Bachelorarbeit

Lorentz-invariante Observablen zur Topquark-Massenmessung am LHC

Munira Khan
Munira.Khan@studium.uni-hamburg.de
Studiengang Physik (B. Sc.)
Matr.-Nr. 6894864
Fachsemester 6

Erstgutachter: Professor Dr. Peter Schleper
Zweitgutachter: Dr. Hartmut Stadie

Abgabefrist: 02.09.2019
Zusammenfassung


Abstract

The mass of the top quark is an important parameter of Standard Model calculations. In this work, an analytic method to decrease the uncertainty on the top quark mass is investigated. To accomplish this, the template fit method is used on simulated Monte Carlo events from the year 2016, based on data of the CMS detector. Through the addition of a new lorentz-invariant observable, which is the invariant mass of the lepton and the b-jet in the semileptonic top decay, the expected uncertainty of the top mass in this analysis could be decreased to a value of 0.45 GeV.
# Contents

1 Motivation ......................................................... 3

2 Standard Model .................................................... 5
2.1 The Top Quark ...................................................... 6
2.2 Top Quark decay ................................................... 7

3 Experimental setup .................................................. 9
3.1 The Large Hadron Collider ......................................... 9
3.2 The CMS experiment ............................................... 9

4 State of affairs in the top quark mass measurement analysis ................. 12
4.1 Reconstruction of $t \bar{t}$ pairs .................................. 12
4.2 Particle Flow algorithm ............................................ 12
4.3 Kinematic Fit ....................................................... 14

5 Analysis Method .................................................... 15
5.1 Distributions of the fitted top mass $m_t^{fit}$ .......................... 15
5.2 Linearisation of fit parameters ..................................... 16

6 Template Fit analysis of the standard case ................................ 19

7 Invariant mass of the lepton and b-jet $m_{lb}$ .......................... 24
7.1 Introduction to invariant mass of the lepton and b-jets $m_{lb}$ ............ 24
7.2 Analysis using the binned version of $m_{lb}$ .............................. 24
7.2.1 One dimensional Template Fit .................................. 25
7.2.2 Two dimensional Template Fit .................................. 29

8 Decorrelation of $m_{lb}$ and $m_t^{fit}$ .................................. 32
8.1 Dependency of $m_{lb}$ to $m_t^{fit}$ ................................... 32
8.2 Distribution of $m_{lb}^{red}$ ............................................ 33
8.3 Analysis using $m_{lb}^{red}$ .............................................. 34
8.3.1 Distribution plotted as Chebyshev polynomial ....................... 34
8.3.2 Distribution plotted into a binned histogram ....................... 36
8.4 Discussion of results ............................................... 38

9 Summary ........................................................... 41
Chapter 1

Motivation

To gain a full understanding of the universe, starting at its smallest components, is the aim of particle physics. Elementary particles and their reactions are investigated in order to learn more about the properties of the universe. The best description of those particles as of current can be found within the Standard Model of particle physics. It has proven to be a reliable model in the past few decades, as many of its theoretical predictions have been confirmed in various experiments.

Since the Standard Model relies on experimentally determined parameters, it is of interest to measure these with a high accuracy. Those parameters include the coupling constants of the forces between the particles as well as the particles masses. At a mass of about 173 GeV, the top quark is the heaviest particle within the Standard Model. Due to the high mass, it has some unique characteristics, putting a special interest on the study of the top quark. It is the only quark which decays before forming bound states through hadronization. This makes it difficult to precisely measure the mass of the top quark. Therefore, it is of interest to find a reliable analysis method to lower the uncertainty on the top quark mass.

The $t\bar{t}$-process at LHC is used for the analysis in this work. The semileptonic channel is investigated, which means that only $t\bar{t}$-processes where one of the W bosons decays into a lepton and a neutrion while the other W boson decays hadronically. The methods from the existing analysis are used, such as the Particle Flow method, a kinematic fit and the so called Monte Carlo template method. As of now, the most precise single measurement of the top quark has been made in the run 1 of the CMS detector[1] with a result of

$$m_t = 172.44 \pm 0.6 \text{ GeV}. \quad (1.1)$$

An approach to improve the accuracy of this method to determine the top mass through the introduction of a new observable. Through the usage of generated event data samples, a template fit is evaluated. This method splits the uncertainty on the top quark into various "nuisances", which refer to different systematic effects that influence the accuracy of the measurement. By varying these nuisances, their impact on the top mass can be constrained. To further improve the template fit, this work aims to include a new observable into the analysis and discuss its impacts on the top mass uncertainty.
After giving an overview on the Standard Model and the top quark in chapter 2, the CMS detector at LHC is introduced in chapter 3. The reconstruction of events is explained in chapter 4 while the description of the analysis method follows in chapter 5. An introduction into the template fit as well as previous results is given in chapter 6. The results of this work are discussed afterwards. The analysis with the newly added observables are presented in the chapters 7 and 8. Finally, a summary of this work is given in chapter 9.
Chapter 2

Standard Model

To answer questions regarding the structure of the universe, it is of interest to start investigating it at its smallest components. Therefore, the aim of particle physics is to study all elementary particles and the associated interactions between them. At present, the best method to describe these is through the Standard Model (SM). While it does not cover gravitational force, it offers explanations for all particles and interactions that have been observed thus far.

Each particle in the SM has its assigned spin, mass and charge. They are classified into fermions and bosons. Visible matter consists of fermions, which have a spin of $\frac{1}{2}$, while bosons have a spin of 1 and are the carriers of the fundamental interactions between fermions. In addition, there is the Higgs boson with a spin 0. An overview of the fermions and bosons in the SM, including their respective masses, charges and spin, can be seen in Figure 2.1.

Fermions can be divided into leptons and quarks. These are then further divided into three generations. For quarks, they can be differentiated as either up-type (featuring the up, charm and top quarks) and down-type quarks (which include down, strange and bottom quarks). It is to be noted that up-type quarks have an electrical charge of $+\frac{2}{3}e$, while down-type quarks have $-\frac{1}{3}e$, with $e$ being the electric charge of an electron.

The remaining fermions are classified as leptons, with charge $-1e$, and their neutrinos, which have a charge of zero. For each fermion, an anti-particle, with conjugated charge but otherwise the same quantum numbers, can be found. Additionally to the electric charge, all quarks have a colour charge (red, green or blue), while their antiparticles, the anti-quarks, carry an anti-colour charge (anti-red, anti-green and anti-blue).

While fermions are the particles that form matter, bosons are responsible for particle interactions. The photon is associated with the electromagnetic force, while the gluon is the mediator of the strong force, that is observed for all particles carrying a colour charge. Both carry no electrical charge and are massless, but the gluon does carry color charge. This leads to the gluons’ ability to interact with other gluons, while photons are unable to do so.

The remaining bosons are the W- and Z-boson as well as the Higgs boson. All of these three have masses, but only the W-boson has an electric charge of $+e$ or $-e$. Together with the Z-boson, it is the carrier of the weak force that is involved in particle decays. Lastly as the only SM particle with a spin of 0, the Higgs-boson is the latest discovery within the SM. It interacts
Figure 2.1: The particles of the Standard Model. The fermions, which are categorized into leptons and quarks, the gauge bosons and the Higgs boson are all listed with their respective quantum numbers.

with all particles that have a mass.

