

Bachelor's Thesis

Monte-Carlo-Simulation und Analyse des $t\bar{t}H$ Prozesses mit dem ATLAS Experiment bei $\sqrt{s} = 13$ TeV

Monte Carlo simulation and analysis of the $t\bar{t}H$ process with the ATLAS experiment at $\sqrt{s} = 13$ TeV

prepared by

Christina Reißel

from Pirmasens

at the 2nd Institut of Physics

Thesis number: II.Physik-UniGö-BSc-2016/11

Thesis period: 4th April 2016 until 27th July 2016

First referee: Prof. Dr. Arnulf Quadt

Second referee: Prof. Dr. Stanley Lai

Abstract

Seit der Entdeckung des Higgs-Bosons im Jahre 2012 stehen insbesondere die Vermessung seiner Eigenschaften sowie Kopplungen an andere Teilchen des Standardmodells im Vordergrund. Zur Untersuchung der Kopplung zwischen Fermionen und Higgs-Boson hat dabei die Produktion eines Higgs-Bosons in Verbindung mit einem Top Quark Paar eine besondere Bedeutung, ebenso wie der Zerfall des Higgs-Bosons in ein Bottom Paar.

Da es bisher noch nicht gelungen ist, den Prozess $t\bar{t}H(H \rightarrow b\bar{b})$ zu beobachten, spielt die Simulation des Prozesses mithilfe von Monte-Carlo-Generatoren eine wichtige Rolle. Die vorliegende Arbeit vergleicht Datensätze, die von verschiedenen, aktuell bei ATLAS genutzten Monte-Carlo-Generatoren modelliert wurden. Besonderes Augenmerk liegt dabei auf der Betrachtung von Unsicherheiten, die aus der Wahl von unterschiedlichen Parametern in der Generierung herrühren. Betrachtet werden hierfür Proton-Proton Kollisionen bei $\sqrt{s} = 13$ TeV.

Abstract

Since the discovery of the Higgs boson in 2012, there is special attention on the investigation and measurement of its properties, such as the coupling between the Higgs boson and other Standard Model particles. To investigate the coupling between the Higgs boson and fermions, the production of a Higgs boson in association with a top quark pair, as well as the decay of the Higgs boson in a bottom pair plays a special role.

Up until today, the $t\bar{t}H(H \rightarrow b\bar{b})$ process has not been observed at the LHC, therefore its simulation with the help of Monte Carlo generators is essential. This thesis compares samples produced by different MC generators currently used in ATLAS. In particular, the modelling uncertainties caused by variations of parameters in the generation are considered. Events produced in proton-proton collisions at $\sqrt{s} = 13$ TeV are taken into account.

Contents

1	Introduction	1
2	Theoretical background	3
2.1	The Standard Model	3
2.1.1	The Top Quark	4
2.1.2	The Higgs boson	5
2.1.3	The $t\bar{t}H$ process	8
3	Experimental Setup	9
3.1	The Large Hadron Collider	9
3.2	The ATLAS Detector	9
4	Monte Carlo simulation	13
4.1	Matrix element generators	13
4.2	Monte Carlo generators	14
4.3	Hadronisation models	15
5	Production of $t\bar{t}H$ samples with the Atlas framework	17
5.1	Event generation	17
6	$t\bar{t}H$ signal modelling uncertainties	21
6.1	Comparison of different MC generators	22
6.2	QCD scale variations	24
6.2.1	Variations of the Hdamp parameter	30
7	Conclusions	33

1 Introduction

Exclusion of theories or measurement of parameters, such as mass or coupling constants, is essential to acquire knowledge in particle physics and need the comparison of theoretical predictions and real collider data. Signal properties are modelled by using Monte Carlo (MC) generators. Based on the appropriate model, they are able to predict various distributions of observables. MC generators use different parameters which cannot be deduced from the theoretical models. Their origin lies in the limitation of numerical models or perturbation theory.

In this thesis, studies of the variations of QCD scale and `hdamp` parameters for the simulation of $t\bar{t}H$ events at $\sqrt{s} = 13$ TeV are shown. The used samples were produced by `MADGRAPH5_AMC@NLO` [1] interfaced with `PYTHIA 8` [2] or `HERWIG ++` [3, 4] and `POWHEG` [5] interfaced with `PYTHIA 8`.

First, the theoretical background will be discussed including a short overview of all particles and their interactions. In particular, the top quark and Higgs boson with their properties will be introduced. Second, a short description of the experiment (LHC and ATLAS) follows. Chapter 4 gives an overview of the event simulation done by MC generators. Afterwards, the production of the `POWHEG + PYTHIA 8` sample is described. This sample is compared to samples produced by other generators and different parameter settings are considered. The final section presents the conclusions.

2 Theoretical background

2.1 The Standard Model

The Standard Model of Particle Physics (SM) is a theoretical model based on quantum field theory, local gauge theory and the Dirac equation which describe the elementary particles and the different forms of interaction between them. The Lagrangian of the SM is renormalisable, invariant and local with respect to the gauge symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$.

Spin-half particles (called fermions) can be divided in two different groups: leptons and quarks. Both groups consist of three generations which means that corresponding fermions of different generations have the same properties such as charge or weak isospin, but differ in mass. An overview of the properties of quarks and leptons is given in Table 2.1. The existence of antifermions can be theoretically deduced from the Dirac equation and was proved in experiments.

The forces between fermions can be described in the frame of quantum field theory as the exchange of spin-1 gauge bosons. There are three forces with corresponding bosons: The electromagnetic force is described via the exchange of a photon. The photon itself is electrically neutral, but only couples to charged particles. This is the reason why there are no self-couplings in quantum electrodynamics. The bosons corresponding to the strong force are the gluons which carry a so-called colour-charge and an anticolour-charge. They only interact with colour-charged particles which means that this interaction affects the

	Leptons		Quarks		Bosons
Electric charge	-1	0	+2/3	-1/3	
1. generation	electron e	neutrino ν_e	up	down	
2. generation	muon μ	neutrino ν_μ	charm	strange	
3. generation	tau τ	neutrino ν_τ	top	bottom	
Interaction	electromagnetic		electromagnetic		photon γ
	weak	weak	strong	weak	gluon g W^\pm, Z

Table 2.1: Summary of the fundamental fermions and their possible ways of interaction.

2 Theoretical background

quarks. Self-couplings of the eight different gluons are possible. The weak interaction, as it is observed for example in β -decays, is the exchange of a massive W^\pm or Z^0 boson. The fourth force, gravity, is not included in the SM.

The masses of W and Z bosons would break the local gauge symmetry of the SM, leading to an inconsistent theory. This problem is solved by the introduction of a spin-0 scalar boson, namely the Higgs boson, and the Higgs mechanism which completes the SM. This will be described further in section 2.1.2.

2.1.1 The Top Quark

The top quark, predicted since the 1970s, was first discovered at TEVATRON in 1995 [6, 7]. It is the up-type quark of the third generation and is the heaviest elementary particle with a mass of $m_t = 173.34 \pm 0.76$ GeV [8]. Because of its high mass, it behaves in a very special way and can be used not only to investigate the SM, but also to search for physics beyond the Standard Model (BSM). The high mass leads to a big decay width and a short lifetime of $\tau_t \approx 5 \times 10^{-25}$ s, such that there are no hadronic bound states of the top quark. Moreover, the lifetime is shorter than the time used for spin decorrelation, so that spin effects propagate to the decay products.

Understanding top quark processes is essential for searching BSM phenomena. On the one hand, the high mass of the top quark offers good possibilities to search for other high massive particles. On the other hand, it leads to high transverse momenta of the various decay products such that the signal of top quark decays could mimic signals of potential new physics.

Production

There are two ways to produce top quarks at hadron colliders: The single-top production and the predominant pair production via annihilation of quark and antiquark or gluon fusion. At LHC, the gluon fusion dominates the production of top quarks because LHC is a matter-matter collider and the center of mass energy is high. Tree level diagrams of $t\bar{t}$ production through gluon fusion can be seen in Figure 2.1. They are indistinguishable with respect to the final states and interfere constructively with each other.

Decay

Because of its mass, the top quark is the only elementary particle that can decay into a real W or Z boson. The dominant decay is $t \rightarrow Wb$ with a branching ratio of over 99.9%, because the decay in s quark and d quark is strongly suppressed by the CKM matrix

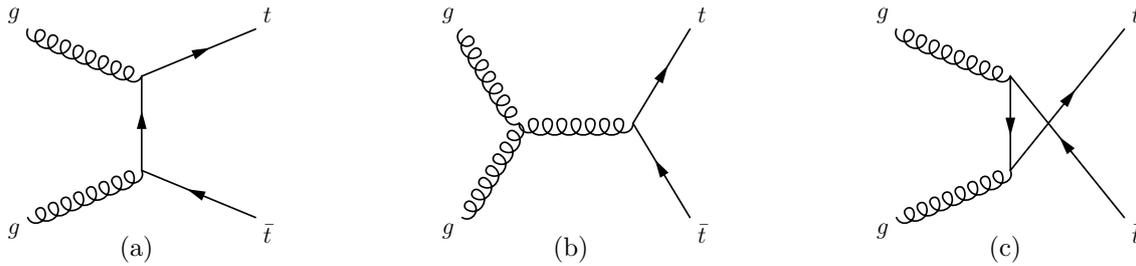


Figure 2.1: Dominant processes for the production of $t\bar{t}$ pairs.

