Department of Physics and Astronomy

University of Heidelberg

Bachelor thesis

in Physics

submitted by

Peter Reimitz

born in Waiblingen (Germany)

summer term 2015

More and Better Fat Jets for a Light Higgs at $\sqrt{s} = 100$ TeV

This Bachelor thesis has been carried out by Peter Reimitz

at the

Institute for Theoretical Physics

under the supervision of

Prof. Tilman Plehn

Zusammenfassung:

In dieser Arbeit untersuchen wir inwiefern ein 100 TeV Hadron-Hadron-Beschleuniger zu einer genaueren Untersuchung von Top-Quark-Paarproduktionen in Zusammenhang mit einem Higgs-Boson beitragen kann. Dieser $t\bar{t}H$ Kanal liefert eine Möglichkeit zur Messung der Top-Yukawa-Kopplung. Da diese Kopplung eine wichtige Rolle beim Test der elektroschwachen Symmetriebrechung spielt, werden wir für den Fall eines 100 TeV Beschleunigers eine Abschätzung der relativen Genauigkeit auf ungefähr 1% machen. Theoretische sowie systematische Unsicherheiten können exzellent kontrolliert werden, mitunter durch glatte signalfreie Regionen und durch Bilden des Verhältnisses $t\bar{t}H/t\bar{t}Z$.

Abstract:

In this thesis, we are going to study the advantages of a 100 TeV hadron-hadron collider for a more precise investigation of the top pair production associated with a Higgs boson. This $t\bar{t}H$ channel makes it possible to measure the top Yukawa coupling. Since measuring this parameter is one of the main tasks in precision tests of the electroweak symmetry breaking, we deliver an estimation of the relative uncertainty of the top Yukawa coupling of about 1% in case of a 100 TeV collider. Theoretical as well as systematic uncertainties can be handled in an excellent way, both from side bands and from taking the ratio $t\bar{t}H/t\bar{t}Z$.

Contents

1 Motivation and Introduction

With the discovery of the Higgs boson during the LHC Run I in 2012 [1, 2], the particle content of the Standard Model of Particles (SM) has been entirely discovered. Nevertheless, especially for the Higgs boson one has to test the SM nature of these particles in upcoming projects hopefully being indicative of SM deviating physics. For that reason, one should push ahead the plans of building new colliders with advanced properties like higher center-of-mass (c.m.) energies. The usefulness of doing that has to be tested with simulations and analyses. In this thesis, the benefit of a 100 TeV collider to investigate the $t\bar{t}H$ channel is examined.

In the following, we focus on the $t\bar{t}H$ production including a semileptonic decay of a top quark pair and a Higgs decay into a bottom quark pair ($H \to b\bar{b}$) generated in a proton-proton collision at $\sqrt{s} = 100$ TeV. Due to its high final state multiplicity, this process is hardly detectable but gives access to the top Yukawa coupling, which is one of the key observables in precision measurements of the electroweak symmetry breaking.

We first generate hard processes with the Monte-Carlo Generator MadGraph5 [10] and shower and hadronize via Pythia8 [9]. The detector simulation part is done with Delphes3.2.0 [12]. A detailed description of how we create step by step 'experiment-like data sets' will be presented in section 3.

Detectors like the ATLAS and the CMS at the LHC measure the four momentum of final particles. In a boosted regime, the decay products of heavy particles like the Higgs boson and the top quark are collimated, with the result that they form so called fat jets. Taking these jets as a starting point for hadronic activity, we investigate their substructure and try to get information about hard processes. That will be described in section 4.

In detail, we use FastJet3 [8] to first build fat jets via the Cambridge/Aachen (C/A) jet algorithm [11]. With a modfied BDRS Higgs tagger [5, 6] and the HEPTopTagger2 [4], we analyse the hard substructure yielding candidates of subjets for a mass reconstruction of the appropriate particle. Following this, we start an analysis (section 5) to isolate the signal from background events. As a result, in section 6 we study the reconstructed mass distribution and determine the statistical significance and uncertainties of the Higgs signal. In addition, we perform a combined Higgs and $t\bar{t}Z$ peak analysis to get a more systematic and theoretical uncertainty reduced value. Finally, these calculations lead to an estimation for a relative uncertainty of the top Yukawa coupling.

2 Theoretical Background

2.1 The Standard Model of Particles

The SM is a quantum field theory describing all interactions between all so far found particles except of the gravitational interaction. Being such a fundamental theory, it is tested very precisely in many ways. Heretofore, in no experiment one could find a clear proof of SM deviating physics. Nevertheless, the SM is limited. For example, it does not include dark matter or dark energy. Hopefully experiments will show SM extending physics in future projects.

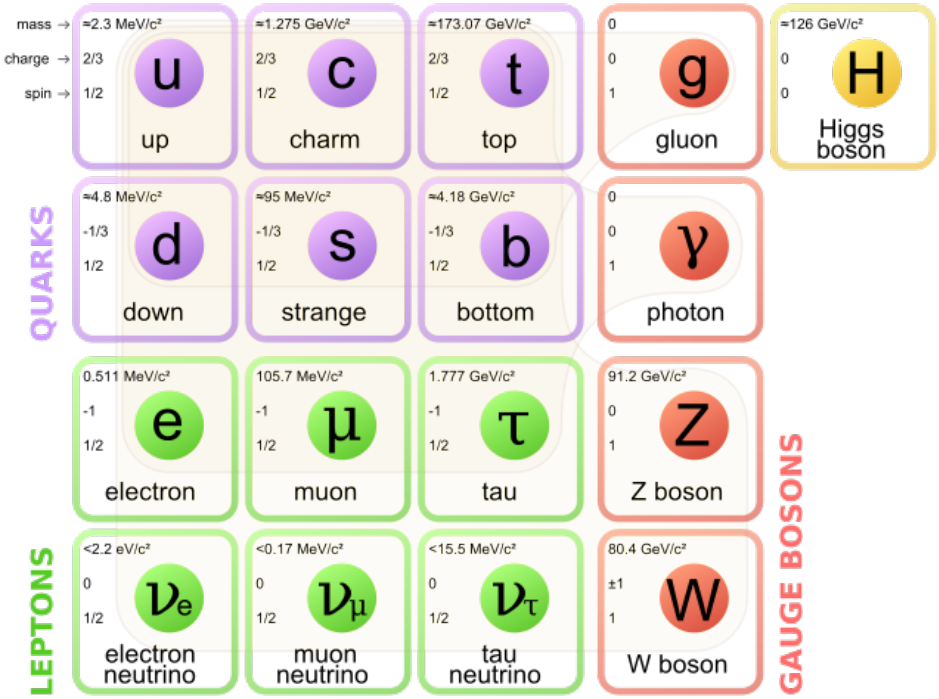

Figure 2.1: Elementary particles of the SM taken from [14].

Particle Content

The SM contains plenty of elementary particles with different properties as seen in Fig. 2.1. For one, there are spin-1/2 particles called fermions consisting of quarks (q) and leptons (l). Both are classified into three generations ordered by increasing mass. All quarks and leptons listed in Fig. 2.1 have antiparticles ($\bar{q}$, $\bar{l}$) with equal mass and opposite signed charge. The six quarks are called up, down, charm, strange, top, and bottom abbreviated as u, d, c, s, t, and b. The electron (e) being the lightest charged particle is stable due to charge conservation. The other two charged leptons can decay, the tau (τ) due to its relatively high mass even

hadronically. Neutrinos (ν_l) are only detected indirectly via other charged particles, missing energy or momentum or scattering.

Beyond fermions we have integer spin particles called bosons. This category includes spin-1 gauge bosons which can be seen as force carriers. Gluons (g) carry colour charge and can interact among themselves and with quarks mediating the strong interaction.

The $W^{\pm}$ and the Z^0 boson (Z) are force carriers of the weak interaction. The Z as an uncharged particle with a high mass compared to other particles has decay channels similar to the Higgs boson (H) as seen in Fig. 2.2. Since we focus on the $H \to b\bar{b}$ decay, we have to add the common $Z \to b\bar{b}$ decay of Z bosons to the background samples. This leads to an additional Z-peak in the m_{bb} distribution which turns out to be very useful as one can see in section 5. The last remaining gauge boson is the photon (γ). It is the exchange particle of the electromagnetic force.

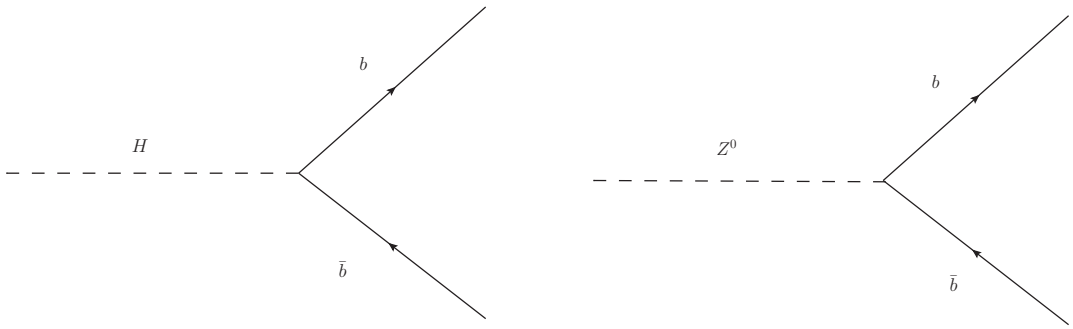

Figure 2.2: Higgs and Z decay into a bottom, antibottom pair.

