Karlsruhe Institute of Technology

Measurement of the branching ratio of $B \to D^{(*)}\tau\nu_\tau$ relative to $B \to D^{(*)}\ell\nu_\ell$ decays with hadronic tagging at Belle

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

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

genehmigte

Dissertation

von

Dipl.-Phys. Matthias Huschle

aus Achern

Tag der mündlichen Prüfung:	23. Januar 2015
Referent:	Prof. Dr. Michael Feindt
Korreferent:	Prof. Dr. Günter Quast

Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
- 1.2 Analysis Outline . . . 3
- 1.3 Previous Analyses . . . 4
  - 1.3.1 Previous Belle Analyses . . . 4
  - 1.3.2 Previous BaBar Analysis . . . 5
  - 1.3.3 Previous Doctoral Thesis . . . 5

2 Theory **7**
- 2.1 $B \to D^{(*)}\tau\nu$ in the Standard Model . . . 7
- 2.2 NP Influences . . . 8
  - 2.2.1 Two-Higgs-Doublet Models . . . 8
  - 2.2.2 NP predictions . . . 9

3 Experimental Setup **13**
- 3.1 The KEKB Accelerator . . . 13
- 3.2 The Belle Detector . . . 14
  - 3.2.1 Overview . . . 14
  - 3.2.2 SVD . . . 15
  - 3.2.3 CDC . . . 15
  - 3.2.4 PID . . . 16
  - 3.2.5 ECL . . . 18
  - 3.2.6 KLM . . . 19
  - 3.2.7 EFC . . . 19
  - 3.2.8 Triggers . . . 20

4 Reconstruction **21**
- 4.1 General Strategy . . . 21
- 4.2 Tag Side Reconstruction . . . 21
  - 4.2.1 Kinematics at B Factories . . . 21
  - 4.2.2 Full Reconstruction Framework . . . 22
    - 4.2.2.1 Reconstruction Principle . . . 22
    - 4.2.2.2 Decay Channels . . . 23
    - 4.2.2.3 Performance . . . 24
    - 4.2.2.4 Continuum Suppression . . . 24
- 4.3 Candidate Combination . . . 25
- 4.4 D^{**} Enriched Sample . . . 25
- 4.5 Reconstruction Cuts . . . 26
- 4.6 Analysis Cuts . . . 27
- 4.7 Best Candidate Selection . . . 29

5 Data Samples **31**
5.1 Four Reconstruction Modes . . . 31
5.2 Real Data . . . 32
5.3 Simulated Data . . . 32
5.3.1 Types of Simulated Data . . . 32
5.3.1.1 Generic Monte Carlo . . . 32
5.3.1.2 Signal Monte Carlo . . . 32
5.3.1.3 $B \to D^{**}\ell\nu$ Monte Carlo . . . 33
5.4 Event Generation . . . 33
5.4.1 Decay Simulation . . . 33
5.4.2 Detector Simulation . . . 34

6 Data Composition **35**
6.1 Missing Mass Distributions . . . 35

7 Correction of Simulated Data **39**
7.1 Tag Side Efficiencies . . . 39
7.2 D_s Component . . . 39
7.3 $D^{(*)}\ell\nu$ components . . . 40
7.4 $D^{**}\ell\nu$ components . . . 41
7.5 Lepton ID Efficiency . . . 41
7.6 Lepton Fake Rates . . . 41
7.7 Wrong Tag Yield . . . 42
7.8 $D^{(*)}$ Meson Corrections and Sidebands . . . 43
7.8.1 Correct $D^{(*)}$ Characteristics . . . 43
7.8.2 Wrong $D^{(*)}$ Characteristics . . . 43
7.8.3 Correction Procedure . . . 43
7.8.4 Weight Application . . . 44
7.8.5 Cut Application and Sideband Sample . . . 45

8 Fitting Variables **49**
8.1 Squared Missing Mass . . . 49
8.2 Other Observables . . . 49
8.3 Neural Network . . . 50
8.3.1 Training . . . 50
8.3.2 Transformation . . . 51
8.4 Sample Split . . . 51

9 Fit Procedure **55**
9.1 Fit Components . . . 55
9.1.1 Probability Density Functions . . . 55
9.1.2 Component Descriptions . . . 55
9.2 Full PDF . . . 58
9.3 Factors . . . 58
9.4 Shapes . . . 59
9.5 Expected Yields . . . 65

10 Fit Validation **67**
10.1 Fit Projections . . . 67
10.2 Cross Validation . . . 78
10.3 Data - Monte Carlo Comparison . . . 80
10.4 Inverse q^2 Sample Comparison . . . 87

10.5 Fits to D^{**} Enriched Sample . . . 88
10.5.1 Sample Reconstruction . . . 88
10.5.2 Sample Fit . . . 88
10.5.3 Compatibility to Reweighted PDFs . . . 93
10.5.4 Conclusion . . . 93
10.6 Comparison to BaBar's D^{**} Sample . . . 93
10.7 Toy Study . . . 97
10.7.1 Description . . . 97
10.7.2 Conclusion . . . 97
10.7.3 Expected Statistical Uncertainty . . . 99
10.7.4 Split Value Validation . . . 99
10.8 New Physics Sensitivity . . . 103
10.8.1 Sensitivity . . . 103
10.8.2 q^2 Tests . . . 103

11 Results **111**
11.1 Fit Result . . . 111
11.2 Fit Projections . . . 111
11.3 Branching Ratios . . . 113

12 Systematic Uncertainties **115**
12.1 Production Parameters . . . 115
12.2 D^{**} Composition . . . 115
12.3 Factors for Simultaneous Fitting . . . 116
12.4 PDF Shapes . . . 116
12.5 Lepton ID Efficiency . . . 116
12.6 Correlations . . . 117

13 Discussion **119**
13.1 $R - R^*$ Plane . . . 119
13.2 Difference to Previous Result . . . 120
13.3 New Physics Search . . . 122
13.3.1 Fit Procedure . . . 122
13.3.2 Results . . . 122
13.3.3 q^2 Distributions . . . 122
13.4 Combination with BaBar Result . . . 122

14 Conclusion **127**

Appendix **131**
A Probability Density Functions . . . 132
B Control Plots . . . 141

1. Introduction

1.1 Motivation

I like ice cream. Unsurprisingly, I already did as a child. The taste, the sensation of the melting sweetness, marvelous! But there is more: By waiting long enough or through dedicated stirring, the ice cream totally changes its consistency, from solid to liquid – a miracle! How did this happen?

From today's point of view, one might be tempted to romanticize the innocent curiosity of children, but it is a trait that can hardly be overrated. It is what makes us dream, envision what possibilities we have, and push their boundaries. For thousands of years man has looked up to the birds in the sky, wondering if he could – maybe – one day fly. Understanding his observation and using it to fulfill his dream may have demystified one of the miracles of his universe, but opened the door to discovering many more of them – and stimulating his dreams of exploring space.

One of the biggest questions, that has always been stimulating man's imagination, is: "Why is the universe the way it is?" A question that was answered countless times in history, but never empirically. The development of modern physics starting in the early 20^{th} century has allowed for the first time to find scientifically consistent explanations for our everyday observations, ranging from the perception of colors to the processes that keep our sun alive. Yet we are not at the core of the universe's existence, and may never be. But we keep on searching – curious, how the next better answer will widen our field of view.

Our current model to answer this question is the *Standard Model* (SM) of particle physics. It has been in place for about four decades and although its parameters have been measured with increasing precision, most discoveries in this time span were already predicted by it. However, there are substantial hints that the SM is not elementary but merely a special case of an underlying mechanism for energy scales within our reach. The most prominent ones come from cosmology: The SM does not provide candidates for dark matter and dark energy, that constitute the vast majority of the energy in the universe. And it has no mechanism to produce the asymmetry of matter and antimatter that we observe in the sheer existence of our and about 170 billion other galaxies. On earth, however, despite countless efforts, the only observed discrepancy to the SM is the oscillation of neutrinos, which requires them to have a mass.

One of the difficulties in searches for *new physics* (NP) is, that to find a solid discrepancy between theoretical prediction and experimental result, we need a good precision in both of

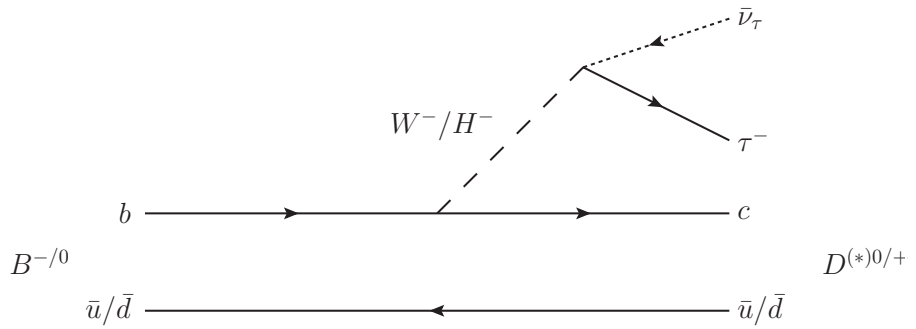


Figure 1.1: Feynman diagram of the B meson decay to $D^{(*)}\tau\bar{\nu}_\tau$. In various NP extensions a charged Higgs boson can take the place of the W boson and interfere with the SM amplitude.

them. Due to their theoretical cleanliness leptonic and semileptonic B meson decays have become a prominent place to search. In various supersymmetric NP models the existence of charged Higgs bosons is postulated. These could replace the $W^\pm$ boson in these weak decays and therefore interfere with the decay amplitude of the SM. The advantage of leptons in these decays is, that a τ lepton is much heavier than the other leptons, but not heavy enough to be strongly suppressed by phase space limitations. As the Higgs bosons couple to the mass of particles, they would affect decays involving τ leptons much more than others. So among the simplest and cleanest channels to search for NP influences of this kind are $B \to \tau\nu$ and $B \to D^{(*)}\tau\nu$. This thesis is about the latter, whose Feynman diagram is shown in Fig. 1.1. A detailed theoretical background is given in Chapter 2.

The latest result from the BaBar collaboration[1] concerning this decay has sparked a lot of discussion in the flavor physics community, as it disagreed with both the standard model *and* the then-favored NP model (two-Higgs-doublet model of type II), by more than 3σ, making it one of the last remaining measurements with evidence for influences of NP.

In this thesis I will provide an independent measurement using data from the Belle experiment at the KEKB e^+e^- collider in Tsukuba, Japan. Due to the neutrinos in the final state, this measurement benefits enormously from the clean environment of lepton colliders and it is highly questionable, if there will be comparable measurements from LHC experiments in the future. So the measurement in this thesis will be the last experimental contribution to this puzzle for several years.

We were children then
Something bid us look to the sky
Saw the hawk that soared
And we saw the sparrow winging his way
Knew that we were meant for something,
Something more
For we would join them
We thought of angels and of jets
Of rockets and capes of crimson

Glass Hammer - "Into Thin Air"

1.2 Analysis Outline

The quantities to be measured in this thesis are the relative decay fractions of $B \to D^{(*)}\tau\nu_\tau$ (the *signal*) with respect to $B \to D^{(*)}\ell\nu_\ell$ (the *normalization*) as defined by Eq. (1.1), where ℓ can be one of the lighter leptons, e or μ.

$$R \equiv \frac{\mathcal{B}(B \to D\tau\nu_\tau)}{\mathcal{B}(B \to D\ell\nu_\ell)} \quad \text{and} \quad R^* \equiv \frac{\mathcal{B}(B \to D^*\tau\nu_\tau)}{\mathcal{B}(B \to D^*\ell\nu_\ell)} \tag{1.1}$$

The B mesons are produced in e^+e^- collisions in pairs with opposing charges at a known center-of-mass (CM) energy. They decay within picoseconds and the end products of their cascading decays are recorded by the Belle detector, which is described in Chapter 3. Fig. 1.2 shows a schematic of a signal decay from the production of the B meson pair. The $B_{\textbf{signal}}$ decays to the requested D meson, a τ lepton and a neutrino. The $D^{(*)}$ meson then decays into several lighter particles. The τ lepton quickly decays to a lighter lepton and two additional neutrinos to fulfill the lepton flavor conservation (hadronic decays of *tau* leptons are not considered in this thesis). Neutrinos barely ever interact with the detector and are therefore invisible to us, which gives us a final state with three invisible particles. Furthermore, in the normalization decay all τ related particles are omitted, but the visible final state particles are the same.

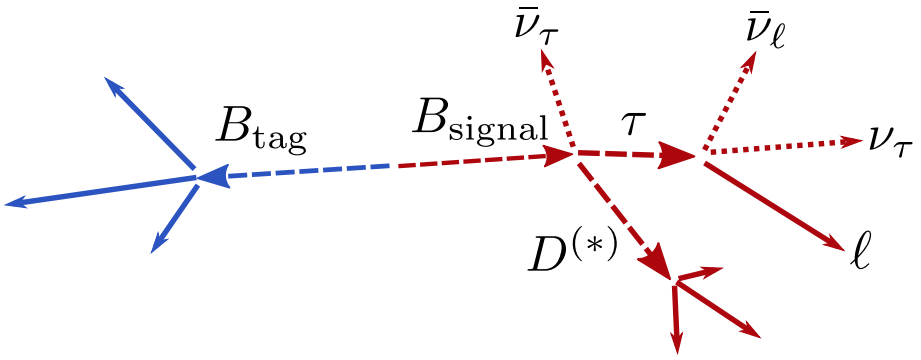


Figure 1.2: Schematic of a signal decay in this analysis. Dashed lines are intermediate states, and dotted lines are particles, that are invisible to the detector.

To still be able to distinguish these two cases, it is necessary to exploit the kinematics of the decay. As the CM energy is known as well as the nominal B meson masses, I reconstruct the other B meson in hadronic decay channels without invisible particles, and thereby derive the four-momentum of the signal B meson. Details of the reconstruction procedure will be described in Chapter 4. This process allows one to extract the *squared missing mass*, which is the invariant mass of the invisible particle system, defined as

$$M^2_{\text{miss}} \equiv \left(p_{\text{CMS}} - p_{B_{\text{tag}}} - p_{D^{(*)}} - p_\ell\right)^2 . \tag{1.2}$$

As the masses of neutrinos are very close to zero, this observable peaks around zero for events where only one neutrino is in the final state, which are the normalization events. For signal events the squared missing mass represents a system of three neutrinos, that takes much higher values.

This analysis uses several different recorded and simulated data samples, whose characteristics are explained in Chapter 5.

Chapter 6 will give an overview of the different components in our sample, besides signal and normalization. The most important background components are candidates containing wrongly reconstructed D or D^* mesons, and candidates coming from semileptonic B meson

decays with higher excitations of D mesons, named D^{**} background, which are the most difficult to distinguish from our signal. As we make extensive use of simulated data, Chapter 7 lists numerous corrections, that have to be applied to retrieve viable information out of it. In this process, the detailed handling of candidates with wrong $D^{(*)}$ mesons is defined, which will be constrained using the sidebands in D mesons mass and $D^* - D$ mass difference of real data.

As the squared missing mass is not sufficient to perform a precise measurement, especially due to the D^{**} background, Chapter 8 discusses other possible observables as fitting dimensions. Several observables are combined in a neural network, dominated by the remaining energy in the electromagnetic calorimeter E_{ECL}. The network is not optimized for separation of signal from all other components, but from D^{**} backgrounds alone. This is accompanied by a rather uncommon fitting strategy: The fitting samples are split at a certain value of M^2_{miss} and fitted simultaneously – the lower region in M^2_{miss} itself and the upper region in the network output. While missing mass provides an excellent estimation for the normalization yields in the lower sample, the higher sample is pure enough to enable a proper separation between signal and D^{**} background. Details on probability density functions, yields, constants and constraints, that are used in the fit, are given in Chapter 9. The overall strategy for parameter estimation is the extended maximum likelihood method, applied simultaneously on all data samples.

Chapter 10 is dedicated to validating the fitting procedure. This comprises simple checks to test the validity of the implementation, but also extended procedures to test the validity of the simulated data, especially concerning the D^{**} background. In its last section, the sensitivity of the fitting procedure with regards to NP effects is examined.

The results of the fitting procedure will be given in Chapter 11, with systematic uncertainties being discussed in Chapter 12. A discussion of the results is provided in Chapter 13, and will not only show the result for itself, but also discuss the compatibility with previous measurements of the BaBar and Belle collaborations, SM predictions, and test a NP hypothesis.

1.3 Previous Analyses

The subject of this thesis has been addressed before by the BaBar and Belle experiments with different analysis strategies. Their outcome and differences to this thesis are given in the following:

1.3.1 Previous Belle Analyses

From 2007 to 2010, there have been several measurements by the Belle Collaboration[2][3] of charged and neutral B decays to $D^{(*)}\tau\nu$ with an inclusive tagging method. In this method, all events are searched for a signature compatible with the signal. Other than in this thesis, some hadronic τ decays were also allowed. As the momenta of B mesons produced at the B factories are tightly constrained, the remainder of the event was then tested for compatibility with a B meson. The momentum of the D meson was used as an additional discriminating variable. The advantage of this method is a higher efficiency compared to the method in this thesis. Its disadvantage is a higher background level and additional systematic uncertainties coming from efficiency effects. In these analyses no normalization sample was reconstructed, so the overall strategies are entirely different and the resulting quantities were the respective branching ratios.

In 2009, the Belle Collaboration published a measurement of $B \to D^{(*)}\tau\nu$[4] performed on a dataset of $657 \times 10^6\, B\bar{B}$ pairs. It was much more similar to this thesis, as it utilized

only leptonic τ decays, reconstructed a normalization sample, and used hadronic tagging. The latter means that the first step in the candidate reconstruction is the search for a B meson decaying hadronically. The signature of the signal decay is only searched for after this *tag* B meson is found. The main difference of the measurements is, that this previous measurement did not assume isospin symmetry, meaning that the extracted R and R^* values were obtained individually for the B^0 and B^+ samples. The measurement in this thesis is going to surpass the previous result by: using a slightly larger data sample; using improved reconstruction methods with higher efficiencies; and applying a more sophisticated fit model, which takes into account many effects that were not considered before. However, some stages of the previous analysis are similar and will later be used for validation purposes.

1.3.2 Previous BaBar Analysis

In 2012 the BaBar collaboration published their final measurement of $B \to D^{(*)}\tau\nu$[5], that claimed evidence for an excess with respect to the SM. It measured the same quantities as the ones presented in this thesis, with similar techniques. At first sight, one might expect the two measurements to be closely related, but there are several significant differences. The BaBar and the Belle detector are both operated on $\Upsilon(4S)$ decays, but have different strengths and weaknesses, which become significant in a measurement as complex as this one. During validation, some aspects will be compared and discussed. As this measurement due to its complexity in both analyses holds a high potential for missing a significant influence, I also consider it to be of advantage, not to imitate their strategy, but develop an independent one.

1.3.3 Previous Doctoral Thesis

In 2013, a new measurement of the Belle collaboration was performed as part of the Ph.D. thesis of Daniel Zander[6]. While his result was approved by the collaboration at that time, some issues were discovered shortly after that could have significant influence on a measurement of this difficulty. The result was revoked before being published by the collaboration and must not be used in scientific work. However, several parts of this analysis still hold validity and were reused in this thesis (e.g. certain corrections to simulated data), while most of the selection and fitting process was entirely restructured. With respect to these previous results, the ones in this thesis should thereby not be seen as an improvement, but a replacement.

2. Theoretical Overview

The goal of this chapter is to shed light on the theoretical aspects of the particular ratios, that are measured in this thesis. It will be shown, why the ratio is preferred to a plain branching ratio measurement, and how NP effects can interfere with the standard model processes. I intend to give a brief overview of the theories, that motivate this analysis, which have been dealt with in much greater detail elsewhere, e.g. [7–9].

2.1 $B \to D^{(*)}\tau\nu$ in the Standard Model

While the experimental handling of semileptonic B decays can be quite challenging in comparison to hadronic decays – due to neutrinos in the final state that can not be effectively detected – the theoretical calculations end up being more precise as the leptons are not affected by non-perturbative quantum chromodynamic (QCD) interactions. The Feynman diagram of the decays $B \to D^{(*)}\tau\nu$ is given in Fig. 1.1. Throughout this thesis, it is assumed, that QCD interactions involving the spectator quark (u, d) are negligible ($\equiv$ *isospin symmetry*). So the decay matrix element involves mainly one hadronic $b \to c$ transition and one leptonic $\tau\bar{\nu}_\tau$ interaction, mediated by a W boson. These two contributions factorize, but the product has to be added up for all helicity states λ_W of the W boson:

$$\mathcal{M}^{\lambda_\tau}_{\lambda_M}(q^2, \theta_\tau) = \frac{G_F V_{cb}}{\sqrt{2}} \frac{m_W^2}{m_W^2 - q^2} \sum_{\lambda_W} \eta_{\lambda_W} L^{\lambda_\tau}_{\lambda_W}(q^2, \theta_\tau) H^{\lambda_M}_{\lambda_W}(q^2) \ . \tag{2.1}$$

Here q^2 is the squared mass of the virtual W boson[1]. θ_τ is defined as the angle between the three-momenta of τ and $D^{(*)}$ in the rest frame of the B meson. $L^{\lambda_\tau}_{\lambda_W}$ represents the leptonic amplitudes, and $H^{\lambda_M}_{\lambda_W}$ the hadronic ones. The charm meson is represented by M, and λ_M represents its helicity state: s for the scalar D meson, and $\pm 1, 0$ for D^*. λ_τ denotes the τ helicity state (± 1), and λ_W can take the values $\pm 1, 0$, and s. η_{λ_W} is a metric factor with $\eta_{\pm 1,0} = 1$ and $\eta_s \approx -1$.

The calculation of the leptonic amplitudes

$$L^{\lambda_\tau}_{\lambda_W}(q^2, \theta_\tau) \equiv \epsilon_\mu(\lambda_W) \langle \tau(p_\tau, \lambda_\tau) \bar{\nu}_\tau(p_\nu) \mid \bar{\tau}\gamma^\mu(1 - \gamma_5)\nu_\tau \mid 0 \rangle \ , \tag{2.2}$$

where $\epsilon_\mu(\lambda_W)$ are polarization vectors of the W boson, is theoretically very clean as it can be done mostly analytically. I omit their explicit form, but for the understanding of the NP

[1] In the experimental approach to these decays from the next chapter on, the symbol q^2 will be redefined to the much better accessible four-momentum transfer from B meson to $D^{(*)}$ meson.

influences, it is important to know, that the leptonic amplitudes with $\lambda_W = s$ vanish for τ with negative helicity ($L_s^- = 0$), and only leptonic amplitudes producing τ with positive helicity have a direct proportionality on m_τ ($L_{\lambda_W}^+ \sim m_\tau$).

The majority of uncertainties in Eq. (2.1) comes from the calculation of the hadronic amplitudes

$$H_{\lambda_W}^{\lambda_M}(q^2) \equiv \epsilon_\mu^*(\lambda_W) \left\langle M(p_M, \lambda_M) \,\middle|\, \bar{c}\gamma^\mu(1-\gamma_5)b \,\middle|\, \overline{B}(p_B) \right\rangle \tag{2.3}$$

that contain several form factors in their explicit representation, which are constrained from measurements in $B \to D^{(*)}\ell\nu$ or HQET calculations. Another source for uncertainties in the branching ratio is the CKM element V_{cb}. However, it is also a factor in the calculation of $B \to D^{(*)}\ell\nu$ decays, which causes them to cancel out in the ratios $R^{(*)}$ of this thesis.

The decay rates can be calculated from these formulas using Fermi's Golden Rule, and the resulting predictions for the ratios R and R^* based on the SM are now quite precise [7]:

$$R_{\text{SM}} = 0.305 \pm 0.012 \tag{2.4}$$

$$R_{\text{SM}}^* = 0.252 \pm 0.004\ . \tag{2.5}$$

There exist several predictions, which are well compatible within their uncertainties and usually deviate because of slightly different inputs (form factors, quark masses). The uncertainties are however much smaller than the ones achievable by experiments. The current averaged experimental values from the BaBar and Belle experiments before this analysis are:

$$R_{\text{exp}} = 0.42 \pm 0.06 \tag{2.6}$$

$$R_{\text{exp}}^* = 0.34 \pm 0.03\ . \tag{2.7}$$

Both experimental values exceed the SM expectation with a combined significance of 3.5σ.

2.2 NP Influences

2.2.1 Two-Higgs-Doublet Models

The SM Higgs field is an SU(2) doublet containing four real components. Linear combinations of the field's components result in three Goldstone bosons, that are absorbed in the interaction with the weak gauge bosons $W^\pm$ and Z, giving them their masses. The fourth degree of freedom results in the Higgs boson itself. The same effect (massive gauge bosons, Higgs particle in the observed mass region) can be easily maintained, if an additional Higgs doublet is added. The four new field components would result in five Higgs particles instead of one: h^0, H^0, A^0, and $H^\pm$.

The extension of the SM by a second Higgs doublet is one major focus in the search for physics beyond the SM [10]. Motivations are (among others):

- **Supersymmetric** models require more than one Higgs doublet. Due to chirality, a single Higgs doublet would be unable to give mass to both up-type and down-type quarks simultaneously.
- A second Higgs doublet allows to construct a global U(1) symmetry of the QCD Lagrangian, which is a possible solution for the **strong CP problem** [11].
- Two-Higgs-Doublet models (2HDM) provide room for new **CP violating** mechanisms, that could help explaining the **baryogenesis**.

One major problem with 2HDM is the possibility for tree-level flavor changing neutral currents (FCNC), as the diagonalization of the mass matrix no longer automatically diagonalizes the Yukawa interactions between Higgs fields and fermions. Tree-level FCNCs have not been observed and must therefore be suppressed somehow. One way to achieve this, is to have each individual class of charged fermions (up-type, down-type, leptons) couple to only one Higgs doublet Φ_i:

type	up-type	down-type	charged lepton
type I (fermiophobic)	Φ_2	Φ_2	Φ_2
type II (MSSM-like)	Φ_2	Φ_1	Φ_1
lepton-specific	Φ_2	Φ_2	Φ_1
flipped	Φ_2	Φ_1	Φ_2

By convention the Higgs doublet that up-type quarks couple to is named Φ_2. 2HDM where fermions can couple to both Higgs doublets (not in the table) and have tree-level FCNC suppressed by other mechanisms (like large scalar masses) are called *type III*. As they lack the constraints on the couplings, they usually have a much bigger parameter space and don't allow predictions, that make them accessible with current experimental setups.

The focus of theoretical 2HDM research has been type II, mainly motivated by Supersymmetry, as the 2HDM in the minimal supersymmetric model (MSSM) is a subset thereof.

2.2.2 NP predictions

Regarding the charged Higgs contribution as modeled in 2HDM of type II, the additional amplitude is [12]:

$$\mathcal{M}^{\lambda_\tau}_{\lambda_M}(q^2,\theta_\tau)\bigg|_{H^\pm} = \frac{G_F V_{cb}}{\sqrt{2}} L'^{\lambda_\tau}\left[\tan^2\beta \frac{m_b m_\tau}{m^2_{H^\pm}-q^2} H'^{\lambda_M}_R + \frac{m_c m_\tau}{m^2_{H^\pm}-q^2} H'^{\lambda_M}_L\right] . \tag{2.8}$$

The $m_{b/c/\tau/H^\pm}$ in this equation are the masses of b quark, c quark, τ lepton, and the charged Higgs $H^\pm$. The factor $\tan\beta$ is the ratio of the vacuum expectation values of the two Higgs doublets and has a central role (besides $m_{H^\pm}$) in the magnitude of NP effects from this model.

The leptonic amplitude in Eq. (2.8) is quite similar to the $W^\pm$ mediated one in Eq. (2.2), omitting the polarization vectors:

$$L'^{\lambda_\tau}(q^2,\theta_\tau) \equiv \langle \tau(p_\tau,\lambda_\tau)\bar{\nu}_\tau(p_\nu) \mid \bar{\tau}(1-\gamma_5)\nu_\tau \mid 0\rangle .$$

Compared to the SM case in Eq. (2.3), the hadronic amplitude is split by chirality:

$$H'^{\lambda_M}_{R,L}(q^2) \equiv \left\langle M(p_M,\lambda_M)\,\middle|\,\bar{c}(1\pm\gamma_5)b\,\middle|\,\overline{B}(p_B)\right\rangle .$$

The upper sign applies for $H'^{\lambda_M}_R$ and the lower one for $H'^{\lambda_M}_L$. Leptonic and hadronic amplitudes of this 2HDM can be expressed in terms of the SM amplitudes in the scalar W polarization as follows:

$$\begin{aligned}
H'^{s}_{R,L} &= \tfrac{\sqrt{q^2}}{m_b-m_c} H^s_s \\
H'^{\pm 1}_R &= 0 \\
H'^{0}_R &= \tfrac{\sqrt{q^2}}{m_b+m_c} H^0_s \\
H'^{\pm 1,0}_L &= -H'^{\pm 1,0}_R \\
L'^{\lambda_\tau} &= \tfrac{\sqrt{q^2}}{m_\tau} L^{\lambda_\tau}_s
\end{aligned}$$

As the Higgs is a scalar, only currents with zero helicity contribute. Due to the conservation of angular momentum, the charged Higgs interference in the case of $M = D^*$ is restricted to longitudinally polarized D^* mesons.

By inserting these equations in Eq. (2.8) and adding Eq. (2.1), the form of Eq. (2.1) can be preserved by replacing the SM scalar hadronic current $H_s^{\lambda_M}$ with the total scalar hadronic current:

$$H_s^{2HDM} \equiv H_s^{\lambda_M} \times \left(1 - \frac{\tan^2\beta \times q^2}{(m_{H^\pm}^2 - q^2)(1 \mp m_c/m_b)} + \frac{q^2}{(m_{H^\pm}^2 - q^2)(m_b/m_c \mp 1)}\right) .$$

The upper sign applies to $M = D$ and the lower one to $M = D^*$. Actual values for q^2 in these decays are in the vicinity of $8\,\mathrm{GeV}^2$. As a light charged Higgs boson has already been excluded experimentally in $B \to X_s\gamma$ decays [13, 14], the assumption $m_{H^\pm}^2 \gg q^2$ provides a simpler form:

$$H_s^{2HDM} \approx H_s^{\lambda_M} \times \left(1 - \frac{\tan^2\beta \times q^2}{m_{H^\pm}^2(1 \mp m_c/m_b)}\right) .$$

Surprisingly, this term does not explicitly depend on m_τ. But as the charged Higgs only affects the scalar amplitudes, it only produces τ leptons with positive helicity ($L_s^- = 0$) via L_s^+, which *is* proportional to the τ mass. As the matrix element enters the decay rate in square, the scale of NP effects in $B \to D^{(*)}\tau\nu$ is enhanced relative to those in $B \to D^{(*)}\ell\nu$ by at least the order of $(m_\tau/m_\mu)^2 \approx 280$. This dependence of the NP influence on the τ helicity also implies, that the overall distribution of the τ helicity will be changed, as well as the distribution of momenta, due to the q^2 dependence.

Reevaluating $R^{(*)}$ using the total scalar hadronic current results in parabolic dependency on $\tan^2\beta/m_{H^\pm}^2$:

$$R_{2HDM}^{(*)} = R_{SM}^{(*)} + A^{(*)}\frac{\tan^2\beta}{m_{H^\pm}^2} + B^{(*)}\frac{\tan^4\beta}{m_{H^\pm}^4} \tag{2.9}$$

The parameters $A^{(*)}$ and $B^{(*)}$ have been calculated along with $R_{SM}^{(*)}$ including uncertainties and correlations from m_c/m_b and form factors in [1].

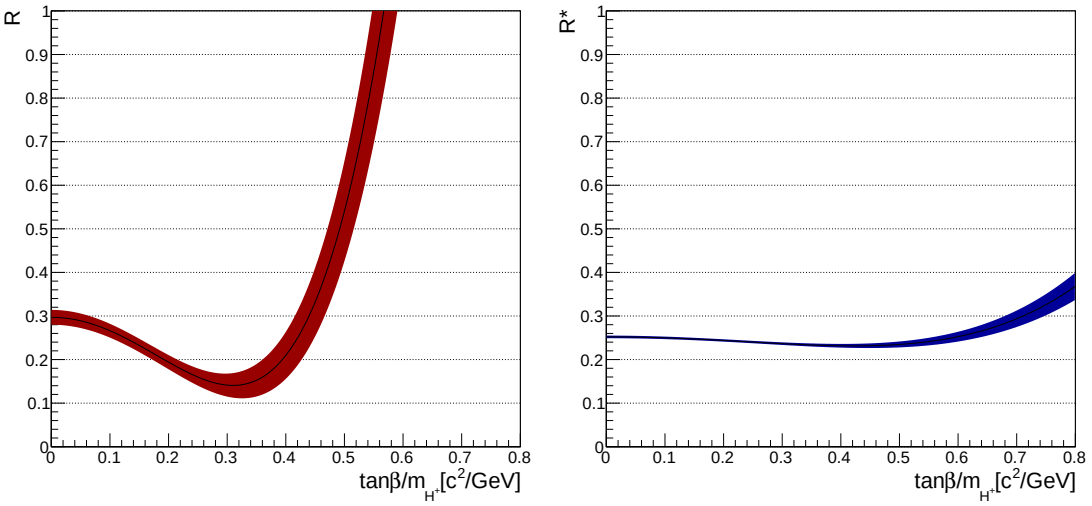


Figure 2.1: Effect of $\tan\beta/m_{H^\pm}$ values on the expectation of R (left) and R^* (right) in the 2HDM of type II. The values at $\tan\beta/m_{H^\pm} = 0$ correspond to the SM expectations. The band indicates the 1σ range.

Using those values, we get the expectations for R and R^* for a given value of $\tan\beta/m_{H^\pm}$ in Fig. 2.1. Both values show destructive interference for lower values of $\tan\beta/m_{H^\pm}$.

But with larger values, the SM expectation is exceeded quite fast. For the model to be compatible with experiments, there must exist a value of $\tan\beta/m_{H^\pm}$, that provides compatible estimations for both R and R^*. In the region for destructive interference in R^* (for $\tan\beta/m_{H^\pm} < 0.6\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$) the value range for R is very wide. For higher values of R^* however, the value of R would have to be extremely high, which should have been visible in earlier measurements. So while a measured deviation only in R would restrict the value of $\tan\beta/m_{H^\pm}$, a significant excess in R^* could hardly be in compatibility with the 2HDM of type II. This is also shown in Fig. 2.2, that shows possible combinations of R and R^* for all values of $\tan\beta/m_{H^\pm}$ in the observed range. The values expand with x^4 for higher values of $\tan\beta/m_{H^\pm}$.

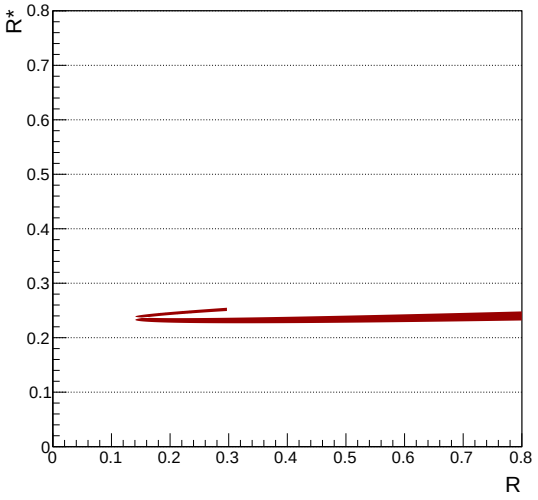


Figure 2.2: Possible value combinations of R and R^* in the 2HDM of type II, including 1σ bands. A variation of R^* to lower values is very limited. Higher values of R^* would require extremely high values of R ($\gtrsim 1.3$).

3. Experimental Setup

This chapter provides an overview of the experimental setup, which consists of the accelerator facility KEKB and the Belle detector with its various subsystems. It is described in great detail in [15].

3.1 The KEKB Accelerator

The data for this thesis was recorded with the Belle detector at the KEKB asymmetric e^+e^- collider, which was located at the KEK high energy research facility in Tsukuba, Japan. The KEKB was a B *factory* of the first generation: the former means, it was designed for the large-scale pair production of B mesons; and the latter, its main purpose was to establish CP violation in the B system and probe its parameters.

The production of B meson pairs is achieved by colliding e^+e^- with a CM energy of $\sqrt{s} = 10.58\,\mathrm{GeV}$. This corresponds to the mass of the $\Upsilon(4S)$ resonance, which is an excited bound $b\bar{b}$ quark state with quantum numbers $J^{PC} = 1^{--}$. It decays in most cases ($\approx 96\%$) to a pair of charged or neutral B mesons. Its rest mass is only about $20\,\mathrm{MeV\,c^{-2}}$ above the $B\overline{B}$ threshold, so there is no room for excitations or additional particles, which provides a very stable and well defined experimental environment.

The probing of CP violation requires excellent time resolution, which is achieved by colliding leptons with asymmetric energies: The high energy ring (HER) carries electrons at 8 GeV, while the low energy ring (LER) carries positrons at 3.5 GeV, resulting in a boost of the CM system of $\beta\gamma = 0.425$ in the laboratory frame in direction of the e^- velocity.

A schematic of the KEKB accelerator is given in Fig. 3.1, where the Belle detector is located at the collision point labeled as "Tsukuba". The storage rings have a circumference of $\approx 3\,\mathrm{km}$. The runtime of the accelerator was from late 1998 until 2010. Over this time, the original design luminosity of $1.0 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$ could be doubled. The recorded integrated luminosity at the $\Upsilon(4S)$ resonance amounts $711\,\mathrm{fb^{-1}}$. This is the data, that is used in this thesis. Furthermore, the accelerator was operated on other energies, including $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$, and a sample with an integrated luminosity of $121\,\mathrm{fb^{-1}}$ was recorded at the energy of the $\Upsilon(5S)$ resonance to study B_s pairs and decays of $\Upsilon(5S)$ to its lower excitation states.

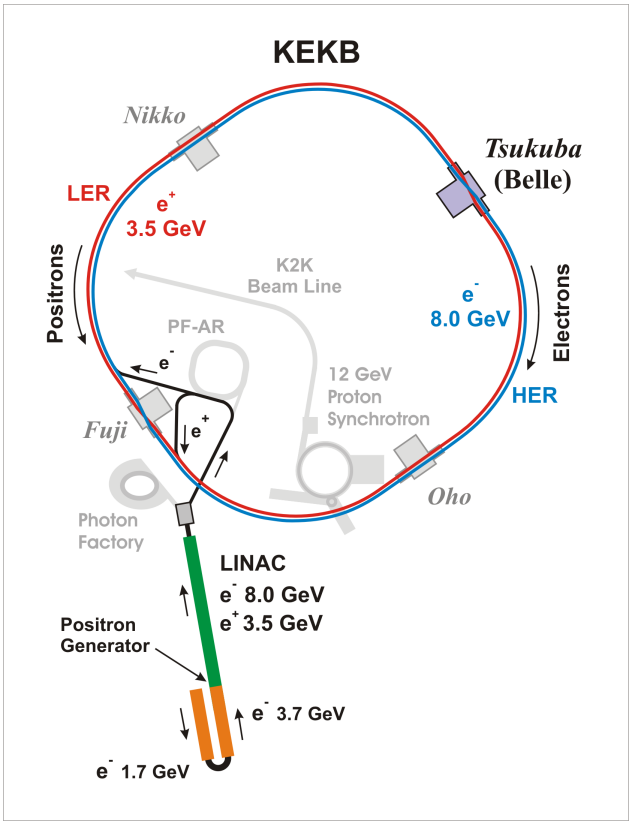


Figure 3.1: Schematic of the KEKB accelerator.

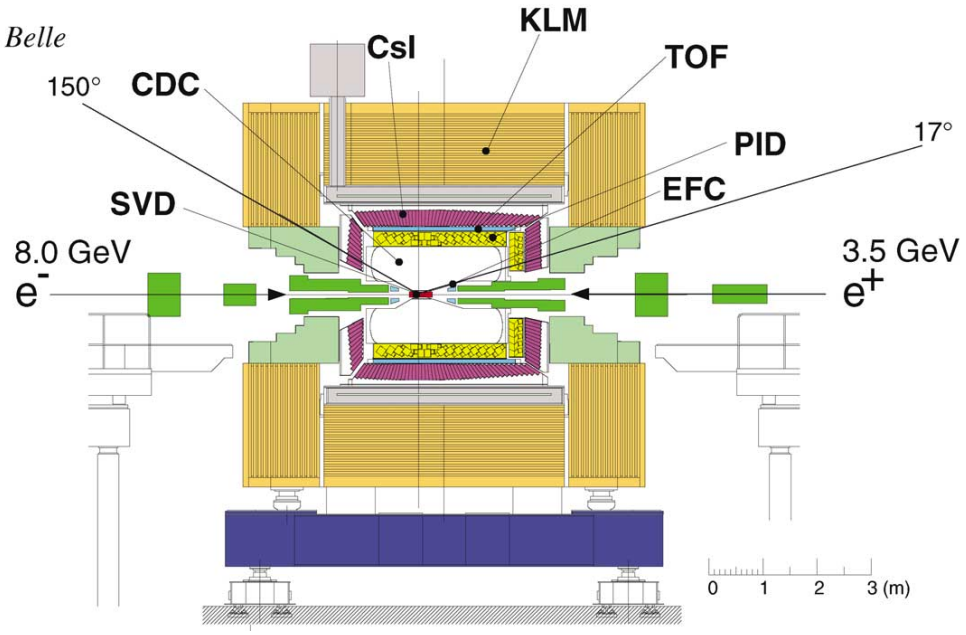


Figure 3.2: Side view of the Belle detector. Source: [15]

3.2 The Belle Detector

3.2.1 Overview

A side view of the Belle detector is shown in Fig. 3.2. It is a 4π multi-purpose detector, containing seven layered subsystems around the interaction point (IP). The innermost detector system around the beam pipe is the silicon vertex detector (SVD), that provides

the fine resolution for decay vertices and aids in the tracking of charged particles. The latter is one of the main tasks of the central wire drift chamber (CDC) in the layer surrounding the SVD. The CDC also provides measurements of energy loss by ionization, which is combined with information from the aerogel Cherenkov counters (ACC, labeled as PID in Fig. 3.2) and the time of flight counters (TOF) to the particle identification(PID). The electromagnetic calorimeter (ECL) records electromagnetic showers to measure particle energies. All these components are surrounded by a superconducting solenoid providing a magnetic field of 1.5 T. Arrays of resistive plate counters are interspersed in the iron yoke to identify K_L mesons and muons (KLM). The detector is slightly asymmetric to adapt to the boost of the CM system, and covers the θ region from 17° to 150° measured to the forward direction. Below and above this angular range there are extreme forward/backward calorimeters (EFC) installed, that provide photon and electron calorimetry at extreme angles, are used for beam and luminosity monitoring, and shield the CDC against beam radiation.

The coordinates in the detector are defined as follows: the x, y, and z axis together form a right-handed coordinate system, with the z axis in direction of the magnetic field, which is also the flight direction of the CM system. The x axis is aligned horizontally, the y axis vertically. In cylindrical coordinates, the r-ϕ plane is the projection on the xy plane, with the azimuthal angle ϕ measured from the positive x axis, and $r = \sqrt{x^2 + y^2}$. The polar angle θ is measured from the positive z axis. The origin of the coordinate system is the nominal IP. The beams are crossing at an angle of ± 11 mrad with a frequency of ≈ 509 MHz, which corresponds to a 2 ns interval.

3.2.2 SVD

The SVD is a crucial element for any time-dependent measurement. The decay length of a B meson in the laboratory system is translated to a decay time in the CM system by the boost of the beam configuration. So the achievable time resolution is limited by the vertex resolution, which is the special domain of the SVD. The average decay length difference between two B mesons is only $\approx 200\,\mu$m, so it should be capable to provide resolutions far below this value. Precise knowledge of vertices helps in any case with the proper reconstruction of particles. The Belle SVD uses several layers of double-sided silicon strip detectors (DSSD). They essentially provide depleted pn-junctions between the two sides. Their strips are aligned perpendicular, so an electron-hole pair, that is created by a passing charged particle, results in a signal on both sides (SVD hit), which provides 2D position information (usually z and ϕ coordinate), that is combined with the knowledge of the position of the DSSD itself.

Two configurations were used for the SVD: With an upgrade in 2003, the radius of the beam pipe was reduced and the old three-layer SVD (SVD1) was replaced by a four-layer SVD (SVD2), that could be mounted closer to the beam pipe, had a better angular coverage, and allowed the reconstruction of charged tracks from SVD hits only. The vertexing resolution improved by $\approx 25\%$. A schematic front- and side view of SVD1 is given in Fig. 3.3.

3.2.3 CDC

The central component to measure the trajectory of charged particles is the CDC. It fills the radial space around the beam pipe between 83 mm and 888 mm with alternating field- and sense wires in a mixture of helium and ethane (50%/50%). Charged particles ionize gas along their path, which is followed by avalanche processes due to the field between the wires, that cause a measurable electric pulse. Alternating layers of wires are oriented under a small angle (*axial* and *stereo* layers) to provide information of transverse and longitudinal trajectory.

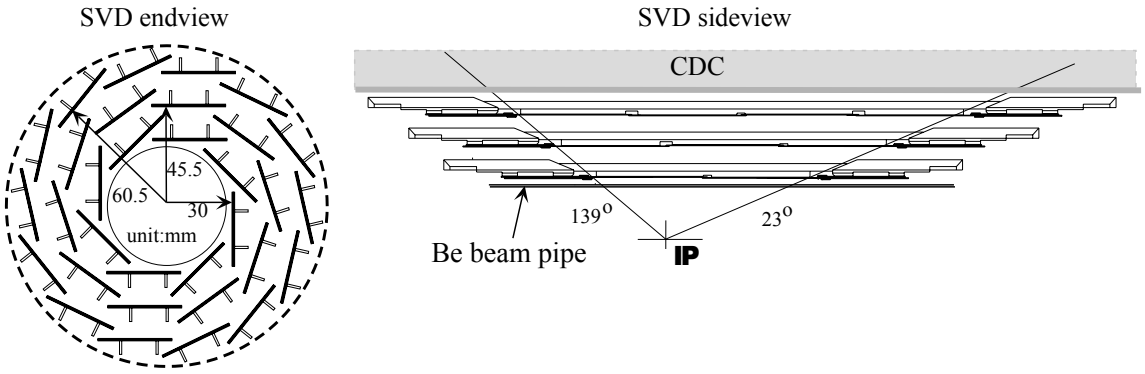


Figure 3.3: Front view (left) and side view (right) of SVD1. Source: [15]

There are several purposes for this information:

Tracking: This is the association of a series of traces in the detector to one particular particle. The CDC has the central role in this process, as it provides adequate drifting space for long series of ionization events without causing particles to change direction due to scattering processes.

Momentum: The paths of charged particles are curved in the CDC, as a 1.5 T magnetic field is maintained. The curvature is used to measure the particles' momenta with high accuracy.

Energy Loss: The strength of the electric pulses at the sense wires is correlated to the energy loss of the ionizing particle. According to the Bethe formula, the characteristic energy loss dE/dx depends on the speed of the charged particle relative to the ionized material. Along with momentum information, this is part of the particle identification (PID), that is explained in Section 3.2.4. The dependency of dE/dx from the particle three-momentum is shown in Fig. 3.4 for various charged particle types.

Trigger: Tracking information of the CDC is used by the trigger system explained in Section 3.2.8.

3.2.4 PID

The identification of final state particles is never unambiguous, as their rest masses can not be measured directly. Instead, the relation between velocity and three-momentum is exploited, to extract likelihoods for different identity hypotheses. This is done in several subcomponents of the Belle detector, as they provide differing accuracies, depending on the particle momentum.

Energy Loss in the CDC

As explained before, the relation of specific energy loss dE/dx due to ionization and particle momentum indirectly contains the velocity-momentum relation and is therefore characteristic for particles with differing rest masses.

ACC

The principle of the ACC is to measure bursts of Cherenkov light, that are emitted by passing particles above the medium's speed of light. The requirements for the material are transparency, low density to avoid energy loss, and a refractive index n, that for a given momentum range ($1\,\mathrm{GeV\,c^{-1}} < p < 4\,\mathrm{GeV\,c^{-1}}$) causes lighter particles ($\pi^{\pm}$) to emit Cherenkov light, while heavier particles ($K^{\pm}$, $p^{\pm}$) are too slow. The silica aerogel used in

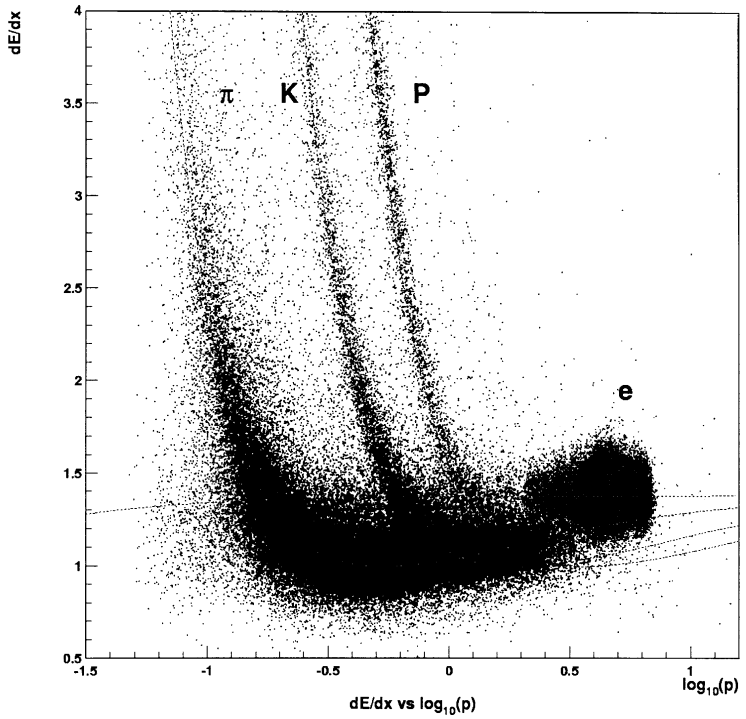


Figure 3.4: Relation of the specific energy loss dE/dx and particle momentum for different charged particles in the CDC. Source: [15]

the ACC provides these properties and is even produced with different refractive indices in the range from 1.01 to 1.03 for different polar angles. The arrangement in the Belle detector is shown in Fig. 3.5.

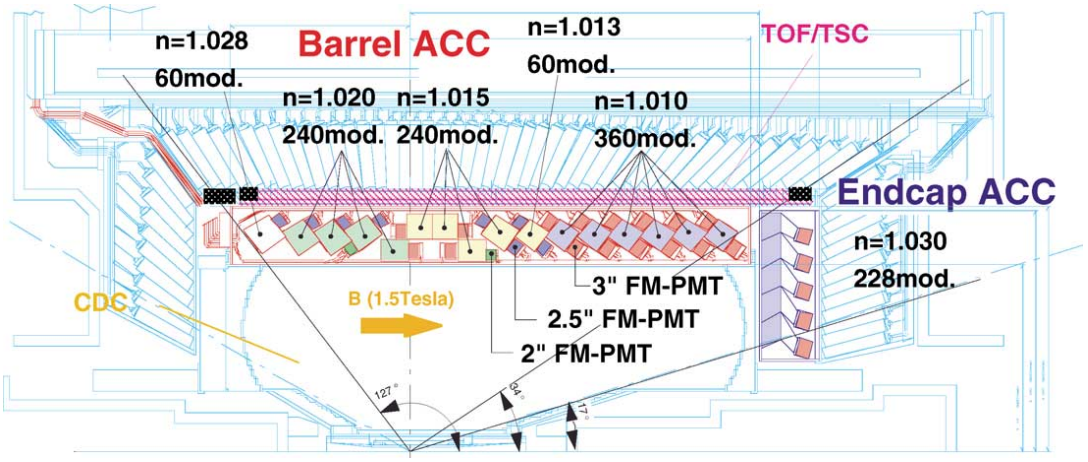


Figure 3.5: The Arrangement of the ACC units in the Belle detector with numbers of counter modules and diameters of the photomultiplier tubes. *FM-PMT* stands for *fine mesh-type photomultiplier tubes*. Source: [15]

TOF

The purpose of the TOF is to measure the time difference from the collision in the IP until a charged particle passes it. This time, combined with the measured momentum and flight path by SVD and CDC, allows the calculation of the particle mass and thereby separating mostly pions from kaons. The particles' momenta should be below $1.2\,\mathrm{GeV\,c^{-1}}$, which makes this subdetector complementary to the ACC. The TOF consist of 64 modules in the

barrel region, that contain plastic scintillator counters attached to photomultiplier tubes. It is also used by the trigger system in Section 3.2.8.

3.2.5 ECL

The ECL records the energy and position of electromagnetic showers, that are caused by cascades of bremsstrahlung and pair production in its material, which are mainly caused by photons and electrons. While electron energies can be measured indirectly by the previous components, there is no momentum or even track information for photons before they reach the ECL. So photons were the main focus in its design. They can come from quite different origins, like very low energetic in $D^* \to D\gamma$ or high energetic in $B \to K^*\gamma$ – and the calorimeter has to provide a good resolution in both cases.

The calorimeter consists of 8736 scintillating CsI(Tl) crystal cells, facing the IP with a small tilt to avoid photons passing undetected in the gap between the cells. Each cell has a length of 30 cm, corresponding to 16.2 radiation lengths, and a square front and back face with an edge length between 4.5 cm and 8.2 cm, depending on their polar angle position. The dimensions of the cells at different polar angles are designed to contain approximately 80% of the energy deposited in the crystal by a photon that enters at the center of the front face. Wider cells could improve this value and thereby energy resolution, but would decrease the position information and two-photon separation. The energy deposit is measured by two photo diodes at the center of each crystal's back face. In its entirety, the ECL covers the polar angle range between 12° and 155°. Its layout is displayed in Fig. 3.6.

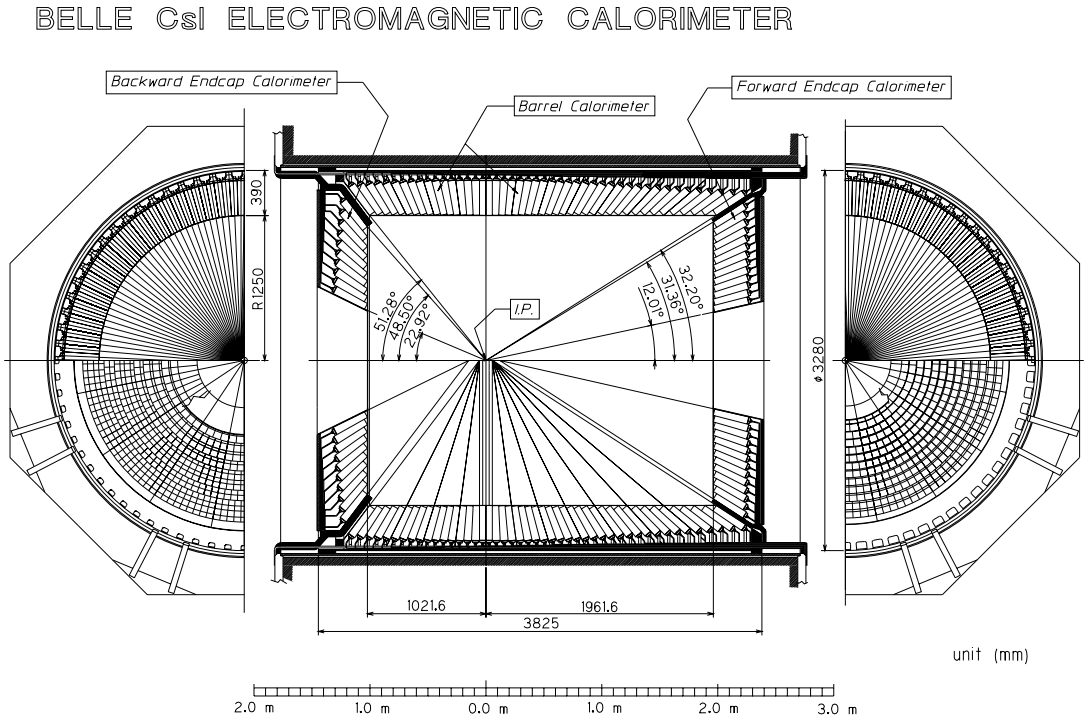


Figure 3.6: Layout of the cells of the electromagnetic calorimeter (ECL) in the barrel and endcap regions. Source: [15]

The ECL also supports the particle identification. The energy deposit per momentum is much higher for electrons than for other charged particles, which simplifies the electron identification.

3.2.6 KLM

The KLM provides identification for K_L^0 mesons and muons by measuring electromagnetic showers. It consists of alternating layers of the iron structure of the detector and resistive plate counters (RPC). While the ECL provides only 0.8 interaction lengths for K_L^0 mesons, the 14 iron layers of the KLM add 3.9 interaction lengths, which allows for measuring the K_L^0 meson's flight direction. However, the particle's energy can not be measured reliably, as the resolution is too low.

Most charged particles other than muons are shielded by the iron structure and don't deposit energy in the KLM. But if they reach the RPCs, they can be distinguished from muons by amount and form of their energy loss. Muons don't interact strongly and therefore travel farther without much transverse scattering in the KLM. As a result, the muon particle identification is by far the purest of any particle type with an efficiency well above 90% and less than 5% misidentification.

For the measurement, the KLM uses glass electrode RPCs, which are parallel electrodes of a high bulk resistivity ($\geq 10 \times 10^{10}\,\Omega\,\mathrm{cm}$), separated by a gas-filled gap. A passing charged particle ionizes the gas, which causes a discharge at the electrodes. Due to the high resistivity, this discharge is local. Perpendicular readout strips in θ and ϕ direction then measure its location. The layout of one RPC superlayer is shown in Fig. 3.7. Two RPCs separated by an insulator form one unit, to provide redundancy and a very high superlayer efficiency $\geqslant 98\%$.

