Bachelor's Thesis

$t\bar{t}H$ Produktion am LHC

$t\bar{t}H$ Production at LHC

prepared by

Timo Dreyer

from Aurich

at the II. Physikalischen Institut

Contents

1. Introduction

The first run of the LHC produced an integrated luminosity of $5.46\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 7\,\mathrm{TeV}$ in 2011 and $22.8\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8\,\mathrm{TeV}$ in 2012. In 2015 the LHC will start its second run in which the centre of mass energy will be increased to $\sqrt{s} = 13\text{-}14\,\mathrm{TeV}$. After the discovery of the Higgs boson in 2012, the focus is on examining all of its properties in detail (e.g. to measure its spin, parity, coupling to fermions and bosons, etc.) in the second run.

The top quark plays an important role for this goal, since it is the heaviest fermion and its Yukawa coupling to the Higgs boson is therefore expected to have the highest value of all fermions. However, a decay of the Higgs boson into top quarks ($H \to t\bar{t}$) is kinematically suppressed, due to the fact that the Higgs mass is below the mass of even one top quark.

Therefore, processes in which the Higgs boson is produced in association with two top quarks ($t\bar{t}H$) are exploited to access the Yukawa coupling. Figure 2.2 shows the Feynman diagram for a $t\bar{t}H$ process in which the Higgs boson is produced via fusion of a virtual top quark pair. Alternatively, the Higgs boson can also be radiated from a top quark line.

Given the high masses of the Higgs boson and top quark, these processes are expected to have very low cross-sections in comparison to processes in which lighter particles are produced. For the analysis of processes with such low cross-sections it is of high importance to have a good understanding of the corresponding background. A way to do this is to model the processes with a Monte Carlo event generator like aMC@NLO [1], which is accurate at next to leading order (NLO) in QCD.

The aim of this thesis is to simulate $t\bar{t}H$ events at an energy of $\sqrt{s} = 14\,\mathrm{TeV}$ with the novel aMC@NLO framework for the simulation of the hard process and either Pythia8 or Herwig++ for the parton showering and hadronisation. This simulated data is then used to make a first prediction for the change of the signal to background separation in the second run of the LHC.

2. Theory

The Standard Model of Particle Physics (SM) describes the fundamental particles and the fundamental forces acting on these particles as quantum field theories. The only exception is the gravitational force, which is not integrated into the SM but is also too weak to play a significant role in particle physics.

2.1. Fundamental Particles in the SM

The fundamental particles are classified into groups according to their properties [2, 3] (see Figure 2.1). Particles with a half integer spin are called *fermions*, particles with an integer spin *bosons*. The fermions are further divided into particles that carry colour charge (*quarks*) and particles that do not (*leptons*). All fundamental fermions can be classified in three *generations*, each containing two *flavours* of quarks (a *down-type* quark with a charge of $-\frac{1}{3}$ and a *up-type* quark with an electric charge of $+\frac{2}{3}$) and two leptons (a neutral neutrino and a corresponding lepton with an electric charge of -1). These pairs are described as *isospin doublets* with third components of the isospin of $-\frac{1}{2}$ and $\frac{1}{2}$.

Since each quark can have three different colour charges and for every fermion exists an anti-particle with the opposite charge, the SM actually contains $2 \cdot 3 \cdot 6 + 2 \cdot 6 = 48$ elementary fermions.

The elementary bosons consist of the *gauge bosons*, which mediate the fundamental forces between particles, and the Higgs boson (see Section 2.3). The gauge bosons for the strong force are the eight massless *gluons*, of which each carries a different combination of colour and anti-colour. For the weak force the gauge bosons are the massive Z^0, W^+ and W^- bosons; and for the electromagnetic force it is the photon.

2.2. Interactions between SM Particles

The fundamental forces included in the SM are listed in table 2.1. Due to its massive gauge bosons, Z^0 and $W^\pm$, the weak force has a limited range; while the strong force has a limited range due to the self-coupling of its massless gauge bosons, the gluons.

Figure 2.1.: Particles of the SM.

These forces are described in the SM as *gauge theories*, i.e. the SM consists of quantum field theories in which the Lagrangian is invariant under certain (local) gauge transformations, namely $SU(3) \times SU(2) \times U(1)$.

The generators of each symmetry group correspond to the gauge bosons of the respective force. Therefore, the strong force ($SU(3)$) has eight, the weak force ($SU(2)$) has three and the electromagnetic force ($U(1)$) has one gauge boson [4].

force	gauge boson	mass [GeV] [3]	range [m] [5]
electromagnetic	γ - photon	$< 10^{-27}$	∞
strong	g - gluon	0	$\leq 10^{-15}$
weak	Z^0	91.1876 ± 0.0021	10^{-18}
	$W^{\pm}$	80.385 ± 0.015	

Table 2.1.: Overview of the fundamental forces in the SM.

2.3. Higgs Mechanism

The description of interactions in the SM via gauge invariant Lagrangians leads to a problem with the weak force, because its mediator particles are not massless as a simple gauge theory would suggest [2]. This issue was solved by the Higgs mechanism, which was suggested (among others) by Peter Higgs in 1964 [6, 7].

The Brout–Englert–Higgs mechanism restates the Lagrangian of the weak interaction in such a way that a new ground state is chosen and thereby introduces a spontaneous symmetry-breaking, which in turn allows for a non-zero mass of the weak gauge bosons,

but also leads to a new massive scalar boson, the *Higgs boson.*

One of the main physics goals for the particle accelerator LHC at CERN was the search for the Higgs boson, since it was the last undiscovered particle of the SM until its discovery in 2012 by the ATLAS and CMS experiments [8, 9].

2.4. Higgs Boson and Top Quark Physics

2.4.1. Higgs Boson Production and Decay

At a hadron collider like the LHC all reactions have to start from gluons or quarks. Since the Higgs boson couples to mass, production channels with heavy particles are more likely. Considering this, there are four main channels for the Higgs boson (see Figure 2.2).
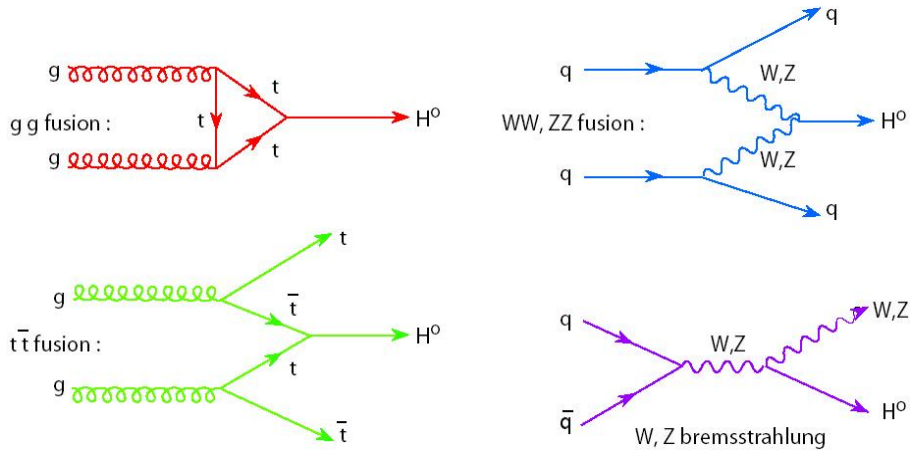

Figure 2.2.: Higgs Boson production channels at the LHC.

The cross-sections of those processes have been computed from theory for different masses of the Higgs boson [10]. With the measured Higgs mass of about 125 GeV [8, 9] this leads to the values given in Table 2.2.

$\sqrt{s}$	gg fusion	WW/ZZ fusion	$W^\pm$ bremstr.	Z^0 bremstr.	$t\bar{t}H$	total
7 TeV	15.31	1.211	0.5729	0.3158	0.08634	17.50
14 TeV	49.85	4.180	1.504	0.8830	0.6113	57.03

Table 2.2.: NLO Cross-sections for Higgs production channels [pb] with uncertainties of up to 19.6 % for gg and $t\bar{t}H$ and 6.4 % for the other channels [10].

The decay of the Higgs boson, like its production, is dominated by its coupling, i.e. in general heavier particles are more likely to be produced than light ones. With its mass of about 125 GeV the heaviest particle anti-particle pair it could decay in is $b\bar{b}$, which is

also the dominant decay channel with a branching ratio of nearly 58 % (see Table 2.3). Following is the WW channel with about 22 %; given that the mass of the Higgs is significantly lower than the mass needed to create two W bosons, roughly 160 GeV, one of the bosons has to be virtual.

$b\bar{b}$	WW	gg	$\tau\tau$	ZZ	$c\bar{c}$	$\gamma\gamma$	$Z\gamma$	$s\bar{s}$	$\mu\mu$
57.8	21.6	8.56	6.37	2.67	2.68	0.230	0.155	0.0440	0.0221

Table 2.3.: Branching ratios for the SM Higgs at $m_H = 125$ GeV in percent with uncertainties of 10 % for gg, 5 % for $t\bar{t}$ and below 2 % for the other channels [10].

2.4.2. Top Quark Decay

The top quark is the heaviest elementary particle with a mass of (174.98 ± 0.76) GeV [11]. In principle it can decay into any other quark plus an electroweak gauge boson, but these processes involve flavour changes of quarks and thus the CKM-matrix elements affect the branching ratios of the different channels. Since $|V_{tb}| \gg |V_{td/ts}|$, the top quark decays predominantly into a bottom quark and a $W^\pm$ boson with a branching ratio of nearly 100 % [3].

The b quark emerging from a top decay will create a jet that can be identified by b-tagging (see Section 3.3.2), while the $W^\pm$ boson can decay into leptons or hadrons. Each lepton channel ($W \to l\,\nu_l$) has a branching ratio of about 11 %, while the hadronic decays have a combined branching ratio of about 68 % [3].

2.4.3. The $t\bar{t}H$ Process

Higgs production processes in association with a $t\bar{t}$ pair (like the $t\bar{t}$ fusion shown in Figure 2.2) are called $t\bar{t}H$ processes. The cross-sections for these events at the LHC are predicted to be small (compare Table 2.2) since a high centre of mass energy is needed to create the three massive particles, i.e. the colliding partons need a high Bjorken x, which becomes unlikely at high collider energies (see Section 3.1).

Despite these difficulties, the $t\bar{t}H$ channel is very interesting since it allows for a direct measurement of the Yukawa coupling of the top quark.

Topology

As discussed before, the top quarks will each decay into a b and a W, while the Higgs will most likely decay into $b\bar{b}$.

To identify an event like this, it has to be differentiated between the different decay channels of the two occurring W bosons, as those will determine the event topology apart from the four direct b-jets from the Higgs and $t\bar{t}$ decay. The different possibilities for a $t\bar{t}$ decay are:

- **dileptonic channel:** Both $W^{\pm}$ bosons decay into a lepton and the corresponding neutrino, which can be seen by two leptons with a high momenta and the missing transverse momentum from the two neutrinos. Usually the decay into τ is *not* included in this channel (see below), therefore this channel occurs in roughly $(2 \times 11\,\%)^2 \approx 5\,\%$ of the cases.

- **lepton+jets channel:** One of the $W^{\pm}$ bosons decays into a lepton (not τ) and the other one hadronically; into a $q\bar{q}'$ pair. This channel appears as one lepton with high momentum and two jets from the two quarks and occurs in about $2\times(22\,\%\times68\,\%) \approx 30\,\%$ of the cases.

- **all-jets channel:** In the most probable channel both $W^{\pm}$ decay hadronically, creating four jets in about $(68\,\%)^2 \approx 45\,\%$ of the cases.

- **τ+X channel:** The remaining $20\,\%$ are decays into τ leptons. Those are excluded from the leptonic channel, since the τ has a very short life time of $(290.6 \pm 1.0) \times 10^{-15}\,\mathrm{s}$ and will decay inside the beam pipe and its high mass of about $1.78\,\mathrm{GeV}$ allows for it to decay into hadrons, e.g. pions, therefore creating narrow jets of its own [3].

Background

The most likely decay for $t\bar{t}H$ is $t\bar{t}H \rightarrow W^+W^-b\bar{b}b\bar{b}$. Due to the high jet background at hadron colliders it is, however, not the easiest to detect. For this reason, the first discovery of the Higgs boson has been made in the channels $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$ and $H \rightarrow WW$ [8, 9]. The most important backgrounds are listed below [12].

- **$t\bar{t}$ +jets:** The production of a $t\bar{t}$ pair with additional jets can be mistaken for a $t\bar{t}H$ event, if the jets are falsely tagged as Higgs decay products. A special case of this background is $t\bar{t}\,b\bar{b}$:

- **$t\bar{t}b\bar{b}$ Production:** The simultaneous production of a $t\bar{t}$ and a $b\bar{b}$ pair is an *irreducible background* to $t\bar{t}H$, since for a single event it is unknown whether the $b\bar{b}$ pair is the product of a Higgs decay or has been produced directly.

- $W^{\pm}$ **+jets:** The production of a W boson with jets is another important background event, especially when the W boson decays leptonically and two of the jets are b-tagged.

- **QCD Multijet events:** Multijet events are problematic, because there is always a chance of jets being falsely b-tagged or mistaken for photons or leptons which may let them occur like $t\bar{t}H$ events.

Process	$\sqrt{s} = 7\text{TeV}$	$\sqrt{s} = 14\text{TeV}$	Order in QCD	
$t\bar{t}H$	$0.08634^{+11.8\%}_{-17.8\%}$	$0.6113^{+14.8\%}_{-18.2\%}$	NLO	[10]
$t\bar{t}$	$172.0^{+3.7\%}_{-4.4\%}$	$953.6^{+3.0\%}_{-4.1\%}$	NNLO+NNLL	[13]

Table 2.4.: NLO cross-sections for $t\bar{t}H$ and $t\bar{t}$ as the main background [fb].

3. Hadron Collider Physics

As the name suggests, hadron colliders collide hadrons, i.e. particles that are made up from two (*mesons*) or three (*baryons*) quarks. Most hadron colliders use protons, which are easy to obtain from hydrogen gas.

3.1. Parton Distribution Functions

Since hadrons are no elementary particles, their internal structure has to be taken into account when conducting experiments with them. This is done with *parton distribution functions* (PDF) and the *Bjorken x*. The Bjorken x describes the fraction of the total momentum of the hadron, which is carried by the parton that interacts in the collision.

Figure 3.1 shows the PDF of a proton at $Q^2 = 100 \, \text{GeV}$, where Q^2 is the transferred momentum in the interaction. Naively one would expect each of the three *valence quarks* (uud) of the proton to carry one third of the total momentum, but due to the exchange of gluons between the valence quarks and quark loops along those gluons (the so called *sea quarks*) other partons do exist and also carry a fraction of the total momentum.

According to the de Broglie equation $\lambda = h/p$, a high momentum p of a particle corresponds to a small wavelength λ (h is Planck's constant). Therefore, high collision energies probe smaller parts of the proton, i.e. at high energies an incoming particle "sees" more partons. Thus, the Bjorken x is typically lower at high energies and the valence quarks become less dominant while the gluons become the most dominant partons.

