mentioned in chapter 6, the D_s mass distribution shows small differences between data and MC. To avoid an indeterminable bias and a systematic uncertainty due to the choice of the D_s signal region, the determination of the signal region using the fit is done on data as well. The D_s mass distribution for data and generic MC, overlaid by the parametrization with the fit defining the signal region, is shown in figures 7.6 and 7.7 for the electron channel and the muon channel, respectively.

After the classification of events as signal region and sideband, another data-MC comparison of the MM^2 distributions of all decay chains and regions is done as a further consistency check regarding the accuracy of the simulated shapes of the MM^2 .

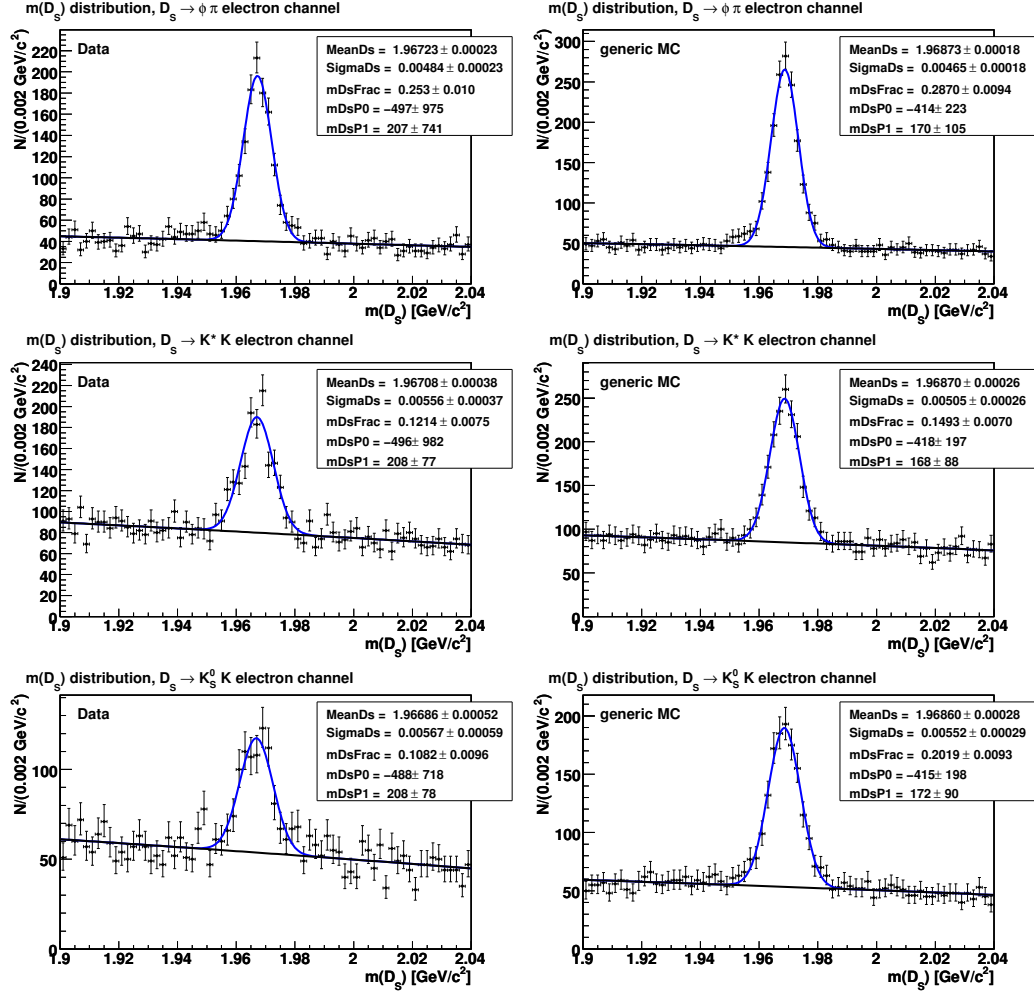


Figure 7.6: D_s mass distributions of the electron channel for data (left) and generic MC (right). Top row: $D_s \rightarrow \phi \pi$ channel, middle row: $D_s \rightarrow K^{*0} K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel. The MC components are scaled to be equivalent to data luminosity.

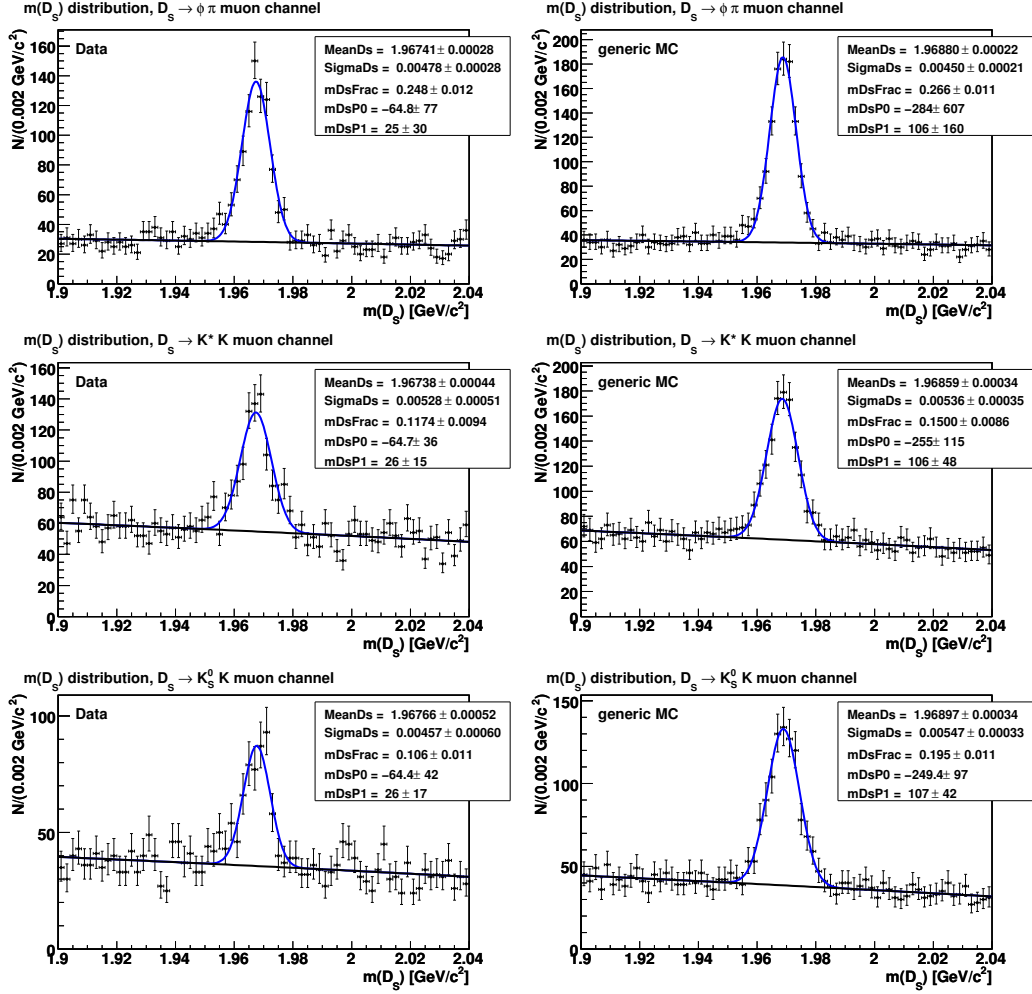


Figure 7.7: D_s mass distributions of the muon channel for data (left) and generic MC (right). Top row: $D_s \rightarrow \phi \pi$ channel, middle row: $D_s \rightarrow K^{*0} K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel. The MC components are scaled to be equivalent to data luminosity.

Data-MC comparison of signal region and sideband MM^2 distributions

The Monte Carlo events have been scaled to be equivalent to data luminosity to alleviate the comparison. It has to be noted, that the analysis has been done as a *blind analysis*, meaning that the development of cuts and fit technique were performed on simulated events only. The validation of the fit is done on Monte Carlo as well. Not until all these steps have been completed successfully, data has been analyzed. This procedure should prevent the analyst from artificially creating a signal leading to wrong results. Therefore, data-MC comparison has been done only for the $m(D_s)$ sideband regions and the signal region excluding the area which is later used for the extraction of the signal yield, $|MM^2| < 1.5 \text{ GeV}^2/c^4$. All plots use the same color code like the ones in chapter 6:

- red for $B\bar{B}$ combinatorial D_s background
- yellow for $B\bar{B}$ True D_s background
- blue for $c\bar{c}$ background
- green for $u\bar{u}/d\bar{d}/s\bar{s}$ background

Figure 7.8 shows the data-MC comparison for the $m(D_s)$ sideband regions for all reconstruction chains and lepton channels. Besides a small global overshoot of the Monte Carlo simulation, there also appears a slight difference in the slope of the distributions in the electron channel. While for bins in the region of $MM^2 < -3.5 \text{ GeV}^2/c^4$ MC contains less entries than data, it changes to an overshoot of MC above this value. Since this slope differences are present only in a region away from the relevant range, later used in the fit, there arise no complications for the further analysis.

Comparisons for the D_s mass signal region have been made in the MM^2 range of $-4.5 \text{ GeV}^2/c^4 < MM^2 < -1.5 \text{ GeV}^2/c^4$. Figure 7.9 shows these distributions. The normalization problem due to the overshoot of True- D_s events in the MC sample, already described in chapter 6, results in larger global differences between data and MC in these plots. Focusing just on possible shape differences, only the $D_s \rightarrow K^{*0}K$ electron channel shows a problem in the region of $-2.8 \text{ GeV}^2/c^4 < MM^2 < -2.0 \text{ GeV}^2/c^4$. The data shows a dip, while the MC increases to a plateau like in all other decay chains.

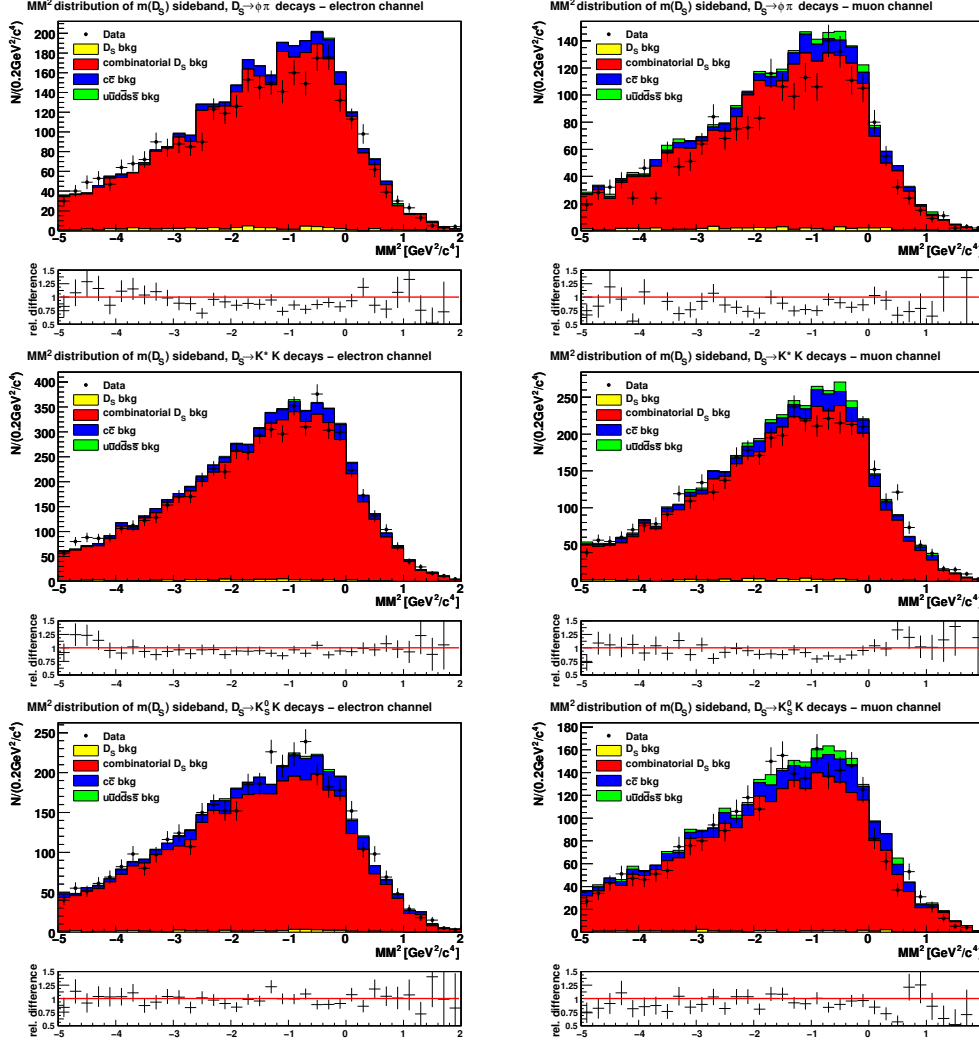


Figure 7.8: MM^2 data-MC comparison of the $m(D_s)$ sidebands for electron channel (left) and muon channel (right). Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel.

Several studies have been made to understand that difference. A comparison of the events inside and outside the dip region and with the other D_s decay chains found no difference in the composition. Furthermore, relaxing the cuts of the reconstruction of this particular channel seemed to reduce the dip, but the comparison of the MM^2 distribution with only event shape cuts applied also showed differences in the corresponding bins. Apparently, it is more likely that the effect is a statistical fluctuation of the background, than that the MC overestimates the entries. The good agreement of the shape in the other decay chains as well as in the $D_s \rightarrow K^{*0}K$ muon channel backs this up.

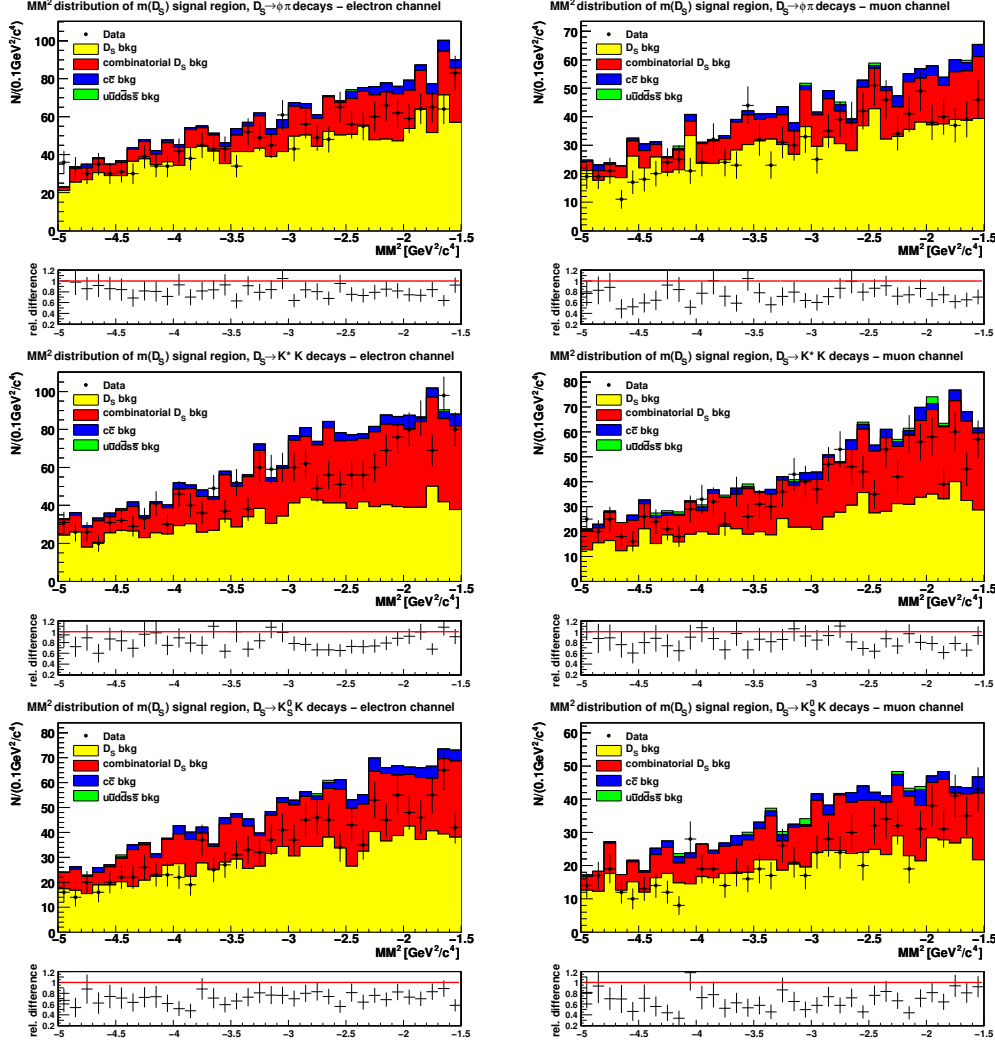


Figure 7.9: MM^2 data-MC comparison of the $m(D_s)$ signal regions for electron channel (left) and muon channel (right). The MM^2 signal region is not shown in the plot. Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel.

Altogether, the comparisons show a satisfying agreement between data and MC and one can conclude, that no signal can accidentally be produced in the sidebands, which was a major concern before examining the comparison plots. If the sideband shows a dip in the relevant region ($|MM^2| < 1.5 \text{ GeV}^2/c^4$), an artificial signal would be created because the assumption that the combinatorial background distributions are the same in the MM^2 signal region and sidebands is no longer satisfied. On the other hand, a peaking structure would also affect the fit result. Such dips or peaks can arise due to changes of selection efficiency for particular events and are only relevant if they have a width similar to the signal width. Since the comparison showed no hints of these kind of structures, the possibility of artificially creating a signal through sideband fluctuations can be excluded.

Parametrization of the background distributions

Since the combinatorial D_s background component is part of the MM^2 distributions for signal and sideband regions, it is possible to parametrize this fraction by the same probability density function simultaneously in both regions. The True- D_s background component, being present only in the signal region, needs to be parametrized separately.

The MM^2 distributions of the True- D_s background are very similar in all D_s reconstruction chains. Thus, a common parametrization of the three distributions is possible. As a first step, several possible parametrizations have been tested through a single channel fit on the $D_s \rightarrow \phi\pi$ electron chain. The tested PDFs were mirrored Landau-functions, Fermi-functions and gaussians, combined with polynomials. Finally, the usage of a Fermi-function (without additional polynomials) to parametrize the True- D_s background fraction turned out to be the best choice, combining a low number of parameters with a good χ^2/NDF value.⁴ The Fermi-function is defined as:

$$f(x) = \frac{1}{e^{((x-M0)/\text{ExpoConst})} + 1}, \quad (7.3)$$

where M0 represents the MM^2 drop-off value and ExpoConst is responsible for the smearing of the Fermi-edge. Figure 7.10 shows the simultaneous fit for the electron and the muon channel for generic MC, scaled to data luminosity, while table 7.1 gives the results of the fits. As can be seen in the figures, there are several bins in almost all histograms, where a peaking structure is visible. Most of them can be ascribed to the chosen binning and have no effect on the fit, which is done as unbinned fit. It has to be noted, that due to the known width of the signal and the usage of a simultaneous fit the influence of fake signals caused by statistical fluctuations is very small.

Table 7.1: Results of the simultaneous fits with the Fermi-function PDF for electron and muon channel.

lepton channel	M0/(GeV ² /c ⁴)	ExpoConst/(GeV ² /c ⁴)	χ^2/NDF
electron	0.303 ± 0.034	0.333 ± 0.018	135.4/118
muon	0.247 ± 0.041	0.346 ± 0.022	131.2/117

First tests of the complete fit on the full sample allowed the D_s background parameters to float within the errors, documented in table 7.1. It turned out that this makes the the fit unusable due to the superposition of two different background PDFs in the signal region, one describing the True- D_s events and one the combinatorial background. Unfortunately, both distributions have a similar, but non-equal shape making it impossible for the fit to distinguish them. A combined description of both fractions with one PDF should not be used since they represent two physically different contributions. Therefore the parameters of the PDF describing the True- D_s fraction need to be fixed in the complete fit. The results of the simultaneous fits of the True- D_s background are used input values.