2.1 The Top Quark

The top quark is the heaviest SM particle. Like the other up-type quarks, it is a fermion with spin of \( \frac{1}{2} \) and has an electric charge of \( \frac{2}{3} \). Together with the bottom quark, it was postulated by Makoto Kobayashi and Toshihide Maskawa in 1973, who predicted a third generation of quarks. This was further confirmed by the experimental discovery of the bottom quark in 1977. However, due to its large mass, the top quark was only discovered later, as its production required a lot of energy. It was discovered at the Tevatron in 1995. Since its discovery, the mass measurement has become more precise. This thesis will discuss an approach to making the top mass measurement even more precise.

Since the aim of particle physics is to completely describe the properties of the smallest particles of the universe, it has to be noted that the SM still relies on experimentally determined parameters. Furthermore, these can be helpful to test the accuracy of the SM. The top quark mass is one of these parameters. Together with the W boson, the top quark mass affects the Higgs mass. This can be used to further investigate the form of the Higgs potential. [2]

Top quark pair production is possible through processes of the strong interaction. It can either be produced through a process with two quarks or two gluons process.
2.2 Top Quark decay

Unlike the other five lighter quarks, the top quark is the only quark that cannot be observed in a bound state. All of the other five quarks are able to form bound states as hadrons. This process is called hadronization. But due its large mass, the top quark has a short life time of about $\tau_{\text{top}} = 5 \cdot 10^{-25}$ s. Hadronization takes place within a time frame of $\tau_{\text{had}} = 3 \cdot 10^{-24}$ s. The fast decay and its inability to hadronize lead to the fact that neither the top quark can be observed directly nor its mass can be measured directly. Instead, it is computed through its decay products.

The top quark decays through weak interactions in two steps. First top quark first decays into a bottom quark and a real W boson. This is possible since the top quark mass is bigger than the sum of the W boson and bottom quark mass ($m_t > m_W + m_b$). The W boson then further splits into either a pair of two quarks or a pair of a lepton and its neutrino, which results in a total of three possible decay channels for the top quarks. They are referred to as dileptonic, all jets and lepton+jets. Each of these have different properties, which are of importance when it comes to the measurement of the top quark.

All three decay channels can be seen in Figure 2.2. The all jets case is also referred to as all hardonic case. There are a total of six jets. These can be separated into two b-tagged jets from both of the bottom quarks and four other jets. There is no missing transversal momentum $p_T$, since there are no neutrinos produced. The correct assignment of jets proves to be difficult when evaluating the decay products. About 44% of the top decay events are part of the all jet category.

In comparison, dileptonic case makes up around 9% of the events. This decay results in two b-jets and two charged leptons. However, since two neutrinos are produced, there are two sources of missing $p_T$, which make the mass measurement difficult.

The best decay to analyse the top quark mass is the lepton+jets case, since it entails one
lepton, two b-jets and two remaining jets. There is only missing $p_T$ from one neutrino, which makes it possible to assign the missing transversal energy to the one neutrino. The produced lepton can either be an electron $e$, a muon $\mu$ or a tau $\tau$. Each of these have a branching fraction of around 15%, leading to a 45% rate for all lepton+jets cases. In this analysis, the semileptonic decay in the muon channel is evaluated.
Chapter 3

Experimental setup

The data used in this thesis is based on simulated proton proton collision event samples. Its experimental setup is based on the Compact Muon Solenoid (CMS) at the Large Hadron Collider (LHC). In this chapter, an overview of LHC will be given and the properties of CMS will be discussed.

3.1 The Large Hadron Collider

Located within the CERN\(^1\) complex, the LHC is a circular particle collider. It is build into the 100 m deep underground tunnel that formerly held the LEP\(^2\) experiment and has a circumference of 26.7 km.

Collisions of protons or Pb-atoms can be observed at four interactions points. At each of these points, detector complexes are located, namely the ATLAS, Alice, LHCb and CMS detectors. A schematic overview of their location as well as of the pre-accelerators of LHC can be seen in Figure 3.1.

The colliding particles are provided from the chain of pre-accelerators, which boost the protons to energies of about 450 GeV. Within the LHC ring, they are then further accelerated up to a final energy of about 6.5 TeV, which leads to the center of mass energy of \(\sqrt{s} = 13\) TeV.

One purpose of the LHC experiments is to search for new physics beyond the standard model and to reach a more precise measurement of the SM parameters. The two detectors CMS and ATLAS focus on these studies, while the other two detectors have other purposes.

3.2 The CMS experiment

The Compact Muon Solenoid (CMS) detector is one of the four detectors at LHC. It is a general-purpose detector, aiming to find phenomena that would go beyond current SM physics as well as improving the precision on current SM parameters, which includes the measurement of the top quark mass that is discussed in this thesis. The detector is 21.6 m long, 15 wide and 15 metres high; it is build around a solenoid magnet which can generate a magnetic field of 3.8 T.\(^3\)

\(^1\)Conseil européen pour la recherche nucléaire, the European Organization for Nuclear Research

\(^2\)Large Electron-Positron Collider

\(^3\)
Its different cylindrical detector layers can be seen in Figure 3.2. The CMS detector determines the type of a particle from the way they interact with each detection layer, thus making it possible to identify the particles and measure their respective momenta and energies. The detection layers are located in the circle around the interaction points.

An accurate momentum measurement is ensured by the tracking system that consists of silicon pixel and silicon strip detectors. They work through the mechanism of a semiconductor, so if a charged particle hits the detector, its trajectory can be measured. Since there is a magnetic field present, the particle tracks of charged particles are bent, which enables momentum measurement.

The tracking system is followed by the calorimeters, where particle energies can be measured. The electromagnetic calorimeter (ECAL) detects the presence of electrons and photons, which are intended to lose all of their energy here through effects like pair production, Compton scattering and the photo effect. Here, pair production plays an important role, as it is the leading effect for the span of high energies.

The ECAL is followed by the hadronic calorimeter (HCAL), which is constructed to surround the collisions. It has the purpose to detect particles that interact through the strong force, such as protons and neutrons.

The HCAL is surrounded by a superconducting solenoid magnet, which generates a magnetic field of around 3.8 T. Finally, muon chambers are located outside of the magnet.
Figure 3.2: Schematic overview of a slice of the CMS detector, showing the way particles interact with the detection layers [3]
Chapter 4

State of affairs in the top quark mass measurement analysis

4.1 Reconstruction of $t\bar{t}$ pairs

To determine the particles that are produced in the reactions within the collisions, the CMS detector records data which are used for the identification and measurement of the particles. The trajectory of the particles, their momenta and their respective energies are of interest. One way to archive this is by applying the *Particle Flow* method. This information makes it possible to identify the decay products and reconstruct the event.

Furthermore, in the reconstruction of a $t\bar{t}$ candidate event in the lepton+jets case, it is convenient to apply a *Kinematic Fit*. The fit ensures that both the top quark and the top anti-quark have the same invariant mass. Additionally, the invariant mass of the W-boson is constrained to its literature value of 80.4 GeV after the fit.

In this chapter, both of these methods will be described in further detail.