3.1.1. Transverse Quantities

Due to the PDFs of the colliding protons the total longitudinal momentum of the collision products is unknown. The total transverse momentum, however, has to be zero since the protons collide head on. This can be used to "detect" particles invisible to the detector, e.g. neutrinos. Their momentum can be determined, if the measured total transverse momentum does not add up to zero.
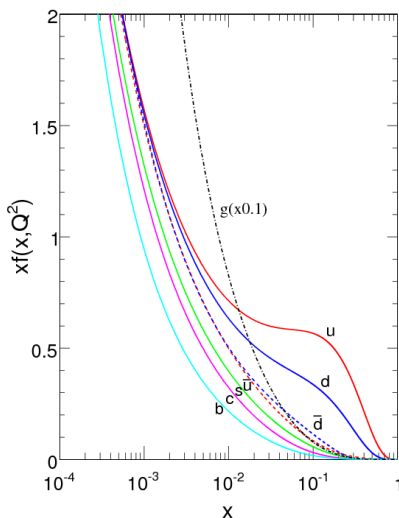
Figure 3.1.: Parton distribution functions at $Q^2=100\,\text{GeV}^2$. The gluon distribution is scaled by a factor of 0.1 [14].

3.1.2. Cross-Sections and Luminosity

A very important quantity for describing interactions of fundamental particles is the *cross-section* σ of the interaction. The geometrical interpretation of the cross-section of an interaction is the area one of the involved particles has to hit for the process to take place.[1] It is usually given in *barn*, where $1\,\text{b} := 10^{-28}\,\text{m}^2$. Since the SM is a quantum theory the cross-section can be interpreted as the probability of a certain interaction.

In combination with information about the total number of interactions $\mathrm{d}N_{tot}$ per total cross-section σ_{tot} per time $\mathrm{d}t$, which is given by the *luminosity* $\mathcal{L} := \frac{1}{\sigma_{tot}}\frac{\mathrm{d}N_{tot}}{\mathrm{d}t}$. The cross-section σ_i gives the expected number of interactions of type i per time:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \sigma_i \cdot \mathcal{L} \tag{3.1}$$

The luminosity $\mathcal{L}$ $[\text{b}^{-1}\text{s}^{-1}]$ as well as the integrated luminosity $\mathcal{L}_{\text{int}} := \int \mathcal{L}(t)\mathrm{d}t$ are characteristics for every specific experimental setup, whereby a higher $\mathcal{L}_{\text{int}}$ corresponds to more produced data.

3.1.3. Rapidity and Azimuth

The symmetry of particle collider experiments suggests a description in a polar coordinate system in which the z-axis is parallel to the beampipe and the origin of the coordinate

---

[1] Therefore, the cross-section of two spheres with radii r_1 and r_2 would be $\pi \cdot (r_1 + r_2)^2$.

system lies in the collision point. The polar angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and the azimuth angle $\phi \in [0, 2\pi]$ follow the usual definition.

However, since the total longitudinal momentum in an event is unknown (see Section 3.1) it is useful to change from the polar angle θ to the *(pseudo-)rapidity* $\eta \in (-\infty, \infty)$:

$$\eta := -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{3.2}$$

This definition has the advantage that rapidity differences $\Delta\eta$ are Lorentz-invariant.

3.2. The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN (European Council for Nuclear Research) is a symmetric proton proton collider [15]. In two beam pipes protons are accelerated in opposite directions up to a design energy of $7\,\text{TeV}$ per beam, i.e. a centre of mass energy of $\sqrt{s} = 14\,\text{TeV}$.

3.3. The ATLAS Detector

The ATLAS detector at the LHC consists of various subdetector systems arranged around a collision point of the LHC accelerator (see figure 3.2). So far the ATLAS experiment has collected 5.46 fb^{-1} of data at $\sqrt{s} = 7\,\text{TeV}$ in 2011 and 22.8 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$ in 2012.

The inner part of ATLAS consists of tracking detectors for charged particles. Surrounding this inner detector are the electromagnetic and the hadronic calorimeters, followed by the muon chambers.

The inner tracking devices are contained in a solenoidal magnetic field, which runs parallel to the beam pipe, thus creating a Lorentz force that is causing charged particles to travel on trajectories which are curved in the plane transverse to the beam pipe. For the muon chambers there is another magnetic field bending the muons in a plane parallel to the beam pipe.

3.3.1. Signal Identification

The reconstruction of a signal from a particle detector and the identification of the event behind it is a complex process [16]. The basic principle, however, is to take a look at the part of the detector in which a passing particle leaves a track.
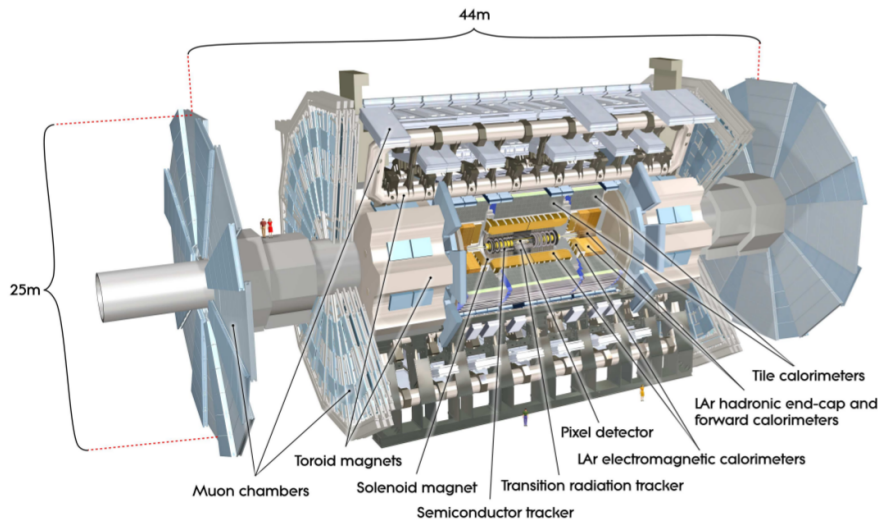
Figure 3.2.: The ATLAS detector.

The tracking devices are only sensitive to ionising, i.e. charged, particles, while neutral particles will pass them without a signal. Combined with the magnetic field, the tracking detectors grand information about the charge and momentum of a passing particle.

Outside the tracking devices are two layers of calorimeters. The inner and thinner one is the electromagnetic calorimeter, in which most electrons and photons will shower and deposit their whole energy. The outer one is the hadronic calorimeter, which is thicker because hadrons have a different *stopping power* $\frac{\mathrm{d}E}{\mathrm{d}x}$ and need more material to be stopped completely.[2]

The muon chambers in the outer layers of ATLAS are basically another layer of tracking devices, but since they are shielded by the calorimeters only muons will reach them and leave a signal.

3.3.2. b-Tagging

Bottom quarks have a relative long life time compared to other heavier particles that can be created in the LHC and therefore tend to fly a few mm away from the primary vertex, where they were created, before decaying at a secondary vertex. By interpolating the tracks of the two created jets backwards, it is possible to detect the secondary vertex, which is used for the identification of created b quarks.

---

[2]It is important to point out that hadrons will also produce showers in the electromagnetic calorimeter and thus deposit energy there as well.

4. Monte Carlo Event Generation in Particle Physics

Monte Carlo event generation is a way to simulate the final states of a chosen particle physics event. A Monte Carlo generator will use pseudorandom number generation to randomly produce an usually large number of events of the desired type. The point of the simulation is that the probability distributions of the events' different observables should be the same as observed in a real experiment.

This kind of simulation is useful for the purpose of testing and calibrating theories like the SM by implementing the parameters of the theory in the simulation and comparing the results of the simulation with data from experiments.

4.1. Basic Principle of Monte Carlo Simulations

In hadron colliders the initial interaction of an event happens between partons of the colliding hadrons, since latter are no fundamental particles. Therefore, the (differential) cross-section $\mathrm{d}\sigma$ of an event $h+h \to X$ (where $h+h$ are the colliding hadrons) is composed of the sum over the (differential) cross-sections $\mathrm{d}\hat{\sigma}_{ab \to X}$ for all possible combinations of colliding partons a and b:

$$\mathrm{d}\sigma = \sum_{a,b} \int \mathrm{d}x_1 \mathrm{d}x_2 \quad f_a(x_1, \mu_F)f_b(x_1, \mu_F) \quad \mathrm{d}\hat{\sigma}_{ab \to X}(\hat{s}, \mu_F, \mu_R) \tag{4.1}$$

where μ_F is the factorisation energy scale, μ_R the renormalisation energy scale and $f_i(x_j, \mu_F)$ are the PDFs for the two colliding hadrons. The PDFs have to be integrated over the available phase space via the Bjorken x of the parton x_j.

The cross-sections $\mathrm{d}\hat{\sigma}_{ab \to X}$ can be calculated by perturbation theory and the integration is done via the Monte Carlo method. The PDFs, on the other hand, have to be inserted in the simulation as an external parameter and are usually obtained by a fit to different measured cross-sections and evolving to higher energy scales via the DGLAP equations.

4.1.1. The Monte Carlo Integration Method

To calculate the cross-section for a given process, it is, in principle, necessary to integrate the probability density over the whole phase space of the process. However, since the phase space for a typical event is very high, this can be a very difficult task.

Instead of directly calculating the value of the integral, the Monte Carlo method chooses random points from the available phase space and evaluates the function on those points. The Integral can than be approximated from the average value of the function at all those points. Consider for example the integral of a function $f(x)$ over the n-dimensional volume V:

$$\int_V f(x)\mathrm{d}^n x \ \approx \ < f >_V \cdot \int_V \mathrm{d}^n x \ \approx \ \frac{1}{N} \sum_{i=1}^{N} f(x_i) \cdot \int_V \mathrm{d}^n x \qquad (4.2)$$

where $< \cdot >_V$ is the average over the volume V. Since the volume over which one needs to integrate is the phase space of the event, each point x_i in this volume corresponds to a certain four-momentum configuration of the particles emerging from that event. The individual points x_i can therefore be used to get a sample with known four-momenta for each particle in each event and are stored for kinematic analysis.

4.1.2. Order of the Calculation

The determination of the parton-level cross-sections $\mathrm{d}\hat{\sigma}_{ab \to X}$ is a process and model dependent QCD calculation. Using the Feynman-calculus, in principle, all possible Feynman-diagrams for the desired process have to be considered. However, the more complex the diagrams become, the more vertices they have, and each vertex in a Feynman-diagram leads to an additional occurrence of the coupling constant of the corresponding interaction in the calculation. In hadron collider physics this is mainly α_S, the coupling constant of QCD.

The (infinite number of) different diagrams can, therefore, be sorted by the number of QCD vertices, i.e. by the order in α_S. Because $\alpha_S \ll 1$, the higher orders can be neglected to get an approximation which is known as *perturbative QCD*. Depending on how many additional vertices are taken into account, the calculation is called a leading-order (LO), next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO), etc. calculation in QCD.

4.1.3. The Energy Scales μ_F and μ_R

The coupling constant α_S of QCD is a *running coupling constant*, i.e. its value depends on the energy scale at which a process happens. Therefore, it is necessary to determine which energy scale is used. Two different scales need to be defined (compare Equation 4.1):

The *factorisation scale μ_F* determines at which energy scale the PDFs are evaluated, while the *renormalisation scale μ_R* sets the energy, at which the coupling constant is calculated.

4.2. Hard Process Generation

The first step to generate a Monte Carlo event is to simulate the *hard process*, which is the fundamental process. In the context of the LHC this is a process of the type $pp \rightarrow X$, where X is the "direct" product of the initial proton collision, in which the PDFs of the protons have to be taken into account, since the protons are no fundamental particles.

In this thesis, the aMC@NLO [1] generator is used to produce the hard processes.

The standard way to store Monte Carlo events are the *Les Houches Files* (LHF), which is a standard format for the files used between the different states of the Monte Carlo event generation [17]. The LHF consist of a file header, containing information about the simulation parameters, followed by a list of all the produced events, in which every single particle is listed with (among others) its PDGid (a code used to identify the type of the particle, see [3]), four-momentum, information about mother- and child-particles and the proper lifetime.

4.3. Parton Showering and Hadronisation

The products X of the hard process, in most cases, are not the particles a detector would measure. Instead they will decay into other particles and interact with remaining partons from the initial hadrons and gluons that are radiated from them. This process will produce a lot of particles and is known as *parton showering* (PS). The initial energy of the collision will be distributed between the new particles which are moving away form the collision point and therefore from each other. Since α_S is a running coupling constant, i.e. it gets higher at low energies, this process will inevitably leave the energy region in which perturbative QCD holds. The non-perturbative QCD process is known as *hadronisation*.

In simulations, these two processes are described by phenomenological models in frameworks like PYTHIA and HERWIG.

(a) String model.   (b) Cluster model.
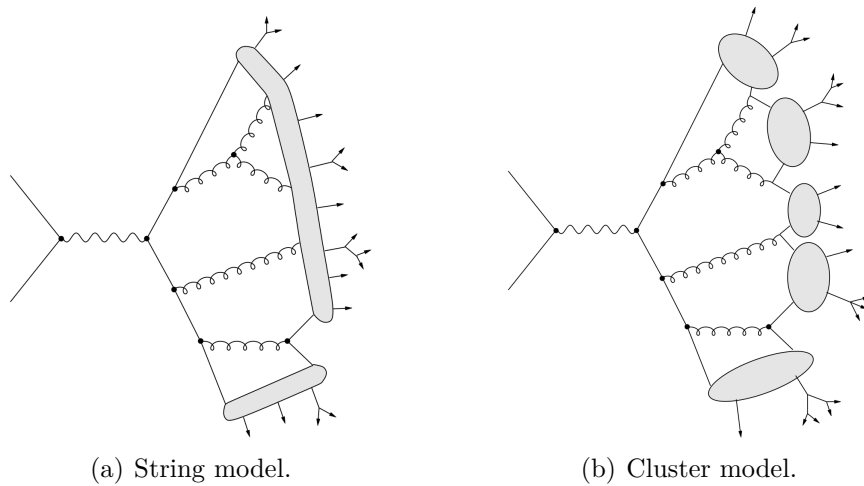
Figure 4.1.: Schematic explanation of the two hadronisation models.

4.3.1. PYTHIA8

PYTHIA8 [18] is a framework that uses the string model for hadronisation.

String Model

A fundamental property of QCD is the so called *confinement*, which states that every stable particle needs to be colourless, i.e. if it is composed of particles that carry colour charge, the charges of all its constituents need to cancel each other out. Therefore, quarks or gluons that arise from a process need to appear at least in pairs with opposite colour charge.

The string model describes these pairs of quarks as connected by a colour string. Gluons can be attached in the middle of the string and create a kink in it (see Figure 4.1(a)). Since the coupling constant α_S of the strong force is a running coupling constant, the binding energy that is stored in the string becomes larger as the string becomes longer, i.e. when the quarks at the end of the string are moving apart from each other. Once the binding energy is large enough, the string can "rip" into two new strings by creating a quark anti-quark pair, which are the new ends of the ripped string.

This way, each new quark will share a string with one of the old quarks and each string will again be colourless. This process of stretching the strings until they rip into two new ones continues as long as the quarks still have enough kinetic energy to stretch the strings. When all kinetic energy is used up, the quarks will form colourless hadrons.

4.3.2. HERWIG++

HERWIG++ [19] uses the cluster model for hadronisation.