⁴The number of degrees of freedom of a fit is abbreviated with 'NDF' all through this thesis.

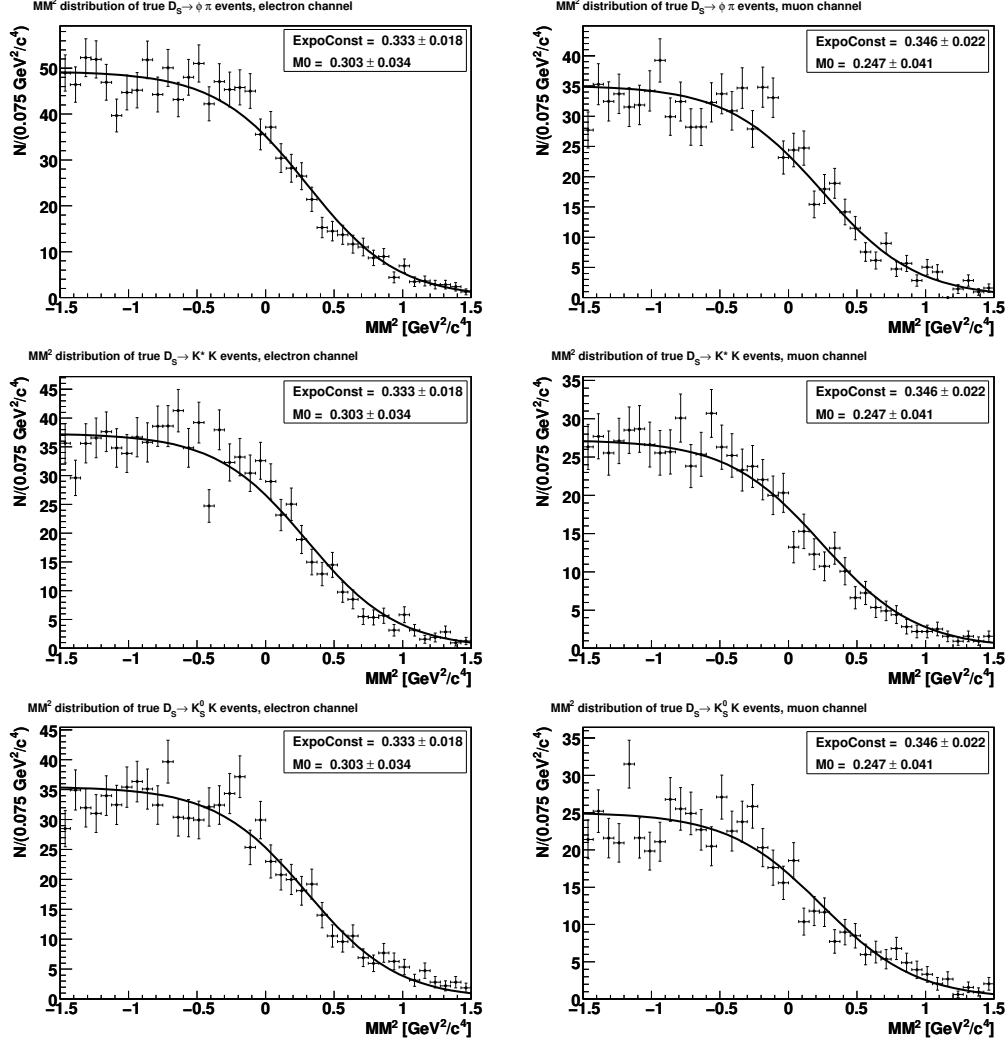


Figure 7.10: Final simultaneous fits with the Fermi-function PDF of the True- D_s background MM^2 distributions for electron channel (left) and muon channel (right). Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel.

The fit errors of the parameters are used later to determine the systematic uncertainty arising from the fixing. Furthermore, the fraction of the True- D_s events in the whole sample is fixed in the final fit as well. Its size can be determined by the fit to the $m(D_s)$ distribution, described at the beginning of this section.

Thus, the parametrization of the background for the signal region (SR) and sideband (SB) for the three D_s decay chains is done with the following composite PDFs for each lepton channel i :

$$\begin{aligned}
PDF_{SR,i}^{\phi\pi} &= f^{\phi\pi} \times PDF_{\text{True-}D_s,i} + (1 - f^{\phi\pi}) \times PDF_{\text{combin. bkg},i} \\
PDF_{SB,i}^{\phi\pi} &= PDF_{\text{combin. bkg},i} = PDF_{SB,i} \\
\\
PDF_{SR,i}^{K^*K} &= f^{K^*K} \times PDF_{\text{True-}D_s,i} + (1 - f^{K^*K}) \times PDF_{\text{combin. bkg},i} \\
PDF_{SB,i}^{K^*K} &= PDF_{\text{combin. bkg},i} = PDF_{SB,i} \\
\\
PDF_{SR,i}^{K_S^0K} &= f^{K_S^0K} \times PDF_{\text{True-}D_s,i} + (1 - f^{K_S^0K}) \times PDF_{\text{combin. bkg},i} \\
PDF_{SB,i}^{K_S^0K} &= PDF_{\text{combin. bkg},i} = PDF_{SB,i}
\end{aligned}$$

with f^{channel} representing the fraction of the True- D_s background in the complete background for the particular D_s reconstruction chain. The parametrization of the combinatorial background is done using two gaussians, whose parameters are floated in the final fit.

Parametrization of the signal component

As already mentioned in section 7.1, the MM^2 distribution shows a gaussian behavior. The distributions of all channels look very similar, only differing by statistical fluctuations. Thus, a simultaneous fit of the three D_s reconstruction chains can be used, in order to reduce the number of fit parameters and to increase the fit stability. The simultaneous fit of the signal fraction is shown in figure 7.11. The fit results are shown in table 7.2. In addition to the fit parameters, the number of fitted signal events in the MM^2 range of $|MM^2| < 1.1$ and the D_s mass signal range is given. For the signal-only fit, a polynomial of first order has been added to the gaussian to account for outlying events, which are too small to be accounted for in the full fit, but dilute the gaussian shape, if no polynomial is used. In the full fit, this polynomial is absorbed by the background components. For the calculation of the signal reconstruction efficiency, only events in the range of $|MM^2| < 1.1$ have been used.

The widths of the gaussians remain fixed to the values, derived by the fits to Signal MC.

Table 7.2: Results of the simultaneous fits to Signal MC for electron and muon channel.

lepton channel	electron channel	muon channel
Mean	$(-0.051 \pm 0.006) \text{ GeV}^2/c^4$	$(-0.082 \pm 0.007) \text{ GeV}^2/c^4$
Width	$(0.313 \pm 0.008) \text{ GeV}^2/c^4$	$(0.299 \pm 0.009) \text{ GeV}^2/c^4$
χ^2/NDF	111.3/108	73.7/90
$N(\text{total})$	3767	2239
$N(\phi\pi)$	1380	808
$N(K^{*0}K)$	895	543
$N(K_S^0K)$	1492	888

This fixation is physics motivated, since the width only appears due to the negligence of the B momentum, which is the same in all B meson decays at the *BABAR* experiment⁵.

⁵Detector resolution effects are negligible compared to the B momentum negligence.

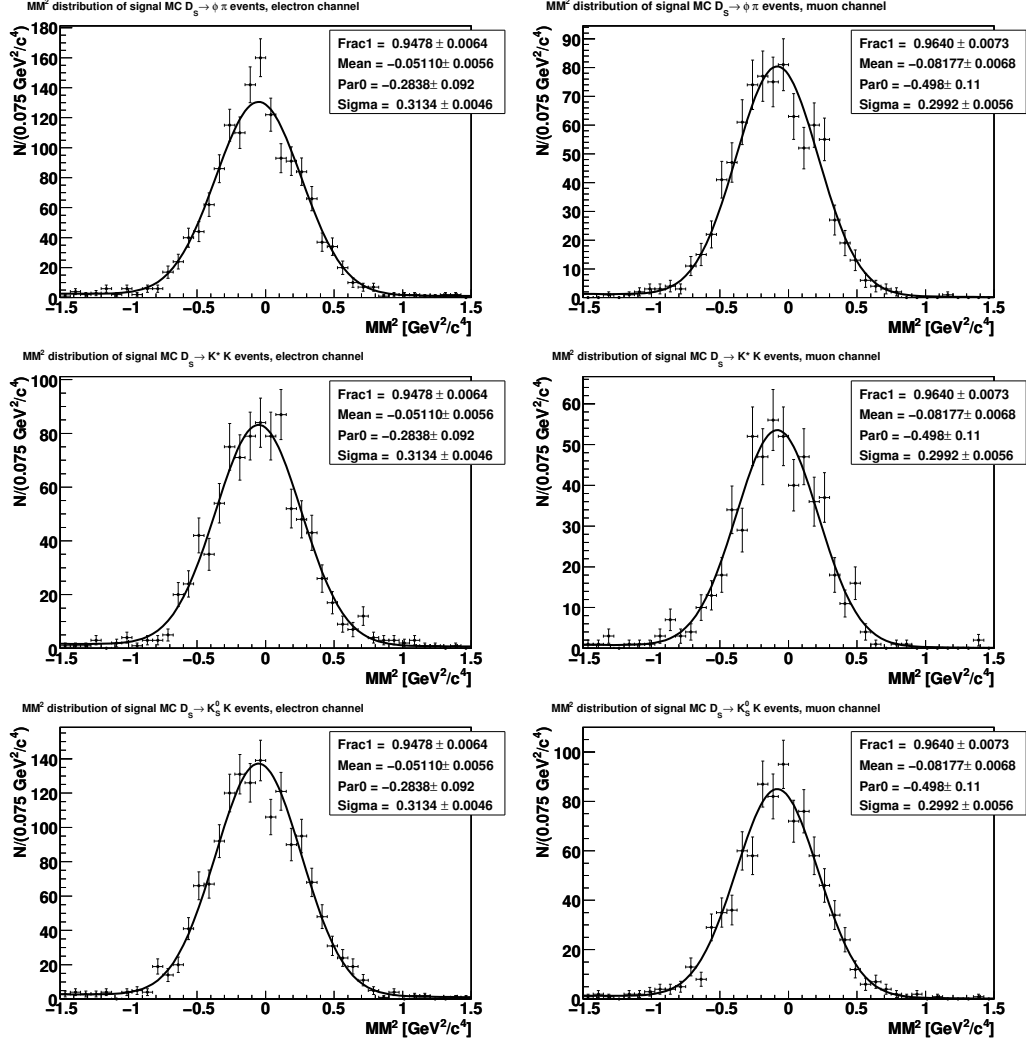


Figure 7.11: Simultaneous fits with a (gaussian + polynomial of 1st order) PDF of the Signal MC MM^2 distributions for electron channel (left) and muon channel (right). Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel.

The observed shift of the mean of the gaussian appears due to various losses, e.g. the production of additional particles or bremsstrahlung [43].

As mentioned in section 2.2.2, the fraction the decays $B \rightarrow D_s^* K \ell \nu_\ell$ compared to the production of a D_s is unknown. If the D_s^* decays through $D_s^* \rightarrow D_s \gamma$, it can contribute to the signal yield. Since the γ cannot be reconstructed in this analysis, an additional particle is missing in the calculation of the MM^2 , which leads to a shift of the mean of its distribution of $\Delta MM^2 \sim 0.2 \text{ GeV}^2/c^4$. Thus, the mean of the gaussian has to be floated in the final fit. The influence of such a contribution of unknown size has been studied by varying the fraction of signal events with a D_s^* to those with a D_s . (See section A.1 in the Appendix for more details.).

It turns out that both decays are indistinguishable for this analysis because the shift is much smaller than the width of the signal distributions, even if both fractions contribute to the signal yield in equal shares. Furthermore, depending on the size of the fraction of the excited composites, the mean of the gaussian distribution shifts to lower MM^2 values, as expected.

Fit setup and constraints on the event yields

To achieve a better stability of the fit, a constraint on the expected signal yields has been inserted. The event yields of the three D_s decay chains are constraint to be equivalent to the same signal branching ratio, taking into account the different efficiencies for the reconstruction and for the branching ratios of the D_s -decays into the final states. Assuming the same signal branching ratio, it is possible to relate the signal yields to each other:

$$N_{\text{exp}}^{\text{decay chain}} \sim \epsilon_{\text{reco}}^{\text{decay chain}} \times \epsilon_{D_s}^{\text{decay chain}} \times \text{branching ratio} \times \mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell)$$

$$N^{K^*K} = N^{\phi\pi} \frac{\epsilon_{\text{reco}}^{K^*K}}{\epsilon_{\text{reco}}^{\phi\pi}} \frac{\epsilon_{\mathcal{BR}}^{K^*K}}{\epsilon_{\mathcal{BR}}^{\phi\pi}} = N^{\phi\pi} R^{K^*K}$$

$$N^{K_S^0K} = N^{\phi\pi} \frac{\epsilon_{\text{reco}}^{K_S^0K}}{\epsilon_{\text{reco}}^{\phi\pi}} \frac{\epsilon_{\mathcal{BR}}^{K_S^0K}}{\epsilon_{\mathcal{BR}}^{\phi\pi}} = N^{\phi\pi} R^{K_S^0K}$$

Using these two constraints, the result of the fit can be expressed as only one number of total signal events, which is calculated by fitting the three subsamples, whose single event yields depend on each other:

$$N^{\text{total}} = N^{\phi\pi} + N^{K^*K} + N^{K_S^0K}$$

$$N^{\phi\pi} = \frac{N^{\text{total}}}{1 + R^{K^*K} + R^{K_S^0K}}$$

$$N^{K^*K} = \frac{N^{\text{total}} \cdot R^{K^*K}}{1 + R^{K^*K} + R^{K_S^0K}}$$

$$N^{K_S^0K} = \frac{N^{\text{total}} \cdot R^{K_S^0K}}{1 + R^{K^*K} + R^{K_S^0K}}$$

Thus, there remains only N^{total} as free parameter, giving directly the total signal yield.

Summarizing the parametrizations, constraints and fixed parameters, the final fit has 10 free parameters:

- 5 parameters for the combinatorial background PDF,
- 3 parameters for the number of background events in the sidebands,
- 1 parameter for the mean of the signal gaussian PDF and
- 1 parameter for the total number of signal events.

The fit is done as an unbinned extended simultaneous maximum likelihood fit. The likelihood function can be written as:

$$\mathcal{L} = \frac{e^{-N^{\text{total}}}}{N!} (N^{\text{total}})^N \prod_i^{N_{\phi\pi, SR}} \prod_j^{N_{\phi\pi, SB}} \prod_k^{N_{K^*0K, SR}} \prod_l^{N_{K^*0K, SB}} \prod_m^{N_{K_S^0K, SR}} \prod_n^{N_{K_S^0K, SB}} \quad (7.4)$$

$$\times PDF_{\phi\pi, SR}(MM_i^2; \alpha_{SR}) PDF_{SB}(MM_j^2; \alpha_{SB}) \quad (7.5)$$

$$\times PDF_{K^*K, SR}(MM_k^2; \alpha_{SR}) PDF_{SB}(MM_l^2; \alpha_{SB}) \quad (7.6)$$

$$\times PDF_{K_S^0K, SR}(MM_m^2; \alpha_{SR}) PDF_{SB}(MM_n^2; \alpha_{SB}) \quad (7.7)$$

In this expression, $N_{D_s \text{ channel, region}}$ represents the number of events in the given D_s decay chain and event region (SR for signal region and SB for sideband region) with $\sum_{D_s \text{ channel, region}} N_{D_s \text{ channel, region}} = N$, N^{total} is the number of expected signal events, $PDF_{\text{channel, SR}}$ represent the probability density functions of the channels signal region (as defined in equation 7.4) and PDF_{SB} is the probability density function of the sidebands. The vectors $\alpha_{SR/SB} = (\alpha_{SR/SB, 1}, \dots, \alpha_{SR/SB, p})$ with p being the total number of free parameters describing the signal and background distributions are the unknown fit parameters in the signal region and sideband PDFs.

As an alternative approach, the usage of a binned simultaneous χ^2 fit of samples, where the sideband components were subtracted before the fit, has been tested. In principle, both methods do the same. The method described in this section contains a sideband subtraction through fitting the sidebands, whereas the subtraction is done before the fit in the alternative method. This method is described in detail in section A.2 in the Appendix. The choice for using the non-sideband subtracted maximum likelihood fit was made because of two reasons. First, there were technical problems with the error propagation due to the sideband subtraction which could not be solved. The second reason was that the expected statistical error of the chosen method⁶ is slightly smaller than the one expected from the sideband subtracted fit.

Nevertheless, there are a couple of aspects which can be concluded from the alternative method. One of them is, that there is no possibility of artificially creating a signal peak by the True- D_s background only, which was a major concern during the development of the technique. Another important conclusion is referred to the results of the fit validation of the default method which will be described below.

7.2.2 Fit validation

A validation of the fit, testing its workability and the reliability of the results has been done. As a first, simple consistency check the fitting procedure has been tested using generic MC with a defined Signal MC contribution, comparing the fit result with the added signal fraction. In addition, a validation based on Toy MC samples has been done to check the correctness of the errors of the fit and to determine possible fit biases, depending on the signal branching ratio.

⁶This method is referred to as 'default method' in this thesis.

Fit validation with generic Monte Carlo

The four generic MC background fractions ($u\bar{u}/d\bar{d}/s\bar{s}$ MC, $c\bar{c}$ MC, $b\bar{b}$ combin. bkg MC, $b\bar{b}_{\text{True-}D_s}$ MC) have all been scaled to data luminosity, while the signal fraction has been scaled using a branching ratio of $\mathcal{BR}_{\text{signal}} = 5 \times 10^{-4}$, which is close to the preliminary Run 1-4 result, as reported in [44]. Figures 7.12 and 7.13 show the fit to all slices and lepton channels, while table 7.3 summarizes the fit results.

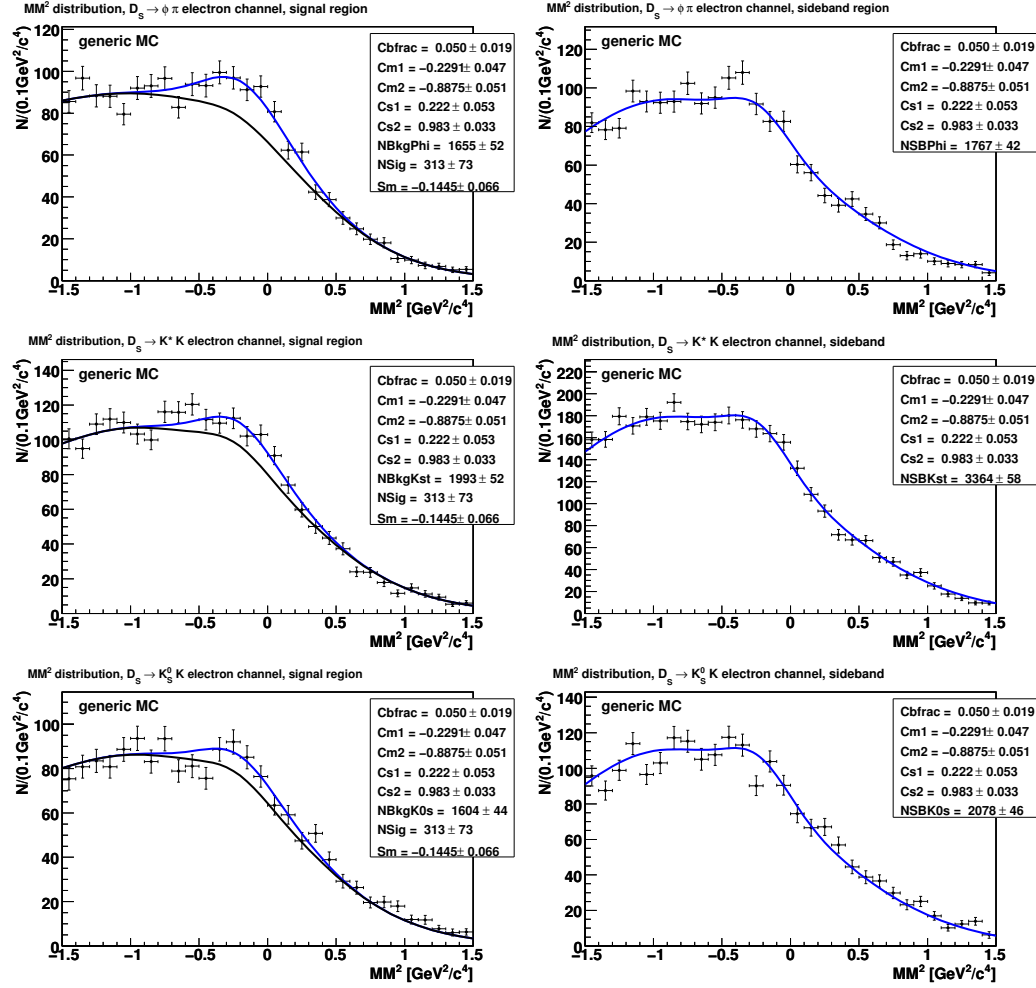


Figure 7.12: Generic MC MM^2 validation fits for the electron channel. Left hand side: signal region fit, right hand side: the sideband fit. Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_s^0 K$ channel.