4.2 Particle Flow algorithm

The particles involved at a collision at the LHC are measured and tracked within the CMS detector complex. As the detector can be divided into several layers, the Particle Flow algorithm combines the information from the detector layers and translates them into physical reaction. This is valuable due to the different response of the detector parts to each of the SM particles. By putting together the traces of each detector layer, the Particle Flow algorithm can determine the type of particle that passed through the detector. The algorithm reconstructs different particles one after the other, so once a particle has been reconstructed, its track is removed from the remaining unsorted signals.

Charged particles are reconstructed first. The trajectory is measured in the tracking detectors. The signals in the muon detectors is used to reconstruct *muons*. This is done to make sure that the muon signals won’t interfere with the signals of the charged hadrons. Furthermore,
muons can be classified as either **standalone muons**, which are identified by using only information of the muon chambers, or as **global muons** that take information from the muon detector and the other trackers into consideration. [2]

After the muon signal, the **electrons** are determined. They lose almost all of their energy in the electromagnetic calorimeter and can be tracked there. Once the tracks of the electrons and muons have been determined, **charged hadrons** can be found through the signals in the hadron calorimeter. The trajectories from the tracking detector are matched in terms of compatible energies to reconstruct the four-momenta of the charged hadron.

Only after all charged particles have been assigned, the momenta for photons and neutral hadrons are found by looking at the signals within both the electromagnetic and the hadron calorimeter. **Photons** only leave signals in the electromagnetic calorimeter, while the **neutral hadrons** leave a signal in both of the calorimeters.

Figure 4.1: Distributions of the top mass and the reconstructed W boson mass. Distributions of the reconstructed W boson mass $m_{W}^{\text{reco}}$ before the kinematic fit (top left) and the respective top quark mass $m_{t}^{\text{reco}}$ (top right) in comparison to the post-fit distributions of $m_{W}^{\text{reco}}$ (bottom left) and $m_{t}^{\text{fit}}$ (bottom right).[1]
4.3 Kinematic Fit

In the measurement of the top quark, it has proven to be of advantage to apply a kinematic fit. It improves the purity of the reconstruction candidates and improves their resolution, so the fit ensures that it matches the hypothesis. In the top quark mass measurement, the fit works through two assumptions:

- the W boson mass is 80.4 GeV, so the combination of two non-b-tagged jets should also be equal to this value
- in the $t\overline{t}$ process, the anti-top quark and the top quark are of equal mass

By applying the fit to the reconstructed values, a fitted top mass $m_t^{\text{fit}}$ is obtained. A cut on the fit probability at $P(x^2) = 0.2$ is applied, which leads to a significantly improved W boson and top mass signal. The improvement through the kinematic fit can be seen in Figure 4.1. Using all decay channels, the most precise top quark measurement has been done in [1], with a total top quark mass result of

$$m_t = 172.44 \pm 0.6 \text{ GeV} \quad (4.1)$$

when using all decay channels. Regarding only the lepton+jets channel, this analysis leads to a top mass value at

$$m_t = 172.14 \pm 0.78 \text{ GeV.} \quad (4.2)$$

The template method used for this is explained in further detail in the following chapter 5.
Chapter 5

Analysis Method

The aim of this work is to improve the measurement strategy for the determination of the top quark mass. By adding a new observable to the analysis, it is investigated whether the accuracy of determining the top quark mass can be positively influenced. The approach of determining such an observable as well as its variations will be discussed.

In this chapter, the first steps of the analysis method will be presented. The distribution of the fitted top mass $m_{t}^{\text{fit}}$ will be described with a fit function. The dependence of fit parameters on the simulated top mass will be shown. Additionally, the template fit method will be introduced and its previous results will be discussed in the following chapter 6.

In chapter 7, the invariant mass of the lepton and b-jets of the top quark decay is plotted. It is referred to as $m_{lb}$ and will be a key observable in this analysis. Finally, $m_{lb}$ will be further modified into a reduced version and the final results will be shown in chapter 8.

5.1 Distributions of the fitted top mass $m_{t}^{\text{fit}}$

Datasets generated with the Monte Carlo method are used. A total of seven datasets with different generated top masses $m_{t}$, using the muon channel, are analysed. The value for $m_{t}$ varies between $166.5$ GeV and $178.5$ GeV. The histograms for the resulting $m_{t}^{\text{fit}}$ are plotted and a fit function has been applied. The used fit parameters can be determined and plotted in dependency of $m_{t}$.

In the first attempt, a modified Gauss function was used by overlapping a Voigtian function with a Crystal Ball function. The results for such a function will be discussed first and the linearisation of its fit parameters will be discussed afterwards.

In Figure 5.1, the distributions for each of the seven generated top masses are pictured. The peak of each function is located around the mass value of the respective data set, which is due to the fact that most events match the correct top mass. The mass distribution also needs to have a certain width according to the relativistic Breit-Wigner distribution. The probability of the production of a particle at its resonant mass peak has the shape of a Breit-Wigner distribution where the width is dependant on the decay width $\Gamma$ of the particle. However the experimental resolution leads to a considerably broader distribution.
Figure 5.1: Distribution of the fitted top mass $m_t^{\text{fit}}$. Each of the seven distribution corresponds to a different generated top mass.

It can also be seen that the function has a long tail in the direction of higher energies until values above 300 GeV. One way to explain this tail is by tracing it back to possible mistakes during the assignments of jets in the datasets.

### 5.2 Linearisation of fit parameters

For each of the seven different generated top masses that can be seen in Figure 5.1, a fit function is plotted. In this section, the first version of the fit, using a Voigtian function that is overlapped with a Crystal Ball distribution, will be discussed. The normalised Voigt function $V(x, \sigma, \gamma)$ is a convolution of the two functions $G(x, \mu, \sigma^2)$ and $L(x, \gamma)$. The Voigt profile is defined as

$$V(x, \sigma, \gamma) = \int_{-\infty}^{\infty} G(x', \sigma)L(x - x', \gamma)dx'$$  \hspace{1cm} (5.1)

with the Gaussian profile $G(x, \mu, \sigma^2)$ and the Lorentzian profile $L(x, \gamma)$

$$G(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$  \hspace{1cm} (5.2)

$$L(x, \gamma) = \frac{\gamma}{\pi(x^2 + \gamma^2)}$$  \hspace{1cm} (5.3)

where $x$ is the value for $m_t^{\text{fit}}$ and $\sigma$ is the width of the profile, while $\mu$ describes its mean.

The Crystal Ball function is described through

$$f(x, \alpha, n, \mu_{CB}, \sigma_{CB}) = N \cdot \begin{cases} 
  \exp\left(-\frac{(x - \mu_{CB})^2}{2\sigma_{CB}^2}\right), & \text{for } \frac{x - \mu_{CB}}{\sigma_{CB}} > -\alpha \\
  A \cdot \left(1 - \frac{x - \mu_{CB}}{\sigma_{CB}}\right)^{-n} & \text{for } \frac{x - \mu_{CB}}{\sigma_{CB}} \leq -\alpha
\end{cases}$$  \hspace{1cm} (5.4a,b)
where \( \alpha, n, \mu_{CB} \) and \( \sigma_{CB} \) are the parameters that are fitted to the data, while \( A \) and \( B \) are constants and \( N \) is a normalization factor.