Cluster Model

Instead of binding the particles into colourless pairs like in the string model, the cluster model clusters the particles in colourless groups. Gluons, however, can not be attached to these groups but are forced to split into quark anti-quark pairs which then can join the clusters (compare Figures 4.1(a) and 4.1(b)).

4.3.3. MADSPIN

Within aMC@NLO it is possible to already include the decay of the top quarks with MADSPIN [20, 21], which allows for the preservation of spin correlations.

4.4. Detector Simulation

To directly compare Monte Carlo data with experimental data, the additional effects of the specific detector have to be taken into account. The showered events therefore have to be processed by a detector simulation, which simulates the detector components.

5. Sample Production

One of the advantages of Monte Carlo data is, that exact information about the particles that occur in the event and the overall progress of the event is available, while in real experimental data the information about the particles needs to be reconstructed from the measurement and this interpretation underlies certain uncertainties.

With Monte Carlo data the choice exists to reconstruct the particles and processes from the observables, as it would be done in experimental data, or to use the additional information about what particles with which properties the event really has. The latter is therefore known as a study at "truth level".

The aim of this thesis is to study the $t\bar{t}H$ process at this truth level. For this purpose, the $t\bar{t}H$ process has been modelled at NLO in QCD using the aMC@NLO framework for the hard process. In this framework, the decay of the top quark can be included using MADSPIN. For the hadronisation of the samples PYTHIA8 and HERWIG++ were used. The simulated samples have not been processed through a simulation of the detector geometry, because this is very CPU-intensive and not necessary for a truth level study.

5.1. Hard Process Simulation with aMC@NLO

For the simulation of the hard process aMC@NLO was used [1]. aMC@NLO is a framework that combines different tools to provide a fully automated process for creating hard events with a single interface. It evaluates one-loop contributions via MADLOOP [22], which uses the OPP method[1] that is implemented in CUTTOOLS, and MADFKS [24] for the other matrix-element contribution.

It is also possible to use MADSPIN [20, 21] to partially decay the particles of the event before a parton shower is used. The MC@NLO method [25] is then used to match the results to different parton shower models.

For this thesis four different datasets, each containing one million events, have been created with aMC@NLO using dynamic factorisation μ_F and renormalisation μ_R scales which are set to the sum of the transverse mass $m_T(i)$ over all final state particles and

---

[1]Named after Ossola, Papadopoulos and Pittau, see [23].

partons i divided by two. All datasets are inclusive in the top quark and Higgs boson decays. The scales and other important parameters are listed in Table 5.1.

Event	$pp \rightarrow t\bar{t}H$
Center of Mass Energy	14 TeV
Number of events	1,000,000
Order in QED/QCD	LO/NLO
PDF	CT10(NLO) [26]
Scales	$\mu_F = \mu_R = \frac{1}{2} \sum_i m_T(i)$ (dynamic)
Higgs boson mass	125 GeV
Top quark mass	172.5 GeV
Bottom quark mas s	4.75 GeV

Table 5.1.: Parameters used for the hard process generation with aMC@NLO.

An example for a used run card can be seen in Section A.3. Additional to these settings, the datasets have been created with matching to either PYTHIA8 or HERWIG++. For each matching two datasets were produced: One using MADSPIN to decay the top and antitop quarks before the matching and one matching directly to the parton shower model.

The four produced datasets are available as Les Houches files on the ATLAS GRID under the names:

```
user.tdreyer.tth_aMCatNLOPythia8_14TeV.TXT.mc14_v2
user.tdreyer.tth_aMCatNLOPythia8_14TeV_madspin.TXT.mc14_v2
user.tdreyer.tth_aMCatNLOHerwigpp_14TeV.TXT.mc14_v2
user.tdreyer.tth_aMCatNLOHerwigpp_14TeV_madspin.TXT.mc14_v2
```

5.2. Parton Showering with PYTHIA8 and HERWIG++

Two different parton shower models were used with the produced $t\bar{t}H$ samples: PYTHIA8 [18] and HERWIG++ [19].

The showering with PYTHIA8 uses the CTEQ6L1 PDFset [26] and the AU2 underlying event tune [18], while the showering with HERWIG++ uses the CT10f4 PDF for hard scattering [26] and CTEQ6L1 with the EE4 tune for showers and multiple partonic scattering.

While the hard process samples produced above are inclusive in the $t\bar{t}$ and Higgs boson decay mode, a cut on $t\bar{t} \rightarrow \ell + jets$ is performed on this level for the samples interfaced with PYTHIA8.

The produced showered datasets are available as root [27] files (with the NTUP_TRUTH structure) on the ATLAS GRID under the names:

```
user.tdreyer.tth_aMCatNLOPythia8_14TeV.NTUP_TRUTH.mc14_v1.140625144647.merge
user.tdreyer.tth_aMCatNLOPythia8_14TeV_madspin.NTUP_TRUTH.mc14_v1.140625145025.merge
user.tdreyer.tth_aMCatNLOHerwigpp_14TeV.NTUP_TRUTH.mc14_v1.140704000045.merge
user.tdreyer.tth_aMCatNLOHerwigpp_14TeV_madspin.NTUP_TRUTH.mc14_v1.140703235828.merge
```

5.3. The Truth Record

The created NTUP_TRUTH files are processed with a selection code that uses the ROOT framework [27] In the Monte Carlo truth record each particle can appear several times (see Figure 5.1). In the selection process various different observables are studied. These include kinematic variables of the particles in the event, information about jets, which are reconstructed using the anti-kt algorithm [28] with a radius parameter of $R = 0.4$.

For variables that are concerning the top, antitop or Higgs, an option is available in the code to select the particle properties before (*hard process*, HP) or after the showering (AS). This is possible, since the parton shower models create a decay tree for every event, in which each particle can appear multiple times (i.e. the particle "decays into itself").

Each particle in this process has a unique number assigned to it for its identification. The additional information about the particle includes its PDGid [3], to identify the type of particle, information about the mother and child particles and a status code that is set by the parton showering model.

Figure 5.1 shows examples for $t\bar{t}H$ events after the parton showering with the status codes of the individual particles. For Pythia8, the codes contain information about the state of the particles. The selection code was therefore using information from the top quarks and Higgs boson with status codes between 21 and 29[2] when the HP option was chosen and the same particles with status codes between 61 and 69[3] when the AS option was chosen.

When showering with Herwig++ (almost) all status codes are set to 11 and therefore no selection regarding these codes is possible. For that reason the selection for Herwig++ has been changed in such a way that it uses the information about mother and child particles. If the HP option is chosen, the particle whose mothers PDGid does not equal its own is selected. I.e. the code uses the particle that does not originate from a decay into itself and is therefore the first particle of this type in the decay chain.

Correspondingly the particle whose child PDGid does not match its own is selected, if the AS option is chosen.

---

[2]Encoding "particles of the hardest subprocess".

[3]Encoding "particles produced by beam-remnant treatment".
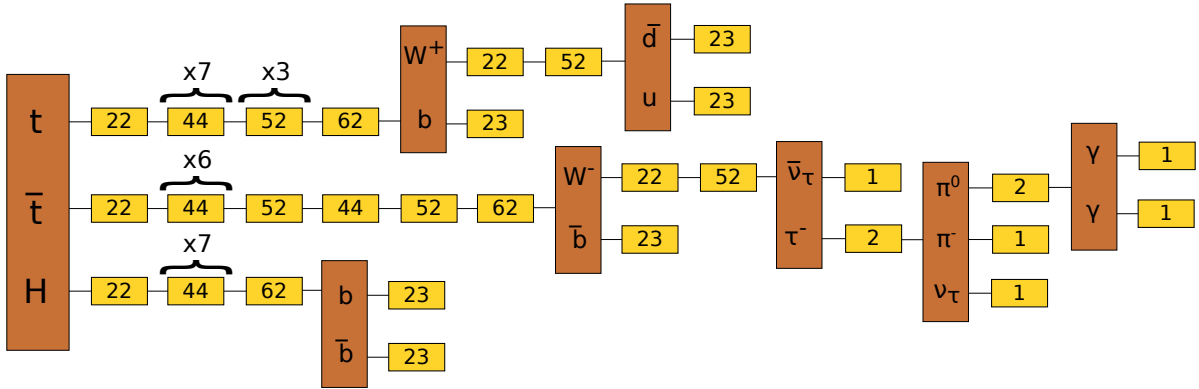
(a) Event record interfaced with PYTHIA8.
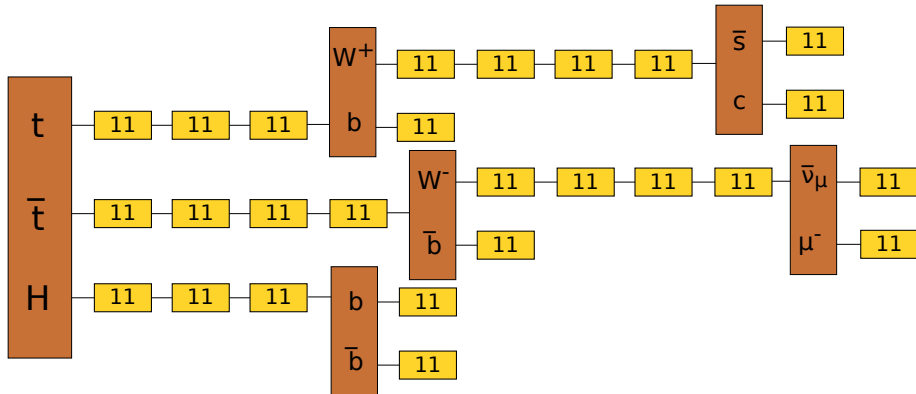
(b) Event record interfaced with HERWIG++.

Figure 5.1.: Examples for status codes in $t\bar{t}H$ events in different parton shower and hadronisation models.

6. Results

The produced Monte Carlo data samples are used to perform the following comparisons:

- The influence of the parton shower model: PYTHIA8 and HERWIG++.

- The top quark is decayed with MADSPIN before applying the parton showering or directly within PYTHIA8 or HERWIG++.

- The separation of the $t\bar{t}H$ signal from the $t\bar{t}$+jets background in different variables.

6.1. Definition of Variables

The transverse momentum p_T and the pseudorapidity η of the top quark, the Higgs boson, the bottom quark and W^+ boson from the top decay, and the ℓ^+ from the leptonic decay of the W^+ boson, as well as from the $t\bar{t}$ system are used for the studies in this thesis. Additional to these kinematic observables, the following other variables were studied:

6.1.1. Jet-related Variables

Only jets with a p_T that is higher than $25\,\mathrm{GeV}$ and an absolute value of the pseudorapidity $|\eta|$ below 2.5 are selected. The number of jets in an event that satisfy those selection criteria is noted as $N_{jets}(p_T > 25\,\mathrm{GeV})$ and their combined transverse momentum p_T as $H_T^{jets}(p_T > 25\,\mathrm{GeV})$.

$\Delta R(jj)$ is a measurement for the difference between two jets and is defined via $(\Delta R)^2 := (\Delta\phi)^2 + (\Delta\eta)^2$. $m(jj)$ and $p_T(jj)$ are the invariant mass and the combined transverse momentum of a dijet system. The two jets of the event that are chosen as this dijet systems are determined in two ways: either the pair of jets with the maximal transverse momentum is used, or the pair with the minimal distance ΔR.

6.1.2. Event Shape Variables

For the analysis of the shape of an event it is useful to define the *sphericity tensor* of the event as follows:

6. Results

$$S^{\alpha\beta} := \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2} \qquad (6.1)$$

where p_i^α is the α component of the ith four-momentum, and only the spatial components are taken into account ($\alpha, \beta = 1, 2, 3$).

The eigenvalues of this tensor satisfy the relation $\lambda_1 + \lambda_2 + \lambda_3 = 1$. If they are named in the way $\lambda_1 \geq \lambda_2 \geq \lambda_3$, the event shape variables *sphericity* $S := \frac{3}{2}(\lambda_2 + \lambda_3) \in [0, 1]$, *aplanarity* $A := \frac{3}{2}\lambda_3 \in [0, \frac{1}{2}]$ and *circularity* $C := 2\frac{\min(\lambda_1, \lambda_2)}{\lambda_1 + \lambda_2}$ can be defined. Another useful event shape variable is the *centrality* C':

$$C' := \frac{\sum p_T}{\sum E_{vis}} \qquad (6.2)$$

where E_{vis} is the *visible Energy* of the jets, i.e. the energy without considering neutrinos and leptons.

6.2. Differences between PYTHIA8 and HERWIG++

The first comparison compares the two different hadronisation models PYTHIA8 and HER-WIG++. The selection of the particles is performed as described in Section 5.3.

6.2.1. Comparison in the Hard Process

To make sure the observed differences of these samples really emerge from the showering in those models and are not caused by the matching that is done in aMC@NLO, the kinematic variables of the $t\bar{t}H$ system are compared in the hard process, i.e. before the showering. A hard process comparison in other variables is not usable in this case, due to the fact that the top quarks and the Higgs boson are the only particles, that are present at the hard process level.

The transverse momenta and the pseudorapidity of the top quark, the $t\bar{t}$ system and the Higgs boson under this condition are shown in Figure 6.1. For the top quark and the Higgs boson no statistically significant difference between the two samples can be observed. In the p_T of the $t\bar{t}$ system, a linear slope in the ratio between the two samples is visible, that tends towards higher transverse momenta for HERWIG++. Accordingly, the corresponding plot of the η of the system shows that the HERWIG++ sample has a higher abundance of low pseudorapidities. This behaviour could arise from the matching which is done by aMC@NLO to prepare the sample for the interfacing with PYTHIA8 or HERWIG++.

(a) Transverse momentum p_T of the top quark.

(b) Pseudorapidity η of the top quark.

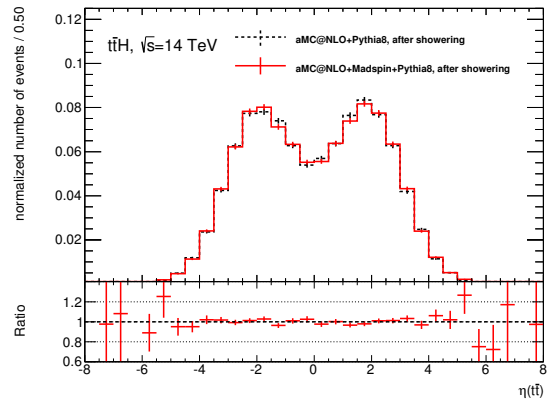(c) Transverse momentum p_T of the $t\bar{t}$ system.

(d) Pseudorapidity η of the $t\bar{t}$ system.

(e) Transverse momentum p_T of the Higgs boson.

(f) Pseudorapidity η of the Higgs boson.

Figure 6.1.: Kinematic distributions for top quark, $t\bar{t}$ system and Higgs boson from the hard process.

6.2.2. Comparison After Showering in Kinematic Variables

To probe the real effect of the parton shower model choice, the datasets need to be compared in variables that are evaluated *after* the showering. Figures 6.2 to 6.4 show the kinematic comparison of the samples with the selection of top quarks and Higgs boson after showering and hadronisation (see Section 5.3).