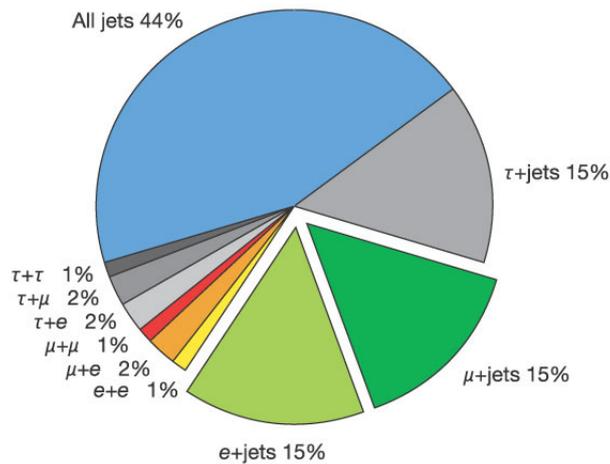


Figure 2.2: Representation of the branching ratios of different top pair decay channels.

elements. The b -quark hadronises and further decays and thus a jet can be measured in the detector. The W boson itself is not stable, either. It can decay into either two quarks or a charged lepton and a neutrino. The possible decays of a $t\bar{t}$ pair and the branching ratios of these processes can be found in Figure 2.2. One can see that the full hadronic decay channel has the highest branching ratio.

2.1.2 The Higgs boson

The Higgs boson and the corresponding Higgs mechanism are an essential part of the SM. Without the Higgs mechanism, the SM would only predict massless particles which move at the speed of light. It is obvious that this prediction does not describe reality. The Higgs boson was predicted in the 1960s and was discovered in 2012 by the ATLAS [9] and CMS [10] experiments at the LHC.

The Higgs boson is a spin-0 scalar particle with a mass of 125.09 ± 0.21 GeV [11] which can be understood as an excitation of the Higgs field in the context of quantum field

2 Theoretical background

theory. This field is described by a doublet of complex scalar fields ϕ with the potential

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2. \quad (2.1)$$

μ and λ stand for two complex constants. The potential must have a finite minimum, so $\lambda > 0$. If $\mu^2 < 0$, the neutral component of the Higgs field has a non-zero vacuum expectation value $v/\sqrt{2}$ with $v = 246$ GeV and the vacuum state is degenerate.

The combination of gauge symmetry of the electroweak model and spontaneous symmetry breaking by choosing a specific vacuum state makes the mass generation of the W and Z boson obvious.

Yukawa Coupling

In the SM, the Higgs boson not only couples to the massive spin-1 bosons, but also to all massive fermions. This interaction is called Yukawa coupling and proportional to the fermion's mass. The Lagrangian density of this interaction is given by

$$\mathcal{L} = -\frac{gv}{\sqrt{2}}(\bar{L}\phi R + \bar{R}\phi^\dagger L) \quad (2.2)$$

with the Yukawa coupling parameter g , the $SU(2)$ doublet L and the $SU(2)$ singlet R . Left-handed chiral fermions are put in the doublet, right-handed fermions are placed in the singlet. If equation (2.2) is compared to the mass dependent part of the Dirac Lagrangian $\mathcal{L} = -m(\bar{L}R + \bar{R}L)$, it becomes clear that the Yukawa coupling of a fermion is

$$g = \frac{\sqrt{2}m}{v}. \quad (2.3)$$

Using the measured value of the top quark mass and the known vacuum expectation value, it can be concluded that the Yukawa coupling of the top quark is approximately one.

Production

At the LHC, Higgs bosons are mainly produced via gluon fusion, vector boson fusion, Higgs-Strahlung and in association with a top quark pair. This thesis will focus on the $t\bar{t}H$ process. Example Feynman diagrams of the various production channels can be seen in Figure 2.3 and the predicted production cross sections for those four processes at $\sqrt{s} = 13$ TeV can be found in Table 2.2.

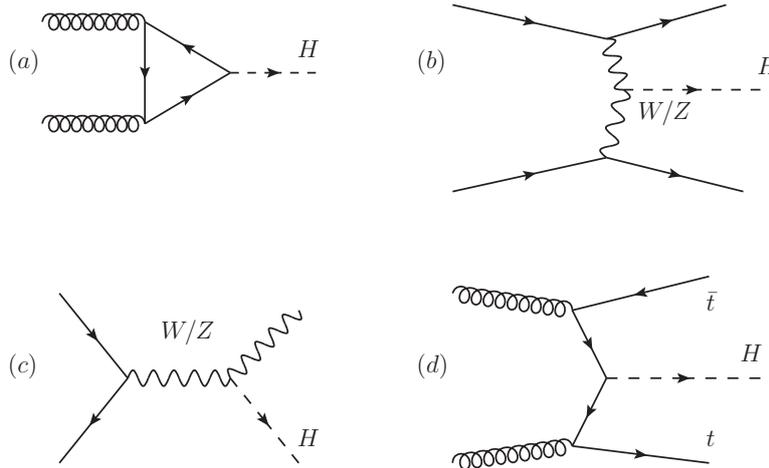


Figure 2.3: Feynman diagrams for the Higgs production via (a) gluon fusion, (b) vector boson fusion, (c) Higgs radiation and (d) in $t\bar{t}H$ processes.

	gluon fusion	vector boson fusion	Higgs-Strahlung	$t\bar{t}H$ process
Cross section [pb]	48.58	3.78	2.26	0.51

Table 2.2: Predicted cross sections of the different Higgs bosons' production channels calculated at $\sqrt{s} = 13$ TeV with a Higgs mass of $m_H = 125$ GeV [12].

Decay

The Higgs boson can decay in fermions as well as in gauge bosons. The branching ratios of the different decay modes can be found in Table 2.3.

In this thesis, the decay of the Higgs boson into $b\bar{b}$ will be considered. As shown above, the coupling between fermions and Higgs boson (Yukawa coupling) is proportional to the mass of the fermion. That is why the branching ratio of $H \rightarrow b\bar{b}$ is the highest because the bottom quark is the heaviest fermion other than the top quark. The Higgs boson cannot decay into a top quark pair, although the Yukawa coupling is higher, because the rest mass of the Higgs boson does not suffice to produce a top quark pair.

Fermions		Bosons	
Decay channel	Branching ratio	Decay channel	Branching ratio
$H \rightarrow b\bar{b}$	0.577	$H \rightarrow gg$	0.082
$H \rightarrow c\bar{c}$	0.029	$H \rightarrow \gamma\gamma$	2.28×10^{-3}
$H \rightarrow \mu\bar{\mu}$	2.19×10^{-4}	$H \rightarrow Z\gamma$	1.54×10^{-3}
$H \rightarrow \tau\bar{\tau}$	0.0632	$H \rightarrow WW$	0.215
		$H \rightarrow ZZ$	0.0264

Table 2.3: Branching ratios for different SM Higgs bosons' decay channels calculated for a Higgs bosons' mass of $m_H = 125$ GeV [13].

2.1.3 The $t\bar{t}H$ process

Since the discovery of the Higgs boson in 2012, its properties have been further investigated. One of the processes including Higgs bosons is the production in association with a top quark pair.

The Yukawa coupling in $t\bar{t}H$ events is expected to be very strong, therefore this process is highly suitable to investigate this coupling further. It could also be a window to new physics.

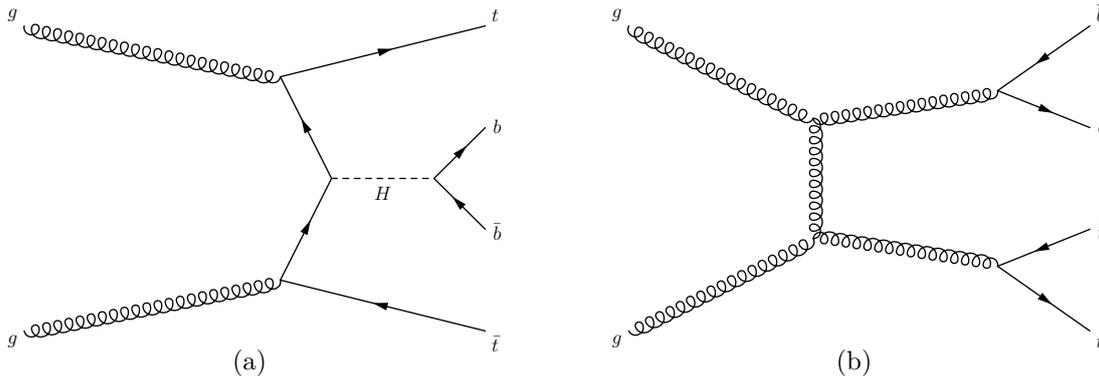


Figure 2.4: Exemplary tree-level diagram for the production of $t\bar{t}H$ with the Higgs boson decaying into a bottom quark pair (a) and the dominant background process (b).

The decay Higgs bosons into $b\bar{b}$ is of special interest, because it is the dominant decay channel. Additionally, the Yukawa coupling to the bottom quark is also very high, so it allows the investigation of another Yukawa coupling. It is very difficult to detect this signal, because it suffers from large backgrounds. Selected Feynman diagrams of the signal process and of the main background process are shown in Figure 2.4. Additionally, the top quarks decay products can be jets, too, which makes it even more difficult to measure the signal. That is why the single-leptonic or dileptonic decay of the top quarks is considered, which comes along with fewer jets one has to measure and distinguish from the background, although there are missing transverse momenta from the neutrinos. There are sophisticated multivariate analysis techniques used to search for signal.