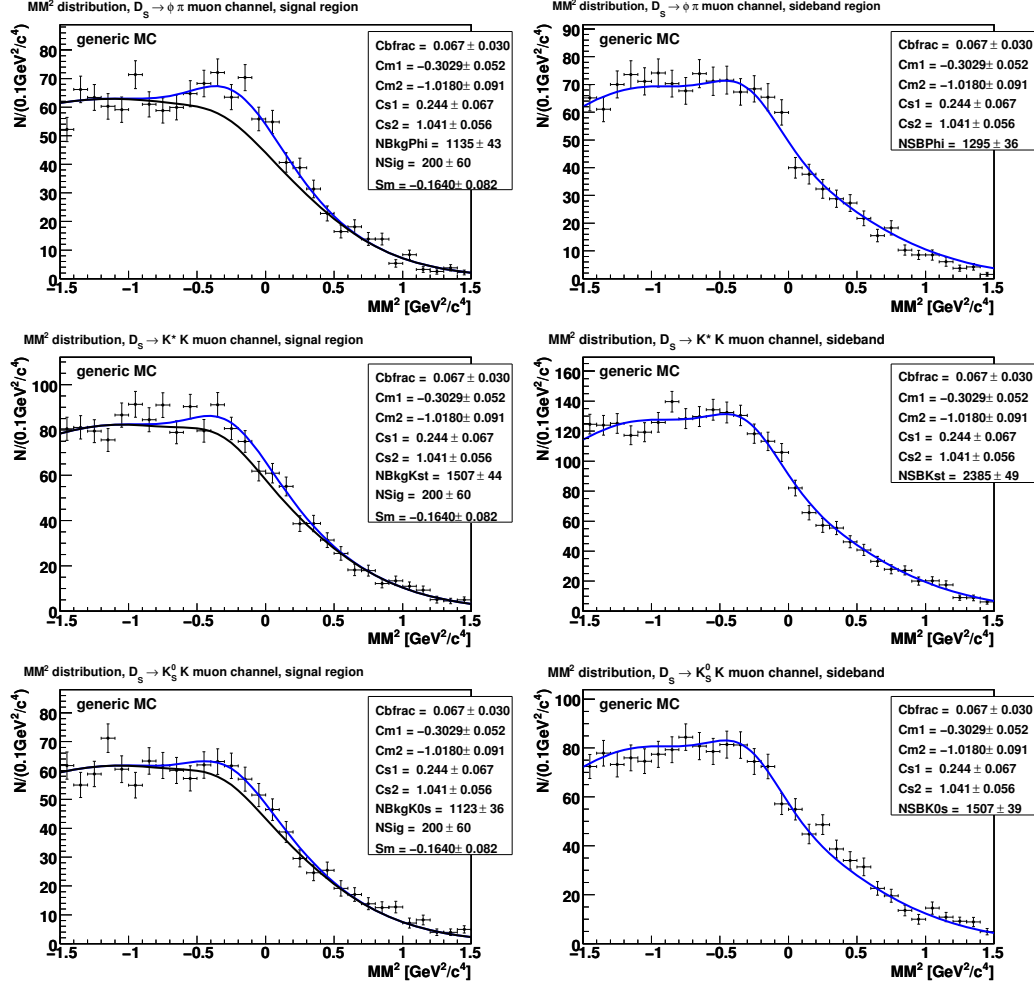


Figure 7.13: Generic MC MM^2 validation fits for the muon channel. Left hand side: signal region fit, right hand side: the sideband fit. Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^*K$ channel, bottom row: $D_s \rightarrow K_S^0K$ channel.

As can be seen in table 7.3, there is a discrepancy between the number of signal events added to the background sample and the number of signal events fitted. The difference decreases, when fixing the signal mean to the Signal MC value, which cannot be done in the final fit due to possible contributions of D_s^* 's in the final state (review subsection 7.2.1 for more details). Thus, a fit bias of the order of 1σ is expected. Extended studies have been made, showing that statistical fluctuations of the combinatorial background component in the K_S^0K channel lead to the shift and create the larger signal yield. Therefore, the results of the fit validation with Toy MC are more meaningful.

Table 7.3: Results of the simultaneous validation fits to full generic MC with a defined signal fraction for electron and muon channel. The numbers in parentheses correspond to the input signal fraction.

lepton channel	electron channel	muon channel
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow \phi\pi$	$[1.957 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.980]$	$[1.958 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.980]$
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow K^{*0}K$	$[1.956 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.982]$
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow K_S^0K$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.982]$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.983]$
Signal Mean	$(-0.144 \pm 0.066) \text{ GeV}^2/c^4$	$(-0.164 \pm 0.082) \text{ GeV}^2/c^4$
χ^2/NDF	223.2/167	183.4/167
$N(\text{total}) (N(\text{total})_{\text{input}})$	$313.8 \pm 73.3 (252)$	$199.8 \pm 60.2 (142)$
$N(\phi\pi) (N_{\text{input}})$	$137.2 \pm 32.1 (106)$	$86.3 \pm 26.0 (61)$
$N(K^{*0}K) (N_{\text{input}})$	$106.1 \pm 24.9 (82)$	$69.1 \pm 20.8 (49)$
$N(K_S^0K) (N_{\text{input}})$	$69.5 \pm 16.3 (64)$	$44.5 \pm 13.4 (32)$

Fit validation with Toy Monte Carlo

In general, generic MC samples are used to develop a fit method and to run first validation checks regarding the reproducibility of input signal fractions. Furthermore, they are used to understand the contributions of single background components and to parametrize them. But they represent only one special sample and cannot deliver information about the behavior or results of the fit, if real data differs from the MC, even if the differences appear only due to statistical fluctuations. To derive information about a possible fit bias and the accuracy of the fit results, in terms of stability and fit errors, more statistically independent samples, looking like real data are needed. Such samples are often referred to as *Toy Monte Carlo*.

The main difference between Toy MC and generic MC is that the latter is a simulation of complete physics processes, the generation of particles, their decays and/or their interaction with the detector material and finally the reconstruction of their tracks, whereas Toy MC is produced using the information of the reconstructed events, only. So, it basically does not care about the underlying physics processes. The two possibilities for producing Toy MC are the simulation of distributions of variables, based on their parametrization and the random selection of events from the full generic MC sample. The aim is to get samples having the same luminosity like data, while being statistically independent. Thus, the amount of generic Monte Carlo limits the number of producible Toy MC samples in the second method. The *BABAR* experiment provides generic MC samples of about three times the data luminosity. Hence, it is only possible to create three completely independent samples. Since one typically needs about 1000 samples, this type of Toy Monte Carlo can only be used for the generation of Signal Toy experiments, where larger generic Monte Carlo samples of several thousand times the data luminosity can be produced at a short timescale. The main advantage of this method is that it contains the information about possible correlations between the distributions (if samples containing more than one variable have to be produced, e.g. when fitting several variables at the same time). This information is not contained when creating the Toy MC samples using the parametrization of the distributions of variables. Using this latter method has the advantage, that no large generic MC samples are necessary and the samples can be produced at a short timescale.

This analysis needs Toy MC samples containing two variables the D_s mass and the MM^2 . Both variables are uncorrelated. Hence, the production of Toy MC samples from the parametrizations can be used. For each of the three components (combinatorial background, True- D_s background and signal), 1000 samples have been simulated. Signal Toy MC events have been produced for each integer branching ratio in the intervall $\mathcal{BR}_{\text{signal}} = [1, \dots, 9] \times 10^{-4}$. For each signal branching ratio, the samples are added and the $m(D_s)$ -fit and the MM^2 -fit ran on the sample. The results of the 1000 fits of the Toy MC samples are filled in two histograms, one showing the distribution of the number of fitted events and the other showing the pull distribution ($\frac{N_{\text{fitted}} - N_{\text{input}}}{\sigma_{\text{fitted}}}$), providing information about the correctness of the error and a fit bias. The pull distribution should show a gaussian shape with a mean at 0, if there is no bias, and a width of ~ 1 (due to the normalization in the definition of the distribution) if the fit errors are correct. Figures 7.14 and 7.15 show the distributions for three signal branching ratios ($\mathcal{BR}_{\text{signal}} = [1, 3, 6] \times 10^{-4}$) only.

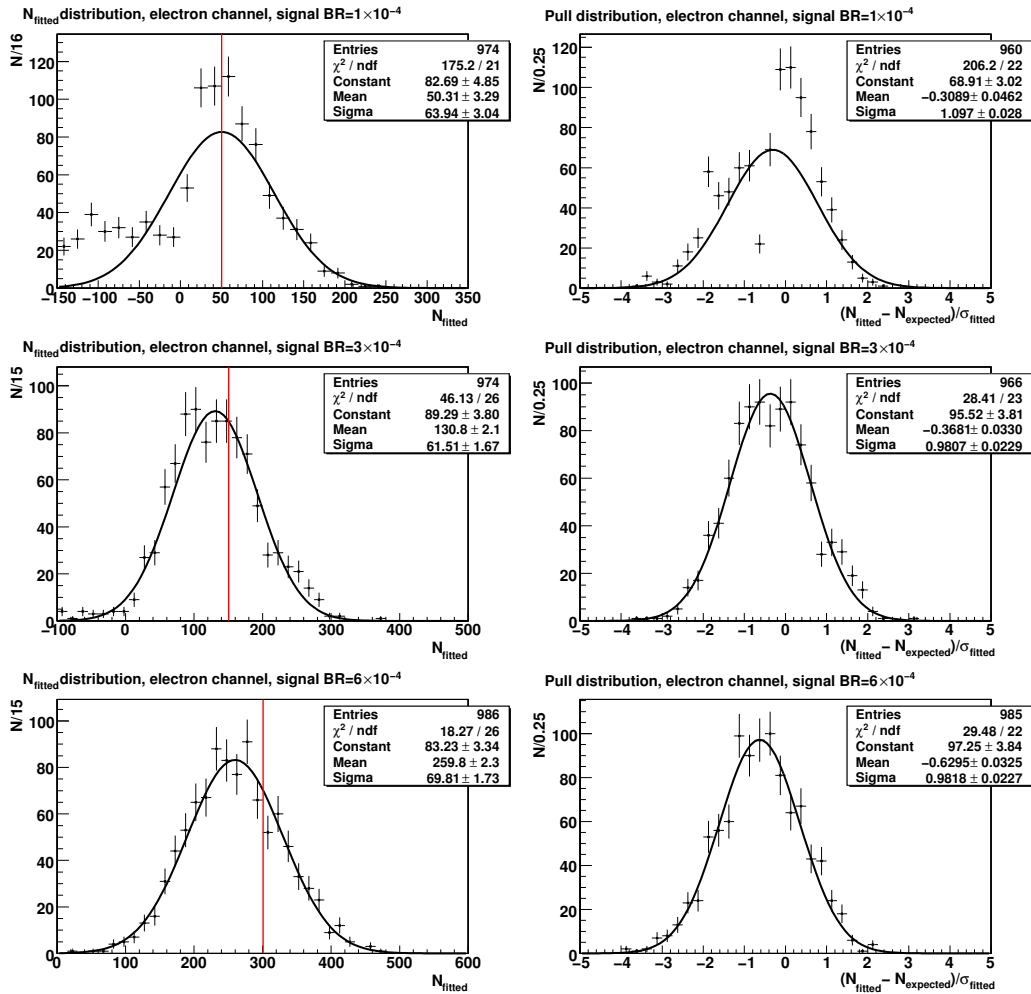


Figure 7.14: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = [1/3/6] \times 10^{-4}$ for the electron channel. Left: N_{fitted} distributions. Right: pull distributions. The red line indicates the input signal size.

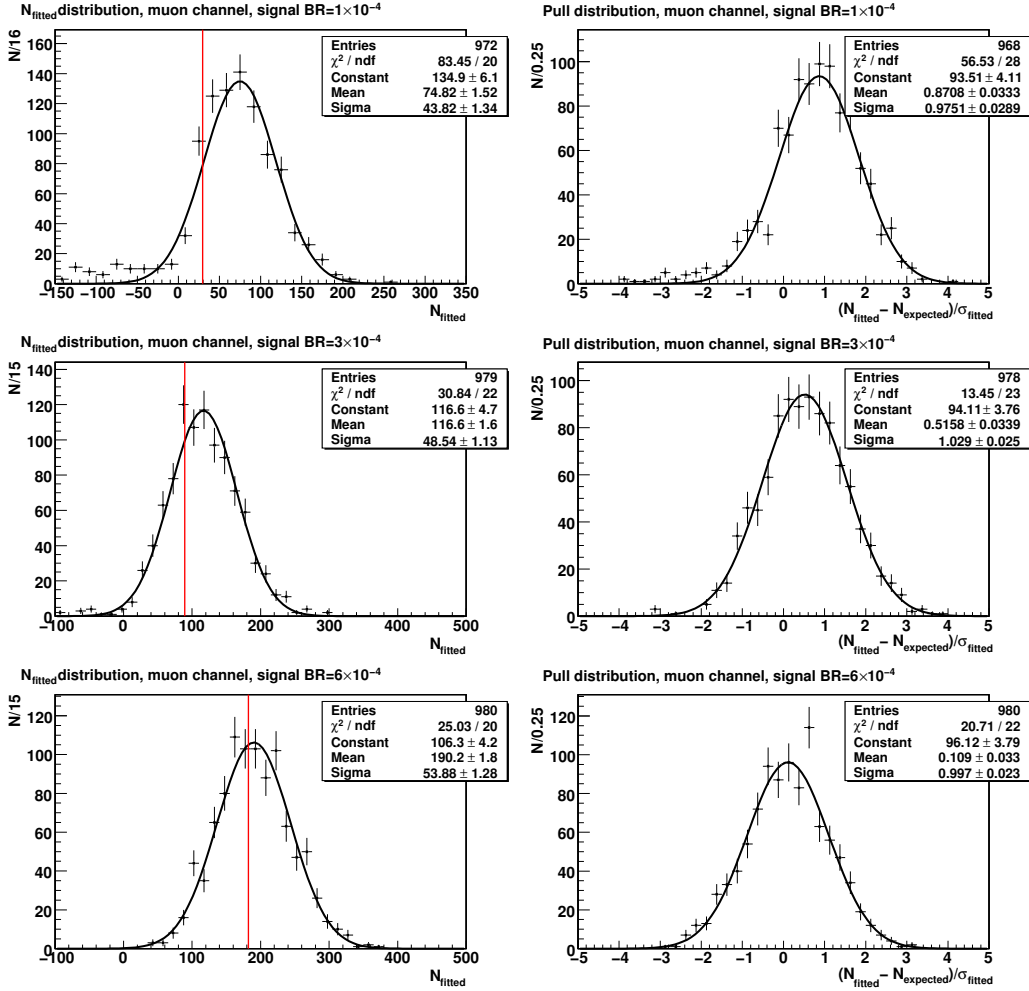


Figure 7.15: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = [1/3/6] \times 10^{-4}$ for the muon channel. Left: N_{fitted} distributions. Right: pull distributions. The red line indicates the input signal size.

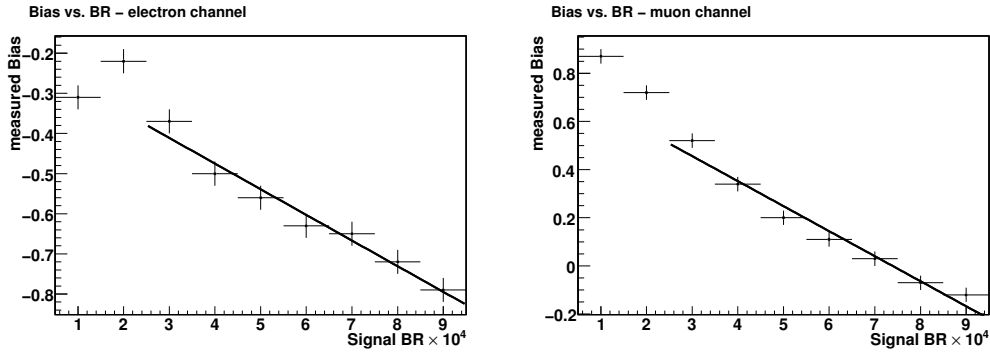
The full set of distributions is shown in figures A.4-A.9 in section A.3 of the appendix, while table 7.4 summarizes the results in terms of fitted mean and width of the N_{fitted} and the pull distributions, as well as the input signal fraction for each branching ratio.

As can be seen in the figures, the N_{fitted} distributions for small signal branching ratios ($\mathcal{BR} = 1 \times 10^{-4}$) do not have a gaussian shape, while the pull distributions exhibit the expected shape considering a variation of the chosen binning in the electron channel. This behavior is hinting to the fit being not reliable for the extraction of small signal yields. With increasing numbers of signal events the N_{fitted} distribution looks more gaussian-like, which leads to the conclusion that the fit is reliable in both lepton channels for signal branching ratios $\mathcal{BR}_{\text{signal}} \geq 3 \times 10^{-4}$.

Table 7.4: Results of the Toy MC fits for electron and muon channel.

Signal \mathcal{BR} $\times 10^4$	electron channel			muon channel		
	N_{fitted}	N_{input}	Pull	N_{fitted}	N_{input}	Pull
1	55	49	-0.309	69	30	0.871
2	94	99	-0.222	86	60	0.720
3	131	149	-0.368	112	90	0.516
4	172	200	-0.500	136	121	0.339
5	214	249	-0.560	163	151	0.203
6	259	301	-0.630	191	180	0.109
7	307	352	-0.646	218	210	0.030
8	350	401	-0.724	247	240	-0.068
9	395	452	-0.793	274	273	-0.123

The pull distributions contain two important information. First, the width of the distributions is approximately 1 in all branching ratios. Thus, one can trust the fit result in terms of the errors of the signal yields in the whole signal range tested. The second interesting point is the appearance of a bias, which depends on the size of the signal branching ratio. The pull mean value decreases linearly with larger signal branching ratios, as can be seen in figure 7.16.

**Figure 7.16:** Bias vs. Signal BR for electron (left) and muon channel (right).

The main source of the bias is the simultaneous parametrization of the True- D_s background fraction, as well as the constraint on the number of signal events in the single D_s reconstruction chains. Nevertheless, it has been decided, that the advantages in terms of stability and reliability of the fit with the used setup exceed the disadvantage of the bias, which has to be corrected after the unblinding. The exact size of the bias is determined after unblinding by using Toy MC samples again.

Furthermore has to be noted, that there is a disagreement between the Toy MC bias and expected bias concluded from the validation with generic MC (see table 7.3). Considering the results of the Toy MC studies of the alternative fit method⁷, it can be deduced, that the difference originates only from including the combinatorial background in the fit.

7.2.3 Significance expectation from Toy Monte Carlo results

Using the information of the expected signal yield in dependence of the fit error for the different signal branching ratios used, an estimation on the expected statistical significance of the measurement is done.

As a lowest order approximation the mean of the number of fitted events, corrected for the bias, is divided by the width of the N_{fitted} distribution, representing the fit error. Figure 7.17 shows the expected significance for the single lepton channels and for a combination of both lepton channels. The aim of observing a 5σ effect, which can be interpreted as the first measurement of the decay, seems to be possible if $\mathcal{BR}_{\text{signal}} \geq 5.5 \times 10^{-4}$ and both lepton channels are combined. The significance of measurements is discussed more in detail in the following section.