Contrary to the previous comparison before showering, after the showering the HER-WIG++ sample shows a general tendency towards lower p_T and higher η. This behaviour can be seen most clearly in the $t\bar{t}$ system (Figures 6.2(a) and 6.2(b)) and the $t\bar{t}H$ system (Figures 6.3(a) and 6.3(b)). Phrased differently, the PYTHIA8 model seems to tend to more "central" $t\bar{t}H$ events than HERWIG++.

However, the difference is less significant for the decay products of the initial particles, as can be seen by comparing the decay chain from the top quark (Figures 6.2(a) and 6.2(b)), over the W^+ boson (Figures 6.3(e) and 6.3(f)) to ℓ^+ (Figures 6.4(a) and 6.4(b)), for which the differences in the two models are indistinguishable from statistical fluctuations.
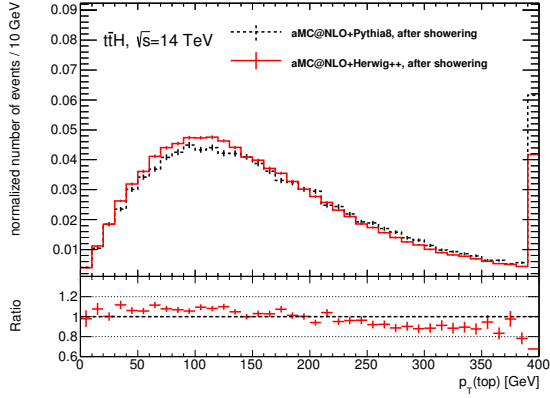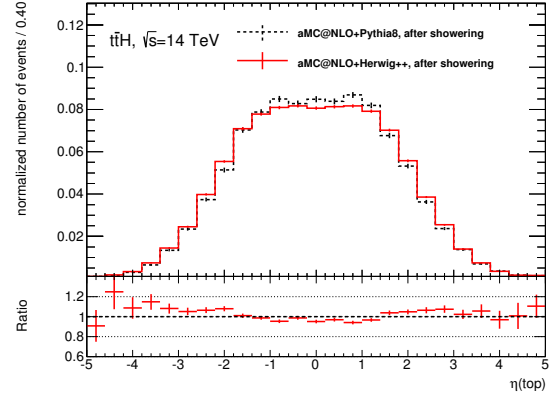
(a) Transverse momentum p_T of the top quark.

(b) Pseudorapidity η of the top quark.
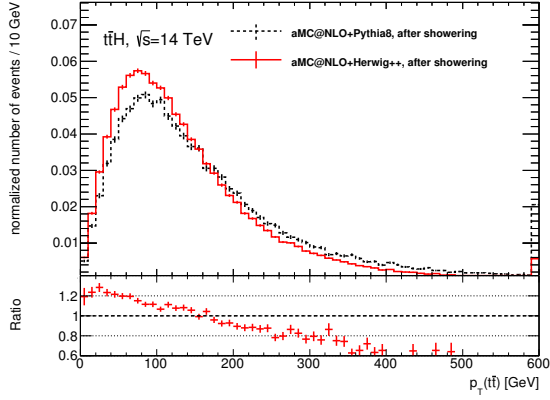
(c) Transverse momentum p_T of the $t\bar{t}$ system.
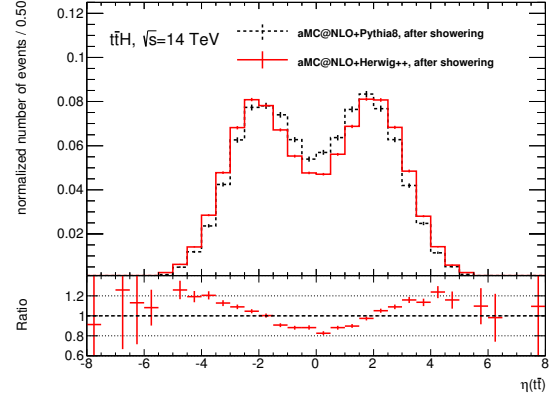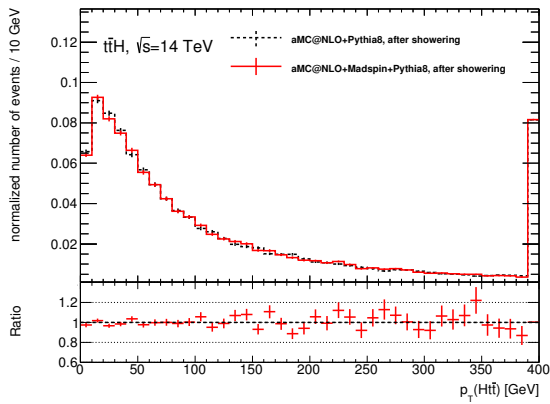
(d) Pseudorapidity η of the $t\bar{t}$ system.
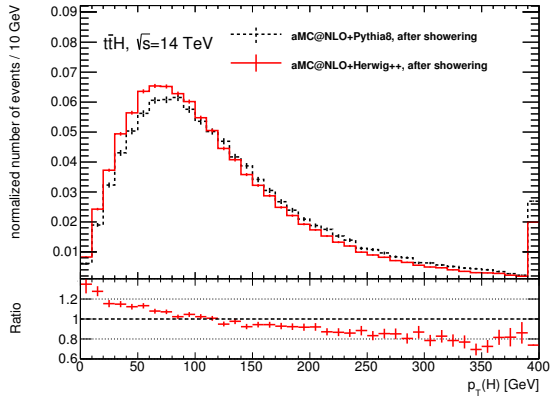
(e) Transverse momentum p_T of the Higgs boson.
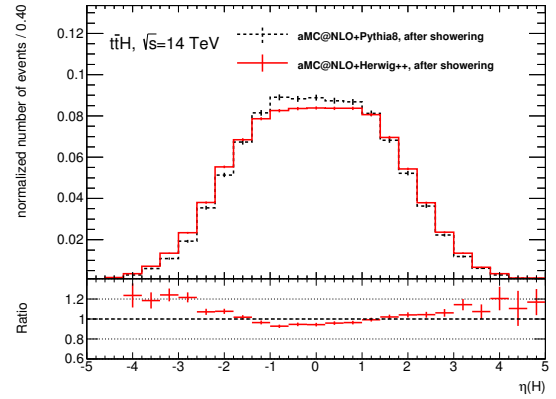
(f) Pseudorapidity η of the Higgs boson.

Figure 6.2.: Kinematic distributions for the top quark, $t\bar{t}$ system and Higgs boson after showering.
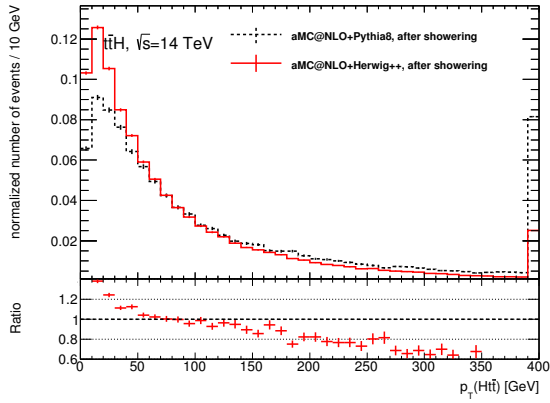
(a) Transverse momentum of the $t\bar{t}H$ system.
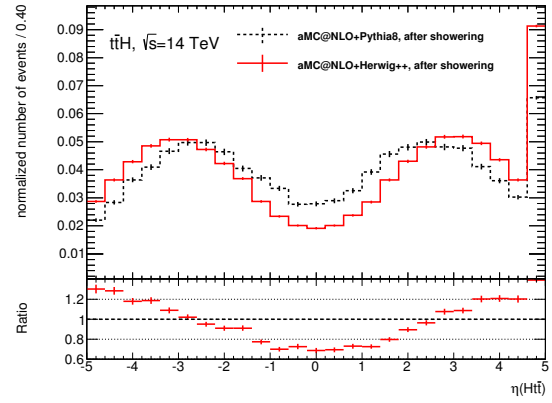
(b) Pseudorapidity η of the $t\bar{t}H$ system.
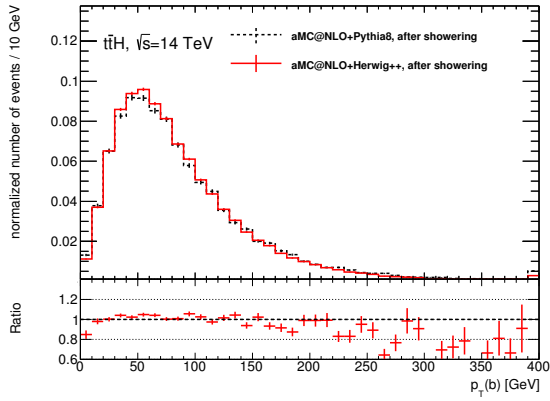
(c) Transverse momentum of the bottom quark.
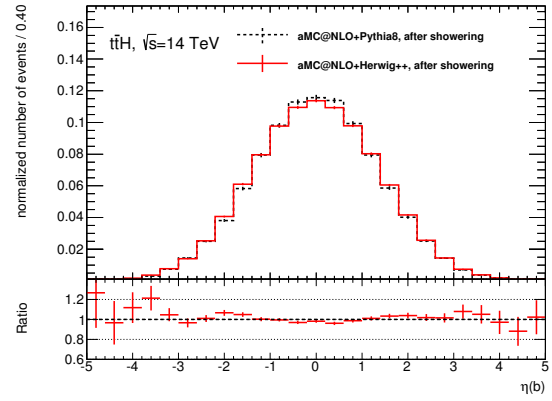
(d) Pseudorapidity η of the bottom quark.
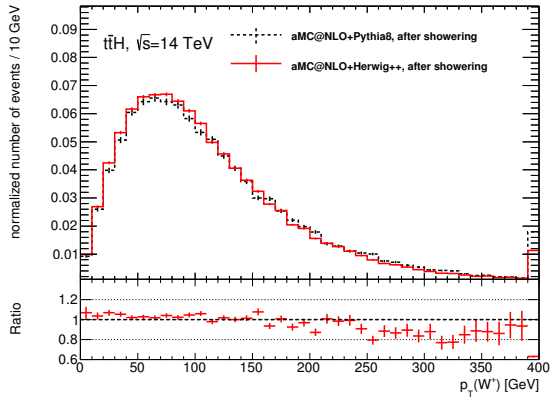
(e) Transverse momentum of the W^+ boson.
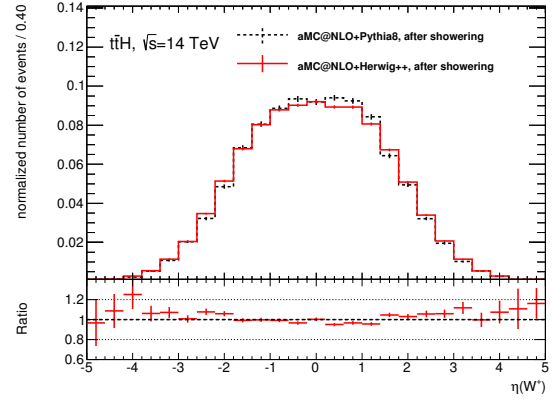
(f) Pseudorapidity η of the W^+ boson.

Figure 6.3.: Kinematic distributions for the $t\bar{t}H$ system, bottom quark and W^+ boson after showering.
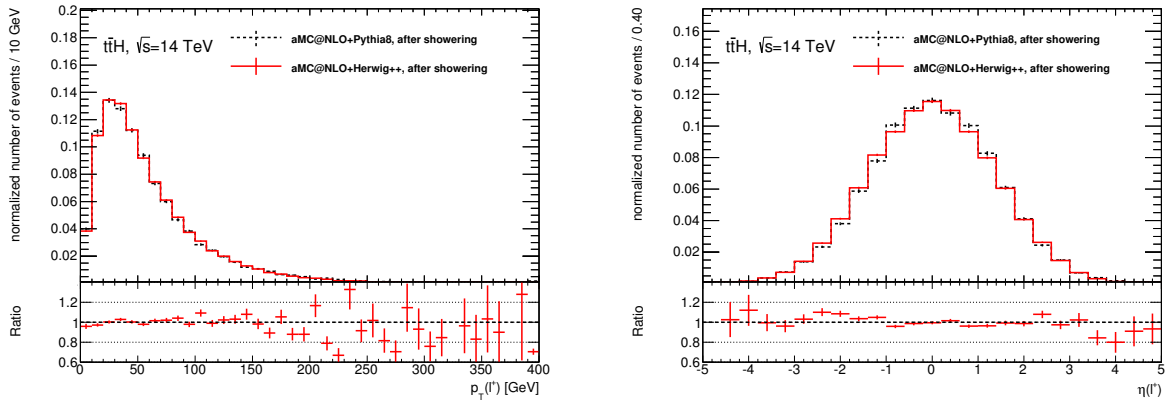
(a) Transverse momentum of the positively charged lepton.

(b) Pseudorapidity η of the positively charged lepton.

Figure 6.4.: Kinematic distributions for the charged lepton after showering.

6.2.3. Comparison After Showering in Jet-related Variables

The number of jets N_{jet} (Figure 6.5(a)) is slightly shifted towards lower jet numbers for HERWIG++, but both models show a clear peak at $N_{jet} = 6$, which is the expected value for $t\bar{t}H$ events given that a selection to $t\bar{t} \to \ell + jets$ and $H \to b\bar{b}$ is applied.

The p_T trend discussed before can also be seen in the transverse momentum H_T^{jets} of all selected jets (Figure 6.5(b)).

In the distribution of $\Delta R(jj)_{\max p_T}$ (Figure 6.5(c)) shows no significant difference between the two shower models, while the distribution of $\Delta R(jj)_{\min \Delta R}$ (Figure 6.5(d)) shows a trend towards higher distances between the jets for HERWIG++.