3 Experimental Setup

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is situated at the European Organization for Nuclear Research (CERN) near Geneva in Switzerland. It is the most powerful particle accelerator in the world and is able to collide protons or lead ions. The construction began in 1998 and since 2009, the accelerator and the corresponding four main experiments provide data for research. ATLAS and CMS are both general purpose detectors, while ALICE focusses on heavy ion physics and LHCb mainly deals with bottom quark physics and low scattering processes.

Before the protons or ions are injected in the main accelerator complex, a ring with 26.7 km circumference which is located approximately 100 m under the surface, they are accelerated by different smaller circular and linear accelerators. The whole complex can be seen in Figure 3.1. LHC belongs to the class of synchrotron colliders, and the high energies are achieved with superconducting magnets. Dipole magnets are used to conduct the beams in the beam pipe, while quadrupole magnets are able to stabilise and focus the beams.

Until now, two periods of data recording took place. Run I occurred from 2010 till 2012 with a centre of mass energy $\sqrt{s} = 7$ TeV in 2010 and 2011 and $\sqrt{s} = 8$ TeV in 2012. After a shutdown of two years, Run II began in 2015 with $\sqrt{s} = 13$ TeV. After a second reconstruction break it is planned that in 2018 the LHC will operate at its design centre of mass energy of $\sqrt{s} = 14$ TeV.

3.2 The ATLAS Detector

ATLAS is one of the two universal detectors at LHC. It is designed and built to investigate the SM, but also to search for BSM physics. The ATLAS detector consists of three main parts: The inner detector, the calorimeters and the muon chambers. Each component is necessary to measure specific properties of the passing particles by using the knowledge of their interaction with matter. A schematic representation of ATLAS can be seen in

3 Experimental Setup

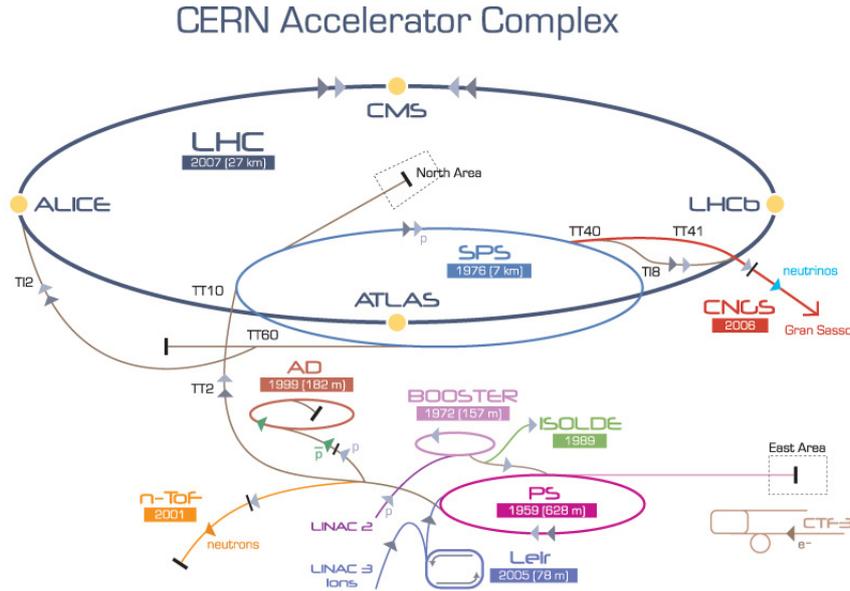


Figure 3.1: Overview of the different accelerators at CERN, including the LHC and the location of the four main experiments ATLAS, CMS, LHCb and ALICE. ©CERN

Figure 3.2.

The inner detector is surrounded by a 2 T solenoidal field and consists of the Silicon Pixel Detector, the Silicon Microstrip Tracker (SCT) and the Transition Radiation Tracker (TRT). A very high resolution of the inner detector is needed for differentiation of pile-up events and tagging for b-jets. The pixel detector and the SCT are based on silicon technology and are used to measure the momenta of charged particles. The resolution σ_p/p of the measurement of the momentum p is proportional to p . The pixel detector consists of hybrid pixels which means that the silicon detector and the readout electronics is connected via bump-bond. This technology provides a high-speed data readout, which is very important because LHC has a bunch crossing frequency of 40 MHz. The inner detectors are exposed to high radiation. Since 2015, the IBL detector is the new innermost layer of the pixel detector which was designed as a new, radiation hard high-speed detector close to the collision point. The inner detector is planned to be replaced in the next break before Run III due to radiation damage. In the TRT, ultra-relativistic charged particles produce transition radiation. The TRT is used to identify electrons and positrons as they are extremely ultra-relativistic due to their small masses.

In the calorimeters, the particles are absorbed to measure their energy. The calorimeters of ATLAS are built as sampling calorimeters which means that there are passive regions in which the particles are stopped and an active region to measure the signals. In the electromagnetic calorimeter, the main effect of energy deposition is a sequence of

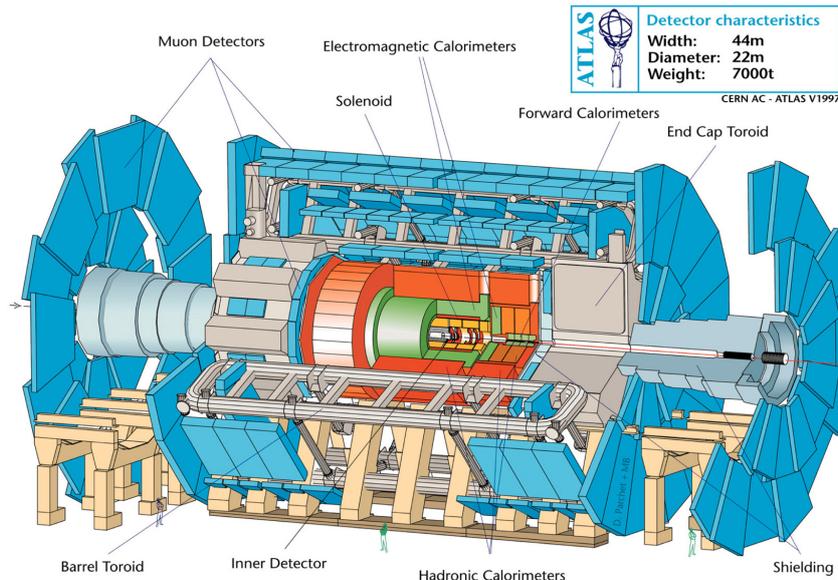


Figure 3.2: Schematic representation of the ATLAS detector. ©CERN

Bremsstrahlung and pair production of electron and positron. It is built of lead or stainless steel as absorber material because of its small radiation length. Liquid argon is used as sampling material which can be ionised by the particles of parton showers. In this calorimeter the energy of charged particles and photons can be measured. As photons carry no charge, they are not registered by the inner detector. To obtain a good resolution, especially referring to photons, the first layer of the electromagnetic calorimeter has a high granularity. In the following hadronic calorimeter, the particles are mainly absorbed because of the strong force. The material of the active and passive regions differ between the barrel and end-cap regions. The resolution σ_E/E of the energy measurement E in the calorimeters is proportional to $1/\sqrt{E}$.

The muon chambers are the outermost region of the ATLAS detector. Because muons are leptons, they are not absorbed in the hadronic calorimeter and because of their mass, their energy loss due to Bremsstrahlung is negligible. That is why the muons are the only charged particles which pass the detector without being stopped before the muon spectrometer (the tau leptons have the same properties, too, but their lifetime is too short to reach the Muon Spectrometers). Due to the toroidal magnet system, the muon's trajectory is curved and the momentum is measured by layers of Monitored Drift Tubes.

Because of the high collision rate, it is impossible to store the information of all occurring events. At ATLAS, a three level trigger system filters out events that are not desired for data analysis, so that the rate of stored events is about 200 MHz.

The focus of this thesis lies on $t\bar{t}H$ events with the top pair decaying semileptonically and

3 Experimental Setup

the Higgs decaying into a bottom antibottom pair. To detect such kind of events, all parts of the detector are needed because of the various decay products. The kinematic properties of leptons, neutrinos and jets have to be measured. Neutrinos cannot be detected, because their interaction cross section with matter is very small. For this reason, measuring the transverse momenta of the other decay products as exactly as possible is important to reconstruct the missing transverse momenta. Finally, as the most likely Higgs decay channel $H \rightarrow b\bar{b}$ is considered in this thesis, b-tagging is essential. B-tagging describes the detection of a jet caused by the production of a bottom quark. The meson including a bottom quark has a relatively long lifetime, so that the hard interaction vertex and the origin of the jet is shifted. The distance between these two vertices can be measured, if the resolution of the tracking detector is high enough and b-jets can be identified.

Also, to differentiate between the signal and the huge background in hadronic decay channels, a good jet detection and reconstruction is needed.

4 Monte Carlo simulation

4.1 Matrix element generators

Matrix element (ME) generators calculate all Feynman diagrams of the given process up to a specific order. Normally, event generators model the cross section to the lowest order of perturbation theory (LO) or to the next-to-leading order (NLO). The event generators used for the following analysis, POWHEG and MG5_AMC are both NLO generators which means that the tree-level Feynman diagram is computed with up to one additional emission of gluons.