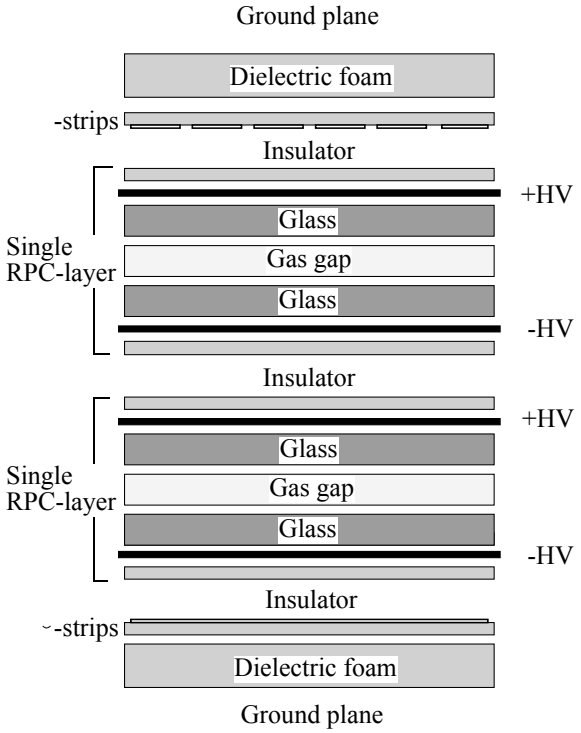


Figure 3.7: Side view of one RPC superlayer in the KLM subdetector. Source: [15]

3.2.7 EFC

The EFC covers extreme polar angles in forward ($6.4° < \theta < 11.5°$) and backward ($163.3° < \theta < 171.2°$) direction to extend the capabilities of the ECL concerning the detection and energy measurement of electrons and photons. Its second purpose is to shield the CDC against beam radiation. As components close to the beam pipe have to withstand high radiation levels, the EFC uses radiation-hard bismuth germanium oxide (BGO) crystals.

Photodiodes are used to measure the scintillation light. Additionally, the EFC is used for beam monitoring and luminosity measurements.

3.2.8 Triggers

The collision rate in the Belle detector is about 509 MHz. Only a small fraction of these spawn physically interesting events, that justify a full detector readout and data storage. The trigger system identifies these events. Interesting events are:

- $e^+e^- \to \Upsilon(XS)$ $(b\bar{b})$
- $e^+e^- \to q\bar{q}$ with $q\,\epsilon\,[u, d, s, c]$ (continuum)
- $e^+e^- \to \ell^+\ell^-$ with $\ell\,\epsilon\,[\mu, \tau]$
- Bhabha scattering
- $e^+e^- \to \gamma\gamma$
- 2γ processes with $p_t \geq 0.1\,\mathrm{GeV\,c^{-1}}$

Bhabha scattering and $e^+e^- \to \gamma\gamma$ are needed for calibration and luminosity monitoring, but are so numerous, that they get scaled down drastically, i.e. the trigger system discards them at a fixed rate. The total rate of remaining physics processes at design luminosity is about 100 Hz. There are also events created by beam background in the rate of about 120 Hz. As the detector was later operated at double the design luminosity, the physics event rate was doubled, while the beam background event rate increased by more than a factor of 2. The limit from the data acquisition system is a data rate of 500 Hz.

The trigger system consists of several stages. The Level-1 trigger receives real-time data from all subdetectors, except SVD, and scans them for a variety of criteria for different types of physics events. This process has to be very fast, with a latency of only 2.2 µs. The Level-1 trigger is therefore implemented mostly on programmable hardware. After this stage, there are still too many false-positives to maintain a good efficiency while writing all data to permanent storage. Therefore one more online trigger stage was added: the Level-3 trigger is implemented as software on an online computing farm. It performs fast track finding algorithms and can therefore veto events, that do not originate close enough to the IP. It provides nearly 100% efficiency for hadronic events, while the data size is cut in half. The events that pass this stage are written to permanent storage. The Level-4 trigger then processes the raw data on an offline computing farm. It applies some soft threshold requirements and provides basic reconstruction steps – like calculating track properties – as first abstraction layer for further processing.

4. Candidate Reconstruction

It is a long way from recording the detector response of individual collision events to a final data sample, that allows the interpretation of particles and observables. However, describing every detail in this chain is beyond the scope of this thesis. The various procedures that lead from recording hits and energy deposits in the detector to actual particle candidates are common for all Belle analyses and are therefore omitted. This chapter describes the various steps that are necessary for forming signal event candidates from said particle candidates.

4.1 General Strategy

As the branching ratio of $B \to D^{(*)}\tau\nu$ (called *signal*) is measured relative to $B \to D^{(*)}\ell\nu$ (called *normalization*), only leptonic τ decays are considered. We therefore lose by construction about 65% of the signal decays, where the τ decays via hadronic modes. As a result the visible final states of signal and normalization are the same: one $D^{(*)}$ meson and one electron or muon. The advantage of this approach is, that various sources for systematic uncertainties cancel out, as they affect the final state particles of signal and normalization in the same way.

We therefore have to distinguish the final states by their kinematics. While the distributions of the momenta of the $D^{(*)}$ meson and the lepton are slightly different for signal and normalization, the main difference comes from the number of neutrinos: the signal decay has three neutrinos in the final state – the normalization only one. As neutrinos usually pass the detector without leaving any traces, we call the undetected neutrinos the *invisible system*[1].

Lepton colliders have the advantage of a precise knowledge of the initial state, so it is possible to measure the properties of the invisible system in an indirect way by examining all visible particles in an event.

4.2 Tag Side Reconstruction

4.2.1 Kinematics at B Factories

The center-of-mass energy at B factories is optimized to produce an $\Upsilon(4S)$ resonance, which is an excited $b\bar{b}$ state, that produces in most cases ($> 96\%$) a $B\overline{B}$ meson pair. The

[1] This expression refers to all undetected particles in an event, so its meaning also comprises other particles in the case of certain background components.

phase space for this decay is rather small, so no additional particles are produced in the immediate decay and the B mesons have low momenta. As the B mesons are also equal in mass, it is a very simple two-body decay, that allows for strong kinematic constraints: A measurement of the properties of one of the B mesons in any available decay chain automatically provides flavor and momentum information of the other B meson. To get the most information out of this *tag* B meson, hadronic decay modes are preferred, as they allow to measure the four-momenta without missing particles like neutrinos. This is especially useful in our decay, as we can calculate the four-momentum of the invisible system by subtracting the four momenta of the tag B meson and the $D\ell$ system of the signal B meson from the known four-momentum of the $\Upsilon(4S)$.

As the initial state is very well known and the two-body decay is kinematically strongly constrained, the knowledge of the four-momentum of a B meson allows for powerful plausibility checks:

- In the center-of-mass system each B meson carries the *beam energy*

  $$E_{\text{beam}} \equiv \frac{1}{2} E_{\text{CMS},e^+e^-}.$$

  The energy-momentum relation of a B meson is then transformed to

  $$M_{\text{bc}} \equiv M_B = \sqrt{E_{\text{beam}}^2 - p_B^2}.$$

  This quantity is called *beam-constrained mass* and provides an indirect measurement of the particle mass based on measured momentum, which is much more precise than a direct measurement which also requires the reconstructed energy. A clear deviation from the nominal mass strongly indicates missed momentum by wrongly identified particles, missed particles, or wrong combinations.

- The directly measured energy of a fully reconstructed B meson should only differ from the beam energy by resolution effects. This difference ΔE is a powerful indicator for missing massive particles.

4.2.2 Full Reconstruction Framework

The Belle collaboration uses the `ekpfullrecon` framework [16] for most analyses that require a hadronic tag B meson. The B mesons from an $\Upsilon(4S)$ decay have opposing flavor and momentum, but are otherwise independent in their decays, so the tag side reconstruction can be used universally. The main idea of the `ekpfullrecon` framework is to implement as many decay channels as possible, while avoiding exponential growth in computing time due to exploding combinatorics.

4.2.2.1 Reconstruction Principle

Individual analyses usually reconstruct only certain final states. But when a huge variety of similar channels is to be reconstructed, many intermediate states overlap. The decay modes are therefore implemented in a hierarchical way: beginning from the lightest particles, all available particles are iteratively combined to heavier particles in all available decay modes, until a B meson is reached. A schematic of this principle is shown in Fig. 4.1.

With this procedure immense calculation efforts due to combinatorics can still occur, but are suppressed by adding quality requirements in every step of the particle combination. Using neural networks, the quality of each particle candidate is estimated by combining various properties, including the qualities of its children. “Bad” candidates are discarded and the cut-off for the quality depends on the purity of a combination mode, so the computing effort per good candidate is optimized.

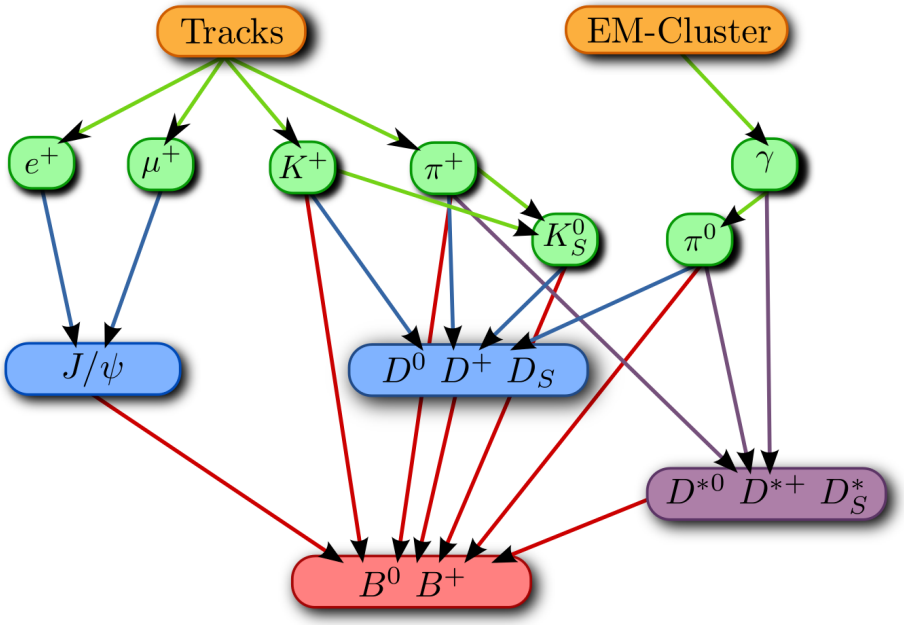


Figure 4.1: Schematic of the hierarchical reconstruction, starting with the lightest particles. Available particles are consecutively combined until a B meson is reached.

4.2.2.2 Decay Channels

The B meson reconstruction channels are listed in Table 4.1, those for intermediate mesons in Table 4.2. The total number of individual reconstruction channels is 1,149. Except for the J/ψ modes, all channels are fully hadronic, so the four-momentum of the tag B meson can be directly reconstructed.

B^+ modes	B^0 modes
$\overline{D}^0\pi^+$	$D^-\pi^+$
$\overline{D}^0\pi^+\pi^0$	$D^-\pi^+\pi^0$
$\overline{D}^0\pi^+\pi^+\pi^-$	$D^-\pi^+\pi^-\pi^+$
$D^-\pi^+\pi^+$	$\overline{D}^0\pi^0$
$\overline{D}^{*0}\pi^+$	$D^{*-}\pi^+$
$\overline{D}^{*0}\pi^+\pi^0$	$D^{*-}\pi^+\pi^0$
$\overline{D}^{*0}\pi^+\pi^+\pi^-$	$D^{*-}\pi^+\pi^-\pi^+$
$\overline{D}^{*0}\pi^+\pi^+\pi^-\pi^0$	$D^{*-}\pi^+\pi^+\pi^-\pi^0$
$\overline{D}^0 D_s^+$	$D^- D_s^+$
$\overline{D}^{*0} D_s^+$	$D^- D_s^{*+}$
$\overline{D}^0 D_s^{*+}$	$D^{*-} D_s^+$
$\overline{D}^{*0} D_s^{*+}$	$D^{*-} D_s^{*+}$
$\overline{D}^0 K^+$	$J/\psi K^0_S$
$J/\psi K^+$	$J/\psi K^+\pi^-$
$J/\psi K^+\pi^+\pi^-$	$J/\psi K^0_S\pi^+\pi^-$
$J/\psi K^+\pi^0$	
$J/\psi K^0_S\pi^+$	

Table 4.1: Tag side B meson reconstruction channels.

D^+ modes	D^0 modes	D^{*+} modes	D^{*0} modes
$K^-\pi^+\pi^+$	$K^-\pi^+$	$D^0\pi^+$	$D^0\pi^0$
$K^0_S\pi^+$	$K^-\pi^+\pi^+\pi^-$	$D^+\pi^0$	$D^0\gamma$
$K^0_S\pi^+\pi^0$	$K^-\pi^+\pi^0$		
$K^-\pi^+\pi^+\pi^0$	$\pi^+\pi^-$		
$K^0_S\pi^+\pi^+\pi^-$	$\pi^+\pi^-\pi^0$		
$K^+K^-\pi^+$	$K^0_S\pi^0$		
$K^+K^-\pi^+\pi^0$	$K^0_S\pi^+\pi^-$		
	$K^0_S\pi^+\pi^-\pi^0$		
	K^+K^-		
	$K^+K^-K^0_S$		

D_s^+ modes	D_s^{*+} modes	J/ψ modes
$K^+K^0_S$	$D_s^+\gamma$	e^+e^-
$K^+\pi^+\pi^-$		$\mu^+\mu^-$
$K^+K^-\pi^+$		
$K^+K^-\pi^+\pi^0$		
$K^+K^0_S\pi^+\pi^-$		
$K^-K^0_S\pi^+\pi^+$		
$K^+K^-\pi^+\pi^+\pi^-$		
$\pi^+\pi^+\pi^-$		

Table 4.2: Reconstruction channels for intermediate states in the decay of the tag side B meson.

4.2.2.3 Performance

The Efficiency of the `ekpfullrecon` module is about 0.18% for B^0 mesons and 0.28% for B^+ mesons in the best case. This gets slightly lowered, as further cuts on M_{bc}, ΔE and the quality indicating network output of the tag side B meson are a reasonable choice in many analyses. If there are several B_{tag} candidates in one event, the choice is postponed until after the signal side selection.

The benefit of having the tag-side B meson reconstructed entirely is not only the constraint on the four-momentum of the signal-side B meson, but also the knowledge, that *all* remaining tracks in the event must belong to it – if they have their physical origin in the e^+e^- collision. This will later be used for additional selection criteria.

4.2.2.4 Continuum Suppression

The class of events, that do not origin in $e^+e^- \to \Upsilon(4S)$ but in the production of lighter quarks (*udcs*) from the e^+e^- annihilation are called *continuum* events. They differ from $B\bar{B}$ events in their topology: Due to the much higher phase-space, the high momenta of the involved particles produce rather jet-like structures in the detector, while $B\bar{B}$ events are more spherical. This is exploited in Fox-Wolfram-Moments [17] (FWM), that describe the event topology. For this analysis the `ekpcontsuppress` module is used, that involves super Fox-Wolfram-Moments [18] (SFWM) in the calculation of the B_{tag} meson quality. The SFWM calculate the FWM not for the event in its entirety, but separately for the B_{tag} candidate and the remainder of the event. By cutting on the quality of the B_{tag} meson candidate, continuum background is no longer an issue in this analysis.

4.3 Candidate Combination

For the measurement of R and R^* we need to reconstruct the normalization modes $B \to D^{(*)}\ell\nu_\ell$ as well as the signal modes $B \to D^{(*)}\tau\nu_\tau$. Due to the exclusively leptonic reconstruction of the τ and the invisibility of the neutrinos, this is done intrinsically when reconstructing B mesons in $D^{(*)}\ell$. The exclusive signal side reconstruction modes are listed in Table 4.3.

The reconstruction of the signal side is also done hierarchically, but the D and D^* meson reconstruction modes (see Table 4.4) are not the same as in the `ekpfullrecon`. Some channels provide a signal-to-background ratio, that is acceptable for the tag side reconstruction, but causes too much pollution on the signal side. Most of these channels would contribute only a small fraction of the signal, but there are exceptions like $D^+ \to K^-\pi^+\pi^+\pi^0$, which could contribute to the signal events in the order of 15% and still dampens the expected precision of the analysis. Besides low branching fraction, the multiplicity especially in connection with π^0 greatly influences the purity in each reconstruction mode.

The signal side reconstruction is attempted in all events, that have a fully reconstructed tag side candidate. A signal side candidate is only valid if his final state particles do not overlap and at least one tag side candidate exists without overlapping final states. It is possible, that there exist several tag side candidates for one signal side candidate. In that case, the tag side candidate with the largest number of charged tracks is chosen, as having unused charged tracks in an event would cause the event to be pruned at a later selection stage. If there exist several compatible tag side candidates with the same number of charged tracks, the one with the better neural network output is chosen.

B^0 modes	B^+ modes
D^-e^+	$\overline{D}^0e^+$
$D^-\mu^+$	$\overline{D}^0\mu^+$
$D^{*-}e^+$	$\overline{D}^{*0}e^+$
$D^{*-}\mu^+$	$\overline{D}^{*0}\mu^+$

Table 4.3: B meson decay channels used for the signal side reconstruction.

D^+ modes	D^0 modes	D^{*+} modes	D^{*0} modes
$K^-\pi^+\pi^+$	$K^-\pi^+$	$D^0\pi^+$	$D^0\pi^0$
$K^0_S\pi^+$	$K^-\pi^+\pi^+\pi^-$	$D^+\pi^0$	$D^0\gamma$
$K^0_S\pi^+\pi^0$	$K^-\pi^+\pi^0$		
$K^0_S\pi^+\pi^+\pi^-$	$K^0_S\pi^0$		
	$K^0_S\pi^+\pi^-$		

Table 4.4: D and D^* meson decay channels used for the signal side reconstruction.

4.4 D^{**} Enriched Sample

As will be explained in Section 5.3.1.3, background from $B \to D^{**}\ell\nu$ decays has a key role in this analysis, and a dedicated sample of signal MC was created to improve its modeling. However, it is necessary to test the validity of this simulated sample by comparing its properties to real data. Unfortunately, the yield in the default reconstruction modes is far from being sufficient for a reliable comparison.

Under the assumption, that most $B \to D^{**}\ell\nu$ background comes from missing a π^0 in a decay $D^{**} \to D^{(*)}\pi^0$, we can investigate the relevant component of these events by explicitly reconstructing an additional π^0 to our default signal side reconstruction modes. The exclusive channels are listed in Table 4.5. This is the D^{**} *enriched sample*, and it will be used in Section 10.5 to validate the D^{**} background description. The assumption of the dominance of missed neutral tracks has been verified with a dedicated Monte Carlo sample and is quantified in [6].

B^0 modes	B^+ modes
$D^- e^+ \pi^0$	$\overline{D}^0 e^+ \pi^0$
$D^- \mu^+ \pi^0$	$\overline{D}^0 \mu^+ \pi^0$
$D^{*-} e^+ \pi^0$	$\overline{D}^{*0} e^+ \pi^0$
$D^{*-} \mu^+ \pi^0$	$\overline{D}^{*0} \mu^+ \pi^0$

Table 4.5: B meson decay channels used for the D^{**} enriched sample.

4.5 Reconstruction Cuts

To keep the number of candidates manageable, a set of soft cuts was applied during the initial signal side reconstruction to exclude the most obvious false positive matches. This allows an easier handling of the dataset and therefore further optimizations. Yet, the cuts are still loose enough to allow the extraction of sideband samples:

1. Sum of the electric charges of B_{tag} and B_{sig} must be 0. Two B^0 mesons with the same flavor are allowed due to mixing.
2. The final state particles of the signal side must be consistent with at least one tag side.
3. $M_{\text{bc,tag}} > 5.22\,\text{GeV}\,\text{c}^{-2}$
4. $-0.05\,\text{GeV} < \Delta E_{\text{tag}} < 0.10\,\text{GeV}$
5. Less than 20 photon candidates in the event to exclude events with enormous combinatorics
6. Particle identification (PID) quality requirements:
   - $e^\pm$ ID > 0.1
   - $\pi^\pm$ ID > 0.1
   - $\mu^\pm$ ID > 0.1
   - $K^\pm$ ID > 0.1
7. `goodKs`$== 1$,
8. $dr < 2.0\,\text{cm}$ and $dz < 4.0\,\text{cm}$ for charged tracks.
9. $E(\gamma) > 50\,\text{MeV}$
10. $E(\gamma \text{ from } \pi^0) > 50/80/80\,\text{MeV}$ (Barrel/Forward/Backward)
11. $\left|M_{\text{trafo}}(\pi^0)\right| < 3.0$
12. $|\vec{p^*}(D^{0/+})| < 3.0\,\text{GeV}\,\text{c}^{-1}$
13. $M(D)$ within a $160\,\text{MeV}\,\text{c}^{-2}$ window of the nominal mass except for the modes below:
    - $1.65\,\text{GeV}\,\text{c}^{-2} < M(D) < 2.15\,\text{GeV}\,\text{c}^{-2}$ for $D^+ \to K^0_S \pi^+ \pi^0$

- $1.70\,\mathrm{GeV\,c^{-2}} < M(D) < 2.10\,\mathrm{GeV\,c^{-2}}$ for $D^0 \to K^-\pi^+\pi^0$
- $1.60\,\mathrm{GeV\,c^{-2}} < M(D) < 2.10\,\mathrm{GeV\,c^{-2}}$ for $D^0 \to K^0_S\pi^0$

14. $\Delta M(D^*, D)$ requirements:
    - $0.125\,\mathrm{GeV\,c^{-2}} < \Delta M(D^*, D) < 0.160\,\mathrm{GeV\,c^{-2}}$ for $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$
    - $0.122\,\mathrm{GeV\,c^{-2}} < \Delta M(D^*, D) < 0.162\,\mathrm{GeV\,c^{-2}}$ for $D^{*0} \to D^0\pi^0$
    - $0.100\,\mathrm{GeV\,c^{-2}} < \Delta M(D^*, D) < 0.190\,\mathrm{GeV\,c^{-2}}$ for $D^{*0} \to D^0\gamma$
15. Not more than 1 remaining track with $dr < 2.0\,\mathrm{cm}$ and $dz < 4.0\,\mathrm{cm}$

Some of these cuts (Items 3, 4, 13 and 14) are chosen very wide on purpose to extract sideband data and will be narrowed down later. This is also the reason for different mass ranges of various D and D^* modes, which differ in mass resolution.

The `goodKs` property in Item 7 is a binary flag, based on quality cuts on the K_S candidate and its daughters. It is a standard cut in analyses of the Belle collaboration.

The cut on the energy of photons from π^0 candidates (Item 10) depends on their angle, as the forward- and backward regions of the detector receive a much higher beam noise level than the barrel region.

The transformed π^0 mass, that is used for the cut in Item 11, is calculated from the deviation of the reconstructed mass from the nominal π^0 mass, divided by the mass resolution. The mass resolution itself is calculated from the angle and momentum of the π^0 candidate. It varies between $5\,\mathrm{MeV\,c^{-2}}$ and $12\,\mathrm{MeV\,c^{-2}}$. So the transformed mass is dimensionless and gives the difference to the nominal mass in standard deviations. As false π^0 candidates are a major issue in this analysis, the cut is chosen rather strict. The calculation of the mass resolution is described in detail in [19] and [20].

4.6 Analysis Cuts

In order to proceed with the analysis, the sample must be further cleared of background events. This is done by optimizing existing cuts and introducing new ones, that are more specific to this analysis:

1. tag side channel dependent $o_{\mathrm{NB,tag\ side}}$ cut, see Table 4.6
2. $5.286\,\mathrm{GeV\,c^{-2}} > M_{\mathrm{bc,tag}} > 5.274\,\mathrm{GeV\,c^{-2}}$
3. $|\Delta E_{\mathrm{tag}}| < 0.05\,\mathrm{GeV}$
4. Particle identification (PID) quality requirements:
    - $e^\pm$ ID > 0.6
    - $\mu^\pm$ ID > 0.9
    - $\pi^\pm$ ID > 0.1
    - $K^\pm$ ID > 0.1
    - $e^\pm$ ID for $\pi^\pm$ and $K^\pm < 0.9$
    - $\mu^\pm$ ID for $\pi^\pm$ and $K^\pm < 0.9$
5. $p^*(\pi^0) > 200\,\mathrm{MeV\,c^{-1}}$ except for slow π^0 from D^*
6. $q^2 > 4.0\,\mathrm{GeV}^2$
7. $-0.2\,\mathrm{GeV^2\,c^{-4}} < M^2_{\mathrm{miss}} < 8.0\,\mathrm{GeV^2\,c^{-4}}$
8. No remaining tracks with $dr < 2.0\,\mathrm{cm}$ and $dz < 4.0\,\mathrm{cm}$

9. No remaining π^0 particles with $E(\gamma) > 50/100/150\,\text{MeV}$ (Barrel/Forward/Backward) and $\left|M_{\text{trafo}}(\pi^0)\right| < 3$

Item 1 is a series of cuts on the quality of the tag side B meson as given by the `ekpfullrecon` framework. They have been optimized by defining a signal region in the dataset using strict cuts on the squared missing mass and $D^{(*)}$ masses, and optimizing the significance ($N_{\text{signal}}/\sqrt{N_{\text{signal}} + N_{\text{background}}}$) of the $D\tau\nu$ signal in this region. The cuts preserve about 85% of the τ signal, depending on the individual channel. The cuts on $M_{\text{bc,tag}}$ and ΔE_{tag} were optimized consecutively in the same way. This procedure was applied after correcting tag side efficiencies on MC, described in Section 7.1.

The cuts on the particle identification (Item 4) are much stricter now. Muons are by far the easiest to identify due to their distinct signature, so their PID requirements are the strictest. Requirements for charged kaons and pions are still weak, as they are the most frequent particles in the detector. However, there is now a requirement for the electron- and muon PID of charged kaon and pion candidates, that favors high-quality lepton hypotheses.

Item 5 sets a lower limit to momenta of π^0 candidates coming from D decays in the center-of-mass system of the event. This must not apply to π^0 candidates from D^* decays, as their momentum is limited by phase-space.

The four-momentum transfer q^2 from the signal side B to the $D^{(*)}$ meson depends highly on the lepton mass. The kinematic boundaries in $B \to D^{(*)}\tau\nu_\tau$ are significantly higher than in $B \to D^{(*)}\ell\nu_\ell$, as shown in Fig. 4.2. Therefore, the cut in Item 6 is used to suppress $B \to D^{(*)}\ell\nu_\ell$ decays, that would otherwise overwhelm the final sample.

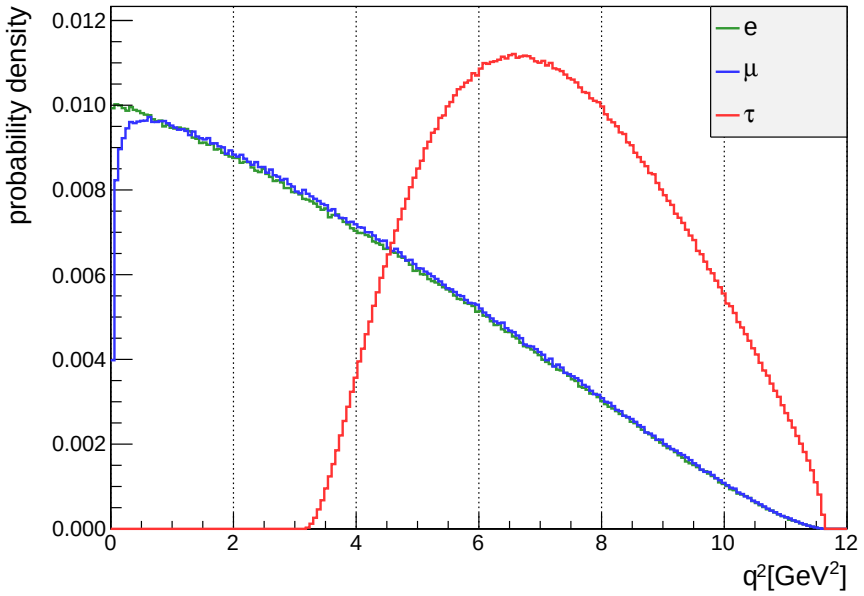


Figure 4.2: Distributions of q^2 in $B^0 \to D^-\ell^+\nu_\ell$, which strongly depend on the mass of the lepton ℓ. Events below $q^2 = 4.0\,\text{GeV}^2$ are excluded from this analysis to suppress approximately half of the normalization events.

The cut on the squared missing mass M^2_{miss} in Item 7 is performed to avoid numeric problems later in the analysis. Events below $M^2_{\text{miss}} = -0.2\,\text{GeV}^2\,\text{c}^{-4}$ are scarce and using them to create probability density functions may introduce trouble with zero-content bins or badly described tails. As they neither contain any signal events, nor improve any background description, it is beneficial to prune them.

Items 8 and 9 improve the quality of the event by requiring, that no detectable particle is missed, that originates from the physical processes in the collision. This would always mean, that the physical process, that actually happened, is different from the one that we assumed in the reconstruction. Unused tracks that originate from the interaction point always veto the candidate. For π^0 candidates, which decay to two photons, the position information is far worse and the probability for random combinations from beam background is much higher. Therefore the veto is weaker and only applies, if the π^0 meets the given quality requirements concerning its reconstructed mass and the energy of its decay products.

Channel	Cut	Channel	Cut
$B^- \to D^{*0}\pi^-$	-3.72	$B^0 \to D^{*+}\pi^-$	-2.40
$B^- \to D^{*0}\pi^-\pi^0$	-5.40	$B^0 \to D^{*+}\pi^-\pi^0$	-5.28
$B^- \to D^{*0}\pi^-\pi^-\pi^+$	-6.12	$B^0 \to D^{*+}\pi^-\pi^+\pi^-$	-4.92
$B^- \to D^0\pi^-$	-3.60	$B^0 \to D^+\pi^-$	-3.24
$B^- \to D^0\pi^-\pi^0$	-5.04	$B^0 \to D^+\pi^-\pi^0$	-4.80
$B^- \to D^0\pi^-\pi^-\pi^+$	-6.12	$B^0 \to D^+\pi^-\pi^+\pi^-$	-6.24
$B^- \to D^{*0}D_s^{*-}$	-5.76	$B^0 \to D^{*+}D_s^{*-}$	-6.12
$B^- \to D^{*0}D_s^-$	-6.36	$B^0 \to D^{*+}D_s^-$	-4.68
$B^- \to D^0D_s^{*-}$	-6.48	$B^0 \to D^+D_s^{*-}$	-4.80
$B^- \to D^0D_s^-$	-7.32	$B^0 \to D^+D_s^-$	-6.24
$B^- \to J/\psi K^-$	-0.96	$B^0 \to J/\psi K_s^0$	-0.48
$B^- \to J/\psi K^-\pi^+\pi^-$	-2.88	$B^0 \to J/\psi K^-\pi^+$	-0.60
$B^- \to D^0K^-$	-5.40	$B^0 \to J/\psi K_s^0\pi^+\pi^-$	-4.32
$B^- \to D^+\pi^-\pi^-$	-5.52	$B^0 \to D^0\pi^0$	-5.40
$B^- \to D^{*0}\pi^-\pi^-\pi^+\pi^0$	-7.20	$B^0 \to D^{*+}\pi^-\pi^-\pi^+\pi^0$	-6.24
$B^- \to J/\psi K^-\pi^0$	-5.76		
$B^- \to J/\psi K_S^0\pi^-$	-1.56		

Table 4.6: Analysis cuts on the tag side B meson network output $\log(o_{\text{NB,tag side}})$.

4.7 Best Candidate Selection

If there is more than one signal candidate in an event, one has to be picked. This is done right before the fitting procedure, after all corrections and additional cuts (see Chapter 7). There exist many methods to choose the "best" candidate in an event, to maximize signal efficiency. However, these usually involve the usage of kinematic properties and/or remaining energy in the detector. As we will make use of several event properties of this type in the fitting procedure, they have the potential to induce a bias from flaws in the efficiency estimation. We therefore do not try to select the best candidate in an event, but select one at random. The disadvantage in efficiency compared to a real best candidate selection is only in the order of 5%.

5. Data Samples

One of the main difficulties of this analysis is the heavy smearing of the kinematics of each event due to several "invisible" particles in the final state. In other words: There is no narrow peak, which could be described by trivial functions, as shown in Fig. 5.1. It is therefore not enough to reconstruct our signature in recorded data. Simulated data is required additionally to examine various physical components. The set of recorded data is described in Section 5.2, while Section 5.3 describes the various sets of simulated data. The methods used to create them is described in Section 5.4.

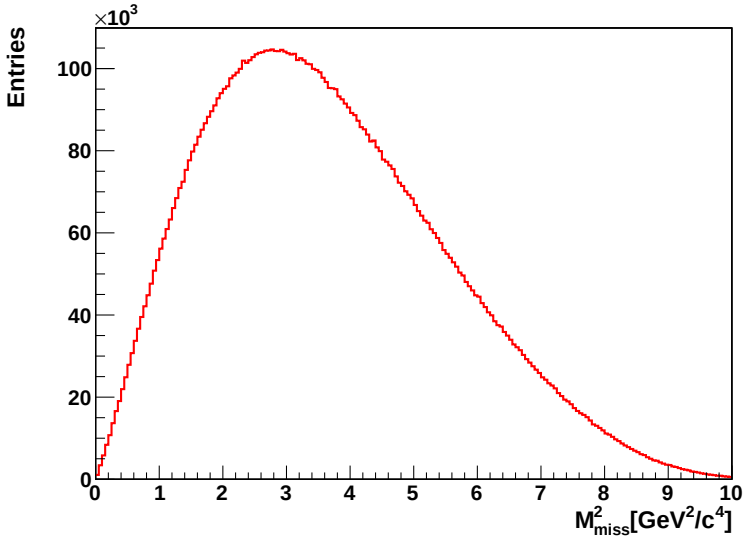


Figure 5.1: Distribution of the generated squared invariant mass of the neutrino system in the decay $\bar{B}^0 \to D^+\tau^-(\to \ell^-\bar{\nu}_\ell\nu_\tau)\bar{\nu}_\tau$. In the case of $\bar{B}^0 \to D^+\ell^-\bar{\nu}_\ell$ a sharp peak at $M^2_{\text{miss}} = 0\,\text{GeV}^2\,\text{c}^{-4}$ is generated. The wide distribution is further complicated by efficiency effects in the reconstruction process.

5.1 Four Reconstruction Modes

Until now we have considered the reconstructed $D^{(*)}\ell$ datasets as a whole, with the general task to separate signal from normalization modes. To properly access the kinematic properties of various components in the dataset, we have to handle the reconstruction modes

separately. While the e and μ modes are very similar, this is not the case for neutral and charged $D^{(*)}$ mesons. The biggest difference in characteristics however is between excited and non-excited D mesons. We therefore treat from now on four reconstruction modes separately: $(D^0\ell^-)$, $(D^{*0}\ell^-)$, $(D^+\ell^-)$, and $(D^{*+}\ell^-)$.

5.2 Real Data

The final parameter estimation of this analysis is done on the full available dataset of the Belle experiment, taken at the $\Upsilon(4S)$ resonance between 1999 and 2010. These 711 fb^{-1} of data correspond to $\approx 772 \times 10^6$ $B\bar{B}$ pairs. Other event types in the dataset are $q\bar{q}$ pairs of up, down, charm, or strange quarks produced in the e^+e^- collision. There are other sets of recorded data, but these were not used in this analysis. Those contain events at energies of $\Upsilon(XS)$ resonances with $X\epsilon\{1, 2, 3, 5\}$, or slightly below the $\Upsilon(4S)$ threshold to study continuum events (explained in Section 5.3.1.1).

5.3 Simulated Data

The estimation procedure requires knowledge about various components, that can not be extracted from the existing data sample itself with sufficient precision. It is therefore necessary to use simulations (called *Monte Carlo* or just *MC*) based on best knowledge of the physical processes in the collision events.

5.3.1 Types of Simulated Data

There are various reasons to use simulated data and they usually have different requirements. So individual sets of simulated data are produced with different focuses:

5.3.1.1 Generic Monte Carlo

This dataset tries to imitate the real data set. This is achieved by generating *streams* of data, that correspond to the total integrated luminosity of the real data and also imitate its composition and decay modes. The streams are structured by their different initial states: besides B^+B^- and $B^0\bar{B}^0$ pairs, there are *continuum* events, where e^+e^- annihilates either to light quarks (*uds* Monte Carlo) or $c\bar{c}$ pairs (*charm* Monte Carlo). For most parts of this analysis five streams of generic MC are used.

The strength of this dataset is, that after processing it in an identical way to the real data, it provides an easy way to estimate efficiencies and distributions of well known components of the data. The downside is, that it must not be trusted blindly, as it has many known imperfections, e.g. outdated decay parameters or flaws in the detector simulation. These must be considered, tested, and corrected before the extracted quantities can be considered reliable, which will be the topic of Chapter 7.

5.3.1.2 Signal Monte Carlo

For some purposes quantities that are extracted from generic Monte Carlo are not precise enough. The most prominent cause is insufficient statistics, but there are other possibilities: Due to high computing resource consumption, the generic Monte Carlo streams were produced only once for all analyses, so production parameters that become outdated over time need to be corrected with specialized data samples. In some cases, branching ratios to certain final states need to be extracted, which can be difficult on generic Monte Carlo, as intermediate resonances can "hide" possible decay chains.

In this analysis the most important sample of signal Monte Carlo is $B \to D^{(*)}\tau\nu$ with increased statistics to extract required distributions. A sample of $B \to D^{(*)}(e/\mu)\nu$ was

also produced to cross-check the efficiencies of the most important components. Corrections on the samples of $B \to D^{(*(*))}\ell\nu$ (see Sections 7.3 and 7.4) were necessary to account for an updated decay-model and were based on accordingly produced signal Monte Carlo. One special case, that is of great importance for this analysis, is the $B \to D^{**}\ell\nu$ Monte Carlo, which is explained in the following section.

5.3.1.3 $B \to D^{**}\ell\nu$ Monte Carlo

One of the most important background components in this analysis is $B \to D^{**}\ell\nu$. If the π^0 coming from $D^{**} \to D^{(*)}\pi^0$ decays is missed – which is not unlikely for particles with low momentum and neutral decay products – the invisible system consists of a neutrino and this π^0, which kinematically resembles our signal decay closely. So a good description of these decays is one of the cornerstones of this analysis.

The generic MC however has several flaws concerning this type of B meson decay:

- Branching ratios of the various D^{**} meson types (D_2^*, D_0^*, D_1, D_1', and the radial excitations $D(2S)$ and $D^*(2S)$, each one occurs charged and neutral) are outdated: Some have been measured more precisely, others had their expectations updated from theoretical constraints or inclusive measurements.
- Only decays to $D^{(*)}\pi$ are allowed: Decays with two or more π are physically possible and in some cases even allow intermediate resonances. But they are not included in the generic Monte Carlo samples.

In the fit we do not rely on the total number of generated $B \to D^{**}\ell\nu$ events. However, the composition of this component heavily influences its overall shape, that is of great interest. So it can make an important difference, whether a specific D^{**} mode has a realistic frequency relative to all others. For this reason new decay tables were created [21] and used to generate $B\bar{B}$ events where one B decays generically and the other one into a $D^{**}\ell\nu$ mode. These events were used to replace all corresponding events in the generic MC entirely.

5.4 Event Generation

The generation of simulated data is performed in two steps: First, the physical event is created using the `EvtGen`[22] package. This is done by simulating the full decay chain, taking into account the properties of intermediate states. Then the response of the detector to the propagating particles is simulated using the `GEANT3`[23] framework.

5.4.1 Decay Simulation

To create a decay chain with the `EvtGen` package, it is necessary to define three components:

Particle properties are defined in *particle definition lists*, that connect a particle name to its properties like mass, width, charge, or spin. As these rarely change, a large default particle definition list is usually used.

Decay models are independent modules, that handle the probability calculations in phase space for transitions from a given particle to given decay products. For simple decays which are only restricted by energy- and momentum conservation, a simple phase space model (`PHSP`) is sufficient. For more complex decays with constraints from angular momenta, there are specialized models, depending on the particles involved, like `TVS_PWAVE` for a **t**ensor particle (with $L = 1$) going to a **v**ector and a **s**calar, which is used for the decay of D_2^* states into $D^*\pi$. The `PHOTOS` module is a

special one, as it is used in addition to another module in an individual decay. It manages final state radiation by triggering the production of an additional photon.

This analysis itself only uses standard decay models, that have been used in many other cases. For tests concerning new physics models however, a custom decay model was used [24]. The default model for $B \to X_c \ell \nu$ decays in generic Monte Carlo is `ISGW2`[25] . It has since been replaced by the more accurate `LLSW`[26] , which will be the subject of Sections 7.3 and 7.4.

Decay tables connect all components, by defining the probability for a particle to decay into certain decay products via a given decay model. These have to be crafted individually for each required sample of signal Monte Carlo. For generic decays (and therefore all of the generic MC) a default decay table is used. It contains some outdated branching ratios, that are of relevance for this analysis and will also be subject of Chapter 7.

For each event, `EvtGen` has to start with an initial particle in a given state, which in case of our Monte Carlo is either an $\Upsilon(4S)$ with the same boost as the real ones produced in collisions at `KEKB`, or a virtual photon carrying the four-momentum of the collision in the case of continuum events. Starting from this particle, a decay time is determined using the particle properties, and the decay tables are used to determine the decay products according to the individual decay probability. The kinematics of the decay products is then determined via the defined decay model. The procedure repeats for each non-stable decay product. The full information of the decay chain then consists of a list of particles and resonances, their mother-daughter relations, positions, and momenta.

5.4.2 Detector Simulation

The output of `EvtGen` is used to simulate the interactions of the generated particles with the Belle detector. For this purpose, the detector with its full geometry, materials, and magnetic fields is modeled in the `GEANT3` framework. It simulates the effects of interactions with the different detector components. Low energetic photons originating from the beam pipe - so called *beam background* - are also added, as they provide a major contribution to the noise in the detector. After the simulated digitalization of the detector response, further processing is identical with real data.

6. Data Composition

This chapter analyzes the expected composition of the dataset. This is done by examining five streams of generic Monte Carlo, including the replacement events for $B \to D^{**}\ell\nu$ decays, as described in Section 5.3. The reconstruction was applied according to Chapter 4.

6.1 Missing Mass Distributions

An observable, that provides a good insight into the kinematics of a reconstructed candidate, is the squared missing mass M^2_{miss}:

$$M^2_{\text{miss}} = \Big[p(\text{Beam}) - \Big(p(B_{\text{tag}}) + p(D^{(*)}) + p(\ell)\Big)\Big]^2$$

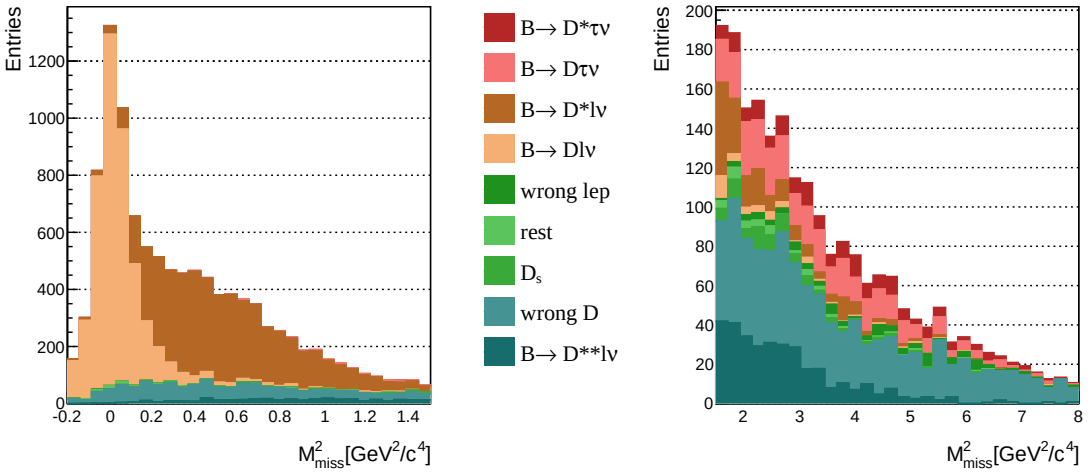
Its distributions are shown separated by reconstruction mode in Figs. 6.1 and 6.2. Due to the dominating peaks in the lower region, the samples are displayed in a lower and an upper M^2_{miss} region, divided at $M^2_{\text{miss}} = 1.5\,\text{GeV}^2\,\text{c}^{-4}$, for better visibility.

The different components in the sample are described in the following:

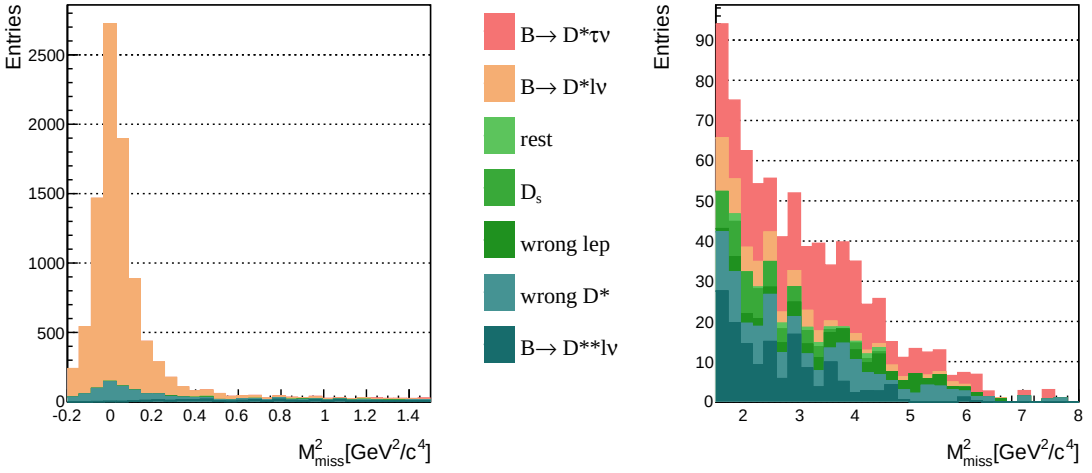
lepton signal: The lepton signal is one of the most prominent components in our data samples. It originates from $B \to D^{(*)}\ell\nu$ decays ($\ell \,\epsilon\, \{e, \mu\}$) and has both visible daughters of the B meson correctly reconstructed. As a consequence, the only missing particle in the reconstruction is the lepton-neutrino, so the squared missing mass peaks around $M^2_{\text{miss}} = 0.0\,\text{GeV}^2\,\text{c}^{-4}$. In all plots within this thesis it is colored light brown.

τ signal: The equivalent component for $B \to D^{(*)}\tau\nu$ decays is the τ signal, that also has the $D^{(*)}$ meson coming directly from the B decay reconstructed correctly and also properly identified the light lepton from the τ decay. So the missing particles in this component are the three neutrinos, building a system of missing particles with a wide distribution of squared missing mass larger than zero and most prominent in regions above $M^2_{\text{miss}} = 1.0\,\text{GeV}^2\,\text{c}^{-4}$. It is much less prominent than the lepton signal. In all plots within this thesis it is colored light red.

lepton cross-feed: In the reconstruction process, it happens quite often, that an excited D meson is reconstructed as an unexcited one. Especially in decays to a D meson and an *uncharged* pion or photon, these can easily elude the detector. These pions and photons are then part of the missing system, which in the case of $B \to D^*\ell\nu$ decays



(a) $D^+\ell^-$

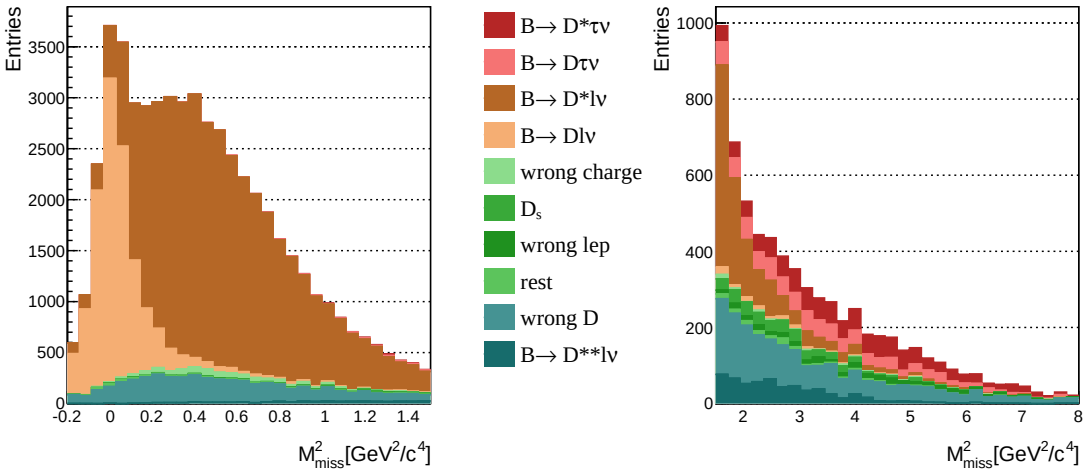


(b) $D^{*+}\ell^-$

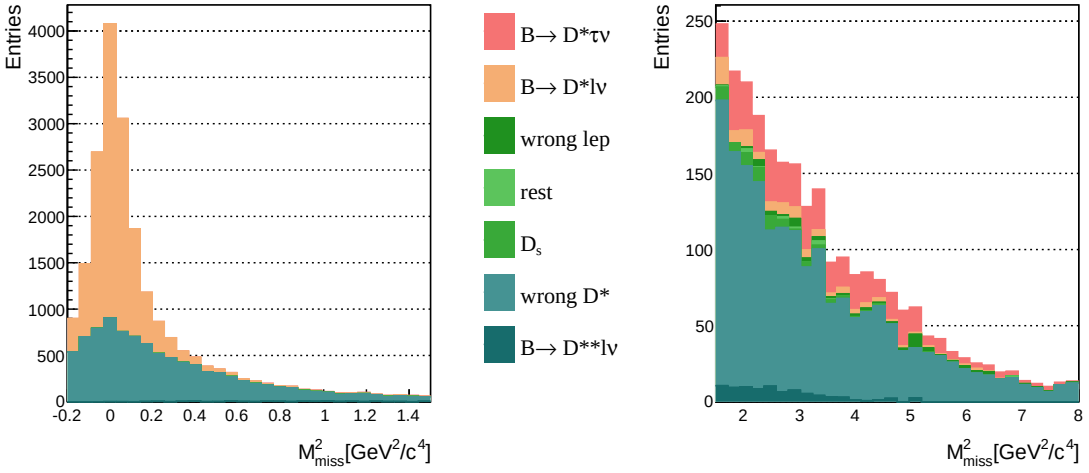
Figure 6.1: The composition of the analysis sample in the reconstruction modes with charged D meson, using five times the expected amount of data. Shown is the M^2_{miss} distribution for $M^2_{\text{miss}} < 1.5\,\text{GeV}^2\,\text{c}^{-4}$ (left) and $M^2_{\text{miss}} > 1.5\,\text{GeV}^2\,\text{c}^{-4}$ (right).

consists of one neutrino and one pion or photon. Such pions and photons usually are low-energetic due to the limited phase-space in the D^* decay, so the resulting distribution of the squared missing mass is not as wide as for the τ signal, but with respect to the lepton signal it is shifted towards higher values and significantly broader with tails into higher regions. The majority of its candidates however lie in the region below $M^2_{\text{miss}} = 1.0\,\text{GeV}^2\,\text{c}^{-4}$. It occurs only in the datasets where a non-excited D meson is reconstructed, but is one of the most prominent components there. The "lepton" in this component's name can be misleading, as it does not refer to the responsible process, but to the original decay, opposing to decays involving τ instead of ℓ. It is colored dark brown in all associated plots.

τ cross-feed: This component is the analogue to the lepton cross-feed, but originating from $B \to D^*\tau\nu$ decays instead. Its shape, yield, and occupation in squared missing mass are quite similar to the τ signal component. It is also only prominent in the datasets with non-excited D mesons. In all plots within this thesis it is colored dark



(a) $D^0\ell^-$



(b) $D^{*0}\ell^-$

Figure 6.2: The composition of the analysis sample in the reconstruction modes with neutral D meson, using five times the expected amount of data. Shown is the M^2_{miss} distribution for $M^2_{\text{miss}} < 1.5\,\text{GeV}^2\,\text{c}^{-4}$ (left) and $M^2_{\text{miss}} > 1.5\,\text{GeV}^2\,\text{c}^{-4}$ (right).

red.

wrong charge cross-feed: Similar to the cross-feed from neutral particles in D^* meson decays, it can happen, that in a decay $D^{*+} \to D^0\pi^+$ the π^+ is missed. For this to happen and not ending up with an event of non-zero total charge, it must contribute to the tag side B meson, so it occurs only, if the B_{tag} meson is wrong. As D^{*0} mesons do not decay to charged pions, it is only relevant in the $D^0\ell^-$ sample. Its squared missing mass distribution resembles the lepton cross-feed, but its yield is much smaller. It is colored bright green in associated plots.

wrong $D^{(*)}$ meson: The majority of background candidates come from random wrong combinations resulting in a wrong D or D^* meson. This can happen by either missing particles in the event or assigning particles to the wrong B meson. This background occurs in all samples, and in the case of $D^*\ell$ samples also comprises candidates that were combined from a correct D meson and a wrong daughter particle. As the

bandwidth of possible wrong combinations is large, this component is very broad in the squared missing mass spectrum up to highest values. For the further procedure, it can be exploited, that correct D mesons peak in the reconstructed mass distributions, respectively in the distributions of mass difference between D^* and D meson. An estimation of the true yield of this component can therefore be made by looking at the sidebands of these distributions. The according procedure is described in Section 7.8. In all plots within this thesis it is colored light blue.

D^{**} background: This component contains candidates that originate from semileptonic B meson decays that include higher excitations of D mesons, which will in the following be called D^{**}. In our simulations these include D_2^*, D_0^*, D_1, D_1', and the radial excitations $D(2S)$ and $D^*(2S)$, each one in charged and neutral state. They decay typically to a D or D^* meson plus one pion, but decays to more particles are also possible. So in the final state these decays provide a real lepton, a real $D^{(*)}$ meson, and (at least) one pion. If the pion is missed in the reconstruction, these decays mimic the signature of the signal decays. As the phase space is much higher than in the case of $D^* \to D\pi$ decays in the lepton cross-feed component, the offset in the squared missing mass is also larger and gives the distribution a greater resemblance to the τ signal. The yield of the D^{**} background is also comparable to the τ signal, so its proper estimation is a key component in this measurement. Contrary to other background components, it is not possible to get reliable constraints concerning the yield from simulated data, as the branching ratios of $B \to D^{**}\ell\nu$ decays are only vaguely known. As this component can not be suppressed either, its yield has to be estimated on real data. It is colored dark blue in all plots in this thesis.

wrong lepton: Besides the $D^{(*)}$ meson, the lepton in our signature can also be wrongly identified. In these cases it is usually a charged kaon or pion, coming from the tag side or from $B \to DK/\pi$ decays. This happens far less frequent than the misidentification of the $D^{(*)}$ meson. If lepton and D meson are both misidentified, the candidate is by definition assigned to the wrong $D^{(*)}$ background. The wrong lepton background is a broad structure in squared missing mass that occurs in all reconstruction samples with low yield. It can be well estimated from simulated data. In all plots within this thesis it is colored dark green.

D_s decay: The final state of the decay chain $B \to D^{(*)}D_s$ with $D_s \to \ell\nu$ is identical to our signal and normalization decays. However, the decay $D_s \to \ell\nu$ is helicity suppressed and only the tauonic D_s decays provide a non-negligible contribution in our samples. Its distribution in the squared missing mass resembles the τ signal, but its yield is rather low and its description on simulated data is considered reliable, as it has been measured directly with the Belle experiment before [27]. It is colored medium green.

rest: The *rest* component contains all background candidates, that are not covered by other components. It contains candidates with well identified final state particles, that do not origin from one of the previously covered sources and may be random combinations of tag- and signal side particles. Its yield is quite low in all samples and is well assessable from simulated data. It is colored light green in the plots in this thesis. In the result plots of the final fit procedure, all non-crucial background components (wrong charge cross-feed, wrong $D^{(*)}$, wrong lepton, D_s and rest) will be combined in the rest component for better visibility.

7. Correction of Simulated Data

The usage of simulated data is crucial for this analysis. The $B \to D^{(*)}\tau\nu$ signal has to be separated from various backgrounds, whose distributions can not be extracted from measured data or simple physical reflections. Furthermore, some component's yield can be estimated much more precisely from simulated data than by fitting recorded data.

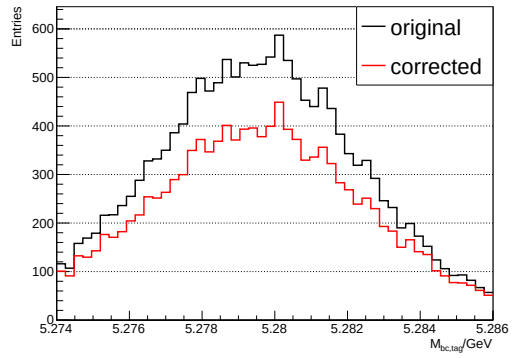
Although the artificial data samples were created with best knowledge of physics parameters at their time, they are far from perfect. One field of imperfections are branching ratios, that get improved frequently by new measurements. Another field is the simulated detector response, that shows deviations from the real behavior. While the overall description of data in the simulations seems to be accurate, the flaws in details have a high potential to infuse biases into this analysis. This chapter lists these flaws and explains ways to quantify and correct them.