For the $m(jj)_{\max p_T}$ and $m(jj)_{\min \Delta R}$ distribution (Figures 6.5(e) and 6.5(e)) HERWIG++ is shifted towards lower masses when compared to PYTHIA8.
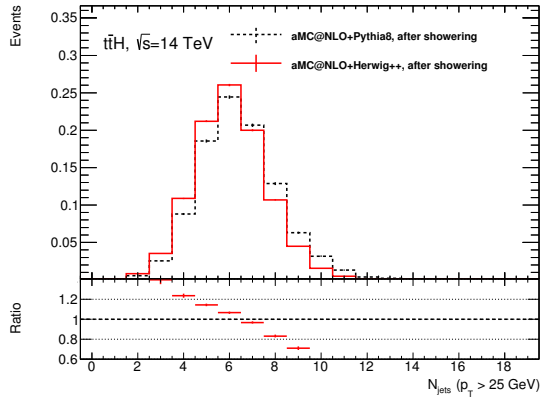
(a) Number of jets with $p_T > 25\,\text{GeV}$ and $|\eta| <$ 2.5.
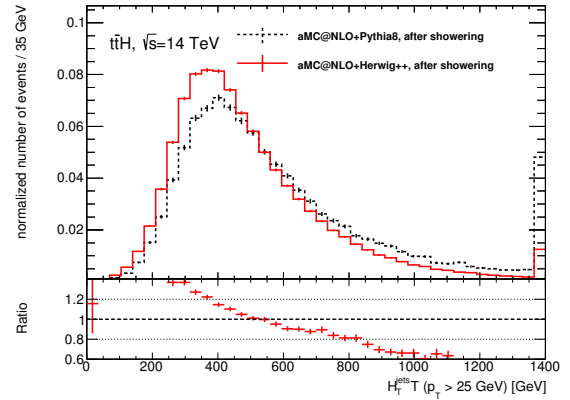
(b) Total transverse momentum of all jets with $p_T > 25\,\text{GeV}$ and $|\eta| < 2.5$.
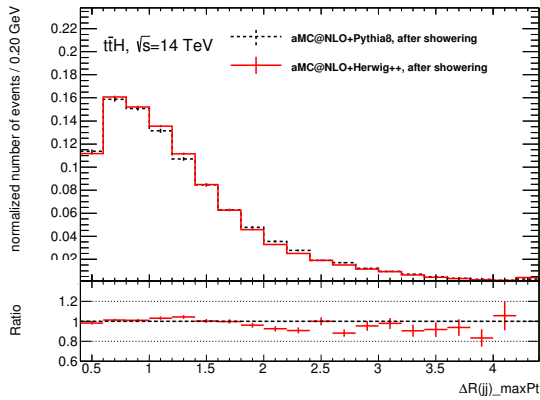
(c) ΔR of the dijet system with the maximal combined p_T.
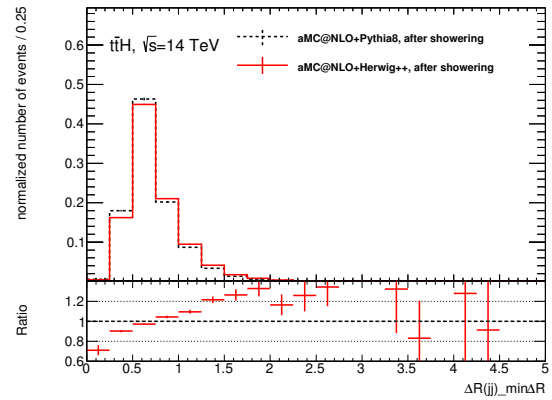
(d) ΔR of the two jets with the smallest distance between them.
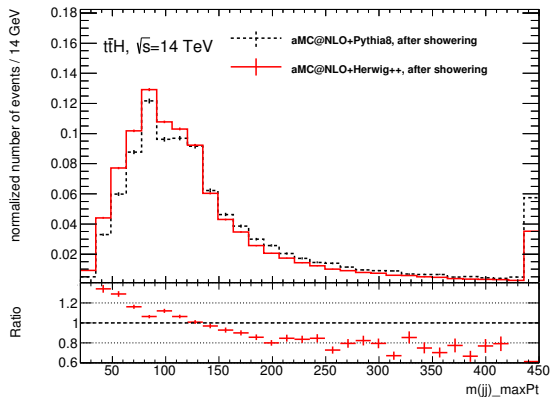
(e) Invariant mass of the dijet system with the maximal combined p_T.
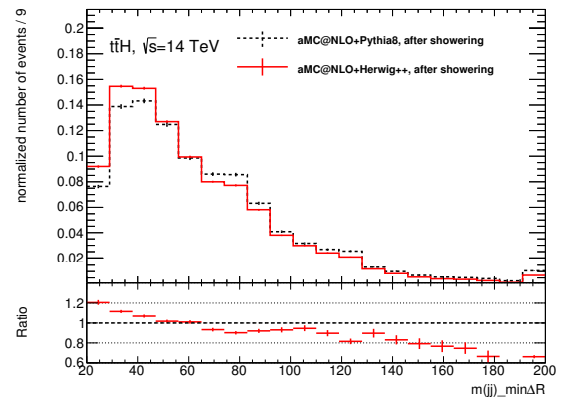
(f) Invariant mass of the two jets with the smallest distance between them.

Figure 6.5.: Comparison of PYTHIA8 and HERWIG++ in jet-related variables after showering.

6.3. Influence of MADSPIN

To determine the effect of MADSPIN on the kinematics of the event, aMC@NLO samples interfaced with the same parton shower and hadronisation model are compared.[1]

A comparison in kinematic variables for the samples showered with PYTHIA8 is shown in the Figures 6.6 and 6.7. No significant difference in the distributions of the samples with and without MADSPIN can be seen.

Since the main argument for MADSPIN is its handling of spin correlations, other variables that are more sensitive to spin effects need to be used for further studies on this topic.
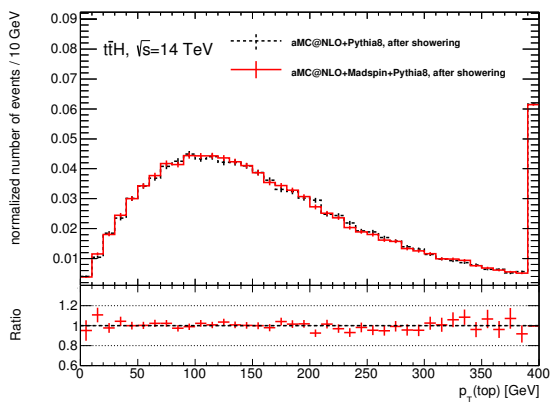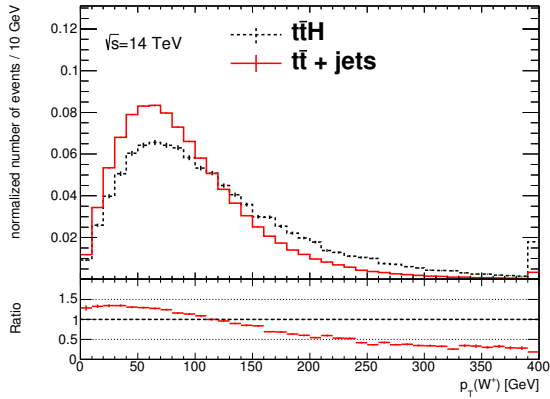
(a) Transverse momentum p_T of the top quark.
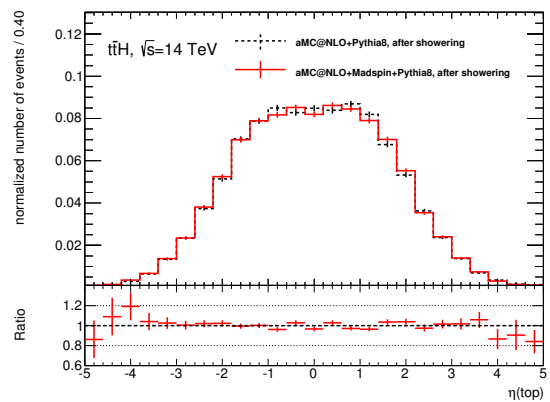
(b) Pseudorapidity η of the top quark.
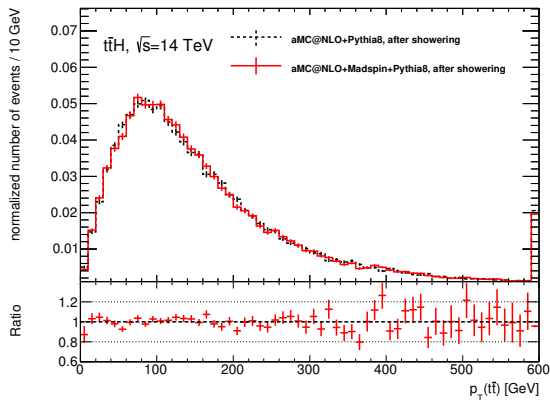
(c) Transverse momentum p_T of the $t\bar{t}$ system.

(d) Pseudorapidity η of the $t\bar{t}$ system.

Figure 6.6.: Comparison of aMC@NLO + PYTHIA8 with or without MADSPIN in kinematic variables of the top quark and the $t\bar{t}$ system after the showering.

---

[1] Since the resulting plots look very similar for PYTHIA8 and HERWIG++, only the ones using PYTHIA8 are presented here, while the ones with HERWIG++ can be found in Section A.4 of the Appendix.
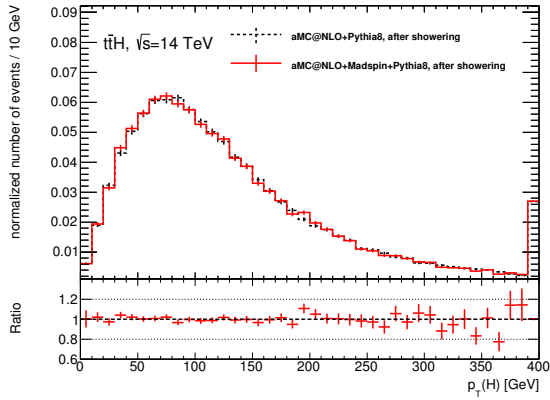
(a) Transverse momentum p_T of the Higgs boson.
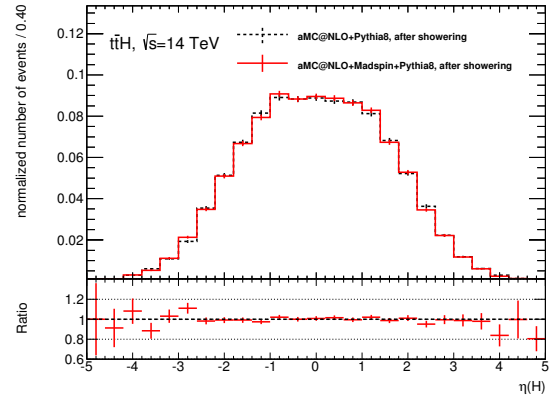
(b) Pseudorapidity η of the Higgs boson.

(c) Transverse momentum p_T of the $t\bar{t}H$ system.
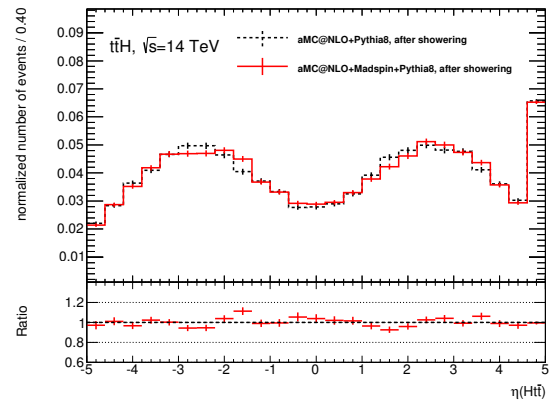
(d) Pseudorapidity η of the $t\bar{t}H$ system.
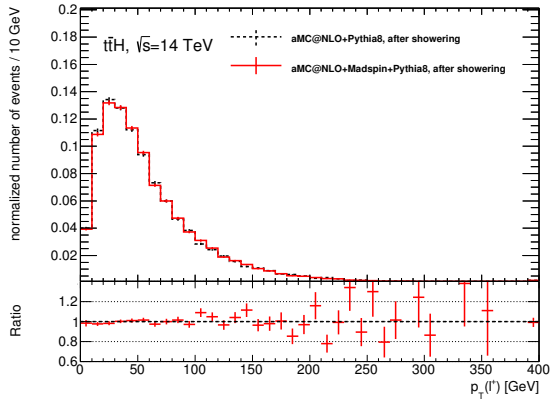
(e) Transverse momentum p_T of the positively charged lepton.
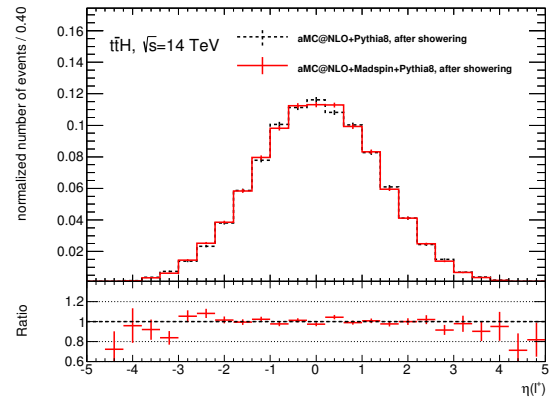
(f) Pseudorapidity η of the positively charged lepton.

Figure 6.7.: Comparison of aMC@NLO + PYTHIA8 with or without MADSPIN in kinematic variables of the Higgs boson, $t\bar{t}H$ system and ℓ^+ (after the showering).

33

6.4. Signal vs. Background Comparison

A comparison of the $t\bar{t}H$ signal against its main background $t\bar{t}+$ jets is also included for $\sqrt{s} = 8\,\text{TeV}$ and $\sqrt{s} = 14\,\text{TeV}$. The $t\bar{t}H$ samples are produced in the way described in Section 5, while the $t\bar{t}+$ jets samples where created with PowHeg as the hard process generator and Pythia6 as the hadronisation model. All samples were, however, processed with variations of the same selection code.

The focus in this Section is the analysis at $\sqrt{s} = 14\,\text{TeV}$, but for comparison the same plots have been created with $\sqrt{s} = 8\,\text{TeV}$ Monte Carlo data (see Figures A.3 to A.6).

6.4.1. Kinematic Variables

In the kinematic variables shown in Figures 6.8 and 6.9 a clear separation between the $t\bar{t}H$ signal and the $t\bar{t}+$ jets background can be observed. The transverse momenta for the latter peak at lower energies and the pseudorapidity distributions are broader. These differences are especially distinct in the $t\bar{t}$ system.

Since the total transverse momentum of the event should vanish (in the lab frame), a non-zero p_T in the $t\bar{t}$ system corresponds to the p_T of the other particles in the event, i.e. the jets in case of $t\bar{t}+$ jets and the Higgs boson in case of $t\bar{t}H$. Figure 6.8(c) for the pseudorapidity η of the $t\bar{t}$ system shows that through the presence of a Higgs boson in the $t\bar{t}H$ event, a more central $t\bar{t}$ system is observed.
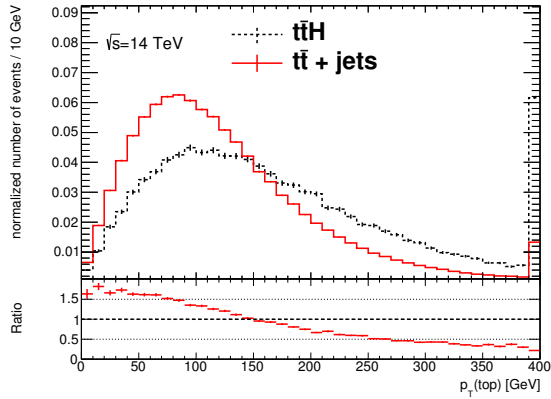
(a) Transverse momentum p_T of the top quark.
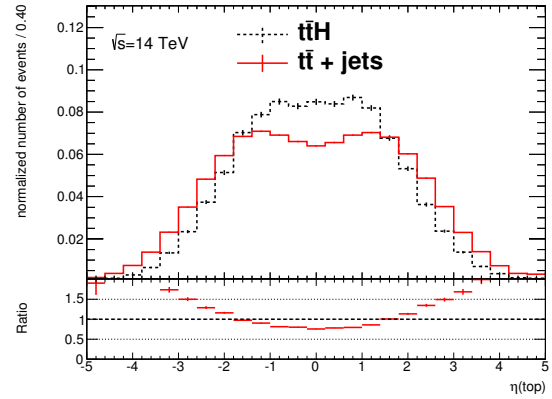
(b) Pseudorapidity η of the top quark.
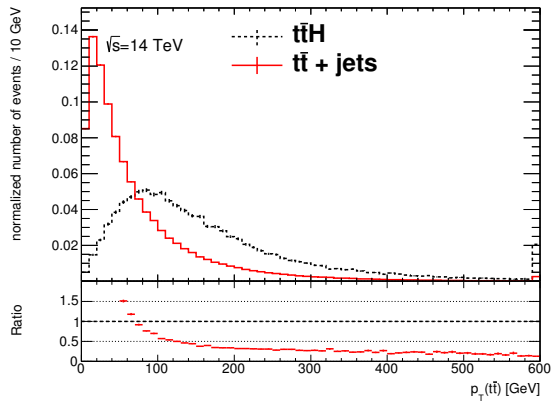
(c) Transverse momentum p_T of the $t\bar{t}$ system.
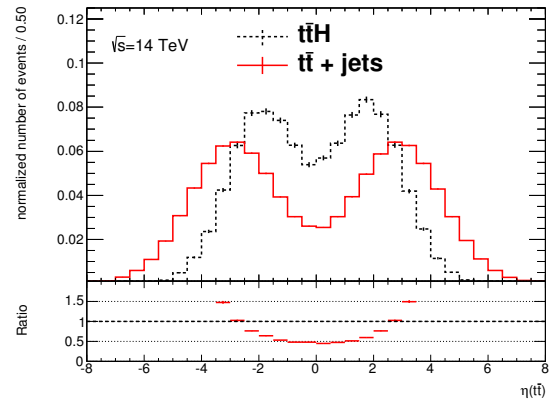
(d) Pseudorapidity η of the $t\bar{t}$ system.