The basic process is the hard scattering resulting from the collision of two partons. The $(n+k)$ -th order approximation of the cross section of a hard process is given by the factorization theorem as [14]

$$\sigma = \sum_{i,j} \int dx_1 dx_2 p_i(x_1, \mu_F^2) p_j(x_2, \mu_F^2) \times \sum_{m=0}^n (\alpha_s(\mu_R^2))^{(m+k)} \sigma_{ij}^{(m)} \left(p_1, p_2, \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2} \right). \quad (4.1)$$

σ_{ij} describes the partonic cross section, $x_{1/2}$ the Björken-x and $p_{i/j}$ the parton density inside hadron $h_{i/j}$. The parton density function (PDF) has to be measured and can not be deduced from QCD. α_S is the strong coupling constant, $p_{1/2}$ the momenta of the partons participating in the hard process and Q the momentum transfer.

Renormalisation scale μ_R

The strong coupling constant α_S depends on the energy. It is low for high energies and vice versa. This behaviour can be deduced from the referring symmetry group which is non-Abelian. The strong coupling constant can be described by the renormalisation group equation

$$\mu_R^2 \frac{d\alpha_s}{d\mu_R^2} = \beta(\alpha_s) \quad (4.2)$$

with the β function. The renormalisation scale parameter μ_R is an unphysical parameter. However, if it is chosen approximately equal to Q , α_S describes the strength of the strong

4 Monte Carlo simulation

coupling in the given process.

Factorization scale μ_F

Theorem (4.1) factorizes the cross section in a part where perturbation theory is applicable and one part where α_s is too large to be developed in terms of perturbation theory. The factorization scale μ_F describes the limit. Above this limit, perturbation theory is valid and the strong coupling constant can be calculated as described before. Under this limit, non-perturbative approaches have to be taken into account.

Normally, the renormalisation scale and the factorization scale are set equal $\mu_R = \mu_F \approx Q$.

hdamp parameter

In POWHEG an additional parameter is used to suppress higher-order effects without changing the calculations of the cross sections at lower orders. To do so, the real emission cross section \mathcal{R} is rescaled by a factor \mathcal{D} which has the following form [15]

$$\mathcal{D} = \frac{h^2}{p_T^2 + h^2} \quad (4.3)$$

with p_T describing the transverse momentum. Because of this construction, \mathcal{D} can take values between zero and one. The real emission cross section \mathcal{R} can be divided into two parts:

1. The term which includes resummation of the approximate higher-order corrections.
2. The part which neglects higher-order effects.

The first one is weighted with the factor \mathcal{D} , the second one with $(1 - \mathcal{D})$. For small transverse momenta, \mathcal{D} is close to one and the higher-orders are included, so that the cross section does not change. For high transverse momenta the second part is mainly considered. This limits the higher-order effects.

The value of \mathcal{D} can be varied in POWHEG by changing the parameter `hdamp` which is equal to h in equation (4.3).

4.2 Monte Carlo generators

Current ME generators are only able to calculate the cross section up to NLO because the computation of higher order corrections would drastically increase the required amount of CPU time. Nevertheless, higher order corrections can not be completely ignored. So

parton showers are used as an approximate method to calculate higher order corrections and match them with the later following hadronisation.

The differential cross section of an $(n + 1)$ -particle state can be expressed in terms of the n -particle cross section (after the integration over the azimuthal angle and with the small-angle approximation):

$$d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{\alpha_s}{2\pi} P_{ba}(z) \quad (4.4)$$

with the momentum transfer-squared t , the momentum fraction z and the appropriate splitting function $P_{ba}(z)$ describing the probability of emitting an additional parton. Equation (4.4) can only be integrated in the limits of Q^2 and Q_0^2 , where Q^2 is the squared momentum transfer of the hard process and $Q_0^2 \approx 1 \text{ GeV}^2$ is defined as the lower bound at which perturbation theory is applicable.

To include not only real parton emission, the Sudakov form factor [14] can be introduced

$$\Delta(t) \equiv \exp \left[- \int_{t_0}^t \frac{dt'}{t'} \int dz \frac{\alpha_s}{2\pi} P(z) \right] \quad (4.5)$$

which describes the probability of evolving from t_0 to t without branching into new partons. To develop the parton cascade, one has to calculate the variables t and x after each step. t_2 (momentum transfer-squared following t_1) is generated by solving

$$\frac{\Delta(t_2)}{\Delta(t_1)} = \chi \quad (4.6)$$

with χ being a uniformly distributed random number out of the interval $[0, 1]$. The corresponding momentum x_2 is set according to the appropriate splitting function $P(z) = P(x_2/x_1)$ with the previous momentum x_1 . If $t_2 > Q^2$, the branching of this parton stops, otherwise the procedure is repeated.

4.3 Hadronisation models

The strong coupling constant depends on the energy which means that decreasing energy leads to increasing α_s . At some point (around $Q \approx 1 \text{ GeV}$), perturbation theory is no longer applicable. To simulate the following processes, mainly two different non-perturbative models are used.

The first one is the string model which is used for example in PYTHIA 8. Each parton in the system is connected to its own colour partner via a so-called string. The potential

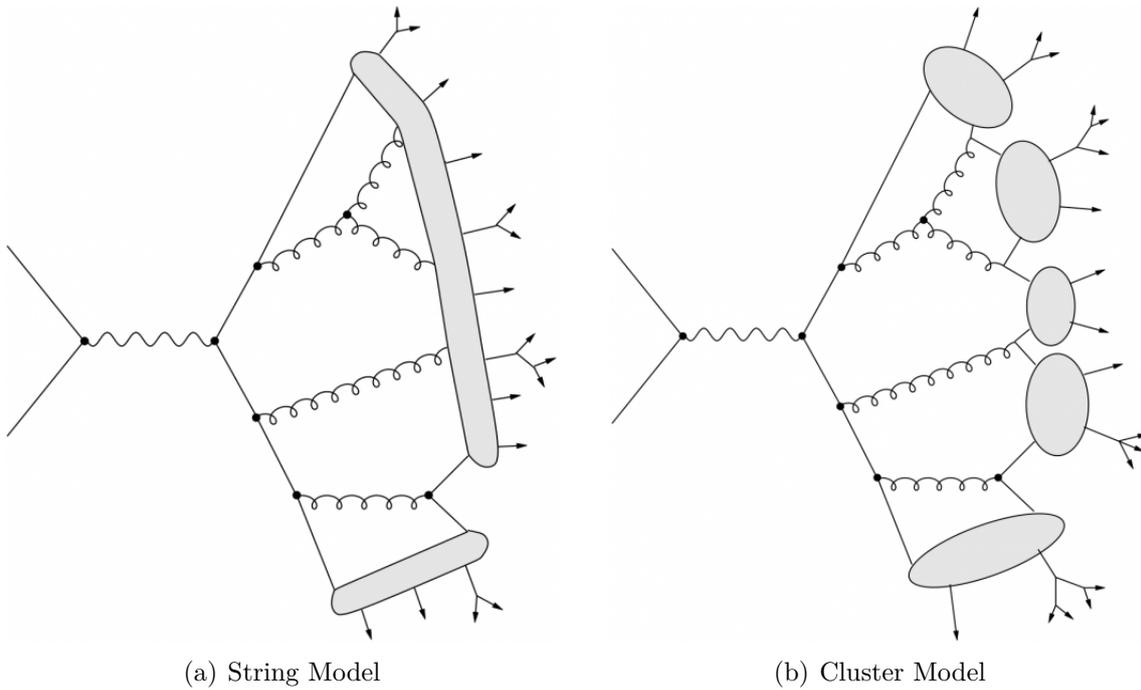


Figure 4.1: Schematic representation of the different models used for hadronisation.

energy in the string is $V = \frac{-4\alpha_s}{3r} + \kappa r$. It is obvious that the energy increases with distance r . If the distance and the potential energy between the partons is big enough to produce a quark/anti-quark pair, the string breaks up. This process ends if there is not enough energy to produce new quark pairs. The partons stay in hadronic bound states. A schematic representation can be seen in Figure 4.1(a).

The second model is the cluster model. It is based on the assumption that partons in a shower with energy below the hadronisation scale form colourless groups. The simplest way of forming colourless groups is the quark/antiquark production from an incoming gluon. The colourless groups decay into hadrons using a simple isotropic quasi-two-body phase space model. A schematical overview can be seen in Figure 4.1(b). This model is used for example in HERWIG ++.

5 Production of $t\bar{t}H$ samples with the Atlas framework

5.1 Event generation

At the moment, there exist two different MC samples of $t\bar{t}H$ events for Run II at $\sqrt{s} = 13$ TeV in ATLAS. In both samples, the hard process is generated at NLO with MG5_AMC. Events are interfaced to PYTHIA 8 or HERWIG ++ for parton showering, hadronisation and the generation of the underlying processes. To avoid over or undercounting of events during matching, the MC@NLO method is used. The top mass is set to 172.5 GeV and the Higgs mass to 125 GeV in both samples. As particle distribution function (PDF) CT10nlo is used for the ME in MG5_AMC and CTEQ6L1 with UE-EE-5 tune for the parton showering in the MG5_AMC/HERWIG ++ sample. For the combination of MG5_AMC and PYTHIA 8 the PDF is set to NNPDF3.0nlo for the hard process and to NNPDF2.31o with the A14 tune option for the parton showering. The nominal set of the factorization and renormalisation scale is $\mu_R = \mu_F = H_T/2$. H_T describes the scalar sum of the transverse masses of all final state particles $H_T = \sum_i m_T(i) = \sum_i \sqrt{m^2(i) + p_{\perp}^2(i)}$. Thanks to the new multiple event weight method in MG5_AMC, it is not necessary to generate new MC event samples to analyze the influence of changes of the QCD scale or PDF settings. An overview of all settings in the already existing MG5_AMC samples can be found in the first two columns of Table 5.1 [16].