7.1 Tag Side Efficiencies

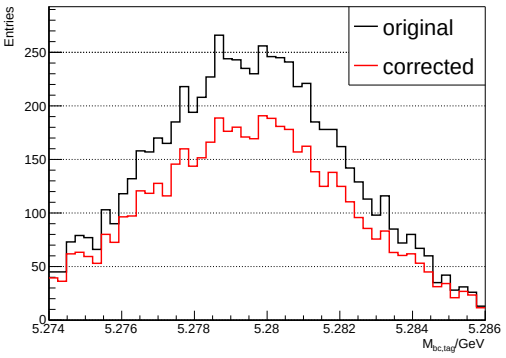
The usage of the `ekpfullrecon` to reconstruct tag side B mesons provides a high variety of decay chains with different multiplicities, which leaves a lot of room for imperfections in the simulation. It was discovered, that on simulated data the amount of correctly reconstructed tags was significantly higher than on real data, depending on the reconstruction mode. While this would alter the efficiency calculations of our signal and normalization in the same way and therefore cancel out, it also affects backgrounds with correct tags, that are estimated via generic Monte Carlo. To counter this, the events are reweighted according to a standard procedure of the Belle collaboration [28]. It uses two different methods to estimate the amount of correctly reconstructed tags per decay mode on data: Fitting general $M_{\mathrm{bc,tag}}$ distributions and fitting M^2_{miss} distributions of semileptonic signal-side decays. The ratios of correct tags on recorded and simulated data range in different channels from 0.35 to 1.1 with an overall ratio of ≈ 0.75. The difference is visualized in the $M_{\mathrm{bc,tag}}$ distributions in Fig. 7.1.

7.2 D_s Component

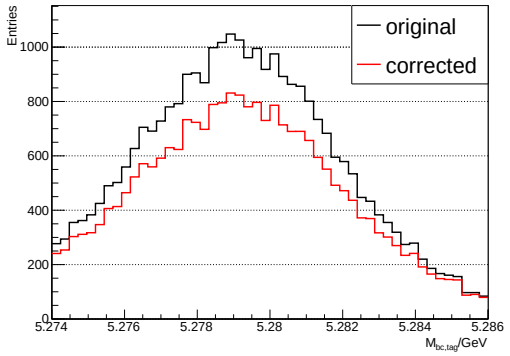
The D_s background is a rather small component in the analysis, but even small imperfections can significantly influence its results. The D_s background stems from $B \to D^{(*)}D_s^{(*)}$ decays with the D_s decaying (semi-)leptonically. Since the determination of the parameters for the Belle generic Monte Carlo sample the branching ratio for $D_s \to \tau\bar{\nu_\tau}$ has been



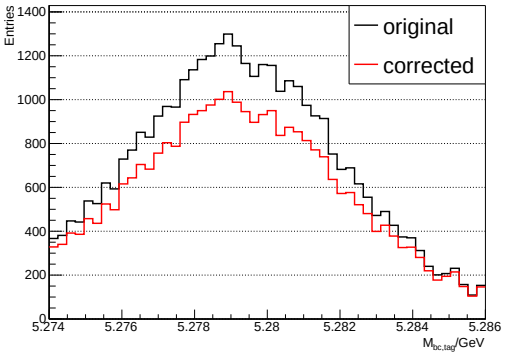
(a) $B^0 \to D^- \ell^+ \nu_\ell$



(b) $B^0 \to D^{*-} \ell^+ \nu_\ell$



(c) $B^+ \to \overline{D}^0 \ell^+ \nu_\ell$



(d) $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$

Figure 7.1: Effect of the tag side efficiency correction on $M_{\mathrm{bc,tag}}$.

measured more precisely. As this decay mimics the final state of our signal, each such event is weighted by the ratio of branching ratios of the Particle Data Group (world average) and generic Monte Carlo $\mathcal{BR}_{\mathrm{PDG}}/\mathcal{BR}_{\mathrm{MC}} = 0.0640/0.0543$.

In semileptonic D_s decays, most branching ratios are very well compatible with the current world average values. The exception are nonresonant $D_s \to KK\ell\nu$ decays. They are simulated additionally to the resonant $D_s \to \phi(\to KK)\ell\nu$ decays, while the Particle Data Group does not give a branching ratio for the nonresonant decay. In the analysis of these decays in [29] it is stated, that the resonant part is overwhelmingly dominating. Our handling is therefore to remove nonresonant decays $D_s \to KK\ell\nu$ from the simulated samples.

7.3 $D^{(*)}\ell\nu$ components

In the standard Belle generic Monte Carlo, the semileptonic $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays are generated according to the heavy quark effective theory (HQET). Reference [30] shows, that the parameters, that were used, are outdated and their current measured values help improving the reproduction of distributions on real data. Therefore an event-by-event reweighting is performed. It is based on a 2-dimensional binning in the momentum transfer q^2 and p^*_ℓ, the lepton momentum in the CM-frame of the decaying B meson, as described in the given reference. The uncertainties on the parameters will be a subject of systematic uncertainties.

7.4 $D^{**}\ell\nu$ components

The same reference also gives reweighting instructions for $B \to D^{**}\ell\nu$ decays from the ISGW model [25], that is used for Monte Carlo production, to the LLSW model [26], that is based on HQET. As $B \to D^{**}\ell\nu$ provides the most important background, proper modelling is crucial for a reliable result. The reweighting here is performed in bins of the recoil w and $\cos\theta_\ell$, the angle between the momenta of D meson and lepton, in the rest frame of the W boson. Uncertainties on the model parameters will also be concerned in the systematic uncertainties.

7.5 Lepton ID Efficiency

The efficiencies of the lepton identification of recorded and simulated data are slightly different. This has been measured for all analyses of the Belle collaboration in [31]. It is based on the process $e^+e^- \to e^+e^-\ell^+\ell^-$, written as $\gamma\gamma \to \ell^+\ell^-$, and was tested in the hadronic environment of $J/\psi \to \ell^+\ell^-$ decays. Corrections are provided in dependence of polar angle θ and momentum p. As the momenta of secondary leptons in the τ signal are slightly lower than in the lepton signal, these corrections can affect the measurement. Using a standard tool [32], the correction factors in Tables 7.1 and 7.2 were calculated. The eID entries are divided into two phases with slightly different detector setup. The μID entries have more phases due to a technical issue in the muon system. The ratios are very well compatible with 1, with eID slightly larger than 1 and μID slightly lower. As we handle e and μ sample only combined, this is on average very close to 1.0, so the correction factors are not applied, but will be subject of systematic uncertainties.

	exp	lepton signal	tau signal	ratio
eID	7 - 27	0.987 ± 0.021	0.963 ± 0.031	1.025 ± 0.039
eID	31 - 65	0.983 ± 0.017	0.968 ± 0.024	1.015 ± 0.030
μID	7 - 27	0.966 ± 0.025	0.949 ± 0.025	1.018 ± 0.037
μID	31 - 39, 45a	0.955 ± 0.019	0.980 ± 0.018	0.985 ± 0.027
μID	41 - 49	0.950 ± 0.020	0.974 ± 0.018	0.976 ± 0.028
μID	51 - 65	0.979 ± 0.019	0.991 ± 0.020	0.989 ± 0.028

Table 7.1: Lepton efficiency corrections for the $D^{(*)+}\ell^-$ samples.

	exp	lepton signal	tau signal	ratio
eID	7 - 27	0.986 ± 0.021	0.964 ± 0.032	1.023 ± 0.039
eID	31 - 65	0.982 ± 0.017	0.969 ± 0.023	1.014 ± 0.029
μID	7 - 27	0.966 ± 0.025	0.948 ± 0.025	1.019 ± 0.037
μID	31 - 39, 45a	0.954 ± 0.020	0.969 ± 0.021	0.984 ± 0.030
μID	41 - 49	0.947 ± 0.020	0.967 ± 0.021	0.979 ± 0.030
μID	51 - 65	0.977 ± 0.020	0.992 ± 0.021	0.985 ± 0.029

Table 7.2: Lepton efficiency corrections for the $D^{(*)0}\ell^-$ samples.

7.6 Lepton Fake Rates

Although the cuts on particle ID (see Section 4.6) result in quite clean samples of charged final state particles, there are always some misidentified. These occur in among all types

of particles, with the most frequent being K^+ mesons misidentified as muons or vice versa. The probabilties for misidentification are not perfectly modeled on simulated data and need to be addressed. A wrong kaon or pion usually results in wrong $D^{(*)}$ background, that is corrected using real data, while a wrong electron or muon needs to be manually corrected. The correction of the lepton fake rate was analyzed for the Belle collaboration [33] by investigating a specialized sample of D^* meson decays, that provide very clean π^+ and K^+ samples and comparing the fake lepton rate on recorded and simulated data. Correction factors are provided in 8 bins of θ ($0 \ldots 151°$) and 11 bins in the momentum of the lepton candidate in the laboratory frame p_{lab} ($0 \ldots 4.0\,\text{GeV}/c$).

7.7 Wrong Tag Yield

While the tag side efficiency bias requires strong corrections, the amount of wrongly reconstructed tag side B candidates in simulated data is much closer to the amount in real data. However, some background components are constrained to Monte Carlo yields and are comprised at least partly of wrong tag candidates. It is therefore necessary to check this difference and apply correction weights, where necessary.

The method used for this is to use a sample in data and MC, where the requirement for $M_{\text{bc,tag}}$ is changed to $5.23\,\text{GeV}\,\text{c}^{-2} < M_{\text{bc,tag}} < 5.25\,\text{GeV}\,\text{c}^{-2}$ to cover a region where correctly reconstructed B_{tag} mesons are suppressed. The raw numbers using all other MC weights are shown in Table 7.3. Two different M^2_{miss} regions were used to check for correlations, but none were found. However, the yields in data and MC are quite different, with MC yields being much larger.

dataset	M^2_{miss} sample	data	MC
$D^+\ell^-$	low	116	163.34
$D^+\ell^-$	high	262	420.35
$D^0\ell^-$	low	163	185.75
$D^0\ell^-$	high	389	475.24
$D^{*+}\ell^-$	low	17	22.23
$D^{*+}\ell^-$	high	29	38.79
$D^{*0}\ell^-$	low	85	92.15
$D^{*0}\ell^-$	high	185	223.10

Table 7.3: Raw yields in a signal suppressed M_{bc} region. Two different M^2_{miss} regions were tested. While these are well compatible, data and MC yields show some large differences.

A closer look into the composition shows, that a large fraction in the order of 75% to 95% of the wrong tags are due to wrongly reconstructed D mesons. The further analysis will correct these from sidebands, so a direct reweighting of these events because of wrong tags is not necessary. To estimate the difference between MC and data in events that are not made up of wrong D mesons, we use the D mass and $D^* - D$ mass difference sidebands (see Section 7.8) to scale down the wrong D events in the Monte Carlo sample. The comparison of the corrected yields is given in Table 7.4.

The fit component with the highest wrong tag yield will be the events with wrong charge, which comes from an interchange of two charged particles between tag and signal B meson. It therefore consists exclusively of candidates with wrong tags. Fortunately, it is only prominent in the $D^0\ell^-$ sample, where the difference between Monte Carlo and real data is $< 1\%$. The other deviations are barely significant with the given statistics, but with up to

dataset	MC (corr.)	data	ratio
$D^+\ell^-$	423.20	378	1.120
$D^0\ell^-$	548.02	552	0.993
$D^{*+}\ell^-$	52.34	46	1.138
$D^{*0}\ell^-$	289.29	270	1.071

Table 7.4: Corrected yields in the signal suppressed M_{bc} region. The two different M^2_{miss} regions were combined. The correction comes from rescaling wrong D events, that will be corrected from sidebands in the analysis procedure. Differences between MC and real data are now much smaller.

14% deviation, they are corrected for by weights and the uncertainties are later treated as systematic uncertainties.

One issue is now, to which events the weight should be applied. It is intended to be a signal side correction, so we have to exclude events, that originate from effects, that are corrected in another way. These are not only wrong D events, but also wrong lepton events. Remaining events with a wrong tag side B meson are then reweighted with the ratio of data to MC.

7.8 $D^{(*)}$ Meson Corrections and Sidebands

The efficiencies for D and D^* reconstruction can be different for recorded data and Monte Carlo and depend on the individual reconstruction channel. As the efficiency correction is determined on the analysis sample itself, it is only performed after all other corrections have been applied.

7.8.1 Correct $D^{(*)}$ Characteristics

An event with a correct D meson can end up in any data component, except for the *wrong* $D^{(*)}$ *background*. In principle, wrong efficiencies would cancel out in the ratios that are measured in this analysis. But as we have background components with correct $D^{(*)}$ mesons that are estimated from Monte Carlo, an efficiency correction is necessary.

7.8.2 Wrong $D^{(*)}$ Characteristics

Wrong $D^{(*)}$ mesons are used indirectly, for estimating the yields of the *wrong D background* components on reconstructed data from the sidebands of $M_D \equiv D$ meson mass or ΔM_{D^*D} $\equiv$ mass difference between D^* and D meson. It is therefore not necessary to reweight wrong D meson candidates beforehand, but the estimation of the yields has to be done separately for each $D^{(*)}$ decay channel, and the sideband regions must also be determined individually.

7.8.3 Correction Procedure

There are two tasks here:

- The *correct* $D^{(*)}$ yield has to be corrected on Monte Carlo by reweighting to match the yield on data.
- The sideband regions must be defined.

For both, we need a model of the M_D distribution, respective ΔM_{D^*D} distribution in the case of D^*. In each $D^{(*)}$ decay mode, the following procedure is applied:

1. Probability density functions for correct and wrong $D^{(*)}$ are created:
   - the functions are based on simulated data
   - they are modeled with kernel estimator functions
2. Probability density functions are fitted to reconstructed data to get the correct $D^{(*)}$ yield:
   - the unbinned maximum likelihood method is used
   - only one free parameter in the fit: correct $D^{(*)}$ fraction
   - only in D channels: if a D^* peak is present (channel $D^+ \to K^0_S\pi^+\pi^0$), the nominal mass of the D^* meson $\pm 60\,\mathrm{MeV\,c^{-2}}$ is excluded
3. Definition of the signal range:
   - it is asymmetric around the nominal mass
   - in both directions: 1.5× root mean square of the one-sided signal distribution is used
4. Definition of the correction weight:
   - It is the ratio of *correct* $D^{(*)}$ yield on reconstructed data vs. Monte Carlo in the signal range.
5. Definition of the sideband region:
   - It is taken from both sides of the M_D (ΔM_{D^*D}) range.
   - It covers at most the same width as the signal region.
   - It may not approach the signal region closer than 3.0× the one-sided root mean square of the signal distribution (asymmetric).
   - Due to zero yields, the lower sidebands are excluded in:
     - $D^0 \to K^-\pi + \pi^0$
     - $D^{*+} \to D^0\pi^+$
     - $D^{*+} \to D^+\pi^0$
     - $D^{*0} \to D^0\pi^0$
   - The D^* peak in the D mode (cross-feed at $(2.01 \pm 0.06)\mathrm{GeV\,c^{-2}}$) is removed from the upper sideband in:
     - $D^+ \to K^0_S\pi^+\pi^0$

An overview of the extracted factors and ranges is given in Table 7.5. A graphical display of the fits and the extracted ranges is given in Figs. 7.2 to 7.5.

7.8.4 Weight Application

This correction is not done on an event-by-event basis. Each candidate, that contains a $D^{(*)}$ meson, that is identified as *correct*, is weighted by the factor given in Table 7.5.

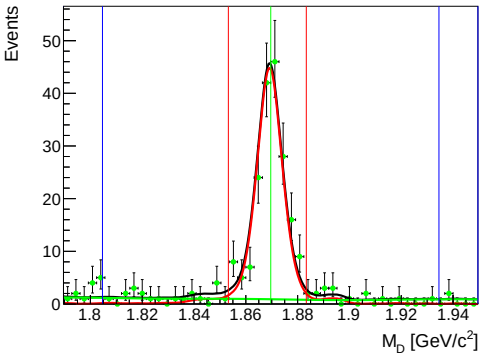
$D^{(*)}$ mode	signal range	lower SB	upper SB	signal corr.
$D^+ \to K^0_S\pi^+$	1.853 - 1.883	1.790 - 1.805	1.935 - 1.950	0.8765
$D^+ \to K^0_S\pi^+\pi^0$	1.798 - 1.983	1.650 - 1.678	2.070 - 2.150	1.0945
$D^+ \to K^-\pi^+\pi^+$	1.852 - 1.887	1.790 - 1.807	1.932 - 1.950	0.8466
$D^+ \to K^0_S\pi^+\pi^+\pi^-$	1.849 - 1.893	1.790 - 1.814	1.931 - 1.950	0.7474
$D^{*+} \to D^+\pi^0$	0.139 - 0.145	-	0.154 - 0.160	0.7630
$D^{*+} \to D^0\pi^+$	0.144 - 0.148	-	0.156 - 0.160	0.9979
$D^0 \to K^0_S\pi^0$	1.772 - 1.889	1.600 - 1.616	1.998 - 2.100	1.0444
$D^0 \to K^-\pi^+$	1.847 - 1.879	1.784 - 1.800	1.928 - 1.944	0.9009
$D^0 \to K^-\pi^+\pi^0$	1.805 - 1.908	-	1.997 - 2.100	0.9927
$D^0 \to K^0_S\pi^+\pi^-$	1.847 - 1.880	1.784 - 1.801	1.928 - 1.944	0.8932
$D^0 \to K^-\pi^+\pi^-\pi^+$	1.845 - 1.883	1.784 - 1.803	1.925 - 1.944	0.9998
$D^{*0} \to D^0\gamma$	0.124 - 0.154	-	0.174 - 0.190	0.8140
$D^{*0} \to D^0\pi^0$	0.140 - 0.147	-	0.156 - 0.162	0.8829

Table 7.5: Ranges and correction factors for M_D and ΔM_{D^*D} distributions. The ranges are given in GeV. The correction factors are multiplicative weights for events with correct $D^{(*)}$.

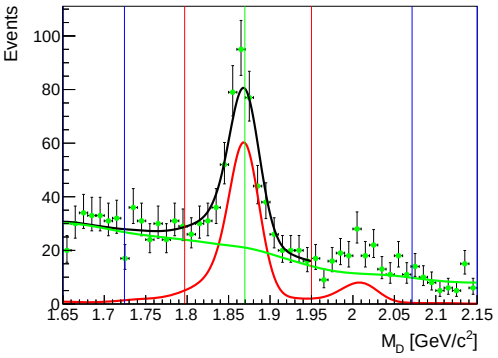
7.8.5 Cut Application and Sideband Sample

The data sample, that is analyzed further, is created by cutting on the signal ranges given in Table 7.5. In the case of D^* mesons, not only a cut on ΔM_{D^*D} is applied, but also a cut on the masses of their D daughters, according to their reconstruction channel.

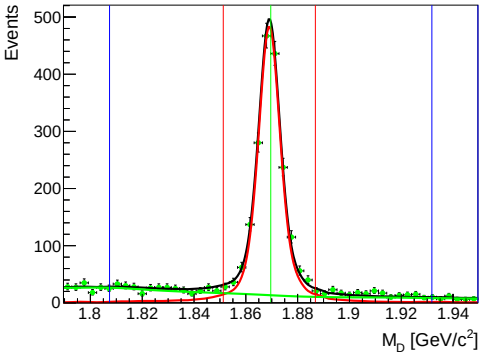
While the sideband sample, that later will be used to determine the *wrong* D yields, consists only of candidates in the sideband ranges, not all of the D^* meson candidates in the sidebands are valid. Their D meson daughter additionally has to fulfill the cuts for its respective signal range. This is done, as the D meson daughters of wrong D^* mesons in the signal range of the analysis sample also fulfill this requirement.



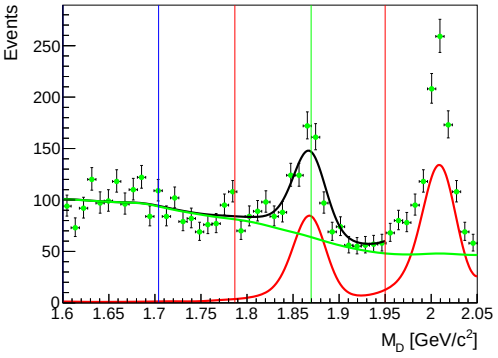
(a) $D^+ \to K^0_S \pi^+$



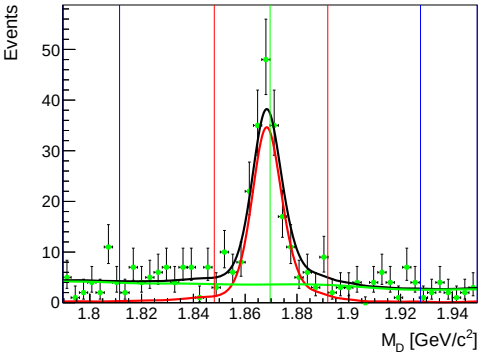
(b) $D^+ \to K^0_S \pi^+ \pi^0$



(c) $D^+ \to K^- \pi^+ \pi^+$

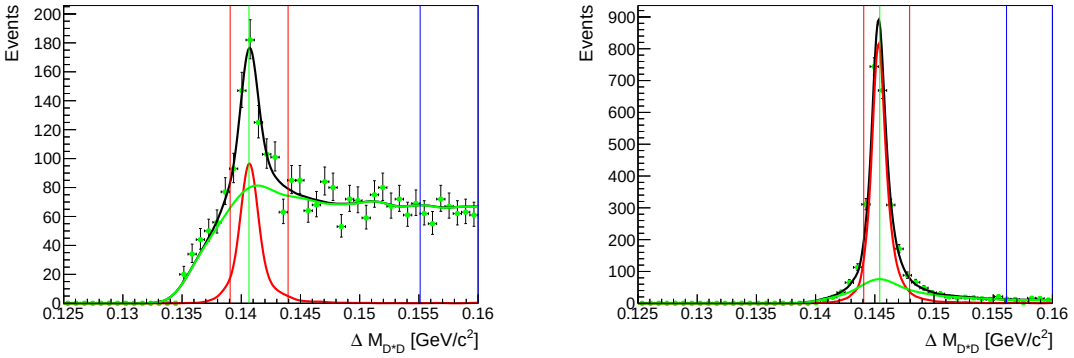


(d) $D^+ \to K^- \pi^+ \pi^+ \pi^0$



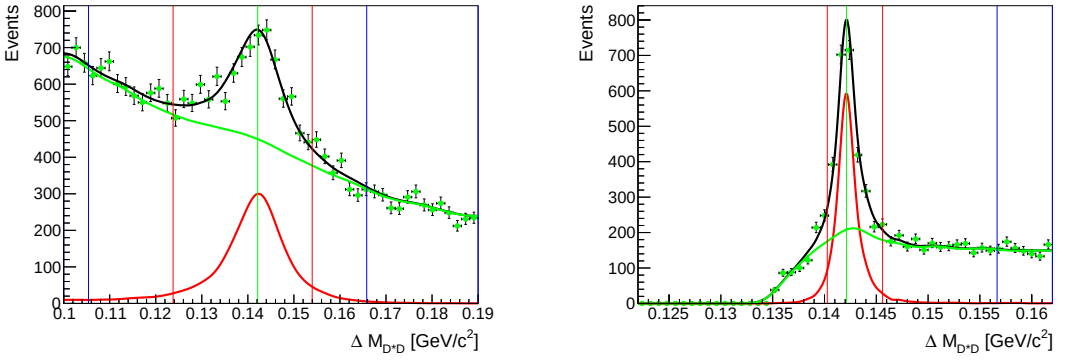
(e) $D^+ \to K^0_S \pi^+ \pi^+ \pi^-$

Figure 7.2: M_D fits to determine correct D^+ yields. The green vertical line indicates the nominal mass, the red lines define the signal region, and the blue lines limit the sideband regions.



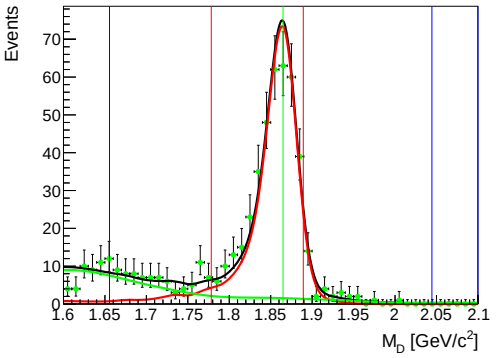
(a) $D^{*+} \to D^{+}\pi^{0}$ (b) $D^{*+} \to D^{0}\pi^{+}$

Figure 7.3: ΔM_{D^*D} fits to determine correct D^{*+} yields. The green vertical line indicates the nominal mass, the red lines define the signal region, and the blue lines limit the sideband regions.

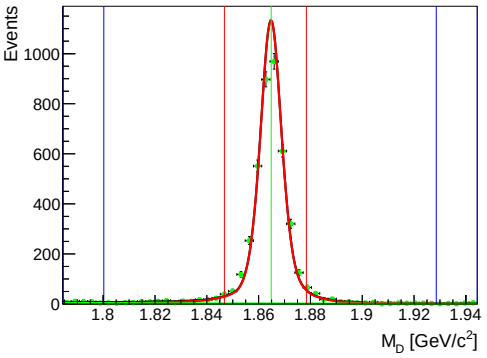


(a) $D^{*0} \to D^{0}\gamma$ (b) $D^{*0} \to D^{0}\pi^{0}$

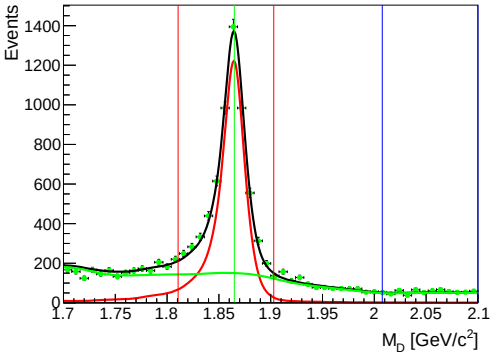
Figure 7.4: ΔM_{D^*D} fits to determine correct D^{*0} yields. The green vertical line indicates the nominal mass, the red lines define the signal region, and the blue lines limit the sideband regions.



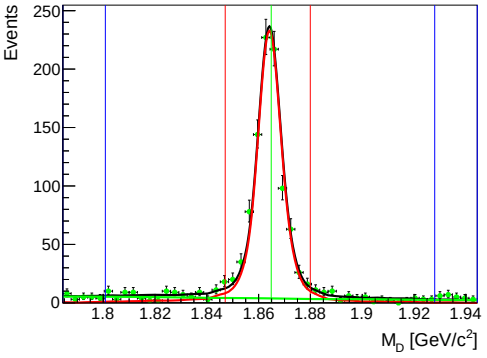
(a) $D^0 \to K^0_S \pi^0$



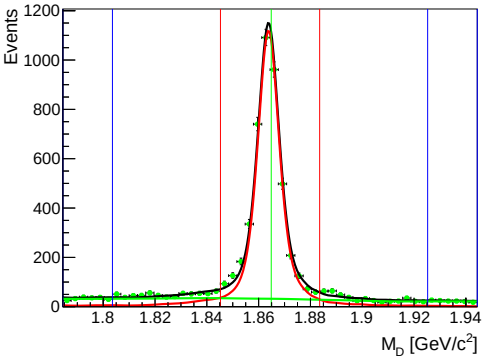
(b) $D^0 \to K^- \pi^+$



(c) $D^0 \to K^- \pi^+ \pi^0$



(d) $D^0 \to K^0_S \pi^+ \pi^-$



(e) $D^0 \to K^- \pi^+ \pi^- \pi^+$

Figure 7.5: M_D fits to determine correct D^0 yields. The green vertical line indicates the nominal mass, the red lines define the signal region, and the blue lines limit the sideband regions.

8. Fitting Variables

The goal of this analysis is to measure the ratios R and R^* as defined in Eq. (1.1). The main task is therefore to distinguish the τ signal from the lepton signal, and both of them from various backgrounds.

The observable final state particles are the same for τ signal and lepton signal, but the invisible particles differ: The former has three neutrinos in its final state, the latter only one. By reconstructing the tag side B meson and with the knowledge of the four-momentum of the beam, the four-momentum of the invisible particle system can be calculated and allows the distinction.

8.1 Squared Missing Mass

To break down the four dimensions of the invisible momentum to a single discriminating variable, the absolute square is calculated, which represents the invariant mass of the invisible system, already introduced as squared missing mass M^2_{miss} in Eq. (1.2). A detailed explanation of the distributions of this observable for different components was already given in Chapter 6. The essence is, that the lepton signal shows a prominent peak around $M^2_{\text{miss}} = 0.0\,\text{GeV}^2\,\text{c}^{-4}$, the lepton cross-feed component (in the samples with non-excited D mesons) is also a distinct component with a shift to slightly higher values, and all other components have wide distributions with huge overlaps among each other.

So the squared missing mass provides some information for the task at hand, but it is by far most powerful to identify lepton signal and lepton cross-feed in its distributions. To be able to also distinguish the τ signal from other background components, we need additional information from other observables.

8.2 Other Observables

There are several other observables, that provide separation power for the τ signal component:

$\boldsymbol{E}_{\mathbf{ECL}}$ is the unassociated energy in the electromagnetic calorimeter (ECL). All energy clusters in the ECL, that can not be associated with either the tag side or the signal side B meson, are summed up in this observable. The implementation is identical to the one in reference [34], where diluting effects like bremsstrahlung are considered before the summation. The idea behind this observable is, that energy deposits in

the ECL from physical processes in the event are likely higher than energy deposits from beam radiation, that is not involved in the collision. Significant energy deposits therefore strongly indicate missing a physical process in the decay. For example, this is the case when one of the two photons of a π^0 decay is not detected and the π^0 eludes the reconstruction. In the case of $B \to D^{**}(\to D\pi^0)\ell\nu$ decays, this results in a matching signature, but one or both photons still contribute to E_{ECL}.

q^2 is the squared four-momentum transfer from the B meson to the D meson, which depends strongly on the mass of the lepton, that is produced besides the D meson.

p_ℓ^* is the absolute measured momentum of the lepton in the rest frame of the B meson. In the case of τ signal, the lepton is produced in a secondary decay and therefore has a reduced momentum compared to the lepton signal.

While q^2 and p_ℓ^* provide sensitivity for the separation of τ signal from lepton signal, their benefit beyond the squared missing mass is quite limited. In fact, they are both strongly correlated to it. E_{ECL} is much better suited for the task of separating the τ signal from various backgrounds. The most important background in this regard is the D^{**} background, as the other ones can be estimated from simulated data or be tightly constrained from other physical relations (as explained in Chapter 6).

8.3 Neural Network

8.3.1 Training

One drawback concerning the usage of these other variables is that they show correlations to the missing mass. However small, they have the potential to severely influence the result, due to the broad nature of the τ signal. As we can not neglect them, we have to find a way to take them into account without throwing away much information. For this reason, we combine these variables in a neural network. Some general quality variables and the squared missing mass itself are also added.

Four separate trainings were performed for the four different reconstruction modes, configured as follows:

Target τ signal

Background D^{**} background, wrong lepton background, wrong charge cross-feed, D_s background, remaining background (see Chapter 6)

Not included wrong $D^{(*)}$ meson background, lepton signal, lepton cross-feed, τ cross-feed

Precut $M^2_{\mathrm{miss}} > 0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$

Each of the above trainings contained the same 8 variables. They are listed together with their significance in Table 8.1. The significance is defined as the correlation of the variable to the target multiplied by the square-root of the sample size [35]. Other variables are not taken into account. The purity over efficiency plots and the network outputs of the trainings can be found in Figs. 8.1 and 8.2.

Due to the setup, the network is optimized to separate τ signal from backgrounds in a region of the squared missing mass distribution, that is only scarcely populated by lepton signal and lepton cross-feed candidates. The motivation for this strategy is explained in Section 8.4.

Variable	Significance [σ]			
	$D^+\ell^-$	$D^{*+}\ell^-$	$D^0\ell^-$	$D^{*0}\ell^-$
Number of remaining π^0 candidates within 5σ of the nominal mass. (compare different requirements on π^0 masses in section 4.6)	6.20	4.47	10.35	7.02
Cosine of the angle between the momentum of the D meson and the line through the D meson decay vertex and the interaction point.	3.24	2.60	0.79	4.94
Remaining energy in the ECL after removal of the clusters used for the tag side or signal side.	19.71	14.37	27.08	16.89
Invariant mass of the missing four-momentum in the event.	8.72	7.56	12.81	10.48
o_{NB} of the tag side B meson.	1.99	0.86	1.80	0.14
Hash identifying the type and decay channel of the tag side B meson.	3.13	4.34	4.74	3.17
Hash identifying the type and decay channel of the D meson.	1.32	2.78	5.18	2.19
Absolute value of the 3-momentum of the lepton in the B_{sig} center of mass frame.	7.87	6.95	8.41	7.13

Table 8.1: Input variables in the NeuroBayes trainings.

8.3.2 Transformation

While the network output could in principle be fitted in this form, a fitting procedure is more robust if it handles smooth distributions. Therefore a transformation is applied on the network output o_{NB}:

$$o_{\mathrm{NB,trafo}} \equiv \log \frac{o_{\mathrm{NB}} - o_{\mathrm{min}}}{o_{\mathrm{max}} - o_{\mathrm{NB}}}$$

The parameters o_{min} and o_{max} are the minimum and maximum network output values in their respective sample. The resulting distributions have smoother shapes and can be well described with bifurcated Gaussian functions.

8.4 Sample Split

The central purpose of the neural network variable is to separate the τ component from all others. However, the sample is dominated by lepton signal and cross-feed, which would be well separable by the squared missing mass itself, in a squared missing mass region that barely contains any τ signal. As a result, training τ versus all other components is not optimal: various backgrounds in the missing mass region of the τ signal would have only

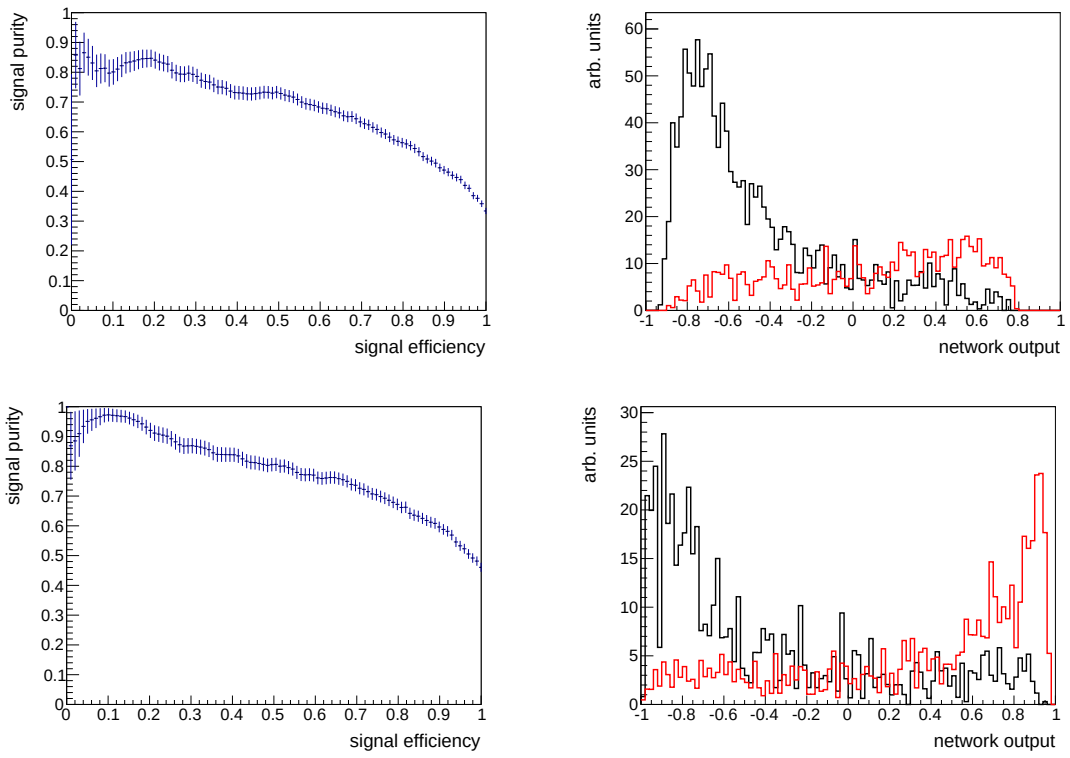


Figure 8.1: Purity over efficiency plots and network output plots for channels $D^{+}\ell^{-}$ and $D^{*+}\ell^{-}$.

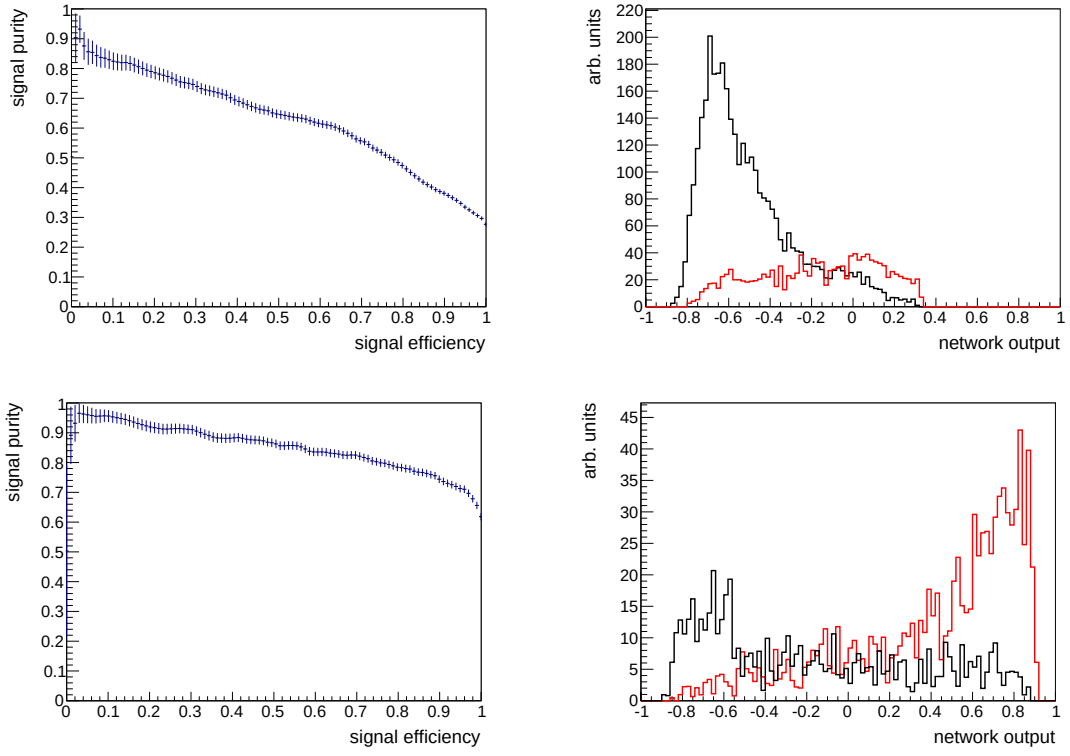


Figure 8.2: Purity over efficiency plots and network output plots for channels $D^{0}\ell^{-}$ and $D^{*0}\ell^{-}$.

small impact on the training, and the components, that dominate at low missing masses, would end up being more difficult to separate from each other.

To improve this, the sample is split at $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$, as shown in Fig. 8.3: The sample with lower missing mass provides an excellent estimation of the yield of the dominant components, lepton signal and cross-feed. Therefore, the missing mass distribution itself will be fitted. As the majority of the τ signal lies above the threshold, its bad separation in missing mass is not critical. The upper region is most important for our measurement goal. To get the best separation power for the τ component, the network is trained with τ as signal and only background components, that are prominent in this region and can not be estimated in a better way, like D^{**} background. In this region the distribution of the transformed neural network output is fitted.

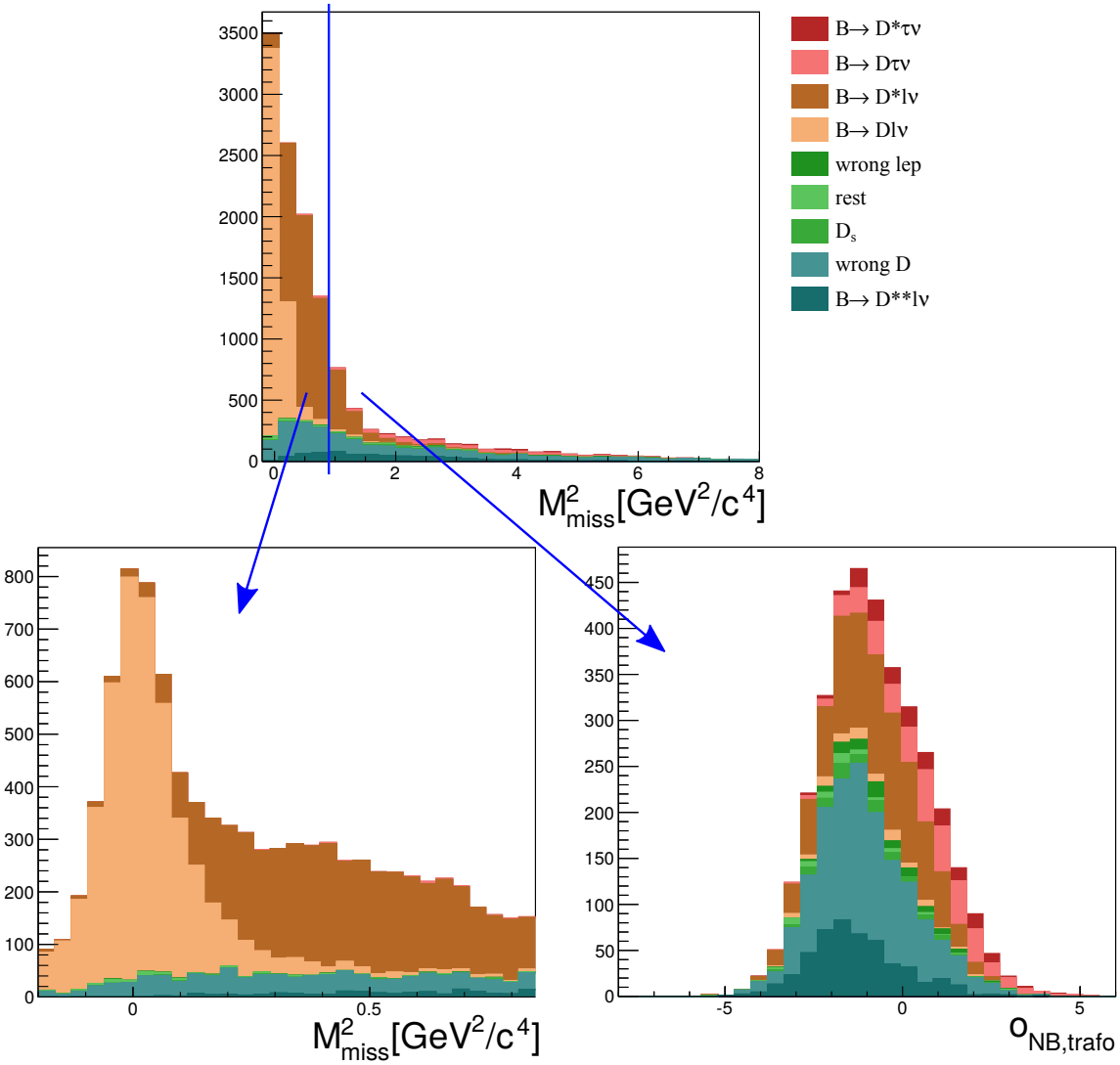


Figure 8.3: Split in M^2_{miss} of the $D^+\ell^-$ sample. The lower squared missing mass area is fitted in the squared missing mass distribution itself, the upper one in the transformed output distribution of the neural network.

9. Fit Procedure

In this analysis we assume isospin symmetry, meaning that the influence of the spectator quark in the semileptonic B meson decays is negligible. The general fit strategy is therefore to fit the B^0 samples ($D^+\ell^-$ and $D^{*+}\ell^-$) and B^- samples ($D^0\ell^-$ and $D^{*0}\ell^-$) simultaneously with the shared parameters R and R^*. Their estimation is performed with the unbinned extended maximum likelihood method.

9.1 Fit Components

We have four samples for the reconstruction modes $D^+\ell^-$, $D^{*+}\ell^-$, $D^0\ell^-$, and $D^{*0}\ell^-$. Each of these is divided by M^2_{miss}, as described in Section 8.4. For each component that could be identified in these samples (see Chapter 6) two probability density functions were determined, one for the lower and higher M^2_{miss} region each.

9.1.1 Probability Density Functions

The shapes of all components were determined from their distributions on generic Monte Carlo samples (5 streams) unless stated otherwise in Section 9.1.2. The probability density functions of M^2_{miss} are represented by smoothed histograms. To generate them, a histogram is filled by the distribution on simulated data, and a smoothing algorithm [36] is applied. The shapes in the $o_{\text{NB,trafo}}$ dimension (see Section 8.3) are parameterized by bifurcated Gaussian functions, that have three parameters each: mean, left width and right width. They were determined for every component by an unbinned maximum likelihood fit of the bifurcated Gaussian function to the distribution on simulated data.

9.1.2 Component Descriptions

The following components are included in the fit:

lepton signal This component has a sharp peak in M^2_{miss} in all four samples, and the yield is a free parameter.

lepton cross-feed This component is a bit broader than the lepton signal, but also contributes strongly in the $D^{+/0}\ell^-$ samples. Its yield can easily be estimated as a free parameter.

wrong charge cross-feed This is a small source of background that only occurs in the $D^0\ell^-$ sample. It could also occur in the $D^+\ell^-$ sample, but is negligibly small there. Its origin is usually a π^+ migration from the signal-side to the tag-side B meson, that leads to a wrong charge on both sides. Its yield is constrained as proportional to the lepton signal in the $D^{*+}\ell^-$ sample, with the proportionality factor taken from simulated data. So it does not add a free parameter to the fit, but one constant factor for the constraint.

τ cross-feed This component cannot be reliably separated by the fit, as its shape is very similar to the τ signal. It has to be determined using other information. The lepton and τ cross-feed candidates are in reality $B \to D^*\ell\nu$ or $B \to D^*\tau\nu$ events, where the excitation of the D^* meson eluded the reconstruction. The same reconstruction error – missing a slow pion or photon – was made for both components. It is therefore reasonable to assume, that the two cross-feed components and their corresponding signal components are connected in the following way:

$$\frac{Y_{D^{0/+}\ell^-,\ell\,\mathrm{CF}}}{Y_{D^{*0/+}\ell^-,\ell\,\mathrm{signal}}} = g_{B^{+/0}} \cdot \frac{Y_{D^{0/+}\ell^-,\tau\,\mathrm{CF}}}{Y_{D^{*0/+}\ell^-,\tau\,\mathrm{signal}}} \,. \tag{9.1}$$

The factor g takes possible efficiency differences between $B \to D^*\ell\nu_\ell$ and $B \to D^*\tau\nu_\tau$ decays into account, most likely kinematic differences due to the large τ mass which results in an altered momentum distribution. The factor g is determined on generic Monte Carlo for the $D^{(*)0}\ell-$ and $D^{(*)+}\ell-$ samples separately using Eq. (9.1), and the τ cross-feed is then determined in the following way:

$$Y_{D^{0/+}\ell^-,\tau\,\mathrm{CF}} = \frac{1}{g_{B^{+/0}}} \cdot Y_{D^{*0/+}\ell^-,\tau\,\mathrm{signal}} \cdot \frac{Y_{D^{0/+}\ell^-,\ell\,\mathrm{CF}}}{Y_{D^{*0/+}\ell^-,\ell\,\mathrm{signal}}} \,. \tag{9.2}$$

The yield of the τ cross-feed is therefore not a free parameter of the fit. It is constrained from the yields of the lepton signal, lepton cross-feed, and τ signal components. The former are free parameters in the fit, and the latter is function of R^* as explained in the following:

τ signal The main focus of this fitting procedure lies on the extraction of R and R^*. Therefore, R and R^* were implemented as free parameters of the fit. In order to link the reconstructed yields of τ signal and lepton signal to the values of R and R^*, their efficiency ratio $f_R^{(*)}$ is needed:

$$R^{(*)} = \frac{\mathcal{N}_{\mathrm{prod}}(B \to D^{(*)}\tau\nu)}{\mathcal{N}_{\mathrm{prod}}(B \to D^{(*)}\ell\nu)} \tag{9.3}$$

$$= \frac{\mathcal{N}_{\mathrm{reco}}(B \to D^{(*)}\tau\nu)}{\mathcal{N}_{\mathrm{reco}}(B \to D^{(*)}\ell\nu)} \cdot \frac{\epsilon_{\mathrm{reco}}(B \to D^{(*)}\ell\nu)}{\epsilon_{\mathrm{reco}}(B \to D^{(*)}\tau\nu)} \tag{9.4}$$

$$= \frac{\mathcal{N}_{\mathrm{reco}}(B \to D^{(*)}\tau\nu)}{\mathcal{N}_{\mathrm{reco}}(B \to D^{(*)}\ell\nu)} \cdot f_R^{(*)} \,. \tag{9.5}$$

The $\mathcal{N}_{\mathrm{reco}}$ refers to the number of reconstructed events, while $\mathcal{N}_{\mathrm{prod}}$ is the number of produced events, and the efficiency ϵ_{reco} is their ratio. The efficiency ratio $f_R^{(*)}$ is calculated for each data sample individually and are named by the excitation state of the reconstructed D meson and the flavor of the reconstructed B meson: f_{R,B^0} in $D^+\ell^-$, f^*_{R,B^0} in $D^{*+}\ell^-$, f_{R,B^+} in $D^0\ell^-$, and f^*_{R,B^+} in $D^{*0}\ell^-$. These factors were determined on generic Monte Carlo and validated on a sample of signal Monte Carlo with equal parts of semitauonic and other semileptonic decays. It is important to remind, that the semileptonic branching ratios in $R^{(*)}$ represent *either e or μ* modes, so the calculation on simulated data has to be done accordingly to avoid an erroneous

factor of 2. The efficiency ratios and $R^{(*)}$ constrain the yields of the τ signal to the yields of the lepton signal by solving Eq. (9.5) for $\mathcal{N}_{\text{reco}}(B \to D^{(*)}\tau\nu)$:

$$\mathcal{N}_{i,\text{reco}}(B \to D^{(*)}\tau\nu) = \mathcal{N}_{i,\text{reco}}(B \to D^{(*)}\ell\nu) \cdot R^{(*)}/f^{(*)}_{R,i} \ , \tag{9.6}$$

where the index i represents the reconstruction sample. However, the depicted values $\mathcal{N}_{i,\text{reco}}$ are not automatically identical with the yields of the τ signal and lepton signal components. In the reconstruction samples without excited D mesons, this assumption is true, but in the case of $D^{*+/0}\ell^-$ many of the produced events in our signal and normalization mode are reconstructed as cross-feed. Mathematically, those events could be omitted in the calculations entirely, but including them improves the uncertainties on the measured yields as well as on the calculated efficiency ratios, that are affected by statistical fluctuations of simulated data.

For the channels with non-excited D mesons, the calculation of the constrained τ signal yield is simple:

$$Y_{D^{+/0}\ell^-,\tau\,\text{signal}} = R \cdot Y_{D^{+/0}\ell^-,\ell\,\text{signal}}/(2f_{R,B^{0/+}}) \ ,$$

where $Y_{i,\ell\,\text{signal}}$ is the lepton signal yield in the respective sample, which counts e and μ modes combined and therefore needs to be corrected by the factor of 2 in the denominator.

For the channels with excited D mesons, the calculation is more complicated, as the efficiency ratio for the combined yield with the respective cross-feed component is used. Therefore the total number of reconstructed $B \to D^*\tau\nu$ events must be calculated, and the yield of τ cross-feed subtracted using its constraint. The final form of the constrained yield is:

$$Y_{D^{*0/+}\ell^-,\tau\,\text{signal}} = \frac{R^* \cdot \left(Y_{D^{*0/+}\ell^-,\ell\,\text{signal}} + Y_{D^{0/+}\ell^-,\ell\,\text{CF}}\right)}{\left(2f^*_{R,B^{+/0}}\right) \cdot \left(1 + \frac{1}{g_{B^{+/0}}} \cdot \frac{Y_{D^{0/+}\ell^-,\ell\,\text{CF}}}{Y_{D^{*0/+}\ell^-,\ell\,\text{signal}}}\right)} \ . \tag{9.7}$$

The τ signal component is special in the sense, that the shapes of the probability density functions are determined on signal Monte Carlo to greatly increase the available statistics. It adds R respective R^* as free parameters and the efficiency ratios $f^{(*)}_R, i$ to the fit.

wrong $D^{(*)}$ background The yield of this background component is determined before the actual fit by using candidates in the M_D ($D^{0/+}$ samples) and ΔM_{D^*D} ($D^{*0/+}$ samples) sideband data. The factors $f^{\text{SB}}_{i,\text{channel}_{D^{(*)}}}$ for sample i are extracted for each $D^{(*)}$ reconstruction channel from generic Monte Carlo:

$$f^{\text{SB}}_{i,\text{channel}_{D^{(*)}}} = \frac{\mathcal{N}^{\text{wrong}\,D^{(*)},\text{fitregion}}_{i,\text{channel}_{D^{(*)}}}}{\mathcal{N}^{\text{all,sidebands}}_{i,\text{channel}_{D^{(*)}}}} \ . \tag{9.8}$$

The $\mathcal{N}$ represent the number of events – either all events in the sidebands, or *wrong* $D^{(*)}$ events in the fit region.

The yield of the *wrong* $D^{(*)}$ background is determined by counting the events in the sidebands of M_D respectively ΔM_{D^*D} on reconstructed data. The yield of the *wrong* $D^{(*)}$ components in the fit samples is calculated as follows:

$$Y_{i,\text{wrong}\,D^{(*)}} = \sum_{\text{channel}_{D^{(*)}}} f^{\text{SB}}_{i,\text{channel}_{D^{(*)}}} \cdot \mathcal{N}^{\text{all,sidebands}}_{i,\text{channel}_{D^{(*)}}} \ . \tag{9.9}$$

The yield is a fixed parameter of the fit.

D^{**} background The yields of this background component are free fit parameters in all four samples.

D_s background This background component has very small yields in all four samples. They are fixed to the standard model expectations using generic Monte Carlo.

wrong lepton background The yields of this component is fixed to MC expectation in all four samples.

rest The yield of the component, that combines all remaining background events is very small in our four samples and fixed to MC expectations.

9.2 Full PDF

The extended Likelihood function to be maximized looks as follows:

$$\mathcal{L} = \prod_i \left[\mathcal{Q}(N_i, K_i) \prod_k \mathcal{P}_i(\mathbf{x}_k) \right] ,$$

where vector $\mathbf{x}_k$ holds the values for M^2_{miss} and $o_{\text{NB,trafo}}$ of candidate k, and i is the reconstruction sample:

$$i \,\epsilon\, \{D^+\ell^-,\, D^0\ell^-,\, D^{*+}\ell^-,\, D^{*0}\ell^-\} .$$

$\mathcal{Q}(N_i, K_i)$ is the Poisson probability to get $K_i \equiv \sum_k 1$ events with an expectation value of $N_i \equiv \sum_j Y_{i,j}$ events, with $Y_{i,j}$ being the total yield of component j in sample i.

The full PDF of reconstruction mode i looks as follows:

$$\mathcal{P}_i(M^2_{\text{miss}}, o_{\text{NB,trafo}}) = \frac{1}{N_i} \cdot \sum_j Y_{i,j} \left[f_{i,j,\text{low}} \mathcal{P}_{i,j,\text{low}}(M^2_{\text{miss}}) + (1 - f_{i,j,\text{low}}) \mathcal{P}_{i,j,\text{high}}(o_{\text{NB,trafo}}) \right] .$$

Table 9.1 summarizes, which components j are non-zero in each reconstruction mode. $f_{i,j,\text{low}}$ is the fraction of events of the component j, that are in the lower M^2_{miss} range. These factors are taken from Monte Carlo and are objective of systematic uncertainties. The probability density functions $\mathcal{P}_{i,j,\text{high}}$ and $\mathcal{P}_{i,j,\text{low}}$ represent the 1-dimensional expected M^2_{miss} distributions in the lower M^2_{miss} region, and $o_{\text{NB,trafo}}$ distributions in the higher M^2_{miss} region.

In sum, the fit has 12 free parameters: D^{**} yields in all four reconstruction modes, lepton signal in all four reconstruction modes, lepton cross-feed in the reconstruction modes with a non-excited D meson, R, and R^*.

9.3 Factors

When performing the simultaneous fitting procedure, the yields of certain components are linked using a numerical factor, as descried in detail in Section 9.1. The numerical values of the used factors are given in Tables 9.2 and 9.6, separated by concerning sample and intersample. Most of the values were determined on 5 streams of generic Monte Carlo. The parameters for the wrong $D^{(*)}$ fit yield were determined from sidebands of real data.