Interactions

Having introduced all particles of the SM, we focus on the interaction between them. As already mentioned, the gauge bosons of the SM namely Z, $W^{\pm}$, and γ are also called force carriers. The reason for it is that these paritcles act as exchange partners in the interactions of the SM. The photon is the mediator of the electromagnetic interaction, the Z and $W^{\pm}$ boson those of the weak interaction and the gluon represents the mediator of the strong interaction. This issue is summarized in Tab. 2.1 . Salam and Weinberg described the electroweak interaction which unifies the electromagnetic and the weak interaction [24, 25].

interaction	mediator	acts on	mass of mediators	involved particles
strong	g	colour charge	0	quarks, gluons
electromagnetic	γ	electric charge	0	electrically charged particles
weak	Z, $W^{\pm}$	flavour	$\approx 10^2\ GeV$	quarks, leptons

Table 2.1: Fundamental forces of the SM and its properties.

In all interactions we preserve energy (E), momentum ($\vec{p}$), angular momentum ($\vec{L}$), electric charge, spin, colour charge, baryon number, and all lepton numbers (L_e, L_μ, L_τ).

Another important fact is that particles can not carry colour isolated (colour confinement) and therefore build compositions of quarks called hadrons. This process is called hadronization. In collider experiments, the produced coloured particles immediately hadronize and form jets that can be observed lately.

Fermion Mass Generation

To see how fermion masses are generated, one has to take a look at the Yukawa Lagrangian of the electroweak SM. This includes a scalar Higgs field ϕ with a potential

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{2.1}$$

and $\lambda > 0$. For $\mu^2 > 0$, we get a nonzero minimum for the potential:

$$|\langle \phi \rangle| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}} \, . \tag{2.2}$$

Now, we can arbitrary choose the following minimum state

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \, . \tag{2.3}$$

As an example of the fermion mass generation, we consider the first family Yukawa Lagrangian:

$$-\mathcal{L}_{\text{Yuk}} = y_e \bar{l}_L \phi e_R + y_u \bar{q}_L \tilde{\phi} u_R + y_d \bar{q}_L \phi d_R + h.c. \, . \tag{2.4}$$

Taking into account the Higgs field from above, the Lagrangian takes the form

$$-\mathcal{L}_{\text{Yuk}} = \frac{y_e v}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) + \frac{y_u v}{\sqrt{2}} (\bar{u}_L u_R + \bar{u}_R u_L) + \frac{y_d v}{\sqrt{2}} (\bar{d}_L d_R + \bar{d}_R d_L) \tag{2.5}$$

from which one can read off the fermion masses:

$$m_i = \frac{y_i v}{\sqrt{2}}, \quad i = e, u, d \, . \tag{2.6}$$

We need this kind of mass generation since direct mass terms are forbidden.

The Top Yukawa Coupling

In the SM, the top Yukawa coupling y_t is much larger compared to all other quark Yukawa couplings. A main problem in particle physics is to understand these large mass ratios. The top Yukawa coupling plays an important role in the breaking of the electroweak symmetry. Hence, it is reasonable to know as much about the top Yukawa coupling as possible which leads to precise measurements of it. In general, the Yukawa couplings are not fixed but dependent on a certain energy scale Q. This is determined by a renormalization group equation. Only limited to the top and Higgs sector, the running of the Higgs self coupling $\lambda\left(Q^2\right)$ reads [26]

$$\frac{d\lambda}{d\log(Q^2)} = \frac{1}{16\pi^2} \left(12\lambda^2 + 6\lambda y_t^2 - 3y_t^4\right) \, , \tag{2.7}$$

Couplings describe the strength of a certain interaction e.g. how strong a Higgs boson couples to a top quark for the top Yukawa coupling. In the SM, y_t directly depends on the top mass m_t. From (2.6), we get for the top quark:

$$y_t = \frac{m_t\sqrt{2}}{v} \tag{2.8}$$

with the vacuum expectation value (VEV) for the Higgs field $v = 246$ GeV. From the top mass $m_t \approx 173$ GeV we get $y_t \approx 1$. As one can see in section 3, the precision of measuring the top Yukawa coupling can be determined by the relative error of the signal rate of the Higgs signal.

2.2 Relevant Particle Processes

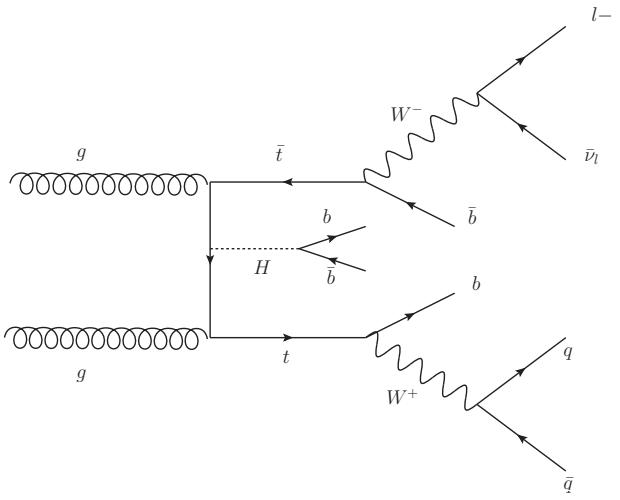

Figure 2.3: Examplary Feynman graph of the Higgs production associated with a top quark pair.

Signal

Compared to the $t\bar{t}H$ channel, it is much easier to search for a Higgs in decay channels like $H \to ZZ^{(*)} \to 4\ell$ or $H \to \gamma\gamma$ although their branching ratios (BR) are low. For example, the decay to photons is well detectable since photons are precisely measurable combined with a smooth background. For that reason, the actual Higgs discovery is based on these two decay channels [2]. Although the $H \to \gamma\gamma$ decay yields the clearest signal, it includes a loop-induced Higgs-photon coupling and therefore does not represent a model-independent measurement of the top Yukawa coupling. The $H \to b\bar{b}$ decay is not model-independent either. But for $t\bar{t}H$, it yields a direct (i.e. tree level) measurement of the top Yukawa coupling. Following this argument, a reliable measurement of the Higgs sector points to the BR dominant

$$pp \to t\bar{t}H \to t\bar{t}b\bar{b} \tag{2.9}$$

decay channel. For a 100 TeV analysis, the largest BR yields an extraordinary statistic which confirms a high precision in this channel. In detail, we will focus on a top pair production associated with a Higgs boson decaying to two bottom quarks as seen in Fig 2.3. We consider a semileptonical top pair to trigger on the isolated lepton and reduce multi-jet combinatorics. Finally, we obtain

$$pp \to t\bar{t}H \to (bjj)\,(\bar{b}\ell\bar{\nu})\,(b\bar{b}), (b\ell\nu)\,(\bar{b}jj)\,(b\bar{b})\ . \tag{2.10}$$

The hadronic top decay will be handled by a fat jet analysis with the HEPTopTagger2 as described below in section 4.1. For the Higgs tagging, we will use a modified BDRS Higgs tagger [5].

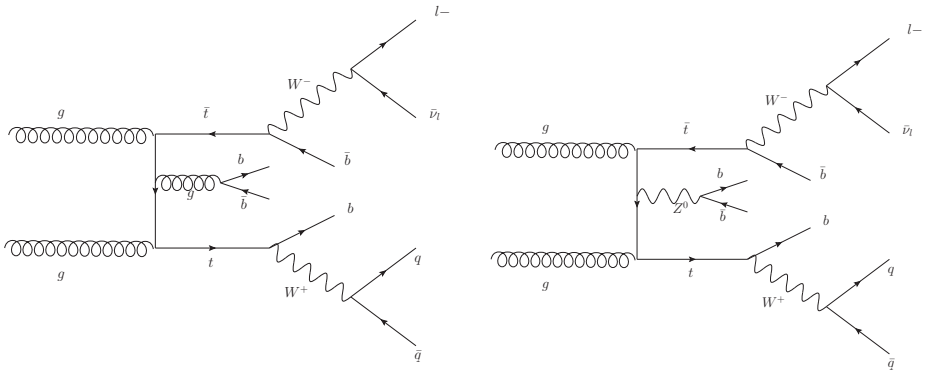

Figure 2.4: Exemplary graphs of $t\bar{t}b\bar{b}$(left) and $t\bar{t}Z$(right) as examples of the background in the $H \to b\bar{b}$ decay channel.

Background

The background mainly consists of

$pp \to t\bar{t}\,b\bar{b}$, the main irreducible QCD background

$pp \to t\bar{t}Z$, including the Z-peak in the m_{bb} distribution

$pp \to t\bar{t}$+jets with fake-bottom tags .

The bottom quark pair production, apart from the top decays, in the $t\bar{t}b\bar{b}$ background is based on gluon radiation as one can see exemplarily in the left diagram of Fig. 2.4. A background like W+jets production can be neglected as discussed in [6].

As already mentioned above, the topology of Feynman graphs of the $t\bar{t}$ decay to $H \to b\bar{b}$ and $Z \to b\bar{b}$ is identic (Fig 2.3 and 2.4). Hence, we will have a clearly visible Z resonance in our final m_{bb} distribution although the cross section is rather small as one can see in the next section. For $t\bar{t}$ + jets we use one hard jet. We see no need for merged samples since we found out that the influence of $t\bar{t} + 2j$ to our analysis is negligible.