The fit parameters can be computed for each of the seven datasets, which leads to seven different values for the respective parameter. Especially the widths \( \sigma \) and mean \( \mu \) of the Voigtian as well as the \( \sigma_{CB} \) and \( \mu_{CB} \) of the Crystal Ball are of interest, as they have a noticeable correlation with the top quark mass. In Figure 5.2 Voigtian mean and width are plotted in dependence of \( m_t^{\text{fit}} \). A linear relation can be seen. The same has been done for the mean and widths for the Crystal Ball function, which can be seen in Figure 5.3. This linear relation leads to the conclusion that the fit parameters are correlated to the top quark mass.

Since they are correlated, it is possible to compute the top quark mass and an error on its value through the fit parameters. This is done in two steps. First, the correct value for each fit parameter, according to the bins in Figure 5.2 and 5.3 of a histogram is determined, using the information of the generated top mass \( m_t \).
<table>
<thead>
<tr>
<th>Top Mass $m_t$ [GeV]</th>
<th>Simulated Top Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>166.502 ± 0.0457</td>
<td>166.5 GeV</td>
</tr>
<tr>
<td>169.481 ± 0.0478</td>
<td>169.5 GeV</td>
</tr>
<tr>
<td>171.436 ± 0.0497</td>
<td>171.5 GeV</td>
</tr>
<tr>
<td>172.501 ± 0.0498</td>
<td>172.5 GeV</td>
</tr>
<tr>
<td>173.591 ± 0.0576</td>
<td>173.5 GeV</td>
</tr>
<tr>
<td>175.559 ± 0.0521</td>
<td>175.5 GeV</td>
</tr>
<tr>
<td>178.428 ± 0.0543</td>
<td>178.5 GeV</td>
</tr>
</tbody>
</table>

Table 5.1: Top masses computed from fit parameters, as well as the simulated top mass of the respective dataset

The function for $\mu$ is defined as following

$$\mu = \mu_0 (1 + S_{mt} \cdot m_t)(1 + S_{\theta_i} \cdot \theta_i)$$

(5.5)

where $\mu_0$ is the start value for $\mu$, $S_{mt}$ refers the to the slope of the $m_t^\text{fit}$ distribution from the Monte Carlo data sets and $S_{\theta_i}$ refers to the slope of the linear distribution of $\mu$, which can be seen in Figure 5.2. In analogue manner, linear function are defined for the other fit parameters. These linear function can then be used for a likelihood fit.

Once the values for all functions of the fit parameters have been found, the top mass $m_t$ is set as unknown. Then $m_t$ is determined, just using the information of the fit parameters functions.

The resulting top masses and their uncertainties can be seen in Table 5.1. The resulting top masses $m_t$ match the generated top masses with a satisfying accuracy, as the errors on the top masses are all in the region of 0.05 GeV. This shows that this approach can determine the top mass without a bias.
Chapter 6

Template Fit analysis of the standard case

Systematic effects and uncertainties in the top quark mass measurement are considered as nuisance parameters. Each nuisance parameter impacts the value of the top quark mass. The nuisance parameters can be grouped into Jet measurement uncertainties, specific b-jet measurement uncertainties, theory uncertainties related to matrix element calculation, parton shower and underlying events, b-quark fragmentations and others. A detailed description is given in [5] and [6].

In this chapter, the template fit method with the two observables \( m_t^{\text{fit}} \) and \( m_W^{\text{reco}} \) as well as a third observable \( R_b \) will be discussed.

The template fit method finds its origin in the precise measurements of the top quark mass at CMS that have used the ideogram method. The ideogram method adds parameters, such as the global energy scale factor JSF, to the computed likelihood. Adding nuisance parameters to the likelihood can therefore lead to learning more about the systematics in the evaluation of the top mass measurement.

In the standard case in one dimension (1D), an error of 0.89 GeV is determined by using only \( m_t^{\text{fit}} \) as variable and varying the nuisances. The jet energy corrections ("JEC") play a dominant role among the uncertainties with "JECUncorrelated" as biggest uncertainty. Additionally to this, two other nuisances, that are related to the jet energy corrections, can be seen among the ten biggest uncertainty of this fit. The full overview on the impact of each nuisance in 1D can be seen in Figure 6.1.

In the two dimensional (2D) Template Fit, the fitted top quark mass \( m_t^{\text{fit}} \) and reconstructed W boson mass \( m_W^{\text{reco}} \) are the standard combination. The observable \( R_b \) is the sum of the transversal momentum of the b-jets divided by the sum of the transversal momentum of the quarks. It is defined as

\[
R_b = \frac{p_T^{b_1} + p_T^{b_2}}{p_T^q + p_T^{q_2}}
\]

(6.1)

The analysis with these three observables will be referred to as the standard case in three dimensions (3D), because it is the combination of variables that have been used in previous works,
Figure 6.1: Nuisance analysis for pseudo-experiments at $m_t = 172.5$ GeV using only $m_t^{\text{fit}}$ as observable (1 dimensional case). On the left side, the fitted average nuisance parameters and their uncertainties can be seen. The nuisance parameters are scaled by their a priori uncertainty. A fit result of $\pm 1$ means that the fit to the data did not lead to a further constraint on the nuisance parameter. Any value below 1 indicates the corresponding uncertainty was reduced from the data itself. On the right side, their impact on the measured top quark mass $m_t$ can be seen.
Figure 6.2: Nuisance analysis for pseudo-experiments at $m_t = 172.5$ GeV using both $m_t^{\text{fit}}$ and $m_t^{\text{reco}}$ as observables (2 dimensional case). For details, check the caption of Figure 6.1.

such as [9] and [6]. Since it is intended to find a new observable for the nuisance analysis, the observables of the standard case will be replaced with $m_{lb}$, which will be introduced in the following chapter 7.

In the 1D fit, seen in 6.1, the jet energy corrections had the largest impact on the top quark mass uncertainty. However, this effect can be decreased in the 2D template fit. By including $m_t^{\text{reco}}$ into the template fit, the resulting error is decreased. Compared to the one dimensional case, the error on the top mass is lowered from $0.89$ GeV to $0.52$ GeV in the 2D case. This is expected, as $m_t^{\text{reco}}$ does not correlate with the top mass and can therefore further constrain the jet energy uncertainty. Especially the jet energy scale could be significantly constrained. The uncorrelated jet energy correction uncertainty drops from being the biggest uncertainty with $0.61$ GeV down to the sixth place with an error of only $0.15$ GeV.

The "FlavourPureBottom" effects are the source of the biggest error in the 2D template fit. This nuisance parameter refers to the uncertainty due to the bottom flavour dependent jet energy scale. Constraining this uncertainty through the addition of a third observable to the template fit would be desired.
Another large uncertainty in this template fit are the statistical uncertainties called "MC-Stat", with an uncertainty of 0.29 GeV in the 2D template fit. "MCStat" refers to a statistical uncertainty that regards multiple minor nuisance parameters. For each of these uncertainties, a linear function similar is generated, in a similar manner as the linearisation of fit parameters was done in the end of the previous chapter. The nuisance "MCStat" was high among the uncertainties in the 1D as well.