Component	$D^0\ell^-$	$D^{*0}\ell^-$	$D^+\ell^-$	$D^{*+}\ell^-$	yield source
ℓ signal	✓	✓	✓	✓	fit
ℓ CF	✓	-	✓	-	fit
τ signal	✓	✓	✓	✓	fit
τ CF	✓	-	✓	-	constrained
wrong D	✓	-	✓	-	M_D sidebands
wrong D^*	-	✓	-	✓	ΔM_{D^*D} sidebands
D^{**}	✓	✓	✓	✓	fit
D_s	✓	✓	✓	✓	MC
wrong ℓ	✓	✓	✓	✓	MC
wrong charge	✓	-	-	-	constrained
rest	✓	✓	✓	✓	MC

Table 9.1: Fit components in each reconstruction sample. "fit" means, the yield is a free parameter in the fit. "constrained" means, the yield depends on other parameters. They are described in the text. "MC" means, the yields are taken from simulated data.

parameter	value	uncertainty	explanation
wrong D yield	350.7	$\pm$20.5	constant fit yield
D_s yield	22.0	$\pm$2.1	constant fit yield
rest yield	23.6	$\pm$2.2	constant fit yield
wrong ℓ yield	20.9	$\pm$2.0	constant fit yield

Table 9.2: Factors used for simultaneous fitting. All parameters concern components in the $D^+\ell^-$ sample.

parameter	value	uncertainty	explanation
wrong D^* yield	180.6	$\pm$31.3	constant fit yield
D_s yield	21.0	$\pm$2.1	constant fit yield
rest yield	4.3	$\pm$0.9	constant fit yield
wrong ℓ yield	13.7	$\pm$1.7	constant fit yield

Table 9.3: Factors used for simultaneous fitting. All parameters concern components in the $D^0\ell^-$ sample.

parameter	value	uncertainty	explanation
wrong charge factor	0.107	$\pm$0.004	to constrain wrong charge background
wrong D yield	1334.9	$\pm$64.3	constant fit yield
D_s yield	111.9	$\pm$4.7	constant fit yield
rest yield	76.5	$\pm$3.9	constant fit yield
wrong ℓ yield	68.7	$\pm$3.7	constant fit yield

Table 9.4: Factors used for simultaneous fitting. All parameters concern components in the $D^{*+}\ell^-$ sample.

9.4 Shapes

The shapes for the individual components described in Section 9.1 were determined on 5 streams of generic Monte Carlo. An exception is the τ signal shape, that was determined on

parameter	value	uncertainty	explanation
wrong D^* yield	2216.7	$\pm$52.9	constant fit yield
D_s yield	20.7	$\pm$2.0	constant fit yield
rest yield	4.2	$\pm$0.9	constant fit yield
wrong ℓ yield	12.9	$\pm$1.6	constant fit yield

Table 9.5: Factors used for simultaneous fitting. All parameters concern components in the $D^{*0}\ell^-$ sample.

parameter	value	uncertainty	explanation
f_{R,B^0}	1.69	$\pm$0.09	efficiency ratio proportionality factor
f_{R,B^+}	1.91	$\pm$0.06	efficiency ratio proportionality factor
f^*_{R,B^0}	3.11	$\pm$0.13	efficiency ratio proportionality factor
f^*_{R,B^+}	3.63	$\pm$0.09	efficiency ratio proportionality factor
g_{B^0}	0.83	$\pm$0.08	τ cross-feed proportionality factor
g_{B^+}	0.69	$\pm$0.04	τ cross-feed proportionality factor

Table 9.6: Factors used for simultaneous fitting. All parameters concern components in several samples.

signal Monte Carlo in order to greatly increase the available statistics in for this component. They can be found in Section A in the Appendix, overlayed with data points that were used in the shape extraction.

For better illustration, the PDFs are additionally plotted in Figs. 9.1 to 9.4 – one page per data sample – using the same color codes as in composition- and result plots. The left plot in each subfigure shows the squared missing mass distribution of the respective component below $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$, the middle plot shows the squared missing mass above this threshold, and the right plot shows the $o_{\mathrm{NB,trafo}}$ distribution also above the threshold. From the two plots of the squared missing mass one can get an impression of how much of the component lies above or below the threshold.

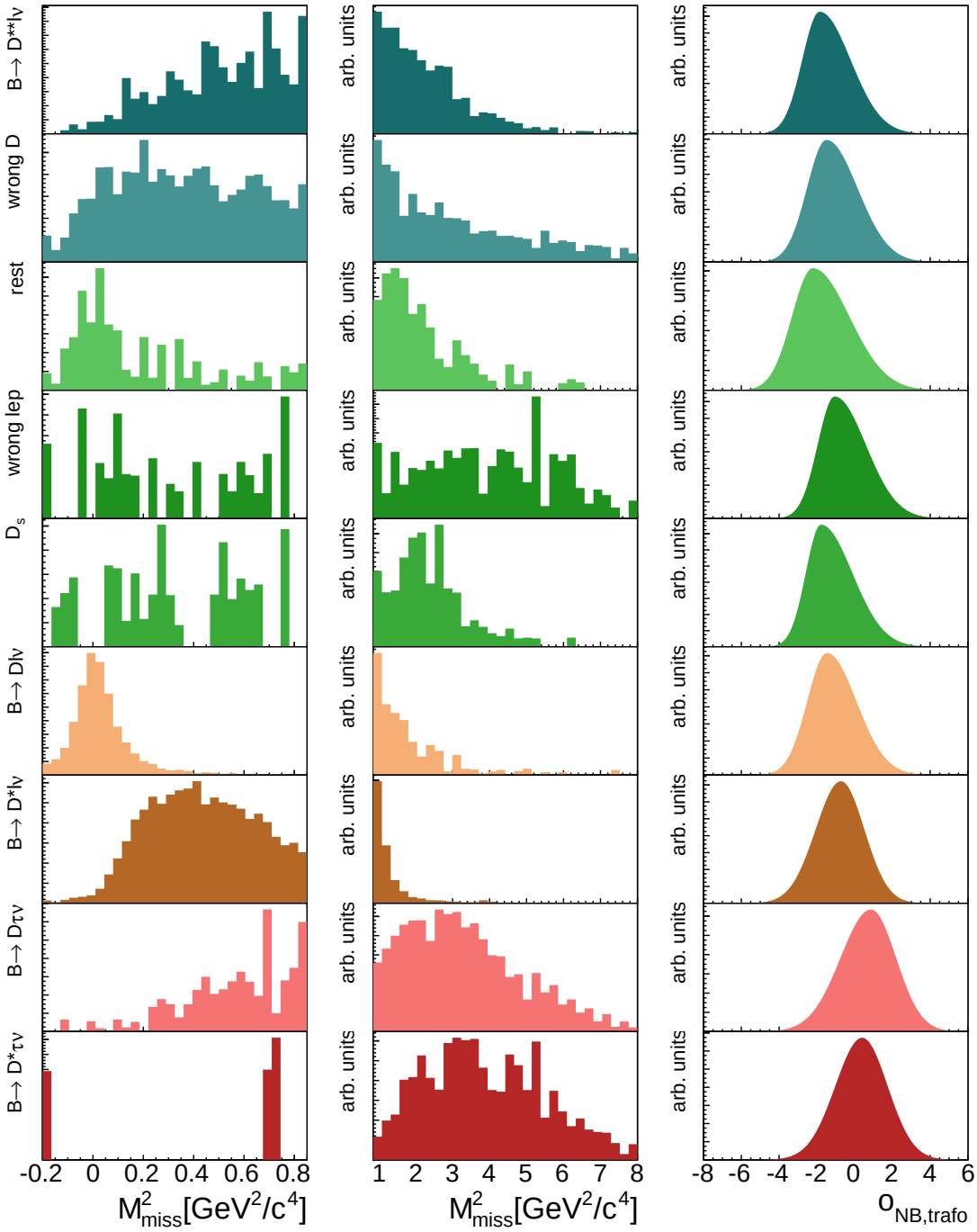


Figure 9.1: Comparison of the probability density functions of all components in the $D^+\ell^-$ sample. The left and center plots give the distribution of the squared missing mass below and above $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$. The right hand plot gives the $o_{\mathrm{NB,trafo}}$ PDF, which is only taken from the upper region.

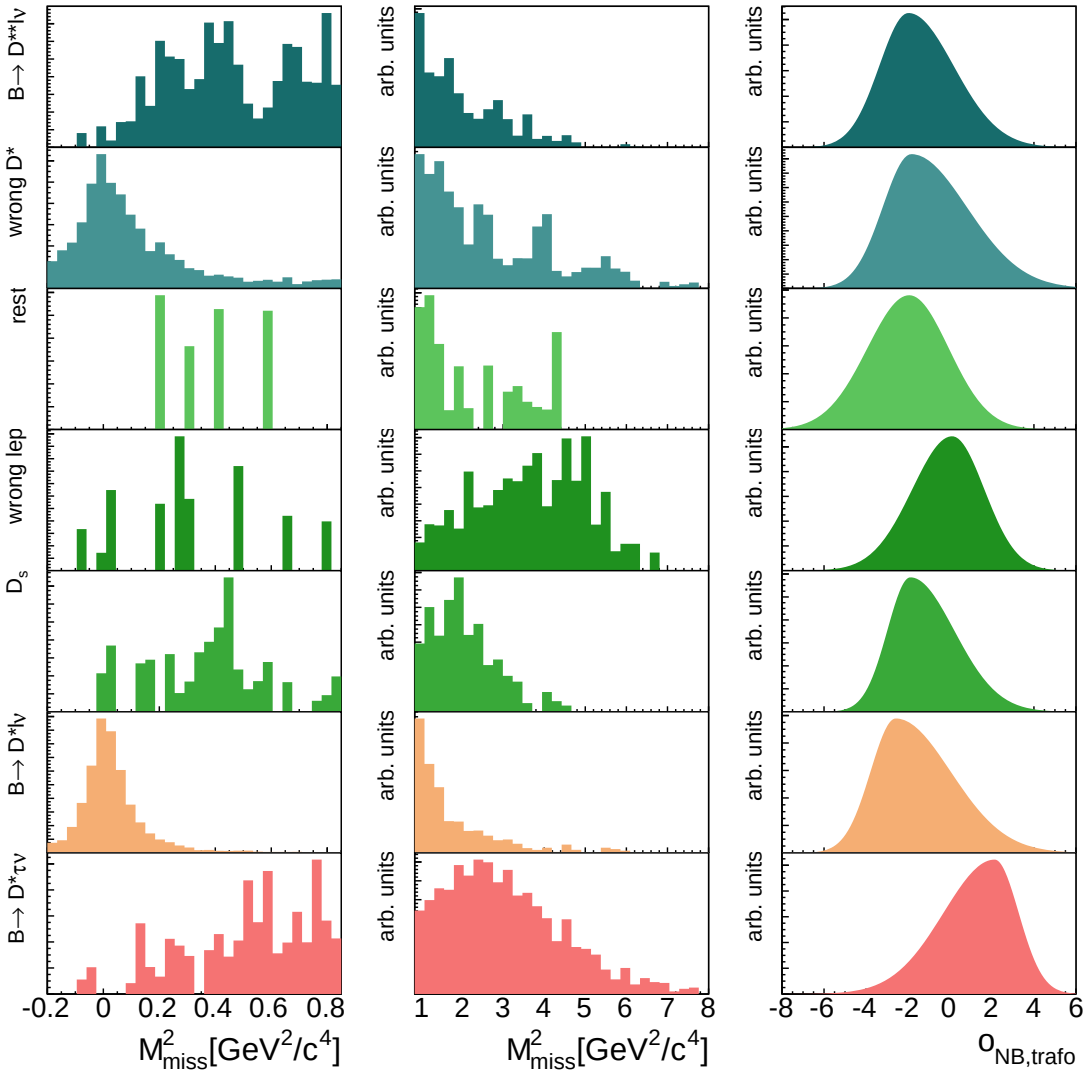


Figure 9.2: Comparison of the probability density functions of all components in the $D^{*+}\ell^-$ sample. The left and center plots give the distribution of the squared missing mass below and above $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$. The right hand plot gives the $o_{\mathrm{NB,trafo}}$ PDF, which is only taken from the upper region.

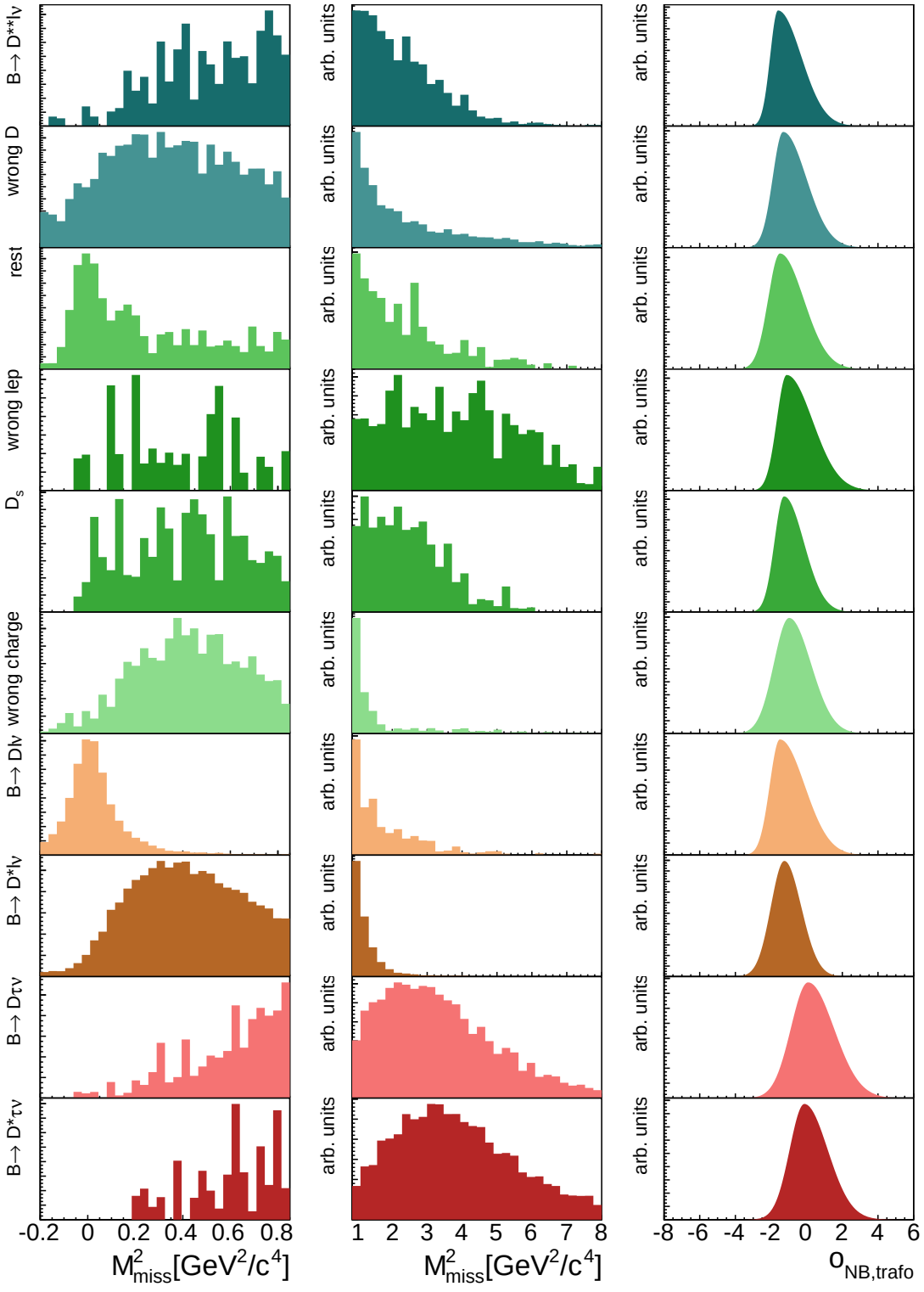


Figure 9.3: Comparison of the probability density functions of all components in the $D^0\ell^-$ sample. The left and center plots give the distribution of the squared missing mass below and above $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$. The right hand plot gives the $o_{\mathrm{NB,trafo}}$ PDF, which is only taken from the upper region.

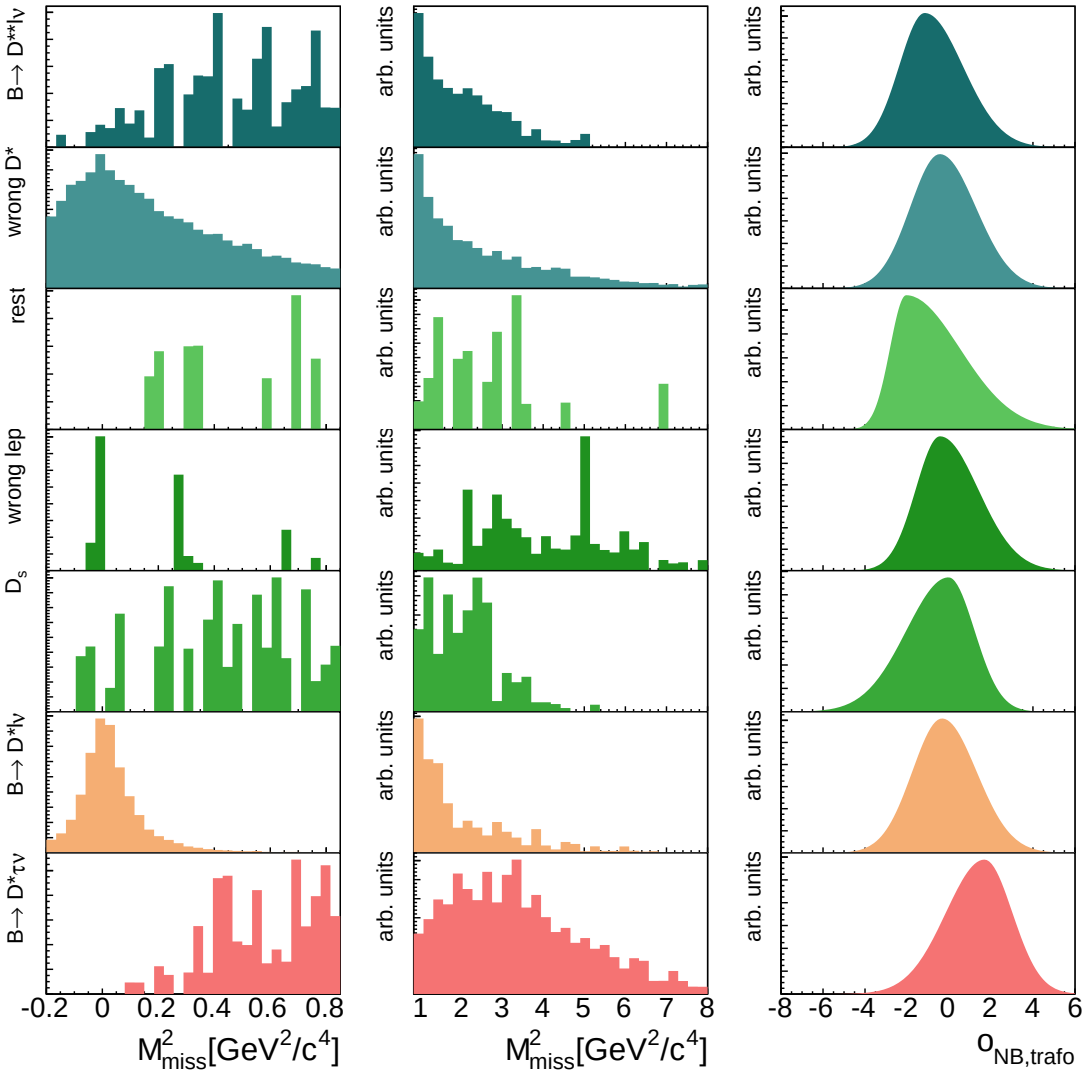


Figure 9.4: Comparison of the probability density functions of all components in the $D^{*0}\ell^-$ sample. The left and center plots give the distribution of the squared missing mass below and above $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$. The right hand plot gives the $o_{\mathrm{NB,trafo}}$ PDF, which is only taken from the upper region.

9.5 Expected Yields

Based on the 5 available complete streams of generic Monte Carlo, the expected yields of the components are calculated. They are listed in Table 9.7. These numbers were produced on a sample where all described reweighting procedures have been applied. The numbers for *wrong* $D^{(*)}$ background are the ones extrapolated from the sidebands of real data.

	$D^+\ell^-$	$D^0\ell^-$	$D^{*+}\ell^-$	$D^{*0}\ell^-$
ℓ signal	869.7	2293.6	1682.1	2279.3
ℓ CF	972.7	7444.1	-	-
τ signal	85.9	200.6	76.8	109.5
τ CF	37.0	247.0	-	-
D^{**}	133.1	208.3	76.0	40.2
wrong $D^{(*)}$	350.7	1334.9	180.6	2216.7
wrong ℓ	20.9	68.7	13.7	12.9
D_s	22.0	111.9	21.0	20.7
wrong charge	-	179.8	-	-
rest	23.6	76.5	4.3	4.2

Table 9.7: Yield expectations for the individual components in the different reconstruction samples.

10. Validation of the Fitting Procedure

10.1 Fit Projections

For better visibility, the background components coming from D_s, *wrong leptons*, *wrong charge*, *wrong* $D^{(*)}$, and *rest* are combined in the green *rest* component in the fit projections. Neither of them has its yield floating in the fit, so the exact contribution of each component is not enlightening. The *wrong* $D^{(*)}$ component is the dominant background in all data samples.

Fit Dimensions

The fit procedure is applied on one stream of simulated data, and the projections of the fitting ranges are given in Figs. 10.1 and 10.2.

Other Dimensions

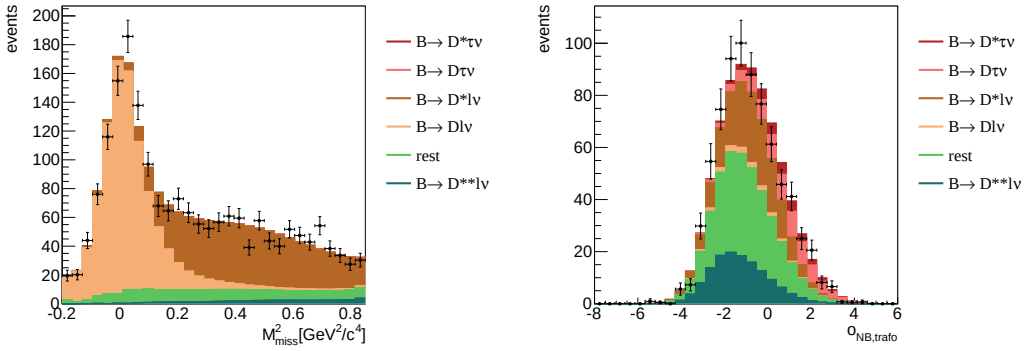
The fit results are also projected on other observables of interest. Figs. 10.3 to 10.6 show the distributions in E_{ECL}, p_ℓ^*, q^2, and M^2_{miss} respectively, including also the M^2_{miss} distribution in the higher M^2_{miss} region. These observables are all significant inputs to the neural network, with E_{ECL} being by far the most influential.

Signal Enhanced Projections

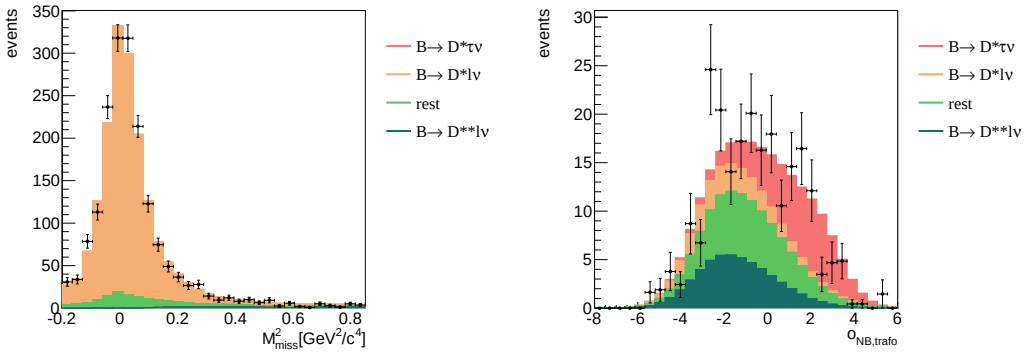
In order to get projections with enhanced τ signal, the fit results are projected again, using a cut of $M^2_{\mathrm{miss}} > 2.0\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$, to suppress components with short tails. Adding an additional cut on $o_{\mathrm{NB,trafo}}$ does not give much improvement, as most observables are highly correlated and cutting away background in $o_{\mathrm{NB,trafo}}$ usually results in removing background events in regions with low τ concentration. Figs. 10.7 to 10.10 show the distributions in E_{ECL}, p_ℓ^*, q^2, and M^2_{miss} respectively.

Conclusion

The fit projections look reasonable in all distributions and do not indicate any obvious technical issues in the fitting procedure. Furthermore, fit projections with increased signal purity can be created, that facilitate the identification of the signal contribution to the observer.

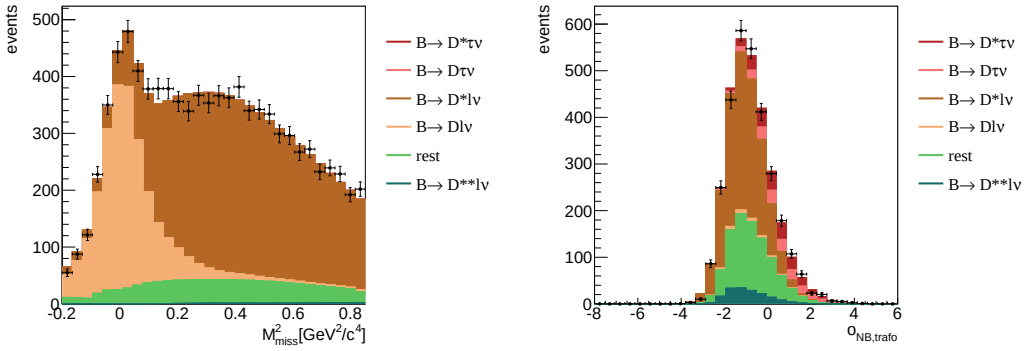


(a) M^2_{miss} in $D^+\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^+\ell^-$

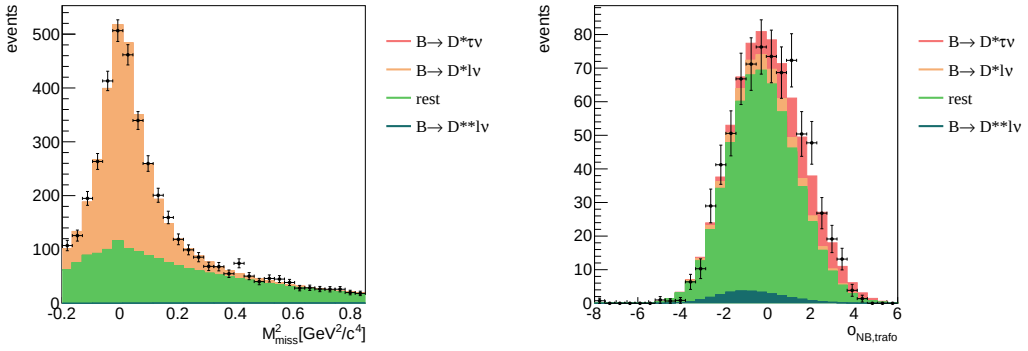


(c) M^2_{miss} in $D^{*+}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*+}\ell^-$

Figure 10.1: Fit projections of the full fit on Monte Carlo for the B^0 reconstruction modes. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.

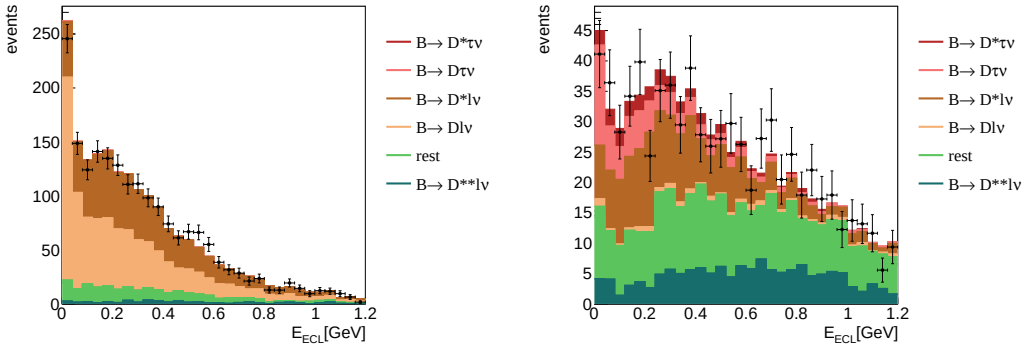


(a) M^2_{miss} in $D^0\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^0\ell^-$



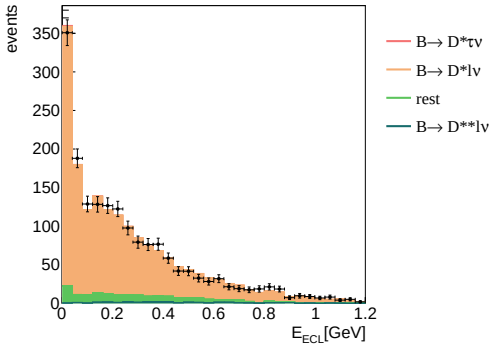
(c) M^2_{miss} in $D^{*0}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*0}\ell^-$

Figure 10.2: Fit projections of the full fit on Monte Carlo for the B^+ reconstruction modes. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.

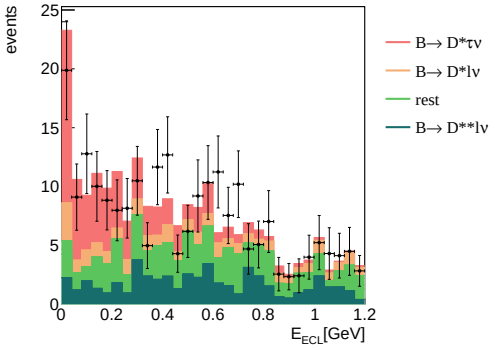


(a) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$

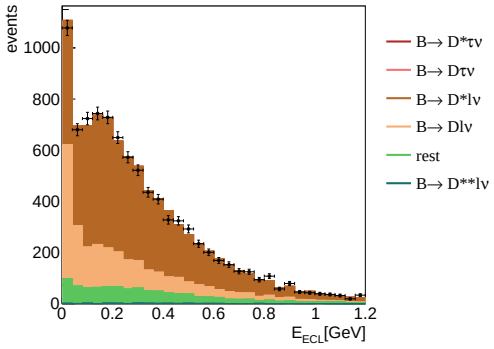
(b) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$



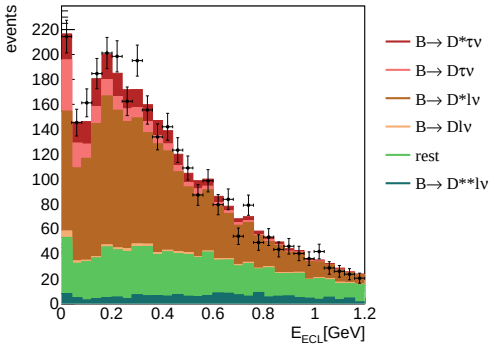
(c) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$



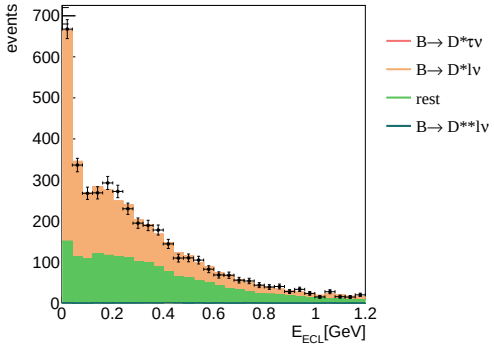
(d) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$



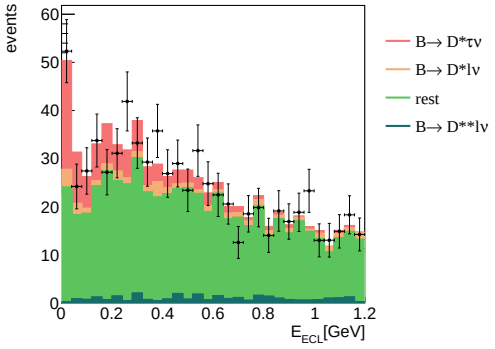
(e) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$



(f) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$

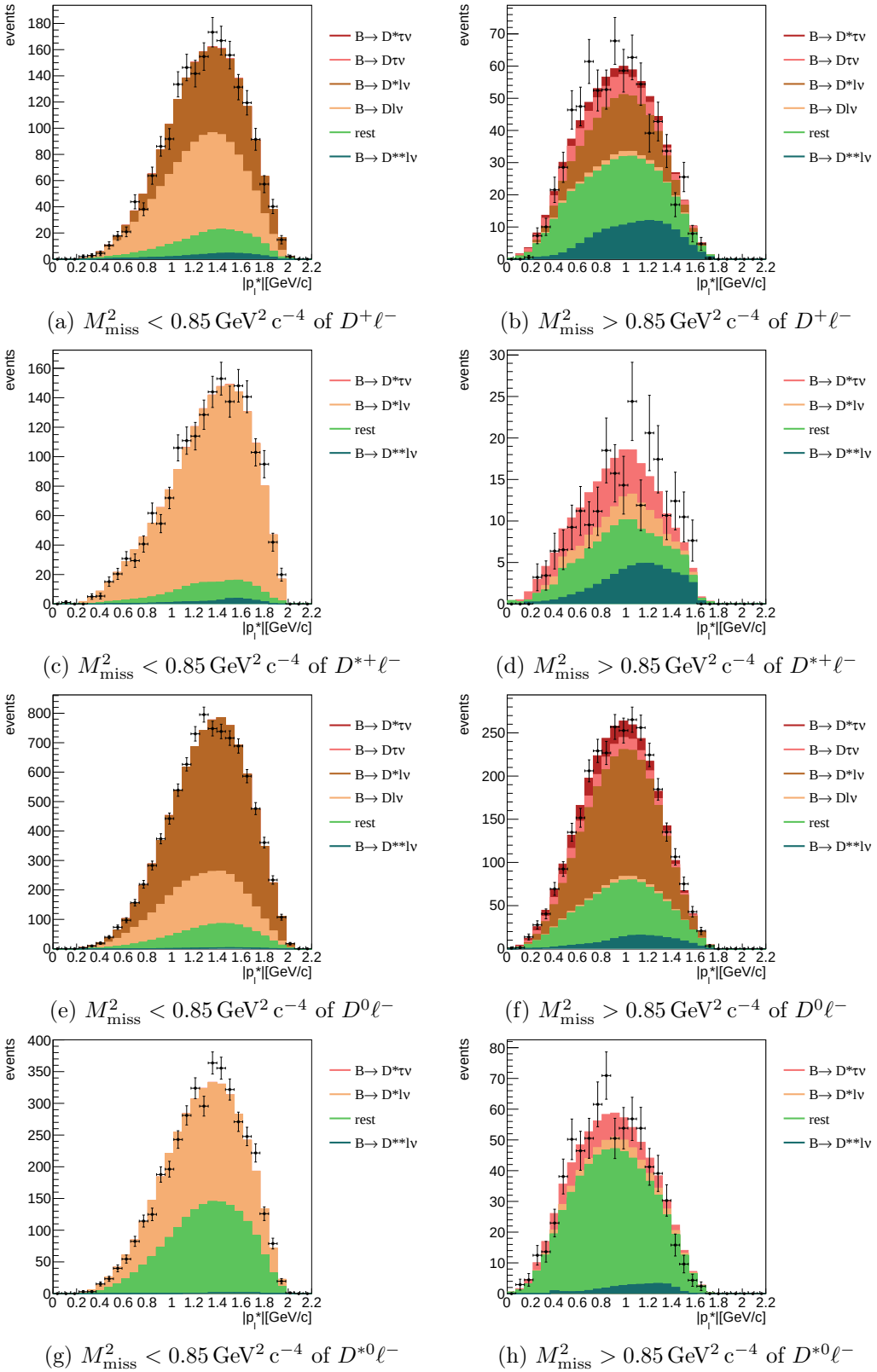


(g) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$



(h) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

Figure 10.3: Fit projections on E_{ECL} of the full fit on Monte Carlo. Left is the lower M^2_{miss} region, right the higher one. Each row is a different dataset.



(a) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$

(b) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$

(c) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$

(d) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$

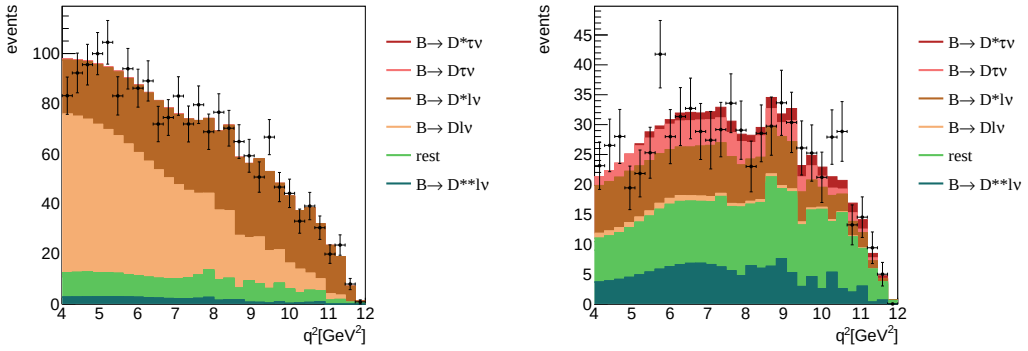
(e) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$

(f) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$

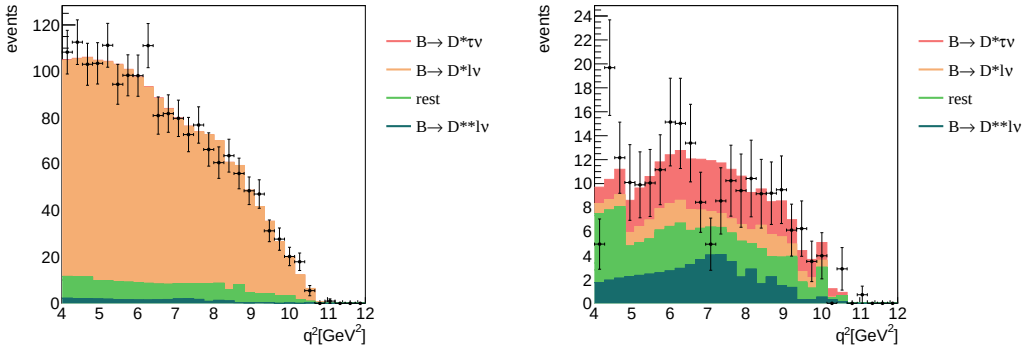
(g) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

(h) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

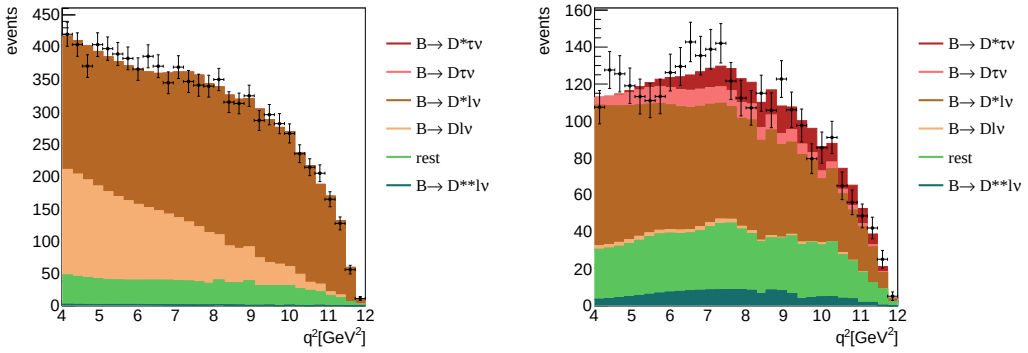
Figure 10.4: Fit projections on p^*_ℓ of the full fit on Monte Carlo. Left is the lower M^2_{miss} region, right the higher one. Each row is a different dataset.



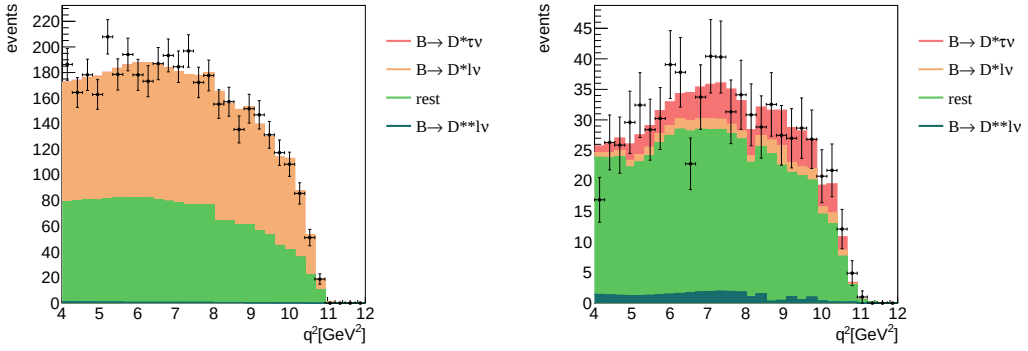
(a) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$ (b) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$



(c) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$ (d) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$

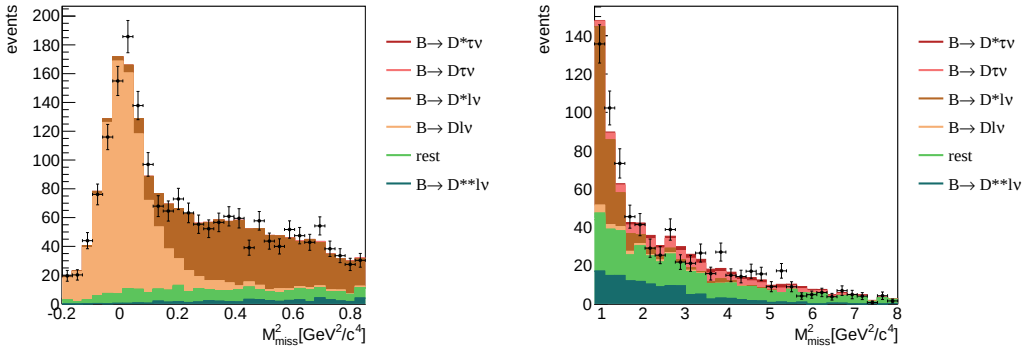


(e) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$ (f) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$



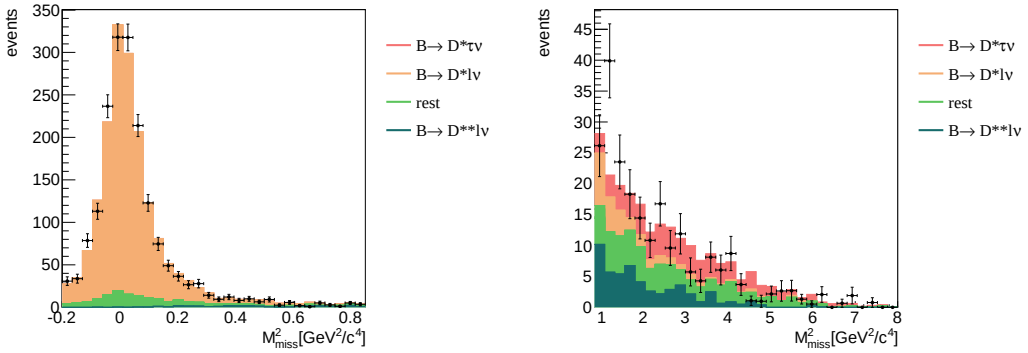
(g) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$ (h) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

Figure 10.5: Fit projections on q^2 of the full fit on Monte Carlo. Left is the lower M^2_{miss} region, right the higher one. Each row is a different dataset.



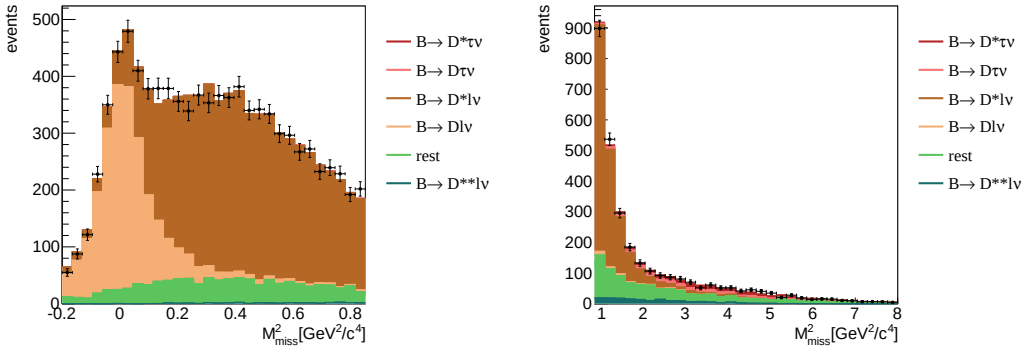
(a) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$

(b) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^+\ell^-$



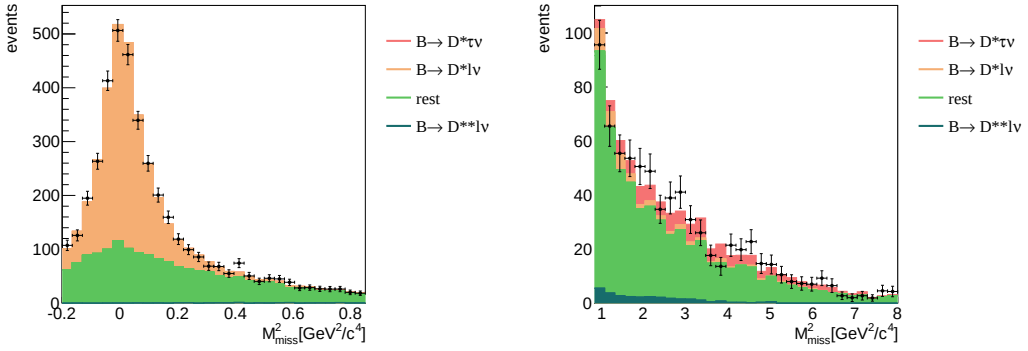
(c) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$

(d) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*+}\ell^-$



(e) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$

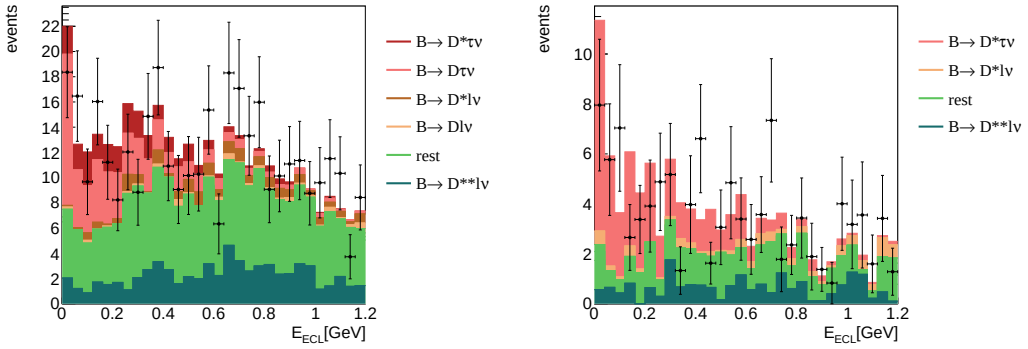
(f) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^0\ell^-$



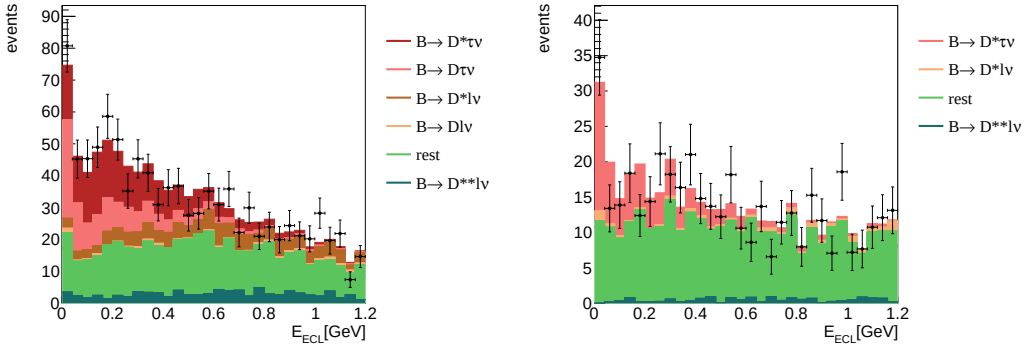
(g) $M^2_{\text{miss}} < 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

(h) $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ of $D^{*0}\ell^-$

Figure 10.6: Fit projections on M^2_{miss} of the full fit on Monte Carlo. Left is the lower M^2_{miss} region, right the higher one. Each row is a different dataset.

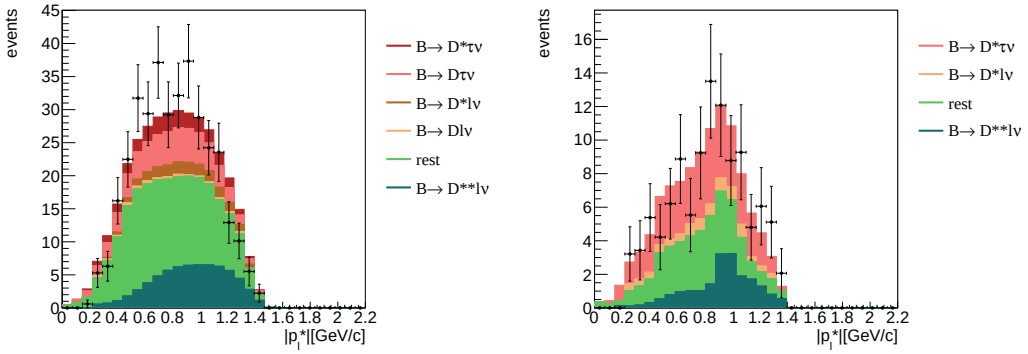


(a) E_{ECL} in $D^{+}\ell^{-}$ (b) E_{ECL} in $D^{*+}\ell^{-}$

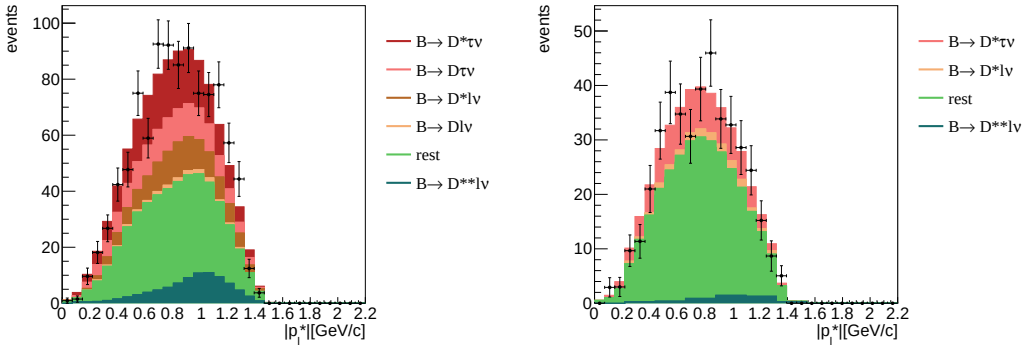


(c) E_{ECL} in $D^{0}\ell^{-}$ (d) E_{ECL} in $D^{*0}\ell^{-}$

Figure 10.7: Fit projections on E_{ECL} of the full fit on Monte Carlo. The four datasets are displayed with a cut on $M^2_{\text{miss}} > 2.0\,\text{GeV}^2\,\text{c}^{-4}$ to enhance the τ signal.

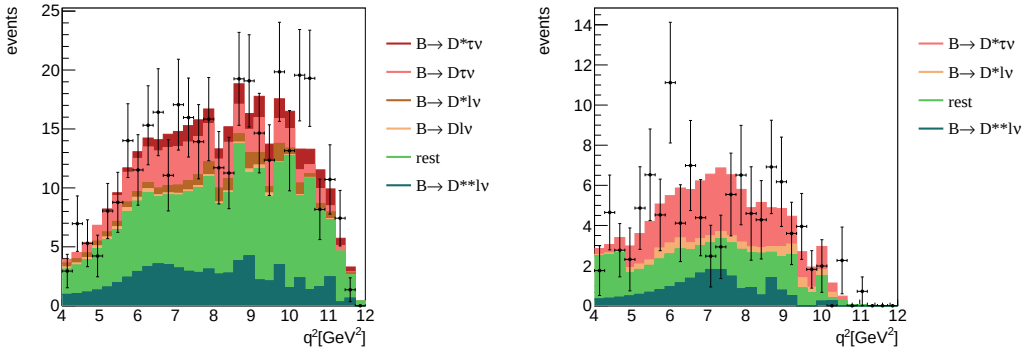


(a) p^*_ℓ in $D^{+}\ell^{-}$ (b) p^*_ℓ in $D^{*+}\ell^{-}$

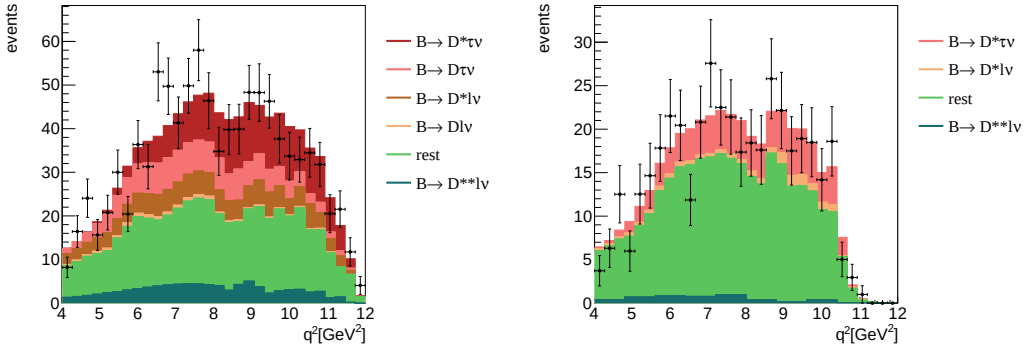


(c) p^*_ℓ in $D^{0}\ell^{-}$ (d) p^*_ℓ in $D^{*0}\ell^{-}$

Figure 10.8: Fit projections on p^*_ℓ of the full fit on Monte Carlo. The four datasets are displayed with a cut on $M^2_{\text{miss}} > 2.0\,\text{GeV}^2\,\text{c}^{-4}$ to enhance the τ signal.

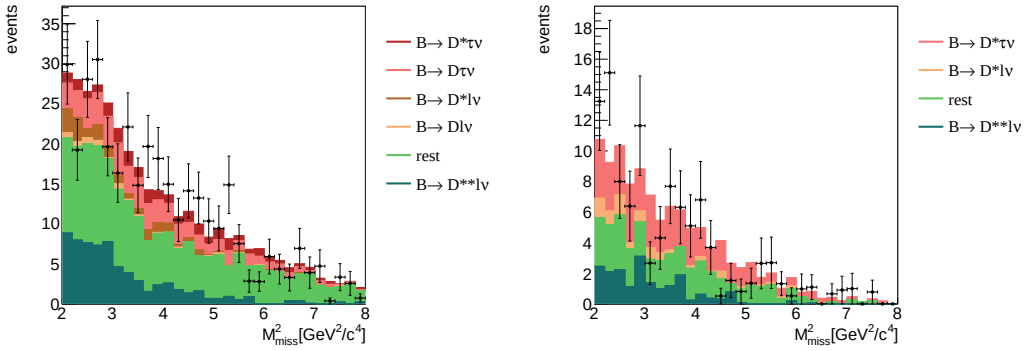


(a) q^2 in $D^+\ell^-$ (b) q^2 in $D^{*+}\ell^-$

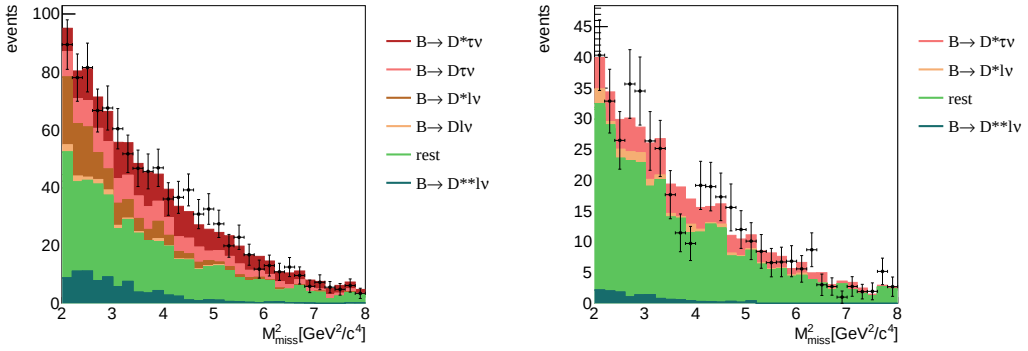


(c) q^2 in $D^0\ell^-$ (d) q^2 in $D^{*0}\ell^-$

Figure 10.9: Fit projections on q^2 of the full fit on Monte Carlo. The four datasets are displayed with a cut on $M^2_{\text{miss}} > 2.0\,\text{GeV}^2\,\text{c}^{-4}$ to enhance the τ signal.