3 Collider Phenomenology

Collider phenomenology is the intersection between theory and experiment in particle collider physics. First, one tries to decode theoretical models and translate them into experimentally observable parameters. Secondly, experimental data have to be handled and interpreted to find their implications to the physics beyond.

3.1 Collider Parameters

In analysing data of an event simulation one must focus on a few collider typical parameters [26, 23]. For a particle with mass m and a four momentum vector $p^\mu = (E, \vec{p}) \equiv p$, in the Minkowski metric the square is defined as

$$p^\mu p_\mu = p^2 = E^2 - \vec{p} \cdot \vec{p} = m^2 \tag{3.1}$$

and characterizes a Lorentz invariant observable. In a collision of two particles $p_1 = (E_1, 0, 0, p_{z,1})$ and $p_2 = (E_2, 0, 0, p_{z,2})$, e.g. a proton-proton collision, this yields the Mandelstam variable

$$s = (p_1 + p_2)^2 = (E_1 + E_2)^2 \tag{3.2}$$

due to $\vec{p_1} + \vec{p_2} = 0$ in the c.m. frame. For the c.m. energy this results in

$$E_{\text{c.m.}} = \sqrt{s} = 2 \cdot E \tag{3.3}$$

assuming that $E_1 = E_2 \equiv E$. For instance, each proton beam of a hadron collider has to carry an energy of 50 TeV in order to reach $\sqrt{s} = 100$ TeV. We want to predict the number of signal and background events as well as the relative error in this analysis. The total event number N is given by

$$N = \epsilon \cdot \sigma \cdot \mathcal{L} \tag{3.4}$$

with the production cross section for a specific event σ listed in Tab. 3.1 and the luminosity $\mathcal{L}$ measured in $ab^{-1} = 10^{46} m^{-2}$. The latter one describes a certain amount of particles crossing a unit area. For a 100 TeV collider, we assume an integrated luminosity of 20 ab^{-1}. In our case, the cross section σ will be multiplied by the fraction of produced events ϵ satisfying the event selection and passing the detector simulation. Obviously, we want to reach a large signal to background ratio $S/B > 1$. For a discovery, it is convention to reach a five sigma excess over background resulting in $S/\sqrt{B} > 5$. According to (3.4) and the fact that the cross section of $t\bar{t}H$ is connected to the top Yukawa coupling via $\sigma \propto y_t^2$, a relative error of the signal rate can be translated into a relative error estimation of the top Yukawa coupling.

Due to the collider geometry, it is useful to describe the kinematics with cylindrical coordinates consisting of the azimuth angle ϕ, $p_L \equiv p_z$ longitudinal to the collider beam direction and the transverse momentum

$$p_T = \sqrt{p_x^2 + p_y^2}\,. \tag{3.5}$$

Later on, p_L will be replaced by another more useful variable y considering that the longitudinal momentum can not be measured and is hence not useful at all. Since the incoming particles being collinear to the proton have no transverse momentum, the final states transverse momenta also sums up to zero:

$$\sum_{\text{finalstate}} \vec{p}_{T,j} = \vec{0} \,. \tag{3.6}$$

The transverse momentum vector is a two dimensional vector in a transverse plane parameterized by the absolute value p_T and the azimuth angle ϕ. In the following analysis cuts, we will often require a minimum p_T. Small values of p_T can be reached by objects directed more or less in beam direction or objects with low momentum at all. In the case of jets, we call the ones with high transverse momentum hard jets. From the rest frame of a particle a longitudinal boost reads

$$\begin{pmatrix} E \\ p_L \end{pmatrix} = \exp\left[y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right] \begin{pmatrix} m \\ 0 \end{pmatrix} \,. \tag{3.7}$$

The rapidity y which is additive under longitudinal boosts can be expressed as

$$y = \frac{1}{2} \cdot \ln\left(\frac{E + p_L}{E - p_L} \right) \,. \tag{3.8}$$

The rapidity becomes zero for $p_L = 0$ and very large for high longitudinal momentum values. Since there is essentially no detector in the longitudinal direction, we set an upper limit for y in the upcoming analysis. Reasonable limits are below $|y| < 4$. For massless particles we have $E = |\vec{p}|$ and after introducing the polar angle θ between the beam axis and the momentum $\vec{p}$ the rapidity y becomes

$$\begin{aligned} y &= \frac{1}{2} \cdot \ln\left(\frac{E + p_L}{E - p_L} \right) \approx \frac{1}{2} \cdot \ln\left(\frac{|\vec{p}| + p_L}{|\vec{p}| - p_L} \right) \\ &= \frac{1}{2} \cdot \ln\left(\frac{1 + \cos(\theta)}{1 - \cos(\theta)} \right) = \frac{1}{2} \cdot \ln\left(\frac{1}{\tan^2(\frac{\theta}{2})} \right) \\ &= -\ln\left(\tan\left(\frac{\theta}{2} \right) \right) \equiv \eta \,. \end{aligned} \tag{3.9}$$

We call η the pseudorapidity. For example, $\theta = 90°$ results in $\eta = 0$ and for $\theta \to 0$ one gets $\eta \to \infty$. Since we deal with very high energies, the approximation $E \approx |\vec{p}|$ i.e. particles with negligible mass can be made very often yielding $y \approx \eta$. With the rapidity y and the azimuth angle ϕ we can introduce another very important parameter in collider physics, the angular distance ΔR which is defined by

$$(\Delta R)^2 = (\Delta y)^2 + (\Delta \phi)^2 \tag{3.10}$$

and for massless particles

$$(\Delta R)^2 = (\Delta \eta)^2 + (\Delta \phi)^2 \,. \tag{3.11}$$

The parameter ΔR is useful to describe distances in the detector plane. So objects with a very small angular distance are very close together in the detector. For that reason, this distance measure is the crucial parameter of jet clustering algorithms described below. Later we will

deal with cone sizes up to $R = 1.8$ building fat jets. For defining jets in general, $R = 0.4$ is already adequate.

3.2 From a Hard Process to a LHC Event

Event Generation with MadGraph5

To study collider physics theoretically, we need to generate events to get event shape observables following Fig. 3.1. For that reason, we use the Monte-Carlo event generator MadGraph5 in order to obtain the cross sections and the hard event generation for our processes. On generator level for a proton-proton collision at $\sqrt{s} = 100$ TeV with NNPDF23 parton densities (PDF) [13] we require:

$$\begin{aligned} p_{T,j} > 10 \text{ GeV}, \quad p_{T,b} > 10 \text{ GeV}, \quad p_{T,\ell} > 10 \text{ GeV}, \\ \Delta R_{jj} > 0.1, \quad \Delta R_{bb} > 0.1, \quad \Delta R_{jl} > 0.1 \, . \end{aligned} \tag{3.12}$$

For the $t\bar{t}+$jets background we require one hard jet with $p_{T,j} > 100$ GeV on the hard matrix element level. Worth mentioning here is that ΔR_{xy} is chosen in the order of the detector resolution and $p_{x,y}$ according to the sensitivity of the detector. Finally, the data is stored in a HepMC-file [19] including information of the process. The cross sections are listed in the following table 3.1.

$t\bar{t}H$	4.2 pb
$t\bar{t}b\bar{b}$	121 pb
$t\bar{t}Z$	1.2 pb
$t\bar{t}j$	2750 pb

Table 3.1: Leading order MadGraph5 cross sections.

As one can see, the $t\bar{t}+$jets background is the dominating sample at this stage.

Parton Shower, Hadronization and Underlying Events with Pythia8

We use Pythia8 [9] to perform the parton shower (PS) and hadronization as well as adding underlying events (UE). This yields more complex multiparticle final states. The physics is derived by phenomenological models using data and experimental proved parameters. In practise, we use the events we obtain from MadGraph5. The partons we include in our process radiate gluons, which can themselves emit further gluons or produce quark-antiquark pairs. This leads to additional quarks and gluons forming a parton shower. One must differ between initial state radiation (ISR) and final state radiation (FSR). ISR is produced by the initial partons and should be removed if possible. FSR, on the contrary, is produced by final partons. We would like to collect this radiation since it contains important momentum information for the reconstructed fat jet. Furthermore, the coloured quarks can not exist freely and according to that they will hadronize to colourless mesons and baryons and subsequently decay since many hadrons are unstable. The hadronization and decay is also included in Pythia8. Besides, the fraction of processes we are interested in at high transverse moment is rather small compared to the total number of processes in a proton-proton collision. Aside

from the gluons leading to the production of our top quark pair, the two protons with their quark and gluon content can interact in several other ways. These remnant events are called underlying events and are finally taken into account. An event simulation at the LHC is shown in Fig. 3.1. The influence of the UE on the mass of a jet with a cone size R is [27]

$$\langle \delta m_j \rangle \simeq \Lambda_{UE} \cdot p_{T,j} \left(\frac{R^4}{4} + \frac{R^8}{4608} + \mathcal{O}(R^{12}) \right) . \tag{3.13}$$

To reduce the influence of the UE, one filters [5] a fat jet and tries to reduces its cone size. These two steps also reduce pile-up effects arising from multiple collisions within a proton bunch since collisions are not performed with single protons. Considering that experimentalists can control the pile-up, we can neglect it in our analysis.