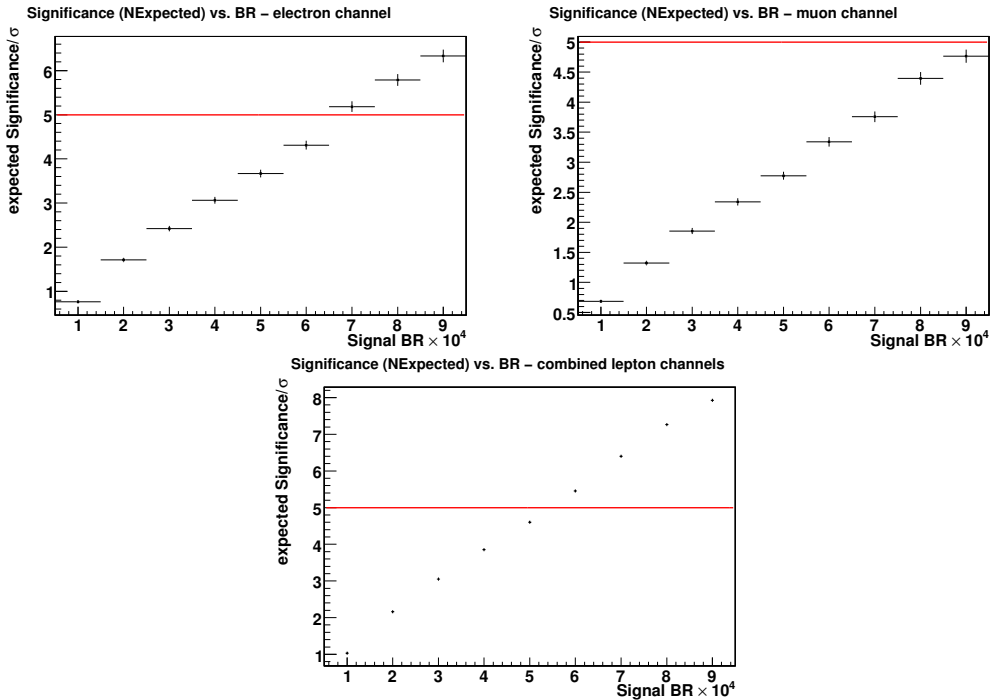


Figure 7.17: Expected significance of the measurement vs. Signal \mathcal{BR} for electron channel only (top left), muon channel only (top right) and the combination of both channels (bottom). The red line indicates a five sigma measurement.

⁷As discussed in section A.2, the observed bias of the alternative fit method determined by Toy MC studies is compatible with the expectation from the generic MC fit.

7.3 Fitting the Data

After the fit technique has been validated successfully, it was deployed to data. Figures 7.18 and 7.19 show the fitted MM^2 distributions of signal region and sideband for both lepton channels.

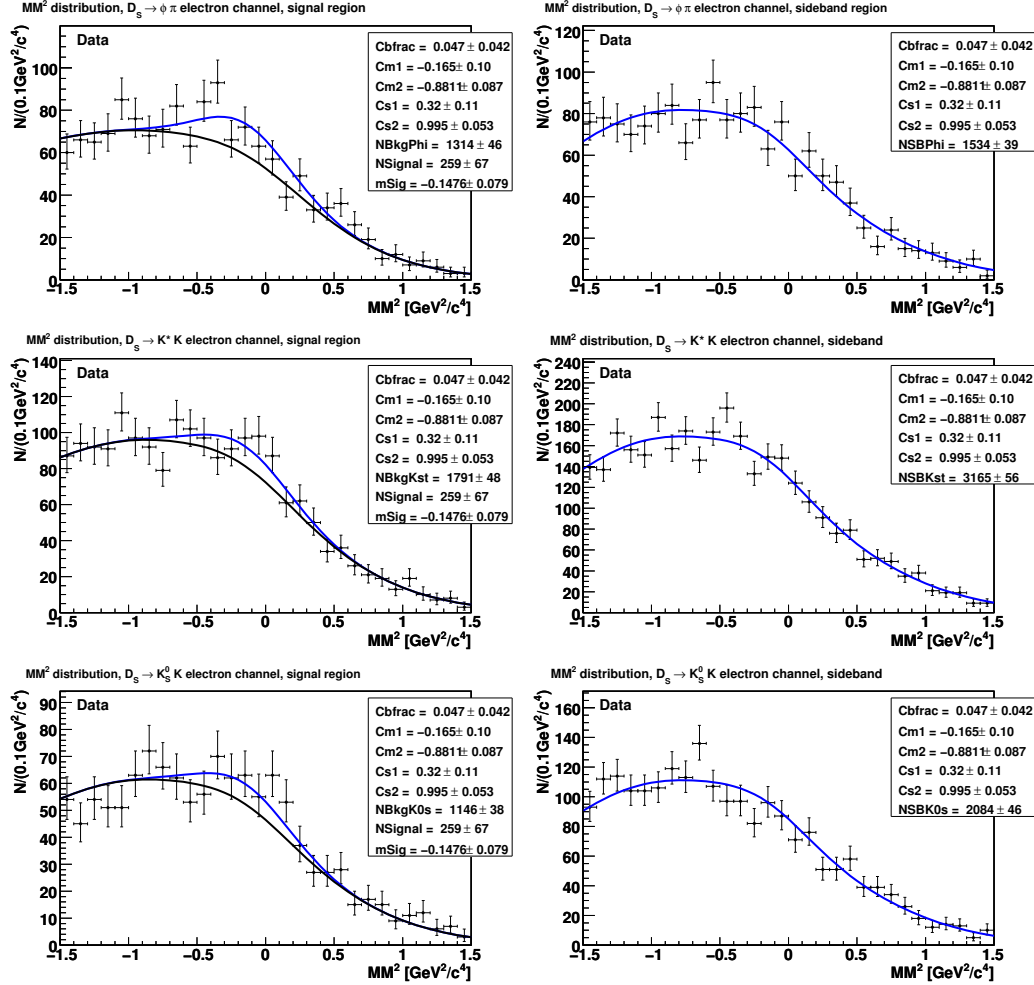


Figure 7.18: Final fits on the electron channel data sample. Left hand side: signal region fit, right hand side: sideband fit. Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_s^0 K$ channel. Black curves indicate the background contribution, while blue curves represent the full fit.

Table 7.5 gives an overview of the fit results. It also provides the statistical significance \mathcal{S} of measuring a signal larger than zero for this measurement. The significance is expressed in numbers of standard deviations (σ), as it represents the difference of the measurement from a distinct value (e.g. $N(\text{Signal}) = 0$) assuming a gaussian distribution of the measured quantity. The number of σ 's can be translated into the probability for obtaining differences between the measured quantity N_{measured} and it's true value N_{true} , that are larger or equal than $|N_{\text{measured}} - N_{\text{true}}|$.

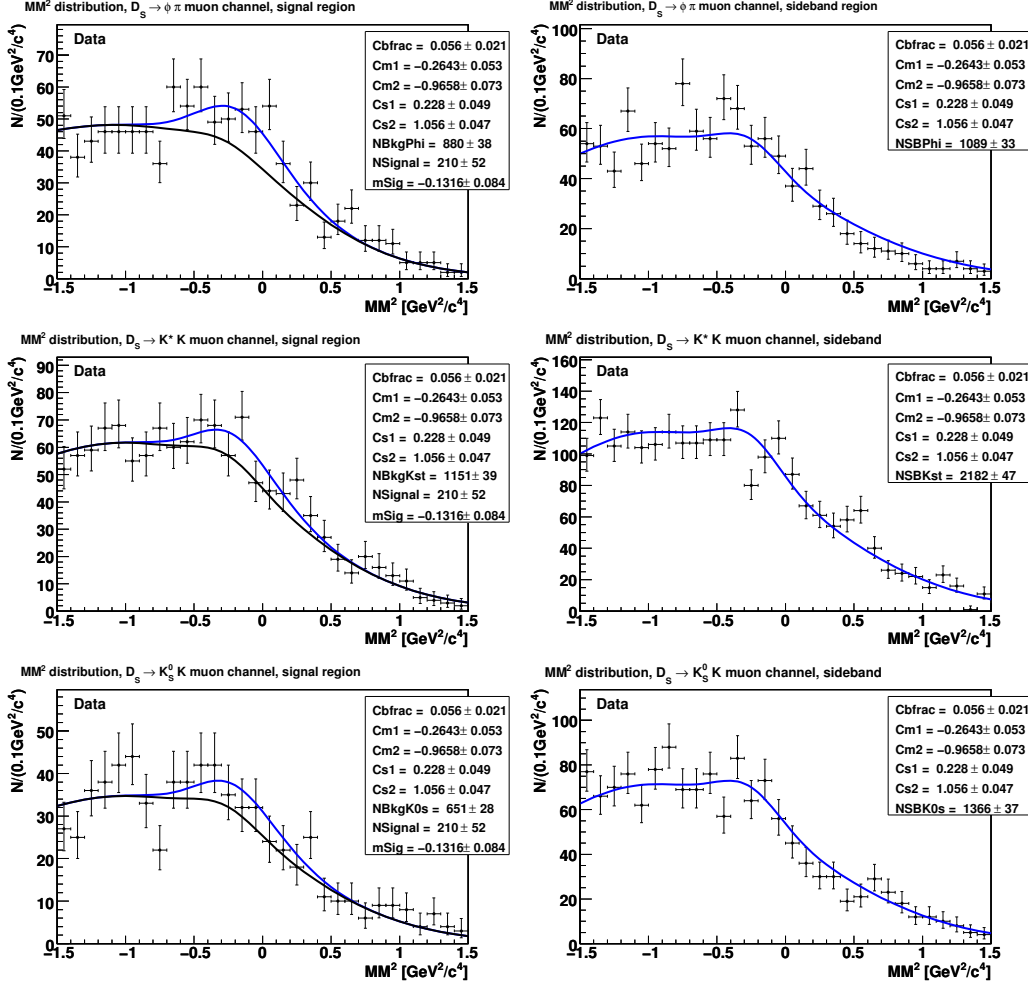


Figure 7.19: Final fits on the muon channel data sample. Left hand side: signal region fit, right hand side: sideband fit. Top row: $D_s \rightarrow \phi \pi$ channel, middle row: $D_s \rightarrow K^* K$ channel, bottom row: $D_s \rightarrow K_s^0 K$ channel. Black curves indicate the background contribution, while blue curves represent the full fit.

Using the hypothesis $N_{\text{true}} = 0$ one obtains the probability that the measurement is only a fluctuation of the background and no “real” signal present (also called error probability $\mathcal{P}_{\text{error}}$). In high energy physics, significances larger than 3σ ($\hat{=} \mathcal{P}_{\text{error}} = 2.7 \times 10^{-3}$) are treated as evidences and significances larger than 5σ ($\hat{=} \mathcal{P}_{\text{error}} = 5.7 \times 10^{-7}$) are interpreted as observations.

The significance of the measurements of this analysis have been calculated using the minimized logarithmic likelihood values of the fit result $-\ln L_{\text{fit}}$ and of a fit imposing a ‘no signal’ hypothesis L_0 . The statistical significance is then given by:

$$S = \sqrt{2 \cdot |-\ln L_0 + \ln L_{\text{fit}}|}. \quad (7.8)$$

Unfortunately, this equation for estimating the significance of measurements is only valid if the fits used to determine the likelihood values differ only by one degree of freedom.

Table 7.5: Results of the simultaneous fits to data for electron and muon channel.

lepton channel	electron channel	muon channel
$m(D_s)_{SR}, D_s \rightarrow \phi\pi$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.979]$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.979]$
$m(D_s)_{SR}, D_s \rightarrow K^{*0}K$	$[1.953 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$	$[1.954 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$
$m(D_s)_{SR}, D_s \rightarrow K_S^0K$	$[1.953 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$	$[1.956 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.979]$
Signal Mean	$(-0.148 \pm 0.079) \text{ GeV}^2/c^4$	$(-0.132 \pm 0.085) \text{ GeV}^2/c^4$
χ^2/NDF	160.9/167	201.9/167
$N(\text{total})$	$259.4^{+67.6}_{-67.2}$	$209.7^{+53.0}_{-52.2}$
$N(\phi\pi)$	$115.7^{+30.2}_{-30.0}$	$92.1^{+23.3}_{-22.9}$
$N(K^{*0}K)$	$85.2^{+22.2}_{-22.1}$	$70.2^{+17.8}_{-17.5}$
$N(K_S^0K)$	$58.5^{+15.3}_{-15.2}$	$47.4^{+12.0}_{-11.8}$
$-\log L_{\text{fit}}$	-44015.1	-26359.4
$-\log L_0$	-44007.3	-26350.6
\mathcal{S}	$(3.6 - 4.0) \sigma$	$(3.8 - 4.2) \sigma$

This means, that all parameters of the PDF describing the signal need to be fixed to MC values or physical parameters. For the mean of the signal this is not possible, as discussed in previous sections. The naive expectation, the difference of the numbers of degrees of freedom being two is also incorrect, as discussed in [45, 46]. The correct approach for evaluating the significance would be:

1. the production of Toy MC samples using the background PDF only,
2. the fit of each of the Toy MC samples with the full fit,
3. counting the number of fit results which report a signal larger than the one observed,
4. translating this probability into an equivalent number of σ 's.

As a matter of fact, such Toy MC studies become very time and computing power consuming, since the error probability of a 4σ effect is already 6.3×10^{-5} . This means, one needs at least 10^7 Toy MC samples for each lepton channel for an accurate calculation. This translates into an expected computing time of 2×500000 hours which is simply not deliverable by the current computing resources.

Thus, table 7.5 can only provide a range of the statistical significance, where the smaller value represents the statistical significance for a difference of the degrees of freedom of two and the larger value gives the statistical significance for $\Delta(NDF) = 1$.

Bias correction

A comparison of the number of fitted events in table 7.5 with the values from the Toy Monte Carlo studies (table 7.4) shows that the fit result for the electron channel is exactly the expected number of signal events, if the signal branching ratio is 6×10^{-4} . At this branching fraction, the fit has a bias of -42 events, leading to a corrected signal yield of $N(\text{total})_{\text{measured}} - N_{\text{Bias}} = N(\text{total})_{\text{corrected}} = 301$. For the muon channel, the data result lies between 6×10^{-4} and 7×10^{-4} .

Thus, the creation and fit of additional Toy MC samples are necessary to determine the exact value of the bias. Samples of signal events, equivalent to signal branching ratios of 6.5×10^{-4} and 6.8×10^{-4} have been produced and analyzed. Figure 7.20 shows the number of fitted events distribution.

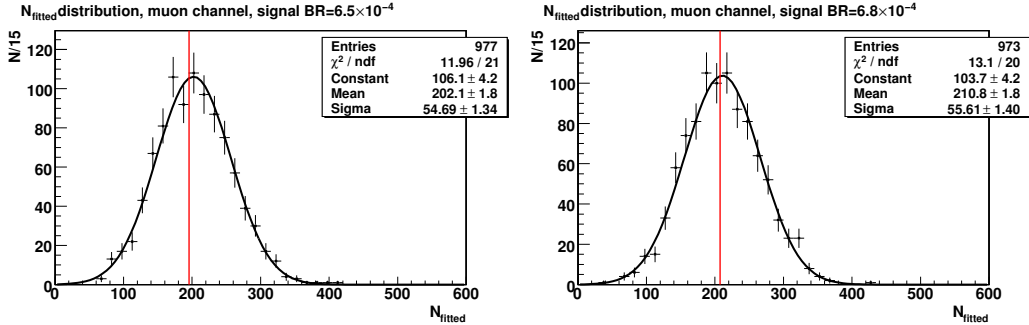


Figure 7.20: N_{fitted} distributions for Toy MC for signal branching ratios of $\mathcal{BR}_{\text{signal}} = 6.5 \times 10^{-4}$ (left) and $\mathcal{BR}_{\text{signal}} = 6.8 \times 10^{-4}$ (right) for the muon channel.

As can be seen, the sample using a signal branching ratio of 6.8×10^{-4} (being equivalent with 206 added signal events) shows a good agreement with the data result, leading to a bias of +4 events. Thus, the corrected signal event yields are 301 events for the electron channel and 206 events for the muon channel.

The combination of both measurements for the determination of the total significance, considering also systematic uncertainties of the measurement, and the calculation of the resulting branching fraction for the signal decay are given in chapter 9. The determination of the systematic uncertainties will be described in the next chapter.

Chapter 8

Systematic uncertainties

The various systematic uncertainties considered can be divided into two main classes: First, the uncertainties arising from the fitting method and second, the uncertainties due to the Monte Carlo simulation. Whereas the former are related to the number of fitted events and therefore need to be considered when calculating the total significance of the measurement, the latter only affect the reconstruction efficiency. The uncertainties arising from the fitting method are given as total numbers of events, while the uncertainties, affecting the signal reconstruction efficiencies of the reconstructed D_s decay chains are given as relative changes.

8.1 Systematic uncertainties due to the fit method

8.1.1 Parametrization of the True- D_s fraction

The uncertainty of the parametrization of the MM^2 distribution of the True- D_s events has been determined by using Toy Monte Carlo. 500 Samples of the two variables, describing the shape of the True- D_s MM^2 distribution, M0 and ExpoConst (see section 7.2.1), have been produced. Both variables are taken distributed like a gaussian with the mean being the fixed value of the fit to the True- D_s background and the width being the fit-error on the mean, as reported in table 7.1. The correlations between the two variables have been taken into account. The variable M0 is produced from a gaussian PDF using the known mean and sigma values. Then, the ExpoConst variable is drawn also from a gaussian PDF using the conditional mean and sigma values, which are defined as:

$$\begin{aligned} m_{\text{ExpoConst, Conditional}} &= m_{\text{ExpoConst}} + \rho \sigma_{\text{ExpoConst}} \left(\frac{v_{\text{M0}} - m_{\text{M0}}}{\sigma_{\text{M0}}} \right) \\ \sigma_{\text{ExpoConst, Conditional}} &= \sqrt{\sigma_{\text{ExpoConst}}^2 \cdot (1 - \rho^2)}. \end{aligned}$$

In this equations, m_i and σ_i are the mean and the error of variable i , ρ is the correlation coefficient between both variables known from the fit and v_{M0} is the value of the M0 variable, drawn from the PDF in the first step. The resulting distributions of the produced samples for electron and muon channel are shown in figure 8.1. They show the expected behavior of an ellipse in the M0-ExpoConst plane. The signal extraction fit has been redone using each of these sets of parameters for the parametrization of the fixed True- D_s fraction.

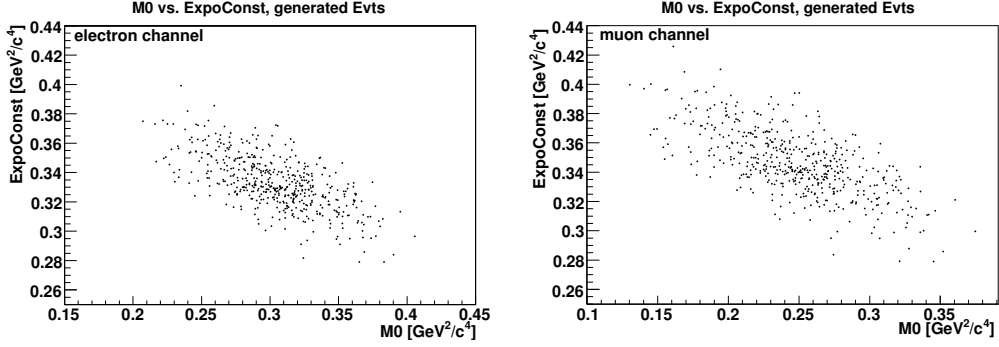


Figure 8.1: 2D distributions of Toy MC for True- D_s background PDF parameters for electron channel (left) and muon channel (right).

The width of the distribution of the number of fitted events is taken as systematic uncertainty. Figure 8.2 shows the resulting distributions with the fit, giving the width. This results in a systematic uncertainty of 19.9 events for the electron channel and 15.9 events for the muon channel.