(a) M^2_{miss} in $D^+\ell^-$ (b) M^2_{miss} in $D^{*+}\ell^-$



(c) M^2_{miss} in $D^0\ell^-$ (d) M^2_{miss} in $D^{*0}\ell^-$

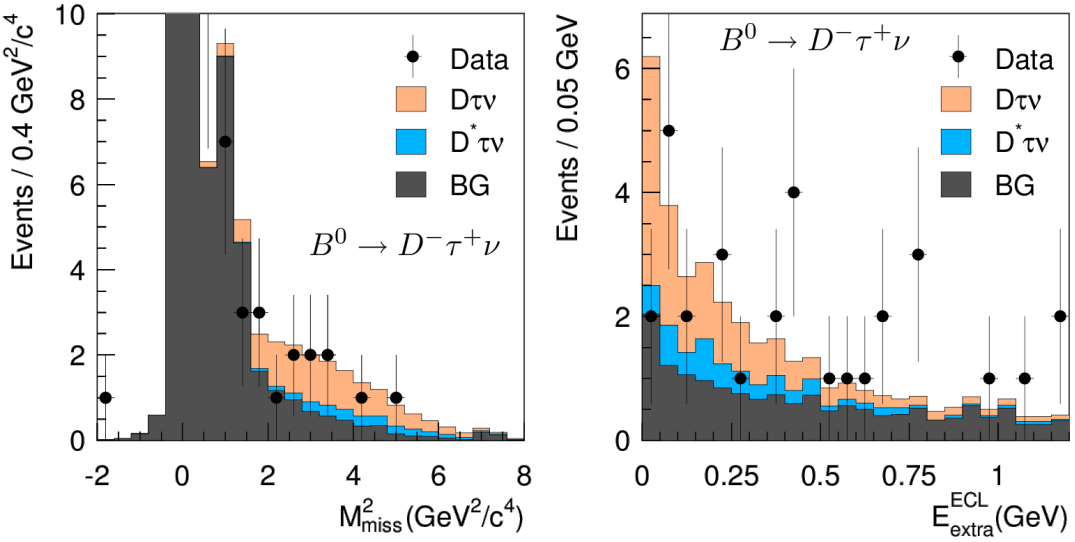
Figure 10.10: Fit projections on M^2_{miss} of the full fit on Monte Carlo. The four datasets are displayed with a cut on $M^2_{\text{miss}} > 2.0\,\text{GeV}^2\,\text{c}^{-4}$ to enhance the τ signal.

Comparison to Previous Belle Analysis

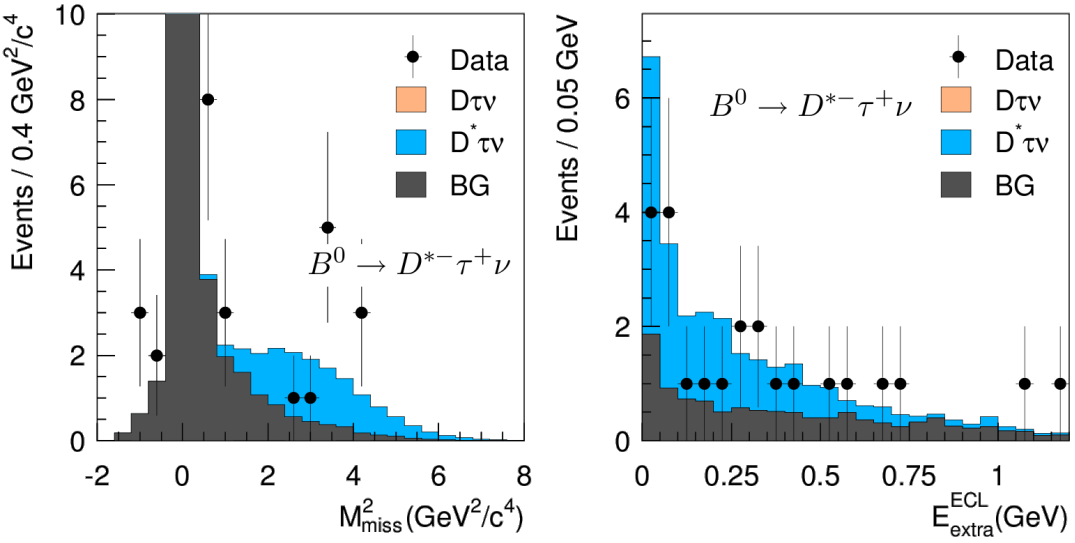
Comparing these distributions to the previous [4] Belle analysis' in Figs. 10.11 and 10.12, it is obvious, that the signal-to-background ratio in this thesis is clearly worse. First of all, it is important to know the differences among the plots. In the missing mass frame, the previous analysis applies additional cuts ($E_{\mathrm{ECL}} < 0.2\,\mathrm{GeV}$ and $p^*_\ell < 1.2\,\mathrm{GeV\,c^{-1}}$) that greatly enhance the tau signal. However, the E_{ECL} plots are comparable, as they have cuts on M^2_{miss} that are comparable to ours (i.e. $M^2_{\mathrm{miss}} > 2.0\,\mathrm{GeV^2\,c^{-4}}$ for D and $M^2_{\mathrm{miss}} > 1.6\,\mathrm{GeV^2\,c^{-4}}$ for D^*). The additional cut on p^*_ℓ has only a very small effect in this M^2_{miss} region, as can be seen in our p^*_ℓ distributions for the high M^2_{miss} sample. So the difference in background yield is still clearly visible.

There are differences in the reconstruction, namely in the γ selection (concerning π^0 in general and $D^* \to D\gamma$) and the signal side channels. Other differences, like the exact signal side modes, are rather small. On the tag side, we use all `ekpfullrecon` channels, which include several channels with higher numbers of pions. The previous analysis only uses these modes, if certain intermediate states (ρ, a_1) are properly identified. So the basic reconstruction of this analysis has a lower purity. It can be varied to the cost of efficiency by varying the cuts on the signal side network, which already have been optimized for significance. The stricter requirements on γ energies can also affect the background yield, especially of the *wrong* $D^{(*)}$ component. I have tested some additional highly specialized cuts on E_γ with a high signal efficiency ($> 90\%$): Increasing the threshold for events with several π^0 on the signal side and generally for $D^* \to D\gamma$ can reduce the *wrong* $D^{(*)}$ component by $15 - 40\%$. But toy studies imply, that this does not improve our expected precision, so the original selection method is kept.

I conclude that the worse signal purity in this thesis is not the effect of erroneous reconstruction, but intentional, as the efficiency gain benefits the expected result.

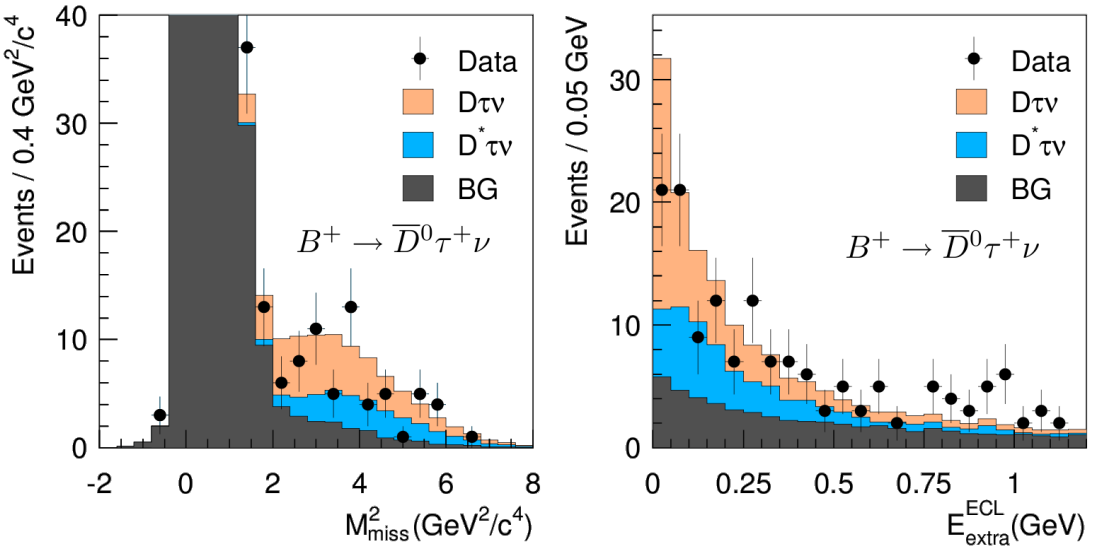


(a) $D^+\ell^-$

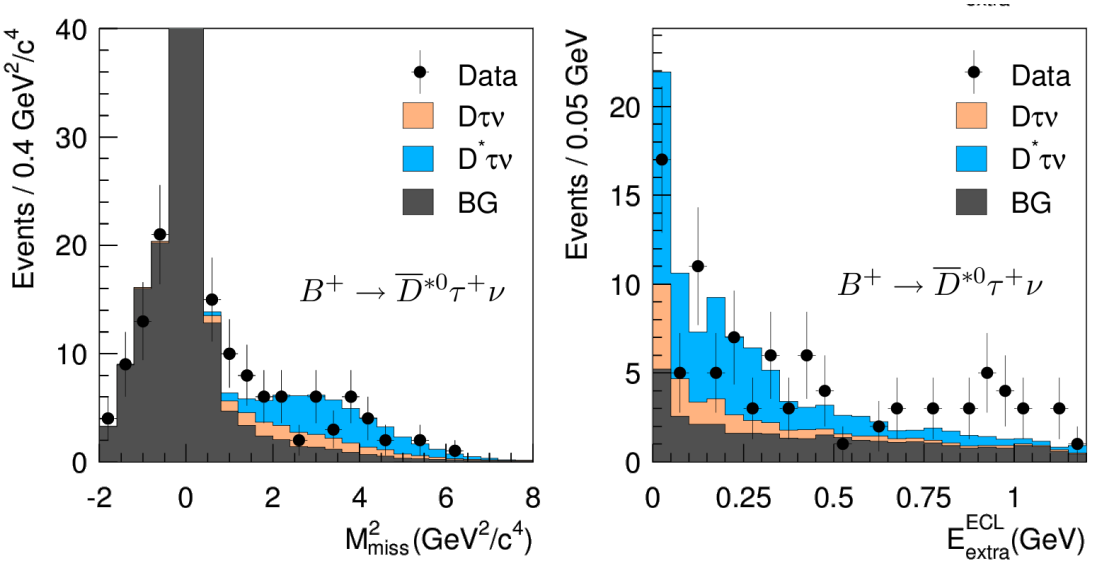


(b) $D^{*+}\ell^-$

Figure 10.11: Fit projections in the neutral B meson samples of the first Belle analysis with hadronic tagging.



(a) $D^0\ell^-$



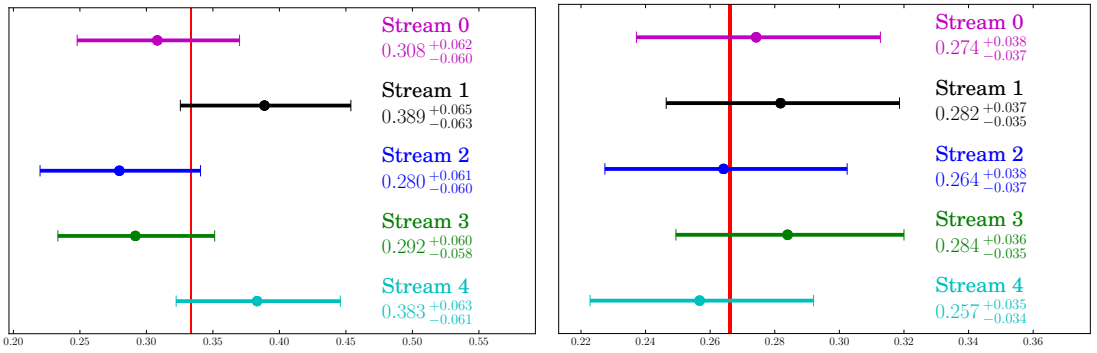
(b) $D^{*0}\ell^-$

Figure 10.12: Fit projections in the charged B meson sample of the first Belle analysis with hadronic tagging.

10.2 Cross Validation

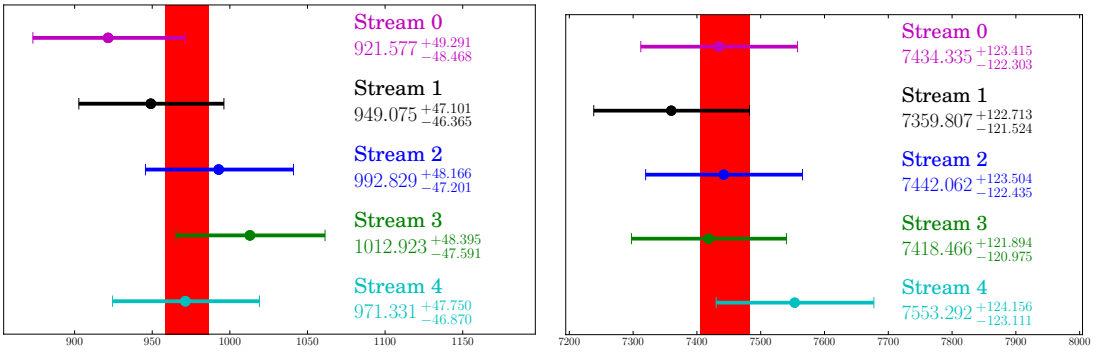
To further validate the fitting procedure, it can be applied on the five streams of Monte Carlo in a way, that one stream is treated as data sample, and the other four are used to extract constants and shapes. Thus we get five results for each free fit parameter and can check, if they are compatible with the truth within their uncertainties.

The results of this procedure are summarized in Figs. 10.13 to 10.16.



(a) R (b) R^*

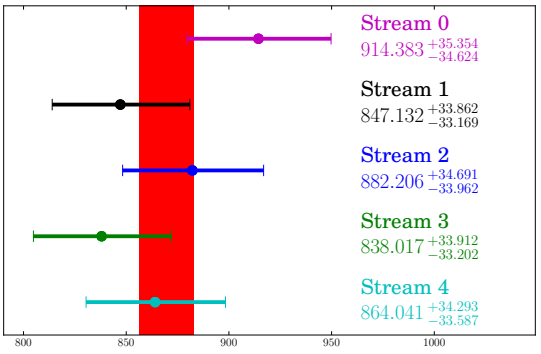
Figure 10.13: Summary of the results of the fits to generic Monte Carlo for R and R^*.



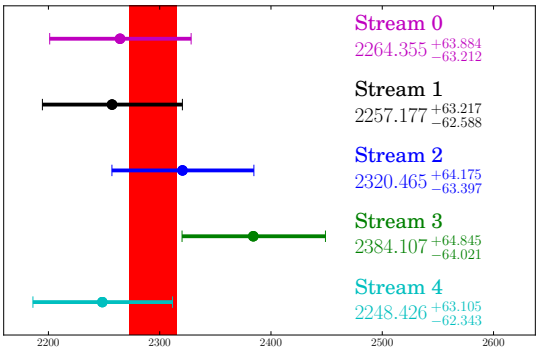
(a) ℓ CF in $D^+\ell^-$ (b) ℓ CF in $D^0\ell^-$

Figure 10.14: Summary of the results of the fits to generic Monte Carlo for ℓ *cross-feed* yields in different samples.

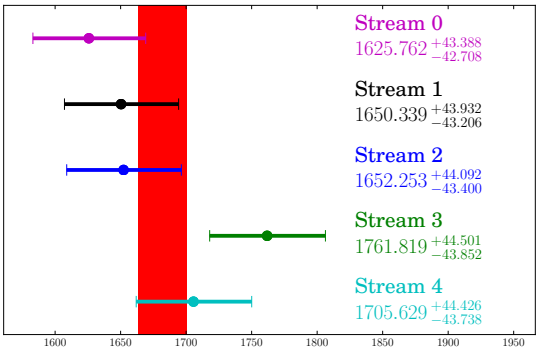
The results mostly follow an expected distribution around the true value. The yields for the D^{**} component deviate a bit more than would be expected from the statistical errors. However, this effect mostly comes from fluctuations of the much larger *wrong* $D^{(*)}$ component, which will be treated in the systematic uncertainties.



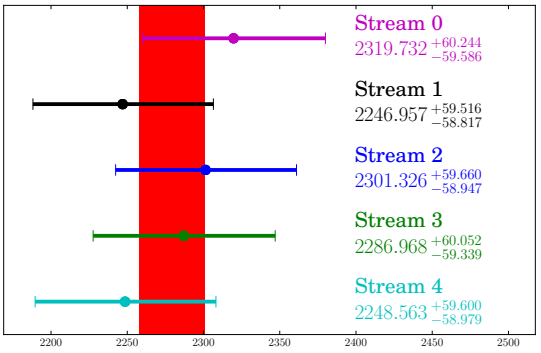
(a) ℓ signal in $D^+\ell^-$



(b) ℓ signal in $D^0\ell^-$

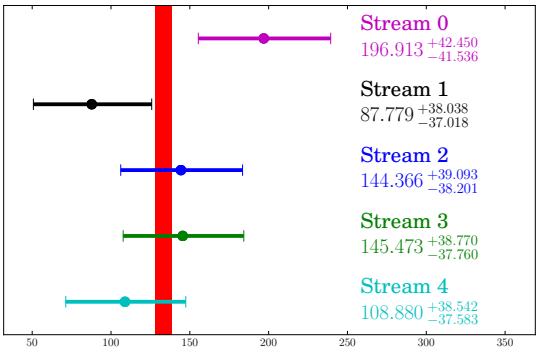


(c) ℓ signal in $D^{*+}\ell^-$

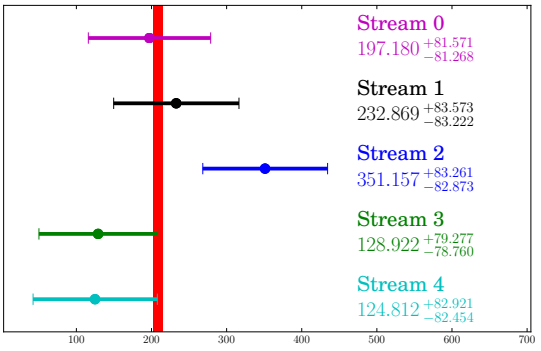


(d) ℓ signal in $D^{*0}\ell^-$

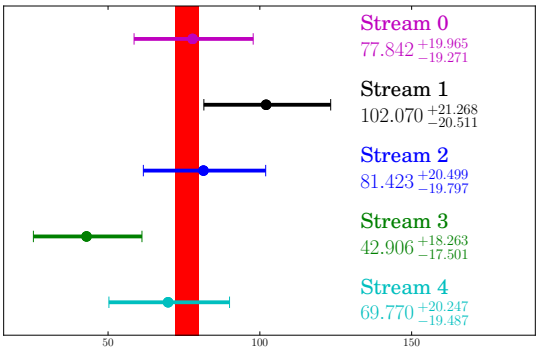
Figure 10.15: Summary of the results of the fits to generic Monte Carlo for ℓ *signal* yields in different samples.



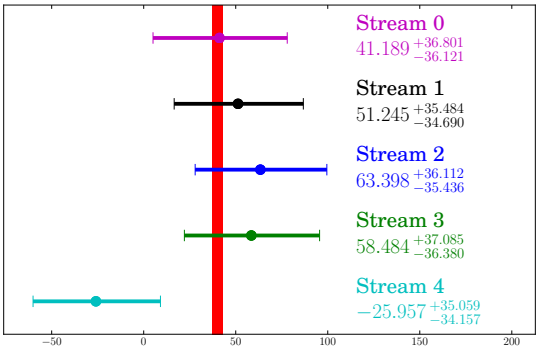
(a) D^{**} background in $D^+\ell^-$



(b) D^{**} background in $D^0\ell^-$



(c) D^{**} background in $D^{*+}\ell^-$

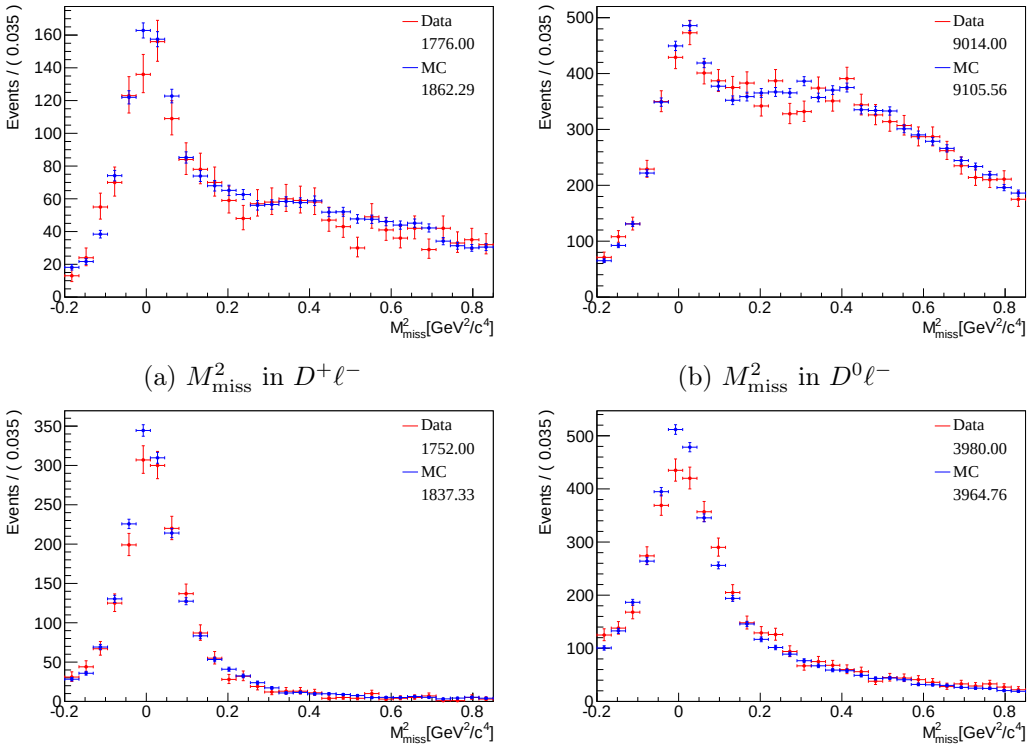


(d) D^{**} background in $D^{*0}\ell^-$

Figure 10.16: Summary of the results of the fits to generic Monte Carlo for D^{**} *background* yields in different samples.

10.3 Data - Monte Carlo Comparison

A fitting procedure, that is working on simulated data can still fail on real data, if there are large differences. Therefore the distributions of the fit variables and three control variables, that are also important inputs to the neural network, are compared on data and 5 streams of Monte Carlo. The comparison is done separately in the lower and higher mass frame, and additionally in the $D^{(*)}$ sidebands. The comparisons are shown in Figs. 10.17 to 10.28, where the histograms for Monte Carlo data are scaled down by factor of 5. While there are visible differences in yields, the comparison does not show significant deviations of the shapes. We are only looking out for any uncovered structures, and the τ *signal* contribution is rather low, so I do not risk to induce a bias by correcting simulated data based on these observations.



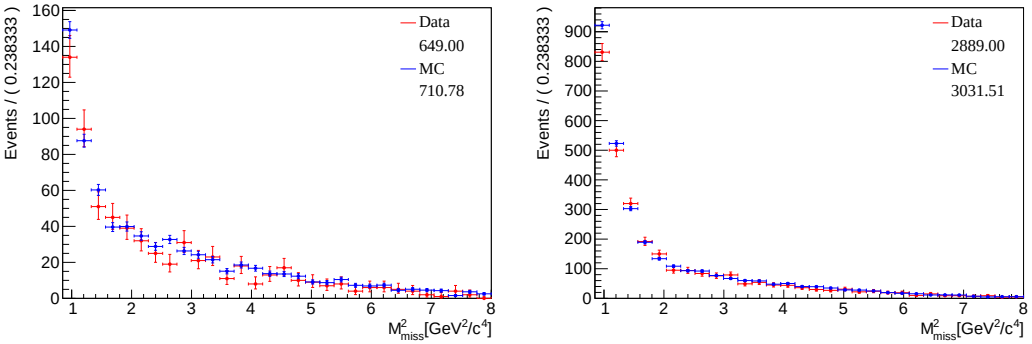
(a) M^2_{miss} in $D^+\ell^-$

(b) M^2_{miss} in $D^0\ell^-$

(c) M^2_{miss} in $D^{*+}\ell^-$

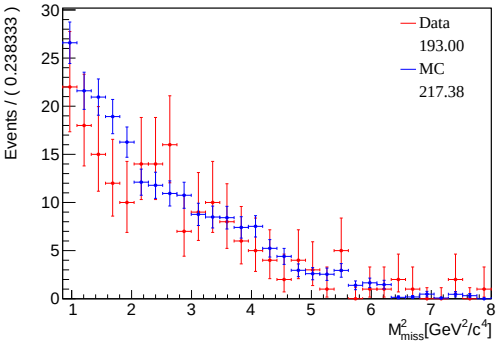
(d) M^2_{miss} in $D^{*0}\ell^-$

Figure 10.17: Comparison of M^2_{miss} distributions of Data and Monte Carlo in the lower M^2_{miss} region.

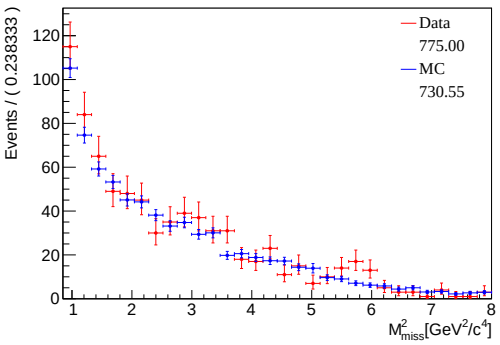


(a) M^2_{miss} in $D^+\ell^-$

(b) M^2_{miss} in $D^0\ell^-$

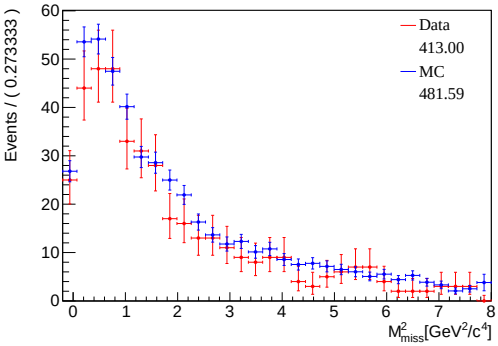


(c) M^2_{miss} in $D^{*+}\ell^-$

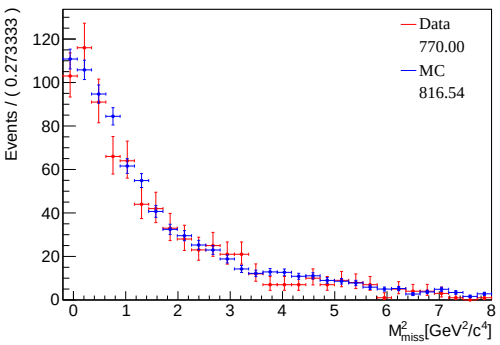


(d) M^2_{miss} in $D^{*0}\ell^-$

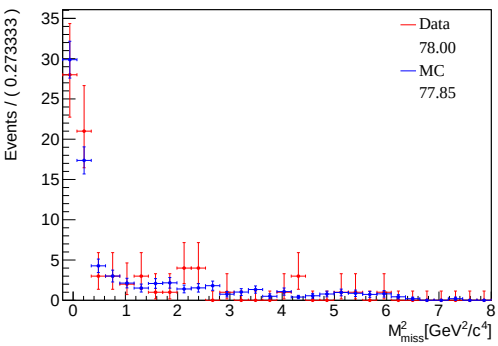
Figure 10.18: Comparison of M^2_{miss} distributions of Data and Monte Carlo in the higher M^2_{miss} region.



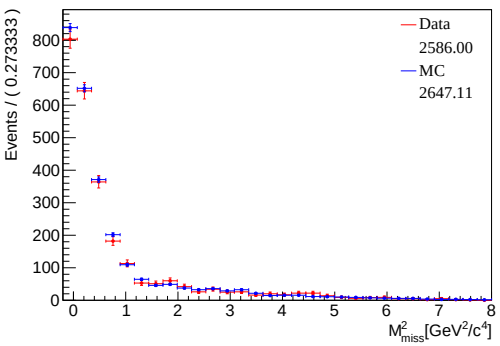
(a) M^2_{miss} in $D^+\ell^-$



(b) M^2_{miss} in $D^0\ell^-$

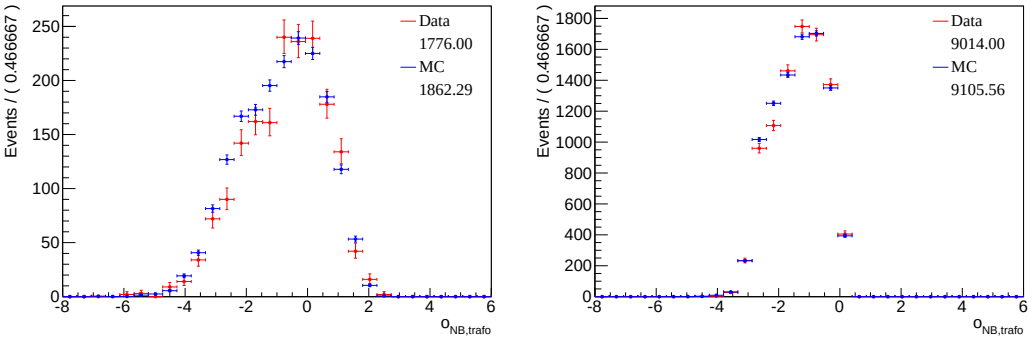


(c) M^2_{miss} in $D^{*+}\ell^-$



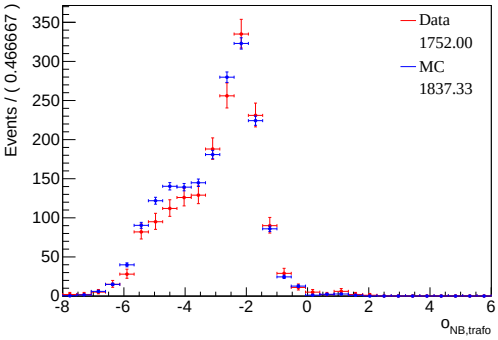
(d) M^2_{miss} in $D^{*0}\ell^-$

Figure 10.19: Comparison of M^2_{miss} distributions of Data and Monte Carlo in the $D^{(*)}$ sidebands.

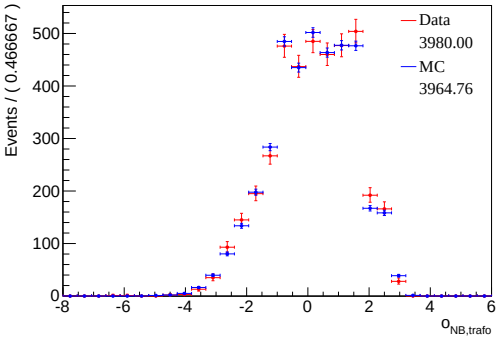


(a) $o_{\mathrm{NB,trafo}}$ in $D^+\ell^-$

(b) $o_{\mathrm{NB,trafo}}$ in $D^0\ell^-$

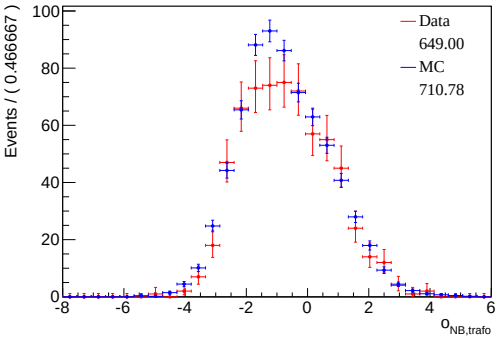


(c) $o_{\mathrm{NB,trafo}}$ in $D^{*+}\ell^-$

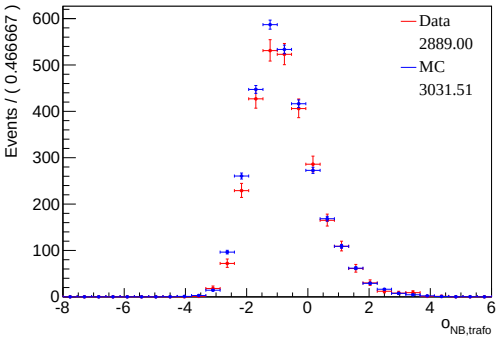


(d) $o_{\mathrm{NB,trafo}}$ in $D^{*0}\ell^-$

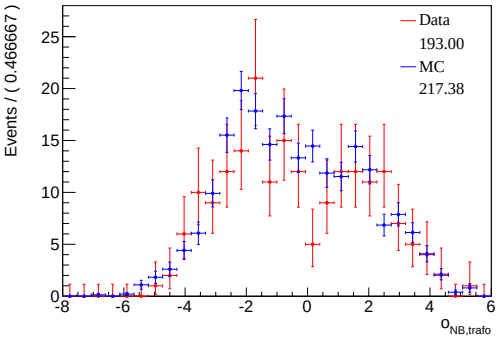
Figure 10.20: Comparison of $o_{\mathrm{NB,trafo}}$ distributions of Data and Monte Carlo in the lower M^2_{miss} region.



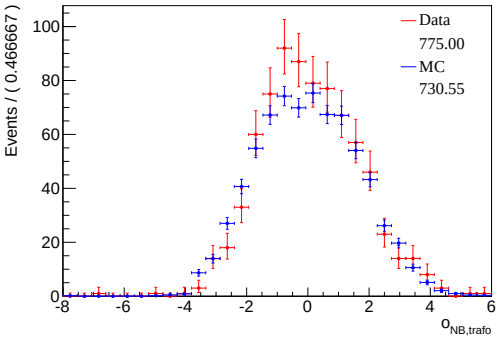
(a) $o_{\mathrm{NB,trafo}}$ in $D^+\ell^-$



(b) $o_{\mathrm{NB,trafo}}$ in $D^0\ell^-$

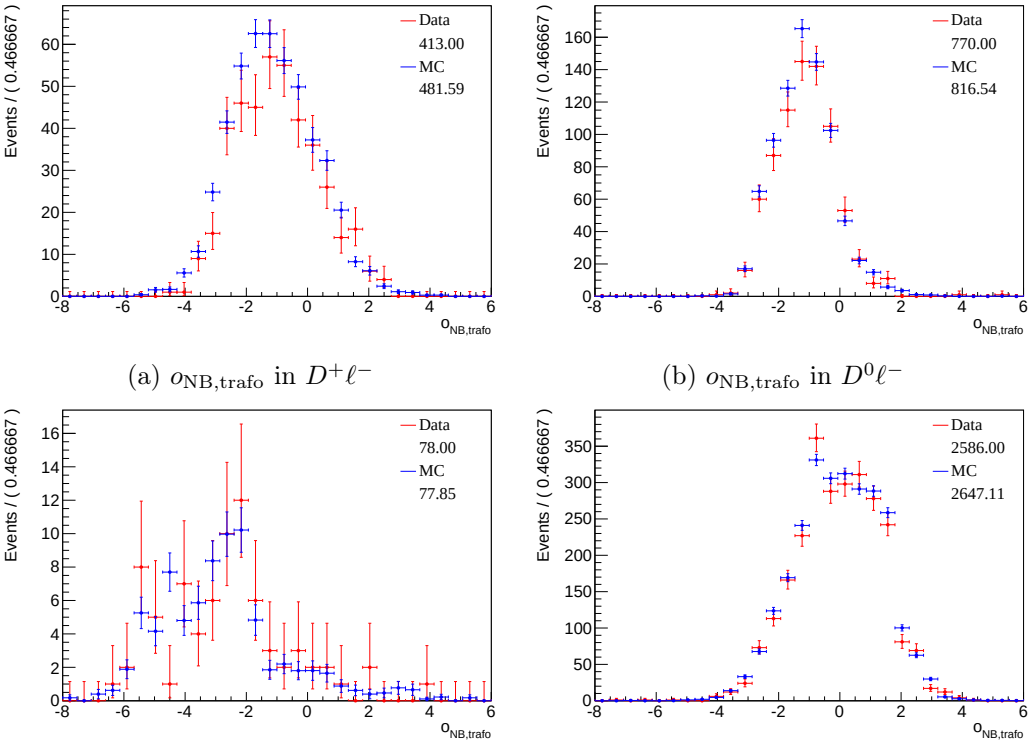


(c) $o_{\mathrm{NB,trafo}}$ in $D^{*+}\ell^-$



(d) $o_{\mathrm{NB,trafo}}$ in $D^{*0}\ell^-$

Figure 10.21: Comparison of $o_{\mathrm{NB,trafo}}$ distributions of Data and Monte Carlo in the higher M^2_{miss} region.



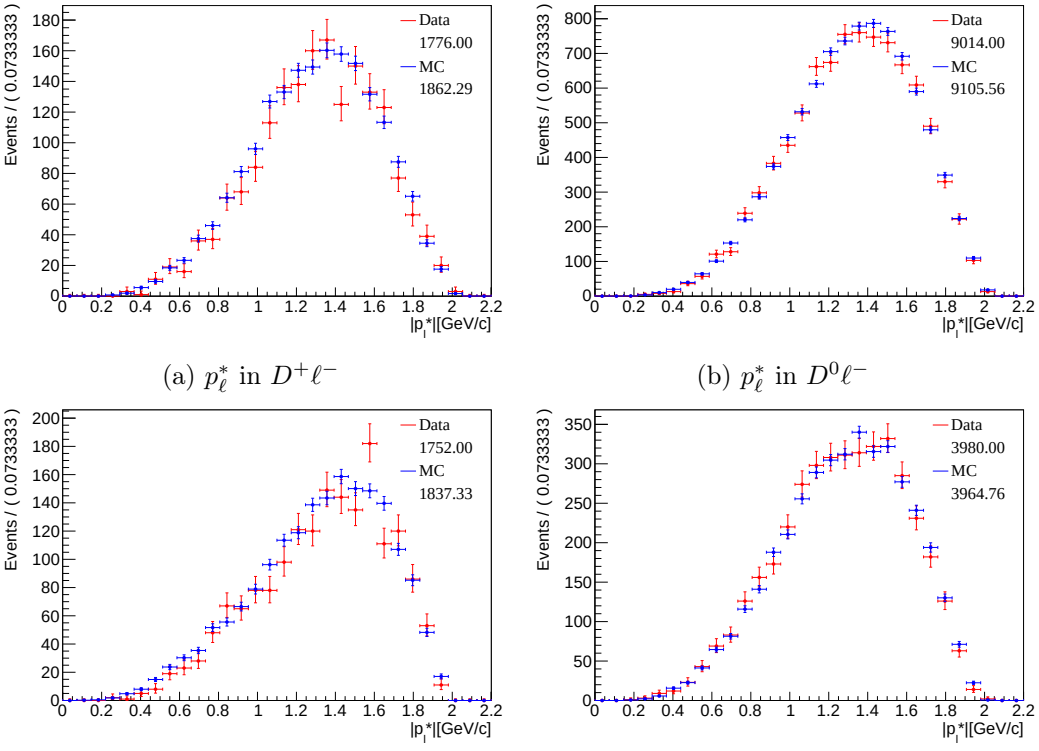
(a) $o_{\mathrm{NB,trafo}}$ in $D^+\ell^-$

(b) $o_{\mathrm{NB,trafo}}$ in $D^0\ell^-$

(c) $o_{\mathrm{NB,trafo}}$ in $D^{*+}\ell^-$

(d) $o_{\mathrm{NB,trafo}}$ in $D^{*0}\ell^-$

Figure 10.22: Comparison of $o_{\mathrm{NB,trafo}}$ distributions of Data and Monte Carlo in the $D^{(*)}$ sidebands.



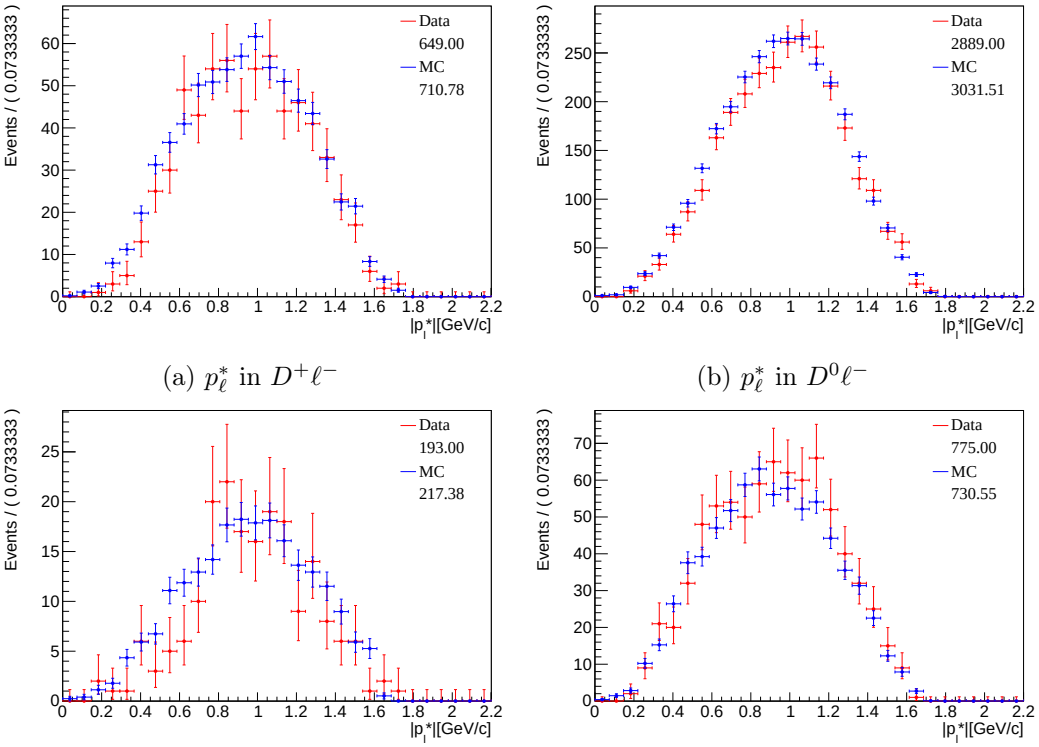
(a) p_ℓ^* in $D^+\ell^-$

(b) p_ℓ^* in $D^0\ell^-$

(c) p_ℓ^* in $D^{*+}\ell^-$

(d) p_ℓ^* in $D^{*0}\ell^-$

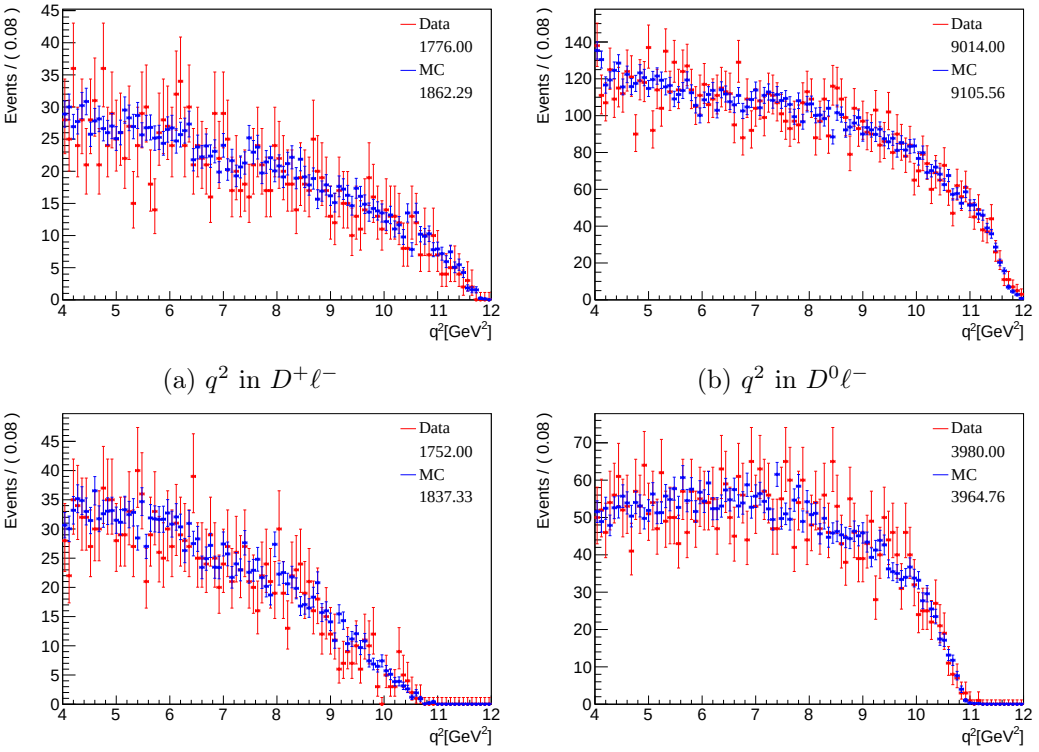
Figure 10.23: Comparison of $|p_\ell^*|$ distributions of Data and Monte Carlo in the lower M^2_{miss} region.



(a) p_ℓ^* in $D^+\ell^-$ (b) p_ℓ^* in $D^0\ell^-$

(c) p_ℓ^* in $D^{*+}\ell^-$ (d) p_ℓ^* in $D^{*0}\ell^-$

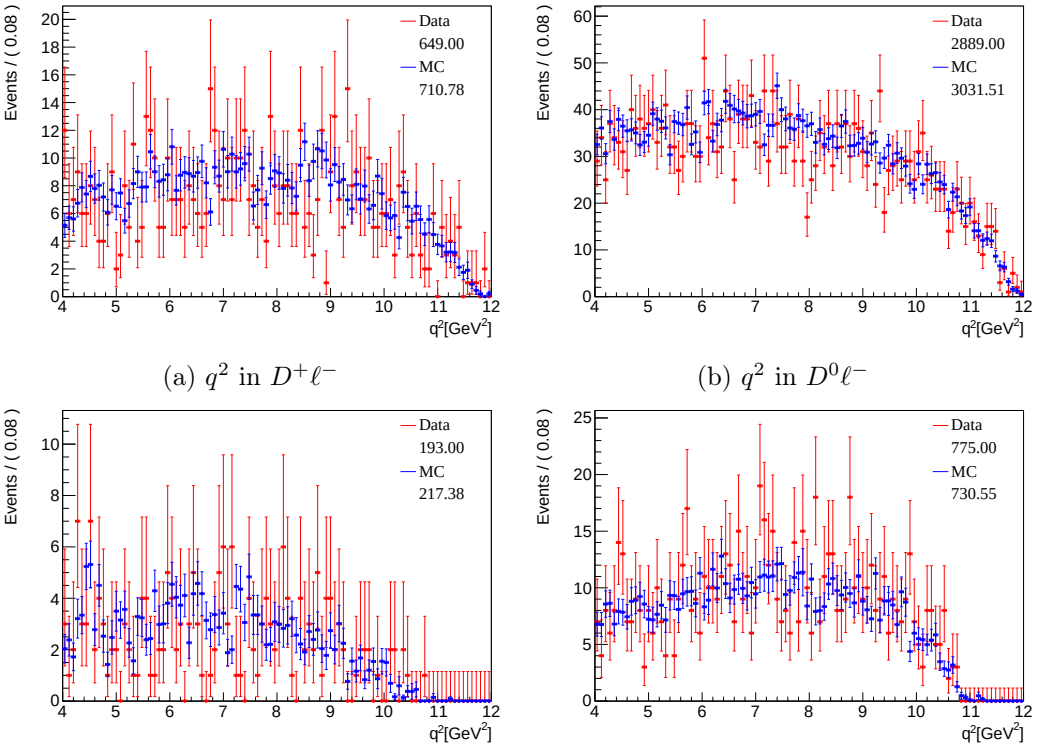
Figure 10.24: Comparison of $|p_\ell^*|$ distributions of Data and Monte Carlo in the higher M^2_{miss} region.



(a) q^2 in $D^+\ell^-$ (b) q^2 in $D^0\ell^-$

(c) q^2 in $D^{*+}\ell^-$ (d) q^2 in $D^{*0}\ell^-$

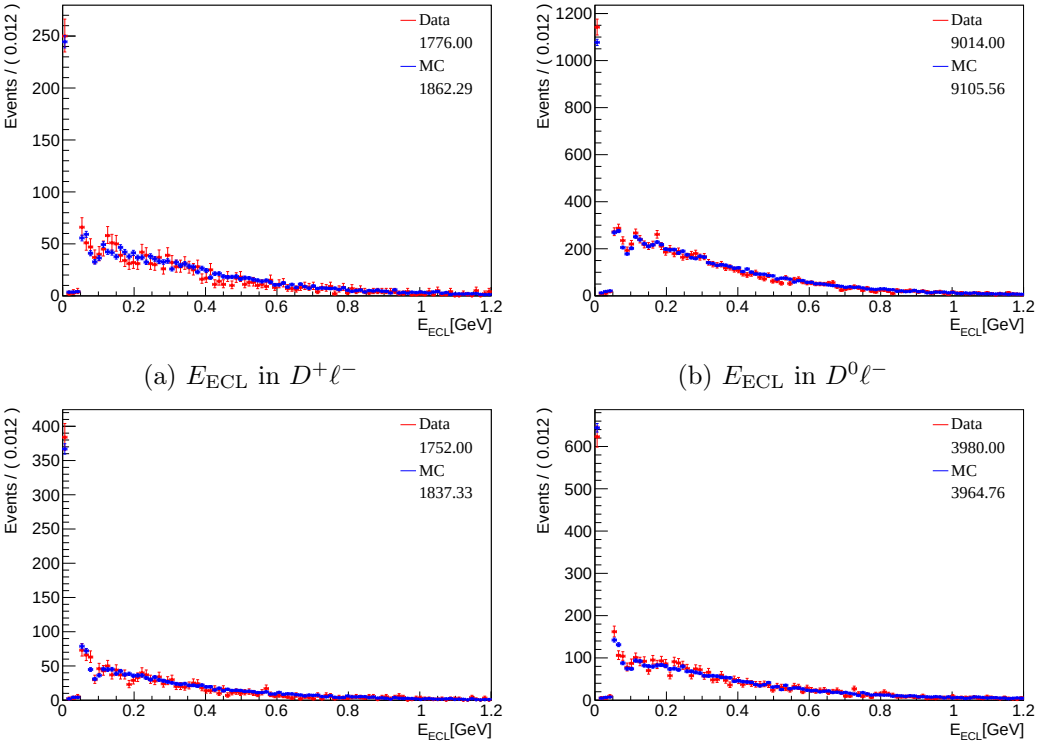
Figure 10.25: Comparison of q^2 distributions of Data and Monte Carlo in the lower M^2_{miss} region.



(a) q^2 in $D^+\ell^-$ (b) q^2 in $D^0\ell^-$

(c) q^2 in $D^{*+}\ell^-$ (d) q^2 in $D^{*0}\ell^-$

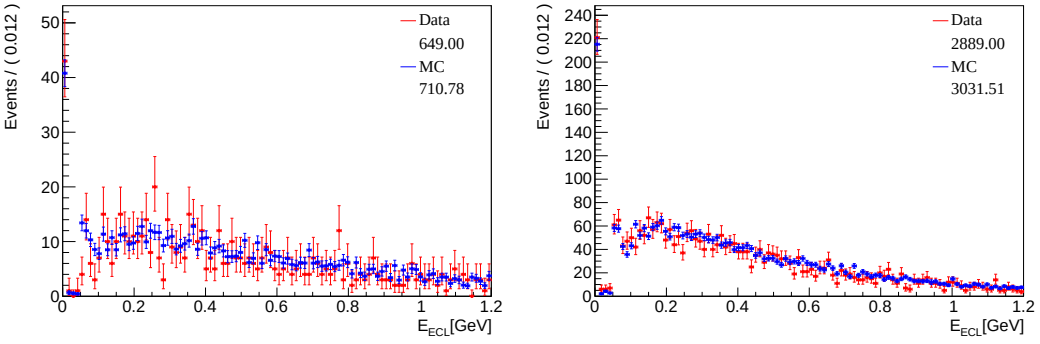
Figure 10.26: Comparison of q^2 distributions of Data and Monte Carlo in the higher M^2_{miss} region.



(a) E_{ECL} in $D^+\ell^-$ (b) E_{ECL} in $D^0\ell^-$

(c) E_{ECL} in $D^{*+}\ell^-$ (d) E_{ECL} in $D^{*0}\ell^-$

Figure 10.27: Comparison of E_{ECL} distributions of Data and Monte Carlo in the lower M^2_{miss} region.



(a) E_{ECL} in $D^{+}\ell^{-}$

(b) E_{ECL} in $D^{0}\ell^{-}$

(c) E_{ECL} in $D^{*+}\ell^{-}$

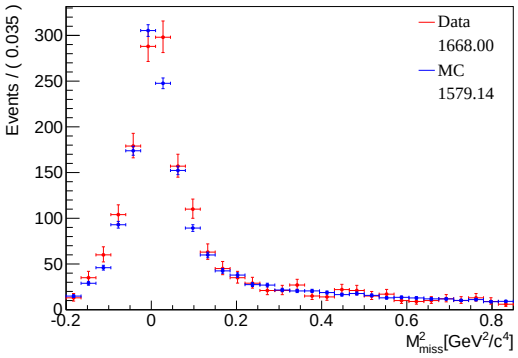
(d) E_{ECL} in $D^{*0}\ell^{-}$

Figure 10.28: Comparison of E_{ECL} distributions of Data and Monte Carlo in the higher M^2_{miss} region.

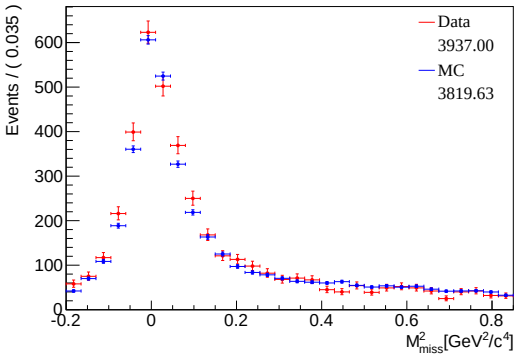
10.4 Inverse q^2 Sample Comparison

Another test of the modeling in Monte Carlo is a comparison of the missing mass distributions in a sample, that is reconstructed with the same cuts except for q^2, which is required to be $< 3.5\,\mathrm{GeV}^2$. The resulting datasets are dominated by the $B \to D\ell\nu$ signal peak, which makes it a good domain to look out for resolution differences.

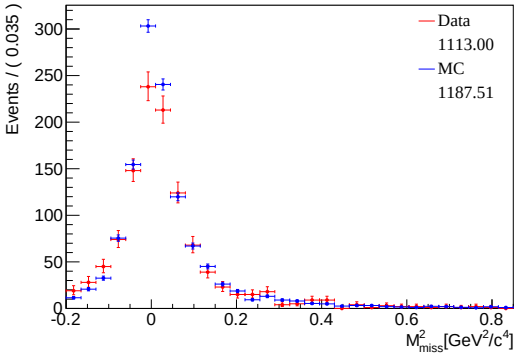
The comparisons are shown in Fig. 10.29. No tendency to over- or underestimate the resolution on simulated data is observed. One interesting finding however is the imbalance of total yields of real and simulated data in the D vs. D^* channels. This indicates, that the probability to misreconstruct a D^* meson as a D meson is different between data and Monte Carlo. As we do not rely on simulated data in this concern, a correction for it is unnecessary.



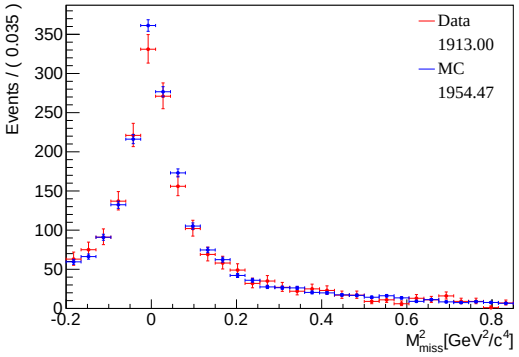
(a) M^2_{miss} in $D^+\ell^-$



(b) M^2_{miss} in $D^0\ell^-$



(c) M^2_{miss} in $D^{*+}\ell^-$



(d) M^2_{miss} in $D^{*0}\ell^-$

Figure 10.29: Comparison of M^2_{miss} distributions of Data and Monte Carlo in the low q^2 region.

10.5 Fits to D^{**} Enriched Sample

One of the major uncertainties in this analysis is the contribution from D^{**} events. Our knowledge about their parameters and branching ratios is fairly limited. These could influence observable distributions and therefore induce a fit bias. Due to limited statistics, we can not entirely verify our modeling of $B \to D^{**}\ell\nu$ decays and have to apply systematic uncertainties later. However, it is possible to check at least the basic compatibility of the shapes, that are extracted from simulated data, with the shapes on real data by using a D^{**} enriched sample.