Compared to that, the statistical uncertainty "DataStat" plays a small role in this template fit with a value at 0.07 GeV. The name "DataStat" refers to the process where an uncertainty is generated for the case where all nuisance parameters are set as constant and are not regarded. This is to estimate the errors on the analyses for the cases when data from experiments instead of generated data is used. Other sources of high uncertainties are the nuisances related to colour reconnection effects, such as "CR QCDBased" and "CR GluonMove".

In the current 3D standard case, the observable $R_b$ is added into the 3D analysis as third variable with the idea to reduce the b-jet energy scale uncertainty. It brings the value of the total top mass uncertainty down to 0.50 GeV. It was intended for $R_b$ to decrease the uncertainty on the "FlavourPureBottom" nuisance. Its effect is to reduce this uncertainty from 0.3 GeV to 0.28 GeV, so that biggest insecurity now are the statistic uncertainties on the MC simulation ("MCStat"). However, the "FlavourPureBottom" remains as the second largest source of uncertainty.
Figure 6.3: Nuisance analysis for pseudo-experiments at $m_t = 172.5$ GeV using both $m_t^{\text{fit}}$, $m_t^{\text{reco}}$ and $R_b$ as observables (3 dimensional case). For details, check the caption of Figure 6.1.
Chapter 7

Invariant mass of the lepton and b-jet $m_{lb}$

7.1 Introduction to invariant mass of the lepton and b-jets $m_{lb}$

An approach to improve the accuracy of the top mass measurement is the search for a new variable that can be used in the template fit. The invariant mass of the lepton and b-quark $m_{lb}$ is used whereas the b-jet according to the lowest $\chi^2$ of the kinematic fit is used to reflect the mass of the $\mu$ and b-jet.

The invariant mass of the lepton and b-quark $m_{lb}$ from the reconstructed decay products is defined by:

$$m_{lb} = \sqrt{(E_{E_{\text{reco}}} + E_{j_{\text{reco}}})^2 - (\vec{p}_{E_{\text{reco}}} + \vec{p}_{j_{\text{reco}}})^2} \quad (7.1)$$

The distribution for $m_{lb}$ can be seen in Figure 7.1. Its shape reflects the top quark mass, as it has a mass peak and a tail in the direction of higher energies. The peak is located at around 100 GeV. It matches the expectation that the peak location should be located at around $\frac{2}{3}$ of the top quark mass. Since the invariant mass of the b-jets and lepton are two of the three particles the top quark decays into, it is expected for $m_{lb}$ to make up typically $\frac{2}{3}$ of the top quark mass.

7.2 Analysis using the binned version of $m_{lb}$

It has been observed that the best results are accomplished by reducing information regarding the shape of the $m_{lb}$ distribution. Therefore, instead of viewing $m_{lb}$ as continuous function and fitting it with an analytic function, its information is compressed into a histogram of seven bins. This will further be referred to as $m_{lb}^{\text{binned}}$.

The histogram for $m_{lb}^{\text{binned}}$ with a much lower binning can be seen in Figure 7.1 alongside the detailed distribution of $m_{lb}$. The distribution for each of the seven generated top mass are plotted. The bin at the peak as well as the very first and the last three bins at the tail all have a similar height, while the second and fourth bin, where the function grows and falls, are of different height. Moreover, the histogram of lower generated top mass $m_t$ are always higher than the bigger ones in the first two bins, but in the fourth bin, they are smaller than ones for higher generated top masses $m_t$. This leads to the conclusion that even compressed into few
Figure 7.1: Distribution the invariant mass of the lepton and b-jets, $m_{lb}$, and its binned counterpart, $m_{lb}^{\text{binned}}$, for different generated top masses. The colours indicate the generated top mass in each of the seven used datasets. The distribution at the value of 172.5 GeV is marked in yellow, while other colours indicate lower or higher masses, as indicated by the legend. The number indicated the value of GeV subtracted or added to 172.5 GeV (e.g. the black distribution belongs to a generated top mass of 166.5 GeV).

bins, $m_{lb}$ reflects properties of the top mass. The observable $m_{lb}^{\text{binned}}$ can now be added to the nuisance analysis.

7.2.1 One dimensional Template Fit

As a first step, the one dimensional analysis which relies only on one variable is regarded. In the standard case, this variable for the 1D analysis is the fitted top quark mass $m_t^{\text{fit}}$. An impact graph of this can be seen in Figure 6.1.

The one dimensional template fit is now run with $m_{lb}$. This is done with the expectation that it will most likely not offer better results than $m_t^{\text{fit}}$. However, it is of interest to still investigate the properties of $m_{lb}$ and to find out how well it compares to $m_t^{\text{fit}}$.

Its impact graph, seen in Figure 7.2, shows that the error on the top mass increases in comparison to the case that uses $m_t^{\text{fit}}$. The total error on the top mass is 0.13 GeV larger than in the standard 1D, as it increases from 0.89 GeV to 1.02 GeV. The comparison of both errors on the top mass as well as the influence of each nuisance on this can furthermore be seen in Figure 7.3. While the total error on the top quark mass increased in the fit that only uses $m_{lb}^{\text{binned}}$, it can be noted that the impact of some nuisances can be reduced by using $m_{lb}^{\text{binned}}$. For example, the nuisances related to colour reconnection, ”CR GluonMove” and ”CR QCDBased” have a smaller impact on the top mass error.

One of the biggest uncertainties for the 1D analysis with $m_{lb}$ comes from the nuisance ”DataStat”, which is also referred to as $\Delta m_t^{\text{stat}}$. It describes the statistical error of the used datasets.
Figure 7.2: 1D Template Fit for pseudo-experiments at $m_t = 172.5$ GeV using only $n_{lb}$ binned as observable (1 dimensional case). On the left side, the fitted average nuisance parameters and their uncertainties can be seen. On the right side, their impact on the measured top quark mass $m_t$ can be seen. (For details, check the caption of Figure 6.1.) The total error on the top mass is 1.02 GeV.
and can be interpreted

\[ \Delta m_t^{\text{stat}} = \frac{\sigma}{\sqrt{N_{\text{events}}}} \]  

with \( \sigma \) the effective resolution on \( m_t \) per event and \( N_{\text{events}} \) as the number of events.

In the 1D standard case using \( m_t^{\text{fit}} \), the statistical uncertainty has a value of only 0.07 GeV, while the 1D case using \( m_{lb} \) has a statistical uncertainty of 0.18 GeV. This leads to values of \( \sigma = 23 \) GeV for the fit with \( m_t^{\text{fit}} \), while the value for the fit using \( m_{lb} \) is at \( \sigma = 60 \) GeV. This uncertainty cannot be constrained by the fit, since it depends on the used datasets. In the case of \( m_{lb} \), it would have been possible to reach an uncertainty close to the case of \( m_t^{\text{fit}} \), if the \( m_{lb} \) template fit used an input of three to four times as many events than the one with \( m_t^{\text{fit}} \) in the 1D template fit.

Figure 7.3: Impact of each nuisance on the top quark mass \( m_t \) when only \( m_t^{\text{fit}} \) as variable is used (right), compared to the case when only \( m_{lb}^{\text{binned}} \) is used (left). The resulting total error is pictured in black, while the coloured lines correspond to the different nuisance parameters, as listed on top.
Figure 7.4: Uncertainty on the nuisance of the used fit for the cases discussed in Figure 7.3.
### 7.2.2 Two dimensional Template Fit

#### Figure 7.5: 2D Nuisance analysis for pseudo-experiments at \( m_t = 172.5 \) GeV using \( m_{lb}^{\text{binned}} \) and \( m_{\text{W reco}}^{\text{binned}} \) as observables. For details, check the caption of Figure 6.1. Compared to the standard 2D case seen in 6.2, \( m_{lb}^{\text{binned}} \) is used instead of \( m_{\text{W reco}}^{\text{binned}} \).