Figure 6.8.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in kinematic variables of the top quark and the $t\bar{t}$ system.
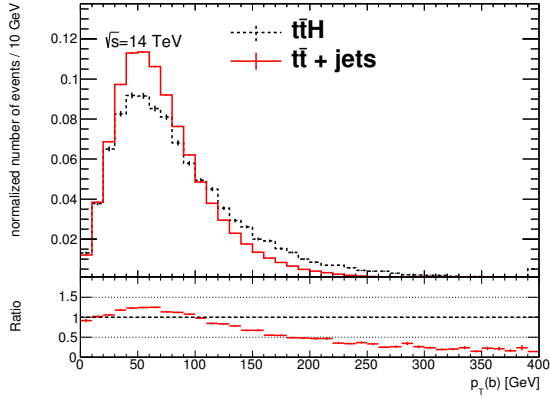
(a) Transverse momentum p_T of the bottom quark.
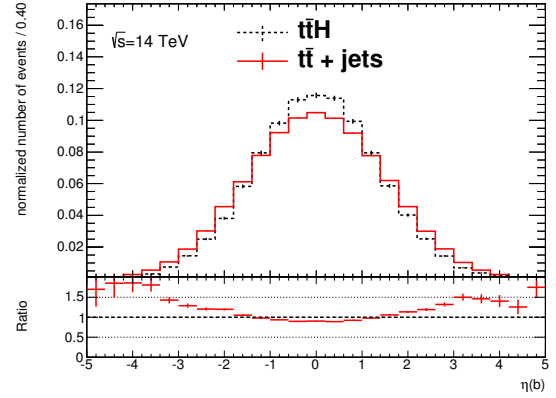
(b) Pseudorapidity η of the bottom quark.

(c) Transverse momentum p_T of the W^+ boson.
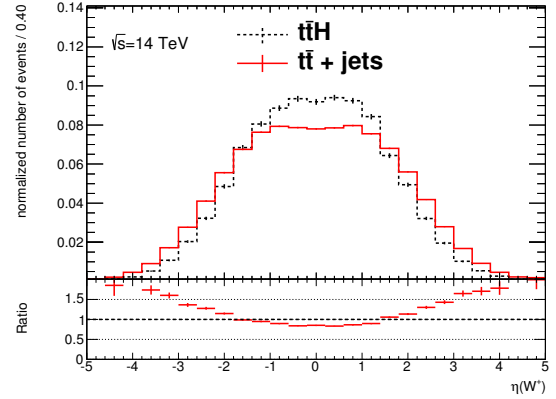
(d) Pseudorapidity η of the W^+ boson.
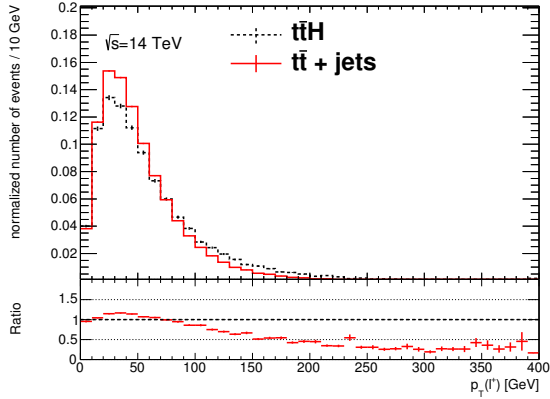
(e) Transverse momentum p_T of the positively charged lepton.
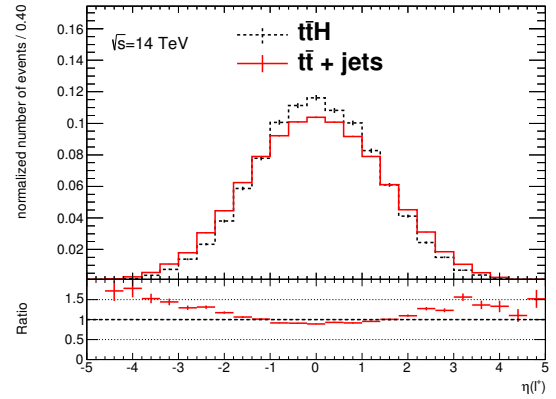
(f) Pseudorapidity η of the positively charged lepton.

Figure 6.9.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in kinematic variables of the bottom quark, W^+ boson and the ℓ^+.

6.4.2. Jet-related Variables

In the distribution of the number of jets (Figure 6.10(a)) an even clearer separation can be seen. While the $t\bar{t}H$ sample has a clear peak at 6 jets, the $t\bar{t}$ + jets sample has a peak at 4 jets. This reflects exactly the expectation for a $t\bar{t} \rightarrow \ell + jets$ and $H \rightarrow b\bar{b}$ selection. This observation justifies that a selection in N_{jets} is usually applied in Higgs physics analysis. The separation can be further improved via b-tagging or if tighter p_T cuts for the jets are applied.

The distribution of H_T^{jets} (Figure 6.10(b)) shows the same behaviour as the other transverse momenta discussed above with a very good separation between signal and background sample. Indeed this observable is exploited in the ATLAS $t\bar{t}H$ selection [29].

Figures 6.10(c) to 6.10(f) show other variables for dijet systems, which also show separation between signal and background. As already pointed out, this could improve if b-tagging is applied and tighter cuts are used, since it is expected that the jets from the top quarks and the Higgs boson carry more momentum than jets from the underlying event.

These observables are chosen similar to variables that were found to be useful in earlier background separation studies for $t\bar{t}H$ at $\sqrt{s} = 8\,\mathrm{TeV}$, for which a neural network has been used. However, since no b-tagging was available for this thesis, it was not possible to use the exact same variables as in [29].
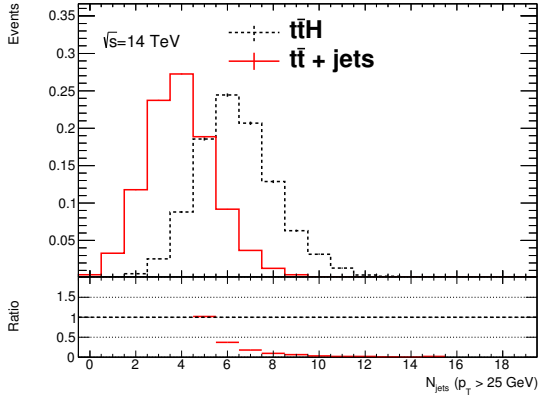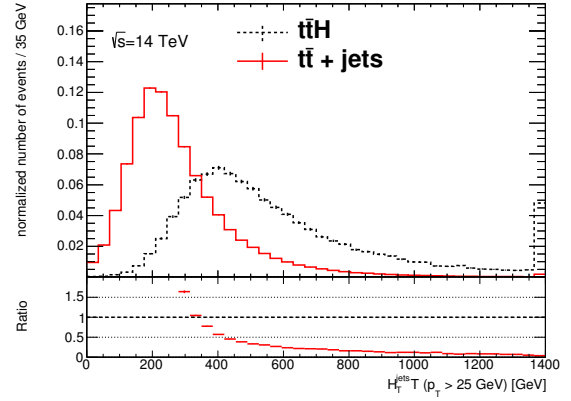
(a) Number of jets with $p_T > 25\,\text{GeV}$ and $|\eta| < 2.5$.

(b) Total transverse momentum of all jets with $p_T > 25\,\text{GeV}$ and $|\eta| < 2.5$.
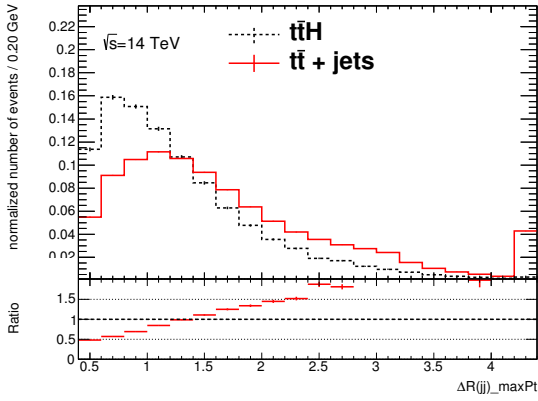
(c) ΔR of the dijet system with the maximal combined p_T.
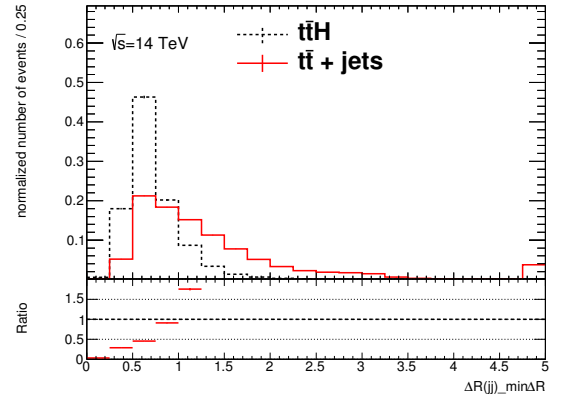
(d) ΔR of the two jets with the smallest distance between them.
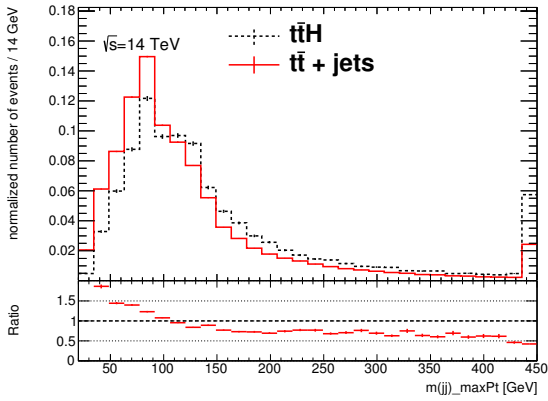
(e) Invariant mass of the dijet system with the maximal combined p_T.
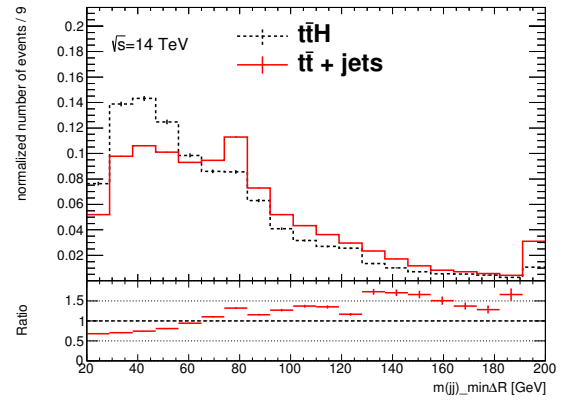
(f) Invariant mass of the two jets with the smallest distance between them.

Figure 6.10.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in jet-related variables.

6.4.3. Event Shape Variables

Event shape variables like sphericity, aplanarity, etc. have been found to be convenient to separate signal and background [29]. This can also be seen in Figure 6.11. In general the $t\bar{t}H$ event is more central and spherical than $t\bar{t}$ + jets events.
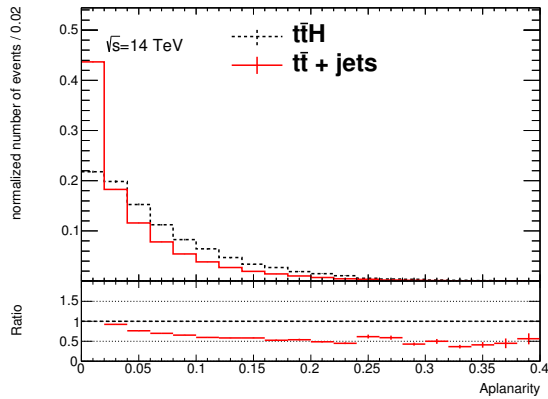
(a) Aplanarity.
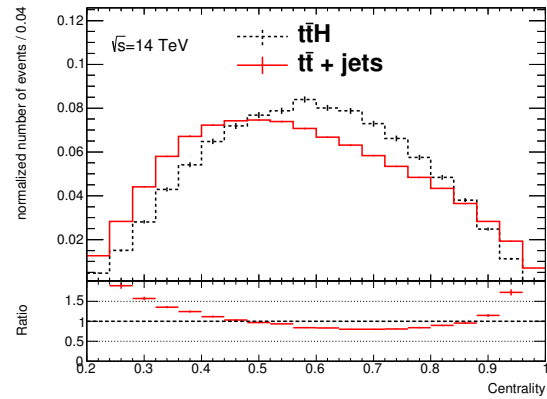
(b) Centrality.
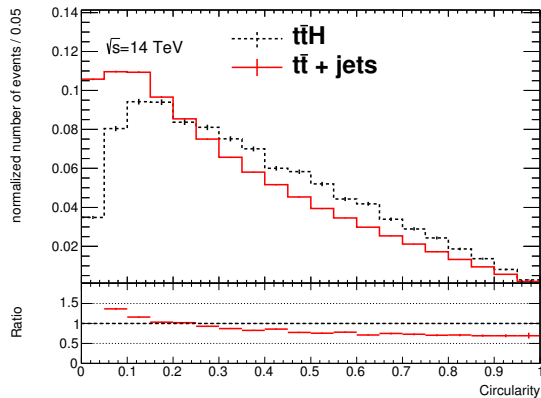
(c) Circularity.
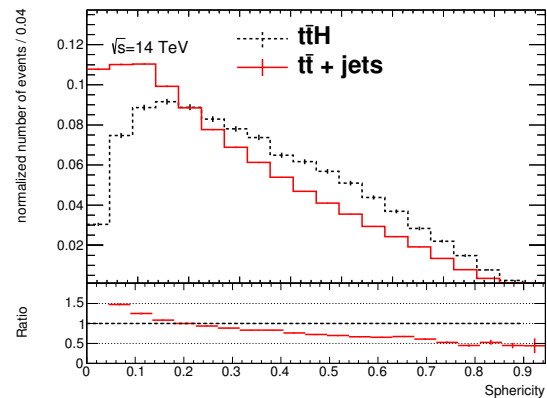
(d) Sphericity.