10.5.1 Sample Reconstruction

The basic assumption is, that most of the $B \to D^{**}\ell\nu$ decays in our fit sample stem from missing a π^0 in the decays $D^{**} \to D^{(*)}\pi^0$. For the enriched sample, the reconstruction modes are therefore the same as for the fit sample, but with an additional π^0. The applied cuts are very similar. Differences are:

- $5.27\,\mathrm{GeV\,c^{-2}} > M_{bc,\mathrm{tag}} > 5.29\,\mathrm{GeV\,c^{-2}}$ to increase D^{**} significance
- looser cuts on the B_{tag} network output
- $M^2_{\mathrm{miss,no}\ \pi^0} > 0.85\,\mathrm{GeV^2\,c^{-4}}$

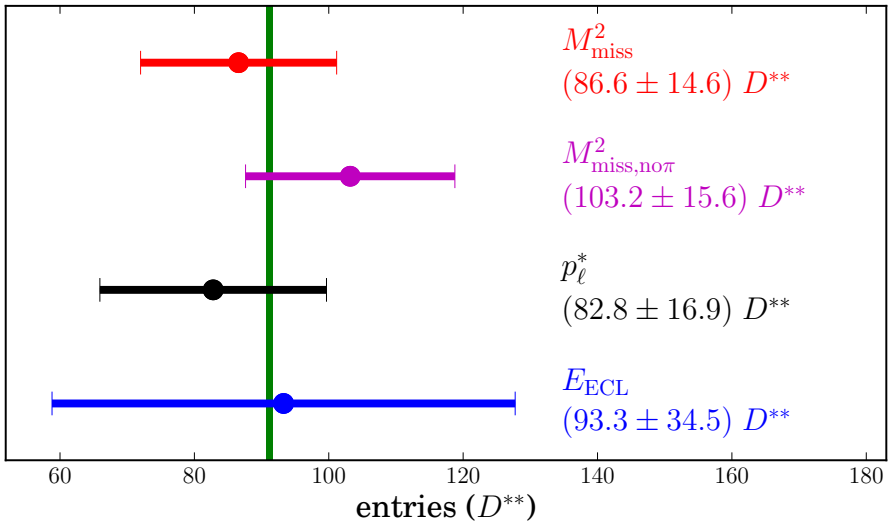
The observable $M^2_{\mathrm{miss,no}\ \pi^0}$ is the missing mass of the candidate, calculated with the additional π^0 excluded. It therefore corresponds to the missing mass, if this event had been reconstructed in the modes of the fit sample with the π^0 missed. The region above $0.85\,\mathrm{GeV^2\,c^{-4}}$ is chosen, as the region below is heavily populated by combinations of $B \to D^{(*)}\ell\nu$ with a random π^0, and is either way not used in the fit procedure.

10.5.2 Sample Fit

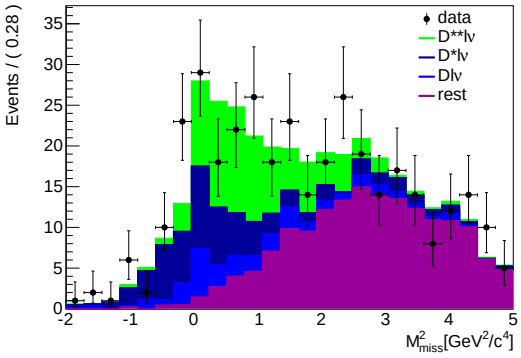
The goal of the fit is to check the compatibility of simulated and real data with respect to the D^{**} description. The method of choice is a series of fits in one dimension to extract the yields of the $B \to D^{**}\ell\nu$ component and see if the results are compatible in different dimensions. Of importance are all observables, that are used in the neural network of the analysis, which are: lepton momentum in the B frame p^*_ℓ, extra energy in the electromagnetic calorimeter E_{ECL}, and squared missing mass without π^0 $M^2_{\mathrm{miss,no}\ \pi^0}$. Additionally the missing mass including the π^0 M^2_{miss} is fitted.

The components that are used in the fits, are: $B \to D^{**}\ell\nu$, $B \to D^*\ell\nu$ with a random π^0, $B \to D\ell\nu$ with a random π^0, and all the rest. The PDFs for the fit procedure are taken from Monte Carlo and transformed to smoothed histogram PDFs. One particularity here is, that not only correctly reconstructed candidates form the component's PDF, but all candidates, that originate from respective B decays. We know, that the yields of $B \to D^{(*)}\ell\nu$ are well modeled on Monte Carlo, which is why their fit yields are constrained to the Monte Carlo expectation with Poisson uncertainties.

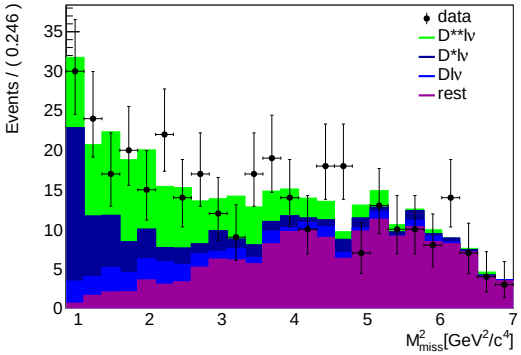
There are four datasets: $D^+\ell^-\pi^0$, $D^{*+}\ell^-\pi^0$, $D^0\ell^-\pi^0$, and $D^{*0}\ell^-\pi^0$. Each observable distribution of each dataset is fitted in an extended unbinned maximum likelihood fit. The fit projections as well as result comparisons are displayed in Figs. 10.30 to 10.33. The green line in the comparison plot is the weighted average. The error bars of the individual fits are reduced by the Poisson error of the individual fit yield under the assumption, that the Poisson error and other uncertainties in the fit result add up squared to the total statistic error. As all fits are applied on the same dataset, fluctuations based on Poisson statistics do not occur.



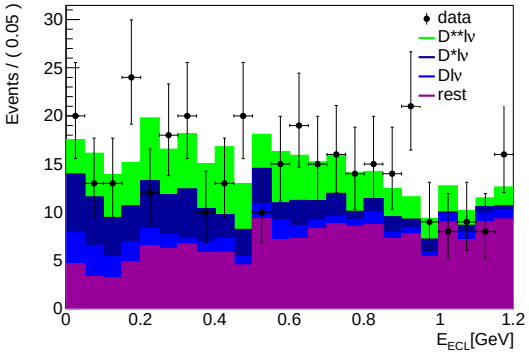
(a) Comparison



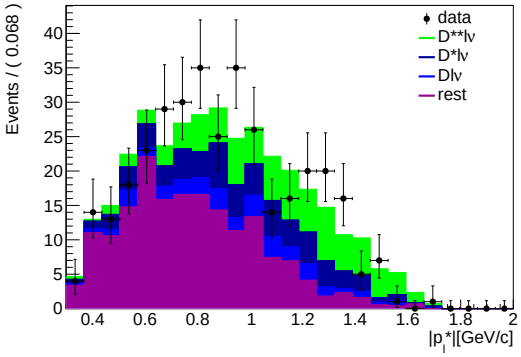
(b) $M^2_{\rm miss}$



(c) $M^2_{\rm miss,no\ \pi^0}$

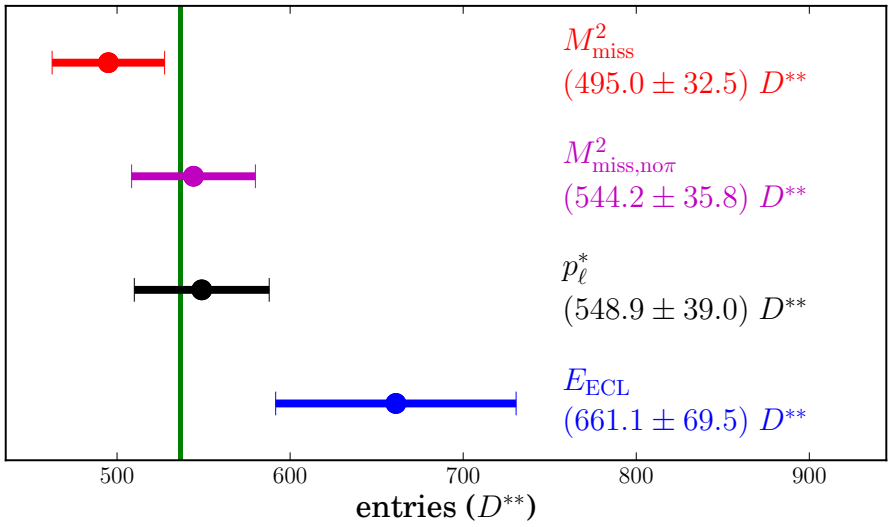


(d) $E_{\rm ECL}$

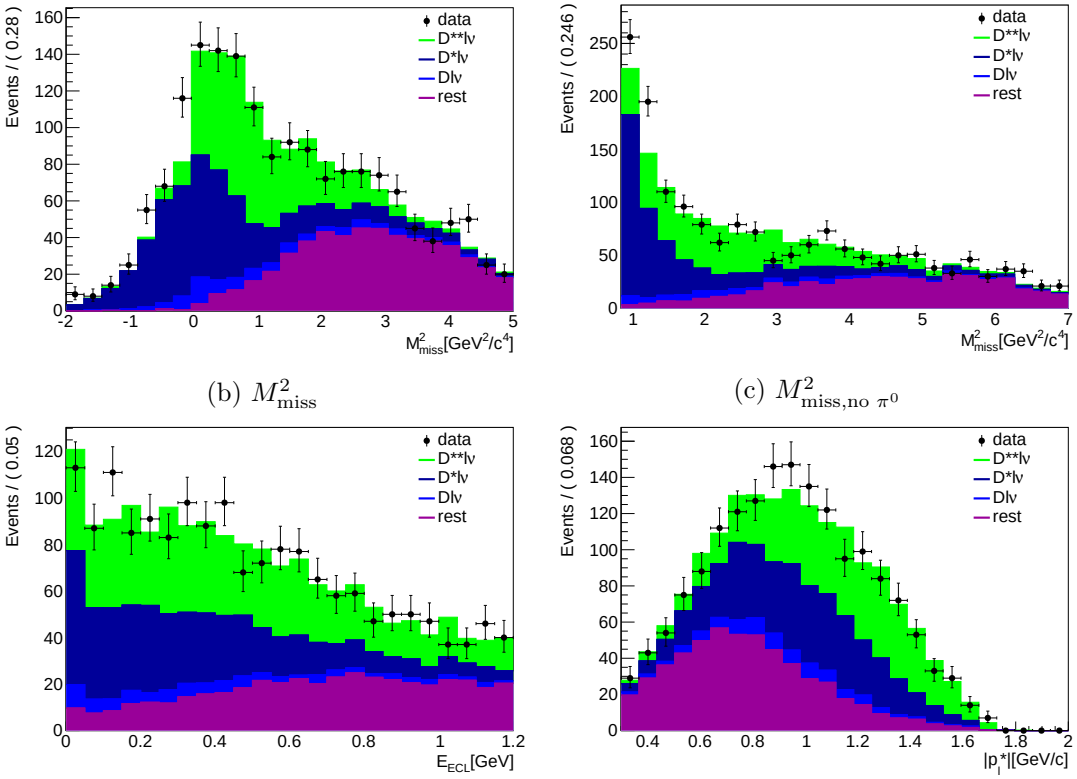


(e) p^*_ℓ

Figure 10.30: Sample: $D^+\ell^-\pi^0$; (a) Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions. (b)-(e) Fit results in four dimensions: real data projected on stacked PDFs.



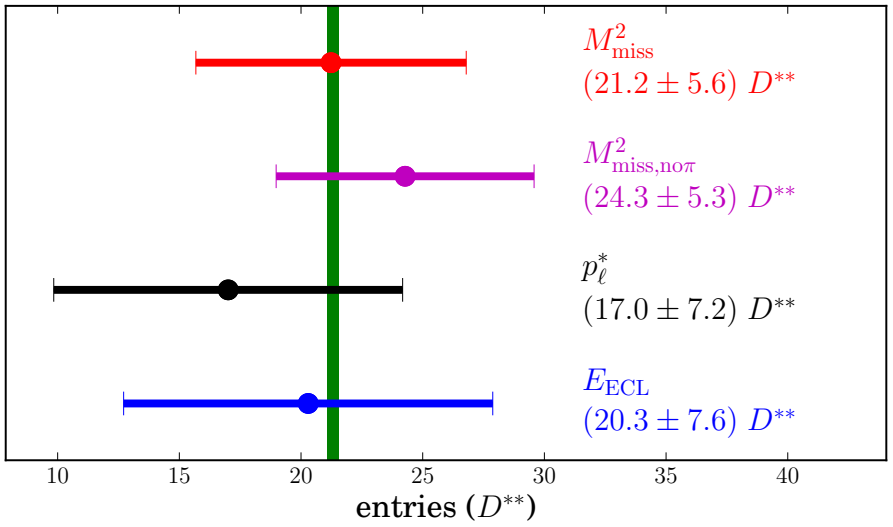
(a) Comparison



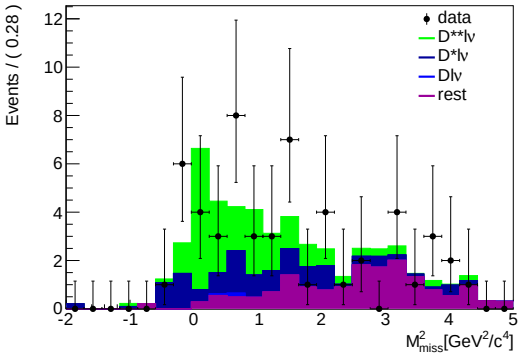
(b) M^2_{miss} (c) $M^2_{\text{miss,no }\pi^0}$

(d) E_{ECL} (e) p^*_ℓ

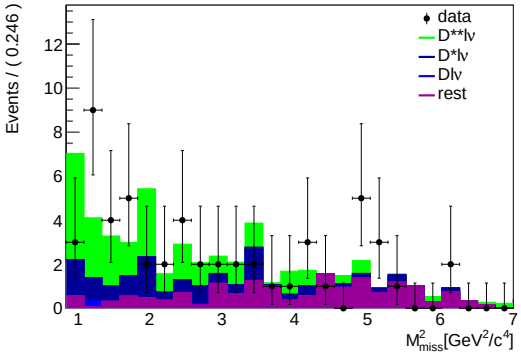
Figure 10.31: Sample: $D^0\ell^-\pi^0$; (a) Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions. (b)-(e) Fit results in four dimensions: real data projected on stacked PDFs.



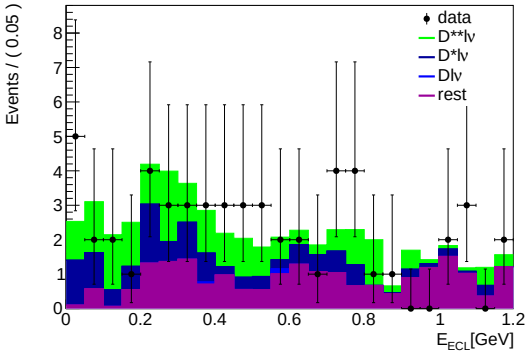
(a) Comparison



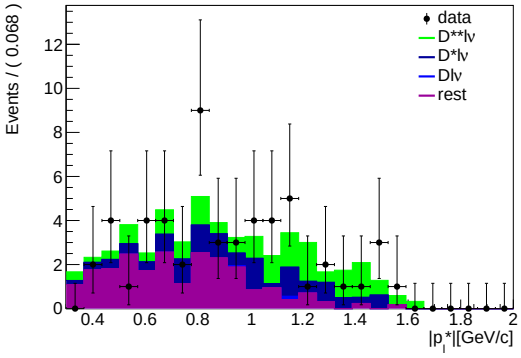
(b) $M^2_{\rm miss}$



(c) $M^2_{\rm miss,no\ \pi^0}$

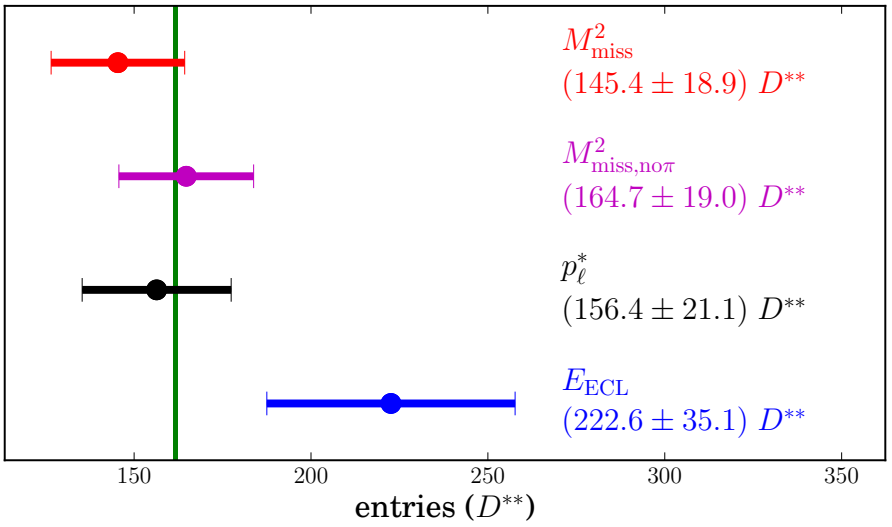


(d) $E_{\rm ECL}$

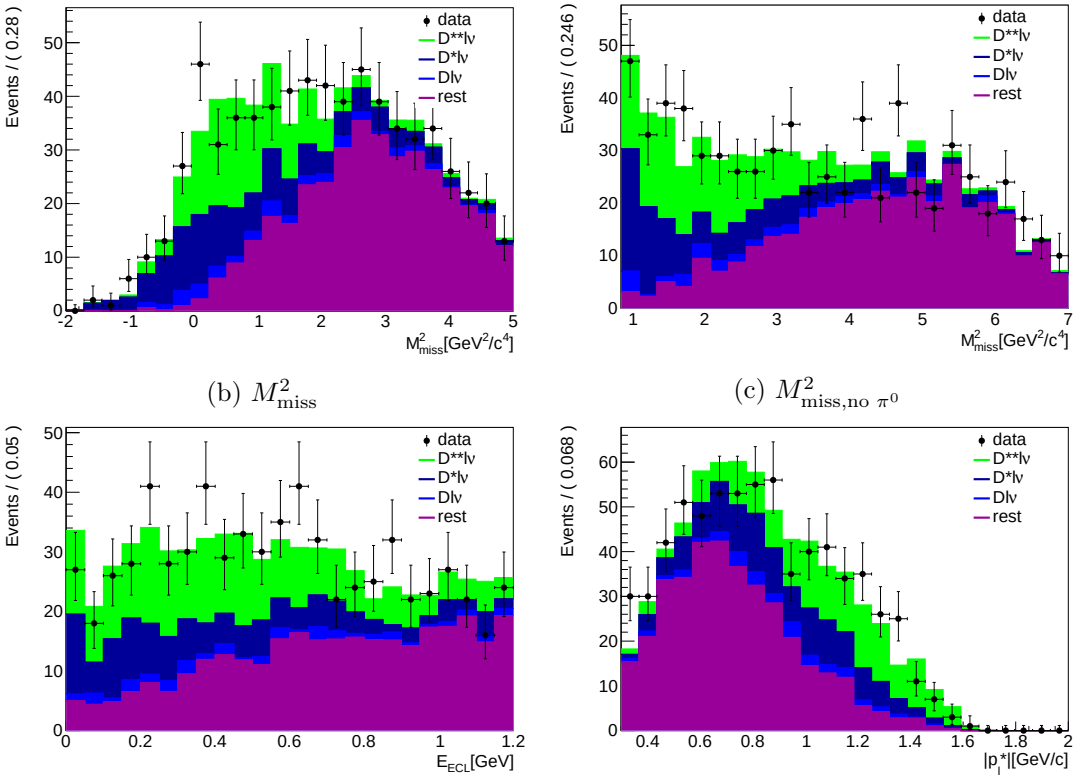


(e) p^*_ℓ

Figure 10.32: Sample: $D^{*+}\ell^-\pi^0$; (a) Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions. (b)-(e) Fit results in four dimensions: real data projected on stacked PDFs.



(a) Comparison



(b) M^2_{miss} (c) $M^2_{\text{miss,no }\pi^0}$

(d) E_{ECL} (e) p^*_ℓ

Figure 10.33: Sample: $D^{*0}\ell^-\pi^0$; (a) Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions. (b)-(e) Fit results in four dimensions: real data projected on stacked PDFs.

10.5.3 Compatibility to Reweighted PDFs

As the parameters for the production of the $B \to D^{**}\ell\nu$ Monte Carlo are not precisely known, several sets of parameters were prepared and implemented as weights, that change the shapes of various distributions. It is therefore interesting, if the fit procedure also works with these altered PDFs. The same datasets were therefore fitted again with all sets of weights. To sum it up, all fits were converging and the corrections to the result with the default weights were rather small, usually in the order of 10% of the statistic uncertainty – even less in the case of D^* modes. This correction was already calculated with the squared sum of the deviation of all weight sets. The individual values are given in the comparison plots in Figs. 10.34 to 10.37.

10.5.4 Conclusion

While the description of the shapes of simulated data is not perfect and suffers from low available statistics, the compatibility of the fit results still indicates, that it is close to real data and can be used in the analysis. Furthermore, the reweighted samples are also well compatible within uncertainties. However, these tests are not suited to extract constraints on the reweighting sets or even composition of the $B \to D^{**}\ell\nu$ component. They do in no way replace later checks for systematic uncertainties stemming from the limited knowledge.

10.6 Comparison to BaBar's D^{**} Sample

The analysis of the BaBar collaboration also uses a D^{**} enriched sample, but in a different way. They use it to extract their D^{**} yields and provide plots that can be compared to ours. However, the selection they use differs from the one in this thesis, so I have to imitate their way of enhancing the D^{**} component. The main difference is the classification method. They use two boosted decision trees: one to suppress $B \to D^{(*)}\ell\nu$, and one to suppress combinatoric and continuum background. This procedure is also applied on our sample using neural networks instead, and optimizing the cuts on the two networks to maximize D^{**} significance.

Due to the systematic uncertainties, the cut sample is no longer suitable to verify our simulated data. The use of this sample is the comparison with the BaBar selection, to see if there are fundamental differences in the datasets, which could indicate problems in the selection process.

The comparison is given in Fig. 10.38. While similar distributions can be reproduced, several things become clear:

- The M^2_{miss} resolution at Belle is slightly worse.
- The analysis sample in this thesis has much less continuum and combinatoric background.
- It has much more cross-feed events.

So while the two experiments are quite similar, they are by no means identical. The differences in the distributions are one cause, that a fragile analysis as the presented one may not be suited to be handled by both collaborations identically.

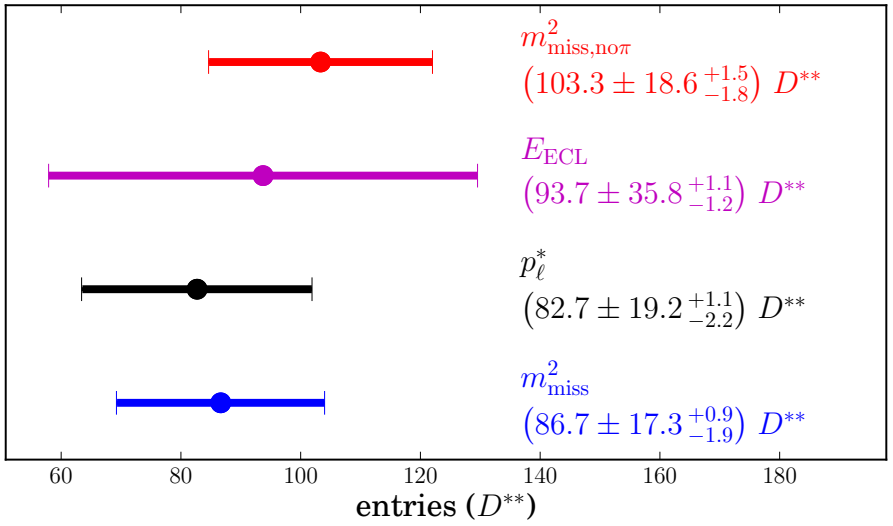


Figure 10.34: Sample: $D^+\ell^-\pi^0$; Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions, including systematic uncertainties from shape reweighting.

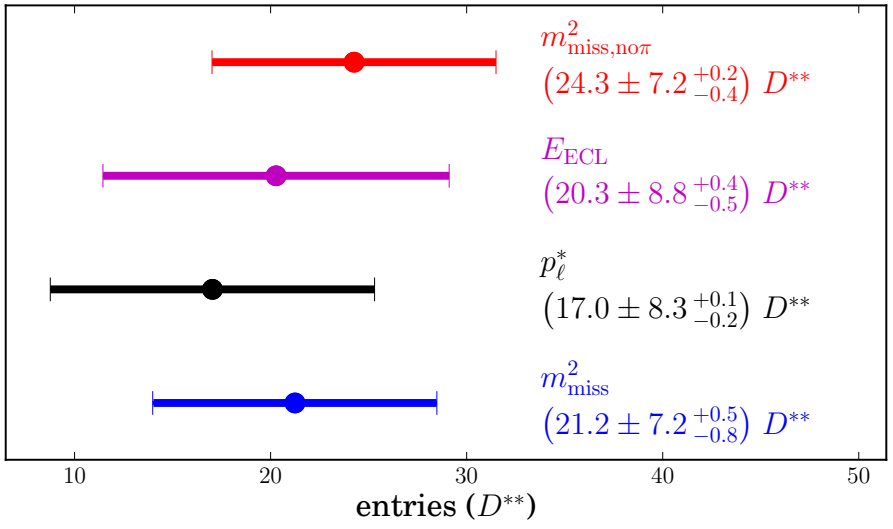


Figure 10.35: Sample: $D^{*+}\ell^-\pi^0$; Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions, including systematic uncertainties from shape reweighting.

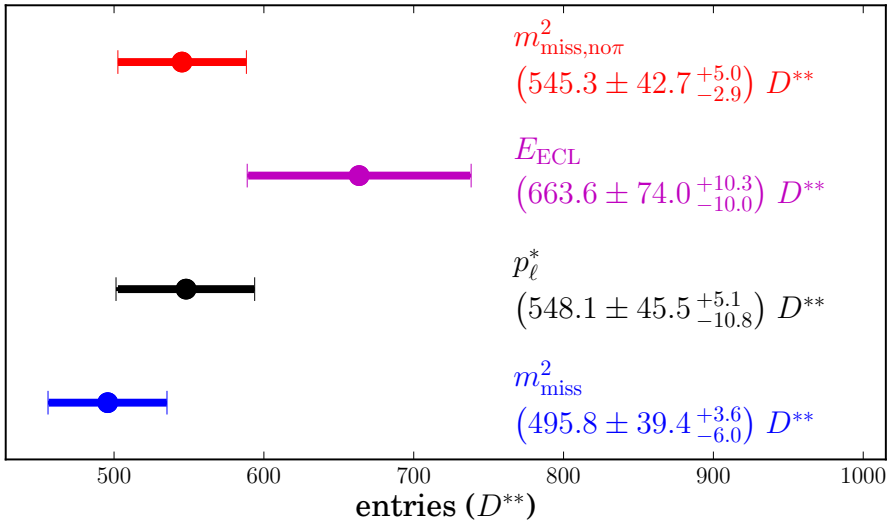


Figure 10.36: Sample: $D^0\ell^-\pi^0$; Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions, including systematic uncertainties from shape reweighting.

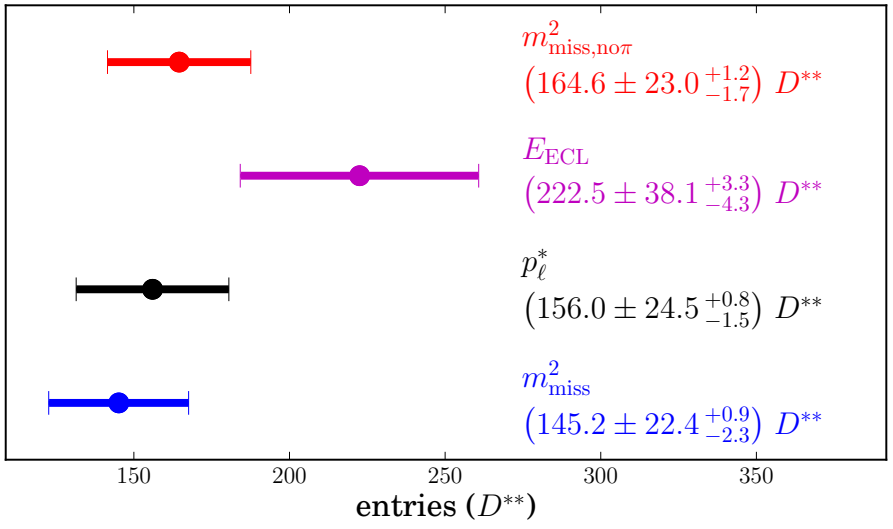
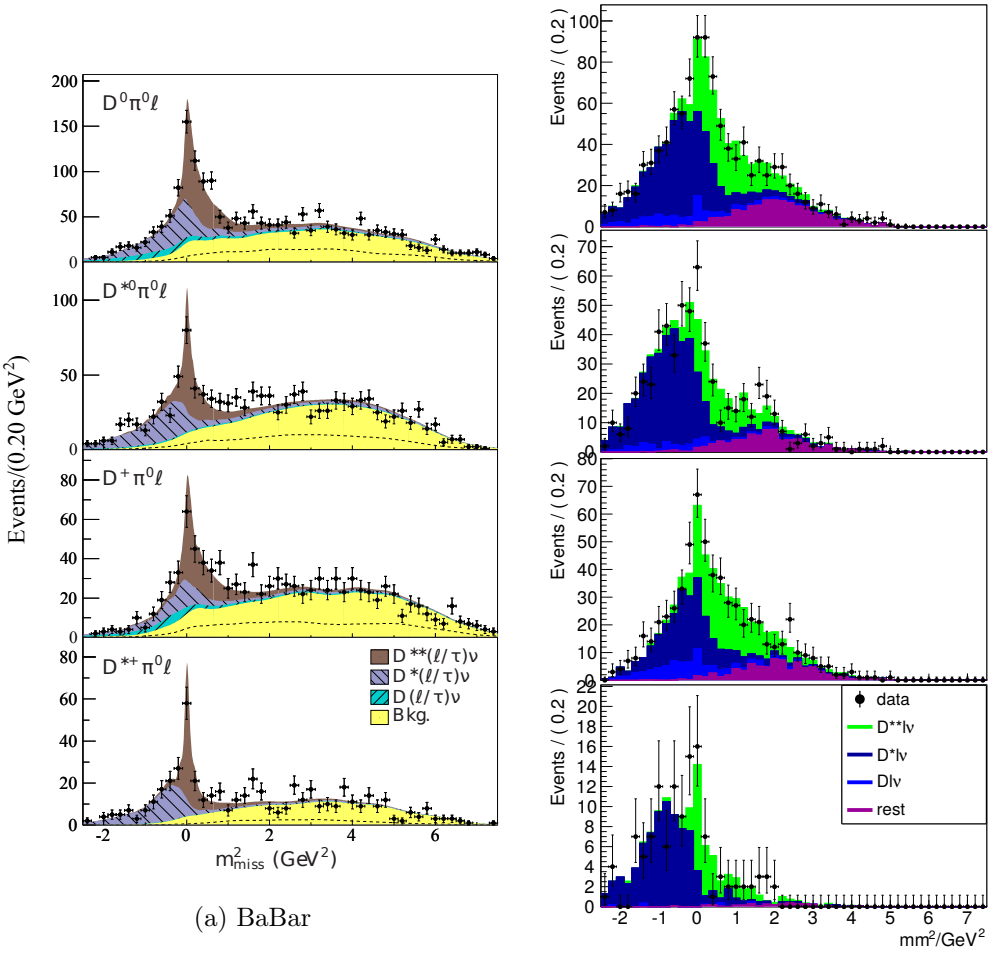


Figure 10.37: Sample: $D^{*0}\ell^-\pi^0$; Comparison of the fit results for the $B \to D^{**}\ell\nu$ yield in the four fit dimensions, including systematic uncertainties from shape reweighting.



(a) BaBar

(b) Belle

Figure 10.38: Comparison of the D^{**} enriched samples of (a) BaBar and (b) Belle with similar selection procedures.

	mean	sigma
$D^+\ell^-$ - ℓ signal	-0.028 ± 0.048	1.050 ± 0.035
$D^+\ell^-$ - ℓ CF	-0.120 ± 0.048	1.061 ± 0.036
$D^+\ell^-$ - D^{**}	-0.053 ± 0.047	1.034 ± 0.035
$D^{*+}\ell^-$ - ℓ signal	0.058 ± 0.051	1.122 ± 0.038
$D^{*+}\ell^-$ - D^{**}	0.005 ± 0.053	1.160 ± 0.039
$D^0\ell^-$ - ℓ signal	-0.047 ± 0.046	1.019 ± 0.033
$D^0\ell^-$ - ℓ CF	0.064 ± 0.047	1.030 ± 0.035
$D^0\ell^-$ - D^{**}	0.076 ± 0.054	1.172 ± 0.041
$D^{*0}\ell^-$ - ℓ signal	-0.106 ± 0.051	1.132 ± 0.037
$D^{*0}\ell^-$ - D^{**}	-0.035 ± 0.054	1.188 ± 0.039
R	-0.049 ± 0.045	1.000 ± 0.033
R^*	-0.019 ± 0.046	1.021 ± 0.034

Table 10.1: Pull mean and sigma overview.

10.7 Toy Study

10.7.1 Description

In order to check the fit model, a toy Monte Carlo study has been performed. Datasets were generated, not based on a physical model, but on the PDFs in Section 9.2, to mimic the real fit data. The shapes of the PDFs were determined on 5 streams of generic MC. The number of events, that were generated in each dataset was selected according to Poisson statistics with the true yield in generic MC as expectation value. The values for the parameters R and R^*, that the τ signal yield is based on, were pulled randomly each run from an even distribution in the interval $[0.2, 0.45]$. For each reconstruction mode ($D^+\ell^-$, $D^{*+}\ell^-$, $D^0\ell^-$, $D^{*0}\ell^-$) the following samples were produced:

- sideband sample (D mass sideband for D modes, $\Delta M(D^*, D)$ sideband for D^* modes)
- M^2_{miss} distribution as low missing mass fit sample
- $o_{\text{NB,trafo}}$ distribution as high missing mass fit sample

The fitting procedure described in Chapter 9 was then applied and the resulting yields and errors for each component saved. A pull distribution, defined as

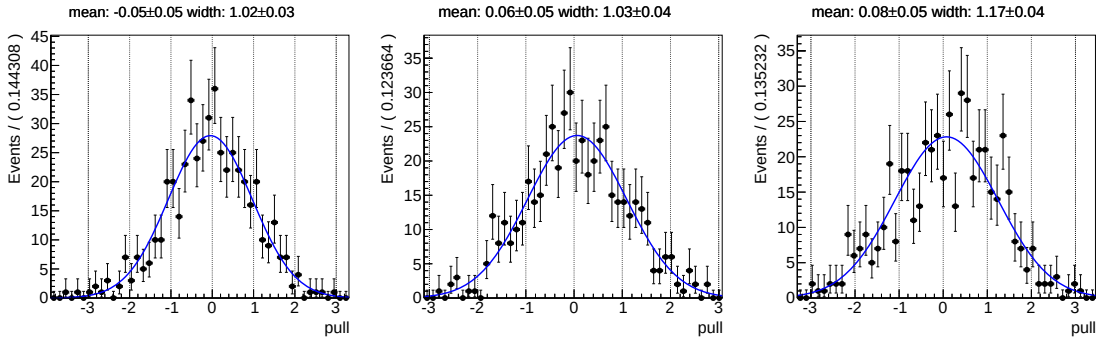
$$\text{pull} = \frac{\text{Yield} - \text{Expectation}}{\text{Error}}$$

was calculated.

Overall, 500 iterations were performed and the resulting distributions for the pull values can be found in Figs. 10.39 to 10.43. Gaussian distributions were fitted to the data points, and the mean and width of them are also noted on the plots. A bias in the fit would manifest itself in the pull distributions: A mean of the fitted Gaussian which would be significantly different from zero would indicate a bias. The overview for fitted mean and sigma values is given in Table 10.1.

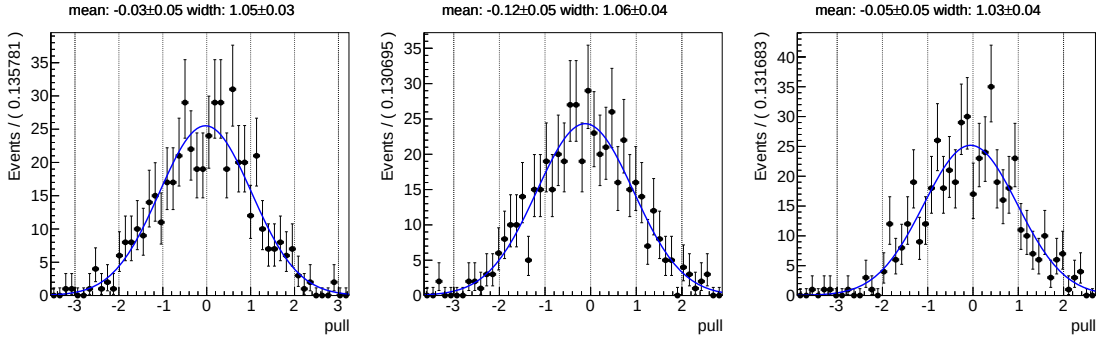
10.7.2 Conclusion

Most pull distributions are well compatible with a Gaussian of mean 0.0 and width 1.0. However, there is a clear tendency to widths larger than 1.0. This is an effect, that can be explained by the systematic uncertainty coming from the fixated yields of wrong D



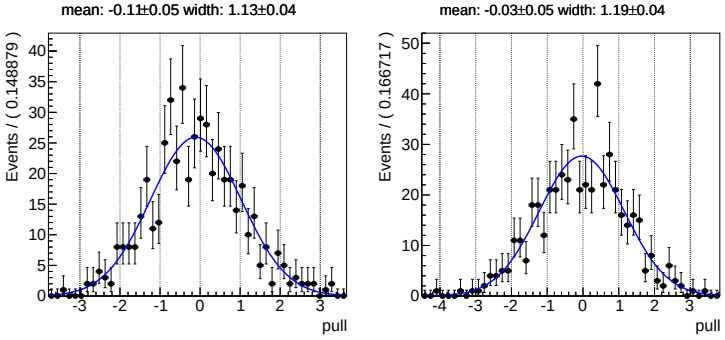
(a) ℓ signal (b) ℓ CF (c) D^{**} background

Figure 10.39: $D^{+}\ell^{-}$ pull distributions for a) lepton signal, b) lepton cross-feed, c) D^{**} components.



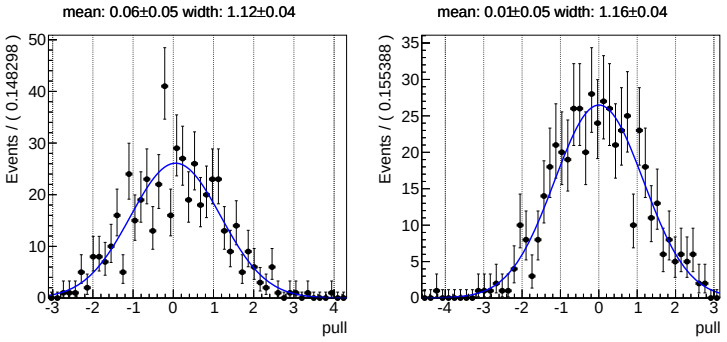
(a) ℓ signal (b) ℓ CF (c) D^{**} background

Figure 10.40: $D^{0}\ell^{-}$ pull distributions for a) lepton signal, b) lepton cross-feed, c) D^{**} components.



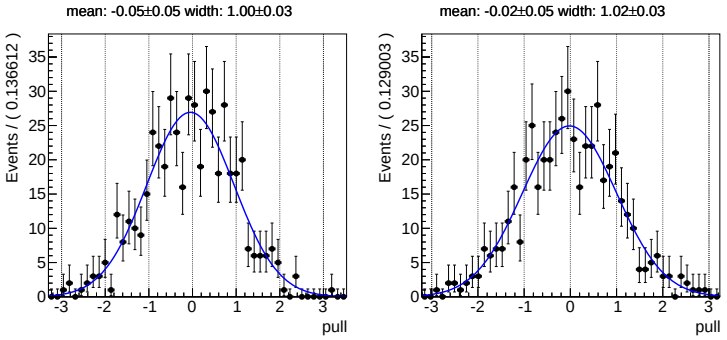
(a) ℓ signal (b) D^{**} background

Figure 10.41: $D^{*+}\ell^{-}$ pull distributions for a) lepton signal, b) D^{**} components.



(a) ℓ signal (b) D^{**} background

Figure 10.42: $D^{*0}\ell^-$ pull distributions for a) lepton signal, b) D^{**} components.



(a) R (b) R^*

Figure 10.43: Pull distributions for a) R, b) R^*.

mesons in each sample. They are an important contribution and determined on D mass sidebands. In the toy study, their yield is generated according to Poisson statistics and can therefore fluctuate around the true value. The quantity of this systematic uncertainty will be discussed later.

Overall, the toy MC study did not reveal any major issues in the fitting procedure.

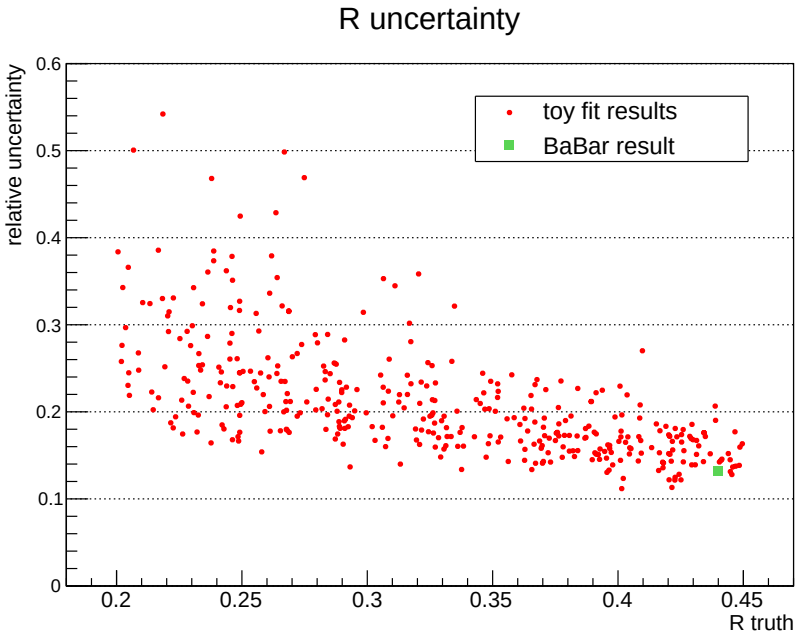
10.7.3 Expected Statistical Uncertainty

An important property of this analysis is the expected relative error on the values of R and R^*. The toy MC studies can provide such estimates. The values can be extracted from the parameter error distributions of the toy MC results. As the relative uncertainty depends strongly on the absolute measured value, the distribution of the relative uncertainties of 500 fits on toy MC samples with varying truth values of R and R^* is given in Fig. 10.44, along with the BaBar result for comparison.

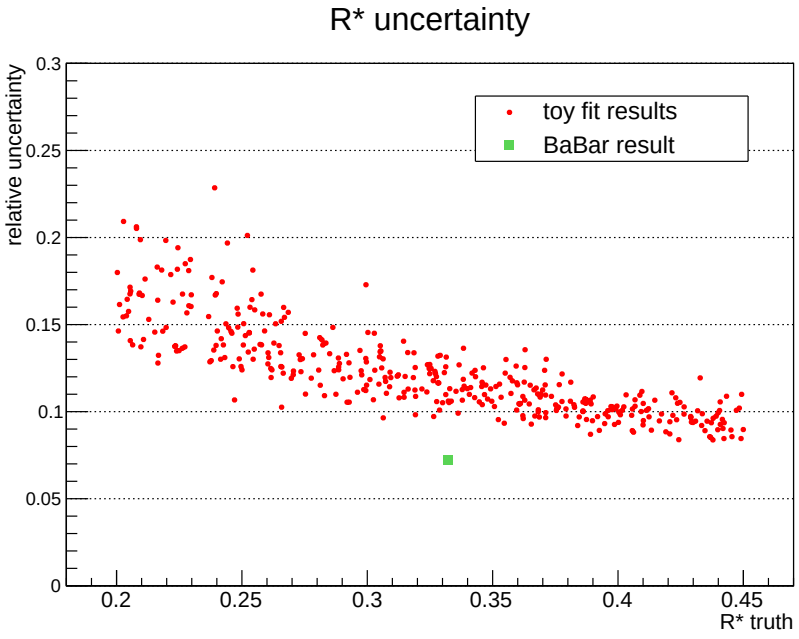
We can see that we expect similar but slightly higher uncertainties compared to the measurement by the BaBar collaboration. The distribution shows that for R a result in the same neighborhood would also produce similar uncertainties. In R^* however, we will not reach their precision, probably due to the larger cross-feed probability in our datasets.

10.7.4 Split Value Validation

It has not been tested yet, what influence the choice of the value, that is used to split the sample in missing mass, has on the fit result. A higher value increases the τ signal purity in the higher missing mass region, but also costs efficiency. One also has to consider, that



(a) R



(b) R^*

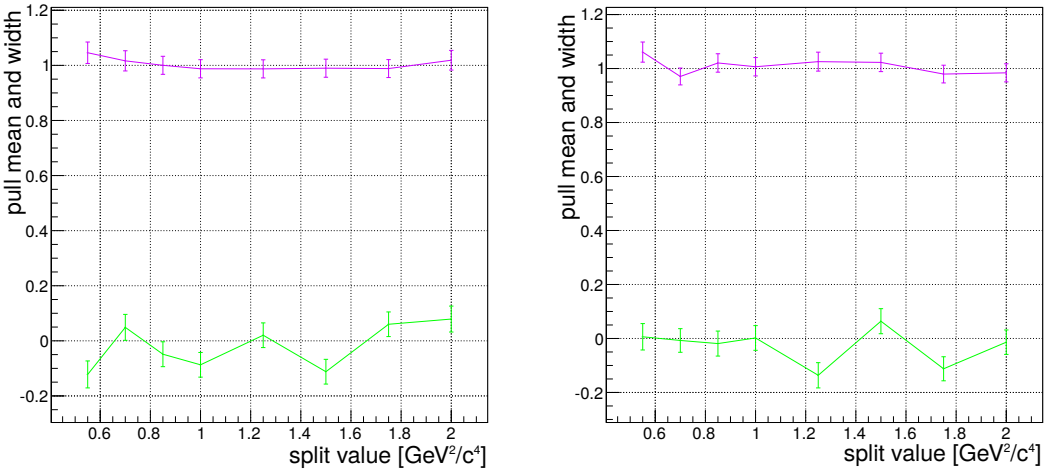
Figure 10.44: Relative uncertainties of the fit procedure performed on 500 toy MC samples in R (a) and R^* (b).

the purity versus the most difficult background from D^{**} can only be improved slightly this way.

A lower value on the other hand increases the efficiency, but can quickly multiply some backgrounds, especially from ℓ *cross-feed* and *wrong* D. The latter also introduces a systematic uncertainty, as it is estimated from sideband yields using ratios from Monte Carlo, that are not perfectly accurate.

The used value $0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$ was chosen by eye and it is tested with the same toy Monte Carlo procedure, what influence this value has on the expected uncertainties. Additionally

to the default fit, the same toy sample is fitted with a variety of differing parameters, accounting for systematic uncertainties coming from shapes (see Section 7.3 and Section 7.4) and from factors that are constrained from MC and prone to limited MC statistics or Poisson fluctuations. Fig. 10.45 show the changes in the pull distributions of the individual fit parameters. Important is here mostly, that there is no tendency, i.e. no additional systematic uncertainty from the choice of the split value, as the means in R and R^* are well compatible with 0.0 at all tested values.



(a) R (b) R^*

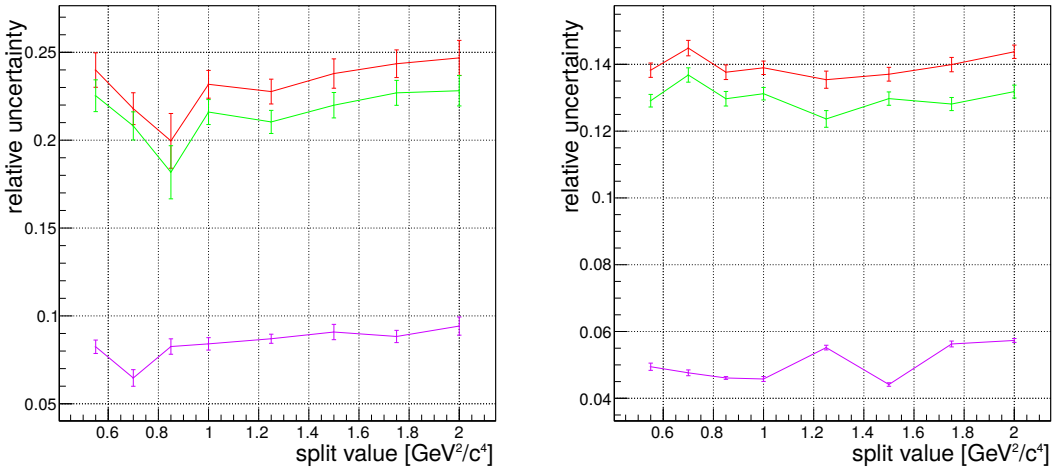
Figure 10.45: Pull means and widths of R and R^* determined with toy studies using different split values.

Split Value and Uncertainties

To assure, that I have selected a proper split value, a look at the expected uncertainties is necessary. The comparison for R and R^* is given in Table 10.2 and in Fig. 10.46. The comparison shows, that there is no best value, but the range between 0.8 GeV^2 and 1.2 GeV^2 seems to be slightly preferred.

parameter	split value [GeV^2 c^{-4}]	stat.	syst.	comb.
R	0.55	0.225	0.082	0.240
R	0.70	0.208	0.065	0.218
R	0.85	0.182	0.083	0.200
R	1.00	0.216	0.084	0.232
R	1.25	0.210	0.087	0.228
R	1.50	0.220	0.091	0.238
R	1.75	0.227	0.088	0.244
R	2.00	0.228	0.094	0.247
R^*	0.55	0.129	0.049	0.138
R^*	0.70	0.137	0.048	0.145
R^*	0.85	0.130	0.046	0.138
R^*	1.00	0.131	0.046	0.139
R^*	1.25	0.124	0.055	0.135
R^*	1.50	0.130	0.044	0.137
R^*	1.75	0.128	0.056	0.140
R^*	2.00	0.132	0.057	0.144

Table 10.2: Relative uncertainties of R and R^* in dependency of the split value.



(a) R (b) R^*

Figure 10.46: Statistic (green), (reduced) systematic (purple) and combined (red) error of R and R^* determined by toy MC studies of different split value.

split value	eff.	eff. NP	loss
0.55	0.0139%	0.0126%	9.7%
0.70	0.0136%	0.0124%	9.1%
0.85	0.0134%	0.0123%	8.2%
1.00	0.0130%	0.0121%	7.3%
1.25	0.0124%	0.0118%	5.1%
1.50	0.0116%	0.0113%	2.6%
1.75	0.0108%	0.0108%	−0.5%
2.00	0.0099%	0.0103%	−4.2%

Table 10.3: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production q^2 value for $B^0 \to D^+\tau\nu$ decays.

10.8 New Physics Sensitivity

10.8.1 Sensitivity

As the sensitivity for the τ yields is far higher in the high M^2_{miss} sample, we have to check, whether my choice to split the fit sample has the potential to suppress effects from NP. This is done by estimating the changes in efficiency coming from different kinematics.

The test is performed in two dimensions independently: lepton momentum p_ℓ and four-momentum transition q^2:

- The produced distribution of the observable is extracted from MC, once for the assumption of validity for the SM, and once for a NP point with $\tan\beta/m_H = 0.5\,\text{c}^2\,\text{GeV}^{-1}$.
- All events of the SM MC that pass the final selection in the high M^2_{miss} sample are used to calculate the efficiency in bins of the observable.
- This binned efficiency is applied to the produced distribution of NP MC.

The procedure is repeated for a variety of possible split values. The overview of the effects on the efficiency for q^2 is given in Tables 10.3 to 10.6. The corresponding distributions and efficiency plots are in Figs. 10.47 and 10.48. They are only shown for the two split values $0.55\,\text{GeV}^2\,\text{c}^{-4}$ and $2.0\,\text{GeV}^2\,\text{c}^{-4}$, as the values in between do not show much different behavior. These two extremal points are therefore suitable to illustrate the effect of varying the split value.

For p_ℓ the respective numbers are in Tables 10.7 to 10.10 and the plots in Figs. 10.49 and 10.50.

It is obvious that the choice of the split value does not have much influence on the sensitivity. While the q^2 distribution is clearly influenced, the efficiency does only favor one distribution over the other by small amounts in the order of up to 10% in both directions. In p_ℓ the effect is even smaller with $0-3\%$, which is due to the smaller influence on the produced distributions.

10.8.2 q^2 Tests

The fit result can be used to create background-subtracted q^2 distributions, which are an additional indicator for the influence of NP. Starting from the q^2 distributions with $M^2_{\text{miss}} > 0.85\,\text{GeV}^2\,\text{c}^{-4}$ in Fig. 10.5, all distributions – except the τ signal – are subtracted from the data points. The resulting distribution is the measured q^2 distribution of the

split value	eff.	eff. NP	loss
0.55	0.0097%	0.0094%	2.7%
0.70	0.0096%	0.0093%	2.7%
0.85	0.0094%	0.0091%	2.5%
1.00	0.0092%	0.0090%	2.4%
1.25	0.0088%	0.0087%	1.9%
1.50	0.0083%	0.0081%	1.6%
1.75	0.0077%	0.0076%	1.1%
2.00	0.0071%	0.0071%	0.2%

Table 10.4: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production q^2 value for $B^0 \to D^{*+}\tau\nu$ decays.

split value	eff.	eff. NP	loss
0.55	0.0223%	0.0223%	0.2%
0.70	0.0221%	0.0222%	−0.2%
0.85	0.0219%	0.0220%	−0.5%
1.00	0.0214%	0.0216%	−1.2%
1.25	0.0204%	0.0210%	−3.0%
1.50	0.0193%	0.0203%	−4.8%
1.75	0.0180%	0.0193%	−7.3%
2.00	0.0165%	0.0183%	−10.8%

Table 10.5: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production q^2 value for $B^- \to D^0\tau\nu$ decays.

split value	eff.	eff. NP	loss
0.55	0.0279%	0.0278%	0.2%
0.70	0.0277%	0.0277%	0.1%
0.85	0.0274%	0.0274%	0.1%
1.00	0.0270%	0.0270%	0.0%
1.25	0.0261%	0.0261%	−0.2%
1.50	0.0251%	0.0252%	−0.4%
1.75	0.0239%	0.0241%	−0.7%
2.00	0.0225%	0.0227%	−1.0%

Table 10.6: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production q^2 value for $B^- \to D^{*0}\tau\nu$ decays.

split value	eff.	eff. NP	loss
0.55	0.0138%	0.0138%	−0.1%
0.70	0.0135%	0.0136%	−0.1%
0.85	0.0133%	0.0133%	−0.1%
1.00	0.0130%	0.0130%	−0.1%
1.25	0.0124%	0.0124%	−0.1%
1.50	0.0117%	0.0117%	−0.2%
1.75	0.0109%	0.0109%	−0.1%
2.00	0.0100%	0.0100%	−0.1%

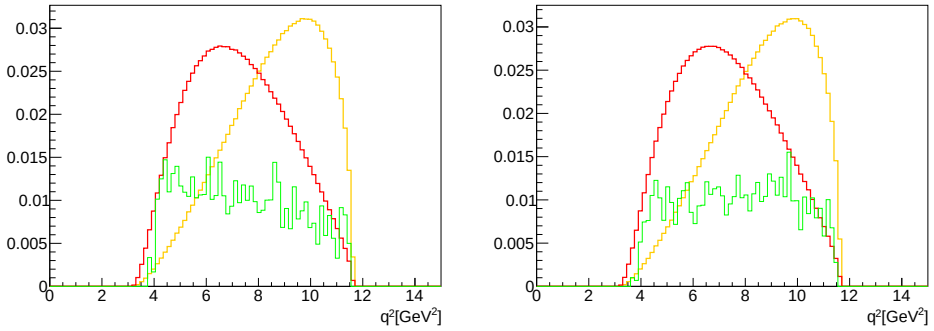
Table 10.7: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production p_ℓ value for $B^0 \to D^+\tau\nu$ decays.

split value	eff.	eff. NP	loss
0.55	0.0097%	0.0096%	0.9%
0.70	0.0096%	0.0095%	0.9%
0.85	0.0095%	0.0094%	0.9%
1.00	0.0093%	0.0092%	0.9%
1.25	0.0089%	0.0089%	0.8%
1.50	0.0084%	0.0084%	0.8%
1.75	0.0079%	0.0079%	0.7%
2.00	0.0073%	0.0072%	0.6%

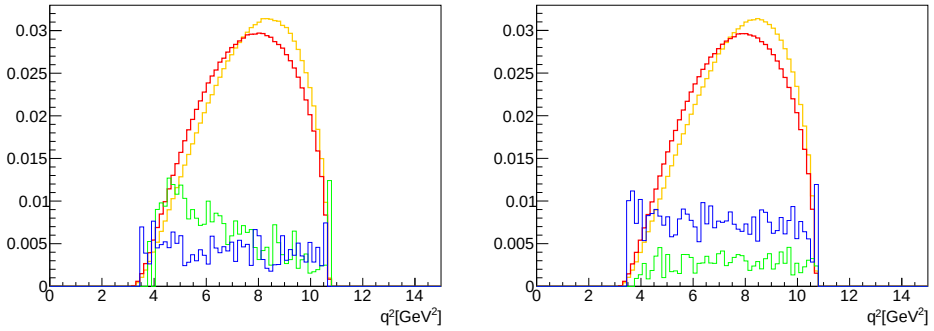
Table 10.8: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production p_ℓ value for $B^0 \to D^{*+}\tau\nu$ decays.

split value	eff.	eff. NP	loss
0.55	0.0222%	0.0215%	3.3%
0.70	0.0220%	0.0213%	3.2%
0.85	0.0218%	0.0211%	3.2%
1.00	0.0214%	0.0207%	3.0%
1.25	0.0205%	0.0199%	2.8%
1.50	0.0195%	0.0191%	2.4%
1.75	0.0183%	0.0179%	2.0%
2.00	0.0168%	0.0165%	1.5%

Table 10.9: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production p_ℓ value for $B^- \to D^0\tau\nu$ decays.