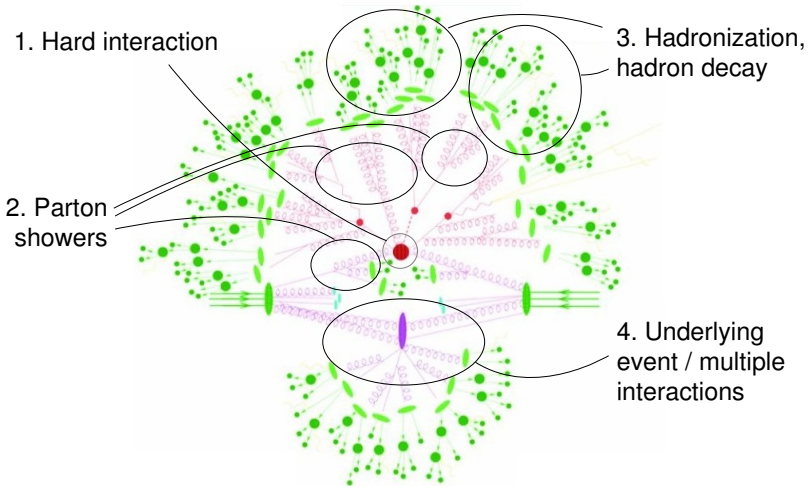

Figure 3.1: Simulating physics at the LHC. Taken from [20, 21].

Fast Detector Simulation with Delphes3

To complete the simulation of the events, we have to take a look at our detector simulation. This will be done by Delphes3 [12]. In general, all properties of our detector like the resolution, the geometry of the detector, effects of the magnetic field as well as a track propagation system or the granularity of the calorimeters will be taken into account by Delphes3 . For the 100 TeV event simulation, we make use of the Snowmass detector card for a 100 TeV collider [7]. For the leptons to be accepted, they have to pass a minimum $p_{T,l} > 10$ GeV. As an isolation criterion, one requires a transverse momentum ratio of $I < 0.1$ within $\Delta R < 0.3$. Finally, we take energy-flow objects for the jet clustering later-on.

3.3 Jets and Fat Jets

Jets or more specifically fat jets play an important part in this analysis. After the parton shower, quarks or gluons in the hard matrix element (ME) build sprays of hadrons since they can not exist freely in the asymptotic limit. With these hadronic depositions, we try to construct jets to find something that refers to the parton. Assuming that each parton turns into a jet, we can map the jet properties to the parton properties i.e. the four-momentum of the jet links to the four-momentum of the original parton or the jet mass reflects the parton mass e.g. a quark mass. The energy and four-momentum depositions measured in hadron-hadron colliders are spread all over the detector. So it is not obvious how to reconstruct these depositions to jets. To get a jet, one has to cluster those depositions together coming from the same direction. The question arises: How do we build a jet?

Jet Algorithms

Hence, we define a jet by a recombination-algorithm. To form parton reflecting jets, we have to find a measure to gain that. The most common algorithms in this field, the purely geometric Cambridge/Aachen-algorithm (C/A) [11] and the (anti-) $k_\perp$-algorithm [29, 28], include a cone size R and the transverse momentum $p_{T,j}$ of a calorimeter tower with respect to another calorimeter tower or the beam axis as a measure. The measures d_{ij} of the above-mentioned algorithms are:

k_T-algorithm: $d_{ij} = \frac{\Delta R_{ij}}{R} \min\left(p_{T,i}, p_{T,j}\right)$, $d_{iB} = p_{T,i}$

C/A-algorithm: $d_{ij} = \frac{\Delta R_{ij}}{R}$, $d_{iB} = 1$

anti-k_T-algorithm: $d_{ij} = \frac{\Delta R_{ij}}{R} \min\left(p_{T,i}^{-1}, p_{T,j}^{-1}\right)$, $d_{iB} = p_{T,i}^{-1}$.

In this algorithm we combine two subjets if the condition $d_{ij} < d_{\text{cut}}$ is fulfilled with a reference scale d_{cut} we give to the algorithm just as R. The procedure is as follows:

1. Identify the minimum $d_{\min} = \min(d_{ij}, d_{iB})$ with all possible combinations (i, j) of subjets
2. Now we have two possibilities:
   a) If $d_{\min} = d_{ij} < d_{\text{cut}}$ merge subjets i and j, keep it as a new subjet i and continue with 1.
   b) If $d_{\min} = d_{iB} < d_{\text{cut}}$ call i beam radiation and remove it from the iteration, continue with 1.
3. Iterate until there is only one jet left or $d_{\min} > d_{\text{cut}}$.

This kind of jet-algorithm we call exclusive jet-algorithm. Fig. 3.2 shows a comparison of the final jet shapes of these algorithms. They are produced by taking parton-level events together with massless objects (ghosts). It illustrates whether they end up in a jet and in which one.

Besides, there also exist inclusive algorithms. In this case d_{iB} acts as a cutoff:

1. Identify the minimum $d_{\min} = \min(d_{ij}, d_{iB})$ with all possible combinations (i, j) of subjets
2. For $d_{\min} = d_{ij}$ merge subjets i and j, keep it as a new subjet i and continue with 1.

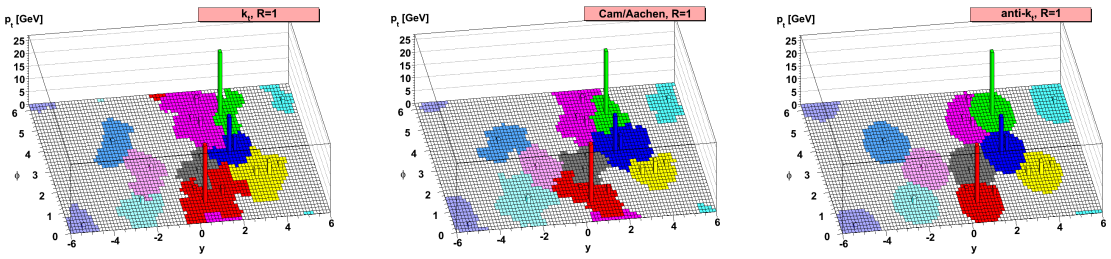

Figure 3.2: Final jet shape for different jet recombination algorithms. Taken from [29].

3. If $d_{\min} = d_{iB}$ take i as a final state subjet and remove it from the iteration, continue with 1.

and ends if there are no subjets left above a certain threshold. Here the jet-beam distance d_{iB} is also the scale of the jet-jet separation.

In our analysis, we want to reconstruct boosted heavy particles namely a Higgs boson and a top quark. These objects correspond to jets with a large geometric size called fat jets. In these fat jets we expect smaller subjets reflecting decay products of the heavy particle. For investigating boosted fat jets, the k_T and the C/A algorithms will work best since their clustering history is similar to the branching history. In our analysis, we make use of the C/A algorithm which is implemented in FastJet3 [8].

4 Hunting Tops and Higgs

In this section we describe the default version of the HEPTopTagger2 and the BDRS Higgs tagger. Especially the new HEPTopTagger2 with its updates achieves significant improvements. For our analysis we only switch on two of these modes for the HEPTopTagger2 described below. These two modes will also be added to the BDRS HiggsTagger resulting in an updated version to recent improvements of fat jet analysis [30].

4.1 The Default HEPTopTagger2

To find a top quark in our data the recently updated HEPTopTagger2 [4] will be used. We only sketch the tagger and describe the important default steps here. A detailed description with additional modes can be found in [4]. The procedure is as follows:

1. **Defining fat jets:** Define a fat jet via the C/A with $R = 1.8$.

2. **Search for hard substructure in a fat jet:** For finding relevant hard substructure inside a fat jet, we apply a recursive procedure.
   a) First, we take two subjets j_1 and j_2 by undoing the last clustering of the fat jet j and order them such that $m_{j_1} > m_{j_2}$.
   b) With a given mass drop criterion f_{drop}, we test if $m_{j_1} > f_{\text{drop}} m_j$.
      i. For that case, we discard j_2 seeing it as coming from an underlying event or soft QCD emission and keep j_1.
      ii. Otherwise we keep both subjets and add it to the relevant substructure for $m_{j_i} < m_{\text{max}}$ (a certain maximum mass) or we continue with a) for j_1 and j_2.

   In conclusion, we only get relevant hard substructure and can search for certain quarks inside. We perform the algorithm with $f_{\text{drop}} = 0.8$ and $m_{\text{max}} = 30$ GeV.

3. **Filtering of subjets:** We iterate through all triplets of three hard subjets. To reduce the contamination of our jet recombination by ISR, UE, and pile-up, we filter the hard substructures. This will be done by a filtering based on [5]. With the angular scale

$$R_{\text{filt}} = \min(0.3, \frac{\Delta R_{jk}}{2}) \, , \tag{4.1}$$

   where ΔR_{jk} is the closest angular distance between two substructures of the triple, we recluster them again in the C/A way. Started with a fat jet, followed by a reclustering into subjets we finally keep the five hardest constituents and calculate their jet mass.

4. **mass window:** Then we recluster these five constituents via C/A to three subjets j_1,j_2 and j_3 ordered by p_T and only continue with those triplets in a mass window of $[150, 200]$ GeV. The three subjets should reflect the three top decay subjets.