As seen in the previous chapter 6, most of the improvement on the error of the top quark mass is done by adding a second variable to \( m_t^{\text{fit}} \). By adding \( m_{\text{W reco}}^{\text{binned}} \) to the analysis, its accuracy was improved by a value of 0.37 GeV (see Figure 6.2 for more details), lowering the uncertainty on the top mass from 0.89 GeV to 0.52 GeV.

Instead of \( m_t^{\text{fit}} \), in this section \( m_{lb}^{\text{binned}} \) will be added as first variable, while the second observable is \( m_{\text{W reco}}^{\text{binned}} \), which is the same as in the standard case. The case using \( m_{lb}^{\text{binned}} \) is compared to the 2D standard case.

Impact graphs are visible in Figure 7.5. Compared to the standard case in 1D, the error on the top quark mass went down from 0.89 GeV to 0.58 GeV, which is an error that is only 0.06 GeV larger than in the standard 2D case. The comparison of the uncertainties on the top quark mass can be seen in Figure 7.6.

Similar to the previously discussed 1D case that used \( m_{lb} \), the uncertainty on the statistical

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \Delta \theta )</th>
<th>( \Delta \theta )</th>
<th>( \Delta \theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{lb}^{\text{binned}} )</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( m_{\text{W reco}}^{\text{binned}} )</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( m_t^{\text{fit}} )</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( m_{t_{\text{binned}}} )</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Note: The table values are placeholders for demonstration purposes.
error has a value of 0.18 GeV, which is larger than in the standard case. Therefore, the resulting uncertainty on the top quark mass got larger. The statistical uncertainty of the Monte Carlo method is also bigger than in the standard case, rising from 0.29 GeV to a value of 0.36 GeV. This puts the systematic uncertainties on top of the list as the biggest source of uncertainty on the top quark mass.

Also important is the "FlavourPureBottom", which could be decreased from 0.28 GeV in the standard case to a value of 0.27 GeV. This systematic effect refers to the uncertainty due to the b-flavour dependant jet energy scale.

Since \( m_{lb} \) is the invariant sum of the lepton and b-jet, it was of special interest to see if \( m_{lb} \) as new observable could further constrain this nuisance. The hypothesis was that due to the properties of \( m_{lb} \) and its direct dependency on the b-jets, it could be possible to further constrain the "FlavourPureBottom". An improvement of 0.01 GeV is not a significant change. Further details on this will be given in the following chapter, where the 3D template fit is discussed.

For \( m_{lb}^{\text{binned}} \) into the nuisance analysis with two observable resulted in a larger error than the standard method. But it is possible to further investigate whether this value can be improved by lowering the statistical error through the usage of more simulated events. It can be speculated that it might prove to be of advantage to use \( m_{lb}^{\text{binned}} \) for cases where the standard 2D case does not deliver satisfying results.
Figure 7.7: Uncertainty on the nuisance of the used fit for the cases discussed in Figure 7.6.
Chapter 8

Decorrelation of $m_{lb}$ and $m_t^{\text{fit}}$

The results of the analysis can further be improved by taking the correlation of $m_{lb}$ with $m_t^{\text{fit}}$ into consideration. In this chapter, their correlation is discussed and a reduced variant of $m_{lb}$ will be introduced.

8.1 Dependency of $m_{lb}$ to $m_t^{\text{fit}}$

The covariance between $m_t^{\text{fit}}$ and $m_{lb}$ can be obtained from Figure 8.1. It is plotted to check whether $m_{lb}$ qualifies as a suitable uncorrelated variable in the analysis. If it is directly correlated to $m_t^{\text{fit}}$, there would be no use in its addition to the analysis, as it wouldn’t bring any new information into the analysis.

![Figure 8.1: Correlation between the top quark mass $m_t^{\text{fit}}$ as well as the mass of the lepton and b-jets, $m_{lb}$. Marked in red is the profile of computed mean values of the correlation in the direction of $m_{lb}$.](image)

The covariance matrix is given through

$$Cov(m_t^{\text{fit}}, m_{lb}) = \begin{pmatrix} 1490.56 & 969.243 \\ 969.243 & 1737.840 \end{pmatrix}$$
with the correlation factor \( C(m_l^{\text{fit}}, m_{lb}) = 0.602 \). This indicates a relation between \( m_{lb} \) and \( m_l^{\text{fit}} \). This is not desired, since the information of \( m_l^{\text{fit}} \) can be used directly in the template fit. The aim for analysis is to maximise the new information on the top mass that is gained from the addition of \( m_{lb} \). To accomplish this, one possible option is to divide \( m_{lb} \) by \( m_l^{\text{fit}} \) and it will be checked whether more precise results can be reached. The ratio of \( m_l^{\text{fit}} \) and \( m_{lb} \) will be introduced as \( m_{lb}^{\text{red}} \).

The correlation of \( m_{lb}^{\text{red}} \) with \( m_l^{\text{fit}} \) is shown in Figure 8.2. With a correlation factor at \( C(m_l^{\text{fit}}, m_{lb}^{\text{red}}) = -0.018 \), it is almost uncorrelated to the top mass.

![Figure 8.2: Correlation between the top quark mass \( m_l^{\text{fit}} \) and the reduced mass of the lepton and b-jets, \( m_{lb}^{\text{red}} \). Marked in red is the profile of computed mean values of the correlation in the direction \( m_{lb}^{\text{red}} \).](image)

**8.2 Distribution of \( m_{lb}^{\text{red}} \)**

Due to the direct relation of \( m_{lb} \) with the top quark mass \( m_t \), a reduced version of it is used in this analysis. It can be computed through the following equation:

\[
m_{lb}^{\text{red}} = \frac{m_{lb}}{m_l^{\text{fit}}} = \sqrt{\frac{(E_{b}^{\text{reco}} + E_{l}^{\text{reco}})^2 - (p_{b}^{\text{reco}} + p_{l}^{\text{reco}})^2}{m_l^{\text{fit}}}} \tag{8.1}
\]

Its distribution can be seen in Figure 8.3. It ranged from 0 to 1 and is fitted with Chebyshev polynomials. The distribution is not symmetrical. At the value of 0.2 it starts to steadily rise, then forms a broad peak before it decreases. The broad peak is located between values of 0.5 and 0.7 for \( m_{lb}^{\text{red}} \).

This can be explained by the properties of \( m_{lb} \). Since it is the sum of the lepton and the b-jets in the top quark decay, it makes up roughly \( \frac{2}{3} \) of the actual top quark mass.