(a) $B^0 \to D^+\tau^-\nu$ (b) $B^- \to D^0\tau^-\nu$



(c) $B^0 \to D^{*+}\tau^-\nu$ (d) $B^- \to D^{*0}\tau^-\nu$

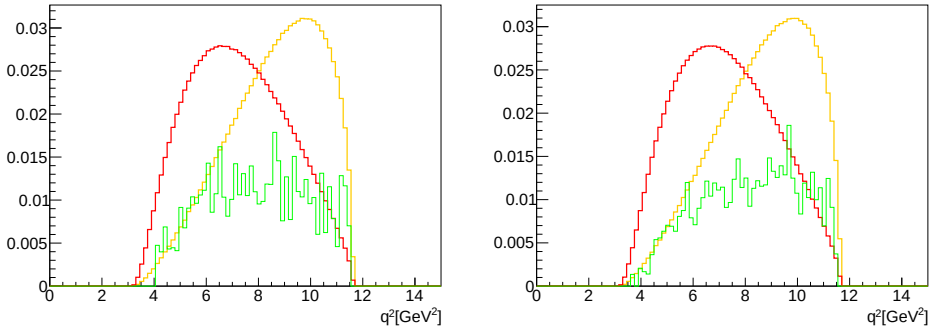
Figure 10.47: Produced q^2 distribution for Standard Model (red) and New Physics (yellow), Reconstruction efficiency (green) and reconstruction efficiency as cross-feed (blue) in the sample with $M^2_{\text{miss}} > 0.55\,\text{GeV}^2\,\text{c}^{-4}$

split value	eff.	eff. NP	loss
0.55	0.0278%	0.0278%	0.0%
0.70	0.0277%	0.0276%	0.0%
0.85	0.0274%	0.0274%	0.0%
1.00	0.0270%	0.0270%	0.0%
1.25	0.0263%	0.0262%	0.1%
1.50	0.0254%	0.0254%	−0.0%
1.75	0.0243%	0.0243%	0.0%
2.00	0.0229%	0.0229%	0.0%

Table 10.10: Efficiencies of New Physics decays vs. Standard Model decays in dependence of production p_ℓ value for $B^- \to D^{*0}\tau\nu$ decays.

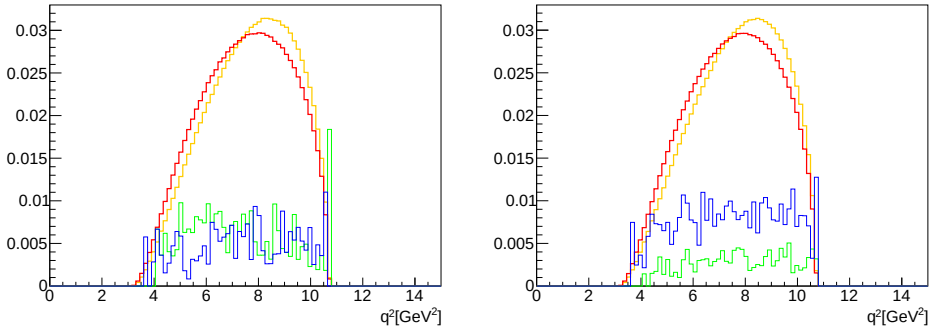
τ signal. Before comparing it to the produced distribution, it is corrected for efficiency, and normalized to the fitted τ signal yield. The distributions are combined for B^+ and B^0 and shown in Fig. 10.51, along with the respective expected distributions, taken from signal MC generator level. A χ^2 test was performed, the results are listed in Table 10.11. The numbers of degrees of freedom differ, because of the zero-entry bins in the D^* samples coming from the kinematic boundaries.

The table shows, that this test is not sensitive to the small differences in the samples with D^* mesons. In the D modes there is a difference, where the NP model is disfavored on the 2σ level from this distribution alone. So while this test itself will likely not provide significant exclusion power, it might help support or oppose the implications from measurements of the branching ratio.



(a) $B^0 \to D^+\tau^-\nu$

(b) $B^- \to D^0\tau^-\nu$



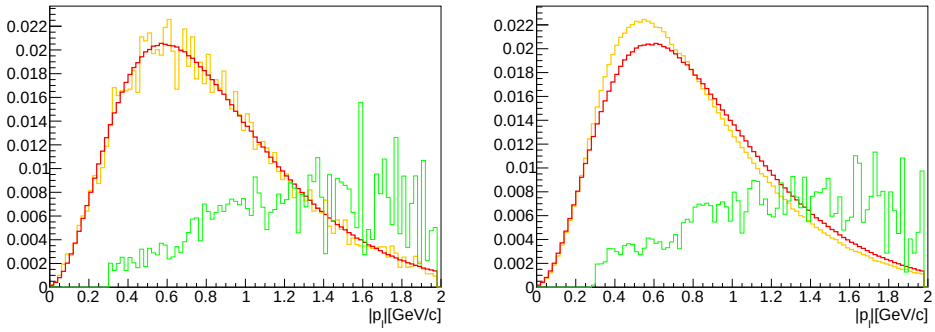
(c) $B^0 \to D^{*+}\tau^-\nu$

(d) $B^- \to D^{*0}\tau^-\nu$

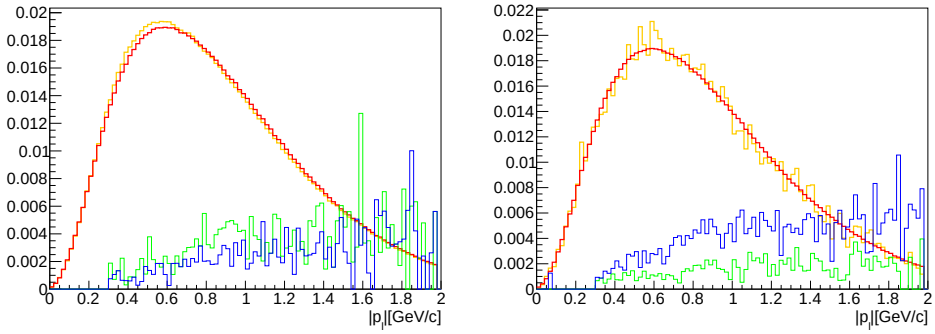
Figure 10.48: Produced q^2 distribution for Standard Model (red) and New Physics (yellow), Reconstruction efficiency (green) and reconstruction efficiency as cross-feed (blue) in the sample with $M^2_{\text{miss}} > 2.00\,\text{GeV}^2\,\text{c}^{-4}$

sample	NDF	χ^2/NDF	p-value
$D\tau$ SM	15	0.89	56.9%
$D\tau$ $\tan\beta/m_{H^+} = 0.5\,\text{c}^2\,\text{GeV}^{-1}$	15	1.71	4.7%
$D^*\tau$ SM	13	0.65	80.3%
$D^*\tau$ $\tan\beta/m_{H^+} = 0.5\,\text{c}^2\,\text{GeV}^{-1}$	13	0.57	86.8%

Table 10.11: Results of the χ^2 tests between measured and expected q^2 distribution under different model assumptions.

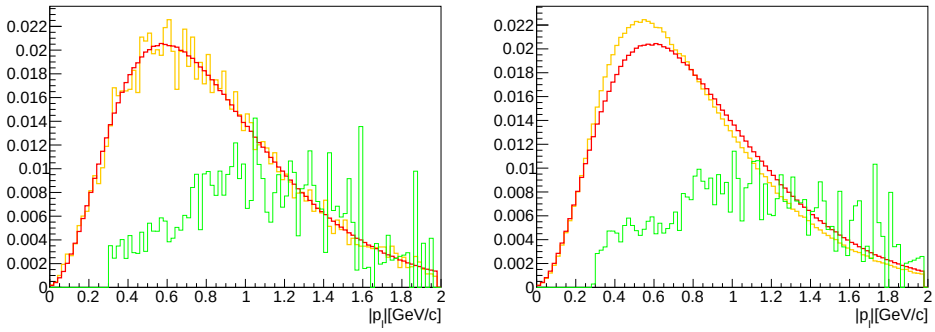


(a) $B^0 \to D^+\tau^-\nu$ (b) $B^- \to D^0\tau^-\nu$

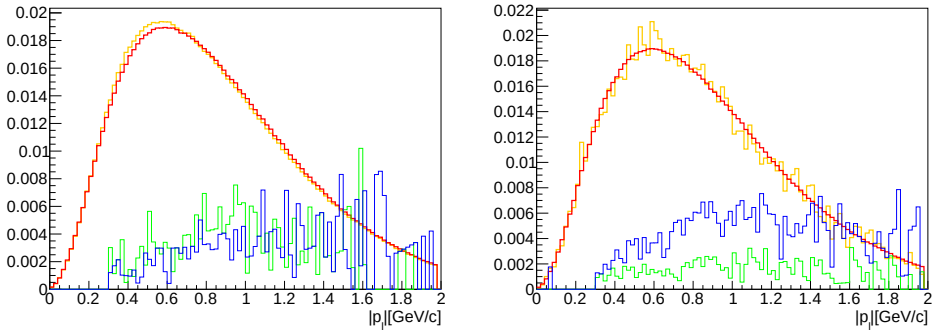


(c) $B^0 \to D^{*+}\tau^-\nu$ (d) $B^- \to D^{*0}\tau^-\nu$

Figure 10.49: Produced p_ℓ distribution for Standard Model (red) and New Physics (yellow), Reconstruction efficiency (green) and reconstruction efficiency as cross-feed (blue) in the sample with $M^2_{\text{miss}} > 0.55\,\text{GeV}^2\,\text{c}^{-4}$

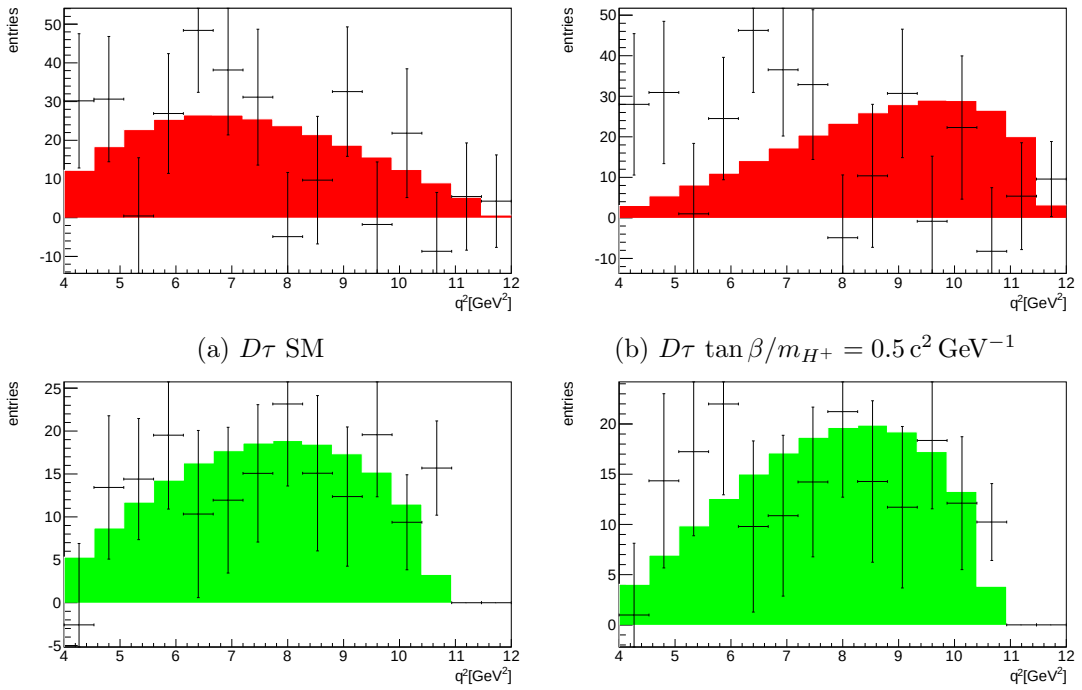


(a) $B^0 \to D^+\tau^-\nu$ (b) $B^- \to D^0\tau^-\nu$



(c) $B^0 \to D^{*+}\tau^-\nu$ (d) $B^- \to D^{*0}\tau^-\nu$

Figure 10.50: Produced p_ℓ distribution for Standard Model (red) and New Physics (yellow), Reconstruction efficiency (green) and reconstruction efficiency as cross-feed (blue) in the sample with $M^2_{\text{miss}} > 2.00\,\text{GeV}^2\,\text{c}^{-4}$



(a) $D\tau$ SM

(b) $D\tau$ $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$

(c) $D^*\tau$ SM

(d) $D^*\tau$ $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$

Figure 10.51: Background-subtracted q^2 distributions of the τ signal in the region of $M^2_{\mathrm{miss}} > 0.85\,\mathrm{GeV}^2\,\mathrm{c}^{-4}$. The distributions have been efficiency corrected and normalized to the fitted yield. The histogram is the respective expected distribution from signal MC. Left: Standard model result, right: New Physics result with $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$.

11. Results

This chapter lists the results of the fitting procedure. An in-depth discussion of these results concerning other measurements and alternative physics models will be given in Chapter 13.

11.1 Fit Result

The previously defined fitting procedure was applied to the real data sample using the `Minuit2` numeric minimization package within the `ROOT` framework. We extract the following result:

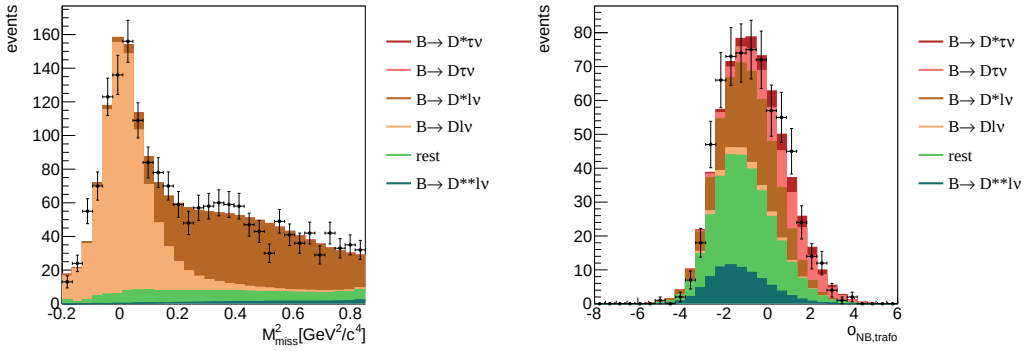
$$R = 0.375^{+0.064}_{-0.063}$$
$$R^* = 0.293^{+0.039}_{-0.037} \ .$$

The results for all free parameters are given in Table 11.1. R and R^* are determined directly in this procedure, assuming isospin symmetry. Due to the constraints concerning the yields for τ and lepton signal and cross-feed in the different samples, all available yields are used in the calculations. Table 11.2 sums up the yields of signal and normalization that were extracted in the fitting procedure.

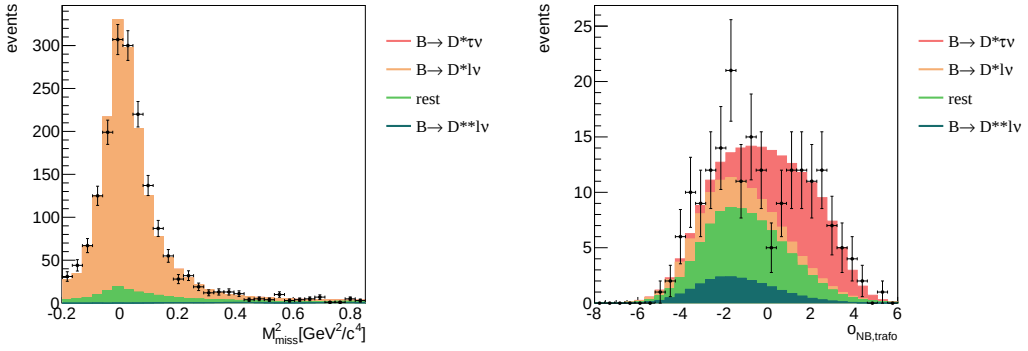
The correlation matrix of the fit parameters is shown in Table 11.4. The correlation between R and R^* is strongly negative (-0.56). The highest other correlations to these parameters are found in the D^{**} components, with 0.1 to 0.2 for R and ≈ 0.3 for R^*. Correlations among other parameters are relatively small except for those between D^{**} components and lepton cross-feed in the same reconstruction mode.

11.2 Fit Projections

The fit projections for the fitting variables are shown in Fig. 11.1 for the B^0 decay modes and in Fig. 11.2 for the B^+ decay modes. There is a good overall agreement between data and fit model within uncertainties. Projections of the fit result on several other observables can be found in Section B in the Appendix.

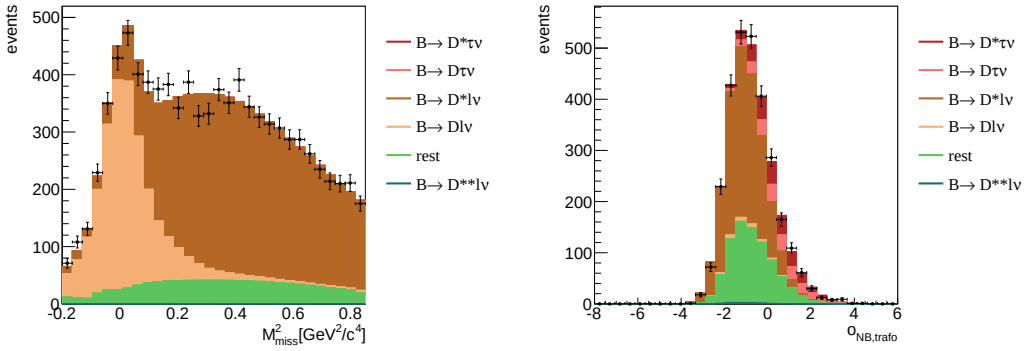


(a) M^2_{miss} in $D^+\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^+\ell^-$

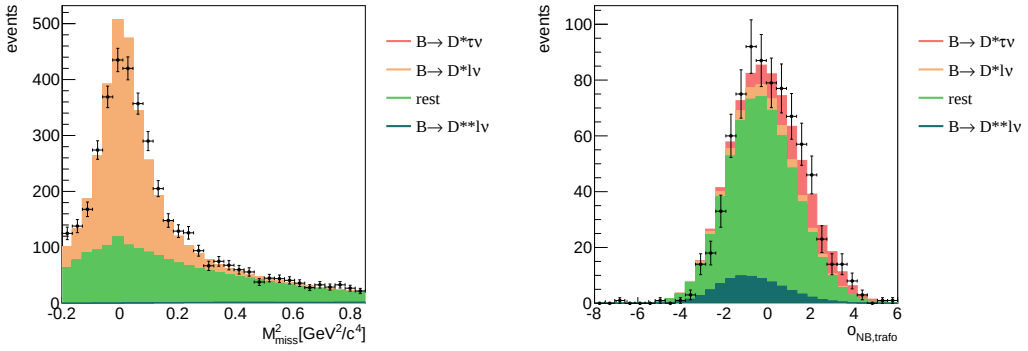


(c) M^2_{miss} in $D^{*+}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*+}\ell^-$

Figure 11.1: Fit projections for the B^0 reconstruction modes. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.



(a) M^2_{miss} in $D^0\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^0\ell^-$



(c) M^2_{miss} in $D^{*0}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*0}\ell^-$

Figure 11.2: Fit projections for the B^+ reconstruction modes. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.

parameter	value	upper error	lower error	MC exp.
R	0.375	+0.064	−0.063	-
R^*	0.293	+0.039	−0.037	-
$Y_{D^+\ell^-,\ell \text{ signal}}$	844	+34	−33	869.7
$Y_{D^+\ell^-,\ell \text{ CF}}$	924	+47	−46	972.7
$Y_{D^+\ell^-,D^{**}}$	108	+38	−37	133.1
$Y_{D^0\ell^-,\ell \text{ signal}}$	2303	+64	−63	2293.6
$Y_{D^0\ell^-,\ell \text{ CF}}$	7324	+122	−121	7444.1
$Y_{D^0\ell^-,D^{**}}$	131	+81	−80	208.3
$Y_{D^{*+}\ell^-,\ell \text{ signal}}$	1609	+43	−43	1682.1
$Y_{D^{*+}\ell^-,D^{**}}$	36	+18	−18	76.0
$Y_{D^{*0}\ell^-,\ell \text{ signal}}$	2188	+60	−59	2279.3
$Y_{D^{*0}\ell^-,D^{**}}$	117	+39	−38	40.2

Table 11.1: Fit results for all free parameters. In the calculation of R^* also the yields of the cross-feed components are used. The last column gives the expected yields as derived from simulated data.

origin	composing yields	full yield
$B \to D\tau\nu_\tau$	$Y_{D^+\ell^-,\tau \text{ signal}}$, $Y_{D^0\ell^-,\tau \text{ signal}}$	320
$B \to D^*\tau\nu_\tau$	$Y_{D^{*+}\ell^-,\tau \text{ signal}}$, $Y_{D^{*0}\ell^-,\tau \text{ signal}}$, $Y_{D^+\ell^-,\tau \text{ CF}}$, $Y_{D^0\ell^-,\tau \text{ CF}}$	503
$B \to D\ell\nu_\ell$	$Y_{D^+\ell^-,\ell \text{ signal}}$, $Y_{D^0\ell^-,\ell \text{ signal}}$	3147
$B \to D^*\ell\nu_\ell$	$Y_{D^{*+}\ell^-,\ell \text{ signal}}$, $Y_{D^{*0}\ell^-,\ell \text{ signal}}$, $Y_{D^+\ell^-,\ell \text{ CF}}$, $Y_{D^0\ell^-,\ell \text{ CF}}$	12045

Table 11.2: Fitted yields of signal and normalization. τ yields are not a direct fit parameter, but constrained as described in Section 9.1.

11.3 Branching Ratios

To calculate the branching ratio of different $B \to D^{(*)}\tau\nu$ modes, we assume the world average branching ratios for $B \to D^{(*)}\ell\nu$, taken from the Particle Data Group[37], and multiply it by the respective $R^{(*)}$ value. The values are given in Table 11.3 and agree very well with the world average values. The errors were combined by calculating the squared relative error on the branching ratios as sum of the squared relative errors of $R^{(*)}$ and the world average semileptonic branching ratio.

decay	semileptonic WA	semitauonic calc.	semitauonic WA
$B^0 \to D^-\tau^+\nu_\tau$	$2.19 \pm 0.12\%$	$0.82 \pm 0.15\%$	$1.03 \pm 0.22\%$
$B^+ \to D^0\tau^+\nu_\tau$	$2.27 \pm 0.11\%$	$0.85 \pm 0.15\%$	$0.77 \pm 0.25\%$
$B^0 \to D^{*-}\tau^+\nu_\tau$	$4.93 \pm 0.11\%$	$1.85 \pm 0.32\%$	$1.84 \pm 0.22\%$
$B^+ \to D^{*0}\tau^+\nu_\tau$	$5.69 \pm 0.19\%$	$2.13 \pm 0.37\%$	$1.88 \pm 0.20\%$

Table 11.3: Branching ratios for semitauonic B meson decays, calculated from the measured values for R and R^*, and the world average values (WA) for the semileptonic branching ratios.

	$global$	R	R^*	$Y_{D^+\ell^-,D^{**}}$	$Y_{D^+\ell^-,\ell\ \mathrm{CF}}$	$Y_{D^+\ell^-,\ell\ \mathrm{signal}}$	$Y_{D^0\ell^-,D^{**}}$	$Y_{D^0\ell^-,\ell\ \mathrm{CF}}$	$Y_{D^0\ell^-,\ell\ \mathrm{signal}}$	$Y_{D^{*+}\ell^-,D^{**}}$	$Y_{D^{*+}\ell^-,\ell\ \mathrm{signal}}$	$Y_{D^{*0}\ell^-,D^{**}}$	$Y_{D^{*0}\ell^-,\ell\ \mathrm{signal}}$
R	0.683	1.000	-0.555	-0.198	-0.032	-0.101	-0.188	0.017	-0.089	0.168	0.056	0.165	0.033
R^*	0.701	-0.555	1.000	0.050	0.004	0.040	-0.176	0.026	-0.025	-0.301	-0.103	-0.300	-0.047
$Y_{D^+\ell^-,D^{**}}$	0.560	-0.198	0.050	1.000	-0.498	0.009	0.061	-0.006	0.024	-0.013	-0.003	-0.015	-0.004
$Y_{D^+\ell^-,\ell\ \mathrm{CF}}$	0.558	-0.032	0.004	-0.498	1.000	-0.196	0.012	-0.001	0.004	-0.006	-0.002	-0.001	0.000
$Y_{D^+\ell^-,\ell\ \mathrm{signal}}$	0.262	-0.101	0.040	0.009	-0.196	1.000	0.026	-0.003	0.011	-0.011	-0.003	-0.012	-0.003
$Y_{D^0\ell^-,D^{**}}$	0.636	-0.188	-0.176	0.061	0.012	0.026	1.000	-0.514	0.089	0.053	0.019	0.054	0.012
$Y_{D^0\ell^-,\ell\ \mathrm{CF}}$	0.592	0.017	0.026	-0.006	-0.001	-0.003	-0.514	1.000	-0.268	0.000	-0.039	-0.011	-0.002
$Y_{D^0\ell^-,\ell\ \mathrm{signal}}$	0.314	-0.089	-0.025	0.024	0.004	0.011	0.089	-0.268	1.000	0.008	0.003	0.008	0.002
$Y_{D^{*+}\ell^-,D^{**}}$	0.366	0.168	-0.301	-0.013	-0.006	-0.011	0.053	0.000	0.008	1.000	-0.175	0.090	0.014
$Y_{D^{*+}\ell^-,\ell\ \mathrm{signal}}$	0.244	0.056	-0.103	-0.003	-0.002	-0.003	0.019	-0.039	0.003	-0.175	1.000	0.031	0.005
$Y_{D^{*0}\ell^-,D^{**}}$	0.337	0.165	-0.300	-0.015	-0.001	-0.012	0.054	-0.011	0.008	0.090	0.031	1.000	-0.139
$Y_{D^{*0}\ell^-,\ell\ \mathrm{signal}}$	0.168	0.033	-0.047	-0.004	0.000	-0.003	0.012	-0.002	0.002	0.014	0.005	-0.139	1.000

Table 11.4: Correlation matrix of the fit parameters. We see a strong negative correlation between R and R^*. Other than that correlations are rather small except for a strong negative correlation between ℓ cross-feed and D^{**} background in the same reconstruction sample.

12. Systematic Uncertainties

One of the main reasons to do a *relative* measurement of the branching ratio and only use leptonic τ decays is to reduce systematic uncertainties. Especially the estimation of detector acceptance behavior is difficult to validate. It comprises effects like the reconstruction efficiency of low-energetic pions, the accuracy of particle identification, or final state radiation. As the measured parameters R and R^* are ratios of yields of the same visible final state, with only small differences in the momenta of the visible particles, these acceptance effects influence numerator and denominator similarly and approximately cancel out. There are however other non-negligible systematic influences, that will be covered in this chapter, with the largest coming from our limited understanding of the D^{**} background and from uncertainties on fitting constants due to Monte Carlo statistics. The systematic uncertainties are summarized in Table 12.2. The result including systematic uncertainties is:

$$R = 0.375^{+0.064}_{-0.063} \pm 0.026$$
$$R^* = 0.293^{+0.039}_{-0.037} \pm 0.015 \ .$$

12.1 Production Parameters

As explained in sections 7.3 and 7.4, the parameters in the Monte Carlo production of $B \to D^{(*(*))}\ell\nu$ are known with limited precision. Different sets of weights are therefore created to test several shape hypotheses with the 1σ values of individual production parameters. The fit procedure is then repeated for each weight set and the deviations to the default set are used as systematic uncertainty. The uncertainties are combined in the summary table as "$D^{(*(*))}\ell\nu$ shapes".

12.2 D^{**} Composition

The D^{**} background events have a strong influence on the extracted amount of τ signal, because they occupy the same region in the M^2_{miss} spectrum. Besides the shape uncertainties that were objected in Section 12.1, there are also uncertainties from the composition: branching ratios of different D^{**} states are only vaguely known. The fit is therefore repeated several times, twice for each D^{**} state for its yield to be varied up and down by its uncertainty. For The $D^*(2S)$ states, no good measurement exists, which is why the yield is varied by 100%. The full overview of the uncertainties is given in Table 12.1. The difference in the fit result is used as a systematic uncertainty. They are combined in the summary table as "D^{**} composition".

12.3 Factors for Simultaneous Fitting

All factors that are used in the simultaneous fitting procedure (see Section 9.3) were evaluated using generic Monte Carlo. They are only known up to a certain precision due to statistical fluctuations. To evaluate the effects of this imprecision to the final result, each factor was varied by its uncertainty. The influence of these factors is shown individually in the summary table. It can be seen from this table, that most factors - especially the yields - have little influence on the final result. However, the uncertainties coming from the factors for the efficiency ratios and the cross-feed probability ratios are among the largest contributions, on par with the D^{**} and shape uncertainties.

12.4 PDF Shapes

The PDFs for all fit components have been determined on simulated data and are therefore sensitive to differences between data and Monte Carlo. Especially components that are similar to the τ signal shape may have a noticeable influence on the result. To evaluate this behavior, the shapes of all components were changed and the fit was repeated.

The fit uses smoothed histogram PDFs in M^2_{miss}, so in the repeated fit, the smoothing was omitted. The change of the result was taken as symmetric systematic uncertainty and is listed as "M^2_{miss} shape" in the summary table.

In the $o_{\text{NB,trafo}}$ frame, it is a bit more complicated: The default fit method uses bifurcated Gaussian functions. The trivial replacement would be histogram PDFs, but this fails, as we have long tails, that can have zero-entry bins on Monte Carlo. If there is a data point in a bin with probability 0, the likelihood is not defined. Instead we use kernel estimator functions, that represent every entry on Monte Carlo with a Gaussian distribution. We use adaptive kernels, that adjust the width of the Gaussian distributions to the local event density, but the base width is still an adjustable parameter. Tests on Monte Carlo show, that this width influences the fit by inducing an asynchronous bias on the measured R and R^* values, mostly due to the tails with high $o_{\text{NB,trafo}}$ values. To counter this, the base width is altered and the value that produces the lowest approximated bias on 5 streams of Monte Carlo is taken. The default value is 1.0, while we use 1.65. Again, the deviation from the default fit value is taken as symmetric systematic uncertainty, labeled as "$o_{\text{NB,trafo}}$ shape" in the summary table. It is among the highest contributions to the systematic uncertainties of the measurement.

12.5 Lepton ID Efficiency

As shown in Section 7.5, the identification efficiencies are slightly different for light leptons and tauons. The modeling on generic MC has been validated on real data and was found to be quite accurate, but some uncertainties remain. These efficiencies affect our result by

state	uncertainty in %
D^*_2	42.3
D^*_0	34.6
D_1	14.9
D'_1	36.2
$D(2S)$	100.0
$D^*(2S)$	100.0

Table 12.1: Uncertainties of D^{**} state yields used to test systematic influences.

modifying the factors for efficiency ratios $f^{(*)}_{R,B^{0/+}}$. The great overlap in Tables 7.1 and 7.2 suggests, that the uncertainties are 100% correlated for all samples. This means, that the yields in the fit are not affected by introducing new dependencies and the effect on R and R^* can be calculated by simply applying a correction factor. This factor is derived by creating a ratio of reconstructed $B \to D^{(*)}\ell\nu$ and $B \to D^{(*)}\tau\nu$ events, as they are used in the calculation of $f^{(*)}_{R,B^{0/+}}$ in Section 9.1.2, but using all samples listed in Tables 7.1 and 7.2. The relative uncertainty on this factor, that comes from the individual uncertainties in the subsamples, is the same as on R and R^*. It is evaluated to 0.5%.

	R [%]	R^* [%]	correlation
$D^{(*(*))}\ell\nu$ shapes	4.2	1.5	0.04
D^{**} composition	1.3	3.0	-0.63
wrong charge factor	0.0	0.0	0.84
$Y_{D^+\ell^-,D_s}$	0.1	0.0	-0.95
$Y_{D^+\ell^-,\mathrm{rest}}$	0.1	0.0	-0.92
$Y_{D^+\ell^-,\mathrm{wrong}D}$	0.4	0.1	-0.99
$Y_{D^+\ell^-,\mathrm{wrong}\ell}$	0.3	0.1	-0.99
$Y_{D^0\ell^-,D_s}$	0.0	0.0	0.81
$Y_{D^0\ell^-,\mathrm{rest}}$	0.0	0.0	0.60
$Y_{D^0\ell^-,\mathrm{wrong}D}$	0.3	0.2	0.96
$Y_{D^0\ell^-,\mathrm{wrong}\ell}$	0.2	0.1	0.98
$Y_{D^{*+}\ell^-,D_s}$	0.1	0.1	-1.00
$Y_{D^{*+}\ell^-,\mathrm{rest}}$	0.0	0.0	-0.99
$Y_{D^{*+}\ell^-,\mathrm{wrong}D^*}$	0.1	0.1	-1.00
$Y_{D^{*+}\ell^-,\mathrm{wrong}\ell}$	0.3	0.5	-1.00
$Y_{D^{*0}\ell^-,D_s}$	0.0	0.0	-0.99
$Y_{D^{*0}\ell^-,\mathrm{rest}}$	0.0	0.0	-0.96
$Y_{D^{*0}\ell^-,\mathrm{wrong}D^*}$	0.1	0.1	-0.83
$Y_{D^{*0}\ell^-,\mathrm{wrong}\ell}$	0.1	0.2	-1.00
g_{B^0}	2.2	2.0	-1.00
g_{B^+}	1.7	1.0	-1.00
f_{R,B^0}	2.5	0.7	-0.98
f_{R,B^+}	1.8	0.4	0.86
f^*_{R,B^0}	1.3	2.5	-0.99
f^*_{R,B^+}	0.7	1.1	0.94
M^2_{miss} shape	0.6	1.0	0.00
$o_{\mathrm{NB,trafo}}$ shape	3.2	0.8	0.00
lepton PID efficiency	0.5	0.5	1.00
Σ	7.1	5.2	-0.32

Table 12.2: Overview of relative systematic uncertainties in percent. The last column gives the correlation between R and R^*.

12.6 Correlations

To check the compatibility of several physics models with our measurement, it is important to involve the correlation of R and R^* in the calculations, as this can significantly increase the exclusion power in the $R-R^*$ plane. Therefore, the correlation of R and R^* concerning the individual systematic influences is also of interest. These are given in the last column of Table 12.2. They were calculated using 500 toy experiments. For each systematic effect the

average covariance was divided by the product of the individual standard deviation of R and R^*, calculated as the square root of the average variance. The shape uncertainties are calculated without toy experiments and assumed to be uncorrelated. For the uncertainties from lepton ID efficiencies 100% correlation is assumed. The total correlation is then computed as the sum of the average covariance in each systematic effect, divided by the product of the total systematic standard deviations of R and R^*.

13. Discussion

This chapter takes a closer look at the implications of the result. The compatibility with the Standard Model prediction and the result from the BaBar collaboration is addressed in Section 13.1, where the $R - R^*$ plane is examined. Section 13.2 focuses on the differences to the result in Reference [6], which was not finally approved by the Belle collaboration, but was obtained on a very similar data set, so the discrepancy must be at least qualitatively explained. Section 13.3 contains further tests for new physics influences in this measurement. The last Section 13.4 gives a naive combination with the result of the BaBar collaboration to estimate the impact on the compatibility of the world average with the Standard Model.

13.1 $R - R^*$ Plane

The fitting procedure as well as the tests for systematic uncertainties confirm a strong negative correlation between R and R^*. It can therefore be misleading to look at the individual uncertainties alone. The 2-dimensional $R-R^*$ likelihood plane provides a better insight into the compatibility of this result with several other assumptions. To create it, the $R-R^*$ plane is scanned and for each point, the fit procedure is applied with the R and R^* values fixed. To add systematic uncertainties the distribution is convoluted numerically with a normalized 2-dimensional normal distribution using the full systematic uncertainties and correlation of R and R^* as given in Table 12.2.

To calculate the exclusion power, the individual likelihood values are compared to the best one. According to Wilks' theorem [38] the quantity $-2 \times (\log \frac{\mathcal{L}_{\text{hypothesis}}}{\mathcal{L}_{\text{best}}})$ for a given hypothesis follows a χ^2 distribution with two degrees of freedom, as $\mathcal{L}_{\text{best}}$ was determined with the two parameters R and R^* floating. For each given hypothesis in the $R - R^*$ plane, the probability to get L_{best} or a better likelihood can be calculated from the χ^2 distribution and transformed to units of sigma. The values for individual hypotheses are given in Table 13.1.

The exclusion power in the $R - R^*$ plane is shown in Fig. 13.1 and various hypotheses are indicated. The mentioned previous model is the result given in Reference [6], which was obtained on an overlapping dataset, so the significant deviation can not come from statistical fluctuations only. This is addressed in the following Section 13.2. It has to be noted, that the errors on other experimental results (BaBar and previous) were not taken into account concerning the compatibility. The errors on the SM prediction are of theoretical nature and in this case very small compared to the experimental ones.

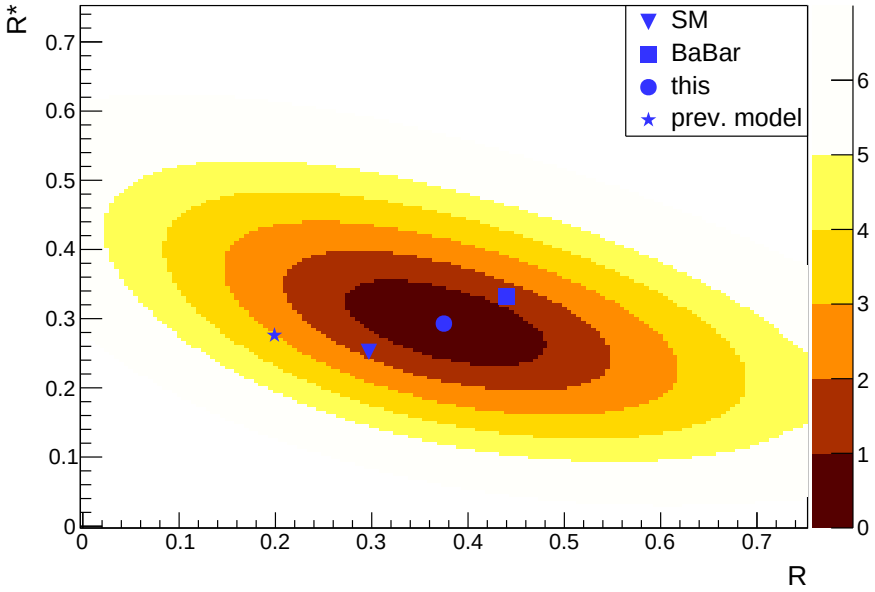


Figure 13.1: Exclusion level of $R - R^*$ value assumptions in sigma, systematic uncertainties included.

hypothesis	χ^2 probability	exclusion
Standard Model	6.5%	$1.8\,\sigma$
BaBar measurement	17.6%	$1.4\,\sigma$
previous result	0.3%	$3.0\,\sigma$

Table 13.1: Exclusion power of the fit result with respect to competing hypotheses. The previous result was obtained on a very similar dataset and the discrepancy is addressed in Section 13.2.

13.2 Difference to Previous Result

The result of the measurement of the same quantity from two years ago [6] is incompatible with the current result by about $3\,\sigma$. Especially the measurement of R differs, by a factor of ≈ 2. As the previous measurement also provides a fit method that is not constrained in the isospin it can be derived, that the main difference comes from a very low value for R^0, which is measured in the $D^{(*)+}\ell^-$ sample only.

Although the selection processes are not identical, there is a large overlap between the final fit samples. About $(70 \pm 10)\%$ of the events in each final fit sample are used in both analyses. So this is unlikely to be the cause for a massively different τ signal yield. There are however other issues of the previous analysis, that were discovered over the course of the development of the new analysis procedure:

Efficiency Ratio Calculations

In the previous fit procedure a wrong value was assumed for R and R^* in the production of generic Monte Carlo. It is used in the calculation of the efficiency ratio (see Section 9.1), that is directly multiplied with the measured ratio. So the old ratios are by construction too low, $\approx 12\%$ for R and $\approx 4\%$ for R^*. This is not visible in the cross validation of the previous analysis, as the results are compared to the same wrong values.

Fit Bias

In the case of the isospin-unconstrained fit, the previous analysis provides cross-validation plots for R in its subsamples (see figure 5.81 (a) in [6]). For the mentioned $D^{(*)+}\ell^-$ sample with particularly low τ signal yield these validation plots indicate an underestimation in all 5 streams of generic Monte Carlo. This is likely connected to the following issue:

Bias in the D^{**} Sample

The cross validation of the D^{**} enriched sample in the same reconstruction mode $D^{(*)+}\ell^-\pi^0$ is shown in in figure 5.23 (a) in [6]. There we have the opposite effect of a systematic overestimation of D^{**} yields. As D^{**} background and τ signal are strongly anti-correlated, this would explain the underestimation of the respective τ signal yield.

Further Issues in the D^{**} Sample

The D^{**} enriched sample was the initial place where the previous analysis procedure was found to be flawed. Several methodical problems surfaced during it's review:

Correlations: The fitting variables in the D^{**} enriched sample were found to be strongly correlated. This was not addressed in the construction of the 2-dimensional probability density functions.

Multiple Candidates: Although the $D\ell$ and the $D\ell\pi^0$ sample were fitted simultaneously, the best candidate selection was applied individually. This resulted in a very high number of events with a candidate in both samples. The constraints in the fit then caused an underestimation of the uncertainties of the derived yields.

Signal Definition: There was a flaw in the definition of the D^{**} signal in the enriched sample. After correction between 20% and 60% of the presumed D^{**} signal disappeared. So the constraint between D^{**} signal in the enriched sample and D^{**} background in the regular sample was not justified.

So the estimation of D^{**} background yields using the enriched sample was invalid. Even without this knowledge, D^{**} background yields were by far the highest systematic uncertainty in the previous result, higher than the new result's systematic uncertainties combined.

Correlations

In the construction of the 2-dimensional probability density function of the regular fit no correlation between the two variables M^2_{miss} and $o_{\text{NB,trafo}}$ was assumed. Later investigations showed, that there are correlations, however small. The problem is, that they became apparent in higher regions of M^2_{miss}, where the τ signal has the best purity and therefore great influence on the fit result. The $o_{\text{NB,trafo}}$ distributions shift slightly to lower values in this region, which causes data to exceed the expectations in regions of low $o_{\text{NB,trafo}}$ and fall short in regions of high $o_{\text{NB,trafo}}$. The τ signal is dominant in the latter and can therefore be easily underestimated.

Conclusion

The flaws that were found in the previous analysis procedure would either directly cause an underestimation of the measured τ signal yields and therefore R and R^*, or they would introduce large systematic errors of unknown direction. In this light the analysis procedure presented in this thesis is assumed to be superior, and the large deviation from the previous result is no longer surprising.

13.3 New Physics Search

13.3.1 Fit Procedure

The presented procedure is also suitable to test the compatibility of the recorded data with the 2-Higgs-doublet model of type II with different sets of parameters. However, it does not suffice to calculate the probability of the predicted values for a set of parameters within our likelihood distribution. As the choice of parameters influences kinematic distributions (as shown in Section 10.8.1), the fit procedure has to be repeated with a sample of signal Monte Carlo that was generated using these parameters, to account for changes in probability density distributions. Furthermore, efficiencies for the τ signal are slightly different, as shown in Section 10.8.1. The loss in efficiency that is listed in Tables 10.3 to 10.6 and 10.7 to 10.10 is used to modify the NP expectation in the exclusion plane.

13.3.2 Results

So far, this procedure was done for the parameter point $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$. The projections of the fit are shown in Figs. 13.2 and 13.3. The resulting values are:

$$R = 0.315^{+0.058}_{-0.057} \pm 0.022$$
$$R^* = 0.297^{+0.039}_{-0.038} \pm 0.015\ ,$$

where the same relative systematic uncertainties as in the default fit are applied.

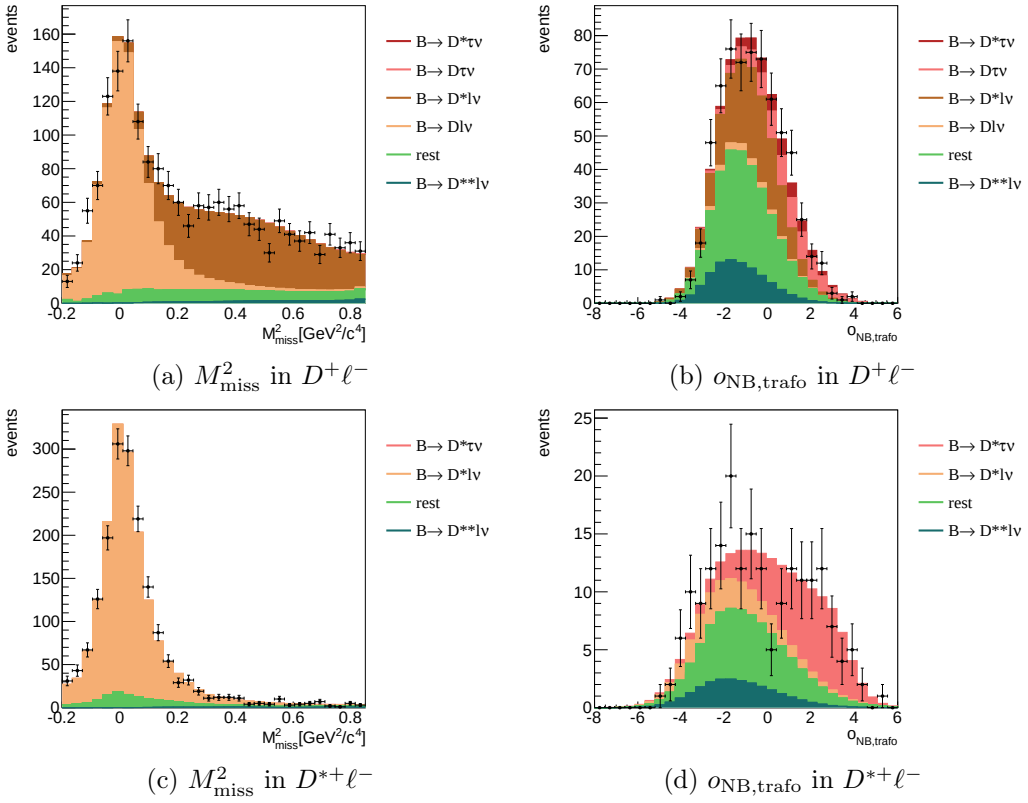
The prediction for the 2HDM of type II according to Eq. (2.9) is $R_{2HDM} = 0.541 \pm 0.115$ and $R^*_{2HDM} = 0.235 \pm 0.007$. Due to the changes in efficiency, the expected values in the exclusion plane are $R_{2HDM} = 0.565 \pm 0.120$ and $R^*_{2HDM} = 0.238 \pm 0.007$. It is displayed in Fig. 13.4 and the central value is excluded by 3.4σ. However, the exclusion is not significant due to the large uncertainty on the predicted R value. The BaBar collaboration gave significant exclusion levels for a large variety of parameter points in the type II 2HDM, but with our result, we will not achieve a similar level of exclusion. The reason is, that R^* is much less affected by model changes than R. With our measured R^* value being much closer to the standard model expectation and having higher uncertainties, the level of exclusion can not be as strong as in the BaBar measurement, that had a 2.7σ discrepancy to the SM expectation in R^* by itself.

13.3.3 q^2 Distributions

The observable, that is most influenced by new physics additions is q^2, the squared four-momentum transition between B meson and $D^{(*)}$ meson in the decay. We apply the method presented in Section 10.8.2 on the fit on data, using the new physics model with $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$. Fig. 13.5 shows the measured background subtracted and efficiency corrected q^2 distributions for the Standard Model and the new physics point. A χ^2 test could now reveal if the theoretical distribution and the one measured on recorded data are incompatible. Unfortunately, in our measurement both hypotheses are well compatible with p-values for the SM distribution of 64% ($D\tau$) and 11% ($D^*\tau$), and for the NP distribution of 53% ($D\tau$) and 49% ($D^*\tau$).

13.4 Combination with BaBar Result

The published result of the BaBar collaboration provides their fitted values and uncertainties, as well as their correlations. Using these, it is possible to create a naive combination of their result with mine. Naive is it in the sense, that the combination is only correct as



(a) M^2_{miss} in $D^+\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^+\ell^-$

(c) M^2_{miss} in $D^{*+}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*+}\ell^-$

Figure 13.2: Fit projections for the B^0 reconstruction modes under assumption of the 2HDM of type II with $\tan\beta/m_{H^+} = 0.5\,\text{c}^2\,\text{GeV}^{-1}$. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.

long as the assumption of a parabolic likelihood holds. To create a proper combination, the real distribution of the likelihood in the $R - R^*$ plane is required.

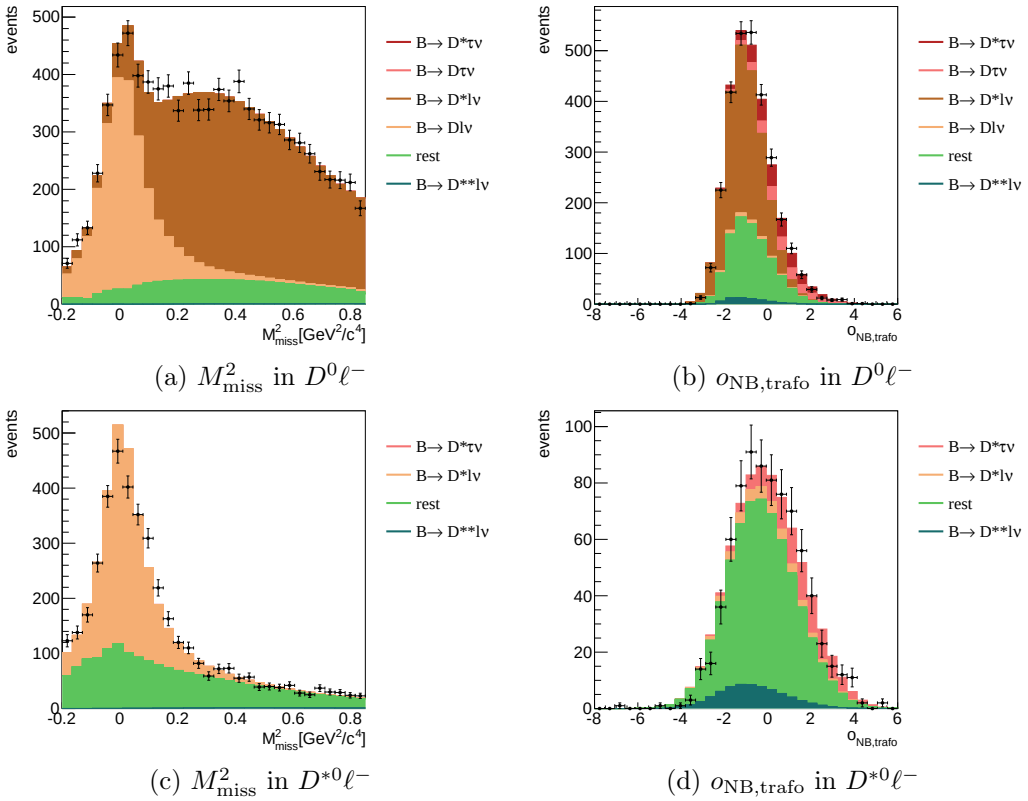
First I created a likelihood distribution according to a bivariate normal distribution with the BaBar result as input, excluding systematic uncertainties:

$$\begin{aligned} R_{BaBar} &= 0.440 \pm 0.058 \\ R^*_{BaBar} &= 0.332 \pm 0.024 \\ \rho_{RR^*} &= -0.45, \end{aligned}$$

where ρ_{RR^*} is the correlation. This distribution is then convoluted with a bivariate normal distribution, that represents their correlated systematic uncertainties:

$$\begin{aligned} \sigma_{R,\text{syst}} &= 0.042 \\ \sigma_{R^*,\text{syst}} &= 0.018 \\ \rho_{RR^*,\text{syst}} &= 0.05. \end{aligned}$$

The procedure is done in a binned $R-R^*$ plane, using the same binning as in Section 13.1. To combine the two measurements, the logarithmic values in both planes are added bin by bin, and the probabilities are recalculated analogous from the resulting logarithmic likelihood distribution. The exclusion plane is shown in Fig. 13.6 and the probabilities are listed in Table 13.2. The discrepancy with the SM point is remarkable: the combined exclusion is now on the 4σ level. The discrepancy is driven by the result of the BaBar collaboration, as my result alone is compatible within 1.8σ. But both measurements deviate clearly in the same direction, which causes the combination to increase the discrepancy to the SM point compared to the BaBar measurement alone.



(a) M^2_{miss} in $D^0\ell^-$ (b) $o_{\text{NB,trafo}}$ in $D^0\ell^-$

(c) M^2_{miss} in $D^{*0}\ell^-$ (d) $o_{\text{NB,trafo}}$ in $D^{*0}\ell^-$

Figure 13.3: Fit projections for the B^+ reconstruction modes under assumption of the 2HDM of type II with $\tan\beta/m_{H^+} = 0.5\,\text{c}^2\,\text{GeV}^{-1}$. Left is the fitted M^2_{miss} frame, right the fitted $o_{\text{NB,trafo}}$ frame.

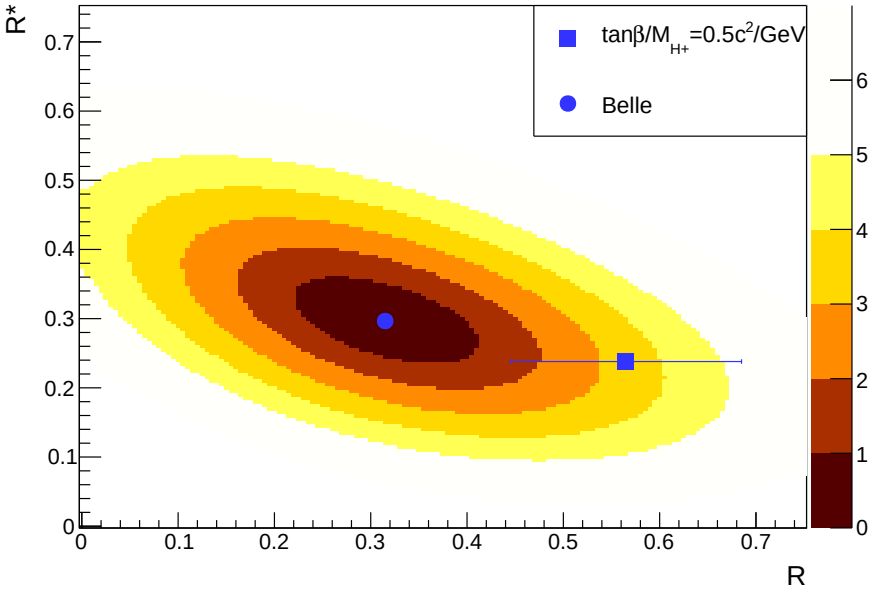
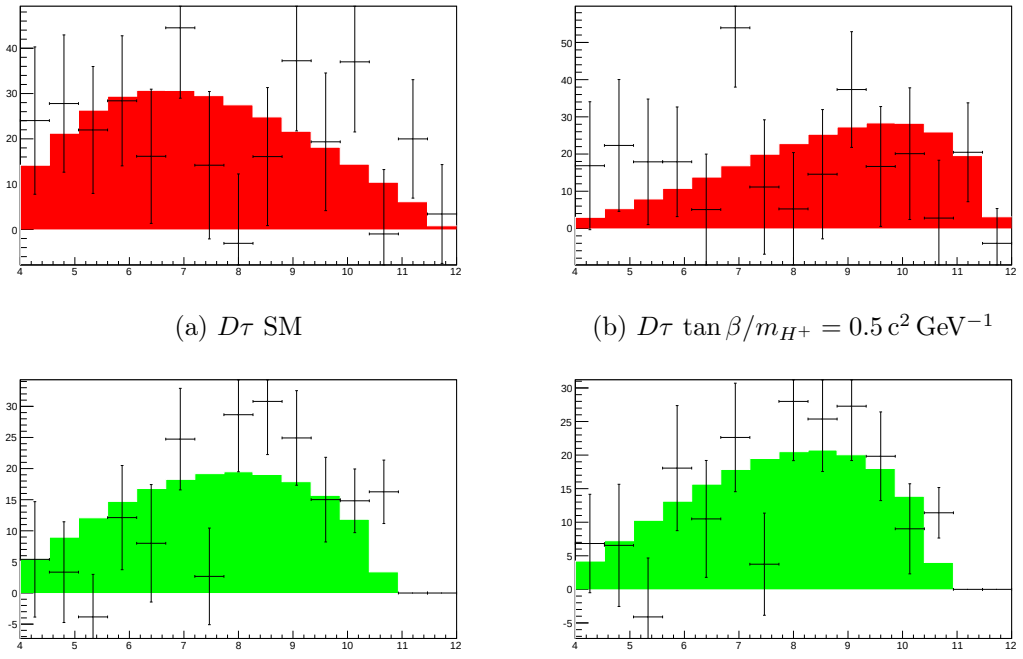


Figure 13.4: Exclusion level of type II 2HDM with $\tan\beta/m_{H^+} = 0.5\,\text{c}^2\,\text{GeV}^{-1}$ in sigma, systematic uncertainties included.