5. **mass plane cuts:** After ordering all j_i by p_T, we calculate all combinations of masses (m_{12}, m_{13}, m_{23}) plus m_{123}. Considering that j_i stands for a top decay product we can set two mass values. For one, m_{123} should be mapped to the top mass m_t. Moreover, the criterion $m_{jk} = m_W$ should hold for one pair (j, k). Assuming approximately massless particles i.e. that $p_i^2 = m_i^2 \approx 0$ in high energy regions, we can write $m_t^2 \equiv m_{123}^2 = m_{12}^2 + m_{13}^2 + m_{23}^2$. With a fixed $m_{123} = m_t$ and $m_W = m_{jk}$ for one (j, k), we remain with two degrees of freedom to fully describe the kinematics. For these two variables, we choose m_{23}/m_{123} and $\arctan(m_{13}/m_{12})$. The distribution of events in relation to the two mass plane variables can be seen in Fig. 4.1. Then for accepting the three subjets as a top candidate, one of the three following criteria has to be satisfied to identify one m_{jk} as m_W:

$$0.2 < \arctan\left(\frac{m_{13}}{m_{12}}\right) < 1.3 \text{ and } R_{\text{min}} < \frac{m_{23}}{m_{123}} < R_{\text{max}} \tag{4.2}$$

$$R_{\text{min}}^2\left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\text{max}}^2\left(1 + \frac{m_{13}}{m_{12}}\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35 \tag{4.3}$$

$$R_{\text{min}}^2\left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\text{max}}^2\left(1 + \frac{m_{12}}{m_{13}}\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35 \tag{4.4}$$

with $R_{\text{min,max}} = (1 \mp f_W)\frac{m_W}{m_t}$ and a parameter f_W which is set to $f_W = 0.15$ by default to reach a range close to m_W. The term $\frac{m_{23}}{m_{123}} > 0.35$ represents the soft cutoff excluding most of the entries in the W + jets and QCD samples in Fig. 4.1.

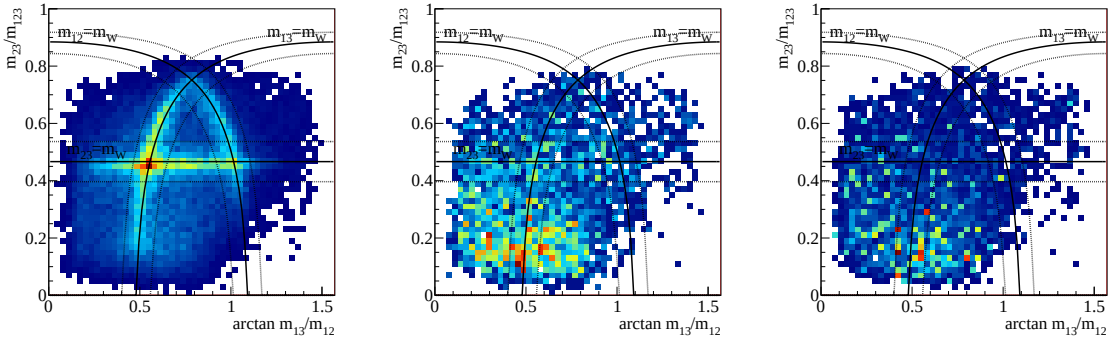

Figure 4.1: Comparison of the mass planes for $t\bar{t}$ (left),W+jets (middle) and pure QCD jets (right) before the mass plane cut. Taken from [3].

6. From all remaining candidates we choose the one with m_{123} closest to the actual top mass m_t.

7. Finally one requires for consistency that the transverse momentum p_T of the reconstructed candidate is larger than 200 GeV.

Additionally, the default HEPTopTagger2 contains the optimalR mode as well as the N-subjettiness mode. They will be described in detail below.

4.2 BDRS HiggsTagger

We use the BDRS HiggsTagger like it is described in [6] and modify it later-on. The tagging procedure is as follows:

1. To begin with, we form C/A fat jets with $R = 1.2$.
2. Similar to the HEPTopTagger2, we search for the hard substructure, this time with a cutoff at $m_{\max} = 40$ GeV and a mass drop threshold of 0.9.
3. At this stage, we try to find relevant pairings of jets for a Higgs mass reconstruction. For that reason, we take all possible pairs of subjets obtained by the hard substructure decomposition and calculate their modified Jade distance given by

$$J = p_{T,1} p_{T,2} (\Delta R_{12})^4. \tag{4.5}$$

   The original Jade distance $J_0 = p_{T,1} p_{T,2} (\Delta R_{12})^2$ is linked to the reconstructed mass $J_0 \approx m_{12}$. The modified Jade distance enhances the weight of the geometric separation. We take the three leading pairings based on the Jade distance and filter it in the end again with $R_{\text{filt}} = \min(0.3, \frac{\Delta R_{jk}}{2})$ including the three hardest substructures.
4. Again, we set a p_T-cut at 200 GeV.

For setting any mass window, we can also require that the reconstructed mass has to be in a certain mass range. If it is helpful or not depends on the analysis. Anyway, in this analysis the mass window criterion has to be switched off. The Higgs subjets represent the bottom quarks of a Higgs decay including an additional radiation as already described above.

4.3 Extensions of the Default Versions

4.3.1 N-subjettiness

An additional way to distinguish between a 'signal jet' and a background based jet is based on N-subjettiness [17]. N-subjettiness could also be used as a tagger with its characteristic jet shape variable τ_N given below in (4.6) . In this thesis, we use it in the BDRS Higgs tagger as a cut criterion to reduce the background as well as in the HEPTopTagger2 for a mild cut to have a handle on the QCD multi-jet background. First, one performs a certain jet clustering, e.g. k_T or C/A on the constituents of the fat jet to be investigated forced to produce exactly N axes. Having these axes, we calculate τ_N which is given by

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \Delta R_{3,k}, ..., \Delta R_{N,k}) \,, \tag{4.6}$$

where $p_{T,k}$ is the transverse momentum of a hard structure k of the fat jet and $\Delta R_{i,k}$ is the angular separation to the i-th recently defined axis. The normalization factor d_0 is given by

$$d_0 = \sum_k p_{T,k} R_0 \tag{4.7}$$

with the original jet radius of the fat jet R_0. This leads to values $\tau_N < 1$. Furthermore, (4.6) gives information about the amount of subjets the fat jet actually consists of. Small values of

τ_N indicate that the radiation is almost aligned to the determined subjet axis. Thus, the fat jet consists of N or less subjets. Besides, a large τ_N implies that the energy flow is not directed along the axes which means that at least $N+1$ axes exist in the fat jet. Taking this into account, the ratio $\frac{\tau_3}{\tau_2}$ is an appropriate value to identify top fat jets according to three hard substructures in the decay process and $\frac{\tau_2}{\tau_1}$ is the right one for finding Higgs fat jets including the two hard decay products $b\bar{b}$. For setting these cuts, we use the already FastJet-implemented version relying on the FastJet Contrib add-on for N-subjettiness [18].

4.3.2 OptimalR Mode

Another improvement is the optimalR mode. The aim of that method is to shrink the radius of a fat jet to an optimum for the same reason as in the filtering above e.g. to reduce the contamination of the jet by pile-up and underlying events. For achieving that, we introduce an algorithmic way [4, 22]. At the step of a tagger algorithm where we already have passed the main tagging criteria, we can take the reconstructed mass m_{rec} as a reference value $m_{\text{rec}}(R_{\text{max}}) \equiv m_{\text{ref}}$. Following this, we reduce the size of the fat jet by a certain value δR to $R_{\text{new}} = R_{\text{old}} - \delta R$ and calculate the mass again. Then again, we calculate the mass of the new fat jet $m(R_{\text{new}})$ and compare it to the reference value m_{ref}. At a certain point the mass will drop down significantly since the reduced radius can not include all characteristic decay jets anymore. As an optimal R, we finally take the value before the drop. For specifying this, one defines the drop through

$$\frac{m_{\text{ref}} - m(R)}{m_{\text{ref}}} > M_{\text{drop}} \Leftrightarrow R < R_{\text{opt}} \, . \tag{4.8}$$

M_{drop} stands for a maximal relative deviation of the new fat jet mass $m(R)$ to the reference mass m_{ref}. Both for the BDRS HiggsTagger and the HEPTopTagger2, we set $M_{\text{drop}} = 0.2$. Additionally, R_{opt} can be predicted from the fat jet kinematics in case of the HEPTopTagger2 [4] and the Higgs tagger as seen in Fig. 4.2. The fit yields a functional correlation [4]

$$R_{bjj} = \frac{327}{p_{T,\text{filt}}} \tag{4.9}$$

in the HEPTopTagger2.

Proceeding in the same way for the Higgs tagger, one can see that its behaviour is similar as seen in Fig. 4.2. We obtain a functional correlation

$$R_{bb} = \frac{250}{p_{T,\text{filt}}} \tag{4.10}$$

this time. This agrees perfectly with the predicted value [5]

$$R_{b\bar{b}} \simeq \frac{1}{\sqrt{z(1-z)}} \frac{m_H}{p_T} \geq \frac{2m_H}{p_T} \tag{4.11}$$

with the momentum fractions z and $1-z$ of the two bottom jets. Additionally, this confirms the choice of $R = 1.2$ as a starting point for the Higgs fat jets. In case of the HEPTopTagger2, we would repeat steps 1 to 4 with a decreasing fat jet radius followed by the drop criterion (4.8). For the BDRS Higgs tagger, we go through the whole tagging procedure before testing the fat

jet radius by relation (4.8).