Furthermore the beginning of the distribution at 0.2 can be explained by looking at the original distribution of \( m_{lb} \) in Figure 7.1. It does not show a remarkable amount of events with \( m_{lb} \) with masses close to zero. The distribution starts to show events at around 20 to 30 GeV for \( m_{lb} \), which is roughly a fifth of the expected top quark mass. Hence, the distribution of \( m_{lb}^{\text{red}} \) is plotted from values starting at 0.2 to 1.
Figure 8.3: Distribution the reduced invariant of the lepton and b-jets, $m_{lb}^{\text{red}}$ and its binned counterpart. For details on the legend, check 7.1. For the binned counterpart, it is visible that visible that the distribution that belong to the smaller top masses (black, green, red) have higher bins at values below a point of 0.5 than the distributions that belong to higher generated top masses (light blue, blue, pink). At values bigger than 0.7, the distributions of lower generated top masses have the lower bins too.

A binned counterpart has also been defined for $m_{lb}^{\text{red, binned}}$. While the dependency on each simulated top mass is not very clear in the $m_{lb}^{\text{red}}$ that was plotted through Chebyshev polynomials, in the binned counterpart is it visible that the distribution with the smaller top masses have higher bins below 0.5, but then decrease faster at $m_{lb}^{\text{red}}$ values above 0.7. The dependence of $m_{lb}^{\text{red}}$ on $m_t$ is however much reduced compared to Figure 7.1.

8.3 Analysis using $m_{lb}^{\text{red}}$

A three dimensional template fit is applied. The first two observables are $m_t^{\text{fit}}$ and $m_W^{\text{reco}}$, according to the standard case. The two versions of $m_{lb}^{\text{red}}$ are used as third observable. The results for the $m_{lb}^{\text{red}}$ distribution fitted with Chebyshev polynomials is compared to the result of its binned counterpart. Furthermore, an overview of all the tested template fits as well as the uncertainties on each nuisance can be seen in Figure 8.7.

8.3.1 Distribution plotted as Chebyshev polynomial

A top mass uncertainty of 0.48 GeV can be accomplished through a 3D template fit using $m_t^{\text{fit}}$ and $m_W^{\text{reco}}$ and $m_{lb}^{\text{red}}$. Compared to the standard case, this result shows an improvement of 0.02 GeV. The impact of each nuisance factor on the top quark mass and the uncertainty of each nuisance can be seen in Figure 8.7.

The four dominant sources of uncertainties for this analysis are the same as in the 3D
Figure 8.4: Impact in 3D Template Fit using $m_{lb}^{red}$ fitted with Chebyshev Polynomials as third variable. The resulting total top quark mass error is 0.48 GeV. For details on the graphic, check the caption of Figure 6.1.
standard template fit. The biggest source of uncertainty are the systematic uncertainties on the Monte Carlo sample. In the analysis using $m_{lb}^{\text{red}}$, this uncertainty is 0.03 GeV higher than in the standard case. This uncertainty is not varied during the template fit and stays as constant, so the resulting top mass error can be even further lowered by e.g. using a larger amount of samples.

"FlavourPureBottom" has the same uncertainty as in the standard case using $R_b$, while the impact of the uncorrelated jet energy scale error got larger by 0.02 GeV. This systematic has already been improved by 0.01 GeV through 2D template fits, which were discussed in the previous chapter. Even in this 3D analysis using $m_{lb}$ combined with $m_t^{\text{fit}}$, the uncertainty on "FlavourPureBottom" is at a value of 0.27 GeV. Whether this is the first hint that this effect can be traced back to a reliable way to constrain the "FlavourPureBottom" nuisance cannot be said for certain just with the results from this work.

The third biggest source of uncertainty on the top quark mass is the QCD based colour reconnection effect. This is the case in this analysis and in the version using $R_b$. But through the addition of $m_{lb}$, this source of uncertainty could be reduced from 0.22 GeV to 0.19 GeV.

### 8.3.2 Distribution plotted into a binned histogram

The 3D template fit method using $m_{lb}^{\text{red}, \text{binned}}$ alongside the observables $m_t^{\text{fit}}$ and $m_t^{\text{reco}}$ leads to a top mass error of 0.45 GeV. This is a 0.03 GeV improvement compared to the template fit with $m_{lb}^{\text{red}}$ as Chebyshev and it is 0.05 GeV more accurate than the analysis that used $R_b$. The impact of each nuisance factor on the top quark mass and the uncertainty of each nuisance can be seen in Figure 8.5.

The uncertainty on the Monte Carlo systematic and "FlavourPureBottom" are similar to the standard case, with the systematic error being 0.01 GeV larger and the effect on "Flavour-PureBottom" getting 0.01 GeV smaller.

The dominant improvement compared to the previous cases can be seen in the colour reconnection effects to QCD based and the uncorrelated jet energy scales. While the previous version of $m_{lb}^{\text{red}}$ lowered the accuracy of the uncorrelated jet correlation, using a binned variant of it can further improve the accuracy on this uncertainty. Its impact on the top mass is 0.15 GeV, while the analysis using a Chebyshev fitted version of $m_{lb}^{\text{red}}$ and the standard case had values of 0.17 and 0.15 GeV, respectively.

The accuracy of this template fit is validated in Figure 8.6. The bias on the measurement of the top quark mass can be seen on the left side of the graphic. The generated top mass is plotted versus resulting top quark mass of the fit. For both the values are shifted by 172.5 GeV for better visibility. In case the fit results in correct values regardless of the generated top mass, a linear relation is expected. As the result of the fit matches this expectation, it can be concluded that the results are unbiased and have no dependency on the generated top mass.

Additionally, the residuum of the resulting top mass can be evaluated, which can be seen in the right graphic. This evaluation is done for the central top mass in the right graph, which corresponds to value around 172.5 GeV. In case the uncertainties have been correctly estimated,
Figure 8.5: Impact in 3D Template Fit using $m_{lb}$ fitted with Chebyshev as third variable, resulting in a top quark mass error of 0.45 GeV. For details, check the caption of Figure 6.1.
the graph is expected to take the shape of a normal distribution, which by definition has a width of 1. This is the case for the plotted data that was created according to the model with the three observables $m_{i}^{\text{fit}}$, $m_{i}^{\text{reco}}$ and $m_{i}^{\text{red}}$, proving this method offers correct results.

8.4 Discussion of results

Three template fits using distributions of $m_{lb}$ have been investigated. The best result, with a total top quark mass error of 0.45 GeV has been archived by using a 3D template fit with $m_{i}^{\text{fit}}$, $m_{i}^{\text{reco}}$ and $m_{i}^{\text{red}}$.

A comparison of all the tested template fits in 2D and 3D can be seen in Figure 8.7. The standard cases in each dimension (1D, 2D and 3D $R_{b}$) are compared to the three dimensional analyses using $m_{lb}$. Through the transition to a 3D nuisance analysis, the uncertainty was further decreased. It has been found that by introducing the observable $m_{lb}^{\text{red}}$, which is found by dividing $m_{lb}$ with the fitted top mass $m_{i}^{\text{fit}}$, the uncertainty on the top mass can be decreased. This can be explained through the fact that the distribution of $m_{lb}$ is correlated to the one of $m_{i}^{\text{fit}}$. Therefore, the introduction of $m_{lb}^{\text{red}}$ helps the purpose of including new information into the analysis. Though it has not been tested if there are more effective ways to decorrelate $m_{lb}$ and $m_{i}^{\text{fit}}$, this method has already delivered satisfying results.