For this thesis, a new MC sample of $t\bar{t}H$ events was produced. The ME of the hard process was calculated with POWHEG [5] which is a generator already used in Run I to generate, for example, top pair events in ATLAS. All masses were set equal to the settings of the MG5_AMC samples. The POWHEG specific `hdamp` parameter (as described in section 4.1) is implemented and set to 235 GeV which corresponds to the sum of the Higgs mass, top mass and antitop mass divided by two. To avoid under and overcounting of events, POWHEG is matched to PYTHIA 8 via the programme `main31`. The PDF is set to `PDF4LHC15_nlo_30_pdfas` for the generation of the ME and to `NNPDF3.0nlo` with the A14 tune for parton showering. The nominal renormalisation and factorization scale is set

5 Production of $t\bar{t}H$ samples with the ATLAS framework

to $\sqrt[3]{m_H^T \cdot m_t^T \cdot m_{\bar{t}}^T}$ with the transverse masses of Higgs, top and antitop. The uncertainties of the changes in QCD scale and PDF choice can be calculated without re-running because POWHEG includes the reweighting technique, too.

To estimate the different uncertainties due to the generator settings, the following variations have been considered and compared to the nominal samples.

MC generators for the matrix element calculation

The uncertainty due to the choice of the MC generator for the ME is estimated by comparing POWHEG and MG5_AMC samples, both interfaced to PYTHIA 8.

Showering, hadronisation and underlying event

The influence of different parton shower implementations and hadronisation models can be analyzed by comparing the samples MG5_AMC + PYTHIA 8 and MG5_AMC + HERWIG ++.

QCD scale variations: renormalisation and factorization scale choice

To analyse the influence of QCD scale variations, the renormalisation and factorization scales were changed by a factor of 1/2 and 2 independently of each other, leading to a total of eight variations. One of these variations is not considered because it is physically meaningless.

hdamp scale variations

Changing the POWHEG specific `hdamp` parameter allows to investigate the influence of this parameter on different kinematic variables. To do so, the `hdamp` parameter was set to $1/2m_t$ for the down variation and $2m_t$ for the up variation.

$t\bar{t}H, \sqrt{s} = 13 \text{ TeV}, m_H = 125 \text{ GeV}, m_t = 172.5 \text{ GeV}$		
ME gen.	MadGraph5_aMC@NLO v2.2.1	Powheg
h damp	+ MadSpin	$m_t + m_H/2 = 235 \text{ GeV}$
PS/UW gen.	-	Pythia8
Ren./Fac. scale	Herwig++	v8.210
ME & PS/UE PDF	v2.7.1	$\sqrt[3]{m_H^T \cdot m_t^T \cdot m_{\bar{t}}^T}$
Tune	$H_T/2$	PDF4LHC15_nlo_30_pdfas & NNPDF2.3lo
Cross section [pb]	CT10nlo & CTEQ6L1	A14
	UE-EE-5	0.511 ± 0.003
	0.447 ± 0.007	0.457 ± 0.001

Table 5.1: Overview of all relevant generator setting in the used samples [16]

6 $t\bar{t}H$ signal modelling uncertainties

The purpose of all studies was the comparison of different MC generators for ME calculation and parton showering, as well as the investigation of the influence of different generator settings on the generation of $t\bar{t}H$ events. For all analysis, the semileptonic decay channel of the top pair is considered ($t\bar{t} \rightarrow qq'l\nu_l$). Additionally, only events in which the Higgs decays in a bottom pair ($H \rightarrow b\bar{b}$) are included in the modelling of uncertainties. For this signature, six jets (including four b-jets) are expected per event. The analysis performed in this thesis is a so-called TRUTH analysis carried out with the programme ROOT [17]. That means, all jets in the event are defined at particle level and can be identified by a unique ID. The top quarks and the Higgs bosons are considered at parton level. In the following studies, only jets with transverse momentum $p_T > 25$ GeV and pseudo-rapidity $|\eta| < 2.5$ will be considered. The same restrictions are applied to H_T .

The samples are mainly normalized in two different ways to closely investigate the influence of cross section and acceptance effects:

1. Normalisation to the appropriate cross section multiplied by the relevant branching ratio. Considering the Higgs decay $H \rightarrow b\bar{b}$ and the semileptonic top pair decay channel, the relevant branching ratio is $0.4389 \cdot 0.5770 \approx 0.25$. This normalisation is only sensitive to cross section effects.
2. Normalisation of the yields to the nominal cross section multiplied by the branching ratio divided by the sum of weights before any selection of the considered variation. This way of normalisation is only sensitive to acceptance effects, i.e. the different theoretical weights of an event before and after the application of cuts.

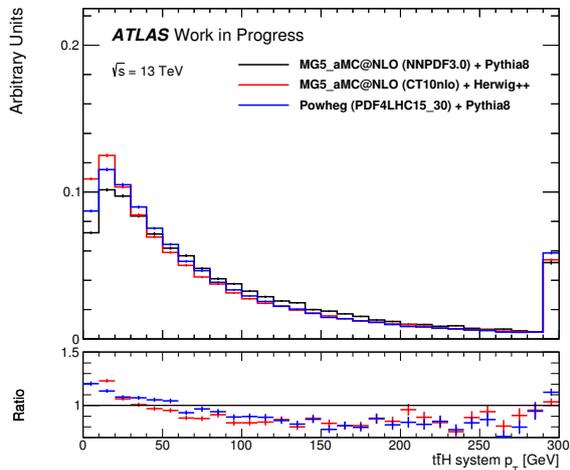
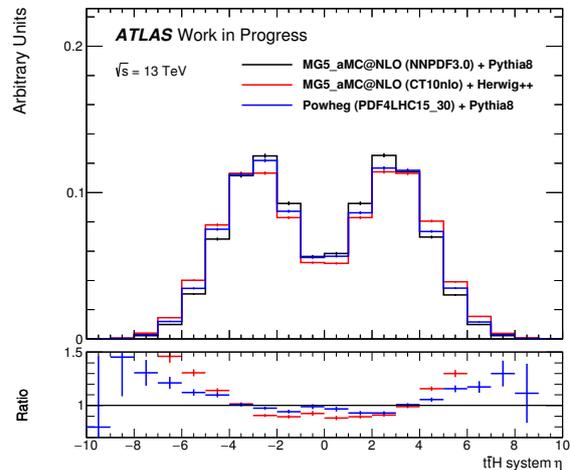
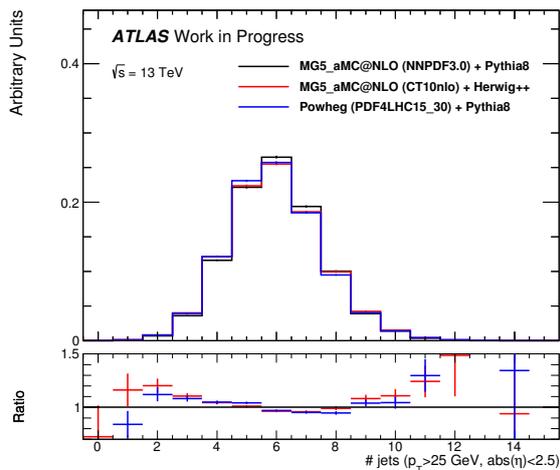
6.1 Comparison of different MC generators

The uncertainty for ME calculation and parton showering as well as hadronisation can be investigated by comparing the different MC generators. The transverse momentum and the pseudorapidity of the $t\bar{t}H$ system, the number of jets and H_T distribution can be seen in Figure 6.1. To compare the results, the samples are normalized to unity. The errors merely arise from the limited statistics in the MC generation.

As can be seen in Figure 6.1(a) and 6.1(b), the MG5_AMC + PYTHIA 8 sample describes a $t\bar{t}H$ system with harder transverse momentum and more central events.

The number of jets differs less in the region between four and eight jets per event, where the modelling uncertainty is smaller than 10%. For the $t\bar{t}H(H \rightarrow b\bar{b})$ six jets are expected, so this region is important. For a small number of jets and huge number of jets, the fraction of events differ up to 50%. But the statistical uncertainties in these regions are larger because only a few events include less than four or more than eight jets. The distributions of the scalar sum of transverse masses H_T show similar behaviour: The fraction of events with $H_T < 300$ GeV generated by POWHEG + PYTHIA 8 and MG5_AMC + HERWIG ++ is larger. The modelling uncertainty amounts up to 30%, but there are large statistical uncertainties.

The differences between the MC Generators are all in all as expected because of the different numerical implementations and scale settings. The modelling uncertainties of the ME generation and parton showering are strongly dependent on the size. For closer investigation, samples with a larger number of events could be considered, so that the statistical uncertainties in the border regions would decrease.

(a) Transverse momentum of $t\bar{t}H$ system.(b) Pseudorapidity of $t\bar{t}H$ system.

(c) Number of jets.

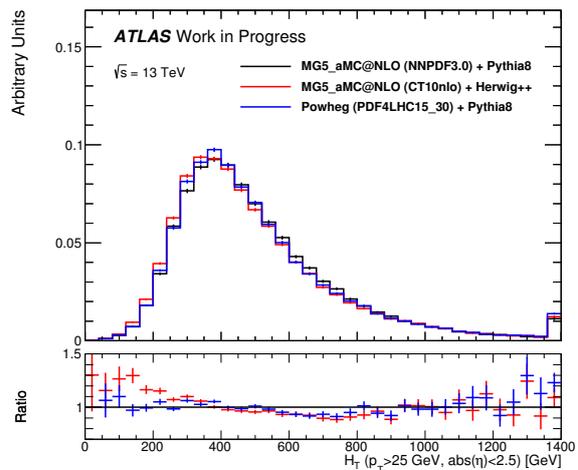
(d) H_T .