4.4 Additional Bottom-tags

After having Higgs candidates including subjets, we can investigate them for a more precise analysis. Therefore, we require b-tags of subjets e.g. the two Higgs substructures or outside any substructure. The first case is done by compairing the subjets with the parton level b quarks. If a subjet has an angular separation ΔR to the parton level b closer than a pre-defined value R_b, we call it b-tagged. Finally, we obtain whether the two Higgs subjets are double tagged, single tagged, or not tagged. In addition, we can also require a third b-tag in the reclustered jets after the Higgs and top tag after removing all Higgs and top hadrons before the C/A reclustering. Beyond the R_b condition, we also require that the b tagged jet has an angular separation $\Delta R_{b,j}$ greater than a certain value to all Higgs and top substructure. Actually, this is no real tagger in terms of physical justification. In an experiment, we do not have information like parton-level bottom quarks. It is just a simple way to ensure that one of the remaining jets includes bottom quark structure.

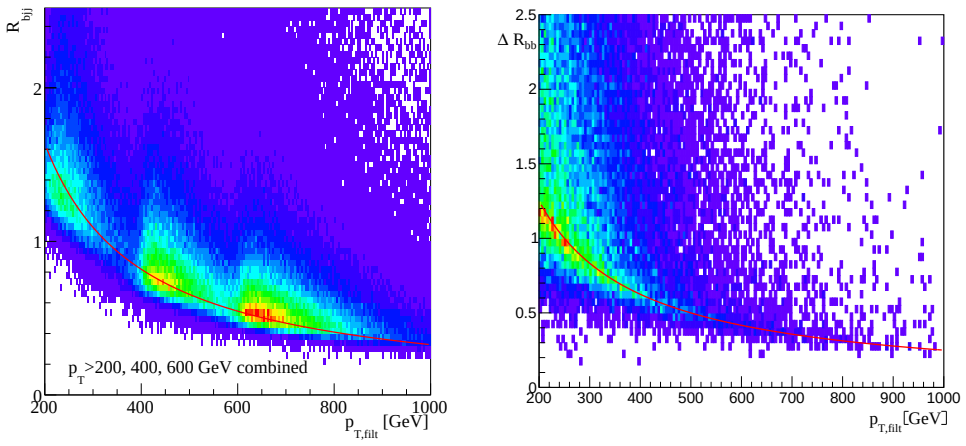

Figure 4.2: R_{opt} fit to a $t\bar{t}$ sample with the parton level distance of the decay products R_{bjj} dependent on $p_{T,\text{filt}} > 200, 400, 600$ GeV (left) and R_{opt} fit to a $t\bar{t}H$ sample with the parton level distance of the decay products ΔR_{bb} dependent on $p_{T,\text{filt}} > 200$ GeV (right). For both, the fat jets are filtered with $N = 10$ and $R = 0.2$. The left one is taken from [4].

5 Analysis of Boosted $t\bar{t}H$ at 100 TeV

The following analysis is based on the original LHC analysis [6] and is done in collaboration with Tilman Plehn and Torben Schell [30]. It will focus on the top pair production associated with a Higgs boson decaying into a bottom pair as seen in Fig. 2.3. Goal of this analysis is to extract signal regions from the underlying background processes we mentioned above in section 2.2 for processes at 100 TeV.

We can already see in Fig. 5.1 that we obtain a harder transverse momentum spectra of all particles for a 100 TeV. Hence the relative fraction of events with $p_{T,t} > m_t$ and $p_{T,H} > m_H$ is larger compared to 13 TeV. With high p_T fat jets, we will be able to work in a boosted regime in which we can run both taggers. The analysis strategy we use can easily been sketched as follows

1. an isolated lepton
2. a tagged top without any b-tag requirement
3. a tagged Higgs with two b-tags inside
4. a continuum b-tag outside the top and Higgs fat-jets .

First, we require one isolated lepton with:

$$p_{T,l} > 15 \text{ GeV}, \qquad |y_l| < 2.5 \,. \tag{5.1}$$

For the top tagger procedure [3, 4, 31, 32], we build fat jets with $R = 1.8$ and $p_{T,j} > 200$ GeV out of the hadronic structure we get from DELPHES3 and require that at least two fat jets

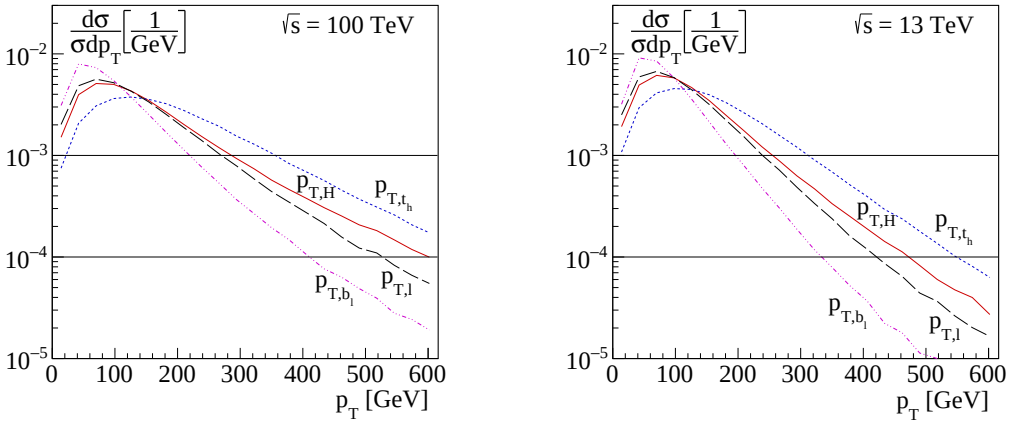

Figure 5.1: Transverse momentum distributions for the $t\bar{t}H$ signal process at a 100 TeV collider (left) and the 13 TeV LHC (right). Taken from [30].

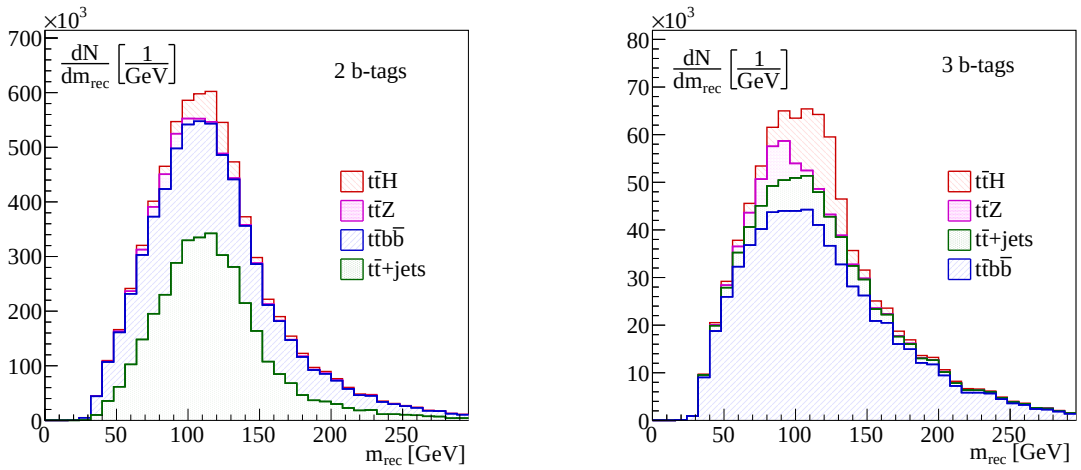

Figure 5.2: Reconstructed m_{bb} for the leading-J substructures in the fat Higgs jet. We require two b-tags inside the fat Higgs jet (left) and an additional continuum b-tag (right). The event numbers are scaled to $\mathcal{L} = 20\ ab^{-1}$. Taken from [30].

$m_{bb} \in [100, 150]$ GeV	2 b-tags	3 b-tags	ratio
$t\bar{t}H$	2.4E+5	6.4E+4	1/3.8
$t\bar{t}b\bar{b}$	1.2E+6	2.4E+5	1/5.0
$t\bar{t}+$ jets	1.9E+6	3.8E+4	1/50
$t\bar{t}Z$	2.3E+4	4.9E+3	1/4.7

Table 5.1: Event rates assuming an integrated luminosity of 20 ab^{-1} [30].

are built since we want to tag one top quark and one Higgs boson. Due to the choice of our process (Fig 2.3) we require exactly one top tag to be tagged in the HEPTopTagger2 which is sensitive to hadronic tops.

We continue our procedure with clustering new fat jets with $R = 1.2$ and again with $p_{T,j} > 200$ GeV after removing all top hadrons from the hadronic content of the event. With these fat jets we go through the BDRS Higgs tagger and obviously require one Higgs tagged jet. At this stage we do not include N-subjettiness or the optimalR mode. Inside the Higgs fat jet we require $|y_j| < 2.5$ and $p_{T,j} > 30$ GeV for the subjets. After that, we ask for two b-tags. We assume a global tagging efficiency of 50% and a mistagging probability of 1%. For calling a subjet tagged, it has to be closer than $R_b = 0.3$ to the parton level b quark. As one can see in the left panel of Fig. 5.2, up to now the $t\bar{t}b\bar{b}$ and $t\bar{t}+$jets backgrounds are of similar size and totally dominate the reconstructed mass spectra. Furthermore, this kind of procedure causes a peak at $m_{bb} \sim 100$ GeV.