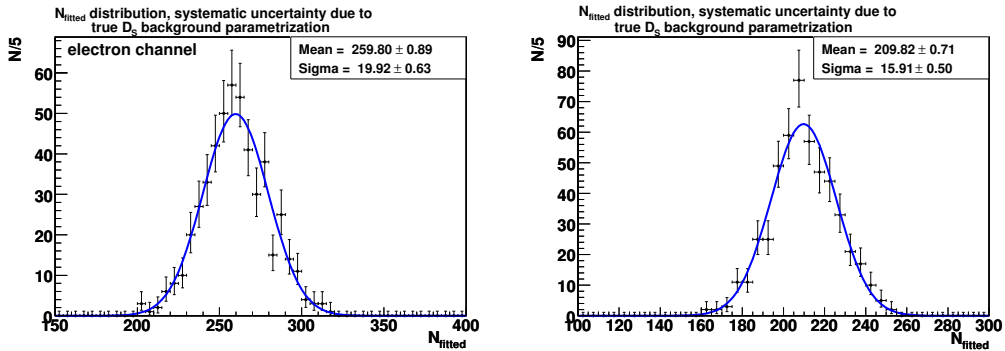


Figure 8.2: Results of Toy MC fits, determining the systematics arising from the fixed True- D_s background PDF parameters for electron channel (left) and muon channel (right).

8.1.2 Fit error of the $m(D_s)$ fits

The errors on the fitted numbers of True- D_s , extracted by the fits to the $m(D_s)$ distributions have been used to evaluate the systematic uncertainty for the fixed size of the True- D_s fraction. The fits have been redone, using $N(D_s)_{\text{fitted}} \pm \sigma(N(D_s)_{\text{fitted}})$ as input, leading to uncertainties of 3.6 events for the electron channel and 3.4 events for the muon channel.

8.1.3 Width of the signal gaussian

Because the width of the signal PDF is not a free parameter and has been fixed in the final fit, the determination of the values used introduces an additional systematic uncertainty to the measurement. Since the values have been determined through fits to Signal Monte Carlo, the fit errors have been used to evaluate the uncertainties. It has to be noted, that they depend on the number of Signal Monte Carlo events, used in this analysis.

If there is more signal MC available, the errors of the fit get smaller. Thus, the values used for the errors on the width, taken from the fits shown in figure 7.11, must be interpreted as conservative estimates only. The MM^2 fits to the full data sample have been redone fixing the signal width to values \pm the Signal MC fit errors. The maximum deviation of the data fit result is taken as the systematic uncertainty of the width of the signal PDF. This leads to uncertainties of 3.9 events for the electron channel and 4.3 events for the muon channel.

8.1.4 Constraints on the single channels signal yields

The constraints on the single channels signal yields have been determined by using the reconstruction efficiency and the D_s branching ratios in the according decay chains. For the latter, the PDG 2008 values are used [8]. A systematic uncertainty on the constraints arises due to the fact that the values used for the constraints may be different from the ones, appearing in data. This results in changes in the ratios of the single signal yields.

To account for this, 500 Toy MC samples of the D_s branching ratios have been produced, using the PDG central value for the mean and the PDG error as width. Each of the three D_s decay chains has been simulated independently. Then, the fractions of $N(\phi\pi)$, $N(K^*0K)$ and $N(K_S^0K)$ in the total number of Signal Monte Carlo events have been rescaled according to the produced Toy MC samples of the branching ratios and the fit has been redone. The widths of the distributions of the number of fitted signal events are taken as systematic uncertainty. Figure 8.3 shows the distribution of the number of the fitted events.

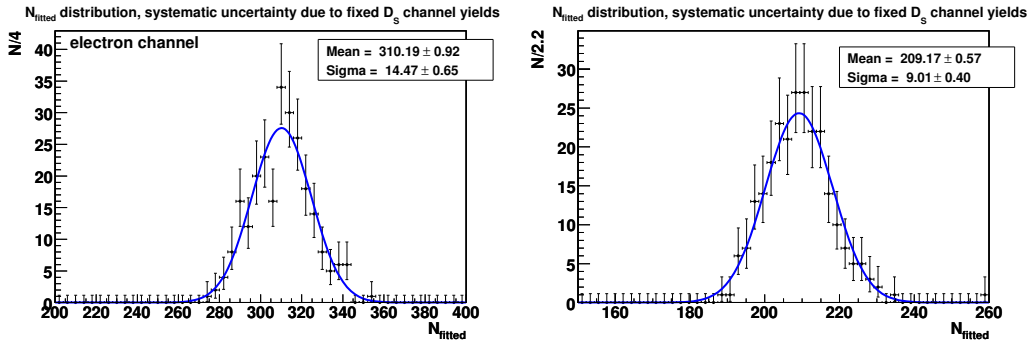


Figure 8.3: Results of Toy MC fits, determining the systematics arising from the constraints on the single channels signal yields for electron channel (left) and muon channel (right).

This leads to systematic uncertainties of 14.5 events for the electron channel and 9.0 events for the muon channel.

8.1.5 Bias correction

The bias correction introduces an additional systematic uncertainty due to the error on the correction factor. As can be seen in figures 7.14 and 7.20, the systematic uncertainty of the bias correction is 2.2 events for the electron channel and 1.8 events for the muon channel. Although this uncertainty is related to the total number of signal events, it doesn't affect the significance of the measurement, because the existence of a fit bias does only changes the total number of fitted events, but not its significance in terms of differences in the logarithmic likelihoods.

8.1.6 Total systematic uncertainties arising from the fitting method

All additive systematic uncertainties evaluated are summarized in table 8.1 and added in quadrature to get the total systematic uncertainty, which is used for calculating the final significance of the signal later on.

Table 8.1: Overview of the additive systematic uncertainties evaluated in this analysis.

source	syst. uncertainty electron channel	syst. uncertainty muon channel
True- D_s fraction parametrization	19.9	15.9
Fit error of the $m(D_s)$ fits	3.6	3.4
Width of the signal PDF	3.9	4.3
Single channels signal yields	14.5	9.0
Total uncer., affecting significance	25.2	19.1
Bias correction	2.2	1.8
Total uncertainty	25.3	19.2

8.2 Systematic uncertainties due to MC simulation and reconstruction

Most of the systematic uncertainties described in this section are relative to the reconstruction efficiency of the particular D_s reconstruction channels. The uncertainty arising from the limited knowledge of the D_s branching fractions into the reconstructed final states is treated separately. Besides that, the error on the number of $B\bar{B}$ events is a global systematic uncertainty, which is equal for all D_s reconstruction and lepton channels.

8.2.1 Signal Monte Carlo model

Since there is no exact decay model for the signal decay, the influence of alternative models on the reconstruction efficiency has been tested. Different models affect the angular distributions of the decay products as well as their momentum distributions. The determination of the Signal Monte Carlo model uncertainty has been done by simulating the signal decay with two alternative decay models. First, a simple four-body decay has been simulated and second, a manipulated $B \rightarrow D^{**}\ell\nu$ decay model (based on the ISGW2 decay model) has been used. The D^{**} is forced to decay into a $D_s K$ final state.¹ The samples have the same size as the Signal MC sample used in the analysis.

The change of the lepton momentum spectra in the different decay models is the most important effect, altering the reconstruction efficiencies (see section 4.2.3). As described in section 5.4, the cut on these momenta is optimized in order to maximize the significance, but the loose cut on the momentum also has the advantage of inserting only small systematic uncertainties into the measurement.

¹The mass has been constrained to be larger than the sum of the masses of the daughters.

Table 8.2 shows the relative changes in the numbers of selected signal events for the two alternative decay models. Differences in the uncertainties between the lepton channels are a relic of the lepton identification by different selectors and algorithms, together with a shift in the momentum distribution due to the higher muon mass.

Table 8.2: Relative change of reconstructed events for the different Signal MC models for electron and muon channel.

D_s decay chain	electron channel		muon channel	
	$\frac{\Delta N_{\text{ISGW2}}}{N_{\text{default}}}$	$\frac{\Delta N_{\text{phasespace}}}{N_{\text{default}}}$	$\frac{\Delta N_{\text{ISGW2}}}{N_{\text{default}}}$	$\frac{\Delta N_{\text{phasespace}}}{N_{\text{default}}}$
$D_s \rightarrow \phi\pi$	7.6%	17.1%	0.2%	3.9%
$D_s \rightarrow K^{*0}K$	3.1%	10.6%	-6.4%	4.9%
$D_s \rightarrow K_S^0 K$	5.9%	15.0%	-2.1%	8.5%

While the ISGW2 based decay model shows only a small deviation to the Goity-Roberts model, the phase space model differs much more. This is expected, since this decay model is not capable of describing weak interactions and just creates the lepton and the neutrino as daughters directly coming from the B meson. This model only serves as a very conservative approach for the estimation of the uncertainty. For both alternative decay models, the electron channel shows a larger difference than the muon channel. The source of this discrepancy is the efficiency of the particle selectors used in this analysis vs. the lepton momentum, which is shown in figure 8.4 [47, 48].

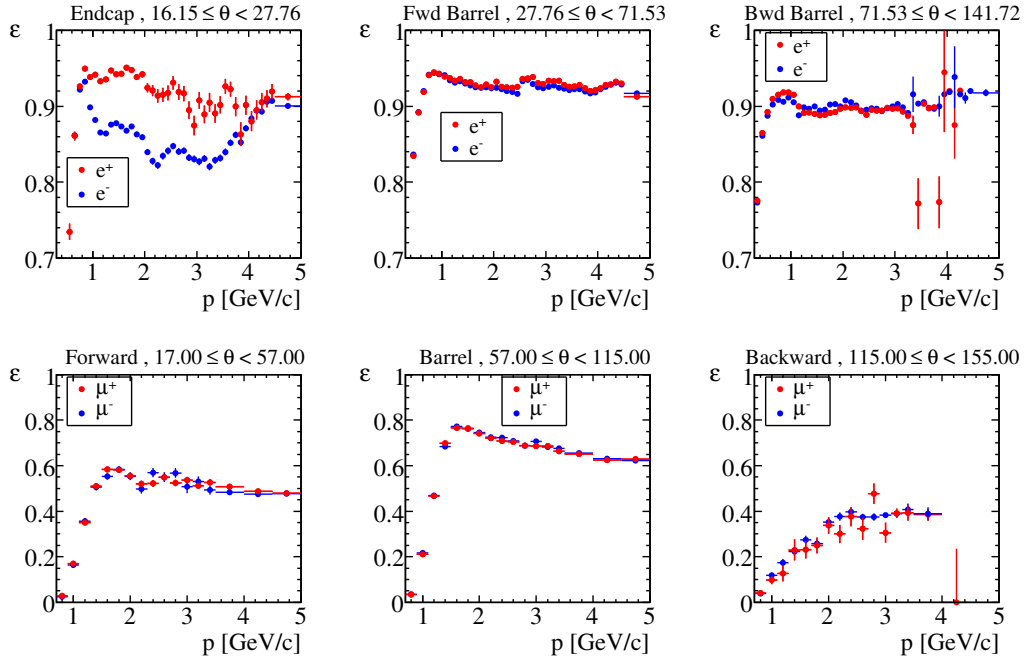


Figure 8.4: Efficiency vs. momentum for different regions of the polar angle Θ for the PidLHElectron selector (top) and the muNNVeryTight selector (bottom).

The lepton momentum distributions for the three decay models are shown in figure 8.5. As can be seen, the main differences between the models appear at lepton momenta below $p_\ell < 1 \text{ GeV}/c$. While the reconstruction efficiency of electrons reaches about 90% in this regions, muons are almost undetectable, as already mentioned in section 3.5. Thus, muons are less sensitive to the model dependent differences of the lepton momenta spectra. Since the ISGW2 model is the better physical motivated decay model for the signal decay, the difference to this data sample is used as systematic uncertainty.

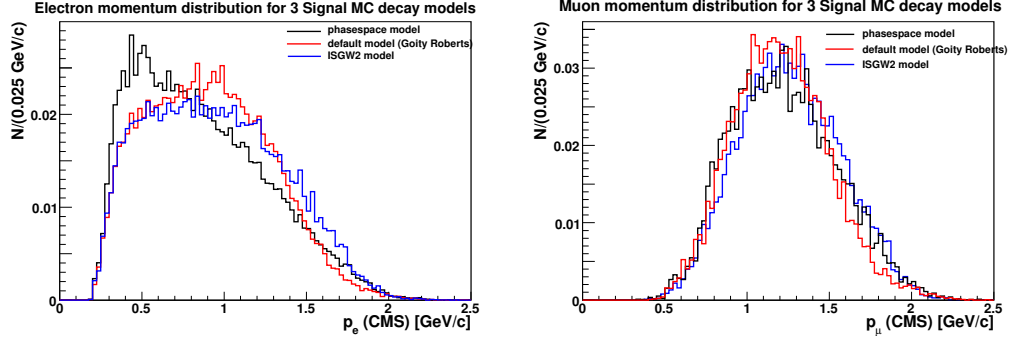


Figure 8.5: Comparison of the lepton momentum distributions between different Signal MC models for the electron channel (left) and muon channel (right).

8.2.2 Signal Monte Carlo statistics

The statistical error of the limited amount of Signal Monte Carlo events leads to a systematic uncertainty of the reconstruction efficiency $\Delta\epsilon^{\text{reco}}$, which has been calculated using the expression

$$\Delta\epsilon_i^{\text{reco}} = \sqrt{\frac{\epsilon_i^{\text{reco}}(1 - \epsilon_i^{\text{reco}})}{N_{\text{produced}}}}.$$

In this equation, ϵ_i^{reco} is the reconstruction efficiency (with i differentiating between the D_s reconstruction chains) and N_{produced} is the number of produced Signal Monte Carlo events. It is derived from the expectation value and the variance of a binomial distribution [46]. The resulting uncertainties are shown in table 8.3.

Table 8.3: Systematic uncertainties arising from the limited amount of Signal MC.

D_s decay chain	syst. uncertainty electron channel	syst. uncertainty muon channel
$D_s \rightarrow \phi\pi$	2.7%	3.5%
$D_s \rightarrow K^{*0}K$	3.3%	4.2%
$D_s \rightarrow K_S^0 K$	2.5%	3.3%

Besides the direct influence by increasing or decreasing the total number of signal events, a different behavior of the single D_s reconstruction chains can lead to an additional uncertainty due to the usage of the reconstruction efficiency, when constraining the signal yields to each other. This effect has been studied by changing the efficiencies of the single reconstruction chains by $\pm 1\sigma$ in all possible variations. It turned out, that the effects are very small, compared to other sources of systematic uncertainties. The resulting numbers are of the order of $\sim 0.5\%$ and will be neglected. Such variations can appear in the estimation of other systematic uncertainties as well, e.g for the systematic uncertainty arising from the particle identification, but will be neglected due to their smallness in the other cases studied, too.

8.2.3 Particle identification

For the determination of the reconstruction efficiency and for a better agreement between data and MC, PidTweaking has been used for the processing of all MC samples. In the past, the particle identification group [34] performed various studies to estimate the systematic uncertainty of the identification by using an alternative set of Control Samples for the derivation of the selection efficiency and the misidentification rate (see section 3.5 for the description of PidTweaking). The results of using the alternative samples were compared to the standard method to establish an uncertainty. This study lead to the following results [49]:

- electron and positron selection agrees by $\sim 1\%$.
- muon alternative tables results agree with the standard tables results by $\sim 5\%$.
- pions and kaons agree very well.

Unfortunately, the usage of the alternative samples results is not approved for other than the particle identification group's analyses yet. Thus, only a conservative estimate of the systematic uncertainties can be done by taking the differences of reconstructed signal events between using and not using PidTweaking. The results in terms of relative errors on the reconstruction efficiencies are given in table 8.4. The resulting uncertainties are in the

Table 8.4: Systematic uncertainties arising from the particle identification.

D_s decay chain	syst. uncertainty electron channel	syst. uncertainty muon channel
$D_s \rightarrow \phi\pi$	0.6%	1.6%
$D_s \rightarrow K^{*0}K$	1.2%	4.9%
$D_s \rightarrow K_S^0K$	3.6%	7.9%

same order of magnitude as the values reported by the particle identification group. The differences between the single channels result from different momenta of the selected particles, as well as from differences in the final states. Again, the differences between the D_s reconstruction chains are neglected due to their smallness (evaluated uncertainty $\sim 0.1\%$).

8.2.4 Track and Photon reconstruction

Uncertainties in the charged track reconstruction efficiencies, fake tracks and tracks from beam background can change the reconstruction of the B candidate and the simulation of event shape variables. Uncertainties in the photon reconstruction only contribute to the latter and should be small.

The uncertainty due to tracking efficiency is evaluated by randomly eliminating tracks from the Charged Tracks list, which is used either directly (in the K_S^0 reconstruction) or indirectly (as input list for several selectors and for the GoodTracksVeryLoose list) in the analysis. The killing probabilities, shown in table 8.5 are taken from the Tracking Efficiency Task Force web page [41] and are valid for the computing environment, used for this analysis.

Table 8.5: Killing probability for charged tracks in different $BABAR$ runs.

Run	systematic uncertainty per track
Run 1	0.749%
Run 2	0.225%
Run 3	0.512%
Run 4	0.707%
Run 5	0.514%

The uncertainty due to photon efficiency is evaluated using the same killing algorithm. The probability of each track for being eliminated is taken from the established π^0 uncertainty [50]. Thus, 1.5% of all clusters have been killed. Table 8.6 shows the resulting systematic uncertainties.

Table 8.6: Systematic uncertainties arising from the charged track (charged) and cluster reconstruction (neutral).

D_s decay chain	syst. uncertainty electron channel		syst. uncertainty muon channel	
	charged	neutral	charged	neutral
$D_s \rightarrow \phi\pi$	0.4%	0.6%	0.1%	0.9%
$D_s \rightarrow K^{*0}K$	0.5%	0.4%	0.2%	0.9%
$D_s \rightarrow K_S^0K$	1.8%	0.5%	2.4%	0.7%

As can be seen, the uncertainties are very small, exceeding 1% only for the charged tracks of the K_S^0K decay chain. The difference between this channel compared to the other two arises from different requirements to the track quality (see section 5.2).

8.2.5 K_S^0 corrections

The efficiency correction applied for the reconstruction efficiency of the K_S^0 also introduces a systematic uncertainty. It has been calculated to be 2.0% for the electron channel and 3.1% for the muon channel. Uncertainties on the other D_s decay chains efficiencies, arising from the change in the ratios of the efficiencies, are negligible ($\sim 0.2\%$).

8.2.6 Radiative corrections - PHOTOS

The final state radiation of all particles is simulated by the PHOTOS package [36]. It features a conservative estimation of the uncertainty of 20% due to radiation corrections. A Signal MC sample of the same size like the original Signal MC sample but with PHOTOS being disabled, has been produced to study the effect of the final state radiation. The final systematic uncertainty to this simulation has been determined by using the difference between the two samples multiplied by the uncertainty of PHOTOS. This leads to systematic uncertainties as shown in table 8.7.

Table 8.7: Systematic uncertainties arising from the simulation of final state radiation by PHOTOS.

D_s decay chain	syst. uncertainty electron channel	syst. uncertainty muon channel
$D_s \rightarrow \phi\pi$	2.0%	2.1%
$D_s \rightarrow K^{*0}K$	2.2%	2.5%
$D_s \rightarrow K_S^0 K$	1.9%	1.9%

8.2.7 B counting

The determination of the number of $B\bar{B}$ pairs through the algorithm, described in reference [51] leads to a systematic uncertainty of 1.1%.

8.2.8 D_s branching ratios

The uncertainties of the branching ratios of the D_s decay chains used in this analysis introduce a systematic uncertainty to the branching ratio efficiency ϵ^{BR} . The values, shown in table 5.1 of section 5.2, lead to relative uncertainties of 15.1% for the $D_s \rightarrow \phi\pi$ decay chain, 15.4% for the $D_s \rightarrow K^{*0}K$ decay chain and 6.0% for the $D_s \rightarrow K_S^0 K$ decay chain. Since it is expected that these uncertainties decrease with further experimental input, their impact on the error of the branching fraction will be quoted separately.

8.3 Total systematic uncertainties relative to efficiencies

All multiplicative systematic uncertainties evaluated are summarized in table 8.8 and added in quadrature to get the total systematic uncertainty.