Figure 6.1: Comparison of different MC samples for different kinematic variables of the $t\bar{t}H$ process generated at $\sqrt{s} = 13$ TeV. The results are normalized to one. The uncertainties only result from the MC generation, so that they are purely statistical errors. The last bin contains the overflow.

6.2 QCD scale variations

The modelling uncertainties of QCD scale variations were calculated for POWHEG + PYTHIA 8 and MG5_AMC + PYTHIA 8. To investigate the different influences of cross section and acceptance effects, the two different ways of normalisation described in section 6 were applied. The different cross sections for the variations of QCD scale can be found in Table 6.1.

Figure 6.2 shows the number of events as a function of different variables which are deduced from events generated by the ME generator MG5_AMC. The events are normalized to only cross section effects. The Figure includes the number of events according to the transverse momentum and pseudo rapidity of the $t\bar{t}H$ system as well as the number of jets per event and H_T . For the selection of the jets, the previously described cuts were used. The impact on the shape of the distributions is negligible. The same variables with the similar cuts are displayed in Figure 6.3, but they are normalized with respect to only acceptance effects. Comparing both sets of plots, one can see that acceptance effects cause smaller modelling uncertainties.

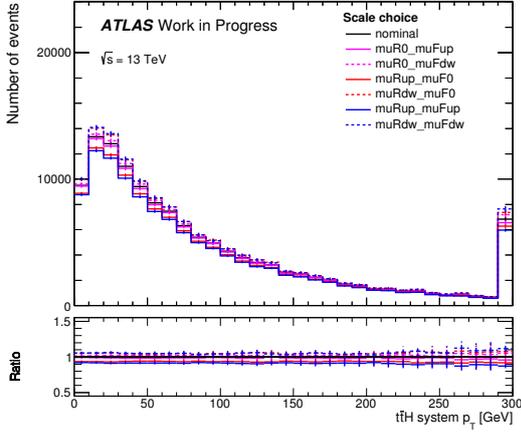
The same procedure was repeated for the events generated with POWHEG. The differential cross section depending on the same variables and normalized to only cross section effects can be seen in Figure 6.4. In contrast to the results obtained with MG5_AMC, the shape of the distributions changes with the variation of the QCD scale choice. In both cases, acceptance effects are smaller than cross section effects. The differential cross section calculated from events generated with POWHEG and normalized to acceptance effects can be seen in Figure 6.5.

For both ME generators, changes on both, renormalisation and factorization scale, in the same direction show the biggest impact on the differential cross section. Furthermore these changes with similar renormalisation and factorization scale are the most relevant ones because the two scales are usually set to the same value.

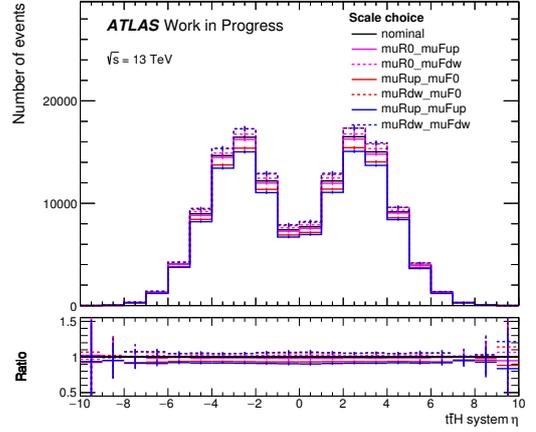
QCD scale		Cross section MG5_AMC	Cross section POWHEG
nominal		115.733 fb	129.408 fb
$\mu_R = \mu_0$	$\mu_F = 2\mu_0$	113.969 fb	126.948 fb
$\mu_R = \mu_0$	$\mu_F = \mu_0/2$	118.032 fb	132.797 fb
$\mu_R = 2\mu_0$	$\mu_F = \mu_0$	108.165 fb	120.774 fb
$\mu_R = \mu_0/2$	$\mu_F = \mu_0$	121.584 fb	135.794 fb
$\mu_R = 2\mu_0$	$\mu_F = 2\mu_0$	105.511 fb	117.354 fb
$\mu_R = \mu_0/2$	$\mu_F = \mu_0/2$	122.157 fb	137.326 fb

Table 6.1: Cross section dependent on the QCD scale choice for ME generators MG5_AMC and POWHEG combined with PYTHIA 8. μ_0 stands for the nominal setting $\mu_0 = H_T/2$ as described in section 5.1.

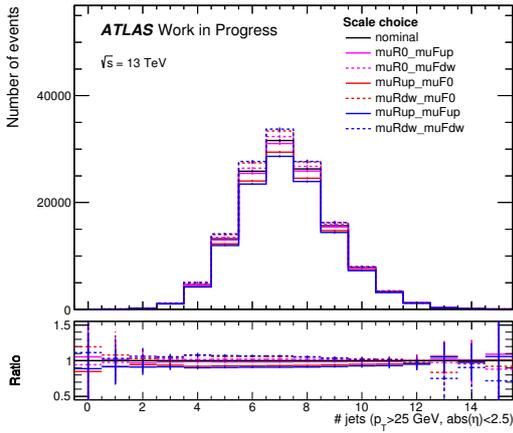
6 $t\bar{t}H$ signal modelling uncertainties



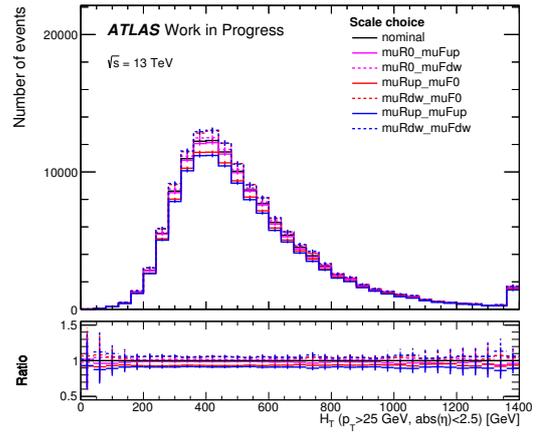
(a) Transverse momentum of $t\bar{t}H$ system.



(b) Pseudorapidity of $t\bar{t}H$ system.

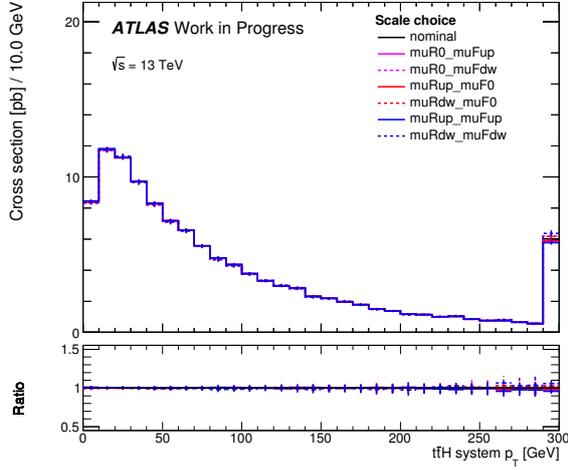
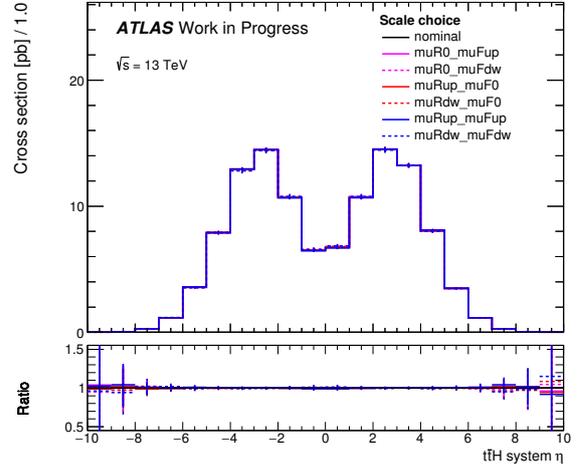
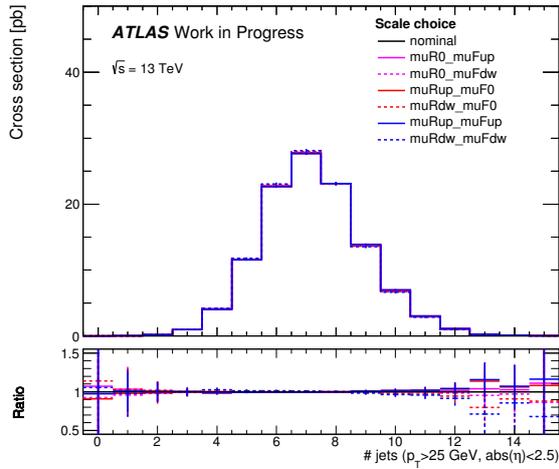


(c) Number of jets.



(d) H_T .

Figure 6.2: Comparison of variations of the QCD scale choice for different kinematic variables of the $t\bar{t}H$ process generated with MG5_AMC + PYTHIA 8 at $\sqrt{s} = 13$ TeV. The results are normalized with respect to cross section effects. The uncertainties are statistical uncertainties due to limited statistics in the MC sample.