To make the background composition simpler and handier, we ask for a third continuum b-tag outside the Higgs and top substructure. The objective of this requirement is to gain control of the $t\bar{t}+$jets background by targeting the b quark from the leptonically decaying top. For this part, we use C/A jets with $R = 0.6$ and $p_{T,j} > 30$ GeV and tag within $|y_b| < 2.5$ demanding an angular separation $\Delta R_{b,j} > 0.4$. This effects a significant improvement as one can see in the right panel of Fig. 5.2. To support this, one can take a look at the correspondig

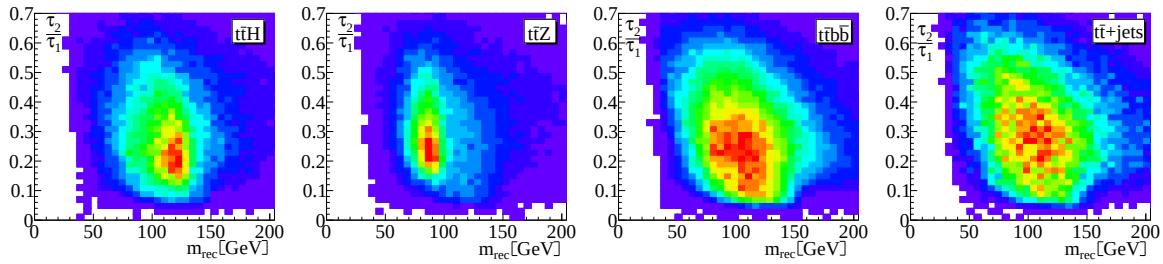

Figure 5.3: Correlation between the reconstructed mass m_{rec} and the N-subjettiness ratio τ_2/τ_1 of the filtered Higgs candidate fat jet for the signal and background samples. The event numbers are scaled to $\mathcal{L} = 20\ ab^{-1}$. Taken from [30].

event rates listed in Tab. 5.1. The $t\bar{t}H$ sample is reduced least of all. In addition while the samples are surpressed about a factor 1/5, the $t\bar{t}+$jets sample is surpressed about a factor 1/10 more than the $t\bar{t}b\bar{b}$ and the $t\bar{t}Z$ samples. As a result, $t\bar{t}b\bar{b}$ now dominates the continuum background. The $t\bar{t}+$jets background is still in the order the Higgs signal and can not be neglected in our analysis.

Now we make use of the two improvements of the HEPTopTagger2 and modify the BDRS Higgs tagger. First, we look at the N-subjettiness behaviour of our samples. Since we focus on the $H \to b\bar{b}$ decay in a Higgs jet, the characteristic N-subjettiness ratio which should get very small has to be τ_2/τ_1. In Fig. 5.3 the correlations between the reconstructed mass and τ_2/τ_1 of the filtered fat jet is given. This entails that a cut which we set to be

$$\frac{\tau_2}{\tau_1} < 0.4 \tag{5.2}$$

seems to be reasonable. As a consequence, the backgrounds can be reduced and our final signal sculpting samples $t\bar{t}H$ and $t\bar{t}Z$ get narrower and sharper mass peaks as seen in Fig. 5.4.

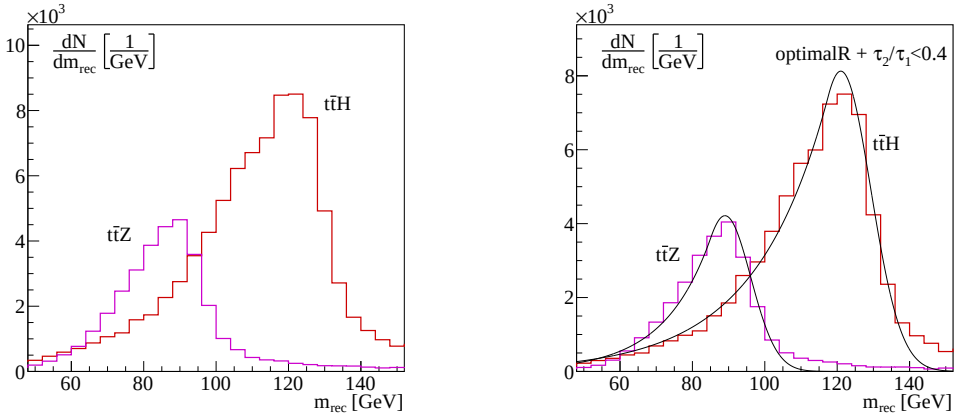

Figure 5.4: Reconstructed m_{bb} of the Higgs and Z candidates in $t\bar{t}H$ and $t\bar{t}Z$ production with the default BDRS tagger (left) and after using optimalR and the N-subjettiness cut $\tau_2/\tau_1 < 0.4$ (right). In the right panel we include the fitted Crystal Ball functions. The event numbers are scaled to $\mathcal{L} = 20\ ab^{-1}$. Taken from [30].

Secondly, we switch on the optimalR mode. Apart from reducing certain contaminations of the fat jet, a smaller fat jet also cuts down the combinatorics of the ensuing mass reconstruction m_{bb}. The optimalR mode we perform in the same way as for a top tag by shrinking the radius of the fat jet in steps of 0.1 till we reach a drop criterion $m_j < 0.8 m_{j,\text{orig}}$ with the original mass of the $R = 1.2$ Higgs fat jet. As described above in section 4.3.2 one can get an calculated $R_{b\bar{b}}$ from a fit to Monte Carlo simulations. A comparison of the calculated and determined value of $R_{b\bar{b}}$ like in a reference value $R_{b\bar{b}} - R_{b\bar{b}}^{\text{calc}}$ does not indicate a distinguishing feature of the $t\bar{t}H$ sample to impel an overall background reduction. However, a shift between the $t\bar{t}Z$ and the $t\bar{t}H$ samples is visible in a $R_{b\bar{b}} - R_{b\bar{b}}^{\text{calc}}$ distribution. So if desired, the Z-peak can be reduced.

6 Results

After adding the two improvements to our analysis combined with a triple b-tag, the result is a mass distribution shown in the left panel of Fig. 6.1. One can see a clearly visible Higgs peak in the range $m_{bb} \in [100, 140]$ GeV in combination with smooth and signal untouched background regions $m_{bb} \in [0, 60]$ GeV and $m_{bb} \in [160, 300]$ GeV. For a signal region $m_{bb} \in [104, 136]$ GeV, we obtain a signal-to-background ratio around $S/B \approx 1/3$ and a Gaussian significance $S/\sqrt{B} = 120$ for an assumed luminosity of $\mathcal{L} = 20\ ab^{-1}$. We calculate the error of the signal events $N_S = 44700$ by including two terms. The first term includes that we can determine N_S from the total event number $N_S + N_B$ using a perfect determination of N_B through the side band. For the second term, we calculate a statistical uncertainty ΔN_B consisting of the side band $m_{bb} \in [160, 300]$ event number $N_{\text{side}} = 135000$ and a relative error of $1/\sqrt{N_{\text{side}}}$ which finally gives us the estimation

$$\begin{aligned} \Delta N_S &= \left[\left(\sqrt{N_S + N_B} \right)^2 + (\Delta N_B)^2 \right]^{1/2} \\ &= \left[\left(\sqrt{N_S + N_B} \right)^2 + \left(\frac{N_B}{\sqrt{N_{\text{side}}}} \right)^2 \right]^{1/2} = 0.013\, N_S \ . \end{aligned} \tag{6.1}$$

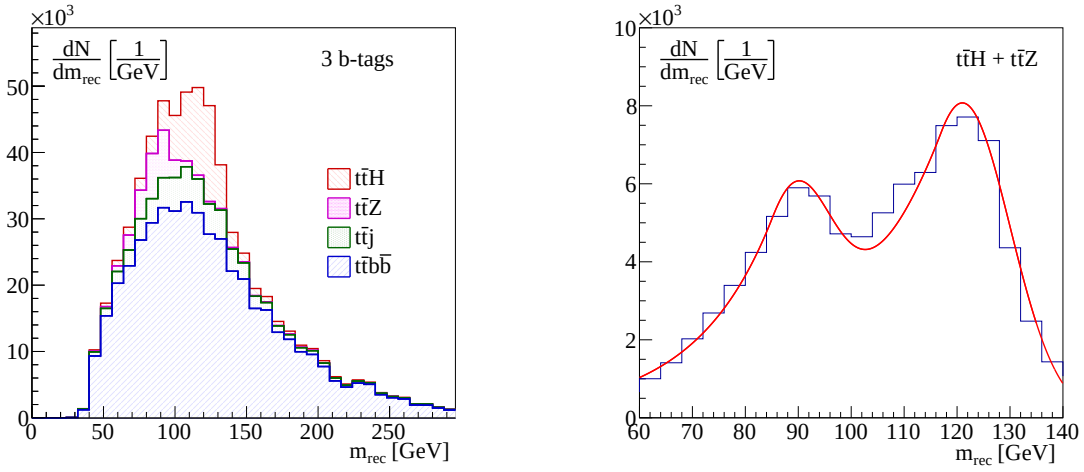

Figure 6.1: Left: Reconstructed m_{bb} for the leading-J substructures in the fat Higgs jet. We require two b-tags inside the fat Higgs jet and a continuum b-tag. Unlike in Fig. 5.2, we apply an N-subjettiness cut and use an optimalR version of the BDRS tagger. Right: Double-peak fit assuming perfect continuum background subtraction. The event numbers are scaled to $\mathcal{L} = 20\ ab^{-1}$. Taken from [30].