Table 8.8: Overview of the multiplicative systematic uncertainties studied in this analysis.

source	syst. uncertainty electron channel			syst. uncertainty muon channel		
	$\phi\pi$	$K^{*0}K$	K_S^0K	$\phi\pi$	$K^{*0}K$	K_S^0K
Signal Monte Carlo model	7.6%	3.1%	5.9%	0.2%	6.4%	2.1%
Signal Monte Carlo statistics	2.7%	3.3%	2.5%	3.5%	4.2%	3.3%
Particle identification	0.6%	1.2%	3.6%	1.6%	4.9%	7.9%
Track reconstruction	0.4%	0.5%	1.8%	0.1%	0.2%	2.4%
Photon reconstruction	0.6%	0.4%	0.5%	0.9%	0.9%	0.7%
K_S^0 corrections	-	-	2.0%	-	-	3.1%
Radiative corrections	2.0%	2.2%	1.9%	2.1%	2.5%	1.9%
Total uncertainty of ϵ^{reco}	8.4%	5.2%	8.1%	4.5%	9.5%	9.9%
B counting	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%
D_s branching ratios	15.1%	15.4%	6.0%	15.1%	15.4%	6.0%

Chapter 9

Results

With the systematic uncertainties evaluated in the last chapter, the determination of the final significance, taking into account the statistical and the additive systematic errors of the measurement, is possible. In addition, the combination of both lepton channels is done to present the final result as a weighted mean of both contributions.

9.1 Total significance of the result

Several additive systematic uncertainties affect the significance of the measurement due to their ability to decrease the number of fitted signal events and increase the probability of the signal being a background fluctuation. Since the multiplicative systematic uncertainties just affect the reconstruction efficiency and/or the translation of the number of fitted events into a branching ratio, they do not change the significance of the number of fitted events at all. The bias correction also only shifts the signal in a defined direction and thus has no influence on the probability of the existence of a signal. The uncertainties, which need to be considered are:

- the parametrization of the True- D_s fraction,
- the fit error on the size of the True- D_s fraction,
- the width of the signal PDF and
- the constraints on the single channels signal yields.

These four contributions are uncorrelated and thus can be added in quadrature to a total systematic uncertainty, relevant for the significance calculation σ_{syst} . To derive a significance of the result using the likelihood difference between the regular signal extraction fit and the one imposing a zero signal hypothesis, the likelihood had been extended by an additional term, containing the information about the systematic uncertainty. The modified negative logarithmic likelihood is:

$$-\ln \mathcal{L}' = -\ln \mathcal{L}(N_{\text{total}} + N_{\text{syst}}) + \frac{N_{\text{syst}}^2}{2\sigma_{\text{syst}}^2} \quad (9.1)$$

with \mathcal{L} being the likelihood, defined in equation 7.7 and N_{syst} being a floating number representing the additive systematic uncertainty.

The additional term represents a gaussian fluctuated contribution from the systematic error of the measurement. It is normalized, so that the regular fit minimizing $-\ln \mathcal{L}'$ reports values of the order of 10^{-4} for N_{sys} with the same minimized likelihood value as the fit using \mathcal{L} . The significance of the measurement can be calculated, using equation 7.8. It has to be noted, that the problem of $\Delta NDF \neq 1$ is present in this calculation, as well. Thus, table 9.1 gives the significance values for $\Delta NDF = 1; 2$, whereas the true value is somewhere in between. The combination of both measurements has been done by simply adding the likelihood differences of both lepton final states. The combined measurement of both lepton

Table 9.1: Significance of the result.

	electron channel	muon channel	combined
$-\log \mathcal{L}'_{fit}$	-44015.1	-26359.4	-70374.5
$-\log \mathcal{L}'_0$	-44008.3	-26351.9	-70360.2
$\Delta(\log \mathcal{L}')$	6.8	7.5	14.3
$\mathcal{S}(\Delta NDF) = 1$	3.7σ	3.9σ	5.4σ
$\mathcal{S}(\Delta NDF) = 2$	3.3σ	3.5σ	5.0σ

final states has a significance between 5.0σ and 5.4σ and thus can be interpreted as the first observation of the signal decay.

Sideband subtracted plots

To provide a better visualization of the signal, sideband subtracted plots are made. In these plots, the combinatorial background from the sideband has been subtracted from the signal region. Then, the fitted PDFs are drawn to illustrate the signal. In principle, the approach for making the plots is the same as the alternative fit method described in section A.2 of the Appendix. The only difference is, that the fit is still done as simultaneous maximum likelihood fit. Figure 9.1 shows the sideband subtracted plots with the fitted PDFs.

As can be seen, there's a clear overshoot of data compared to the True- D_s background, visualized by the blue curves, in the signal region of all channels. The red curves, which represent the PDF for True- D_s background plus a signal fraction, describe the data points better, indicating a strong evidence for the presence of a signal.

9.2 Calculation of the branching ratio

The calculation of the branching ratio from the signal yields of both lepton channels is done using

$$\mathcal{BR} = \frac{N(\text{total})_{\text{corrected}}}{2 \cdot N_{B\bar{B}} \cdot 1/2} \cdot \frac{1}{\sum_i \epsilon_i^{\mathcal{BR}} \epsilon_i^{\text{reco}}}. \quad (9.2)$$

In this equation, $N(\text{total})_{\text{corrected}} = 301/206$ is the number of bias-corrected signal events for electron and muon final states respectively and $N_{B\bar{B}} = 376.924 \times 10^6$ represents the number of $B\bar{B}$ pairs analyzed. The ratio of charged B mesons in the whole data set is taken to be $1/2$, since the ratio of $\Upsilon(4S)$ decaying into charged B mesons to the one into neutral B mesons is compatible with 1. The efficiencies of the reconstruction ϵ_i^{reco} and the D_s decay branching ratios $\epsilon_i^{\mathcal{BR}}$ are multiplied for each each D_s decay chain i and then added up to give the total detection efficiency.

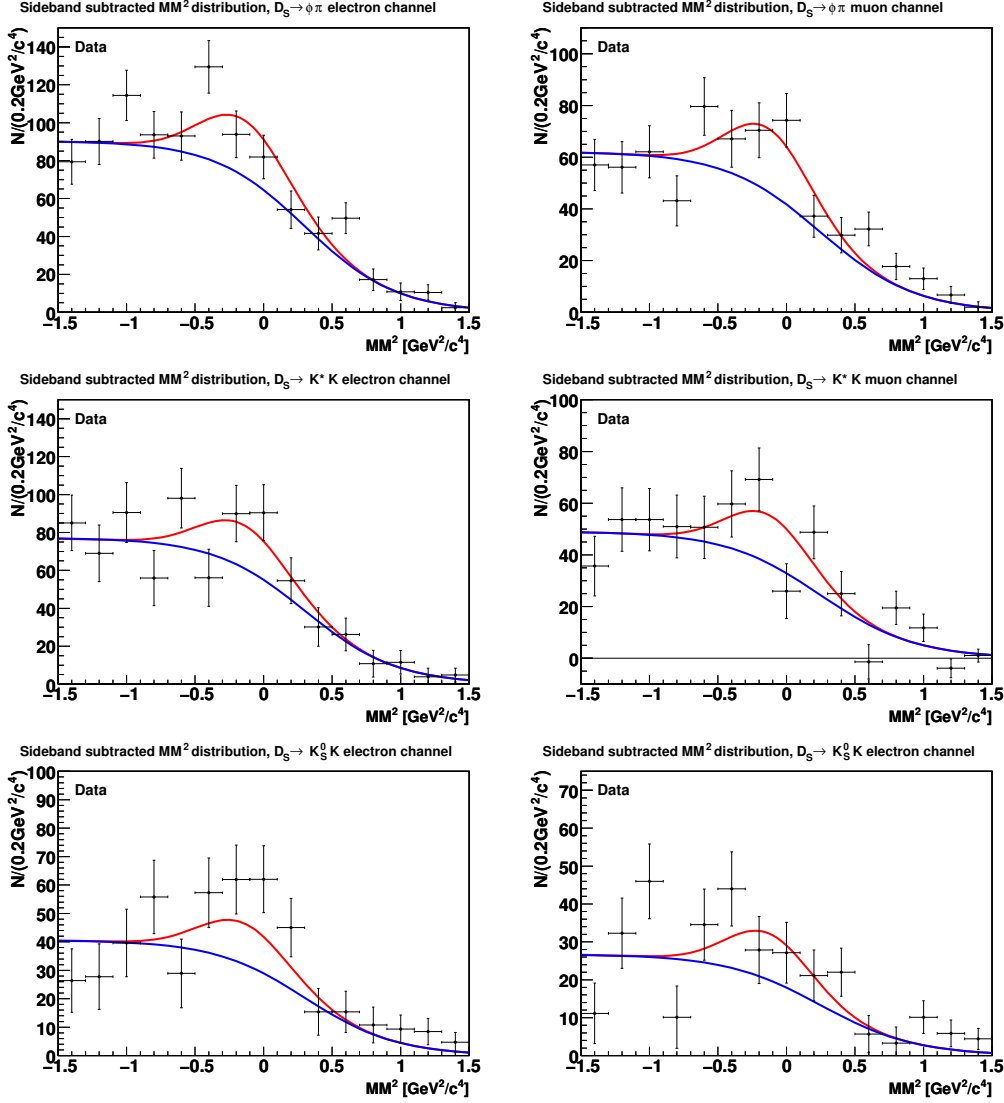


Figure 9.1: Sideband subtracted MM^2 distributions with fitted PDFs for electron (left hand plots) and muon channel (right hand plots). Blue curves represent only the True- D_s background, while the red curves describe the signal as well. Top row: $D_s \rightarrow \phi\pi$ channel, middle row: $D_s \rightarrow K^{*0}K$ channel, bottom row: $D_s \rightarrow K_S^0 K$ channel.

The reconstruction efficiencies are calculated from the number of reconstructed and produced Signal MC events $\epsilon_i^{\text{reco}} = \frac{N_i^{\text{Signal MC, reco}}}{N_i^{\text{Signal MC, produced}}}$. The branching ratio efficiencies are calculated by multiplying the branching ratios of the decays of the D_s and its daughters [8], since all of them are set to 100% in the Signal Monte Carlo. The multiplicative systematic uncertainties of the branching ratios were propagated to errors on the \mathcal{BR} and added in quadrature with the additive uncertainties. As mentioned before, the uncertainties arising from the errors of the D_s branching ratios are given exclusively to provide the opportunity to rescale the results in case of improved measurements reduce the errors on these quantities.

The final branching ratios of the single channels measured by this analysis are:

$$\begin{aligned}\mathcal{BR}(B^- \rightarrow D_s^+ K^- e^- \bar{\nu}_e) &= (5.81_{-1.30}^{+1.30}(\text{stat.}) \pm 0.54(\text{syst.}) \pm 0.49(\mathcal{BR}(D_s))) \times 10^{-4}, \\ \mathcal{BR}(B^- \rightarrow D_s^+ K^- \mu^- \bar{\nu}_\mu) &= (6.68_{-1.69}^{+1.72}(\text{stat.}) \pm 0.69(\text{syst.}) \pm 0.56(\mathcal{BR}(D_s))) \times 10^{-4},\end{aligned}$$

where the first error is the statistical error, the second error is the systematic uncertainty and the third error represents the uncertainty due to the limited knowledge of the branching ratios of the decays of the D_s and its daughters. The slight differences of the branching ratios central values between the actual result and the expectations from Toy MC calculations are consistent with the errors arising from the bias correction.

Combination of both lepton channels

The results of both lepton channels can be combined using reference [46]:

$$\begin{aligned}\mathcal{BR}_{\text{combined}} &= \frac{\mathcal{BR}_e/\sigma_e^2 + \mathcal{BR}_\mu/\sigma_\mu^2}{1/\sigma_e^2 + \mathcal{BR}_\mu/\sigma_\mu^2} \text{ and} \\ \sigma_{\mathcal{BR}} &= \frac{1}{1/\sigma_e^2 + 1/\sigma_\mu^2}.\end{aligned}$$

In the equations, $\mathcal{BR}_{\text{combined}}$ and $\sigma_{\mathcal{BR}}$ represent the branching fraction and error of the combined lepton channels. $\mathcal{BR}_{e/\mu}$ and $\sigma_{e/\mu}$ are the branching fraction and error of the electron channel and muon channel respectively. Only the statistical and systematic errors have been used for this calculation and have been added in quadrature. The combination of the two lepton states leads to the final result of this analysis:

$$\boxed{\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = (6.13_{-1.24}^{+1.26} \pm 0.51(\mathcal{BR}(D_s))) \times 10^{-4}.}$$

Chapter 10

Conclusions

The decay $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell$ has been analyzed using final states with electrons and muons. The D_s has been reconstructed in the three decay chains $D_s \rightarrow \phi(K^+ K^-)\pi$, $D_s \rightarrow \bar{K}^{*0}(K^\pm \pi^\mp)K$ and $D_s \rightarrow K_S^0(\pi^+ \pi^-)K$. Multi Layer Perceptron neural networks were used for optimizing the candidate selection in terms of deriving selection criteria with the highest possible significance. A simultaneous extended maximum likelihood fit of the Missing Mass squared of the reconstructed particles with respect to the nominal B mass has been performed, resulting in the first observation of the signal decay with a significance larger than 5σ . The branching ratios for the electron, muon and combined lepton final states were calculated to

$$\begin{aligned}\mathcal{BR}(B^- \rightarrow D_s^+ K^- e^- \bar{\nu}_e) &= (5.81_{-1.30}^{+1.30}(\text{stat.}) \pm 0.54(\text{syst.}) \pm 0.49(\mathcal{BR}(D_s))) \times 10^{-4}, \\ \mathcal{BR}(B^- \rightarrow D_s^+ K^- \mu^- \bar{\nu}_\mu) &= (6.68_{-1.69}^{+1.72}(\text{stat.}) \pm 0.69(\text{syst.}) \pm 0.56(\mathcal{BR}(D_s))) \times 10^{-4}, \\ \mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) &= (6.13_{-1.24}^{+1.26} \pm 0.51(\mathcal{BR}(D_s))) \times 10^{-4},\end{aligned}$$

where the errors are statistical, systematic and due to the limited knowledge of the branching ratios of the D_s and its daughters, respectively. The first error of last branching ratio represents the total statistical and systematic error of the combined measurement.

The measurement is in very good agreement with the result, derived by a earlier unpublished analysis of the same decay, which used a smaller fraction of the $BABAR$ data. The final result of this analysis was $\mathcal{BR}(B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}_\ell) = (5.51 \pm 1.94) \times 10^{-4}$ [44], where the total error represents the addition of the statistical error, the systematic uncertainty and the uncertainty arising from the limited knowledge of the D_s decay fractions. Thus, the total error of the measurement of the signal decay has been reduced by about $\sim 30\%$. In addition, the result of this thesis also agrees with the upper limit for this decay, published by the ARGUS collaboration in 1993 [17] and with the theoretically predicted branching ratio range. Nevertheless, the signal branching ratio is too small to completely explain the $\sim 2\%$ difference between the total inclusive semileptonic branching fraction and the sum of all exclusively known contributions and further analyses are needed to fill this gap.

Although the full $BABAR$ data set is about 30% larger than the used fraction, it is expected, that this result will be the final work of $BABAR$ on this decay. An analysis using the full data set would reduce the statistical error of the measurement by only $\sim 15\%$.

Appendix A

A.1 Study of D_s^* contributions to the signal

Besides the signal decay, also the decay with an excited D_s^* (decaying into $D_s\gamma$ with a branching fraction of $\mathcal{BR}(D_s^* \rightarrow D_s\gamma) = (94.2 \pm 0.7)\%$) in the final state can contribute to the signal. Due to the uncertainty, whether the signal decay takes place dominantly as a non-resonant decay or a resonant decay and, if the latter, through which D^{**} resonances this happens, no reliable estimation of the fraction of decays with a D_s^* in the final state can be derived. Since the γ is not reconstructed in the analysis, the MM^2 distribution is shifted in such decays, because there's one more missing particle in the $B \rightarrow D_s^*K\ell\nu$ decay. This shift doesn't affect the signal extraction, since the position of the signal is a free parameter in the fit. It has been studied, whether it is possible to separate the decays with a D_s^* from those without. An additional sample of signal Monte Carlo with excited D_s^* in the final state has been produced and reconstructed. The MM^2 distributions of both samples have been compared. The fractions of the two samples contributing to the total signal have been varied from 0% to 100% in 10% intervals. The maximal possible admixture, a 50% : 50% composition is shown in figure A.1 for all D_s reconstruction chains and both lepton channels. As can be seen, both gaussian distributions overlap each other too much for being able to separate them. Even if there's no background at all, the resolution is not good enough. The change of the fraction also causes a larger width of the signal distribution. However, the deviation is smaller than the error, arising from the simultaneous fit to the signal MC MM^2 distributions. Thus, the influence of D_s^* contributions on the signal is covered by the systematic uncertainties of the fit error on the signal width.

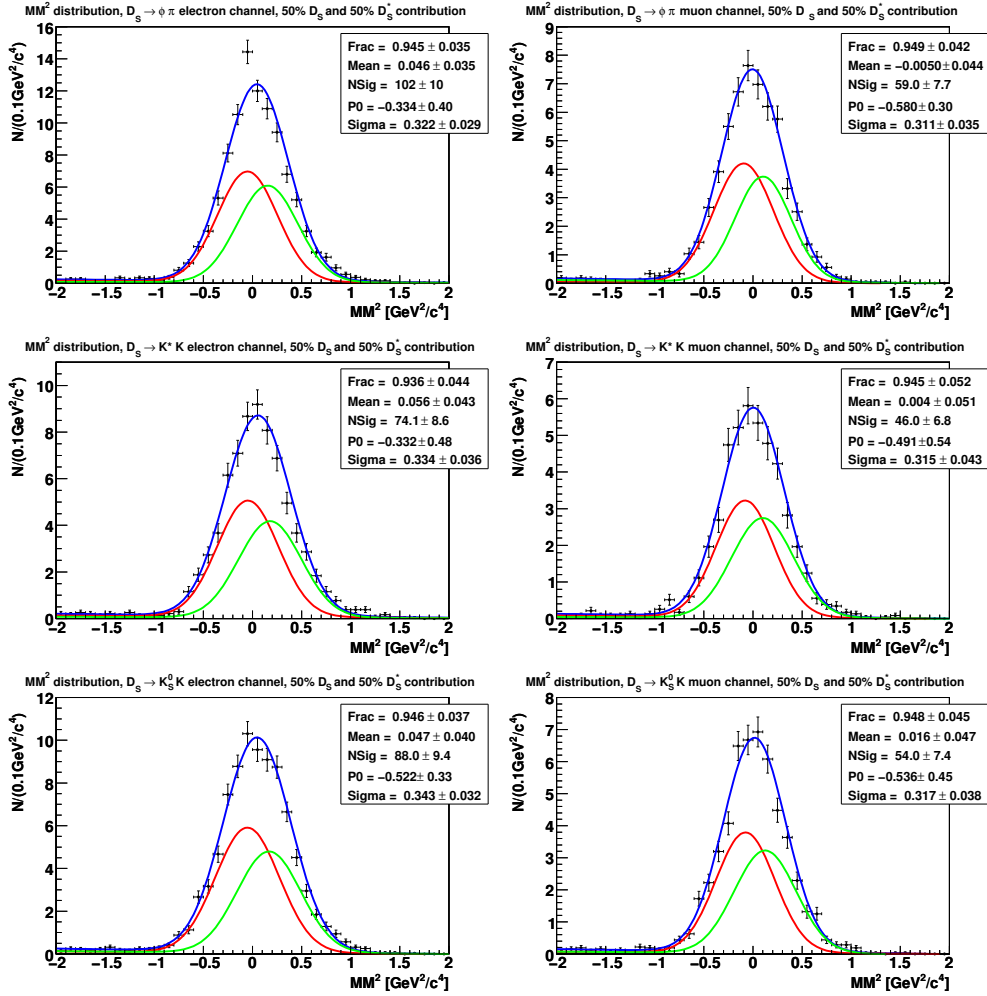


Figure A.1: Comparison of MM^2 distributions for decays with D_s and D_s^* in the final state. Left plots show the distributions for the electron channel, right plots the ones for the muon channel. Top row: $\phi\pi$ decay chain, middle row $K^{*0}K$ decay chain, bottom row K_S^0K decay chain. The data has been scaled to a signal branching ratio of 5×10^{-4} . Red curves represent the contribution from decays with D_s in the final state, green curves show the contribution from decays with D_s^* in the final state and the blue curves are the fit to the shown data and also the sum of both components.