(a) Transverse momentum of $t\bar{t}H$ system.(b) Pseudorapidity of $t\bar{t}H$ system.

(c) Number of jets.

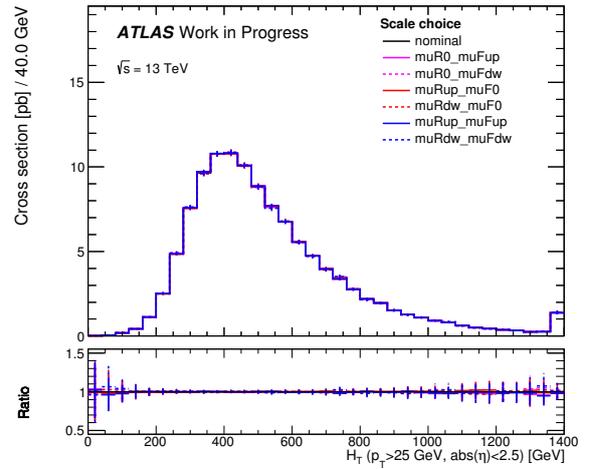
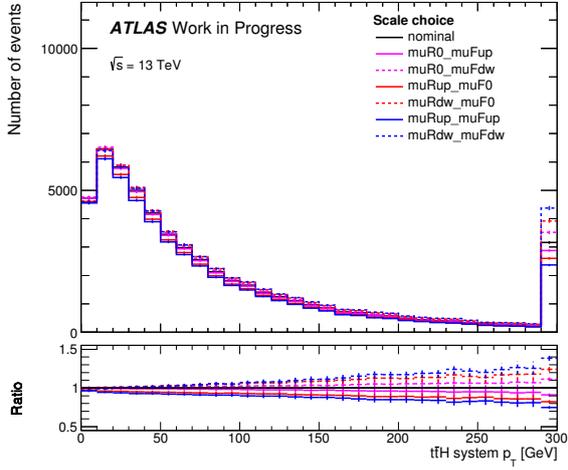
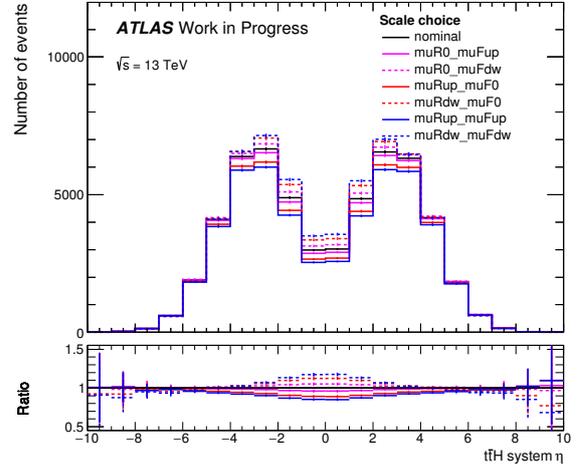
(d) H_T .

Figure 6.3: Comparison of variations of the QCD scale choice for different kinematic variables of the $t\bar{t}H$ process generated with MG5_AMC + PYTHIA 8 at $\sqrt{s} = 13$ TeV. The results are normalized with respect to acceptance effects. The uncertainties only originate from the MC generation.

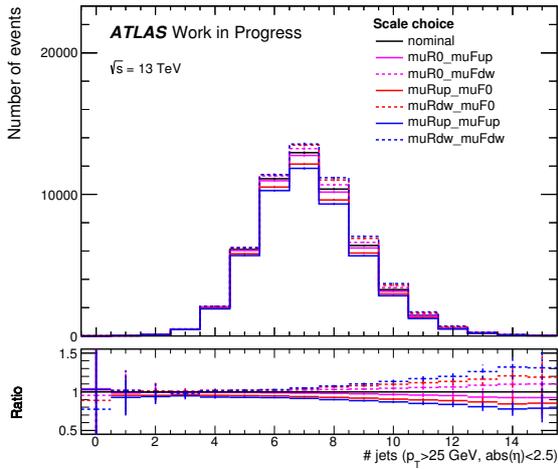
6 $t\bar{t}H$ signal modelling uncertainties



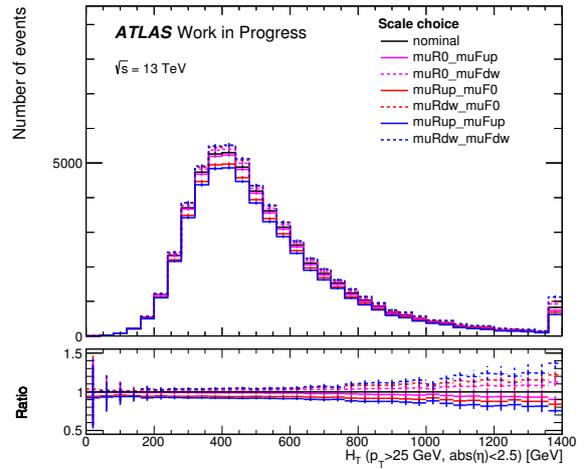
(a) Transverse momentum of $t\bar{t}H$ system.



(b) Pseudorapidity of $t\bar{t}H$ system.

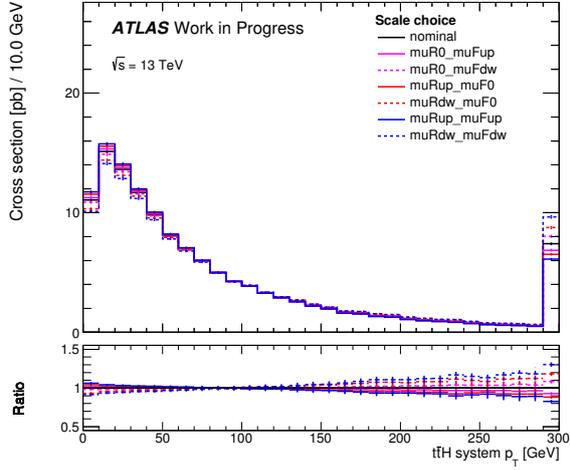
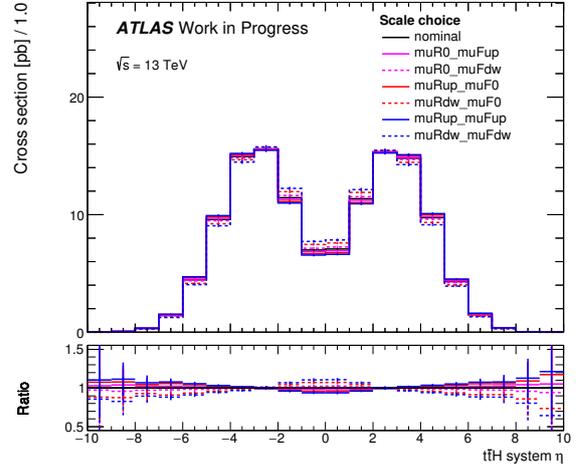
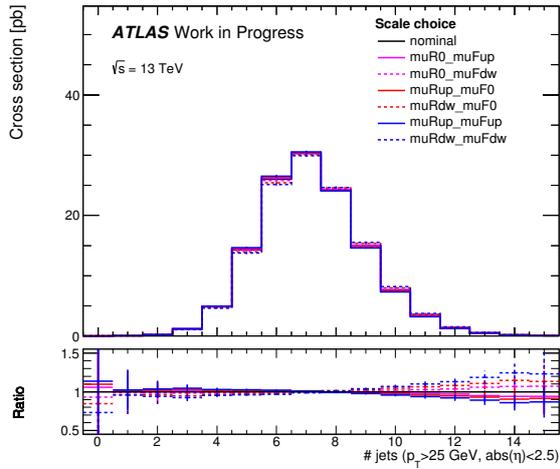


(c) Number of jets.



(d) H_T .

Figure 6.4: Comparison of variations of the QCD scale choice for different kinematic variables of the $t\bar{t}H$ process generated with POWHEG + PYTHIA 8 at $\sqrt{s} = 13$ TeV. The results are normalized with respect to cross section effects. The errors result from statistical uncertainties caused by the MC generation.

(a) Transverse momentum of $t\bar{t}H$ system.(b) Pseudorapidity of $t\bar{t}H$ system.

(c) Number of jets.

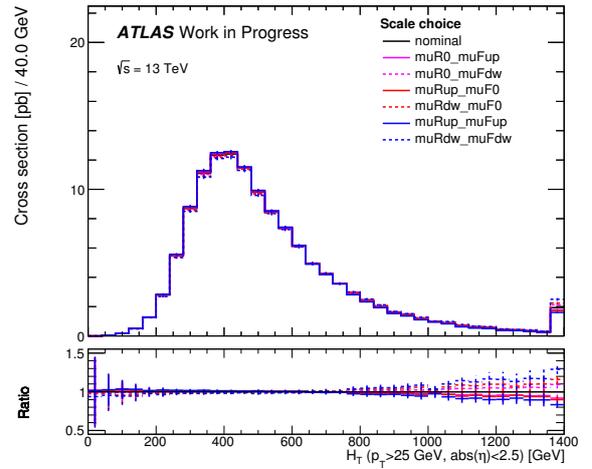
(d) H_T .

Figure 6.5: Comparison of variations of the QCD scale choice for different kinematic variables of the $t\bar{t}H$ process generated with POWHEG + PYTHIA 8 at $\sqrt{s} = 13$ TeV. The uncertainties are statistical errors caused by the MC generation. The results are normalized with respect to acceptance effects.