Figure 6.11.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in event shape variables.

7. Conclusions

For this thesis project, the whole process chain of a Monte Carlo analysis at truth level has been performed:

1. Production of the hard process $pp \rightarrow t\bar{t}H$ with the aMC@NLO framework.

2. Parton showering and hadronisation of the samples using the Pythia8 and Herwig++ models.

3. Selection, preparation and comparison of variables from the different samples.

The total number of events per sample after the selection to $t\bar{t} \rightarrow \ell + jets$ and $H \rightarrow b\bar{b}$ is shown in Table A.1, while the cross-sections, that were obtained from the event generation with aMC@NLO can be seen in Table A.2.

The produced samples are the first $t\bar{t}H$ Monte Carlo events at $\sqrt{s} = 14\,\text{TeV}$ in NLO available for the ATLAS experiment and therefore will be used further for other analysis. It could be shown that jet-related variables and event shape variables can be used to obtain a separation of the $t\bar{t}H$ signal from its main background $t\bar{t} + \text{jets}$. With tighter cuts on the jets and b-tagging information it should be possible to further improve the separation.

A. Appendix

A.1. Number of Events in Samples

Process	$\sqrt{s}$	Event generator	MadSpin	PS model	#Events
$t\bar{t}$ + jets	8 TeV	PowHeg	no	Pythia6	≈2,400,00
$t\bar{t}$ + jets	14 TeV	PowHeg	no	Pythia6	≈450,000
$t\bar{t}H$	8 TeV	aMC@NLO	no	Pythia8	≈200,000
$t\bar{t}H$	14 TeV	aMC@NLO	no	Pythia8	≈200,000
$t\bar{t}H$	14 TeV	aMC@NLO	yes	Pythia8	≈200,000
$t\bar{t}H$	14 TeV	aMC@NLO	no	Herwig++	≈490,000
$t\bar{t}H$	14 TeV	aMC@NLO	yes	Herwig++	≈490,000

Table A.1.: Total number of events in the used samples.

A.2. $t\bar{t}H$ Cross-sections obtained from Event Production

$\sqrt{s}$	PS model	σ [fb]
8 TeV	Pythia8	$120.5 \pm 0.19(\text{stat.})^{+5.6\%}_{-9.5\%}(\text{scale})$
8 TeV	Herwig++	$120.3 \pm 0.15(\text{stat.})^{+5.3\%}_{-9.4\%}(\text{scale})$
14 TeV	Pythia8	$544.1 \pm 0.8(\text{stat.})^{+6.0\%}_{-8.8\%}(\text{scale})$
14 TeV	Herwig++	$545.5 \pm 0.8(\text{stat.})^{+5.7\%}_{-8.7\%}(\text{scale})$

Table A.2.: $pp \rightarrow t\bar{t}H$ cross-sections obtained from the aMC@NLO simulation at NLO in QCD. The statistical uncertainty refers to statistics in the event production.

A.3. aMC@NLO Runcard

```
<LesHouchesEvents version="1.0">
  <header>
  <MG5ProcCard>
#*************************************************************
#*                    MadGraph5_aMC@NLO                     *
#*                                                          *
#*              *                        *                  *
#*                *         * *        *                    *
#*                  * * * * 5 * * * *                       *
#*                *         * *        *                    *
#*              *                        *                  *
#*                                                          *
#*                                                          *
#*          VERSION 2.1.0              2014-02-21           *
#*                                                          *
#*    The MadGraph5_aMC@NLO Development Team - Find us at   *
#*    https://server06.fynu.ucl.ac.be/projects/madgraph    *
#*                                                          *
#*************************************************************
#*                                                          *
#*             Command File for MadGraph5_aMC@NLO           *
#*                                                          *
#*      run as ./bin/mg5_aMC  filename                      *
#*                                                          *
#*************************************************************

set group_subprocesses Auto
set ignore_six_quark_processes False
set loop_optimized_output True
set gauge unitary
set complex_mass_scheme False
import model sm
define p = g u c d s u~ c~ d~ s~
define j = g u c d s u~ c~ d~ s~
define l+ = e+ mu+ ta+
define l- = e- mu- ta-
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
import model loop_sm
generate p p > h t t~ [QCD]
output Proc_ttH_14TeV_NLO_Pythia8_v1
  </MG5ProcCard>
```

A.4. Mᴀᴅsᴘɪɴ **Comparison with** Hᴇʀᴡɪɢ++
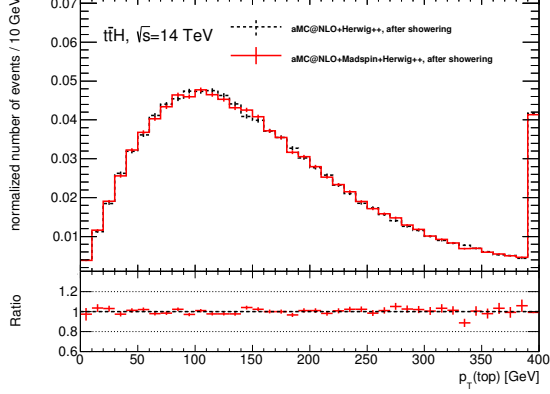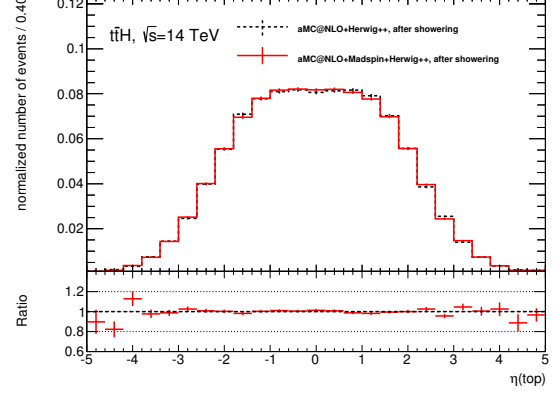
(a) Transverse momentum p_T of the top quark.
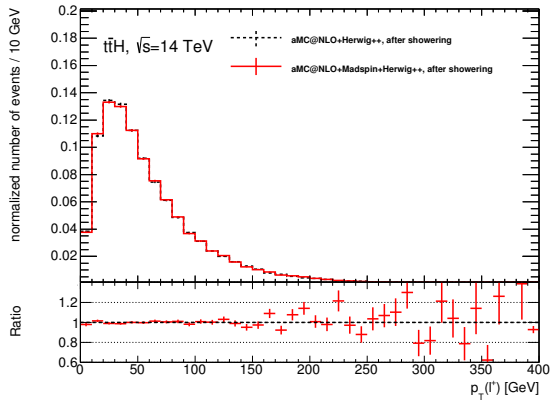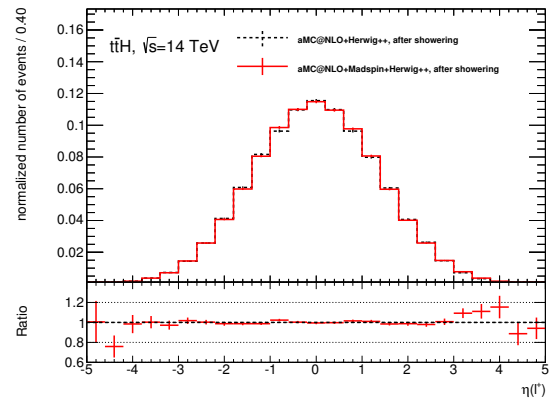
(b) Pseudorapidity η of the top quark.
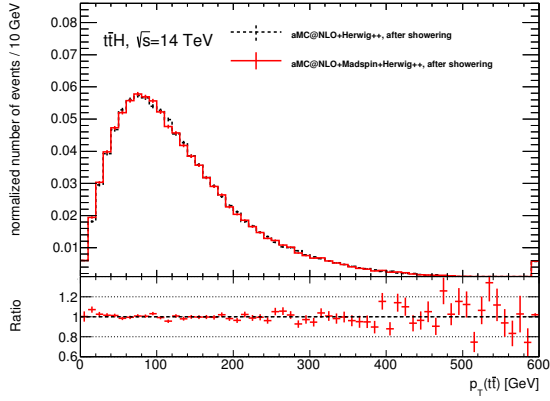
(c) Transverse momentum p_T of the $t\bar{t}$ system.
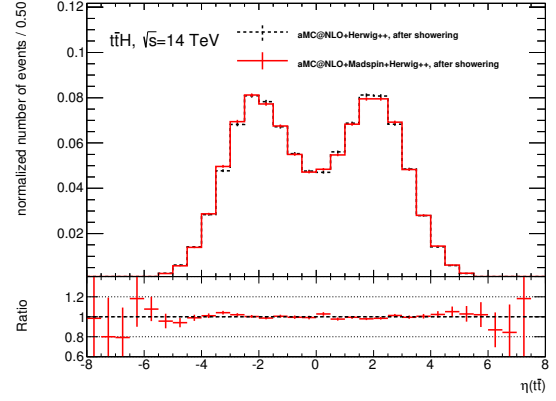
(d) Pseudorapidity η of the $t\bar{t}$ system.
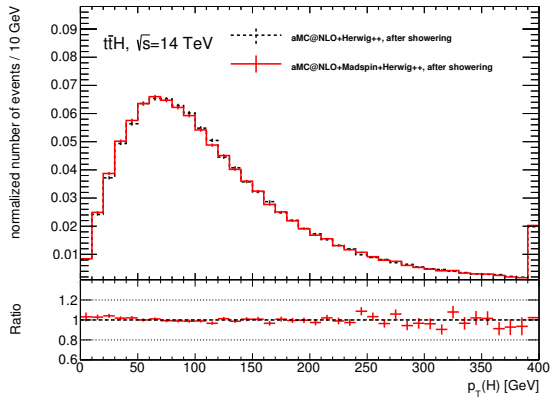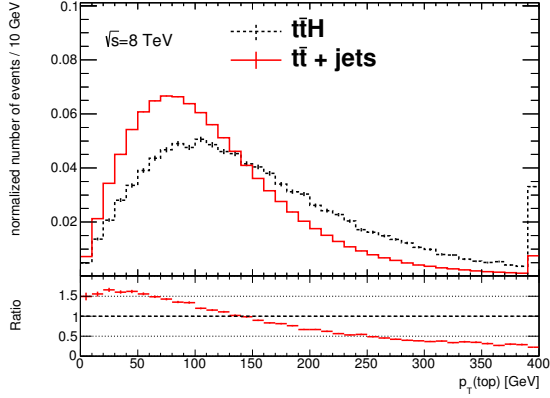
(e) Transverse momentum p_T of the Higgs boson.
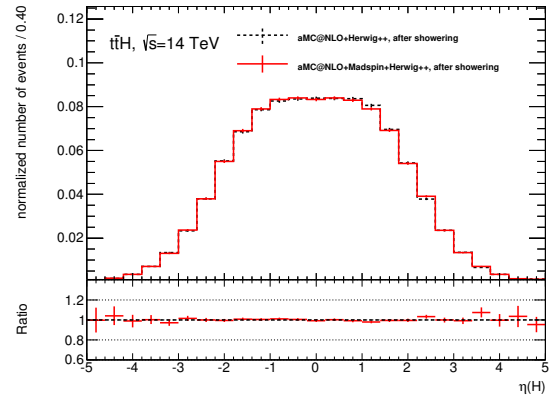
(f) Pseudorapidity η of the Higgs boson.

**Figure A.1.:** Comparison of aMC@NLO + Hᴇʀᴡɪɢ++ with or without Mᴀᴅsᴘɪɴ (after the showering).

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

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

(c) Transverse momentum p_T of the positively charged lepton.

(d) Pseudorapidity η of the positively charged lepton.

Figure A.2.: Comparison of aMC@NLO + Herwig++ with or without MadSpin (after the showering).

A.5. Signal vs. Background Comparison at 8 TeV

A.5.1. Kinematic Variables

(a) Transverse momentum p_T of the top quark.
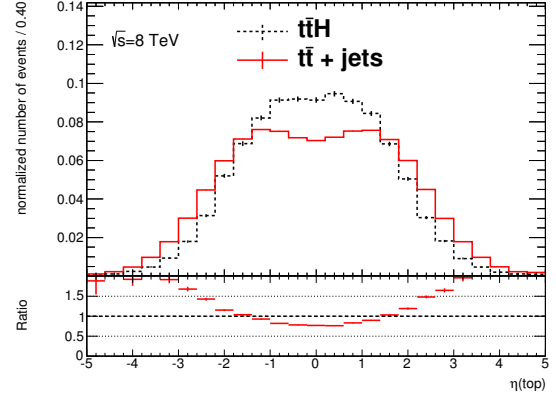
(b) Pseudorapidity η of the top quark.
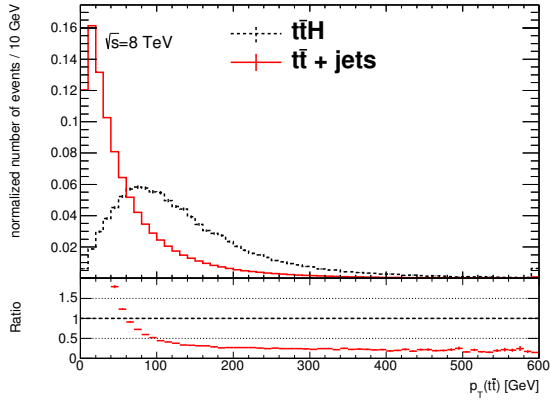
(c) Transverse momentum p_T of the $t\bar{t}$ system.
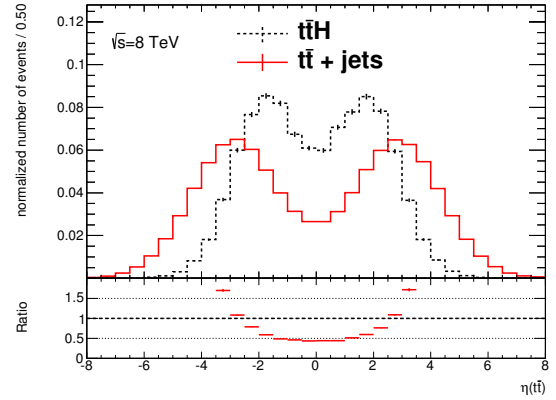
(d) Pseudorapidity η of the $t\bar{t}$ system.