A.2 Fit of sideband subtracted data

An alternative method for the extraction of the signal yield has been developed, using a $m(D_s)$ sideband subtracted data sample as input. The determination of signal region and sideband in the D_s mass distribution is done the same way as for the default method, but before running the MM^2 fit, the sideband data is scaled according to the relative size of the chosen signal region range. Then, it is subtracted from the signal region data and a simultaneous (in the three D_s decay chains) χ^2 fit is done, determining the fractions of the signal in the full data sample. The PDFs for signal and True- D_s background are:

$$\begin{aligned} PDF^{\phi\pi} &= f^{\phi\pi} \times PDF_{\text{Sig}} + (1 - f^{\phi\pi})PDF_{\text{True-}D_s \text{ bkg}} \\ PDF^{K^*0K} &= f^{K^*0K} \times PDF_{\text{Sig}} + (1 - f^{K^*0K})PDF_{\text{True-}D_s \text{ bkg}} \\ PDF^{K_S^0K} &= f^{K_S^0K} \times PDF_{\text{Sig}} + (1 - f^{K_S^0K})PDF_{\text{True-}D_s \text{ bkg}}, \end{aligned}$$

where f represents the fraction of signal events in the fitted sample. The parametrizations of the single components are the same as in the default method, a Fermi-function for the True- D_s background and a gaussian for the signal component. The signal width is constrained to MC values, as well as the fractions of the single D_s decay chains are constraint to each other using the same expression as in the default fit. Since the fit has to be done as a χ^2 fit, the final number of signal events needs to be calculated using the signal fraction parameter, multiplied with total the number of events in the samples. One variable of the Fermi-function, ExpoConst (describing the slope of the Fermi-edge), has to be fixed to the background generic MC value in order to get a reasonable agreement between input signal fraction¹ and fit result. Figure A.2 shows the generic MC validation plots for the fits, while table A.1 gives an overview of the fit results.

Table A.1: Results of the simultaneous validation fits to the sideband subtracted generic MC with a defined signal fraction for electron and muon channel.

lepton channel	electron channel	muon channel
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow \phi\pi$	$[1.957 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$	$[1.957 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.980]$
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow K^*0K$	$[1.956 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$	$[1.956 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.981]$
$m(D_s)_{\text{sig. reg.}}, D_s \rightarrow K_S^0K$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.983]$	$[1.955 \leq \frac{m(D_s)}{\text{GeV}/c^2} \leq 1.982]$
Signal Mean	$(-0.148 \pm 0.041) \text{ GeV}^2/c^4$	$(-0.206 \pm 0.052) \text{ GeV}^2/c^4$
χ^2/NDF	146.8/117	87.1/87
$N(\text{total}) (N(\text{total})_{\text{input}})$	$312.6 \pm 45.6 (235)$	$211.2 \pm 40.5 (138)$
$N(\phi\pi) (N_{\text{input}})$	$137.0 \pm 17.9 (104)$	$91.6 \pm 15.4 (61)$
$N(K^*0K) (N_{\text{input}})$	$106.2 \pm 14.4 (78)$	$72.5 \pm 13.6 (46)$
$N(K_S^0K) (N_{\text{input}})$	$69.4 \pm 13.3 (53)$	$47.0 \pm 11.6 (31)$

Two interesting things can be noted, considering the fit results. First, the number of fitted events and the shift of the mean of the signal components gaussian agree very well with the results, obtained by the default fit method. There remains a bias in this method as well. The source of the bias is a shift of the second parameter of the Fermi-function, describing the background, which results in a shift of the complete function to lower values. Since the signal is located on the decreasing flank, the shift introduces a bias.

¹For the validation, $\mathcal{BR}_{\text{signal}} = 5 \times 10^{-4}$ has been used.

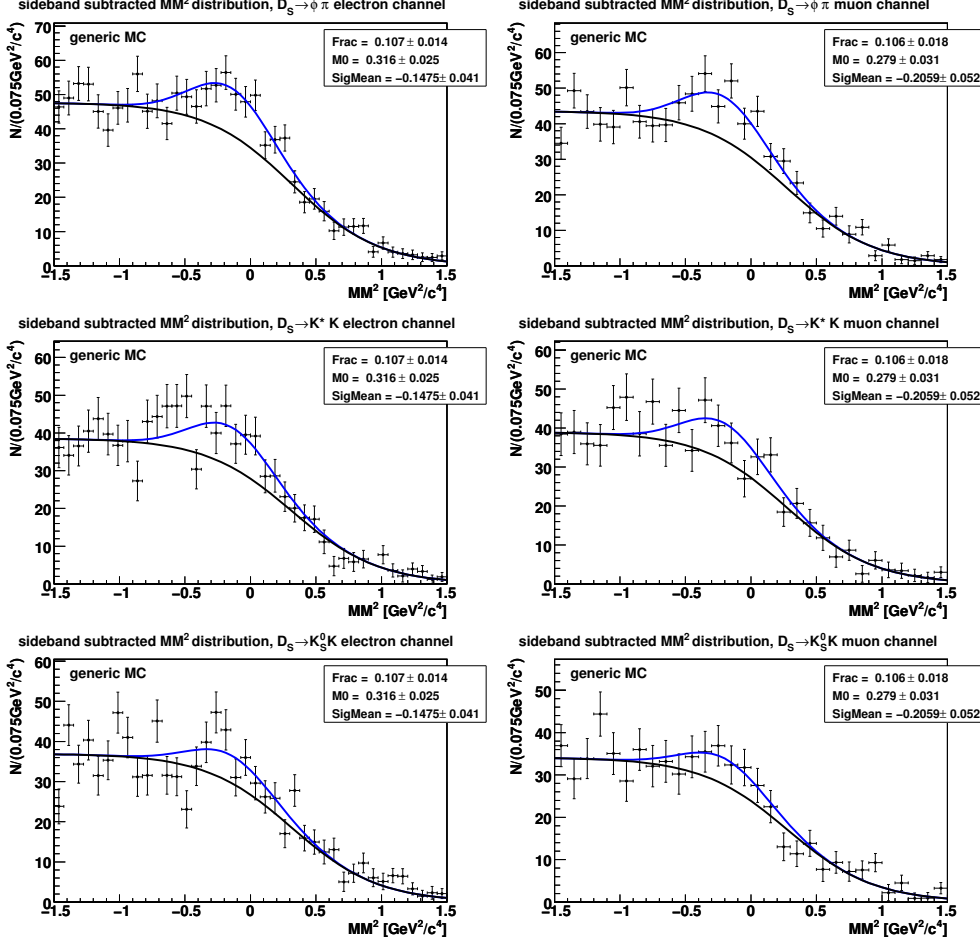


Figure A.2: Simultaneous χ^2 fits to sideband subtracted generic MC for electron channel (left column) and muon channel (right column). Top row: $\phi\pi$ decay chain, middle row $K^{*0}K$ decay chain, bottom row $K_S^0 K$ decay chain. Blue curves represent the full fit PDF, while black curves are the background PDFs only.

Nevertheless, it has been decided to stay with only one fixed parameter for the reason of lower systematic uncertainties due to the background parametrization².

Second, the fit errors on the signal yields are very small, actually too small (the expected error is $\Delta N_{\text{Signal}} \sim 75$). This might be a relic of the error propagation during the construction of the sideband subtracted data samples, but is not completely understood. Thus, a more reliable estimation on the expected fit error is given by the Toy MC studies.

The Toy MC samples, generated for validating the default fit method, are used for the validation of the alternative method as well. The studies have been limited to only one signal branching ratio, $\mathcal{BR}_{\text{signal}} = 5 \times 10^{-4}$. Figure A.3 shows the events distributions and pulls for both lepton channels.

²The sideband subtracted fit was first the default method, so these thoughts were quite relevant at some time during the analysis.

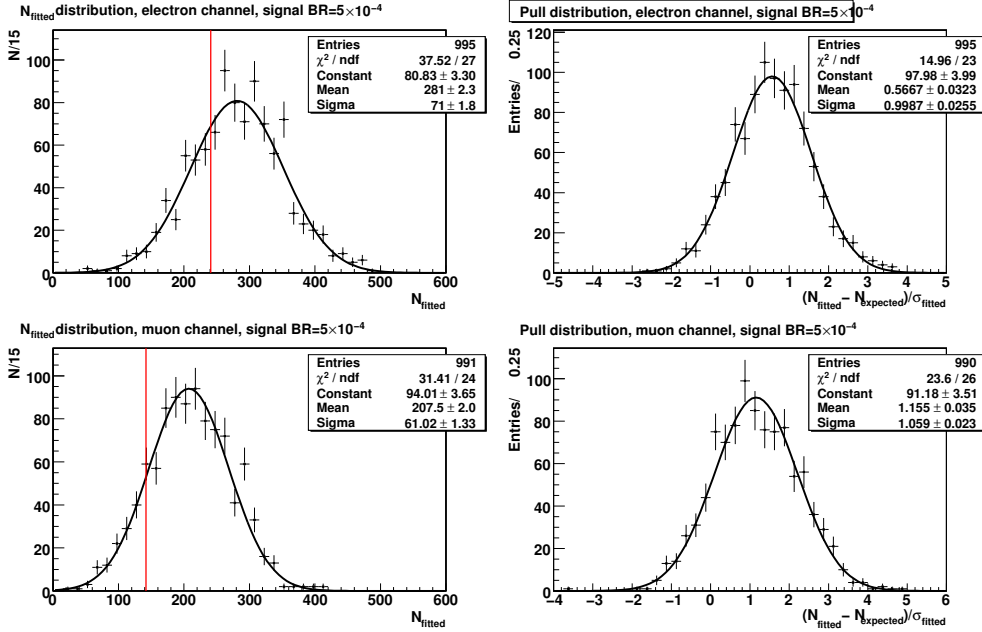


Figure A.3: Toy MC results for of the fit to the sideband subtracted MC sample electron channel (top row) and muon channel (bottom row). Left plots show the N_{fitted} distributions with signal added according to a signal branching ratio of $\mathcal{BR}_{\text{Signal}} = 5 \times 10^{-4}$, while right plots show the pull distributions. The red line indicates the input signal size.

As can be seen, the bias in this method is compatible with the expectation from the generic MC fit. It can be concluded, that the disagreement between Toy MC bias and the expectation from generic MC validation in the default method originates only from including the combinatorial background in the fit. Toy studies with the sideband subtracted fit showed, that there appears no bias in the measurement, when both parameters describing the True- D_s background remain fixed in the fit.

Despite the problem with the fit bias, the non-sideband subtracted fit method has been chosen as default fit method. The main reasons for this were the expectation of a slightly smaller fit error, compared to the alternative fit method³, the problematic fit errors in the generic MC validation and the slightly lower signal yield because of the subtraction of the events. Especially the latter is an important point, since a decrease in the signal yield would result in a less significant measurement and possibly eliminate the opportunity to claim the first observation of the signal decay.

³The expected fit error of the alternative fit method has been concluded from the width of the Toy MC number of fitted events distribution.

A.3 Toy MC validation plots

The full set of Toy MC validation plots, distribution of fitted events and pull distribution for each simulated signal \mathcal{BR} are shown in figures A.4-A.9.

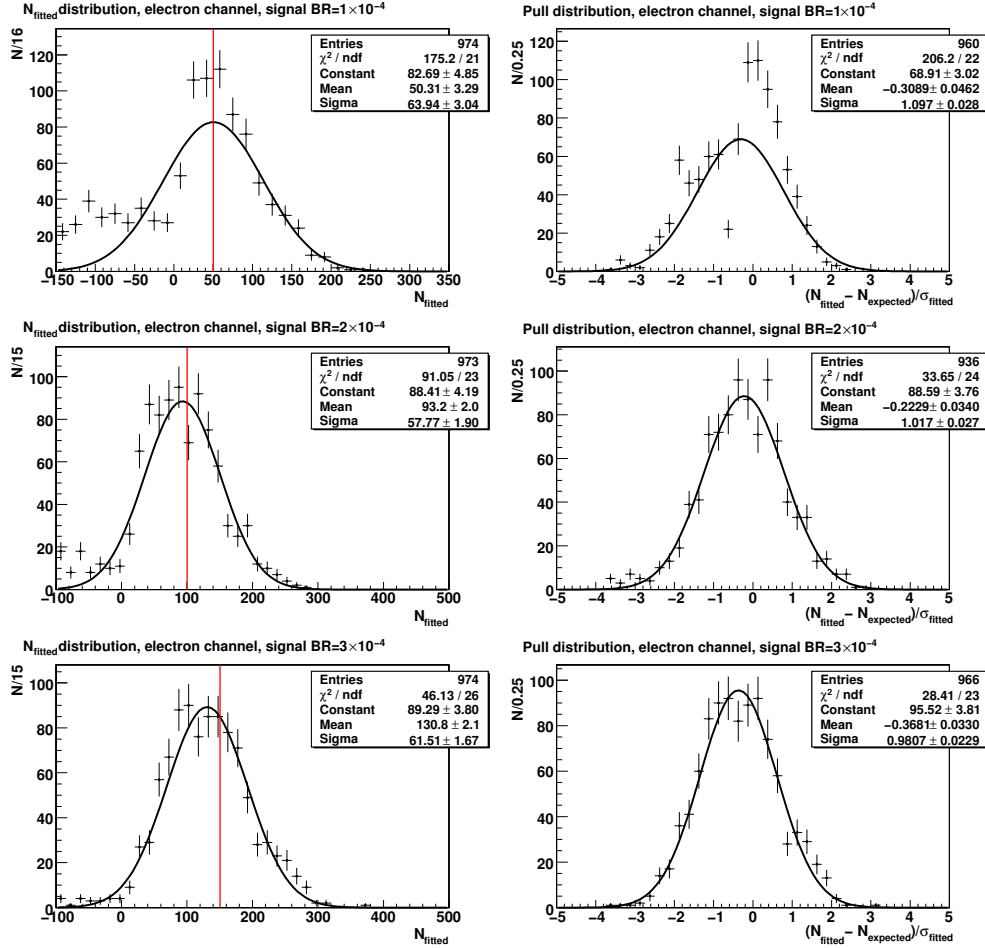


Figure A.4: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = [1/2/3] \times 10^{-4}$ for electron channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

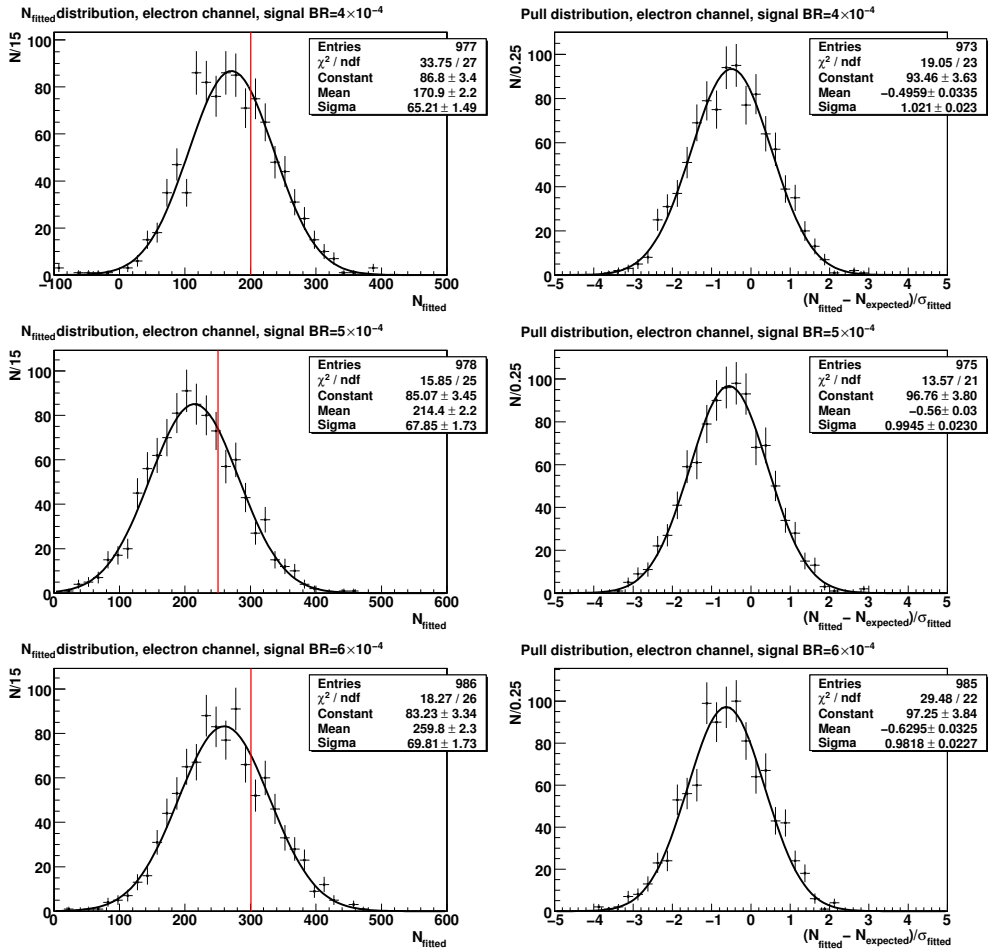


Figure A.5: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{signal}} = [4/5/6] \times 10^{-4}$ for electron channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

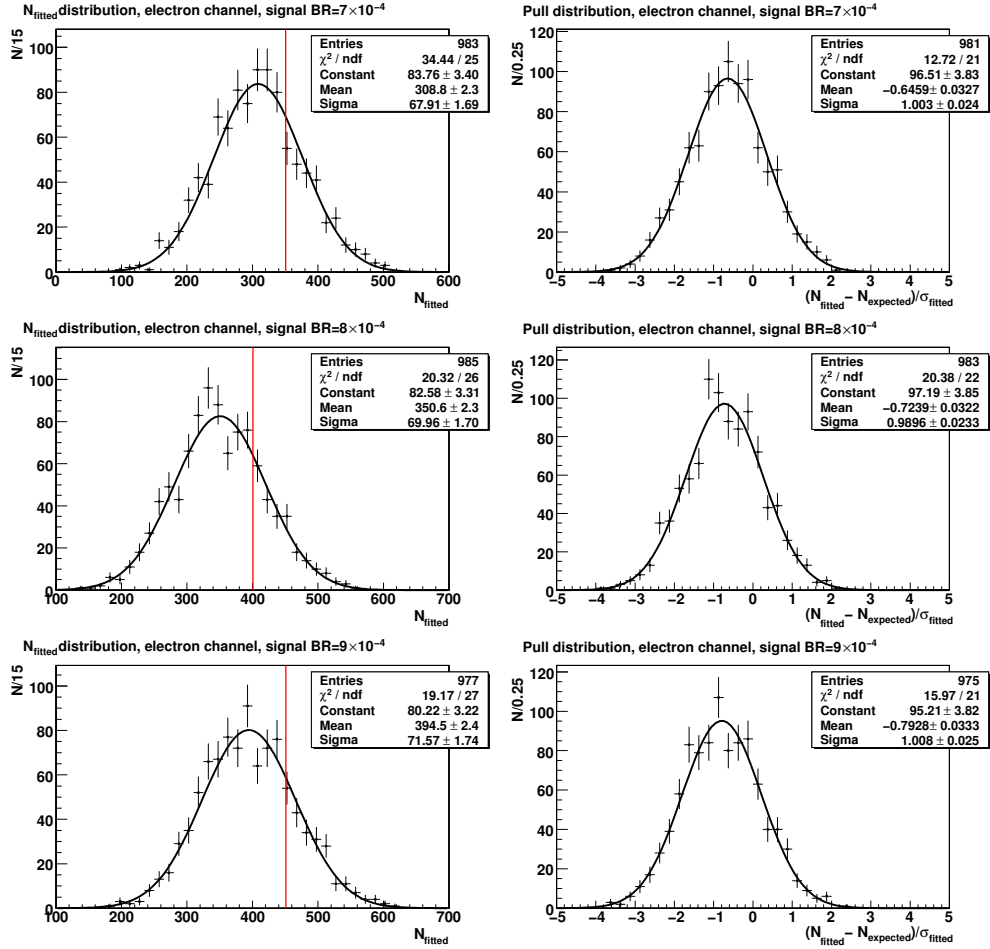


Figure A.6: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = [7/8/9] \times 10^{-4}$ for electron channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