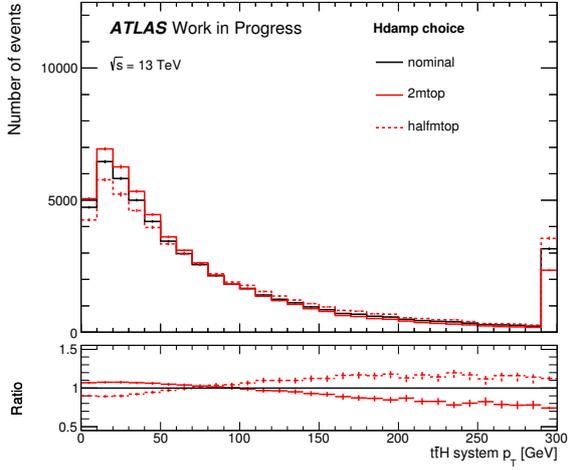
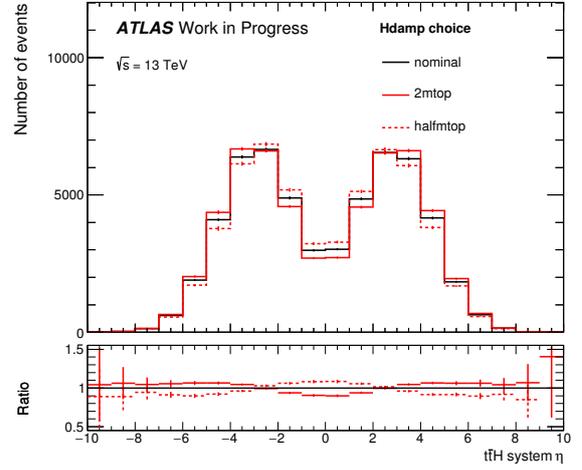
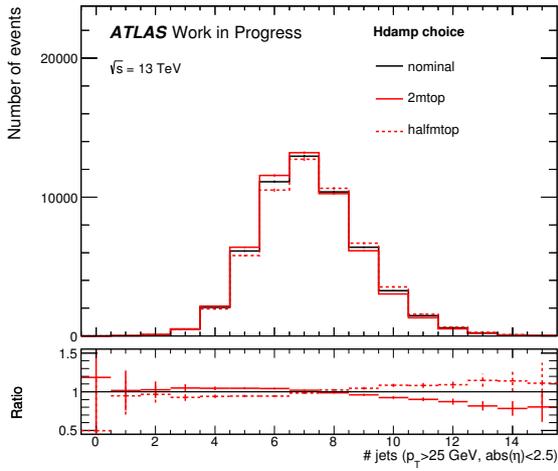
6.2.1 Variations of the Hdamp parameter

The differences resulting from the variation of `hdamp` scale were calculated for the POWHEG + PYTHIA 8 sample. Cross sections for the different variations can be found in Table 6.2.

<code>hdamp</code> scale	Cross section
nominal	129.408 fb
$2 \cdot m_{\text{top}}$	129.826 fb
$1/2 \cdot m_{\text{top}}$	128.718 fb

Table 6.2: Cross section depending on the `hdamp` scale choice for POWHEG matched with PYTHIA 8.

Figure 6.6 depicts the number of events depending on kinematic variables of the $t\bar{t}H$ system, number of jets per event and the scalar sum of transverse masses H_T normalized with respect to only cross section effects. Decreasing `hdamp` parameter, the $t\bar{t}H$ system is harder and more central in η and vice versa. This can be explained by the definition of the `hdamp` parameter (4.3) in section 4.1. The modelling uncertainty reaches a maximal value of 20% between the variation and nominal distribution. Additionally, the shape of the distributions changes under `hdamp` variations.

(a) Transverse momentum of $t\bar{t}H$ system.(b) Pseudorapidity of $t\bar{t}H$ system.

(c) Number of jets.

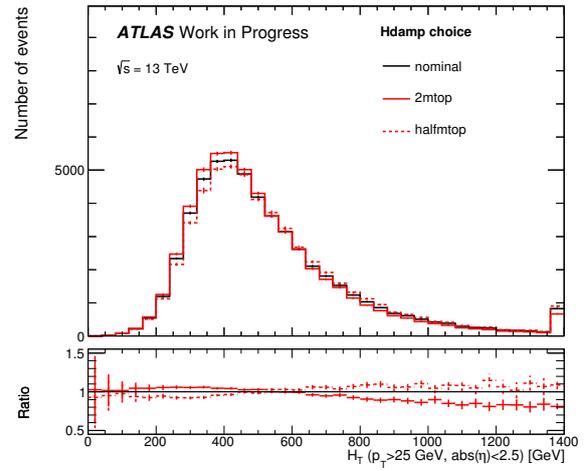
(d) H_T .

Figure 6.6: Comparison of variations of the `hdamp` choice for different kinematic variables of the $t\bar{t}H$ process generated with POWHEG + PYTHIA 8 at $\sqrt{s} = 13$ TeV. The results are normalized with respect to cross section effects. The errors result from statistical uncertainties.

7 Conclusions

The Monte-Carlo description of $t\bar{t}H$ events with the Higgs boson decaying into a bottom quark pair has been studied, focussing on the modelling uncertainties caused by the choice of different ME generators, parton shower programmes and QCD scales as well as the impact of `hdamp` in the sample of POWHEG + PYTHIA 8.

It was shown that the modelling uncertainties caused by the choice of different generators are dependent on the size of the kinematic variable. The largest differences could be observed in the border regions. Events with kinematic properties in the border regions are less likely to occur, so there were fewer events considered which led to larger statistical uncertainties. One could perform further investigations of these events with larger samples to reduce statistical uncertainties. However, the border regions can be excluded in the analysis by using cuts, so depending on the analysis, a further investigation of these regions is not necessary because all events with these properties are dropped out.

To investigate the influence of QCD scale choice, the POWHEG + PYTHIA 8 sample was compared to the sample generated by MG5_AMC + PYTHIA 8. The shape of the distributions of the kinematic variables of the MG5_AMC + PYTHIA 8 sample is stable under variations of the QCD scale. In contrast to this, the distributions described by POWHEG change under the influence of QCD scale variations. The reason for this change of shape has to be investigated further. The largest difference between the nominal distribution and a variation could be observed, if the renormalisation and factorization scale were both varied in the same direction. Acceptance effects have a smaller impact on the modelling uncertainties than the cross section effects. Considering the results of QCD scale variations, one can conclude how an analysis has to be set up to minimise the impact of modelling uncertainties.

Moreover, the impact of `hdamp` variations have been considered in a sample produced by POWHEG + PYTHIA 8. The results show the behaviour that could be expected by the definition of the `hdamp` parameter.

All in all, it can be concluded that the comparison between data and simulation is not only limited because of the uncertainties caused by the measurement, but also by the modelling uncertainties resulting from different settings of parameters. Careful

7 Conclusions

consideration is essential when applying the results from MC simulation to later data analysis.

Bibliography

- [1] J. Alwall, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07**, 079 (2014)
- [2] T. Sjöstrand, et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015)
- [3] M. Bahr, et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58**, 639 (2008)
- [4] J. Bellm, et al., *Herwig++ 2.7 Release Note* (2013), arXiv:hep-ph/1310.6877
- [5] C. Oleari, *The POWHEG-BOX*, Nucl. Phys. Proc. Suppl. **205-206**, 36 (2010)
- [6] S. Abachi, others (DØ Collaboration), *Observation of the top quark*, Phys. Rev. Lett. **74**, 2632 (1995)
- [7] F. Abe, others (CDF Collaboration), *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74**, 2626 (1995)
- [8] ATLAS, CDF, CMS, DØ Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass* (2014), arXiv:hep-ex/1403.4427
- [9] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716**, 1 (2012)
- [10] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716**, 30 (2012)
- [11] ATLAS and CMS Collaboration, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114**, 191803 (2015)
- [12] B. Mellado Garcia, et al., *CERN Report 4: Part I Standard Model Predictions* (2016), URL <https://cds.cern.ch/record/2150771>

Bibliography

- [13] J. R. Andersen, et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties* (2013), [arXiv:hep-ph/1307.1347](https://arxiv.org/abs/hep-ph/1307.1347)
- [14] R. K. Ellis, W. J. Stirling, B. R. Webber, *QCD and collider physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8**, 1 (1996)
- [15] S. Alioli, et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06**, 043 (2010)
- [16] ATLAS Collaboration, *Modelling of the $t\bar{t}H$ and $t\bar{t}V$ ($V = W, Z$) processes for $\sqrt{s} = 13$ TeV ATLAS analyses*, Technical Report ATL-PHYS-PUB-2016-005, CERN, Geneva (2016)
- [17] R. Brun, F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389**, 81 (1997)

Acknowledgment

First and foremost, I would like to express my sincere gratitude to Prof. Dr. Arnulf Quadt for giving me the chance to work on such an interesting topic at the 2nd Institute of Physics. I really enjoyed the working at his institute, having the opportunity to meet all the wonderful people there and to learn a lot about physics.

I would also like to thank Prof. Dr. Stanley Lai for the stimulating discussions and for taking over the part of the second referee.

My sincere thanks also go to Dr. Maria Moreno Llacer and Jannik Geisen for their assistance and dedication during all the phases of completion of this thesis. I am very grateful for the (almost) 24/7 service they provided for answering all my questions.

Last but not least, I want to thank my friends and my family for their love, encouragement and understanding, especially during last few months.

Erklärung

nach §13(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestanden Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 23. September 2016

(Christina Reißel)