(a) $D\tau$ SM

(b) $D\tau$ $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$

(c) $D^*\tau$ SM

(d) $D^*\tau$ $\tan\beta/m_{H^+} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$

Figure 13.5: Background-subtracted q^2 distributions of the τ signal in the region of $M^2_{\mathrm{miss}} > 0.85\,\mathrm{GeV}^2$. The distributions have been efficiency corrected and normalized to the fitted yield. The histogram is the respective expected distribution from signal MC. Left: Standard Model result, right: New Physics result with $\tan\beta/m_{H^+} = 0.5\,\mathrm{GeV}^{-1}$.

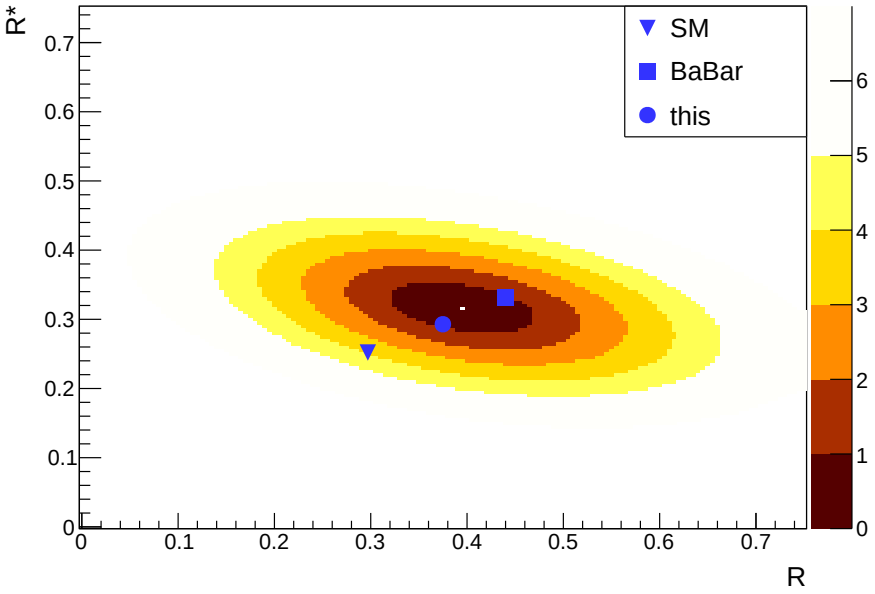


Figure 13.6: Exclusion level of $R - R^*$ value assumptions in sigma. Naive combination with the result of the BaBar collaboration, systematic uncertainties included.

hypothesis	χ^2 probability	exclusion
Standard Model	0.006%	$4.0\,\sigma$
BaBar measurement	41.1%	$0.8\,\sigma$
this Belle measurement	47.5%	$0.7\,\sigma$

Table 13.2: Exclusion power of the combined fit results of BaBar and Belle with respect to competing hypotheses.

14. Conclusion

In summary, this thesis presents the measurement of the branching ratio of $B \to D^{(*)}\tau\nu$ relative to $B \to D^{(*)}\ell\nu$ decays – where ℓ is either e or μ – using hadronic tagging channels and leptonic τ decays exclusively. It is performed on the full dataset on the $\Upsilon(4S)$ resonance of the Belle collaboration. The goal was to test the compatibility of this experimental data with the SM, as well as NP extensions, especially models with two Higgs doublets of type II. Both of them have shown discrepancies to the corresponding measurement by the BaBar collaboration.

The main challenge of the analysis is the special final state of $B \to D^{(*)}\tau(\to \ell\nu\nu)\nu$, containing three neutrinos that are virtually undetectable in the experiment. Using the full available event information and the clean environment of a lepton collider, kinematic observables were constructed, that contain information on the invisible particles. Yet, the signal decay is still distributed over a broad range, that is also populated by various background processes.

Although there are previous measurements of the same quantities by the BaBar and Belle collaborations, I redesigned the analysis strategy entirely. I pointed out significant differences between the experimental data of the BaBar and Belle collaboration – that are usually considered very similar – and showed that the BaBar analysis strategy is not suitable to produce a robust result with Belle data. I also disproved the validity of the strategy in the previous Belle analysis – that was based on the very same dataset as the one in this thesis – despite it having already passed an internal review process.

The measurement in this thesis is highly optimized and surpasses its predecessors in several ways. The various signal and background components of the dataset have been identified and corrected in simulated data for imperfections individually and in greater detail than before. For the most important background component, $B \to D^{**}\ell\nu$ decays, a dedicated set of simulated data was created to drastically reduce its systematic uncertainties. I conducted a series of tests to verify the validity of the simulation on a dedicated sample of recorded data. All previous analyses had the yield of the $B \to D^{**}\ell\nu$ background either externally constrained or neglected and my analysis is the first one, that measures it simultaneously on the same data sample as the $B \to D^{(*)}\tau\nu$ signal. The fitting procedure was developed blindly and avoids the usage of correlated 2-dimensional PDFs, that rely strongly on accurate correlation modeling and are prone to numeric problems from statistical limitations. Instead, I could show, that the results can be extracted by fitting the combined 1-dimensional PDFs to distributions in two kinematically different samples of the

dataset simultaneously, but in two different observables. This procedure and its sensitivity to NP influences was again thoroughly tested.

The results for the observables R and R^* from the fitting procedure are:

$$R = 0.375^{+0.064}_{-0.063} \pm 0.026$$
$$R^* = 0.293^{+0.039}_{-0.037} \pm 0.015 \ .$$

While the combined deviation of 1.8σ from the SM expectation is by itself not significant, my method shows an excess in both parameters, which supports the reported excess in the measurement by the BaBar collaboration, and will likely increase the tension of the world average values with the SM predictions. A preliminary combination with said result is in discrepancy on the 4.0σ level.

I have also performed the fitting procedure under the assumption of a two-Higgs-doublet model of type II with the parameter choice $\tan\beta/m_{H^\pm} = 0.5\,\mathrm{c}^2\,\mathrm{GeV}^{-1}$. The exclusion level is driven by the deviation in R^*, as the value of R covers a wide range and has larger uncertainties. My result for R^* is well compatible with expectations within uncertainties, which does not allow strong constraints on the possible values of $\tan\beta/m_{H^\pm}$. A combination of these tests with the analysis of the BaBar collaboration is more complicated than in the case of the SM fit model, as the NP expectations are corrected by efficiency effects, that differ strongly between the experiments. My tests for compatibility of measured q^2 distributions in the $B \to D^{(*)}\tau\nu$ decays with the expectations in SM and 2HDM, that exploit the shifted q^2 dependence by the charged Higgs contribution, do not provide any restriction on the parameter space.

Looking forward, my result will be presented at the winter conferences in early 2015 as the new official Belle result, with a peer-reviewed article following shortly after. The only other measurement in this field, that is expected within the next years, is a measurement by the Belle collaboration of $B \to D^{(*)}\tau\nu$ with hadronic τ decays, which suffers from increased uncertainties in theory and experiment. So my result will influence the focus of theoretical research on 2HDM for the next years, and the door for multi-Higgs-doublet models is still wide open. The next big experimental leap will come with data taking at Belle II. Not only will an analysis similar to mine be performed with a multiple of the current dataset, but the increased statistics will also allow access to other sensitive properties of these decays, like the τ polarization.

I ascend
I grow blind and I blunder
Bitter cold
Does away with the wonder
There is very little air here
Where I stand upon the summit of
All creation
I will close my eyes
And drift away
At last I learned to fly
And found the secret name of longing
Climb, oh we must climb
For we were born for something higher
Than we dream

Glass Hammer - “Into Thin Air”

Bibliography

[1] J. Lees *et al.* (The BaBar Collaboration), "Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons", Phys.Rev. **D88**, 072012 (2013) arXiv:1303.0571 [hep-ex].

[2] A. Matyja *et al.* (Belle Collaboration), "Observation of $B^0 \to D^{*-}\tau^+ nu_\tau$ decay at Belle", Phys.Rev.Lett. **99**, 191807 (2007) arXiv:0706.4429 [hep-ex].

[3] A. Bozek *et al.* (Belle Collaboration), "Observation of $B^+ \to \bar{D}^{*0}\tau^+\nu_\tau$ and Evidence for $B^+ \to \bar{D}^0\tau^+\nu_\tau$ at Belle", Phys.Rev. **D82**, 072005 (2010) arXiv:1005.2302 [hep-ex].

[4] I. Adachi *et al.* (Belle Collaboration), "Measurement of $B \to D^{(*)}\tau\nu$ using full reconstruction tags", (2009) arXiv:0910.4301 [hep-ex].

[5] J. Lees *et al.* (BaBar Collaboration), "Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays", Phys.Rev.Lett. **109**, 101802 (2012) arXiv:1205.5442 [hep-ex].

[6] D. Zander, "Full Reconstruction And Search For Charged Higgs Effects In Semi-Tauonic B Decays", KIT, Diss., 2013, Dr. (KIT, 2013), https://ekp-invenio.physik.uni-karlsruhe.de/record/48271.

[7] M. Tanaka and R. Watanabe, "New physics in the weak interaction of $\overline{B} \to D^{(*)}\tau\overline{\nu}$ ", Phys. Rev. D **87**, 034028 (2013).

[8] S. Fajfer, J. Kamenik, and I. Nišandžić, "$B \to D^*\tau\overline{\nu}_\tau$ sensitivity to new physics", Phys. Rev. D **85**, 094025 (2012).

[9] U. Nierste, S. Trine, and S. Westhoff, "Charged-Higgs-boson effects in the $B \to D\tau\nu_\tau$ differential decay distribution", Phys. Rev. D **78**, 015006 (2008).

[10] G. Branco *et al.*, "Theory and phenomenology of two-Higgs-doublet models", Phys.Rept. **516**, 1–102 (2012) arXiv:1106.0034 [hep-ph].

[11] R. Peccei and H. Quinn, "CP Conservation in the Presence of Pseudoparticles", Phys. Rev. Lett. **38**, 1440–1443 (1977).

[12] M. Tanaka, "Charged higgs effects on exclusive semi-tauonic b decays", English, Zeitschrift für Physik C Particles and Fields **67**, 321–326 (1995).

[13] A. Limosani *et al.* (Belle Collaboration), "Measurement of Inclusive Radiative B-meson Decays with a Photon Energy Threshold of 1.7-GeV", Phys.Rev.Lett. **103**, 241801 (2009) arXiv:0907.1384 [hep-ex].

[14] J. Lees *et al.* (BaBar Collaboration), "Measurement of B($B \to X_s\gamma$), the $B \to X_s\gamma$ photon energy spectrum, and the direct CP asymmetry in $B \to X_{s+d}\gamma$ decays", Phys.Rev. **D86**, 112008 (2012) arXiv:1207.5772 [hep-ex].

[15] A. Abashian *et al.*, "The belle detector", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **479**, 117–232 (2002), Detectors for Asymmetric B-factories.

[16] M. Feindt *et al.*, "A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories", Nucl. Instrum. Methods Phys. Res., Sect. A **654**, 432–440 (2011).

[17] G. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes", Phys. Rev. Lett. **41**, 1581–1585 (1978).

[18] S. Lee *et al.* (Belle Collaboration), "Evidence for $B^0 \to \pi^0\pi^0$", Phys. Rev. Lett. **91**, 261801 (2003).

[19] H. Hayashii, "Measurement of the branching fraction in the $\tau^- \to \pi^-\pi^0\nu_\tau$ decay", Belle Note 622, 2003.

[20] S. Ryu, "Measurement of branching fractions of τ lepton decays containing one or more K^0_S", Belle Note 1166, 2013.

[21] The Belle CKM Group, *$D^{**}\ell\nu$ MC production*, `http://belle.kek.jp/secured/wiki/doku.php?id=physics:ckm:xcmodel`.

[22] D. J. Lange, "The EvtGen particle decay simulation package", Nucl. Instrum. and Meth. **A462**, 152 (2001).

[23] R. Brun *et al.*, *Geant3 user's guide*, tech. rep. (CERN DD/EE/84-1, 1987).

[24] The Belle CKM Group, *New semileptonic B Monte Carlo*, `http://belle.kek.jp/secured/wiki/doku.php?id=physics:ckm:newbsemileptonicmc`.

[25] D. Scora and N. Isgur, "Semileptonic meson decays in the quark model: an update", Phys. Rev. D **52**, 2783–2812 (1995).

[26] A. K. Leibovich *et al.*, "Semileptonic B decays to excited charmed mesons", Phys.Rev. **D57**, 308–330 (1998) arXiv:`hep-ph/9705467 [hep-ph]`.

[27] A. Zupanc *et al.* (Belle Collaboration), "Measurements of branching fractions of leptonic and hadronic D_s^+ meson decays and extraction of the D_s^+ meson decay constant", JHEP **1309**, 139 (2013) arXiv:`1307.6240 [hep-ex]`.

[28] The Belle CKM Group, *Hadronic tag calibration with charm semileptonic decays*, Nov. 2012, `http://belle.kek.jp/secured/wiki/doku.php?id=physics:ckm:htag`.

[29] B. Aubert *et al.* (BABAR Collaboration), "Study of the decay $D_s^+ \to K^+K^-e^+\nu_e$", Phys. Rev. D **78**, 051101 (2008).

[30] C. Oswald *et al.*, "Modelling of semileptonic $B_{(s)}$ decays", Belle Note 1335, 2014.

[31] L. Hinz, "Lepton id efficiency correction and systematic error", Belle Note 954, 2006.

[32] The Belle PID Joint Group, *Lepton id*, Jan. 2013, `http://belle.kek.jp/group/pid_joint/`.

[33] P. Urquijo, "Semi-inclusive semileptonic b decays", Belle Note 1251, 2012.

[34] T. K.Hara, Y.Horii, "Measurement of $B \to \tau\nu$ with hadronic tagging method", Belle Note 1245, 2012.

[35] phi-t GmbH, *The NeuroBayes User's Guide*, Apr. 2010, `http://neurobayes.phi-t.de/index.php/public-information/documentation`.

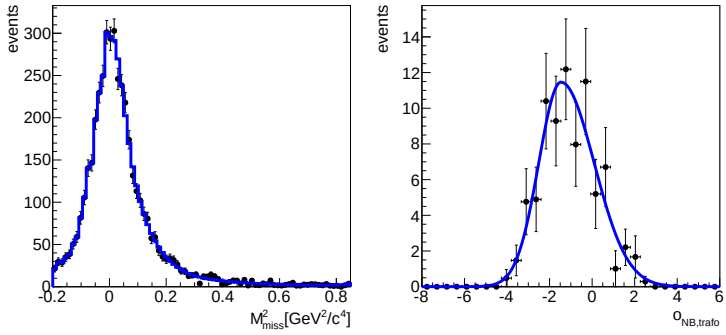
[36] V. Blobel, *Smoothing of Poisson distributed data*, `http://www.desy.de/~blobel/splft.f`.

[37] K. Olive *et al.* (Particle Data Group), "The review of particle physics", Chin. Phys. C **38**, 090001 (2014).

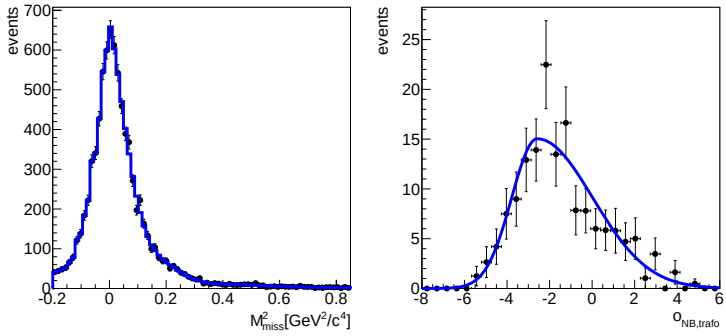
[38] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", The Annals of Mathematical Statistics **9**, 60–62 (1938).

Appendix

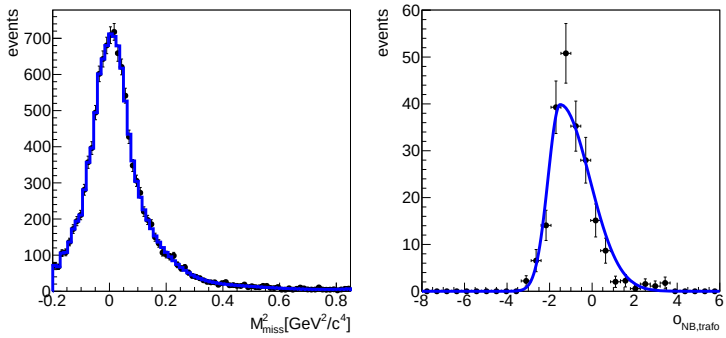
A Probability Density Functions



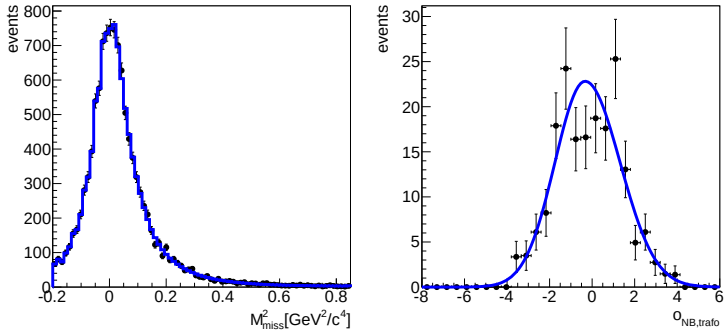
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

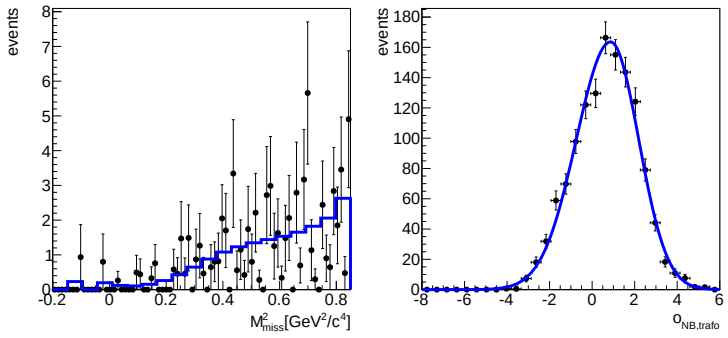


(c) $D^0\ell^-$

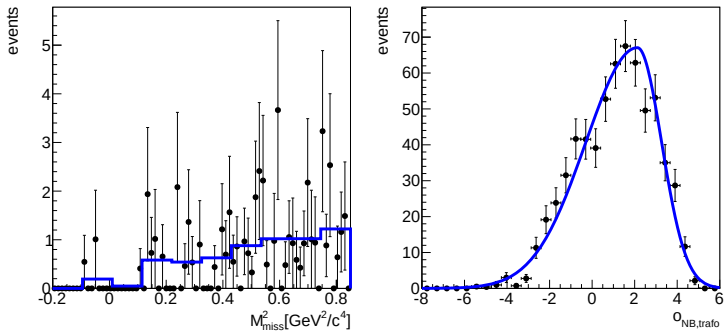


(d) $D^{*0}\ell^-$

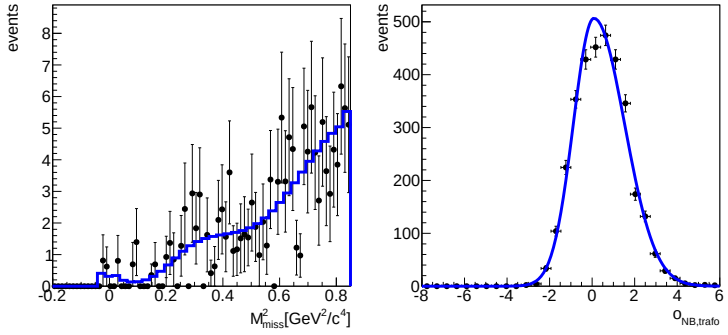
Figure A.1: Lepton signal shapes determined on 5 streams of generic Monte Carlo.



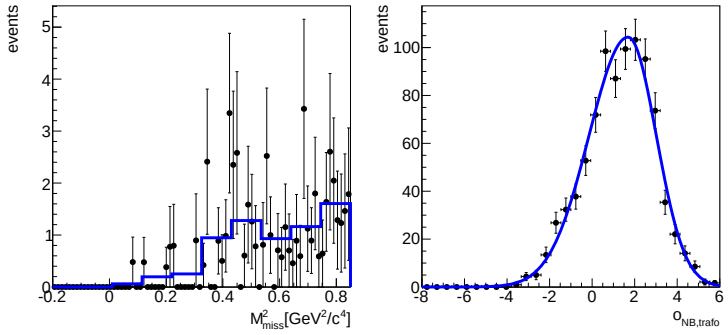
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

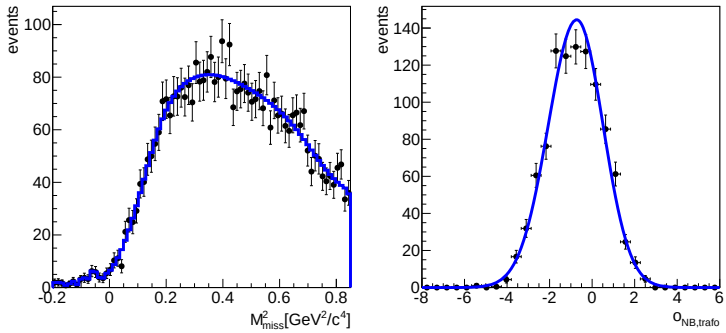


(c) $D^0\ell^-$

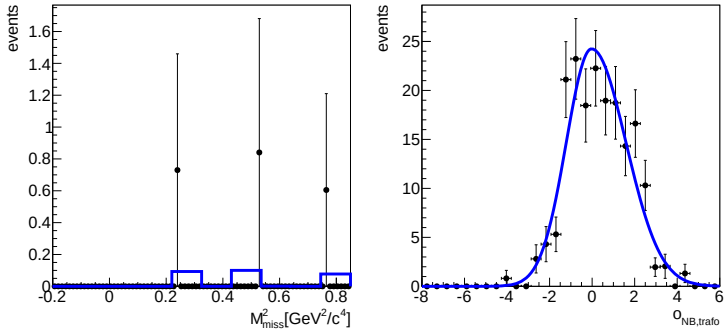


(d) $D^{*0}\ell^-$

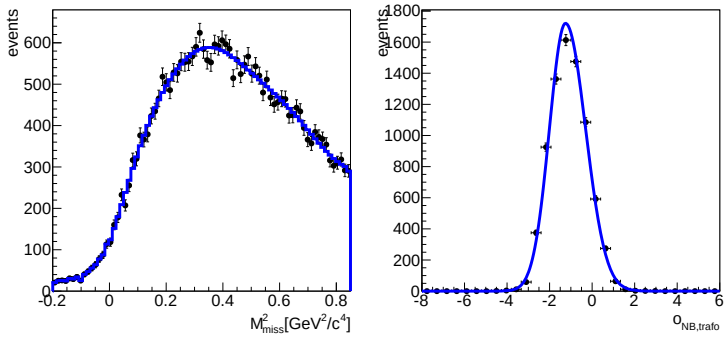
Figure A.2: τ signal shapes determined on signal Monte Carlo.



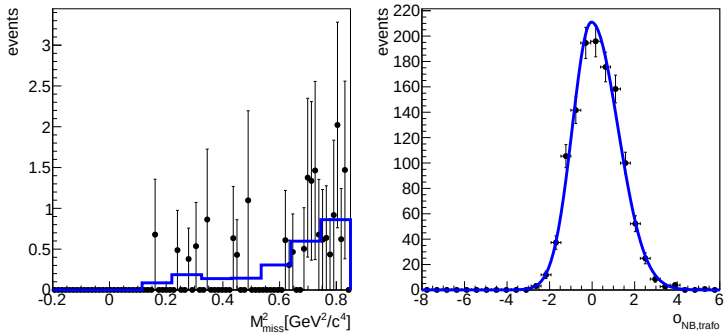
(a) lepton cross-feed in $D^+\ell^-$



(b) τ cross-feed in $D^+\ell^-$

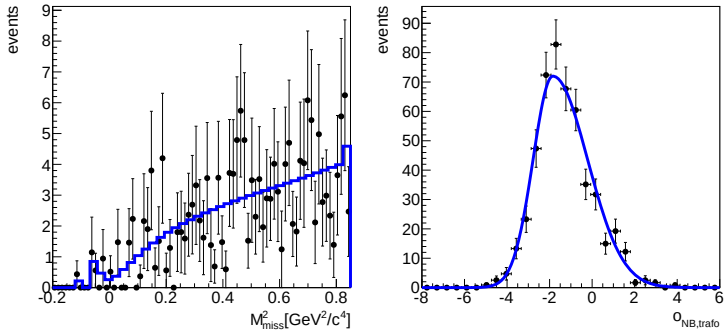


(c) lepton cross-feed in $D^0\ell^-$

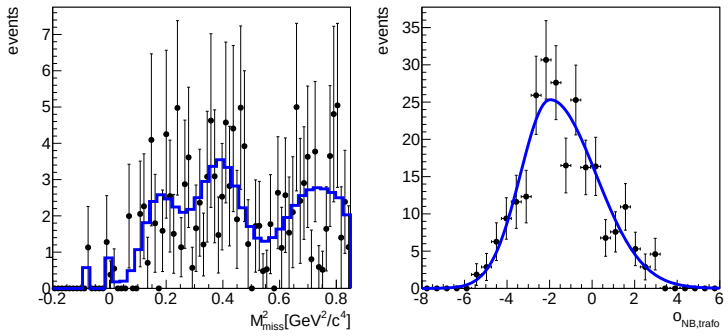


(d) τ cross-feed in $D^0\ell^-$

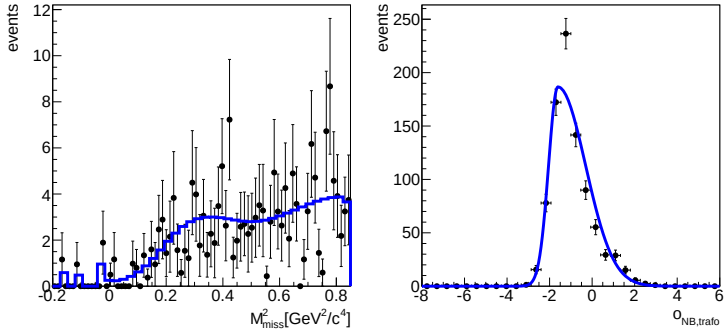
Figure A.3: Lepton and τ cross-feed shapes determined on 5 streams of generic Monte Carlo.



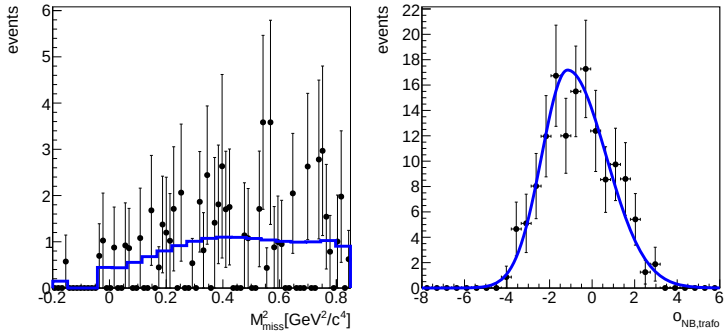
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

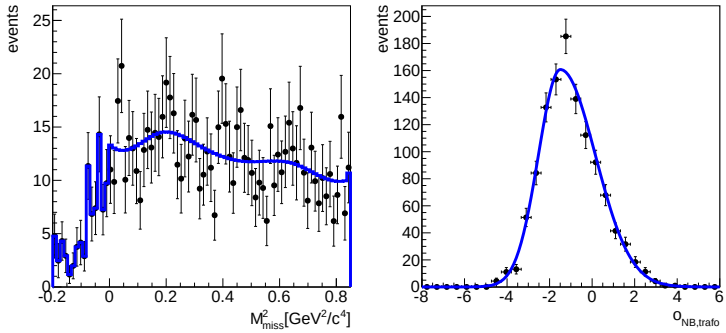


(c) $D^0\ell^-$

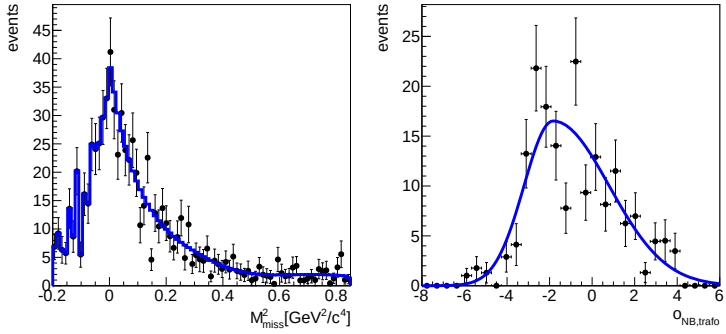


(d) $D^{*0}\ell^-$

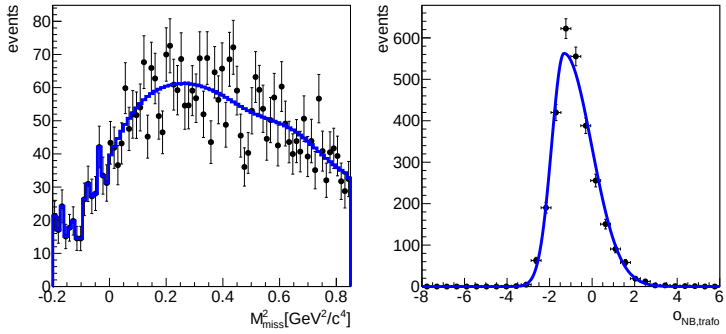
Figure A.4: D^{**} background shapes determined on 5 streams of generic Monte Carlo.



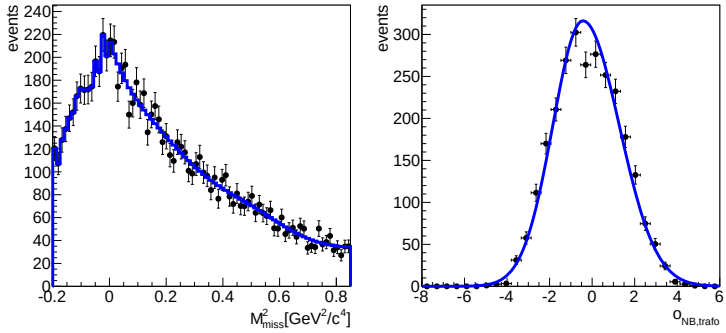
(a) wrong D in $D^+\ell^-$



(b) wrong D^* in $D^{*+}\ell^-$

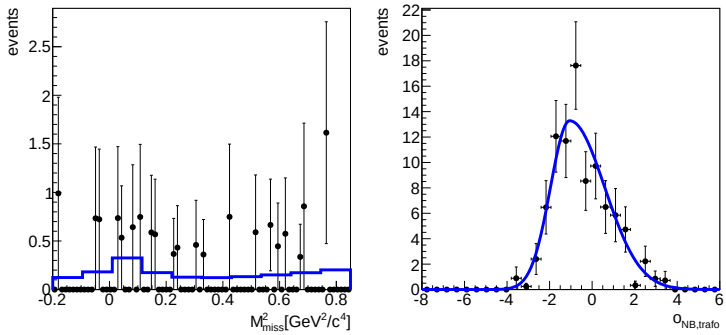


(c) wrong D in $D^0\ell^-$

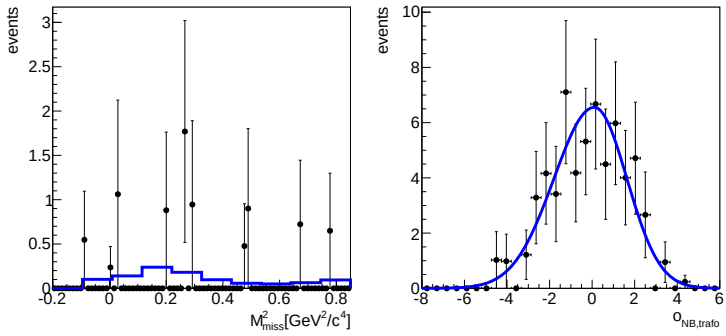


(d) wrong D^* in $D^{*0}\ell^-$

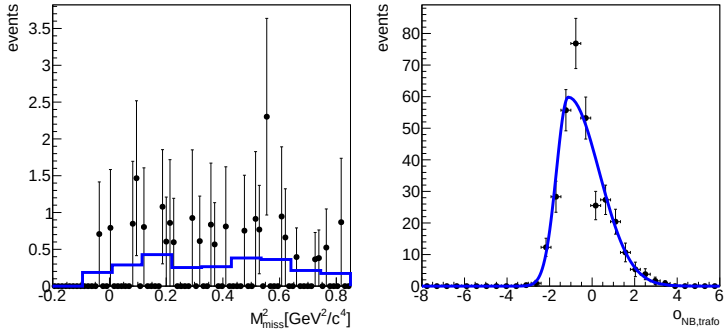
Figure A.5: Wrong $D^{(*)}$ background shapes determined on 5 streams of generic Monte Carlo.



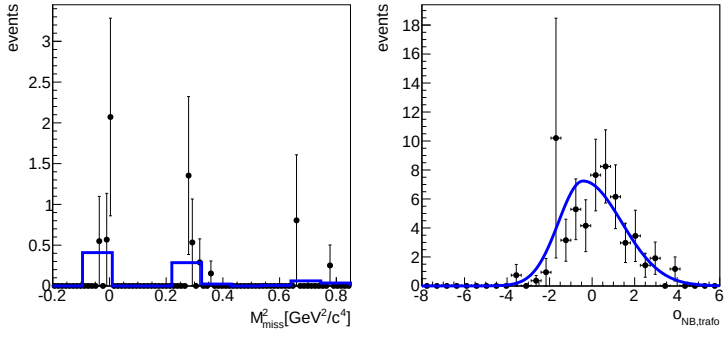
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

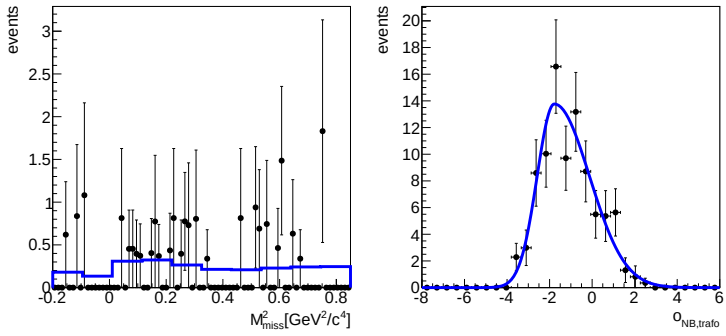


(c) $D^0\ell^-$

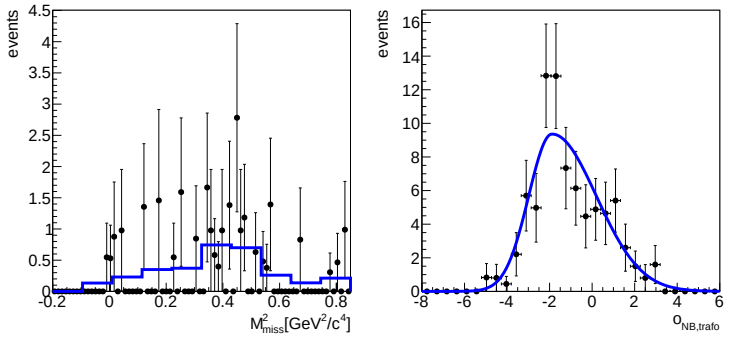


(d) $D^{*0}\ell^-$

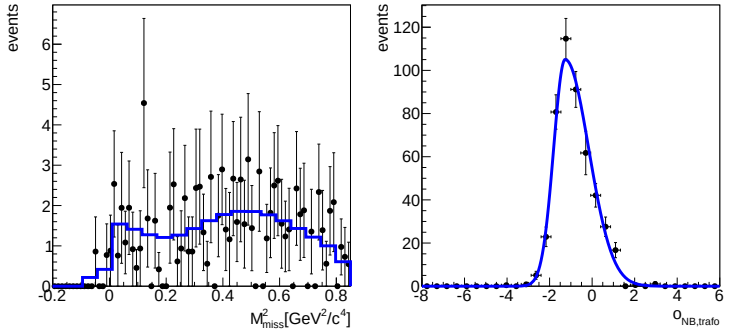
Figure A.6: Wrong lepton background shapes determined on 5 streams of generic Monte Carlo.



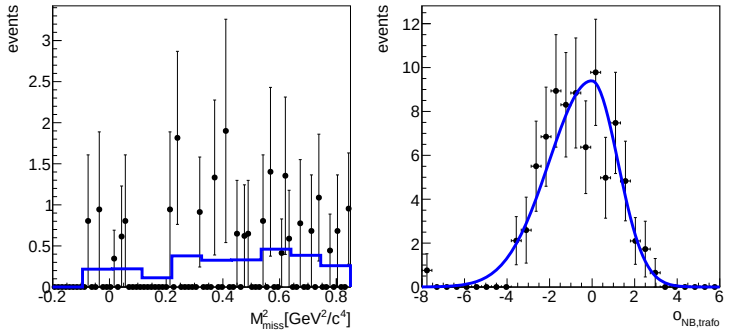
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

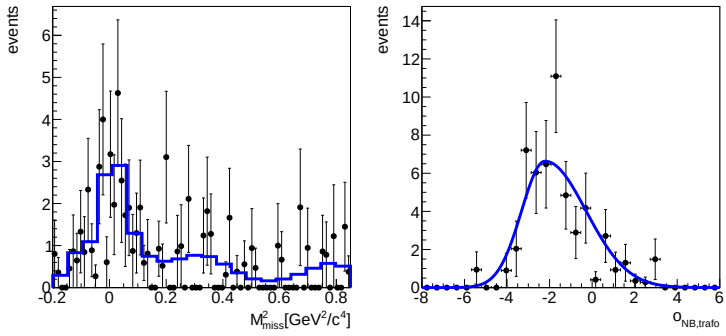


(c) $D^0\ell^-$

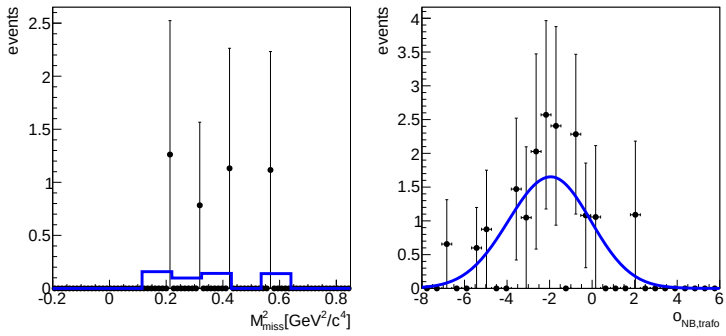


(d) $D^{*0}\ell^-$

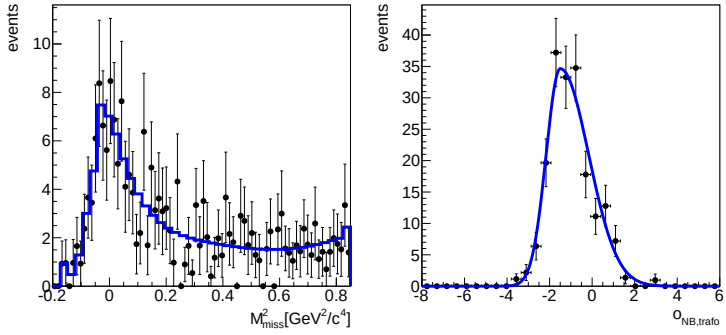
Figure A.7: D_S background shapes determined on 5 streams of generic Monte Carlo.



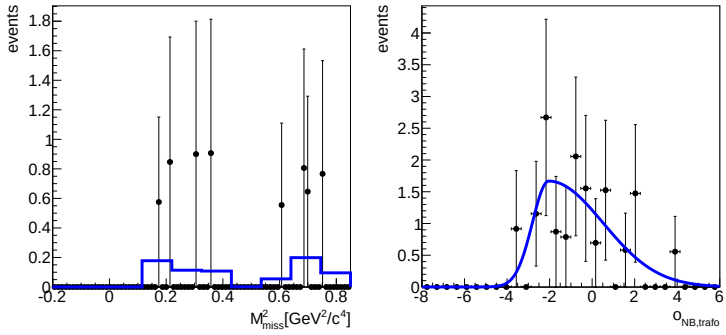
(a) $D^+\ell^-$



(b) $D^{*+}\ell^-$

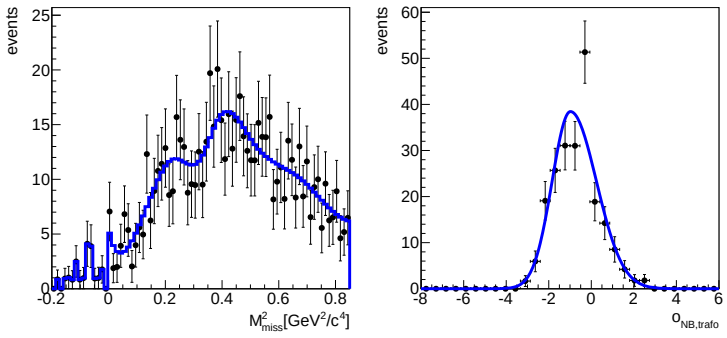


(c) $D^0\ell^-$



(d) $D^{*0}\ell^-$

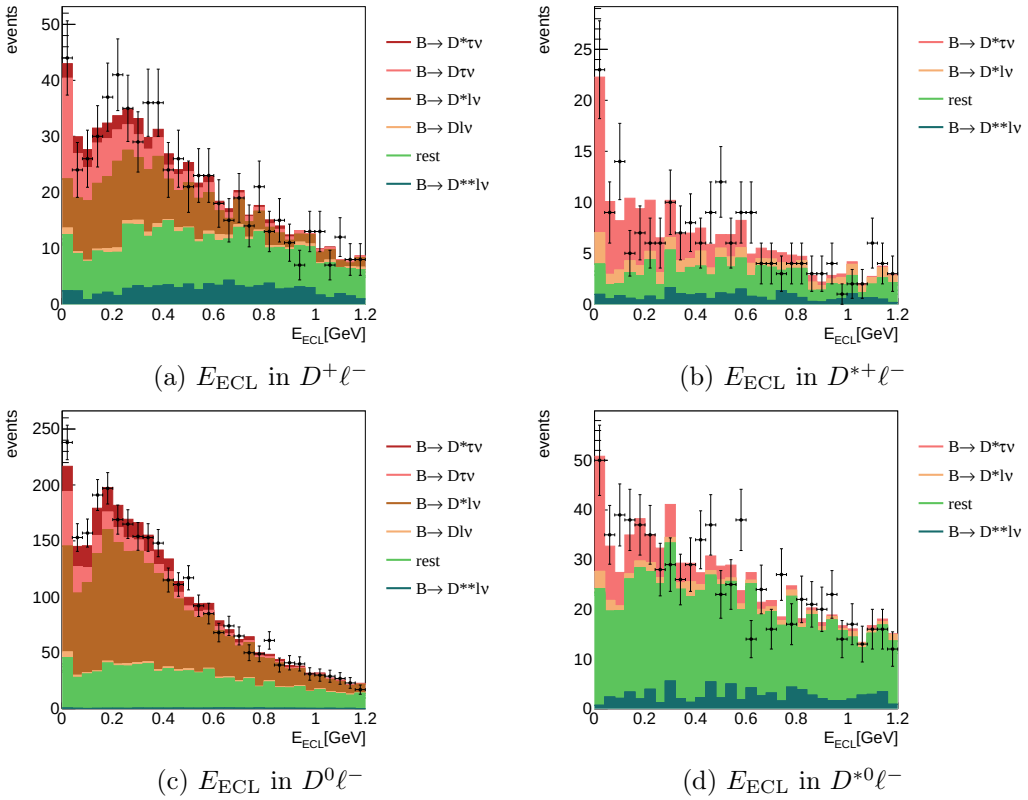
Figure A.8: Remaining background shapes determined on 5 streams of generic Monte Carlo.



(a) $D^0\ell^-$

Figure A.9: Wrong charge cross-feed background shapes determined on 5 streams of generic Monte Carlo.

B Control Plots



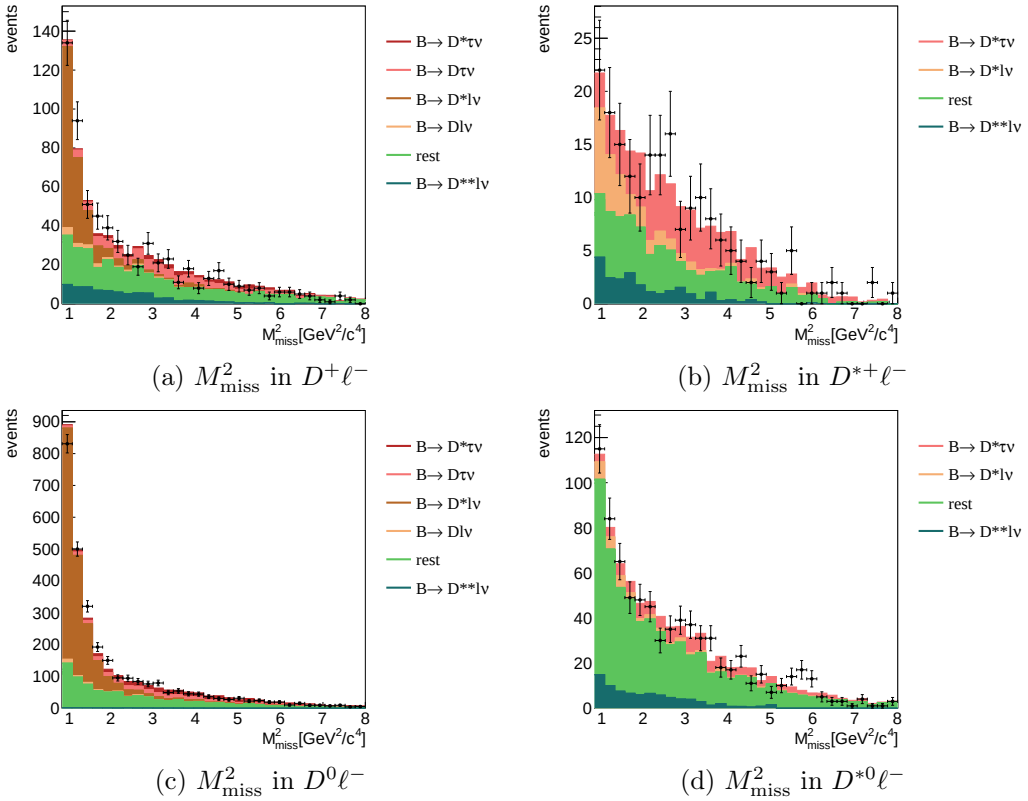
(a) E_{ECL} in $D^{+}\ell^{-}$

(b) E_{ECL} in $D^{*+}\ell^{-}$

(c) E_{ECL} in $D^{0}\ell^{-}$

(d) E_{ECL} in $D^{*0}\ell^{-}$

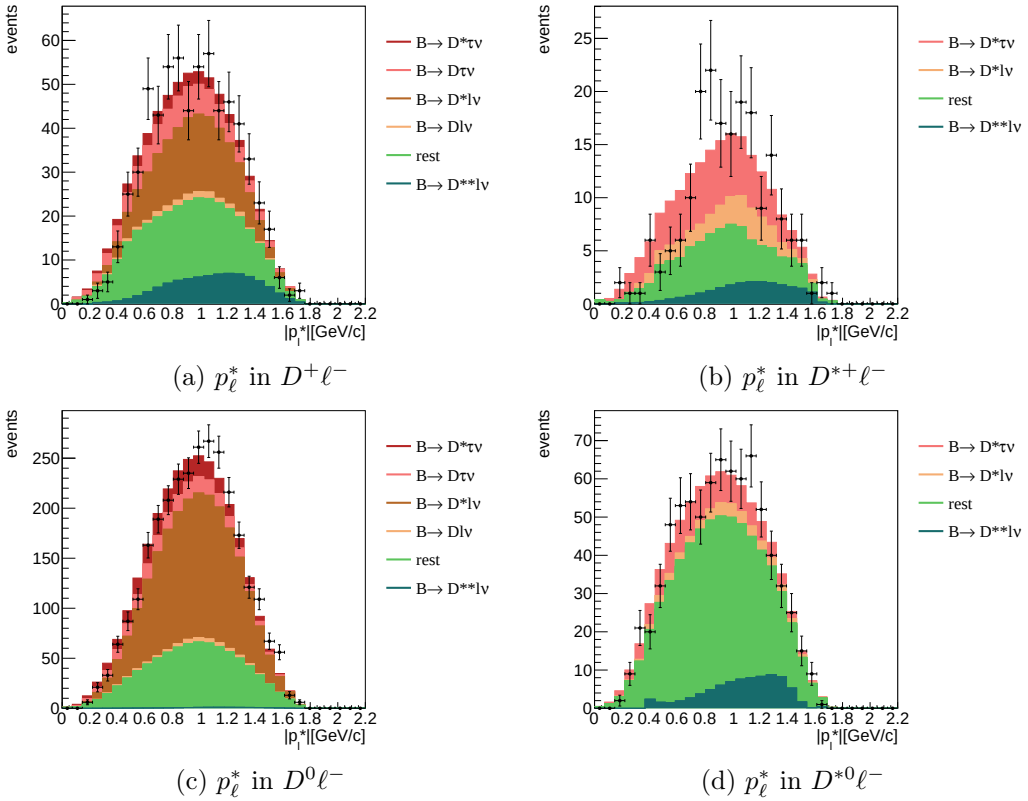
Figure B.10: Fit projections for the E_{ECL} in all four reconstruction samples. The region above $M^{2}_{\mathrm{miss}} = 0.85\,\mathrm{GeV}^{2}\,\mathrm{c}^{-4}$ is used.



(a) M^2_{miss} in $D^+\ell^-$ (b) M^2_{miss} in $D^{*+}\ell^-$

(c) M^2_{miss} in $D^0\ell^-$ (d) M^2_{miss} in $D^{*0}\ell^-$

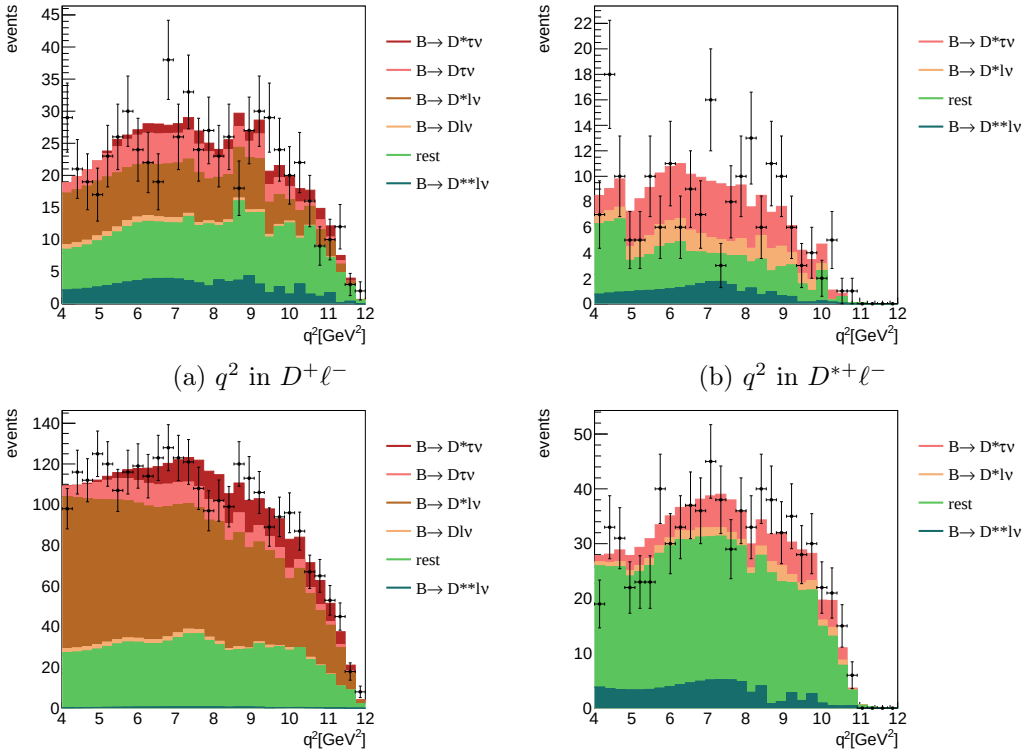
Figure B.11: Fit projections for the squared missing mass in all four reconstruction samples. The region above $M^2_{\text{miss}} = 0.85\,\text{GeV}^2\,\text{c}^{-4}$ is used.



(a) p^*_ℓ in $D^+\ell^-$ (b) p^*_ℓ in $D^{*+}\ell^-$

(c) p^*_ℓ in $D^0\ell^-$ (d) p^*_ℓ in $D^{*0}\ell^-$

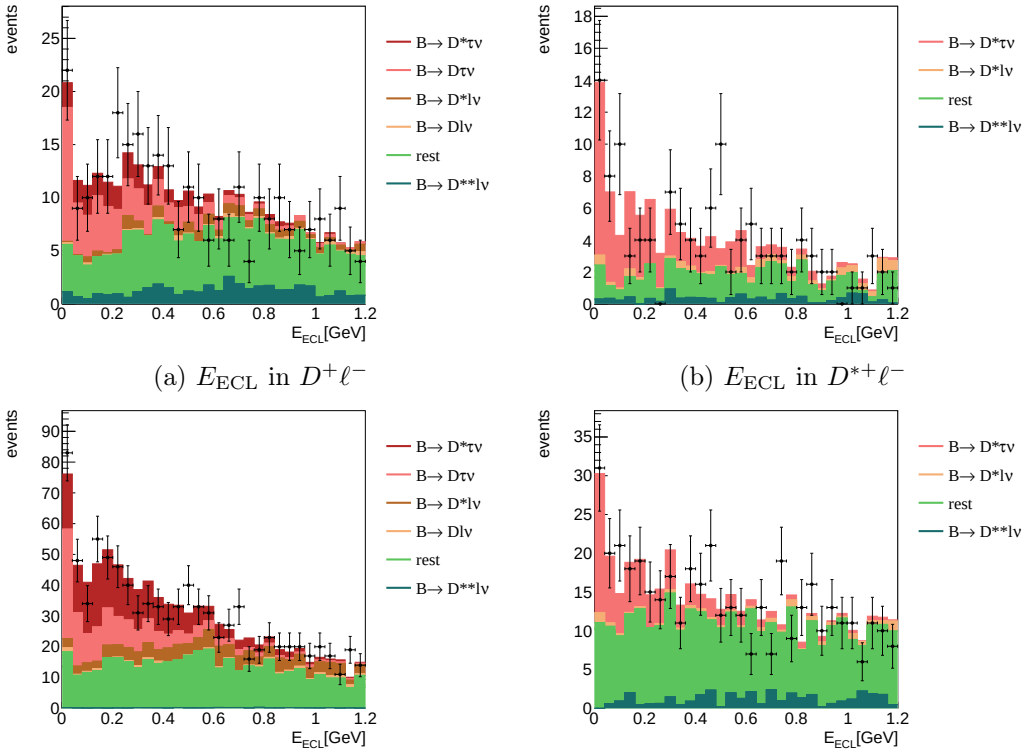
Figure B.12: Fit projections for p^*_ℓ in the B meson rest frame in all four reconstruction samples. The region above $M^2_{\text{miss}} = 0.85\,\text{GeV}^2\,\text{c}^{-4}$ is used.



(a) q^2 in $D^+\ell^-$ (b) q^2 in $D^{*+}\ell^-$

(c) q^2 in $D^0\ell^-$ (d) q^2 in $D^{*0}\ell^-$

Figure B.13: Fit projections for q^2 in all four reconstruction samples. The region above $M^2_{\text{miss}} = 0.85\,\text{GeV}^2\,\text{c}^{-4}$ is used.



(a) E_{ECL} in $D^+\ell^-$ (b) E_{ECL} in $D^{*+}\ell^-$

(c) E_{ECL} in $D^0\ell^-$ (d) E_{ECL} in $D^{*0}\ell^-$

Figure B.14: Fit projections for E_{ECL} in all four reconstruction samples. The region above $M^2_{\text{miss}} = 2.0\,\text{GeV}^2\,\text{c}^{-4}$ is used.

Danksagung

Mein besonderer Dank gilt Prof. Michael Feindt, nicht nur für fachlichen Beistand, sondern auch für das entgegengebrachte Vertrauen, das mir viele Türen geöffnet hat. Prof. Günter Quast danke ich sehr für die herrlich unkomplizierte Übernahme des Korreferats.

Auch die Post-Docs des Instituts – Thomas Kuhr, Anže Zupanc, Martin Heck und Pablo Goldenzweig – haben durch ihre stete Bereitschaft zu fachlichem Rat wichtigen Anteil am Gelingen dieser Doktorarbeit.

Ich habe in den letzten beinahe fünf Jahren am gesamten EKP viele Menschen als sehr angenehme Kollegen kennen gelernt. Der tagtägliche Umgang ist von großartigem Teamgeist geprägt. Für ihre Unterstützung und die gute Zeit danke ich der ganzen B-Gruppe, insbesondere Bastian Kronenbitter und Manuel Heider.

Ein sehr großer Dank gebührt meiner Famile, besonders meinen Eltern Maria und Johann-Georg, sowie meinen Brüdern Thomas, Johannes und Christoph. Auch wenn ihnen bewusst ist, dass mich der eingeschlagene Weg immer seltener in heimische Gefilde treibt, erfuhr ich in jeder Phase meines Studiums uneingeschränktes Wohlwollen, starken Rückhalt und unerschütterliches Vertrauen.