Figure A.3.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background.
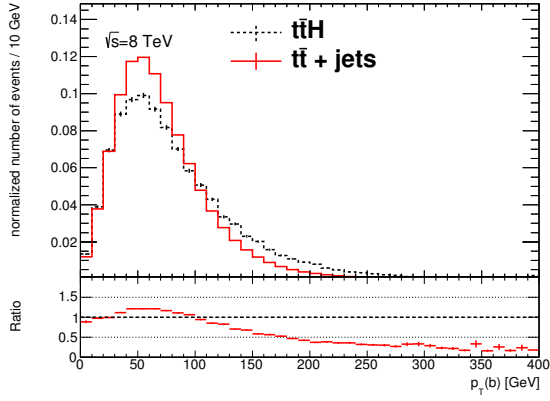
(a) Transverse momentum p_T of the bottom quark.
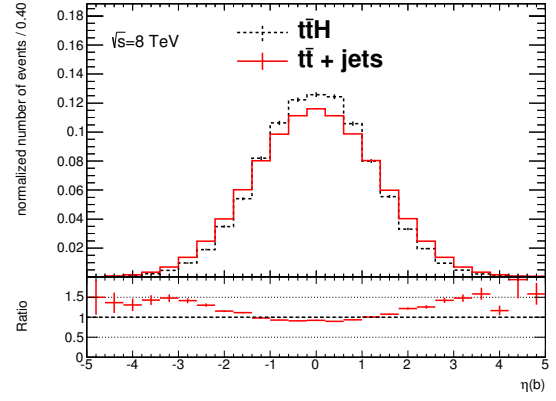
(b) Pseudorapidity η of the bottom quark.
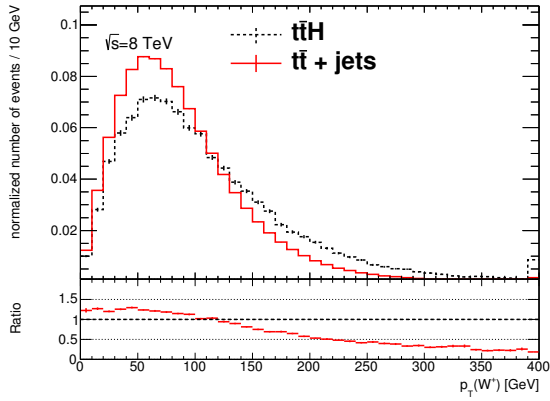
(c) Transverse momentum p_T of the W^+ boson.
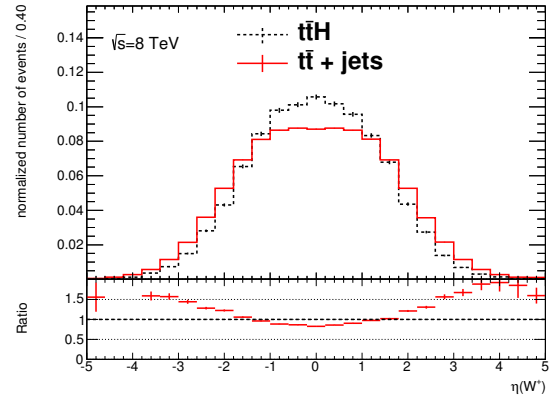
(d) Pseudorapidity η of the W^+ boson.
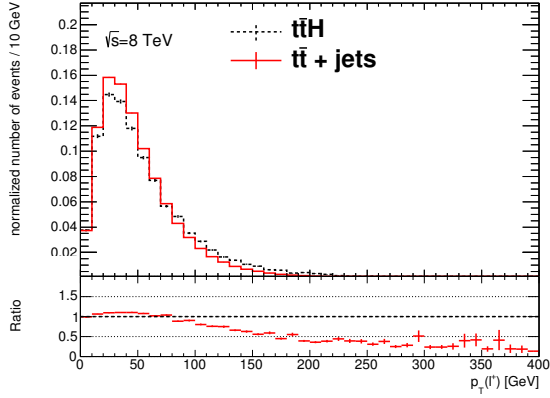
(e) Transverse momentum p_T of the positively charged lepton.
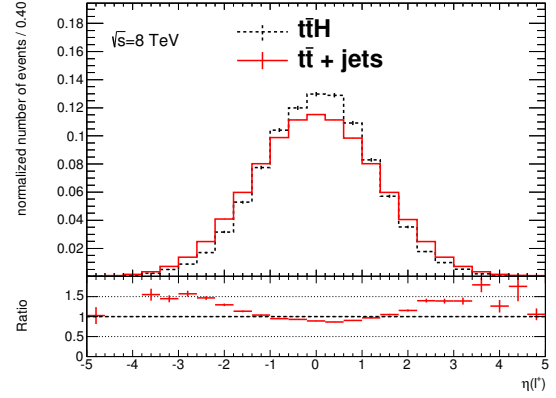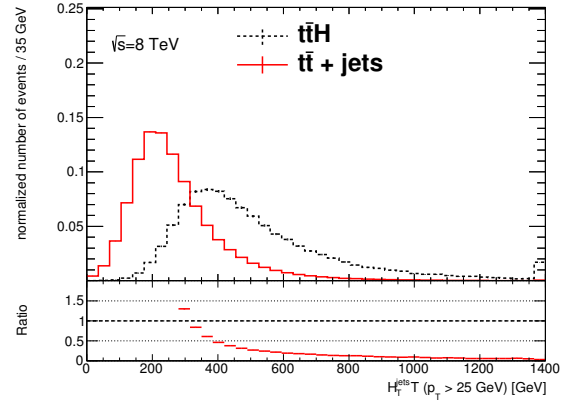
(f) Pseudorapidity η of the positively charged lepton.

Figure A.4.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background.
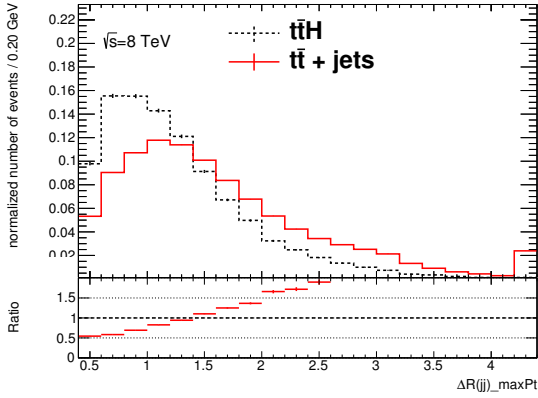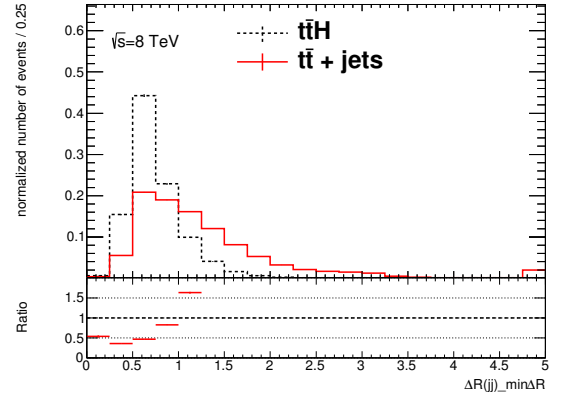
A.5.2. Jet-related Variables
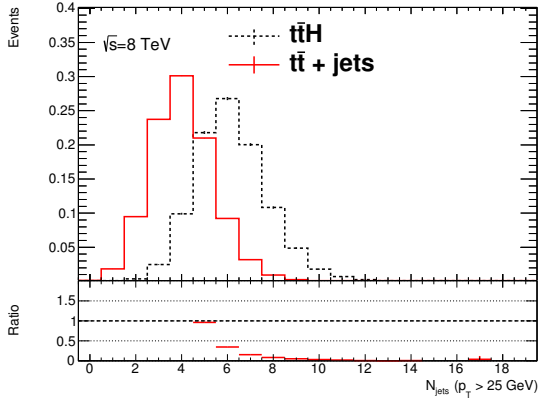
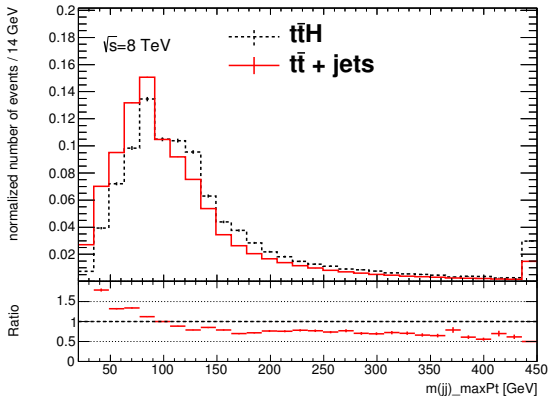
(a) Number of jets with $p_T > 25$ GeV and $|\eta| <$ 2.5.

(b) Total transverse momentum of all jets with $p_T > 25$ GeV and $|\eta| < 2.5$.
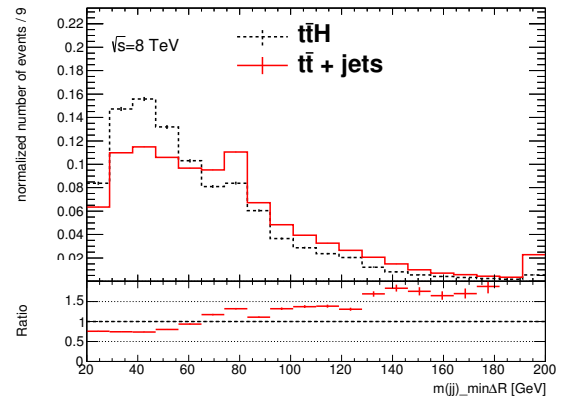

(c) ΔR of the dijet system with the maximal combined p_T.

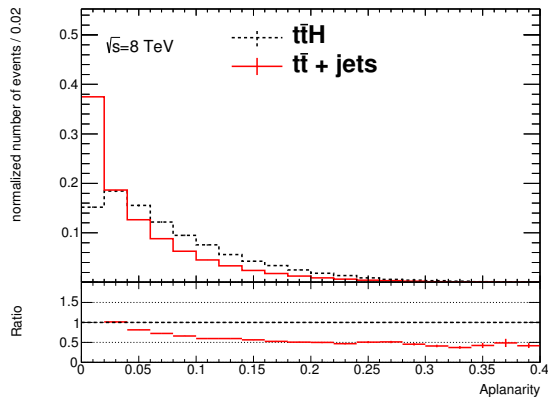(d) ΔR of the two jets with the smallest distance between them.

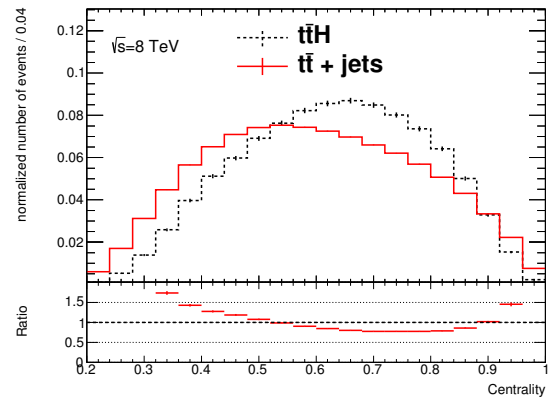(e) Invariant mass of the dijet system with the maximal combined p_T.

(f) Invariant mass of the two jets with the smallest distance between them.

Figure A.5.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background.
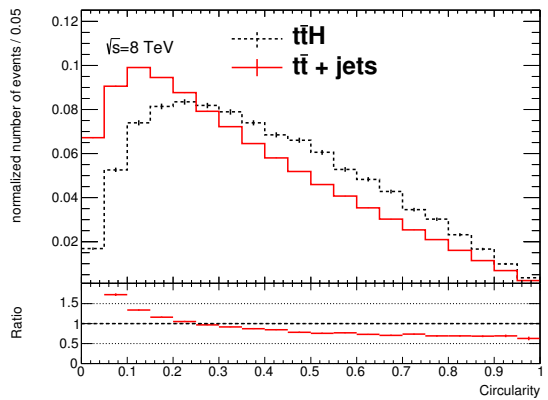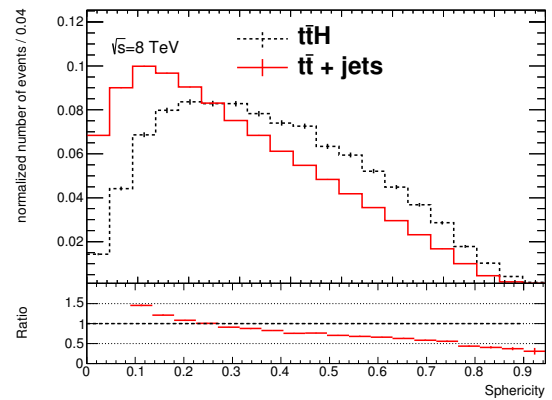
A.5.3. Event Shape Variables

(a) Aplanarity.

(b) Centrality.

(c) Circularity.

(d) Sphericity.

Figure A.6.: Comparison between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background.

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* (2014), `arXiv:hep-ph/1405.0301`

[2] D. Griffiths, *Introduction to Elementary Particles*, Wiley-VCH, Weinheim (2008), second, Revised Edition

[3] J. Beringer, et al. (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D **86**, 010001 (2012)

[4] F. Halzen, A. D. Martin, *Quarks & Leptons*, John Wiley & Sons (1984)

[5] D. H. Perkins, *Introduction to High Energy Physics*, Cambridge University Press, Cambridge (2000), fourth Edition

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964)

[7] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 1156 (1966)

[8] 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)

[9] 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)

[10] S. Dittmaier, et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables* (2011), `CERN-2011-002`

[11] V. M. Abazov, et al. (D0 Collaboration), *Precision measurement of the top-quark mass in lepton+jets final states* (2014), `arXiv:hep-ex/1405.1756`

Bibliography

[12] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector* (2012), `ATLAS-CONF-2012-135,ATLAS-COM-CONF-2012-162`

[13] M. Czakon, P. Fiedler, A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys.Rev.Lett. **110**, 252004 (2013)

[14] S. Chatrchyan, et al. (CMS Collaboration), *Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC*, Phys.Rev. **D84**, 112002 (2011)

[15] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, P. Proudlock, *LHC Design Report*, CERN, Geneva (2004)

[16] R. Mankel, *Pattern recognition and event reconstruction in particle physics experiments*, Rep.Prog.Phys. **67**, 553 (2004)

[17] J. Alwall, et al., *A standard format for Les Houches Event Files* `arXiv:hep-ph/0609017`

[18] ATLAS Collaboration, *Further ATLAS tunes of Pythia6 and Pythia8* (2011), `ATL-PHYS-PUB-2011-014`

[19] M. Bahr, et al., *Herwig++ Physics and Manual*, Eur.Phys.J. **C58**, 639 (2008)

[20] P. Artoisenet, R. Frederix, O. Mattelaer, R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **1303**, 015 (2013)

[21] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **0704**, 081 (2007)

[22] V. Hirschi, R. Frederix, S. Frixione, M. V. Garzelli, F. Maltoni, et al., *Automation of one-loop QCD corrections*, JHEP **1105**, 044 (2011)

[23] G. Ossola, C. G. Papadopoulos, R. Pittau, *Reducing full one-loop amplitudes to scalar integrals at the integrand level*, Nucl.Phys. **B763**, 147 (2007)

[24] R. Frederix, S. Frixione, F. Maltoni, T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, JHEP **0910**, 003 (2009)

[25] R. Frederix, S. Frixione, *Merging meets matching in MC@NLO*, JHEP **1212**, 061 (2012)

[26] H.-L. Lai, et al., *New parton distributions for collider physics*, Phys.Rev. **D82**, 074024 (2010)

[27] R. Brun, F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl.Instrum.Meth. **A389**, 81 (1997)

[28] M. Cacciari, G. P. Salam, G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804**, 063 (2008)

[29] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS detector at the LHC.* (2014), `ATLAS-CONF-2014-011`

Acknowledgements

Erklärung nach §13(8) 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 nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 30. September 2014

(Timo Dreyer)