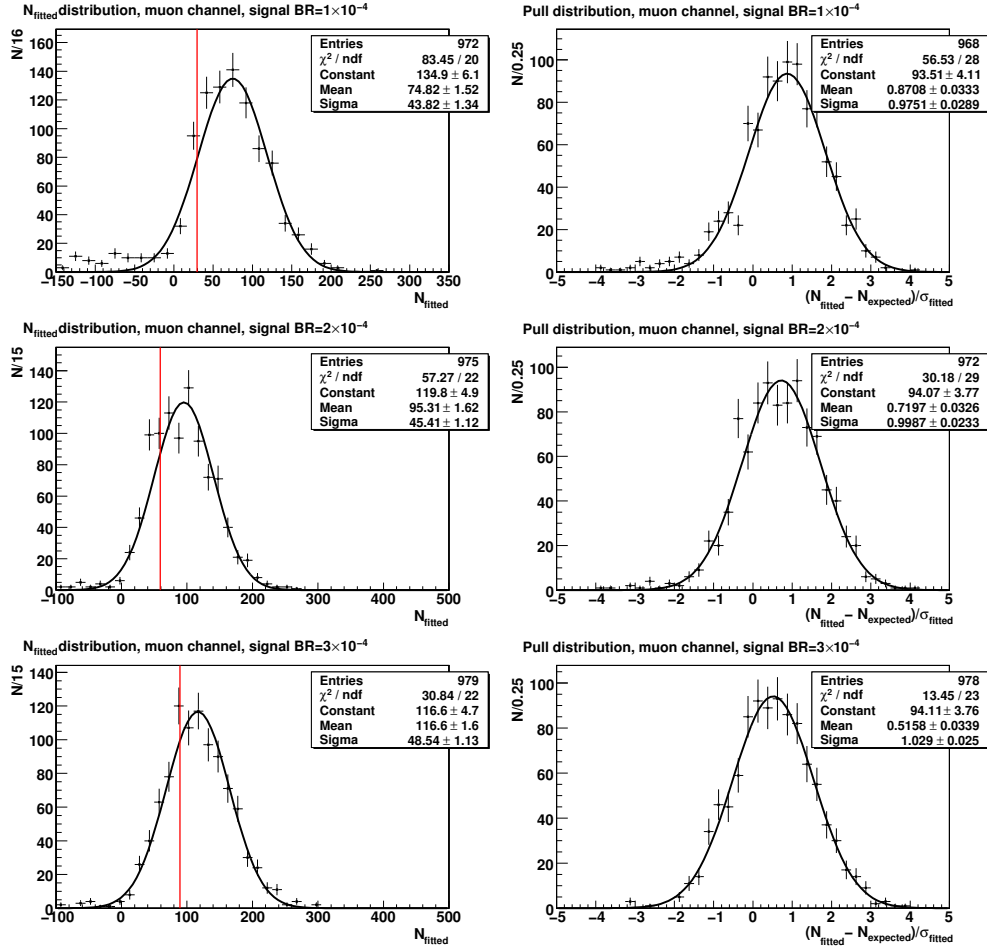


Figure A.7: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = [1/2/3] \times 10^{-4}$ for muon channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

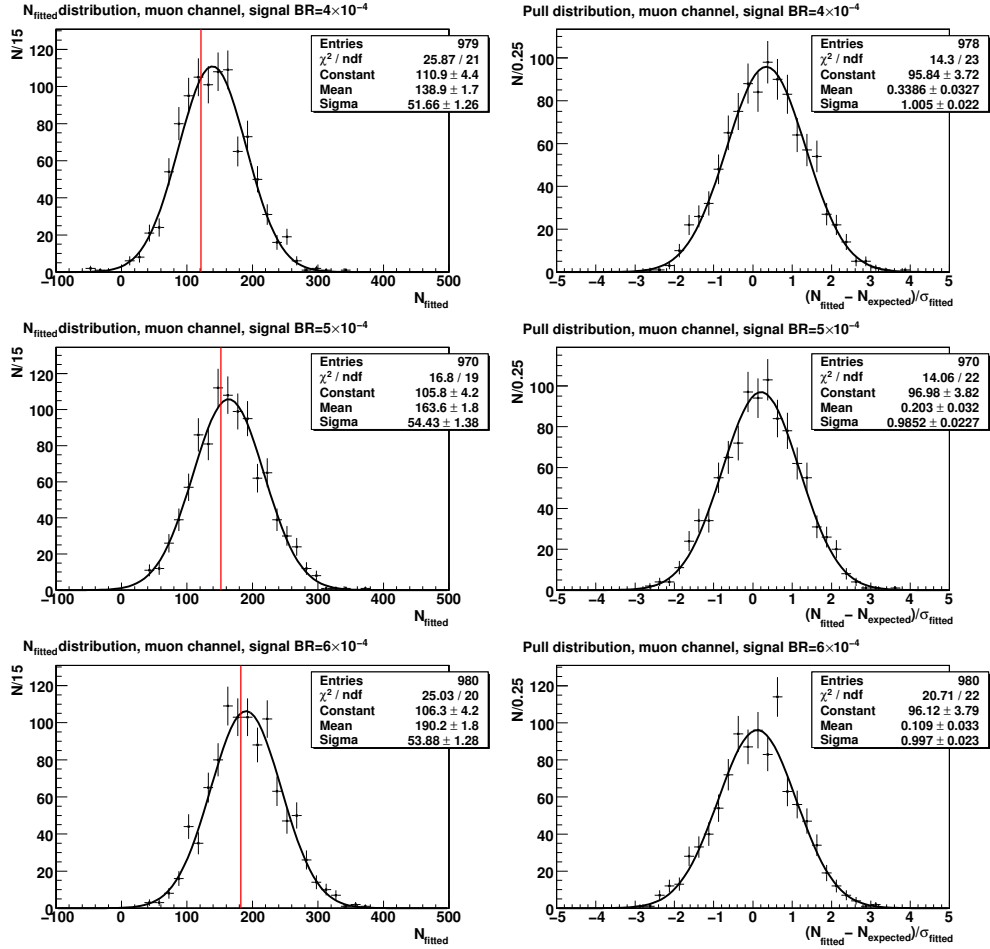


Figure A.8: Toy MC results for signal branching ratio of $\text{BR}_{\text{Signal}} = [4/5/6] \times 10^{-4}$ for muon channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

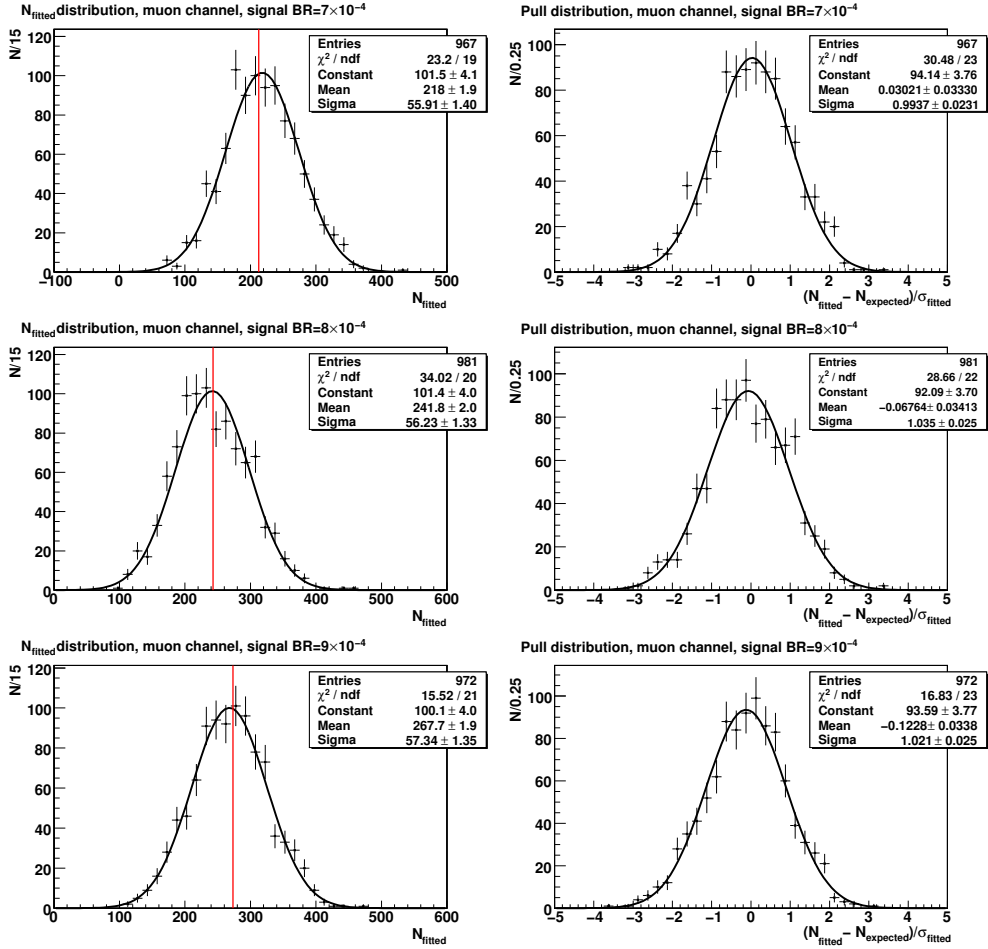


Figure A.9: Toy MC results for signal branching ratio of $\mathcal{BR}_{\text{Signal}} = [7/8/9] \times 10^{-4}$ for muon channel. Left plots show the N_{fitted} distributions, while right plots show the pull distributions. The red line indicates the input signal size.

References

- [1] A. D. Sakharov
Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe
ZhETF Pis'ma 5, **1** 32
1967
- [2] J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay
Evidence for the 2π Decay of the K_2^0 Meson
Phys. Rev. Lett. **13** 138
1964
- [3] T. Müller
Gravitation und Quantentheorie
PhD thesis, Eberhard-Karls-Universität Tübingen
2001
- [4] O. Nachtmann
Elementarteilchenphysik Phänomene und Konzepte
vieweg Verlag
1991
- [5] P. W. Higgs
Broken symmetries and the masses of Gauge Bosons
Phys. Rev. Lett. **13**, 508
1964
- [6] P. W. Higgs
Spontaneous Symmetry Breakdown without Massless Bosons
Phys. Rev. **145**, 1156
1966
- [7] L. Wolfenstein
Parametrization of the Kobayashi-Maskawa Matrix
Phys. Rev. Lett. **51**, 1945
1983
- [8] C. Amsler *et al.* [Particle Data Group]
Review Of Particle Physics
<http://pdg.lbl.gov>
2008

- [9] M. Neubert
Heavy Quark-Symmetry
Phys. Rept. **245**, 256
1994
- [10] P. F. Harrison, H. R. Quinn [BABAR Collaboration]
The BABAR Physics Book
SLAC
1998
- [11] B. Aubert *et al.* [BABAR Collaboration]
Measurement of the Branching Fractions of Exclusive $\bar{B} \rightarrow D/D^/D^{(*)}\pi l^- \bar{\nu}_l$ Decays in Events Tagged by a Fully Reconstructed B Meson*
arXiv: 0708.1738
2007
- [12] B. Aubert *et al.* [BABAR Collaboration]
Determination of the Form Factors for the Decay $B^0 \rightarrow D^{-} l^+ \nu_l$ and of the CKM Matrix Element $|V_{cb}|$*
arXiv: 0705.4008
2008
- [13] B. Aubert *et al.* [BABAR Collaboration]
*Measurement of the Decay $B^- \rightarrow D^{*0} e^- \bar{u}$*
arXiv: 0707.2655
2007
- [14] K. Abe *et al.* [BELLE Collaboration]
Determination of $|V_{cb}|$ using the semileptonic decay $\bar{B}^0 \rightarrow D^{+} e^- \bar{u}$*
Phys. Lett. **B526** 247
2002
- [15] T. Brandt
Electron Identification and Measurement of the Inclusive Semileptonic Branching Fraction of B Mesons at the BABAR Experiment
PhD thesis, Technische Universität Dresden
2001
- [16] B. Aubert *et al.* [BABAR Collaboration]
Determination of the Branching Fraction for $B \rightarrow X_c \ell \nu$ Decays and of $|V_{cb}|$ from Hadronic-Mass and Lepton-Energy Moments
hep-ex 0404017
2007
- [17] H. Albrecht *et al.* [ARGUS Collaboration]
Search for Rare B Meson Decays into D_s^+ Mesons
ZPhys. C60, 11
1993
- [18] D. Liventsev *et al.* [BELLE Collaboration]
Measurement of the branching fractions for $B^- \rightarrow D^{()+} \pi^- \ell^- \nu_\ell$ and $B^0 \rightarrow D^{(*)0} \pi^+ \ell^- \nu_\ell$*
Phys. Rev. D **72**, 051109
2005

-
- [19] B. Aubert *et al.* [*BABAR* Collaboration]
Measurement of the branching fractions of exclusive $\bar{B} \rightarrow D^{()}(\pi)\ell^{-}\nu_{\ell}$ decays in events with a fully reconstructed B meson*
Phys. Rev. Lett. **100**, 151802
2008
- [20] J. L. Goity, W. Roberts
Soft Pion Emission In Semileptonic B -Meson Decays
hep-ph 9406236
1994
- [21] N. Isgur, D. Scora
Semileptonic meson decays in the model: An update
Phys. Rev. D **52**, 2783
1995
- [22] N. Isgur, D. Scora, B. Grinstein, M. B. Wise
Semileptonic B and D decays in the quark model
Phys. Rev. D **39**, 799
1989
- [23] B. Aubert *et al.* [*BABAR* Collaboration]
The $BABAR$ detector
Nucl. Instrum. Methods **A479**, 1
2002
- [24] G. Benelli *et al.*
The $BABAR$ LST detector high voltage system: design and implementation
IEEE Nucl. Sci. Symp. Conf. Rec. 2, 1145
2005
- [25] B. Aubert *et al.* [*BABAR* Collaboration]
Study of time-dependent CP -violating asymmetries and flavor oscillations in neutral B decays at the $\Upsilon(4S)$
Phys. Rev. D **66**, 032003
2002
- [26] [*BABAR* Collaboration]
BtaCandidate Lists
<http://www.slac.stanford.edu/BFROOT/www/doc/workbook/eventinfo/BtaCandInfo/CandLists.html>
2007
- [27] Charged Particle AWG [*BABAR* Collaboration]
New BtaCandidate Lists
<http://www.slac.stanford.edu/BFROOT/www/Physics/Charged.html#Lists>
2008
- [28] D. J. Lange
The EvtGen particle decay simulation package
Nucl. Instrum. Methods **A462**, 152
2001

- [29] T. Sjostrand
High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4
Comput. Phys. Commun. **82**, 74
1994
- [30] S. Agostinelli *et al.* [The GEANT4 Collaboration]
GEANT4 - a simulation toolkit
Nucl. Instrum. Methods **A 506**, 250
2003
- [31] G. Mancinelli, S. Spanier [BABAR Collaboration]
Kaon Selection at the BABAR Experiment
BABAR Analysis Document 116
2001
- [32] A. Mohapatra, J. Hollar, H. Band [BABAR Collaboration]
Studies of A Neural Net Based Muon Selector for the BABAR Experiment
BABAR Analysis Document 474
2004
- [33] A. V. Telnov [BABAR Collaboration]
A new method of track-level dE/dx calibration for the BaBar Drift Chamber and Silicon Vertex Tracker
BABAR Analysis Document 1500
2007
- [34] A. Gaz, BABAR PID Group [BABAR Collaboration]
BABAR Particle ID
<http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/pid.html>
2009
- [35] D. N. Brown [BABAR Collaboration]
MC Production - Home Page -
<http://www.slac.stanford.edu/BFROOT/www/Computing/Offline/Production/>
2009
- [36] W. Richter-Was *et al.*
QED bremsstrahlung in semileptonic B and leptonic τ decays
Phys. Lett. **B303** 163
1993
- [37] G. Punzi
Sensitivity of searches for new signals and its optimization
hep-ph: 0308063
2003
- [38] G. C. Fox, S. Wolfram
Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes
Phys. Rev. Lett. **41**, 1581
1978

-
- [39] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. Toerne, H. Voss [TMVA Group]
Toolkit for Multivariate Data Analysis with ROOT
<http://tmva.sourceforge.net/>
2009
- [40] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. Toerne, H. Voss [TMVA Group]
TMVA 4 Users Guide
<http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf>
2009
- [41] D. L. Perna [BABAR Collaboration]
Tracking Efficiency Task Force
<http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2007-R22.html>
2007
- [42] K. S. Cranmer
Kernel estimation in high-energy physics
hep-ex: 0011057
2001
- [43] F. Simonetto
Partially reconstructed $B^0 \rightarrow D^ \ell \nu$ decays*
BABAR Analysis Document 1377
2006
- [44] H. Jasper
Search for the semileptonic decay $B^- \rightarrow D_s^+ K^- \ell^- \bar{\nu}$
BABAR Analysis Document 1299
2007
- [45] F. Porter [BABAR Collaboration]
Hypothesis Tests, Significance
http://www.slac.stanford.edu/BFROOT/www/Organization/Workshops/2008/Babar_Physics_Analysis_School/feb14-thurs-b/L0802B.pdf
2008
- [46] G. Cowan
Statistical Data Analysis
Clarendon Press, Oxford
1998
- [47] [BABAR Collaboration]
Details of PidLHElectronSelector
http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Selectors/r22a-old/PidLHElectronSelector_details.html
2007
- [48] [BABAR Collaboration]
Details of NNVeryTightMuonSelection
http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Selectors/r22a-old/NNVeryTightMuonSelection_details.html
2007

REFERENCES

- [49] M. Purohit, R. M. White
PID Systematic errors
BABAR Analysis Document 1465
2007
- [50] M. T. Allen, M. T. Naisbit, A. M. Roodman
A Study of π^0 Efficiency
BABAR Analysis Document 870
2004
- [51] C. Hearty
Hadronic Event Selection and B-Counting for Inclusive Charmonium Measurements
BABAR Analysis Document 30
2000

Acknowledgments

At first, I'd like to thank my adviser Prof. Dr. Bernhard Spaan for the opportunity to become a part of the international *BABAR* collaboration, as well as to the Lehrstuhl für Experimentelle Physik 5 at the Technische Universität Dortmund. His helpful ideas, patience and especially his confidence and support helped me a lot during the last almost 5 years.

I am also very grateful to Dr. Jochen Dingfelder, who I first met during one of my stays at SLAC National Laboratory in Palo Alto/Menlo Park, California. As convener of the semileptonic Analysis Working Group, but also in private communications he provided valuable advice in physics issues and shared his experience in making the analysis ready for being published in a scientific journal. The publication will finally complete this work.

Furthermore, I have to thank the Bundesministerium für Bildung und Forschung for the great financial support, which made a German contribution to the *BABAR* experiment possible and allowed me several journeys to SLAC as well as the participation on various physics meetings.

During my time as PhD student in Dortmund and at SLAC, I met a lot of new people, found many new friends - many of them contributed in a way, they might even not have noticed to this thesis. Especially I'd like to thank the Rostock, Dresden and Karlsruhe guys I met at SLAC: We've had a lot of informative discussions and it was always fun. Of course, I also want to thank the whole Dortmund group for lots of support in many ways. My officemates Enrico and Moritz were always open for inspiring conversations about physics, C++ programming and the world beyond. The "Halo"-night (and tabletop game and role-playing) guys, as well as the super-diploma students provided a lot of distraction (and cookies) when I needed it. Klaus answered all my questions about statistics and helped me a lot in the final part of the analysis. So, once again: Thanks to all of you!

Separately, I'd like to mention our Norwegian colleague Magnus and the M&M's, Mirco and his wife Maika. We spend a lot time together talking about important things, as well as nonsense, laughing, playing, drinking, watching movies, and making some kind of sport. I enjoyed every minute of that time and wish to thank you for spending it with me.

Finally, I have to thank my whole family, especially my parents and my sister for their continuous and full support: Vielen, vielen Dank! Ohne euch wäre ich nicht so weit gekommen.