It is moreover interesting to note that the binned variant of $m_{lb}^{\text{red}}$ has a lower error than its continuous counterpart. This might be a consequence of the fact that by taking a distribution with less bins, the dependency of the top mass gets stronger, while this dependency is less striking in the distribution that was fitted with a Chebyshev function.
Figure 8.7: The impact of each nuisance on the mass of the top quark $m_t$ for each approach with a newly added observable. The standard case for each dimension is marked as 1D, 2D and 3D $R_b$, while 2D $m_{lb}$ refers to the two dimensional template fit using $m_{lb}^{\text{fit}}$. Furthermore, the results of the template fits using the reduced version of $m_{lb}$ (3D $m_{lb}$ as Cheb) parametrized via Chebyshev polynomials is compared to the case when $m_{lb}$ is regarded as a binned distribution (3D $m_{lb}$ in Bins)
Figure 8.8: Uncertainty on the nuisance of the used fit for the scenarios discussed in Figure 8.7.
Chapter 9

Summary

As the heaviest discovered Standard Model particle, the top quark plays an important role for current research within particle physics. Its mass is a vital parameter for calculations within the Standard Model as well as searches for new physics.

Figure 9.1: The error of the determined top masses for the toy experiments in all tested template fits. Each bar represents the resulting error of the top mass. The one dimensional standard case can be seen in black. The 2D bar show the interval of the error when the information for \( m_{W}^{\text{reco}} \) (red) is added into the analysis, while the 2D case combining \( m_{W}^{\text{reco}} \) with \( m_{t\bar{b}} \) (dark blue). Finally, the 3D bars show the impact that the addition of \( m_{t\bar{b}}^{\text{Cheb}} \) (pink) as well as \( m_{t\bar{b}}^{\text{binned}} \) (light blue) bring to the analysis. Their values can be compared with the standard analysis in 3D using \( R_{b} \) (yellow)

Due to the large mass of the top quark and its colour charge, there are various factors contributing to the high uncertainty on the top quark compared to the masses of e.g. leptons or gauge bosons. The aim of this work was to obtain a higher accuracy on the measured top mass.
To accomplish this, template fits have been used to constrain the uncertainties on the top mass.

Datasets generated through Monte Carlo methods were used to improve the analysis method, aiming to reduce its uncertainties. These simulated distributions of the top quark mass are fitted with Chebyshev polynomials. Using the fit parameters of the polynomials makes it possible to estimate the top quark mass.

Template fits with nuisances can be used to determine the uncertainty on the top mass. Through the usage of several one dimensional distributions of observables related to the top quark mass, the template fit analyses the effects contributing to the error on the top mass. In this work, the invariant mass of the lepton and bottom quark from the top quark decay $m_{lb}$ is added into the template fit.

Three variants of the template fit using distributions of $m_{lb}$ have been investigated. In Figure 9.1, a total overview of the top mass error in each version of the template fit and its uncertainty can be seen. The best results have been archived by using 3D template fits with $m^\text{fit}_t$, $m^\text{reco}_W$ and $m^\text{red}_{lb}$, where $m^\text{red}_{lb}$ refers to the quotient $\frac{m_{lb}}{m^\text{fit}_t}$. This has been done with two variants of $m^\text{red}_{lb}$, one which was fitted with Chebyshev polynomials and another that was fitted with binned histograms.

While previous three dimensional template fit analyses have led to a resulting top quark mass uncertainty of 0.5 GeV, the addition of the observable $m_{lb}$ has further constrained this error. Through a decorrelation of $m_{lb}$ from $m^\text{fit}_t$, a new observable $m^\text{red}_{lb}$ has been introduced. This has lead to a total uncertainty of the top mass of 0.48 GeV for the variant where $m^\text{red}_{lb}$ is plotted through Chebyshev polynomials. Previous analyses using $m^\text{fit}_t$, $m^\text{reco}_W$ and a combination of transversal impulses called $R_b$ have lead to a resulting expected top mass error of 0.50 GeV.

In addition, this result can even be further approved by splitting the distribution of $m^\text{red}_{lb}$ into a binned histogram. Using a combination of $m^\text{red, binned}_{lb}$, $m^\text{fit}_t$ and $m^\text{reco}_W$ leads to a top mass uncertainty of 0.45 GeV.
Eidesstattliche Versicherung

Ich bin mit einer Einstellung in den Bestand der Bibliothek des Fachbereiches einverstanden.

Hamburg, den 30. August 2019
## List of Figures

2.1 Standard Model .................................................. 6
2.2 Feynman diagramm of the top quark decay ...................... 7
3.1 Overview on LHC .................................................. 10
3.2 Overview on the CMS detector .................................. 11
4.1 Distribution of $m_t$ and $m_W$ prior and after the kinematic fit ........................................ 13
5.1 Distribution of $m_t^{\text{fit}}$ ...................................... 16
5.2 Linearisation for the fit parameters of the Voigtian ............. 17
5.3 Linearisation for the fit parameters of the Crystal Ball .......... 17
6.1 Standard case Impact in 1D Template Fit ......................... 20
6.2 Standard case Impacts in 2D Template Fit ....................... 21
6.3 Standard case Impacts in 3D Template Fit ....................... 23
7.1 Distribution of $m_{lb}$ and $m_{lb}^{\text{binned}}$ ..................... 25
7.2 Impact 1D using $m_{lb}^{\text{binned}}$ ................................ 26
7.3 Impact Comparison in the 1D Template Fit analysing $m_{t}^{\text{fit}}$ and $m_{lb}^{\text{binned}}$ ...................... 27
7.4 Nuisance Comparison in 1D using $m_{t}^{\text{fit}}$ and $m_{lb}^{\text{binned}}$ ................................. 28
7.5 Impact 2D using $m_{lb}^{\text{binned}}$ .................................. 29
7.6 Impact Comparison in the 2D Template Fit using $m_{t}^{\text{fit}}$ and $m_{lb}^{\text{binned}}$ ...................... 30
7.7 Nuisance Comparison in 2D using $m_{t}^{\text{fit}}$ and $m_{lb}^{\text{binned}}$ ...................... 31
8.1 Correlation of $m_{t}^{\text{fit}}$ and $m_{lb}$ .................................. 32
8.2 Correlation of $m_{lb}^{\text{red}}$ and $m_{lb}^{\text{binned}}$ ...................... 33
8.3 Distribution of the reduced $m_{lb}$ ................................. 34
8.4 Impact in 3D Template Fit using $m_{lb}^{\text{red}}$ fitted with Chebyshev Polynomials as third variable .......... 35
8.5 Impact in 3D Template Fit using $m_{lb}^{\text{red,binned}}$ third variable ................. 37
8.6 Bias on the measurement and the residuum of the resulting top mass ........................................ 38
8.7 Impact Comparison .................................................. 39
8.8 Nuisance Comparison ................................................. 40
9.1 Summary of resulting top mass uncertainties in each Template Fit ........................................ 41
Bibliography


[5] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using events in the dilepton final state in pp collisions at $\sqrt{s} = 13$ TeV," CMS PAS TOP-17-001


[9] Christoph Garbers, Johannes Lange, Peter Schleper and Hartmut Stadie. "A profiled likelihood approach to measure the top quark mass." CMS Analysis note, CMS AN-18-261, 2019