By only regarding the first term, we would have $\Delta N_S = 0.010 N_S$. According to this fact, the precision of a Yukawa coupling measurement is in the range of around 1%.

Nevertheless, we still remain with systematic and theoretical uncertainties we can not estimate very well. For that reason, we perform a combined fit to the Z and Higgs peak with known masses based on the assumption of a perfect background substraction. We regard separated simulations for the Higgs and Z peak and fit a Crystal Ball function [33] to both samples as one can see in the right panel of Fig. 5.4. The exponent of the non-Gaussian part of the Crystal Ball function is limited to 50. Knowing both masses, we fix the peak position of each fit. With the information gained by these fits, we perform a double-peak fit. The only parameter we do not fix is the scaling factor of each peak. As a result of the combined fit as seen in Fig. 6.1, we get a relative size $N_H/N_Z = 2.80 \pm 0.03$ of both peaks. This indeed provides an investigation of the top Yukawa coupling with $\sim 1\%$ precision. As decribed in [30], it represents a value where many systematic and theoretical uncertainties cancel.

7 Conclusion and Outlook

In the former sections, we have used the recently updated HEPTopTagger2 for a boosted $t\bar{t}H$ analysis at 100 TeV. We added two significant updates to the BDRS Higgs tagger to perform an up-to-date fat jet analysis for the boosted Higgs jets. These two taggers were applied to investigate the $t\bar{t}(H \to b\bar{b})$ decay channel which is useful for a precise measurement of the top Yukawa coupling.

With a very simple analysis strategy [6] including a trigger lepton, one fat jet for the hadronic Higgs decay as well as one fat jet for the hadronic top decay and a continuum b-tag we reached a remarkable excess of the Higgs signal and a clearly visible Z peak. A smooth and signal untouched side band leads to a well executable background subtraction. The additional Z peak enables us to cancel many systematic and theoretical uncertainties by a combined fit and relative event size calculation of the Higgs signal and the peaked $t\bar{t}Z$ background.

In conclusion, we arrive at a promising relative uncertainty estimation of the signal of $\Delta N_S/N_S = 0.013$. We also performed a double fit yielding a calculated ratio of $N_H/N_Z = 2.80 \pm 0.03$ for the two peaks. These two ways allow us to state that a probe of the top Yukawa coupling at a 100 TeV collider can be measured with a precision of $\sim 1\%$. This is an order of magnitude improvement compared to the limits $\Delta y_t/y_t \approx 10\%$ of LHC determinations [35] due to statistical and theoretical uncertainties. As a result, the thesis gives kind of a foretaste of what can be reached by a 100 TeV collider.

Acknowledgements

First of all, I would like to thank Tilman Plehn for giving me the opportunity to work in his group and to get an insight into current research. Furthermore, I want to thank Torben for helping me whenever I needed it and guiding me through this project. I really appreciate your dedication to help me at any stage of the thesis. Last but not least thanks to Tilman Plehn's group as well as Jörg Jäckel's group for a lovely working atmosphere.

Bibliography

[1] P. W. Higgs, Phys. Lett. **12**, 132 (1964); P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964); F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[2] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012), S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012).

[3] T. Plehn, M. Spannowsky, M. Takeuchi and D. Zerwas, JHEP **1010** (2010) 078 [arXiv:1006.2833 [hep-ph]].

[4] G. Kasieczka, T. Plehn, T. Schell, T. Strebler and G. P. Salam, arXiv:1503.05921 [hep-ph].

[5] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, arXiv:0810.0409 [hep-ph].

[6] T. Plehn, G. P. Salam and M. Spannowsky, Phys. Rev. Lett. **104** (2010) 111801 [arXiv:0910.5472 [hep-ph]].

[7] J. Anderson, A. Avetisyan, R. Brock, S. Chekanov, T. Cohen, N. Dhingra, J. Dolen and J. Hirschauer *et al.*, arXiv:1309.1057 [hep-ex].

[8] M. Cacciari, G. P. Salam and G. Soyez, Eur. Phys. J. C **72** (2012) 1896 M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006);

[9] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna and S. Prestel *et al.*, Comput. Phys. Commun. **191** (2015) 159

[10] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, JHEP **1106** (2011) 128 [arXiv:1106.0522 [hep-ph]]; J. Alwall *et al.*, JHEP **1407** (2014) 079 [arXiv:1405.0301 [hep-ph]].

[11] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, JHEP **9708**, 001 (1997); M. Wobisch and T. Wengler, In "Hamburg 1998/1999, Monte Carlo generators for HERA physics" 270-279. [hep-ph/9907280].

[12] M. Selvaggi, J. Phys. Conf. Ser. **523** (2014) 012033.

[13] R. D. Ball *et al.* [NNPDF Collaboration], Nucl. Phys. B **877**, 290 (2013).

[14] Wikimedia Commons, `http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg` (May 31, 2015).

[15] C. Berger, "Elementarteilchenphysik: Von den Grundlagen zu den modernen Experimenten," 3. Auflage, Springer Verlag, 2014.

[16] B. Povh, K. Rith, C. Scholz, F. Zetsche, W. Rodejohann, "Teilchen und Kerne", 9. Auflage, Springer-Verlag, 2013.

[17] J. Thaler and K. Van Tilburg, JHEP **1103** (2011) 015 [arXiv:1011.2268 [hep-ph]]; J. Thaler and K. Van Tilburg, JHEP **1202**, 093 (2012); [arXiv:1108.2701 [hep-ph]]. I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, Phys. Rev. Lett. **105** (2010) 092002 [arXiv:1004.2489 [hep-ph]].

[18] `http://fastjet.hepforge.org/contrib/`.

[19] `http://lcgapp.cern.ch/project/simu/HepMC/206/HepMC2_user_manual.pdf`

[20] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, JHEP **0902** (2009) 007 [arXiv:0811.4622 [hep-ph]].

[21] J. Alwall, Michelson lectures at Case Western Reserve, April 13-16, 2009, `http://www.phys.cwru.edu/events/mppl/Johan/Michelson-lecture-3.pdf`

[22] T. Strebler, Master thesis, ETH Zürich (2014).

[23] T. Han, "Collider phenomenology: Basic knowledge and techniques," hep-ph/0508097.

[24] S. Weinberg, "A Model of Leptons," Phys. Rev. Lett.19, 1264 (1967).

[25] A. Salam, "Weak and Electromagnetic Interactions," Conf. Proc. C **680519** (1968) 367.

[26] T. Plehn, Lect. Notes Phys. **886** (2015).

[27] M. Dasgupta, L. Magnea and G. P. Salam, JHEP **0802** (2008) 055 [arXiv:0712.3014 [hep-ph]].

[28] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, "Longitudinally invariant Kt clustering algorithms for hadron hadron collisions," Nucl. Phys. B **406**, 187 (1993).

[29] M. Cacciari, G. P. Salam and G. Soyez, "The Anti-k(t) jet clustering algorithm," JHEP **0804** (2008) 063 [arXiv:0802.1189 [hep-ph]].

[30] M. L. Mangano, T. Plehn, P. Reimitz, T. Schell and H. S. Shao, arXiv:1507.08169 [hep-ph].

[31] K. Agashe *et al* Phys. Rev. D **77**, 015003 (2008); G. Brooijmans, ATL-PHYS-CONF-2008-008 and ATL-COM-PHYS-2008-001, Feb. 2008 J. Thaler and L. T. Wang, JHEP **0807**, 092 (2008); D. E. Kaplan, K. Rehermann, M. D. Schwartz and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008); L. G. Almeida *et al* , Phys. Rev. D **79**, 074017 (2009); L. G. Almeida *et al* , Phys. Rev. D **79**, 074012 (2009).

[32] M. H. Seymour, Z. Phys. C **62**, 127 (1994); T. Plehn and M. Spannowsky, J. Phys. G **39**, 083001 (2012); A. Abdesselam *et al.*, Eur. Phys. J. C **71**, 1661 (2011); A. Altheimer *et al.*, J. Phys. G **39**, 063001 (2012).

[33] M. Oreglia, SLAC-236 (80,REC.APR. 81), J. Gaiser, SLAC-255 (82,REC.JUN.83) T. Skwarnicki, DESY-F31-86-02.

[34] BwUniCluster (`https://www.bwhpc-c5.de`), "This work was performed on the computational resource bwUniCluster funded by the Ministry of Science, Research and the Arts Baden-Württemberg and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC."

[35] M. Klute, R. Lafaye, T. Plehn, M. Rauch and D. Zerwas, Europhys. Lett. **101**, 51001 (2013);

[36] D. Binosi, J. Collins, C. Kaufhold and L. Theussl, Comput. Phys. Commun. **180** (2009) 1709 [arXiv:0811.4113 [hep-ph]].

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 23.10.2015 .....